Statistical Analysis Plan: I8G-MC-LMDC (v6)

Assessment of Safety, Tolerability, and Efficacy of LY3303560 in Early Symptomatic Alzheimer's Disease

NCT03518073

Approval Date: 11 Oct 2021

1. Statistical Analysis Plan: I8G-MC-LMDC: Assessment of Safety, Tolerability, and Efficacy of LY3303560 in Early Symptomatic Alzheimer's Disease

LY3303560

Alzheimer's disease

Multicenter, randomized double-blind, placebo-controlled, Phase 2 study comparing up to 5600 mg of LY3303560 with placebo over 104 weeks in approximately 225 patients with early symptomatic AD.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I8G-MC-LMDC
Phase 2

SAP electronically signed and approved by Lilly on date provided below:

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Appendix **Page**

3. Revision History

SAP Version 1 was approved on Oct. 10, 2018 prior to unblinding. SAP Version 2 was approved on Nov. 19, 2019 prior to efficacy unblinding. SAP version 3 was approved on Aug. 11, 2020 prior to interim analysis 2. SAP version 4 was approved April 21, 2021 prior to interim analysis 3. SAP version 5 was approved August 31, 2021 prior to the primary outcome lock. SAP version 6 was approved prior to the primary outcome lock and included the following changes:

- Fixed typo in description of number of missing items allowed in the ADAS-Cog13 scale (sec. 6.3)
- Added the Bayesian DPM thresholds for the Multiplicity Graph for iADRS LY5600 vs. placebo and iADRS LY1400 vs. placebo and a description calculating all of the thresholds included in the Multiplicity Graph (sec. 6.11.2)
- Updated ECG timepoints (sec. 6.13.6)
- Added two subgroup analyses: baseline tau grouped into thirds and baseline MMSE grouped into thirds, and updated subgroup definitions of disease severity based on MMSE (sec. 6.14)

4. Study Objectives

4.1. Primary Objective

The primary objective of protocol I8G-MC-LMDC (LMDC) is to test the hypothesis that LY3303560 administered for 100 weeks will decrease the decline in cognitive and/or functional outcomes in patients with early symptomatic Alzheimer's Disease (AD) relative to placebo as measured by the integrated Alzheimer's Disease Rating Scale (iADRS) score from baseline to 104 weeks.

4.2. Secondary Objectives

The secondary objectives of LMDC are to assess the effect of LY3303560 versus placebo on clinical progression, on brain aggregated tau deposition, and on attenuating downstream markers of the neurodegenerative process in patients with early symptomatic AD and evaluate the safety and tolerability of LY3303560.

Clinical progression will be assessed by change from baseline to 104 weeks in cognition and/or function as measured by the following:

- Alzheimer's Diasease Assessment Scale—Cognitive subscale (ADAS-Cog13) score
- Alzheimer's Disease Cooperative Study—instrumental Activities of Daily Living scale (ADCS-iADL) score
- Clinical Dementia Rating Scale – Sum of Boxes (CDR-SB) score
- Mini-Mental State Examination (MMSE) score

Brain aggregated tau deposition will be assessed by change from baseline through 104 weeks as measured by flortaucipir F 18 PET scan.

Attenuation of downstream markers of the neurodegenerative process in AD will be assessed by change from baseline through 104 weeks as measured by volumetric magnetic resonance imaging (vMRI).

Safety assessments used to evaluate the safety and tolerability of LY3303560 include the following:

- Spontaneously reported adverse events
- Clinical laboratory tests
- Vital signs and body weigh measurements
- 12-lead ECGs
- Physical and neurological exams
- Anti-drug antibodies
- Safety MRIs

- Columbia Suicide Severity Rating Scale (C-SSRS)

4.3. Exploratory Objectives

The exploratory objectives described in the protocol for study LMDC are in the following table:

Exploratory Objectives	Exploratory Endpoints
To assess the effect of LY3303560 versus placebo on clinical progression in patients with early symptomatic AD.	Change in dependence level derived from ADCS-ADL scale scores from baseline to 104 weeks.
To assess the effect of LY3303560 versus placebo on clinical progression in patients with early symptomatic AD.	Change in cognition from baseline to 104 weeks as measured by the change in: • CogState Brief Battery (CBB)
To assess peripheral PK and presence of anti-LY3303560 antibodies over 104 weeks.	• Maximum serum concentration of LY3303560 at steady state ($C_{max,ss}$). • Anti-drug antibodies (ADA) against LY3303560 including treatment-emergent ADA and neutralizing antibodies.
To assess the effect of LY3303560 versus placebo on clinical progression as measured by Digital Clock Drawing test (DCTClock) in patients with early symptomatic AD.	Change in DCTClock results from baseline through 104 weeks.
To assess the initial effect of LY3303560 versus placebo on plasma total tau and phospho181tau concentrations in patients with early symptomatic AD.	Change from baseline to steady state (16 weeks) in plasma tau and phospho181tau levels.
To assess the steady state effects of LY3303560 versus placebo on plasma tau and phospho181tau concentrations in patients with early symptomatic AD.	Change in plasma tau and phospho181tau concentrations at steady state through 104 weeks.
To assess the utility of DCTClock and plasma phospho181tau in screening phase efficiency for trials of patients with early symptomatic AD.	Associations of DCTClock and plasma phospho181tau with MMSE, CBB, screen failure categories, and baseline data from enrolled subjects.

Additional exploratory objectives not included in the protocol are described in the following table:

To assess the initial effect of LY3303560 versus placebo on plasma ptau217 and Neurofilament light chain (NfL) concentrations in patients with early symptomatic AD.	Change from baseline to steady state (16 weeks) in plasma ptau217 and NfL.
To assess the steady state effects of LY3303560 versus placebo on ptau217 and NfL concentrations in patients with early symptomatic AD.	Change in ptau217 and NfL at steady state through 104 weeks.

5. Study Design

5.1. Summary of Study Design

Study LMDC is a multicenter, randomized, double-blind, placebo-controlled, Phase 2 study of LY3303560 in patients with early symptomatic AD and low-to-medium cerebral tau burden. A combination of visual and quantitative assessments of tau PET is used to identify eligible patients:

- Visual assessment of AD+ (increased activity in posterior temporal regions only) AND $1.10 < \text{SUVR} \leq 1.46$
- Visual assessment of AD++ (increased activity in temporal and parietal regions) AND $\text{SUVR} \leq 1.46$

The maximum possible duration of the study is 121 weeks that includes a screening period of up to 8 weeks, a treatment period of 100 weeks, a 4-week post last dose assessment, and an immunogenicity and safety follow-up period of up to 13 weeks, following the last dose of the study drug at Week 100. Subjects who meet entry criteria will be randomized in a 1:1:1 ratio to one of the following treatments:

- **1400 mg LY3303560:** LY3303560 1400 mg IV infusion Q4W for 100 weeks (Visit 2 [Week 0] to Visit 27 [Week 100]).
- **5600 mg LY3303560:** LY3303560 5600 mg IV infusion Q4W for 100 weeks (Visit 2 [Week 0] to Visit 27 [Week 100]).
- **Placebo**: IV placebo infusion Q4W for 100 weeks (Visit 2 [Week 0] to Visit 27 [Week 100]).

The first safety review by a Data Monitoring Committee will be conducted after approximately 30 patients have received 3 doses and completed the 4 week assessment period (i.e. just prior to their 4th dose). Enrolment will halt after these patients have received one dose of study treatment, however those patients already dosed will continue their allocated treatment. If any of these patients discontinue from study treatment for reasons other than safety, they may be replaced, upon discussion between the investigators and Lilly.

The primary hypothesis being tested is that LY3303560 administered for 100 weeks will result in a significant slowing in cognitive/functional decline compared with placebo as measured by the change from baseline to the end of the double-blind period (Week 104) on the integrated Alzheimer's Disease Rating Scale (iADRS), in subjects with early symptomatic AD (where early symptomatic AD refers to the combination of 2 stages: prodromal AD [mild cognitive impairment (MCI)-AD] and mild dementia due to AD; Alaka et al. 2015) who have low-to-medium tau burden.

5.2. Determination of Sample Size

Based on the Disease Progression Model (DPM), 360 enrolled patients, a 30% dropout rate, and a 3-month delay of treatment effect onset, this sample size will provide more than 95% power to

demonstrate that at least one of the active treatment arms has slowed cognitive/functional decline. This claim of slowing cognitive/functional decline is based on showing that at least one of the active treatment arms has a ≥0.6 posterior probability of slowing of iADRS progression over placebo by at least 4 points (that is, 25% slowing relative to placebo, based on a mean change from baseline in the placebo group of 16 points and mean change from baseline on at least one of the active treatment arms of at most 12 points) after 104 weeks (approximately 24 months) on study treatment. This power calculation comes from simulations using a placebo decline of 19.56 points and 33% slowing on 1400 mg LY3303560 and 50% slowing on 5600 mg LY3303560 relative to placebo at 104 weeks. The variance-covariance matrix in the simulations had a variance at Week 104 of 436.85 (equal to a standard deviation of 20.9). If both active treatment arms are placebo-like with no efficacy, the probability of passing the efficacy criterion specified above (that is, false positive) is approximately 2.5%. To test the hypothesis of a disease progression benefit, we calculate the posterior probability of superiority in cognitive/functional slowing, and if it is above a pre-specified threshold (which controls the experiment-wise type I error at approximately 2.5%), then a claim of cognitive slowing will be made. The simulation for the power calculation and sample size determination was carried out in FACTS Version 5.5.

5.3. Method of Assignment to Treatment

Patients who meet all criteria for enrollment will be assigned a study (patient) number at Visit 1 and randomized to double-blind treatment at Visit 2. Patients will be randomized to LY3303560 or Placebo in a 1:1:1 ratio. Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web response system (IWRS). For the first, approximately, 30 patients, the IWRS will be programmed to guarantee balance between the arms for the first interim analysis for safety; this is referred to as the burn-in period. After the burn-in period, patient randomization will then follow the dynamic allocation (minimization) method of Pocock and Simon (1975) to balance the treatment arms using investigative site as a factor. This is to ensure balanced patient assignment between treatment arms within each site at the end of the study.

5.4. Interim Analyses

An external DMC is authorized to evaluate results from unblinded interim analyses for the assessment of safety, to recommend any modifications to the study (such as stopping the study or dropping an arm), and to assess efficacy to inform future development. Operational details and any efficacy decision rules will be provided in the DMC charter and the interim efficacy SAP. Study sites will receive information about interim results ONLY if relevant for the safety of their patients. Unblinding details will be specified in the separate unblinding plan document.

The first unblinded DMC review for safety will be conducted once approximately 30 patients have completed 3 months of exposure to study treatment (i.e., after completing 4 weeks following the third dose of study treatment). No patients beyond this initial group of approximately 30 patients will be dosed until the DMC review has been completed and the recommendation is to continue the study. This initial group will continue treatment during the assessment period.

The purpose of the second interim analysis is to evaluate efficacy results from unblinded interim analyses to inform future development, as well as to evaluate safety. This second unblinded DMC review will be conducted after approximately all patients have had the opportunity to complete 12 months of exposure to study treatment (i.e., 4 weeks after the 12th dose at Visit 15 [Week 52]). Treatment will continue during the second interim analysis. With regard to safety, the outcomes of the DMC reviews could be to: 1) continue as planned, 2) modify the protocol, including additional interim analyses, and 3) stop the study. The outcomes with regard to efficacy will be detailed in the interim efficacy SAP.

The purpose of the third interim analysis is to evaluate safety. This third unblinded DMC review will be conducted after approximately all patients have had the opportunity to complete 18 months of exposure to study treatment (i.e., 4 weeks after the 18th dose at Visit 22 [Week 80]). Treatment will continue during the third interim analysis. With regard to safety, the outcomes of the DMC reviews could be to: 1) continue as planned 2) modify the protocol, including additional interim analyses and 3) stop the study

To facilitate PK/PD interim analyses and to initiate the final population PK/PD model development processes, a limited number of preidentified individuals may also gain access to the unblinded data, as specified in the unblinding plan. This will be conducted at the same time as the second DMC. In preparation for the final analysis, access may also be granted after all patients complete 88 weeks of treatment (i.e., Visit 24). Information that may unblind the study during the analyses will not be reported to study sites or blinded study team until the study has been unblinded.

Quarterly, blinded trial level safety reviews will be carried out as documented in the trial level safety review plan.

6. A Priori Statistical Methods

6.1. General Considerations

The protocol calls for a Data Monitoring Committee (DMC) charged with making decisions regarding patient safety and interim efficacy. This analysis plan describes analyses for the final clinical study report and all interim safety analyses for the DMC.

All analyses will follow the intention-to-treat (ITT) principle unless otherwise specified. An ITT analysis is an analysis of data by the groups to which subjects are assigned by random allocation, even if the subject does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol. Consistent with ITT studies the estimand for Study LMDC is the Treatment Policy estimand. Safety analyses will group subjects based on the actual treatment received. If a patient receives a treatment different than the randomized treatment for the duration of the placebo-controlled period, then that different treatment is the actual treatment; otherwise, the actual treatment is the planned treatment.

Unless otherwise noted, all pairwise tests of treatment effects will be conducted at a 2-sided alpha level of 0.05; 2-sided confidence intervals (CIs) will be displayed with a 95% confidence level. All tests of interactions between treatment and other factors will be conducted at an alpha level of 0.05.

Unless otherwise noted no formal statistical hypothesis testing will be made during the first safety interim. Any reported p-values that may be reported as part of standardized output are merely for information purposes only.

Unless otherwise noted baseline is defined as the last measurement prior to dosing. When change from baseline is assessed, subjects will only contribute to the analysis if both a baseline and a post-baseline measurement are available. Endpoint is the last non-missing post-baseline measurement within the time period for the given analysis. For mixed-effect model for repeated measures (MMRM) models, observations collected at nonscheduled visits will not be included in the analyses (Andersen and Millen 2013). For analyses using last observation carried forward (LOCF), the last nonmissing post-baseline observation (scheduled or unscheduled) will be used to calculate change from baseline.

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol and the justification for making the change will be described within this SAP and clinical study report. Additional exploratory analyses of the data will be conducted as deemed appropriate.

6.1.1. *COVID-19*

Study LMDC was fully enrolled at the time the COVID-19 pandemic had impacted the countries participating in LMDC (March 1, 2020). The following summaries including treatment comparisons will be made to assess the impact of the pandemic on the study:

- Number of patients who discontinued the study early due to the pandemic

- Number of patients who used phone visits to replace site visits due to the pandemic
- Number of patients who used video visits to replace site visits due to the pandemic
- Number of patients who had out-of-window visit intervals due to the pandemic

All patients who are impacted by COVID as described above will have their iADRS data censored from that time forward and the DPM and MMRM will be assessed. In other words, patients' data after a temporary discontinuation of study drug because of COVID will be censored in order to assess if COVID impacted the DPM and MMRM analyses of iADRS.

6.2. Adjustments for Covariates

The repeated measures models will include the fixed, categorical effects of baseline score, pooled site, treatment, visit, treatment-by-visit interaction, baseline-by-visit interaction, concomitant AChEI and/or memantine use at baseline (yes/no), and age at baseline.

When an analysis of covariance (ANCOVA) model is used to analyze a continuous efficacy variable, the model will contain the main effects of treatment and appropriate baseline value included as a covariate. When an ANCOVA model is used to analyze a continuous safety variable, the model will contain the main effects of treatment, age, and appropriate baseline value included as a covariate.

6.3. Handling of Dropouts or Missing Data

6.3.1. *Handling Missing Data from Patient Dropouts*

A likelihood-based mixed effects model for repeated measures will be used to handle missing data. The model parameters are simultaneously estimated using restricted likelihood estimation incorporating all of the observed data. Estimates have been shown to be unbiased when the missing data are missing at random and when there is ignorable non-random missing data.

Repeated measures and disease progression model analyses will only use data from visits where the data was scheduled to be collected (Andersen and Millen 2013). When patients discontinue from the study early, there may be efficacy or safety data measurements at visits where the variables were not scheduled to be collected. This data will be used in all other analyses.

6.3.2. *Handling Missing Items in Calculating Totals*

If any of the individual items for ADAS-Cog or ADCS-ADL are missing or unknown, every effort will be made to obtain the score for the missing item or items.

For ADAS-Cog_{13}, if fewer then 4 of a total of 13 items are missing, the total score (maximum =85) will be imputed as follows: the total from the remaining items will be multiplied by a factor that includes the maximum score for the missing items. For example, if the first item, "Word-Recall Task," which ranges from a score of 0 through 10 (maximum = 10), is missing, and the second item "Commands," which ranges from a score of 0 to 5 (maximum = 5), is missing, then the multiplication factor = 85/(85 - [10 + 5]) = 85/70 = 1.21. Thus, the total score for this example will be the sum of the remaining 11 items multiplied by 1.21. The

imputed number will be rounded up to the nearest integer. If more than 4 items are missing, the total score for ADAS-Cog_{13} at that visit will be considered missing.

For the ADCS-iADL, if <30% of the items are missing, the total score will be imputed (maximum = 59). The sum of the nonmissing items will be prorated to the sum of total items. The imputed number will be rounded up to the nearest integer. If the nearest integer is greater than the maximum possible score, the imputed score will be equal to the maximum score. If >30% of the items are missing, the total score for ADCS-iADL at that visit will be considered missing. The same imputation technique will be applied to the ADCS-ADL total score. Note that, depending on the specific item responses that are missing, it is possible to have an imputed total score for both the ADCS-iADL and the ADCS-ADL, an imputed total score for one but not the other, or both total scores missing.

The same imputation technique will be applied to the CDR-SB. If only 1 box (of 6) of the CDR is missing, the sum of the boxes will be imputed by prorating the sum from the other 5 boxes. If the score from more than 1 box is not available, the CDR-SB at that visit will be considered missing.

The iADRS score is calculated as follows: iADRS score = [−1(ADAS - Cog_{13}) + 85] + ADCS-iADL (Wessels et al. 2015). If either ADAS-Cog_{13} or ADCS-iADL is missing, iADRS score will be considered missing.

For all other scales, if any item is missing, any total or sum involving that item will be considered missing.

6.3.3. *Handling Missing Date Information*

For previous medications and medical history if parts of dates are missing, the following imputations will be performed:

- For start dates
    a. if DAY is unknown, it will be set to 01
    b. if MONTH is unknown, it will be set to JAN (01)
    c. if Day and Month are unknown, it will be set to 01 and JAN (01)
    d. if year is unknown, then start date is missing
- For end dates
    a. if DAY is unknown, it will be set to 30
    b. if MONTH is unknown, it will be set to DEC (12)
    c. if Day and Month are unknown, it will be set to 30 and Dec (12)
    d. if year is unknown, then end date is missing

6.4. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites internationally. In the event that there is an inadequate number of subjects (defined as 1 or 0 randomized subjects per treatment group) at a site for the planned analyses, the following strategy will be implemented. Data from all such sites will be pooled. The pooling will be done first within a country. If the resulting pool within a country is still inadequate (1 or 0 randomized subjects to 1 or more treatment arms), no further pooling will be performed. The pooled site variable for each patient will be the country code if site was pooled; otherwise, it will be the site number. A listing including country, investigator site with address, number of patients enrolled (randomized) by each site, and unique patient IDs will be presented.

6.5. Multiple Comparisons/Multiplicity

A graphical strategy may be used for testing key secondary hypotheses to protect against Type I error of falsely rejecting a null hypothesis (Section 6.11.2.). The use of a prespecified analysis plan that employs Bretz' graphical approach will provide strong control of the study-wise Type I error rate for the primary and key secondary hypotheses at level α=0.05 (Bretz et al. 2009, 2011).

6.6. Analysis Populations

For purposes of analysis, populations are defined in Table 6.6.1 and Table 6.6.2. These tables also list the study measures that will be summarized and/or analyzed for each population.

Table 6.6.1. Analysis Populations for Study I8G-MC-LMDC

Population	Description
Entered	All participants who sign informed consent
Randomized	All entered patients who are randomized to study treatment
Evaluable Efficacy	All randomized patients with at least one post-baseline iADRS13 result (protocol refers to this as the Full Analysis Set)
Safety	All randomized participants who take at least 1 dose of double-blind study treatment. Participants will be analyzed according to the actual treatment group received (actual treatment group is defined as the treatment received at Visit 2).
Per-Protocol	All subjects in the Evaluable Efficacy population who also: • signed the inform consent form • had an assessment of the primary endpoint at each scheduled visit completed • had no violations of inclusion/exclusion criteria • had no study dosing algorithm violation (such as if subjects randomized to treatment A were given treatment B or subjects randomized to treatment A never received the assigned study drug) • had no unqualified raters and no raters with substantial scoring errors for the primary measure • were not considered non-compliant with regard to study drug (a subject is considered non-compliant when missing or incomplete infusions exceed 20%, OR if missed more than 3 consecutive infusions at any point)

Completers	All randomized subjects who have treatment disposition status of complete.

Table 6.6.2 Efficacy and Safety Measures Summarized and/or Analyzed for Each Analysis Population

Population	Variables Assessed
Entered	Listings
Randomized	Tables and Listings for patient characteristics, baseline severity, and patient disposition
Evaluable Efficacy	Tables, Listings and/or Figures of the following: iADRS, ADAS-Cog_{13}, ADCS-ADL (basic, instrumental and total), CDR-SB, MMSE, CBB, DCTClock, plasma total tau, plasma p-tau (181 and 217), plasma NfL, flortaucipir parameters, volumetric MRI measurements, and concomitant medications.
Safety	Tables, Listings and/or Figures of the following: Compliance, adverse events, laboratory results, vital signs, weight, ECG, safety MRIs, and C-SSRS.
Per-Protocol	Tables, Listings and/or Figures of iADRS, ADAS-Cog_{13}, ADCS-ADL (basic, instrumental and total), CDR-SB, MMSE
Completers	Tables, Listings and/or Figures of iADRS, ADAS-Cog_{13}, ADCS-ADL (basic, instrumental and total), CDR-SB, MMSE, plasma total tau, plasma p-tau (181 and 217), NfL, flortaucipir parameters, and volumetric MRI measurements

6.7. Patient Disposition

Because this is a long-term study in a patient population that is elderly with multiple comorbidities, patient withdrawal is of particular concern. Additional efforts will be undertaken to reduce patient withdrawals and to obtain information on patients who are initially categorized as lost to follow-up.

From the randomized population, the percentage of patients withdrawing from each treatment group will be compared between groups using Fisher's exact test. Comparisons using Fisher's exact test will be done for the overall percentage of patients who withdraw and also for each specific reason for withdrawal.

The median time to discontinuation will also be compared between treatment groups using the Kaplan-Meier product limit estimator. For any-cause discontinuation as well as discontinuation due to adverse event or death, comparisons of time-to-discontinuation will be conducted using the Kaplan-Meier product limit estimator and the associated log-rank test.

6.8. Patient Characteristics

Baseline characteristics will be summarized for the randomized population by treatment group and overall. Summaries will include descriptive statistics for continuous and categorical measures. Fisher's exact test or Pearson's chi-square test will be used for treatment-group comparisons of categorical data. For continuous data, analysis of variance (ANOVA), with independent factors for treatment, will be used. Patient characteristics to be presented include:

- age
- gender

- race
- ethnicity
- height
- body weight
- body mass index (weight (kg) / [height (m)]2)
- tobacco use
- alcohol use
- years of education
- work status
- time since onset of first AD symptoms
- tau PET burden (various measures)
- time since diagnosis
- APOE4 carrier status (carrier [ε2/ε4, ε3/ε4, ε4/ε4], noncarrier [ε3/ε3, ε2/ε2, ε3/ε2])
- APOE4 genotype (ε2/ε4, ε3/ε4, ε4/ε4, no ε4)
- having 1 or more first degree relatives with AD
- AChEI and/or memantine use at baseline

Baseline severity of impairment as measured by ADAS-Cog_{13}, ADCS-ADL total score and instrumental (ADCS-iADL) and basic subscores (ADCS-bADL), CDR Sum of Boxes, MMSE, Cogstate Brief Battery (CBB), and Digital Clock Drawing Test (DCTClock). Baseline characteristics and baseline severity will also be listed.

6.9. Treatment Compliance

Because dosing occurs at study visits, patients who attend all visits and successfully receive LY3303560 or placebo infusions are automatically compliant with this treatment. Any infusion at which 75% (approximately 190 mL) or more of the infusion solution is given will be considered a complete infusion.

Summary statistics for LY3303560 treatment compliance will be provided for the total number of complete infusions received, duration of complete infusion, and volume of complete infusion by treatment group. Frequencies and percentages of reasons why infusion was stopped will also be presented.

6.10. Concomitant Therapy

Prior medications are defined as those that stop before randomization (the day prior to the first administration of study drug). Concomitant medications are defined as those being taken on or after randomization (the day prior to the first administration of study drug). A summary of concomitant medications will be presented as frequencies and percentages for each treatment group. Fisher's exact test will be used to test for treatment differences between groups. If the start or stop dates of therapies are missing or partial to the degree that determination cannot be made of whether the therapy is prior or concomitant, the therapy will be deemed concomitant. A summary table will also be provided for concomitant AChEI/memantine medications.

Medications will be coded using the World Health Organization (WHO) drug dictionary. Concomitant medications will be listed.

6.11. Efficacy Analyses

6.11.1. *Primary Outcome and Methodology*

The primary objective of this study is to test the hypothesis that IV infusion of LY3303560 will slow the cognitive and/or functional decline of AD as measured by the composite measure iADRS compared with placebo in patients with early symptomatic AD. This will be assessed using a Disease Progression Model (DPM).

The iADRS at each scheduled visit (according to the SoA) during the treatment period will be the dependent variable. The DPM is as follows:

$$Y_{ij} = \gamma_i + e^{\theta_{T_i}} \sum_{v=0}^{j} \alpha_v + \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_{ij}, i = 1, 2, \ldots k; j = 1, 2, \ldots l$$

where Y_{ij} denotes the clinical outcome at visit *j* for participant *i* and clinical outcome score at baseline (prior to treatment) is Y_{i0}. γ_i (*i*=1, 2, …, *k*) represents the random participant effects. T_i denotes the treatment arm for participant i. T_i takes a value of 1 if a participant is randomized to the low dose, a value of 2 for high dose, and a value of 0 for placebo. e^{θ_T} is the disease progression ratio (DPR) for treatment *T* and $e^{\theta_T} = 1$ for placebo. e^{θ_1} and e^{θ_2} represent the DPR of the low dose and the high dose, respectively. Furthermore, α_v is the change in mean cognitive score for placebo from visit *v*-1 to *v*. ε_{ij} is the error term. Additional covariates $\boldsymbol{x}_i$ will include pooled site ID, age, and baseline AChEI/memantine use (yes/no).

The DPM model assumes the treatment effect of LY3303560 is proportional to placebo over the course of the study. This proportionality assumption is similar to the assumption made in proportional hazards modeling of time to event data. The DPM includes generally diffuse priors on all parameters. The precision value for the prior distribution of all parameters is set to a small value; therefore, the prior distributions on all parameters have very little impact on the posterior distributions. No information or knowledge of the effect of LY3303560 from previous studies will be incorporated into the prior distributions, and the inference will be based on Study LMDC only. Diffuse priors will be used.

For subject level random effect $\gamma_i \sim N(\mu_\gamma, \sigma_\gamma^2)$, a diffuse normal prior is assumed for μ_γ, the mean of the subject level random effect and a diffuse inverse gamma prior is considered for σ_γ^2, the variance of the subject level random effect:

$$\mu_\gamma \sim N(0, 10^2),\ \sigma_\gamma^2 \sim IG(0.005, 0.005).$$

A diffuse normal prior distribution is assumed for each mean change in disease progression from visit $v-1$ to visit v in the placebo, α_v:

$$\alpha_v \sim N(1.25, 900)$$

A normal prior distribution $N(0,4)$ is assumed for the log(DPR) parameter of each active treatment arm (i.e., θ_1 and θ_2). For the error term $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$, an inverse gamma prior is assumed for the model error variance:

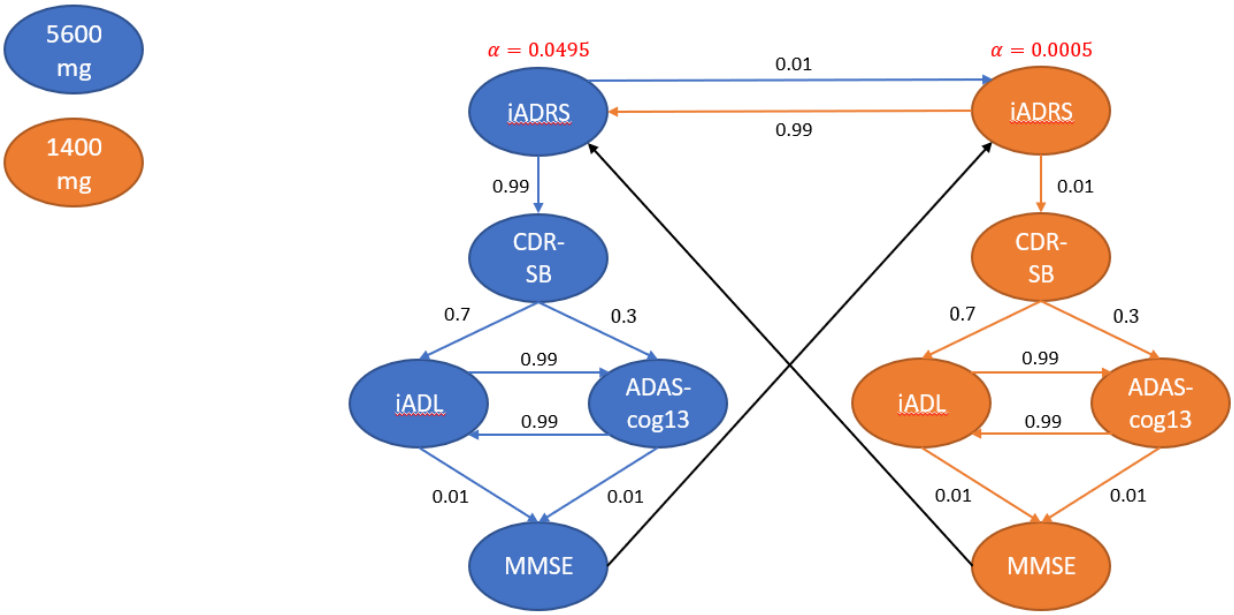
$$\sigma_\epsilon^2 \sim IG(0.005,0.005).$$

Diffuse normal prior of the form N(0,100) is used for each baseline covariate (pooled site ID, age and baseline AChEI/memantine use (yes/no)) in the model.A DPR less than 1 favors LY3303560 and corresponds to a slowing of disease progression with LY3303560 in comparison to placebo; similarly, a DPR greater than 1 favors placebo. The DPM will be fit to the data and Bayesian inferences will be summarized including posterior distribution of DPR and posterior probabilities of various DPR thresholds of interest (for example, 0.75 which translates to 25% slowing of disease progression with LY3303560 group versus placebo). To test the hypothesis of a disease progression benefit, we calculate the posterior probability of superiority in cognitive/functional slowing and if it is above a pre-specified threshold (which controls the false positive rate at the values in the graphical testing scheme), then a claim of cognitive/functional slowing will be made. The null hypothesis is that the DPR between the LY3303560 group versus placebo equals 1.

In addition, Bayesian posterior probability of active treatment arm being superior to placebo by at least a margin of interest will also be calculated assuming a noninformative prior.

6.11.2. *Gated Secondary Efficacy Analyses*

Bretz's graphical approach may be utilized to provide strong control of the study-wise type I error rate for the primary and key secondary hypotheses at a two-sided alpha level of 0.05 (Bretz et al. 2009, 2011). The DPM analyses will be conducted on the iADRS, CDR-SB, ADAS-Cog13, iADL and MMSE scores, and statistical significance will be determined based on the following multiplicity graph of hypotheses regarding the IV infusion of LY3303560 slowing the cognitive and/or functional decline of AD:



iADRS = DPM analysis of iADRS

CDR-SB = DPM analysis of CDR-SB

iADL = DPM analysis of iADL

ADAS-Cog13 = DPM analysis of ADAS-Cog13

MMSE = DPM analysis of MMSE

The preceding graphical testing scheme is outlined using alpha levels in a frequentist testing paradigm. The primary efficacy analysis for this trial is the Bayesian DPM. The following describes how probability thresholds will be developed for each test (and potential test) in the graph that control the false positive rate at the appropriate corresponding alpha level. The tests for each scale will all be of the form Pr(DPR < 0.75) > X, where the value of X is determined via simulation to control the test at the corresponding alpha level.

To test the hypothesis of a disease progression benefit, the posterior probability of superiority in cognitive/functional slowing will be calculated, and if it is above a pre-specified threshold (which controls the false positive rate at approximately 2.5%), then a statistical significant claim of cognitive/functional slowing will be made. Success will be declared at a graphical node if the posterior probability of at least 25% slowing is greater than a certain threshold, where the threshold is determined via simulation to control the false positive rate under the null scenario for a given scale. The null scenario used in these simulations to determine the posterior probability thresholds will be the overall mean trajectory from the blinded LMDC trial data, conducted at a date late in the trial but prior to unblinding. As a sensitivity analysis of the selected posterior probability thresholds, additional simulations may be run after unblinding using the unblinded data to understand the impact of the choice of variance-covariance matrix and mean placebo trajectory on the posterior probability thresholds used to control the false positive rate.

Assuming the null scenario of 0% slowing in treated arms relative to the placebo, synthetic data are sampled from a multivariate normal distribution for each endpoint that accounts for the correlation between visits after adjusting for concomitant drug use, pooled site, and age. The mean trajectory and variance-covariance matrix for each scale is the overall (i.e. treatment-naïve) least-squares mean (from frequentist MMRM analyses adjusting for pooled site, baseline AChEI/memantine use and age) at each visit in the blinded trial data. The Bayesian DPM is fit to each simulated dataset and the posterior probability of at least 25% slowing is calculated. The threshold is the value at which approximately $(\alpha/2 \times 100)$% of the simulated posterior probabilities are larger. The cutoff for the primary analysis of LY5600 is 0.51 (Pr(DPR < 0.75) > 0.51); the cutoff for LY1400 is 0.99928 (Pr(DPR < 0.75) > 0. 99928). The other thresholds will be calculated in the same way.

6.11.3. *Additional Exploratory Analyses of the Primary Outcome*

6.11.3.1. Mixed Model Repeated Measures (MMRM)

The change from baseline score on the iADRS at each scheduled postbaseline visit (according to the Schedule of Activities) during the treatment period will be the dependent variable. The model for the fixed effects will include the following terms: baseline score, pooled site, treatment, visit, treatment-by-visit interaction, baseline-by-visit interaction, concomitant AChEI and/or memantine use at baseline (yes/no), and age at baseline. Visit will be considered a categorical variable. The null hypothesis is that the contrast between the LY3303560 group versus placebo at the Visit 28, Week 104 (4 weeks after the final dose of LY3303560) equals zero. An unstructured covariance matrix will be used to model the within-subject variance–covariance errors. If the unstructured covariance structure matrix results in a lack of convergence, the following tests will be used in sequence: heterogeneous Toeplitz covariance structure, heterogeneous autoregressive covariance structure, heterogeneous compound symmetry covariance structure, and compound symmetry covariance structure. The Kenward–Roger approximation will be used to estimate the denominator degrees of freedom.

The primary time point for treatment comparison will be at Week 104. The treatment group contrast in least-squares mean change from baseline score on the iADRS scale and its associated p-value and 95% CI will be calculated for treatment comparisons of LY3303560 versus placebo using the MMRM model specified above.

6.11.3.2. Natural Cubic Spline

In addition to the DPM and MMRM models described above, the iADRS will be modeled using natural cubic splines (NCS) (Chambers and Hastie 1992). The postbaseline score on the iADRS at each scheduled postbaseline visit (according to the Schedule of Activities) during the treatment period will be the dependent variable, and the mean for each treatment group over the entire double-blind duration of the study will be modeled. The NCS model provides a type of smoothing function to the data, can adequately estimate longitudinal trajectories under a variety of shapes (e.g., linear, quadratic, etc.) for each treatment group, and is thought to be a more

natural parameterization than modeling the mean change from baseline over time. The degrees of freedom of the model can be prespecified to establish the level of smoothing of the data. The number and location of the "knots" are utilized to parse out different time periods where the data may transition from one shape to another to provide an adequate fit. The primary time point for treatment comparison will be at Week 104. The variance-covariance structure assumptions of the NCS model are the same as the MMRM model and the covariates used in the model would remain unchanged. The model would be estimated by restricted maximum likelihood using `SAS PROC MIXED` or the `gls` function from the `nlme` package in the `R`.

6.11.3.3. Minimization (Dynamic Allocation) Assessment

Minimization (dynamic allocation): Because treatment assignment followed a minimization procedure rather than a randomization algorithm, a randomization test (i.e., permutation test) along with the MMRM analyses to confirm the asymptotic inference for the MMRM analysis will be conducted (Proschan et al. 2011). The main features of the randomization test would be: 1) the patients' covariates, responses and enrollment order would be considered fixed, 2) the sharp null hypothesis would be assumed (i.e., the patients responses would be assumed exactly the same under LY3303560 or placebo), 3) the exact minimization algorithm and exact pooling site algorithm would be reproduced in order to generate a null distribution of the test statistic from the MMRM, and 4) the p-value from the generated null distribution would be obtained by comparing the observed data p-value to the percentiles of the generated distribution. Because of the sample size of the study, the exact null distribution will be impossible to enumerate, but a Monte Carlo approach will be used with a large number of realizations (approx 10,000) to simulate the null distribution.

6.11.3.4. Delta Adjustment Tipping Point Analysis

Sensitivity to departures from the missing-at-random (MAR) assumption will be investigated using a tipping point analysis (Carpenter and Kenward 2013). This method is a sensitivity analysis in multiple imputation under the missing-not-at-random (MNAR) assumption that searches for a tipping point that reverses the study conclusion. Departures from MAR in the LY3303560 treatment groups will be assessed assuming that patients who discontinue the study have, on average, efficacy outcomes after discontinuation that are worse by some amount δ compared to other similar patients with observed data (ie, compared to a value which would have been assumed under an MAR model). A series of analyses will be performed with increasing values of δ until the analysis conclusion of a statistically significant treatment effect no longer holds. The value of δ that overturns the primary results will represent a tipping point. An interpretation of clinical plausibility of the assumption underlying the tipping point will be provided.

Mean changes from baseline in iADRS scores will be analyzed based on data observed while the patient remains on study as well as data imputed using multiple imputation (MI) methodology for time points at which no value is observed. Imputed values in the LY3303560 treatment groups will first be sampled from an MAR-based multiple imputation model and then δ-adjusted as described below.

Missing-at-random-based imputations will be generated for iADRS scores at each time point, and then a value of $\delta = \{\Delta\}$ will be added to all imputed values in the LY3303560 treatment groups prior to analyzing multiply imputed data. This approach assumes that the marginal mean of imputed patient measurements is worse by δ at each time point after discontinuation compared to the marginal mean of patients with observed data at the same time point. Analyses will be conducted with values of δ starting from 0 with increments of 0.10 until the null hypothesis can no longer be rejected.

6.11.3.5. Bayesian Analysis of Shared Control

Sensitivity of comparative inference for slowing the cognitive and/or functional decline of AD based on the shared control may be accomplished via Bayesian mixture modeling. Supplementing the iADRS analyses with placebo data from studies I5T-MC-AACG, I8D-MC-AZES, and H8A-MC-LZAX will be explored and potentially based on matching baseline tau PET scan results (Viele, 2014).

6.11.3.6. Slope Analyses

Slopes of the iADRS will be assessed using an MMRM analysis. The change from baseline score at each post-baseline visit during the treatment period will be the dependent variable. The model will include the fixed, categorical effects of treatment, concomitant AChEI or memantine use at baseline (yes/no), pooled site, and continuous effects of baseline score, time, time-by-treatment interaction, and age at baseline. Time will be assumed to be a continuous variable calculated as number of days between baseline and each postbaseline visit (ie, [visit-baseline]+1) during the treatment period. The actual visit dates will be used to calculate number of days (time). The null hypothesis is that the contrasts of slopes of LY3303560 dose groups versus placebo equal zero.

A quadratic slopes model will also be fitted to these same scales. The quadratic model will include the linear component of time (TIME) and a quadratic component of time (TIME*TIME), the linear component of time and treatment interaction (TIME*TREATMENT) and quadratic component of time and treatment interaction (TIME*TIME*TREATMENT).

6.11.3.7. Completer Analyses

The primary efficacy outcome, iADRS, from the dataset of those patients who remained in the study and on treatment through Week 104 ("completers") will be analyzed using an ANCOVA. The change from baseline at Week 104 will be the dependent variable. The model will include the fixed, categorical effects of treatment, concomitant AChEI use at baseline (yes/no), pooled site, and the continuous effects of baseline iADRS score and age at baseline. The null hypothesis is that the differences in least-squares means between the LY3303560 dose groups versus placebo at Week 104 equal zero.

6.11.3.8. Per Protocol Analyses

The primary efficacy outcome, iADRS, from the per-protocol dataset will be analyzed using the DPM from the primary analysis. The change from baseline at each scheduled postbaseline visit will be the dependent variable. The model for the fixed effects will include the following terms

(same as primary efficacy analysis): baseline score, pooled site, treatment, visit, treatment-by-visit interaction, baseline-by-visit interaction, concomitant AChEI and/or memantine use at baseline (yes/no), and age at baseline. The null hypothesis is that the differences in least-squares means between the LY3303560 dose groups versus placebo at Week 104 equal zero.

6.11.4. *Other Secondary Efficacy Analyses*

The additional clinical and outcome measurements listed below will be analyzed separately using an MMRM analysis. The change from baseline at each scheduled postbaseline visit will be the dependent variable. The model for the fixed effects will include the following terms (same as primary efficacy analysis): baseline score, pooled site, treatment, visit, treatment-by-visit interaction, baseline-by-visit interaction, concomitant AChEI and/or memantine use at baseline (yes/no), and age at baseline. The null hypothesis is that the differences in least-squares means between the LY3303560 dose groups versus placebo at Week 104 equal zero. The outcomes that will be analyzed are:

- Change from baseline as obtained from the ADAS-Cog_{13}
- Change from baseline in ADCS-ADL total score
- Change from baseline in ADCS-iADL score
- Change from baseline in ADCS-bADL score
- Change from baseline in CDR-SB
- Change from baseline in MMSE

6.12. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

6.12.1. *Analysis of flortaucipir PET*

To evaluate the change from baseline in tau imaging parameters, an MMRM analysis will be used to compare change from baseline in SUVr at 104 weeks in the Evaluable Efficacy dataset. The model will include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, as well as continuous effects of baseline SUVr and age at baseline. Visit will be considered a categorical variable with values equal to the visit numbers at which tau imaging is assessed. The null hypothesis is that the difference in LSM between the LY3303560 dose groups and placebo equals zero.

To assess the relationship of biomarker with cognition and function with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to Week 104 for the composite summary standard uptake value ratio (SUVr) normalized to bimodal white matter and with change from baseline to Week 104 for iADRS, ADAS-Cog_{13}, ADCS-ADL, MMSE, and CDR-SB. Correlation analyses will be conducted using only patients who have the clinical outcome and SUVr result at Week 104 and include patients from all 3 dose groups.

6.12.2. *Analysis of NfL*

To evaluate the change from baseline in neurofilament light chain (NfL) concentrations, an MMRM analysis will be used to compare change from baseline between treatments up to 104 weeks of treatment in the Evaluable Efficacy dataset. All NfL values will be log transformed

prior to calculating baseline and change values. The model will include the fixed, categorical effects of treatment, visit, sex, and treatment-by-visit interaction, as well as continuous effects of log baseline, age, and log baseline-by-treatment interaction. Visit will be considered a categorical variable with values equal to the visit numbers at which NfL is assessed. The null hypothesis is that the difference in LSM between the LY3303560 dose groups and placebo equals zero.

6.12.3. *Analysis of Plasma Tau*

To evaluate the change from baseline in plasma tau concentrations (total tau and phospho tau 181), an MMRM analysis will be used to compare change from baseline between treatments up to 104 weeks of treatment in the Evaluable Efficacy dataset. All plasma tau values will be log transformed prior to calculating baseline and change values. The model will include the fixed, categorical effects of treatment, visit, sex, and treatment-by-visit interaction, as well as continuous effects of log baseline, age, and log baseline-by-treatment interaction. Visit will be considered a categorical variable with values equal to the visit numbers at which tau is assessed. The null hypothesis is that the difference in LSM between the LY3303560 dose groups and placebo equals zero.

6.12.4. *Analysis of vMRI*

Analyses of the following volumetric MRI (vMRI) parameters will be conducted (right + left for all but whole brain volume and ventricular volume):

- Hippocampal volume (cm^3)
- Entorhinal cortex (cm^3)
- Inferior parietal lobe (cm^3)
- Isthmus cingulate (cm^3)
- Lateral parietal lobe (cm^3)
- Medial temporal lobe (cm^3)
- Precuneus (cm^3)
- Prefrontal lobe (cm^3)
- Superior temporal lobe (cm^3)
- Cortical (cm^3)
- Whole temporal lobe (cm^3)
- Atrophy of total whole brain volume (cm^3)
- Enlargement of Ventricular volume (cm^3)
- White Matter (cm3)
- White Matter Hypo Intensities (cm3)

All of the above volumes are corrected for intracranial volume. To evaluate the changes in vMRI data after treatment, an MMRM analysis will be used to compare change from baseline to 104 weeks in the Evaluable Efficacy dataset. The change from baseline to the endpoint visit will be the dependent variable. The model will include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, as well as continuous effects of baseline vMRI, baseline intracranial volume (ICV) and age. The null hypothesis is that the difference in LSM between the

LY3303560 active dose groups and placebo equal zero. A similar analysis will be performed for completers.

To assess the relationship of vMRI with cognition and function with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to Week 104 for vMRI parameters with change from baseline to Week 104 for iADRS, ADAS-Cog_{13}, ADCS-ADL, MMSE, and CDR-SB; this will be performed using all patients who have the clinical outcome and vMRI result at Week 104; adjusted for age, baseline intracranial volume (ICV), and sex; and include patients from both treatment groups.

6.13. Safety Analyses

6.13.1. *Extent of Exposure*

Maximum days of exposure will be calculated for each patient (date of last dose – date of first dose +1). Accounting for missed infusions, minimum days of exposure will be collected for each patient (number of infusions*28). Summary statistics will be provided for the total number of days and patient-years of exposure by treatment. Additionally, the number of infusions will be summarized by treatment. Study drug treatment assignment will be listed.

6.13.2. *Adverse Events*

Treatment-emergent adverse events (TEAEs) will be defined as events that first occurred or worsened after the randomization date (Visit 2 date). Should there be insufficient data for AE start date, stop date, and time to make this comparison, the AE will be considered a post-baseline event and eligible for being treatment-emergent. The MedDRA lower-level term (LLT) will be used in the treatment-emergent computation. The maximum severity for each lower-level term (LLT) during the baseline period will be used as baseline.

An overview of AEs, including the number and percentage of patients who died, suffered serious adverse events (SAEs), discontinued due to AEs and who suffered TEAEs, will be provided. Comparison between treatments will be performed using Fisher's Exact Test.

Summaries of AEs by decreasing frequency of PT within SOC will be provided for the following:

- Preexisting conditions
- TEAEs
- TEAEs by maximum severity
- TEAEs occurring in greater than or equal to 2% of patients by PT
- Serious adverse events
- Adverse events and death reported as reason for study treatment discontinuation

These summaries will include number and percentages of patients with TEAEs. Treatment comparisons will be carried out using Fisher's Exact Test.

Preexisting conditions, TEAEs, SAEs, and discontinuations due to AEs will be listed.

6.13.3. *Deaths, Other Serious Adverse Events, and Other Notable Adverse Events*

An overview of AEs, including the number and percentage of patients who died or suffered SAEs during the study, discontinued due to AEs and who suffered TEAEs, will be provided. Comparison between treatments will be performed using Fisher's Exact Test.

6.13.4. *Clinical Laboratory Evaluation*

Laboratory measurements will be analyzed using continuous data (change from baseline) and categorical or ordinal data (proportion of treatment-emergent abnormalities). If there are multiple records of laboratory measurements at baseline or postbaseline visit, the last record will be used. Summaries and analyses of continuous data (change from baseline) will be performed using both conventional and International System of Units (SI units).

Measures of central tendency for planned lab analytes' raw measurements and change from baseline (in CN and SI units) will be summarized with boxplots. Boxplots will display results semi-annually (visits 8, 15, 22, and 28) and for the last visit (LOCF) and will include summary tables of N, mean, median, quartiles, min, max, standard deviation, and p-value (for change scores). If there are considerable missing visits, the measures of central tendency may be based on MMRM analyses.

Treatment differences in the proportion of patients with treatment-emergent high or treatment-emergent low or treatment-emergent abnormal laboratory values at (1) anytime and (2) semi-annually (visits 8, 15, 22, and 28) will be assessed using Fisher's exact test. Treatment-emergent high or low laboratory abnormality will be based on SI unit. For each laboratory analyte, only patients who were low or normal at all baseline assessments and have at least 1 post-baseline will be included in the denominator when computing the proportion of patient with treatment-emergent high. Similarly, only patients who were high or normal at all baseline assessments and have at least 1 post baseline will be included in the denominator when computing the proportion of patient with treatment-emergent low. In addition, treatment differences in the proportion of patients who have normal baselines with a change to abnormal high or abnormal low values at any post-baseline visits will be summarized.

For urinalysis parameters, baseline to post-baseline shifts will be summarized at each protocol-specified visit. Likelihood ratio chi-square tests will be used to compare increase, no change, and decrease shifts in urinalysis parameters between treatment groups at each visit.

The proportion of patients with treatment-emergent clinically significant changes from a low value or normal value at all baselines at any time in ALT and total bilirubin will be summarized by treatment group. Clinically significant changes of interest at any time are: ALT ≥3 x upper limit of normal (ULN) and total bilirubin ≥2 x ULN, AST ≥3 x ULN, ALT ≥5 x ULN, ALT ≥10 x ULN, and total bilirubin ≥2 x ULN. Additionally, Hy's Law analysis will be conducted by comparing treatment groups with regard to the proportion of patients with (ALT ≥3 x ULN OR AST ≥3 x ULN) AND total bilirubin ≥2 x ULN at any time. Comparisons between treatment groups will be carried out using Fisher's Exact test. When criteria are met for hepatic evaluation

and completion of the hepatic safety case report form (CRF), investigators are required to answer a list of questions pertaining to the patient's history, relevant pre-existing medical conditions, and other possible causes of liver injury. A listing of the information collected on the hepatic-safety CRF will be generated.

6.13.5. *Vital Signs and Other Physical Findings*

Vital sign measurements and weight will be analyzed using continuous data (change from baseline) and categorical data (proportion of potentially clinically significant changes) using the Safety Dataset.

If there are multiple records of vital sign or weight measurements at baseline or postbaseline visit, the last record will be used. Summary statistics will be presented for observed values at baseline and for change from baseline results at each scheduled postbaseline visit. Systolic and diastolic blood pressure and pulse (collected in sitting position), orthostatic diastolic and orthostatic systolic blood pressures and orthostatic pulse, temperature, and weight by treatment group for all patients in the safety population will be summarized.

Measures of central tendency for vital sign or weight raw measurements and change from baseline will be summarized with boxplots. Boxplots will display results semi-annually (visits 8, 15, 22 and 28) and for the last visit (LOCF) and will include summary tables of N, mean, median, quartiles, min, max, standard deviation, and p-value (for change scores). If there are considerable missing visits, the measures of central tendency may be based on MMRM analyses.

In order to assess outliers and potentially clinically significant changes from baseline, the number and percent of patients meeting criteria for treatment-emergent abnormalities in vital signs and weight at any time during study will be summarized. Treatment group comparisons will be performed using Fisher's exact test. Baseline is defined as the entire screening period (visits 1 and 2), and all baseline assessments must not meet abnormality criteria to be included in the analysis. Abnormal criteria for post-baseline vital signs and weight are presented in Appendix 1. Any vital sign or weight meeting the criteria will be considered abnormal. Treatment differences in the proportion of patients with treatment-emergent abnormal high or low vital signs and weight will be assessed between treatment groups using Fisher's exact test at (1) any time (2) semi-annually (visits 8, 15, 22, and 28).

A listing of treatment-emergent abnormal vital signs and weight will also be presented.

6.13.6. *Electrocardiograms*

ECG measurements will be analyzed using continuous data (change from baseline) and categorical data (proportion of treatment-emergent abnormalities) using the Safety Dataset.

The ECG measurements are derived from three 10 second readings taken every 30 seconds. These 3 readings are to be averaged prior to analysis. Additionally, whenever ECG is measured in triplicate, the average of these readings will be used in the analysis. If there are multiple records after averaging ECG triplicates within a visit, the last record of averages will be used.

The analysis will be done for the following ECG measurements: heart rate, PR, QT, QTc, and RR intervals and QRS duration. All analyses of QTc will be carried out using the Fridericia correction (QTcF) method. Measures of central tendency for ECG raw measurements and change from baseline will be summarized with boxplots. Boxplots will display results semi-annually (visits 8, 15, 22, and 28) and for the last visit (LOCF) and will include summary tables of N, mean, median, quartiles, min, max, standard deviation, and p-value (for change scores). If there are considerable missing visits, the measures of central tendency may be based on MMRM analyses.

In order to assess outliers and potentially clinically significant changes from baseline, the number and percent of patients meeting criteria for treatment-emergent abnormalities in ECGs will be summarized. Treatment group comparisons will be performed using Fisher's exact test. Baseline is defined as the entire screening period (visits 1 and 2) , and all baseline assessments must not meet abnormality criteria to be included in the analysis. Incidence of treatment-emergent abnormal ECGs will be assessed by comparisons at (1) anytime and (2) semi-annually (visits 8, 15, 22, and 28).

Abnormal ECG criteria and criteria for abnormal QTcF prolongation are presented in Appendix 2.

Treatment-emergent high ECG parameters (heart rate, PR interval, QRS duration, QT and QTcF intervals) are the values which are low or normal at all baseline visits and fall into the high abnormal categories post-baseline. Similarly, treatment-emergent low ECG parameters (heart rate, PR interval, QRS duration) are the values which are high or normal at all baseline visits and fall into the low abnormal categories above.

In addition, treatment differences in the proportion of patients who have normal baselines with a change to abnormal high or abnormal low values at any post-baseline visits will be summarized.

6.13.7. *Safety MRIs*

To evaluate white matter changes over time, a shift table will be created from the following categories:

- 0 = No lesions
- 1 = Focal lesions
- 2 = Beginning confluence of lesions
- 3 = Diffuse involvement of entire region

A listing of MRI data will also be presented.

6.13.8. *Immunogenicity*

The frequency and percentage of subjects with preexisting (baseline) ADA, ADA at any time after baseline, and TE-ADAs to LY3303560 will be summarized. If no ADAs are detected at baseline, TE-ADAs are defined as those with a titer 2-fold (1 dilution) greater than the MRD of the assay. For samples with ADA detected at baseline, TE-ADA are defined as those with a 4-fold (2 dilutions) increase in titer compared to baseline. For the TE-ADA subjects, the

distribution of maximum titers will be summarized. The frequency of subjects with neutralizing antibodies (subset of the TE-ADA patients) will also be summarized.

6.13.9. *Columbia Suicide Severity Rating Scale*

Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent occurring during treatment, based on the Columbia-Suicide Severity Rating Scale (C-SSRS), will be summarized by treatment. In particular, for each of the following events, the number and percent of patients with the event will be enumerated by treatment: completed suicide, nonfatal suicide attempt, interrupted attempt, aborted attempt, preparatory acts or behavior, active suicidal ideation with specific plan and intent, active suicidal ideation with some intent to act without specific plan, active suicidal ideation with any methods (no plan) without intent to act, nonspecific active suicidal thoughts, wish to be dead, and self-injurious behavior without suicidal intent. Although not suicide-related, the number and percent of patients with non-suicidal self-injurious behavior occurring during the treatment period will also be summarized by treatment.

In addition, the number and percent of patients who experienced at least one of various composite measures during treatment will be presented and compared. These include suicidal behavior (completed suicide, non-fatal suicidal attempts, interrupted attempts, aborted attempts, and preparatory acts or behavior), suicidal ideation [active suicidal ideation with specific plan and intent, active suicidal ideation with some intent to act without specific plan, active suicidal ideation with any methods (no plan) without intent to act, non-specific active suicidal thoughts, and wish to be dead], and suicidal ideation or behavior.

The number and percent of patients who experienced at least one of various comparative measures during treatment will be presented and compared. These include treatment-emergent suicidal ideation compared to recent history, treatment-emergent serious suicidal ideation compared to recent history, emergence of serious suicidal ideation compared to recent history, improvement in suicidal ideation at endpoint compared to baseline, and emergence of suicidal behavior compared to all prior history.

Specifically, the following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definitions of the composite and comparative endpoints, and to enable clarity in the presentation of the results.

Category 1 – Wish to be Dead
Category 2 – Non-specific Active Suicidal Thoughts
Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5 – Active Suicidal Ideation with Specific Plan and Intent
Category 6 – Preparatory Acts or Behavior
Category 7 – Aborted Attempt
Category 8 – Interrupted Attempt
Category 9 – Actual Attempt (non-fatal)
Category 10 – Completed Suicide

Self-injurious behavior without suicidal intent is also a C-SSRS outcome (although not suicide-related) and has a binary response (yes/no).

Composite endpoints based on the above categories are defined below.

- Suicidal ideation: A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS.
- Suicidal behavior: A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS.
- Suicidal ideation or behavior: A "yes" answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.

The following outcome is a numerical score derived from the C-SSRS categories. The score is created at each assessment for each patient and is used for determining treatment emergence.

- Suicidal Ideation Score: The maximum suicidal ideation category (1-5 on the C-SSRS) present at the assessment. Assign a score of 0 if no ideation is present.

Comparative endpoints of interest are defined below. "Treatment emergence" is used for outcomes that include events that first emerge or worsen. "Emergence" is used for outcomes that include events that first emerge.

- Treatment-emergent suicidal ideation compared to recent history:
  An increase in the maximum suicidal ideation score during treatment (Visits Y1-Y2) from the maximum suicidal ideation category during the screening and lead-in periods (C-SSRS scales taken at Visits X1-X2). Recent history excludes "lifetime" scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale.
- Treatment-emergent serious suicidal ideation compared to recent history: An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment (Visits Y1-Y2) from not having serious suicidal ideation (scores of 0-3) during the screening and lead-in periods (C-SSRS scales taken at Visits X1-X2). Recent history excludes "lifetime" scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale.
- Emergence of serious suicidal ideation compared to recent history:
  An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment (Visits Y1-Y2) from no suicidal ideation (scores of 0) during the screening and lead-in periods (C-SSRS scales taken at Visits X1-X2). Recent history excludes "lifetime" scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale.
- Improvement in suicidal ideation at endpoint compared to baseline:
  A decrease in suicidal ideation score at endpoint (the last measurement during treatment; Visits Y1-Y2) from the baseline measurement (the measurement taken just prior to treatment; (Visit X2). This analysis should only be performed for a non-lifetime baseline measurement (i.e., having improvement from the worse event over a lifetime is not clinically meaningful). A specific point in time can be used instead of endpoint.
- Emergence of suicidal behavior compared to all prior history:

The occurrence of suicidal behavior (Categories 6-10) during treatment (Visits Y1-Y2) from not having suicidal behavior (Categories 6-10) prior to treatment (Visits X1-X2). Prior to treatment includes "lifetime" and/or "screening" scores from the Baseline C-SSRS scale, Screening C-SSRS scale, or Baseline/Screening C-SSRS scale, and any "Since Last Visit" from the Since Last Visit C-SSRS scales taken prior to treatment.

Patients who discontinued from the study with no postbaseline C-SSRS value will be considered unevaluable for analyses of suicide-related events. Only evaluable patients will be considered in the analyses. Fisher's exact test will be used for treatment comparisons.

6.14. Subgroup Analyses

To assess the effects of various demographic and baseline characteristics on treatment outcome, subgroup analyses for the primary endpoint, iADRS:

- APOE4 Carrier Status – Carrier defined as E2/E4, E3/E4, or E4/E4 genotype; No-Carrier defined as all other genotypes
- Disease status at baseline based on MMSE total score
    - MCI or mild AD (MCI: 27-30; mild AD: 20-26)
    - Grouped into thirds
- Tau burden at baseline based on eligibility tau PET SUVR
    - very low tau: SUVR < 1.10; low tau: 1.10 <= SUVR < 1.23; medium tau: 1.23 <= SUVR < 1.46
    - Grouped into thirds

The primary outcome measure will be modeled using a DPM approach. This general model will include terms for baseline, treatment, pooled site, visit, concomitant AChEI/memantine use at baseline (yes/no), baseline age, treatment by visit, subgroup by treatment, subgroup by visit, and treatment by visit by subgroup. Redundant terms will be dropped from the model in those cases where the subgroup of interest is overlapping with this general model. In order to run these analyses, at least 20 patients are required in each strata-treatment combination.

6.15. Protocol Violations

Listings of patients with significant protocol violations will be provided for the Randomized population. The following list of significant protocol violations will be determined from the clinical database and from the clinical/medical group:

- Informed consent violation detected as a missing date of informed consent.
- Did not have an assessment of either the ADAS-Cog or ADL at any of the visits at which the scales were scheduled to be assessed.
- Incomplete infusions (any infusion at which less than 75%, approximately 190 mL, of the infusion solution is given).

The following list of significant protocol violations will be determined by clinical/medical group:

- Protocol violations of inclusion/exclusion criteria.

- Had a study dosing algorithm violation (such as if patients randomized to treatment A were given treatment B or patients randomized to treatment A never received the assigned study drug.)
- Unqualified raters for the ADAS-Cog or ADL.

Other protocol violations reported through the monitoring process will be reviewed by the study team and if judged to be significant, will be added to the final reported listing.

6.16. Interim Analyses and Data Monitoring

An external DMC is authorized to evaluate results from unblinded interim analyses for the assessment of safety, to recommend any modifications to the study (such as stopping the study or dropping an arm), and inform future development. The DMC will have the responsibility to review accumulating unblinded study data and make recommendations to protect the safety of patients. Each member of the DMC is a recognized expert in the fields of Alzheimer's Disease, geriatric neurology, geriatric psychiatry, or biostatistics. All members will be external to the Sponsor. The approved DMC charter enumerates the roles of the DMC members, the frequency with which it meets, and the structure of their meetings. Study sites will receive information about interim results ONLY if relevant for the safety of their patients.

6.16.1. *Interim Analysis 1*

The objective of the first interim analysis (IA1) is to assess the viability of continuing with at least 1 dose of LY3303560 with regard to safety. An unblinded DMC review for safety will be conducted once approximately 30 patients have completed 3 months of exposure to study treatment (i.e., after completing 4 weeks following the third dose of study treatment). No patients beyond this initial group of approximately 30 patients will be dosed until the DMC review has been completed and the recommendation is to continue the study. This initial group will continue treatment during the assessment period. No statistical adjustments will be made to account for this interim since it is a safety only interim and there is no possibility of stopping for efficacy.

The following analyses will be included in IA1:

- Summary of Patient Disposition
- Summary of Patient Compliance
- Summary of Patient Demographics
- Summary of Patient Baseline Severity
- Summary of Patient Exposure
- Summary of Concomitant Medications
- Summary of Concomitant AChEI/Memantine Medications
- Summaries of Adverse Events (SAEs, DCAE's, TEAE's)

- Listing of Adverse Events
- Summaries of Laboratories (Categorical and Mean Change from Baseline)
- Listing of Laboratories
- Summaries of Vital Signs and Weight (Categorical and Mean Change from Baseline)
- Listing of Vitals Signs and Weight
- Summaries of ECGs (Categorical and Mean Change from Baseline)
- Listing of ECGs
- Summary of Safety MRIs
- Listing of MRIs
- Summary of C-SSRS
- Listing of C-SSRS

6.16.2. *Safety Summaries for Interim Analysis 2*

The efficacy analyses used in Interim Analysis 2 (IA2) will be detailed in the Interim Efficacy SAP. In addition to the listings and summaries reviewed at IA1, the following analyses will be included in Interim Analysis 2 (IA2):

- Listing of patients missing site assessments and infusions because of COVID-19
- Listing of patients discontinuing the study because of COVID-19
- Dot plot of visits (x-axis) by patient (y-axis) with color denoting missed infusions
- Summary of ADA
- ANCOVA of change in whole brain volume from vMRI

6.16.3. *Interim Analysis 3*

A third unblinded DMC review will be conducted for safety when approximately 100% of patients, have completed 18 months of exposure to study treatment (that is, 4 weeks after the 18th dose at Visit 21 [Week 76]). The following safety analyses will be included:

- Summary of Patient Disposition
- Summary of Patient Demographics
- Summary of Patient Exposure
- Summary of Concomitant Medications
- Summaries of Adverse Events (SAEs, DCAE's, TEAE's)
- Listing of Serious Adverse Events

- Summaries of Laboratories (Categorical and Mean Change from Baseline)
- Listing of Abnormal Laboratories
- Summaries of Vital Signs and Weight (Categorical and Mean Change from Baseline)
- Summaries of ECGs (Categorical and Mean Change from Baseline)
- Summary of Safety MRIs
- Listing of MRIs
- Summary of C-SSRS
- Listing of C-SSRS
- Summary of ADA
- ANCOVA of change in whole brain volume from vMRI

6.17. Planned Exploratory Analyses

6.17.1. *Exploratory Efficacy Analyses*

6.17.1.1. MMRM Analyses

The additional clinical and outcome measurements listed below will be analyzed separately using an MMRM analysis. The change from baseline at each scheduled postbaseline visit will be the dependent variable. The model for the fixed effects will include the following terms (same as primary efficacy analysis): baseline score, pooled site, treatment, visit, treatment-by-visit interaction, baseline-by-visit interaction, concomitant AChEI and/or memantine use at baseline (yes/no), and age at baseline. The null hypothesis is that the differences in least-squares means between the LY3303560 dose groups versus placebo at Week 104 equal zero. The outcomes that will be analyzed are:

- Change from baseline in CogState Brief Battery (CBB)
- Change from baseline in Digital Clock Drawing Test (DCTClock)

6.17.1.2. Ordinal Regression Analyses

Treatment differences in level of dependence derived from the ADL scale will be assessed using logistic ordinal regression analysis. The logistic ordinal regression model will include independent variables for treatment and concomitant AChEI and/or memantine use at baseline (yes/no). The null hypothesis is that the contrast of LY3303560 groups versus placebo equals zero.

6.17.2. *Exploratory Bioanalytical and PK/PD Analyses*

6.17.2.1. Analysis of Neurofilament Light Chain (NfL)

To evaluate the change from baseline in Neurofilament Light chain (NfL), an MMRM analysis will be used to compare change from baseline at 104 weeks in the Evaluable Efficacy dataset. The model will include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, as well as continuous effects of baseline log(NfL) and age at baseline. The NfL values will be log transformed prior to calculating the changes from baseline. Visit will be considered a categorical variable with values equal to the visit numbers at NfL is assessed. The null hypothesis is that the difference in LSM between the LY3303560 dose groups and placebo equals zero.

Change from baseline and annualized change from baseline analyses will be conducted on NfL. The annualized change will be compared between the treatment groups with an ANCOVA on the full efficacy dataset. The ANCOVA model will include the fixed effect of treatment as well as continuous effects of baseline NfL value and age at baseline. The null hypothesis is that the difference in LSM between the LY3303560 dose groups and placebo equal zero.

To assess the relationship of biomarker with cognition and function with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to Week 104 for the NfL and with change from baseline to Week 104 for iADRS, ADAS-Cog_{13}, ADCS-ADL, MMSE, and CDR-SB. Correlation analyses will be conducted using only patients who have the clinical outcome and SUVr result at Week 104 and include patients from all 3 dose groups.

6.17.2.2. Analysis of Plasma Tau

To evaluate the change in plasma tau analytes (including assayed plasma total tau and p-tau) after treatment, an MMRM will be used to compare change from baseline to 104 weeks. The plasma tau values will be log transformed prior to calculating the changes from baseline. This analysis will be run separately for each plasma tau parameter using the Evaluable Efficacy dataset. The model will include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, as well as the continuous effect of baseline log(plasma tau). Visit will be considered a categorical variable with values equal to the visit numbers at which plasma tau is assessed. The null hypothesis is that the difference in LSM between the LY3303560 dose groups and placebo equal zero. A similar analysis will be performed for completers.

To assess the relationship of plasma tau with cognition and function with treatment, Spearman's rank correlation coefficient will be obtained on change in plasma tau from baseline to Week 104 and with change from baseline to Week 104 for iADRS, ADAS-Cog_{13}, ADCS-ADL, MMSE, and CDR-SB. Correlation analyses will be conducted using only patients who have the clinical outcome and plasma tau result at Week 104.

6.17.2.3. PK/PD Analyses

Compartmental modeling of LY3303560 PK data using nonlinear mixed effects modeling or other appropriate methods may be explored, and population estimates for clearance and central volume of distribution may be reported. Depending on the model selected, other PK parameters

may also be reported. Exploratory graphical analyses of the effect of dose level or demographic factors on PK parameters may be conducted. If appropriate, data from other studies of LY3303560 may be used in this analysis.

The PK/PD relationships between plasma LY3303560 concentration and SUVr, cognitive endpoints, or other markers of PD activity may be explored graphically. The relationship between the presence of antibodies to LY3303560 and PK, PD, safety, and/or efficacy may be assessed graphically. If warranted, additional analysis may be explored to evaluate potential interactions for ADA, PD, and other endpoints (PET scan, safety, etc.).

6.18. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

Summary of adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Adverse Events are summarized: by treatment group, by MedDRA preferred term.

- An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (TEAE).
- An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

7. References

Alaka K, Campbell G, Fleisher A, Hake A, Irizarry M, Miller B, Pontecorvo M, Siemers E, Sims J, Yaari R. The nosology of Alzheimer's disease and application of concepts to amyloid-targeting treatments: an internal summary of Lilly thinking. White Paper. 2015.

Andersen SW and Millen BA. On the practical application of mixed effects models for repeated measures to clinical trial data. *Pharm Stat*. 2013;12(1):7-16.

Bates DM and Chambers JM. "Nonlinear Models," Chapter 10 of Statistical Models in S. J. M. Chambers and T. J. Hastie, Eds., Wadsworth & Brooks/Cole, 1992.

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med*. 2009;28(4):586-604.

Bretz F, Posch M , Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913.

Carpenter JR, Kenward MG. Multiple Imputation and Its Application. New York, NY: John Wiley & Sons Ltd; 2013.

Kahle-Wrobleski K, Andrews JS, Belger M, Gauthier S, Stern Y, Rentz DM, Galasko D. Clinical and economic characteristics of milestones along the continuum of Alzheimer's disease: transforming functional scores into levels of dependence. J Prev Alzheimers Dis. 2015;2(2):115-120.

Kahle-Wrobleski K, Andrews JS, Belger M, Ye W, Gauthier S, Rentz DM, Galasko D. Dependence levels as interim clinical milestones along the continuum of Alzheimer's disease: 18-month results from the GERAS Observational Study. J Prev Alzheimers Dis. 2017;4(2):72-80.

McLaughlin T, Feldman H, Fillit H, Sano M, Schmitt F, Aisen P, Leibman C, Mucha L, Ryan JM, Sullivan SD, Spackman DE, Neumann PJ, Cohen J, Stern Y. Dependence as a unifying construct in defining Alzheimer's disease severity. *Alzheimers Dement.* 2010;6(6):482-493.

Proschan M, Brittain E, Kammerman L. Minimize the Use of Minimization with Unequal Allocation. *Biometrics* 2011; 67: 1135-1141.

Spackman DE, Kadiyala S, Neumann PJ, Veenstra DL, Sullivan SD. The validity of dependence as a health outcome measure in Alzheimer's disease. *Am J Alzheimers DisOther Demen.* 2013;28(3):245-252.

Viele, Kert et al. "Use of Historical Control Data for Assessing Treatment Effects in Clinical Trials." Pharmaceutical statistics 13.1 (2014): 41–54. PMC. Web. 9 Oct. 2018.

Wessels AM, Siemers ER, Yu P, Andersen SW, Holdridge KC, Sims JR, Sundell K, Stern Y, Rentz DM, Dubois B, Jones RW, Cummings J, Aisen PS. A Combined Measure of Cognition and Function for Clinical Trials: The Integrated Alzheimer's Disease Rating Scale (iADRS). *J Prev Alzheimers Dis*. 2015;2(4):227-241.

8. Appendices

Appendix 1. Potentially Clinically Significant Changes in Vital Signs and Weight

Vital Sign Parameter (Unit)	Postbaseline Low Criteria	Postbaseline High Criteria
Sitting systolic blood pressure (mmHg)	Absolute value ≤90 and ≥20 decrease from baseline	Absolute value ≥160 and ≥20 increase from baseline
Sitting diastolic blood pressure (mmHg)	Absolute value ≤50 and ≥10 decrease from baseline	Absolute value ≥100 and ≥10 increase from baseline
Sitting pulse (bpm)	Absolute value <50 and ≥15 decrease from baseline	Absolute value >100 and ≥15 increase from baseline
Weight	≥7% decrease	≥7% increase
Vital Sign Parameter (Unit)	**Postbaseline Criteria for Abnormality**	
Orthostatic systolic blood pressure (mmHg)	≥20 mmHg decrease in systolic blood pressure (supine to standing) (i.e., supine minus standing ≥20)	
Orthostatic diastolic blood pressure (mmHg)	≥10 mmHg decrease in diastolic blood pressure (supine to standing) (i.e., supine minus standing ≥10 mm Hg)	
Orthostatic pulse (bpm)	≥30 increase in bpm (standing to supine) (i.e., standing minus supine ≥30)	
Temperature	Absolute value ≥38.3°C and ≥1.1°C increase from baseline (Absolute value ≥101°F and ≥2°F increase from baseline)	

Abbreviation: bpm = beats per minute.

Appendix 2. Potentially Clinically Significant Changes in ECGs

ECG Parameter	Low Criteria	High Criteria
Heart Rate	<50 bpm	>100 bpm
PR Interval	<120 msec	≥220 msec
QRS Duration	<60 msec	≥120 msec
QTcF Interval		
Males	<330 msec	≥450 msec
Females	<340 msec	≥470 msec
Males and females		> 500 msec

Abbreviations: bpm = beats per minute; ECG = electrocardiogram; QTcF = Fridericia-corrected QT interval.

Criteria for Prolonged ECG QTcF Interval

QTcF Delta Changes (msec)

>30

>60

>75

Abbreviation: QTcF = Fridericia-corrected QT interval.

Signature Page for VV-CLIN-008357 v1.0

Approval	PPD Statistician 11-Oct-2021 18:39:05 GMT+0000

Signature Page for VV-CLIN-008357 v1.0