STATISTICAL ANALYSIS PLAN
PHASE 4

VERSION: Final

DATE OF PLAN:

02-AUG-2020

BASED ON:

Protocol Amendment No. 2 Final Version Date: 14 May 2019

Protocol Amendment No. 1 Final Version Date: 13 September 2018

Original Final Version Date: 18 July 2018

STUDY DRUG:

NBI-98854

PROTOCOL NUMBER:

NBI-98854-TD4002

STUDY TITLE:

A Phase 4, Double-Blind, Placebo-Controlled, Randomized Withdrawal Study to Evaluate the Persistence of Effect and Safety of Valbenazine for the Treatment of Tardive Dyskinesia

SPONSOR:

Neurocrine Biosciences, Inc.
12780 El Camino Real
San Diego, CA 92130

858-617-7600

This study is being conducted in compliance with good clinical practice, including the archiving of essential documents.

SIGNATURE PAGE

This document has been prepared by:

[redacted] [redacted] Date

This document has been reviewed and accepted by:

[redacted] [redacted] Date

[redacted] [redacted] Date

TABLE OF CONTENTS

LIST OF TABLES

1. LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Term
AE	Adverse Event
AIMS	Abnormal Involuntary Movement Scale
ALT	Alanine Aminotransferase
ALP	Alkaline Phosphatase
AST	Aspartate Transaminase
ATC	Anatomical Therapeutic Chemical Classification
BMI	Body Mass Index
BPRS	Brief Psychiatric Rating Scale
BUN	Blood Urea Nitrogen
CI	Confidence Interval
C-SSRS	Columbia-Suicide Severity Rating Scale
eCRF	Electronic Case Report Form
eTMF	Electronic Trial Master File
ECG	Electrocardiogram
EQ-5D-5L	EuroQol 5 Dimensions 5 Levels
ET	Early Termination
GGT	Gamma-glutamyl Transferase
IPD	Important Protocol Deviation
IWRS	Interactive Web Response System
LOCF	Last observation carried forward
LS	Least Squares
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed-effect Model with Repeated Measures
NBI	Neurocrine Bioscience, Inc.
PCS	Potentially Clinically Significant
PE	Persistence of Effect
PT	Preferred Term
PRO	Patient Reported Outcome
QTcF	Fridericia's Correction of QT Interval
SAE	Serious Adverse Event

SAP	Statistical Analysis Plan
SD	Standard Deviation
SDS	Sheehan Disability Scale
SEM	Standard Error of the Mean
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TD	Tardive Dyskinesia
ULN	Upper Limit of Normal
VAS	Visual Analogue Scale
WHO	World Health Organization

2. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the planned analyses and data displays that will be prepared to summarize the data from the Phase 4 study described in Neurocrine Biosciences, Inc. (NBI) Protocol NBI-98854-TD4002.

This SAP was developed in accordance with ICH E9 guidance. All decisions regarding the final analysis, as defined in this SAP document, will be made prior to database lock and unblinding of the study team to the study data. Further information related to study design and methodology can be found in the protocol.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1. Study Objectives

3.1.1. Primary Objective

The primary objective of this clinical study is:

- To evaluate the persistence of effect of valbenazine in subjects with tardive dyskinesia (TD) who receive placebo in a double-blind, randomized withdrawal period following open-label treatment with valbenazine.

3.1.2. Secondary Objectives

The secondary objectives of this clinical study are:

- To evaluate the relationship between subject clinical characteristics and persistence of effect of valbenazine during the double-blind, placebo-controlled treatment period.
- To evaluate the effect of valbenazine on measures of quality of life and disability when administered once daily for up to 16 weeks.
- To evaluate the safety and tolerability of valbenazine administered once daily for up to 16 weeks.

4. STUDY DESIGN

4.1. Summary of Study Design

This is a Phase 4, randomized, double-blind, placebo-controlled study to evaluate the persistence of effect of valbenazine 40 mg and 80 mg. Approximately 120 medically stable male and female subjects with clinical diagnoses of schizophrenia or schizoaffective disorder with neuroleptic-induced TD or mood disorder with neuroleptic-induced TD will be enrolled.

The study includes an initial open-label treatment period for 8 weeks, followed by a double-blind, placebo-controlled treatment period for 8 weeks, for a total of up to 16 weeks of treatment. A final study visit will be conducted at Week 20 or upon early termination.

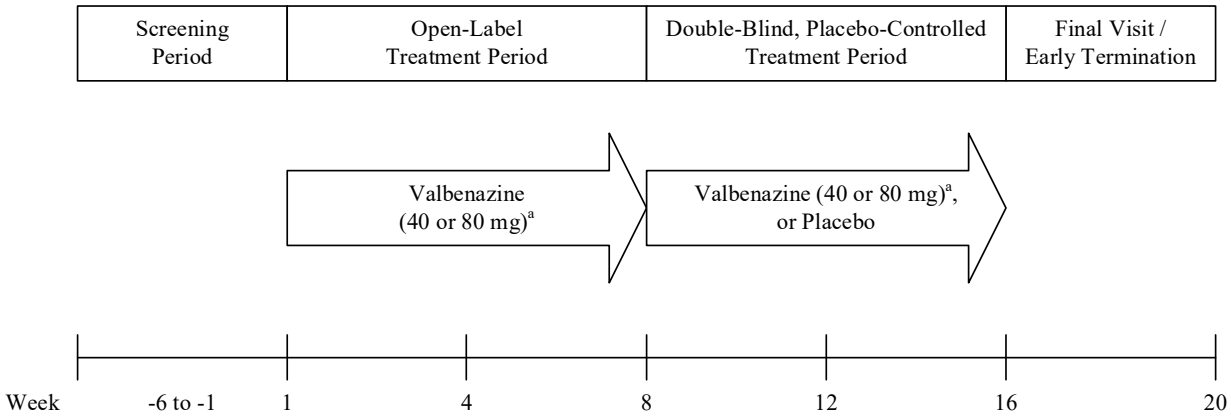
Eligible subjects will be enrolled in the study on Day 1. Valbenazine will be self-administered at home (in the presence of the subject's caregiver, if applicable) beginning on Day 1; the subject will be directed to take their dose at about the same time each day. During the open-label treatment period, subjects will receive 40 mg for the first week followed by 80 mg for 7 weeks.

At the end of Week 8, subjects will be randomized 1:1 to valbenazine or placebo. Subjects randomized to valbenazine will continue with the same dose as the last dose they received during the open-label treatment period. Randomization will be stratified based on the use of concomitant antipsychotic medication; this will include up to 40 subjects who are not using antipsychotic medications (ie, have not used antipsychotic medication for at least 60 days prior to screening and have no plans for resuming antipsychotic treatment during the course of the study) and the remaining subjects will be using concomitant antipsychotic medications.

At any time during treatment (open-label and placebo-controlled treatment periods), subjects who are unable to tolerate the 80 mg dose will have their dose decreased to 40 mg (during the double-blind, placebo-controlled treatment period, this will be done in a blinded manner; subjects receiving placebo will continue to receive placebo). Subjects who are unable to tolerate the 40 mg dose will be discontinued from the study.

A schematic of the study design is shown in Figure 1.

Figure 1: Study Design Schematic



[a] During the open-label treatment period, subjects will receive 40 mg for the first week followed by 80 mg for 7 weeks. At any time during treatment, subjects who are unable to tolerate the 80 mg dose will have their dose decreased to 40 mg. Subjects randomized to valbenazine will continue with the same dose as the last dose they received during the open-label treatment period.

4.2. Sample Size Considerations

Approximately 120 subjects will be enrolled in this study. The 1:1 randomization to placebo or valbenazine during the double-blind, placebo-controlled treatment period will therefore provide a sample size of approximately 60 subjects in the placebo arm. The standard deviation (SD) for the Abnormal Involuntary Movement Scale (AIMS) dyskinesia total score change from the end of the open-label treatment period (end of Week 8) to visits during the double-blind, placebo-controlled treatment period in the placebo treatment arm is estimated to be 4.0 based on results from the previously reported Phase 3 study NBI-98854-1304. With this SD and a sample size of 60, the overall width of a two-sided 95% confidence interval (CI) for the AIMS dyskinesia total score mean change will be approximately 2.0. The width of the confidence interval will increase to approximately 2.2 if the placebo treatment arm sample size is 50 and to approximately 2.5 if the sample size is 40.

4.3. Randomization

At the end of Week 8, subjects will be randomized 1:1 to valbenazine or placebo. Subjects randomized to valbenazine will continue with the same dose as the last dose they received during the open-label treatment period. Randomization will be stratified based on the use of concomitant antipsychotic medication (yes vs. no).

4.4. Clinical Assessments

Assessments of persistence of effect after randomization will be based on:

- AIMS
- Adverse events (AEs) of TD

- Discontinuation from study during the double-blind, placebo-controlled treatment period due to lack of efficacy

Health-related quality of life and disability assessments include:

- EuroQol 5 Dimensions 5 Levels (EQ-5D-5L)
- Sheehan Disability Scale (SDS)

Safety assessments include:

- AEs
- Clinical laboratory tests (hematology, clinical chemistry, and urinalysis)
- Vital signs (including orthostatic blood pressures and pulse)
- Physical examinations
- 12-lead electrocardiograms (ECGs)
- Suicidal ideation and behavior – evaluated using the Columbia-Suicide Severity Rating Scale (C-SSRS)
- Brief Psychiatric Rating Scale (BPRS)

Blood samples for plasma valbenazine and metabolite concentration analyses are also collected. Subjects/caregivers will be asked to record and provide dosing times from the evening before the treatment period visits when these samples are collected.

The schedule of assessments can be found in the protocol.

5. PLANNED ANALYSES

5.1. Interim Analyses

An interim analysis is not planned for this study.

5.2. Final Analyses

Final analyses, as specified in the protocol and in this SAP, will be performed after the study database has been locked and treatment code has been unblinded.

6. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING

All analyses described in this plan are considered *a priori* analyses in that they have been defined prior to locking the study database and unblinding the treatment group assignments. Analyses defined subsequent to locking the database and unblinding will be considered *post hoc* analyses and will be applied as exploratory methodology. Any *post hoc* analyses will be clearly identified in the clinical study report.

6.1. General Statistical Procedures

Descriptive statistics will be used to summarize the data from this study. The term "descriptive statistics" refers to the number of subjects, mean, median, SD, standard error of the mean (SEM), minimum, and maximum for numerical variables; and refers to the number and/or percentage of subjects for categorical variables. Two-sided 95% confidence intervals will be presented for selected variables. Additional descriptive statistics may be presented for selected variables.

Summary statistics will be displayed using the following decimal precision rules: the minimum and maximum will have the same number of decimal places as the data; the mean, median, SD, and SEM will have one more decimal place than the data being summarized. Percentages will be displayed using one decimal place; percentages for 0 counts will be omitted. These rules may be modified if warranted, based on practical considerations.

Some analyses may be combined in the same summary table. Graphical displays of the analyses described in this SAP may also be produced.

6.2. Analysis Sets

6.2.1. Definition of Analysis Sets

For purposes of defining analysis sets, "enrolled subjects" refers to subjects enrolled into the study at the Day 1 visit (ie, are not screen failures) as specified by the subject enrollment electronic case report form (eCRF).

6.2.1.1. Safety Analysis Set

The safety analysis set will include all enrolled subjects who receive at least one dose of open-label study drug and have any safety data collected after the first dose of study drug. The safety analysis set will be used for summaries of safety data during the open-label period and for all summaries of plasma concentration data. A single treatment group (Open-label Valbenazine) will generally be used for analyses of the open-label period.

6.2.1.2. Randomized Safety Analysis Set

The randomized safety analysis set will include all randomized subjects who receive at least one dose of randomized study drug and have any safety data collected after the first dose of randomized study drug. The randomized safety analysis set will be used for all summaries of safety data during the placebo-controlled period. Data from the preceding open-label period may be included in by-visit summaries. Subjects will be grouped by randomized treatment.

6.2.1.3. Persistence of Effect Analysis Set

The persistence of effect (PE) analysis set will include all subjects who are randomized to a treatment group at Week 8, take at least one dose of randomized study drug and have at least one value for change from randomization baseline in the AIMS total score during the placebo-controlled period. Treatment assignment for all summaries and analyses using the PE analysis set will be based on the randomization schedule.

6.2.1.4. Patient Reported Outcome Analysis Set

The patient reported outcome (PRO) analysis set will include all enrolled subjects who take at least one dose of study drug and have at least one health-related quality of life and/or disability assessment collected after baseline. Descriptions of the analysis groups are provided in Section 10.1.

6.3. Baseline Definition

Two definitions of baseline will be used:

Study baseline: The assessments collected at the Day 1 study visit will serve as the study baseline value for all assessments. If a value is not available at the Day 1 visit, then the last measurement collected on or prior to the date of the first dose of open-label study drug will serve as study baseline.

Randomization baseline: The assessments collected at the randomization visit (Week 8) will serve as the randomization baseline value for select assessments. If a value is not available at the randomization visit, then the last measurement collected after the first dose of open-label study drug and on or prior to the date of the first dose of randomized study drug will serve as randomization baseline.

6.4. Derived and Transformed Data

6.4.1. Study Day

Study day is calculated relative to the date of the Day 1 visit. If the date of interest occurs on or after the Day 1 visit, then the study day will be calculated as: date of interest – date of Day 1 visit + 1. If the date of interest occurs prior to the Day 1 visit, then the study day will be calculated as: date of interest – date of Day 1 visit.

6.4.2. Change from Baseline

Change from baseline is calculated as (postbaseline value – baseline value).

Percent change from baseline is calculated as (change from baseline/baseline value * 100).

If either the baseline or postbaseline value is missing, the change from baseline and the percent change from baseline will also be missing. The percent change from baseline will also be missing if the baseline value is equal to zero.

6.4.3. Handling of Early Termination Visit Data

An early termination (ET) visit occurs when a subject discontinues from the study prior to completing the scheduled Week 20 visit. The data collected at ET visits will be included in summary tables and figures in accordance with the ET visit mapping scheme described in this section.

An ET visit will be mapped to the next scheduled study visit if it occurs within 14 days prior to and 13 days after the expected study day of the next scheduled visit (with the requirement that the scheduled visit prior to the ET visit was actually completed by the subject). ET visits occurring after the Week 16 visit will be mapped to the Week 20 follow-up visit if it occurs after day 125.

Early termination visit data which are not mapped to a scheduled visit will not be included in by-visit analyses and summaries. They will be included in any analyses that look across all available assessments during the treatment period, including unscheduled visits. They will also be included in any applicable by-subject data listings.

Table 2 displays the allowable study day range for each scheduled visit for ET visit mapping purposes.

Table 2: Allowable Study Day Range for Early Termination Visit Mapping

Scheduled Visit	Target Study Day	Visit Window (Study Day Range)
Week 4	28	14-41
Week 8	56	42-69
Week 12	84	70-97
Week 16	112	98-125
Week 20	140	>125

6.5. Handling of Missing Data

6.5.1. Missing Outcome Data

Unless otherwise specified, missing values will not be imputed for descriptive statistics.

For persistence of effect summaries based on the AIMS dyskinesia total score, the missing values for change from randomization baseline will be handled in a linear mixed-effect model with repeated measures (MMRM), where the values are assumed to be missing at random.

Missing data for other selected endpoints will be imputed using last observation carried forward (LOCF).

6.5.2. Missing Dates

6.5.2.1. First and Last Dose Dates

Missing and incomplete (“partial”) dates for first and last dose dates will be imputed for the purpose of estimating exposure and defining treatment periods. Missing dates will not be

imputed for subjects when the subject is known to have not taken at least one dose of study drug, as documented by the site in the dosing eCRF.

The imputation rules for first open-label dose date are as follows:

- If the date is completely missing or if both the day and month are missing, the date will be imputed as the date when the first study drug kit was assigned;
- If only the day is missing, the date will be imputed as the date when the first study drug kit was assigned if the month and year match the month and year of the kit assignment date; if the month and year occur after the kit assignment date, the missing day will be imputed as the first day of the month.

A similar logic will be used for the first dose of randomized study drug, using the kit dispensed at the time of randomization.

If the date of the last dose of study drug is missing, then the last dose date will be imputed as the earliest of:

- the Week 16 visit date,
- study discontinuation date for subjects who discontinue before Week 16,
- the date when the last study drug kit(s) was assigned + the number of doses dispensed with the kit(s).

6.5.2.2. Missing Start Dates for Adverse Events and Prior and Concomitant Medications

Missing and incomplete dates for AEs and concomitant medications will be imputed for the purpose of estimating the time of the event or medication usage in relationship to study treatment. Any data listings will display the original dates as reported in the database.

The imputation rules for AE start dates are as follows:

- If the date is completely missing, the date will be imputed as the date of the first dose of study drug;
- If only the day is missing, the date will be imputed as the date of the first dose of study drug if the month and year of the AE start date match the month and year of the first dose of study drug; otherwise, the missing day will be imputed as the first day of the month;
- If both the day and month are missing, the date will be imputed as the date of the first dose of study drug if the AE start date is in the same year as the first dose of study drug; otherwise, the missing day and month will be imputed as 01 January;
- If any of the above imputations result in a start date that is later than an existing (not imputed) end date for the event, the start date will be imputed as the end date.

The imputation rules for concomitant medication start dates are as follows:

- If the date is completely missing, the date will be imputed as 01 January in the year of the subject's screening vital signs assessment;
- If only the day is missing, the date will be imputed as the date of the first dose of study drug if the concomitant medication start date is in the same month and year as the first

dose of study drug; otherwise, the missing day will be imputed as the first day of the month;

- If both the day and month are missing, the date will be imputed as the date of the first dose of study drug if the concomitant medication start date is in the same year as the first dose of study drug; otherwise, the missing day and month will be imputed as 01 January;
- If any of the above imputations result in a start date which is later than an existing (not imputed) medication stop date, the start date will be imputed as the stop date.

7. STUDY POPULATION

A summary of the number and percentage of subjects included in (and excluded from, as applicable) each analysis set (Section 6.2.1) will be provided. The number and percentage of subjects excluded from each analysis set by reason for exclusion will also be provided.

A summary of subject disposition will be prepared that displays the number of subjects who were enrolled, who received at least one dose of study drug, who were randomized to each treatment group, who were randomized and completed the placebo-controlled withdrawal period, who were randomized and completed the follow-up period, and who were not randomized and discontinued study participation. The number of subjects who discontinued from the study will also be displayed by reason for discontinuation.

The analysis set and subject disposition summaries will be based on all enrolled subjects and will include the following analysis groups:

- subjects enrolled but not randomized;
- subjects randomized to placebo;
- subjects randomized to valbenazine;
- all subjects.

A listing of randomized subjects will also be provided and will include subject ID, informed consent date, enrollment date, randomization date, concomitant antipsychotic use (yes vs. no), and randomized treatment group.

A separate summary of randomization by study site will be presented. This summary will display the number of subjects randomized to each treatment group by site.

7.1. Protocol Deviations

Protocol deviations described in the study-specific Protocol Deviation Plan will be entered into the study electronic Trial Master File (eTMF) system and used to identify important protocol deviations (IPDs). IPDs are protocol deviations that might significantly affect the completeness, accuracy and/or reliability of the study data or that might significantly affect a subject's rights, safety, or well-being. An assessment of IPDs will be performed by a committee composed of NBI Clinical Development project team members prior to database lock. This committee will review a listing of all protocol deviations reported in the study database and determine which deviations are IPDs.

A summary of the number and percentage of subjects with IPDs by deviation category will be provided using all enrolled subjects. The following analysis groups will be used: subjects enrolled but not randomized; subjects randomized to placebo; subjects randomized to valbenazine; all subjects.

All major protocol deviations entered into the study eTMF will be presented in a data listing. Any IPDs will be flagged in the listing.

7.2. Demographic and Baseline Characteristics

Demographic and baseline characteristics data will be summarized using descriptive statistics for continuous variables, and frequency counts and percentages for categorical variables. Summaries will be presented for each of the analysis sets described in Section 6.2.1. An additional "All Subjects" column will be included, where appropriate.

Demographics include:

- Age (years)
- Sex
- Ethnicity
- Race

Study baseline subject characteristics include:

- Primary clinical diagnosis (schizophrenia or schizoaffective disorder with neuroleptic-induced TD; mood disorder with neuroleptic-induced TD)
- Age at first diagnosis of schizophrenia, schizoaffective disorder, or mood disorder (years)
- Age at TD diagnosis (years)
- Height (measured at screening; cm)
- Weight (presented in both pounds and kilograms)
- Body mass index (BMI; calculated using height collected at the screening visit and baseline weight; kg/m^2)
- CYP2D6 genotype status
- AIMS dyskinesia total score
- Concomitant antipsychotic use (yes vs. no; as indicated by the investigator in IWRS and included for randomized analysis groups only)

7.3. Medical History and Medical Conditions Present at Entry

Medical history will be summarized in frequency tables (number and percentage of subjects) by Medical Dictionary for Regulatory Activities (MedDRA) System Organ Class (SOC) and Preferred Term (PT), with SOCs and PTs within each SOC sorted alphabetically. Separate summaries will be presented for the safety analysis set and the randomized safety analysis set. The randomized safety analysis set will be presented by treatment and will include an additional "All Subjects" column.

7.4. Study Drug Dosing, Compliance, and Dose Reductions

7.4.1. Compliance

Subjects will bring all unused study drug and empty study drug packaging material to the center at each study visit for drug accountability and reconciliation by study center personnel. A

compliance check will be performed by counting the capsules returned at each study visit. The site will then enter whether the subject's dosing compliance since the previous visit was ≥80% into the eCRF.

The number and percentage of subjects who are dosing compliant at each visit will be summarized using the safety analysis set (through Week 8) and the randomized safety analysis set (Week 8 through Week 16).

7.4.2. Dose Reductions

Subjects will receive 40 mg for the first week of treatment and 80 mg until the end of Week 8 in subjects randomized to receive placebo and until the end of Week 16 in subjects randomized to remain on valbenazine. If a subject is unable to tolerate the 80 mg dose, the daily dosage will be reduced to 40 mg. Subjects unable to tolerate the 40 mg dose (or placebo) will be discontinued from the study.

The number and percentage of subjects with a dose reduction at any time during the open-label treatment period will be presented using the safety analysis set. The number and percentage of subjects with a dose reduction at any time during the placebo-controlled treatment period will be presented by treatment group using the randomized safety analysis set.

8. PLASMA CONCENTRATION DATA

The plasma concentrations of valbenazine (NBI-98854) and its metabolite NBI-98782 will be summarized with descriptive statistics by visit and by the last valbenazine dose (40 mg or 80 mg) received prior to that visit. Concentrations below the lower limit of quantification will be set equal to zero for all plasma concentration summaries. The lower limits of quantification are as follows: (a) NBI-98854: 1.00 ng/mL and (b) NBI-98782: 0.100 ng/mL.

In addition to the descriptive statistics noted in Section 6.1, the following descriptive statistics will be included in the plasma concentration summary tables: (a) the number of plasma concentration values greater than or equal to the lower limit of quantification, (b) the geometric mean, and (c) the geometric coefficient of variation (%).

The safety analysis set will be used for all plasma concentration summaries. Data for subjects receiving placebo at the visit will not be included.

9. PERSISTENCE OF EFFECT OF VALBENAZINE

9.1. General Considerations

Unless otherwise specified, the PE analysis set will be used for all analyses described in this section. Results will be displayed by randomized treatment group (Randomized Valbenazine vs Randomized Placebo).

9.2. Abnormal Involuntary Movement Scale

The severity of TD will be assessed using the AIMS. The AIMS examination will be administered by the investigator (or designee) in accordance with the AIMS administration procedure. The AIMS administration will be video recorded (approximately 10 minutes) following standardized guidelines. The investigator (or designee) will score items 8-10 and items 11-12 (binary dental status items). Items 1-7 will be scored by two blinded Central AIMS Video Raters using consensus scoring.

9.2.1. AIMS Individual Items

The score for Items 1 through 7 ranges from 0 (no dyskinesia) to 4 (severe dyskinesia) and includes facial and oral movements (Items 1 to 4), extremity movements (Items 5 to 6), and trunk movements (Item 7). Items 8, 9 and 10 rate global judgments: Items 8 (severity of abnormal movements) and 9 (incapacitation due to abnormal movements) range from 0 (none, normal) to 4 (severe) and Item 10 being scored based only on the subject's report of his/her awareness of abnormal movements from 0 (no awareness) to 4 (aware, severe distress). Items 11 and 12 are yes/no questions concerning problems with teeth and/or dentures.

Descriptive statistics of the individual AIMS items 1 through 10 will be presented for study baseline, randomization baseline and for each study visit occurring thereafter. Change from randomization baseline will be presented for the post-randomization visits.

9.2.2. AIMS Dyskinesia Total Score

The AIMS dyskinesia total score is defined as the sum of the scores of AIMS items 1 through 7. If any of the seven items have a missing value, the total score for that subject/visit will be set equal to missing. The AIMS dyskinesia total score can therefore range from 0 to 28, with higher scores indicating greater severity.

The change from randomization baseline in AIMS dyskinesia total score during the placebo-controlled treatment period will be analyzed using a linear mixed-effect model of repeated measures (MMRM) that includes the fixed effects of treatment, visit, treatment-by-visit, concomitant antipsychotic use (yes vs. no), and study baseline AIMS dyskinesia total score as a covariate. Within-subject variability will be accounted for using a random effect with an unstructured covariance matrix. The least squares (LS) mean for each treatment group and the associated 95% confidence interval (CI) will be presented for each visit. The primary measurement of persistence of effect will be based off the CI for the LS mean change from randomization baseline in the placebo group at Week 16.

Descriptive statistics of the AIMS dyskinesia total score will be presented for study baseline, randomization baseline and for each study visit occurring thereafter. Descriptive statistics for

change from randomization baseline and change from study baseline will also be presented for the relevant postbaseline visits. The descriptive statistics for the change from randomization baseline will include a 2-sided 95% confidence interval for the mean.

Mean (±SEM) values of the observed values at each visit will be summarized in a line graph by treatment group. A similar graph will be presented for the changes from randomization baseline for the post-randomization visit. The LS means (±SEM) from the MMRM analysis results will be summarized in a similar line graph.

9.2.3. AIMS Responder Analysis

An AIMS responder is defined, on a per-visit basis, as a subject whose AIMS dyskinesia total score is reduced by at least 50% from study baseline. Descriptive statistics will be presented by visit and treatment group for the number and percentage of subjects classified as AIMS responders.

9.3. AEs of TD and Discontinuations

Additional measures of persistence of effect that will be summarized include AEs of TD and early discontinuations from the study due to lack of efficacy during the double-blind, placebo-controlled treatment period.

The summary will include the number and percentage of subjects who:

1) experience an AE of TD during the placebo-controlled treatment period;
2) discontinue from the study due to lack of efficacy during the placebo-controlled treatment period;
3) meet either of the previous criteria.

An AE of TD is defined as AEs with a MedDRA preferred term of "Tardive dyskinesia" or "Dyskinesia".

9.4. Subgroup Analyses

The relationship between subject clinical characteristics and the persistence of effect of valbenazine during the double-blind, placebo-controlled treatment period will be evaluated.

The following clinical characteristics will be investigated:

- Primary psychiatric diagnosis (schizophrenia or schizoaffective disorder vs. mood disorder)
- Age group (<45 years vs. ≥45 years; <65 years vs. ≥65 years)
- Sex (male vs. female)
- Concomitant use of antipsychotic medications (yes vs. no, as indicated by investigator in IWRS)
- Treatment response status (yes vs. no, as indicated by AIMS Responder status [Section 9.2.3] at randomization baseline)

Each subgroup factor along with its 2-way and 3-way interactions with treatment group and visit will be added to the MMRM model specified for the AIMS dyskinesia total score in Section 9.2.2. The LS mean for each combination of treatment group and subgroup stratum and the associated 95% CIs will be presented by visit. Descriptive statistics for the observed AIMS dyskinesia total scores and changes from randomization will also be presented by visit and treatment for each stratum of the subgroups. The MMRM results will not be presented if a subgroup has <10 subjects in any treatment and subgroup stratum combination or if the model fails to converge.

If needed to investigate a potential signal, descriptive statistics of the individual AIMS items or other assessments of persistence of effect may also be presented by subgroup.

10. HEALTH-RELATED QUALITY OF LIFE AND DISABILITY

10.1. General Considerations

The PRO analysis set will be used for all analyses described in this section. Study baseline will be used as baseline for all assessments. By-visit summaries will include the following analysis groups:

- Open-Label Valbenazine – all subjects in the PRO analysis set. Data will be presented through Week 8.
- Randomized Valbenazine – all subjects in the PRO analysis set who were randomized to valbenazine at Week 8.
- Randomized Placebo – all subjects in the PRO analysis set who were randomized to placebo at Week 8.

In addition to observed summaries, missing data at Week 8 and Week 16 will be imputed using LOCF:

- Week 8: last postbaseline assessment collected during the open-label treatment period
- Week 16: last post-randomization collected during the placebo-controlled treatment period

10.2. EuroQol 5 Dimensions 5 Levels

The EQ-5D-5L is a general, single index measure for describing and valuing health (Herdman et al, 2011).

The first part assesses health in terms of 5 dimensions: Mobility, Self-Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The subject indicates his/her health state by checking the box next to the most appropriate statement. The scores for the 5 dimensions can be combined into a 5-digit number that describes the patient's health state. The health state index score will be calculated from the individual health profiles using the United States value set (Pickard et al, 2019) and ranges from -0.573 to 1.0, with higher scores indicating higher health utility.

The second part of the questionnaire consists of a visual analogue scale (VAS) on which the patient rates his/her perceived health from 0 (the worst imaginable health) to 100 (the best imaginable health).

The EQ-5D-5L health state index and VAS scores will be summarized with descriptive statistics at baseline and at each postbaseline visit. Both observed values and changes from baseline will be summarized. Frequency counts and percentages for the responses of the 5 individual dimensions will also be presented at baseline and at each scheduled postbaseline visit.

10.3. Sheehan Disability Scale

The SDS is a brief, validated measure of functional impairment in a number of psychiatric disorders to measure the effect of treatment on disability (Leon et al, 1997). It includes 3 self-

rated items designed to measure how work, social life, and family life are impaired by current psychiatric symptoms. Each item includes an 11-point analog scale that uses visual-spatial, numeric, and verbal descriptive anchors to represent the degree of disruption from 0 (none at all) to 10 (extremely). It also assesses the number of days a subject was unable to work/attend school and the number of days a subject was underproductive in the past week.

The SDS total score is the sum of the 3 impairment items and will only be calculated for subjects who rate all three items.

The SDS individual items and total scores will be summarized with descriptive statistics at baseline and at each postbaseline visit. Both observed values and changes from baseline will be summarized.

11. SAFETY AND TOLERABILITY

11.1. General Considerations

11.1.1. Format for By-Visit Summaries

Study baseline will be used as baseline for all safety assessments. Descriptive statistics of observed and change from baseline values will be presented at baseline and at each scheduled postbaseline visit through Week 20. Results will be displayed by the following analysis groups:

- Open-Label Valbenazine – all subjects in the safety analysis set. Data will be presented through Week 8.
- Randomized Valbenazine – all subjects in the randomized safety analysis set who were randomized to valbenzine.
- Randomized Placebo – all subjects in the randomized safety analysis set who were randomized to placebo.

11.1.2. Format for Adverse Events and Other Event-Based Summaries

Adverse events and assessment-based abnormality summaries, such as potentially clinically significant (PCS) laboratory values or "Yes" responses in the C-SSRS items, will be displayed by treatment period as followed:

Open-Label Treatment Period: includes AEs and events that begin after the first dose of study drug, excluding any events with a start date or assessment date after the first dose of randomized study drug. Results will be displayed in a single treatment group (Open-Label Valbenazine) using the safety analysis set.

Placebo-Controlled Treatment Period: includes AEs and events that begin after the first dose of randomized study drug. Results will be displayed by randomized treatment group (Randomized Placebo or Randomized Valbenazine) using the randomized safety analysis set.

11.2. Adverse Events

Adverse events are recorded in the eCRF. Each AE will be coded to SOC and PT using MedDRA (Version 21.0).

A treatment-emergent adverse event (TEAE) is an AE not present prior to the initiation of study drug dosing or is an already present event that worsens either in intensity or frequency following the initiation of study drug dosing. Investigators will be asked to respond “Yes” or “No” on the eCRF as to whether the AE started after the subject took the first dose of study drug. An AE with a response of “Yes” will be classified as a TEAE. If the investigator’s response is missing, then the treatment-emergent status will be derived based on the AE onset date relative to the date of the subject’s first dose of study drug. If the AE onset date is unknown, it will be assumed that the AE is a TEAE. TEAEs will be assigned to a treatment period based on the onset date. Missing or incomplete AE onset dates or study drug dosing dates will be imputed as described in Section 6.5.2.

The frequency tables will include the number and percentage of unique subjects experiencing each event at least once during the specified treatment period. Separate summaries will be presented for each treatment period defined in Section 11.1.2.

Two versions of the primary TEAE frequency tables will be presented:

- Frequency of TEAEs by SOC and PT, with SOCs and PTs within each SOC sorted by decreasing frequency (number of unique subjects) in the Valbenazine treatment group.
- Frequency of TEAEs by PT, with PT sorted by decreasing frequency (number of unique subjects) in the Valbenazine treatment group.

The number and percentage of subjects with severe TEAEs will also be summarized. These tables will include both SOC and PT, sorted in the same method as the primary TEAE table. The first line of the table will display the number and percentage of subjects with at least one severe TEAE.

AE overview summary tables will be provided for each treatment period which summarize the number and percentage of subjects with any TEAE, any TEAE leading to dose reduction, any TEAE leading to study discontinuation, any serious TEAE, and any TEAE resulting in death. The summary table will also include the maximum TEAE intensity (mild, moderate, severe) reported for each subject during the specified treatment period.

11.2.1. Adverse Events Resulting in Premature Discontinuation from Study

Summary tables of TEAEs resulting in premature discontinuation from study will be presented. The number and percentage of subjects with a TEAE resulting in study discontinuation will be presented by PT within SOC (presented in the same method as the primary TEAE table). More than one AE can contribute to study discontinuation per subject. The first line of the table will display the number and percentage of subjects with at least one TEAE leading to study discontinuation.

A listing of TEAEs resulting in premature study discontinuation will be provided which includes subject ID, last study drug dose level received prior to the onset date of the TEAE(s) leading to discontinuation, study day of the discontinuation, and other relevant information from the AE eCRF or subject characteristics.

11.2.2. Adverse Events Resulting in Study Drug Dose Reductions

Summary tables of TEAEs resulting in study drug dose reductions will be presented. The number and percentage of subjects with a TEAE resulting in a dose reduction will be presented by PT within SOC (presented in the same method as the primary TEAE table). More than one AE can contribute to a dose reduction per subject. The first line of the table will display the number and percentage of subjects with at least one TEAE leading to dose reduction.

11.2.3. Deaths and Other Serious Adverse Events

Summary tables of serious adverse events (SAEs) will be presented. The tables will include the frequency of SAEs presented by PT within SOC (presented in the same method as the primary TEAE table).

Separate listings of SAEs and fatal TEAEs will also be provided. Each listing will include subject ID, last study drug dose level received prior to the onset date of the SAE or fatal TEAE, study day of the SAE or fatal TEAE, and any additional relevant information from the AE eCRF or subject characteristics.

11.3. Clinical Laboratory Data

By-visit summaries

The clinical chemistry and hematology data will be summarized with descriptive statistics at baseline and at each postbaseline visit as described in Section 11.1.1. Both observed values and changes from baseline will be summarized.

Shift tables

Shift tables will be presented for selected clinical laboratory variables based on the reference range-based categories of "Low," "Normal," or "High." A clinical laboratory variable value will be assigned to one of these three categories according to the reference ranges provided by the central clinical laboratory.

Shift tables will be presented by treatment period:

- Shifts from baseline to Week 8 (or last available assessment in the open-label treatment period) using the safety analysis set.
- Shifts from baseline to Week 16 (or last available assessment in the placebo-controlled period) using the randomized safety analysis set.

Each shift table will have three rows and three columns, with rows reflecting the reference range category at baseline, and columns reflecting the reference range at postbaseline. A "Total" row and "Total" column will also be included. Subjects with a missing baseline value or who do not have postbaseline data will not be included in the tables for that variable. The number and percentage of subjects in each shift category will be displayed in the table; percentages will be based on the number of subjects included in the table.

Shift tables will be presented for the following clinical laboratory variables: aspartate transaminase (AST), alanine transaminase (ALT), alkaline phosphatase (ALP), gamma-glutamyl transferase (GGT), total bilirubin, creatine kinase, creatinine, blood urea nitrogen (BUN), white blood cell count, absolute neutrophil count, hemoglobin, and platelet count.

PCS values

Summaries of sponsor-defined PCS values will be presented for selected laboratory variables. The analytes and criteria for identifying PCS laboratory values are provided in Table 3.

The number and percentage of subjects with PCS laboratory values that are reported at any postbaseline visit (scheduled or unscheduled, including repeat values) will be summarized. Results will be presented for each of the treatment periods described in Section 11.1.2.

Table 3: Potentially Clinically Significant Criteria for Clinical Laboratory Variables

Variable	PCS Threshold
ALT	>3 x ULN (upper limit of normal)
AST	>3 x ULN
Creatine kinase	>5 x ULN
GGT	>3 x ULN
Total bilirubin	>1.5 x ULN
White blood cell count	≤2.8 x 1000/µL
Absolute neutrophil count	<1.5 x 1000/µL
Creatinine	>1.5 x baseline value or > 1.5 x ULN
BUN	>30 mg/dL (> 10.71 mmol/L)

11.4. Vital Signs

By-visit summaries

The vital signs data, including orthostatic blood pressures and heart rate (calculated as standing value minus supine value), will be summarized with descriptive statistics at baseline and at each postbaseline visit as described in Section 11.1.1. Both observed values and changes from baseline will be summarized.

PCS values

Summaries of sponsor-defined PCS values will be presented for systolic blood pressure, diastolic blood pressure, and heart rate. The criteria for identifying PCS vital signs values are provided in Table 4. Both supine and standing values will be included in the identification and summary of PCS values.

The number and percentage of subjects with PCS vital signs values that are reported at any postbaseline visit (scheduled or unscheduled, including repeat values) will be summarized. Results will be presented for each of the treatment periods described in Section 11.1.2.

Table 4: Potentially Clinically Significant Criteria for Vital Signs Variables

Variable Name	PCS – Low if:			PCS – High if:		
	Observed Value is:	AND	Decrease from Baseline is:	Observed Value is:	AND	Increase from Baseline is:
Systolic Blood Pressure	<90 mmHg		≥20 mmHg	>180 mmHg		≥20 mmHg
Diastolic Blood Pressure	<50 mmHg		≥10 mmHg	>105 mmHg		≥15 mmHg
Heart Rate	<50 bpm		≥15 bpm	>120 bpm		≥15 bpm

11.5. Body Weight

By-visit summaries

The body weight data (in units of kilograms) will be summarized with descriptive statistics at baseline and at each postbaseline visit as described in Section 11.1.1. Both observed values and changes from baseline will be summarized.

11.6. Electrocardiogram

The triplicate values of the quantitative ECG variables (heart rate, PR interval, QRS duration, QT interval, and Fridericia's correction of QT interval [QTcF]) measured at each visit will be averaged for the purpose of analysis. For the categorical ECG interpretation variable (the investigator's assessment of the ECG as "Normal", "Abnormal, not Clinically Significant", or "Abnormal, Clinically Significant"), which is also reported for each replicate, the value that represents the greatest degree of abnormality will be used in all summary tables. If less than 3 values are recorded at an assessment, then the average of the available value(s) for the quantitative variables and the greatest degree of abnormality values(s) for the interpretation variable will be used.

The quantitative ECG variables will be summarized with descriptive statistics at baseline and at each postbaseline visit as described in Section 11.1.1. Both observed values and changes from baseline will be summarized. Frequency counts and percentages for the ECG interpretation variable categories will also be presented at baseline and at each scheduled postbaseline visit.

Categorical summaries will be presented for the QT and QTcF interval data. For these summaries, a subject's highest reported postbaseline value (including values reported at unscheduled visits) will be used to determine in which category(s) the subject will be counted. The averaged triplicate values will be used when determining each subject's highest reported values. Results will be presented for each of the treatment periods described in Section 11.1.2 and may include post-treatment assessments.

For the first summary, the number and percentage of subjects whose highest reported QT or QTcF postbaseline value meets the following thresholds will be summarized:

- Greater than 450 msec

- Greater than 480 msec
- Greater than 500 msec

The second categorical summary will display the number and percentage of subjects whose largest QT or QTcF increase from their baseline value meets the following thresholds:

- Greater than 30 msec
- Greater than 60 msec

11.7. Columbia-Suicide Severity Rating Scale

The C-SSRS data will be presented in the following summaries:

- Screening/lifetime assessment
- Screening/past 3 month assessment
- Baseline (Day 1) assessment
- Open-Label Treatment Period
- Placebo-Controlled Treatment Period

Definitions for the treatment periods are described in Section 11.1.2, and may include post-treatment assessments.

Each summary will display the number and percentage of subjects who report "Yes" to specific C-SSRS items or categories of items (a category is assigned a "Yes" value if a "Yes" is reported for any item in the category). These C-SSRS items and categories are as follows:

- Suicidal Ideation Items
  - (1) Wish to be dead
  - (2) Non-specific active suicidal thoughts
  - (3) Active suicidal ideation with any methods (not plan) without intent to act
  - (4) Active suicidal ideation with some intent to act, without specific plan
  - (5) Active suicidal ideation with specific plan and intent
- Suicidal Ideation Category: Any of items (1) through (5)
- Suicidal Behavior Items
  - (6) Preparatory acts or behavior
  - (7) Aborted attempt
  - (8) Interrupted attempt
  - (9) Non-fatal suicide attempt
  - (10) Completed suicide
- Suicidal Behavior Category: Any of items (6) through (10)
- Suicidal Ideation or Behavior Category: Any of items (1) through (10)

For the treatment period summaries, each subject's C-SSRS responses for all postbaseline assessments during the specified treatment period will be evaluated, and a "Yes" response at any postbaseline assessment for a particular item or category during the specified treatment period will be considered as a "Yes" for that item or category for the subject.

In addition to the summaries described above, shift tables comparing postbaseline suicidal ideation scores to baseline scores will be presented for each treatment period. The shift table scores are defined as the following:

0 = No suicidal ideation
1 = Wish to be dead
2 = Non-specific active suicidal thoughts
3 = Active suicidal ideation with any methods (not plan) without intent to act
4 = Active suicidal ideation with some intent to act, without specific plan
5 = Active suicidal ideation with specific plan and intent

The shift tables will display the number and percentage of subjects within each cell of a 6 x 6 table, with the rows representing the study baseline score and the columns representing the maximum score recorded across all postbaseline assessments (including both scheduled and unscheduled visits) during the specified treatment period. Subjects missing either a baseline score or all postbaseline scores will not appear in the table.

11.8. Brief Psychiatric Rating Scale

The BPRS is a clinician-rated tool designed to assess the severity of psychopathology in patients with schizophrenia and other psychotic disorders. The BPRS includes 18 items that address somatic concern, anxiety, emotional withdrawal, conceptual disorganization, guilt feelings, tension, mannerisms and posturing, grandiosity, depressive mood, hostility, suspiciousness, hallucinatory behaviors, motor retardation, uncooperativeness, unusual thought content, blunt affect, excitement, and disorientation.

The severity of each of the 18 items of the BPRS is rated on a scale of 1 (not present) to 7 (extremely severe) (total score range: 18 to 126). Higher scores represent greater symptom severity.

The BPRS total score will be summarized with descriptive statistics at baseline and at each postbaseline visit as described in Section 11.1.1. Both observed values and changes from baseline will be summarized.

11.9. Prior and Concomitant Medications

Prior medications and concomitant medications will be summarized by World Health Organization (WHO) Drug Anatomical Therapeutic Chemical Classification (ATC) Level 3 category (or Level 2 if there is not an applicable Level 3 category) and preferred name.

Medications will be assigned to one or more reporting periods based on the medication start and stop dates relative to study drug dosing dates.

- Prestudy/screening: medications with a start date prior to open-label study drug dosing
- During the open-label treatment period or posttreatment: medications ongoing at the time of first open-label study drug dosing or with a start date after first dose of open-label study drug, excluding medications started after the first dose study drug in the placebo-controlled period

- During the placebo-controlled period or posttreatment: medications ongoing at the time of first dose of study drug in the placebo-controlled period or with a start date after the first dose of study drug in the placebo-controlled period.

A given medication can be assigned to multiple study periods in the tabular summaries, depending on its start and end dates.

The number and percentage of subjects using medications in each WHO Drug ATC category (Level 3/preferred name) will be summarized by the reporting periods defined in the previous paragraph. A subject may take the same medication more than once or multiple medications for a subject may be classified under the same ATC level or preferred name. A subject is counted only once for each level of medication classification within a summary.

12. REFERENCES

Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011 Dec;20(10):1727-36.

Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. Int J Psychiatry Med. 1997;27(2):93-105.

Pickard S, Law E, Jiang R, Pullenayegum E, Shaw J, Xie F, Oppe F, Boye K, Chapman R, Gong C, Balch A, Busschbach J. United States Valuation of EQ-5D-5L Health States Using an International Protocol. VALUE HEALTH. 2019; 22(8):931–941