STATISTICAL ANALYSIS PLAN
PHASE 2

VERSION: Final

DATE OF PLAN:

02-Aug-2019

BASED ON:

Protocol Amendment No. 2 Final Version: 03 December 2018
Protocol Amendment No. 1 Final Version: 08 June 2018

Original Version Date: 18 December 2017

STUDY DRUG:

NBI-98854

PROTOCOL NUMBER:

NBI-98854-TS2005

STUDY TITLE:

A Phase 2, Double-Blind, Placebo-Controlled, Randomized Withdrawal Study to Evaluate the Safety and Efficacy of NBI-98854 in Pediatric Subjects with Tourette Syndrome

SPONSOR:

Neurocrine Biosciences, Inc.

2780 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 and/or reviewed by:

[redacted]
Neurocrine Biosciences, Inc

[redacted]
Signature
[redacted]
Date

[redacted]
Neurocrine Biosciences, Inc

[redacted]
Signature
7-Aug-2019
Date

This document has been reviewed and accepted by:

[redacted]
Neurocrine Biosciences, Inc

[redacted]
Signature
07-Aug 2019
Date

[redacted]
Neurocrine Biosciences, Inc

[redacted]
Signature
8/7/19
Date

TABLE OF CONTENTS

LIST OF TABLES

1. LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Term
ADHD	Attention-Deficit Hyperactivity Disorder
ADHD Rating Scale-5	Attention-Deficit Hyperactivity Disorder Rating Scale-5: Home Version
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine transaminase
ANCOVA	Analysis of covariance
AST	Aspartate transaminase
ATC	Anatomical Therapeutic Chemical Classification
BLQ	Below the lower limit of quantification
BMI	Body mass index
BUN	Blood urea nitrogen
C&A-GTS-QOL	Gilles de la Tourette Syndrome – Quality of Life Scale for Children and Adolescents
CDRS-R	Children's Depression Rating Scale - Revised
CGI-S	Clinical Global Impression of Movement Severity
CGI-Tics-Severity	Clinical Global Impression of Tics-Severity scale
CGI-TS-Improvement	Clinical Global Impression of Tourette Syndrome-Improvement scale
CMH	Cochran-Mantel-Haenszel
C-SSRS	Columbia-Suicide Severity Rating Scale
CY-BOCS	Children's Yale-Brown Obsessive Compulsive Scale
DSM-IV or -5	Diagnostic and Statistical Manual of Mental Disorders, 4th or 5th Editions
ECG	Electrocardiogram
eCRF	Electronic case report form
ESRS-A	Extrapyramidal Symptom Rating Scale-Abbreviated
ET	Early termination
FAS	Full analysis set
GGT	Gamma-glutamyl transferase
IPD	Important protocol deviation
IWRS	Interactive Web Response System
KM	Kaplan-Meier

LOCF	Last observation carried forward
LS	Least-squares
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed-effect model repeated measures
NRI	Nonresponder imputation
PCS	Potentially clinically significant
PT	Preferred term
PUTS	Premonitory Urge for Tics Scale
QTcF	Fridericia's correction of QT interval
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SEM	Standard error of the mean
SOC	System organ class
TEAE	Treatment-emergent adverse event
TS	Tourette Syndrome
TTS	Total Tic Score
ULN	Upper limit of normal
VAS	Visual analog scale
WHO	World Health Organization
YGTSS	Yale Global Tic Severity Scale

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 2 study described in Neurocrine Biosciences, Inc. (NBI) Protocol NBI-98854-TS2005.

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 maintenance of efficacy of NBI-98854 during a double-blind, placebo-controlled withdrawal period in pediatric subjects with Tourette Syndrome (TS) who have responded to open label NBI-98854 treatment.

3.1.2. Secondary Objective

The secondary objectives of this study are:

- To evaluate the effect of NBI-98854 on measures of TS symptoms during a placebo-controlled withdrawal period.
- To evaluate the safety and tolerability of NBI-98854 administered once daily for up to 36 weeks.

3.2. Primary Estimand

The primary estimand is the difference between NBI-98854 and placebo in the distribution of Kaplan-Meier Product limit estimates of times to loss of treatment response during the withdrawal portion of this study in pediatric subjects with TS whose tic behaviors are sufficiently controlled and who have tolerated NBI-98854 during up to 12 weeks of open-label dose optimization treatment and who have entered the randomized withdrawal portion of this study. An on-treatment strategy will be used for the intercurrent events of study discontinuation for reasons other than the lack of efficacy or a TEAE of worsening of tics: time to loss of treatment response for subjects who do not have a loss of treatment response prior to discontinuation will be censored at the time of discontinuation.

Additional details regarding the definition of loss of treatment response and analysis methods are provided in Section 9.3.1.

4. STUDY DESIGN

4.1. Summary of Study Design

This is a Phase 2, double-blind, placebo-controlled, randomized withdrawal study to evaluate the safety and maintenance of efficacy of an optimized dose of NBI-98854 in pediatric subjects with TS. The study includes an initial open-label treatment period for 6 weeks, a blinded randomization period for 6 weeks, followed by a double-blind, placebo-controlled withdrawal period for 24 weeks, for a total of up to 36 weeks of treatment. Follow-up assessments will be conducted at the end of Week 40 after a 4-week washout of study drug, or upon early termination. NBI-98854 will be titrated to the subject's optimal dose in the range of 20 mg to 60 mg for subjects <50 kg and 40 mg to 80 mg for subjects ≥50 kg during the initial open-label treatment period.

Approximately 180 male and female pediatric subjects, 6 to 17 years of age, with a Diagnostic and Statistical Manual of Mental Disorders, 4th or 5th Editions (DSM-IV or -5) diagnosis of TS will be enrolled into the 6-week open-label, dose-optimization treatment period. It is estimated that approximately 90 of these subjects will meet the criterion for treatment response after dose optimization, and will then be randomized in a 1:1 ratio to placebo or NBI-98854 during the 6-week blinded randomization period, before entering the 24-week, double-blind, placebo-controlled withdrawal period. Treatment response will be based on the physician investigator's clinical assessment of sufficient control of tic behaviors.

4.2. Sample Size Considerations

The sample size for this study is based on the projected number of subjects expected to respond to treatment with NBI-98854 and not on a formal statistical power calculation.

4.3. Randomization

At the end of Weeks 8, 10, and 12, subjects whose tics have responded to NBI-98854 treatment (the physician investigator will determine treatment response based on a clinical assessment of sufficient control of tic behaviors), and who are tolerating their optimized dose of NBI-98854 based on physician investigator assessment, will be randomized in a 1:1 ratio to receive placebo or continue with their optimized dose of NBI-98854. Assessment of treatment response and tolerability will occur at each of the blinded randomization period visits at Weeks 8, 10, and 12; therefore, a subject who did not achieve a treatment response at Week 8 could achieve treatment response at Week 10 or Week 12 and be randomized. The subject responders, investigators, and Sponsor will be blinded to the visit week when randomization occurs (programmed on the Interactive Web Response System [IWRS]). Of the estimated 90 subjects who are anticipated to be treatment responders, approximately 30 will be randomized at each of the randomization visits. Randomization will be stratified based on subject's weight group at baseline (<50 kg vs. ≥50 kg).

4.4. Clinical Assessments

Efficacy assessments for TS include:

- Yale Global Tic Severity Scale (YGTSS),

- Clinical Global Impression of Tics-Severity scale (CGI-Tics-Severity),
- Clinical Global Impression of Tourette Syndrome-Improvement scale (CGI-TS-Improvement),
- Premonitory Urge for Tics Scale (PUTS), and
- Gilles de la Tourette Syndrome – Quality of Life Scale for Children and Adolescents (C&A-GTS-QOL).

Safety assessments include:

- adverse event (AE) monitoring,
- clinical laboratory tests (hematology, clinical chemistry, and urinalysis),
- serum prolactin,
- hemoglobin A1c,
- vital sign measurements,
- physical examinations, including weight,
- 12-lead electrocardiogram (ECG),
- Columbia-Suicide Severity Rating Scale (C-SSRS, Children's Version),
- Children's Depression Rating Scale - Revised (CDRS-R),
- Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS),
- Attention-Deficit Hyperactivity Disorder Rating Scale-5: Home Version (ADHD Rating Scale-5), and
- Extrapyramidal Symptom Rating Scale-Abbreviated (ESRS-A).

Blood samples for plasma drug 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

No interim analysis is 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.

"Withdrawal Period" for purposes of analyses includes the placebo-controlled withdrawal period and any post-randomization data collected in the blinded randomization period.

6.1. General Statistical Procedures

Descriptive and inferential statistical methods will be used to summarize the data from this study. The term "descriptive statistics" refers to the number of subjects, mean, median, standard deviation (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. Additional descriptive statistics may be presented for selected variables. The term "inferential statistics" refers to hypothesis tests which will be performed to assess differences between the NBI-98854 treatment group and the placebo group for selected efficacy variables during the withdrawal period. All hypothesis tests will be tests of the null hypothesis of no difference between the treatment groups being compared versus the two-sided alternative hypothesis that there is a difference. The level of significance (Type I error) for declaring statistical significance will be 0.05.

Descriptive statistics will generally be presented separately for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups. The treatment groups for all 3 weight group categories are NBI-98854 and Placebo. The NBI-98854 treatment group includes all possible study drug doses (ranging from 20 mg to 80 mg). Select tables will include an "All Subjects" treatment group, which combines the NBI-98854 and Placebo treatment groups. Select tables will separate the NBI-98854 treatment group by dispensed dose. A single treatment group (NBI-98854) will generally be used for summaries that include pre-randomization data.

Separate summaries may be combined in the same table for reporting purposes.

Inferential statistics will only be presented for the pooled weight groups.

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. P-values will be displayed using 4 decimal places. These rules may be modified if warranted, based on practical considerations.

6.2. Analysis Sets

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

6.2.1. Definition of Analysis Sets

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 open-label study drug. The safety analysis set will be used for all summaries of pre-randomization safety data (e.g., AEs and clinical laboratory data) and all plasma concentration data.

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 after their respective randomization visit (Week 8, 10, or 12) 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 (e.g., AEs and clinical laboratory data) that include post-randomization assessments. For subjects randomized to remain on active treatment, select summaries will include both pre- and post-randomization data.

6.2.1.3. Full Analysis Set

The full analysis set (FAS) will include all subjects who are randomized to a treatment group and have at least one post-randomization visit where loss of treatment response is able to be assessed (i.e., subject has a change from randomization baseline for both YGTSS Total Tic Score [TTS] and CGI-Tics-Severity or discontinues from the study during the withdrawal period due to either lack of efficacy or a treatment-emergent adverse event [TEAE] of worsening of tics). The FAS will be used for summaries and analyses of efficacy data. Treatment assignment for all summaries and analyses using the FAS will be based on the randomization schedule.

6.2.2. Summary of Analysis Sets

A summary of the number and percentage of subjects included in (and excluded from, as applicable) each analysis set will be provided. The number and percentage of subjects excluded from each analysis set by reason for exclusion will also be provided. Non-randomized and randomized subjects (presented by treatment group) will be summarized separately. An additional "All Subjects" column will be included. Summaries will be presented for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups using all enrolled subjects.

6.2.3. Application of Analysis Sets

Summaries of subject disposition and randomization, analysis set inclusion/exclusion status, and important protocol deviations (IPDs) will include all enrolled subjects. All other summaries by analysis set are identified in Table 2.

Table 2: Data Summaries by Analysis Set

Data Summary/Analysis	Analysis Set		
	Safety	Randomized Safety	FAS
Demographics	X		X
Baseline subject characteristics	X		X
Medical history	X	X	
Study drug exposure	X	X	
Study drug dose reductions	X	X	
Study drug dosing			X
Study drug compliance			X
Plasma concentrations	X		
Loss of treatment response (Section 9.3)			X
YGTSS			X
CGI-Tics-Severity			X
CGI-TS-Improvement			X
PUTS			X
C&A-GTS-QOL			X
Adverse events	X	X	
Clinical laboratory data	X	X	
Hemoglobin A1c	X	X	
Serum prolactin	X	X	
Vital signs	X	X	
Weight	X	X	
ECG	X	X	
C-SSRS	X	X	
CY-BOCS	X	X	
CDRS-R	X	X	
ESRS-A	X	X	
ADHD rating scale-5	X	X	
Prior and concomitant medications	X	X	

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 Day 1 study visit (excluding any assessments that occur after the first dose of open-label study drug) will serve as study baseline.

Randomization baseline: The assessments collected at the randomization visit (Week 8, 10, or 12) will serve as the randomization baseline value for select post-randomization assessments. If a value is not available at the randomization visit, then the last measurement collected on or prior to the randomization date (excluding any assessments that occur after 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/or 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 40 visit or is not eligible for randomization by Week 12. 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 7 days prior to and 6 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) for visits that are expected to occur every 2 weeks and within 14 days prior to and 13 days after for visits that are expected to occur every 4 weeks.

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 3 and Table 4 display the allowable study day range for each scheduled visit for ET visit mapping purposes for non-randomized and randomized subjects, respectively.

Table 3: Allowable Study Day Range for Early Termination Visit Mapping in Non-Randomized Subjects

Scheduled Visit	Target Study Day	Visit Window (Study Day Range)
Week 2	14	7-20
Week 4	28	21-34
Week 6	42	35-48
Week 8	56	49-62
Week 10	70	63-76
Week 12	84	77-90
Follow-up[1]	98-112	>90

[1] Subjects who are nonresponders at the Week 12 visit and who were not previously randomized will be discontinued from the study and asked to return 2 to 4 weeks after their final dose of study drug for an ET visit.

Table 4: Allowable Study Day Range for Early Termination Visit Mapping in Randomized Subjects

Scheduled Visit	Target Study Day	Visit Window (Study Day Range)
Week 10[1]	70	63-76
Week 12[2]	84	77-90
Week 14	98	91-104
Week 16	112	105-118
Week 18	126	119-132
Week 20	140	133-146
Week 24	168	154-181
Week 28	196	182-209
Week 32	224	210-237
Week 36	252	238-265
Week 40	280	>265

[1] Only possible for subjects randomized at Week 8.
[2] Only possible for subjects randomized at Weeks 8 or 10.

6.4.4. Post-randomization Timepoints

Timepoints for post-randomization summaries will be defined based on approximate time since randomization. Table 5 provides a list of the study visits that contribute to each defined timepoint based on the subject's randomization visit (Week 8, 10, or 12).

Table 5: Timepoint Definitions for Post-Randomization Summaries

Timepoint	Study Visit[1]		
	Randomized at Week 8	Randomized at Week 10	Randomized at Week 12
Randomization Baseline[2]	Week 8	Week 10	Week 12
2 Weeks Post-randomization	Week 10	Week 12	Week 14
4 Weeks Post-randomization	Week 12	Week 14	Week 16
6 Weeks Post-randomization	Week 14	Week 16	Week 18
8 Weeks Post-randomization	Week 16	Week 18	Week 20
10 Weeks Post-randomization	Week 18	Week 20	
12-14 Weeks Post-randomization	Week 20	Week 24	Week 24
16-18 Weeks Post-randomization	Week 24	Week 28	Week 28
20-22 Weeks Post-randomization	Week 28	Week 32	Week 32
24-26 Weeks Post-randomization	Week 32	Week 36	Week 36
28 Weeks Post-randomization	Week 36		
Follow-up	Week 40	Week 40	Week 40

[1] Includes mapped ET visits for post-randomization timepoints.

[2] Randomization Baseline as defined in Section 6.3, which may include assessments collected prior to the randomization visit.

6.5. Handling of Missing Data

6.5.1. Missing Efficacy Endpoints

Missing values for outcome measures will generally not be replaced with imputed values except as noted in Section 6.4.3 for the ET visit data mapped to scheduled visits for data summary purposes.

Any imputation methods used for the efficacy endpoints are discussed in the Efficacy Section (Section 9).

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 Day 1 visit date or enrollment date, whichever occurs later;
- If only the day is missing, the date will be imputed as the Day 1 visit date if the month and year of the first dose date match the month and year of the Day 1 visit; if the month and year of the first dose date occur after the Day 1 visit date, the missing day will be imputed as the first day of the month.

The imputation rules for first randomized dose date are as follows:

- If the date of the first dose of randomized study drug is completely missing or if both the day and month are missing, then the randomization date will be used.
- If only the day is missing, the date will be imputed as the randomization date if the month and year of the first dose of randomized study drug date match the month and year of randomization; if the month and year occur after the randomization date, the missing day will be imputed as the first day of the month.

The imputation rules for last dose date are as follows:

If the date of the last dose of study drug is missing in non-randomized subjects, then the last dose date will be imputed as the earliest of:

- the Week 12 visit date,
- study discontinuation date for subjects who discontinue before Week 12,
- the last visit (up to Week 10) prior to discontinuation + 17 or the date of an unscheduled dose reduction +17, whichever occurs later.

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

- the Week 36 visit date,
- study discontinuation date for subjects who discontinue before Week 36,
- the last visit prior to discontinuation + number of doses dispensed at the visit or the date of an unscheduled dose reduction + number of doses dispensed, whichever occurs later.

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

Missing and incomplete ("partial") 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

7.1. Subject Disposition

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 medication, 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 ended study participation. The number of subjects who discontinued from the study will also be displayed by reason for discontinuation.

These summaries will be presented for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups using all enrolled subjects. The following groups will be used: subjects enrolled but not randomized; subjects randomized to placebo; subjects randomized to NBI-98854; all subjects.

A listing of randomized subjects will also be provided and will include subject ID, informed consent/assent date, enrollment date, randomization date, baseline weight group, and randomized treatment group.

7.2. Protocol Deviations

Protocol deviations described in the study-specific Protocol Deviation Plan will be entered into the study database and used to identify IPDs. Important protocol deviations 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 and unblinding of the randomized treatment assignments. This committee will review a listing of all protocol deviations reported in the study database and determine which deviations are IPDs. Important protocol deviations include, but are not limited to, the following:

- Failure to obtain informed consent from the subject prior to performing any study procedures.
- Deviations from key inclusion/exclusion criteria.
- Use of prohibited concomitant medications.
- Error in drug dispensing which results in a subject not receiving intended randomized treatment.
- Significant deviation from protocol-specified dosing regimen.

A summary of the number and percentage of subjects with IPDs by deviation category will be provided. Separate summaries will be presented for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups using all enrolled subjects. The following groups will be used: subjects enrolled but not randomized; subjects randomized to placebo; subjects randomized to NBI-98854; all subjects.

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

7.3. 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. Separate summaries will be presented for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups using both the safety analysis set and the FAS. The FAS summary will be presented by treatment and will include an additional "All Subjects" column.

Demographics include:

- Age (years)
- Age category (child [ages 6-11], adolescent [ages 12-17])
- Sex
- Ethnicity
- Race

Study baseline subject characteristics include:

- Age at TS diagnosis (years)
- Baseline value of TTS
- Height (measured at screening; cm)
- Weight (presented in both pounds and kilograms)
- Body mass index (BMI; measured at screening; kg/m^2)
- CYP2D6 genotype status

7.4. 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 each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups using 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.

7.5. Study Drug Dosing and Compliance

7.5.1. Exposure to Study Drug

The following exposure durations will be calculated:

- Pre-randomization: first randomized dose date – first open-label dose date in subjects who are randomized; last dose date – first open-label dose date + 1 in subjects who are not randomized, or who are randomized but never received a dose of randomized study drug.
- Post-randomization: last dose date – first randomized dose date + 1

- Entire study: last dose date – first open-label dose date + 1

Duration of exposure (number of days) will be summarized with descriptive statistics. A single treatment group (NBI-98854) will be used for the pre-randomization exposure using the safety analysis set. Post-randomization exposure will be summarized by treatment group using the randomized safety analysis set. Entire study exposure will be summarized for subjects in the randomized safety analysis set who are randomized to NBI-98854.

The number and percentage of subjects within the following exposure categories will also be presented:

Pre-randomization

- >0 to <2 weeks
- ≥2 to <4 weeks
- ≥4 to <6 weeks
- ≥6 to <8 weeks
- ≥8 to <10 weeks
- ≥10 to <12 weeks
- ≥12 weeks

Post-randomization

- >0 to <2 weeks
- ≥2 to <4 weeks
- ≥4 to <6 weeks
- ≥6 to <8 weeks
- ≥8 to <10 weeks
- ≥10 to <12 weeks
- ≥12 to <16 weeks
- ≥16 to <20 weeks
- ≥20 to <24 weeks
- ≥24 to <28 weeks
- ≥28 weeks

Entire Study

- >0 to <8 weeks
- ≥8 to <10 weeks
- ≥10 to <12 weeks
- ≥12 to <16 weeks

- ≥16 to <20 weeks
- ≥20 to <24 weeks
- ≥24 to <28 weeks
- ≥28 to <32 weeks
- ≥32 to <36 weeks
- ≥36 weeks

Separate summaries will be presented for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

7.5.2. Dose Adjustments and Dosages

Descriptive statistics will be presented for the study drug dosage received immediately prior to randomization. Results will be presented by treatment group using the FAS and will include an additional "All Subjects" column. For subjects randomized to NBI-98854, descriptive statistics will also be presented for study drug dosages at the Week 36 visit and for the last dose received prior to study completion or early discontinuation. The number and percentage of subjects receiving each dosage will also be presented.

The number and percentage of subjects with a dose reduction at any time on or prior to randomization, or early discontinuation in non-randomized subjects, will be presented using the safety analysis set. The number and percentage of subjects with a dose reduction at any time after randomization will be presented by treatment group using the randomized safety analysis set.

All summaries described in this section will be presented for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

7.5.3. 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 in each treatment group who are dosing compliant will be presented for each post-randomization timepoint (defined in Table 5). Separate summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups using the FAS.

8. PLASMA CONCENTRATION DATA

The plasma concentrations of NBI-98854 and its active metabolite NBI-98782 will be summarized with descriptive statistics by visit and by the last dosage received prior to the blood sample being drawn (20, 40, 60, or 80 mg). The dose at the Week 40 visit will reflect the last dose the subject received during the study treatment period (i.e., the dosage at the subject's Week 36 visit). These summary tables will be presented for each baseline weight group (<50 kg vs. ≥50 kg) only. There will not be a pooled weight groups summary.

The summary tables will also be generated for CYP2D6 poor metabolizers vs. non-poor metabolizers within each baseline weight group.

Concentrations below the lower limit of quantification (BLQ) 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.

The following additional 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 (%).

Plasma concentrations of each analyte will be summarized with box plots by NBI-98854 dosage at each visit. These will be presented for each baseline weight group (<50 kg vs. ≥50 kg).

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. EFFICACY

9.1. General Considerations

Unless otherwise specified, the FAS will be used for all efficacy analyses.

9.2. Statistical Models

9.2.1. Mixed-Effect Model Repeated Measures (MMRM)

The primary analysis of numerical efficacy variables during the withdrawal period will be a mixed-effect model repeated measures (MMRM) analysis, which includes the changes from randomization baseline (or other dependent variable) to each post-randomization timepoint, as defined in Table 5. Unless otherwise specified, the model will include the following timepoints: 2 Weeks Post-randomization, 4 Weeks Post-randomization, 6 Weeks Post-randomization, and 8 Weeks Post-randomization. The model will include the randomization baseline value as a covariate, and study baseline weight group (<50 kg, ≥50 kg), treatment group (NBI-98854, Placebo), timepoint, treatment group-by-timepoint interaction, and randomization baseline value-by-timepoint interaction as fixed effects. Subject will be included as a random effect. Study site will not be included in the model, as there is a large number of sites, with most sites anticipated to enroll a small number of subjects.

Treatment group comparisons of the NBI-98854 treatment group vs. Placebo at each timepoint will be performed by constructing linear contrasts (or equivalent programming code) for differences between treatment group least-squares (LS) means. Nominal (raw) two-sided p-values for testing the statistical significance of these differences and associated 95% confidence intervals will be reported in summary tables.

The MMRM analysis will be implemented with the PROC MIXED procedure of [redacted], using the restricted maximum likelihood method, an unstructured within-subject covariance matrix, and denominator degrees of freedom from the Kenward-Roger method. If convergence is not obtained with the unstructured covariance matrix, a Toeplitz covariance structure, followed by the heterogeneous autoregressive covariance structure, followed by the compound symmetry will be used.

9.2.2. Analysis of Covariance

An analysis of covariance (ANCOVA) model will be used for select numerical endpoints. The model will include the randomization baseline value as a covariate, and study baseline weight group (<50 kg, ≥50 kg) and treatment group (NBI-98854, Placebo) as fixed effects.

Unless otherwise specified, missing values at Week 36 will be imputed using last observation carried forward (LOCF) using the following rules:

- If a subject discontinues study prior to Week 36, the last post-randomization observation on or prior to discontinuation will be used.
- If subjects are missing the Week 36 value, then the last post-randomization observation prior to Week 36 will be used.

9.3. Analysis of the Primary Efficacy Endpoint

9.3.1. Primary Efficacy Analysis

The primary efficacy endpoint is time to loss of treatment response over the course of the withdrawal period (through Week 36). Loss of treatment response during the withdrawal period will be defined as:

- 2 consecutive visits with 1) an increase in the YGTSS TTS of greater than 35% or 7 points from the randomization baseline (Section 6.3) and 2) an increase in CGI-Tics-Severity score of ≥2 points from the randomization baseline. If these 2 criteria are met at a subject's last observed visit during the withdrawal period (e.g., Week 36 or ET), then it will be considered as loss of treatment response. Or,
- Discontinuation from study due to either lack of efficacy or a TEAE of worsening of tics.

Time to loss of treatment response will be measured in days from randomization to the date of loss of treatment response, defined as the earliest of:

- The date of the first visit of the 2 consecutive visits where the loss of treatment response criterion based on the YGTSS TTS and CGI-Tics-Severity is met.
- The date of the last visit if the loss of treatment response criterion based on the YGTSS TTS and CGI-Tics-Severity is met at a subject's last observed visit during the withdrawal period (e.g., Week 36 or ET).
- The discontinuation date for subjects who discontinue from the study due to either lack of efficacy or a TEAE of worsening of tics.

The primary analysis of the primary efficacy endpoint will be a log-rank test stratified by baseline weight group. Kaplan-Meier (KM) estimates of loss of treatment response for each treatment group and median, upper, and lower quartiles and associated confidence intervals will be calculated where possible. Time to loss of treatment response for subjects who do not have a loss of treatment response prior to completion or discontinuation from the withdrawal period will be censored at their last observed visit during the withdrawal period (i.e., Week 36 visit date or ET visit).

KM plots of the time to loss of treatment response will be presented by treatment. For the primary efficacy analysis, KM plots will also be presented by treatment within each baseline weight group.

These analyses will be implemented with the PROC LIFETEST procedure of [redacted]

The hazard ratio (NBI-98854 vs. Placebo) and its 95% confidence interval for time to loss of treatment response will be calculated using the PROC PHREG procedure in [redacted], with baseline weight group included in the model.

9.3.2. Sensitivity Analyses of the Primary Efficacy Results

Modifications of the primary efficacy analysis (described in Section 9.3.1):

The log-rank test and KM estimates for loss of treatment response will be repeated with randomization visit (Week 8, 10, 12) included as a stratification factor.

As supportive analyses, the log-rank test and KM estimates for loss of treatment response will also be calculated using the following strategies for handling early discontinuations:

- Discontinuations due to AEs are treated as loss of response. Discontinuations for reasons other than AEs or lack of efficacy still result in censored time to loss of response.
- Discontinuations for any reason are treated as loss of response.

KM plots of the time to loss of treatment response using the described methods will be presented by treatment.

Percentage of subjects with loss of treatment response:

Descriptive statistics will be presented by treatment group for the number and percentage of subjects with loss of treatment response at any time during the withdrawal period. Subjects who discontinue the study prior to the Week 36 visit will be classified as losing treatment response. Summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups. For the pooled weight groups summary, an analysis comparing the NBI-98854 treatment group to the Placebo treatment group will be performed using the Cochran-Mantel-Haenszel (CMH) procedure, with baseline weight group as a stratification variable.

9.4. Analysis of the Secondary Efficacy Endpoints

9.4.1. YGTSS TTS

Change from randomization baseline to the 8 weeks post-randomization timepoint and to the Week 36 study visit in the YGTSS TTS as generated by the certified site rater using the RaterStation™ are secondary endpoints. Changes from randomization baseline to other timepoints or visits are exploratory efficacy endpoints.

The TTS is defined as the sum of the YGTSS motor tic severity score and phonic (vocal) tic severity score. The motor tic severity score is calculated as the sum of the scores for the 5 motor tic items (number, frequency, intensity, complexity, and interference). The score for each item can range from 0 to 5, for a maximum total score of 25. The vocal (phonic) tic severity score is calculated similarly. The TTS value can range from 0 to 50, with higher scores representing greater severity. If any one of the 5 items for the motor or vocal tic severity score is not scored (i.e., has a missing value), the associated severity score will be set equal to missing. If any of these items has a missing value at a given visit, the TTS value for the visit will also be set equal to missing.

Descriptive statistics will be presented by treatment group for the observed values at study baseline, randomization baseline, and at each timepoint (as defined in Table 5) that occurs after randomization. Change from randomization baseline will be included for the post-randomization timepoints. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The change from randomization baseline at each timepoint through 8 weeks post-randomization will be analyzed using an MMRM analysis for the pooled weight groups, as described in Section 9.2.1.

The change from randomization baseline to the Week 36 study visit will be analyzed for the pooled weight groups using an ANCOVA model, as described in Section 9.2.2. Missing values at Week 36 will be imputed using LOCF.

Mean (±SEM) values of the observed values at each timepoint will be summarized in line graphs by treatment group. Similar graphs will be presented for the changes from randomization baseline. These graphs will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The LS means (±SEM) from the MMRM analysis results will be summarized in similar line graphs for the pooled weight groups.

9.4.2. CGI-Tics-Severity

Change from randomization baseline to the 8 weeks post-randomization timepoint and to the Week 36 study visit in the CGI-Tics-Severity score are secondary endpoints. Changes from randomization baseline to other timepoints or visits are exploratory efficacy endpoints.

Each of the CGI-Tics-Severity response categories will be assigned a numerical score as follows:

- Normal, not at all ill = 1
- Borderline ill = 2
- Mildly ill = 3
- Moderately ill = 4
- Markedly ill = 5
- Severely ill = 6
- Among the most extremely ill patient = 7

Descriptive statistics will be presented by treatment group for the observed numerical scores at study baseline, randomization baseline, and at each timepoint (as defined in Table 5) that occurs after randomization. Change from randomization baseline will be included for the post-randomization timepoints. Frequency counts using the response categories will also be presented. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The change from randomization baseline at each timepoint through 8 weeks post-randomization will be analyzed using an MMRM analysis for the pooled weight groups, as described in Section 9.2.1.

The change from randomization baseline to the Week 36 study visit will be analyzed for the pooled weight groups using an ANCOVA model, as described in Section 9.2.2. Missing values at Week 36 will be imputed using LOCF.

Mean (±SEM) values of the observed values at each timepoint will be summarized in line graphs by treatment group. Similar graphs will be presented for the changes from randomization baseline. These graphs will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The LS means (±SEM) from the MMRM analysis results will be summarized in similar line graphs for the pooled weight groups.

9.5. Analysis of the Exploratory Efficacy Endpoints

9.5.1. YGTSS

9.5.1.1. YGTSS TTS Responder Analysis

A TTS responder is defined, on a per-visit basis, as a subject whose TTS value is reduced by at least 30% from study baseline at the specified postbaseline visit.

Descriptive statistics will be presented by treatment group for the number and percentage of subjects classified as TTS responders at each timepoint (defined in Table 5). These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

Descriptive statistics will also be presented for the number and percentage of subjects classified as TTS responders at the Week 36 study visit. Missing values will be imputed using nonresponder imputation (NRI). These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups. For the pooled weight groups summaries, an analysis comparing the NBI-98854 treatment group to the Placebo treatment group will be performed using the CMH procedure, with baseline weight group as a stratification variable.

9.5.1.2. YGTSS Impairment Score

The YGTSS Impairment score can range in value from 0 to 50, with higher scores representing more severe impairment.

Descriptive statistics will be presented by treatment group for the observed values at study baseline, randomization baseline, and at each timepoint (as defined in Table 5) that occurs after randomization. Change from randomization baseline will be included for the post-randomization timepoints. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The change from randomization baseline at each timepoint through 8 weeks post-randomization will be analyzed using an MMRM analysis for the pooled weight groups, as described in Section 9.2.1.

The change from randomization baseline to the Week 36 study visit will be analyzed for the pooled weight groups using an ANCOVA model, as described in Section 9.2.2. Missing values at Week 36 will be imputed using LOCF.

Mean (±SEM) values of the observed values at each timepoint will be summarized in line graphs by treatment group. Similar graphs will be presented for the changes from randomization baseline. These graphs will be presented for the pooled weight groups.

The LS means (±SEM) from the MMRM analysis results will be summarized in similar line graphs for the pooled weight groups.

9.5.1.3. YGTSS Global Tic Severity Score

The YGTSS Global Tic Severity score is the sum of the TTS and the YGTSS Impairment score. The YGTSS Global Tic Severity score at a given subject visit will be set equal to missing if

either of the TTS or Impairment scores are missing. The YGTSS Global Tic Severity score value can range from 0 to 100.

Descriptive statistics will be presented by treatment group for the observed values at study baseline, randomization baseline, and at each timepoint (as defined in Table 5) that occurs after randomization. Change from randomization baseline will be included for the post-randomization timepoints. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The change from randomization baseline at each timepoint through 8 weeks post-randomization will be analyzed using an MMRM analysis for the pooled weight groups, as described in Section 9.2.1.

The change from randomization baseline to the Week 36 study visit will be analyzed for the pooled weight groups using an ANCOVA model, as described in Section 9.2.2. Missing values at Week 36 will be imputed using LOCF.

Mean (±SEM) values of the observed values at each timepoint will be summarized in line graphs by treatment group. Similar graphs will be presented for the changes from randomization baseline. These graphs will be presented for the pooled weight groups.

The LS means (±SEM) from the MMRM analysis results will be summarized in similar line graphs for the pooled weight groups.

9.5.2. CGI-TS-Improvement

9.5.2.1. CGI-TS-Improvement

Each of the CGI-TS-Improvement response categories will be assigned a numerical score as follows:

- Very much improved = 1
- Much improved = 2
- Minimally improved = 3
- Not changed = 4
- Minimally worse = 5
- Much worse = 6
- Very much worse = 7

Descriptive statistics will be presented by treatment group for the observed numerical scores at randomization baseline and at each timepoint (as defined in Table 5) that occurs after randomization. Frequency counts using the response categories will also be presented. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The observed values at each timepoint through 8 weeks post-randomization will be analyzed using an MMRM analysis for the pooled weight groups. The model will be similar to that described in Section 9.2.1 with the following changes: observed values are used instead of changes from randomization baseline. The numerical score at randomization will be used as the covariate.

The observed values at the Week 36 study visit will be analyzed for the pooled weight groups using an ANCOVA model, as described in Section 9.2.2. Missing values at Week 36 will be imputed using LOCF.

Mean (±SEM) values of the observed values at each timepoint will be summarized in line graphs by treatment group. These graphs will be presented for the pooled weight groups.

The LS means (±SEM) from the MMRM analysis results will be summarized in similar line graphs for the pooled weight groups.

9.5.2.2. CGI-TS-Improvement Responder Analysis

A subject is classified as a CGI-TS-Improvement responder at a given visit if their CGI-TS-Improvement score is either a "1" ("very much improved") or a "2" ("much improved") at the visit.

Descriptive statistics will be presented by treatment group for the number and percentage of subjects classified as CGI-TS-Improvement responders at each timepoint (defined in Table 5). These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

Descriptive statistics will also be presented for the number and percentage of subjects classified as CGI-TS-Improvement responders at the Week 36 study visit. Missing values will be imputed using NRI. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups. For the pooled weight groups summaries, an analysis comparing the NBI-98854 treatment group to the Placebo treatment group will be performed using the CMH procedure, with baseline weight group as a stratification variable.

9.5.3. Premonitory Urge for Tics Scale

The PUTS is an instrument for quantifying the premonitory urge phenomena associated with tics. It consists of 9 items, each of which is scored on a 4-point scale:

- 1 = not at all true
- 2 = a little true
- 3 = pretty much true
- 4 = very much true

The PUTS total score is calculated as the sum of the scores for the 9 items. The PUTS total score value can range from 9 to 36. If any one of the 9 items is not scored (i.e., has a missing value), the PUTS total score will be set equal to missing.

Descriptive statistics will be presented by treatment group for the observed values at study baseline, randomization baseline, and at each timepoint (as defined in Table 5) that occurs after randomization. Change from randomization baseline will be included for the post-randomization timepoints. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

The change from randomization baseline at each timepoint through 8 weeks post-randomization will be analyzed using an MMRM analysis for the pooled weight groups, as described in Section 9.2.1.

The change from randomization baseline to the Week 36 study visit will be analyzed for the pooled weight groups using an ANCOVA model, as described in Section 9.2.2. Missing values at Week 36 will be imputed using LOCF.

Mean (±SEM) values of the observed values at each timepoint will be summarized in line graphs by treatment group. Similar graphs will be presented for the changes from randomization baseline. These graphs will be presented for the pooled weight groups.

The LS means (±SEM) from the MMRM analysis results will be summarized in similar line graphs for the pooled weight groups.

9.5.4. Gilles de la Tourette Syndrome-Quality of Life Scale for Children and Adolescents

The C&A-GTS-QOL has two parts:

(1) a 27-item scale which includes 4 factors (subscales), with each of the 27 items scored on a 5-point scale (1=never, 2=rarely, 3=sometimes, 4=often, 5=always); and

(2) a visual analog scale (VAS) which ranges in value from 0 to 100, with 0 representing extremely unhappy/dissatisfied with life, and 100 representing extremely happy/satisfied with life.

The C&A-GTS-QOL total score is calculated as the sum of the scores for the 27 items. If any one of the 27 items is not scored (i.e., has a missing value), the total score will be set equal to missing.

Scores will be calculated for each of the 4 C&A-GTS-QOL factors in a similar fashion. These factors and the corresponding item numbers are as follows:

- Psychological (16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27)
- Physical and activities of daily living (1, 2, 3, 4, 5, 6, 26)
- Obsessive-compulsive (7, 8, 9, 10, 15)
- Cognitive (11, 12, 13, 14)

The total score and each factor will be normalized to a value that can range from 0 to 100 by using the following formula:

$$100 \times \frac{observed\ score - minimum\ possible\ score}{maximum\ possible\ score - minimum\ possible\ score}$$

Two versions of this instrument are used in this trial: 1 version for children aged 6 to 12 years and 1 version for adolescents aged 13 to 18 years. Scoring is identical for both versions and all results will be combined for analyses.

Descriptive statistics will be presented by treatment group for the observed C&A-GTS-QOL normalized total scores, the 4 normalized factor scores, and the VAS scores at study baseline, randomization baseline, and at each timepoint (as defined in Table 5) that occurs after randomization. Change from randomization baseline will be included for the post-randomization timepoints. These summaries will be presented by baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

For the normalized total scores, the change from randomization baseline at each timepoint through 8 weeks post-randomization will be analyzed using an MMRM analysis for the pooled weight groups, as described in Section 9.2.1.

For the normalized total scores, the change from randomization baseline to the Week 36 study visit will be analyzed for the pooled weight groups using an ANCOVA model, as described in Section 9.2.2. Missing values at Week 36 will be imputed using LOCF.

Mean (±SEM) values of the observed normalized total scores at each timepoint will be summarized in line graphs by treatment group. Similar graphs will be presented for the changes from randomization baseline. These graphs will be presented for the pooled weight groups.

The LS means (±SEM) from the MMRM analysis results will be summarized in similar line graphs for the pooled weight groups.

10. SAFETY

Unless otherwise specified, summaries for each analysis described in this section will be presented separately for each baseline weight group (<50 kg vs. ≥50 kg) and for the pooled weight groups.

10.1. General Considerations

10.1.1. By-visit Summaries

The following table formats will be used when summarizing by-visit data:

Pre-randomization: descriptive statistics of observed and change from study baseline values will be presented at study baseline and at each scheduled postbaseline visit through Week 12. A Follow-up visit, as described in Table 3, will also be included. Results will be presented for a single treatment group (NBI-98854) using the safety analysis set. Assessments occurring after randomization in subjects who are randomized will be excluded. For select assessments (clinical laboratory data, vital signs, body weight, and ECG data), a summary at "Pre-randomization Endpoint" will also be included.

"Pre-randomization Endpoint" will be defined as:

- For randomized subjects: the last value at any scheduled visit occurring after the first dose of open-label study drug and on or prior to the date of the first dose of randomized study drug
- For non-randomized subjects: the Week 12 visit data if available; otherwise the last value at any scheduled visit or ET visit occurring after the first dose of open-label study drug and on or prior to the Week 12 visit or discontinuation (whichever occurs first)

Post-randomization: descriptive statistics of observed and change from study baseline values will be presented at study baseline, randomization baseline, and at each timepoint (as defined in Table 5) that occurs after randomization. Results will be presented by treatment group (NBI-98854 or Placebo) using the randomized safety analysis set. Study baseline will be used when calculating change from baseline. For select assessments (clinical laboratory data, vital signs, body weight, and ECG data), a summary at "Post-randomization Endpoint" will also be included.

"Post-randomization Endpoint" will be defined as:

- The Week 36 visit data for subjects who complete the withdrawal period and have available data
- The last value at any scheduled visit or ET visit occurring after the first dose of randomized study drug and on or prior to the Week 36 visit or discontinuation (whichever occurs first) for subjects who discontinue study or otherwise have missing data

Entire Study: for select assessments (clinical laboratory data, vital signs, body weight, and ECG data), descriptive statistics of observed and change from study baseline values will be presented at study baseline and at each scheduled postbaseline visit through Week 40. Results will be presented only for subjects who are randomized to NBI-98854, using the randomized safety analysis set. A summary at "Post-randomization Endpoint" will also be included.

10.1.2. Treatment Period Definitions

Table 6 provides a description of the treatment periods that will be used to describe AEs and assessment-based abnormalities (such as laboratory potentially clinically significant [PCS] values or "Yes" responses in the C-SSRS items).

Table 6: Description of Treatment Periods for Event Data

Event Type	Component	Treatment Period		
		Pre-randomization	Post-randomization	Entire Study
All	Analysis Set	Safety Analysis Set	Randomized Safety Analysis Set	Randomized Safety Analysis Set
	Treatment Groups	NBI-98854	NBI-98854 Placebo	NBI-98854
Adverse Events	Start of Period	AE start date on or after the date of the first dose of open-label study drug	AE start date after the date of the first dose of randomized study drug	AE start date on or after the date of the first dose of open-label study drug
	End of Period	AE start date on or prior to the date of the first dose of randomized study drug for subjects who are randomized AE start date on or prior to the date of the last dose of study drug + 30 for subjects who are not randomized or who are randomized but never receive a dose of randomized study drug	AE start date on or prior to the date of the last dose of study drug + 30	AE start date on or prior to the date of the last dose of study drug + 30
Abnormal Assessments	Start of Period	Assessment date after the date of the first dose of open-label study drug	Assessment date after the date of the first dose of randomized study drug	Assessment date after the date of the first dose of open-label study drug
	End of Period	Assessment date on or prior to the date of the first dose of randomized study drug for subjects who are randomized End of study participation for subjects who are not randomized or who are randomized but never receive a dose of randomized study drug	End of study participation	End of study participation

10.2. Adverse Events

Adverse events are recorded in the eCRF. Each AE will be coded to SOC and PT using MedDRA (Version 12.0).

A 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. The determination of whether an AE is treatment-emergent is based on the AE onset date relative to the date of the subject's first dose of study drug. If the AE onset date and date of the first dose of study drug are the same, or if the AE onset date is unknown, it will be assumed that the AE is a TEAE.

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 Table 6.

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 NBI-98854 treatment group. If more than one NBI-98854 treatment group is included, the group that combines all doses or entire study will be used when sorting;
- Frequency of TEAEs by PT, with PT sorted by decreasing frequency (number of unique subjects) in the NBI-98854 treatment group. If more than one NBI-98854 treatment group is included, the group that combines all doses or entire study will be used when sorting.

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 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.

10.2.1. Adverse Events Resulting in Premature Discontinuation from Study

The number and percentage of subjects with a TEAE resulting in study discontinuation will be presented by PT within SOC for each treatment period (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 weight group, subject ID, treatment group, last treatment received prior to the onset time of the TEAE(s) leading to discontinuation, study day of the discontinuation, and other relevant

information from the AE eCRF. Note that "last treatment received prior to the onset time of the TEAE[s] leading to discontinuation" reflects the actual dosage received prior to the AE.

10.2.2. Adverse Events Resulting in Study Drug Dose Reductions

The number and percentage of subjects with a TEAE resulting in a dose reduction will be presented by PT within SOC for each treatment period (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.

10.2.3. Deaths and Other Serious Adverse Events

Summary tables of serious adverse events (SAEs) will be presented for each treatment period. 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 weight group, subject ID, treatment group, last treatment received prior to the onset time of the SAE or fatal TEAE, study day of the SAE or fatal TEAE, and any additional relevant information from the AE eCRF.

10.2.4. Summaries by NBI-98854 Dose

To help understand any potential dose-related safety signals, select summaries will be presented where subjects are analyzed by the NBI-98854 dosage dispensed at the randomization visit (regardless of any subsequent dose reductions):

- Placebo
- NBI-98854 20 mg
- NBI-98854 40 mg
- NBI-98854 60 mg
- NBI-98854 80 mg
- NBI-98854 All Doses

These summaries will include all subjects in the randomized safety analysis set, and will be presented for the following previously described tables:

- AE overview table
- TEAEs by SOC and PT
- SAEs by SOC and PT

These summaries will only be presented for the post-randomization treatment period.

10.3. Clinical Laboratory Data

Repeat clinical laboratory samples may be collected at any time during this study due to either missing or abnormal results. The general rule for summarizing these data is to include the

original sample results in any by-visit summary tables and graphs. Repeat or unscheduled assessments will be included in any lab abnormality summaries.

By-visit summaries

The observed and change from study baseline hematology, clinical chemistry, hemoglobin A1c, and prolactin data will be summarized with descriptive statistics by visit/timepoint, as described in Section 10.1.1.

The prolactin data will also be summarized for each sex separately.

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.

The following shift tables will be presented:

- Shift from study baseline to "Pre-randomization Endpoint" (as defined in Section 10.1.1) using the safety analysis set
- Shift from study baseline to "Post-randomization Endpoint" (as defined in Section 10.1.1) using the randomized safety analysis set, presented by treatment group

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 category at postbaseline. A "Total" row and "Total" column will also be included. Subjects with a missing baseline value or who do not have the specified 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.

Potentially clinically significant (PCS) values

Summaries of sponsor-defined PCS values will be presented for the following clinical laboratory variables: ALT, AST, ALP, creatine kinase, GGT, total bilirubin, white blood cell count, absolute neutrophil count, creatinine, and BUN. The criteria for identifying PCS clinical laboratory values are provided in Table 7.

The number and percentage of subjects with PCS values that are reported at any postbaseline visit (scheduled or unscheduled) will be summarized. Results will be presented for each of the treatment periods described for abnormal assessments in Table 6.

Table 7: Potentially Clinically Significant Criteria for Clinical Laboratory Variables

Variable	PCS Threshold
ALT	>3 x ULN (upper limit of normal)
AST	>3 x ULN
ALP	>2.5 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)

Figures

Scatter plots of selected variables will be created which display the following:

- Study baseline vs. "Pre-randomization Endpoint" (as defined in Section 10.1.1) using the safety analysis set
- Study baseline vs. "Post-randomization Endpoint" (as defined in Section 10.1.1) using the randomized safety analysis set, presented by treatment group

Subjects with a baseline and postbaseline value will be included. Each plot will include a 45 degree ("y=x") reference line. The plots will be generated for ALT, AST, creatine kinase, GGT, total bilirubin, and prolactin.

Boxplots will be presented for the prolactin data, as described in the by-visit summaries above. Separate plots will also be presented for each sex.

10.4. Vital Signs

The observed and change from study baseline vital signs data, including orthostatic blood pressures and heart rate (calculated as standing value minus supine value), will be summarized with descriptive statistics by visit/timepoint, as described in Section 10.1.1.

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 for children (6 to 11 years of age at study baseline) and adolescents (12 to 17 years of age at study baseline) are provided in Table 8 and Table 9, respectively.

The number and percentage of subjects with PCS values that are reported at any postbaseline visit (scheduled or unscheduled) will be summarized. Results will be presented for each of the treatment periods described for abnormal assessments in Table 6.

Table 8: Potentially Clinically Significant Criteria for Vital Signs Variables in Children (6 to 11 years of age)

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	N/A		≥20 mmHg	>130 mmHg		≥20 mmHg
Diastolic Blood Pressure	N/A		≥10 mmHg	>85 mmHg		≥10 mmHg
Heart Rate	N/A		≥15 bpm	>130 bpm		≥10 bpm

Table 9: Potentially Clinically Significant Criteria for Vital Signs Variables in Adolescents (12 to 17 years of age)

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	N/A		≥20 mmHg	>145 mmHg		≥20 mmHg
Diastolic Blood Pressure	N/A		≥10 mmHg	>90 mmHg		≥10 mmHg
Heart Rate	N/A		≥15 bpm	>110 bpm		≥10 bpm

Both supine and standing values of blood pressures and heart rate will be included in the identification and summary of PCS values. Study baseline will be used when determining any increases or decreases from baseline.

10.5. Body Weight

The observed and change from study baseline body weight data (in units of kilograms) will be summarized with descriptive statistics by visit/timepoint, as described in Section 10.1.1.

10.6. 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 in triplicate, 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/greatest abnormality of the available value(s) will be used.

The observed and change from study baseline values for the quantitative ECG variables will be summarized with descriptive statistics by visit/timepoint as described in Section 10.1.1. Frequency counts and percentages for the ECG interpretation variable categories will also be summarized.

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.

Two categorical summaries will be presented for the QT and QTcF intervals (each interval will be summarized separately). 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 study baseline value meets the following thresholds:

- Greater than 30 msec
- Greater than 60 msec

The number and percentage of subjects in each of the specified categories using assessments collected at any postbaseline visit (scheduled or unscheduled) will be summarized. Results will be presented for each of the treatment periods described for abnormal assessments in Table 6.

10.7. Columbia-Suicide Severity Rating Scale Children's Versions

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

- Screening/lifetime assessment
- Screening/past 1 year assessment
- Study baseline assessment
- Pre-randomization treatment period
- Post-randomization treatment period

- Entire study treatment period

The screening and study baseline summaries will include the following columns:

- NBI-98854: all subjects in the safety analysis set
- NBI-98854: all subjects in the randomized safety analysis set who are randomized to NBI-98854
- Placebo: all subjects in the randomized safety analysis set who are randomized to placebo

Each treatment period 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) at any time during the specified treatment period, using the treatment periods and analysis sets defined for abnormal assessments in Table 6.

The 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 (not collected for the Screening/past 1 year assessment)
  - (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)

In addition to the summaries described above, shift tables comparing postbaseline suicidal ideation scores to study 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.

10.8. Extrapyramidal Symptom Rating Scale-Abbreviated

The ESRS-A assesses 4 types of movement disorders: parkinsonism, dystonia, akathisia, and dyskinesia. The ESRS-A consists of four subscales, one for each type of movement disorder. The ESRS-A contains 10 items to evaluate parkinsonism, 6 items to evaluate dystonia, 6 items to evaluate dyskinesia, and 2 items to evaluate akathisia. Each item score can range from 0 to 5, for a maximum possible parkinsonism score of 50, maximum possible dystonia score of 30, maximum possible dyskinesia score of 30, and maximum possible akathisia score of 10. A Clinical Global Impression of Movement Severity (CGI-S) is also completed for each type of movement disorder, and is also is scored on a 0 to 5 scale.

The subscale scores for each type of movement disorder (parkinsonism, akathisia, dystonia, and dyskinesia) will be calculated as the sum of the scores of the individual items comprising each subscale. The overall total score will be calculated as the sum of each of the subscale scores. If any one of the items is not scored (i.e., has a missing value), the associated subscale score and total score will be set equal to missing. The CGI-S scores will be summarized separately from the subscale scores.

The observed and change from study baseline values for each subscale score, total score and CGI-S score will be summarized with descriptive statistics by visit/timepoint, as described in Section 10.1.1.

10.9. Children's Yale-Brown Obsessive-Compulsive Scale

The CY-BOCS is a semi-structured interview designed to rate the severity of obsessive and compulsive symptoms in children.

The CY-BOCS obsession subtotal score is calculated as the sum of the scores for items 1 through 5 of the CY-BOCS scale (excluding item 1b), and the CY-BOCS compulsion subtotal score is calculated as the sum of the scores for items 6 through 10 of the CY-BOCS scale (excluding item 6b). The CY-BOCS total score is the sum of the obsession and compulsion subtotal scores. Each item score ranges from 0 to 4, with a maximum possible obsession subtotal score of 20, a maximum possible compulsion subtotal score of 20, and a maximum possible total score of 40. If any one of these 10 items is not scored (i.e., has a missing value), the associated subtotal score and total score will be set equal to missing.

The observed and change from study baseline values for the obsession subtotal, compulsion subtotal and total scores will be summarized with descriptive statistics by visit/timepoint, as described in Section 10.1.1.

10.10. Children's Depression Rating Scale, Revised

The CDRS-R is a 17-item, semi-structured interview to determine the severity of depression in children.

The CDRS-R total score is calculated as the sum of the 17 items making up the CDRS-R. Each item score ranges from 1 to 7 with the exception of items 4, 5, and 16, which range from 1 to 5.

The maximum possible total score is 113. If any one of the 17 items is not scored (i.e., has a missing value), the total score will be set equal to missing.

The observed and change from study baseline values for the total score will be summarized with descriptive statistics by visit/timepoint, as described in Section 10.1.1.

10.11. Attention-Deficit Hyperactivity Disorder Rating Scale-5: Home Version

The ADHD Rating Scale-5: Home Version will be used to determine the frequency and severity of ADHD symptoms and impairments over the 2 weeks prior to each visit.

The scale consists of 2 symptom subscales: Inattention (9 items) and Hyperactivity-Impulsivity (9 items). The Inattention subscale and Hyperactivity-Impulsivity subscale scores are each derived by summing up the 9 relevant item scores. Each item score ranges from 0 to 3, therefore each subscale score has a range of 0 to 27. The Total Scale raw score is defined as the sum of the Inattention and Hyperactivity-Impulsivity subscale scores, and has a range of 0 to 54. If any one of the 18 items is not scored (i.e., has a missing value), the affected subscale score(s) and Total Scale raw score will be set equal to missing.

The scale also assesses 6 domains of impairment that are common among children and adolescents with ADHD: relationships with family members, peer relationships, academic functioning, behavioral functioning, homework performance, and self-esteem. Each domain impairment item ranges from 0-3 and is assessed after completing each of the 2 subscales (i.e., each impairment item is assessed twice). The score for each domain is defined as the higher (worst) of the two scores. A total impairment score is defined as the sum of the 6 impairment items (using the higher of the two scores) and has a range of 0 to 18. If any one of the 12 individual items is not scored (i.e., has a missing value), the total impairment score will be set equal to missing.

Two versions of this instrument are used in this trial: 1 version is for children aged 5 to 10 years and 1 version is for adolescents aged 11 to 17 years. Scoring is identical for both versions and all results will be combined for analyses.

The observed and change from study baseline values for the Inattention and Hyperactivity-Impulsivity subscale scores, the Total Scale raw score, and the Total Impairment score will be summarized with descriptive statistics by visit/timepoint, as described in Section 10.1.1.

10.12. 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, two, or three study periods based on the medication start and stop dates relative to study drug dosing and randomization.

- Prestudy/screening: medications with a start date prior to open-label study drug dosing
- During the pre-randomization treatment 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 of randomized study drug

- During randomized treatment or posttreatment: medications ongoing at the time of first randomized study drug dosing or with a start date after the first dose of randomized study drug.

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 study period. 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.

The following columns will be included in the summaries:

- NBI-98854: all subjects in the safety analysis set (note: omitted from the "during randomized treatment or posttreatment" summary)
- NBI-98854: all subjects in the randomized safety analysis set who are randomized to NBI-98854
- Placebo: all subjects in the randomized safety analysis set who are randomized to placebo

11. APPENDICES

11.1. Sample [redacted] Code

This appendix contains sample [redacted] code that will be used for the efficacy analyses described in this SAP. The exact code may vary from the samples provided.

The following variables are used:

- trt = treatment group (NBI-98854 or Placebo)
- weight = baseline weight group (<50 kg or ≥50 kg)
- subjid = subject identification number
- dur = time of event (loss of treatment response) or censoring
- status = indication of whether or not the observation was censored ("0"=censored)
- timepoint = visit/timepoint number
- val = numerical value of variable being analyzed (typically, change from baseline)
- base = baseline value of variable being analyzed

Stratified log-rank test

The following code will be used to test treatment effects on time of loss of treatment response stratified by baseline weight group:

Kaplan-Meier estimates

The following code will be used to generate KM estimates of loss of treatment response for each treatment group (such as, median, upper, and lower quartiles and associated confidence intervals; KM plots):

Hazard ratios

The following code will be used to estimate the hazard ratio (NBI-98854 vs. Placebo) and its 95% confidence interval for time to loss of treatment response, adjusting for baseline weight group:

MMRM

After subsetting the dataset to observed records at the specified postbaseline timepoints, the following code will be used to implement the MMRM analysis:

ANCOVA

After subsetting the dataset to records at the specified postbaseline visit and imputation method, the following code will be used to implement the ANCOVA analysis: