Published in final edited form as:

Clin Trials. 2019 April ; 16(2): 111–119. doi:10.1177/1740774518816323.

Design and Sample Size Considerations for Alzheimer's Disease Prevention Trials Using Multistate Models

Ron Brookmeyer, Ph.D. and

Department of Biostatistics, University of California, 650 Charles Young Drive, Los Angeles, CA South 90095, telephone 310-825-2187

Nada Abdalla, M.S.

Department of Biostatistics, University of California, Los Angeles, CA 90095

Abstract

Background/Aims—Clinical trials for Alzheimer's disease (AD) have been aimed primarily at persons who have cognitive symptoms at enrollment. However, researchers are now recognizing that the pathophysiological process of AD begins years if not decades prior to the onset of clinical symptoms. Successful intervention may require intervening early in the disease process. Critical issues arise in designing clinical trials for primary and secondary prevention of AD including determination of sample sizes and follow-up duration. We address a number of these issues through application of a unifying multistate model for the preclinical course of AD. A multistate model allows us to specify at which points during the long disease process the intervention exerts its effects.

Methods—We used a nonhomogeneous Markov multistate model for the progression of AD through preclinical disease states defined by biomarkers, mild cognitive impairment and AD dementia. We used transition probabilities based on several published cohort studies. Sample size methods were developed that account for factors including the initial preclinical disease state of trial participants, the primary endpoint, age-dependent transition and mortality rates and specifications of which transitions rates are the targets of the intervention.

Results—We find that AD prevention trials with a clinical primary endpoint of mild cognitive impairment or AD dementia will require sample sizes of the order many thousands of individuals with at least 5 years of follow-up which is larger than most AD therapeutic trials conducted to date. The reasons for the large trial sizes include the long and variable preclinical period that spans decades, high rates of attrition among elderly populations due to mortality and losses to follow-up and potential selection effects whereby healthier subjects enroll in prevention trials. A web application is available to perform sample size calculations using the methods reported here.

Conclusion—Sample sizes based on multistate models can account for the points in the disease process when interventions exert their effects and may lead to more accurate sample size determinations. We will need innovative strategies to help design Alzheimer's disease prevention trials with feasible sample size requirements and durations of follow-up.

Keywords

Alzheimer's disease; multistate models; prevention trials; sample size

Introduction

The prevalence of Alzheimer's disease (AD) in the United States will more than double by midcentury because of aging of the population.¹ Currently, only five drugs for the treatment of Alzheimer's dementia have been approved by the Food and Drug Administration and while they temporarily relieve symptoms they do not stop the progressive brain damage.² In recent years a number of promising drugs failed to show clinical benefit in double blind placebo controlled trials in persons with mild to moderate AD dementia.^{3, 4} One explanation for those disappointing findings is that the drugs were administered too late in the disease course, well after the occurrence of irreversible brain damage. To date, clinical trials for Alzheimer's disease have been aimed primarily at persons who already have some cognitive symptoms. However, researchers are recognizing that the pathophysiological process of Alzheimer's disease begins years if not decades prior to onset of clinical symptoms. Successful intervention may require intervening early in the disease process.^{5, 6}

Secondary prevention of Alzheimer's disease refers to slowing disease progression in persons who have some brain pathology who not have cognitive symptoms. Several secondary prevention trials are now underway or in development.⁷ Primary prevention refers to stopping or delaying Alzheimer's associated brain pathology from developing in the first place. Some researchers believe primary prevention will ultimately be the most effective prevention strategy.⁶

A critical issue in designing Alzheimer's disease prevention trials is specification of the primary study endpoint. Candidate primary endpoints include Alzheimer's dementia, or mild cognitive impairment (MCI), Pre-clinical biomarkers, such levels of amyloid deposition or neurodegeneration, have also been proposed. However, an intervention's effect on a preclinical biomarker may not translate to an effect on a clinical endpoint. Persons with preclinical disease identified by biomarkers may in fact never develop clinical symptoms during their lifespan because of the long and variable preclinical period and the high mortality rates among elderly populations.⁸

Another critical design question is whether enrollment criteria should be restricted to those persons at higher risk of developing clinical disease. Enrollment of high risk persons could decrease required sample sizes and study duration that are necessary to achieve adequate statistical power and targets those persons most likely to benefit from trial participation. However, if enrolled individuals have preclinical disease that is too advanced, thereby rendering the intervention ineffective, then the trial could result in a negative study and a potentially useful intervention if it were administered earlier would have been discarded. The Anti-Amyloid Treatment for Asymptomatic AD Trial (A4 Trial) is enrolling persons with evidence of amyloid accumulation (amyloidosis) who do not have any clinical symptoms and randomizing them to placebo or an anti-amyloid agent.⁹

Sample sizes of major cardiovascular prevention trials have been large. The Multiple Risk Factor Intervention Trial which evaluated interventions to lower mortality from coronary heart disease enrolled nearly 13,000 individuals with average follow-up of seven years.¹⁰ The Physician Health Study to evaluate the effects of aspirin on reducing cardiovascular disease mortality enrolled approximately 22,000 physicians with average follow-up of five years.¹¹ To date, there have been no Alzheimer's disease prevention trials either underway or in development with sample sizes that large or follow-up durations that long.

Sample sizes and follow-up requirements for Alzheimer's prevention trials could be expected to be large for several reasons. If the intervention is only effective at the early stages of AD pathophysiology, long follow-up periods and large sample sizes will be required to observe sufficient numbers of clinical events because the preclinical period could span decades. Elderly populations have higher rates of mortality increasing trial attrition rates. Persons who enroll in prevention trials might be healthier and have lower disease progression rates. Selection of unusually healthy populations in prevention trials has led to underpowered studies.¹²

Here we propose methods for determining sample sizes and follow-up in Alzheimer's prevention trials based on a multistate model for preclinical disease that accounts for the points along the disease process when the intervention exerts its effects. Multistate models have been used to aid in the design of cancer clinical trials¹³ but not to specify the transitions affected by interventions. We evaluate sample size requirements based on design variables such as the choice of the primary endpoint, follow-up duration, enrollment criteria, and the specific transition rates affected by interventions.

Methods

Multistate Model for Preclinical Alzheimer's disease

We used a multistate model for the progression of AD through preclinical disease states, mild cognitive impairment and AD dementia (Figure 1). The model is based on the National Institute of Aging-Alzheimer's Association frameworks for the preclinical stages of AD and biological definition of Alzheimer's disease.^{14, 15} The biological definition requires presence of β -amyloid ($A\beta$) plaques to meet criteria for AD dementia. The model postulates that a pathway that leads to AD dementia (solid (red) arrows in Figure 1) is sequential progression through the following states: normal (state 1); asymptomatic amyloidosis which refers to $A\beta$ deposition (state 2); amyloidosis and neurodegeneration (state 4); mild cognitive impairment due to AD with both amyloidosis and neurodegeneration present (state 5); and AD dementia (state 7). Evidence supporting an alternative pathway leading to AD dementia (state 7) has also been described in which neurodegeneration arises prior to amyloidosis and is indicated by the dashed (blue) arrows from states 1 to 3 and from states 3 to 4.¹⁶ State 6 in Figure 1 represents a pathway to other non-AD types of dementia (state 8).

Amyloidosis can be detected by specific biomarkers for $A\beta$ accumulation such as positron emission tomography (PET) amyloid imaging or low $A\beta$ 42 in the cerebrospinal fluid (CSF). Neurodegeneration can be detected by biomarkers including elevated CSF tau, neuronal

dysfunction based on fluorodeoxyglucose (FDG) PET, or hippocampal atrophy/cortical thinning on volumetric magnetic resonance imaging (MRI).

We implemented the multistate model of Figure 1 as a discrete time nonhomogeneous Markov model in which the transition rates depend on current age. The transition probability $p_{ij}(a, g)$ is the probability that a person who is age a , gender g and in state i will transition to state j one time unit later (we used time units of 0.25 years). We express $p_{ij}(a, g)$ for $i \neq j$ as

$$p_{ij}(a, g) = r_{ij}(a) (1 - d_i(a, g))$$

where the first factor, $r_{ij}(a)$, represents the probability that a person who is in state i at age a transitions to state j in the following time unit in the absence of competing risks of death. The second factor $d_i(a, g)$ represents the probability a person of gender g (male or female) who is alive at age a and in state i does not survive the subsequent time unit. The probability of remaining in the same state in the following time unit is $p_{ii}(a, g) = 1 - \sum_{i \neq j} p_{ij}(a, g)$

We used transition probabilities $r_{ij}(a)$ from several published cohort studies including the Mayo Clinic Study of Aging¹⁷ and an international multi-center cohort study¹⁸ (see Supplementary Material and reference¹). AD incidence rates derived from the multistate model are consistent with a systematic review of the worldwide literature of AD incidence.¹ We used U.S. 2014 death rates by age and gender.¹⁹ Studies²⁰ have reported that the hazard ratio of death for persons with MCI compared to those without MCI was 1.65 and we multiplied the background death rates by that factor to obtain the death rates for persons with mild cognitive impairment (state 5).

Modeling Intervention Effects

We modeled the effects of an intervention by multiplying the transition probabilities $r_{ij}(a)$ by proportionality constants θ_{ij} . The proportionality constants θ_{ij} characterize the effectiveness of the interventions and specify which transition rates are the targets of the intervention.

We considered several intervention scenarios. We considered a disease modifying secondary prevention intervention, such as an anti-amyloid agent, that lowered the probability of transitioning from amyloidosis to neurodegeneration. We modeled the effects of such an intervention by choosing values for θ_{24} less than 1 (for example, $\theta_{24} = 0.50$ or 0.67) and setting other values for θ_{ij} equal to 1. We considered a primary prevention intervention that lowered the probabilities of developing amyloidosis in the first place and modeled its effects choosing θ_{12} and θ_{34} less than 1 (for example, $\theta_{12} = \theta_{34} = 0.50$ or 0.67) and setting other values for θ_{ij} equal to 1.

Sample Size Considerations

We considered a randomized prevention trial in which equal numbers of individuals are assigned to the control and intervention arms using simple randomization (rather than stratified designs). Individuals are followed until they reach the primary study endpoint (e.g., AD dementia or mild cognitive impairment due to AD). Each person will be followed for a

maximum of T years; however, some persons may die prior to completion of follow-up. We first consider the case when the only reason for incomplete follow-up (attrition) is death. At the end of this section we consider the situation of incomplete follow-up for reasons other than death.

The probabilities that an individual in the control and intervention arms experience the primary endpoint during follow-up (which implies they have not died first) are called P_1 and P_2 , respectively. If the only reason for attrition is death, then the probabilities P_1 and P_2 are “absolute risks.” Absolute risks refer to the probabilities of developing a clinical condition over a time period and are calculated in the presence of risks of death rather than under a scenario in which risks of death are eliminated. Absolute risks depend on mortality rates.^{21–23} The importance of absolute risk in clinical practice and public health has been discussed in the literature.^{21–23}

We assume the death rates are the same in the two randomized arms of the trial. If there is no effect of the intervention, then the θ_{ij} 's are all equal to 1, which implies $P_1 = P_2$. We test the null hypotheses $H_0: P_1 = P_2$ using a two sample test for binomial probabilities. We can determine sample sizes to have a desired level of statistical power at specified alternative values P_1 and P_2 using a two sided test of level α . The values of the probabilities P_1 and P_2 depend on multiple factors: choice of primary endpoint; follow-up time T ; intervention effect θ_{ij} 's; and the mix of preclinical disease states, ages, and genders at enrollment. The Supplementary Material describes how P_1 and P_2 are determined from these factors. We used a built-in function in the R statistical computing package (`power.prop.test`)²⁴ to determine the total trial sample size from both groups N ($N/2$ in each group).

We considered the impact of ages at enrollment on sample size requirements. We assumed the numbers of males and females trial participants are approximately equal and the distributions of ages at enrollment are the same for males and females. Let $w(a)$ represent the proportion of participants enrolled at age a . We considered three age distributions: a young age distribution where $w(a) = 0.25$ at ages $a = 60, 65, 70$ and 75 (and 0 at other ages) which has a mean age of 67.5; a uniform age distribution where $w(a) = 0.167$ at ages $a = 60, 65, 70, 75, 80$ and 85 which has a mean age of 72.5; an old age distribution where $w(a) = 0.25$ at ages $a = 70, 75, 80$ and 85 which has a mean age of 77.5.

We evaluated the impact on sample sizes for several types of trials with different enrollment criteria for preclinical disease: trial that enrolled persons with amyloidosis who are in state 2; trials that enrolled persons without any brain pathology who are in the normal state 1; and trials that did not screen with biomarkers for preclinical disease state but rather enrolled a random sample of persons in preclinical states (1 through 4).

Selective enrollment of persons in good health and exclusion of persons who have chronic illness or other disabilities may affect the power of prevention trials.¹² We investigated the sensitivity of sample sizes to healthy selection bias. Suppose the death rates for persons who enroll in prevention trials are $\lambda d(a, g)$ where the factor λ accounts for possible healthy selection bias on mortality. If trial enrollees are especially healthy we would expect λ to be less than 1. It is also conceivable enrollees in prevention trials have lower disease

progression rates. Suppose the transition probabilities for trial enrollees are $pr_{ij}(a)$ where the factor ρ accounts for possible selection effects on disease transition probabilities. We considered various values for λ and ρ to evaluate the sensitivity of sample sizes to healthy selection effects.

We considered the impact of incomplete follow-up (attrition) for reasons other than death (i.e., lost- to follow-up). Let ψ be the probability of being lost to follow-up in a time unit for reasons other than death. If we replace $d_i(a, g)$ by $d_i(a, g) + \psi$, then the methods described in this section can also be used to calculate required sample sizes accounting for all reasons for attrition. Note that if there is attrition for reasons other than death, then the probabilities P_1 and P_2 can no longer be interpreted as absolute risks because persons lost to follow-up still have an opportunity to reach the primary endpoint.

A Shiny web application that performs the sample size calculations described in this paper is available.²⁵ The underlying R source code is provided in the app.

Results

We considered a trial to evaluate an intervention that slows progression from state 2 (amyloidosis) to state 4 ($\theta_{24} = .50$ or 0.67 and all other $\theta_{ij} = 1.0$). The intervention is “disease modifying” in that it delays the onset of clinical endpoints (e.g. MCI and AD dementia). The trial enrolls persons with amyloidosis (state 2) who are followed for a maximum of $T=5$ years or 7 years. The primary endpoint is MCI (state 5). Table 1 shows total required sample sizes (N) (both groups combined) to have 90% power with a two sided test significance level .05. The distributions of ages at enrollment have a strong effect on sample size requirements. Lower sample size requirements arise with older populations because transition rates $r_{ij}(a)$ increase with age and that effect overwhelms higher death rates in older populations. For example, the sample size requirements for a 5 year study with $\theta_{24}=0.5$ are 6505, 3274 and 2318 with young, uniform and old age distributions respectively.

Table 2 shows required sample sizes for a trial with the same design parameters as Table 1 except the primary endpoint is AD dementia (state 7) rather than MCI (state 5). The sample sizes in Table 2 are greater than in Table 1 by factors of approximately 3.3 and 2.6 for 5 year and 7 years studies respectively. Sample sizes greater than 10,000 are required in many situations in Table 2. If the intervention’s effect is strong ($\theta_{24}=0.5$) and the duration of follow-up is 7 years sample sizes are reduced to under 10,000.

We considered a trial to evaluate an intervention that prevents both the occurrence of amyloidosis (primary prevention) and progression from amyloidosis to neurodegeneration (secondary prevention) with MCI as the primary endpoint (state 5). We specified the effects of such an intervention by $\theta_{12} = \theta_{24} = \theta_{34} = 0.50$. First, we considered enrolling persons who are in the normal state (state 1). As shown in Table 3, with only $T=5$ years, enrollment of normal individuals (state 2) with a young age distribution resulted in a sample size requirement of nearly 23,000; with $T=7$ years sample sizes were less than 10,000 for the three age distributions. We considered enrolling a random sample of persons in preclinical states 1–4 without any biomarker pre-screening. We find that extremely large sample sizes

are required with random sampling of persons in pre-clinical disease states. The reason for these large sample sizes is because without biomarker prescreening many persons enrolled could be in state 4 and thus would not be helped by this particular intervention.

We considered using a preclinical biomarker as the primary endpoint rather than a clinical endpoint. Table 4 shows the sample size requirements for a primary prevention trial in which normal individuals (state 1) are followed for the primary endpoint of amyloidosis and neurodegeneration (state 4). We considered a disease modifying intervention that prevents the occurrence of amyloid which was modeled by setting θ_{12} and θ_{34} equal to 0.50. We considered $T=3, 5$ and 7 years. Sample size requirements with preclinical biomarker endpoint were considerably less than with clinical endpoints. Even with only three years of follow-up, sample size requirements with young, uniform and old age distributions were 6533, 3455 and 2488, respectively. However, the cautionary warning with using a preclinical biomarker as the primary endpoint is that an intervention's effect on biomarker (surrogate) endpoint does not necessarily imply an effect on a clinical endpoint such as MCI and AD dementia.

Sensitivity analyses were performed to evaluate the impact on sample sizes of particularly healthy enrollees. A sensitivity analysis was performed using various values of λ and ρ for a prevention trial that enrolled persons with amyloidosis (state 2) with $T=5$ years, a primary endpoint of MCI (state 5) and $\theta_{24}=0.50$ (Table 5). If healthy section effects decrease death rates by 50% but do not change the transition probabilities ($\lambda=0.50$ and $\rho=1.0$), we find that required sample sizes actually decrease slightly compared to the case $\lambda=1.0$ and $\rho=1.0$. The reason for the lower sample sizes is that lower death rates increase the likelihood that persons will have a primary event during follow-up. However, if the transition probabilities are also decreased say by 50% ($\lambda=0.50$ and $\rho=0.50$), the sample sizes increase very substantially.

We evaluated the impact of attrition for reasons other than death (i.e., lost- to follow-up) on sample size requirements. We assumed an annual lost to follow-up probability of $\psi=0.05$ and recalculated the sample sizes in Table 1. The samples sizes with $\psi=0.05$ per year compared to those in Table 1 increased on average across the scenarios considered by approximately 26.1% with $T=5$ years and 32.6% with $T=7$ years. We assumed that the lost-to follow-up rate ψ did not depend on age or gender, but if ψ increased sufficiently with advancing age, then the smaller sample size advantage of older populations compared to younger populations observed in Tables 1–5 could disappear.

Discussion

Advances in our understanding of preclinical Alzheimer's disease are opportunities to improve the design of prevention trials.²⁶ The multistate model provides a framework for evaluating sample size requirements that account for which transitions rates are affected by intervention, the primary endpoint, follow-up duration, enrollment criteria, attrition from mortality, attrition from losses to follow-up and healthy selection effects.

The effect sizes of interventions can be expressed as either the θ_{ij} 's or alternatively as the ratio of absolute risks ($R=P_2/P_1$). The θ_{ij} 's specify exactly at which points during the disease process that the intervention is exerting its effects. The two effect size measures, R and the θ_{ij} 's are not the same numerically. To illustrate, consider the situation in Table 3 with $T=5$ years and the uniform age distribution where $\theta_{12}=\theta_{24}=\theta_{34}=0.50$. If trial participants were in the normal state we find R is equal to 0.40. The interpretation is that the intervention could prevent about 60% of incident MCI's over 5 years. If trial participants were a random sample of enrollees from the preclinical states, we find R is equal to .93 which indicates that the intervention prevents only about 7% of incident MCIs over 5 years. The ratio R depends on the θ_{ij} 's, the death rates, the transition rates in the control group and the distribution of ages and preclinical disease states at enrollment.

Sample size calculations for Alzheimer's trials have varied widely in the literature, in part from uncertainties in the various inputs into the calculations.²⁷ An important input into our sample size calculations were the transition rates. We utilized transition rates based on the largest studies of their kind to date. Nevertheless, we recognize that these rates may not be applicable to all study populations. We relied on the Mayo Clinic Study of Aging which is a population based cohort study in Olmsted County Minnesota but is not ethnically diverse. Data from additional studies will lead to improved sample size estimates for prevention trials.

We considered primary endpoints that were binary states in the multistate model in which persons either reach the primary endpoint (e.g., MCI, or AD dementia) during follow-up or they do not. An alternative is to base the primary endpoint on a battery of neuropsychological tests summarized by a quantitative composite score.^{28, 29} The analysis of the composite scores could be based on the mean change or rate of change of the scores using longitudinal mixed effects models. Important work has been undertaken to develop methods for defining and optimizing composite scores from a battery of tests.^{29, 30} An advantage of using a composite score is it may be an earlier indicator of an intervention's effect than endpoint such as MCI which may lead to trials with smaller sample sizes or follow-up durations. However, if an intervention has an effect on a composite cognitive score it does not imply it would also delay onset of MCI or AD dementia. Important questions with an approach based on either individual or composite neuropsychological test scores are how large of a difference in scores or rates of change between groups would represent a meaningful clinical effect and how to translate that effect size to be understandable. Useful work with the approach has been undertaken to evaluate the value of enriching trial populations with participants who have risk factors for disease progression such as apolipoprotein $\epsilon 4$ carrier status, or preclinical disease state.^{28, 31}

Inclusion of additional states in the multistate model such as subtle cognitive and behavioral decline prior to MCI¹⁴ would provide an early indication of an intervention's effect. Incorporation of a state based on biomarkers for tau pathology¹⁵ would increase the specificity that an individual is indeed in the Alzheimer's continuum. Subdivision of disease states into quantitative categorical levels such as by amyloid load may also be useful. Figure 1 is a progressive disease model. However, back transitions may occur, as for example if

persons with mild cognitive impairment transition back to an asymptomatic state or if an anti-amyloid drug promotes transition from amyloidosis back to normal.

If enrollment criteria are based on preclinical disease state, then biomarker screening must be performed to determine trial eligibility. Large numbers of potential trial participants will need to be screened to identify the required numbers of participants. For example, consider a prevention trial that enrolls persons with amyloidosis (state 2). The prevalence of amyloidosis (state 2) among person without clinical disease is approximately 20.6% (assuming a uniform age distribution between ages 60 and 85; see Supplementary Material). Thus, the numbers of preclinical individuals who would need to be screened for biomarkers in order to enroll the required sample sizes is approximately $1/.206=4.85$ times the numbers shown in Table 1. Biomarker screening to identify the required numbers of trial participants is an important consideration in assessing the cost and feasibility of designs of AD prevention trials.

An important question is whether supplementary information on apolipoprotein (APOE) $\epsilon 4$ carrier status, in addition to that on preclinical biomarkers would be useful. Some studies have suggested that APOE $\epsilon 4$ carrier status increases risk of amyloidosis, although it is unclear if APOE $\epsilon 4$ carriers are at additional risk once amyloidosis occurs.³² In any case in large randomized prevention trials the expectation is that the groups will be approximately balanced on known and unknown risk factors.

Alzheimer's disease prevention trials with primary endpoints of mild cognitive impairment or AD dementia will require sample sizes of the order of many thousands of individuals with at least 5 years of follow-up. These numbers are larger than most AD therapeutic trials conducted to date. Prevention trials of that magnitude will stretch limited resources especially as more candidate interventions are developed that require rigorous evaluations. We will need innovative strategies to help design prevention trials with feasible sample size requirements. Such strategies could include enriching prevention trials with participants who have preclinical disease, or using earlier endpoints in the disease process such as subtle cognitive changes before mild cognitive impairment. Sample sizes based on multistate models can account for the points in the disease process when interventions exert their effects and may lead to more accurate sample size determinations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This research was funded by a grant from the National Institutes of Health (1R21AG055361).

References

1. Brookmeyer R, Abdalla N, Kawas CH, Corrada MM. Forecasting the prevalence of preclinical and clinical Alzheimer's disease in the United States. *Alzheimer's & Dementia* 2018; 14:121–129.
2. Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimer's Research & Therapy* 2014; 6:37.

3. Amanatkar HR, Papagiannopoulos B, Grossberg GT. Analysis of recent failures of disease modifying therapies in Alzheimer's disease suggesting a new methodology for future studies. *Expert review of Neurotherapeutics* 2017 ;17:7–16. [PubMed: 27223100]
4. Cummings J Lessons learned from Alzheimer disease: Clinical trials with negative outcomes. *Clinical and Translational Science* 2018;11:147–52. [PubMed: 28767185]
5. Selkoe DJ. Preventing Alzheimer's disease. *Science* 2012; 337:1488–92. [PubMed: 22997326]
6. McDade E, Bateman RJ. Stop Alzheimer's before it starts. *Nature News* 2017; 547:153–155.
7. Carrillo MC, Brashear HR, Logovinsky V, Ryan JM, Feldman HH, Siemers ER, Abushakra S, Hartley DM, Petersen RC, Khachaturian AS, Sperling RA. Can we prevent Alzheimer's disease? Secondary "prevention" trials in Alzheimer's disease. *Alzheimer's & dementia: The Journal of the Alzheimer's Association* 2013; 9:123–31.
8. Brookmeyer R and Abdalla N. Estimation of lifetime risks of Alzheimer's disease dementia using biomarkers for preclinical disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 2018; 10.1016/j.jalz.2018.03.005
9. Sperling RA, Rentz DM, Johnson KA, Karlawish J, Donohue M, Salmon DP, Aisen P. The A4 study: stopping AD before symptoms begin? *Science Translational Medicine* 2014; 6:228fs13-.
10. Multiple Risk Factor Intervention Trial Research Group. Multiple Risk Factor Intervention Trial: risk factor changes and mortality results. *JAMA* 1982; 248:1465–77. [PubMed: 7050440]
11. Steering Committee of the Physicians' Health Study Research Group*. Final report on the aspirin component of the ongoing Physicians' Health Study. *New England Journal of Medicine* 1989; 321(3):129–35. [PubMed: 2664509]
12. Ederer F, Church TR, Mandel JS. Sample sizes for prevention trials have been too small. *American Journal of Epidemiology* 1993; 137(7):787–96. [PubMed: 8484370]
13. Xia F, George SL, Wang X. A Multi-State Model for Designing Clinical Trials for Testing Overall Survival Allowing for Crossover after Progression. *Statistics in Biopharmaceutical Research* 2016; 8(1):12–21. [PubMed: 27239255]
14. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, ... & Park DC Toward defining the preclinical stages of AD: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for AD. *Alzheimer's & Dementia* 2011; 7: 280–292
15. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Elliott C. NIA-AA research framework: towards a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 2018; 14:535–562.
16. Knopman DS, Jack CR, Wiste HJ, Weigand SD, Vemuri P, Lowe VJ, ... & Roberts RO (2013). Brain injury biomarkers are not dependent on β -amyloid in normal elderly. *Annals of Neurology* 2013; 73(4), 472–480 [PubMed: 23424032]
17. Jack CR, Therneau TM, Wiste HJ, Weigand SD, Knopman DS, Lowe VJ, ... & Senjem ML Transition rates between amyloid and neurodegeneration biomarker states and to dementia: a population-based, longitudinal cohort study *The Lancet Neurology* 2016; 15: 56–64. [PubMed: 26597325]
18. Vos SJ, Verhey F, Frölich L, Kornhuber J, Wiltfang J, Maier W, Peters O, Rütther E, Nobili F, Morbelli S, Frisoni GB Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. *Brain* 2015; 138:1327–1338. [PubMed: 25693589]
19. Human Mortality Database University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org (download 12/09/2015).
20. Vassilaki M, Cha RH, Aakre JA, Therneau TM, Geda YE, Mielke MM, Knopman DS, Petersen RC, Roberts RO. Mortality in mild cognitive impairment varies by subtype, sex, and lifestyle factors: the mayo clinic study of aging. *Journal of Alzheimer's Disease* 2015; 45:1237–45
21. Gail MH. Personalized estimates of breast cancer risk in clinical practice and public health. *Statistics in Medicine* 2011; 30:1090–104. [PubMed: 21337591]
22. Pfeiffer R and Gail MH Absolute Risk: Methods and Applications in Clinical Management and Public Health, Monographs on Statistics and Applied Probability 154, Boca Raton FL: CRC Press; 2017.

23. Seshadri S, Wolf PA. Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham Study. *The Lancet Neurology* 2007; 6:1106–14. [PubMed: 18031707]
24. R Core Team. R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria 2017 <https://www.R-project.org/>
25. Abdalla N and Brookmeyer R. Sample Sizes for Alzheimer's Disease Prevention Trials Available at <https://alzheimers-stats.shinyapps.io/samplesizes/> (accessed 10/22/2018).
26. Sperling R, Mormino E, Johnson K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. *Neuron* 2014; 84:608–22. [PubMed: 25442939]
27. Ard MC, Edland SD. Power calculations for clinical trials in Alzheimer's disease. *Journal of Alzheimer's Disease* 2011 1 1; 26(s3):369–77.
28. Insel PS, Mattsson N, Mackin RS, Kornak J, Nosheny R, Tosun-Turgut D, Donohue MC, Aisen PS, Weiner MW, Alzheimer's Disease Neuroimaging Initiative. Biomarkers and cognitive endpoints to optimize trials in Alzheimer's disease. *Annals of Clinical and Translational Neurology* 2015; 2:534–47. [PubMed: 26000325]
29. Donohue MC, Sun CK, Raman R, Insel PS, Aisen PS, North American Alzheimer's Disease Neuroimaging Initiative, Japanese Alzheimer's Disease Neuroimaging Initiative. Cross-validation of optimized composites for preclinical Alzheimer's disease. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 2017; 3:123–9.
30. Edland SD, Ard MC, Li W, Jiang L. Design of pilot studies to inform the construction of composite outcome measures. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 2017; 3:213–8.
31. Grill JD, Di L, Lu PH, Lee C, Ringman J, Apostolova LG, Chow N, Kohannim O, Cummings JL, Thompson PM, Elashoff D. Estimating sample sizes for predementia Alzheimer's trials based on the Alzheimer's Disease Neuroimaging Initiative. *Neurobiology of Aging* 2013; 34:62–72. [PubMed: 22503160]
32. Morris JC, Roe CM, Xiong C, Fagan AM, Goate AM, Holtzman DM, Mintun MA. APOE predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. *Annals of Neurology* 2010; 67:122–31. [PubMed: 20186853]

Figure1. The multistate model for the progression of Alzheimer's disease. Persons progress through preclinical states including amyloidosis and neurodegeneration, mild cognitive impairment (MCI), and Alzheimer's dementia.

Table 1.

Sample sizes (N) for a secondary prevention trial enrolling persons with amyloidosis (state 2) with MCI (state 5) as the primary endpoint for three age distributions at enrollment. Follow-up durations are $T=5$ years and $T=7$ years. The intervention's effects are $\theta_{24}=0.50$ and $\theta_{24}=0.67$. P_1 and P_2 are the absolute risks of occurrence of the primary endpoint within T years. Power=0.90 and two sided significance level=.05

	$\theta_{24}=0.50$						$\theta_{24}=0.67$					
	$T=5$ years			$T=7$ years			$T=5$ years			$T=7$ years		
AGE	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N
Young	0.023	0.013	6505	0.047	0.026	3520	0.023	0.016	17452	0.047	0.034	9671
Uniform	0.051	0.029	3274	0.088	0.052	2118	0.051	0.037	9030	0.088	0.066	6026
Old	0.071	0.040	2318	0.121	0.072	1524	0.071	0.052	6408	0.121	0.091	4352

Age distributions are: young is $w(a)=0.25$ at ages $a=60, 65, 70$ and 75 ; uniform is $w(a)=0.167$ at ages $60, 65, 70, 75, 80$ and 85 ; old is $w(a)=0.25$ at ages $70, 75, 80$ and 85 . Incomplete follow-up is due only to death. N is the sample size from both groups combined ($N/2$ per group).

Table 2.

Sample sizes (N) for a secondary prevention trial enrolling persons with amyloidosis (state 2) with AD dementia (state 7) as the primary endpoint for three age distributions at enrollment. Follow-up durations are $T=5$ years and $T=7$ years. The intervention's effects are $\theta_{24}=0.50$ and $\theta_{24}=0.67$. P_1 and P_2 are the absolute risks of occurrence of the primary endpoint within T years. Power=0.90 and two sided significance level=.05.

	$\theta_{24}=0.50$						$\theta_{24}=0.67$					
	$T=5$ years			$T=7$ years			$T=5$ years			$T=7$ years		
AGE	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N
Young	0.007	0.004	20727	0.018	0.010	8840	0.007	0.005	54951	0.018	0.013	23906
Uniform	0.014	0.008	10983	0.031	0.018	5633	0.014	0.010	29719	0.031	0.022	15635
Old	0.020	0.011	7898	0.042	0.024	4148	0.020	0.014	21424	0.042	0.031	11563

Age distributions are: young is $w(a)=0.25$ at ages $a=60, 65, 70$ and 75 ; uniform is $w(a)=0.167$ at ages $60, 65, 70, 75, 80$ and 85 ; old is $w(a)=0.25$ at ages $70, 75, 80$ and 85 . Incomplete follow-up is due only to death. N is the sample size from both groups combined ($N/2$ per group).

Table 3.

Sample sizes (N) for a prevention trials enrolling normal persons at enrollment (state 1) and a random sample of persons in the preclinical states (without biomarker screening) for three age distributions at enrollment. MCI (state 5) is the primary endpoint. Follow-up durations are $T=5$ years and $T=7$ years. The intervention's effects are $\theta_{12} = \theta_{24} = \theta_{34} = 0.50$. P_1 and P_2 are the absolute risks of occurrence of the primary endpoint within T years. Power=0.90 and two sided significance level=.05.

	Enroll normals (state 2)						Enroll sample of preclinical states 1–4					
	$T=5$ years			$T=7$ years			$T=5$ years			$T=7$ years		
AGE	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N
Young	0.003	0.001	22907	0.009	0.003	8358	0.035	0.031	83579	0.056	0.047	29220
Uniform	0.010	0.004	9088	0.024	0.011	4176	0.093	0.086	72588	0.128	0.116	29694
Old	0.015	0.007	6252	0.035	0.016	2905	0.135	0.125	51714	0.184	0.167	21607

Age distributions are: young is $w(a)=0.25$ at ages $a=60, 65, 70$ and 75 ; uniform is $w(a)=0.167$ at ages $60, 65, 70, 75, 80$ and 85 ; old is $w(a)=0.25$ at ages $70, 75, 80$ and 85 . Incomplete follow-up is due only to death. N is the sample size from both groups combined ($N/2$ per group).

Table 4.

Sample sizes (N) for a primary prevention trial enrolling normal persons (state 1) with a preclinical biomarker primary endpoint defined as amyloidosis and neurodegeneration (state 4)). Follow-up durations are $T=3, 5$, and 7 years. The intervention's effects are $\theta_{12} = \theta_{34} = 0.50$. P_1 and P_2 are the absolute risks of occurrence of the primary endpoint within T years. Power=0.90 and two sided significance level=.05.

	$T=3$ years			$T=5$ years			$T=7$ years		
AGE	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N
Young	0.020	0.011	6533	0.056	0.030	2484	0.104	0.057	1377
Uniform	0.040	0.021	3455	0.095	0.052	1534	0.156	0.088	982
Old	0.055	0.029	2488	0.128	0.071	1120	0.205	0.117	734

Age distributions are: young is $w(a)=0.25$ at ages $a=60, 65, 70$ and 75; uniform is $w(a)=0.167$ at ages 60, 65, 70, 75, 80 and 85; old is $w(a)=0.25$ at ages 70, 75, 80 and 85. Incomplete follow-up is due only to death. N is the sample size from both groups combined ($N/2$ per group).

Table 5.

Sensitivity analysis of sample sizes to healthy selection effects as characterized by λ (mortality relative risk) and ρ (transition relative risk). Sample sizes are for a secondary prevention trial that enrolls persons with amyloidosis (state 2) and with MCI (state 5) as the primary endpoint. Follow-up duration is $T=5$ years. The intervention's effect is $\theta_{24}=0.50$. P_1 and P_2 are the absolute risks of occurrence of the primary endpoint within T years. Power=0.90 and two sided significance level=.05.

	$\lambda=1.0, \rho=1.0$			$\lambda=0.5, \rho=1.0$			$\lambda=0.5, \rho=0.67$			$\lambda=0.5, \rho=0.5$		
AGE	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N	P_1	P_2	N
Young	0.023	0.013	6505	0.025	0.013	6073	0.012	0.006	11914	0.007	0.004	20008
Uniform	0.051	0.029	3274	0.061	0.035	2804	0.031	0.017	5020	0.018	0.010	8025
Old	0.071	0.040	2318	0.086	0.049	1959	0.044	0.024	3511	0.026	0.014	5611

Age distributions are: young is $w(a)=0.25$ at ages $a=60, 65, 70$ and 75 ; uniform is $w(a)=0.167$ at ages $60, 65, 70, 75, 80$ and 85 ; old is $w(a)=0.25$ at ages $70, 75, 80$ and 85 . Incomplete follow-up due only to death. N is the sample size from both groups combined ($N/2$ per group).