STATISTICAL ANALYSIS PLAN

Title: Open label extension study for continued safety and efficacy evaluation of azeliragon in patients with mild Alzheimer's disease

Protocol: TTP488-303, Amendment 1, 9 August 2017

Study Drug: TTP488

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

Date: 21 September 2018

Signatures:

Imogene Dunn 21 SEP 2018

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

21 Sep 2018

Aaron Burstein, PharmD Date
SVP, Clinical Development
vTv Therapeutics LLC

Table of Contents

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 (azeliragon), 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 the open-label extension study for TTP488-301.

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

- To describe the approach of statistical analysis in order 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).

1.1 Study Objectives

Primary objective:

- To collect long-term safety and tolerability data.

Secondary objectives:

- 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 measures of behavior, cognition, function, resource utilization, and quality of life.

1.2 Study Rationale

The primary purpose of this trial is to allow participants from TTP488-301 who completed the 18-month study who may have received either TTP488 or placebo to receive TTP488 open-label.

1.3 Subject Population

The study population includes AD patients who completed TTP488-301 (entry MMSE for TTP488-301 was between 21 and 26, inclusive; CDR-global of 0.5 or 1; no requirements were imposed for entry into TTP488-303) who have a reliable caregiver 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 must have sufficient subject interaction to be able to provide meaningful input into the rating scales administered in this study where caregiver 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 that they meet caregiver criteria. Subjects were required to be at least 50 years of age to participate in TTP488-301 and TTP488-303.

1.4 Randomization

The open-label extension study was not randomized.

It is noted that TTP488-301 was randomized, and the randomization scheme for TTP488-301 is used to identify what will be called "treatment groups" in this SAP. The completers of TTP488-301 who are enrollees of TTP488-303 will be designated as

- Crossing over from placebo to azeliragon or
- Continuing on azeliragon.

Randomization in TTP488-301 had balanced allocation (1:1) between active and placebo. Dropouts were not allowed to be replaced.

1.5 Study Design

This is an open-label extension study in patients with mild Alzheimer's disease who have successfully completed participation in the azeliragon Phase 3 TTP488-301 (STEADFAST) trial. Patients may receive azeliragon 5 mg/day for up to 2 years.

Participants were required to continue to be on SoC of a stable background acetylcholinesterase inhibitor therapy and / or memantine and continue the therapy for the duration of the trial, unless the investigator judges that the dose needs to be reduced or stopped due to a safety and/or tolerability reason.

Subjects were expected to participate in approximately 9 outpatient visits including baseline which is the same visit at the Month 18 visit in the TTP488-301 study.

Protocol TTP488-301 included two studies with a common infrastructure (A-study and B-study). Participants from either study are allowed to roll over into this open-label extension study.

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

This study was terminated prematurely before participation outside of the US and Canada began; therefore, this study is conducted at multiple sites in the US and Canada.

The overall structure of the study includes the following periods:

Baseline Study Period: (Month 18 visit of TTP488-301): At the Month 18 visit of TTP488-301, subjects who met eligibility requirements were offered enrollment into this open-label extension study. Subjects had baseline assessments for efficacy and safety measures, including PK assessments.

Treatment Period: (Day 1 (Month 18 visit of TTP488-301) — Month 24): Subjects were treated with open-label azeliragon 5 mg QD. Efficacy and safety measures were taken. Blood samples were taken and subjects have assessments for safety and PK.

1.6 Sample Size and Power Considerations

This open-label extension permited enrollment of subjects who completed TTP488-301. The total sample size is limited by the enrollment into TTP488-301 and completion of TTP488-301. The protocol for TTP488-301 planned for a total of 800 patients to be enrolled.

No power considerations apply to this open-label study.

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.

1.8 Conformance to Regulatory Standards

An objective of this SAP is to comply with regulatory standards.

The planned statistical analysis of this study is intended to comply with regulatory expectations and standards,.

1.9 Modifications from the Statistical Section of the Protocol

There are no modifications from the Statistical Section of the Protocol.

2 Statistical Hypotheses

There are no formal statistical hypotheses for this open-label study.

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 population of analysis 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 assessment.

- The per protocol set (PPS) includes all subjects in the FAS, 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., taking less than 50% of prescribed study medication during participation).

  Final determinations of the PPS will be made at the final data review meeting (DRM) held in accordance with ICH E9 prior to the lock data.

- The safety set (SAF) includes all patients who received at least one dose of study medication. The SAF is used for all safety analyses.

The FAS is used for efficacy analysis; the SAF will be used for safety analyses.

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 FAS.

4 Variables of Analysis

4.1 Efficacy Variables

Efficacy evaluation will include as efficacy variables of analysis ADAS-cog and CDR-sb (co-primary variables of analysis), and MMSE, ADCS-ADL, and NPI; other efficacy measure may be evaluated using descriptive statistics, including other efficacy markers as identified in TTP488-301. It is noted that, although the primary objective of the study is safety and tolerability, efficacy evaluation is expected to be informative and useful. This SAP defines the primary and supportive variables of analysis to guide the statistical evaluation of the efficacy data.

4.1.1 Primary Efficacy Variable

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

- Mean change from Baseline in ADAS-cog
- Mean change from Baseline in CDR-sb.

.

4.1.2 Other Efficacy Variables of Analysis

Additional efficacy variables of analysis are as follows:

- Individual item responses on the Resource Utilization in Dementia (RUD) questionnaire and the total caregiver time spent assisting the patient based on the RUD.
- Change from Baseline in the Dementia Quality of Life (DEMQOL)-Proxy total score.
- Change from Baseline in the Neuropsychiatric Inventory (NPI) total score.
- Change from Baseline on the Mini-mental State Examination (MMSE) total score.
- Change from Baseline in Alzheimer's Disease Cooperative Study Activities of Daily Living Scale (ADCS-ADL) by visit.
- Change from baseline in Continuous Oral Word Association Task (COWAT) by visit.
- Change from Baseline in Category Fluency Test (CFT) by visit.
- Change from baseline in Trail-making Test (TMT) by visit.

4.1.3 Efficacy Scales

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

4.1.3.1 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. In TTP488-303, the ADAS-cog will be conducted at Months 3, 6, 9, 12, 15, 18, 21 and 24 or in the event of early termination. The ADAS-cog should always be administered prior to other cognitive measures.

4.1.3.2 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. In TTP488-303, the MMSE will be administered at Months 6, 12, 18 and 24 or in the event of early termination.

4.1.3.3 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. In TTP488-303, the COWAT will be administered at Months 6, 12, 18 and 24 or in the event of early termination.

4.1.3.4 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. In TTP488-303, the CFT will be administered at Months 6, 12, 18 and 24 or in the event of early termination.

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

- Range: 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. In TTP488-303, the ADCS-ADL will be administered at Months 6, 12, 18 and 24 or in the event of early termination.

4.1.3.6 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 Months 6, 12, 18 and 24 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.

4.1.3.7 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. In TTP488-303, the NPI will be administered at Months 6, 12, 18 and 24 or in the event of early termination.

4.1.3.8 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. In TTP488-303, this scale will be administered at Months 6, 12, 18 and 24 or in the event of early termination.

4.1.3.9 Dementia Quality of Life (DEMQOL) – Proxy:

- Range: 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. In TTP488-303, this scale will be administered at Months 6, 12, 18 and 24 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.

4.1.3.10 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 (eg, A to 1; 1 to B; B to 2; 2 to C). Lower scores (seconds) indicate greater cognitive function. In TTP488-303, the TMT will be administered at Months 6, 12, 18 and 24 or in the event of early termination.

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 are reviewed by medically qualified individuals of vital signs, laboratory data, electrocardiograms (ECG), and all safety data from all other sources 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 with treatment-emergent SAEs.
- Subjects with TEAEs that result in study termination.

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

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.) . In TTP488-303, the C-SSRS will be conducted at Months 3, 6, 9, 12, 15, 18, 21 and 24 or in the event of early termination.

- 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.

Endpoints:

Composite endpoints based on the above categories are defined below.

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

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

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

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

4.2.3 Vital Signs

Vital signs measurements consist of blood pressure and pulse rate. . In TTP488-303, the vital signs will be conducted at Months 3, 6, 9, 12, 15, 18, 21 and 24 or in the event of early termination. 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 clinically important 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. In TTP488-303, the clinical laboratory blood samples will be collected at Months 6, 12, 18 and 24 or in the event of early termination.

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

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 function test (LFT)

4.3 Pharmacokinetics and Pharmacodynamics Variables

Blood samples for pharmacokinetic (PK) analysis will be collected prior to dosing at Baseline (Month 18 of TTP488-301), Months 6, 12, 18 and 24 or at Early Termination.

Blood samples for plasma Aβ (total, 1-40 and 1-42) biomarkers analysis will be collected prior to dosing on Baseline (Month 18 of TTP488-301), Months 6, 12, 18, and 24 or at Early Termination.

4.4 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.

5 Statistical Methodology

General statistical methods will be applied as appropriate for out-patient, open-label clinical studies. Safety evaluations will rely on descriptive statistics. Randomization groups from TTP488-301 will be examined to delineate between subjects who received placebo during TTP488-301 and those who received TTP488 during TTP488-301. For the purpose of this SAP, these groups will be considered "treatment groups" acknowledging that all subjects receive open-label azeliragon in this study.

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.

Baseline will be the latest available data point prior to start of treatment in TTP488-303.

Data displays will include the following designations for treatment groups:

- Subjects randomized to azeliragon in TTP488-301
- Subjects randomized to placebo in TTP488-301
- All subjects.

SAS Version 9.4 or later will be used. Medical dictionary for Regulatory Activities (MedDRA) Version 18.0 or later will be used for coding adverse events. Medications will be coded using WHO Drug Dictionary (WHODD) Version March 2009 or later.

5.1 Statistical Methodology for Efficacy Analysis

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

Statistical analysis will be considered descriptive analysis. No alpha is allocated for this study.

5.1.1 Statistical Methodology for Primary Efficacy Analysis

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

Analysis of covariance (ANCOVA) will be done to compare those patients who were randomized to azeliragon in TTP488-301 with those who were randomized to placebo. The ANCOVA will include baseline measure as a covariate and treatment arm from TTP488-301; a main-effects model will be used.

Descriptive analysis will be applied to those patients who were treated with placebo in TTP488-301 to compare the profile during TTP488-301 with that of the profile in TTP488-303.

Descriptive statistics will be applied to the total patient population without regard for prior treatment.

Analyses will include data from the 18-month period from TTP488-301 with continuation for the available visit data from TTP488-303.

Additional supportive analyses will be done for which data displays are not planned.

Descriptive summaries will be produced of the observed values and change from Baseline in efficacy 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.

5.1.2 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.

Subgroup analyses may 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 ANCOVA analysis similar to that described for the co-primary endpoints with baseline plasma concentrations of Aβ species as a continuous covariate will be performed.

For change from Baseline in MMSE, an ANCOVA 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 ADCS-ADL, an ANCOVA analysis similar to that described for the co-primary variables of interest will be performed, with baseline ADCS-ADL as a continuous covariate.

For change from Baseline in NPI, an ANCOVA 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.

5.1.3 Subgroup Analyses

Subgroups may be defined post-hoc.

5.2 Disposition, Demography, and Baseline Characteristics

A tabulation of subject disposition will be presented, including the number 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 enrolled 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
- MMSE category (mild or moderate)
- MMSE, ADAS-cog, CDR-sb, CDR-global, ADCS-ADL
- ApoE status
- Education level
- Background AD medications (ACHEI, Memantine, Both).

5.3 Analysis of Safety Data

5.3.1 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 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 4.2.1.

Supportive listings will be constructed that includes the subject identification, the treatment group (based on randomization from TTP488-301 and the pooled group of enrollees from TTP488-303), TEAEs, MedDRA terms, seriousness, severity, causality, elapsed time to onset, duration, and outcome.

Methodology for AESIs will follow the same methodology for TEAEs.

5.3.2 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.

5.3.3 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.

5.3.4 Liver 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.

5.3.5 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 Fridericia'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 Pharmacokinetics Analysis

5.4.1 Methodology for PK Analysis

Data resulting from blood sampling for trough concentrations of azeliragon and its metabolites TTP1494, TTP2266, and TTP2123 will be collected and analyzed using descriptive statistical methods.

No formal inferential statistics will be applied to the pharmacokinetic data. Potential relationships between azeliragon concentrations and treatment effects may be examined.

5.5 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.

5.6 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

Two baselines will be used in this study: (1) the latest assessment prior to the first dose of double-blind study drug, and (2) the latest assessment prior to the first dose of open-label study drug. For safety evaluations, the Baseline assessment for all measurements will be the latest available, valid measurement taken prior to the initiation of study medication in TTP488-303.

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.

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 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.

6.6 Values Below LLQ

Concentrations for TTP488 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 9.4 or later.

9 Data Displays

The plans for displaying data will be appropriate for open-label extension studies.

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

- Summary of Accountability (number 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 PD Data
- Summary of Efficacy Data (ADAS-cog, CDR-sb, CDR-global, MMSE, ADCS-ADL, NPI)
- 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 azeliragon, TTP1494, TTP2266, TTP2123 Concentration Data
- Listing of plasma Aβ (total, 1-40 and 1-42)

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. Biometrikia, Volume 63, Issue 3 (Dec., 1976), 581-592.
- 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.

- 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.

12 Schedule of Time and Events

SCHEDULE OF ACTIVITIES

Protocol Activity	Baseline	Treatment								Early Term
	Month 18 of TTP488-301 study	Month 3	Month 6	Month 9	Month 12	Month 15	Month 18	Month 21	Month 24	
Study Days		90	180	270	360	450	540	630	720	
Window		±14d	±14d	±14d	±14d	±14d	±14d	±14d	±14d	
Sign informed consent / provide assent	X*									
Registration	X*									
Review Inclusion/Exclusion Criteria	X*									
Body weight	X		X		X		X		X	X
Brief Neuro & Physical Exams	X	X	X	X	X	X	X	X	X	X
Review Concomitant Medications	X	X	X	X	X	X	X	X	X	X
Blood Pressure and Pulse Rate (sitting or supine)	X	X	X	X	X	X	X	X	X	X
12 Lead ECG	X		X		X		X		X	X
Adverse Events Assessment	X	X	X	X	X	X	X	X	X	X
Study Drug Dispensing / Accountability	X*	X	X	X	X	X	X	X	X	X
Study Drug Dosing	X*	→	→	→	→	→	→	→		
MMSE	X		X		X		X		X	X
ADAS-cog	X	X	X	X	X		X		X	X
CDR	X		X		X		X		X	X
NPI	X		X		X		X		X	X
ADCS-ADL	X		X		X		X		X	X
COWAT	X		X		X		X		X	X
CFT	X		X		X		X		X	X
TMT	X		X		X		X		X	X
Columbia Suicide Severity Scale (C-SSRS)	X	X	X	X	X	X	X	X	X	X
RUD-lite	X		X		X		X		X	X
DEMQOL-proxy	X		X		X		X		X	X
Hematology	X		X		X		X		X	X
Blood Chemistry (incl HbA1c)	X		X		X		X		X	X
Pharmacokinetic Blood Sampling	X		X		X		X		X	X
Pharmacodynamic blood sampling (Aβ)	X		X		X		X		X	X

*procedures unique to OLE. All other baseline procedures are required TTP488-301 Month 18 procedures and are not repeated as part of the TTP488-303 baseline visit.