Power Calculations for Two-Wave, Change from Baseline to Follow-Up Study Designs

M. Colin Ard[1] and Steven D. Edland[1,2] *

[1]*Department of Neurosciences;* [2]*Department of Family and Preventive Medicine Division of Biostatistics, University of California San Diego, La Jolla, California, USA*

Abstract: Change in a quantitative trait is commonly employed as an endpoint in two-wave longitudinal studies. For example, early phase clinical trials often use two-wave designs with biomarker endpoints to confirm that a treatment affects the putative target treatment pathway before proceeding to larger scale clinical efficacy trials. Power calculations for such designs are straightforward if pilot data from longitudinal investigations of similar duration to the proposed study are available. Often longitudinal pilot data of similar duration are not available, and simplifying assumptions are used to calculate sample size from cross-sectional data, one standard approach being to use a formula based on variance estimated from cross sectional data and correlation estimates abstracted from the literature or inferred from experience with similar endpoints. An implicit assumption of this standard approach is that the variance of the quantitative trait is the same at baseline and follow-up. In practice, this assumption rarely holds, and sample size estimates by this standard formula can be dramatically anti-conservative. Even when longitudinal pilot data for estimating parameters required in sample size calculations are available, sample size calculations will be biased if the interval from baseline to follow-up is not of similar duration to that proposed for the study being designed. In this paper we characterize the magnitude of bias in sample size estimates when formula assumptions do not hold and derive alternative conservative formulas for sample size required to achieve nominal power.

Keywords: Sample Size, Phase 2, Phase II, Clinical Trial, Rate of Change, Compound Symmetry.

1. INTRODUCTION

Quantitative traits are commonly employed as endpoints in phase II clinical trials. In Alzheimer's disease, recent efforts have focused on changes in biomarkers as a means of studying disease progression and drug efficacy, and as an aid to diagnosis and early decision-making in the evaluation of clinical interventions for chronic, debilitating and degenerative disease [1-4]. The ongoing process of identifying and validating biomarkers for use as surrogate endpoints, and more generally the quality and informativeness of research outcomes derived from the analysis of such endpoints, require that these trials be adequately powered.

Power and sample size calculations for two-wave (baseline and one follow-up) designs using quantitative endpoints are straightforward if pilot data from longitudinal investigations of similar duration to the proposed study are available. In this case the variance of change scores, $\sigma^2_{d(t)}$, can be estimated from pilot

data, and supplied to the usual two-sample normal approximation formula for calculating sample size requirements for a *t*-test of no treatment effect against a non-directional alternative hypothesis:

$$n_t = \frac{2\sigma^2_{d(t)}\left[\Phi^{-1}\left(\alpha/2\right) + \Phi^{-1}\left(\beta\right)\right]^2}{\Delta^2_{(t)}} \qquad (1)$$

where n_t is the number of subjects required per study arm for a trial of duration t, $\sigma^2_{d(t)}$ is the within-group variance of the change in measurements from baseline to follow-up for a study of duration time t, $\Phi^{-1}(p)$ is the quantile function of the standard normal distribution, α and β are the Type I and Type II error rates, and $\Delta_{(t)}$ is the magnitude of the between-group difference in mean change from baseline to follow-up time t that the investigator wishes to be able to detect with probability $1 - \beta$.

Longitudinal pilot data are required to estimate the variance term, $\sigma^2_{d(t)}$, in equation (1). Lacking longitudinal pilot data or published estimates, it may be possible to abstract related parameters from the literature for estimating $\sigma^2_{d(t)}$. For example, values of variance at baseline and at follow-up, and the correlation of baseline to follow-up, may be available in published reports. Calling the variance at baseline $\sigma^2_{(0)}$, the variance at follow-up $\sigma^2_{(t)}$, and the correlation between baseline and follow-up $\rho_{(0t)}$, $\sigma^2_{d(t)}$ is then calculated as:

$$\sigma_{d(t)}^2 = \sigma_{(0)}^2 + \sigma_{(t)}^2 - 2\rho_{(0t)}\sigma_{(0)}\sigma_{(t)} \qquad (2)$$

and the value obtained from equation (2) can then be used in place of $\sigma_{d(t)}^2$ in equation (1).

If $\sigma_{(t)}^2 = \sigma_{(0)}^2$, that is, if there is compound symmetry of the covariance matrix of the baseline and follow-up (time *t*) measurements, then equation (2) reduces to:

$$\sigma_{d(t)}^2 = 2\left(1 - \rho_{(0t)}\right)\sigma_{(0)}^2 \qquad (3)$$

As was the case with equation (2), the result of equation (3) can be substituted directly into equation (1), but this is valid only when $\sigma_{(t)}^2 = \sigma_{(0)}^2$.

Equation (3) is presented, for example, in Meinert [5] (equation 9.14), Lachin [6], Overall and Doyle [7], and Overall & Starbuck [8]. In all cases the formula is presented with no discussion of what the assumption that $\sigma_{(t)}^2 = \sigma_{(0)}^2$ implies about the data generating process, or of the consequences of making this assumption in error. This is potentially problematic, because applying equation (3) when compound symmetry does not hold can lead to dramatic underestimation of the required sample size. For chronic progressive disease in particular, persons tend to progress at different rates, so that trajectories of progression tend to fan apart and the variance of measurements of disease severity increase as the cohort is followed over time [9]. If the interval from baseline to follow-up is sufficient, there may be a substantial increase in variance from baseline to follow-up and a substantial underestimation of required sample size by equation (3).

A similar danger arises when the researcher has direct access to longitudinal pilot data or is able to abstract relevant parameters from a published longitudinal study, but the duration *t* of the planned trial exceeds the duration, *s* say, of the available pilot data. Here, again, if longitudinal trajectories fan apart, then $\sigma_{d(s)}^2 < \sigma_{d(t)}^2$ whenever $s < t$, and sample size estimates obtained by naïve use of $\sigma_{d(s)}^2$ in place of $\sigma_{d(t)}^2$ will be anti-conservative. In the absence of a model describing the functional dependence of the per-occasion variances and between-occasion correlations on the measurement time, no sample size formula is available in this scenario.

In this paper we illustrate analytically and by way of example the potential magnitude of underestimation in sample size calculations by naively applying equation

(3) when the assumption that $\sigma_{(t)}^2 = \sigma_{(0)}^2$ does not hold (Section 2). We also demonstrate that even when longitudinal pilot data are available for estimating parameters used in sample size calculations, these estimates and resulting sample size calculations can be biased if the interval from baseline to follow-up is not of similar duration to that proposed for the trial (Section 3). Section 3 will also derive conservative upper bounds on required sample size useful for sensitivity analyses when adequate pilot data are not available for planning a future trial.

2. BIAS IN SAMPLE SIZE CALCULATIONS WHEN THE $\sigma_{(t)}^2 = \sigma_{(0)}^2$ ASSUMPTION DOES NOT HOLD

Letting n_t represent the true required sample size per arm for a study of length *t* and denoting the estimate obtained using equation (3) as $n_t^{(3)}$, for a given effect size and Type I and Type II error probabilities, the extent to which $n_t^{(3)}$ underestimates n_t, expressed as a percentage of the true required sample size is given by:

$$\%Underestimation = 100\% \times \frac{n_t - n_t^{(3)}}{n_t} = 100\%$$

$$\times \frac{\left(\sigma_{(t)} - \sigma_{(0)}\right)\left[\sigma_{(t)} - \left(2\rho_{(0t)} - 1\right)\sigma_{(0)}\right]}{\sigma_{(0)}^2 + \sigma_{(t)}^2 - 2\rho_{(0t)}\sigma_{(0)}\sigma_{(t)}} \qquad (4)$$

Note that $2\rho_{(0t)} - 1 \leq 1$, so that the numerator terms in equation (4) are always greater than zero, and hence $n_t^{(3)}$ underestimates the true required sample size, whenever $\sigma_{(t)}^2 > \sigma_{(0)}^2$.

Example

To illustrate, we present examples of sample size calculations for a hypothetical phase II clinical trial in Alzheimer's disease. Longitudinal data excerpted from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort study are used to inform power calculations. ADNI is a joint government, private pharmaceutical, and non-profit initiative to determine promising markers of early AD progression and aid researchers and clinicians in developing effective therapies. The ADNI cohort is a longitudinal observational study that is by design representative of subjects eligible for recruitment to Alzheimer's disease treatment trials [10]. Baseline and 12-month follow-up data were accessed from this publically available dataset [10] on September 29, 2009. Estimates of $\sigma_{(0)}^2$, $\sigma_{(t)}^2$, and $\rho_{(0t)}$ for lateral inferior ventricular volume and for the Alzheimer's Disease Assessment Scale cognitive subscale (ADAS-

cog [11]), two potential endpoints for a phase II trial of an Alzheimer's disease treatment, are summarized in Table **1**. Subjects with an Alzheimer diagnosis at baseline and data at both baseline and 12-month time points were included in the calculations. The variance is larger at follow-up than at baseline for both outcomes, as expected for longitudinal measurements of a chronic progressive disease. The parameter estimates in Table **1** were used to calculate sample size to obtain 80% power to detect $\Delta_{(t)}$ equal to a 25% reduction of ventricular enlargement relative to that observed in the ADNI cohort placebo condition pilot data, and likewise to calculate sample size to obtain 80% power to detect a 25% slowing of rate of progression on the ADAS-cog relative to placebo, using both the naïve but biased estimate of $\sigma^2_{d(t)}$ (equation (3)) and the appropriate estimate of $\sigma^2_{d(t)}$ (equation (2)). Sample size estimates resulting from naively using equation (3) to calculate $\sigma^2_{d(t)}$ are approximately 50% smaller than the sample sizes required to obtain nominal power as calculated using equation (2) (Table **1**). For example, naïve use of equation (3) to estimate $\sigma^2_{d(t)}$ in equation (1) would lead us to conclude that a sample size of 96 per arm would be sufficient to ensure power = 0.8 to detect a 25% reduction in the rate of ventricular volume increase, whereas the actual required sample size, derived using equation (2), would be 195 subjects per arm. For the ADAS-cog data, these values are 356 and 719, respectively. These sample calculations are for illustrative purposes only. Power calculations for an Alzheimer treatment trial involve additional considerations, including but not limited to calculating bootstrap confidence intervals

about sample size estimates acknowledging the uncertainty in population parameters supplied to the power formula [12, 13].

3. POTENTIAL BIAS WHEN PILOT DATA ARE INCONSISTENT WITH PLANNED TRIAL DESIGN

If researchers have access to pilot data or summary statistics from a study of comparable duration to the planned trial, a valid estimate of the required sample size can be determined using equation (2) to calculate $\sigma^2_{d(t)}$. Often, however, pilot data are from a study of duration s, with $s < t$. This mismatch in trial duration is significant because typically $\sigma^2_{(s)} < \sigma^2_{(t)}$ when s is shorter than t, and therefore sample size calculations tend to be anti-conservative. Letting $n_t^{[s]}$ denote the estimated required sample size per arm calculated from equation (1) with $\sigma^2_{d(s)}$ used in place of $\sigma^2_{d(t)}$, the magnitude of underestimation expressed as a percentage of actual required sample size n_t is given by:

$$\%Underestimation = 100\% \times \frac{n_t - n_t^{[s]}}{n_t} = 100\%$$

$$\times \frac{\sigma^2_{(t)} - \sigma^2_{(s)} - 2\sigma_{(0)}\left(\rho_{(0t)}\sigma_{(t)} - \rho_{(0s)}\sigma_{(s)}\right)}{\sigma^2_{(0)} + \sigma^2_{(t)} - 2\rho_{(0t)}\sigma_{(0)}\sigma_{(t)}} \quad (5)$$

While the specification of conditions that are both necessary and sufficient for equation (5) to take values greater than zero is somewhat more involved than was the case with equation (4), it can be shown, for example, that $n_t^{[s]}$ will be anti-conservative for data

Table 1: Parameter and Sample Size Requirement Estimates for ADAS-cog and Ventricular Volume Endpoints, as Calculated from ADNI Pilot Data

Statistic	Endpoint	
	ADAS-cog	**Inf. Lat. Ventricles**
Pilot Data Sample Size	163	149
Baseline Mean	18.3	4.2
12-Month Change	4.2	0.6
$\sigma^2_{(0)}$	38.6	4.3
$\sigma^2_{(12)}$	92.6	6.0
$\rho_{(0,12)}$	0.68	0.98
n_{12}	719	195
$n_{12}^{(3)}$	356	96
%Underestimation	50.5%	51.1%

Notes: Inf. Lat. Ventricles is the sum of the volumes of the left and right inferior lateral ventricles in cm^3; $\sigma^2_{(0)}$ = Baseline variance; $\sigma^2_{(12)}$ = Month 12 variance; $\rho_{(0,12)}$ = Baseline-to-Month 12 correlation; n_{12} = required sample size per arm to detect a 25% reduction in 12-Month Change with power = 0.8 and two-tailed significance level = 0.05; $n_{12}^{(3)}$ = estimated sample size per arm to detect a 25% reduction in 12-Month Change with power = 0.8 and two-tailed significance level = 0.05 as calculated by equation (3); *%Underestimation* = underestimation of the required sample size per arm by use of equation (3), expressed as a percent.

generated by the familiar mixed effects model with linear trajectories fanning apart over time and i.i.d. residual error, or whenever both: 1) $\sigma^2_{(0)} < \sigma^2_{(s)} < \sigma^2_{(t)}$ and 2) $0 \leq \rho_{(0t)} \leq \rho_{(0s)}$ (Appendix).

Lacking pilot data consistent with the planned trial, there are several alternative approaches. If longitudinal data with more than two waves of observation are available, then power calculations can be performed under the assumption of a linear mixed effects model with correlated random intercepts and slopes [12], a flexible yet parsimonious analytic framework for longitudinal data that is well-suited to situations where variances are expected to increase (and correlations are expected to decrease), as the study duration is extended. However, if only two-wave pilot data are available and the duration of pilot data observation is different (typically shorter than), the duration proposed for the planned trial, the variance components from the correlated random intercepts and slopes variant of the linear mixed effects model are not identifiable, and calculation of the required sample size per arm is not possible. Nonetheless, the linear mixed effects model can still be relied upon to motivate solutions for the required sample size per arm that are guaranteed to be conservative under certain conditions. We offer here two power formulas that provide conservative overestimates of required sample size when the duration of pilot data is less than the duration of the planned trial.

Conservative Formula 1

Under the linear mixed effects model, the variance of change from baseline to follow-up $\sigma^2_{d(t)}$ for a study of length t can be expressed as a function of the variance of change from baseline to follow-up $\sigma^2_{d(s)}$ for a study of length s and the mixed model residual error variance σ^2_e as follows (Appendix):

$$\sigma^2_{d(t)} = \frac{t^2}{s^2} \sigma^2_{d(s)} - 2\frac{t^2 - s^2}{s^2} \sigma^2_e \qquad (6)$$

For $t > s$, the second term is greater than or equal to zero, and therefore:

$$\dot{\sigma}^2_{d(t)} = \frac{t^2}{s^2} \sigma^2_{d(s)} \qquad (7)$$

is an overestimate of $\sigma^2_{d(t)}$, and power calculations by equation (1) with $\dot{\sigma}^2_{d(t)}$ used in place of $\sigma^2_{d(t)}$ will be conservative overestimates of required sample size to achieve power $(1 - \beta)$.

Conservative Formula 2

Alternatively, under the linear mixed effects model, the variance of change from baseline to follow-up $\sigma^2_{d(t)}$ for a study of length t can be expressed as a function of $\sigma^2_{d(s)}$, $\sigma^2_{(s)}$ = the variance of the quantitative measure at follow-up time s, $\sigma^2_{(0)}$ = the variance at the baseline observation, and σ_{ab} = the covariance of the random intercept and slope coefficients (Appendix):

$$\sigma^2_{d(t)} = \sigma^2_{d(s)} + \left(\frac{t^2 - s^2}{s^2}\right)\left(\sigma^2_{(s)} - \sigma^2_{(0)}\right) - 2\left(\frac{t^2 - s^2}{s}\right)\sigma_{ab} \qquad (8)$$

For $\sigma_{ab} > 0$ and $t > s$ the final term is greater than or equal to zero, and therefore:

$$\ddot{\sigma}^2_{d(t)} = \sigma^2_{d(s)} + \left(\frac{t^2 - s^2}{s^2}\right)\left(\sigma^2_{(s)} - \sigma^2_{(0)}\right) \qquad (9)$$

is an overestimate of $\sigma^2_{d(t)}$, and power calculations by equation (1) with $\ddot{\sigma}^2_{d(t)}$ used in place of $\sigma^2_{d(t)}$ will yield conservative overestimates of required sample size to achieve power $(1 - \beta)$. Note that the solution is exact, i.e., $\ddot{\sigma}^2_{d(t)} = \sigma^2_{d(t)}$, when $\sigma_{ab} = 0$.

Since both equation (7) and equation (9) yield exact or conservative solutions for required sample size under the linear mixed effects model assumption when the covariance of the random intercept and slope coefficients is non-negative, the smallest sample size estimate of the two calculated sample sizes, i.e., the one using $\min[\dot{\sigma}^2_{d(t)}, \ddot{\sigma}^2_{d(t)}]$, is to be preferred under these conditions.

4. DISCUSSION

Considerable costs can be incurred when time and resources are allocated to an otherwise well conceived but underpowered study. Investigators may therefore wish instead to opt for a conservative solution to the problem at hand, i.e., one that results in a trial with greater than nominal power. In the absence of longitudinal pilot data, there is little information and little recourse for powering longitudinal trials. We recommend against using equation (3) however, and rather suggest using equation (2) to estimate the variance for sample size calculations when balanced two-wave pilot data is available, because equation (2) requires you to explicitly acknowledge the unknown parameters and therefore provide estimates for these parameters to the sample size formula. For the case where longitudinal pilot data of length $s < t$ are available, we have provided conservative formulas for

sample size calculation for trials of length t that are valid under the modest assumption that the data derive from a mixed effects model with linear longitudinal trajectories that are fanning apart. Anti-conservative trials that ignore these concerns risk unnecessary false negative findings and lost opportunities for drug development.

ACKNOWLEDGEMENTS

APPENDIX

Conditions Under which $\sigma^2_{d(s)}$ Leads to Anti-Conservative Sample Size Estimation

When Data are Generated by a Linear Mixed Effect Model

We begin by introducing notation for two-wave data generated by the mixed effects model commonly applied to longitudinal data from clinical trials and cohort studies. Let $y_{i\tau}$ be observations on some random variable of interest that follow the linear mixed effects model with bivariate normal random intercept and slope coefficients and i.i.d. residual error:

$$y_{i\tau} = \alpha + \beta\tau + a_i + b_i\tau + e_{i\tau} \tag{A1}$$

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \sigma^2_a & \sigma_{ab} \\ \sigma_{ab} & \sigma^2_b \end{pmatrix}\right), \quad e_{i\tau} \sim N(0, \sigma^2_e) \tag{A2}$$

Here α and β represent the fixed intercept and slope terms (rather than the Type I and Type II error probabilities), respectively, a_i and b_i are their random counterparts for subject i, and the $e_{i\tau}$ are i.i.d. residual error terms specific to subject i at time τ. For two-wave data, τ will typically only take on two values by design, *e.g.*, zero at baseline and t at follow-up for a trial of length t, or zero at baseline and s at follow-up for a trial of length s. It follows from equations (A1) and (A2) that for any τ:

$$\begin{aligned} \sigma^2_{(\tau)} &= Var\{y_{i\tau}\} \\ &= Var\{a_i + b_i\tau + e_{i\tau}\} \\ &= \sigma^2_a + \tau^2\sigma^2_b + 2\tau\sigma_{ab} + \sigma^2_e \end{aligned} \tag{A3}$$

Also, for any $\tau \neq 0$:

$$\begin{aligned} \sigma^2_{d(\tau)} &= Var\{y_{i\tau} - y_{i0}\} \\ &= Var\{b_i\tau + e_{i\tau} - e_{i0}\} \\ &= \tau^2\sigma^2_b + 2\sigma^2_e \end{aligned} \tag{A4}$$

To prove that $\sigma^2_{d(s)}$ leads to anti-conservative sample size estimates under this model, simply note that by (A4) $\sigma^2_{d(s)} < \sigma^2_{d(t)}$, and therefore $n_t^{[s]}$ calculated using $\sigma^2_{d(s)}$ in equation (1) will underestimate n_t when $s < t$ and $\sigma^2_b > 0$. The $\sigma^2_b > 0$ condition simply means that trajectories are fanning apart, as is typically observed in longitudinal studies of chronic, progressive disease.

When $\sigma^2_{(0)} < \sigma^2_{(s)} < \sigma^2_{(t)}$ and $0 \leq \rho_{(0t)} \leq \rho_{(0s)}$

To show that equation (5) is greater than zero, and hence that $n_t^{[s]}$ is anti-conservative, under the general conditions that 1) $\sigma^2_{(0)} < \sigma^2_{(s)} < \sigma^2_{(t)}$ and 2) $0 \leq \rho_{(0t)} \leq \rho_{(0s)}$, we need to prove that the numerator of equation (5) is positive under these conditions, which is equivalent to showing that:

$$\frac{\sigma_{(s)} + \sigma_{(t)}}{2\sigma_{(0)}} > \frac{\rho_{(0t)}\sigma_{(t)} - \rho_{(0s)}\sigma_{(s)}}{\sigma_{(t)} - \sigma_{(s)}} \tag{A5}$$

The proof follows immediately by noting that, given the assumptions, the left-hand side of inequality (A5) is > 1, and the right-hand side of (A5) is ≤ 1.

Derivation of Equation (6)

Using the model and notation above, equation (6) follows by writing:

$$\sigma_{d(t)}^2 - \frac{t^2}{s^2}\sigma_{d(s)}^2 = t^2\sigma_b^2 + 2\sigma_e^2 - \frac{t^2}{s^2}\left(s^2\sigma_b^2 + 2\sigma_e^2\right)$$

$$= -2\frac{t^2 - s^2}{s^2}\sigma_e^2 \tag{A6}$$

Derivation of Equation (8)

Similarly, equation (8) follows by noting that:

$$\sigma_{d(t)}^2 - \sigma_{d(s)}^2 - \left(\frac{t^2 - s^2}{s^2}\right)\left(\sigma_{(s)}^2 - \sigma_{(0)}^2\right)$$

$$= t^2\sigma_b^2 + 2\sigma_e^2 - \left(s^2\sigma_b^2 + 2\sigma_e^2\right) - \left(\frac{t^2 - s^2}{s^2}\right)$$

$$\times \left[\left(\sigma_a^2 + s^2\sigma_b^2 + 2s\sigma_{ab} + \sigma_e^2\right) - \left(\sigma_a^2 + \sigma_e^2\right)\right] \tag{A7}$$

$$= \left(t^2 - s^2\right)\sigma_b^2 - \left(\frac{t^2 - s^2}{s^2}\right)\left(s^2\sigma_b^2 + 2s\sigma_{ab}\right)$$

$$= -2\left(\frac{t^2 - s^2}{s}\right)\sigma_{ab}$$

REFERENCES

[1]   Hampel H, Mitchell A, Blennow K, *et al*. Core biological marker candidates of Alzheimer's disease - perspectives for diagnosis, prediction of outcome and reflection of biological activity. J Neural Trans 2004; 111: 247-72. http://dx.doi.org/10.1007/s00702-003-0065-z

[2]   Matthews B, Siemers ER, Mozley PD. Imaging-based measures of disease progression in clinical trials of disease modifying drugs for Alzheimer disease. Am J Geriat Psychiatr 2003;11: 146-59.

[3]   Mohs RC, Kawas C, Carrillo MC. Optimal design of clinical trials for drugs designed to slow the course of Alzheimer's disease. Alzheimers Dement 2006; 2: 131-9. http://dx.doi.org/10.1016/j.jalz.2006.04.003

[4]   Thal LJ, Kantarci K, Reiman EM, *et al*. The role of biomarkers in clinical trials for Alzheimer disease. Alz Dis Assoc Dis 2006; 20: 6-15. http://dx.doi.org/10.1097/01.wad.0000191420.61260.a8

[5]   Meinert CL. Clinical trials: design, conduct, and analysis. 2nd ed. New York: Oxford University Press 2012.

[6]   Lachin JM. Introduction to sample-size determination and power analysis for clinical trials. Control Clin Trials 1981; 2: 93-113. http://dx.doi.org/10.1016/0197-2456(81)90001-5

[7]   Overall JE, Doyle SR. Estimating sample sizes for repeated measurement designs. Control Clin Trials 1994; 15: 100-23. http://dx.doi.org/10.1016/0197-2456(94)90015-9

[8]   Overall JE, Starbuck RR. Sample-size estimation for randomized pre-post designs. J Psychiatr Res 1979; 15: 51-5. http://dx.doi.org/10.1016/0022-3956(79)90006-2

[9]   Edland SD. Blomqvist revisited: how and when to test the relationship between level and longitudinal rate of change. Stat Med 2000; 19: 1441-52. http://dx.doi.org/10.1002/(SICI)1097-0258(20000615/30)19:11/12<1441::AID-SIM436>3.0.CO;2-H

[10]   Petersen RC, Aisen PS, Beckett LA, *et al*. Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. Neurol 2010; 74: 201-9.

[11]   Mohs RC, Knopman D, Petersen RC, *et al*. Development of cognitive instruments for use in clinical trials of antidementia drugs: Additions to the Alzheimer's disease assessment scale that broaden its scope. Alz Dis Assoc Dis 1997; 11: S13-S21. http://dx.doi.org/10.1097/00002093-199700112-00003

[12]   Ard MC, Edland SD. Power calculations for clinical trials in Alzheimer's disease. J Alzheimers Dis 2011; 26: 369-77.

[13]   McEvoy LK, Edland SD, Holland D, *et al*. Neuroimaging enrichment strategy for secondary prevention trials in Alzheimer disease. Alz Dis Assoc Dis 2010; 24: 269-77.