Statistical Analysis Plan for Study M23-515

A Randomized, Double-blind, Placebo-controlled, Dose-finding Study to Evaluate the Safety, Efficacy, Pharmacokinetics, and Pharmacodynamics of ABBV-552 in Participants with Mild Alzheimer's Disease

Date: 20 September 2024

Version 3.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for ABBV-552 Study M23-515, A Randomized, Double-blind, Placebo-controlled, Dose-finding Study to Evaluate the Safety, Efficacy, Pharmacokinetics, and Pharmacodynamics of ABBV-552 in Participants with Mild Alzheimer's Disease (AD).

Study M23-515 examines the efficacy and safety of ABBV-552 in participants with dementia due to mild AD.

The analyses of pharmacokinetic endpoints and plasma biomarkers will not be covered in this SAP.

The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

Unless noted otherwise, all analyses will be performed using SAS Version 9.4 (SAS Institute Inc., Cary, NC 27513) or later.

This SAP includes changes to analyses described in the protocol. Details are outlined in Section 14.1.

2.0 Study Objectives and Design

2.1 Study Objectives

This is a proof-of-concept, dose-finding study to evaluate the safety, efficacy, pharmacokinetics (PK), and pharmacodynamics (PD) of ABBV-552 once daily for the treatment of dementia due to mild AD.

The primary clinical hypothesis is that at least 1 dose of ABBV-552 (i.e., 1 mg, 5 mg, or 15 mg) is superior to placebo on the primary endpoint in participants with mild AD.

The administration of ABBV-552 once daily will be well tolerated and provide clinical benefit for participants with mild AD.

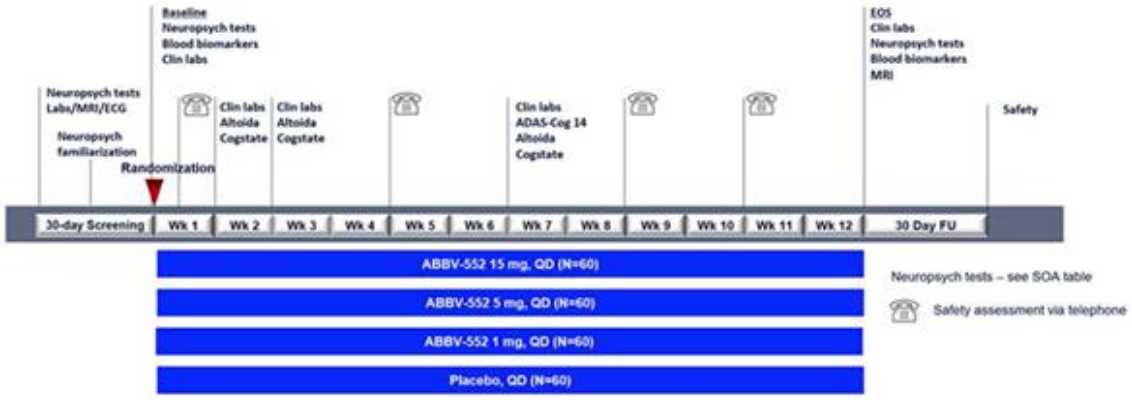
2.2 Study Design Overview

The study will consist of a Screening Period of up to 30 days, a Double-Blind Treatment Period of 12 weeks, and a Safety Follow-up Period of 30 days.

There will be a total of 12 scheduled study visits as shown in Figure 1. The total study duration is approximately 20 weeks.

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



2.3 Treatment Assignment and Blinding

Participants will be randomized to placebo, ABBV-552 1 mg QD, ABBV-552 5 mg QD, or ABBV-552 15 mg QD, in a 1:1:1:1 ratio. Randomization will be stratified by whether or not the participant is currently receiving a stable dose of symptomatic treatment for AD at Baseline/Day 1.

2.4 Sample Size Determination

It is assumed that the treatment difference of Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) 14-item score change from Baseline at Week 12

between each ABBV-552 treatment arm and placebo is 2.85 with a pooled standard deviation of 5.7 (effect size of 0.5) based on data from prior cholinesterase inhibitor and AbbVie studies. The standard deviation of 5.7 was estimated from pooled previous AbbVie symptomatic treatment in mild/moderate AD studies (M10-984, M10-985, M10-822, and M12-033). Assuming a 20% discontinuation rate, 60 participants randomized per group will provide about 80% power to detect the above treatment difference in ADAS-Cog 14 score at the 2-sided 10% significance level. East version 6.5.3 was used for the sample size calculation.

3.0 Endpoints

3.1 Primary Endpoint(s)

The primary endpoint is the change from Baseline in the ADAS-Cog 14 score at Week 12.

3.2 Additional Efficacy Endpoint(s)

Additional efficacy endpoints include:

- Change from Baseline in the Cogstate Computerized Battery at specified timepoints, including Weeks 1, 2, 6, and 12.
- Change from Baseline in the ADAS-Cog 14 score at Week 6.
- Change from Baseline in the Mini-Mental State Examination (MMSE) score at Week 12.
- Change from Baseline in the Clinical Dementia Rating sum of boxes (CDR-SB) score at Week 12.
- Change from Baseline in the Alzheimer's Disease composite score (ADCOMS) at Week 12.
- Change from Baseline in Neuropsychiatric Inventory (NPI) scores at Week 12.

3.3 Safety Endpoint(s)

Safety evaluations include AE monitoring, weight, vital sign measurements, ECG variables, C-SSRS, and clinical laboratory testing (hematology, chemistry, and urinalysis) as measures of safety and tolerability for the entire study duration.

3.4 Other Endpoint(s)

The analysis plan for endpoints related to brain magnetic resonance imaging (MRI) is provided in a separate document and is outside of the scope of this SAP.

4.0 Analysis Sets

The following analysis sets will be used.

The Intent-to-Treat analysis set includes all randomized participants and participants will be grouped according to treatment as randomized.

The modified Intent-to-Treat (mITT) analysis set includes all randomized participants who received at least 1 capsule of study drug and have at least 1 post-baseline ADAS-Cog 14 score. The participants will be grouped according to treatment as randomized. The mITT analysis set will be used for all efficacy analyses, including baseline efficacy summaries.

The safety analysis set consists of all participants who received at least 1 capsule of study drug. The participants will be grouped according to treatment actually received the majority of the time. This will be defined based on the number of days each treatment is received. The safety analysis set will be used for all safety analyses and demographic analyses, unless otherwise noted.

5.0 Participant Disposition

A summary of participant accountability by country and investigator will be provided where the number of participants in each of the following categories will be tabulated for each treatment group for the Intent-to-Treat Analysis Set:

- Participants randomized in the study;
- Participants who took at least one dose of study treatment;
- Participants who completed study treatment;
- Participants who prematurely discontinued study treatment;
- Participants in each analysis population, as applicable.

The summary of participant accountability by investigator also will include the number of participants who screened and the number of participants who screen failed and will be summarized for all participants.

The number and percentage of participants in the safety analysis set who prematurely discontinued study treatment will be summarized by reason for not completing study treatment overall and by treatment group.

The number and percentage of participants in the safety analysis set who did not complete the protocol defined follow-up period will be summarized by reason for not completing study overall and by treatment group.

6.0 Study Treatment Duration and Compliance

For the safety analysis set, duration of treatment will be summarized for each treatment group and for total ABBV-552 group. Duration of treatment is defined for each Participant as last dose date minus first dose date + 1. Duration of treatment will be summarized using the number of participants treated, mean, standard deviation, median, minimum and maximum. In addition, the number and percentage of participants in each treatment duration interval (1 to 28, 29 to 56, 57 to 84, > 84) will be summarized.

Treatment compliance will be summarized for the entire treatment period by treatment group and total ABBV-552 group for the safety analysis set. Treatment compliance is defined as the number of capsules actually taken divided by the number of capsules that should have been taken (number of capsules that should have been taken equals 2 per day multiplied by days where participant is on study treatment regardless of treatment

interruption). Percent compliance will be summarized with mean, SD, median, min, and max, as well as categorically (< 80%, ≥ 80% to ≤ 120%, > 120%).

7.0 Participant Characteristics

Categorical variables will be summarized with the number and percentage of participants. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

Demographics and baseline or disease characteristics will be summarized descriptively, overall and by treatment group for the safety analysis set. Unless otherwise specified, baseline is defined as the last non-missing value prior to the first administration of study treatment.

Continuous demographic variables include age, weight, height, and body mass index (BMI).

Categorical demographic variables include:

- Sex
- Ethnicity
- Race
- Age (< 65, ≥ 65 - < 75, ≥ 75 - < 85, ≥ 85)
- Weight (< 60 or ≥ 60 kg) by sex
- BMI (< 25 or ≥ 25 kg/m^2) by sex
- Region (Europe, North America, or rest of world)
- Nicotine user (current, former, never, unknown)
- Alcohol user (current, former, never, unknown).

Disease characteristics include the time since symptom (cognitive impairment) onset and time since Alzheimer's disease diagnosis. This will be summarized in years as a continuous variable.

The history of symptomatic treatment for AD will be summarized categorically as: previously treated and stopped, previously treated and treatment is ongoing, and never treated. For those who were previously treated and stopped, the reason for discontinuation will be summarized.

The following will be summarized for the mITT analysis set. The number and percentage of participants on symptomatic treatment for AD as reported on the concomitant medication eCRF at baseline will be summarized. The number of participants that start, stop or change doses of symptomatic treatment for AD during the treatment will be summarized.

7.2 Medical History and Prior and Concomitant Medications

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of participants in each medical history category (by MedDRA system organ class (SOC) and preferred term) will be summarized overall, all active treatment, and by treatment group for the safety analysis set. The SOC will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Participants reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

Prior and concomitant medications will be summarized separately. The number and percentage of participants taking prior and concomitant medications will be summarized by generic drug name, based on the World Health Organization (WHO) Drug Dictionary overall, all active treatment, and by treatment group for the safety analysis set. The actual

version of the WHO Drug Dictionary will be noted in the statistical tables and clinical study report.

7.3 Protocol Deviations

For each of the following protocol deviation categories and across all categories, the number and percentage of randomized participants with at least one protocol deviation will be summarized overall, all active treatment, and by treatment group for the safety analysis set:

- Participant entered into the study even though did not satisfy entry criteria;
- Participant developed withdrawal criteria during the study but was not withdrawn;
- Participant received wrong treatment or incorrect dose of study treatment;
- Participant took prohibited concomitant medication.

A listing of participants with protocol deviations will be provided.

8.0 Handling of Potential Intercurrent Events for the Primary Endpoint

The primary efficacy endpoint of change from baseline in the ADAS-Cog 14 score (defined in Section 3.1) will be analyzed based on the mITT analysis set and the following methods will be used to address potential intercurrent events:

- Efficacy data which is missing due to a participant's discontinuation from the study, or blind broken will be assumed missing at random and will not be included in the statistical analyses.
- Efficacy data collected after a start, stop or dose change of AD medication during the study will not be included in the primary analysis and will be assumed missing at random.

9.0 Efficacy Analyses

9.1 General Considerations

Unless otherwise noted, all efficacy analyses will be conducted on the mITT analysis set.

The primary analysis will be conducted following a database lock after all participants have completed Visit 11 (Week 12) or prematurely discontinued study drug.

Where hypothesis testing is specified, it will be based on a 2-sided $\alpha = 0.1$ level.

Continuous variables will be analyzed using mixed-effects model for repeated measures (MMRM) method or analysis of covariance (ANCOVA) as noted. For continuous variables analyzed using an MMRM or ANCOVA, summary statistics by visit will also be provided by treatment group.

For analyses that include the stratification factor in the model (on symptomatic treatment for AD at baseline – yes or no), any participant who is randomized within an incorrect stratum will be analyzed according to the actual stratum to which the participant belongs according to the concomitant medication eCRF, unless otherwise indicated.

9.2 Handling of Missing Data

Missing data for the primary endpoint and the continuous analysis of other efficacy endpoints that are measured at multiple time points post-baseline will be addressed in a model-based fashion via an MMRM analysis with no data imputation prior to creating the model.

Efficacy data collected after a start, stop or dose change of AD medication during the study will not be included in the primary, secondary, or additional efficacy analyses and will be assumed missing at random unless otherwise noted.

For continuous efficacy endpoints measured at one time point post-baseline, missing data will not be imputed, unless otherwise noted.

9.3 Primary Efficacy Endpoint and Analyses

9.3.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the change from Baseline in the ADAS-Cog 14 total score at Week 12.

9.3.2 Main Analysis of Primary Efficacy Endpoint

The attributes of the estimand corresponding to the primary efficacy objective are summarized in Table 1.

Table 1. Summary of the Estimand Attributes Corresponding to the Primary Efficacy Objective

Estimand Label	Attributes of the Estimand				
	Treatment	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Primary estimand	Placebo, ABBV-552 1 mg, 5 mg, and 15 mg QD	Change from Baseline in ADAS-Cog 14 score at Week 12	mild Alzheimer's Disease participants	Efficacy data which is missing due to a participant's discontinuation from the study will be assumed missing at random and will not be included in the statistical analyses. Efficacy data collected after a start, stop or dose change of AD medication during the study will not be included in the primary analysis and will be assumed missing at random	LS mean difference between each ABBV-552 dose and placebo in ADAS-Cog 14 score change from baseline at Week 12 based on an MMRM

The primary analysis of change from baseline in the ADAS-Cog 14 score will be conducted using an MMRM including treatment group, visit, treatment-by-visit interaction, and the randomization stratum as fixed effects and the baseline ADAS-Cog 14 score as a covariate. The unstructured covariance structure will be used to estimate the

within-participant variance-covariance. Denominator degrees of freedom will be computed using the Kenward-Roger method. The primary comparison time point is Week 12, and the group mean treatment difference between each dose and placebo at Week 12 will be based on contrasts from this model. If the model fails to converge with an unstructured covariance structure, the compound symmetry (CS) covariance structure will be substituted. If the model still fails to converge, data will be analyzed as observed using an ANCOVA with treatment and the randomization stratum as fixed effects and baseline ADAS-Cog 14 score as a covariate.

P-value for testing the group mean difference between each dose and placebo groups will be calculated.

9.3.3 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoint(s)

9.3.3.1 Sensitivity Analyses

The sensitivity analyses of the primary endpoint will include imputation using the copy-reference approach, and an alternate handling of intercurrent events.

The copy-reference approach is one type of pattern-mixture models (PMM), under which data could be missing-not-at-random (MNAR), with repeated analyses combined via the reference based multiple imputation (MI) procedure.[1] This approach is to assess the robustness of the MMRM analysis to possible violation of the missing-at-random (MAR) assumption in the primary analysis.

Step 1. Missing values will first be imputed by the Markov Chain Monte Carlo (MCMC) in order to attain monotone missingness. The MCMC imputation assumes MAR for intermittent missing data. The MCMC method will be implemented using SAS PROC MI, with the use of option IMPUTE = MONOTONE in the MCMC statement. The rest of the missing data will then follow a monotone missing pattern.

Step 2. Implementation of the copy-reference method are as follows:

1. The reference-based approach uses the placebo group as the reference. The missing values in the reference group are imputed using the observed data overall in that group under the MAR assumption. The missing pattern is defined by the participant's last visit with a non-missing value. The mean vector and the covariance matrix of the multivariate normal distribution are estimated for the reference group.
2. For ABBV-552 treatment groups, missing values are imputed based on the distribution estimated from the reference group (placebo group).

The first PROC MI will be performed 100 times using MCMC method for partial imputation of the data with a non-monotone missing pattern. The output dataset will then be used as the input dataset for the next PROC MI. Note that the output dataset already contains 100 copies of the original dataset. With the next invocation of MI procedure, the missing data will be filled in (Steps 1 and 2) for the existing copies. This is achieved with the use of NIMPUTE=1 and a BY _Imputation_ statement. Finally, each of the 100 imputed datasets will be analyzed using an analysis of covariance (ANCOVA) model. The model includes treatment group and randomization stratum as main effects, and baseline ADAS-Cog 14 score as a covariate. The LS mean difference and corresponding SE are estimated from the model comparing each ABBV-552 treatment group with the placebo group. The ANCOVA analysis results from 100 completed datasets are combined for overall estimation and inference using Rubin's rule[2] to produce a pooled estimate of LS mean difference and its SE. This copy-reference approach sensitivity analysis will be performed on mITT analysis set and safety analysis set.

Two additional sensitivity analyses of the primary endpoint will be provided by alternatively handling the intercurrent events:

1. Efficacy data collected after a start, stop or dose change of AD medication during the study will be included in the primary analysis. The same MMRM as specified in the primary analysis will be used.

2. Efficacy data collected after participant discontinued study treatment will be excluded in the primary analysis. The same MMRM as specified in the primary analysis will be used.

9.3.3.2 Supplementary Analyses

The group mean difference between each treatment group and placebo at Week 6 will be based on the same MMRM model noted in Section 9.3.1.

The change from baseline in ADAS-Cog 14 effect size will be calculated for each treatment group at Weeks 6 and 12.

9.4 Additional Efficacy Endpoints and Analyses

9.4.1 CDR

The change from baseline in the CDR-SB will be analyzed at Week 12 using an ANCOVA including treatment group and the randomization stratum as main effects and the baseline CDR-SB as a covariate.

The effect size for each treatment group over placebo will be provided.

A shift table from baseline to Week 12 of the number and percentage of participants in each category of the CDR global score, as well as missing, will be provided by treatment group.

9.4.2 MMSE

The change from baseline in the MMSE total score will be analyzed at Week 12 using an ANCOVA including treatment group and the randomization stratum as main effects and the baseline MMSE total score as a covariate.

The effect size for each treatment group over placebo will be provided.

9.4.3 Cogstate Brief Battery (CBB) and International Shopping List Test (ISLT)

The CBB consists of the following components: detection test [DET], identification test [IDN], one card learning [OCL], and one back test [ONB]. The ISLT consists of the following: ISLT-Immediate Recall and ISLT-Delayed Recall. Together, the CBB and ISLT are referred to as the Cogstate Computerized Battery (CCB).

The change from baseline for the CCB will be analyzed using an MMRM including treatment group, visit, treatment-by-visit interaction, and the randomization stratum as fixed effects and the baseline score as a covariate. The unstructured covariance structure will be used to estimate the within-participant variance-covariance. Denominator degrees of freedom will be computed using the Kenward-Roger method. The group mean treatment difference between each dose and placebo at Weeks 1, 2, 6, and 12 will be based on contrasts from this model. If the model fails to converge with an unstructured covariance structure, the compound symmetry (CS) covariance structure will be substituted. If the model still fails to converge, data will be analyzed as observed using an ANCOVA model with treatment and the randomization stratum as main effects and the baseline of the measure being analyzed as a covariate.

This will be calculated for the primary outcome measures (e.g., reaction time, accuracy, correct responses) for each component of the CCB as well as the global and domain-specific (attention domain, learning/working memory, and episodic memory) composite scores. The primary outcome measures and calculation details of the composite scores may be found in Appendix C.

Additionally, the effect sizes for the change from baseline for each component of the CCB will be calculated at Weeks 1, 2, 6 and 12.

9.4.4 Neuropsychiatric Inventory (NPI)

The change from baseline in the NPI will be analyzed at Week 12 using an ANCOVA including treatment group and the randomization stratum as main effects and the baseline MMSE as a covariate.

This will be done for the total (frequency x severity) and caregiver distress scores for each behavioral domain, as well as the total NPI score and the total distress score.

The effect size for each treatment group over placebo will be provided.

9.4.5 Alzheimer's Disease Composite Score (ADCOMS)

The Alzheimer's Disease Composite Score[3] is a function of components of the ADAS-Cog, MMSE and CDR-SB, and ranges from 0 to 1.97. It is calculated by weighting the score for each of the following by the coefficient, and taking the sum:

Table 2. Derivation of ADCOMS

		PLS coefficient
ADAS-Cog	Delayed word recall	0.008
	Orientation	0.017
	Word recognition	0.004
	Word finding difficulty	0.016
MMSE	Orientation time	0.042
	Drawing	0.038
CDR-SB	Personal care	0.054
	Community affairs	0.109
	Home and hobbies	0.089
	Judgment and problem solving	0.069
	Memory	0.059
	Orientation	0.078

The change from baseline in the ADCOMS score will be analyzed at Week 12 using an ANCOVA including treatment group and the randomization stratum as main effects and the baseline ADCOMS score as a covariate.

The effect size for each treatment group will be provided.

9.5 Efficacy Subgroup Analyses

The primary endpoint will be conducted for the following subsets of participants. Within each subgroup category, data will be divided into mutually exclusive subsets and then a separate analysis will be performed for each subset (e.g., for the "Sex" category, a separate analysis will be performed for all male participants and the same analysis will be performed for all female participants.)

- On symptomatic treatment for AD at baseline: Yes, No
- Sex
- Age: < 65 years or ≥ 65 years and < 75 years or ≥ 75 years

Other subgroup analyses (e.g., subgroups based on biomarkers) may be conducted given sufficient data.

The MMRM will include treatment group, visit, treatment-by-visit interaction, and the randomization stratum (for the age and sex subgroups only) as fixed effects and the baseline ADAS-Cog 14 score as a covariate. The unstructured covariance structure will be used to estimate the within-participant variance-covariance. Denominator degrees of freedom will be computed using the Kenward-Roger method. The group mean treatment difference between each dose and placebo at Week 12 will be based on contrasts from this model. If the model fails to converge with an unstructured covariance structure, the compound symmetry covariance structure will be substituted. If the model still fails to converge, descriptive statistics based on the univariate summaries will be provided.

10.0 Safety Analyses

10.1 General Considerations

Safety data will be summarized for the Safety Analysis Set. Safety summaries will be presented by treatment group, including a total group for all participants on active study treatment (ABBV-552 1 mg, 5 mg, 15 mg) unless otherwise noted.

10.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of final database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each Participant for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a Participant, the highest severity and level of relationship to investigational product will be reported.

10.2.1 Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as any AE with an onset date that is after the first dose of study treatment and no more than 30 days after the last dose of study treatment. Events where the onset date is the same as the study treatment start date are assumed to be treatment-emergent. All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of participants experiencing treatment-emergent AEs will be summarized.

10.2.2 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of participants experiencing at least one event for each of the following AE categories:

- Any treatment-emergent AE
- Any treatment-emergent AE related to study treatment according to the investigator
- Any treatment-emergent AE Grade 3 or higher
- Any serious treatment-emergent AE
- Any treatment-emergent AE leading to discontinuation of study treatment
- Any treatment-emergent AE leading to interruption of study treatment
- Any treatment-emergent AE leading to death
- All deaths

10.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study treatment as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum grade and SOC and PT; and

by Participant number and SOC and PT. Specific adverse events will be counted once for each Participant for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a Participant, the highest severity and level of relationship to investigational product will be reported.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the total active group.

10.2.4 Deaths, Serious Adverse Events, and Adverse Events Leading to Study Treatment Discontinuation

Treatment-emergent serious adverse events (SAEs), TEAEs leading to premature discontinuation of study treatment, and TEAEs leading to death will be summarized by SOC and PT.

Tabular listings will be provided for all deaths, TESAEs, and TEAEs leading to premature discontinuation of study treatment.

10.3 Analysis of Laboratory Data

The clinical laboratory tests defined in the protocol operations manual (e.g., hematology, clinical chemistry, and urinalysis) will be summarized.

Each laboratory variable will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for select laboratory variables, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented for the mean change from baseline within each treatment group. The select laboratory variables include alanine aminotransferase (ALT), aspartate aminotransferase (AST), bilirubin, alkaline phosphatase, hemoglobin, hematocrit, neutrophils, lymphocytes, monocytes, eosinophils, absolute platelet count, blood urea nitrogen (BUN), creatinine and glomerular filtration rate (GFR).

Changes in laboratory parameters will be tabulated using shift tables the by National Cancer Institute (NCI) Common Terminology Criteria (CTC) version 4.03. Shift tables from baseline to worst post-baseline value and baseline to final treatment period value will be presented by treatment group.

Laboratory abnormalities meeting NCI-CTC grade 3 and 4 will be summarized.

Laboratory values will also be categorized as low, normal, or high based on the normal ranges of the laboratory used for each sample. Shift tables from low or normal baseline to high post-baseline, and from normal or high baseline to low post-baseline will be presented by treatment group. This will be repeated for the shift to final treatment period value based on the lab normal ranges.

Laboratory abnormalities will be evaluated based on Potentially Clinically Significant (PCS) criteria (Appendix B). For each laboratory PCS criterion, the number and percentage of participants who have a laboratory value meeting the criteria will be summarized. Listings will be provided to summarize Participant-level laboratory data for participants meeting PCS criteria.

The following criteria will be used to assess for potential hepatotoxicity. The number and percentage of participants meeting each of the following criteria will be summarized by treatment group, and a listing of ALT, AST, bilirubin, and alkaline phosphatase at each time for all participants that met any of the criteria below at any time will be produced.

- ALT > 3 × ULN, > 5 × ULN, > 10 × ULN, > 20 × ULN
- AST > 3 × ULN, > 5 × ULN, > 10 × ULN, > 20 × ULN
- TBL > 1.5 × ULN, > 2 × ULN
- ALT and/or AST > 3 × ULN and bilirubin > 1.5 × ULN
- ALT and/or AST > 3 × ULN and bilirubin > 2 × ULN
- ALT > 3 × ULN and bilirubin > 1.5 × ULN
- ALT > 3 × ULN and bilirubin > 2 × ULN
- Alkaline phosphatase > 1.5 × ULN

10.4 Analysis of Vital Signs and Weight

Vital sign measurements of systolic and diastolic blood pressure, pulse rate, and body temperature will be summarized.

Each vital sign variable and weight will be summarized for all time points (starting with Baseline) with the number of non-missing observations, mean and standard deviation, median, minimum and maximum. Mean change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 95% confidence interval will be presented within each treatment group.

Vital sign variables and weight will be evaluated based on Potentially Clinically Significant (PCS) criteria (Appendix B). For each vital sign PCS criterion, the number and percentage of participants who have a vital sign value meeting the criteria will be summarized. Listings will be provided to summarize Participant-level vital sign and weight data for participants meeting PCS criteria.

10.5 Other Safety Analyses

10.5.1 ECG

The number and percentage of participants with Normal and electrocardiogram (ECG) readings will be summarized at each visit ECG is performed. The number and percentage of participants with baseline Normal or Abnormal ECG and shifted to post-baseline (worst) Abnormal or Normal will be summarized in a shift table.

Descriptive statistics for ECG parameters (heart rate, RR interval, PR interval, QRS duration, QT interval, and QT interval corrected for heart rate [QTc]) at baseline, postbaseline, and changes from baseline values at each assessment time point to the end of study will be presented by treatment group. The QTc will be calculated using both the Bazett and Fridericia corrections (QTcB and QTCF).

The number and percentage of participants with PCS ECG values according to Table B-5 at any time post-baseline will be summarized. Listings will be provided to summarize Participant-level ECG data for participants meeting PCS criteria.

10.5.2 C-SSRS

For C-SSRS, given sufficient data, the number and percentage of participants with suicidal ideation and/or suicidal behavior as recorded on the C-SSRS will be summarized by treatment group. The summary of responses for most severe suicidal ideation and most sever suicidal behavior in the participant's lifetime history, in the past 2 years, in the past 12 months, during the double-blinded treatment period, and during the follow-up period will be presented by each treatment group, active total and overall. A listing will also be prepared that includes all participants with 1 or more affirmative responses.

10.6 Safety Subgroup Analyses

The AE Overview and TEAE by SOC and PT summaries will be provided for the subset of participants taking symptomatic medication for AD at baseline and those who are not.

11.0 Other Analyses

All planned analyses are listed in the previous sections.

12.0 Interim Analyses

One unblinded interim analysis is planned for this study, which will take place when 50% of participants (approximately 30/treatment group) have completed the Week 6 visit (including those who discontinue prior to the Week 6 visit) and the database has been locked for the interim analysis.

The interim analysis will include the following endpoints and analyses:

- Demographics and baseline characteristics
- Primary Endpoint:
    - Change from baseline to Week 6 in ADAS-Cog 14 total score

- Other Endpoints:
  - Change from baseline to Week 6 in ADAS-Cog 14 total score
  - Change from baseline to Week 6 in Cogstate Computerized Battery (CCB) Composite and Individual Scores
- Summary of Safety:
  - AE overview, TEAE by SOC and PT, SAE by SOC and PT
  - Summary of potentially clinically significant laboratory tests and vital sign values

13.0 Overall Type-I Error Control

This is a Phase 2b dose-ranging and hypothesis generating study; therefore, there will be no control of Type I error for testing multiple doses, multiple endpoints, or interim analyses in this study.

14.0 Version History

Table 3. SAP Version History Summary

Version	Date	Summary
1.0	12 July 2023	Initial version
2.0	2 May 2024	Update to clarify the definition of analysis sets. Update to clarify the definition of number of capsules that should have been taken. Update to exclude efficacy data collected after blind broken. Update to replace posterior probability with p-value statistic to efficacy analysis. Added sensitivity analysis to primary endpoint using safety analysis set. Removed weight-by-sex summary analysis. Update to modify C-SSRS analysis. Removed lab/vital sign comparison between treatment groups
3.0	20 September 2024	Updated analysis for vital signs. Added language for conducting additional subgroup analysis given sufficient data.

14.1 Changes to Planned Analyses in the Protocol

There are currently no changes to the planned analyses in the protocol.

15.0 References

1. Carpenter JR, Roger JH, Kenward MG. Analysis of Longitudinal Trials with Protocol Deviation: A Framework for Relevant, Accessible Assumptions, and Inference Via Multiple Imputation. J Biopharm Stat. 2013;23(6):1352-71.
2. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley; 1987.
3. Wang J, Logovinsky V, Hendrix SB, et al. ADCOMS: a composite clinical outcome for prodromal Alzheimer's disease trials. J Neurol Neurosurg Psychiatry. 2016;87(9):993-9.

Appendix A. List of SAP Signatories

Name	Title	Role/Functional Area
[redacted]	[redacted]	Author, Clinical Statistics
		Clinical Statistics
		Statistical Programming
		Medical/Scientific Monitor

Appendix B. Potentially Clinically Significant Criteria for Safety Endpoints

The criteria for Potentially Clinically Significant (PCS) laboratory findings are described in Table B-1, Table B-2, and Table B-3, PCS criteria for vital sign and weight findings are described in Table B-4, and PCS criteria for ECG are described in Table B-5.

Table B-1. Criteria for Potentially Clinically Significant Hematology Values

Hematology	PCS Low	PCS High
White blood cells	$< 3.0 \times 10^9$/L	$> 16.0 \times 10^9$/L
Lymphocytes	$< 0.8 \times 10^9$/L	$> 12.0 \times 10^9$/L
Neutrophils	$< 1.5 \times 10^9$/L	$> 13.5 \times 10^9$/L
Monocytes	N/A	$> 2.5 \times 10^9$/L
Eosinophils	N/A	$> 1.6 \times 10^9$/L
Basophils	N/A	$> 1.6 \times 10^9$/L
Red blood cells	$\leq 3.5 \times 10^{12}$/L	$\geq 6.4 \times 10^{12}$/L
Hemoglobin – Females	≤ 95 g/L	≥ 175 g/L
Hemoglobin – Males	≤ 115 g/L	≥ 190 g/L
Hematocrit – Females	≤ 32%	≥ 54%
Hematocrit – Males	≤ 37%	≥ 60%
Platelets	$\leq 75 \times 10^9$/L	$\geq 700 \times 10^9$/L

N/A = Not applicable; PCS = potentially clinically significant; ULN = upper limit of normal

Note: A post-baseline value must be more extreme than the baseline value to be considered a Potentially Clinically Significant finding.

Values from CTCAE v4.03.

Table B-2. Criteria for Potentially Clinically Significant Chemistry Values

Blood Chemistry	PCS Low	PCS High
Alanine aminotransferase (ALT)	N/A	3 × ULN
Aspartate aminotransferase (AST)	N/A	3 × ULN
Alkaline phosphatase (ALP)	N/A	3 × ULN
Total bilirubin	N/A	> 1.5 × ULN
Blood urea nitrogen (BUN)	N/A	≥ 10.7 mmol/L
Creatinine	N/A	≥ 176.8 umol/L
Sodium	≤ 126 mmol/L	≥ 156 mmol/L
Potassium	≤ 3 mmol/L	≥ 6 mmol/L
Chloride	≤ 90 mmol/L	≥ 118 mmol/L
Bicarbonate	≤ 16 mmol/L	≥ 35 mmol/L
Glucose	≤ 2.2 mmol/L	≥ 9.7 mmol/L
Calcium	≤ 2 mmol/L	≥ 3 mmol/L
Phosphorus	≤ 0.6 mmol/L	≥ 1.7 mmol/L
Albumin	≤ 25 g/L	≥ 625 g/L
Total protein	≤ 45 g/L	≥ 100 g/L

N/A = Not applicable; PCS = potentially clinically significant; ULN = upper limit of normal

Note: A post-baseline value must be more extreme than the baseline value to be considered a Potentially Clinically Significant finding.

Values from CTCAE v4.03.

Table B-3. Criteria for Potentially Clinically Significant Urinalysis Values

Urinalysis	PCS Low	PCS High
Glucose	N/A	≥ ++++ (+4)
Ketones	N/A	≥ ++++ (+4)
Protein	N/A	≥ ++ (+2)

N/A = Not applicable; PCS = potentially clinically significant; ULN = upper limit of normal

Note: A post-baseline value must be more extreme than the baseline value to be considered a Potentially Clinically Significant finding.

Values from CTCAE v4.03.

Table B-4. Criteria for Potentially Clinically Significant Vital Sign Values

Vital Signs	Very Low (VL)	Very High (VH)
Systolic Blood Pressure (mmHg)	< 90 or decreased ≥ 30 from baseline	> 160 or increased ≥ 40 from baseline
Diastolic Blood Pressure (mmHg)	< 50 or decreased ≥ 20 from baseline	> 100 or increased ≥ 30 from baseline
Pulse (bpm)	< 60 or decreased ≥ 30 from baseline	> 100 or increased ≥ 30 from baseline
Temperature (°C)	< 36 or decreased ≥ 1.1 from baseline	> 38 or increased ≥ 1.1 from baseline
Weight (kg)	Decreased ≥ 7% from baseline	Increased > 7% from baseline

bpm = beats per minute

Note: A post-baseline value must be more extreme than the baseline value to be considered a Potentially Clinically Significant finding.

Table B-5. Criteria for Potentially Clinically Significant Electrocardiograms

Parameter	Criteria
QRS duration	≥ 150 msec
PR interval	≥ 250 msec
QTcF interval	> 450 msec, > 480 msec, > 500 msec
QTcF interval	Increase from baseline > 60 msec
QTcF interval	Increase from baseline > 30 msec but ≤ 60 msec
QTcB interval	> 500 msec
QTcB interval	Increase from baseline > 60 msec

Note: A post-baseline value must be more extreme than the baseline value to be considered a Potentially Clinically Significant finding.

Appendix C. CCB Z-Score and Composite Score Calculation Document

1 Summary

This document details the algorithms for the composite scores to be computed as a part of the M23-515 study.

2 Tests Included in the Cogstate Data Transfer and Their Primary Outcome Measures

Cogstate Test	Cognitive Domain Assessed	Primary Outcome Measure	Interpretation of Primary Outcome Score
Detection Test (DET)	Psychomotor Function	Reaction Time	Lower score = better performance
Identification Test (IDN)	Attention	Reaction Time	Lower score = better performance
One Card Learning (OCL)	Visual Learning	Accuracy	Higher score = better performance
One Back Test (ONB)	Working Memory	*[Primary Outcome]* Reaction Time	Lower score = better performance
		[Additional Outcome] Accuracy	Higher score = better performance
International Shopping List Test – Immediate Recall (ISLT)	Verbal Learning	Correct Responses	Higher score = better performance
International Shopping List Test – Delayed Recall (ISRL)	Memory	Correct Responses	Higher score = better performance

3 Calculation of Z-scores Relative to Normative Data

Cogstate data transfer will include standard z-scores relative to normative data. For ONB, there will be two standard z-scores available. The standard z-score will be computed using the ONB *speed* measure whereas the alternate standard z-score will be computed using the ONB *accuracy* measure.

The formula is provided below for reference.

$$z - Score\ (z_{ijt}) = \frac{(x_{ijt} - \bar{x}_{1t})}{\sigma_{1t}} * Multiplicand$$

Where:

t = test indicator

i = indexes subject *i*

j = indexes the *jth* assessment for subject *i*

x = cognitive score

$\bar{x}_{1t}$ = mean performance score of the age-matched normative sample for test t

σ_{1t} = standard deviation of the age-matched normative sample for test t

Multiplicand equals 1 for tests in which a higher score is indicative of better cognitive performance (i.e., OCL, ONB accuracy, ISLT, ISRL) and -1 for tests in which a lower score is indicative of better cognitive performance (i.e., DET, IDN, ONB speed).

4 Global Composite Score

4.1 Tests Included in the Global Composite

- Detection Test (DET)
- Identification Test (IDN)
- One Card Learning (OCL)
- One Back Test (ONB) – Accuracy
- International Shopping List Test – Immediate Recall (ISLT)
- International Shopping List Test – Delayed Recall (ISRL)

Note: Cogstate recommends using ONB – accuracy to compute composite scores for the M23-515 study given its indication (i.e., Alzheimer's Disease)

4.2 Global Composite Score Computation

1) For a given subject at a given assessment, determine if a minimum of 3 valid test scores (of the tests listed in Section 4.1), one representation from each cognitive domain, are available.

    - Attention: minimum of DET or IDN
    - Learning/Working Memory: minimum of OCL or ONB accuracy
    - Episodic Memory: minimum of ISLT (note: ISRL will be invalid automatically if ISLT is invalid)

2) If the condition in Step 1 is not satisfied, set the composite to missing at that assessment. If the condition in Step 1 is satisfied, proceed with the following steps:

3) Calculate the mean of the available z-scores (z-score for each test can be found in the transfer) to compute the Global Composite score for a given subject at a given assessment. *Note: Use Alt Standard Score Z (ALTSTDSZ) for ONB accuracy z-score.*

5 Attention Domain Composite Score

5.1 Tests Included in the Attention Domain Composite

- Detection Test (DET)
- Identification Test (IDN)

5.2 Attention Domain Composite Score Computation

1) For a given subject at a given assessment, determine if valid test scores are available for all tests (of the tests listed in Section 5.1).

2) If the condition in Step 1 is not satisfied, set the composite to missing for that assessment. If the condition in Step 1 is satisfied, proceed with the following steps:

3) Calculate the mean of the z-scores (z-score for each test can be found in the transfer) to compute the Attention Domain Composite score for a given subject at a given assessment.

6 Learning/Working Memory Composite Score

6.1 Tests Included in the Learning/Working Memory Composite

- One Card Learning (OCL)
- One Back Test (ONB) – Accuracy

Note: Cogstate recommends using ONB – accuracy to compute composite scores for the M23-515 study given its indication (i.e., Alzheimer's Disease)

6.2 Learning/Working Memory Composite Score Computation

1) For a given subject at a given assessment, determine if valid test scores are available for all tests (of the tests listed in Section 6.1).

2) If the condition in Step 1 is not satisfied, set the composite to missing for that assessment. If the condition in Step 1 is satisfied, proceed with the following steps:

3) Calculate the mean of the available z-scores (z-score for each test can be found in the transfer) to compute the Learning/Working Memory Composite score for a given subject at a given assessment. *Note: Use Alt Standard Score Z (ALTSTDSZ) for ONB accuracy z-score.*

7 Episodic Memory Domain Composite Score

7.1 Tests Included in the Episodic Memory Domain Composite

- International Shopping List Test – Immediate Recall (ISLT)
- International Shopping List Test – Delayed Recall (ISRL)

7.2 Episodic Domain Composite Score Computation

1) For a given subject at a given assessment, determine if valid test scores are available for all tests (of the tests listed in Section 7.1).

2) If the condition in Step 1 is not satisfied, set the composite to missing for that assessment. If the condition in Step 1 is satisfied, proceed with the following steps:

3) Calculate the mean of the z-scores (z-score for each test can be found in the transfer) to compute the Episodic Memory Domain Composite score for a given subject at a given assessment.

Document Approval

Study M23515 - Statistical Analysis Plan Version 3 - 20Sep2024 (E3 16.1.9)

Version: 1.0 **Date:** 23-Sep-2024 **Company ID:** 20240923-0900f9f6884b1823-1.0-en

Signed by:	Date:	Meaning of Signature:
[redacted]	23-Sep-2024 16:02 UTC	Approver
	23-Sep-2024 02:05 UTC	Approver - Statistics
	20-Sep-2024 15:06 UTC	Approver - Statistics
	20-Sep-2024 14:50 UTC	Author