

UC San Diego

UC San Diego Previously Published Works

Title

Power formulas for mixed effects models with random slope and intercept comparing rate of change across groups

Permalink

<https://escholarship.org/uc/item/1jj9c0sq>

Journal

The International Journal of Biostatistics, 18(1)

ISSN

2194-573X

Authors

Zhao, Yu
Edland, Steven D

Publication Date

2022-05-01

DOI

10.1515/ijb-2020-0107

Peer reviewed



Yu Zhao and Steven D. Edland*

Power formulas for mixed effects models with random slope and intercept comparing rate of change across groups

<https://doi.org/10.1515/ijb-2020-0107>

Received March 23, 2020; accepted December 16, 2020; published online January 18, 2021

Abstract: We have previously derived power calculation formulas for cohort studies and clinical trials using the longitudinal mixed effects model with random slopes and intercepts to compare rate of change across groups [Ard & Edland, Power calculations for clinical trials in Alzheimer’s disease. *J Alzheim Dis* 2011;21:369–77]. We here generalize these power formulas to accommodate 1) missing data due to study subject attrition common to longitudinal studies, 2) unequal sample size across groups, and 3) unequal variance parameters across groups. We demonstrate how these formulas can be used to power a future study even when the design of available pilot study data (i.e., number and interval between longitudinal observations) does not match the design of the planned future study. We demonstrate how differences in variance parameters across groups, typically overlooked in power calculations, can have a dramatic effect on statistical power. This is especially relevant to clinical trials, where changes over time in the treatment arm reflect background variability in progression observed in the placebo control arm plus variability in response to treatment, meaning that power calculations based only on the placebo arm covariance structure may be anticonservative. These more general power formulas are a useful resource for understanding the relative influence of these multiple factors on the efficiency of cohort studies and clinical trials, and for designing future trials under the random slopes and intercepts model.

Keywords: clinical trial; linear mixed effects model; power; sample size; study subject attrition.

1 Introduction

Ref. [1] have previously described sample size formulas for longitudinal studies with study subject dropout for the mixed model repeated measures analysis comparing change from baseline to last visit across groups. Missing data due to study subject dropout in clinical trials and cohort studies is common and reduces statistical power to detect treatment effects or differences in change across groups. We here derive power formulas for longitudinal studies with study subject dropout for a different model, the mixed effects model with random slopes and intercepts comparing mean slope across groups. We demonstrate how power formulas under this model can be used to power a future trial of arbitrary design (arbitrary number and interval between follow-up observation) regardless of the design of pilot study informing power calculations. We expand and generalize previously published mixed effects model power formulas (e.g. [2, 3]) to fully accommodate differences in length and interval between longitudinal observations, different allocation ratios, and different study subject attrition rates. We also derive a formula that accommodates different covariance structures across groups. Differences in covariance are typically ignored, but may be critical to clinical trials, where changes over time in the treatment arm reflect the normal background variability in progression observed in the placebo control arm plus variability in response to treatment, meaning that power calculations based only on the placebo arm

*Corresponding author: Steven D. Edland, Division of Biostatistics, School of Public Health & Human Longevity Science, University of California San Diego, 9500 Gilman Dr, 92093-0021 La Jolla, USA; and Department of Neurosciences, School of Medicine, University of California San Diego, La Jolla, USA, E-mail: sedland@ucsd.edu. <https://orcid.org/0000-0002-1153-7335>

Yu Zhao, Division of Biostatistics, School of Public Health & Human Longevity Science, University of California San Diego, 9500 Gilman Dr, 92093-0021 La Jolla, USA

covariance structure may be anticonservative. To our knowledge, this is the first presentation of power formulas for the mixed effects model with random slopes and intercepts that accommodates differences in model variance parameters across groups. We note that a substantial literature describes many of these features for mixed model repeated measures analyses assuming compound symmetric or autoregressive covariance of repeated measures [1, 3–5]. While compound symmetric and autoregressive covariance structures are mathematically more tractable, in our experience these models are not appropriate for repeated measures of chronic progressive conditions. We demonstrate by example that compound symmetric and autoregressive covariance structures typically are not appropriate for modeling chronic progressive conditions. In the interest of clarity, in this paper we focus exclusively on the model with covariance structure imposed by random slopes and intercepts most appropriate for chronic progressive outcome measures.

2 Background, the mixed effects model

The parameterization of the mixed effects model with random slopes and intercepts used in this derivation is the familiar Laird and Ware mixed effects model parameterization with estimation and hypothesis testing by restricted maximum likelihood (REML). We use the notation of [6] to represent within group longitudinal observations y_i on subject i as

$$y_i = X_i\alpha + Z_i b_i + e_i \quad (1)$$

where α are the fixed effect intercept and slope describing the mean longitudinal trajectory, $b_i \sim \mathcal{N}(0, D)$ are random, subject-specific intercepts and slopes, and $e_i \sim \mathcal{N}(0, R_i)$ is residual variation about the individual trajectories. When convenient, we will represent the elements of D as $\begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0, b_1} \\ \sigma_{b_0, b_1} & \sigma_{b_1}^2 \end{pmatrix}$. In the derivation below, $X_i = Z_i$ are subject specific design matrices composed of a column of ones and a column of times at which measurements y_i were made. To simplify presentation we maintain large sample normality assumptions in all that follows, and we do not consider covariates beyond t_i . Consistent with prior literature [2, 3], we assume that data are missing at random and that the covariance parameters are known.

Ref. [7] showed that $V(\hat{\alpha})$, the asymptotic variance of maximum likelihood estimates of α , is independent of $\hat{\alpha}$ and derived its value. Under model (1), y is normally distributed with mean $X\alpha$ and variance-covariance V . The likelihood function is

$$l = (2\pi)^{-\frac{1}{2}n} |V|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y - X\alpha)' V^{-1} (y - X\alpha)\right) \quad (2)$$

The log likelihood, apart from a constant is

$$L = -\frac{1}{2} \log |V| - \frac{1}{2} (y - X\alpha)' V^{-1} (y - X\alpha) \quad (3)$$

By the \sqrt{n} -consistency and asymptotic efficiency of MLE, $\hat{\alpha}$ the maximum likelihood estimate of α follows

$$\sqrt{n}(\hat{\alpha} - \alpha) \xrightarrow{d} \mathcal{N}(0, I^{-1}(\alpha)) \quad (4)$$

where $I(\alpha)$ is the information matrix which equals to $\mathbb{E}\left(\frac{\partial^2 L}{\partial \alpha_i \partial \alpha_k}\right)$. For the log likelihood (3), after taking the partial derivative and expectation,

$$I(\alpha) = X' V^{-1} X \quad (5)$$

Thus the asymptotic variance of $\hat{\alpha}$ is

$$V(\hat{\alpha}) = (X' V^{-1} X)^{-1} \quad (6)$$

We can further simplify this as

$$\begin{aligned}
 V(\hat{\alpha}) &= (X'V^{-1}X)^{-1} \\
 &= \left((X'_1, X'_2, \dots, X'_n) \begin{pmatrix} V_1^{-1} & 0 & 0 & 0 \\ 0 & V_2^{-1} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & V_n^{-1} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \right)^{-1} \\
 &= \left(\sum_i (X'_i V_i^{-1} X_i) \right)^{-1}
 \end{aligned} \tag{7}$$

where

$$V_i = \text{Cov}(y_i) = Z_i D Z_i' + R_i \tag{8}$$

In particular, the lower right diagonal of $V(\hat{\alpha})$ is the variance of the mean slope estimate which is required for sample size formulas to power clinical trials comparing mean slope in treatment versus control. The components of $V(\hat{\alpha})$ can be estimated by REML [6].

Two specific cases of Eq. (7) are useful for illustrative purposes. If we are dealing with balanced data, then X_i and V_i are constant across subjects, and Eq. (7) reduces to simply

$$V(\hat{\alpha}) = (nX'_i V_i^{-1} X_i)^{-1} \tag{9}$$

A similar clinical trial with missing observations due to missed clinical exams or study subject dropout would not have constant V_i and X_i , but instead would have a finite set of design and variance matrix pairs. Letting k index this set, the variance of the fixed effect estimates for a clinical trial with missing data is then equal to

$$V(\hat{\alpha}) = \left(\sum_k n_k (X'_k V_k^{-1} X_k) \right)^{-1} = \left(n \sum_k p_k (X'_k V_k^{-1} X_k) \right)^{-1} \tag{10}$$

where the n_k are counts of subjects in each set and sum to n , and $p_k = n_k/n$.

3 Power formulas derived

3.1 Power formula, balanced design with no dropout

For the balanced design with no dropout, standard power formulas apply. E.g., for equal allocation to arms, sample size to detect a difference in mean slope Δ between treatment and control is

$$N/Arm = 2(z_{\alpha/2} + z_{\beta})^2 \left[(X'_i V_i^{-1} X_i)^{-1} \right]_{22} / \Delta^2 \tag{11}$$

This formula can be used given an estimate of $V_i = \text{Cov}(y_i)$ obtained from pilot data or a previously completed trial of comparable design.

A more generally applicable formula can be derived given the usual assumption of independent residual error ($R_i = \sigma_e^2 I$). Under this assumption, it can be shown (Appendix A) that $[(X'_i V_i^{-1} X_i)^{-1}]_{22} = (\sigma_{b_1}^2 + \sigma_e^2 / \sum (t_j - \bar{t})^2)$ [8], and Eq. (11) reduces to

$$N/Arm = 2(z_{\alpha/2} + z_{\beta})^2 (\sigma_{b_1}^2 + \sigma_e^2 / \sum (t_j - \bar{t})^2) / \Delta^2 \tag{12}$$

where $\sum (t_j - \bar{t})^2$ is the sum over the measurement time vector $t = (t_1, t_2, \dots, t_m)'$ of the squared differences t_j minus mean time.

Equation (12) is more generally applicable because it only requires estimates of σ_ϵ^2 and $\sigma_{b_1}^2$, which can be obtained by REML fit to longitudinal pilot data of arbitrary design. That is, future studies can be powered using prior study data that do not necessarily have the same duration or interval between follow-up as the planned future study [9]. Equation (12) also provides a heuristic illustration of the influence of study design on power – longer trials or trials with more longitudinal observations increase power by reducing the influence of σ_ϵ^2 on overall variance.

3.2 Power formula, balanced design with dropout

Another important example, following Lu et al., is the case of study subject dropout during a cohort study or clinical trial, also referred to as study subject attrition (SSA). SSA implies a subset of the dropout patterns indexed by k in Eq. (10), restricting to the $m - 1$ longitudinal dropout patterns composed of subjects whose last visit is at t_k , $k = 2$ through m inclusive. Given the independent residual errors assumption and equal allocation to arms, under SSA the sample size is calculated by

$$N/Arm = 2(z_{\alpha/2} + z_\beta)^2 \left(\left[\left(\sum p_k (X_k' V_k^{-1} X_k) \right)^{-1} \right]_{22} / \Delta^2 \right) \quad (13)$$

where the sum is over the $m - 1$ dropout patterns defined by SSA, $p_k (X_k' V_k^{-1} X_k)$ are as in Eq. (10), and V_k are matrices with off diagonal elements u, v equal to $\sigma_{b_0}^2 + (t_u + t_v)\sigma_{b_0, b_1} + t_u t_v \sigma_{b_1}^2$ and diagonal elements u, u equal to $\sigma_{b_0}^2 + 2t_u \sigma_{b_0, b_1} + t_u^2 \sigma_{b_1}^2 + \sigma_\epsilon^2$. As before, the parameters $\sigma_{b_0}^2$, σ_{b_0, b_1} , and $\sigma_{b_1}^2$ of D and the residual error variance σ_ϵ^2 are estimated by REML fit to representative prior longitudinal data.

Power formulas accommodating study subject attrition such as Eq. (13) and [1] are technically anti-conservative because they ignore information lost by the occasional missed interim visit, although this bias is typically small. If missing interim visit data is a concern, then applying Eq. (13) over all sets of missing data patterns will ensure true nominal type I error rates are maintained.

3.3 Power formula, unequal allocation, unequal study subject attrition, and unequal variance across groups

Formulas (12) through (13) assume that variance parameters and study subject attrition rates are the same in the two groups being compared and the number of subjects in each group is equal. We may require a formula that accommodates different study subject attrition rates across groups, and/or unequal allocation to groups [1]. It would also be useful to have a formula that accommodates different variance parameters across groups. Letting $Term_1$ and $Term_2$ indicate the values $[\left(\sum p_k (X_k' V_k^{-1} X_k) \right)^{-1}]_{22}$ calculated separately for $group_1$ and $group_2$, and given the independent identically distributed residual error assumption, sample size for $group_1$ can be calculated by

$$N_{group_1} = (z_{\alpha/2} + z_\beta)^2 (Term_1 + \lambda Term_2) / \Delta^2 \quad (14)$$

where λ is the sample size ratio across groups ($N_{group_2} = N_{group_1} / \lambda$). The derivation of Eq. (14) is straightforward, and follows from the observation that the variance of the difference in fixed effects slope estimates equals the sum of the individual slope estimate variances. Factoring out $1/N_{group_1}$ from this sum leaves the quantity $(Term_1 + \lambda Term_2)$, and power as a function of N_{group_1} follows.

3.4 Modeling under the unequal variance across groups assumption

It is given that using Eq. (14) with unequal variance parameters to power a study presumes the analysis plan for the study explicitly models the covariance structure of the two groups. For most applications, including

clinical trials, σ_ϵ^2 is assumed constant across groups. Sample syntax explicitly modeling the remaining, within group random effects parameters determining the covariance structure of repeated measures is included in Appendix B.

4 Example

Given representative pilot data it is a simple matter to estimate the variance terms required for the power formulas. For example, Table 1 is the output from a mixed effect model fit to longitudinal ADAS-cog scores observed in the ADCS trial of a folic acid/B6/B12 compound to slow the progression of Alzheimer's disease [10] ($n = 330$ subjects and $m = 7$ observations per subject) using the software provided with the standard mixed effects model text *Mixed-Effects Models in S and S-PLUS* [11]. The correlation of repeated measures estimated by the random slopes and random intercepts REML model fit (Table 2) mirrors the empirical correlation calculated from the same sample data, confirming that this model well represents the covariance structure of longitudinal repeated measures of a chronic progressive condition. In contrast, the commonly assumed compound symmetric and autoregressive covariance structures are constant on the diagonals and inconsistent with these longitudinal data of a chronic progressive condition.

From Table 1, the estimated standard deviation of slopes $\hat{\sigma}_{b_1}$ is 3.964 and the estimated standard deviation of residual errors $\hat{\sigma}_\epsilon$ is 3.705 (Table 1). Assuming equal variance across arms, and using these values in Eq. (12), the sample size required to detect a 25% slowing of cognitive decline ($\Delta = 0.25 \cdot 4.06$) with 80% power and a type I error rate of 5% for an 18 month trial with observations every three months is 360 subjects/arm. For comparison, a 24 month trial with observations every three months would require 296 subjects per arm using Eq. (12). Note that it is not necessary for the design of the pilot study (i.e., the number of observations and interval between observations) to match the design of the future trial, we only require that there are sufficient pilot data to estimate the variance parameters $\sigma_{b_1}^2$ and σ_ϵ^2 .

5 Validation by computer simulation

To evaluate the performance of Eq. (12) through (14) we have performed computer simulations assuming data following the model fit obtained in the Example above. We first performed simulations assuming a clinical trial with balanced design with six post-baseline time points with no loss to follow-up and with equal variance within arms consistent with Eq. (12). Simulating a series of clinical trials with sample size from 200 to 600 subjects per arm with effect size equal to a 25% reduction in the mean rate of decline observed in placebo (25% of the mean 4.06 points per year rate of decline observed in the pilot data (Table 1)) with 10,000 simulations per sample size simulated, we found that simulated power closely tracks the power predicted by Eq. (12) (top line, Figure 1).

Table 1: Sample model fit using the R package nlme and R function lme.

```
lme(y ~ time, random = ~ time|id)
Random effects formula: ~time | subject
      StdDev  Corr
(Intercept)  7.432548 (Intr)
Time        3.964215 0.465
residual    3.705466
Fixed effects:  ADAS ~ time
      Value Std.Error
(Intercept) 17.745024 0.4321112
Time        4.057879 0.2672020
Number of observations: 2310/Number of groups: 330
```

Table 2: Correlation matrices estimated using data from the ADCS Folate/B6/B12 clinical trial. The correlation matrix imposed by a random effect model fit (RE, bottom panel) closely mirrors the empirical correlation matrix (top panel).

Empirical correlation matrix

$$\begin{pmatrix} 1 & 0.83 & 0.82 & 0.79 & 0.80 & 0.80 & 0.76 \\ 0.83 & 1 & 0.84 & 0.82 & 0.82 & 0.85 & 0.80 \\ 0.82 & 0.84 & 1 & 0.85 & 0.85 & 0.85 & 0.84 \\ 0.79 & 0.82 & 0.85 & 1 & 0.88 & 0.85 & 0.84 \\ 0.80 & 0.82 & 0.85 & 0.88 & 1 & 0.90 & 0.88 \\ 0.80 & 0.85 & 0.85 & 0.85 & 0.90 & 1 & 0.89 \\ 0.76 & 0.80 & 0.84 & 0.84 & 0.88 & 0.89 & 1 \end{pmatrix}$$

Correlation matrix estimated assuming RE

$$\begin{pmatrix} 1 & 0.81 & 0.80 & 0.80 & 0.78 & 0.77 & 0.76 \\ 0.81 & 1 & 0.83 & 0.83 & 0.82 & 0.81 & 0.81 \\ 0.80 & 0.83 & 1 & 0.85 & 0.85 & 0.85 & 0.84 \\ 0.80 & 0.83 & 0.85 & 1 & 0.87 & 0.87 & 0.87 \\ 0.78 & 0.82 & 0.85 & 0.87 & 1 & 0.88 & 0.89 \\ 0.77 & 0.81 & 0.85 & 0.87 & 0.88 & 1 & 0.90 \\ 0.76 & 0.81 & 0.84 & 0.87 & 0.89 & 0.90 & 1 \end{pmatrix}$$

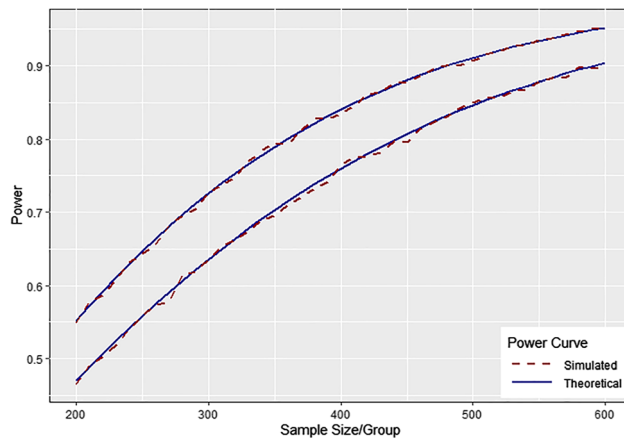


Figure 1: Theoretical power curves versus power estimated by computer simulation given no study subject attrition (top curve) and give 5% attrition per follow-up visit (bottom curve) (10,000 simulations per sample size, two-sided test, type I error $\alpha = 0.05$).

To validate the power formula for data with study subject attrition described in Eq. (13), we simulated data under equivalent conditions, except that for each simulation we randomly dropped 5% of the initial sample from each arm at t_2 through t_7 . We similarly found that simulated power closely tracks the power predicted by Eq. (13) power formula (bottom line, Figure 1). Study power decreases when there is study subject attrition (Figure 1).

To validate the power formula for data with unequal allocation to groups described in Eq. (14), we simulated data with 5% study subject attrition at each follow-up visit as above, but let the allocation ratio λ vary from one to two. Simulated power closely tracks the power predicted by Eq. (14) power formula (Figure 2). Predictably [12, 13], power is maximized when λ equals one, and declines as the allocation ratio deviates from one (Figure 2).

To validate Eq. (14) power formula when covariance structures differ across groups, we simulated data as done in the top line of Figure 1, but increased σ_{b_1} by 50% in one of the groups. Simulated power closely tracks the power predicted by Eq. (14) power formula (Figure 3). The top line from Figure 1 is included in Figure 3 for reference. Figure 3 illustrates the potential for anticonservative power calculations in the clinical trial setting when variance parameters used in power calculations are informed by prior placebo arm data and assumed to be constant across arms.

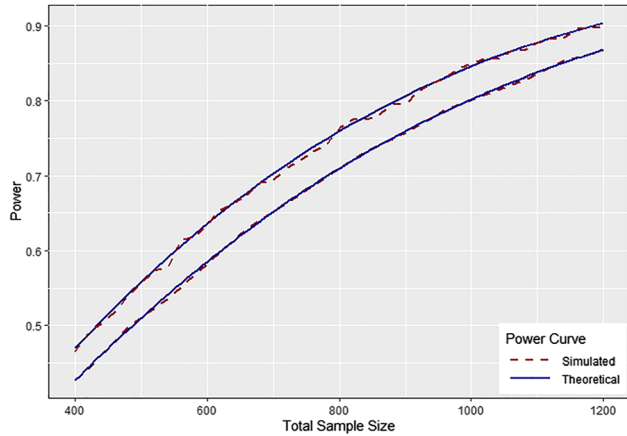


Figure 2: Theoretical powers curve versus power estimated by computer simulation given 5% study subject attrition per visit, and allocation ratio $\lambda = 1$ (top curve) and $\lambda = 2$ (bottom curve) (10,000 simulations per sample size, two-sided test, type I error $\alpha = 0.05$).

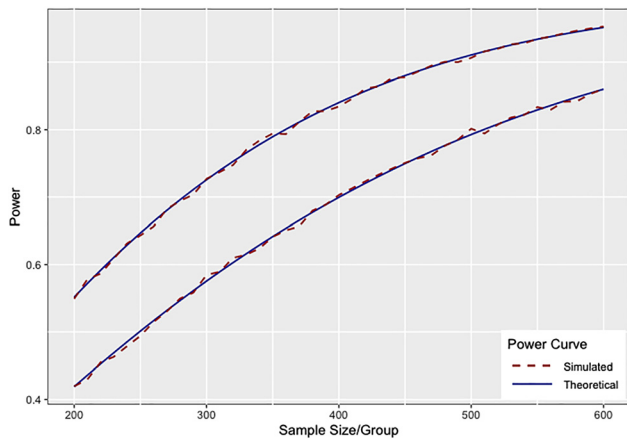


Figure 3: Theoretical power curves versus power estimated by computer simulation given equal variance of random slopes (top line) and given σ_{b_1} is increased by 50% in one of the groups (bottom line) (10,000 simulations per sample size, two-sided test, type I error $\alpha = 0.05$).

6 Discussion

There are limitations to the Laird and Ware model as parameterized in Eq. (1), because this model depends on the assumption that mean trajectories are linear as a function of time. This assumption may be violated, particularly in clinical trials of treatments with potential acute treatments effect beyond simple alteration of rate of progress of disease. In this circumstance mixed model repeated measure analysis [1] or model robust alternatives such as generalized estimating equations [14] would be preferred. In our experience the linearity assumption is often appropriate for chronic progression conditions, especially when the interval of observation under study is small relative to the full trajectory of disease.

We further note that the formulas presented here assume variance parameters are known, as is typical of the power formula literature [1–3, 5, 15]. However, variance parameters may be uncertain if sample size in pilot studies used to estimate the variance parameters is small or if pilot data are not perfectly representative of the future investigation being powered. There is a literature on characterizing power when variance parameter estimates are uncertain (e.g. [16]). However, these methods apply to narrow applications that do not include random effects models. We recommend sensitivity analyses using a range of plausible variance parameters to ensure that planned future investigations are adequately powered. If the prior data informing power calculations are available, sensitivity analyses may be informed by bootstrap estimates of the uncertainty of variance parameter estimates (e.g., [17]). We have also used computer simulations to explore the adequacy of pilot study sample size to inform future trials in other applications [18].

The formulas derived here are useful for determining the relative efficiency of different study designs using the mixed effects model to test for differences in mean rate of change between groups. We have described how efficiency can vary by the number and interval between observations, the study subject attrition rate, the allocation ratio, and by differences in variance parameters between groups. Increasing the length of observation or number of observations increases statistical power, although with diminishing returns depending on the magnitude of residual error variance of the outcome measure under study (see Eq. (12)). Study subject attrition can also meaningfully impact statistical power and should be accounted for in study design (see Eq. (13) and, e.g., Figure 1).

Regarding recruitment allocation ratios, if all other conditions are equal across groups, then altering the allocation ratio from one-to-one reduces statistical power for given study sample size [12]. Altering the allocation ratio has been proposed to improve statistical power when there is differential attrition rates across clinical trial arms [1]. More commonly, allocation ratios are altered to increase the probability of randomization to the active treatment in the hope of increasing clinical trial recruitment rates. While this approach may increase recruitment rates, it also implies more subjects will have to be recruited to achieve target statistical power, and trade-offs between clinic trial cost and time to completion should be considered carefully when planning a trial with unequal randomization to arms [13].

Finally, we describe how statistical power depends on variance parameters which may vary across groups (Eq. (14)). This consideration is typically overlooked, but may be especially relevant to clinical trials, where rate of progression in the active treatment arm is a function of both underlying variability in rate of progression and variability in response to treatment. Given that response to treatment is unlikely to be constant across subjects, we can anticipate that the variance of random slopes in the treatment arm will be larger than variance in the control arm if there is a treatment effect. Hence, power calculations based only on the covariance within placebo data will be anticonservative. Typically pilot data for clinical trials are from placebo arm data of a previous trial or registry trial with no treatment arm. A conservative power calculation assumption under these circumstances would be to use an inflation factor for $\sigma_{b_1}^2$ within the treatment arm in (14) to be more likely to achieve nominal power in the planned trial.

Formulas (12), (13), and (14) are implemented in the R package *longpower* [19], and will be useful tools for planning future cohort studies and clinical trials as well as for comparing the influence of the many factors affecting the efficiency of such investigations. Areas of additional research include modifying power calculation methods in anticipation of evolving guidelines on statistical analysis plans for clinical trials in the presence of missing not at random data [20], and generalizing power formulas to more directly address the stochastic nature of covariance parameter estimates typically used in practice.

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This work was supported by The Shiley Marcos Alzheimer's Disease Research Center and National Institute on Aging (AG049810, AG062429, AG066088).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

Appendix A

To derive the variance term in Eq. (12), we need to find the bottom right corner of $(X_i'V_i^{-1}X_i)^{-1}$. As derived by [21],

$$\begin{aligned} V_i^{-1} &= (X_iDX_i^T + \sigma_e^2I)^{-1} \\ &= \sigma_e^{-2}I - \sigma_e^{-2}X_i(X_i'X_i)^{-1}X_i' + X_i(X_i'X_i)^{-1}\left(\sigma_e^2(X_i'X_i)^{-1} + D\right)^{-1}(X_i'X_i)^{-1}X_i' \end{aligned}$$

Substituting and collecting terms,

$$(X_i' V_i^{-1} X_i)^{-1} = \sigma_e^2 (X_i' X_i)^{-1} + D$$

and

$$\left[(X_i' V_i^{-1} X_i)^{-1} \right]_{22} = \sigma_e^2 / \sum (t_j - \bar{t})^2 + \sigma_{b_1}^2$$

Appendix B

The random effects model with random slopes and intercepts can be performed with the `lmer` function within the R package `lmerTest` [22]. To test for differences in slopes between groups under the assumption of equal covariance structure in the two groups, the `lmer` model call is

```
lmer (Y ~ GROUP * TIME + (TIME | ID))
```

where `ID` indexes individual subjects, `GROUP` is a 0, 1 variable indicating placebo (0) and active treatment (1), and `TIME` are times of repeated observations on the dependent variable `Y`.

To test for differences in slopes between groups under the assumption of unequal covariance structure in the two groups, as implemented in power Formula (14), the `lmer` model call is

```
TIME_0 <- ifelse (GROUP == 0, TIME, 0)
TIME_1 <- ifelse (GROUP == 1, TIME, 0)
lmer (Y ~ TIME * GROUP + (TIME_0 | ID) + (TIME_1 | ID))
```

References

1. Lu K, Luo X, Chen P-Y. Sample size estimation for repeated measures analysis in randomized clinical trials with missing data. *Int J Biostat* 2008;4:9.
2. Hedeker D, Gibbons RD, Wateraux C. Sample size estimation for longitudinal designs with attrition: comparing time-related contrasts between two groups. *J Educ Behav Stat* 1999;24:70–93.
3. Wang C, Hall CB, Kim M. A comparison of power analysis methods for evaluating effects of a predictor on slopes in longitudinal designs with missing data. *Stat Methods Med Res* 2015;24. <https://doi.org/10.1177/0962280212437452>.
4. Muller KE, LaVange LM, Ramey SL, Ramey CT. Power calculations for general linear multivariate models including repeated measures applications. *J Am Stat Assoc* 1992;87:1209–26.
5. Rochon J. Sample size calculations for two-group repeated-measures experiments. *Biometrics* 1991;47:1383–98.
6. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982;38:963–74.
7. Searle SR. Large sample variances of maximum likelihood estimators of variance components using unbalanced data. *Biometrics* 1970;26:505–24.
8. Snijders TAB, Bosker RJ. Standard errors and sample sizes for two-level research. *J Educ Stat* 1993;18:237–59.
9. Ard MC, Edland SD. Power calculations for clinical trials in Alzheimer's disease. *J Alzheim Dis* 2011;21:369–77.
10. Aisen PS, Schneider LS, Sano M, Diaz-Arrastia R, van Dyck CH, Weiner MF, et al. High-dose B vitamin supplementation and cognitive decline in Alzheimer disease: a randomized controlled trial. *J Am Med Assoc* 2008;300:1774–83.
11. Pinheiro JC, Bates DM. *Multivariate statistical modelling based on generalized linear models*. New York: Springer-Verlag; 2000.
12. Meinert CL. *Clinical trials design, conduct, and analysis*. New York: Oxford University Press; 1986.
13. Vozdolska R, Sano M, Aisen P, Edland SD. The net effect of alternative allocation ratios on recruitment time and trial cost. *Clin Trials* 2009;6:126–32.
14. Jung S, Ahn C. Sample size estimation for gee method for comparing slopes in repeated measurements data. *Stat Med* 2003; 30:1305–15.
15. Tu XM, Zhang J, Kowalski J, Shults J, Feng C, Sun W, et al. Power analyses for longitudinal study designs with missing data. *Stat Med* 2007;26:2958–81.
16. Browne RH. On the use of a pilot sample for sample size determination. *Stat Med* 1995;14:1933–40.
17. McEvoy LK, Edland SD, Holland D, Hagler DJ, Roddey JC, Fennema-Notestine C, et al. Neuroimaging enrichment strategy for secondary prevention trials in Alzheimer disease. *Alzheimer Dis Assoc Disord* 2010;24:269–77.

18. Edland SD, Ard MC, Li W, Jiang L. Design of pilot studies to inform the construction of composite outcome measures. *Alzheimer's Dementia* 2017;3:213–8.
19. Donohue MC. Longpower: Power and sample size calculations for longitudinal data; 2019.
20. ICH Working Group. Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials E9(R1), European Medical Agency. R package version 1.0–19; 2019. Available from: https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf.
21. Leeuw JD, Kreft I. Random coefficient models for multilevel analysis. *J Educ Stat* 1986;11:57–85.
22. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Software* 2017;82: 1–26.