1.0 TITLE PAGE

[redacted]

MT10109L-006

A Multicenter, Double-Blind, Randomized, Placebo-Controlled, Parallel-Group Study to Evaluate the Safety and Efficacy of MT10109L (NivobotulinumtoxinA) for the Treatment of Lateral Canthal Lines with or Without Concurrent Treatment of Glabellar Lines

STATISTICAL ANALYSIS PLAN - Clinical Study Report

Final: 18 NOV 2020

2.0 TABLE OF CONTENTS

3.0 LIST OF ABBREVIATIONS

AE	adverse event
AESI	adverse events of special interest
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
CMH	Cochran-Mantel-Haenszel
CRF	case report form
ECG	electrocardiogram
EU	European Union
FLO-11	11-item Facial Line Outcomes questionnaire
FLSQ	Facial Line Satisfaction Questionnaire
FWS	Facial Wrinkle Scale with Photonumeric Guide
GL	glabellar lines
ITT	intent to treat
LCL	lateral canthal lines
MCMC	Markov chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
PDSOT	possible distant spread of toxin
Q1	25^{th} percentile
Q3	75^{th} percentile
QTc	QT interval corrected for heart rate
QTcB	QT interval corrected for heart rate using the Bazett formula ($QTcB = QT/(RR)^{1/2}$)
QTcF	QT interval corrected for heart rate using the Fridericia formula ($QTcF = QT/(RR)^{1/3}$)
RBC	Red blood cell
SAP	statistical analysis plan

SAS	Statistical Analysis Software
TEAE	treatment-emergent adverse event
UFL	upper facial line

4.0 INTRODUCTION

This statistical analysis plan (SAP) provides a more technical and detailed elaboration of the statistical analyses of the efficacy and safety data as outlined and/or specified in the most recent protocol of Study MT10109L-006 (version dated Oct 2020). Specifications of tables, figures, and data listings are contained in a separate document.

This is a multicenter, randomized, double-blind, placebo-controlled, parallel-group Phase 3 study. Participants are adults at least 18 years of age with bilaterally symmetrical moderate to severe lateral canthal lines (LCL) at maximum smile [redacted] and moderate to severe glabellar lines (GL) at maximum frown (assessed by the investigator).

The length of this study will be approximately 12 months. On Day 1, participants will be randomly assigned in a 2:2:1 ratio to receive MT10109L 24 U (MT10109L in LCL area, placebo in GL area), MT10109L 44 U (24 U in LCL area and 20 U in GL area), or placebo (placebo in both LCL and GL areas). [redacted]

After the first treatment on Day 1, all participants will be evaluated for safety and efficacy at follow-up visits occurring on Day 7, 14, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330, and 360 [redacted]

Table 4-1. Schedule of Evaluations: Study MT10109L-006

	All Participants						Only Participants Meeting Retreatment Criteria[a]				All Participants
Study Period	Screening[b]	Randomization Treatment (Day 1)[c]	Follow-up (Day 7 and Day 14)	Days 30, 60, 90, 120, 150	Day 180	Days 210, 240, 270, 300, 330	Retreatment 1 and Retreatment 2 (from Day 180 through Day 330)[c,d,e]	Follow-up (Day 7 and Day 14 after each retreatment)	Follow-up (Day 30 after each retreatment)[e]	Follow-up (Day 90 after each retreatment)[e]	End of Study (Day 360 or Early Exit)
Visit Windows	*Days −28 to −1*	-	*± 3 days*	*± 7 days*	*± 7 days*	*± 7 days*	-	*± 3 days*	*± 7 days*	*± 7 days*	*± 7 days*
Consent/ Authorization	X										
Demographics, Weight/ Height	X										
Inclusion/ Exclusion Criteria	X	X									
Medical/ Surgical History	X										
[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]		[redacted]									
[redacted]			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FWS Assessment of LCL (Investigator and Participant) and GL (Investigator Only)[f]	X[g]	X[g]	X	X	X	X	X	X	X	X	X
Standardized Facial Photography		X	X	X	X						

	All Participants						Only Participants Meeting Retreatment Criteria[a]				All Participants
Study Period	Screening[b]	Randomization Treatment (Day 1)[c]	Follow-up (Day 7 and Day 14)	Days 30, 60, 90, 120, 150	Day 180	Days 210, 240, 270, 300, 330	Retreatment 1 and Retreatment 2 (from Day 180 through Day 330)[c,d,e]	Follow-up (Day 7 and Day 14 after each retreatment)	Follow-up (Day 30 after each retreatment)[e]	Follow-up (Day 90 after each retreatment)[e]	End of Study (Day 360 or Early Exit)
Visit Windows	*Days −28 to −1*	-	*± 3 days*	*± 7 days*	*± 7 days*	*± 7 days*	-	*± 3 days*	*± 7 days*	*± 7 days*	*± 7 days*
Physical Examination	X										
[redacted]		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Vital Sign Measurements[i]	X	X		X[j]			X		X		X
Pregnancy Test (Urine)[k]		X					X				X
12-lead ECG[l]	X		X	X[j]	X		X		X		X
Collection of Blood Samples for Hematology and Chemistry Testing	X			X[j]							
[redacted]		[redacted]		[redacted]			[redacted]		[redacted]	[redacted]	[redacted]
Adverse Events	X	X[n]	X	X	X	X	X[n]	X	X	X	X
Concomitant Medications	X	X	X	X	X	X	X	X	X	X	X
Concurrent Procedures	X	X	X	X	X	X	X	X	X	X	X
Study Intervention Injection		X					X				

Note: Participants completing this study through Day 360 will be eligible to enroll into a 2-year open-label extension study (Study MT10109L-004).

ECG = electrocardiogram; FLO-11 = 11-item Facial Line Outcomes Questionnaire; FLSQ = Facial Line Satisfaction Questionnaire; FWS = Facial Wrinkle Scale with Photonumeric Guide; GL = glabellar lines; LCL = lateral canthal lines (or crow's feet lines)

[a] Participants who meet retreatment criteria will receive up to 2 blinded study interventions of the same study intervention that they received on Day 1.

[b] All screening procedures must be completed up to 28 days before Day 1, and the results must be available to the investigator before randomization on Day 1

[redacted]

[d] Retreatment 1 cannot be administered earlier than the Day 180 visit; Retreatment 2 cannot be administered earlier than 84 days since administration of Retreatment 1 and no later than Day 330.

[e] If a procedure or assessment has already been performed as part of a scheduled visit, it must not be repeated.

[f] LCL at maximum smile and at rest; GL at maximum frown and at rest.

[g] To be randomized into the study, participants must meet Inclusion Criterion 2.01 both at the Screening Visit and prior to randomization at the Day 1 Visit.

[h] Neurologic assessment includes the Focused Symptom Questionnaire and Focused Neurologic Examination. On Day 1, both assessments must be performed prior to randomization.

[i] Vital sign measurements are pulse rate, respiration rate, and blood pressure. Participants are to be seated for at least 2 minutes before measurements are collected.

[j] Only on Days 30 and 120

[k] Women of childbearing potential must have a negative pregnancy test. Investigator may ask participant to perform urine pregnancy test at any visit. At each visit, the investigator must discuss compliance with contraceptive use with women of childbearing potential.

[redacted]

[n] Participants must be observed for ≥ 30 minutes after each study intervention for adverse events.

5.0 OBJECTIVES

The objectives of this study are to evaluate the safety and efficacy of MT10109L versus placebo in the treatment of participants with moderate to severe LCL with or without concurrent treatment of GL.

6.0 PARTICIPANT POPULATIONS

6.1 INTENT-TO-TREAT POPULATION

The Intent-to-Treat (ITT) Population will consist of all randomized participants. Efficacy analyses for US FDA will be based on the ITT population.

[redacted]

6.3 SAFETY POPULATION

The Safety Population will consist of all participants who received at least 1 injection of study intervention. The safety analyses will be based on the safety population and will be performed with participants analyzed by their actual treatment received.

7.0 PARTICIPANT DISPOSITION

The number and percentage of participants in 3 of the study populations (Safety, ITT, [redacted]) will be summarized by treatment group and overall; the number of participants screened and randomized will be summarized by study center.

The number of participants screened will be summarized overall. The number and percentage of participants who were randomized, treated, or who completed the study and prematurely discontinued will be presented for each treatment group and pooled across treatment groups. A frequency table showing participant disposition (continuing, entered to the next cycle, completed, discontinued) will also be provided for all analysis populations for each treatment cycle.

Tabulation of the numbers and percentages of participants in each exit status category (i.e., adverse event, pregnancy, lost to follow-up, personal reasons, protocol violations and other) will be provided for each treatment group for entire study and by treatment cycle for all analysis populations. Discontinued participants will be listed along with the corresponding reason(s) for early withdrawal from the study.

8.0 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

Demographic parameters and baseline characteristics will be summarized by treatment group for ITT [redacted]. Continuous variables will be summarized by number of participants and mean, SD, median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of participants.

Demographic parameters (age; age group (< 65 and ≥ 65); race; ethnicity; sex).

Baseline characteristics (weight; height; and body mass index, calculated as weight [kg]/(height [m])^2, FWS scores of LCL severity at maximum smile and rest (assessed by investigators and participants), GL severity at maximum frown and at rest (assessed by investigators only), [redacted] The distribution of the baseline characteristics will be summarized by treatment group.

The number and percentage of participants with abnormalities in medical and surgical histories in each system organ class (SOC) and preferred term (PT) will be summarized by treatment group for the ITT Population.

Prior medications include all medications taken any time prior to the day 1 baseline visit, whether or not medication is continuing beyond the baseline visit.

Concomitant medications encompass all medicinal products that the participant was taking prior to the day 1 baseline visit which are ongoing at the visit, in addition to all medications that have a start date on or after the day 1 visit date.

WHODrug Global B3 202003 will be used to classify prior and concomitant medications by therapeutic class and drug name.

Both prior and concomitant medications will be coded by drug name and therapeutic class. The use of prior and concomitant medications will be summarized by drug class and drug name in each treatment group for the ITT [redacted].

9.0 EXTENT OF EXPOSURE

Participants' exposure to study intervention will be summarized by total duration of treatment exposure and by cycle [redacted] and by number of treatments received.

10.0 EFFICACY ANALYSES

For US FDA, the primary and secondary efficacy analyses will be based on ITT population. [redacted]

[redacted]

Missing values for the primary measures will be imputed using multiple imputation methods up to day 180 for Treatment Cycle 1. Multiple imputation will be done for the ITT population. [redacted]

The evaluation of the equality of the proportions of responders will be based on Cochran-Mantel-Haenszel (CMH) test stratified by investigator-assessed baseline LCL severity at maximum smile. For analyses using multiple imputations, relative risk and 95% confidence interval for relative risk will be presented. Otherwise Wald confidence intervals for proportions of responders and difference in the proportion of responders will be presented.

Continuous descriptive statistics include: N1 (number of participants with non-missing values at both baseline and the specified post-baseline analysis visit), mean, standard deviation (SD), median, minimum, and maximum.

Categorical variables will be summarized by number and percentage of participants.

Efficacy analyses will be done for all visits. The primary timepoint is day 30 of Treatment Cycle 1.

The 95% confidence intervals for the treatment difference in response rates and relative risk will be summarized using Proc MIANALYZE procedure. The p-value for CMH test will be presented as well. For CMH test, majority rule will be used to determine the significance of treatment-by-baseline-severity interaction. If the Breslow-Day tests for at least 3 out of the 5 MI dataset have significant p-values, then the treatment-by-baseline-severity interaction is considered significant.

10.1 PRIMARY EFFICACY PARAMETERS

US FDA

For the US FDA, the composite primary efficacy endpoint is proportion of participants with a $\geq$ 2-grade improvement from baseline on the Facial Wrinkle Scale with Photonumeric Guide (FWS) according to investigator and participant assessments of LCL severity at maximum smile at Day 30 of Treatment Cycle 1 after a single intramuscular (IM) injection of MT10109L or placebo in the LCL.

The primary analysis will be performed on the ITT population.

The following hypothesis will be used to compare the MT10109L groups with placebo:

- Null hypothesis: MT10109L 24 U and placebo are equally effective in reducing LCL severity at maximum smile as measured by the proportion of responders with a ≥ 2-grade improvement from baseline based on both the investigator- and participant-rated FWS at Day 30 of Treatment Cycle 1.
- Alternative hypothesis: MT10109L 24 U and placebo are not equally effective in reducing LCL severity at maximum smile as measured by the proportion of responders with a ≥ 2-grade improvement from baseline based on both the investigator- and participant-rated FWS at Day 30 of Treatment Cycle 1.

[redacted]

[redacted]

10.2 SECONDARY EFFICACY PARAMETERS

For US FDA:

- The duration of LCL treatment effect, estimated as the median time to return to *moderate* or *severe* LCL at maximum smile in participants who achieved a rating of a ≥ 2-grade improvement from baseline in LCL severity at maximum smile at Day 30 of Treatment Cycle 1 according to investigator assessments using the FWS (this will not be included in the hierarchical testing strategy for multiplicity control)

- The proportion of responders for investigator assessments of LCL severity at maximum smile using the FWS, where a responder is defined as achieving a ≥ 2-grade improvement from baseline at maximum smile at Day 30 of Treatment Cycle 1

- The proportion of responders for investigator assessments of LCL severity at maximum smile using the FWS, where a responder is defined as achieving a rating of *none* or *mild* at maximum smile at Day 30 of Treatment Cycle 1

- The proportion of participants reporting *mostly satisfied*/*very satisfied* on the Facial Line Satisfaction Questionnaire (FLSQ) follow-up version Item 5 for LCL at Day 60 of Treatment Cycle 1

- The proportion of responders for investigator assessments of LCL severity at rest using the FWS among participants who were rated at least *mild* at rest at baseline, where a responder is defined as achieving a ≥ 1-grade improvement from baseline at Day 30 of Treatment Cycle 1

[redacted]

Analyses of the secondary efficacy variables will be performed for all study visits using observed data, with the primary timepoint at day 30 of the Treatment Cycle 1 (except FLSQ Item 5, for which the primary time-point is at day 60 of the Treatment Cycle 1).

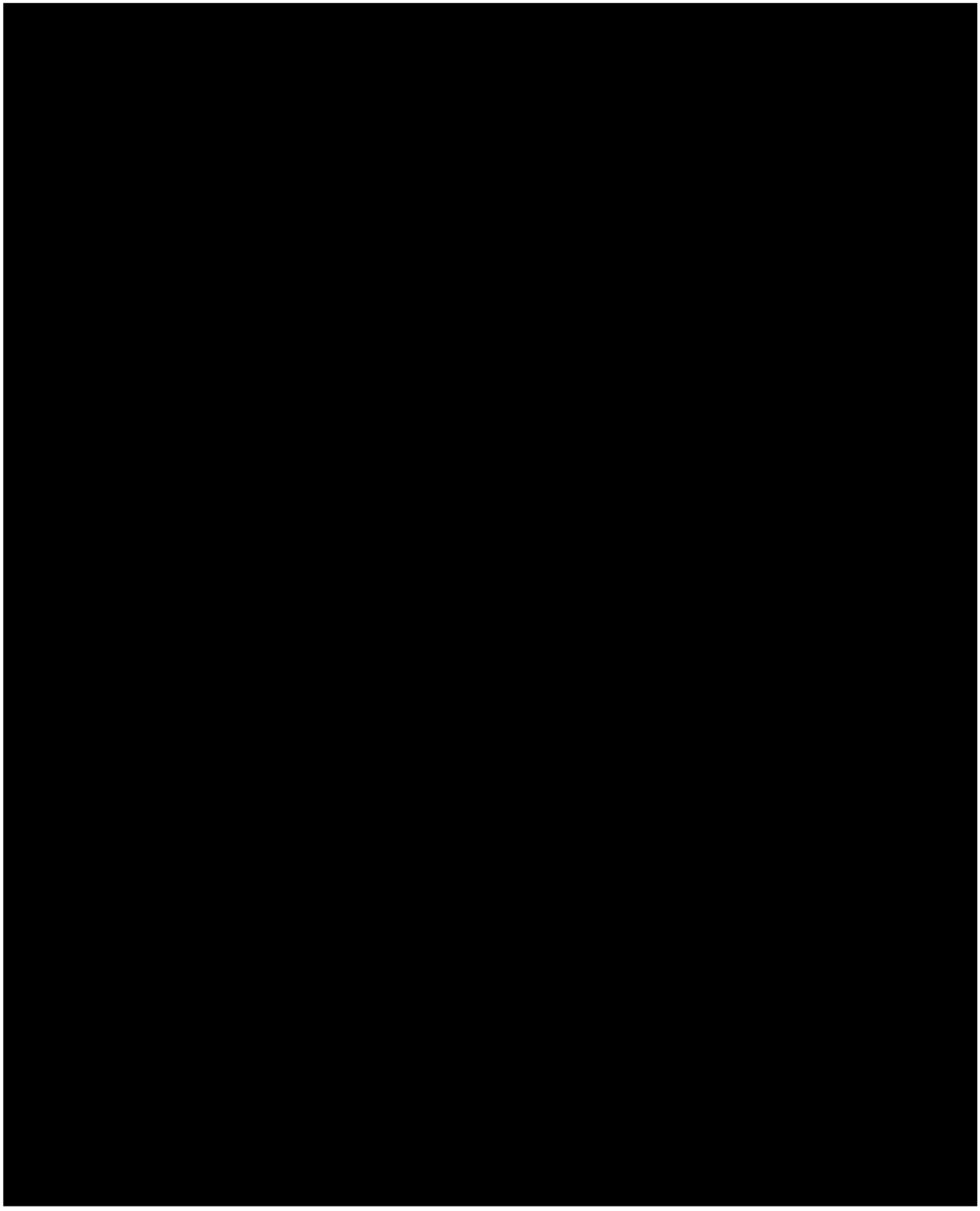
The duration of LCL treatment effect is defined as the time to return to *moderate* or *severe* LCL at maximum smile in those participants achieving a responder status (for US FDA, responder is defined as participants who achieved a rating of a ≥ 2-grade improvement from baseline in LCL severity at maximum smile; [redacted] at Day 30 of Treatment Cycle 1 according to investigator assessments using the FWS. It will be estimated using the Kaplan-Meier method. The 25th percentile, median, and 75th percentile estimates will be calculated. Duration of effect is a key secondary efficacy variable; however, as it is not part of the

hypothesis testing, it will not be included in the hierarchical testing strategy for multiplicity control.

[redacted]

The proportion of responders will be analyzed using the CMH test stratified by baseline LCL severity at maximum smile as assessed by the investigator. The continuous variables and ordinal variables will be analyzed using analysis of covariance (ANCOVA) with treatment as the main effect, baseline LCL severity, and corresponding baseline value as factors.

[redacted]



11.0 SAFETY ANALYSES

The safety analysis will be performed using the Safety Population. The safety parameters will include adverse events (AEs), [redacted] vital sign, electrocardiographic (ECG), Hematology and Chemistry laboratory and Immunogenicity analyses. For each safety parameter of the vital sign, ECG, Hematology and Chemistry laboratory parameters, the last non-missing safety assessment before the first dose of study treatment will be used as the baseline for all analyses of that safety parameter. Continuous variables will be summarized by number of participants and mean, SD, median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of participants.

11.1 ADVERSE EVENTS

AEs will be coded from the verbatim text into preferred term (PT) and the primary system organ class (SOC) by using the MedDRA dictionary. In general, adverse events (AEs) data will be analyzed and presented for:

1. TEAEs: An adverse event will be considered a treatment-emergent adverse event (TEAE) if: : 1) The adverse event began on or after the date of the first study intervention; or 2) The adverse event was present before the date of the first study intervention, but increased in severity or became serious on or after the date of the first study intervention An adverse event that occurs more than 30 days after the study exit will not be counted as a TEAE.

2. TEAEs by cycle: TEAEs by cycle is defined as 1) a TEAE began on or after the date of the most recent of study intervention to the date of next study intervention, or to study exit, or 2) The adverse event was present in the last cycle (or pretreatment), but increased in severity or became serious on or after the date of the most recent study intervention. An adverse event that occurs more than 30 days after the study exit will not be counted as a TEAE.

[redacted]

In each of the analysis periods (entire study, or by treatment cycle), a specific TEAE will only count once per participant, associated with its worst severity during the time period of interest. Unless stated otherwise, the method of analyses described in this section will be applied to each of the screening/baseline and study treatment periods.

Adverse events will be summarized by treatment group for the entire study and by treatment cycle in descending order of incidence rate.

Three incidence rate tables will be presented for summarizing all TEAEs :

1) by descending order of incidence rate

2) by primary SOC and PT

3) by SOC, PT, and severity.

Serious TEAE and treatment-related TEAE will be summarized by primary SOC and PT for entire study. Treatment-related TEAE will also be summarized by related to study drug or related to study procedure. TEAE leading to study discontinuation will be summarized by primary SOC and PT for entire study. [redacted]

A participant listing will be generated for all AEs, SAEs, treatment-related AEs and AEs leading to study discontinuation.

[redacted]

[redacted]

[redacted]

11.2 CLINICAL LABORATORY PARAMETERS

Study baseline is defined to be the data measured before dosing on Day 1. Descriptive statistics (n, mean, standard deviation, minimum, median, and maximum) for clinical laboratory values at baseline, and change from baseline at post-baseline visits will be summarized, in SI units (System of International Units), for all continuous clinical laboratory parameters specified below.

The clinical laboratory parameters include the following:

Hematology: Hemoglobin, hematocrit, red blood cell count (RBC), RBC morphology, white blood cell count (WBC), neutrophils, lymphocytes, monocytes, basophils, eosonophils, and platelets

Chemistry: glucose, creatinine, urea nitrogen, total bilirubin, albumin, aspartate aminotransferase (AST/SGOT), alanine aminotransferase (ALT/SGPT), alkaline phosphatase, uric acid, sodium, potassium, bicarbonate (carbon dioxide content), chloride, phosphorus, calcium, magnesium, and total protein.

A participant level listing for hematology and blood chemistry will be provided.

11.3 VITAL SIGNS

Study baseline is defined to be the data measured before dosing on Day 1. Baseline and change from study baseline data on blood pressure (mm Hg), pulse rate (beats/min) and respirations (breaths per minute) will be summarized as descriptive statistics for each time point and each visit for each treatment group.

11.4 ELECTROCARDIOGRAM

Descriptive statistics for ECG parameters (eg, heart rate, PR interval, RR interval, QRS duration, QT interval, corrected QT [QTc] intervals) at baseline, and changes from baseline at all post-baseline timepoints will be presented for all participants in the Safety Population. For each parameter, only participants who had both baseline post-baseline assessments will be included in the summary. QTc interval will be calculated using both Bazett ($QTcB = QT/(RR)^{1/2}$) and Fridericia ($QTcF = QT/(RR)^{1/3}$) corrections; if RR is not available, it will be replaced with 60/heart rate in the correction formula.

Data listings which include ECG basic parameters and ECG abnormalities will be produced.

11.5 IMMUNOGENICITY ANALYSIS

A 2-stage assay approach will be used for the detection of binding antibodies against MT10109L and neutralizing antibodies against MT10109L in participants' serum. In Stage 1, serum samples will be screened for the presence of binding antibodies using the validated ELISA in a 3-tier format (screening, confirmation, and titering). The screen positive serum samples will be subsequently immunodepleted to confirm that the binding antibodies are specific to MT10109L and titered to assess the extent of antibodies present. In Stage 2, only samples testing positive in the ADA confirmatory assay will be evaluated in the mouse protection assay.

Immunogenicity results, manifested as the presence of binding antibodies to MT10109L, will be summarized in a table. Neutralizing antibodies will be summarized as a tabulation and will be reported as positive, negative, insufficient, or inconclusive as result.

[redacted]

11.7 PREGNANCY TESTS

Urine pregnancy tests are performed prior to treatment and at exit visit for females of childbearing potential. Participants with positive pregnancy test results will be listed by treatment group, including urine sample collection date, and days since most recent treatment.

[redacted]

12.0 INTERIM ANALYSIS

No interim analysis is planned for this study.

13.0 DETERMINATION OF SAMPLE SIZE

Approximately 375 participants will be randomized into the study in a 2:2:1 ratio yielding approximately 150 participants in each of the MT10109L 24 U and 44 U groups and 75 participants in the placebo group in Treatment Cycle 1. The sample size of 375 allows for an adequate safety database of participants treated with MT10109L for these indications. [redacted]

For the primary analysis, an estimated sample size of 214 participants will give a power of greater than 95% to detect a difference in responder rates between the MT10109L 24 U and placebo groups, assuming a 5% drop-out rate by Day 30. [redacted]

0.

[redacted]

14.0 STATISTICAL SOFTWARE

Statistical analyses will be performed using version 9.3 (or newer) of SAS on a Linux operating system.

15.0 DATA HANDLING CONVENTIONS

15.1 VISIT TIME WINDOWS

For ITT and mITT analyses, the following visit windows will be used to determine the visit day assignment, based on the observed number of days relative to the day of the injection (Table 15.1-1):

Table 15.1-1. Visit Time Windows (ITT and mITT Analyses)

Visit Type	*Target Day*	*Visit Window (Day)*
Baseline	Day 1	Prior to Day 1 or on Day 1 (prior to treatment)
Treatment 1 Day 7	Day 7	Day 2-10
Treatment 1 Day 14	Day 14	Day 11-22
Treatment 1 Day 30	Day 30	Day 23-45
Treatment 1 Day 60	Day 60	Day 46-75
Treatment 1 Day 90	Day 90	Day 76-105
Treatment 1 Day 120	Day 120	Day 106- 135
Treatment 1 Day 150	Day 150	Day 136- 165
Treatment 1 Day 180[a]	Day 180	Day 166 -195 or date of 2nd treatment
Treatment 1 Day 210[a]	Day 210	Day 196 – 225 or date of 2nd treatment
Treatment 1 Day 240[a]	Day 240	Day 226 – 255 or date of 2nd treatment
Treatment 1 Day 270[a]	Day 270	Day 256 – 285 or date of 2nd treatment
Treatment 1 Day 300[a]	Day 300	Day 286 – 315 or date of 2nd treatment
Treatment 1 Day 330[a]	Day 330	Day 316 – 345 or date of 2nd treatment
Treatment 1 Day 360[b]	Day 360	Day 346 – 375
Treatment 1 Late exit[c]		Day 376 – exit
Treatment 2		Day 1
Treatment 2 Day 7	Day 7	Day 2-10
Treatment 2 Day 14	Day 14	Day 11-22
Treatment 2 Day 30	Day 30	Day 23-45
Treatment 2 Day 60	Day 60	Day 46-75
Treatment 2 Day 90[a]	Day 90	Day 76-105 or date of 3rd treatment
Treatment 2 Day 120[a]	Day 120	Day 106- 135 or date of 3rd treatment
Treatment 2 Day 150[a]	Day 150	Day 136- 165 or date of 3rd treatment
Treatment 2 Day 180[b]	Day 180	Day 166 –195
Treatment 2 Late exitc		Day 196 – exit
Treatment 3		Day 1
Treatment 3 Day 7	Day 7	Day 2-10

Visit Type	*Target Day*	*Visit Window (Day)*
Treatment 3 Day 14	Day 14	Day 11-22
Treatment 3 Day 30	Day 30	Day 23-45
Treatment 3 Day 60	Day 60	Day 46-75
Treatment 3 Day 90	Day 90	Day 76-105
Treatment 3 Late exit[c]		Day 106 – exit

a Treatment 2 or Treatment 3 can be administered if retreatment criteria are met

b If participant does not qualify for retreatment, he/she will be followed up to day 360 in treatment cycle 1 or up to day 180 in treatment cycle 2.

c Late exit might be due to COVID-19.

If there are values from multiple visits in a given window, the value collected from the visit closest to the target day will be used to represent the window. [redacted]

15.2 MULTIPLE IMPUTATION METHOD

SAS Proc MI procedure will be used to generate 5 imputation data sets. Seed for all Proc MI procedure is pre-specified as [redacted].

Specifically, the missing data [redacted] at visits from baseline up to Day 180 (Period 1) will be imputed using the multiple imputation method as follows: Each efficacy variable will be imputed stratified by treatment group, using the Markov chain Monte Carlo (MCMC) method. Although the MCMC method assumes multivariate normality, inferences based on multiple imputation can be robust to departures from the multivariate normality if the amount of missing information is not large. [redacted] The imputed values will then be rounded to integer values.

The following is a sample SAS code to impute missing data for FWS score.

```
proc mi data=in out=out seed=626122 nimpute=5 minimum=. . . .-0.49 -0.49 -0.49  -0.49  -0.49
-0.49  -0.49  -0.49  -0.49  maximum=. . . . 3.49 3.49 3.49 3.49 3.49 3.49 3.49 3.49 3.49;

by trtcd;

[redacted]
;

run;
```

If the imputation causes an error message in SAS, the minimum and maximum boundary for proc MI will be removed, then in the output dataset, a data step will push those out of range imputed values to the closest value (0 or 3).

After imputation, the changes from baseline values will be calculated, from which the responder status based on FWS scores will be derived for each post baseline visit.

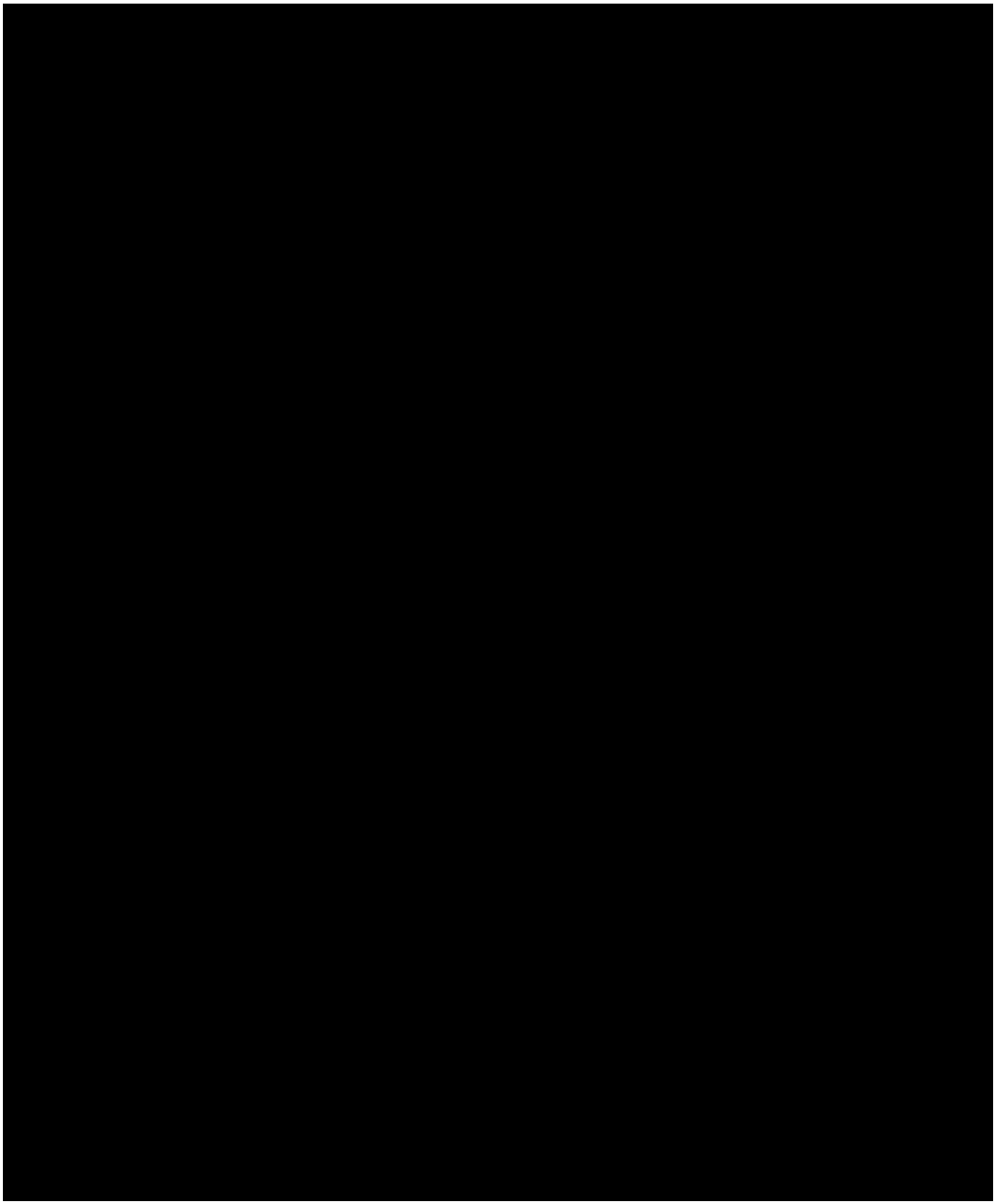
Each of the 5 imputation data sets will be analyzed individually, pooled results of the relative risk estimate (including point estimate and 95% CI) and CMH p-value will be obtained using Proc MIANALYZE procedure, respectively.

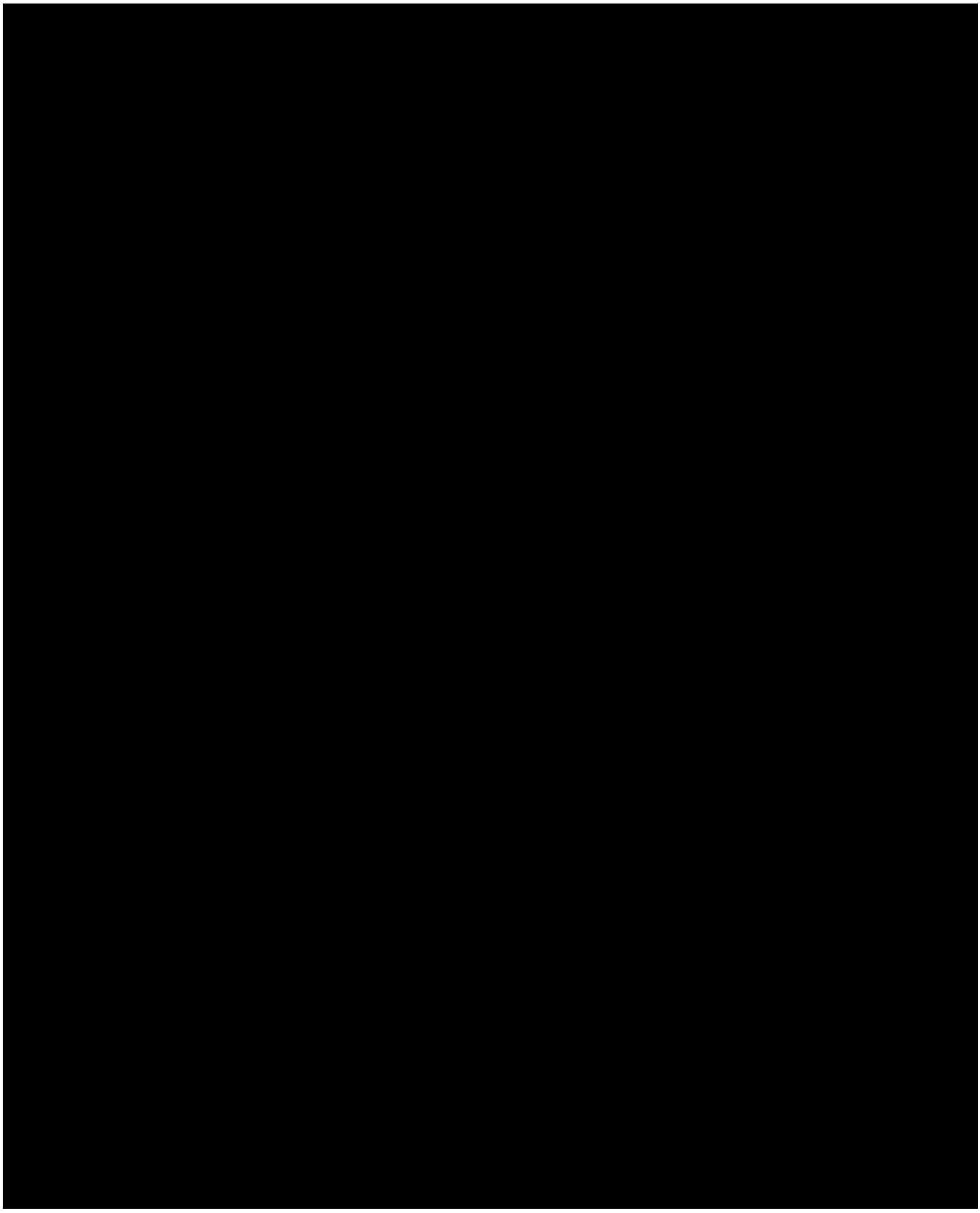
Log of the relative risk stratified by baseline FWS and corresponding variance will be used as input for Proc MIANALYZE. The 95% CI (L, U) for relative risk will be used to calculate the variance for the log relative risk [variance= ((log(U)-log(L))/2*1.96)**2] using normal approximation. The pooled result for log relative risk will be transformed back by antilog to obtain pooled point estimate and 95% CI for relative risk.

To obtain pooled CMH p-value, the Wilson-Hilferty transformation (Wilson 1931, O'Kelly 2014) will be used. Under the null hypothesis, the transformed statistic is approximately normally distributed:

$$\text{wh\_CMH}^{(m)}=\sqrt[3]{\frac{cmh^{(m)}}{k}} \sim \text{Normal}\left(1-\frac{2}{9k}, \frac{2}{9k}\right)$$

Where $cmh^{(m)}$ is the chi-square statistics each with k degrees of freedom associated with the CMH statistic from m = 1, …, 5 imputed datasets. In this case (cross table of two treatment arms and responder status), k=1. This statistic will be transformed to z test statistics. The point estimate of the mean and standard error = 1.0 will be passed to PROC MIANALYZE to obtain the combine p-value for CMH test.





[redacted]

15.5 REPEATED OR UNSCHEDULED ASSESSMENTS OF SAFETY PARAMETERS

If a participant has repeated assessments before the start of the first treatment, the results from the final nonmissing assessment made prior to the start of the study treatment will be used as baseline. If end-of-study assessments are repeated or if unscheduled visits occur, the last nonmissing postbaseline assessment will be used as the end-of-study assessment for generating summary statistics.

15.6 MISSING SEVERITY ASSESSMENT FOR ADVERSE EVENTS

If severity is missing for an AE that started before the date of the first dose of study treatment, an intensity of mild will be assigned. If severity is missing for an AE that started on or after the date of the first dose of study treatment, an intensity of severe will be assigned. The imputed values for severity assessment will be used for the incidence summary; the values will be shown as missing in the data listings.

15.7 MISSING CAUSAL RELATIONSHIP TO STUDY INTERVENTION FOR ADVERSE EVENTS

If the causal relationship to the study intervention is missing for an AE that started on or after the date of the first dose of study treatment, a causality of yes will be assigned. The imputed values for causal relationship to study treatment will be used for the incidence summary; the values will be shown as missing in the data listings.

15.8 MISSING DATE INFORMATION FOR ADVERSE EVENTS

Partial adverse event onset date will be imputed as follows: 1) if day is missing but month is not, impute the date as the first day of the month; 2) if both day and month are missing, impute the date as 01 Jan; 3) if imputed onset date is before the first treatment, yet the corresponding adverse event was not observed pre-treatment, then impute the onset date as the first treatment date. Imputed partial adverse event onset date will only be used to determine the adverse event onset cycle.

Other partial adverse event dates will not be imputed. All partial dates will be listed "as is" in the data listings.

15.9 MISSING DATE INFORMATION FOR PRIOR OR CONCOMITANT MEDICATIONS

There will be no imputation for missing or partially reported medication start and/or end dates. However, reported partial information will be utilized in classification of prior and concomitant medications, when appropriate. For example, a medication may be classified as a pre-study medication if the partial end date provided is determined to definitively have occurred prior to the study treatment date (for example, the partial end date provided is "2017", and Day 1 occurred on August 29, 2018).

If start and/or stop dates for medications are only partially reported but can be classified as having occurred prior to Day 1, then the medications will be included in this summary of prior medications.

If stop dates for medications are only partially reported and cannot be definitively classified as having occurred prior to Day 1, then the medications will be included in this summary of concomitant medications.

16.0 DEVIATION FROM PROTOCOL

[redacted].

17.0 REFERENCES

Dmitrienko A, Molenberghs G, Chuang-Stein C, Offen W. Gatekeeping strategies. In: Analysis of clinical trials using SAS: a practical guide. Cary, NC: SAS Institute; 2005. p. 104-108.

O'Kelly M, Ratitch B. Clinical trials with missing data: A guide for practitioner. Statistics in Practice. Wiley;2014.

Wilson EB, Hilferty, MM. The distribution of chi-squared. Proceedings of the National Academy of Sciences, Washington. 1931;17:12,684 – 688.

18.0 HISTORY OF CHANGES

Amendment 1:

Date	Section(s)	Description
11/18/2020	3.0	Added abbreviation AESI.
11/18/2020	4.0	Changed to base on protocol dated Oct 2020.
11/18/2020	7.0	Deleted word "cumulative" to remove confusion.
11/18/2020	10.1	Clarified missing as non-responder imputation will be conducted up to Day 180 in Cycle 1.
11/18/2020	11.1	Added summary for AESI.
[redacted]	[redacted]	[redacted]
11/18/2020	15.1	Revised window for exit. Added text to explain how to handle multiple visits in same window.
11/18/2020	15.2	Added data step in a specific scenario when value is not in range after iteration when conducting proc MI. Added Wilson-Hilferty transformation detail for CMH test.
[redacted]	[redacted]	[redacted]
11/18/2020	16.0	Added Deviation from Protocol section.
11/18/2020	17.0	Added references.

Electronic Signatures

User	Date	Justification
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]