Supplemental Text, Tables, and Figures

Analysis of Delayed Start Randomized Controlled Trials

There has been substantial interest in recent years in distinguishing between treatments that alter the underlying disease process (disease-modifying treatments) and those that only affect downstream consequences while allowing the disease process to continue (symptomatic treatments) [1]. Delayed-start (sometimes referred to as randomized start) trials have been proposed as a means of discriminating between these two types of effects [2,3]. Essentially, the thinking is that if a treatment effect is disease modifying, then during phase 1 of the trial, participants randomized to treatment will have a decreased rate of worsening on the primary outcome, as compared to those randomized to placebo. During phase 2 of the trial (when participants originally randomized to placebo also receive the intervention), the rate of progression of both groups will be similar, but the groups will differ consistently in their outcome means. By contrast, if a treatment is merely symptomatic, while the rates of worsening will differ during phase 1, participants originally randomized to placebo will "catch up" to the group that has been receiving the intervention from the start of the trial (Figure S1.) [3]. Though not shown in figure S1, with a symptomatic treatment one might also see the two trajectories converge at the end of phase 2 through changes in both treatment group trajectories.

Testing for a disease-modifying treatment within this framework entails two different hypothesis tests; first one tests whether those in the "early on" arm have less worsening than those in the "delayed on" arm over both phases. This can be visualized as the vertical distance between the mean trajectories at the end of phase 2. Second, one tests whether the slopes of the two arms are equivalent during phase 2 with a preselected margin of noninferiority [2].

In practice, this conceptualization may be overly simplistic, and it relies on a number of assumptions that may not be tenable. First, it assumes that the trajectories of worsening in both

arms are linear, though several cohort studies have reported accelerating decline in AD [4,5]. Further, it assumes that the effect of the intervention on either the underlying disease process or symptomatology does not vary as a function of when in the course of disease the intervention occurs. Assuming that the treatment arms do not vary with regard to disease stage at baseline, this would mean that those in the delayed on arm would receive the intervention at a later stage of the disease than those randomized to the early on arm. There is a precedent for such stage-specific effects; the timing of exposure to non-steroidal anti-inflammatory treatments has been proposed to explain the discrepancy between observational studies demonstrating an association between NSAID use and decreased risk of AD and AD treatment trials demonstrating null effects [6–8].

The choice of length of follow-up for both phases can also substantially affect the interpretation of the results. If there is narrowing of the between-arm difference during phase 2 relative to phase 1, but not enough for the two arms to meet, it may be difficult to determine whether the treatment is disease-modifying, or if there has simply not been enough time for the delayed on arm to catch up. There is also an inherent assumption that there will be a between-arm difference on the outcome at the end of phase 1, and the margin of noninferiority is chosen *a priori* based on the expected difference. To address this, use of a pre-specified fraction of the between-arm difference at the end of phase 1 has been proposed, rather than an absolute difference [9].

Linear Spline Models for Multiphase Trials

In this case, no between-arm differences were observed at the end of phase 1, and so the typical approaches for analyzing data from delayed start trials have limited applicability to our data. Those findings may indicate that DBS-f is not efficacious for mild AD. If that is the case, we would expect to see in phase 2 that the mean trajectories of both arms would continue to be

identical. If the rate of worsening for both arms is constant across the two phases of the trial, then the trajectories would appear to be two straight lines on top of each other. If the rate of worsening increases during the second year, but equally for both arms, then the trajectories would resemble two broken sticks. Such trajectories can be modeled by including a linear spline "knot" at twelve months, thus allowing for a different rate of worsening after the first year. If the rate of worsening in the early on arm improves in phase 2 relative to its rate of worsening in phase 1, this might imply that DBS-f is efficacious, but that prolonged treatment is required in order to observe the effect. To examine these possibilities, it is necessary to model outcome trajectories as a function of treatment arm and phase. Rates of worsening can then be compared both within arms between phases, and between arms. We will fit a mixed effects model with a random intercept such that $E[y_{ij}]= \beta_0+\beta_1 time+\beta_2 arm+\beta_3 time*arm+\beta_4 time^{12+} + \beta_5 time^{12+}*arm$, where y_{ij} is the outcome for the *i*th participant at the *j*th timepoint, β_1 is the expected rate of change for the delayed on arm during phase 1, β_3 is the expected difference in rate of change between the two arms during phase 1, β_4 is the expected difference in rate of change for participants in the delayed on arm between phase 1 and phase 2, and β_5 is the expected difference in change in rate of change between phase 1 and phase 2, between treatment arms. Figure S2 demonstrates how various scenarios would manifest themselves through this model.

Since *post-hoc* subgroup analyses of phase 1 data suggested possible age effects, we will also fit a model that adjusts for age group at baseline (above or below 65) and that includes interactions between age and each of the slope terms, such that $E[y_{ij}]= \beta_0+\beta_1 time+\beta_2 arm+\beta_3 older+\beta_4 older*arm+\beta_5 time*arm+\beta_6 older*time+\beta_7 older*time*arm+ \beta_8 time^{12+} +\beta_9 time^{12+}*arm+\beta_{10} time^{12+}*older+\beta_{11} time^{12+}*older*arm$, where y_{ij} is the outcome for the *i*th participant at the *j*th timepoint, β_1 is the expected rate of change for younger participants in the

delayed on arm during phase 1, β_5 is the expected difference in rate of change for younger participants between the two arms during phase 1, β_6 is the expected difference in rate of change in the delayed on arm during phase 1 between older and younger participants, and β_7 is the expected difference arm effect between older and younger participants during phase 1, β_8 is the expected difference in rate of change for younger participants in the delayed arm between phase 1 and phase 2, β_9 is the expected difference in between phase difference in slope between treatment arms, β_{10} is the expected difference in between phase difference in slope between older and younger participants, and β_{11} is the expected difference between older and younger participants in the between-phase differences between treatment arms.

Figure S1. Schematic of Disease Modifying and Symptomatic Treatment Effects

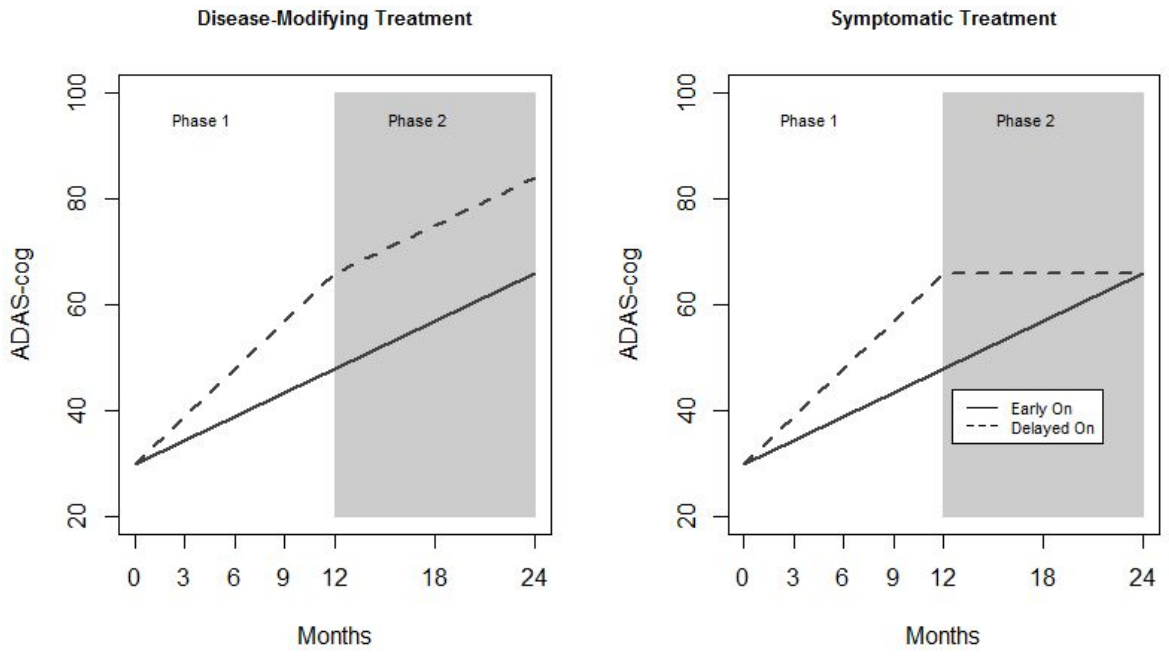

Figure S1 shows hypothetical trajectories for early on and delayed on study arms. In the both cases, the trajectories diverge during phase 1, when the early on arm is receiving the intervention and the delayed on arm is not. During phase 2, when the delayed on arm also receives the intervention, the trajectory of the delayed on arm changes. In the case of a disease-modifying treatment, the trajectory of the delayed on arm changes to match that of the early on arm, but the outcome means continue to differ between arms. In the case of a symptomatic treatment, the delayed on arm catches up to the early on group, and there are no between-arm differences in outcome means at the end of phase 2.

Figure S2. Schematic of of Longitudinal Mixed Effects Linear Spline Models

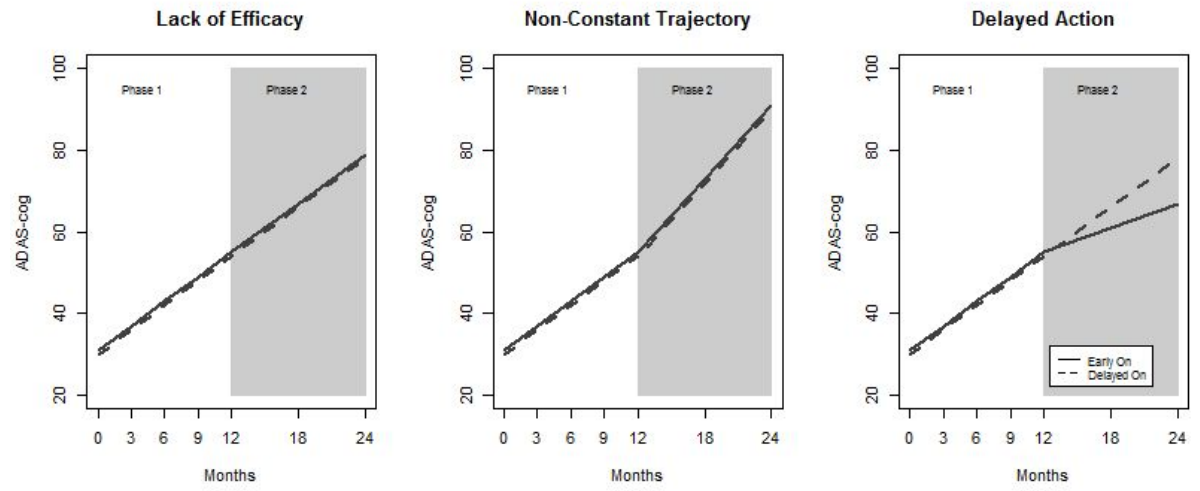

Figure S2. Shows scenarios that can be modeled via the mixed effects longitudinal model $E[y_{ij}]= \beta_0+\beta_1 time + \beta_2 arm+ \beta_3 time*arm+\beta_4 time^{12+} +\beta_5 time^{12+}* arm$, where y_{ij} is the outcome for the *i*th participant at the *j*th timepoint, β_1 is the expected rate of change for the "delayed on" arm during phase 1, β_3 is the expected difference in rate of change between the two arms during phase 1, β_4 is the expected difference in rate of change for people in the "delayed on" arm between phase 1 and phase 2, and β_5 is the expected difference in change in rate of change between phase 1 and phase 2, between treatment arms. In the figure on the left, β_1 is positive, showing worsening over time in both arms, but β_3 is zero, denoting no between-arm difference in rate of change, β_4 is zero, denoting no change in trajectory for the delayed on arm between phases, and β_5 is zero, denoting no difference in between-phase change in trajectory between arms. The figure in the middle is identical to the figure on the left, except now β_4 is also positive, denoting an increase in rate of worsening for both arms. In the figure on the right, β_4 is zero, denoting no change in trajectory for the delayed on arm between phases, but β_5 is negative, denoting slowing of worsening in phase 2 in the early on arm only.

Supplemental Results

Table S1 shows parameter estimates and standard errors for each of the terms in the longitudinal mixed effect models. As time was measured in months, the model-based change scores for each treatment arm and phase shown in the main text were calculated as linear combinations of the slope terms multiplied by 12. For example, the expected change in ADAS-cog 13 during phase 2 for individuals in the early on arm is calculated as $\beta_1+\beta_3+\beta_4+\beta_5$; and the standard errors are functions of the standard errors of each term and their covariances.

Table S2 shows observed mean change scores as a function of treatment arm and phase. The model-based change scores in the main paper should approximate these change scores, but note while the regression model allows for calculation of an expected change score even when there is missing data, that is not possible when calculating change scores from observed data. As such, calculation of observed mean change scores excludes individuals who are missing values at either the beginning or end of a given phase, and therefore should not be considered intention to treat analyses.

Tables S3 and S4 contain parameter estimates and observed change scores by treatment arm and phase for the *post hoc* age group secondary analyses.

Table S1. Parameter Estimates for Primary Analysis[1]				
	ADAS-cog 13[2]	CDRsob[3]	CVLT[4]	NPI[5]
Intercept (β_0)	30.14 (2.00)[6]	4.24 (.55)	12.73 (1.13)	4.12 (1.77)
Between Group Difference in Intercept (β_2)	1.58 (1.49) *.29*	.37 (.51) *.46*	-2.08 (1.82) *.25*	.92 (1.59) *.56*
Phase 1 Slope (β_1)	.69 (.15) *<.01*	.22 (0.05) *<.01*	-.36 (.12) *<.01*	.53 (.19) *<.01*
Between-Group Difference in Phase 1 Slope (β_3)	-.42 (.22) *.85*	-.02 (.06) *.81*	.01 (.14) *.96*	-.07 (.23).77
Between-Phase Difference in Slope (β_4)	-.18 (.24) *.46*	.00 (.08) *>.99*	.05 (.17) *.78*	-.42 (.23) *.07*
Between Group Difference in Between-Phase Differences in Slope (β_5)	-.01 (.35) *.99*	-.04 (.10) *.73*	.09 (.24) *.71*	-.13 (.34) *.71*

[1] Estimates are from a longitudinal mixed effects with a random intercept and adjustment for site effects (not shown).
[2] ADAScog-13: Alzheimer's Disease assessment scale-cognitive subscale 13
[3] CDRsb: Clinical Dementia Rating sum of boxes
[4] CVLT: California Verbal Learning Test, second edition, sum of first five trials
[5] NPI: Neuropsychiatric Inventory, total score
[6] Estimates are presented as β (Standard Error) *p-value*, where p-value is from the test of whether the regression coefficient is equal to 0.

Table S2. Observed change scores by treatment arm and phase for primary analysis				
	Delayed On		Early On	
	Phase 1	Phase 2	Phase 1	Phase 2
ADAS-cog 13	7.952(1.929)	6.158(1.93)	8.7(2.121)	6.294(1.878)
CDRsb	2.643(.706)	2.694(.675)	2.526(.414)	2.30(.676)
CVLT	-3.952(1.188)	-4.6(1.29)	-4.9(1.3)	-4.154(1.775)
NPI	5.571(1.788)	1.158(1.911)	6.7(2.03)	.412(1.879)
Note: values are observed change per year (SE)				

Table S3. Parameter Estimates for Secondary Subgroup Analyses by Age[1]				
	ADAS-cog 13[2]	CDRsb[3]	CVLT[4]	NPI[5]
Intercept (β_0)	29.11 (2.15)[6]	4.60 (.55)	14.60 (1.60)	2.87 (2.97)
Between Arm Difference in Intercept (β_2)	8.84 (2.28) *<.01*	.04 (.04) *.29*	-9.52 (4.00) *.02*	-.21 (2.72) *.94*
Between Age Difference in Intercept (β_3)	2.48 (2.62) *.34*	-.16 (.82) *.85*	-3.25 (2.34) *.16*	2.10 (2.84) *.46*
Between Age Difference in Between Arm Difference in Intercept (β_4)	10.02 (3.31) *<.01*	-1.96 (1.00) *.05*	10.38 (5.32) *.05*	1.61 (3.44) *.64*
Phase 1 Slope (β_1)	.88 (.31) *<.01*	.038 (.04) *.29*	-.70 (.27) *.01*	0.62 (.46) *.18*
Between Arm Difference in Phase 1 Slope (β_5)	.55 (.42) *.19*	.22 (.07) *<0.01*	.26 (.30) *.38*	.40 (.53) *.45*
Between Age Difference in Phase 1 Slope (β_6)	-.26 (.35) *.46*	.25 (.07) *<.01*	.47 (.29) *.10*	-.13 (.49) *.80*
Between Age Difference in Between Arm Difference in Phase 1 Slope (β_7)	-.84 (.47) *.07*	-.33 (.10) *<0.01*	-.35 (.33) *.29*	-.66 (.57) .25
Between Phase Difference in Slope (β_8)	-.25 (.58) *.67*	.26 (.14) *.05*	.37 (.37) *.33*	-.41 (.65) *.53*
Between Arm Difference in Between-Phase Differences in Slope (β_9)	-1.21 (.66) *.07*	-.18 (.19) *.34*	-.43 (.46) *.35*	-.44 (.85) *.61*
Between Age Difference in Between-Phase Differences in Slope (β_{10})	.10 (.64) *.87*	-.37 (.16) *.02*	-.45 (.42) *.28*	-.01 (.68) *.99*
Between Age Difference in Between Arm Differences in Between Phase Differences in Slope (β_{11})	1.69 (.74) *.02*	.22 (.21) *.30*	.68 (.53) *.20*	.50 (.89) .58

[1] Estimates are from a longitudinal mixed effects with a random intercept and adjustment for site effects (not shown).
[2] ADAScog-13: Alzheimer's Disease assessment scale-cognitive subscale 13
[3] CDRsb: Clinical Dementia Rating sum of boxes
[4] CVLT: California Verbal Learning Test, second edition, sum of first five trials
[5] NPI: Neuropsychiatric Inventory, total score
[6] Estimates are presented as β (Standard Error) *p-value*, where p-value is from the test of whether the regression coefficient is equal to 0.

Table S4. Observed change scores by treatment arm and phase, stratified by age				
	Phase 1			
	Delayed On		Early On	
outcome	<65	>=65	<65	>=65
ADAS-cog 13	8.33(4.536)	7.8(2.121)	18.667(4.128)	4.429(1.382)
CDRsb	.500(.289)	3.5(.894)	3.4(.828)	2.214(.468)
CVLT	-6.167(2.96)	-3.067(1.173)	-5.833(2.774)	-4.5(1.489)
NPI	5.833(3.27)	5.467(2.206)	13.833(4.665)	3.643(1.619)
	Phase 2			
ADAS-cog 13	6.2(5.435)	6.143(1.938)	.5(2.958)	8.077(2.086)
CDRsb	3.8(1.765)	2.27(.664)	5.00(2.082)	1.625(.568)
CVLT	-4.25(1.887)	-4.727(1.668)	-7.333(3.383)	-3.2(2.065)
NPI	1.4(4.106)	1.071(2.237)	6.25(3.637)	-1.385(1.999)
Note: values are observed change per year (SE)				

Table S5. Descriptive Baseline Statistics by Age Group (<65, >65)			
	<65 (n=12)	>65 (n=30)	t/[χ^2] (df) p[1]
Male (n, %)	7 (58)	12 (40)	1.16(1) *.281*
Age (yrs)	57.65 (4.86)	72.35 (3.73)	
Time Since Diagnosis (yrs)	2.05 (1.70)	2.45 (1.75)	-.67 (20.91) *.508*
College Education (n, %)	12 (100)	19 (63)	5.96 (1) *0.015*
ADAScog13[2]	27.00 (4.18)	28.23 (3.78)	-.89 (18.61) *.386*
CDRsb[3]	4.33 (1.51)	3.62 (1.44)	1.41 (19.48) *.176*
CVLT[4]	20.92 (8.41)	21.10 (7.59)	-.07 (18.57) *.984*
NPI[5]	2.58 (2.91)	2.80 (3.40)	-.208 (23.64) *.837*

[1] t-test or χ^2 test (degrees of freedom (Satterthwaite for t-tests)), *p-value*
[2] ADAScog-13: Alzheimer's Disease assessment scale-cognitive subscale 13
[3] CDRsb: Clinical Dementia Rating sum of boxes
[4] CVLT: California Verbal Learning Test, second edition, sum of first five trials
[5] NPI: Neuropsychiatric Inventory, total score

References

[1] Cummings JL, Doody R, Clark C (2007) Disease-modifying therapies for Alzheimer disease: challenges to early intervention. *Neurology* **69**, 1622–1634.

[2] D'Agostino RBS (2009) The Delayed-Start Study Design. *N. Engl. J. Med.* **361**, 1304–1306.

[3] Leber P (1996) Observations and suggestions on antidementia drug development. *Alzheimer Dis. Assoc. Disord.* **10 Suppl 1**, 31–35.

[4] Wilson RS, Segawa E, Boyle PA, Anagnos SE, Hizel LP, Bennett DA (2012) The natural history of cognitive decline in Alzheimer's disease. *Psychol. Aging* **27**, 1008–1017.

[5] Tschanz JT, Corcoran CD, Schwartz S, Treiber K, Green RC, Norton MC, Mielke MM, Piercy K, Steinberg M, Rabins PV, Others (2011) Progression of cognitive, functional, and neuropsychiatric symptom domains in a population cohort with Alzheimer dementia: the Cache County Dementia Progression study. *Am. J. Geriatr. Psychiatry* **19**, 532–542.

[6] Szekely CA, Thorne JE, Zandi PP, Ek M, Messias E, Breitner JCS, Goodman SN (2004) Nonsteroidal anti-inflammatory drugs for the prevention of Alzheimer's disease: a systematic review. *Neuroepidemiology* **23**, 159–169.

[7] Wang J, Tan L, Wang H-F, Tan C-C, Meng X-F, Wang C, Tang S-W, Yu J-T (2015) Anti-inflammatory drugs and risk of Alzheimer's disease: an updated systematic review and meta-analysis. *J. Alzheimers. Dis.* **44**, 385–396.

[8] Leoutsakos J-MS, Muthen BO, Breitner JCS, Lyketsos CG, ADAPT Research Team (2012) Effects of non-steroidal anti-inflammatory drug treatments on cognitive decline vary by phase of pre-clinical Alzheimer disease: findings from the randomized controlled Alzheimer's Disease Anti-inflammatory Prevention Trial. *Int. J. Geriatr. Psychiatry* **27**, 364–374.

[9] Liu-Seifert H, Andersen SW, Lipkovich I, Holdridge KC, Siemers E (2015) A novel approach to delayed-start analyses for demonstrating disease-modifying effects in Alzheimer's disease. *PLoS One* **10**, e0119632.