Published in final edited form as:

Contemp Clin Trials. 2011 May ; 32(3): 412–417. doi:10.1016/j.cct.2011.01.002.

How Many Measurements for Time-Averaged Differences in Repeated Measurement Studies?

Song Zhang and

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

Chul Ahn

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

SUMMARY

In many studies, investigators have perceived the number of repeated measurements as a fixed design characteristic. However, the number of repeated measurements is a design choice that can be informed by statistical considerations. In this paper, we investigate how the number of repeated measurements affects the required sample size in longitudinal studies with scheduled assessment times and a fixed total duration. It is shown that the required sample size always decreases as the number of measurements per subject increases under the compound symmetry (CS) correlation. The magnitude of sample size reduction, however, quickly shrinks to less than 5% when the number of measurements per subject increases beyond 4. We then reveal a counterintuitive property of the AR(1) correlation structure, under which making additional measurements from each subject might increase the sample size requirement. This observation suggests that practitioners should be cautious about assuming the AR(1) model in repeated measurements studies, whether in experimental design or in data analysis. Finally, we show that by introducing measurement error into the AR(1) model, the counterintuitive behavior disappears. That is, additional measurements per subject result in reduced sample sizes.

Keywords

time-averaged response; compound symmetry; AR(1); sample size; measurement error

1 Introduction

In controlled clinical trials, subjects are randomly assigned to one of the treatment groups, evaluated at baseline, and then re-evaluated at intervals across a treatment period. In most trials, the length of the treatment period and the number of repeated measurements are pre-determined. In the design of repeated measurements studies, investigators are often confronted with the challenge of determining the total number of subjects and the number of repeated measurements per subject across a given treatment period.

In many comparative trials, investigators adopt the repeated measurements design in the hope that by obtaining extra measurements from each subject, the required sample size can

Correspondence should be sent to: Chul Ahn, Ph.D., Department of Clinical Sciences, UT Southwestern Medical Center, 5323 Harry Hines Blvd, E5.506, Dallas, TX 75390, Chul.Ahn@UTSouthwestern.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

be reduced. Nonetheless, the number of repeated measurements per subject has received little attention since this may be perceived as a pre-specified characteristic of a particular trial design (Vickers, 2003). Vickers (2003) illustrated that, using ANCOVA (the analysis of covariance), negligible power improvement is achieved for time-averaged differences assuming the CS correlation among repeated measurements when the number of post-treatment assessments is increased beyond four. There has been little guidance in the methodologic literature about determining the optimal number of repeated measurements.

Two approaches are widely used to estimate the sample size for repeated measurements studies. One approach is based on comparing the time-averaged difference in the outcome variable between the treatment groups. The other approach is based on comparing the rates of change in the outcome variable over the study period. Diggle *et al.* (2002) provided the sample size formulas to compare the time-averaged responses and the rates of change in studies with continuous outcomes, assuming the compound symmetry (CS) correlation among observations from the same subject and no missing data.

In this paper, we restrict our attention to studies where researchers make inference based on the difference in time-averaged responses between treatment groups over a treatment period of a fixed duration. This type of analysis is frequently used when the outcome to be measured varies over time. For example, blood pressure levels vary depending on the amount of food taken, sleep, and exercise, etc. Aronow and Ahn (1994) showed that the mean maximal decrease in postprandial systolic blood pressures was 15 ± 6 mm Hg in 499 elderly persons in a long-term health care facility. The mean maximal decrease in postprandial systolic blood pressure occurred 15 minutes after eating in 13% of residents, 30 minutes after eating in 20% of residents, 45 minutes after eating in 26% of residents, 60 minutes after eating in 30% of residents, and 75 minutes after eating in 11% of residents. If researchers only take one measurement from each subject and compare the mean blood pressure levels between two treatment groups, the experiment may have a poor performance due to substantial within-subject variation. We will investigate, to make inference based on time-averaged differences, how the decision on the number of repeated measurements affects the required sample size under different correlation structures.

The rest of the paper is organized as follows. In Section 2, we explore the relationship between sample size and the number of repeated measurements under the CS and the AR(1) correlation structures, with and without measurement error. Real application examples are presented in Section 3. The final section is devoted to discussion.

2 Statistical Method

Suppose in a clinical trial we use Y_{ij} to denote the continuous response measurement obtained at time t_j ($j = 1, \dots, m$) from subject i ($i = 1, \dots, 2n$). Thus all subjects follow the same measurement schedule, m is the number of repeated measurements per subject, and n is the number of subjects per group. To make inference about the difference in time-averaged responses between the two groups, we assume the following statistical model (Diggle *et al.*, 2002),

$$Y_{ij} = \beta_0 + \beta_1 x_i + \varepsilon_{ij}, \quad (1)$$

where x_i is the indicator of the group assignment for subject i , parameter β_0 models the intercept effect, parameter β_1 models the difference in time-averaged responses between the two groups, and ε_{ij} denotes the residual error. The inference about β_1 is of our primary interest. It is usually assumed that $E(\varepsilon_{ij}) = 0$ and $Var(\varepsilon_{ij}) = \sigma^2$. Different correlation structures

can be assumed for the error terms within the same subject. For example, the AR(1) model assumes an exponentially decaying pattern of correlation according to the (spatial or temporal) distance between the repeated measurements while the CS model assumes a constant correlation between two distinct measurements regardless of the distance. It is also assumed that the measurements are independent among different subjects.

Under the CS correlation structure, it is assumed that the correlation between ε_{ij} and ε_{ik} is ρ for $j \neq k$. Then the variance of the individual time-averaged response from subject i is

$$\text{Var}\left(\sum_{j=1}^m Y_{ij}/m\right) = \frac{\sigma^2}{m} [1 + (m-1)\rho]. \quad (2)$$

To achieve a testing power of $1 - \gamma$ with a significance level of α , the required sample size per group can be estimated by the following formula (Diggle *et al.*, 2002),

$$n = 2(z_{1-\alpha/2} + z_{1-\gamma})^2 \sigma^2 [1 + (m-1)\rho] / (md^2), \quad (3)$$

where $z_{1-\alpha/2}$ is the $100(1 - \alpha/2)$ th percentile of the standard normal distribution, and d is the clinically meaningful difference to be detected. Equation (3) is an extension to the sample size formula for a single measurement study, with additional terms including the number of repeated measurements (m) per subject and the variance inflation factor $[1 + (m-1)\rho]$, to account for the degree of clustering among observations from the same subject. It has also been used in sample size calculation for cluster randomization trials (Donner and Klar, 2000). Liu and Wu (2005) provided an extension of Equation (3) to accommodate unbalanced clinical trials.

It is obvious from (3) that, under the CS correlation, the sample size n increases with correlation ρ . Furthermore, by showing that the second derivative of n with respect to m is negative (details omitted), we can establish that collecting more measurements from each subject always lead to saving in sample size under CS. In this study we assess the impact of an additional measurement (from m to $m+1$) on sample size requirement by

$$V_{m+1,m} = \frac{n(m) - n(m+1)}{n(1)},$$

where $n(m)$ denotes the required sample size under m repeated measurements per subject. We interpret $V_{m+1,m}$ as the marginal saving in sample size with an additional measurement from each subject, normalized by the sample size under the single measurement design. By plugging Equation (3) into $V_{m+1,m}$, we see that only two factors affect the marginal sample size saving: the correlation ρ and the current number of measurements m . The other design configurations such as power $1 - \gamma$, type I error α , variance σ^2 , and target difference d , are irrelevant.

In Table 1 we list the values of $V_{m+1,m}$ under various combinations of m and ρ . For example, when the correlation between observations is 0.8, increasing the number of measurements from two to three reduces the sample size requirement by 3.3%. Table 1 indicates that obtaining additional measurements is more effective in reducing sample size for smaller

values of ρ . The maximum sample size reduction is achieved under $\rho=0$. As a general rule, the required sample size (n) decreases as the number of repeated measurements (m) increases under the CS correlation. However, the marginal saving decreases rapidly as the number of repeated measurements increases. For $\rho \geq 0$, researchers can achieve at most a 5% marginal reduction in sample size when the number of measurements is increased beyond four.

Under the AR(1) correlation structure, it is assumed that the correlation between ε_{ij} and ε_{ik} is $\rho^{|k-j|/(m-1)}$, where ρ is the correlation between the first and the last measurements and the measurement times are equidistant. Under the AR(1), the variance of the individual time-averaged response from subject i is

$$\text{Var} \left(\sum_{j=1}^m Y_{ij} / m \right) = \frac{\sigma^2}{m^2} \left\{ m+2 \left[\sum_{j=1}^{m-1} (m-j) \rho^{j/(m-1)} \right] \right\} \quad (4)$$

Then the sample size formula for the time-averaged differences is given by

$$n = 2(z_{1-\alpha/2} + z_{1-\gamma})^2 \sigma^2 \left\{ m+2 \left[\sum_{j=1}^{m-1} (m-j) \rho^{j/(m-1)} \right] \right\} / (m^2 d^2). \quad (5)$$

As a special case, the AR(1) correlation is equivalent to the CS correlation when $m=2$. For $m > 2$, Equation (5) is too complicated for us to derive the analytical relationship between sample size (n) and the number of repeated measurements (m).

In Table 2, we present the results of a numerical study to explore the impact of additional measurements under different correlation (ρ) levels. Note that the AR(1) and CS are equivalent under $m = 2$ or $\rho = 0$ or $\rho = 1$. As a result, the first and the last columns and the first row of Table 2 are the same as those of Table 1. One important observation from Table 2 is that, under the AR(1) correlation structure, making additional measurements does not always reduce sample size requirement. As a matter of fact, the additional measurements beyond $m = 2$ leads to an increased sample size for $0.4 \leq \rho \leq 0.8$. The reduction in sample size is only 2.35% when $\rho = 0.2$. This observation might be counterintuitive, but it is determined by the theoretical property of the AR(1) model. Overall et al. (1998) also found the counterintuitive observation that the increased frequency of measurements during a fixed study period tended to decrease power for comparing the rates of changes under AR(1). It also sends an important message to practitioners that researchers should be cautious about assuming an AR(1) model in repeated measurements studies, whether in clinical trial design or in data analysis. For example, if a clinician assumes an AR(1) structure with a moderate level of correlation, then any design scheme that have more than 2 measurements per subject will be considered theoretically sub-optimal. Note that when the data suffers from missingness, increasing the number of measurements might be beneficial under an AR(1) correlation (Zhang and Ahn, 2010).

Under the AR(1) correlation structure with measurement error, we assume that the randomness in the outcome measurements comes from two distinct sources. The first is associated with the random fluctuation in the attribute of the subject. It is usually assumed to be associated with a certain correlation structure such as CS or AR(1) within a subject, while independent between subjects. The second source is measurement error. It is caused by

random factors involved in the measuring procedure itself, and is thus independent across all subjects and all measurements. The statistical model can be written as

$$Y_{ij} = \beta_0 + \beta_1 x_i + \eta_{ij}, \text{ with } \eta_{ij} = \varepsilon_{ij} + \xi_{ij}.$$

Here ε_{ij} corresponds to the first source of randomness. It is defined similarly as in Model (1). The parameter ξ_{ij} represents measurement error, for which we assume independence across any combination of (i, j) and $\text{Var}(\xi_{ij}) = \sigma_E^2$.

Suppose we assume a CS correlation structure for ε_{ij} with correlation ρ . It can be shown that, with the additional measurement error term ξ_{ij} , the correlation structure of η_{ij} is again CS.

The correlation between η_{ij} and η_{ik} is $\rho/(1+R)$, with $R = \sigma_E^2/\sigma^2$ being the ratio of variances. Thus there is no need to discuss the “CS+measurement error” model separately.

It is a different story if we assume the AR(1) correlation structure for ε_{ij} . With the introduction of measurement error ξ_{ij} , the covariance of η_{ij} and η_{ik} remains the same as that of ε_{ij} and ε_{ik} . That is, $\text{Cov}(\eta_{ij}, \eta_{ik}) = \text{Cov}(\varepsilon_{ij}, \varepsilon_{ik}) = \sigma^2 \rho^{|k-j|/(m-1)}$. The variance of η_{ij} , however, includes an additional term σ_E^2 to account for the additional randomness induced by measurement error, $\text{Var}(\eta_{ij}) = \sigma^2 + \sigma_E^2$. Consequently, the correlation between η_{ij} and η_{ik} is then $\rho^{|k-j|/(m-1)/(1+R)}$. Note that the correlation no longer has the exponentially decaying pattern as in AR(1). The sample size formula for the time-averaged differences is

$$n = 2(z_{1-\alpha/2} + z_{1-\gamma})^2 \sigma^2 \left\{ m(1+R) + 2 \left[\sum_{j=1}^{m-1} (m-j) \rho^{j/(m-1)} \right] \right\} / (m^2 d^2). \quad (6)$$

Plugging (6) into the equation of $V_{m+1,m}$, it is obvious that the factors influencing the marginal saving in sample size include the current number of repeated measurements m , the correlation ρ , and the variance ratio R .

Tables 3–5 shows the reduction in sample size with additional measurements under AR(1) correlation and measurement error, with $R = 0.5, 1$, and 2 , respectively. First, comparing with Table 2, we notice that under the assumed values of R , the required sample size is always reduced with more measurements per subject. This characteristic might be intuitively appealing to practitioners. Second, as we have explained previously, the addition of measurement error modifies the correlation coefficient in the original AR(1) model by a factor of $1/(1+R)$. Thus larger values of R result in smaller correlation. The extreme case is that the measurements become independent as $R \rightarrow +\infty$. Thus more sample size reduction is associated with larger R . Figure 1 shows the sample size requirement under different combinations of (m, ρ, R) with the x -axis indicating the number of repeated measurements (m) and the y -axis indicating the correlation (ρ). For easy comparison, the z -axis shows the relative sample size, compared with the sample size under $m=1$. Figure 1 suggests that more saving in sample size can be achieved as the variance ratio, $R = \sigma_E^2/\sigma^2$ increases.

3 Example

3.1 Example 1

Suppose that an investigator plans to examine the effect of a new drug on heart rate by evaluating the time-averaged difference in heart rate between subjects taking the standard drug and those taking the new drug. The study will measure heart rate at baseline and then follow the heart rate for 120 minutes. A pilot study showed that for subjects taking the standard drug, the heart rates have an AR(1) correlation structure with an average of 80.0 ± 10.5 . The correlation between the baseline (time=0) and the endpoint (120 minute) measurement is $\rho = 0.5$. It is assumed that the heart rate measurements obtained from the proposed trial follows the same correlation structure. Heart rates will be measured at equal intervals. The investigator would like to estimate the sample size needed to detect a 7% reduction from the subjects taking the new drug with a 5% significance level and 80% power. From an average heart rate of 80, a 7% reduction is equal to a difference of 5.6. Using Equation (5), the estimated sample sizes are 56, 42, 42, 43 and 43 subjects per group for $m=1, 2, 3, 4$, and 5, respectively. The same results can be obtained from the sample size software PASS 2008, under “Inequality Tests for Two Means in a Repeated Measures Design”. For the design of $m=1, 2, 3, 4$, and 5 repeated measurements, the correlations between adjacent measurements (required by PASS 2008) are 1, 0.5, 0.7071, 0.7937, and 0.8409, respectively, which are calculated by $\rho^{1/(m-1)}$. For illustration, the following is the summary statement produced from PASS 2008. “Group sample sizes of 43 and 43 achieve 80% power to detect a difference of 5.600 in a design with 5 repeated measurements having an AR(1) covariance structure when the standard deviation is 10.500, the correlation between observations on the same subject is 0.841, and the alpha level is 0.050.”

It has been assumed that the required sample size always decreases as the number of repeated measurements increases. This example uses a commercial software to demonstrate that making additional measurements from each subject might actually increase the sample size requirement. For AR(1) models with a moderate correlation, it is statistically sub-optimal to take more than two measurements per subject.

3.2 Example 2

Investigators plan to evaluate the effect of a new drug on total cholesterol levels by investigating if the time-averaged total cholesterol levels are significantly different between the placebo and a new drug. Assessment of cholesterol level will be carried out at baseline and at months 3 and 6 ($m=3$). Previous studies have found an average total cholesterol level of 260 ± 50 mg/dL for individuals taking placebo with the AR(1) correlation of baseline-to-endpoint correlation of 0.4. It is expected that the observations from the same subject will have the AR(1) correlation. Taking measurements at baseline, 2, 4 and 6 months ($m=4$) will result in a 0.9% increase in sample size compared with that under $m=3$ measurements per subject with no measurement error. If there is a measurement error with the ratio of

$R = \sigma^2 / \sigma_F^2 = 1$, we can achieve a 3.7% reduction in sample size. If the correlation among observations follows a CS correlation structure with $\rho=0.4$, the sample size reduction will be 5.0%.

4 Discussion

In this paper we investigate how to select the number of measurements per subject in repeated measurement studies. We compare the sample size requirement under different numbers of repeated measurements using simple statistical formulas and Tables. Under the CS correlation, the required sample size always decreases as the number of repeated measurements increases. The relative reduction in sample size, however, rapidly decreases

as the number of repeated measurements increases. When the number of measurements per subject is increased beyond 4, sample size reduction is at most 5% under the CS correlation. When the repeated measurements follow the AR(1) structure with a moderate level of correlation, the design schemes of more than two measurements per subject are not recommended because the additional measurements either lead to negligible sample size reduction or slight increase in sample size. This paper shows that the required sample size may increase under heterogeneous correlation structure such as AR(1) as the number of repeated measurements increases. When measurement error is introduced into the AR(1) model, the required sample size decreases as the number of repeated measurements increases for variance ratio of $R=0.5$, 1 and 2. More simulation studies are needed to investigate sample size reduction under other values of R .

We identify the optimal number of repeated measurements based on a “statistical” criterion of estimating the sample sizes. We suggest that the preferred design is to make $m = 2$ measurements per subject under the AR(1) correlation. In many longitudinal treatment studies, however, researchers follow the individuals at intervals not only to estimate the treatment effect, but also to monitor safety and other endpoints over time. The statistically preferred design with only $m = 2$ observations would prevent the investigators from addressing safety and other endpoints. This is a problem with the statistically preferred design in general, in that it focuses on a narrow statistical objective without considering other aims. That is, when a longitudinal study is designed, the number of repeated measurements needs to be determined not only by the statistically preferred design, but also by regulatory implications. If investigators blindly adopt the statistically preferred design, investigators might be prevented from addressing these important questions. One solution is to specify the minimal number of repeated measurements that is required by clinical goals, and then apply the statistical method to identify the optimal number of repeated measurements. Alternatively, it might be interesting to discuss with regulatory agencies about the possibility that researchers obtain more measurements to satisfy regulatory requirements, but indicate *a priori* that only the statistically optimal number of measurements will be included in statistical analysis.

Investigators are also challenged by budget constraints in designing a repeated measurement study, where they need to decide how to make a full use of limited financial resources. Zhang and Ahn (in press) derived the optimal combination of n and m that yields the smallest variance of the treatment effect under budget constraints by comparing the rate of changes over the study period. We will continue to work on how the number of repeated measurements affects the required sample size by taking account of missing data and various correlation structures under budget constraints.

The marginal reduction in sample size offers a rational basis to determine the number of measurements per subject in repeated measurement studies with continuous outcomes. We will investigate the appropriate number of repeated measurements for the rates of change under different correlation structures with or without measurement error.

Acknowledgments

This work was supported in part by NIH grants UL1 RR024982, P30CA142543, P50CA70907, and DK081872.

References

- Aronow W, Ahn C. Postprandial hypotension in 499 elderly persons in a long-term health care facility. *Journal of the American Geriatric Society*. 1994; 42:930–932.
- Davis CS. Semi-parametric and non-parametric methods for the analysis of repeated measurements with applications to clinical trials. *Statistics in Medicine*. 1991; 10:1959–1980. [PubMed: 1805321]

- Diggle, PJ.; Heagerty, P.; Liang, KY.; Zeger, SL. Analysis of longitudinal data. 2. Oxford University Press; 2002.
- Donner, A.; Klar, N. Design and analysis of cluster randomization trials in health research. Oxford University Press; 2000.
- Liu H, Wu T. Sample size calculation and power analysis of time-averaged difference. Journal of Modern Applied Statistical Methods. 2005; 2:1–9.
- Overall JE, Shobaki G, Anderson CB. Comparative evaluation of two models for estimating sample sizes for tests on trends across repeated measurements. Controlled Clinical Trials. 1998; 19:188–197. [PubMed: 9551283]
- Vickers AJ. How many repeated measures in repeated measures designs? Statistical issues for comparative trials. BMC Medical Research Methodology. 2003; 1:1–9.
- Zhang S, Ahn C. Effects of Correlation and Missing Data on Sample Size Estimation in Longitudinal Clinical Trials. Pharmaceutical Statistics. 2010; 9:2–9. [PubMed: 19053162]
- Zhang S, Ahn C. Adding subjects or adding measurements in repeated measurements under financial constraints, *Statistics in Biopharmaceutical Research*. (in press).

Figure 1.

The sample size requirement under different combinations of (m, ρ, R) . The x -axis indicates the number of repeated measurements (m). The y -axis indicates the correlation (ρ). The z -axis shows the sample size relative to the one under $m=1$.

Table 1

Relative Reduction in Sample Size for Time-Averaged Responses by Increasing the Number of Visits from m to $m + 1$ under the CS Correlation.

		ρ									
		$m \rightarrow m + 1$	0	0.2	0.4	0.5	0.6	0.8	1		
1	$\rightarrow 2$	50.0%	40.0%	30.0%	25.0%	20.0%	10.0%	0.0%	0.0%		
2	$\rightarrow 3$	16.67%	13.33%	10.00%	8.33%	6.67%	3.33%	0.0%	0.0%		
3	$\rightarrow 4$	8.33%	6.67%	5.00%	4.17%	3.33%	1.67%	0.0%	0.0%		
4	$\rightarrow 5$	5.00%	4.00%	3.00%	2.50%	2.00%	1.00%	0.0%	0.0%		
5	$\rightarrow 6$	3.33%	2.67%	2.00%	1.67%	1.33%	0.67%	0.0%	0.0%		
6	$\rightarrow 7$	2.38%	1.90%	1.43%	1.19%	0.95%	0.48%	0.0%	0.0%		
7	$\rightarrow 8$	1.79%	1.43%	1.07%	0.89%	0.71%	0.36%	0.0%	0.0%		
8	$\rightarrow 9$	1.39%	1.11%	0.83%	0.69%	0.56%	0.28%	0.0%	0.0%		
9	$\rightarrow 10$	1.11%	0.89%	0.67%	0.56%	0.44%	0.22%	0.0%	0.0%		

Table 2

Relative Reduction in Sample Size for Time-Averaged Responses by Increasing the Number of Visits from m to $m + 1$ under the AR(1) Correlation.

		ρ						
$m \rightarrow m + 1$		0	0.2	0.4	0.5	0.6	0.8	1
1 $\rightarrow$ 2	50.00%	40.0%	30.0%	25.0%	20.0%	10.0%	0.0%	
2 $\rightarrow$ 3	16.67%	2.35%	-0.33%	-0.87%	-1.09%	-0.86%	0.0%	
3 $\rightarrow$ 4	8.33%	-0.33%	-0.87%	-0.89%	-0.82%	-0.49%	0.0%	
4 $\rightarrow$ 5	5.00%	-0.54%	-0.67%	-0.63%	-0.55%	-0.31%	0.0%	
5 $\rightarrow$ 6	3.33%	-0.48%	-0.50%	-0.45%	-0.38%	-0.21%	0.0%	
6 $\rightarrow$ 7	2.38%	-0.39%	-0.38%	-0.34%	-0.28%	-0.15%	0.0%	
7 $\rightarrow$ 8	1.79%	-0.32%	-0.30%	-0.26%	-0.22%	-0.11%	0.0%	
8 $\rightarrow$ 9	1.39%	-0.26%	-0.24%	-0.21%	-0.17%	-0.09%	0.0%	
9 $\rightarrow$ 10	1.11%	-0.22%	-0.19%	-0.17%	-0.14%	-0.07%	0.0%	

Table 3

Relative Reduction in Sample Size for Time-Averaged Responses by Increasing the Number of Visits from m to $m + 1$ under the AR(1) correlation in the Presence of Measurement Error with $R=\sigma^2_E/\sigma^2=0.5$.

		ρ						
$m \rightarrow$	$m + 1$	0	0.2	0.4	0.5	0.6	0.8	
1 $\rightarrow$ 2		50.0%	43.3%	36.7%	33.3%	30.0%	23.3%	
2 $\rightarrow$ 3		16.67%	7.12%	5.33%	4.97%	4.83%	4.98%	
3 $\rightarrow$ 4		8.33%	2.56%	2.20%	2.18%	2.23%	2.45%	
4 $\rightarrow$ 5		5.00%	1.31%	1.22%	1.24%	1.30%	1.46%	
5 $\rightarrow$ 6		3.33%	0.79%	0.78%	0.81%	0.85%	0.97%	
6 $\rightarrow$ 7		2.38%	0.53%	0.54%	0.57%	0.61%	0.69%	
7 $\rightarrow$ 8		1.79%	0.38%	0.40%	0.42%	0.45%	0.52%	
8 $\rightarrow$ 9		1.39%	0.29%	0.31%	0.33%	0.35%	0.40%	
9 $\rightarrow$ 10		1.11%	0.22%	0.24%	0.26%	0.29%	0.32%	

Note that when $\rho = 1$, the correlation structure is equivalent to CS with $\rho = 0.5$.

Table 4

Relative Reduction in Sample Size for Time-Averaged Responses by Increasing the Number of Visits from m to $m + 1$ under the AR(1) correlation in the Presence of Measurement Error with $R=\sigma^2_E/\sigma^2=1$.

		ρ							
$m \rightarrow m + 1$		0	0.2	0.4	0.5	0.6	0.8		
1 $\rightarrow$ 2	50.0%	45.0%	40.0%	37.5%	35.0%	30.0%			
2 $\rightarrow$ 3	16.67%	9.51%	8.17%	7.90%	7.79%	7.90%			
3 $\rightarrow$ 4	8.33%	4.00%	3.73%	3.72%	3.76%	3.92%			
4 $\rightarrow$ 5	5.00%	2.23%	2.16%	2.18%	2.23%	2.35%			
5 $\rightarrow$ 6	3.33%	1.43%	1.42%	1.44%	1.47%	1.56%			
6 $\rightarrow$ 7	2.38%	0.99%	1.00%	1.02%	1.05%	1.11%			
7 $\rightarrow$ 8	1.79%	0.73%	0.75%	0.76%	0.78%	0.84%			
8 $\rightarrow$ 9	1.39%	0.56%	0.58%	0.59%	0.61%	0.65%			
9 $\rightarrow$ 10	1.11%	0.45%	0.46%	0.47%	0.49%	0.52%			

Table 5

Relative Reduction of Sample Size for Time-Averaged Responses by Increasing the Number of Visits from m to $m + 1$ under the AR(1) correlation in the Presence of Measurement Error with $R = \sigma^2_E / \sigma^2 = 2$.

		ρ							
		0	0.2	0.4	0.5	0.6	0.8		
$m \rightarrow$	$m + 1$								
1	2	50.0%	46.7%	43.3%	41.67%	40.0%	36.7%		
2	3	16.67%	11.89%	11.00%	10.82%	10.75%	10.82%		
3	4	8.33%	5.45%	5.27%	5.26%	5.28%	5.39%		
4	5	5.00%	3.15%	3.11%	3.12%	3.15%	3.23%		
5	6	3.33%	2.06%	2.06%	2.07%	2.09%	2.15%		
6	7	2.38%	1.46%	1.46%	1.47%	1.49%	1.54%		
7	8	1.79%	1.08%	1.09%	1.10%	1.12%	1.15%		
8	9	1.39%	0.84%	0.85%	0.86%	0.87%	0.90%		
9	10	1.11%	0.67%	0.68%	0.69%	0.69%	0.72%		