Statistical Analysis Plan I8D-MC-AZET (Version 3)

A Randomized, Double-Blind, Placebo-Controlled and Delayed-Start Study of LY3314814 in Mild Alzheimer's Disease Dementia (The DAYBREAK Study)

NCT02783573

Approval Date: 17-Sep-2018

1. Statistical Analysis Plan: I8D-MC-AZET: A Randomized, Double Blind, Placebo Controlled and Delayed-Start Study of LY3314814 in Mild Alzheimer's Disease Dementia (The DAYBREAK Study)

LY3314814 (AZD3293) Mild AD

Study I8D-MC-AZET is a multicenter, randomized, parallel-group, double-blind, placebo-controlled Phase 3 study of 2 fixed dose levels of LY3314814 in subjects with mild AD

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol I8D-MC-AZET
Phase 3

Approval Date: 29-May-2018 GMT

Approval Date: 17-Sep-2018 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

SAP Version 1 was approved prior to unblinding on 31-Aug-2016. SAP Version 2 was approved 29-May-2018 prior to unblinding for a potential interim futility analysis triggered if study I8D-MC-AZES failed an interim analysis. Version 3 of the SAP will be approved prior to final unblinding for the clinical summary report and includes the following changes:

1. Added a statement about the futility of the study to the objectives section.
2. Two of the secondary objectives in the placebo-controled period were removed.
3. In section 4.3 a statement about the futility of the study was added to explain the adjustments to the delayed-start objectives.
4. The study design section was updated to describe discontinuations resulting from the study stopping early.
5. Added a note about the way to interpret statistical results in a cancelled study (see section 6.1).
6. The placebo-controled summary structure was added to section 6.1.1.
7. A revised delayed-start summary plan was added to section 6.1.2 to account for the cancellation of the study.
8. Added new missing data rules for the RBANS measure to generate consistency with the I8D-MC-AZES study.
9. Text referencing multiplicity adjustments has been deleted or edited as appropriate (see section 6.5).
10. Deleted population definitions for populations that will no longer be analysed (see section 6.6).
11. The study populations planned to be used in summaries and analyses have been updated throughout the SAP to account for the early stopping of the study.
12. Deleted table of analyses AZET.SAP.2 and renumbered tables.
13. Delayed-start summary and analysis descriptions have been deleted throughout the SAP where appropriate.
14. Kaplan-Meier summaries of disposition have been deleted (see section 6.7).
15. ADAS-Cog11 was added to the baseline summaries.
16. Planned output for prior medications and anticholinergic medications have been removed (see section 6.10).
17. Deleted the graphical method for hypothesis testing of key secondary objectives and revised the description of secondary analyses.
18. Deleted all slope analyses.
19. Deleted all time-to-progression analyses.
20. Deleted all exploratory analyses.
21. Deleted all sensitivity analyses.
22. Deleted compliance summaries and analyses.
23. Added a note about p-values in the delayed-start safety summaries.
24. Applied some editing in section 16.3.2 to clarify the handling of TEAEs.
25. Changed the reporting of common AEs from 5% to 2%.

26. In AEs of special interest, revised the skin disorders cluster.
27. Subgroup analyses of TEAEs are deleted.
28. Added a summary of PCS laboratory results.
29. The threshold for weight gain or loss was change from 4% to 7%.
30. Added a shift table for white matter changes detected through MRIs.
31. The description of the CSSRS analysis has been completely revised (see section 6.13.8).
32. Added summaries to the skin evaluation analyses.
33. Added detail to the eye examination analysis description.
34. Deleted subgroup analyses of plasma Aβ.
35. Deleted correlation calculations between imaging results and ADAS-Cog_{11} and ADAS-Cog_{12}.
36. Revised the description of amyloid PET analysis.
37. Revised the description of the FDG PET analysis and added the perfusion analysis.
38. Revised the description of TAU PET analysis.
39. Revised the description of vMRI analysis
40. Deleted a selection of subgroup analyses.
41. Deleted appendix of drugs with anticholinergic properties.
42. There will no longer be any per-protocol population analyses so the wording of section 6.16 has been amended to reflect the protocol violations that will be listed for the final report.
43. Added two references.
44. Minor editing for grammar and clarity throughout.

4. Study Objectives

AZET (The DAYBREAK Study) is a randomized, double-blind, placebo-controlled study of LY3314814 in patients with mild Alzheimer's Disease (AD) dementia and abnormal levels of amyloid, consisting of a 78 week Placebo-Controlled period followed by a 78 week Delayed-Start period. Study AZET was deemed to be futile and stopped on June 12, 2018. Because of this futility, the original planned delayed-start analysis of evaluating disease modification of LY3314814 by summarizing both efficacy and safety across both study periods became moot. Analyses of study AZET will be separated by study period: placebo-controlled study period and the delayed start study period.

4.1. Primary Objective – Placebo-Controlled Period

The primary objective of the Placebo-Controlled period of AZET is to test the hypothesis that LY3314814, administered orally at doses of 20 and 50 mg daily for 78 weeks, will slow the decline of AD relative to placebo in patients with mild AD dementia, as measured by change from baseline to the end of the Placebo-Controlled period on the ADAS-Cog13.

4.2. Secondary Objectives – Placebo-Controlled Period

4.2.1. Secondary Clinical Efficacy Objectives

The secondary clinical efficacy objectives of the Placebo-Controlled period of AZET include:

- Evaluation of the efficacy of LY3314814 in patients with mild AD dementia on functional, clinical, and cognitive outcomes (ADCS-iADL, FAQ, iADRS, CDR-SB, CDR-Global, NPI, and MMSE) at the end of the Placebo-Controlled period (Week 78).
- Evaluation of the relationship between the treatment effect of LY3314814 and time on functional, clinical, and cognitive outcomes (ADAS-Cog13, ADCS-iADL, FAQ, CDR-SB, and iADRS) at points other than the end of the Placebo-Controlled period.

4.2.2. Secondary Biomarker Objectives

The secondary biomarker objectives of the Placebo-Controlled period of this study include evaluation of the effect of LY3314814 on:

- CSF Aβ PD markers;
- CSF markers of neurodegeneration;
- Brain amyloid burden;
- Regional cerebral blood flow (rCBF);
- Brain aggregated tau levels;
- Brain metabolism; and
- Brain atrophy.

4.3. Secondary Objectives – Delayed-Start Period

AZET also includes a Delayed-Start period with separate objectives for efficacy.

Study AZET was deemed to be futile and stopped on June 12, 2018. Because of this futility, the original planned delayed-start analysis of evaluating disease modification of LY3314814 by summarizing both efficacy and safety across both study periods became moot.

Secondary analyses of the Delayed-Start period include assessment of the:

- Difference in mean change from baseline on the primary outcome measure at Weeks 117, 130, 143, and 156 of the Delayed-Start period; and
- Difference in mean change from baseline on the MMSE, iADRS, ADCS-iADL, and FAQ at Weeks 104, 117, 130, 143, and 156 of the Delayed-Start period.

4.4. Pharmacokinetic Objective

The pharmacokinetic objective of this study is to assess the population PK of LY3314814 and metabolite AZ13569724 in patients with mild AD dementia.

4.5. Safety Objectives

The safety objectives of AZET span the Placebo-Controlled and Delayed-Start periods, and include evaluation of the safety and tolerability of LY3314814 in patients with mild AD dementia as evaluated by spontaneously reported AEs, clinical laboratory tests, vital signs and body weight measurements, 12-lead ECGs, physical and neurological examinations, eye examinations, skin examinations, serial MRIs, and periodic administrations of the C-SSRS.

5. Study Design

5.1. Summary of Study Design

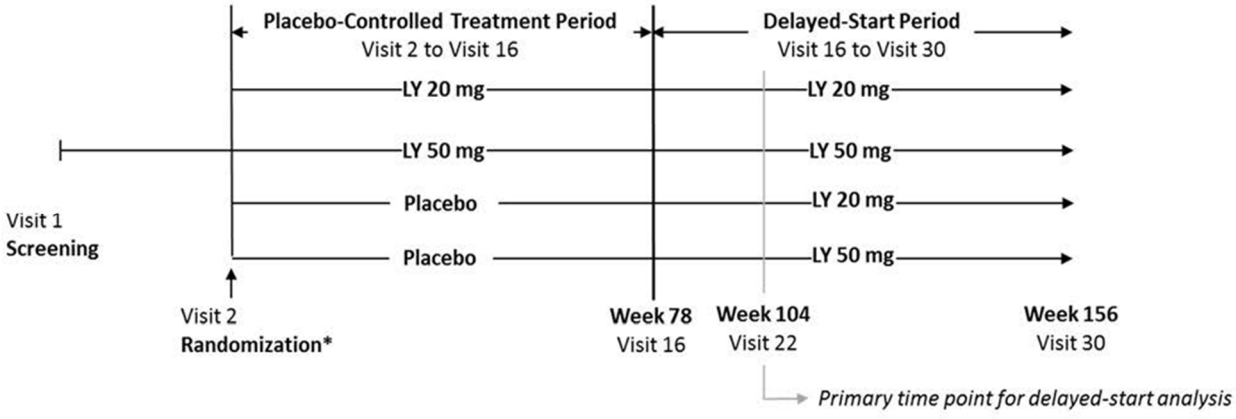
Study I8D-MC-AZET (AZET) is an international, multicenter, randomized, parallel-group, 78 week double-blind (DB), placebo-controlled study of 2 fixed dose levels of LY3314814 (20 mg or 50 mg) in patients with mild AD dementia and abnormal levels of amyloid, followed by a 78 week Delayed-Start period.

Study AZET is the second Phase 3 study that was initiated after a safety interim analysis of an ongoing Phase 2/3 study [I8D-MC-AZES] in an early AD (mild cognitive impairment [MCI] due to AD and mild AD dementia) population. Study AZET will only enroll patients with mild AD. This study is 18 months (78 weeks) in duration for the primary outcome. Studies of this duration, in mild AD dementia patients, are believed to be of sufficient length to detect a difference from the placebo decline rate (Liu-Seifert et al. 2015).

The purpose of the Delayed-Start period of AZET is to assess whether LY3314814 has an effect independent of acute symptomatic effects otherwise consistent with disease modification or slowing of disease progression. In other words, the objective of the Delayed-Start period is to determine if the patients originally randomized to receive placebo and switched to LY3314814 at the start of the Delayed-Start period could achieve the same outcome after receiving 6 months of active therapy (at Week 104) as patients originally randomized to receive LY3314814 in the double-blind period. This objective will be evaluated using all available data through Week 104 by the time of database lock of the Placebo-Controlled period. After the study was stopped for futility, all ongoing patients were asked to come in for a futility-based discontinuation visit. Patients at this visit followed the schedule of events for an early discontinuation visit.

Figure AZET.SAP.1 illustrates the study design.

Figure AZET.SAP.1 Illustration of Study Design for Clinical Protocol I8D-MC-AZET.



Abbreviations: LY=LY3314814

*At Visit 2, patients randomized to placebo in the Placebo-Controlled period will also be randomly assigned to a Delayed-Start treatment group to begin at Week 78. This will be achieved by randomizing patients to one of four

sequences in a randomization ratio of 2:2:1:1 (LY3314814 20 mg: LY3314814 50 mg: Placebo then LY3314814 20 mg; Placebo then LY3314814 50 mg). The primary analysis of the Placebo-Controlled period will occur at Visit 16 (Week 78). The primary analysis of the Delayed-Start period will occur at Visit 22 (Week 104).

Subjects who meet all inclusion criteria and no exclusion criterion will be randomly assigned to 1 of 4 treatment groups in a 2:2:1:1 ratio: 20 mg LY3314814 – 20 mg LY3314814, 50 mg LY3314814 – 50 mg LY3314814, placebo – 20 mg LY3314814, or placebo – 50 mg LY3314814 (the first treatment representing the Placebo-Controlled period, and the second treatment the Delayed-Start period).

Following the Placebo-Controlled period, patients initially randomized to placebo will switch to either the 20 or 50 mg dose of LY3314814. Allocation to placebo or drug during the Placebo-Controlled period, and allocation to dose arm in the Delayed-Start period will be blinded for patient and site for the duration of the study. Patients randomized to either the 20 or 50 mg dose of LY3314814 in the Placebo-Controlled period will be denoted as early-start patients; patients randomized to placebo, delayed start patients.

During the Delayed-Start period, the initial 6 months of LY3314814 therapy for those completing the Week 104 visit will be compared to those subjects who started LY3314814 therapy at randomization (Visit 2) to test whether there is a prespecified treatment effect that cannot be achieved with the later start of treatment. The time points beyond the primary analysis at 6 months of the Delayed-Start period will be analyzed upon last patient visit for Visit 30 to determine the robustness of early treatment over the full 18-month Delayed-Start period.

5.2. Determination of Sample Size

Approximately 1899 patients will be randomized to achieve approximately 1424 patients who will complete the Placebo-Controlled period.

This sample size has been calculated to provide at least 90% power to detect a treatment difference in a key functional secondary outcome, that is, the change in ADCS-iADL at 78 weeks. Based on estimates of treatment change and standard deviation from the mild APOE4 carriers (APOE4 used as a surrogate for amyloid positivity to more closely correspond to the patient population to be enrolled in this study) in the pooled solanezumab studies LZAM/LZAN, an expected mean change (SD) in ADCS-iADL score is 2.07 (9.78) which equates to an effect size of 0.212.

A total of 554 patients completing the study per arm would be required to provide 90% power at $\alpha = 0.025$. Assuming a 25% dropout rate and a 50% information rate for dropouts, 633 (554/(1-0.25%*50%)) patients will need to be randomized per arm (633 in combined placebo arms during the Placebo-Controlled period). This equates to a total randomized sample size of 1899.

This calculated sample size provides greater than 90% power to evaluate the primary outcome for the Placebo-Controlled period. Based on estimates of treatment change and standard deviation from the mild APOE4 carriers in the pooled solanezumab studies LZAM/LZAN, an expected mean change (SD) in ADAS-Cog13 is 2.91 (10.84), corresponding to an effect size of

0.269. A total of 633 patients randomized per arm will yield over 99% power for this analysis. Additionally, 554 patients completing the study per arm would provide over 98% power.

Using the three stage hypothesis testing approach to delayed start analysis, as outlined in Liu-Seifert and colleagues (2015), this sample size will provide approximately 80% power at an alpha level of 0.1 for each dose when all patients have the opportunity to reach the 6 month time point of the Delayed-Start period. At the 18 month time point of the Delayed-Start period, this sample size will provide approximately 50% power. At the time of submission based on current enrollment projections, it is forecasted that 33-40% of patients will not have had the opportunity to reach the 6-month time point. Accounting for this decreased sample size at the time of submission, the power of this Delayed Start analysis is reduced from 80% to approximately 75%.

These power calculations are based on the delayed-start results from the pooled solanezumab studies LZAM/LZAN among APOE4 carriers. Treatment differences at the end of the Placebo-Controlled and Delayed-Start periods and the corresponding variance and covariance estimates were used to calculate the power empirically. For the 6 month time point of the Delayed-Start period, the assumed early discontinuation (EDC) rate used for these calculations was 30%, and 40% for the 18 month Delayed-Start time point. These EDC assumptions were used to adjust the randomized sample sizes using the following formula:

Effective Sample Size = (Randomized Sample Size) * (1 – 0.50*EDC)

The effective sample size assumes that EDC patients will contribute half the information that completing patients contribute. The effective sample sizes for the 6 month power calculation were 531 and 266, early-start arm and delayed-start arm, respectively; for 18 months, 500 and 250, early-start arm and delayed-start arm, respectively.

5.3. Method of Assignment to Treatment

Patients who meet all criteria for enrollment will be randomized at Visit 2. Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IxRS). The IxRS will be used to assign the package containing double-blind investigational product to each patient. Site personnel will confirm that they have located the correct package by entering a confirmation number found on the packages into the IxRS.

To achieve between-group comparability for site, the randomization will be stratified by site.

This is a double-blind study, with a Placebo-Controlled period and a period without placebo (Delayed-Start period). Both patients and site personnel are to remain blinded to dose and prior allocation of placebo throughout the study.

To preserve the blinding of the study, a minimum number of Lilly personnel will see the randomization table and treatment assignments before the study is complete.

The stratified randomization schedule will be prepared by CCI . Randomization codes will be allocated in blocks within strata (investigative sites) to LY3314814 20 mg in both the Placebo-Controlled and Delayed-Start periods, LY3314814 50 mg in both the Placebo-

Controlled and Delayed-Start periods, placebo followed by LY3314814 20 mg, or placebo followed by LY3314814 50 mg in a ratio of 2:2:1:1. If a subject withdraws from the study, the subject's screening and randomization code will not be reused, and the subject will not be allowed to reenter the study. The IxRS will be used to assign a dosing regimen to each patient.

6. A Priori Statistical Methods

6.1. General Considerations

For treatment group comparisons of the Placebo-Controlled period, the two placebo treatment groups (Placebo-20 mg and Placebo-50 mg) will be pooled to form one placebo-treated control group. This group will be labelled as the placebo treatment group (PLA) in presentations summarizing the Placebo-Controlled period.

As the study was stopped for futility, the planned statistical tests lose their scientific validity and multiplicity is no longer a concern. All reported p-values will not be adjusted for multiplicity. No test will be interpreted as statistically significant, and tests resulting with low p-values will only be considered as potential results of interest.

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.

6.1.1. Placebo-Controlled Period

Data summaries will include a total column and columns describing the treatments received only in the Placebo-Controlled period: 20mg, 50mg, or PLA. When change from baseline is assessed, subjects will be included in the analysis only if both a baseline and a post-baseline measurement are available. Unless otherwise defined, a baseline measurement is the last non-missing observation collected during Visits 1 and 2. End point is the last non-missing post-baseline measurement prior to the Delayed-Start period. For analyses using last observation carried forward (LOCF), the last non-missing post-baseline measurement (scheduled or unscheduled) will be used to calculate change from baseline.With the early termination of this study, all results are considered to be suggestive or hypothesis-generating rather than the originally planned confirmatory interpretation of the test or comparison.

6.1.2. Delayed-Start Period

Unless otherwise defined, a baseline measurement is the last non-missing observation collected prior to the Delayed Start period. No between treatment p-values will be reported because (1) all patients received active treatment in the Delayed Start period, and (2) previous treatments in the Placebo-Controlled period were randomized at baseline and do not guarantee patients entering the Delayed Start period are without selection bias. Data summaries will include a total column and columns describing the treatments received across both study periods: 20mg, 50mg, PL/20 and PL/50.

When change from baseline is assessed, subjects will be included in the analysis only if both a baseline and a post-baseline measurement are available. Endpoint is the last non-missing post-baseline measurement. For analyses using last observation carried forward (LOCF), the last non-missing post-baseline measurement (scheduled or unscheduled) will be used to calculate change from baseline.

6.2. Adjustments for Covariates

Repeated measures models will include the fixed, categorical effects of treatment, visit (treated as a categorical variable), treatment-by-visit interaction, pooled site, and concomitant AChEI and/or memantine use at baseline (yes/no), as well as the continuous, fixed covariate of baseline, baseline-by-visit interaction and age at baseline.

When an analysis of covariance (ANCOVA) model is used to analyze a continuous efficacy or safety variable, the model will contain the main effects of treatment, pooled site, concomitant AChEI and/or memantine use at baseline (yes/no), and the appropriate baseline value as a covariate.

6.3. Handling of Dropouts or Missing Data

6.3.1. Handling Missing Data from Subject 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 analyses will only use data from visits where the data was scheduled to be collected (see Andersen and Millen 2013). When subjects discontinue from the study early, there may be efficacy or safety data measurements at visits where the variables were not scheduled to be collected. These data will be used in all other analyses.

6.3.2. Handling Missing Items in Calculating Totals

All total and subscale scores for safety, efficacy, and health outcomes measures will be derived from individual items. If any of the individual items are missing or unknown, every effort will be made to obtain the score for the missing item or items.

For ADAS-Cog13, if <30% (4 or fewer of a total of 13 items) of the items are missing, the total score (maximum = 85) will be imputed as follows: The total from 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 0-5 (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-Cog13 at that visit will be considered missing. The same imputation technique will be applied to any other ADAS-Cog subscore, if tested.

For the FAQ and ADCS-iADL, if 30% or fewer of the items are missing, the total score will be imputed. 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 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.

For the RBANS, if <30% of the sub-items are missing (ie, no more than 3 of the 12 sub-items), the item score will be imputed. For the missing subtest, the scaled score from the other subtest within that index will be used to impute the missing scaled score, which is then converted to a raw score. If List Recognition is missing, the scaled score mean for List Recall, Story Recall, and Figure Recall should be used to impute the missing value. If two sub items are missing within the same index and/or if >30% of the sub-items are missing, the total score for the RBANS at that visit will be considered missing.For all other scales, if any item is missing, any total or sum involving that item will be considered missing.

6.4. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites internationally. A listing will be prepared including country, investigator site with address, number of subjects enrolled (randomized) by each site and unique subject ID.

Sites with fewer than 6 evaluable patients (see Section 6.6.) will be pooled for the statistical analysis. All sites with fewer than 6 evaluable patients on the primary endpoint will be pooled within country and considered as a single site for analysis. If this results in a pooled site still having fewer than 6 evaluable patients, this site will be pooled together with the next smallest site, if one exists, in that country; otherwise, no further pooling is needed. These pooled sites will be used for any analysis that has site as a fixed effect in the model. The actual site numbers will be included in the listings.

6.5. Multiple Comparisons/Multiplicity

As the study was stopped for futility, multiplicity is no longer a concern. All reported p-values will not be adjusted for multiplicity.

6.6. Analysis Populations

For purposes of analysis, populations are defined in Table AZET.SAP.1.

Table AZET.SAP.1. Analysis Populations for Study I8D-MC-AZET

Population	Description
Entered	All participants who sign informed consent
Randomized	All entered patients who are randomized to study treatment.
Evaluable – PC	All randomized patients with a baseline and at least one post-baseline scale result; only placebo-controlled visits included (1 – 16)
Safety – PC	All randomized participants who take at least 1 dose of double-blind study treatment. Participants will be analyzed according to the treatment group to which they were randomized. Only placebo-controlled visits included (1 – 16)
Safety – DS	All randomized participants who take at least 1 dose of double-blind study treatment. Participants will be analyzed according to the treatment group to which they were randomized. All visits included (1 – 30)

Abbreviations: PC = Placebo-Controlled; DS = Delayed-Start

6.7. Subject Disposition

Because this is a long-term study in an elderly patient population with multiple comorbidities, patient discontinuation is of particular concern. Efforts will be undertaken to reduce patient discontinuations and to obtain information on patients initially categorized as lost to follow-up.

From the randomized population, the percentage of patients withdrawing from each treatment group will be summarized. Patients discontinuing treatment due to the sponsor's decision to end the phase 3 program, following the futility analysis, will have "sponsor decision" as reason for discontinuation. From the Safety – PC population, the percentage of subjects discontinuing 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 discontinue and also for each specific reason for discontinuation. From the Safety – DS population, the percentage of subjects discontinuing from each treatment group will be summarized.

6.8. Subject Characteristics

Baseline subject 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 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. Subject characteristics that will be summarized include:

- age (continuous and categorized into 55-64, 65-74, and 75+)
- gender
- race (and ethnicity)

- height
- body weight
- body mass index (weight (kg) / [height (m)]2)
- region
- tobacco use
- alcohol use
- caffeine use
- years of education
- work status
- having 1 or more first degree relatives with AD
- AChEI and/or memantine use at baseline
- time since diagnosis
- amyloid inclusion method (PET, CSF, historical PET)
- Mini Mental State Examination (MMSE) score at Visit 1
- Baseline severity of impairment as measured by ADAS-Cog13, ADAS-Cog11 ADCS-ADL total score and instrumental (ADCS-iADL) and basic subscores (ADCS-bADL), iADRS, CDR Sum of Boxes, RBANS, NPI, RUD-Lite, EQ-5D Proxy, and FAQ.
- APOE4 genotype (ε2/ε4, ε3/ε4, ε4/ε4, no ε4) and carrier status (carrier [ε2/ε4, ε3/ε4, ε4/ε4], noncarrier [ε3/ε3, ε2/ε2,ε3/ε2])

6.9. Treatment Compliance

Treatment compliance will be assessed in the Safety – PC population. Subjects who consume at least 80% of the prescribed daily dose during this study will be considered compliant. The percentage of compliant subjects will be compared across treatment groups using Fisher's Exact test.

6.10. Concomitant Therapy

Concomitant medications will be summarized for the Safety – PC population. 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. Concomitant medications will be listed by subject with both the Placebo-Controlled and Delayed-Start periods combined.

A summary table will also be provided for AChEI/memantine medications. Medications will be coded using the current version of the World Health Organization (WHO) drug dictionary.

6.11. Study Partners

Study partners will be summarized for the Safety – PC population. The protocol states every effort should be made to keep the same study partner throught the duration of this trial. However, changes may be unavoidable. The percentage of subjects with the same study partner will be compared across treatment groups using a Fisher's Exact test. Additionally, study partner changes will be categorized (0 changes, 1 change, and more than 1 change) and compared using Pearson's chi-square test.

6.12. Efficacy Analyses

6.12.1. Placebo-Controlled Treatment Period

6.12.1.1. Primary Efficacy Analysis

The primary objective of this study and of the Placebo-Controlled period is to test the hypothesis that at least one dose of LY3314814 will slow the clinical decline of AD as compared to placebo in patients with mild AD dementia. This objective will be assessed on the Evaluable – PC population using an MMRM analysis of the ADAS-Cog13 in patients with mild AD at baseline, in which the specific hypothesis is that the decline from baseline at the end of the placebo-controlled treatment period (78 weeks) for at least one dose of LY3314814 will be significantly less than that for placebo.

The change from baseline score on the ADAS-Cog13 at each scheduled postbaseline visit (according to the Study Schedule) during the Placebo-Controlled period will be the dependent variable. The model for the fixed effects will include terms for: 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 with values equal to the visit numbers at which the ADAS Cog13 was assessed.

The null hypothesis is that the contrast between both LY3314814 dose groups versus placebo at the last visit equals zero. An unstructured covariance matrix will be used to model the within-subject variance-covariance structure. If the unstructured covariance 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.

6.12.1.2. 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 terms for the 8 independent effects listed previously (Section 6.12.1.1) The null hypothesis is that the differences in least-squares means between the LY3314814 dose groups versus placebo at Week 78 equal zero. The outcomes that will be analyzed are:

- Change from baseline as obtained from the ADAS-Cog_{11}.
- Change from baseline in ADCS-ADL total score, the instrumental ADL subscale, and the basic ADL subscale.
- Change from baseline in CDR Sum of Boxes
- Change from baseline in MMSE
- Change from baseline in FAQ
- Change from baseline in behavioral disturbance as measured by NPI (12 item) total score (frequency multiplied by severity).
- Change from baseline in RBANS Total, Immediate Memory, Delayed Memory, Visuospatial/Construction, Language, and Attention scores.

6.12.2. Delayed-Start Period

6.12.2.1. Primary Analysis

Study AZET was deemed to be futile and stopped on June 12, 2018. Because of this futility, the original analysis to test the delayed-start hypothesis to evaluate disease modification by LY3314814 assessed by ADAS-Cog_{13} was moot. Additionally, very few patients have exposure in the Delayed-Start period. Because of the limited number of patients and exposures in the Delayed-Start period, only safety data will be summarized from this study period. No efficacy or biomarker data will be summarized from this study period. These data will be included in listings covering both the Placebo-Controlled period and the Delayed-Start period.

6.13. Safety Analyses

Safety presentations will be based on the Safety-PC and Safety-DS populations, as relevant. No between treatment p-values will be calculated for analyses of the Safety-DS population.

6.13.1. Extent of Exposure

Days of exposure will be calculated for each subject: (Date of last dose – date of first dose + 1). Summary statistics will be presented for the total number of days of exposure by treatment. Study drug treatment assignment will be listed.

6.13.2. Adverse Events

Adverse events will be summarized and compared across treatment groups in the Placebo-Controlled period and the entire study. Treatment-emergent adverse events (TEAEs) will be defined as those events that first occurred or worsened after the date of randomization (Visit 2).

Should there be insufficient data to make this determination (i.e., missing or incomplete AE start date, stop date, or time), the AE will be considered treatment-emergent. The treatment-emergent period ends on the last day of treatment plus 5 days (these 5 days constitute at least 5 half-lives of lanabecestat). Events occurring within the study but during a period of treatment interruption will only be treated as treatment-emergent if they occur within 5 days of the last dose prior to the treatment interruption.

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 condition
- TEAEs
- TEAEs by maximum severity
- TEAEs occurring in at least 2% of subjects by preferred term
- Serious AEs
- AEs reported as reason for study treatment discontinuation

These summaries will include number and percentages of subjects experiencing the events. Treatment comparisons will be carried out using Fisher's Exact Test. For TEAEs by maximum severity, Severity ="Severe" and "More Severe" will be combined into "Severe" category and the treatment comparison will be performed using Fisher's Exact Test.

All reported preexisting conditions and adverse events (with designations for TEAEs, SAEs, and discontinuations due to AEs) will be listed.

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

Overviews of adverse events will be presented for the Placebo-Controlled period and the Delayed-Start period, respectively, including within each period the number and percentage of subjects who died, experienced serious adverse events, experienced adverse events leading to study discontinuation, and experienced treatment-emergent adverse events. Treatment groups will be compared in the Placebo-Controlled period using Fisher's Exact Test.

Events of the major adverse cardiovascular event type (MACE) that are blindly evaluated by external independent consultants will be summarized and listed.

In addition, the proportion of subjects within specific clusters of TEAEs will be summarized and treatment comparisons will be conducted using Fisher's Exact Test. Clusters will be created from MedDRA High Level Group Terms (HLGTs) and MedDRA SMQ's.

Table AZET.SAP.2. Adverse Events of Special Interest

AE Groups of Interest (Clusters)	MedDRA High Level Group Terms (HLGTs)
Nervous System Disorders	Neuromuscular Disorders HLGT Demyelination SMQ Peripheral Neuropathy SMQ
Eye Disorders	Retinal disorders SMQ
Skin Disorders	Sub-group A: Epidermal and Dermal Conditions HLGT (excluding sub-group B terms) Sub-group B (Hypopigmentation-related events): Hypopigmentation disorders HLT Pigmentation changes, NEC HLT Preferred terms: 'hair depigmented', 'eyelash discolouration', 'iris hypopigmentation', 'eye colour change', 'lip colour altered', 'lip discolouration', 'hair colour changes', 'achromotrichia aquired', 'poliosis'
Liver Disorders	Drug related hepatic disorders - comprehensive search SMQ
Cardiovascular-type events – Arrhythmic	Arrhythmia related investigations, signs and symptoms SMQ Cardiac arrhythmia terms (incl bradyarrhythmias and tachyarrhythmias) SMQ TdP/QT prolongation SMQ
Cardiovascular-type events - Ischemic	Ischaemic heart disease SMQ
Cardiovascular-type events - Stroke	Central nervous system vascular disorders SMQ
Cardiovascular-type events – including orthostatic hypotension	Decreased and Nonspecific Blood Pressure Disorders and Shock HLGT

6.13.4. Clinical Laboratory Evaluations

Laboratory measures will be analyzed both as continuous data (change from baseline) and categorical or ordinal data (proportions of treatment-emergent abnormalities).

If there are multiple records of laboratory measurements at a baseline or post-baseline visit, the latest record will be used.

Summaries and analyses of continuous data (change from baseline) will use both conventional and International System of Units (SI units).

For the Placebo-controlled period, change from baseline to post-baseline visits will be compared between treatment groups using an MMRM model. For each lab analyte, the rank-transformation will be applied to the change from baseline for all subjects and all visits prior to analysis. Similarly, an independent rank transformation will be applied to the baseline values prior to analysis. The model for the fixed effects will include terms for: baseline, treatment, visit, and the treatment-by-visit interaction. This analysis will be done separately for each lab analyte. For the Delayed-Start period, mean change from baseline to post-baseline visits will be summarized.

For the Placebo-Controlled period, treatment differences in the proportions of subjects with treatment-emergent high or treatment emergent low or treatment-emergent abnormal laboratory values at (1) anytime and (2) each post-baseline visit will be assessed using Fisher's Exact Test. Treatment-emergent high or low abnormalities will be calculated using SI units. For each lab analyte, only those subjects who were low or normal at baseline and have at least one post-baseline measurement will be included in the denominator when calculating the proportion of subjects with treatment-emergent high values. Similarly, only subjects who were high or normal at baseline and have at least one post-baseline measurement will be included in the denominator when calculating the proportion of subjects with treatment-emergent low values. In addition, treatment differences in the proportion of subjects who were normal at baseline and change to abnormal high or abnormal low values at any post-baseline visit will be summarized. For the Delayed-Start period, the proportions will be summarized.

A second categorical analysis will be conducted on laboratory analytes. This analysis is considered a PCS analysis and will use limits typically wider than the first categorical analysis. Abnormal criteria for these treatment-emergent PCS changes are presented in **Error! Reference source not found.**.

For urinalysis parameters, baseline to post-baseline shifts will be summarized at each visit. For the Placebo-Controlled period, 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.

For all laboratory analytes, frequencies of subjects with notable changes (that is, increases or decreases of a prespecified amount unique to each analyte) from baseline to each postbaseline visit were also summarized for all subjects and stratified by low, normal, or high at baseline.

The proportion of subjects with treatment-emergent clinically significant changes from a low value or normal value at all baseline at any time in ALT and total bilirubin will be summarized by treatment group. Clinically significant changes of interest are: ALT $\geq$ 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. Comparisons between treatment groups will be carried out using Fisher's Exact test in the Placebo-Controlled period. 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).

If there are multiple records of vital sign or weight measurements at baseline or a 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 (measurement after 5 minutes in the supine position minus that after 2 and 5 minutes in the standing position), temperature, and weight by treatment group for all subjects in the safety population will be summarized.

With the large number of visits at which vital signs are scheduled to be measured, the MMRM model is not suitable for the change from baseline comparison of treatments. In the Placebo-Controlled period change from baseline to each post-baseline visit at which vital signs are taken will be assessed using an ANCOVA model with treatment and as independent factors and baseline value and age as covariates in the model. This analysis will be done separately for each vital sign parameter and weight. In the Delayed-Start period, these changes will be summarized.

The incidence of treatment-emergent abnormal high or low vital signs and weight will be presented by treatment group and visit. Treatment-emergent vital sign evaluations are defined for evaluations collected after the initiation of study medication. Abnormal criteria for postbaseline vital signs and weight are presented in Appendix 5. Any vital sign or weight meeting the criteria will be considered abnormal. In the Placebo-Controlled period, treatment differences in the proportion of subjects 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 and (2) by post-baseline visit.

For each vital sign at each post-baseline visit, only subjects who had a baseline result and had a nonmissing result at that post-baseline visit will be included in the denominator when computing the proportion of subjects with treatment-emergent high, low, or abnormal values.

Summary and analyses of change from baseline in weight will be provided. The proportion of subjects with a weight gain or loss of greater than or equal to 7 percent of baseline body weight

will be compared between treatment groups using Fisher's Exact test at each visit and at any time in the Placebo-Controlled period..

A listing of treatment-emergent abnormal vital signs and weight will also be presented by subject and visit across the Placebo-Controlled and Delayed-Start periods.

6.13.6. Electrocardiograms

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

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. These summaries will include data from each visit ECG measures are performed. In the Placebo-Controlled period, change from baseline to each post-baseline visit at which ECG measurements are taken will be assessed using an MMRM model. The model for the fixed effects will include terms for the following independent effects: baseline ECG score, treatment, visit, treatment-by-visit interaction, and age at baseline. This analysis will be done separately for each ECG parameter. These change from baseline will be summarized in the Delayed-Start period.

In the Placebo-Controlled period, incidence of treatment-emergent abnormal ECGs will be assessed by comparisons at (1) anytime and (2) each post-baseline visit between treatment groups with Fisher's Exact test. For analyses of treatment-emergent abnormal ECGs, baseline will be considered as all visits before the initiation of drug dose. For the Delayed-Start period, these incidences will be summarized.

Abnormal ECG criteria and criteria for abnormal QTcF prolongation are presented in Appendix 6. A second summary of QTc intervals will compare incidence of treatment-emergent increases at any time in four different groups: < 30msec, 30 <= and < 60 msec, > 60 msec, and > 75 msec. This second summary will correct the QT intervals using Fridericia correction (QTcF) and using a large clinical trial population based correction factor (QTcLCTPB).

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 the Placebo-Controlled period, treatment differences in the proportion of subjects who have normal baselines with a change to abnormal high or abnormal low values at any post-baseline visits will be summarized. In the Delayed-Start period, these proportions will be summarized.

6.13.7. Analyses of MRI Data

To evaluate any changes in MRI data following treatment, Pearson's chi-square tests will be used to compare frequencies of responses in the MRI parameters in the Placebo-Controlled period. In the Delayed-Start period, these frequencies will be summarized.

Frequencies and percentages of the following amyloid-related imaging abnormality – edema (ARIA-E, also known as vasogenic edema) and ARIA – hemorrhage (ARIA-H, also known as microhemorrhage) parameters will be summarized:

- ARIA-E:

o Severity (questionable, mild, moderate, severe, or no presence)

o Status compared to the previous MRI(s) (questionable, unchanged, increased, decreased, no longer present)

- ARIA-H:

o Number of ARIA-H (1, 2 to 5, 6 to 10, >10, or no presence),

o Baseline to endpoint changes (increase in size of pre-existing ARIA-H, increase in number of ARIA-H, no change, partial resolution, or complete resolution)

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 by subject and visit across the Placebo-Controlled and Delayed-Start periods.

6.13.8. Additional Safety Concerns

6.13.8.1. 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 in the Placebo-Controlled period.

6.13.8.2. Skin Examination

Skin color will be reported at baseline using Fitzpatrick Scale Rating. The frequencies of the Fitzpatrick Scale Rating (I, II, III, IV, V, and VI) will be displayed by treatment group.

Any hypopigmentation will be assessed by location, percentage of body surface area involvement, degree (partial/decreased pigmentation to complete depigmentation), and other findings in or around the hypopigmentation area (eg, redness or induration). A static physician's

global assessment (sPGA) will be used to determine the subject's overall hypopigmentation severity at a given timepoint using a visual analog scale (VAS) ranging from 0 to 100. In addition, subjects noted to have evidence of hypopigmentation will be asked to record how bothersome they find the hypopigmentation on a VAS ranging from 0 to 100. Additionally, the percentage of patients with emergence of greater than expected hair hypopigmentation will be summarized for patients 'no' at baseline if question is available on worksheet.Frequency tables and summary statistics for the continuous parameters will be presented by treatment group for the Safety-PC and Safety-DS populations.

In order to display any changes/deteriorations during treatment, the following will be reported: number of subjects for whom no hypopigmentation was observed at baseline, but for whom at least once after randomization hypopigmentation was observed during treatment (Placebo-Controlled or Delayed-Start); summary statistics for the difference of maximum value during treatment minus baseline value for percentage BSA of hypopigmentation,; shift table of baseline vs. maximum value during treatment for degree of overall lesion severity. Summary statistics for the change in overall severity (sPGA) and "how bothered is the patient" will be reported for patients with emergence of hypopigmentation, increased severity of hypopigmentation, or increased BSA during the study.

6.13.8.3. Eye Examination

Frequency tables will be produced for all time points for performance of eye examination, visual acuity examination, intraocular pressure examination, and slit lamp exam status and dilated fundus exam status (normal, abnormal – clinically not significant, and abnormal – clinically significant). Clinically significant abnormalities will be displayed together with the corresponding specifications of abnormalities in separate individual data listings including data from the Placebo-Controlled and Delayed-Start periods.

Summary statistics will be produced for the following continuous parameters: left eye total visual acuity score, right eye total visual acuity and both eyes score (scores expressed as logMAR calculated as the negative log (base 10) of the decimal scores) as well as left eye intraocular pressure and right eye intraocular pressure (both in mmHg). For visual acuity, "count fingers" will be given a decimal score of 0.01 and a logMAR of 2 (reference http://www.hicsoap.com/publications/ProperMethodforCalculating.pdf). "Light perception" and "no light perception" cannot be assigned decimal or LogMAR values and so are treated as missing in the mean change summary tables. Visual acuities of patients with these values at any time during the study will be provided in a separate listing.

In order to display any changes/deteriorations during treatment, the following will be reported: number of subjects with any abnormal finding (slit lamp examination or dilated fundus examination) documented during treatment that was not already present at baseline; summary statistics for the difference of maximum value during treatment minus baseline value for left eye total visual acuity score, right eye total visual acuity score, left eye intraocular pressure (mmHg), and right eye intraocular pressure (mmHg); and worst assessment of overall eye examination results during treatment with possible entries "unchanged", "new", "improved", and "worsened".

6.14. Bioanalytical Methods

6.14.1. Analysis of Plasma Aβ

6.14.1.1. Placebo-Controlled Period

To evaluate the change in plasma Aβ analytes (including assayed plasma $A\beta_{1-40}$ and $A\beta_{1-42}$) after treatment, an MMRM will be used to compare change from baseline to 71 weeks. This analysis will be done separately for each plasma Aβ parameter. The model for the fixed effects will include terms for the following independent effects: baseline plasma Aβ, treatment, visit, and treatment-by-visit interaction. Visit will be considered a categorical variable with values equal to the visit numbers at which plasma Aβ is assessed. The null hypothesis is that the difference in LSM between the LY3314814 (previously known as AZD3293) dose groups and placebo equal zero.

6.14.2. Analysis of CSF Data

6.14.2.1. Placebo-Controlled Period

To evaluate the change in CSF biomarkers (including total CSF $A\beta_{1-40}$, total CSF $A\beta_{1-42}$, CSF total tau, and CSF p-tau from lumbar puncture) after treatment, an ANCOVA will be used. The change from baseline score to the LOCF endpoint during the treatment period will be the dependent variable. The model for the fixed effects will include terms for: baseline CSF and treatment. The null hypothesis is that the difference in LSM between the LY3314814 dose groups versus placebo equal zero. Analyses of additional CSF analytes including RBC and WBC will be done in a similar manner.

Annualized change in CSF biomarkers for each patient will be calculated using the change in CSF at the last post-baseline visit. The annualized change will be compared between the treatment groups with an ANCOVA. The ANCOVA model will include the following independent variables: baseline CSF value and treatment. The null hypothesis is that the difference in LSM between the LY3314814 dose groups and placebo equal zero.

6.14.3. Analyses of Amyloid PET Data

6.14.3.1. Placebo-Controlled Period

The PET images acquired in the study will be processed using previously mentioned methods (Clark et al. 2012). At baseline, standard uptake value ratio (SUVr) will be calculated using as a ratio of the composite summary region that is an average of 6 different cortical regions (anterior cingulate, posterior cingulate, medial orbital frontal, lateral temporal, lateral parietal, precuneus) with whole cerebellum as a reference region. However, post-baseline SUVr values will be calculated using 2 different reference regions whole cerebellum and a correction factor using atlas based white matter (AWM). The SUVr with whole cerebellum will be calculated as a ratio of composite summary region to whole cerebellum as a reference region, similar to the calculation at baseline. The SUVr values using AWM correction factors will be calculated by dividing the composite summary ratio by an AWM correction factor. This correction factor is a

ratio of SUV values of AWM to whole cerebellum from baseline to post-baseline. The complete listing of AV45 regions is provided in **Error! Reference source not found.**.

The change from baseline to the post-baseline visit of the composite summary standard uptake value ratio (SUVr) of AV-45 (amyloid imaging) normalized (based on AVID guidelines) will be done using an analysis of covariance (ANCOVA) model with fixed effects of baseline AV-45 result, and treatment. The null hypothesis is that the difference in LSM between the LY3314814 dose groups versus placebo equal zero.

Annualized change in the composite summary SUVr of AV-45 for each patient will be calculated using the change at the last post-baseline visit. The annualized change will be compared between the treatment groups with an ANCOVA. The ANCOVA model will include the following independent variables: baseline AV-45 value and treatment. The null hypothesis is that the difference in LSM between the LY3314814 dose groups and placebo equal zero.

6.14.4. Analyses of FDG PET Data

6.14.4.1. Placebo-Controlled Period

The primary analysis of the fluorodeoxyglucose positron emission tomography (FDG PET) scans will follow the established methods of Landau et al (2011). Composite FDG SUVr will be calculated using the following regions: the left and right parietal (angular gyrus), left and right posterior cingulate, and the left and right temporal lobes (Landau et al, 2011). Two composite summary standard uptake value ratios (SUVr) of FDG PET normalized to the pons + vermis will be assessed: (1) Composite Meta and (2) Composite Meta Automated Anatomical Labeling atlas (AAL).

The change from baseline to the post-baseline visit of the composite summary standard uptake value ratio (SUVr) of FDG PET (fluorodeoxyglucose positron emission tomography) will be done using an analysis of covariance (ANCOVA) model with fixed effects of baseline biomarker result, treatment, and baseline biomarker result taken prior to study drug initiation (Yes or No). The null hypothesis is that the difference in LSM between the LY3314814 dose groups versus placebo equal zero.

Annualized change in the composite summary SUVr for each patient will be calculated using the change at the last post-baseline visit. The annualized change will be compared between the treatment groups with an ANCOVA. The ANCOVA model will include the following independent variables: baseline biomarker value, and treatment. The null hypothesis is that the difference in LSM between the LY3314814 dose groups and placebo equal zero.

The analyses described in this section will be repeated in application to the florbetapir (AV-45) perfusion analysis.

6.14.5. Analyses of Tau PET Data

6.14.5.1. Placebo-Controlled Period

To evaluate the change from baseline in tau imaging parameters, an MMRM analysis will be used to compare change from baseline in SUVr at 78 weeks in the PC Evaluable population. 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 LY3314814 dose groups and placebo equals zero.

Change from baseline and annualized change from baseline analyses will be conducted on SUVrs computed from the MUBADA region with the bimodal white matter serving as the reference region.The annualized change will be compared between the treatment groups with an ANCOVA in the PC Evaluable population. The ANCOVA model will include the fixed effect of treatment as well as continuous effects of baseline AV-1451 value and age at baseline. The null hypothesis is that the difference in LSM between the LY3314814 dose groups and placebo equal zero.

6.14.6. Analyses of vMRI Data

6.14.6.1. Placebo-Controlled Period

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

- Hippocampal volume (mm^3)
- Entorhinal cortex (mm^3)
- Inferior parietal lobe (mm^3)
- Isthmus cingulate (mm^3)
- Lateral parietal lobe (mm^3)
- Medial temporal lobe (mm^3)
- Precuneus (mm^3)
- Prefrontal lobe (mm^3)
- Superior temporal lobe (mm^3)
- Atrophy of Total whole brain volume (cm^3)
- Enlargement of Ventricular volume (cm^3)

All of the above volumes are corrected for intracranial volume. To evaluate the changes in vMRI data after treatment, an ANCOVA model will be used to compare change from baseline to 78 weeks in the PC Evaluable population. The change from baseline to the endpoint visit will be the dependent variable. The model for the fixed effects will include terms for the following independent effects: baseline vMRI value and treatment. The null hypothesis is that the difference in LSM between the LY3314814 dose groups and placebo equal zero.

Annualized change in vMRI for each patient will be calculated using the change in vMRI at the last post-baseline visit. The annualized change will be compared between the treatment groups with an ANCOVA model. The ANCOVA model will include the following independent variables: baseline vMRI value and treatment. The null hypothesis is that the difference in LSM between the LY3314814 dose groups and placebo equal zero.

6.15. Subgroup Analyses

To assess the effects of various demographic and baseline characteristics, subgroup analyses will be performed for the primary endpoint for the Placebo-Controlled period based on the following variables:

- Age - ≥ 55 and < 72 or ≥ 72 and < 85
- APOE4 carrier status (carrier [ε2/ε4, ε3/ε4, ε4/ε4]/non-carrier)
- Region – US or OUS

The primary outcome measure for the Placebo-Controlled period will be modeled using a mixed model repeated measures (MMRM) approach. This general model will include terms for baseline, treatment, visit, concomitant AChEI/memantine use at baseline (yes/no), age (except when age is the subgroup being assessed), treatment by visit, subgroup by treatment, subgroup by visit, and treatment by visit and subgroup. Redundant terms will be dropped from the model in cases where the subgroup of interest overlaps with this general model.

All subgroup analyses presented will be exploratory.

6.16. Protocol Violations

Listings of subjects with significant protocol violations will be provided for the Randomized Population.

The following list of significant protocol violations will be programmed against the clinical database:

- Informed consent violation detected as a missing date of informed consent
- Did not have an assessment of the primary endpoint at any of the visits at which the scale was scheduled to be assessed
- Not compliant with study drug calculated as taking less than 80% or greater than 120% of study drug while the subject was expected to be on treatment

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

- Protocol violations of inclusion/exclusion criteria

- 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.)
- Unqualified raters for the primary endpoint

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.

6.17. Interim Analyses and Data Monitoring

An independent data monitoring committee (IDMC) will have the responsibility to review accumulating unblinded study data on a periodic basis and make recommendations to protect the safety of study participants. Each member of the IDMC is a recognized expert in the fields of AD, neurology, cardiology, or biostatistics. All members will be external to the Sponsor. The approved IDMC charter enumerates the roles of the IDMC members, the frequency with which it meets, and the structure of their meetings. A statistical analysis plan (SAP) for analyses associated with the IDMC outlines the specific analyses that the IDMC will review.

An interim futility analysis may be conducted by the IDMC, for example, if study I8D-MC-AZES (AZES) is stopped for futility at an interim analysis. The purpose of this potential interim analysis for AZET is to potentially stop the study for futility.

Only the IDMC is authorized to evaluate unblinded interim efficacy and safety analyses. Study sites will receive information about interim results only if they need to know for the safety of their patients.

Study AZET will not be stopped early for efficacy.

6.17.1. Potential Futility Interim Analysis if Study AZES Proves Futile

A futility interim analysis may be performed if study I8D-MC-AZES fails its futility interim analysis.

6.17.1.1. Data Included in Futility Analysis

The Bayesian predictive probability of statistical significance for the 50 mg dose vs. placebo and the 20 mg dose vs. placebo at Week 78 will be calculated for the ADAS-Cog13 and iADL endpoints. The Bayesian predictive probability will be calculated from a Bayesian joint model of ADAS-Cog_{13} and ADCS-iADL scores using diffuse priors.

For the patients included in the interim analysis data set, all available scores from ADAS-Cog13 and ADCS-iADL will be utilized. This will potentially include data from an individual patient at weeks 0, 13, 26, 39, 52, 65, and 78 for ADAS-Cog13, and weeks 0, 26, 52, and 78 for ADCS-iADL. A response vector is formulated for the observed scores from all patients in the interim analysis set, including the week 0 baseline assessment, for both the ADAS-Cog13 and the ADCS-iADL. If a patient has a missing value at a visit, or has not progressed in the trial to a visit, a missing value indicator (an NA) is included in the response vector for all missing visits.

The purpose of including the 'NA' values at missing visits is that it enables the prediction of the missing values which are provided in a very convenient fashion in the Bayesian framework.

6.17.1.2. Joint Model

The primary purpose of the joint model is to obtain a distribution of predicted values of ADAS-Cog13 and iADL at week 78 for patients that do not have observed values at week 78. The key features of the model are that it assumes linearity for both ADAS-Cog13 and ADCS-iADL, and that it includes subject specific intercepts and slopes for both ADAS-Cog13 and ADCS-iADL. The linearity assumption is reasonable based on current understanding of the progression of Alzheimer's Disease and is necessary due to the fact that fewer patients will have data available at week 78 at the timing of this futility analysis. The predicted values at week 78 are obtained assuming the same trend observed at this futility analysis will continue for all data collected through the end of the trial. The model is a 'joint' model since the overall error terms for ADAS-Cog13 and ADCS-iADL are allowed to be correlated in a 2 by 2 variance-covariance matrix, and the subject-specific random slopes and intercepts for ADAS-Cog13 and ADCS-iADL are allowed to be correlated in a 4 by 4 variance-covariance matrix. No structure is imposed on either of the variance-covariance matrices. The OpenBugs code that will be used to fit the joint model is provided below.

```
## OpenBugs code for the Bayesian joint model
{
    for (i in 1:Ntotal) {
        Y[i, 1:2] ~ dmnorm(mu[i, 1:2], omega[, ])
        mu[i, 1] <- b[subject[i], 1] + (b[subject[i], 2] + Beta5 *
            LY1[i] + Beta6 * LY2[i]) * time[i]
        mu[i, 2] <- b[subject[i], 3] + (b[subject[i], 4] + Beta7 *
            LY1[i] + Beta8 * LY2[i]) * time[i]
    }
    for (j in 1:Nsubj) {
        b[j, 1:4] ~ dmnorm(Beta.re[j, ], omega.re[, ])
        Beta.re[j, 1] <- Beta1
        Beta.re[j, 2] <- Beta2
        Beta.re[j, 3] <- Beta3
        Beta.re[j, 4] <- Beta4
    }
    Beta1 ~ dnorm(30, 0.005)
    Beta2 ~ dnorm(6.9, 0.005)
    Beta3 ~ dnorm(44, 0.005)
    Beta4 ~ dnorm(-6.9, 0.005)
    Beta5 ~ dnorm(0, 0.005)
    Beta6 ~ dnorm(0, 0.005)
    Beta7 ~ dnorm(0, 0.005)
    Beta8 ~ dnorm(0, 0.005)
    omega[1:2, 1:2] ~ dwish(R[, ], 2)
    sig[1:2, 1:2] <- inverse(omega[, ])
    R[1, 1] <- 1
    R[1, 2] <- 0
    R[2, 1] <- 0
    R[2, 2] <- 1
```

```
    omega.re[1:4, 1:4] ~ dwish(R.re[, ], 4)
    sig.re[1:4, 1:4] <- inverse(omega.re[, ])
    R.re[1, 1] <- 10
    R.re[1, 2] <- 0
    R.re[1, 3] <- 0
    R.re[1, 4] <- 0
    R.re[2, 1] <- 0
    R.re[2, 2] <- 10
    R.re[2, 3] <- 0
    R.re[2, 4] <- 0
    R.re[3, 1] <- 0
    R.re[3, 2] <- 0
    R.re[3, 3] <- 10
    R.re[3, 4] <- 0
    R.re[4, 1] <- 0
    R.re[4, 2] <- 0
    R.re[4, 3] <- 0
    R.re[4, 4] <- 10
}
```

The model includes generally diffuse priors on all parameters. For the parameters that estimate the mean baseline value for ADAS-Cog13 and ADCS-iADL (Beta1 and Beta3, respectively), the location parameter is defined to what is anticipated as the average baseline value for each scale. For the parameters that estimate the rate of decline for placebo at the end of the trial for ADAS-Cog13 and ADCS-iADL (Beta2 and Beta4, respectively), the location parameter is defined to what is anticipated as the overall rate of decline of placebo at week 78. 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 calculated predicted probability. The variable Ntotal has a value of 1 to 13,293 (1899 potential total patients x 7 visits). The variable Nsubj has a value of 1 to 1899. The variables LY1 and LY2 represent indicator variables for patients on the 20 mg and 50 mg doses, respectively. The variable time includes values of 0, 0.167, 0.33, 0.5, 0.667, 0.833, and 1, corresponding to the data collected at weeks 0, 13, 26, 39, 52, 65, and 78, respectively.

6.17.1.3. Calculating the Predicted Probability of ADAS-Cog13 and iADL

The joint model will be fit and a distribution of predicted values at week 78 for ADAS-Cog13 and iADL will be obtained. The single fitting of the joint model provides the predicted values at week 78 needed to create the predicted probability metric. The process used to generate the predicted probability is outlined in the following steps:

1. After a sufficient burn-in period, keep the collection of samples from the first iteration of the MCMC chain and obtain the week 78 predicted scores for patients that don't have an observed value for ADAS-Cog13 or iADL.
2. Merge the week 78 predicted scores with the values from patients that have observed data at week 78 to create a complete dataset of 1899 patients.
3. Calculate the change from baseline score as: week 78 score – Week 0 baseline score.
4. Fit the following ANCOVA model: Change from baseline = baseline + 20 mg Treatment Indicator + 50 mg Treatment Indicator.

5. Store the test statistics corresponding to the contrast of the 50 mg dose vs. placebo and the 20 mg dose vs. placebo.
6. Repeat steps 1-5 50,000 times to generate 50,000 test statistics for the 50 mg vs. placebo contrast and the 20 mg vs. placebo (note that updated predicted values—where baseline or week 78 values were not observed—will be utilized for each iteration of the MCMC chain).
7. There are 4 predicted probabilities calculated, one for each scale at each doses. The predicted probability is calculated as the proportion of the 50,000 times that the change score on treatment is statistically significantly less than the change score of the placebo dose (at a 1-sided 0.0125 alpha level).

6.17.1.4. Futility Threshold

Each of the four predictive probabilities will be compared to 0.465. In order to pass the futility threshold, the predictive probabilities for ADAS-Cog13 and iADL must be greater than 0.465 for at least one dose. A futility threshold of 0.465 has the following operating characteristics:

- If the assumed effect size is 0, 90.1% of the trials that would fail if the study were run to completion would be stopped by this futility threshold;
- If the assumed effect size is 0.11, 21.4% of trials that would succeed if the trial were run to completion would be stopped by this futility threshold.

6.18. Clinical Trial Registry Analyses

Analyses provided for the CTR requirements will be a 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 subjects/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).

7. References

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

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

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. Biometrical Journal. 2011;53(6):894-913.

Clark CM, Pontecorvo MJ, Beach TG, Bedell BJ, Coleman RE, Doraiswamy PM, Fleisher AS, Reiman EM, Sabbagh MN, Sadowsky CH, Schneider JA, Arora A, Carpenter AP, Flitter ML, Joshi AD, Krautkramer MJ, Lu M, Mintun MA, Skovronsky DM; AV-45-A16 Study Group. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study [published correction appears in *Lancet Neurol*. 2012;11(8):658]. *Lancet Neurol*. 2012;11(8):669-678.

Landau SM, Harvey D, Madison CM, Koeppe RA, Reiman EM, Foster NL, Weiner MW, Jaqust WJ; Alzheimer's Disease Neuroimaging Initiative. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiol Aging*. 2011;32(7):1207-1218.

Liu-Seifert H, Siemers E, Holdridge K, Andersen S, Lipkovich I, Carlson C, Sethuraman G, Hoog S, Hayduk R, Doody R, Aisen P. Delayed-start analysis: mild Alzheimer patients in solanezumab trials, 3.5 years. Alzheimers Dement. 2015a;1(2):111-121.

Liu-Seifert H, Andersen SW, Lipkovich I, Holdridge KC, Siemers E. A novel approach to delayed-start analyses for demonstrating disease-modifying effects in Alzheimer's disease. PLoS ONE. 2015b;10(3):Article No: e0119632.

Mohs RC, Knopmann D, Petersen RC, Ferris SH, Ernesto C, Grundman M, Sano M, Bieliauskas L, Geldmacher D, Clark C, Thal LJ. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord. 1997;11(Suppl 2):S13-S21.

8. Appendices

Appendix 1. Complete list of Florbetapir Parameters

SUVr will be obtained for the regions listed below normalized (based on AVID guidelines) and to cerebellar gray matter:

composite summary	lateral temporal cortex left[a]	pons
caudate left	lateral temporal cortex right[a]	putamen right
caudate right	mean cerebellum gray matter	putamen left
cerebellar cortex left	mean whole cerebellum	rectus left
cerebellar cortex right	mesial temporal cortex left	rectus right
cerebellar white matter	mesial temporal cortex right	subcortical white matter
cingulum anterior left[a]	occipital cortex left	temporal cortex left
cingulum anterior right[a]	occipital cortex right	temporal cortex right
cingulum posterior left[a]	orbitofrontal cortex left	thalamus left
cingulum posterior right[a]	orbitofrontal cortex right	thalamus right
frontal cortex left[a]	parietal cortex left[a]	
frontal cortex right[a]	parietal cortex right[a]	

[a]Regions used in calculation of the composite summary SUVr.

Appendix 2. Complete list of FDG PET Parameters

SUVr will be obtained for the regions listed below normalized to whole cerebellum and to cerebellar gray matter:

composite summary	lateral temporal cortex left[a]	pons
caudate left	lateral temporal cortex right[a]	putamen right
caudate right	mean cerebellum gray matter	putamen left
cerebellar cortex left	mean whole cerebellum	rectus left
cerebellar cortex right	mesial temporal cortex left	rectus right
cerebellar white matter	mesial temporal cortex right	subcortical white matter
cingulum anterior left[a]	occipital cortex left	temporal cortex left
cingulum anterior right[a]	occipital cortex right	temporal cortex right
cingulum posterior left[a]	orbitofrontal cortex left	thalamus left
cingulum posterior right[a]	orbitofrontal cortex right	thalamus right
frontal cortex left[a]	parietal cortex left[a]	
frontal cortex right[a]	parietal cortex right[a]	

[a]Regions used in calculation of the composite summary SUVr.

Appendix 3. Potentially Clinically Significant Laboratory Values

Parameter	SI Unit	Low PCS Criteria	High PCS Criteria
Hematology (whole blood)			
Hemoglobin (male)	mml/L-Fe	<6.8266	>11.1708
Hemoglobin (female)	mml/L-Fe	<6.206	>10.5502
Hematocrit	Proportion of 1.0	<0.3	>0.50 (F); >0.55 (M)
Leukocyte (WBC Count)	10^9/L	≤2.8	≥15
Neutrophils	10^9/L	≤1.5	NA
Platelet Count	10^9/L	≤75	≥700
Chemistry (serum or plasma)			
ALT (SGPT)	U/L	NA	≥3 X ULN
AST (SGOT)	U/L	NA	≥3 X ULN
Total Bilirubin	umol/L	NA	≥1.5 ULN
BUN	mmol/L	NA	≥1.2 ULN
Creatinine Kinase (CK)	U/L	NA	≥3ULN
Sodium	mmol/L	≤125	≥155
Potassium	mmol/L	≤3.0	≥5.5
Calcium	mmol/L	≤0.7 ULN	≥1.2 ULN
Alkaline Phosphatase	U/L	NA	≥3ULN
Albumin	g/L	≤26	≥60
Chloride	mmol/L	≤85	≥120
Glucose (random)	mmol/L	≤0.3 ULN	≥1.5 ULN
Serum Creatinine	umol/L	NA	>1.5 ULN
TSH	mIU/L	below normal range	above normal range
Urinalysis			
Hb/RBCs/Blood		NA	≥ + 2
Protein/Albumin		NA	≥ + 2
Glucose		NA	≥ + 2

Abbreviations: ALT/SGPT = alanine aminotransferase/serum glutamic pyruvic; AST/SGOT = aspartate aminotransferase/serum glutamic oxaloacetic transaminase; BUN = blood urea nitrogen; Hb = heart beat, PCS = potentially clinically significant; RBC = red blood cells; TSH = thyroid stimulating hormone; ULN = upper limit of normal; WBC = white blood cells.

Appendix 4. Potentially Clinically Significant 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 decrease in systolic blood pressure (supine to standing) (ie, supine minus standing ≥20)	
Orthostatic Diastolic Blood Pressure (mmHg)	≥10 decrease in diastolic blood pressure (supine to standing) (ie, supine minus standing ≥10)	
Orthostatic Pulse (bpm)	≤ -30 decrease (supine to standing) (ie, supine minus standing ≤ -30)	
Temperature	Absolute value ≥38.3°C and ≥1.1°C increase from baseline (Absolute value ≥101°F and ≥2°F increase from baseline)	

Abbreviations: bpm=beats per minute, NA=not applicable.

Appendix 5. Potentially Clinically Significant ECGs

Parameter	Unit	Low PCS Criteria	High PCS Criteria
QRS Interval	msec	NA	≥120
PR Interval	msec	<100	≥220
Heart Rate	bpm	<45	≥120
QTcF	msec	<320	>500
QTcF interval: change from baseline	>60 msec at any time after randomization		

Abbreviation: PCS = potentially clinically significant.