Published in final edited form as:

Pharm Stat. 2010 January ; 9(1): 2–9. doi:10.1002/pst.359.

Effects of correlation and missing data on sample size estimation in longitudinal clinical trials

Song Zhang and Chul Ahn^{*,†}

Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX, USA

Abstract

In longitudinal clinical trials, a common objective is to compare the rates of changes in an outcome variable between two treatment groups. Generalized estimating equation (GEE) has been widely used to examine if the rates of changes are significantly different between treatment groups due to its robustness to misspecification of the true correlation structure and randomly missing data. The sample size formula for repeated outcomes is based on the assumption of missing completely at random and a large sample approximation. A simulation study is conducted to investigate the performance of GEE sample size formula with small sample sizes, damped exponential family of correlation structure and non-ignorable missing data.

Keywords

damped exponential correlation; missing data; rates of changes; non-ignorable missingness

1. INTRODUCTION

In controlled clinical trials, subjects are often evaluated at baseline and intervals across a treatment period. A common objective of longitudinal clinical trials is to investigate whether the rates of changes in an outcome variable are significantly different between two treatment groups. Generalized estimating equation (GEE) has been widely used to compare the rates of changes between treatment arms due to its robustness to misspecification of the true correlation structure and randomly missing data [1].

Sample size calculation is an important issue in the design of longitudinal clinical trials to compare the rates of changes between two treatment groups. Several authors present sample size estimation for repeated measurements studies using GEE. Liu and Liang [2] propose a general sample size formula for studies with correlated observations using GEE. Their sample size formula is applicable for both clustered data and longitudinal data. Their method is based on the generalized score test and the resulting score test statistic follows an asymptotically non-central χ^2 distribution under the alternative hypothesis. No closed-form sample size formula is available except for some special cases. Rochon [3] proposes a sample size formula using a non-central version of the Wald χ^2 test statistics. The sample size formula is provided to handle monotone missing data patterns. Jung and Ahn [4] derive an explicit closed-form sample size formula using GEE for the comparison of the rates of changes between two treatment groups.

*Correspondence to: Chul Ahn, Department of Clinical Sciences, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., E5.506, Dallas, TX 75390, USA.

†Chul.Ahn@UTSouthwestern.edu

Their formula handles both independent and monotone missing data patterns and general correlation structures.

The sample size formula of Jung and Ahn [4] is based on an asymptotic theory and the assumption of missing completely at random (MCAR), see Rubin [5]. In this paper, we investigate the performance of the sample size formula with small sample sizes, damped exponential family of correlation structures and non-ignorable missing data.

2. SAMPLE SIZE ESTIMATION

For subject i ($1 \leq i \leq n$), let y_{ij} denote a continuous response variable at the measurement time t_{ij} ($j=1, \dots, K_i$). We consider the linear model

$$y_{ij} = \beta_1 + \beta_2 r_i + \beta_3 t_{ij} + \beta_4 r_i t_{ij} + \varepsilon_{ij} \quad (1)$$

where serial correlation may exist among error terms $\varepsilon_{i1}, \dots, \varepsilon_{iK_i}$, which have mean zero and variance σ^2 . Here, r_i is the treatment indicator taking 0 for control group and 1 for treatment group. Observations are assumed to be MCAR.

Suppose that each subject is scheduled to be assessed at K time points. Let $\sigma^2 = \text{var}(\varepsilon_{ij})$ and $\rho_{jj'} = \text{corr}(\varepsilon_{ij}, \varepsilon_{ij'})$ for $j \neq j'$ and $\rho_{jj} = 1$. Let $p_j = E(\delta_{ij})$ be the proportion of subjects with assessment at t_j , and $p_{jj'} = E(\delta_{ij}\delta_{ij'})$ be the proportion of subjects with assessments at both t_j and $t_{j'}$ ($p_{jj} = p_j$).

With the two-sided α and power $1-\gamma$, the required sample size to test $H_0: \beta_4 = 0$ versus $H_a: \beta_4 \neq 0$ is given by

$$n = \frac{\sigma^2 s_t^2 (z_{1-\alpha/2} + z_{1-\gamma})^2}{\beta_4^2 \mu_0^2 \sigma_r^2 \sigma_t^4} \quad (2)$$

where $\mu_0 = \sum_{j=1}^K p_j$, $\mu_k = \mu_0^{-1} \sum_{j=1}^K p_j t_j^k$ for $k=1, 2$, $\sigma_t^2 = \mu_2 - \mu_1^2$, $\bar{r} = E(r_i)$, $\sigma_r^2 = \bar{r}(1 - \bar{r})$ and $s_t^2 = \sum_{j=1}^K \sum_{j'=1}^K p_{jj'} \rho_{jj'} (t_j - \mu_1)(t_{j'} - \mu_1)$. See Jung and Ahn [4] for details.

3. SIMULATION STUDIES

The adequacy of the sample size formula will be assessed through simulation for large effect sizes and non-ignorable missing situations. We estimate the required number of subjects (n) using equation (2), and then generate for each of 5000 simulated samples n subjects from a multivariate normal distribution with the damped exponential correlations and missing data patterns. We consider two missing patterns. One is independent missing, where $p_{jj'} = p_j p_{j'}$, and the other is monotone missing, where $p_{jj'} = p_{j'}$ for $j < j'$ ($p_1 \geq \dots \geq p_K$). We set the maximum number of measurements for each subject at $K=6$, and use the following vectors for the probability of assessment at each time point:

$$\begin{aligned}
P_1 &= (1, 0.82, 0.79, 0.76, 0.73, 0.7) \\
P_2 &= (1, 0.94, 0.88, 0.82, 0.76, 0.7) \\
P_3 &= (1, 1, 1, 0.9, 0.8, 0.7) \\
P_4 &= (1, 1, 1, 1, 1, 1)
\end{aligned} \tag{3}$$

Note that P_1 , P_2 and P_3 describe the scenarios where an increasing number of subjects miss visits over time, with the dropout rate at the end of study being 0.3. We use P_4 to denote no missing data. In sample size estimation, a common practice is to use $n = n_0/(1 - q)$, where n_0 is the sample size estimate under no dropout and q is the expected dropout rate. Following this procedure, the final sample size estimates under P_1 , P_2 and P_3 would all be $n = n_0/0.7$, even though different amounts of information have been lost in the data. Simply adjusting the sample size estimate by $n = n_0/(1 - q)$ is a crude and conservative method. As we would demonstrate in the simulation study, the sample sizes estimated by (2) under P_1 , P_2 and P_3 are usually much less than the adjusted sample size under P_4 by a factor of 0.7.

We estimate the sample size using the sample size formula given in equation (2) with the prespecified missing data patterns, missing probabilities, damped exponential correlation structures, $\sigma^2=1$, type I error $\alpha=0.05$, power $1-\gamma=0.8$, $\beta_{40} = 0.1$ and $\bar{r} = 0.5$. Once the sample size (n) is estimated, we generate 5000 replicated samples of n subjects using $\beta = (0.3, 0.3, 0.5, 0.1)$ with correlated measurement errors generated from the multivariate normal distribution. Note that the sample size estimate does not depend on the values of β_1 , β_2 and β_3 .

3.1. Damped exponential family of correlation structures

We examine the effect of correlation using a damped exponential family of correlation structures proposed by Munoz *et al.* [6]. The correlation between two observations separated by s units of time is modeled by $\rho_{j,j+s} = \rho^c$, where $c = s^\theta$. Here, ρ is the correlation between observations separated by one unit and θ is a damping parameter. The damped exponential correlation structure provides a rich family of correlation structures. Compound symmetry (CS) and first-order autoregressive (AR(1)) correlation structures are obtained by setting $\theta=0$ and 1. The correlation structures of the repeated measurements change from CS to AR(1) in graded steps as θ increases from 0 to 1. We investigate the effect of the correlation structures on sample size estimate using $\theta=0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ and 1 and $\rho=0.1, 0.25$ and 0.5.

Table I presents the sample size estimates obtained to produce 80% power along with the empirical powers computed from the simulated samples under independent and monotone missing data patterns, respectively. As θ increases, sample size increases since the dependency decays when two measurement times get farther apart. The estimated sample size decreases as ρ increases. The empirical power obtained using the estimated sample size is generally close to the nominal value of 80%. Furthermore, given ρ and θ , the estimated sample sizes decrease in the order from P_1 to P_4 . Comparing the estimated sample sizes under independent missingness and monotone missingness, we found that the difference increases with ρ . For example, under P_2 and the CS correlation structure, the difference in sample sizes is $199-197=2$ when $\rho=0.1$, $175-169=6$ when $\rho=0.25$ and $135-124=11$ when $\rho=0.5$. Table I also reports that adjusting for missing data by $n_0/(1-q)$ is too conservative. For example, under monotone missing pattern, for $\rho=0.25$ and $\theta=1$, the estimated sample sizes under P_1 , P_2 and P_3 are 270, 266 and 262, respectively. If we use the naive adjustment method, the required sample size would be $\frac{224}{0.7}=320$.

To better understand the impact of dropout on sample size and power, we conduct a simulation study with the dropout rate at the end of study equal to 0.5. The vectors of assessment probability are defined as follows:

$$\begin{aligned}
P_1^* &= (1, 0.7, 0.65, 0.6, 0.55, 0.5) \\
P_2^* &= (1, 0.9, 0.8, 0.7, 0.6, 0.5) \\
P_3^* &= (1, 1, 1, 0.83, 0.67, 0.5) \\
P_4^* &= (1, 1, 1, 1, 1, 1)
\end{aligned} \tag{4}$$

The comparison of Tables I and II indicates that the dropout rate has a great impact on the sample size. For a given missing pattern, ρ and θ , a higher dropout rate leads to a much larger sample size. The differences between independent and monotone missing patterns are more pronounced under the higher dropout rate. For example, under P_2^* and the CS correlation structure, the difference in sample sizes is $240-234=6$ when $\rho=0.1$, $220-206=14$ when $\rho=0.25$ and $187-159=28$ when $\rho=0.5$.

3.2. Small sample size

The GEE method is based on a large sample approximation. A large value of β_4 will lead to a small sample size estimate. We investigate the performance of sample size formula when the estimated sample size is small. We set $\beta_{40} = 0.2$, and then estimate the required sample size n to achieve 80% power using the assessment probability vectors given in (3) and the damped exponential correlation structure under independent and monotone missing data patterns. Five thousand samples of repeated measurements data are generated. Table III reports the estimated sample sizes and the corresponding empirical powers calculated from the simulated samples under independent and monotone missing data patterns. According to formula (2), the sample size estimate is proportional to $1/\beta_{40}^2$. Thus, the estimated sample sizes in Table III are about one-fourth of those estimated in Table I. The empirical powers are generally close to the nominal 80% power.

3.3. Non-ignorable missing

The sample size formula given in equation (2) is constructed under the MCAR assumption, which has been assumed in Sections 3.1 and 3.2. In this section, we investigate the performance of the sample size formula under non-ignorable missingness, where the missing probability depends on unobserved outcomes [7]. Specifically, we assume that a higher outcome value leads to a lower chance of being observed. We define $\mu_{0j} = \beta_1 + \beta_3 t_j$, $j=1, \dots, K$, to be the mean response at time j for the control arm. The probability of followup for an individual i at time j is

$$q_{ij} = p_j \frac{v_{ij}}{\sum_{i=1}^n v_{ij}/n} \tag{5}$$

with $v_{ij} = 1 - a(\Phi(y_{ij}; \mu_{0j}, 1) - 0.5)$. Here $\Phi(\cdot; \mu, s^2)$ is the cumulative normal distribution function with mean μ and variance s^2 , p_j is the overall assessment probability at time j and a is a tuning parameter controlling the impact of outcome on followup. For example, when $a=0$, we have $q_{ij} = p_j$ for $i=1, \dots, n$. Model (5) clearly shows the dependence of q_{ij} on y_{ij} . When $a>0$, larger values of y_{ij} lead to smaller values of v_{ij} or a lower followup rate q_{ij} . For the control group, y_{ij} centers around μ_{0j} , which means that v_{ij} fluctuates around 1. On the other hand, with $\beta_2=0.3$ and $\beta_4=0.1$, the treatment group on average has outcomes greater than μ_{0j} or on average $v_{ij}<1$ for a positive value of a . As a result, the treatment group tends to have a lower followup rate. We consider four possible values of the tuning parameter $a=(0.5, 0.75, 1, 1.25)$, and the same four vectors of overall followup probability specified in (3).

We generate the simulated data by imposing non-ignorable missingness, and evaluate the performance of the sample size formula in terms of empirical powers when the MCAR assumption is violated. We compute the empirical powers as the proportion of samples rejecting $H_0: \beta_4=0$ among 5000 simulated samples. Table IV gives the sample size estimates and empirical powers under monotone missing data patterns. Larger values of a lead to greater loss in power. The simulation results under independent missing data patterns are similar to those under monotone missing data patterns. The table is not included here for independent missing data patterns. In both monotone and independent missing data patterns, we did not observe obvious trend in the effects of θ, ρ or assessment probability. Table IV gives the impact of non-ignorable missingness. Specifically, a great deal of empirical powers are below the nominal value 0.8, which is especially true for $a=1.25$. Note that the missing mechanism assumed in (5) leads to higher dropout in subjects with above-average measurements. That is, the large measurement value yields high dropout. Such a mechanism causes an underestimation of β_{40} and a smaller empirical power than the nominal level. Figure 1 plots the summary of assessment probabilities under the non-ignorable missingness with independent missing pattern, averaged over ρ and θ . The baseline assessment probability is $P_1=(1, 0.82, 0.79, 0.76, 0.73, 0.7)$. Under the assumed missing mechanism (5), the treatment group suffers higher dropout than the control group. Thus, each value of a corresponds to two curves, the upper curve for the control group and the lower curve for the treatment group. This group disparity grows as the value of a increases.

4. DISCUSSION

The GEE method has been widely used for the analysis of longitudinal clinical trials data since it does not depend on the restrictive symmetry assumption and does not require complete data for all subjects. The GEE sample size formula is based on the missing completely at random (MCAR) assumption and a large sample approximation. We evaluate the performance of the sample size estimate with the damped exponential family of correlation structures, small sample sizes and non-ignorable missing mechanisms. AR(1) and CS have been the formal correlation structures most often used by researchers for the sample size (power) estimation of longitudinal clinical trials due to their simplicity and distinctness of correlation pattern. We investigate the effect of correlation structures on sample size estimates using the damped exponential correlation structures, in which AR(1) and CS are special cases. Simulation results suggest that the sample size formula given in equation (2) yields empirical powers close to nominal levels for various correlation structures. When the information concerning the appropriate correlation pattern is absent, a conservative approach for testing the rates of changes between treatment groups is to adopt the AR(1) model instead of the CS model. Sample size formula also yields empirical powers close to the nominal level when the estimated total sample size ranges from 23 to 75, that is, when the estimated sample size in each group ranges from 12 to 38. The sample size formula for repeated outcomes is based on the assumption of MCAR. A simulation study is conducted to investigate the performance of the sample size formula under non-ignorable missingness. The empirical powers are also close to nominal levels under nonignorable missing mechanism even though the larger values of the tuning parameter a reduce the power of the study. However, more extensive simulation study is needed to evaluate the performance of GEE sample size formula with other non-ignorable missing mechanisms. The GEE sample size formula generally yields empirical powers close to the nominal levels under damped exponential correlation structures and small sample sizes. The computer program to estimate the sample size can be obtained from the author.

Acknowledgments

This work was supported in part by NIH grant UL1 RR024892. We thank two anonymous reviewers for their constructive comments and helpful suggestions.

References

1. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13–22.
2. Liu G, Liang KY. Sample size calculation for studies with correlated observations. *Biometrics* 1997;53:937–947. [PubMed: 9290224]
3. Rochon J. Application of GEE procedures for sample size calculations in repeated measures experiments. *Statistics in Medicine* 1998;17:1643–1658. [PubMed: 9699236]
4. Jung SH, Ahn C. Sample size estimation for comparing slopes in repeated measurements data. *Statistics in Medicine* 2003;22:1305–1315. [PubMed: 12687656]
5. Rubin DB. Inference and missing data. *Biometrika* 1976;63:81–92.
6. Munoz A, Carey V, Schouten J, Segal M, Rosner B. A parametric family of correlation structures for the analysis of longitudinal data. *Biometrics* 1992;48:733–742. [PubMed: 1420837]
7. Little, RJA.; Rubin, DB. *Statistical analysis with missing data*. Wiley; New York: 1987.

Figure 1.

Summary of assessment probability under non-ignorable missingness with independent missing pattern, averaged over ρ and θ . The baseline assessment probability is $P_1=(1, 0.82, 0.79, 0.76, 0.73, 0.7)$. Each value of a corresponds to two curves, the upper curve for the control group and the lower curve for the treatment group. The disparity between groups grows as the value of a increases.

Table 1

Sample size n (empirical power) based on P_i ($i=1, \dots, 4$) as defined in (3) with effect size $\beta_{40} = 0.1$, type I error = 0.05, nominal power = 0.8, damped correlation $\rho_{i,j+s} = \rho^c$, where $c = s^\theta$, ρ is the correlation between observations separated by one unit and θ is a damping parameter, independent or monotone missing pattern with an equal allocation of patients to the treatment and control groups, and variance of measurement error $\sigma^2 = 1$.

ρ	P	$\theta = 0$	CS	$\theta = 0.25$	$\theta = 0.5$	$\theta = 0.75$	$\theta = 1$	AR(1)
(a) Independent missing								
0.1	P_1	201	(0.804)	224 (0.802)	233 (0.805)	236 (0.807)	237 (0.809)	
	P_2	197	(0.805)	219 (0.806)	229 (0.806)	232 (0.803)	233 (0.802)	
	P_3	193	(0.791)	216 (0.807)	226 (0.797)	229 (0.806)	230 (0.812)	
	P_4	162	(0.815)	184 (0.811)	195 (0.809)	197 (0.811)	198 (0.805)	
0.25	P_1	174	(0.801)	214 (0.803)	242 (0.802)	256 (0.805)	261 (0.813)	
	P_2	169	(0.805)	210 (0.804)	238 (0.808)	253 (0.801)	259 (0.803)	
	P_3	165	(0.807)	206 (0.801)	235 (0.810)	250 (0.802)	256 (0.807)	
	P_4	135	(0.801)	175 (0.806)	203 (0.804)	218 (0.799)	224 (0.797)	
0.5	P_1	129	(0.803)	174 (0.807)	220 (0.801)	259 (0.801)	285 (0.796)	
	P_2	124	(0.801)	170 (0.814)	216 (0.809)	256 (0.808)	283 (0.806)	
	P_3	120	(0.806)	167 (0.806)	214 (0.789)	254 (0.807)	282 (0.794)	
	P_4	90	(0.805)	136 (0.797)	182 (0.800)	222 (0.807)	249 (0.811)	
(b) Monotone missing								
0.1	P_1	202	(0.801)	226 (0.786)	237 (0.811)	240 (0.809)	240 (0.797)	
	P_2	199	(0.807)	222 (0.798)	232 (0.805)	235 (0.793)	236 (0.801)	
	P_3	195	(0.802)	218 (0.806)	228 (0.798)	231 (0.795)	232 (0.796)	
	P_4	162	(0.799)	184 (0.804)	195 (0.806)	197 (0.808)	198 (0.799)	
0.25	P_1	177	(0.810)	219 (0.805)	248 (0.799)	264 (0.800)	270 (0.793)	
	P_2	175	(0.804)	216 (0.788)	245 (0.803)	260 (0.810)	266 (0.795)	
	P_3	172	(0.803)	212 (0.794)	241 (0.801)	256 (0.798)	262 (0.800)	
	P_4	135	(0.808)	175 (0.814)	203 (0.806)	218 (0.806)	224 (0.808)	
0.5	P_1	134	(0.812)	182 (0.804)	230 (0.814)	272 (0.796)	300 (0.818)	
	P_2	135	(0.811)	182 (0.794)	229 (0.787)	269 (0.798)	297 (0.799)	
	P_3	132	(0.809)	179 (0.801)	225 (0.798)	265 (0.804)	293 (0.804)	

ρ	P	$\theta=0$	CS	$\theta=0.25$	$\theta=0.5$	$\theta=0.75$	$\theta=1$	AR(1)
	P_4	90	(0.814)	136 (0.808)	182 (0.809)	222 (0.803)	249 (0.802)	

Table II

Sample size n (empirical power) based on P_i^* ($i=1, \dots, 4$) as defined in (4) with effect size $\beta_{40} = 0.1$, type I error=0.05, nominal power=0.8, damped correlation $\rho_{i,j+s} = \rho^c$, where $c = s^\theta$, ρ is the correlation between observations separated by one unit and θ is a damping parameter, independent or monotone missing pattern with an equal allocation of patients to the treatment and control groups, and variance of measurement error $\sigma^2=1$.

ρ	P	$\theta=0$	CS	$\theta=0.25$	$\theta=0.5$	$\theta=0.75$	$\theta=1$	AR(1)			
(a) Independent missing											
0.1	P_1	245	(0.800)	267	(0.801)	277	(0.788)	280	(0.812)	280	(0.804)
	P_2	234	(0.801)	257	(0.798)	267	(0.803)	270	(0.797)	271	(0.792)
	P_3	125	(0.799)	248	(0.805)	259	(0.790)	262	(0.805)	263	(0.801)
	P_4	162	(0.796)	184	(0.796)	195	(0.803)	197	(0.795)	198	(0.804)
0.25	P_1	216	(0.804)	256	(0.802)	284	(0.800)	298	(0.805)	304	(0.802)
	P_2	206	(0.805)	247	(0.810)	276	(0.801)	291	(0.800)	297	(0.808)
	P_3	197	(0.810)	238	(0.799)	268	(0.800)	284	(0.808)	290	(0.799)
	P_4	135	(0.805)	175	(0.809)	203	(0.808)	218	(0.814)	224	(0.806)
0.5	P_1	169	(0.800)	215	(0.803)	260	(0.807)	299	(0.796)	326	(0.804)
	P_2	159	(0.801)	206	(0.806)	253	(0.797)	293	(0.800)	321	(0.809)
	P_3	150	(0.807)	197	(0.802)	245	(0.812)	286	(0.807)	315	(0.796)
	P_4	90	(0.812)	136	(0.708)	182	(0.805)	222	(0.790)	249	(0.805)
(b) Monotone missing											
0.1	P_1	249	(0.788)	273	(0.789)	284	(0.802)	287	(0.807)	287	(0.801)
	P_2	240	(0.802)	263	(0.803)	273	(0.806)	276	(0.806)	277	(0.797)
	P_3	230	(0.800)	253	(0.798)	263	(0.807)	266	(0.798)	267	(0.796)
	P_4	162	(0.804)	184	(0.803)	195	(0.810)	197	(0.803)	198	(0.799)
0.25	P_1	227	(0.799)	270	(0.801)	300	(0.798)	316	(0.805)	321	(0.806)
	P_2	220	(0.802)	261	(0.805)	290	(0.808)	306	(0.803)	312	(0.799)
	P_3	209	(0.807)	250	(0.808)	279	(0.800)	295	(0.806)	301	(0.799)
	P_4	135	(0.803)	175	(0.803)	203	(0.801)	218	(0.796)	224	(0.802)
0.5	P_1	191	(0.802)	240	(0.801)	289	(0.809)	331	(0.810)	360	(0.800)
	P_2	187	(0.800)	234	(0.802)	282	(0.803)	322	(0.810)	351	(0.798)
	P_3	174	(0.799)	221	(0.812)	268	(0.798)	309	(0.801)	337	(0.802)

ρ	P	$\theta=0$	CS	$\theta=0.25$	$\theta=0.5$	$\theta=0.75$	$\theta=1$	AR(1)
	P_4	90	(0.808)	136 (0.804)	182 (0.804)	222 (0.803)	249 (0.808)	

Table III

Sample size n (empirical power) based on P_i ($i=1, \dots, 4$) as defined in (3) with effect size $\beta_{40} = 0.2$, type I error = 0.05, nominal power = 0.8, damped correlation $\rho_{i,j+s} = \rho^c$, where $c = s^\theta$, ρ is the correlation between observations separated by one unit and θ is a damping parameter, independent or monotone missing pattern with an equal allocation of patients to the treatment and control groups, and variance of measurement error $\sigma^2 = 1$.

ρ	P	$\theta = 0$	$\theta = 0.25$	$\theta = 0.5$	$\theta = 0.75$	$\theta = 1$
(a) Independent missing						
0.1	P_1	51 (0.821)	56 (0.801)	59 (0.817)	59 (0.809)	60 (0.805)
	P_2	50 (0.820)	55 (0.818)	58 (0.818)	58 (0.799)	59 (0.814)
	P_3	49 (0.811)	54 (0.818)	57 (0.800)	58 (0.813)	58 (0.814)
	P_4	41 (0.816)	46 (0.805)	49 (0.813)	50 (0.816)	50 (0.811)
0.25	P_1	44 (0.807)	54 (0.810)	61 (0.818)	64 (0.807)	66 (0.824)
	P_2	43 (0.826)	53 (0.818)	60 (0.814)	64 (0.809)	65 (0.810)
	P_3	42 (0.825)	52 (0.811)	59 (0.813)	63 (0.809)	64 (0.810)
	P_4	34 (0.827)	44 (0.808)	51 (0.808)	55 (0.811)	56 (0.810)
0.5	P_1	33 (0.824)	44 (0.813)	55 (0.804)	65 (0.809)	72 (0.811)
	P_2	31 (0.823)	43 (0.816)	54 (0.820)	64 (0.801)	71 (0.807)
	P_3	30 (0.810)	42 (0.807)	54 (0.815)	64 (0.815)	71 (0.819)
	P_4	23 (0.834)	34 (0.807)	46 (0.808)	56 (0.816)	63 (0.814)
(b) Monotone missing						
0.1	P_1	51 (0.806)	57 (0.813)	60 (0.822)	60 (0.804)	60 (0.816)
	P_2	50 (0.812)	56 (0.804)	58 (0.806)	59 (0.813)	59 (0.810)
	P_3	49 (0.807)	55 (0.804)	57 (0.801)	58 (0.806)	58 (0.808)
	P_4	41 (0.816)	46 (0.807)	49 (0.807)	50 (0.815)	50 (0.816)
0.25	P_1	45 (0.811)	55 (0.816)	62 (0.810)	66 (0.811)	68 (0.820)
	P_2	44 (0.817)	54 (0.799)	62 (0.811)	65 (0.805)	67 (0.807)
	P_3	43 (0.819)	53 (0.810)	61 (0.815)	64 (0.810)	66 (0.800)
	P_4	34 (0.819)	44 (0.812)	51 (0.810)	55 (0.808)	56 (0.814)
0.5	P_1	34 (0.817)	46 (0.818)	58 (0.813)	68 (0.807)	75 (0.811)
	P_2	34 (0.812)	46 (0.812)	58 (0.815)	68 (0.815)	75 (0.801)
	P_3	33 (0.805)	45 (0.811)	57 (0.811)	67 (0.812)	74 (0.819)

ρ	P	$\theta = 0$	CS	$\theta = 0.25$	$\theta = 0.5$	$\theta = 0.75$	$\theta = 1$	AR(1)
	P_4	23	(0.808)	34 (0.816)	46 (0.813)	56 (0.810)	63 (0.811)	

Table IV

Sample size n (empirical power) based on P_i ($i=1, \dots, 4$) as defined in (3), where sample size n is for effect size $\beta_{40} = 0.1$, type I error = 0.05, nominal power = 0.8, damped correlation $\rho_{j_i+s} = \rho^c$, where $c = s^\theta$, ρ is the correlation between observations separated by one unit and θ is a damping parameter, non-ignorable missing with $a=0.50, 0.75, 1.00$ and 1.25 under monotone missing data pattern with an equal allocation of patients to the treatment and control groups, and variance of measurement error $\sigma^2=1$.

ρ	P	a	$\theta=0$	CS	$\theta=0.25$	$\theta=0.5$	$\theta=0.75$	$\theta=1$	AR(1)
0.1	P_1	0.50	202	(0.798)	226 (0.802)	237 (0.801)	240 (0.803)	240 (0.805)	
		0.75	—	(0.806)	— (0.799)	— (0.795)	— (0.797)	— (0.801)	
		1.00	—	(0.792)	— (0.805)	— (0.806)	— (0.806)	— (0.804)	
	P_2	1.25	—	(0.802)	— (0.796)	— (0.795)	— (0.798)	— (0.799)	
		0.50	199 (0.803)	222 (0.799)	232 (0.804)	235 (0.799)	236 (0.796)		
		0.75	—	(0.806)	— (0.791)	— (0.805)	— (0.795)	— (0.790)	
0.25	P_3	1.00	—	(0.800)	— (0.807)	— (0.794)	— (0.794)	— (0.792)	
		1.25	—	(0.799)	— (0.794)	— (0.792)	— (0.786)	— (0.792)	
		0.50	195 (0.798)	218 (0.802)	228 (0.785)	231 (0.792)	232 (0.795)		
	P_4	0.75	—	(0.794)	— (0.787)	— (0.783)	— (0.793)	— (0.805)	
		1.00	—	(0.789)	— (0.784)	— (0.790)	— (0.794)	— (0.783)	
		1.25	—	(0.783)	— (0.778)	— (0.779)	— (0.793)	— (0.781)	
0.50	P_1	—	162 (0.805)	184 (0.813)	195 (0.812)	197 (0.806)	198 (0.814)		
		0.50	177 (0.804)	219 (0.803)	248 (0.804)	264 (0.783)	270 (0.803)		
		0.75	—	(0.795)	— (0.800)	— (0.812)	— (0.794)	— (0.809)	
	P_2	1.00	—	(0.798)	— (0.804)	— (0.789)	— (0.800)	— (0.806)	
		1.25	—	(0.793)	— (0.787)	— (0.800)	— (0.791)	— (0.799)	
		0.50	175 (0.803)	216 (0.801)	245 (0.802)	260 (0.803)	266 (0.804)		
1.00	P_3	0.75	—	(0.803)	— (0.797)	— (0.796)	— (0.801)	— (0.806)	
		1.00	—	(0.781)	— (0.797)	— (0.789)	— (0.791)	— (0.798)	
		1.25	—	(0.781)	— (0.792)	— (0.799)	— (0.790)	— (0.787)	
	P_4	0.50	172 (0.805)	212 (0.803)	241 (0.793)	256 (0.812)	262 (0.790)		
		0.75	—	(0.791)	— (0.790)	— (0.790)	— (0.795)	— (0.802)	
		1.00	—	(0.792)	— (0.787)	— (0.776)	— (0.791)	— (0.790)	
0.75	P_4	1.25	—	(0.785)	— (0.786)	— (0.793)	— (0.782)	— (0.798)	
		—	135 (0.795)	175 (0.803)	203 (0.807)	218 (0.810)	224 (0.806)		
		—	—	—	—	—	—		

ρ	P	a	$\theta=0$ CS	$\theta=0.25$	$\theta=0.5$	$\theta=0.75$	$\theta=1$ AR(1)
0.5	P_1	0.50	134 (0.810)	182 (0.793)	230 (0.801)	272 (0.801)	300 (0.813)
		0.75	— (0.797)	— (0.804)	— (0.804)	— (0.797)	— (0.796)
		1.00	— (0.794)	— (0.812)	— (0.782)	— (0.778)	— (0.801)
		1.25	— (0.787)	— (0.783)	— (0.783)	— (0.795)	— (0.787)
	P_2	0.50	135 (0.793)	182 (0.790)	229 (0.803)	269 (0.800)	297 (0.799)
		0.75	— (0.786)	— (0.793)	— (0.792)	— (0.799)	— (0.797)
		1.00	— (0.780)	— (0.788)	— (0.793)	— (0.795)	— (0.790)
		1.25	— (0.779)	— (0.790)	— (0.771)	— (0.799)	— (0.796)
	P_3	0.50	132 (0.790)	179 (0.791)	225 (0.797)	265 (0.792)	293 (0.794)
		0.75	— (0.796)	— (0.784)	— (0.773)	— (0.791)	— (0.787)
		1.00	— (0.786)	— (0.788)	— (0.791)	— (0.783)	— (0.798)
		1.25	— (0.765)	— (0.780)	— (0.774)	— (0.780)	— (0.775)
	P_4	—	90 (0.813)	136 (0.806)	182 (0.808)	222 (0.803)	249 (0.802)