

Featured Article

A simulation study comparing slope model with mixed-model repeated measure to assess cognitive data in clinical trials of Alzheimer's disease

Yun-Fei Chen^{a,*}, Xiao Ni^a, Adam S. Fleisher^a, Wei Zhou^a, Paul Aisen^b, Richard Mohs^a

^aDepartment of Neuroscience, Eli Lilly and Co., Indianapolis, IN, USA

^bDepartment of Neurology, University of Southern California, Los Angeles, CA, USA

Abstract

Introduction: In clinical trials of Alzheimer's disease, a mixed-model repeated measure approach often serves as the primary analysis when evaluating disease progression; a slope model may be secondary.

Methods: Longitudinal change from baseline (14-item version of Alzheimer's Disease Assessment Scale–Cognitive Subscale) was simulated for treatment/placebo from multivariate normal distributions with the variance-covariance matrix estimated from solanezumab trial data. Type I error, power, and bias were based on 18-month treatment contrast. Sample sizes included 500 and 1000 patients/arm.

Results: The slope model was more powerful in most scenarios. Mixed-model repeated measure was relatively unbiased in parameter estimation. The slope model yielded unbiased estimates whenever the underlying trajectory was not detectably different from linear. Both methods led to similar type I error.

Discussion: In clinical trials of Alzheimer's disease, mixed-model repeated measure analysis with relaxed assumptions on disease progression seems to be preferred. The slope model might be more powerful if the trajectory has little departure from linearity.

© 2018 Eli Lilly and Company. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords:

Clinical trials efficiency; Mixed-effects models; Mixed-model repeated measure (MMRM); Quantitative review

1. Background

Progression of a chronic disease such as Alzheimer's disease (AD), by definition, involves kinetics or dynamics of cognitive change relative to time, or the trajectory and shape

Y.-F.C., A.S.F., and W.Z. are full-time employees and minor shareholders of Eli Lilly and Company. X.N. was an employee of Eli Lilly and Company at the time of this research and is currently an employee of Novartis. P.A. has served as a consultant to the following companies: NeuroPhage, Elan, Eisai, Bristol-Myers Squibb, Eli Lilly and Company, Merck, Roche, Amgen, Genentech, Abbott, Pfizer, Novartis, AstraZeneca, Janssen, Medivation, Ichor, Toyama, Lundbeck, Biogen Idec, iPerian, Probiobdrug, Somaxon, Biotie, Cardeus, Anavex, AbbVie, and Cohbar. P.A. receives research support from Eli Lilly and Company, NIH (NIA U01-AG10483 [PI], NIA U01-AG024904 [Coordinating Center Director], NIA R01-AG030048 [PI], and R01-AG16381 [Co-I]). R.M. was an employee of Eli Lilly and Company at the time of this research and retired in 2015.

*Corresponding author. Tel.: +1-317-292-3994; Fax: +1-317-276-7100.

E-mail address: chen_yun-fei@lilly.com

<https://doi.org/10.1016/j.trci.2017.12.002>

2352-8737/© 2018 Eli Lilly and Company. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

of a curve. In clinical trials of drugs intended for the treatment of AD, comparing mean changes (baseline to endpoint) between treatment groups using a mixed-model repeated measure (MMRM) approach often serves as the primary analysis.

The MMRM analysis is a “semiparametric” approach, which treats time as a factor, or a categorical variable, and estimates the mean change from baseline in the outcome in each group treating baseline performance as a covariate [1]. The primary efficacy analysis is pivoted against a single endpoint (e.g., 18 months). Mallinckrodt et al. demonstrated that mixed-effects models, particularly the MMRM with unstructured mean and within-subject error correlation, provide more accurate estimates of treatment effect and its standard error than last observation carried forward analysis of covariance when data are missing at random [2]. Slope model, in contrast, assumes a linear progression model and may often serve as an alternative secondary analysis, which compares the slopes between treatment groups and treats time as a continuous variable.

The objective of this simulation study is to investigate the fixed (treatment) effect via MMRM and slope model using the same unstructured variance-covariance matrix. Herein, to make a fair comparison with the MMRM model, we use the term “slope model” to refer to a linear mixed-effects model that is linear in time without adding random slope or intercept in the model. The slope model described in this analysis uses a single parameter and is based on a simple and intuitive parametric trajectory model that can capture dynamics based on data from multiple visits. Several related types of slope models could also be considered, and similar inferences could be obtained by modeling unadjusted (not change from baseline) scores, using a model with patient-level random effects for slope and intercept. When changes from baseline are modeled, baseline Mini-Mental Status Examination and/or baseline Alzheimer’s Disease Assessment Scale–Cognitive Subscale (ADAS-Cog) are important covariates for rate of progression [3].

In AD, it is often hypothesized that for a disease-modifying drug assumed to slow disease progression, the treatment group compared with the placebo group should shift the slope of decline on a given clinical outcome. It is important to note, however, that this is only one hypothesis among many regarding the accrual of treatment benefit that might be observed under an efficacious disease-modifying therapy. The gradual accrual of apparent treatment benefit (e.g., as would arise with diverging slopes) may be suggestive of permanent benefit, but continued and gradual accrual of apparent treatment benefit is neither sufficient nor necessary to establish the permanency of the benefit [4]. With this in mind, it can be interpreted that a slope model may provide a more intuitive and clinically meaningful way of demonstrating a disease-modifying effect than MMRM. Although MMRM analysis is an approach accepted by regulatory agencies to examine treatment efficacy, the slope model is required by the European Medicines Agency and has been proposed as an alternative approach, given its usefulness in consideration of possible disease-modifying effects [5].

Typically, a disease-modifying intervention is considered to be one that can slow disease progression by altering the neurobiology of the disease. While AD placebo trajectories are generally nonlinear because of an evident placebo effect occurring in the first 12 weeks or a finer time resolution assessment (e.g., every 6 weeks) [3], the disease trajectory often appears linear after the 1- to 2-year time course of initial improvement [6]. An expert group advocated the use of longer trials for disease modification coupled with slope models and biomarkers, specifically recommending that trials of 18-month duration be used [7]. This group also suggested that slope models be used from the perspective that diverging slopes of decline between drug and placebo groups can provide evidence for disease modification. The merits of slope model include it being a simpler model with a clear clinical interpretation, pertinent to the disease progression and modification concept, and potential efficiency gain. As is typical of more parsimonious models, a more favorable bias-variance trade-off may poten-

tially be obtained, whereby the negative consequences of increased model bias are offset by the benefit of stabilized (reduced variance) estimation. One risk of the slope model is an incorrect model specification due to a strong linear assumption that could lead to a bias in estimation.

Due to the nature of AD, clinical trials are often plagued with high rates of missing data and highly variable clinical assessments underscoring the importance of efficient study design and analysis. In a chronic condition like AD, a linear model for progression is probably not an unreasonable approximation within a short window. Given the value of a slope model as a secondary analysis, it would be valuable to benchmark against the more general MMRM analysis and evaluate the trade-off, as well as the risk, of bias under varying degrees of departure from linearity. Here, we conducted a simulation study to compare the slope model and MMRM analysis based on various scenarios to better understand the performance of each method.

2. Methods

2.1. Study design

The design of EXPEDITION2 (NCT00904683) has been described previously [8]. Briefly, EXPEDITION2 was a multinational, randomized, double-blind, placebo-controlled, phase 3 study of solanezumab, an immunoglobulin G subclass 1 anti-amyloid monoclonal antibody that binds to the mid-domain of the amyloid- β peptide and is thought to increase clearance of soluble amyloid- β . Solanezumab was given intravenously 400 mg every 4 weeks into outpatients with mild-to-moderate AD dementia. Patients were at least 55 years of age and met criteria for probable AD dementia based on National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer’s Disease and Related Disorders Association criteria [9]. Patients with Mini-Mental Status Examination [10] scores of 16 to 26 were allowed to participate. Mild AD dementia was defined as screening visit Mini-Mental Status Examination scores of 20 to 26; moderate AD dementia was defined as screening visit scores of 16 to 19. Randomization to treatment was stratified by AD severity to ensure a balance of treatment assignment within both the mild and moderate AD dementia patient groups. Patients were allowed to continue treatment with stable doses of standard-of-care AD treatments (e.g., acetylcholinesterase inhibitors and memantine) throughout the studies.

Institutional review boards at all participating sites approved the study. The study was conducted in accordance with ethical principles of Good Clinical Practice and the Declaration of Helsinki and its guidelines.

2.2. Statistical methods

Simulations were performed to compare the statistical properties of the slope model and MMRM analysis. Longitudinal change from baseline of the ADAS-Cog 14 for six

postrandomization visits (up to 18 months) were simulated for both treatment and placebo groups from multivariate normal distributions with the variance-covariance matrix estimated from the solanezumab EXPEDITION2 trial data.

To assess the performance of both methods under varying degrees of departure from linear trajectories for disease progression, various models for the mean profiles were used to generate the longitudinal data. Table 1 details the experimental factors that comprised these models. Sample sizes included 500 and 1000 patients per arm. Linear, quadratic, and integrated two-component prediction models [11] with varying shape parameters corresponding to gradual departure from linearity were simulated (Fig. 1) based on Monte Carlo sampling of 500. Data simulation was executed with no missing data and a 20% dropout. The integrated two-component prediction model equation used was

$$y_{ijk} = (\theta_i + s_{ij} + \varepsilon_{ijk}) \frac{1 - e^{p_i t_{ijk}}}{1 - e^{p_i d}}$$

In this equation, y_{ijk} represents an observation (change from baseline) from subject j in the treatment arm i at time k . Parameter θ_i is the i th treatment mean effect by the end of time d , where d represents the treatment duration. Parameter p_i determines the shape of the i th treatment time course, and t_{ijk} is the time covariate. The parameter ij is an individual-level random effect, and the parameter ε_{ijk} is the residual error. Based on EXPEDITION2 data in patients with mild AD, the true between-treatment difference (treatment minus placebo) at 18 months was set to be a common

Table 1
Experimental factors involved in simulation models

Experiment factor*	Levels
Sample size/arm	<ul style="list-style-type: none"> • 500 • 1000
Placebo_18 month_change	<ul style="list-style-type: none"> • 7
Trt_18 month_change	<ul style="list-style-type: none"> • 7 (NULL case, for type 1 error) • 5.5
Variance-covariance matrix	<ul style="list-style-type: none"> • Estimated variance matrix from EXPEDITION2
Trajectory models	<ol style="list-style-type: none"> 1. Linear 2. Quadratic 3. ITP model ($P = .01$, almost linear[†]) 4. ITP model ($P = .05$) 5. ITP model ($P = .1$) 6. ITP model ($P = .2$, quite curvy) 7. ITP model ($P = .3, .4, .5$) 8. ITP model ($P = -.01$, close to linear) 9. ITP model ($P = -.05$) 10. ITP model ($P = -0.1$) 11. ITP model ($P = -.2$, quite curvy)
Missing data	<ul style="list-style-type: none"> • No missing data • 20% dropout
Monte Carlo samples	<ul style="list-style-type: none"> • 500

Abbreviations: ITP, integrated two-component prediction; Trt, treatment.

*Not a full factorial design.

[†]ITP standards for integrated two-component prediction model; P value is the shape parameter.

value of -1.5 for all disease progression models. Both treatment and placebo arms were simulated using the same shape parameter for integrated two-component prediction models.

Type I error, power, mean squared error, and bias (difference between estimation and the truth) were summarized based on the estimand of the treatment contrast at 18 months to compare the performance of the slope model as compared with MMRM. Both models included simulated ADAS-Cog 14 change from baseline data for six post-randomization visits as dependent variables, and treatment, time, and treatment and time interaction terms as independent variables, where time was considered a categorical variable in MMRM and a continuous variable for the slope model. This estimand was chosen for bias evaluation to examine longitudinal change based on previous guidance [12].

All statistical analyses were conducted using SAS, version 9.2.

3. Results

Fig. 2A–D displays model-fitted and true treatment disease progression trajectories for the placebo and treated groups over the course of 18 months with various scenarios using the slope model method. When the true trajectory was close to linear, the estimates at the end of 18 months were close to the true value for each placebo and treated group. When the true trajectory deviated from linearity, more bias was introduced as the estimates from slope model departed from the true values at the end of 18 months for both groups. Disease progression trajectories for the MMRM method (not shown) were unbiased and did not deviate from the true values.

Using the effect size and variance component estimates from the longitudinal models, clinical trials were simulated to reflect the changes observed over time in the EXPEDITION2 study. Holding the sample size constant at 500 participants per arm (Fig. 3), the slope model was more powerful than the MMRM analysis in most scenarios (70% vs. 57% for a shape parameter of -0.05 , 73% vs. 59% for a shape parameter of -0.2 , 62% vs. 59% for a shape parameter of 0.05 , and 38% vs. 59% for a shape parameter of 0.2), although the power advantage was moderate.

The MMRM analysis was relatively unbiased in parameter estimation in all scenarios, whereas the slope model yielded unbiased estimates whenever the underlying trajectory was not detectably different from linear (see the relative bias table in Fig. 3). For shape parameters ranging from -0.4 to -0.05 , the slope model was more powerful than MMRM analysis, but resulting bias increased. However, compared to MMRM, the slope model was increasingly biased when the true underlying progression departed from linearity.

The type I error rates at the nominal level, $\alpha = 0.05$, were compared between MMRM and slope model. Simulations showed that the type I error for both MMRM and slope model are comparable within a maximum deviation of 0.02 from the nominal level.

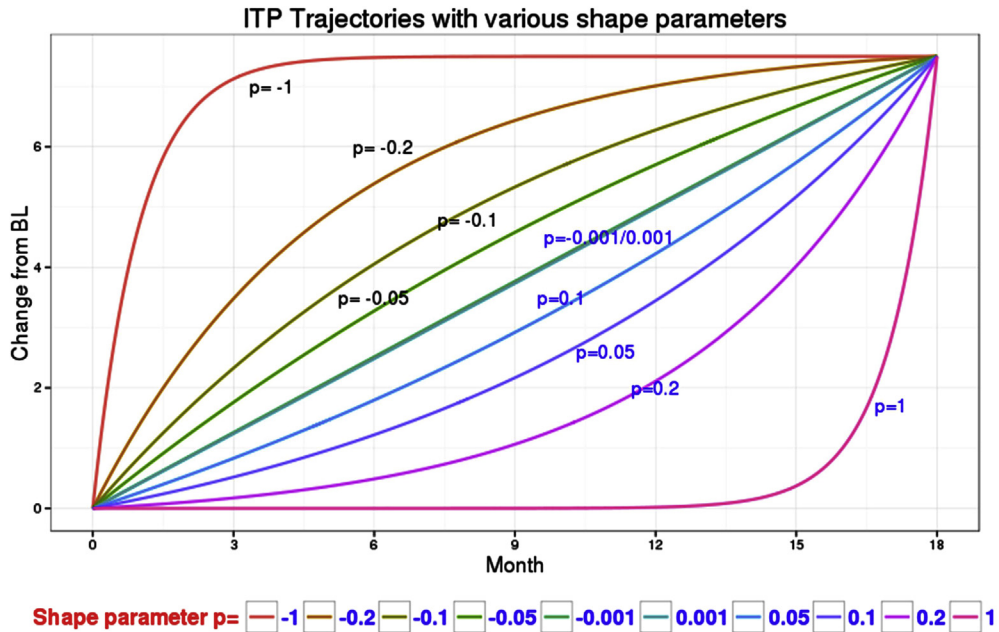


Fig. 1. ITP trajectories with various shape parameters. Abbreviations: BL, baseline; ITP, integrated two-component prediction.

While sample size increased from 500 to 1000 patients per arm, the power increased for both MMRM analysis and slope model, although the power advantage for slope model

decreased. A simulation example with the linear trajectory scenario in comparing power, relative bias, and mean squared error using 500 versus 1000 participants per arm was shown in Fig. 3.

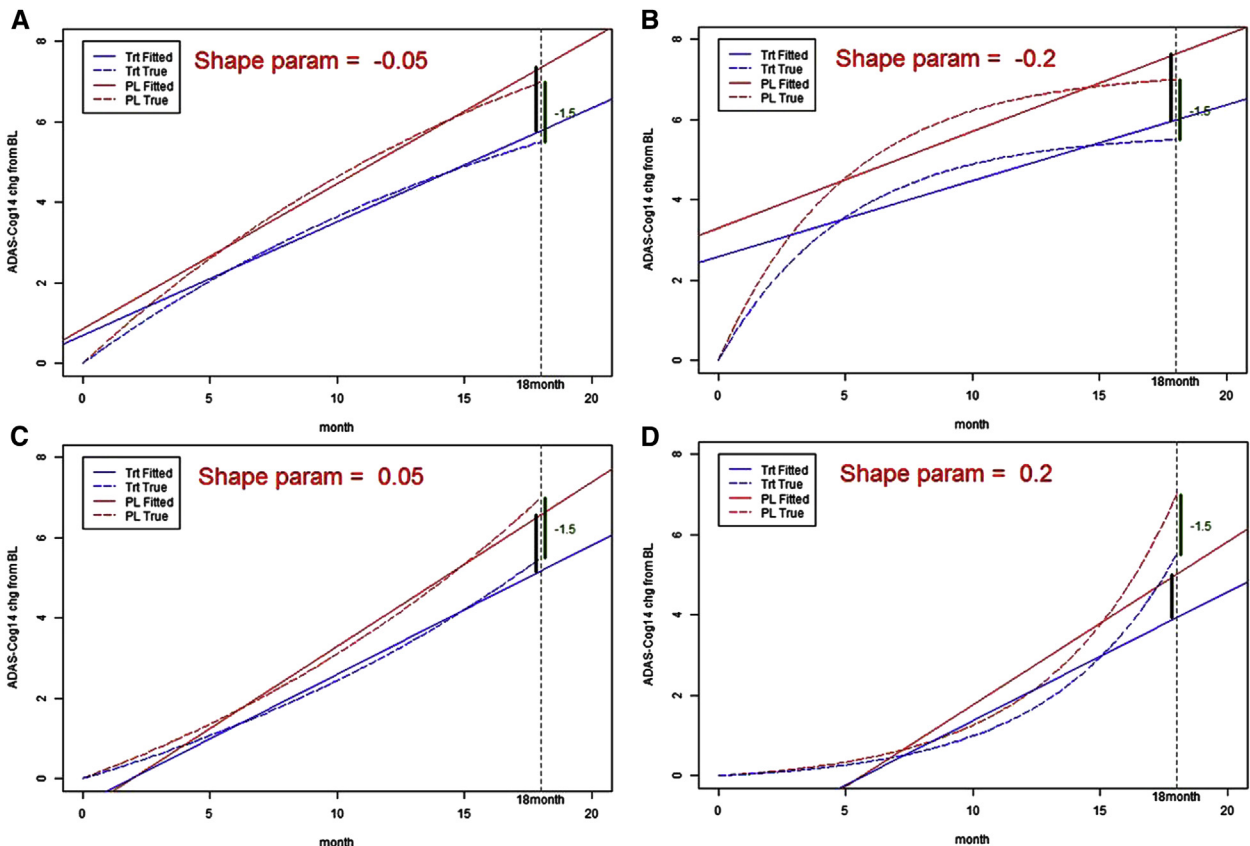


Fig. 2. (A–D) Fitted and true treatment (LY2062430 [solanezumab]) disease progression trajectories, at shape parameters of -0.05 (A), -0.2 (B), 0.05 (C), and 0.2 (D), over the course of 18 months for the slope model method. Abbreviations: ADAS-Cog 14, Alzheimer's Disease Assessment Scale–Cognitive Subscale 14; BL, baseline; Chg, change; LY, LY2062430 (solanezumab); Param, parameter; PL, placebo; Trt, treatment.

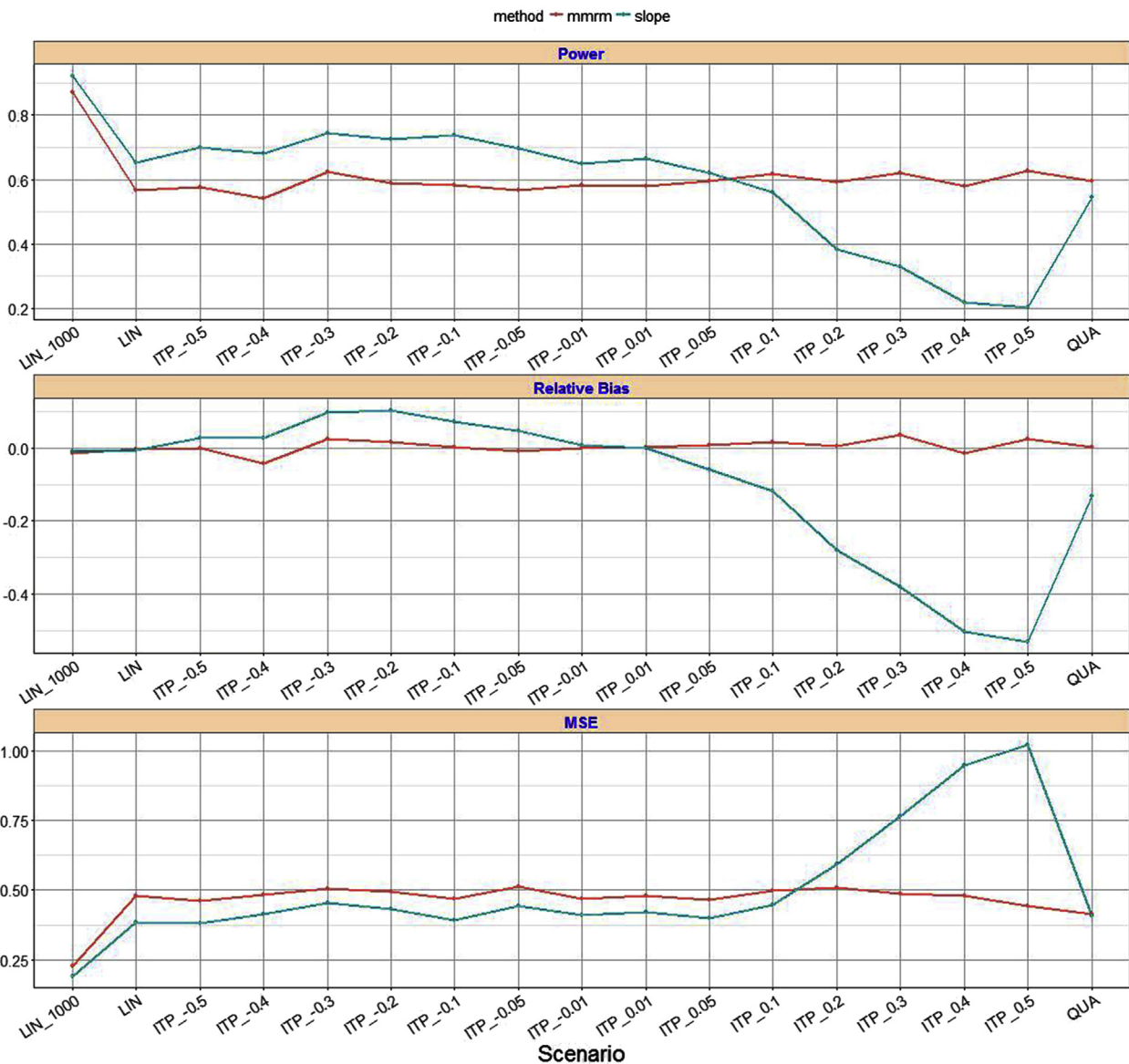


Fig. 3. Power and relative bias comparison between MMRM and slope model. Abbreviations: ITP, integrated two-component prediction; LIN, linear; MMRM, mixed-model repeated measure; MSE, mean squared error; QUA, quadratic.

In comparison between no missing data and 20% dropout data using 500 patients per arm, the power decreased moderately with 20% dropout for both MMRM and slope model (see Fig. 4 for selected scenarios). MMRM remained stable in power performance across various scenarios, whereas slope model varies the power performance. Similarly, the relative bias and mean squared error increased moderately for both MMRM analysis and slope model with 20% dropout data.

4. Discussion

Type I error rates and power based on the simulations were evaluated to determine the optimal method of analysis. Simulations showed that both MMRM and slope model led to similar type I error around the nominal level, but that slope model had a moderate power advantage over MMRM analysis.

The slope model is based on a stronger model assumption and therefore suffered increased bias when the true disease progression model is nonlinear. The MMRM analysis with relaxed assumptions on disease progression seems to be a preferred approach with more general applicability, whereas the slope model is more powerful if the true model is linear or near linear.

Reporting outcomes in terms of slope (rate of decline) may be more appropriate for emphasizing long-term outcomes in degenerative diseases, since slowing rate of decline can be clinically meaningful to patients and physicians. However, the specification of the statistical model for the slope model is not always straightforward. The natural progression of AD may be approximated with a linear model typically within a period of 18 months (e.g., length of clinical trial). However, it is still unclear whether this linearity assumption remains true in a clinical trial in which a potentially disease-

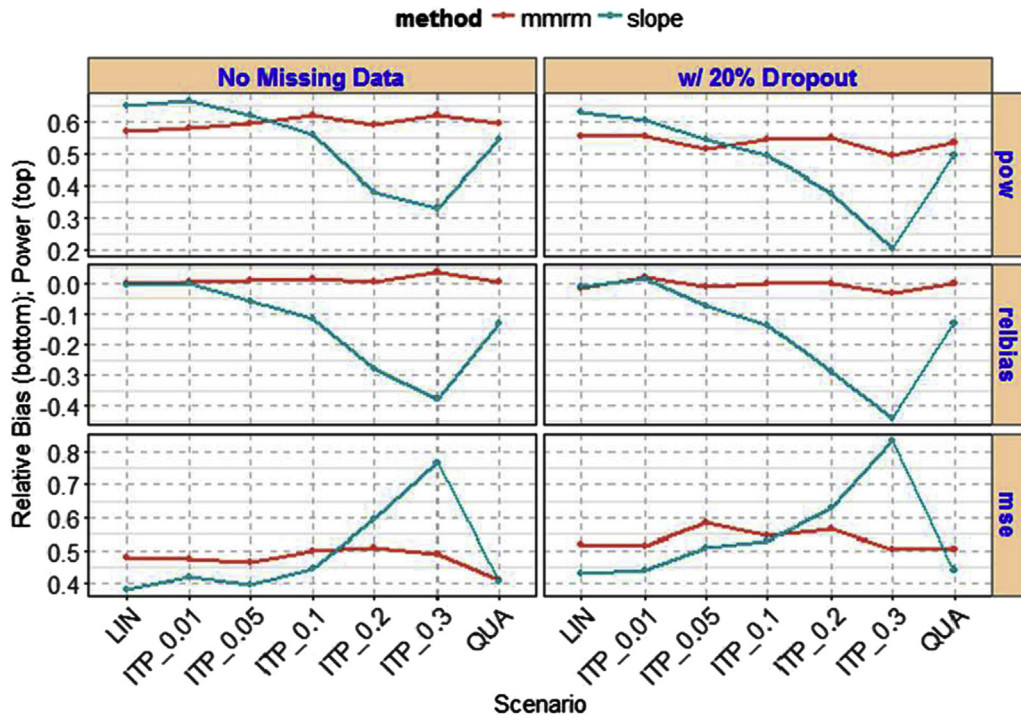


Fig. 4. Comparison of missing data and 20% dropout between MMRM and slope model. Abbreviations: ITP, integrated two-component prediction; LIN, linear; MMRM, mixed-model repeated measure; MSE, mean squared error; pow = power (comparison); QUA, quadratic; relbias, relative bias (comparison).

modifying treatment effect is also a variable, and it is unknown whether the linearity assumption is constant over the course of treatment [13]. Importantly, there are both individual and aggregate aspects to the correctness or incorrectness of the linearity assumption. The simulation model considered here does not reflect the boundedness of the ADAS-Cog instrument, as some patients will progress toward the ceiling of the instru-

ment (90 for ADAS-Cog 14) over an 18-month duration in a population that is mild-to-moderate at baseline. This is likely to influence the statistical properties being assessed.

An analysis roadmap (Fig. 5) may be a useful tool to determine which type of analysis to use when examining clinical data, to perform a model assumption check, to adapt a final analysis plan based on interim data, and to use separate trial

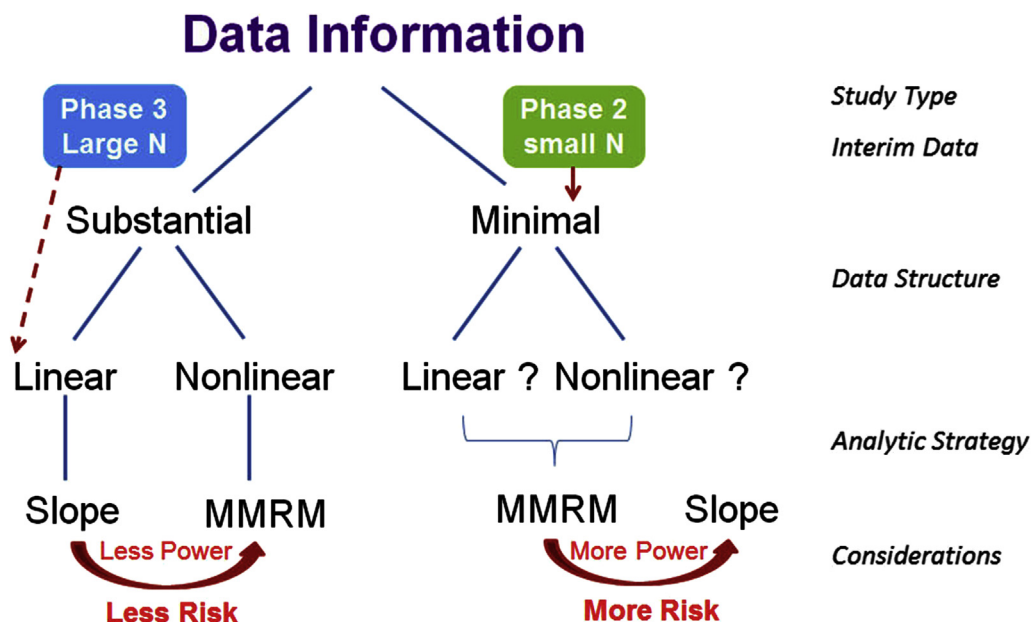


Fig. 5. Proposed analysis roadmap. Abbreviations: MMRM = mixed-model repeated measure.

data to inform method selection. In principle, when the amount of prior data information is substantial in either supporting linear or nonlinear trajectories (e.g., from a phase II or interim study), investigators can plan to use either a slope or MMRM model in the corresponding phase III studies. While planning the analysis for phase III AD studies, investigators often utilize a sample size that is quite large. For benefit and risk trade-off, although a linear trajectory was suggested from prior information, MMRM might be a preferred choice because it can provide an unbiased estimate to eliminate potential risk with a small trade-off in power performance when compared to a slope model. When the amount of available data information is minimal to provide sufficient evidence of linear or nonlinear trajectories, MMRM analysis might be a more conservative approach than the slope model. Conversely, when investigators are planning for a phase II data analysis with a smaller number of patients, the slope model might be a more practical option that can be used to enhance the study power with the risk of increased bias.

The analysis roadmap and general guidance given in this article may also be complemented in the context of particular development programs by routine simulation-based assessment of trial design options. An extensively qualified quantitative simulation tool that allows researchers to model clinical trials in mild-to-moderate AD is publicly available and can be used by researchers to optimize the design of new trials [14].

Some limitations to this study should be considered. Data in this simulation are limited to a single drug in a population selected to specifically assess an anti-amyloid treatment effect for a specific mechanism of action and disease. In addition, this study only assesses mild-to-moderate AD dementia, and results may not be consistent with findings in other stages of AD, with potential variation in disease progression and treatment effects. Therefore, this analysis may not be generalizable to all clinical trial populations and has inherent cohort biases. Replication in other cohorts, disease states, and treatment trials will be important. Finally, the longitudinal shape of AD treatment effect was not explored in this study. A potential direction for future research would be a fixed linear trajectory for placebo while allowing the nonlinear variations in the treatment arm.

In conclusion, MMRM analysis was shown to perform well under various AD disease progression trajectories. The slope model has a modest power advantage over MMRM in linear progression models. However, further research is needed to test linearity and to further quantify the degree of departure from linearity to determine an optimal method of analysis to examine disease progression.

Acknowledgments

Funding was provided by Eli Lilly and Company. Eli Lilly and Company was the funding source and was involved in all stages of the study conduct and analysis. The authors would like to thank Craig Mallinckrodt of Eli Lilly and Company. The authors would also like to thank

Shannon E. Gardell, PhD, and Angela Lorio, BA, of INC Research/inVentiv Health for their assistance with the preparation of this manuscript. Finally, the authors would like to thank one of the reviewers of the manuscript for helpful and detailed suggestions leading to an improved manuscript.

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed the literature using traditional (e.g., PubMed) sources, meeting abstracts, and presentations regarding the comparison of a slope model and endpoint analysis. Although this type of comparison is not yet widely studied, there have been some recent publications examining it; these relevant citations have been appropriately cited.
2. Interpretation: The results of the slope model and mixed-model repeated measure analysis of the EXPEDITION2 study suggest that mixed-model repeated measure analysis with relaxed assumptions on disease progression seems to be the preferred approach with more general applicability to clinical trial design, whereas a slope model might be applicable only if the true model has very little departure from linearity.
3. Future directions: Further research is needed to test linearity and to quantify the degree of departure from linearity observed in this analysis to provide a more detailed comparison of mixed-model repeated measure and a slope model and the application of the two methods in clinical research.

References

- [1] Donohue MC, Aisen PS. Mixed model of repeated measures versus slope models in Alzheimer's disease clinical trials. *J Nutr Health Aging* 2012;16:360-4.
- [2] Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using mixed-effects models. *J Biopharm Stat* 2001;11:9-21.
- [3] Ito K, Corrigan B, Romero K, Anziano R, Neville J, Stephenson D, et al. Understanding placebo responses in Alzheimer's disease clinical trials from the literature meta-data and CAMD database. *J Alzheimers Dis* 2013;37:173-83.
- [4] ALZFORUM. Rusty Unleashed: Forget Disease Modification, Go for Big Effect. Clinical Trials on Alzheimer's Disease. Part 2 of 11. Available at: <http://www.alzforum.org/news/conference-coverage/rusty-unleashed-forget-disease-modification-go-big-effect>. Accessed November 20, 2017.
- [5] Schneider LS, Sano M. Current Alzheimer's disease clinical trials: methods and placebo outcomes. *Alzheimers Dement* 2009; 5:388-97.

- [6] Rogers JA, Polhamus D, Gillespie WR, Ito K, Romero K, Qiu R, et al. Combining patient-level and summary-level data for Alzheimer's disease modeling and simulation: a β regression meta-analysis. *J Pharmacokinet Pharmacodyn* 2012;39:479–98.
- [7] Vellas B, Andrieu S, Sampaio C, Wilcock G. Disease-modifying trials in Alzheimer's disease: a European task force consensus. *Lancet Neurol* 2007;6:56–62.
- [8] Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, et al. Alzheimer's Disease Cooperative Study Steering Committee; solanezumab study group. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N Engl J Med* 2014;370:311–21.
- [9] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* 1984;34:939–44.
- [10] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.
- [11] Fu H, Manner D. Bayesian adaptive dose-finding studies with delayed responses. *J Biopharm Stat* 2010;20:1055–70.
- [12] National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials, Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press; 2010.
- [13] European Medicines Agency. Discussion Paper on the Clinical Investigation of Medicines for the Treatment of Alzheimer's Disease and Other Dementias, 2014. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500176827.pdf. Accessed November 20, 2017.
- [14] ADsim simulation tool. Available at: <https://bitbucket.org/metrumrg/alzheimers-disease-progression-model-adascog/overview>. Accessed November 20, 2017.