STATISTICAL ANALYSIS PLAN

Title: Randomized, double-blind, placebo-controlled, multi-center registration trial to evaluate the efficacy and safety of TTP488 in patients with mild Alzheimer's disease receiving acetylcholinesterase inhibitors and/or memantine

Protocol: TTP488-301, Amendment 7, 28 November 2017

Study Drug: Azeliragon (TTP488)

Sponsor: vTv Therapeutics LLC (formerly TransTech Pharma, LLC)

Date: 23 March 2018

Signatures:

Imogene Dunn 23 March 2018

Imogene Dunn, Ph.D. Date
SVP, Biometrics and Regulatory Affairs
vTv Therapeutics LLC

23 MARCH 2018

Larry Altstiel, MD Date
EVP Clinical Development
vTv Therapeutics LLC

Table of Contents

1 Introduction 1
1.1 Study Objectives 2
1.2 Study Rationale 3
1.3 Subject Population 3
1.4 Randomization 3
1.5 Study Design 4
1.6 Sample Size and Power Considerations 4
1.7 Early Stopping, Data Monitoring, and Interim Analysis 5
1.8 Conformance to Regulatory Standards 6
1.9 Modifications from the Statistical Section of the Protocol 6
1.9.1 Modification: Elimination of Bazett's correction for QT-interval. 7
2 Statistical Hypotheses 8
3 Populations of Analysis 11
4 Variables of Analysis 12
4.1 Efficacy Variables 12
4.1.1 Primary Efficacy Variable 12
4.1.2 Key Secondary Variables of Analysis (Imaging) 12
4.1.3 Secondary Variables 12
4.1.4 Other Efficacy Variables of Analysis 13
4.1.5 Efficacy Scales 14
4.2 Safety Variables 17
4.2.1 Treatment-emergent Adverse Events 17
4.2.2 The Columbia Suicide Severity Rating Scale (C-SSRS) 18
4.2.3 Vital Signs 19
4.2.4 Clinical laboratory assessments 19
4.2.5 Brain MRI: Safety 20
4.3 Pharmacokinetics Variables 21
4.4 Pharmacodynamics Variables 21
4.5 Adverse Events of Special Interest 21
5 Statistical Methodology 22
5.1 Pooling of Investigator Sites 22
5.2 Statistical Methodology for Efficacy Analysis 22
5.2.1 Statistical Methodology for Primary Efficacy Analysis 23
5.2.2 Statistical Methods for Secondary Variables 27
5.2.3 Subgroup Analyses 28
5.2.4 MRI and PET Variables 28
5.3 Disposition, Demography, and Baseline Characteristics 29
5.4 Analysis of Safety Data 29

5.4.1 Adverse Events 29
5.4.2 Vital Signs 30
5.4.3 Routine Clinical Laboratory Measurements 30
5.4.4 Liver Enzyme Function Tests 30
5.4.5 Electrocardiography 30
5.4.6 Methodology for Brain MRI (Safety) Analysis 31
5.5 Pharmacokinetics Analysis 31
5.6 Pharmacodynamics Analysis 32
5.7 Exploratory Analysis 32
5.8 Concomitant Medications 32
6 Data Conventions 33
6.1 Definition of Baseline 33
6.2 Missing Data 33
6.3 Dropouts 34
6.4 Visit Windows 34
6.5 Early Termination Assessments and Follow-up Assessments 34
6.6 Unscheduled Assessments 34
6.7 Values Below LLQ 35
7 Analysis Deviations 36
8 Software 37
9 Data Displays 38
10 Values of Potential Clinical Concern 40
11 References 41
12 Schedule of Time and Events 43

1 Introduction

Alzheimer's disease (AD) is a neurodegenerative disease and the leading cause of dementia in the aging population. Neuropathological changes in AD consist of the formation and deposition of amyloid plaques and neurofibrillary tangles. The principal components of amyloid plaques are aggregated and insoluble forms of amyloid beta (Aβ) peptides ending predominantly at amino acid residues 40 and 42. Furthermore, surrounding the amyloid plaque are astrocytes expressing a calgranulin protein, S100b, which is a cytokine associated with chemoattractant activity for monocytes and activation of inflammatory cells of the myeloid lineage. The reactive microglia that surrounds plaques increases the expression of pro-inflammatory cytokines and complement receptors.

TTP Translational Technology®, vTv Therapeutics' proprietary drug discovery engine, was employed to develop TTP488, which is an orally bioavailable antagonist of the Receptor for Advanced Glycation Endproducts (RAGE). This product is being developed as a potential treatment for AD.

This study is planned using an enrichment strategy, following the draft regulatory guidance depicted in the FDA guidance *Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products* (draft guidance, December, 2012). This study features a conditional sequential approach to hypothesis testing to control alpha, using Bauer closed procedures, which are referenced in the FDA adaptive design guidance: FDA guidance *Adaptive Design for Clinical Trials for Drugs and Biologics* (draft guidance, February 2010) referenced above), and principles for subgroup analyses. The statistical analysis is planned to conform to FDA guidance provided in Draft Guidance *Alzheimer's Disease: Developing Drugs for the Treatment of Early Stage Disease* (CDER, 2013), review templates provided by the FDA, and *Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* (CDER&CBER, 1998).

The goals of this statistical analysis plan (SAP) are the following:

- To describe the approach of statistical analysis to support scientifically sound study conclusions.
- To conform to regulatory guidance to facilitate the use of the statistical evaluation for support in the regulatory review of study results for purposes of product registration.
- To guide the analysis and reporting of study results to enable the production of valid and accurate depictions of study data.

The approach will conform to FDA guidance as described in Guidance for Industry Statistical Principles for Clinical Trials (1998).

This final draft SAP reflects finalization of all attributes of the analysis plan that pertain to or affect alpha and key estimation activities, including initial feedback from the FDA. In accordance with ICH E9, the SAP was finalized prior to any unblinding, and following the final blind data review meeting (BDRM) when the per-protocol set (PPS) was characterized. The

SAP will be updated, if warranted, based on future feedback from FDA reviewers; any changes to the SAP after finalization will be tracked in amendments similar to the process for the study protocol.

Please note that this SAP describes the statistical analyses of two studies with a common infrastructure (A-study and B-study). The statistical analysis will be performed independently as two separate studies (the A-study and the B-study) for the co-primary variables: change from baseline in ADAS-cog and change from baseline in CDR-sb. For the key secondary variable (and for all other secondary variables), all study subjects in both the A- and B- studies (A-study and B-study subjects combined) will be used for inferential purposes. All clinical efficacy variables will be summarized statistically for descriptive purposes for A-study and B-study separately (as well as combined). The randomization process supports the independent analysis of the A-study and the B-study, and also supports the statistical analysis of patients from both studies (combined) for the key secondary and other secondary variables. Safety analyses will also be performed on all patients in both studies.

1.1 Study Objectives

The objectives of the primary, key secondary, and secondary endpoints are presented below without a hierarchical order of evaluation. The hierarchical order of the evaluation of the endpoints is provided in Section 4.1 (efficacy variables). The primary objective presented below will be evaluated separately for the two studies (A-study and B-study). The key secondary and all other secondary objectives will be evaluated for the combined data from the A-study and B-study.

Primary objective:

- To evaluate the efficacy of azeliragon on cognitive [Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog)] and global function [Clinical Dementia Rating Scale Sum of Boxes (CDR-sb)] measures in patients with mild AD.

Key secondary objective:

- To examine the effect of azeliragon on MRI volumetrics (e.g., whole brain volume, ventricular volume, hippocampal volume)

Secondary objectives:

- To evaluate the safety and tolerability of 18 months of treatment with azeliragon plus standard of care (SoC) relative to placebo plus standard of care.
- To evaluate the time course of the effect of azeliragon on the cognitive (ADAS-cog) and global functional outcome (CDR-sb) measures.
- To evaluate the efficacy of azeliragon on behavioral symptoms as assessed by the Neuropsychiatric Inventory (NPI).
- To evaluate the efficacy of azeliragon on cognition as assessed by the Mini-Mental State Examination (MMSE).

- To evaluate the efficacy of azeliragon on the Alzheimer's Disease Cooperative Study - Activities of Daily Living Scale (ADCS-ADL).
- To evaluate the efficacy of azeliragon on cognition as assessed by the Continuous Oral Word Association Task (COWAT), Category Fluency Test (CFT) and Trail Making Test (Versions A and B).
- To evaluate the effect of azeliragon on health status, health care resource utilization, and health-related quality of life as assessed by Resource Utilization in Dementia (RUD Lite), and Dementia Quality of Life (DEMQOL-proxy)
- To evaluate the effect of azeliragon on regional and global brain glucose metabolism using FDG-PET.
- To evaluate the effect of azeliragon on plasma Aβ biomarkers.
- To evaluate azeliragon exposure-response relationships for selected efficacy, safety and biomarker endpoints (e.g., ADAS-cog, ADCS-ADL, adverse events, vital signs, plasma concentrations of Aβ).

1.2 Study Rationale

The primary purpose of this trial is to evaluate the efficacy and safety of azeliragon in patients with mild AD. Prior studies (TTP 488-203) have shown promise for azeliragon in this population.

1.3 Subject Population

The study population includes mild AD patients (entry MMSE between 21 and 26, inclusive; CDR-global score of 0.5 or 1) who have a reliable caregiver/informant with regular contact (i.e., 10 hours a week as combination of face-to-face visits and telephone contact acceptable) who will facilitate the subject's full participation in the study. Caregivers/informant must have sufficient subject interaction to be able to provide meaningful input into the rating scales administered in this study where caregiver/informant input is required, in particular the CDR, and evidence of this should be documented in source documentation. Participants who reside in assisted living facilities are permitted, provided they meet caregiver/informant criteria. Subjects are required to be at least 50 years of age.

1.4 Randomization

A total of approximately 800 patients will be randomized into this study. Subjects will be enrolled and randomized according to a fixed randomization scheme blocked by study investigative site. Severity of AD will be determined by MMSE (21-26) at screening. ADAS-cog at Screening will be used as a stratification variable in the randomization scheme:

- Lower ADAS-cog: Screening ADAS-cog 19 or less.
- Higher ADAS-cog: Screening ADAS-cog more than 19.

Randomization will have balanced allocation (1:1) between active plus SoC and placebo plus SoC. Dropouts will not be replaced.

1.5 Study Design

This trial is a multi-center, randomized, double-blind, placebo-controlled, parallel-group study evaluating the efficacy and safety of 18 months of treatment with azeliragon relative to placebo conducted in approximately 800 mild AD patients who are on background standard of care (SoC). Patients will be randomly assigned to 1 dose level (5 mg/day plus SoC) or placebo plus SoC in a 1:1 randomization (approximately n=400/group).

This protocol describes two statistically independent studies with a common infrastructure. Statistical analysis for clinical efficacy will be analyzed as two identical separate studies (A-study and B-study). Analysis for safety and all secondary and exploratory endpoints will be analyzed for both the A and B study as a whole (i.e., the two studies combined).

A-study is conducted at multiple sites in the U.S and Canada.

B-study is conducted at multiple sites in the U.S., Canada, Ireland, UK, South Africa, New Zealand, and Australia.

The overall structure of the study includes the following periods:

Screening period: Screening procedures occur within 60 days of start of study drug administration, and they include obtaining informed consent and evaluations to determine eligibility for study participation.

Baseline Study Period (Day -7 — Day 1, pre-dose): Subjects will have baseline assessments for efficacy and safety measures, including vital signs, safety labs, adverse events, electrocardiography and concomitant medications, as well as imaging and PK/PD assessments.

Treatment Period (Day 1 — Month 18): Subjects are treated with double-blind study medication. Efficacy and safety measures are taken. Blood samples are taken and subjects have assessments for safety and PK.

Follow-up period (Month 21): Subjects return to the site for final safety assessments including vital signs, adverse events, electrocardiography and concomitant medications are taken. Blood draws are taken for safety evaluations and for pharmacokinetics evaluations.

1.6 Sample Size and Power Considerations

The protocol (Section 9.1) includes the following statements about sample size and power considerations:

> The primary objective is to evaluate the effect of azeliragon plus SoC on cognition and functional outcome by comparing azeliragon plus SoC with placebo plus SoC. The co-primary endpoints are change from Baseline on the ADAS-cog total score at Month 18 and change from Baseline on the CDR-sb total score at Month 18. The study includes AD patients with mild disease as defined by having a baseline MMSE of 21 through 26.

This trial includes two independent sub-studies, each of which is supported by randomization independently, because randomization by investigative site is used. Each sub-study is powered using alpha = 0.05.

Approximately 160 subjects (80 per treatment group) will be enrolled and participate in the FDG-PET sub-study. The sample size for this sub-study was selected empirically.

Sample size calculations are done to power each study (A-study and B-study) separately. It is anticipated that each A-study and B-study will randomize approximately N=400.

Assuming a standard deviation of the change from Baseline to Month 18 in ADAS-cog of 9 (based on Study TTP488-203), using alpha = 0.05, a total sample size of 282 patients in balanced allocation (141 patients per group) provides at least 90% power to detect a difference between treatment groups of 3.5 (based on data from mild AD patients in Study TTP488-203) using a 2-sided, 2-sample t-test. Assuming a dropout rate of 20% or less, randomization of 354 patients provides adequate statistical power for this study to demonstrate superiority of azeliragon plus SoC over placebo plus SoC based on the ADAS-cog.

Assuming a standard deviation of the change from Baseline to Month 18 in CDR-sb of 2.4, using alpha = 0.05, a total sample size of 304 patients in balanced allocation (152 patients per group) provides at least 90% power to detect a difference between treatment groups of 0.9 (based on data from mild AD patients in Study TTP488-203) using a 2-sided, 2-sample t-test. Assuming a dropout rate of 20% or less, randomization of 380 patients provides adequate statistical power for this study to demonstrate superiority of azeliragon plus SoC over placebo plus SoC based on CDR-sb.

1.7 Early Stopping, Data Monitoring, and Interim Analysis

There is no interim analysis in this study, except for medical and safety monitoring by the independent data monitoring committee (IDMC). This study has no plan for early stopping.

Except for the ongoing, blinded medical monitoring of accruing data and except for the ongoing review by the IDMC, no interim analysis is planned for this study. This study has no plan for interim analysis.

As a conservative measure, an alpha adjustment is applied to accommodate up to 10 analyses by the IDMC because the safety evaluation will include review of ADAS-cog data from the safety perspective. Alpha = 0.0001 is apportioned to each of no more than 10 analyses by the IDMC that includes review of ADAS-cog data, blinded unless there is a data-driven need expressed by the IDMC to be unblinded.

With the exception of the IDMC, no unblinding is planned in this study for any reason other than emergency unblinding for medical imperatives.

It is emphasized that there is no interim analysis in this study for efficacy. The interim analyses described in this section refer to the analyses done by the IDMC for safety monitoring only.

1.8 Conformance to Regulatory Standards

An objective of this SAP is to comply with regulatory standards. The following guidance documents were specifically consulted in the preparation of this SAP:

- ICH E9: Statistical principles for clinical trials (September 1998); ICH and FDA
- ICH E9(R1): Estimands and sensitivity analysis in clinical trials (draft; 16 June 2017); ICH and FDA.
- ICH E14: Clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs (October 2005); ICH and FDA
- E14 Q&A (R3) ICH guideline E14: the clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs (R3)—questions and answers (26 January 2016); EMA
- Guideline on adjustment for baseline covariates in clinical trials (26 February 2015); EMA.
- Guideline on medicinal products for the treatment of Alzheimer's disease and other dementias (24 July 2008); EMA.
- Good Review Practice: Statistical review template (30 July 2012). Office of Biostatistics; CDER; MAPP 6010.4
- Good Review Practice: Clinical review template (10 December 2010). Office of New Drugs; CDER; MAPP 6010.3
- Guidance: Providing clinical evidence of effectiveness for human drug and biological products (May 1998); CDER; CBER.
- Good Review Practice: Reviewer guidance: Conducting a clinical safety review of a new product application and preparing a report on the review (February 2005); CDER.
- Draft guidance: Developing drugs for the treatment of early stage disease (February 2013); CDER.
- Draft guidance: Adaptive design clinical trials for drugs and biologics (February 2010); CDER, CBER.
- Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design (18 October 2007); EMA
- Draft guidance: Enrichment strategies for clinical trials to support approval of human drugs and biological products (December 2012); CDER, CBER, CDRH.
- Guidance: Drug-induced liver injury: premarketing clinical evaluation (July 2009); CDER, CBER.
- Guidance: Patient-reported outcome measures: Use in medical product development to support labeling claims (December 2009); CDER, CBER, CDRH.
- ICH and MSSO. (2017, September). MedDRA Introductory Guide Version 20.1.

The planned statistical analysis of this study is intended to comply with regulatory expectations and standards as depicted in the above-listed guidance documents.

1.9 Modifications from the Statistical Section of the Protocol

One modification from the Statistical Section of the Protocol is reflected in this SAP, based on agreement with the Neurology Products Division of the FDA:

(1) Elimination of the use of Bazett's correction for QT-interval analysis, based on the recommendation depicted in ICH E14 Q&A R3 (dated 10 December 2015).

1.9.1 Modification: Elimination of Bazett's correction for QT-interval.

The Q&A, Revision 3 (10 December 2015), published for ICH E14 indicates that Bazett's correction is not informative and discourages its use, and this guideline recommends Fridericia's correction.

2 Statistical Hypotheses

As depicted in the study protocol, a 2-stage conditional sequence of statistical hypothesis tests, will be used. The testing sequence will be as follows:

- Stage 1: Four concurrent analyses will be done:
  - A-study analysis of co-primary variables:
    - ADAS-cog is analyzed in A-study.
    - CDR-sb is analyzed in A-study
  - B-study analysis of co-primary variables:
    - ADAS-cog is analyzed in B-study.
    - CDR-sb is analyzed in B-study
- Stage 2: Conditional on statistical significance of both A-study and B-study in both co- primary variables, testing will continue:
  - MRI assessment is analyzed in the study (as a whole) Stage 2:
- Stage 3: Conditional on statistical significance in both stage 1 and stage 2 testing will continue:
- All other secondary endpoints will be evaluated in a hierarchical manner in the combined total of subjects in the 2 studies

For simplicity in notation, thresholds for significance are indicated as "$p<0.05$"; however, "0.05" is representative of the alpha allocated to the analysis. The overall study-wise alpha is 0.05, but particular analyses may have alpha less than 0.05.

STAGE 1:

Hypothesis tests for A-study are as follows:

The first set of hypotheses to be tested is as follows:

- H_{01}: $\mu^{(\text{ADAS-cog, azeliragon, A-study})} = \mu^{(\text{ADAS-cog, P, A-study})}$
- H_{11}: $\mu^{(\text{ADAS-cog, azeliragon, A-study})} \neq \mu^{(\text{ADAS-cog, P, A-study})}$,

where $\mu^{(\text{ADAS-cog, azeliragonTTP488, A-study})}$ and $\mu^{(\text{ADAS-cog, P, A-study})}$ denote the true mean changes from Baseline at Month 18 on the ADAS-cog total score for the A-study azeliragon 5 mg/day added to SoC and placebo added to SoC groups, respectively.

The second set of hypotheses to be tested is as follows:

- H_{02}: $\mu^{(\text{CDR-sb, azeliragon, A-study})} = \mu^{(\text{CDR-sb, P, A-study})}$

- H_{12}: $\mu^{(\text{CDR-sb, azeliragon, A-study})} \neq \mu^{(\text{CDR-sb, P, A-study}).}$

where $\mu^{(\text{CDR-sb, azeliragon, A-study})}$ and $\mu^{(\text{CDR-sb, P, A-study})}$ denote the true mean changes from Baseline at Month 18 on the CDR-sb score for the A-study azeliragon 5 mg/day and placebo groups, respectively.

Hypothesis tests for B-study are as follows:

The first set of hypotheses to be tested in B-study (the third set of hypotheses for the study as a whole) is as follows:

- H_{03}: $\mu^{(\text{ADAS-cog, azeliragon, B-study})} = \mu^{(\text{ADAS-cog, P, B-study})}$
- H_{13}: $\mu^{(\text{ADAS-cog, azeliragon, B-study})} \neq \mu^{(\text{ADAS-cog, P, B-study})}$,

where $\mu^{(\text{ADAS-cog, azeliragon, B-study})}$ and $\mu^{(\text{ADAS-cog, P, B-study})}$ denote the true mean changes from Baseline at Month 18 on the ADAS-cog total score for the B-study azeliragon 5 mg/day added to SoC and placebo added to SoC groups, respectively.

The second set of hypotheses to be tested in B-study (the fourth set of hypotheses for the study as a whole) is as follows:

- H_{04}: $\mu^{(\text{CDR-sb, azeliragon, B-study})} = \mu^{(\text{CDR-sb, P, B-study})}$
- H_{14}: $\mu^{(\text{CDR-sb, azeliragon, B-study})} \neq \mu^{(\text{CDR-sb, P, B-study}).}$

where $\mu^{(\text{CDR-sb, azeliragon, B-study})}$ and $\mu^{(\text{CDR-sb, P, B-study})}$ denote the true mean changes from Baseline at Month 18 on the CDR-sb score for the B-study azeliragon 5 mg/day and placebo groups, respectively.

STAGE 2:

If and only if both hypothesis tests in A-study AD subjects and also B-study AD subjects for both co-primary variables (ADAS-cog and CDR-sb) have $p<0.05$ (i.e., $p<0.05$ for ADAS- cog demonstrating statistical superiority of azeliragon 5 mg added to SoC versus placebo added to SoC and also $p<0.05$ for CDR-sb demonstrating statistical superiority of azeliragon 5 mg added to SoC versus placebo added to SoC (i.e., statistical significance of the test of the null hypothesis for each A-study and B-study independently), testing will proceed to the fifth set of hypotheses to the expansion of the population of subjects from A-study and B-study combined.

Consistent with the FDA draft Guidance for Industry: *Alzheimer's Disease: Developing Drugs for the Treatment of Early Stage Disease* (February, 2012), toward the goal of demonstrating disease modification, analysis of MRI data (hippocampal volume as the key secondary endpoint) will be done conditional on significance of the above-named hypothesis tests:

The next set of hypotheses to be tested is as follows:

- H_{05}: $\mu^{(MRI, azeliragon, All)} = \mu^{(MRI, P, All)}$
- H_{15}: $\mu^{(MRI, azeliragon, All)} \neq \mu^{(MRI, P, All)}$,

where $\mu^{(MRI, azeliragon, Alld)}$ and $\mu^{(MRI, P, All)}$ denote the true mean changes from Baseline at Month 18 on MRI volume measures for all AD azeliragon 5 mg/day and placebo groups, respectively.

STAGE 3:

If all hypotheses defined in both Stage 1 and Stage 2 demonstrate statistical superiority of azeliragon TTP488 5 mg added to SoC versus placebo added to SoC, the secondary efficacy endpoints will be tested in hierarchical order. The order of testing is defined in Section 4.1 (efficacy variables). The outline of the hierarchical testing is:

- First, the secondary endpoint ranked as #1 will be compared between azeliragon TTP488 5 mg added to SoC versus placebo added to SoC in the study as a whole (study A and study B combined). If $p<0.05$, statistical significance will be declared.

- Secondly, if the secondary endpoint ranked as #1 showed a statistically significant difference, the secondary endpoint ranked as #2 will be compared between azeliragon 5 mg added to SoC versus placebo added to SoC in the study as a whole (study A and study B combined). If $p<0.05$, statistical significance will be declared.

The hierarchical testing will continue with the consecutive secondary endpoints, as long as the previous secondary endpoint is declared as statistically significant.

3 Populations of Analysis

In accordance with guidance from the International Conference of Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guideline ICH E9 Statistical Principles for Clinical Trials (1998), the following data sets will be used for all statistical analysis:

- The full analysis set (FAS) includes all randomized subjects who receive any study medication and have at least one post-Baseline.

- The per protocol set (PPS) includes all subjects in the FAS who receive at least 6 months of study drug and attend at least 6 scheduled visits, except for those who are excluded because of major protocol violations, where a major protocol violation is one that may affect the interpretation of study results (e.g., non-compliance with study medication requirements as evaluated using TTP488 plasma concentration below the limit of quantification at last study visit during 18-month treatment period). The intended criterion of "6 scheduled visits" includes screening and baseline and at least 4 scheduled clinic visits after first dose of study medication.

  Final determinations of the PPS will be made at the masked data review meeting (BDRM) held in accordance with ICH E9 prior to the lock data. The complete specification (made blinded to treatment arm) will be documented in the minutes to the blind data review meeting held prior to lock the data.

- The safety set (SAF) includes all subjects who receive any study medication.

For the purpose of understanding the influence of dropouts on study conclusions, the FAS will be partitioned into subjects who complete (completers) and subjects who do not complete (dropouts). For purposes of this study, a subject is a completer if the subject has received treatment for 18 months and has efficacy data for 18 months (evaluation is made at the endpoint level, so if a subject is missing ADAS-cog at Month 18 but has CDR-sb at Month 18, this subject is included in analyses for which the data are available).

It is noted that if the FAS and PPS do not differ by more than 15% in numbers of individuals, analysis may be limited to the FAS (ITT analysis), and the PP analysis will not be conducted.

4 Variables of Analysis

4.1 Efficacy Variables

Efficacy evaluation will include the co-primary endpoints (ADAS-cog and CDR-sb), key secondary measures, and other efficacy markers.

4.1.1 Primary Efficacy Variable

The primary analysis will include assessment of the following variables of analysis.

- Mean change from Baseline to Month 18 in ADAS-cog
- Mean change from Baseline to Month 18 in CDR-sb.

Key supportive variables will include mean change from Baseline to all other Visits in ADAS-cog and mean change from Baseline to all other Visits in CDR-sb. Time to threshold worsening will be done as supportive analyses for change in ADAS-cog with thresholds of 7 points as an indicator of progression, with other cut-points from 1 to 20 as exploratory.

4.1.2 Key Secondary Variables of Analysis (Imaging)

The key secondary endpoint is:

- Change from Baseline in hippocampal volume at Month 18.

PET variables are considered to be exploratory and will be defined separately.

4.1.3 Secondary Variables

The following secondary efficacy endpoints will be evaluated in hierarchical order:

- Secondary endpoint ranked as #1: The change from baseline to final visit in the ADCS-ADL scale through Month 18.
- Secondary endpoint ranked as #2: Change from Baseline in the Neuropsychiatric Inventory (NPI) total score at Month 18.
- Secondary endpoint ranked as #3: The change from baseline to final visit in the Mini-Mental State Examination (MMSE) through Month 18.
- Secondary endpoint ranked as #4: Change from Baseline on the Trail Making Test Version A time at Month 18.
- Secondary endpoint ranked as #5: Change from Baseline on the Trail Making Test Version B time at Month 18.
- Secondary endpoint ranked as #6: Change from Baseline on the Continuous Oral Word Association Task (COWAT) score at Month 18.
- Secondary endpoint ranked as #7: Change from Baseline on the Category Fluency Test (CFT) score at Month 18.

- Secondary endpoint ranked as # 8: Change from Baseline in the Dementia Quality of Life (DEMQOL) -Proxy total score at Month 18.

4.1.4 Other Efficacy Variables of Analysis

Additional efficacy variables of analysis are as follows:

- Responder analysis at Months 3, 6, 9, 12, 15 and 18 based on ADAS-cog (increase from baseline of 7 points or more) and responder analysis at Months 3, 6, 12, and 18 based on the CDR-sb (increase from baseline of 1 point or more). It is noted that a responder is a subject who has not declined, and thresholds are used for classification of having declined.

- Change from Baseline in ventricular volume at Month 18.

- Change from Baseline in whole brain volume at Month 18.

- Time to loss of one global stage on the Clinical Dementia Rating (CDR) score through Month 18.

- Change from Baseline on the ADAS-cog at Months 3, 6, 9, 12 and 15 and at Months 3, 6 and 12 for the CDR-sb and CDR subscale scores.

- Individual item responses on the Resource Utilization in Dementia (RUD) questionnaire and the total caregiver time spent assisting the patient based on the RUD at Months 6, 12, and 18.

- Proportion of participants who enter intermediate or long-term residential care (RUD) at Months 6, 12, and 18.

- Change from Baseline in the Dementia Quality of Life (DEMQOL) - Proxy total score at Months 6 and 12.

- Change from Baseline in the Neuropsychiatric Inventory (NPI) total score and subscale scores at Months 3, 6 and 12.

- Change from Baseline on the Mini-mental State Examination (MMSE) total score at Months 3, 6 and 12.

- Change from Baseline on the Continuous Oral Word Association Task (COWAT) score at Months 3, 6 and 12.

- Change from Baseline on the Category Fluency Test (CFT) score at Months 3, 6 and 12.

- Change from Baseline on the Trail Making Test (Versions A and B) time at Months 3, 6 and 12.

- Change from Baseline in the extent of brain glucose hypometabolism at Month 18.
- Change from Baseline in the severity of brain glucose hypometabolism at Month 18

4.1.5 Efficacy Scales

The following instruments that are used in this study are subject to statistical analysis:

4.1.5.1 Alzheimer's Disease Assessment Scale - Cognitive Subscale 70 point (ADAS-cog):

- Range: The scale range is 0 to 70 with higher scores indicating greater cognitive impairment.
- Brief description: The ADAS-cog is a structured scale (approximately 40 minutes to complete) that evaluates memory, orientation, attention, reasoning, language and constructional praxis (Rosen, 1984). The ADAS-cog scoring range for the version used in this study is from 0 to 70. The ADAS-cog will be conducted at Baseline and at Months 3, 6, 9, 12, 15 and 18 or in the event of early termination. The ADAS-cog should always be administered prior to other cognitive measures.

4.1.5.2 Mini Mental State Examination (MMSE):

- Range: The scale range is 0 to 30 with lower scores indicating greater cognitive impairment.
- Brief description: The MMSE is a brief 30-point test that is used to assess cognition (Folstein, 1975). It is commonly used to screen for dementia. In the time span of about 10 minutes, it samples various functions, including arithmetic, memory and orientation. The MMSE will be administered at Screening, Baseline, and at Months 3, 6, 12 and 18 or in the event of early termination. Participants with scores of 21-26 will be eligible for participation in the study.

4.1.5.3 Controlled Oral Word Association Test (COWAT):

- Range: The score is the number of correct words. Lower numbers are associated with greater impairment.
- Brief description: The COWAT is a measure of verbal fluency in which the participant is asked to generate orally as many words as possible that begin with the letters “F”, “A”, and “S”, excluding proper names and different forms of the same word (Borkowski, 1967, Loonstra 2001). For each letter, the participant is allowed one minute to generate the words. A comparable form of the test includes use of “C”, “F” and “L,” which may be used at alternate visits to reduce practice effect due to longitudinal testing on this task. Performance is measured by the total number of correct words produced summed across the three letters. Perseverations (i.e.,

repetitions of a correct word) and intrusions (i.e., words not beginning with the designated letter) are noted. Although fluency tests are sensitive to language dysfunction and deterioration of semantic knowledge, they can also reflect an inability to initiate systematic retrieval of information in semantic storage.

4.1.5.4 Category Fluency Task (CFT):

- Range: The score is the number of acceptable words produced. Lower numbers are associated with greater impairment.

- Brief description: The Category Fluency Test (CFT) is a measure of the subject's working memory and executive function. The CFT evaluates the spontaneous production of words beginning with a given letter, or of a given category, within 60 seconds. In this study, the participant is asked to orally produce as many animals as possible in one minute. Performance is measured by the number of acceptable words produced.

4.1.5.5 Alzheimer's Disease Cooperative Study Activities of Daily Living Scale (ADCS-ADL):

- Range: The scale range is 0-78, with lower scores indicating greater impairment.

- Brief description: The ADCS-ADL is an activity of daily living inventory developed by the ADCS to assess functional performance in participants with AD (Galasko et al., 1997). Informants are queried via a structured interview format as to whether participants attempted each item in the inventory during the preceding 4 weeks, as well as their level of performance. The ADCS-ADL will be administered at Baseline, and at Months 3, 6, 12 and 18 or in the event of early termination.

4.1.5.6 Clinical Dementia Rating Scale (CDR):

- Global CDR Range:
  - 0 = normal; healthy individuals
  - 0.5 = questionable dementia
  - 1 = mild dementia
  - 2 = moderate dementia
  - 3 = severe dementia.

- CDR sum-of-boxes (CDR-sb) range: CDR-sb scores range from 0 to 18 with higher scores indicating greater cognitive impairment.

- Brief description: The CDR scale is used as a global measure of dementia and is completed by a clinician in the setting of detailed knowledge of the individual patient collected from interviews with the patient and caregiver (Berg, 1988). The CDR describes 5 degrees of impairment in performance on each of 6 categories including memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. Much of the information will therefore already have been

gathered, either as part of normal clinical practice or as part of a research study. The interview takes approximately 40 minutes to administer.

The scores for each category can also be summed and this is known as the sum of box score (CSR-SB). The CDR will be conducted at Baseline and at Months 3, 6, 12 and 18 or in the event of early termination. To avoid patient fatigue, the CDR should be performed on a separate visit day than the other cognitive tests.

4.1.5.7 Neuropsychiatric Inventory (NPI):

- Range: Scores range from 0-144 with higher scores indicating a greater presence of neuropsychiatric symptoms.

- Brief description: The NPI is a well-validated, reliable, multi-item instrument to assess psychopathology in AD based on an interview with the caregiver (Cummings et al, 1994, Cummings, 1997). The interview is also relatively brief (15 minutes). It evaluates both the frequency and severity of 12 behavioral areas including delusions, hallucinations, dysphoria (depression) anxiety, agitation/aggression, euphoria, disinhibition, irritability, lability, apathy, aberrant motor behavior, appetite and eating changes and night-time behaviors.

  Frequency assessments range from 1 (occasionally, less than once per week) to 4 (very frequently, once or more per day or continuously) as well as severity (1= mild, 2 = moderate, 3 = severe).

  Distress is rated by the study partner or caregiver and ranges from 0 (no distress) to 5 (very severe or extreme). Distress is captured separately from the total score.

  The overall score and the score for each subscale are the product of severity and frequency. The NPI will be administered at Screening, Baseline, and at Months 3, 6, 12 and 18 or in the event of early termination.

4.1.5.8 Resource Utilization in Dementia (RUD):

- Resource questionnaire with limited statistical evaluation as an instrument.

- Brief description: The RUD is a validated and reliable questionnaire which assesses the health care resource utilization (HCRU) of the patient and caregiver and measures the level of formal and informal care (Wimo et al, 2012). The RUD consists of items about caregiver time, caregiver work status, caregiver HCRU (e.g., hospitalization, ER visits, health care professional visits and medication), patient living accommodation, and patient HCRU. It takes approximately 15-20 minutes to complete and is usually interview administered by any health care professional. This scale will be administered at Baseline and at Months 6, 12 and 18 or in the event of early termination.

4.1.5.9 Dementia Quality of Life (DEMQOL) – Proxy:

- Range: The scale range is 31—124, with lower scores indicate greater impairment.
- Brief description: The DEMQOL-Proxy questionnaire is a validated and reliable questionnaire that is interview administered and completed by the caregiver about the patient's health related quality of life (Smith et al, 2005). It consists of 31 items representing 5 domains (daily activities and looking after oneself, health and well-being, cognitive functioning, social relationships, and self-concept) and takes approximately 20 minutes to complete. Higher scores indicate better health-related quality of life. This scale will be administered at Baseline and at Months 6, 12 and 18 or in the event of early termination. Item 32 is a global quality of life question that elicits an overall assessment of the patient's general quality of life.

4.1.5.10 Trail-making Tests (TMT) "A" and "B":

- Range: 150 seconds maximum for "A" and 300 seconds maximum for "B."
- Brief description: The Trail-making tests "A" and "B" score is the number of seconds needed to complete the trails with maximum values specified (Reitan, 1958). Part A consists of 25 circles numbered 1 through 25 semi-randomly distributed over a white sheet of 8 1/2" X 11" paper. The participant is instructed to connect the circles with a pencil line as quickly as possible in ascending numerical order. Part B also consists of 25 circles, but these circles are either numbered (1 through 13) or contain letters (A through L). Now the participant must connect the circles while alternating between numbers and letters in an ascending order (e.g., A to 1; 1 to B; B to 2; 2 to C). Lower scores (seconds) indicate greater cognitive function.

4.2 Safety Variables

Safety is monitored in this study by collection of adverse events, vital signs, electrocardiography, and clinical laboratory measures. It is noted that all untoward events or experiences are reported as adverse events regardless of whether they are identified by clinical observation, subject reporting, physical examination, clinical laboratory test result, electrocardiography, or any other examination or test. It is noted that listings of vital signs, laboratory data, electrocardiograms (ECG), and all safety data from all other sources are reviewed by medically qualified individuals to ensure that safety signals are identified and reported in the analysis of the safety data from this study.

4.2.1 Treatment-emergent Adverse Events

Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA) Version 16.0 or above. Adverse event coding will be done to the lowest level term (LLT). Adverse events will be summarized by System Organ Class (SOC) and preferred terms (PT).

Definitions:

- A treatment-emergent adverse event (TEAE) is an event that is observed or reported after administration of study medication that was not present prior to study medication administration or an event that represents the exacerbation of a pre-existing event.
- An adverse withdrawal is a subject who withdrew from the study due to an adverse event.
- A serious adverse event (SAE) is an AE that is classified as serious according to the criteria specified in the study protocol.

Safety and tolerability variables based on adverse events include the following:

- Proportions of subjects with TEAEs by Preferred Term and decreasing frequency of AEs
- Proportions of subjects with TEAEs by System Organ Class and Preferred Term
- Proportions of subjects with related TEAEs
- Proportions of subjects with severe TEAEs
- Proportions of subjects who died.
- Proportions of subjects with treatment-emergent SAEs other than death.
- Subjects with TEAEs that result in study termination.

4.2.2 The Columbia Suicide Severity Rating Scale (C-SSRS)

- Range: 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).

    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.

- Brief description: The Columbia Suicide Severity Rating Scale (C-SSRS) is a joint interview with the caregiver and patient that systemically assesses suicidal ideation and suicidal behavior (Posner et al, 2007). This scale will be administered at Screening Visit to evaluate life time suicide attempt, suicide behaviors, and other non-suicidal self-injuries. Positive responses on the C-SSRS will be mapped to the Columbia Classification Algorithm of Suicide Assessment (C-CASA) for classification and reporting using a standard algorithm.)

4.2.3 Vital Signs

Vital signs measurements consist of blood pressure and pulse rate. Variables of analysis will include means, mean changes over time, and proportions of subjects meeting criteria for potential clinical concern in vital signs:

- Mean and mean changes in systolic and diastolic blood pressures
- Mean and mean changes in pulse
- Proportions of subjects with treatment-emergent values or changes of potential clinical concern.

4.2.4 Clinical laboratory assessments

Routine clinical laboratory data will be summarized with descriptive statistics on assessment values and change from Baseline. Routine clinical laboratory data will also be evaluated on the basis of laboratory-defined reference ranges and criteria defining potential clinical concern for values or changes, as defined in this document (Section 10). Routinely collected laboratory safety data that are not required by the protocol will be considered source data, and will not be captured for inclusion into the study database.

Definition:

- A treatment-emergent abnormal value (TEAV) is a laboratory value that is abnormal after administration of study medication that was normal prior to study medication administration.
- A TEAV of potential clinical concern is a value or change that meets criteria specified in this SAP in Section 10.

Laboratory analytes in this study include:

- Hematology
    - Hemoglobin
    - Hematocrit
    - Erythrocyte (RBC) count
    - Platelet count
    - Total leukocyte (WBC) count

  - MCV
  - Leukocyte differential (percent and total)
    - Total neutrophils
    - Eosinophils
    - Monocytes
    - Basophils
    - Lymphocytes

- Clinical chemistry
  - BUN
  - Creatinine
  - Glucose
  - HbA1c
  - Calcium
  - Sodium
  - Potassium
  - Chloride
  - Bicarbonate
  - AST
  - ALT
  - GGT
  - LDH
  - Total bilirubin
  - Alkaline phosphatase
  - Uric acid
  - Albumin
  - Total protein
  - Phosphorus

Safety variables of analysis include:

- Subjects with TEAVs
- Subjects with TEAVs of potential clinical concern (based on criteria in this SAP)
- Subjects with a TEAV in a liver enzyme function test (LFT).

4.2.5 Brain MRI: Safety

MRI scans will be performed at Screening and Month 18 or in the event of early termination. The primary intent of the MRI assessments was to support eligibility for the trial and key secondary objectives of evaluating the effect of TTP488 on MRI brain volumetric measures.

Variables of analysis will include number (%) of subjects with:

- Presence of vasogenic edema as categorized as:
  - No vasogenic edema present;
  - Questionable presence of vasogenic edema;
  - Vasogenic edema present, mild severity;
  - Vasogenic edema, moderate severity,
  - Vasogenic edema present, severe severity.

- The status of vasogenic edema at Month 18 compared to baseline will be evaluated as the number (%) of subjects categorized as vasogenic edema either present or absent on prior MRI; if present, further categorized as:
    - Vasogenic edema present and unchanged in size,
    - Vasogenic edema present and increased in size,
    - Vasogenic edema present and decreased in size.

4.3 Pharmacokinetics Variables

Blood samples for TTP488 and metabolite (TTP2123, TTP1494, TTP2266) plasma concentrations will be collected prior to dosing at Months 3, 6, 12, 18 and 21 or at Early Termination. All participants with at least one dose of study medication will be included in the pharmacokinetic analysis. Variable of analysis will be plasma trough concentration at each time point.

4.4 Pharmacodynamics Variables

Blood samples for plasma biomarkers analysis will be collected prior to dosing on Baseline, Months 3, 6, 12, 18, 21 or at Early Termination.

All participants with at least one dose of study medication will be included in the pharmacodynamic analyses. For individual endpoints, participants must have at least 1 post-dose pharmacodynamic measurement for the given endpoint. For change from Baseline, participants must also have a Baseline value. Variables of analysis will be Plasma Aβ (total, 1-40 and 1-42) concentration and change from baseline over time.

4.5 Adverse Events of Special Interest

Adverse events related to signs and symptoms of AD potentially pertain to efficacy as well as safety of the compound. Analysis of prior study data suggest the potential abatement of the signs and symptoms of AD that are associated with worsening of AD. Specifically, TEAEs classified by MedDRA into the SOC for psychiatric disorders (at the SOC level within the hierarchy of MedDRA) and specifically high level group terms (HLGTs) of anxiety disorders and symptoms (at the HLGT level within the hierarchy of MedDRA) will be variables of interest for comparison.

Standardized MedDRA queries (SMQs) will be used for dynamic analysis of any TEAE that emerges of potential interest.

5 Statistical Methodology

General statistical methods will be applied as appropriate for out-patient clinical studies. Safety evaluations will rely on descriptive statistics. Randomization groups will be examined for homogeneity between treatment groups.

Continuous variables will be summarized using mean, median, standard deviation, minimum, maximum, and number of subjects available for analysis. Categorical variables will be summarized using frequency, proportion, and number of subjects available for analysis.

The primary analysis is completely specified in the protocol, and that analysis is the intended analysis to be advanced as primary. Additionally, supportive analyses, sensitivity analyses, and secondary analyses will be done with respect to the primary endpoint.

SAS version 9.3 or later will be used.

5.1 Pooling of Investigator Sites

Blind data review (recommended by ICH E9) has revealed many sites with few subjects randomized. Understanding the influence of site effects on analysis conclusions is essential. Analysis by site is recommended in ICH E9 for key variables of analysis. Site pools are defined as follows:

- Sites with 12 or more subjects randomized will stand alone
- Sites with exactly 11 subjects randomized will be pooled together for Pool 11
- Sites with exactly 10 subjects randomized will be pooled together for Pool 10
- Sites with exactly 9 subjects randomized will be pooled together for Pool 9
- Sites with exactly 8 subjects randomized will be pooled together for Pool 8
- Sites with exactly 7 subjects randomized will be pooled together for Pool 7
- Sites with exactly 6 subjects randomized will be pooled together for Pool 6
- Sites with exactly 5 subjects randomized will be pooled together for Pool 5
- Sites with exactly 4 subjects randomized will be pooled together for Pool 4
- Sites with exactly 3 subjects randomized will be pooled together for Pool 3
- Sites with exactly 2 subjects randomized will be pooled together for Pool 2
- Sites with a single subject randomized will be pooled together for Pool 1.

Site or site pool will be included in analysis and displays using site or site pool as defined above.

Forest plots will be generated by site to the extent feasible to describe key results by site.

5.2 Statistical Methodology for Efficacy Analysis

Efficacy evaluations will include the co-primary endpoints (ADAS-cog and CDR-sb), key secondary endpoint, secondary endpoints and other efficacy variables.

The primary analysis of efficacy endpoints will be done on the FAS with supportive and sensitivity analyses done to examine robustness against analysis assumptions and handling of missing data.

Statistical analysis will be done on the A-study and the B-study, independently, for the co-primary variables: change from baseline in ADAS-cog and change from baseline in CDR-sb. The A-study includes the first approximately 400 patients randomized, and the B-study includes the remaining approximately 400 patients randomized. For all other secondary variables and subgroup analyses, the study as a whole (A-study and B-study combined) will be used for inferential purposes. It is noted that randomization supports independent analysis of A-study and B-study.

The primary analysis will be done for the A-study and the B-study separately, and each will use alpha = 0.049, which includes an adjustment of 0.001 for interim analyses performed by the Independent Data Monitoring Committee (IDMC), thereby preserving the overall study-wise alpha = 0.05.

This study is randomized using site randomization; the study randomization also includes a randomization stratum for severity of AD using a cut-point of 19 on the ADAS-cog at the time of randomization (screening values are used, due to the expectation that a confirming evaluation would be done). Efficacy analyses will be done including randomization stratum as a main effect.

5.2.1 Statistical Methodology for Primary Efficacy Analysis

The primary analysis will done on change from baseline in ADAS-cog and on change from baseline in CDR-sb.

The primary analysis will use the ITT methodology and a main-effects (for covariates) mixed-models for repeated measures (MMRM) methodology. The MMRM will utilize an unstructured covariance where the number of parameters is “t(t+1)/2” where “t” is the dimension of the covariance matrix, using PROC MIXED in SAS. In the unlikely event of lack of convergence, Toeplitz structure will be employed where the number of parameters is “t” where “t” is the dimension of the covariance matrix. Time is considered as a class variable. Analysis will include treatment, time, treatment-by-time interaction, Baseline as covariate, and baseline stratum as a covariate, and subject as a random effect (acknowledging that subject is a source of random variation). This method provides the primary treatment comparison at Month 18. Supportive modeling will also include MMRM main-effects model with treatment, time, and subject.

The primary analysis will also be done on randomization stratification subgroups defined by (1) baseline ADAS-cog 19 or less and (2) baseline ADAS-cog more than 19.

Multiple imputations (MI) will be used as an alternative for coping with missing data as supportive analyses with 100 invocations (acknowledging that more invocations are needed with more missing data). Monte Carlo methods are planned.

The MI analysis assuming Missing at Random (MAR) will be done as follows:

- First, for patients with intermittent missing values, before performing the MI, a monotone missingness pattern will be created. Intermittent missing values will be imputed using the Markov Chain Monte Carlo (MCMC) methodology which assumes a multivariate normal distribution over all variables included in the imputation model. The imputation model will be performed by treatment group and will include age, sex, ApoE status, education level, body weight (or BMI), baseline stratum, baseline and observed values at Months 3, 6, 9, 12, 15 and 18. Months 9 and 12 are not included in the model for endpoints which were not scheduled to assessed at those visits. The MI procedure in SAS will be used for this purpose and this first MI step is planned to be repeated 100 times, creating several different datasets with a monotone missing data structure. A unique seed value will be used in the MI procedure and document to allow for replication. The imputation is based on the missing at random (MAR) assumption, i.e. the missing data are assumed to follow the same model as the other patients in their respective treatment arm.

- After this, the remaining missing data will be imputed using a method for monotone missingness, also based on the MAR assumption. Thus, for each of the 100 created datasets with a monotone missing data pattern, the MI procedure in SAS will used to impute missing values based on a sequential procedure reflecting the monotone missing data pattern. A seed value of 446786 will be used for this second invocation of the MI procedure. Patients with the first missing value occurring at Month 3 will have their missing Month 3 value replaced by an imputed value from a regression model by treatment group with age, sex, ApoE status, education level, BMI, baseline stratum, and baseline as potentially predictive variables.

- In the next step, patients with the first missing value occurring at Month 6 will have their missing Month 6 value replaced by an imputed value from a regression model with Month 3 data, in addition to the variables listed above.

- The same procedure will be repeated for all other visits.

- The imputed datasets generated with the approach described above do contain only non-missing values and are used as input in the model for the supportive analysis of the primary endpoint. MMRM models similar to what are described above will thus be run on each of the generated imputed datasets and the difference between the treatment groups at Month 18 will be estimated. The MMRM model will be similar to the primary analysis. Finally, the MIANALYZE procedure in SAS will be applied to combine the results from these several datasets to derive an overall estimate of the treatment difference at Month 18. In addition to the estimates, corresponding 95% confidence intervals and p-values will be calculated.

Another MI analysis will be performed with the assumption that the data in the active treatment group are missing not at random (MNAR). A tipping point-based assumption will be used, i.e. the trajectories of the patients in the active treatment group after withdrawal are assumed to be worse by an amount of delta. After the MI using the MAR assumption, as defined above has been done, the amount of delta will be added to each imputed value in the active treatment group. Successively harsher deltas will be imposed on the imputed values in the active group, starting with an ADAS-cog increment (worsening) of 0.25 points or a CDR-sb increment

(worsening) of 0.1 points. The delta is further increased in the steps of 0.25 points (e.g., 0.50. 0.75) for ADAS-cog or in the steps of 0.1 points (e.g., 0.2, 0.3) for CDR-sb until the statistical significance is lost, i.e., until the p-value becomes more than 0.049. For the placebo group, the MI using MAR assumption will be used.

Further MI analysis assuming MNAR will be done using the Copy Reference approach. Placebo-group based assumption will be used, i.e. the trajectories of the patients are assumed to follow the placebo group after the discontinuation. Methods similar to those described above will be used to generate monotone missing data. After this, the missing data will be imputed sequentially for each visit (e.g., Month 3, Month 6). Only the data from the placebo group will be used for the imputation.

For the ANCOVA models, data from baseline and Month 18 will be used. Analysis will include treatment, baseline stratum as fixed factors and Baseline as covariate. For the LOCF analysis, the last available measurement will be used to impute the Month 18 data in case of missing data.

Rank analogues (e.g., Wilcoxon analysis, rank ANCOVA, and van Elteren analysis when strata are a part of the approach) will be used for additional supportive analyses. An observed-cases analysis will be done by visit.

An essential component of the thorough analysis includes an assessment of the impact of missing data on study conclusions. As a part of assessing this impact, sensitivity analysis will include an endpoint analysis (reduction to last-observation-carried-forward, justified in this study based on the monotonic progression of the natural course of AD and assuming no more dropouts in the group treated with azeliragon plus SoC than in the group treated with placebo plus SoC), a completers analysis (observed cases at Month 18), and observed cases by assessment time.

Interaction terms will be examined in the supportive analyses. The primary model will not include interaction terms with covariates. In the event of a significant interaction term, the impact on analysis conclusion will be examined. A key sensitivity analysis will use the ITT methodology and a main-effects model for ANCOVA adjusting for baseline ADAS-cog (CDR-sb) using last-observation-carried-forward (LOCF) methods for missing data, justified based on known profiles of ADAS-cog (CDR-sb) in AD patients, which is conservative under the assumption that there are not more dropouts in the active-treated group than the placebo group. Rank ANCOVA will be done as a supportive analysis.

Additional unplanned sensitivity analyses may be performed.

Descriptive summaries will be produced of the observed values and change from Baseline in co-primary variables by treatment group at each individual time point and at endpoint (final on-treatment assessment for each subject).

For statistical analyses, 95% confidence intervals will be produced for the least-squares means (LSM) in each treatment group, as well as the LSM differences as compared to placebo plus SoC. For MMRM and ANCOVA, two-sided p-values will be displayed for the comparison against placebo plus SoC.

Because there are limited numbers of subjects per center, there will inadequate power to explore impact of the center effect on study conclusion. Descriptive statistics by center will be presented.

The differences in the least squares means between Azeliragon and placebo, p-values, standard errors of the differences, and corresponding 95% confidence intervals for the differences will be presented. In addition, the least-squares means and corresponding 95% confidence intervals will be presented for each treatment group.

5.2.1.1 Responder Analyses

Survival techniques will be applied to data assessing time to increase in ADAS-cog of 7 points (based on meta-analysis published by Vellas, et al; 2007). For assurance that the results are not an artifact of the cut-point, analysis will be done for time to increases of each point from 1 to 20. Logrank and Wilcoxon tests will be done as descriptive analyses. Hazard ratios will be examined as descriptive analyses.

A forest plot will be produced to evaluate the sensitivity to the cut-point of the responder analysis.

5.2.1.2 Other Descriptive Analysis

Descriptive summaries will be produced of the observed values and change from Baseline in co-primary variables by treatment group at each individual time point and at endpoint (final on-treatment assessment for each subject).

For statistical analyses, 95% confidence intervals will be produced for the least-squares means (LSM) in each treatment group, as well as the LSM differences as compared to placebo plus SoC. For MMRM and ANCOVA, two-sided p-values will be displayed for the comparison against placebo plus SoC.

5.2.1.3 Investigator Sites

Because there are limited numbers of subjects per center, there is inadequate power to explore impact of the center effect on study conclusions. Descriptive statistics by center or center-pool will be presented.

5.2.1.4 Per Protocol Set

If the PPS differs from the FAS by more than 15%, the analyses will be replicated on the PPS. If the PPS and the FAS do not differ by more than 15%, analysis may not be done on the PPS.

5.2.1/5 Adjustments for multiplicity

Multiplicity of the primary efficacy analyses is controlled by using a conditional sequence of hypotheses. The study will be considered to demonstrate statistical significance if the primary analysis has a resulting p-value less than 0.049 for both primary endpoints for both A-study and B-study. No other adjustment for multiplicity is required.

5.2.2 Statistical Methods for Secondary Variables

Secondary and exploratory endpoints that are measurement variables will use similar statistical methodology to the methodology used for primary analysis. Stratification for A-study and B-study will be done as for primary variables. Efficacy variables that are proportions will be analyzed using Mantel-Haenszel test, controlling for Stratum, if appropriate. For the analysis of the categories for responders, a Cochran-Mantel-Haenszel test will be used. Methodologies for variables that are proportions will include construction of confidence intervals for each group and for the difference between groups. For analysis, Fisher's exact test will be used for single-population analysis, and a Mantel-Haenszel test will be used for analyses combining data across strata.

Responders based on the ADAS-cog total score, will be assessed using cumulative percentage plots. Similarly, responders based on the CDR-sb score will be assessed using cumulative percentage plots. The y-axis will show the cumulative percentage of patients who achieved the specific measure of improvement in the ADAS-cog (or CDR-sb) total score shown on the x-axis with a separate cumulative percentage curve for each treatment group.

The proportion of participants who experience loss of one global stage on the CDR will be assessed at each Visit using logistic regression with treatment (as a categorical/class variable), and baseline CDR global score (as a continuous variable). Summary measures from the analysis will include the odds ratio, 95% confidence interval for the odds ratio, and p-value for treatment comparison. Additionally, the proportion of participants in each CDR global stage will be presented.

Kaplan-Meier plots will be provided for each time-to-event variable. For each day, the "survival probability" will be computed by the product-limit method. Time-to-event endpoints may be modelled using Cox proportional hazards models with fixed effects for treatment and baseline measures as a covariate for supportive purposes. The estimated hazard ratio between Azeliragon and placebo will be presented with associated p-values and 95% confidence intervals.

Time to loss of one global stage on the global CDR will include Kaplan-Meier curves. Survival estimates for treatment groups will be compared using the logrank statistic. Observations will be treated as right censored for participants who complete the study without a loss of one stage on the global CDR score or who discontinue the study prematurely.

Subgroup analyses will be done as identified subsequently; data-driven subgroup analyses may be done as exploratory analyses.

For change from Baseline in plasma concentrations of Aβ species, an MMRM analysis similar to that described for the co-primary endpoints with baseline plasma concentrations of Aβ species as a continuous covariate will be performed. Transformation using logarithms may be needed, depending on the distribution.

For change from Baseline in MMSE, an MMRM analysis similar to that described for the co-primary endpoints will be performed, with baseline MMSE as a continuous covariate.

For change from Baseline in NPI, an MMRM analysis similar to that described for the co-primary endpoints will be performed, with baseline NPI as a continuous covariate.

Individual item responses on the RUD questionnaire and the total caregiver time spent assisting the patient based on the RUD questionnaire will be summarized by treatment group.

For change from baseline in DEMQOL-Proxy total score, an ANCOVA analysis with Baseline DEMQOL-Proxy as a continuous covariate will be performed.

If the distribution of any of the above parameters, key secondary, or other secondary endpoints does not appear to be normally distributed, the rank analogues will be utilized.

5.2.3 Subgroup Analyses

Subgroups will be defined based on randomization stratification and background or baseline characteristics.

The following subgroup analyses are planned:

- ApoE status: subjects will be classified according they have 1 or 2 copies of E4 (one group) or zero copy of E4 (one group). Efficacy analyses will be done for these subgroups.
    - E4=yes
    - E4=no

- Randomization stratification is done on the presenting ADAS-cog value (which is the screening assessment, due to the intention for the score to be evaluated for validity based on external review of the scoring). Subgroups will be defined by the randomization stratum of ADAS-cog 19 or less (one group) or ADAS-cog more than 19 (one group). Efficacy analysis will be done for these subgroups:
    - ADAS-cog 19 or less
    - ADAS-cog more than 19

- Years since diagnosis is calculated based on the date of diagnosis of AD. Subjects with date of diagnosis within 3 months or those with date of diagnosis more than 5 years have questionable status for having declining mild AD on stable medication for AD for at least 3 months. Efficacy analysis will be done for these subgroups:
    - Years since diagnosis between 3 months and 3 years
    - Years since diagnosis between 3 years and 5 years
    - Years since diagnosis more than 5 years.

5.2.4 MRI and PET Variables

For imaging variables, MRI variables are specified; PET variables are considered to be exploratory and will be specified separately. Statistical methodology will include stratification for A-study and B-study, acknowledging that the analysis of imaging data is intended to be for

the combined studies. Methodology for MRI will be similar to that for other secondary variables of analysis. Statistical methodology for PET variables will be described in the CSR.

Statistical methodology will include main-effects ANCOVA on change from baseline (adjusted for skull size). Supportive analysis will include Rank ANCOVA.

5.3 Disposition, Demography, and Baseline Characteristics

A tabulation of subject disposition will be presented, including the number screened, the number randomized in each population group, the number dosed in each population group, the number who withdrew prior to completing the study, and reasons for withdrawal.

Demographic and baseline characteristics (disease history, medical history, and prior treatments for AD) will be summarized for all randomized patients and for the FAS. No formal statistical comparisons will be performed. Summaries of continuous variables will include number of patients, mean, median, minimum, maximum, and standard deviation. Summaries of categorical variables will include numbers of patients in each category.

The variables to be summarized will include:

- Age, gender, race, and ethnicity
- Weight, height, BMI
- Years since diagnosis of AD
- Baseline MMSE
- ApoE status
- Education level
- Background AD medications (ACHEI, Memantine, Both).

5.4 Analysis of Safety Data

5.4.1 Adverse Events

Adverse events reported in this study will be coded using MedDRA®, Version 16.0, or later. Coding will be to the lowest level terms (LLT). The verbatim text, the preferred term, and the primary system organ class (SOC) will be listed in subject listings. Summaries that include frequencies and proportions of patients reporting AEs will include the preferred terms and the SOCs. Summaries will include TEAEs by severity and by relationship to study medication.

Adverse event summaries will be constructed displaying AEs in decreasing order of total frequency according to the numbers of subjects reporting the AE (not the number of reports).

- Number (percent) of subjects reporting TEAE by treatment and overall in accordance with variables listed in Section 4.1.1.

Supportive listings will be constructed that includes the subject identification, the treatment group, TEAEs, MedDRA terms, seriousness, severity, causality, elapsed time to onset, duration, and outcome.

Methodology for AESIs will follow the same methodology for TEAEs.

5.4.2 Vital Signs

Subjects with vital signs meeting the criteria in Section 10 of this SAP for values of potential clinical concern will be flagged and summarized. Proportions of subjects with vital signs of potential clinical concern will be examined for the treatment groups.

No formal inferential statistics will be applied to the vital signs data.

For each post-Baseline assessment, descriptive statistics are provided for the assessment value and the change from Baseline to the assessment with mean, median, standard deviation, minimum, and maximum at each assessment time.

5.4.3 Routine Clinical Laboratory Measurements

Because "exacerbations" of pre-existing abnormalities in laboratory analytes are examined using clinical judgment and are reported as TEAEs, additional analysis on TEAVs is limited to subjects with values that are normal prior to dosing and abnormal after dosing. Potentially clinically significant abnormalities in laboratory analytes are to be reported as TEAEs and summarized as clinical TEAEs.

Subjects with clinical laboratory data meeting the criteria in Section 10 of this SAP for values of potential clinical concern will be flagged and summarized. Proportions of subjects with clinical laboratory values of potential clinical concern will be examined for the treatment groups.

For each post-Baseline assessment, descriptive statistics are provided for the assessment value and the change from Baseline to the assessment with mean, median, standard deviation, minimum, and maximum at each assessment time.

5.4.4 Liver Enzyme Function Tests

Liver function tests have additional analysis views for this study in accordance with current regulatory guidance. To explore the potential for drug-induced liver injury consistent with *Guidance for Industry "Drug-induced liver injury: premarketing clinical evaluation"* (CDER, CBER, July 2009), subjects will be summarized and listed who meet the following criteria:

(1) Elevations in either AST or ALT of at least 3-times the upper limit of normal, and
(2) An accompanying abnormal bilirubin of at least 2-times the upper limit of normal.

5.4.5 Electrocardiography

The heart rate, QT, PR, and QRS intervals will be recorded at each assessment time. No formal inferential statistics will be applied to the ECG data.

For each post-Baseline assessment, descriptive statistics are provided for the assessment value and the change from Baseline to the assessment with mean, median, standard deviation, minimum, and maximum at each assessment time.

In addition, the number of subjects with corrected and uncorrected QT values >500 msec will be summarized. Values from individual tracings within the triplicate measurement on Day 1, time 0 hour that are >500 msec will not be included in the categorical analysis unless the average from that triplicate measurement is also >500 msec.

Electrocardiography data will be summarized by Baseline and Change from Baseline to each scheduled assessment time with descriptive statistics.

Corrections to QT intervals will be made by Fridercia's method (QTcF). Categorical analysis will be done consistent with ICH E14, "Clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs" (October 2005).

In accordance with ICH E14, subjects will be categorized and summarized as described above according to:

- Absolute QTc interval prolongation:
  - QTc interval > 450
  - QTc interval > 480
  - QTc interval > 500
- Change from Baseline in QTc interval:
  - QTc interval increases from Baseline > 30
  - QTc interval increases from Baseline > 60.

5.4.6 Methodology for Brain MRI (Safety) Analysis

The primary intent of the MRI assessments was to support eligibility for the trial and key secondary objectives of evaluating the effect of TTP488 on MRI brain volumetric measures. In those subjects with MRI assessments safety evaluation was performed by a central neuroradiologist to evaluate presence and severity of ARIA-E (vasogenic edema), presence and extent of ARIA-H (microhemorrhages), presence of white matter disease, documentation of any other MRI abnormalities.

ARIA-H will be evaluate based on whether microhemorrhages are present, number present, status compared to prior MRI, number of new microhemorrhages since previous MRI, number of microhemorrhages on previous MRI that have resolved, number of microhemorrhages that have increased in size and number that have decreased in size.

5.5 Pharmacokinetics Analysis

Data resulting from blood sampling for trough concentrations of TTP488 and metabolites (TTP2123, TTP1494, TTP2266) will be collected and analyzed for changes over the 18-month treatment period.

No formal inferential statistics will be applied to the pharmacokinetic data.

5.6 Pharmacodynamics Analysis

Plasma Aβ (total, 1-40 and 1-42) will be analyzed for changes over the 18-month treatment period.

5.7 Exploratory Analysis

The analyses described in this analysis plan are intended to be done. It is noted that additional analyses may be done for this study. These additional analyses may or may not be reported.

5.8 Concomitant Medications

Concomitant medications will be summarized by drug category and listed.

6 Data Conventions

The following analysis conventions will be used in the statistical analysis.

6.1 Definition of Baseline

For safety evaluations, the Baseline assessment for all measurements will be the latest available, valid measurement taken prior to the initiation study medication, except for electrocardiography (ECG) values: The Baseline assessment for ECG include a triplicate assessment with the baseline value intended to be the arithmetic mean of the values taken at the Baseline assessment that are taken in triplicate.

6.2 Missing Data

In general, missing clinical data will not be imputed, except where explicitly defined. In the event of study withdrawal, a subject's final assessment will be used as the endpoint value, if applicable. For Trail Making Test A & B, if the assessment was not done due to cognitive reasons, the max score would be imputed. Note that the reason for "not-done" assessment is captured in the eCRF, and the specify field provides reason why. If the reason relates to cognitive reasons, the maximum time for Trails A (150 seconds) or for Trails B (300 seconds) is used.

Data are considered to be "on-treatment" if the assessment or collection follows the first administration of study medication and if the assessment occurs within 45 days following the final administration of study medication, justified based on the long half-life of this drug.

Dates with missing fields will not have days imputed.

In accordance with the recommendations of the report from the National Academy of Science (NAS) panel "*The Prevention and Treatment of Missing Data in Clinical Trials*," (National Research Council, 2010), sensitivity analyses will be done to ensure that study conclusions are robust against missing data. Last observation carried forward (LOCF) and baseline observation carried forward (BOCF) methods may be used as appropriate in sensitivity analysis. Missing data types will be examined, and statistical methodologies for appropriate sensitivity analyses will be finalized during blind data review as patterns emerge as the study progresses. Primary methodologies following the ITT principle are not subject to change.

The three types of missing data (Rubin, 1976) are considered to examine robustness of analysis conclusions against missingness and type of missingness:

- MCAR (missing completely at random) assumes the event of missingness is independent both of observable values and of unobservable values, and the missingness is entirely at random.

- MAR (missing at random) assumes that the missingness can be fully accommodated by accounting for variables where data are available.

- MNAR (missing not at random) the missingness may be related to the reason the data are missing.

The robustness of analysis conclusions against missingness is managed through supportive and sensitivity analyses, employing techniques MAR MI for placebo and MAR MI with delta adjustment (jump-to-placebo) for the group randomized to TTP488 (O'Kelly and Ratitch, *Clinical Trials with Missing Data: A Guide for Practitioners*, 2014).

6.3 Dropouts

If a subject withdraws from the study, if the date of an adverse event is not available and a determination of whether or not the event is treatment emergent cannot be made, by convention the event will be considered to be treatment emergent.

6.4 Visit Windows

Assessments taken outside of protocol allowable windows will be displayed according to the case report form (CRF) assessment recorded by the investigator.

6.5 Early Termination Assessments and Follow-up Assessments

When a patient withdraws from the study prematurely for any reason, an early termination (ET) visit is to be performed, except in the case when a patient withdraws consent (and the protocol specifically states that no further data would be collected). An ET visit should be "on treatment," and efficacy data collected at the ET visit are suitable for use in statistical analysis. Efficacy data from an ET visit will be mapped to a scheduled visit based on the study day of the ET visit using the convention that the ET is mapped to the nearest scheduled visit by date. If a subject has more than one efficacy value for a visit, the assessment closest to the target day for the visit will be used for analysis.

A follow-up visit is planned for subjects 3 months following the last dose of study medication, noting that subjects who roll-over into the open-label extension study will not have the follow-up visit, because the last day of dosing for TTP488-301 is one day before the planned first dose of medication under TTP488-303 (the protocol for the open-label extension study). Data collected at the follow-up visit are not considered "on treatment," and the data are not suitable for use in statistical analysis (as the data can be tainted by other treatments).

6.6 Unscheduled Assessments

Extra assessments (laboratory data or vital signs associated with non-protocol clinical visits or obtained in the course of investigating or managing adverse events) will be included in listings, but not summaries. If more than one laboratory value is available for a given visit, the observation identified by the investigator to be the scheduled assessment will be used in summaries; all observations will be presented in listings.

Data from unscheduled assessments will be included in longitudinal data summaries (e.g., minimum, maximum over time). Data planned at scheduled times for which unscheduled assessments are taken will not be reflected in by-visit summaries.

6.7 Values Below LLQ

Concentrations for TTP488 or metabolites below the lower limit of quantification (LLQ) for the assay will be set to zero, for the purpose of analysis.

7 Analysis Deviations

All statistical analyses and summary information will be generated according to this analysis plan. Any deviation from this plan will be documented in the clinical study report. Exploratory analyses are permitted per protocol; those analyses will not be considered to be deviations.

8 Software

All analyses will be done using SAS Version 9.3 or later.

9 Data Displays

The plans for displaying data will conform to this guideline and also to ICH E3 (*Structure and Content of clinical Study Reports)*, and to ICH E9 (*Statistical Principles for Clinical Trials*).

The following excerpt from ICH E3 (1996) applies to this SAP:

> "The guidance provided below is detailed and is intended to notify the applicant of virtually all of the information that should routinely be provided so that post-submission requests for further data clarification and analyses can be reduced as much as possible. Nonetheless, specific requirements for data presentation and/or analysis may depend on specific situations, may evolve over time, may vary from drug class to drug class, may differ among regions and cannot be described in general terms; it is therefore important to refer to specific clinical guidelines and to discuss data presentation and analyses with the reviewing authority, whenever possible. Detailed written guidance on statistical approaches is available from some authorities.
>
> "Each report should consider all of the topics described (unless clearly not relevant) although the specific sequence and grouping of topics may be changed if alternatives are more logical for a particular study" (p. 2).

In accordance with FDA and ICH guidance, tables, figures, and key data listings are planned. Exploratory and data-driven data displays may also be generated.

Planned displays will include, but are not limited to, the following:

- Summary of Accountability (number screened, enrolled, completing, withdrawn, by reason; listing of withdrawals with study day and reason)
- Summary of Demography and Background Data
- Summary of Adverse Events
- Summary of Adverse Events by severity and by relatedness
- Summary of Vital Signs (change over time; criteria of potential clinical concern)
- Summary of Laboratory Data, including LFTs (change over time, TEAV, and criteria of potential clinical concern)
- Summary of PK Data
- Summary of Plasma Amyloid-beta Data
- Summary of Efficacy Data (ADAS-cog, CDR-sb, CDR-global, MMSE, ADCS-ADL, NPI)
- Summary of MRI Data
- Summary of PET Data
- Listing of Demography
- Listing of Dosing
- Listing of all Adverse Events
- Listing of all Adverse Events that resulted in Death
- Listing of all Serious Adverse Events
- Listing of all Adverse Events resulting in Study Termination

- Listing of Treatment-emergent Abnormal Values of Potential Clinical Concern for Clinical Laboratory Data
- Listing of Vital Signs Data
- Listing of QTc data that meet criteria of potential concern as described in ICH E14
- Listing of Concentration Data
- Listing of MRI safety Data
- Listing of MRI volumetric Data
- Listing of FDG-PET SUVR Data

In addition to the data summaries and listings described above, the set of displays will include clinical laboratory normal ranges (with units) and also a display of the MedDRA mapping glossary, which will display the verbatim text as reported by the investigator, the LLT, PT, and SOC.

10 Values of Potential Clinical Concern

Parameter	Threshold (study-specific threshold calculated)
Hematology	
Hemoglobin	<8.5 g/dL
Hematocrit	<25 %
WBC (Leukocytes)	<0.6 x LLN (**<2.1 10^3/uL**)
Platelets	<0.5 x LLN (**<70 10^3/uL**)
Total Neutrophils (Abs)	<0.8 x LLN (**<0.8 10^3/uL**)
Lymphocytes (Abs)	<0.6 x LLN (**<0.6 10^3/uL**)
Chemistry	
Total bilirubin	>2.5 x ULN (>**2.7 mg/dL**)
AST	>3 x ULN (**>102 U/L**)
ALT	>3 x ULN (**>123 U/L**)
Alk Phosphatase	>3 x ULN (**>348 U/L**)
Creatinine	>2 mg/dL
BUN	>1.4 x ULN (**30.8 mg/dL**)
Glucose	≤70 mg/dL or ≥270 mg/dL
Uric acid	>12 mg/dL
Sodium	<125 mmol/L or >155mmol/L
Potassium	<3.0 mmol/L or >6.0 mmol/L
Bicarbonate	<16 mmol/L
Calcium	<6.0 mg/dL and >13.0 mg/dL
Triglycerides	>750 mg/dL
Electrocardiogram	
Heart Rate	<50 bpm **and** >/=25% decrease from baseline >100 bpm **and** >/=25% increase from baseline
PR interval	≥200 msec **and** ≥25% increase from baseline
QRS interval	≥200 msec **and** ≥25% increase from baseline
QTc interval	≥500 msec ≥30-60 msec increase >60 msec increase
Vitals (sitting)	
SBP	≥190 mmHg (entry is <180 mmHg) Systolic <80 mm Hg Decrease from baseline ≥30 mmHg
DBP	≥120 mmHg (entry is <105) <50 mm Hg

Parameters not provided with specific thresholds for clinical concern: RBC count, eosinophils, basophils, monocytes, direct or indirect bilirubin (just total as shown above), chloride, albumin, total protein, insulin, total cholesterol, HDL, LDL, urinalysis, pulse rate from VS (rely on ECG heart rate).

11 References

- Berg, L. Clinical Dementia rating (CDR). Psychopharmacol Bull. 1988; 24:637-9.
- Folstein, MF, Folstein, SE; McHugh, PR. Mini-mental State. A practical method for grading the cognitive state of participants for the clinician. Journal of Psychiatric Research 1975; 12:189-98.
- Little, RJA, and Rubin, DB. (1987). Statistical analysis with missing data. New York: John Wiley and Sons.
- O'Kelly, M. Ratitch, B., 2014. *Clinical Trials with Missing Data: A Guide for Practitioners.* https://books.google.com/books?isbn=1118762533
- Rubin,l D.B. (1976, December). Inference and Missing data. Biometrika, Volume 63, Issue 3 (Dec., 1976), 581-592.
- Morris, J.C., Heyman, A., Mohs, R.C., Hughes, J.P., van Belle, G., Fillenbaum, G., Mellitis, E.D., Clark, C. "The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part 1. Clinical and neuropsychological assessment of Alzheimer's disease." Neurology 1989; Sep; 39 (9): 1159-65.
- Reitan R. Validity of the trail making test as an indicator of organic brain disease. Perceptual and Psychomotor Skills 1958; 8:271 6.
- Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. American Journal of Psychiatry 1984; 141:1356-64.
- Vellas, B., Andrieu, S., Canter, C. Dartigues, JF, Gauthier, S. Long term changes in ADAS-cog: what is clinically relevant for disease modifying trials in Alzheimer's. J Nutr Health Aging. 2007: 11(4);338-41.
- ICH. (1996, July 30). ICH Harmonized tripartite guideline: Structure and content of clinical study reports (E3). International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use.
- ICH. (1997, July 17). ICH Harmonized tripartite guideline: General considerations for clinical trials (E8). International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use.
- ICH. (1998, February 5). ICH Harmonized tripartite guideline: Statistical principles for clinical trials (E9). International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use.
- CDER and CBER. (1998, May). Guidance for Industry: Providing clinical evidence of effectiveness for human drug and biological products.
- ICH. (2005, October). ICH Harmonized tripartite guideline: Clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs(E14).
- International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use.
- CDER and CBER. (2009, July). Guidance for industry: "Drug-induced liver injury: premarketing clinical evaluation.
- CDER and CBER. (2010, February). Draft Guidance: Adaptive design clinical trials for drugs and biologics.
- CDER, Office of New Drugs. (2010, December 12). Manual of policies and procedures: Clinical review template.

- CDER, Office of Biostatistics. (2012, July 30). Manual of policies and procedures: Good Review Practice: Statistical Review Template.
- CDER, CBER, and CDRH. (2012, December). Draft guidance: Enrichment strategies for clinical trials to support approval of human drugs and biological products.
- CDER. (2013, February). Guidance for industry: "Alzheimer's disease: Developing drugs for the treatment of early stage disease."
- ICH. (2015, December). E14 Implementation Working Group: ICH E14 Guideline: The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for anon-antiarrhythmic drugs. Questions and Answers (R3).
- ICH and MSSO. (2017, September). MedDRA Introductory Guide Version 20.1.

12 Schedule of Time and Events

SCHEDULE OF ACTIVITIES

Protocol Activity	Screening	Baseline	Treatment									Early Term	Follow-up[a]
	Day -60 to Day 0	Day -7 to Day 0	Month 0	Month 1	Month 2	Month 3	Month 6	Month 9	Month 12	Month 15	Month 18		Month 21
Study Days			1	30	60	90	180	270	360	450	540		630
Window			±4d	±4d	±4d	±7d	±7d	±7d	±7d	±7d	±7d		± 7d
Sign informed consent / provide assent	X												
Screening Confirmation Form[b]	X												
Registration/Randomization		X											
Demographic information	X												
Review Inclusion/Exclusion Criteria	X	X											
Medical History	X												
Drug, Alcohol, Tobacco Use History	X												
Complete Neuro Exam & Physical Exam	X												
Rosen-Modified Hachinski Ischemic Score	X												
Height	X												
ApoE genotyping		X											
Telephone contact				X	X								
Body weight	X	X				X	X	X	X	X	X	X	
Brief Neuro & Physical Exams		X				X	X	X	X	X	X	X	X
Review Concomitant Medications	X	X		X	X	X	X	X	X	X	X	X	X
Blood Pressure and Pulse Rate (supine)	X	X				X	X	X	X	X	X	X	X
12 Lead ECG[d]	X	X				X	X	X	X	X	X	X	X
Adverse Events Assessment		X		X	X	X	X	X	X	X	X	X	X
Dispense Study Drug		X				X	X	X	X	X			
Study Drug Dosing			X	→	→	→	→	→	→	→	X		
MMSE	X	X				X	X		X		X	X	X
Study Days			0	30	60	90	180	270	360	450	540		630
Window			±4d	±4d	±4d	±7d	±7d	±7d	±7d	±7d	±7d		± 7d
ADAS-cog	X	X				X	X	X	X	X	X	X	X
CDR	X	X				X	X		X		X	X	X
NPI		X				X	X		X		X	X	X
ADCS-ADL		X				X	X		X		X	X	X
COWAT		X				X	X		X		X	X	X
CFT		X				X	X		X		X	X	X
Trail Making Test (Versions A and B)		X				X	X		X		X	X	X
Columbia Suicide Severity Scale (C-SSRS)[c]	X	X				X	X	X	X	X	X	X	X

Protocol Activity	Screening	Baseline	Treatment									Early Term	Follow-up[a]
	Day -60 to Day 0	Day -7 to Day 0	Month 0	Month 1	Month 2	Month 3	Month 6	Month 9	Month 12	Month 15	Month 18		Month 21
RUD Lite		X					X		X		X	X	
DEMQOL-proxy		X					X		X		X	X	
FSH (females only), hepatitis, VDRL, Thyroid function tests, Vitamin B12	X												
Hematology	X	X				X	X	X	X	X	X	X	X
Blood Chemistry (incl HbA1c)	X	X				X	X	X	X	X	X	X	X
Urinalysis	X	X				X	X	X	X	X	X	X	X
Urine Drug Screen	X												
Brain MRI	X										X	X	
Brain FDG-PET		X							X		X		
Blood Sample for Plasma Retention and Storage		X				X	X		X		X	X	X
Pharmacokinetic Blood Sampling		X				X	X		X		X	X	X
Pharmacodynamic blood sampling (Aβ)		X				X	X		X		X	X	X

a Follow-up visit required for those participants completing the study but choosing to not participate in an open label extension study.

b Following completion of Screening visit assessments but prior to scheduling baseline visit, obtain approval from the sponsor for the patient to participate in the study.

c C-SSRS is administered to subject jointly with care-giver.

d ECGs at Baseline visit performed in triplicate