Published in final edited form as:

Stat Biopharm Res. 2012 August 30; 4(3): 216–227. doi:10.1080/19466315.2011.634757.

Optimum Design of Disease-modifying Trials on Alzheimer's Disease

Chengjie Xiong,

Division of Biostatistics, Washington University, St. Louis, MO 63110 (314) 362 3635 (Office)

Jingqin Luo,

Division of Biostatistics, Washington University, St. Louis, MO 63110 (314) 362 3718 (Office)

Feng Gao,

Division of Biostatistics, Washington University, St. Louis, MO 63110 (314) 362 3682 (Office)

Ling Chen, and

Division of Biostatistics, Washington University, St. Louis, MO 63110 (314) 747 2373 (Office)

Yan Yan

Department of Surgery and Division of Biostatistics, Washington University, St. Louis, MO 63110 (314) 362 9290 (Office)

Chengjie Xiong: chengjie@wubios.wustl.edu; Jingqin Luo: rosy@wubios.wustl.edu; Feng Gao: feng@wubios.wustl.edu; Ling Chen: ling@wubios.wustl.edu; Yan Yan: yany@wubios.wustl.edu

Abstract

Randomized start and withdrawal designs have been recently proposed to test the disease-modifying agents on Alzheimer's disease (AD). This article provides methods to determine the optimum parameters for these designs. A general linear mixed effects model is proposed. This model employs a piecewise linear growth pattern for those in the delayed treatment or early withdrawal arm, and incorporates a potential correlation on the rates of change on efficacy outcome before and after the treatment switch. Based on this model, we formulate the disease-modifying hypothesis by comparing the rate of change on efficacy outcome between treatment arms with and without a treatment switch, and develop a methodology to optimally determine the sample size allocations to different treatment arms as well as the time of treatment switch for subjects whose treatment is changed. We then propose an intersection-union test (IUT) to assess the disease-modifying efficacy, and study the size and the power of the IUT. Finally, we employ two recently published symptomatic trials on AD to obtain pilot estimates to model parameters, and provide the optimum design parameters including total and individual sample size to different arms as well as the time of treatment switch for future disease-modifying trials on AD.

Keywords

Alzheimer's disease; disease-modifying trials; randomized start design; intersection-union test

1. Introduction

There are two types of therapeutic trials in the search of agents that can treat people with Alzheimer's disease (AD): symptomatic and disease-modifying trials. The former includes

these for symptomatic agents with a primary objective of improving cognition, function, and global measures or deferring decline over a short period of time. The latter consists of those for disease-modifying agents which strive to show that the course of AD has been altered and the rate of disease progression has been slowed (Cummings 2006, Aisen 2006, Citron 2004, Mani 2004). Currently, clinical trials on AD have been almost entirely focused on symptomatic trials for which the standard randomized and placebo controlled parallel designs have been used on patients with AD or on subjects at risk of AD (Kryscio et al. 2004, Ringman et al. 2009, Andrieu et al. 2006). All FDA-approved treatments to AD so far have been symptomatic in nature, and their effectiveness has not been established for the long term and disease-modifying benefit of treating AD.

Although clinical trials for disease-modifying agents have been widely discussed in the AD research community (Leber 1997; Sampaio 2006; Whitehouse et al. 1998, Cummings et al. 2007), the analytic and design complications of such trials remain poorly understood. Because of the novel analytic and design features involved in designing such trials, it is important to obtain optimum design parameters to guide the future clinical trial design for testing disease-modifying compounds on AD.

Complex trial designs have been proposed to allow definite distinctions between symptomatic and disease-modifying clinical trials on AD (Cummings 2006, Aisen 2006, Citron 2004). These designs in general require the switch of treatments in the middle of longitudinal follow-up for at least a proportion of subjects originally randomized to either placebo or active treatment. One such design is the randomized start design (Mani 2004). All patients in the design eventually will receive the active treatment, but are randomized to two treatment groups that begin the active drug at different times. During the initial time period of the study one group receives active drug and the other placebo. After an interval of time sufficient to demonstrate a difference in performance on the efficacy measure between the two groups, the placebo group switches to the active drug. If the patients who begin active drug late ‘catch up’ with those who begin the active drug at baseline, the treatment effect is assumed to be symptomatic. If there is no ‘catch-up’, it is assumed that the effect of the drug is disease-modifying. Often, in order to preserve the blinding of patients and investigators to the active drug, a second randomization may be performed to the initial placebo group so that a proportion of patients will maintain on placebo throughout the trial. Figure 1 presents the expected longitudinal cognitive growth profiles of a randomized start design.

Another design to identify disease modification is the randomized withdrawal design (Mani 2004). This design involves an initial period of double-blinded, placebo-controlled, parallel-arm treatment that is sufficient in duration to establish a difference in effect between the active drug and placebo. Following this period, all those who initially receive active drug are switched to placebo. Both initial groups are then assessed in parallel over a further period of time. If the group that is withdrawn from active drug then regresses on the measure of efficacy to, or towards, the level of the group that receives only placebo, a purely symptomatic effect is assumed. On the other hand, if the group withdrawn from the active drug maintains some gains on the efficacy measure relative to the placebo group, it is assumed that the drug has some effects on the biology of the disease. In order to preserve the blinding of patients and investigators to the active drug, a second randomization may be performed to the initial active drug group so that a proportion of patients will maintain on active drug throughout the trial.

Optimum disease-modifying clinical trials on AD must be based on reasonable statistical models that appropriately fit the longitudinal cognitive or biomarker changes specific to the randomized start and the randomized withdrawal designs. This paper aims first to provide a

general linear mixed effects model to analyze data from clinical trials for disease-modifying agents on AD and hence to lay the analytic foundation to optimally determine design parameters such as the treatment switch time and sample size allocations into different treatment arms. Further, using reported statistics from published symptomatic trials on AD, we obtain estimates to the crucial model parameters and demonstrate the practical application of our proposed methodology to optimally design future disease-modifying trials on AD in this paper.

2. A Longitudinal Model

Efficacy assessments for disease-modifying trials on AD are in general based on three types of outcomes (Cummings 2008): cognitive measures, activity of daily living, and biomarkers. The most commonly used primary cognitive outcome in therapeutic trials of AD has been the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog) (Rosen et al. 1984), whose score ranges from 0 to 70. Traditionally, the ADAS-cog has been treated as a continuous measure in the analyses of clinical trials for symptomatic agents on AD. Recently, many modalities of biomarkers have shown promising ability to track the disease progression, including magnetic resonance imaging (MRI)-based brain volumes (Storandt et al. 2009), diffusion tensor imaging (DTI)-based measures of white matter microstructure (Head et al. 2004), cerebrospinal fluid (CSF, Fagan et al. 2006), and molecular imaging of cerebral fibrillar amyloid with positron emission tomography (PET) using the [^{11}C] benzothiazole tracer, Pittsburgh Compound-B (PIB, Mintun et al. 2006). General linear mixed effects models have been very successful to fit the longitudinal data from many of these efficacy outcomes (Johnson et al. 2009, Storandt et al. 2006). In the following, we propose a general linear mixed effects model to analyze disease-modifying efficacy of clinical trials on AD, i.e., those with a randomized start design or a randomized withdrawal design. We will focus on trials with a randomized start design, though it is straightforward to generalize the proposed methodology to trials with a randomized withdrawal design.

Let Y be the primary efficacy outcome (i.e., either the ADAS-cog or a measure on activity of daily living or a biomarker) tested at time points $t_1, t_2, \dots, t_k$ in a disease-modifying trial with a randomized start design. Let $Y_j^u = (y_{j1}^u, y_{j2}^u, \dots, y_{jk}^u)^t$ be the vector of longitudinal measurements of the j -th subject from the treatment group u . We use $u=tt$ and pp to represent the group of subjects who are in the treatment arm and placebo arm throughout the trial, respectively, and let $u=pt$ represent the group of subjects who initially receive the placebo and then switch to the active treatment. We assume that for either the placebo or the treatment arm, their effects on the longitudinal changes in the mean response can be modeled by a linear trend over time and therefore the slope over time can be used to describe the rate of change. We further assume that when subjects switch from the placebo to the active treatment, the only effect of this switch on the longitudinal growth pattern is through the rate of change. Figure 1 presents the expected rate of cognitive progression for subjects under different treatment arms. The major objective here is first to compare the rate of change (i.e., the slope) over time between the treatment and the placebo before the treatment switch to establish the symptomatic efficacy of the treatment, and then to compare the rate of change between subjects receiving the treatment throughout the trial and those receiving the delayed treatment. These comparisons are complicated by the fact that they must include analyses on longitudinal efficacy data from subjects who switch from the placebo to the treatment in the middle of the trial.

We begin by analyzing longitudinal efficacy outcome for treatment arms that do not change over time, i.e., $u=tt$ or pp , by assuming a standard two-stage random effects model (Diggle et al. 2002) with an individual linear growth curve for ADAS-cog or a biomarker measure for subject j ,

$$Y_{ji}^u = \beta_{0j}^u + \beta_{1j}^u t_i + e_{ji}^u, \quad (1)$$

where β_{0j}^u and β_{1j}^u are subject-specific baseline level (i.e., at $t=0$) and the rate of longitudinal change of the efficacy variable, respectively, and e_{ji}^u 's are assumed to be independent and identically distributed as a normal distribution with mean 0 and variance σ_e^2 . It is clear that the proposed model can easily accommodate covariates, but such covariates are suppressed from notation for simplicity. Notice also that the proposed model and subsequent optimum designs can be extended to accommodate within-patients correlations on the error terms e_{ji}^u 's. Correlations such as autoregressive (AR) structure can be easily implemented in these analyses. Across subjects within a treatment arm, the subject-specific rates of change (β_{1j}^u 's) are further assumed to follow another normal distribution with mean β_1^u and variance σ_u^2 for group $u=tt$ or pp , and are independent of e_{ji}^u 's σ

Next, for the delayed treatment arm, we assume that subjects switch from the placebo to the active treatment (i.e., $u=pt$) at time t_{k0} ($1 < k_0 < k$), i.e., one of the measurement times. Similar to the piecewise random coefficients models proposed in (Xiong et al. 2008), we assume that the immediate and only effect of treatment switch is that the expected progression of the efficacy outcome for subjects right after the switch follows another rate of change that may be different from that for those in the active treatment arm throughout the trial (i.e., $u=tt$). Therefore, the longitudinal growth profile of the efficacy outcome for the j -th subject can be modeled by

$$Y_{ji}^{pt} = \beta_{0j}^{pt} + \beta_{1j}^b (t_i - t_{k0})^- + \beta_{1j}^a (t_i - t_{k0})^+ + e_{ji}^{pt}, \quad (2)$$

where $(t_i - t_{k0})^+ = t_i - t_{k0}$ if $t_i > t_{k0}$, and 0 otherwise, and $(t_i - t_{k0})^- = t_i - t_{k0}$ if $t_i < t_{k0}$, and 0 otherwise, and β_{1j}^b and β_{1j}^a are the subject-specific rate of change before and after the treatment switch time, respectively. We assume that e_{ji}^{pt} 's are independent and identically distributed as a normal distribution with mean 0 and variance σ_e^2 . Across subjects, we assume that $(\beta_{1j}^b, \beta_{1j}^a)$ follows a bivariate normal distribution with mean $(\beta_1^{pp}, \beta_1^{pt})$ and covariance matrix

$$\begin{pmatrix} \sigma_{pp}^2 & \rho \sigma_{pp} \sigma_u \\ \rho \sigma_{pp} \sigma_u & \sigma_u^2 \end{pmatrix},$$

and is independent of e_{ji}^{pt} 's. Notice that here we assume that for subjects whose treatment is delayed, their expected rate of change before the treatment switch is exactly the same as that for the group of subjects receiving the placebo throughout the trial. However, we allow a different expected rate of change (i.e., β_1^{pt}) after the treatment switch from those who receive the treatment throughout the trial.

2.1. Disease-modifying Hypotheses

The comparative nature of β_1^{tt} , β_1^{pp} , and β_1^{pt} determines whether the novel treatment is disease-modifying. More specifically, before the treatment switch, it is expected that the symptomatic efficacy for treating AD will be established. This implies that the expected

decline patterns (see Figure 1) for the treated and the placebo arms, if assumed linear, should separate well before the time when the placebo is switched to the active treatment. This implies that $\beta_1^{tt} > \beta_1^{pp}$. After the treatment switch, the efficacy for modifying the disease can only be established by the fact that the subjects whose treatment is delayed (i.e., $u=pt$) can not 'catch up' those who have been treated throughout the trial. Mathematically, this occurs if and only if $\beta_1^{tt} \geq \beta_1^{pt}$. Therefore, an appropriate statistical hypothesis for establishing the disease-modifying efficacy of the novel treatment is $\beta_1^{tt} \geq \beta_1^{pt} > \beta_1^{pp}$. In order to test this hypothesis, the major interest is in the estimation of mean rates of change from the active treatment and the placebo, i.e., $(\beta_1^{pp}, \beta_1^{tt})$, as well as the rate of change after the treatment switch, i.e., β_1^{pt} . For each treatment arm that does not change over time (i.e., $u=pp$ or $u=tt$), let n_u be the sample size within group u . The simple least squares estimate $\widehat{\beta}_{1j}^u$ to the subject-specific rate of change in the outcome measure for subject j within treatment group u is given by

$$\widehat{\beta}_{1j}^u = \frac{\sum_{i=1}^k (t_i - \bar{t}) Y_{ji}^u}{\sum_{i=1}^k (t_i - \bar{t})^2}, \quad (3)$$

where $\bar{t} = \sum_{i=1}^k t_i / k$. It is straightforward to derive the variance for the least square estimate $\widehat{\beta}_{1j}^u$ to the rate of change as

$$\delta_u^2 = \sigma_e^2 \left\{ \sum_{i=1}^k (t_i - \bar{t})^2 \right\}^{-1} + \sigma_u^2. \quad (4)$$

Notice that $\widehat{\beta}_{1j}^u$ follows a normal distribution with mean β_1^u and variance δ_u^2 . Let $\bar{\beta}_1^u$ be the mean estimated rate of change for subjects receiving either active treatment or placebo throughout the trial (i.e., $u=tt$ and pp). It is clear that $\bar{\beta}_1^u$ is an unbiased estimator of β_1^u , and follows a normal distribution with the variance given by $\sigma_{\bar{\beta}_1^u}^2 = \delta_u^2 / n_u$.

For subject j who begins with the placebo and then switches to the active treatment, i.e., $u=pt$, a similar estimate to the rate of change before and after the treatment change time t_{k_0} ($1 < k_0 < k$) is

$$\widehat{\beta}_{1j}^b = \frac{\sum_{i=1}^{k_0} (t_i - \bar{t}_p) Y_{ji}^{pt}}{\sum_{i=1}^{k_0} (t_i - \bar{t}_p)^2}, \quad (5)$$

and

$$\widehat{\beta}_{1j}^a = \frac{\sum_{i=k_0}^k (t_i - \bar{t}_t) Y_{ji}^{pt}}{\sum_{i=k_0}^k (t_i - \bar{t}_t)^2}, \quad (6)$$

respectively, where $\bar{t}_p = \sum_{i=1}^{k_0} t_i / k_0$ and $\bar{t}_t = \sum_{i=k_0}^k t_i / (k - k_0 + 1)$. The covariance matrix of $(\widehat{\beta}_{1j}^b, \widehat{\beta}_{1j}^a)$ is

$$\Sigma_{pt} = \begin{pmatrix} \sigma_{pt1}^2 & \rho_{pt}\sigma_{pt1}\sigma_{pt2} \\ \rho_{pt}\sigma_{pt1}\sigma_{pt2} & \sigma_{pt2}^2 \end{pmatrix}, \quad (7)$$

where

$$\sigma_{pt1}^2 = \sigma_e^2 \left\{ \sum_{i=1}^{k_0} (t_i - \bar{t}_p)^2 \right\}^{-1} + \sigma_{pp}^2, \quad \sigma_{pt2}^2 = \sigma_e^2 \left\{ \sum_{i=k_0}^k (t_i - \bar{t}_t)^2 \right\}^{-1} + \sigma_{tt}^2,$$

and

$$\rho_{pt} = \frac{\rho\sigma_{pp}\sigma_{tt} + (t_{k_0} - \bar{t}_p)(t_{k_0} - \bar{t}_t)\sigma_e^2 \left[\sum_{i=1}^{k_0} (t_i - \bar{t}_p)^2 \sum_{i=k_0}^k (t_i - \bar{t}_t)^2 \right]^{-1}}{\sigma_{pt1}\sigma_{pt2}}.$$

In the special case with an evenly spaced longitudinal design among the repeated measures, if $t_k - t_1$ and k are chosen, it is straightforward to prove that

$$\sum_{i=1}^{k_0} (t_i - \bar{t}_p)^2 = \frac{(t_{k_0} - t_1)^2 k_0(k_0 + 1)}{12(k_0 - 1)},$$

and

$$\sum_{i=k_0}^k (t_i - \bar{t}_t)^2 = \frac{(t_k - t_{k_0})^2 (k - k_0 + 1)(k - k_0 + 2)}{12(k - k_0)}.$$

Notice that the estimated rates of change before and after the treatment switch, i.e.,

$(\widehat{\beta}_{1j}^b, \widehat{\beta}_{1j}^a)$, follow a bivariate normal distribution with mean rates $(\beta_1^{pp}, \beta_1^{pt})$ and covariance matrix Σ_{pt} . Let $(\widehat{\beta}_1^{pp}, \widehat{\beta}_1^{pt})$ denote the mean estimated rate of change before and after the treatment switch. $(\widehat{\beta}_1^{pp}, \widehat{\beta}_1^{pt})$ is an unbiased estimator to $(\beta_1^{pp}, \beta_1^{pt})$ with a covariance matrix given by Σ_{pt}/n_{pt} , where n_{pt} is the sample size for $u=pt$.

Now that there are two unbiased estimators to β_1^{pp} , one from subjects in the placebo arm throughout the trial (i.e., $\widehat{\beta}_1^{pp}$), and the other from subjects whose treatment is switched (i.e., $\widehat{\beta}_1^{pp}$ before the switch). For any constant weight $0 < c < 1$, let $\widehat{\beta}_{1c}^{pp} = c\widehat{\beta}_1^{pp} + (1-c)\widehat{\beta}_1^{pp}$ be another unbiased estimator to β_1^{pp} . The variance of $\widehat{\beta}_{1c}^{pp}$ is given by

$$\sigma_{\widehat{\beta}_{1c}^{pp}}^2 = \frac{c^2 \delta_{pp}^2}{n_{pp}} + \frac{(1-c)^2 \sigma_{pt1}^2}{n_{pt}}. \quad (8)$$

Let $\alpha = \beta_1^{tt} - \beta_1^{pt}$ and $\theta = \beta_1^{pt} - \beta_1^{pp}$. Let $\hat{\alpha} = \hat{\beta}_1^{tt} - \hat{\beta}_1^{pt}$ and $\hat{\theta} = \hat{\beta}_1^{pt} - \hat{\beta}_1^{pp}$ be the corresponding estimates. $(\hat{\alpha}, \hat{\theta})$ follows a bivariate normal distribution with mean (α, θ) and covariance matrix given by

$$\Omega = \begin{pmatrix} \sigma_{\hat{\alpha}}^2 & \sigma_{\hat{\alpha}\hat{\theta}} \\ \sigma_{\hat{\alpha}\hat{\theta}} & \sigma_{\hat{\theta}}^2 \end{pmatrix}, \quad (9)$$

where

$$\sigma_{\hat{\alpha}}^2 = \frac{\delta_{tt}^2}{n_{tt}} + \frac{\sigma_{pt2}^2}{n_{pt}}$$

$$\sigma_{\hat{\theta}}^2 = \frac{c^2 \delta_{pp}^2}{n_{pp}} + \frac{(1-c)^2 \sigma_{pt1}^2 + \sigma_{pt2}^2 - 2\rho_{pt}(1-c)\sigma_{pt1}\sigma_{pt2}}{n_{pt}}$$

and

$$\sigma_{\hat{\alpha}\hat{\theta}} = \frac{(1-c)\rho_{pt}\sigma_{pt1}\sigma_{pt2} - \sigma_{pt2}^2}{n_{pt}}.$$

To test the disease-modifying efficacy of the active treatment, we propose to test the null hypothesis $H_0: \alpha < 0$ or $\theta < 0$ against the alternative $H_1: \alpha \geq 0$ and $\theta > 0$. The null hypothesis is the union of two null hypotheses $H_{0\alpha}: \alpha < 0$ and $H_{0\theta}: \theta < 0$, and the alternative is the intersection of two alternative hypotheses $H_{1\alpha}: \alpha \geq 0$ and $H_{1\theta}: \theta > 0$. For each individual set of null and alternative hypotheses, let $z^\alpha = \hat{\alpha}/\sigma_{\hat{\alpha}}$ and $z^\theta = \hat{\theta}/\sigma_{\hat{\theta}}$ be the test statistic for testing the corresponding individual hypothesis. If $\alpha = 0$ or $\theta = 0$, the corresponding test statistic follows a standard normal distribution. To test the null hypothesis $H_0: \alpha < 0$ or $\theta < 0$ against the alternative $H_1: \alpha \geq 0$ and $\theta > 0$, an intersection-union test (IUT, Berger and Sinclair 1984, Berger 1989, Liu and Berger 1995) rejects the null hypothesis when both $z^\alpha > M$ and $z^\theta > M$ for some constant M . In order for the test to have a size of γ ($0 < \gamma < 1$), M has to be chosen such that

$$\sup_{H_0} \Pr(z^\alpha > M, z^\theta > M) = \gamma.$$

Notice that

$$p(\alpha, \theta, M) = \Pr(z^\alpha > M, z^\theta > M) = \int_{m_\alpha}^{\infty} \int_{m_\theta}^{\infty} \frac{(2\pi)^{-1}}{\sqrt{|\Omega_1|}} \exp\left[-0.5Z' \Omega_1^{-1} Z\right] dz^\alpha dz^\theta \quad (10)$$

where $Z = (z^\alpha, z^\theta)$, $m_\alpha = M - \alpha/\sigma_{\hat{\alpha}}$ and $m_\theta = M - \theta/\sigma_{\hat{\theta}}$ and

$$\Omega_1 = \begin{pmatrix} 1 & \sigma_{\hat{\alpha}\hat{\theta}}/(\sigma_{\hat{\alpha}}\sigma_{\hat{\theta}}) \\ \sigma_{\hat{\alpha}\hat{\theta}}/(\sigma_{\hat{\alpha}}\sigma_{\hat{\theta}}) & 1 \end{pmatrix}.$$

If $\theta = 0$, $\alpha > 0$, then $\lim_{\alpha \rightarrow \infty} m_\alpha = -\infty$. It follows that

$$\lim_{\alpha \rightarrow \infty} P(\alpha, \theta, M) = \Pr(z^\theta > M | \theta = 0).$$

Therefore, when $M = z_\gamma$, the 100γ upper percentile of the standard normal distribution, the rejection region $z^\alpha > M$ and $z^\theta > M$ for the IUT provides the size γ for testing $H_0: \alpha < 0$ or $\theta = 0$ against the alternative $H_1: \alpha = 0$ and $\theta > 0$, and the corresponding power function for the IUT is given by

$$p(\alpha, \theta) = p(\alpha, \theta, z_\gamma).$$

Thus, the sample sizes required to achieve a statistical power of $(1 - \eta)$ ($0 < \eta < 1$) are the solutions to n_{tt} , n_{pp} , and n_{pt} such that $P(\alpha, \theta) = 1 - \eta$.

Notice that the total spacing or the duration of the trial (i.e., $t_k - t_1$), the number of repeated measures on the outcome variable (i.e., k), the spacing of the repeated measures, and the time when the delayed treatment group switches from placebo to the active treatment (i.e., t_{k0}) all impact the statistical power and therefore the sample sizes. Notice also that in the test of $H_0: \alpha < 0$ or $\theta = 0$ against the alternative $H_1: \alpha = 0$ and $\theta > 0$, the estimate to β_1^{pp} , i.e., $\widehat{\beta}_{1c}^{pp} = c\widehat{\beta}_1^{pp} + (1-c)\widehat{\beta}_1^{pt}$, depends on the constant c . The optimum test on the treatment efficacy in this family of test statistics relies on the optimum estimate to β_1^{pp} . Let $n = n_{tt} + n_{pp} + n_{pt}$ be the total sample size. Let $\lambda_u = n_u/n$ be the proportion (i.e., allocation) of sample size to each treatment group $u = pp, tt$, and pt . It is clear that $\lambda_{pp} + \lambda_{tt} + \lambda_{pt} = 1$. We optimize the choice of c by minimizing the variance of $\sigma_{\widehat{\beta}_{1c}^{pp}}^2$ as given by Equation (8), i.e.,

$$c = \frac{\sigma_{pt1}^2}{\sigma_{pt1}^2 + \frac{\lambda_{pt}}{\lambda_{pp}} \delta_{pp}^2}. \quad (11)$$

Thus, the proposed IUT with the optimum c provides the most powerful test within the family.

2.2. Optimum Design Parameters

Even if the optimum weight c is chosen to obtain the minimum variance estimator to β_1^{pp} , the estimators to both $\alpha = \beta_1^{tt} - \beta_1^{pt}$ and $\theta = \beta_1^{pt} - \beta_1^{pp}$ depend on the choice of sample sizes n_{pp} , n_{tt} , and n_{pt} . Given a total sample size of n for the clinical trial, another important design issue is how to allocate the sample size into different treatment arms, i.e., $\lambda_u, u = pp, tt, pt$, so that the estimates to the efficacy parameters can be optimized. In order to find the optimum sample size allocations $\lambda_u, u = pp, tt, pt$, we compute the ‘total variance’ of $\hat{\alpha}$ and $\hat{\theta}$ as

$$\sigma_{total}^2 = \sigma_{\hat{\alpha}}^2 + \sigma_{\hat{\theta}}^2.$$

It is straightforward from Equation (9) that

$$n\sigma_{total}^2 = \frac{\delta_{tt}^2}{\lambda_{tt}} + \frac{2\sigma_{pt2}^2}{\lambda_{pt}} + \frac{\delta_{pp}^2(\sigma_{pt1}^2 - 2\rho_{pt}\sigma_{pt1}\sigma_{pt2})}{\lambda_{pp}\sigma_{pt1}^2 + \lambda_{pt}\delta_{pp}^2}. \quad (12)$$

To find the optimum sample size allocations $\lambda_u, u = pp, tt, pt$, the total variance needs to be minimized with respect to the choices of $\lambda_u > 0, u = pp, tt, pt$, and subject to $\lambda_{pp} + \lambda_{tt} + \lambda_{pt} = 1$. In general, when the between-subject variance (i.e., σ_u^2) is large compared to the within-subject variance (i.e., σ_e^2), for example, when

$$\delta_{pp}^2 \geq \frac{\sigma_{pt1}^2(\sigma_{pt1}^2 - 2\rho_{pt}\sigma_{pt1}\sigma_{pt2})}{1 + \sigma_{pt1}^2 - 2\rho_{pt}\sigma_{pt1}\sigma_{pt2}},$$

the derivative of $n\sigma_{total}^2$ with respect to λ_{pp} becomes positive, indicating that the total variance $n\sigma_{total}^2$ is an increasing function of λ_{pp} . Therefore, the total variance $n\sigma_{total}^2$ is minimized with respect to λ_{pp} when $\lambda_{pp} = 0$, i.e., no subjects will be randomized to the placebo throughout the trial. This fact highlights the focus of disease-modifying trials with a randomized start design on the other two treatment arms, i.e., the delayed and non-delayed treatment arms. Although the second stage randomization in a randomized start design is introduced among subjects who receive placebo at baseline, the main purpose is to preserve the blinding of subjects and investigators to the active drug, i.e., the subjects who are randomized into the placebo again from the second randomization are not the major focus of the design, albeit they have to participate in the efficacy analyses based on the ‘intent-to-treat’ principle (Montori and Guyatt 2001, Heritier et al. 2003). Because of the reason, it is practical for the investigators to pre-specify a small portion of subjects to be randomized to placebo from the second randomization, and design the randomized start design based on the specification. Let λ_0 be a small pre-specified portion of subjects who will be randomized to $u=pp$ to maintain blinding, the optimum sample size allocations are then the solutions for minimizing $n\sigma_{total}^2$ subject to $\lambda_{pt} + \lambda_{tt} = 1 - \lambda_0$. Mathematically, the optimum sample size allocation to group $u=pt$ is the solution of $\lambda_{pt} (0 < \lambda_{pt} < 1 - \lambda_0)$ to the following equation:

$$f(\lambda_{pt}) = \frac{\delta_{tt}^2}{(1 - \lambda_0 - \lambda_{pt})^2} - \frac{2\sigma_{pt2}^2}{\lambda_{pt}^2} - \frac{(\sigma_{pt1}^2 - 2\rho_{pt}\sigma_{pt1}\sigma_{pt2})\delta_{pp}^4}{(\lambda_0\sigma_{pt1}^2 + \lambda_{pt}\delta_{pp}^2)^2} = 0. \quad (13)$$

Function $f(\lambda_{pt})$ is a strictly increasing function of λ_{pt} because its derivative with respect to λ_{pt} is positive. Given that $f(\lambda_{pt})$ approaches $-\infty$ and ∞ , respectively, when λ_{pt} approaches 0 and $1 - \lambda_0$. The solution to Equation (13) uniquely exists and involves solving a 4-th degree polynomial. The closed form solution to λ_{pt} from the 4-th degree polynomial is very complex but can be found in Abramowitz and Stegun (1972). In practice, standard software can be used to find the solution. After the optimum sample size allocation to group $u=pt$ is obtained, the optimum sample size allocation to group $u=tt$ is given by $\lambda_{tt} = 1 - \lambda_0 - \lambda_{pt}$.

In designing clinical trials for testing potential disease-modifying agents on AD, if the linear growth model or piecewise linear growth model is a valid statistical model and that the logistic and practical factors allow, an increase of either the study duration or the frequency of repeated measures will in general decrease the within-subject variability and improve the precision of parameter estimates or the statistical power in the test on the rate of change over time. Although the choice of measurement times, $t_1, t_2, \dots, t_k$, should theoretically be chosen

to minimize the variance of the estimated difference on the rate of change across treatment groups, many economic, logistic, or subject-specific factors may dictate the choice. In addition, the validity of the assumed statistical model also constrains the choice of trial duration $t_k - t_1$ in the sense that a linear growth or a piecewise linear growth for the delayed treatment group over time might not be a reasonable assumption with a very long study duration. Similarly, the number of repeated measures in a longitudinal study might also be constrained by many practical factors and it may be impossible for the designers of the study to freely choose the number of repeated measures. As a result, many clinical trials are restricted to relatively short duration with a pre-determined number of repeated measures which is not chosen statistically based on an optimal design. Given that the duration of the trial (i.e., $t_k - t_1$) and the number of measurements per subject (i.e., k) are typically chosen by some non-statistical reasons, and assuming an evenly spaced longitudinal design among the repeated measures, the total variance as given in (12) can be simplified as a function of treatment switch time k_0 only through

$$g(k_0) = \frac{24\sigma_e^2(L^2\lambda_{pt})^{-1}}{(k-k_0)(k-k_0+1)(k-k_0+2)} + \frac{\delta_{pp}^2[\sigma_{pp}^2 - 2\rho\sigma_{pp}\sigma_{tt} + \sigma_e^2B]}{\lambda_{pp}[12\sigma_e^2/((k_0-1)k_0(k_0+1)L^2)] + \lambda_{pp}\sigma_{pp}^2 + \lambda_{pt}\delta_{pt}^2}$$

where L is the time interval between 2 adjacent measurements, and

$$B = \frac{1}{L^2} \left\{ \frac{12k_0^{-1}}{(k_0-1)(k_0+1)} + \frac{72k_0^{-1}(k_0+1)^{-1}}{(k-k_0+1)(k-k_0+2)} \right\}.$$

The optimum treatment switch time k_0 therefore must minimize $g(k_0)$ over the possible options $k_0 = 2, 3, \dots, k-1$. Assuming a randomized start trial with a total of 10% subjects randomized to the placebo throughout the trial and that $\sigma_{pp}^2 = \sigma_{tt}^2 = \sigma^2$, and given an evenly spaced measurements design (i.e., quarterly) with a total of k measurements per subject, Table 1 shows the optimum treatment switch time (i.e., k_0 , $1 < k_0 < k$) and the optimum allocations (in %) of the total sample size to the other two treatment arms (i.e., λ_{pt} and λ_{tt}) as a function of k , the ratio of between and within subject variances $r = \sigma^2 / \sigma_e^2$, and the correlation ρ between the rates of change before and after the treatment switch. The optimum weight c in the total variance is assumed in Table 1.

Finally, although randomized start design allows investigators to pre-specify a small portion of subjects to be randomized to placebo throughout the trial, other factors may affect such design. For example, differences between “phases” of the trial in rates of learning effects on cognitive tests and differential influences of study effects as well as placebo effects all might suggest the need to allocate more patients to placebo in the second phase of the trial in order to more reliably estimate these effects. Patients, care givers, and investigators would know, even if blinded to exact timing of the treatment switch, that the odds of being on placebo are low at the end of the trial. This alteration in expectation could also invalidate the linearity assumption. Further, when the percentage of subjects allocated to pp is small, confounding between treatment and site becomes a concern in the design, especially for the sites with relatively small number of subjects. This concern might be partially alleviated by the fact that the estimate to the slope of pp is dependent not only on subjects allocated to pp , but also on those from pt before their treatment is switched.

3. Designing Future Disease-modifying Trials on AD

AD is a progressive neurodegenerative disorder of the brain characterized by an insidious onset of memory deterioration, progressive cognitive deterioration, emergence of neuropsychiatric symptoms and behavioral disturbances, impairment of activities of daily living, and loss of independent function. During the past decade or so, several compounds have been approved by the FDA to enhance cognition and global function of AD patients, and recent advances in understanding AD pathogenesis has led to the development of numerous compounds that might modify the disease process. A wide array of anti-amyloid and neuroprotective therapeutic approaches are under investigation on the basis of the hypothesis that amyloid beta (A β) protein plays a pivotal role in disease onset and progression and that secondary consequences of A β generation and deposition, including tau hyperphosphorylation and neurofibrillary tangle formation, oxidation, inflammation, and excitotoxicity, contribute to the disease process (Salloway et al. 2008). Interventions in these processes with agents that reduce amyloid production, limit aggregation, or increase removal might block the cascade of events comprising AD pathogenesis. Reducing tau hyperphosphorylation, limiting oxidation and excitotoxicity, and controlling inflammation might be beneficial disease-modifying strategies. Potentially neuroprotective and restorative treatments such as neurotrophins, neurotrophic factor enhancers, and stem cell-related approaches are also under investigation (Salloway et al. 2008). It is anticipated that these promising agents and treatments will soon be tested for their ability to modify the disease process of AD through well designed clinical trials.

Here we provide optimum design parameters for future clinical trials of disease-modifying agents on AD by applying our proposed methodology to a variety of design scenarios. We assume a randomized start design in which 10% or 20% subjects will be randomized to receiving placebo throughout the trial and then optimize the sample size allocations to the treatment arm and the delayed treatment arm. We also optimize the time of treatment switch for the delayed treatment arm. The optimum weight c as given by Equation (11) is used in the estimate of β_1^{pp} and subsequently in the IUT of the disease-modifying efficacy. Finally, we assume that the efficacy outcome will be assessed quarterly in future disease-modifying trials on AD.

Because most reported symptomatic trials on AD used ADAS-cog as the primary efficacy outcome measure, we assume that the future disease-modifying trials will also use the same cognitive outcome as the primary efficacy endpoint. Further, although the design and analysis of disease-modifying trials on AD have been extensively discussed in the AD literature (Cummings 2006, Aisen 2006, Citron 2004, Mani 2004), no disease-modifying trials on patients with AD have been reported. We therefore choose to obtain necessary estimates to important model parameters through several recently reported symptomatic trials on AD. These parameters include between and within subject variances σ_{pp}^2 (or σ_u^2) and σ_e^2 .

Essentially all published symptomatic trials on AD that used ADAS-cog as the primary efficacy endpoint reported the efficacy analyses using the change of ADAS-cog score from the baseline. These published symptomatic treatment trials on AD followed patients for a duration ranging from 4 weeks to 1 year (Qizilbash et al. 1998; Kaduszkiewicz et al. 2005), and therefore the reported variance for the change from baseline on ADAS-cog score also spanned a wide range (Qizilbash et al. 1998; Kaduszkiewicz et al. 2005). None of the published symptomatic trials directly reported estimates to individual variance components σ_{pp}^2 and σ_u^2 associated with the rates of cognitive change (i.e., β_{1j}^{pp}) or the within-subject variance σ_e^2 as given in Model (1) and (2). We therefore propose to combine statistics from

multiple published symptomatic trials on AD to obtain estimates that are needed for the optimum design of a future disease-modifying trial on AD.

We conducted a comprehensive literature review on symptomatic clinical trials on AD, and located two trials that were reasonably large in sample size and long in follow-up duration and also specifically reported the variance associated with the change of ADAS-cog score from the baseline for the placebo arm. Aisen et al. (2003) reported the effects of Rofecoxib or Naproxen in treating AD for a 1-year trial from which 111 subjects were randomized to placebo. Rogers et al. (1998) reported the effects of Donepezil in treating AD for a 24-weeks trial from which 162 subjects were randomized to placebo. Because of variable length of longitudinal follow-up for these trials, the reported variance associated with the change of ADAS-cog score from the baseline is a function of the length of follow-up. However, if model (1) is appropriate, i.e., assuming a linear growth pattern of ADAS-cog over time, the annual rate of change on ADAS-cog (i.e., the slope) can be estimated (most times through extrapolations because of the less than 1 year follow-up) by the reported mean difference from baseline divided by the follow-up time (in years). Therefore the standard deviation σ_{pp} for the annual rate of change (i.e., β_{1j}^{pp}) can be estimated by the reported standard deviation on the change from baseline divided by the number of years in follow-up. Let $D(t_k - t_1)$ be the duration of a reported symptomatic trials on AD. Because only statistics on the change score of ADAS-cog from the baseline were reported in the published symptomatic trials on AD, we propose to link the reported statistics with our proposed model (1) as if the published trials were conducted with only two time points, i.e., the baseline and the final measurements on ADAS-cog. Whereas these published trials did assess subjects at more than two time points, our proposed approach is the only practical one because of the fact that no statistics has been reported on the efficacy at middle time points between the baseline and the final assessments from these publications. Assuming that $\sigma_{pp}^2 = \sigma_u^2 = \sigma^2$, we therefore have $\sigma_{\Delta}^2/D^2 = 2\sigma_e^2/D^2 + \sigma^2$, where σ_{Δ}^2 is the reported variance for the change score of ADAS-cog from the baseline in the placebo arm. For the trial reported by Aisen et al. (2003), a sample of 111 subjects were randomized to the placebo, the mean 1-year change from baseline on ADAS-cog is 5.7 points with an estimated $\sigma_{\Delta} = 8.2$ points. Therefore $8.2^2 = 2\sigma_e^2 + \sigma^2$. For the trials reported by Rogers et al. (1998), a sample of 162 subjects were randomized to the placebo, and the estimated mean change of ADAS-Cog from baseline in a 24-week (i.e., 0.46-year) follow-up is 1.82 with an estimated standard deviation of $\sigma_{\Delta} = 6.06$. Therefore, $6.06^2/0.46^2 = 2\sigma_e^2/0.46^2 + \sigma^2$. Solving these two equations, we obtain $\sigma^2 = 38.71$ and $\sigma_e^2 = 14.27$.

Now that we have obtained estimates to between- and within- subject variances σ_{pp}^2 and σ_e^2 associated with the annual rate of change in model (1), we search for the optimum design parameters for future disease-modifying trials on AD. Assuming a randomized start design for a 2-year or 3-year clinical trial with quarterly assessments and $\sigma_{pp}^2 = \sigma_u^2 = \sigma^2$, for a selected set of effect sizes typically reported in the literature (i.e., both $\alpha = \beta_1^{tt} - \beta_1^{pt}$ and $\theta = \beta_1^{pt} - \beta_1^{pp}$), Table 2 presents the sample sizes for different treatment arms (i.e., $u=pt$ and $u=tt$) required to detect the effect sizes with a statistical power of 80%. These individual sample sizes are based on the optimum sample size allocations (in %) to group $u=pt$ and $u=tt$ (i.e., λ_{pt} and λ_{tt}). 10% or 20% subjects are assumed to be randomized into placebo arm throughout the trial for preserving the blinding of the trial. Table 2 also presents the optimum treatment switch time (i.e., k_0 , $1 < k_0 < k$). The optimum test statistic in the IUT was used (i.e., with optimum weight c) in Table 2. A correlation of 0.5 (i.e., ρ) between the rates of change before and after the treatment switch (i.e., $(\beta_{1j}^b, \beta_{1j}^a)$) from the delayed

treatment group was assumed in Table 2. We also computed the sample sizes for individual treatment arms ($u=tt$ or pt) in a randomized start design with different correlations (0.1 and 0.8) between the rates of change before and after the treatment switch, and found that results in Table 2 are not changed significantly. Power function (10) used in Table 2 was evaluated by SAS function PROBBNRM (SAS, 1990). In real world multi-site clinical trials, randomization is often stratified by investigative sites. Therefore, the optimum allocations in Table 1 may not always be practical. An easy assessment on the loss of power or efficiency when using practical allocations can be obtained by comparing the ‘total variance’ of $\hat{\alpha}$ and $\hat{\theta}$, σ_{total}^2 , as given by Equation (12) to that obtained with the optimum allocations.

4. Discussion

The looming public health crisis due to AD mandates a fast development of novel disease-modifying treatments for the disease. Unlike symptomatic trials for which a single randomization at baseline is generally implemented, disease-modifying trials require an initial randomization followed by a re-randomization of patients in either the placebo or treatment arm. In order to design optimum clinical trials for establishing the disease-modifying efficacy of potential novel treatments, we proposed a general linear mixed effect model to analyze the rate of change for efficacy outcome variables in a randomized start trial on AD. Based on this model, we first formulated the appropriate disease-modifying hypothesis by comparing the rate of change in efficacy outcome between the treated group throughout the trial and the delayed treatment group. Because of the second stage randomization to the subjects who are initially randomized to the placebo, a third treatment arm in which subjects are randomized to the placebo throughout the trial is available. The third treatment arm complicated the statistical test of disease-modifying efficacy because of the need to use all data to estimate the rate of change for subjects receiving the placebo. After obtaining an optimum estimate to the rate of change for placebo by combining data from subjects receiving placebo throughout the trial and from those before receiving the delayed treatment, we developed a methodology to optimally determine the sample size allocations to different treatment arms as well as the time for treatment switch for subjects whose treatment is delayed. After the design parameters were optimally chosen, we proposed an intersection-union test to assess the efficacy of potential disease-modifying agents on AD. We studied the size and the power of the IUT, and provided a method of determining the sample sizes to adequately power the test of disease-modifying efficacy.

The randomized start and the randomized withdrawal designs considered here are by far the most popular choices for disease-modifying trials on AD (Leber 1997; Sampaio 2006; Whitehouse et al. 1998, Cummings et al. 2007). These designs differ from the standard crossover designs (Chi 1992) in the sense that the former allows some of the subjects receiving only one treatment throughout the trials. Jarjoura (2003) considered the efficiency of a clinical trial design which did allow crossing control to treatment (i.e., similar to our designs), but no optimal design parameters such as sample size allocations and treatment switch times were provided in their work. Our analytic approaches also differ from those of other authors (Jarjoura 2003). Here we assumed a random intercept and random slope model (Laird and Ware 1982) for the repeatedly measured continuous efficacy outcome, and derived statistical tests and optimal design parameters based on this model. For individuals whose treatment was delayed, our model assumed a piecewise linear pattern. More importantly, our model allowed potentially differential rate of change for subjects receiving delayed treatment as compared to those receiving the treatment throughout the trial, as well as a correlation on the rates of change before and after the treatment switch. Finally, we point out that there are many possible extensions the proposed model can be potentially useful: non-linear progression, inclusion of covariates such as baseline disease status that has direct association with the subsequent rate of change in AD, correlated or non-normal

errors, specific dependence on the rates of progression with and without delayed active treatment, delayed effect of treatment, and multiple switches of treatments over time. Some of these extensions can be straightforward. For example, the proposed model can easily accommodate covariates as well as within-patients correlations on the errors terms e_{ji}^u 's (e.g., autoregressive (AR) structure). Some other extensions require more careful evaluations before implementation. For example, linear and piecewise linear decline over time is an important assumption in our derived optimum designs for disease-modifying trials on AD. Without such an assumption, our proposed designs are likely no longer optimum. If pilot longitudinal data exist, linearity should be carefully assessed against the existing data over a relatively long follow-up typically required in disease-modifying trials on AD. If no pilot longitudinal data are available, appropriate interim analyses can be designed in such clinical trials to allow an assessment of longitudinal pattern and adjustments on the study design. Future work is also needed to assess the sensitivity of the linear assumption in the proposed optimum design, especially in terms of bias, power, and Type I error rate.

In order to design optimum future disease-modifying trials on AD, we conducted a literature review on published symptomatic trials on AD, and located two recently reported symptomatic trials on AD that were relatively large in sample size and long in follow-up and also reported the variance associated with the change of ADAS-cog score from the baseline for the placebo arm (Aisen et al. 2003; Rogers et al. 1998). Given that none of the published trials directly reported the estimates to between- and within-subject variances σ_{pp}^2 and σ_e^2 in model (1), we proposed a novel approach to obtain pilot estimates to these important model parameters by linking our proposed model (1) to the reported statistics. More specifically, we solved a system of equations that were derived from the reported variances associated with the change of ADAS-cog score from the baseline in these two trials. After obtaining estimates to between- and within-subject variances σ_{pp}^2 and σ_e^2 , we computed the optimum design parameters (i.e., sample size allocations, and treatment switch time) for future disease-modifying trials on AD, and provided the sample sizes into different treatment arms required to detect a selected set of effect sizes with a statistical power of 80%. Our results show that clinical trials for disease-modifying agents on AD can be adequately powered and optimized. The proposed methods of sample size determination provide evidence that much larger sample sizes are in general required to adequately power disease-modifying trials when compared to symptomatic trials on AD, even when the treatment switch time and the test statistic for efficacy are optimally chosen. Finally, a disease-modifying trial, by definition, requires longer follow-up than symptomatic trials because the former needs to first establish symptomatic efficacy (before the treatment switch) and then the disease-modifying efficacy (after the treatment switch). Although we have presented the design of disease-modifying trials on AD with a relatively long follow-up (i.e., 2 or 3 years) in Table 2, the proposed analytic approaches can be used for proof-of-concept studies with a much shorter follow-up. Similarly, because a disease-modifying trial needs to first establish symptomatic efficacy, it is no surprise that a disease-modifying trial requires larger sample size than symptomatic trials as demonstrated by our findings. However, we point out that the large sample sizes needed for disease-modifying trials on AD are at least also partly due to the fact that ADAS-cog is subject to a large variation over time and therefore may not be an ideal efficacy outcome in these trials. Much more sensitive and reliable novel biomarkers will be needed to design future disease-modifying trials on AD with a much smaller sample size. Many recently reported promising biomarkers on AD, such as MRI-based brain volumes (Storandt et al. 2009), DTI-based measures of white matter microstructure (Head et al. 2004), CSF-based biomarkers (Fagan et al. 2006), and molecular imaging of cerebral fibrillar amyloid with PET using the [^{11}C] benzothiazole tracer, Pittsburgh Compound-B (PIB, Mintun et al. 2006), are potential candidates of efficacy outcomes for future disease-modifying trials on AD.

Acknowledgments

Dr. Xiong's work was partly supported by National Institute on Aging grants NIH/NIA R01 AG029672, NIH/NIA R01 AG034119, AG003991, AG005681, AG026276, and by the Alzheimer's Association grant NIRG-08-91082.

References

- Abramowitz, M.; Stegun, IA., editors. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. New York: Dover; 1972. Solutions of Quadratic Equations; p. 17-18.
- Aisen PS. Commentary on "Challenges to demonstrating disease-modifying effects in Alzheimer's disease clinical trials". *Alzheimer's & Dementia*. 2006; 6:272-274.
- Aisen PS, Schafer KA, Grundman M, et al. Effects of Rofecoxib or Naproxen vs. Placebo on Alzheimer disease progression: a randomized controlled trial. *JAMA*. 2003; 289:2819-2826. [PubMed: 12783912]
- Andrieu S, Rascol O, Lang T, et al. Methodological issues and statistical analyses. *J Nutr Health Aging*. 2006; 10:116-117. [PubMed: 16554944]
- Berger RL. Uniformly more powerful tests for hypotheses concerning linear inequalities and normal means. *Journal of the American Statistical Association*. 1989; 84:192-199.
- Berger RL, Sinclair DF. Testing hypotheses concerning unions of linear subspaces. *Journal of the American Statistical Association*. 1984; 79:158-163.
- Chi EM. Analysis of cross-over trials when within-subject errors follow an AR(1) process. *Biom J*. 1992; 34:359-365.
- Citron M. Strategies for disease modification in Alzheimer's disease. *Net Rev Neurosci*. 2004; 5:677-685.
- Cummings JL. Challenges to demonstrating disease-modifying effects in Alzheimer's disease clinical trials. *Alzheimer's & Dementia*. 2006; 6:263-271.
- Cummings JL. Optimizing phase II of drug development for disease-modifying compounds. *Alzheimers Dement*. 2008; 4:S15-20. [PubMed: 18631992]
- Cummings JL, Doody R, Clark C. Disease-modifying therapies for Alzheimer disease: challenges to early intervention. *Neurology*. 2007; 69:1622-1634. [PubMed: 17938373]
- Diggle, PJ.; Heagerty, P.; Liang, K-Y.; Zeger, SL. Analysis of Longitudinal Data. 2. New York: Oxford University Press; 2002.
- Fagan AM, Mintun MA, Mach RH, et al. Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid A β_{42} in humans. *Ann Neurol*. 2006; 59:512-519. [PubMed: 16372280]
- Head D, Buckner RL, Shimony JS, Williams LE, Akbudak E, Conturo TE, McAvoy M, Morris JC, Snyder AZ. Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: evidence from diffusion tensor imaging. *Cereb Cortex*. 2004; 14(4):410-23. [PubMed: 15028645]
- Jarjoura D. Crossing controls to treatment in repeated-measures trials. *Controlled Clinical Trials*. 2003; 24:306-323. [PubMed: 12757996]
- Johnson DK, Storandt M, Morris JC, Galvin JE. Longitudinal study of the transition from healthy aging to Alzheimer's disease. *Arch Neurol*. 2009; 66:1254-1259. [PubMed: 19822781]
- Kaduszkiewicz H, Zimmermann T, Beck-Bornholdt H-P, van den Bussche H. Cholinesterase inhibitors for patients with Alzheimer's disease: systematic review of randomised clinical trials. *BMJ*. 2005; 331(7512):321-327. [PubMed: 16081444]
- Kryscio RJ, Mendiondo MS, Schmitt FA, Markesbery WR. Designing a large prevention trial: statistical issues. *Stat Med*. 2004; 23:285-296. [PubMed: 14716729]
- Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*. 1982; 38:963-974. [PubMed: 7168798]
- Leber P. Slowing the progression of Alzheimer's disease: methodologic issues. *Alzheimer Dis Assoc Disord*. 1997; (Suppl 5):S10-S20. [PubMed: 9348423]
- Liu H, Berger RL. Uniformly more powerful, one-sided tests for hypotheses about linear inequalities. *Annals of Statistics*. 1995; 23:55-72.

- Mani RB. The evaluation of disease modifying therapies in Alzheimer's disease: a regulatory viewpoint. *Stat Med.* 2004; 23:305–314. [PubMed: 14716731]
- Mintun MA, LaRossa GN, Sheline YI, et al. [¹¹C] PIB in a nondemented population: Potential antecedent marker of Alzheimer disease. *Neurology.* 2006; 67:446–452. [PubMed: 16894106]
- Montori VM, Guyatt GH. Intention-to-treat principle. *CMAJ.* 2001; 165:1339–41. [PubMed: 11760981]
- Heritier SR, Gebski VJ, Keech AC. Inclusion of patients in clinical trial analysis: the intention-to-treat principle. *Med J Aust.* 2003; 179:438–40. [PubMed: 14558871]
- Qizilbash N, Whitehead A, Higgins J, Wilcock G, Schneider L, Farlow M. for the Dementia Trialists' Collaboration. Cholinesterase inhibition for Alzheimer disease: A meta-analysis of the Tacrine trials. *JAMA.* 1998; 280:1777–1782. [PubMed: 9842955]
- Ringman JM, Grill J, Rodriguez-Agudelo Y, Chavez M, Xiong C. Prevention Trials in Persons At-Risk for Dominantly-Inherited Alzheimer's Disease: Opportunities and Challenges. *Alzheimer's & Dementia.* 2009 in press.
- Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. Donepezil Study Group. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. *Neurology.* 1998; 50(1):136–145. [PubMed: 9443470]
- Rosen WG, et al. A new rating scale for the Alzheimer's disease. *Am J Psychiatry.* 1984; 141:1356–1364. [PubMed: 6496779]
- Salloway S, Mintzer J, Weiner MF, Cummings JL. Disease-modifying therapies in Alzheimer's disease. *Alzheimer's Dement.* 2008; 4(2):65–79. [PubMed: 18631951]
- Sampaio C. Alzheimer disease: disease modifying trials. Where are we? Where do we need to go? A reflective paper. *J Nutr Health Aging.* 2006; 10:113–115. [PubMed: 16554943]
- SAS Institute Inc. SAS Language: Reference, Version 6. 1. Cary, NC: SAS Institute Inc; 1990.
- Storandt M, Grant EA, Miller JP, Morris JC. Longitudinal course and neuropathological outcomes in original versus revised MCI and in PreMCI. *Neurology.* 2006; 67:467–473. [PubMed: 16894109]
- Storandt M, Mintun MA, Head D, Morris JC. Cognitive decline and brain volume loss as signatures of cerebral amyloid- β peptide deposition identified with Pittsburgh compound B. *Arch Neurol.* 2009; 66:1476–1481. [PubMed: 20008651]
- Whitehouse PJ, Kittner B, Roessner M, et al. Clinical trial designs for demonstrating disease-course-altering effects in dementia. *Alzheimer Dis Assoc Disord.* 1998; 12:281–294. [PubMed: 9876956]
- Xiong, C.; Zhu, K.; Yu, K. Statistical modeling in biomedical research: longitudinal data analysis. In: Rao, CR.; Miller, JP.; Rao, DC., editors. *Epidemiology and Medical Statistics.* Amsterdam: Elsevier B.V; 2008.

Figure 1.
Expected Cognitive Progression for Testing Disease Modifying Agents on AD

Table 1

Optimum time of treatment switch (k_0) and optimum sample size allocations (λ_{tb} , λ_{pt}) to treatment groups (in %), given a minimum of 10% subjects assigned to the placebo throughout the trial, as a function of the number of quarterly spaced measures (k), the correlation on the rates of change before and after the treatment switch (ρ), the ratio of between and within subject variances $r_s = \sigma^2 / \sigma_e^2$.

k	k_0	ρ	r_s	$(\lambda_{tb}, \lambda_{pt})$ (in %)
9	4	0.1	1	0.336, 0.564
9	4	0.1	5	0.344, 0.556
9	4	0.1	9	0.344, 0.556
9	4	0.5	1	0.359, 0.541
9	5	0.5	5	0.370, 0.530
9	5	0.5	9	0.371, 0.529
9	5	0.9	1	0.388, 0.512
9	5	0.9	5	0.404, 0.496
9	5	0.9	9	0.406, 0.494
10	5	0.1	1	0.338, 0.562
10	5	0.1	5	0.344, 0.556
10	5	0.1	9	0.345, 0.555
10	5	0.5	1	0.363, 0.537
10	5	0.5	5	0.371, 0.529
10	5	0.5	9	0.372, 0.528
10	5	0.9	1	0.393, 0.507
10	5	0.9	5	0.405, 0.495
10	5	0.9	9	0.407, 0.493
12	6	0.1	1	0.341, 0.559
12	6	0.1	5	0.345, 0.555
12	6	0.1	9	0.345, 0.555
12	6	0.5	1	0.367, 0.533
12	6	0.5	5	0.371, 0.529
12	6	0.5	9	0.372, 0.528
12	6	0.9	1	0.399, 0.501
12	6	0.9	5	0.407, 0.493
12	6	0.9	9	0.407, 0.493

(time unit 1=length of two adjacent measures)

Table 2

Sample sizes for individual treatment arms ($u=tt$ or pt) in a randomized start design required to detect a selected set of effect sizes in differences of slopes (i.e., $(\alpha, \theta)=(\beta_1^{tt}-\beta_1^{pt}, \beta_1^{pt}-\beta_1^{pp})$, slopes are the annual rate of change in the scale of ADAS-cog score) with a statistical power of 80% from the IUT, as well as the optimum treatment switch time (i.e., $1 < k_0 < K$). 10% or 20% subjects are randomized into placebo arm throughout the trial for preserving the blinding of the trial. A correlation of 0.5 (i.e., ρ) between the rates of change before and after the treatment switch (i.e., $(\beta_{1,j}^b, \beta_{1,j}^d)$) from the delayed treatment group is assumed. Quarterly efficacy assessments are assumed.

Duration of the Trial (in years)	k_0	Effect size (α, θ)	% of subjects to $u=pp$	N for $u=tt$	N for $u=pt$	Total N
2	4	(0.5, 0.5)	10	2166	3925	6768
2	4	(0.5, 1.0)	10	1758	3187	5494
2	4	(0.5, 1.5)	10	1758	3186	5493
2	4	(1.0, 0.5)	10	1381	2503	4315
2	4	(1.0, 1.0)	10	541	981	1690
2	4	(1.0, 1.5)	10	445	807	1391
2	4	(1.5, 0.5)	10	1370	2482	4280
2	4	(1.5, 1.0)	10	379	686	1183
2	4	(1.5, 1.5)	10	241	436	752
2	4	(0.5, 0.5)	20	2174	3623	7246
2	4	(0.5, 1.0)	20	1820	3032	6065
2	4	(0.5, 1.5)	20	1820	3033	6065
2	4	(1.0, 0.5)	20	1325	2209	4418
2	4	(1.0, 1.0)	20	544	906	1812
2	4	(1.0, 1.5)	20	459	765	1529
2	4	(1.5, 0.5)	20	1308	2180	4359
2	4	(1.5, 1.0)	20	371	618	1237
2	4	(1.5, 1.5)	20	242	403	806
3	6	(0.5, 0.5)	10	1844	2898	5269
3	6	(0.5, 1.0)	10	1683	2645	4809
3	6	(0.5, 1.5)	10	1683	2645	4809
3	6	(1.0, 0.5)	10	974	1531	2784
3	6	(1.0, 1.0)	10	461	725	1318

Duration of the Trial (in years)	k_0	Effect size (α, θ)	% of subjects to $u=pp$	N for $u=tt$	N for $u=pt$	Total N
3	6	(1.0, 1.5)	10	421	662	1204
3	6	(1.5, 0.5)	10	931	1462	2659
3	6	(1.5, 1.0)	10	290	457	830
3	6	(1.5, 1.5)	10	205	322	586
3	6	(0.5, 0.5)	20	1815	2869	5856
3	6	(0.5, 1.0)	20	1679	2654	5417
3	6	(0.5, 1.5)	20	1679	2654	5417
3	6	(1.0, 0.5)	20	932	1472	3005
3	6	(1.0, 1.0)	20	454	717	1464
3	6	(1.0, 1.5)	20	420	664	1355
3	6	(1.5, 0.5)	20	880	1392	2840
3	6	(1.5, 1.0)	20	281	445	908
3	6	(1.5, 1.5)	20	202	319	651

(k_0 = the number of quarterly spaced measures from baseline, including baseline)