

Research Article

Conservative Sample Size Determination for Repeated Measures Analysis of Covariance

Timothy M. Morgan* and L. Douglas Case

Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA

Corresponding author

Timothy M. Morgan, Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC 27157, Tel: 336-716-1009; Fax: 336-716-6427; E-mail: tomorgan@wfubmc.edu

Submitted: 19 June 2013

Accepted: 02 July 2013

Published: 05 July 2013

Copyright

© 2013 Morgan and Case

OPEN ACCESS

Keywords

- Sample size
- Repeated measures
- Analysis of covariance

Abstract

In the design of a randomized clinical trial with one pre and multiple post randomized assessments of the outcome variable, one needs to account for the repeated measures in determining the appropriate sample size. Unfortunately, one seldom has a good estimate of the variance of the outcome measure, let alone the correlations among the measurements over time.

We show how sample sizes can be calculated by making conservative assumptions regarding the correlations for a variety of covariance structures. The most conservative choice for the correlation depends on the covariance structure and the number of repeated measures. In the absence of good estimates of the correlations, the sample size is often based on a two-sample t-test, making the 'ultra' conservative and unrealistic assumption that there are zero correlations between the baseline and follow-up measures while at the same time assuming there are perfect correlations between the follow-up measures.

Compared to the case of taking a single measurement, substantial savings in sample size can be realized by accounting for the repeated measures, even with very conservative assumptions regarding the parameters of the assumed correlation matrix. Assuming compound symmetry, the sample size from the two-sample t-test calculation can be reduced at least 44%, 56%, and 61% for repeated measures analysis of covariance by taking 2, 3, and 4 follow-up measures, respectively.

The results offer a rational basis for determining a fairly conservative, yet efficient, sample size for clinical trials with repeated measures and a baseline value.

BACKGROUND

It is not unusual for a clinical trial to include multiple assessments of the outcome, both before and particularly after randomization. These repeated measures serve multiple purposes, from reducing the within person variability to allowing an evaluation of the change in the outcome over time. Mathews et al. [1] suggested using summary measures to capture the clinically relevant information available in the repeated measures, and several authors [2-9] have discussed various aspects of such an approach. Obvious summary measures include the mean of the repeated measures, the area under the curve, and the within patient slope.

Frison and Pocock [2] suggest replacing the repeated measures with pre and post-randomization means of the outcome variables and using analysis of covariance to assess the treatment

main effect when the main interest is in the difference in average responses. They provide a formula for calculating the sample size for a clinical trial with both pre and post-randomization repeated measures. The sample size depends on the variance of the outcome variable as well as the correlation among the repeated measures. Estimates for these parameters are sometimes sought in the literature. While it is often difficult to obtain good estimates for the variance, it is even more difficult to obtain good estimates of the correlations between time points for the same population proposed and with the same time spacing between observations. In the absence of a good estimate for the correlations, one sometimes conservatively assumes that the correlation between baseline and post-randomization measures is zero and the correlation among the post-randomization measures is one, and calculates a sample size for a simple two-sample comparison of means [10]. While this produces an ultraconservative estimate for

the variance of the statistic, it is usually unreasonable to assume the post randomized outcomes will be perfectly correlated while having absolutely no correlation with the baseline values.

The use of repeated measures increases the power of clinical trials to detect treatment differences in mean levels of the outcome measure over time. The power decreases with increasing correlations among post-randomization measures and with decreasing correlations between the pre-randomization and post-randomization measures. By reconciling these competing interests, we show that correlations can be chosen that maximize the sample size for different numbers of repeated measures and different covariance structures. This paper will present conservative estimates of the correlation between outcome measures under different assumptions regarding the covariance structure and give the ratio of the conservative sample size to the ultraconservative sample size as a function of the number of repeated post measures (k). Even under the most conservative assumption for a given covariance structure, we show that the sample size is greatly reduced, relative to the ultra-conservative sample size, when just two follow-up measurements are taken.

METHODS

An approximate formula for the sample size required for a two-sample t-test is given by

$$N \approx 4\sigma^2(z_{1-\alpha/2} + z_{1-\beta})^2 / \Delta^2, \quad (1)$$

Where σ^2 is the common variance in the two groups, Δ is the difference in group means, and $z_{1-\alpha}$ is the 100(1- α)th percentile of the standard normal distribution. When the variance is estimated, the sample size formula is based on the non-central t-distribution, but even for studies of modest size ($N > 25$), the sample size is nearly proportional to σ^2 .

When one has a baseline measure of the outcome, analysis of covariance (ANCOVA) is used for the analysis, and the variance of the treatment effect depends on the correlation between the pre and post-randomization measures. The asymptotic variance of the difference in adjusted mean values of the outcome measure is equal to $4\sigma^2(1-\rho^2)/N$, where σ^2 is the variance of the post-randomized outcome measure and ρ is the pre/post correlation. The required sample size can be obtained approximately from (1) with σ^2 replaced by $\sigma^2(1-\rho^2)$ [11,12].

$$N \approx 4\sigma^2(1-\rho^2)(z_{1-\alpha/2} + z_{1-\beta})^2 / \Delta^2$$

When one has multiple post-randomization measurements, repeated measures (RM) ANOVA can be used for the analysis, and the average treatment main effect is the difference in the mean values of the outcome averaged over the post-randomization period. The variance of the treatment effect depends on the variance of the outcome variable during the post randomization period and the correlations between time points, and is given by

$$4 \sum_{i=1}^k \sum_{j=1}^k \rho_{ij} S_i S_j / Nk^2,$$

where ρ_{ij} is the correlation between outcome measures at times i and j , k is the number of follow-up repeated measures, and S_i is the standard deviation of the outcome measured at time i . Replacing

the variance in (1) with $\sum_{i=1}^k \sum_{j=1}^k \rho_{ij} S_i S_j / k^2$, an approximate sample size formula for RM-ANOVA is given by

$$N \approx 4 \sum_{i=1}^k \sum_{j=1}^k \rho_{ij} S_i S_j (z_{1-\alpha/2} + z_{1-\beta})^2 / (k\Delta)^2.$$

When one has a baseline measure and multiple post-randomization measurements, RM-ANCOVA is used for the analysis, and the treatment main effect is the difference in the mean values of the outcome averaged over the post-randomization period adjusted for the baseline levels. The variance of the treatment effect depends on the variance of the outcome variable, the correlation between the baseline and post-randomization means, and the post-randomization correlations, and is given by

$$N \times \text{Var}(\text{statistic}) = 4 \left(\frac{\sum_{i=1}^k V_i + 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^k \rho_{ij} S_i S_j}{k^2} - \frac{(\sum_{i=1}^k \rho_{0i} S_0 S_i)^2}{(S_0 k)^2} \right), \quad (2)$$

where V_i is the variance at time i , ρ_{ij} is the correlation between outcome measures at times i and j ($i \neq j > 0$), and ρ_{0i} is the correlation between outcome measures at times 0 and i ($i > 0$). Replacing $4\sigma^2$ in (1) with (2) gives an approximate sample size formula for RM-ANCOVA. If we assume homogeneity of variances with all variances equal to V , (2) can be simplified [13] as:

$$N \times \text{Var}(\text{statistic}) = 4V \left[\frac{1 + (k-1)\bar{\rho}_{ij}}{k} - \bar{\rho}_{0i}^2 \right],$$

where $\bar{\rho}_{ij}$ is the mean of the $k(k-1)/2$ correlations between the outcome measures and $\bar{\rho}_{0i}$ is the mean of the k correlations between the baseline measure and the outcome measures.

The required sample size is approximately proportional to the variance of the proposed test statistic. Let the variance of the statistic be $N^{-1}V_s$, where N is the total sample size. In the absence of a good estimate for ρ_{ij} , one often uses an ultra-conservative assumption that the correlations between baseline and post-randomization measures are zero and the correlations among the post-randomization measures are one, in which case $V_s = 4\sigma^2$ and one would calculate a sample size for a simple two-sample comparison of means [10]. Define the Variance Ratio (VR) as the variance given in (2) compared to the variance obtained using ultraconservative assumptions, ($V_s = 4V$). The degree to which the variance of the test statistic can be reduced is given by 1-VR.

RESULTS

Compound Symmetry (CS)

Compound symmetry (CS) is often assumed as a covariance structure between repeated measures as it is the variance structure assumed in a random effects model. Assuming CS with a common variance for the dependent variable, V , and a common correlation between the time periods, the variance of the adjusted mean values of the dependent variable over the k follow-up times is given by [2] (see Appendix):

$$V_s = \frac{4V[1+(k-1)\rho] - k\rho^2}{k}, \text{ and}$$

$$VR = [1+(k-1)\rho - k\rho^2]/k.$$

The VR for simple ANCOVA ($k=1$) is $1-\rho^2$ and decreases with increasing positive values of ρ ; however, the VR for k post randomized repeated measures is $(1+(k-1)\rho)/k$ and increases with increasing values of ρ . The variance ratio for RM-ANCOVA is a combination of these effects that increases with ρ for small values and then decreases for larger values (Figure 1).

The VR is maximized when ρ is taken to be:

$$\rho_{\max} = \frac{k-1}{2k}.$$

The value of VR evaluated at ρ_{\max} is

$$VR_{\max} = \left[1 + \frac{(k-1)^2}{4k} \right] / k \\ = \frac{(k+1)^2}{4k^2}.$$

The values of ρ_{\max} and VR_{\max} under CS are given in Table 1, and it is seen that the VR decreases with k . Relative to the ultraconservative (two-sample test), the sample size can be reduced 56% when k is 3 by making a reasonable assumption of CS and then using a conservative estimate of ρ . The VR decreases with k .

Autoregressive

Sometimes the CS assumption is overly restrictive and the correlation decreases with the length of the interval between time points. To allow for decreasing correlations the farther apart the time periods, an autoregressive (AR) covariance structure can be assumed, where the correlation between the repeated measures at periods i and j is equal to $\rho^{|i-j|}$. Assuming an AR matrix with a common variance for the dependent variable, V , and the correlations between the time periods being powers of ρ , the variance of the adjusted mean values of the dependent variable over the k follow-up times is given by $4V \times VR$, where (Appendix):

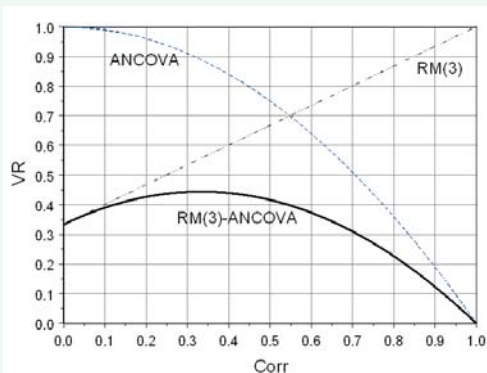


Figure 1 Variance Ratio (VR) as a function of the correlation between measures assuming compound symmetry for three designs 1) Analysis of covariance, ANCOVA, with one outcome measure and one baseline measure, 2) Repeated measures with three outcome measures, RM(3) and no covariate, and 3) Repeated measures analysis of covariance with three outcomes and one covariate, RM-ANCOVA(3).

$$VR = \left[k + 2 \sum_{i=1}^{k-1} (k-i) \rho^i \cdot \frac{(\rho - \rho^{k+1})^2}{(1-\rho)^2} \right] / \left[k^2 (1-\rho)^2 \right] \\ = \left(k + \left\{ 2 \left[(k-1)\rho - k\rho^2 + \rho^{k+1} \right] - (\rho - \rho^{k+1})^2 \right\} / (1-\rho)^2 \right) / k^2. \quad (3)$$

The values of ρ_{\max} that maximize (3) and the VR_{\max} under AR are given in Table 1. It can be seen that the sample size can be reduced 43% relative to the ultraconservative sample size when k is 3 by making a reasonable assumption of AR and then using a conservative estimate of ρ . The VRs under the AR structure are greater than the VRs under the CS structure because under the AR assumption the average of the correlations between baseline and the outcome measures is less than the average of the correlations between the outcome measures.

Dampened autoregressive

While the autoregressive structure allows for the correlation between measures to decline the farther apart in time they are, that model imposes a fairly drastic decrease in time. If $\rho=0.6$, the correlations for measures 1, 2, 3, 4, and 5 units apart are 0.60, 0.36, 0.22, 0.13, and 0.08, respectively. A more reasonable structure that still allows the pairwise correlations to decrease with time is the dampened autoregressive structure where $\rho_{ij} = \rho^{|i-j|} \theta$. If the dampening factor θ is selected to be 0.5, there is still a significant, but more reasonable, decrease in the pairwise correlations with time. For $\rho=0.6$ and $\theta=0.5$, the correlations for measures 1, 2, 3, 4, and 5 units apart are 0.60, 0.49, 0.41, 0.36, and 0.32, respectively. The values of ρ_{\max} and VR_{\max} for the dampened AR model with $\theta=0.5$ are given in Table 1. It can be seen that the sample size can be reduced 51% relative to the ultraconservative sample size when k is 3 by making a reasonable assumption of AR and then using a conservative estimate of ρ .

Toeplitz

If we expect the correlation between time points to differ depending on the spacing between time periods but do not want to assume the correlations are a power function of each other, we could model the correlations using a banded Toeplitz structure where the correlation between periods i and j is equal to $\rho_{|i-j|}$. This structure allows for $k+1$ parameters for the covariance matrix of the baseline and k repeated measures. Assuming banded Toeplitz correlations with a common variance for the dependent variable, the VR of the estimated adjusted mean values of the dependent variable over the k follow-up times is given by (Appendix):

$$VR = \left[k + 2 \sum_{i=1}^{k-1} (k-i) \rho_i - \left(\sum_{i=1}^k \rho_i \right)^2 \right] / k^2. \quad (4)$$

The correlations that maximize (4) are equal to 1 for $i \leq k/2$ and 0 otherwise. The VR_{\max} is equal to $(k+1)/(2k)$. The values of VR_{\max} are given in Table 1. Even assuming a less restrictive banded Toeplitz correlation matrix, the sample size can be reduced 33% when $k=3$.

Heterogeneous Compound Symmetry (CSH)

In order to consider the effect of heterogeneous variances over time, a CSH model will be assumed and compared to the results for CS, which assumes the variances at all times are equal. Assuming a CSH model with the standard deviation of the outcome

Table 1: The correlation (ρ_{max}) that maximizes the variance of the statistic or the required sample size and the variance ratio (VR) as a function of the number of repeated measures (k).

K	CS		AR		Dampened AR		Toepelitz	
	ρ_{max}	VR	ρ_{max}	VR	ρ_{max}	VR	ρ_{max}	VR
2	0.2500	0.5625	0.3981	0.6216	0.3253	0.5925	0 or 1	0.7500
3	0.3333	0.4444	0.5529	0.5297	0.4465	0.4887	0 or 1	0.6667
4	0.3750	0.3906	0.6416	0.4884	0.5154	0.4421	0 or 1	0.6250
5	0.4000	0.3600	0.7001	0.4650	0.5617	0.4159	0 or 1	0.6000
10	0.4500	0.3025	0.8336	0.4211	0.6769	0.3677	0 or 1	0.5500
∞	0.5000	0.2500	1.0000	0.3811	1.0000	0.3267	0 or 1	0.5000

measure at time i is S_{ij} , the ρ_{max} and VR_{max} estimated adjusted mean values of the dependent variable over the k follow-up times are given by (Appendix):

$$\rho_{max} = \frac{\left(\sum_{i=1}^{k-1} \sum_{j=i+1}^k S_i S_j \right)}{\left(\sum_{i=1}^k S_i \right)^2}$$

$$VR_{max} = \frac{\left[\sum_{i=1}^k V_i + \left(\sum_{i=1}^{k-1} \sum_{j=i+1}^k S_i S_j \right)^2 / \left(\sum_{i=1}^k S_i \right)^2 \right]}{\left(\sum_{i=1}^k S_i \right)^2}$$

Values of ρ_{max} and VR_{max} are provided for various degrees of heterogeneity in Table 2. In that table, the standard deviation of the dependent variable at baseline (time 0) is S and the standard deviation at time periods i is equal to $R^i S$ so the standard deviation increases by a multiplicative factor R at each subsequent time period or the variance increases by a factor R^2 . For even moderate heterogeneity where the variance increases 50% at each subsequent time period, there is a negligible effect on VR_{max} .

CONCLUSIONS

The samples size required to meet specified design criteria depends on the variance of the proposed test statistic. The variance of the test statistic depends on the variance of the outcome measure(s) and, for a study with repeated measures, also depends on the correlation between the repeated measures. For a simple two-sample t-test, the variance of the statistic only depends on the variance of the outcome measure. For designs that make use of repeated measures of the outcome variable, including a baseline pre-randomized value and/or multiple repeated post-randomized values, the variance of the statistic for the main group effect for ANCOVA, RM-ANOVA, and RM-ANCOVA depends on the correlations between the repeated observations as well as the variance of the outcome. The true variance is never known. It is often difficult to obtain good estimates of the variance of the proposed outcome variable measured in a population with similar eligibility criteria, as the proposed study. As difficult as it is to obtain good estimates for the variance, obtaining good estimates of the correlations is much more difficult. Even when you can find a study in the literature that uses the same outcome measure as that being proposed and has similar eligibility criteria; it is hard to find one that has repeated measures taken at the same time intervals as those for the proposed study. Even when such a study is found, it is rare that the correlations between the repeated measures are published.

For simple ANCOVA, the variance and hence the sample size is proportional to $(1-\rho^2)$. When a good estimate of ρ is not available, the advantage of having a baseline covariate is ignored by conservatively assuming $\rho=0$ and then calculating the sample size based on a simple two sample t-test. For simple RM-ANOVA with k repeated outcome measures, the variance, and hence the sample size, is proportional to $[1+(k-1)\rho]/k$. When a good estimate of ρ is not available, the advantage of having multiple outcome measures is ignored by conservatively assuming $\rho=1$ and then calculating the sample size based on a simple two sample

Table 2: The correlation (ρ_{max}) that maximizes the variance of the statistic or the required sample size and the variance ratio (VR) as a function of the number of repeated measures (k) and the ratio (R) of the neighboring standard deviations in the heterogeneous model. (R^2 is the ratio of the neighboring variances).

K	R	R^2	ρ_{max}	VR
2	0.8	0.64	0.2469	0.5671
	0.9	0.81	0.2493	0.5635
	1.0	1.0	0.2500	0.5625
	1.1	1.21	0.2494	0.5633
	1.2	1.44	0.2479	0.5656
	1.3	1.69	0.2457	0.5689
	1.5	2.25	0.2400	0.5776
3	2.0	4.00	0.2222	0.6049
	0.8	0.64	0.3279	0.4518
	0.9	0.81	0.3321	0.4461
	1.0	1.0	0.3333	0.4444
	1.1	1.21	0.3323	0.4458
	1.2	1.44	0.3297	0.4493
	1.3	1.69	0.3258	0.4545
4	1.5	2.25	0.3158	0.4681
	2.0	4.00	0.2857	0.5102
	0.8	0.64	0.3674	0.4002
	0.9	0.81	0.3733	0.3928
	1.0	1.0	0.3750	0.3906
	1.1	1.21	0.3736	0.3924
	1.2	1.44	0.3699	0.3971
5	1.3	1.69	0.3645	0.4038
	1.5	2.25	0.3508	0.4215
	2.0	4.00	0.3111	0.4746

t-test. For RM-ANCOVA, the variance of the statistic for the main effect of group depends on the correlations between the outcome variables, and the variance decreases with increasing correlations between the baseline measure and the outcome measures while it increases with increasing correlations between the post-randomized measures of the outcome variables. To be the most conservative, which we have termed being ultraconservative, you must unrealistically assume there is absolutely no correlation between the outcome variable measured at baseline and the post randomized measures of the outcome measures and at the same time assume there is perfect correlation between the outcome variables at the different post randomized values. Under this ultraconservative assumption, the statistic for the main effect of group in the RM-ANCOVA design reduces to the two sample t-test. In this paper, we have conditioned on some reasonable assumptions about the form of the covariance matrix of the repeated measures and determined the correlation(s) that maximize the variance of the statistic for the intervention main effect to produce conservative determinations of the sample size. Depending of the assumed structure of the covariance matrix, this paper gives the appropriate factor, VR, which one would multiply the sample size derived from a two-sample t-test by to obtain a reasonably conservative sample size determination.

All too often the justification for a sample size is given before the primary statistical method to assess the treatment effect is given. Suppose a study desires to have 90% power at the 5% two-sided level of significance to detect a 10mm difference in systolic blood pressure (SBP) between the two randomized groups when the standard deviation of 20mm. Given these design criteria and assumption regarding the variance, one often sees the sample size justified as 172 subjects based on a two sample t-test. However, suppose the main test for treatment effect will be based on the test for the main effect in repeated measures ANCOVA. This will allow the trial to have greater than 90% power to detect the 10mm difference between the groups. If you stated the primary test statistic first, you would word the sample size justification something like the following:

The sample size determination can be computed using formula for a two sample t-test replacing the variance of the outcome measure by N times the variance ratio of the estimated main effect in a repeated measures ANCOVA model. Given we do not have good estimates of the correlations between the outcome measures, we will be (ultra) conservative and assume the correlation between baseline and follow-up measures is 0 and the correlation between is 1 in which case the statistic is a two-sample t-test that requires a sample size of 172 subjects. Given these ultra-conservative assumptions will not be true the repeated measure ANCOVA design will provide an unknown power greater than 90%.

However, if we rely on the ultra-conservative assumptions, there is no decrease in the 'required' number of subjects from using a two-sample t-test which requires no baseline assessment and only one post randomized assessment compared to the repeated measures ANCOVA which requires (k+1) outcome assessments. Without having a good estimate of the correlations between time points, one can realize some of the savings that should be possible by using a covariate and repeated follow-

up assessments by assuming some structure to the covariance matrix. In this case the sample size determination section can be worded as follows:

The sample size based on the test for main effect from a 3 repeated measures ANCOVA design can be computed using formula for a two sample t-test replacing the variance of the outcome measure by N times the variance ratio of the estimated main effect in a repeated measures ANCOVA model. Assuming a compound symmetry covariance structure and using the value of the common correlation that maximizes the variance, the required sample size is $172 \times 0.444 = 77$.

For the correlation structures considered where the correlation stay the same or decreased across time, the greatest reduction in the maximal variance of the statistic is when we assume a CS structure for the covariance matrix of the repeated measures. When ρ is unknown and the sample size is based on the most conservative value of ρ , the actual power will be greater or equal to the designed power if the CS assumption holds. While a CS structure is often assumed at the design stage of planning a study, it is possible the CS assumption does not hold and another structure is more appropriate. In such a case, it is possible that the actual power will be less than the desired power. Consider for example the power of a study designed to have adequate power for the worst possible correlation under CS when in fact the correlations between time periods decrease somewhat the farther part in time? For the case of $k=3$ and the sample size is chosen to achieve 90% power at the 5% two-sided level of significance using VR_{max} assuming CS, if the correlation matrix is really a dampened autoregressive matrix with power of $\frac{1}{2}$, the true power of the study will be greater than 90% if the true correlation between neighboring times is ≤ 0.235 or ≥ 0.631 and has its lowest power of 87% when this correlation is 0.446. Similarly, if the true correlation matrix is autoregressive, the true power will be greater than 90% if the true correlation between neighboring times is ≤ 0.245 or ≥ 0.765 and has its lowest power of 84% when this correlation is 0.553.

This paper has looked at the required sample size for statistics that are normally distributed. When the variance of the statistic is known, the sample size formula is given by [1]. When the variance is estimated and the statistic has a t-distribution, the quantiles for the normal distribution in (1) would be replaced by quantiles for a non-central t-distribution and iterative methods would be used to solve for N. For large N, the sample size formula using the non-central t is still proportional to the variance of the statistic. For small sample sizes it is not. The results presented can still be used to determine N by using a sample size program for a simple two sample t-test but multiplying the variance by the VR provided. While we assume the statistic is normally distributed, we do not assume the data have a normal distribution. The initial assumption that the variance of the statistic is equal under the null and alternative hypotheses was made for simplicity in notation. Even if the variance is dependent on the mean and is different under the null and alternative hypotheses, the VR is the same in each case.

REFERENCES

1. Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial

- measurements in medical research. *BMJ*. 1990; 300: 230-235.
2. Frison L, Pocock SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Stat Med*. 1992; 11: 1685-1704.
 3. Matthews JN. A refinement to the analysis of serial data using summary statistics. *Stat Med*. 1993; 12: 27-37.
 4. Everitt BS. The analysis of repeated measures: a practical review with examples. *The Statistician* 1995; 44: 113-135.
 5. Frison LJ, Pocock SJ. Linearly divergent treatment effects in clinical trials with repeated measures: efficient analysis using summary statistics. *Stat Med*. 1997; 16: 2855-2872.
 6. Dawson JD. Sample size calculations based on slopes and other summary statistics. *Biometrics*. 1998; 54: 323-330.
 7. Schouten HJ. Planning group sizes in clinical trials with a continuous outcome and repeated measures. *Stat Med* 1999; 18: 255-264.
 8. Ahn C, Overall JE, Tonidandel S. Sample size and power calculations in repeated measurement analysis. *Comput Methods Programs Biomed*. 2001; 64: 121-124.
 9. Senn S, Stevens L, Chaturvedi N. Repeated measures in clinical trials: simple strategies for analysis using summary measures. *Stat Med*. 2000; 19: 861-877.
 10. Rochon J. Sample size calculation for two-group repeated-measures experiments. *Biometrics*. 1991; 47: 1383-1398.
 11. Fleiss JL. *The Design and Analysis of Clinical Experiments (Appendix: Sample Size Determination)*. New York, New York: Wiley; 1986. p. 369-371.
 12. Borm GF, Franssen J, Lemmens WAJG. A simple sample size formula for analysis of covariance in randomized clinical trials. *J Clin Epidemiol*. 2007; 60: 1234-1238.
 13. Vickers AJ. How many repeated measures in repeated measures designs? Statistical issues for comparative trials. *BMC Med Res Methodol*. 2003; 3: 1-9.

Cite this article

Morgan TM, Case LD (2013) Conservative Sample Size Determination for Repeated Measures Analysis of Covariance. *Ann Biom Biostat* 1: 1002.