Statistical Analysis Plan for Study M21-307

Phase 3 Multicenter, Randomized, Double-blind, Placebo-controlled Study of BOTOX (Botulinum Toxin Type A) for the Prevention of Migraine in Subjects with Episodic Migraine

Date: 03 June 2024

Version 4.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for BOTOX (Botulinum Toxin Type A) Study M21-307 "Phase 3 Multicenter, Randomized, Double-blind, Placebo-controlled Study of BOTOX (Botulinum Toxin Type A) for the Prevention of Migraine in Subjects with Episodic Migraine." This study consists of a 4-week screening/baseline phase, followed by a 24-week double-blind treatment phase, and a 24-week open-label phase.

Study M21-307 examines the efficacy and safety of BOTOX in subjects with Episodic Migraine (EM).

The analyses of biomarkers (discussed in Protocol Section 3.6) will not be covered in this SAP.

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

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

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

2.0 Study Objectives and Design

2.1 Study Objectives

The objective for the double-blind phase is to evaluate the efficacy and safety of 2 dose levels of BOTOX compared with placebo as migraine prevention in EM subjects (defined as subjects with 6 to 14 migraine days and < 15 headache days per 28-day period).

The objective for the open-label phase is to evaluate the long-term safety of BOTOX as migraine prevention in EM subjects.

Clinical Hypotheses

- At least 1 dose level of BOTOX is more effective than placebo, as measured by the difference between treatment groups in the change from baseline in the frequency of migraine days per 28-day period
- BOTOX has an acceptable safety profile

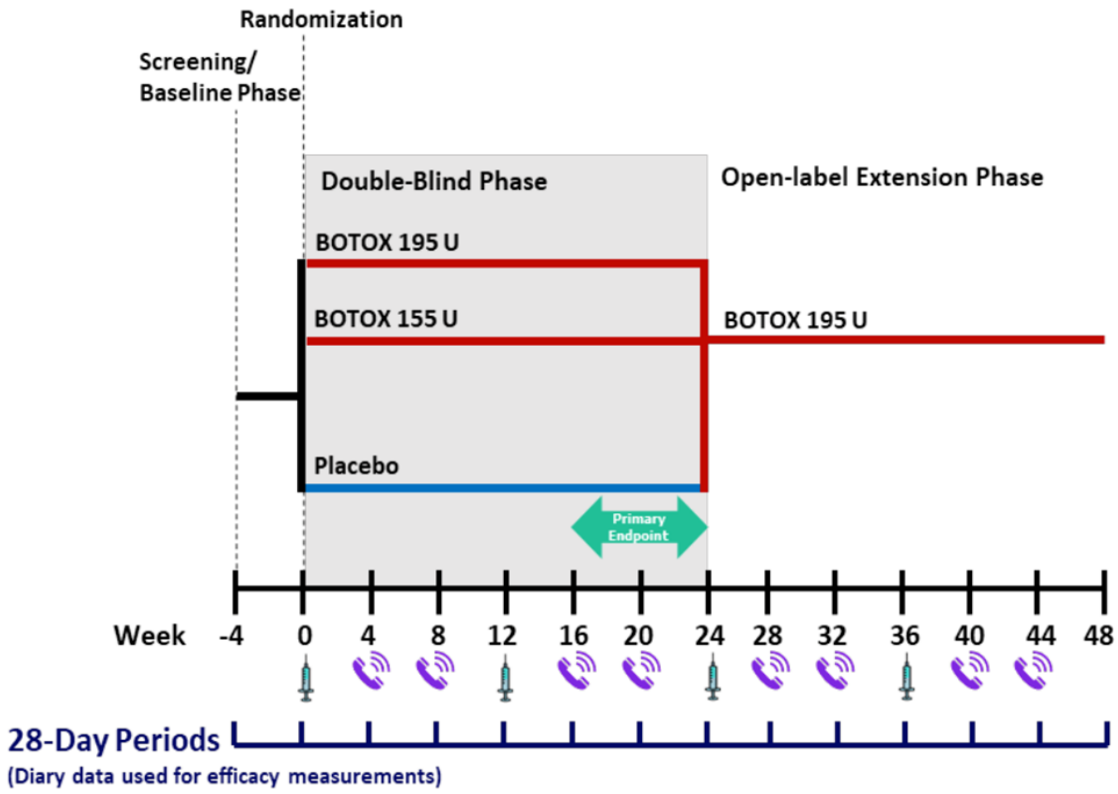
The estimand corresponding to the primary efficacy endpoint is the difference in the change from baseline in the frequency of monthly migraine days across Months 5 and 6 between each BOTOX dose group and placebo in the intent-to-treat (ITT) population.

2.2 Study Design Overview

This is a multicenter, randomized, double-blind placebo-controlled, parallel-group study with an open-label extension phase in adults with EM. Approximately 777 subjects who meet study enrollment eligibility criteria will be randomized in this study. Subjects in this study will be adults 18 to 65 years of age with a history of migraine headache disorder meeting ICHD-3 diagnostic criteria for migraine with aura or migraine without aura (1.1 and 1.2) for ≥ 12 months, with onset before 50 years of age. Specifically, subjects will have 6 to 14 migraine days and < 15 headache days per month in each of the 3 months prior to the screening visit (Visit 1) and during the 4-week screening/baseline phase. See Protocol Section 4.1 for additional information about the enrollment requirements.

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



Screening/baseline phase

This 4-week period consists of 1 screening visit (V1) and ends with the randomization visit (V2, on study Day 1). The baseline diary will include Day -28 through Day -1, which ends the day prior to V2, even if the screening period runs longer than 4 weeks.

Double-blind, placebo-controlled phase

This 24-week period consists of 7 visits (V2, V3, V4, V5, V6, V7, V8). V2 (Day 1) is the randomization visit and subjects will be administered the first blinded treatment. Telephone visits will occur at Week 4 (V3), Week 8 (V4), Week 16 (V6), and Week 20 (V7); a clinic visit will occur at Week 12 (V5) and subjects will be administered the

second blinded treatment. A clinic visit will occur at Week 24 (V8) to end the double-blind placebo-controlled (DBPC) phase and begin the open-label phase.

Open-label phase

This 24-week period consists of 7 visits (V8, V9, V10, V11, V12, V13, V14). At Week 24 (V8), all eligible subjects will receive BOTOX 195 U at 12-week intervals for up to 2 treatment cycles. Telephone visits will occur at Week 28 (V9), Week 32 (V10), Week 40 (V12), and Week 44 (V13); clinic visits will occur at Week 36 (V11) and Week 48 (V14).

The total duration of the study will be up to 52 weeks. The screening visit window is Day -35 to Day -28 and all post-randomization visits have a window of ± 3 days. Diary months will be the 4-week periods relative to the most recent injection (i.e., Days 1-28, 29-56, and 57-84). The third month of a given injection cycle could be truncated early, because it will end the day prior to a subsequent injection cycle, if that subsequent injection occurs prior to Day 85.

2.3 Treatment Assignment and Blinding

Subjects will be randomized to BOTOX 195 U, BOTOX 155 U, or placebo, in a 1:1:1 ratio for a 24-week Double-Blind Treatment Phase.

Randomization will be stratified by [redacted] (yes/no) and ≥[redacted] monthly migraine days at baseline (yes/no) within country. Country is essentially an undeclared stratification factor.

Enrollment will be monitored to target at least [redacted]% of randomized subjects having taken at least [redacted] with proven efficacy (i.e., to target at most [redacted]% without such prior treatment). The list of such medications is in Section 8.1 of the Operations Manual, which is Protocol Appendix F. Similarly, enrollment restrictions will target at least [redacted]% of randomized subjects having ≥[redacted] monthly migraine days during baseline (i.e., to target at most [redacted]% with <[redacted] such days). The population of subjects who

have failed [redacted] preventive treatments will also be monitored throughout the trial. The target is to enroll no less than [redacted]% of the study population who have failed [redacted] preventive treatments. This may result in enrollment restrictions for subjects who failed fewer than [redacted] preventive treatments (i.e., to target no more than [redacted]% of the study population). The study will also aim to enroll no more than [redacted]% of subjects meeting criteria for medication overuse (i.e., to target no less than [redacted]% without overuse) per the investigator's discretion, based on regular overuse of acute headache medication, as defined in the Operations Manual Section 8.2, in the 3 months prior to Visit 1, or history of medication overuse headache diagnosis per ICHD-3 8.2 (ICHD 2018).[1] The study will aim to enroll approximately [redacted]% males (i.e., approximately [redacted]% females) in order to obtain a representative proportion of males with EM (Katsarava 2004).[2] Enrollment restrictions may be implemented to achieve these enrollment targets.

The first day of the blinded study drug administration is defined as Day 1. That will be the case even if it unexpectedly occurs on a day subsequent to randomization, and even though the randomization day would have been deemed Day 1 for purposes of study eligibility decisions and baselines. Although also not expected, if a subject is randomized without being treated, the randomization day will be Day 1.

Subjects who complete the double-blind, placebo-controlled phase may be eligible to continue to the open-label treatment phase of the study at Week 24 (V8). All subjects in the open-label phase will receive BOTOX 195 U at 12-week intervals for up to 2 treatment cycles.

2.4 Sample Size Determination

For the primary endpoint of change from baseline in the frequency of monthly migraine days to the primary time point (8 weeks ending with Week 24), power and sample size calculations assume a 2-sided false-positive error level of alpha 0.05, with multiple comparisons accommodated by gatekeeping.

Under a 1:1:1 randomization ratio, 259 subjects in each of the BOTOX dose groups and the placebo group at the primary time point will provide 90% power to detect a mean difference of [redacted] days between the groups for tests at alpha 0.05 and to detect a mean difference of [redacted] days for any tests done at alpha 0.025. These power calculations used a standard deviation estimate of [redacted], based on data for the month ending with Day 180 in EM Phase 2 studies for the subgroup of subjects with 8 to 14 migraine/probable migraine days at baseline (AbbVie data on file, Study 191622-037/509). Thus, the sample size is based on an effect size of [redacted].

With approximately 125 investigational sites planned, it is recommended that no site randomize more than approximately 12 subjects (i.e., twice the proportional share of 1 more site 2 × [777/126]). Due to enrollment challenges a maximum number of 24 subjects (around 3% of 777 subjects) randomized per site was allowed.

Calculations were done by using the commercial software nQuery Advisor® version 7.0, using a 2-sample Student's t-test with equal variances.

3.0 Endpoints

3.1 Primary Endpoint

The primary efficacy endpoint is the change from baseline in the frequency of monthly migraine days across Months 5 and 6, which are defined as the 28-day daily diary periods ending with Days 56 and 84 after the second study treatment intervention day with BOTOX or placebo injections, for subjects in the intent-to-treat (ITT) population. Baseline is defined as the number of migraine days during the last 28 days prior to the randomization date.

3.2 Secondary Endpoints

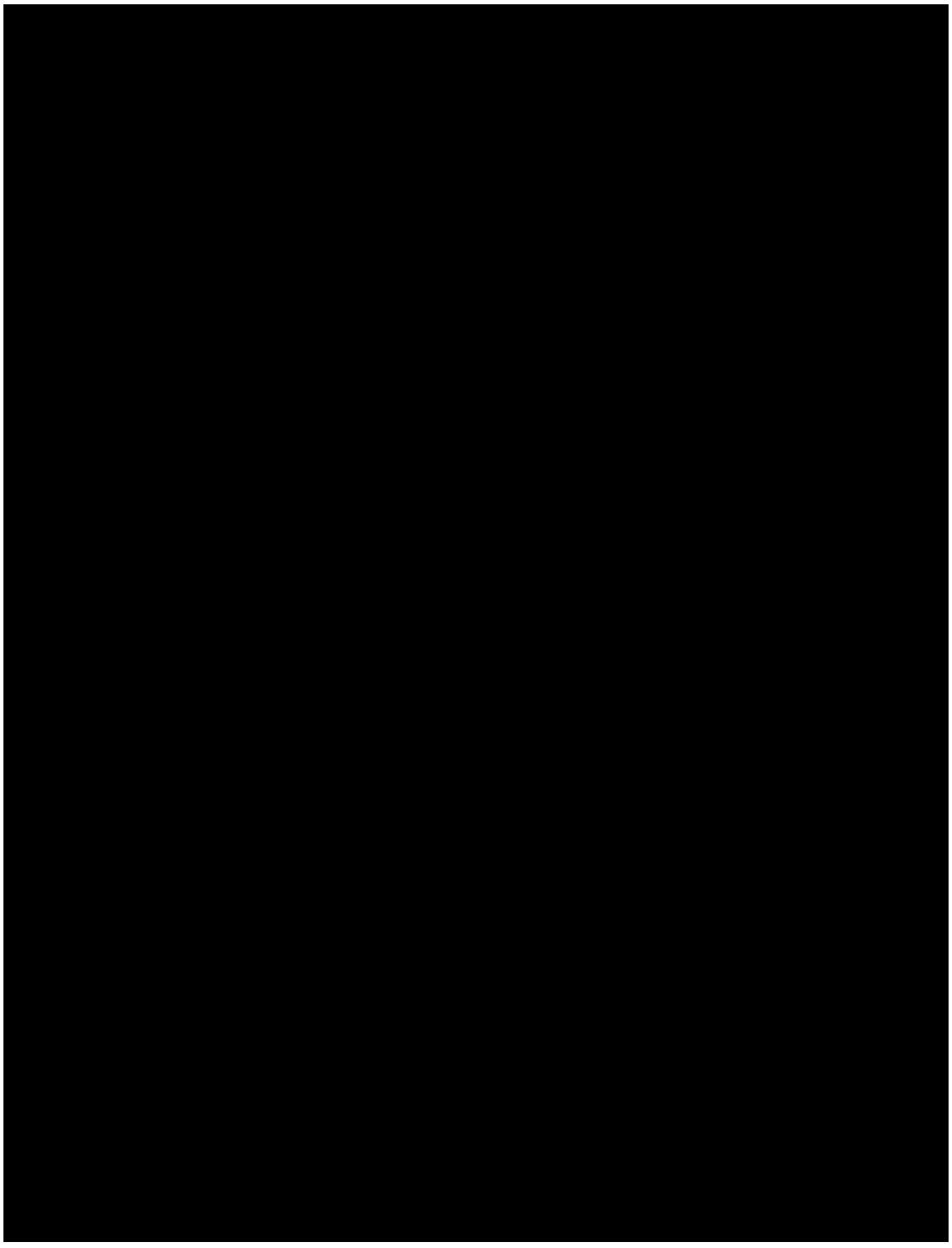
The secondary endpoints of the Double-Blind Treatment Phase will be included in multiplicity adjustment of the Type I error to control the familywise error rate (FWER) at 2-sided significance level of 0.05 for the entire study. Each dose comparison will be

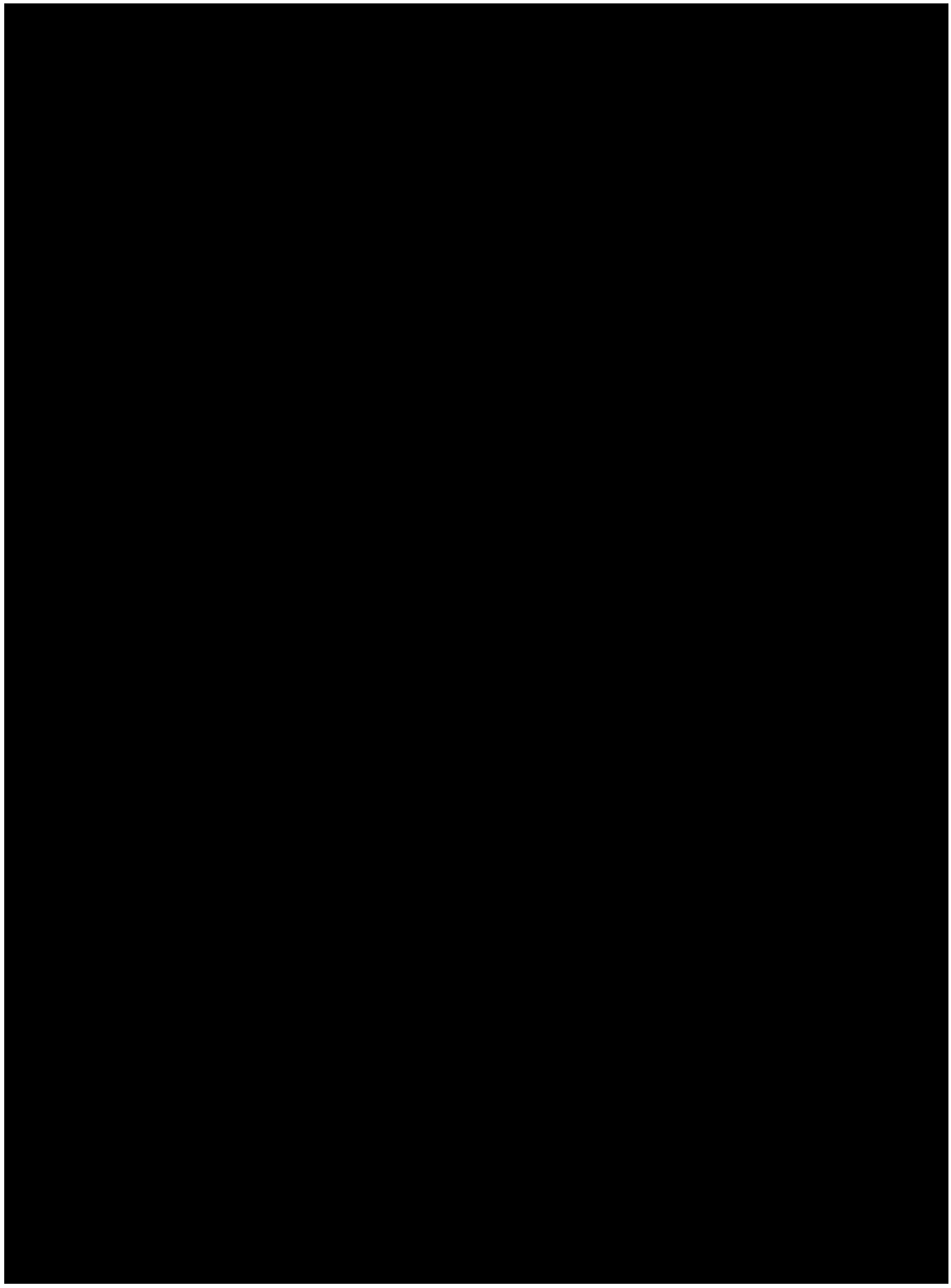
analyzed across primary and secondary endpoints in the order described in Section 13.0, using a ranked-order serial gatekeeping approach to control the overall type I (false positive) error rate.

1. Change from baseline in the frequency of monthly headache days across Months 5 and 6
2. Responder status of 50% reduction from baseline in the frequency of monthly migraine days across Months 5 and 6
3. Change from baseline in the frequency of monthly acute headache medication days across Months 5 and 6
4. Change from baseline in Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQ v2.1) Role Function Restrictive (RFR) domain score at Month 6
5. Change from baseline in the Activity Impairment in Migraine Diary (AIM-D) Physical Impairment (PI) domain score across Months 5 and 6 (total 6-item Headache Impact Test [HIT-6] score for EU)

3.3 Additional Efficacy Endpoints

The primary and secondary efficacy endpoints included in multiplicity adjustment of the FWER are listed in Section 3.1 and Section 3.2, respectively. The primary and secondary efficacy measures will also be evaluated at monthly time points. The additional efficacy endpoints are:





3.4 Safety Endpoints

Safety endpoints are:

- Adverse events (AEs), serious AEs (SAEs) and possible distant spread of toxin (PDSOT)
- clinical laboratory testing (hematology, chemistry, and urinalysis)
- Vital sign measurements
- The Columbia-Suicide Severity Rating Scale (C-SSRS) scores

For clinical laboratory and vital sign parameters, the last non-missing safety assessment before the date of first dose of the study treatment will be used as the baseline for all analyses of that safety parameter.

All AEs and other safety data will be presented in the listings at the subject level.

4.0 Analysis Populations and Data Conventions

4.1 Analysis Populations

The following population sets will be used for the analyses.

The Intent-to-Treat (ITT) population includes all randomized subjects. This will be used as the full analysis set (FAS) population. The ITT population will be used for all efficacy and baseline analyses. Subjects will be included in the analysis according to the treatment group to which they were randomized (as randomized), regardless of study intervention received. (All subjects will have non-missing baseline data for the primary efficacy variable because it is a condition of randomization, and therefore, the ITT population is not being modified with such a constraint.)

The Safety Analysis Set includes all subjects who received any amount of study treatment. Subjects will be included in the analysis according to the study treatment that they actually received (as treated). A subject's actual treatment will be determined by the first dose of study treatment.

4.2 Visit Windows

For data to be summarized on a per visit basis, the visit windows used for analysis are defined below.

Analysis Visit	Scheduled Visit	Nominal Day (Days Post Injection)	Case Report Form Window (Days Post Injection)	Diary Window (Days Post Injection)
Week -4	Visit 1	-28	Baseline Phase	-28 to -1
Week -4.1	Visit 1.1	n/a	Baseline overflow	< -28
Day 1	Visit 2	0	0	0
Week 4	Visit 3	28	1 to 42	1 to 28
Week 8	Visit 4	56	43 to 70	29 to 56
Week 12	Visit 5	84	> 70	57 to 84
Week 12.1	Visit 5.1	> 84	n/a	> 84
Week 16	Visit 6	28	1 to 42	1 to 28
Week 20	Visit 7	56	43 to 70	29 to 56
Week 24	Visit 8	84	> 70	57 to 84
Week 24.1	Visit 8.1	> 84	n/a	> 84
Week 28	Visit 9	28	1 to 42	1 to 28
Week 32	Visit 10	56	43 to 70	29 to 56
Week 36	Visit 11	84	> 70	57 to 84
Week 36.1	Visit 11.1	> 84	n/a	> 84
Week 40	Visit 12	28	1 to 42	1 to 28
Week 44	Visit 13	56	43 to 70	29 to 56
Week 48/Exit	Visit 14	84	> 70	57 to 84
Week 48.1/Exit	Visit 14.1	> 84	n/a	> 84

Note: Visit windows will be based on the time since the most recent injection (0 represents Day 1; negative value means the number of days prior to randomization, i.e., 1 represents the day prior to randomization (Day 1)).

In the protocol, the screening visit window is Day -35 to Day -28 and all post-randomization visits have a window of ± 3 days. The analysis windows are aligned with, but wider than, those specified by the schedule of visits and procedures in the protocol in order to include in analyses all study data that fall outside of protocol visit windows. In the event that two visits for a subject occur within the same visit window, if values observed on different study days or values are observed on the same study day with different time stamp, the last observation chronologically for each study phase (baseline, double-blind and open-label) will be used. If values are observed on the same study day without time stamp or values are observed on same study day with the same time stamp,

the average value will be used for analysis. For laboratory variables with qualitative values, if it cannot be determined which is most recent, then the worst measurement will be used. This rule will be applied separately for each variable for non-missing data only.

Diary data will be assigned per the diary window regardless of the actual visit schedule. Specifically, the diaries will be grouped into 28-day periods for baseline determination prior to randomization and for post treatment injection. This includes the baseline period, defined as the last 28 days of the baseline period (Day -28 to Day -1), and diary months will be the 4-week periods relative to the most recent injection (i.e., Days 1-28, 29-56, and 57-84). The third month of a given injection cycle could be truncated early, because it will end the day prior to a subsequent injection cycle, if that subsequent injection occurs prior to Day 85.

Also, some variables are not scheduled to be recorded at all visits. Visit windows for such visits will be concatenated into the next visit window. For example, for variables that are not to be recorded at Week 4 and Week 8, the analysis window for Week 12 will be widened to include the window for Week 4 and Week 8 (i.e., from Day 2 through the day of the Week 12 injection visit).

4.3 Data Conventions

4.3.1 Baseline and Changes from Baseline

The baseline for diary variables will be the last 28 days of the baseline period. The last non-missing measurement prior to randomization will serve as the baseline for other data, including health outcomes measures (e.g., MSQ and HIT-6). The last non-missing measurement prior to the date of first dose of study treatment will serve as the baseline for laboratory evaluations, vital signs, and C-SSRS.

Changes from baseline will not be tested for statistical significance within treatment groups since the purpose of the study is to make between-treatment comparisons.

4.3.2 Direction of Differences

Treatment differences will be reported as higher dose minus lower dose (e.g., BOTOX 195 U minus BOTOX 155 U, BOTOX 155 U minus placebo and BOTOX 195 U minus placebo). In the open-label phase, all subjects will receive BOTOX 195 U, so the higher minus lower dose will be determined by the DBPC treatment component of the treatment regimen. Thus, treatment differences will mean BOTOX 195 U/BOTOX 195 U minus BOTOX 155 U/BOTOX 195 U, BOTOX 155 U/BOTOX 195 U minus Placebo/BOTOX 195 U, BOTOX 195 U/BOTOX 195 U minus Placebo/BOTOX 195 U (e.g., "BOTOX 195 U/BOTOX 195 U" represents subjects who receive BOTOX 195 U in both the double-blind and open-label phases of the study. "Placebo/BOTOX 195 U" represents subjects who cross over from placebo in the double-blind phase to BOTOX 195 U in the open-label phase.) Changes from baseline will be reported as follow-up minus baseline. A negative score for the change from baseline in frequency of headache days indicates improvement and a positive score indicates worsening.

4.3.3 Substitution for Missing Times and Dates

In general, there will be no substitution of missing time or date. However, any partial information will be utilized to its full extent wherever sensible. For example, a medication may be classified as a prior medication if the partial information of the medication ending date with only month and year permits a determination that the medication ended prior to the injection date of the study medication.

The rules of the substitution of missing times or dates for AEs will follow the Standard Programming Guidelines (SPG) version 2.1.

4.3.4 Descriptive and Inferential Statistical Methods

Data will be summarized with descriptive statistics and/or response frequencies for each treatment group. Descriptive statistics will include sample size, mean, standard deviation, median, Q1, Q3, minimum and maximum. Plots will be provided to enhance the summary of results.

4.3.5 Ordinal and Continuous Variables

The primary analysis of ordinal/continuous variables that are counts based on diary data and some patient-reported outcomes (e.g., MSQ v2.1 and HIT-6), including all primary and secondary variables, will be done by the mixed model for repeated measures (MMRM) of the change from baseline, with the baseline counts or score as a covariate, as indicated in Section 9.3.3, Section 9.4.2 and Section 9.5. The MMRM model will be supplemented by an ANCOVA analysis of the rank of the change from baseline for each post-baseline timepoint, with the unranked baseline score as covariate.

4.3.6 Nominal Variables

For variables where the data is essentially binomial, comparisons between treatment groups will be done with a logistic regression, with the unranked baseline score as covariate. For example, for the binomial response variable of a 50% decrease in migraine days, the covariate will be the frequency of migraine days during baseline.

4.3.7 Rounding

A data value that does not lie directly on a response category (e.g., due to prorating, averaging or imputation) will be rounded to the nearest category. However, the average across Months 5 and 6 will be rounded to tenth.

Probability values will be rounded to three decimal points. If the result is a value of 0.000, it will be displayed as < 0.001. If the result is a value of 1.000, it will be displayed as > 0.999.

5.0 Subject Disposition

The total number of participants who were screened, enrolled (randomized), and treated will be summarized. Enrollment failure subjects (i.e., subjects who consented to participate in the study but were not randomized) and the associated reasons for failure to randomize (e.g., eligibility criteria not met) will be tabulated for all screened subjects.

A summary of subject accountability by country and investigator will be provided where the number of subjects in each of the following categories will be tabulated for each treatment group:

- Subjects enrolled (randomized) in the study;
- Subjects who were randomized broken out by randomization stratification factors ([redacted] [yes/no] and ≥[redacted] monthly migraine days at baseline [yes/no]);
- Subjects who took at least one dose of study treatment;
- Subjects who completed study treatment;
- Subjects who prematurely discontinued study treatment;
- Subjects in each analysis population, as applicable.

The number and percentage of subjects in the ITT population who prematurely discontinued study treatment will be summarized by primary reason for not completing study treatment overall and by treatment group for double-blind phase and open-label phase, respectively.

The number and percentage of subjects in the ITT population who prematurely discontinued study will be summarized by primary reason for not completing study overall and by treatment group for double-blind phase and open-label phase, respectively. For analysis purposes, subjects who complete the Week 24 visit and remain in the study regardless of whether or not they receive further treatment, will be considered as having entered the open-label phase.

"Actual" stratification will be re-derived using data from eDiary and eCRF. A summary table and list of participants with an inconsistent randomization stratum against IRT will be provided.

6.0 Study Treatment Duration and Compliance

For the Safety Analysis Set, duration of treatment exposure will be summarized for each treatment group and for total BOTOX group. Duration of treatment exposure is defined for each subject as last dose date minus first dose date + the minimum of [84 days or (exit visit date last dose date plus 1)]. For this purpose, the duration is assumed to last 12 weeks after the last dose unless the subject exits the study sooner. This applies to the placebo group as well as to BOTOX groups. Duration of treatment exposure will be summarized using the number of subjects treated, mean, standard deviation, median, Q1, Q3, minimum and maximum. The number and percentage of subjects with treatment durations will be summarized in 12-week intervals (e.g., weeks: < 12, 12 to < 24, 24 to < 36, 36 to < 48 and ≥ 48). In addition, the number and percentage of subjects treated in each treatment cycle (Day 1, Week 12, Week 24, and Week 36) will be summarized. Also, the number and percentage of subjects who are treated in at least 1, 2, 3 and 4 injection cycles will be summarized.

In addition to each 28-day diary window, daily diary compliance rates will be calculated, summarized as a continuous variable and compared between treatments for each injection cycle and over the whole study for the intent-to-treat population. Diary compliance will be determined by the subject's recording of data for a given day, either on that given day or for that given day during recall data entry the next day.

7.0 Subject Characteristics

Categorical variables will be summarized with the number and percentage of subjects. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, Q1, Q3, minimum and maximum).

7.1 Demographics and Baseline Characteristics

Demographics and baseline or disease characteristics will be summarized descriptively, overall and by treatment group for the ITT population. Unless otherwise specified, baseline is defined as the last non-missing value prior to randomization.

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

Categorical demographic variables include:

- Sex
- Race (White, Black or African American, Asian, American Indian or Alaska Native, and Native Hawaiian or Other Pacific Islander)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Region 1 (US, Ex-US)
- Region 2 (North America, Europe, Other)
- Age (< 40, ≥ 40 years)

The age of onset and time since onset of episodic migraine will be summarized as continuous variables. They will be determined from the onset date of episodic migraine recorded as part of the medical history as compared to the age at screening and the date of the subject's screening visit. The age of onset will also be summarized for ordered categories (< 12, 12-17, 18-39 and ≥ 40). Similarly, the time since onset will also be summarized for ordered categories (< 10, 10-20 and > 20). Baseline distributions will also be summarized for primary and secondary efficacy variables.

7.1.1 [redacted]

The [redacted] measures overall [redacted] and subtypes and will be assessed at screening to characterize baseline [redacted] Cutaneous [redacted] affects individuals with migraine and is associated with the frequency, severity, disability, and associated symptoms of migraine. The [redacted] includes 12 questions about the frequency of various [redacted]

symptoms in association with headache attacks. Each item is scored as 0 (i.e., very rarely or does not apply to me), 1 (less than half the time), or 2 (half the time or more). The [redacted] will be administered on the handheld device provided to the subjects and will be administered only at the screening visit (Visit 1).

The total [redacted] score will be summed from answers to the 12 questions, with total scores ranging from 0 to 24. If responses to at least half (i.e., 6 or more) of the 12 questions are recorded for a subject, the subject's score will be prorated and rounded to the nearest whole number to account for questions with missing response. If there are responses to fewer than 6 questions, the total score for the subject will be set to missing and not be included in summary statistics.

The total [redacted] score will also be collapsed to a 4-point ordinal scale of [redacted] grade: no [redacted] (0-2), mild (3-5), moderate (6-8), and severe (9-24).

The total [redacted] score at baseline will be summarized as a continuous variable, using observed data. The [redacted] grade's response frequencies will also be displayed.

7.1.2 [redacted]

The [redacted] is a self-administered questionnaire in which subjects will be asked the following 3 questions to characterize their baseline [redacted] experience:

1. Have you ever had any of these [redacted] experiences during the hour before the start of your headache pain?
2. In the past 12 months, how often have you had one or more of these [redacted] experiences during the hour before the start of your migraine or severe headache pain?
3. Have you had at least 2 headaches in your lifetime with this/these [redacted] experience(s)?

The summary statistics, including response frequencies as well as distribution parameters of aura experiences of Question (1), will be displayed for treatment groups.

The responses to the Question (2) above will be summarized on a 5-point scale: 0 Never, 1 Rarely, 2 Less than half the time, 3 Half the time or more and 4 With all or almost every headache. The score ranges from 0 to 4, where a lower score indicates a better condition. The summary statistics, including response frequencies as well as distribution parameters, will be displayed.

For question (3), subjects will be dichotomized into two groups (yes/no) for at least 2 headaches in lifetime with this/these [redacted] experience(s), for use in subgroup analyses.

7.2 Medical History and Prior and Concomitant Medications

7.2.1 Medical History

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

Incidences of medical conditions recorded at the Week -4 and Day 1 exams will be summarized and compared between treatments according to medical history-code categories as binary variables.

7.2.2 Prior and Concomitant Medications

Prior and concomitant medications will be summarized separately. A prior medication is defined as any medication taken prior to the date of the first dose of study drug. A concomitant medication is defined as any medication that started prior to the date of the

first dose of study drug and continue to be taken after the first dose of study drug or any medication that started on or after the date of the first dose of study drug. Prior and concomitant medications will be summarized separately. The number and percentage of subjects taking prior and concomitant medications will be summarized by generic drug name, based on the World Health Organization (WHO) Drug Dictionary. The actual version of the WHO Drug Dictionary will be noted in the statistical tables and clinical study report.

7.2.2.1 Prior Medication

Frequency tabulations will be displayed for pre-study headache prophylaxis medications that were to have been stopped at least 28 days prior to the first study visit (Week -4). The tabulations will display the medication incidence by treatment group, grouped by chemical name within drug class.

In addition, subjects will be dichotomized into two groups (yes/no) for a history of use of headache pain prophylactic medication and compared as a binomial variable. The two groups will include (1) naïve subjects, who did not report any history of such medication and (2) history subjects, who did report such history. The reason for discontinuation of prophylactic medication will also be displayed, with reasons including (1) inadequate response, (2) safety/tolerability (including adverse event, intolerability, or toxicity), (3) discontinued to begin new treatment regimen, (4) other and (5) unknown. Any given subject can be counted in more than one of these discontinuation-reason categories.

7.2.2.2 Study Medication

The volume of study medication injected will be summarized totaled across the head/neck injection areas and separately for each head/neck injection area. This will be done for each injection cycle and totaled across injection cycles.

7.2.2.3 Concomitant Medications and Concurrent Procedures

Frequency tabulations will be displayed for concomitant medication incidence by treatment group, grouped by chemical name within drug class. For the primary analysis,

concomitant medications will be summarized for those taken at any time from Day 1 injection through Week 24. For the final analysis, concomitant medications will be summarized for those taken at any time from Day 1 injection through Week 48.

Frequency tabulations will be displayed for concurrent procedures incidence by treatment group. For the primary analysis, concurrent procedures will be summarized for those done at any time from Day 1 injection through Week 24. For the final analysis, concurrent procedures will be summarized for those done at any time from Day 1 injection through Week 48.

7.3 Protocol Deviations

Protocol deviations include eligibility criteria violations, receipt of wrong treatment or incorrect dose of study treatment, development of withdrawal criteria without being withdrawn, and use of prohibited concomitant medications. A listing of subjects with protocol deviations will be provided.

For each of the following protocol deviation categories and other protocol deviation categories specified in study-specific issue management plan and across all categories, the number and percentage of randomized subjects with at least one protocol deviation will be summarized overall and by treatment group:

- Subject entered into the study even though did not satisfy entry criteria;
- Subject developed withdrawal criteria during the study but was not withdrawn;
- Subject received wrong treatment or incorrect dose of study treatment;
- Subject did not receive the Week 12 treatment but continued in the study beyond Week 12 visit;
- Subject took prohibited concomitant medication.

8.0 Handling of Potential Intercurrent Events for the Primary and Secondary Endpoints

The estimand corresponding to the primary objective of this study is displayed in the table below.

ESTIMAND	
Target Study Population: Subjects with EM as constrained by the study inclusion and exclusion criteria; most notably < 15 headache days and 6 to 14 migraine days during the 28-day baseline period that ends the day before randomization. The population of inference will be further constrained by baseline ranges that are narrower or less inclusive than the eligibility criteria (e.g., too few males compared with the target population epidemiology).	Endpoint of Interest: Change from baseline in the frequency of monthly migraine days across Months 5 and 6, which are defined as the 28-day daily diary periods ending with Days 56 and 84 after the second study treatment intervention day with BOTOX or placebo injections, for subjects in the ITT population.
Intercurrent Events: Study discontinuation, especially those due to lack of efficacy or due to AEs. Also, any significant protocol deviation that could affect the primary endpoint, such as taking rescue medication or not receiving the Week 12 treatment. A sensitivity analysis will impute scores differently for such intercurrent events (e.g., reversion toward baseline).	Population-Level Summary of Variable: Between-treatment comparison of the change from baseline in monthly migraine days across Months 5 and 6.

The primary analysis assumes that missing data are missing completely at random or missing at random. A sensitivity analysis will use the same model as the primary analysis after imputation of missing data by the modified last observation carried forward (mLOCF) rate-change method, which adjusts the most recent observation by the average change since that observation within the same treatment regimen. An additional sensitivity analysis, which is aimed at intercurrent events, will instead impute missing data by reversion toward baseline for subjects who have intercurrent events that could affect the primary endpoint.

9.0 Efficacy Analyses

9.1 General Considerations

All efficacy analyses will be conducted on the ITT population, in general. All efficacy analyses discussed in this section will be performed as the primary analysis for the double-blind phase of the study and also as the final analysis for the open-label phase.

A 2-sided test with a p-value ≤ 0.05 will be considered as statistically significant for the treatment group comparisons of each BOTOX dose (BOTOX 195 U and BOTOX 155 U) versus placebo. Interaction effects will be examined at the 0.10 level by including treatment group-by-country interaction in a separate expansion of the primary Mixed-Effect Model Repeated Measurement (MMRM) model.

The primary analysis will be performed after all ongoing subjects have completed the double-blind phase and entered a final open-label phase. Subsequently, a final analysis will be performed at the end of the open-label phase. A database lock will occur, and randomization data will be released when all ongoing subjects have completed the double-blind phase through the Week 24 visit. The planned primary analysis will be performed at that time. The final database will be locked when all subjects have completed the open-label phase of the study. For analysis purposes, subjects who complete the Week 24 visit and remain in the study regardless of whether or not they receive further treatment, will be considered as having entered the open-label phase. The planned final analysis will be performed after the final database is locked. The final analysis will include data from both the double-blind phase and the open-label phase of the study.

Baseline for efficacy analyses is defined as the last non-missing efficacy assessment before randomization. For the primary efficacy variable and similar variables based on headache diary data, baseline will be the count of affected days during the 4 weeks ending the day before randomization. The last non-missing measurement prior to randomization will serve as the baseline for other data, including health outcomes measures.

For analyses by stratification factors used for randomization, all subjects will be analyzed according to the actual stratum to which the subject belongs, instead of randomized stratum, unless otherwise indicated.

9.2 Handling of Missing Data

The primary efficacy variable is the frequency of monthly migraine days across Months 5 and 6, which are defined as the 28-day daily diary periods ending with Days 56 and 84 after the second study treatment intervention day, compared to the last 28 days of the baseline phase.

All subjects will have non-missing baseline data for the primary efficacy variable, because it is a condition of randomization, and therefore, the ITT population is not being modified with such a constraint.

In order to have data for all subjects in the ITT population for the ITT analysis, missing values (due to having reported less than 14 days of headache data in a 28-day period) will be handled using the following methods for the primary efficacy analysis. Details can be found in Section 9.3.3 and Section 9.3.4.

- Mixed model for repeated measures (MMRM) based on observed data collected at baseline and each month of the double-blind, placebo-controlled phase (i.e., months ending with Weeks 4, 8, 12, 16, 20, and 24). A few subjects (less than 1%) who did not have post-baseline scores will not be included in the primary MMRM due to having missing values for the dependent variable.
- A sensitivity analysis will use the same MMRM model after imputation of missing data by the mLOCF rate-change method, which adjusts the most recent observation by the average change since that observation within the same treatment regimen
- A sensitivity analysis will use the same MMRM, after imputation of missing values by mLOCF, but with imputation instead by reversion toward baseline for intercurrent events that could affect the primary endpoint.

- A sensitivity analysis will use the same MMRM model, after imputation of missing data by the multiple imputation (MI) method.

9.3 Primary Efficacy Endpoint and Analyses

9.3.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the change from baseline in the frequency of monthly migraine days across Months 5 and 6, which are defined as the 28-day daily diary periods ending with Days 56 and 84 after the second study treatment intervention day with BOTOX or placebo injections, for subjects in the intent-to-treat (ITT) population. Baseline is defined as the last non-missing number of migraine days during the last 28 days prior to the randomization date.

9.3.2 Collection and Derivation of Primary Efficacy Assessment

On a daily basis, subjects are to report information on the cumulative duration of any headaches that day and the previous day, along with associated headache specific characteristics and symptoms and use of any acute headache pain medication. (Note: Headache characteristics and symptoms are described in Section 9.3.2.1.) Subjects will be able to report headache data, including absence of headache, for each report date and for the day immediately preceding the report date. This is defined as a 1-day "missing-recall" window.

On a weekly basis, the subject will also be asked to enter the number of migraine attacks he or she experienced over the past 7 days, at the end of each 7-day period since the most-recent injection. The counts for a given week will be prorated to 28-day counts before calculating the change from baseline.

In practice, there may or may not be an exact 4-week (i.e., 28-day) duration between two consecutive visits. For data analysis purposes, the number of headache days during the last 28 days of the baseline phase will serve as the "baseline" for calculating change from baseline for 28-day periods subsequent to each injection. If the baseline phase starting

with the screening visit and preceding the first injection exceeds 28 days, then the baseline period will only include the last 28 days. For 28-day periods that occur at the end of any injection cycle, the 28 days ending with Day 84 post injection will be counted, unless truncated by an early injection visit; any days beyond Day 84 and prior to re-injection will be in an overflow period that will not be included in efficacy analyses.

In order to be randomized, a subject is supposed to have been in the baseline phase for at least 28 days and must have reported diary data for at least 20 days (including missing recall) during the 28-day baseline period. If less than 28 days of baseline data are reported, the number of headache days and other such counting variables for "baseline" will be prorated according to the method in Section 9.3.2.5. Subsequent to each injection, the number of headache days will be counted in successive and non-overlapping 4-week (i.e., 28-day) diary windows. Headache attacks that continue into a subsequent 4-week period will be counted (with recorded severity and duration) as occurring on the headache days contained in each period.

For primary and secondary efficacy variables, in order to have data for all subjects in the ITT population for the ITT analysis, missing diary values (due to having reported less than 14 days of headache data in a 28-day period) will be estimated according to the primary imputation methods discussed in Section 9.3.2.7.

9.3.2.1 Definition of Headache, Migraine, and Probable Migraine Days

A headache day is determined by the recording in the electronic diary of headache specific characteristics and/or symptoms for a calendar day. In general, the duration of headache must be at least [redacted] hours that day in order to be analyzed as a headache day, but this definition might be altered for some sensitivity analyses (to [redacted] hours or [redacted] minutes, for example). The duration constraint is removed if an acute headache medication is taken that day.

A migraine day is a headache day of at least [redacted] hours duration containing at least one set of headache characteristics and symptoms that fit the following definition of a migraine,

where a "set" refers to headache data reported for the same diary report. The four migraine characteristics are (1) unilateral location (rather than bilateral), (2) pulsating/ throbbing quality (rather than pressing/squeezing), (3) moderate or severe pain intensity (rather than mild) and (4) aggravated by or causing avoidance of routine physical activity (e.g., walking or climbing stairs). The three symptoms of migraine are (1) nausea and/or vomiting, (2) both photophobia and phonophobia and (3) typical aura (i.e., visual, sensory, or speech/language) accompanying or within 60 minutes before headache begins. A diary report must simultaneously include at least 2 of the 4 characteristics AND at least 1 of the 3 symptoms in order for a headache day to be counted as a migraine day. The duration constraint is removed if an acute migraine medication is taken that day.

A "probable migraine" day is a headache day of at least [redacted] hours duration for which a diary report must simultaneously include (1) at least 2 of the 4 characteristics and none of the symptoms or (2) exactly 1 of the 4 characteristics and at least 1 of the 3 symptoms. The duration constraint is removed if an acute migraine medication is taken that day.

A "migraine/probable migraine" day is a headache day of at least [redacted] hours duration that is either a migraine day or a probable migraine day. The duration constraint is removed if an acute migraine medication is taken that day. While days of migraine, probable migraine and migraine/probable migraine will be calculated from the data, migraine/probable migraine days are referred to as migraine days in the protocol.

9.3.2.2 Missing Headache Diary Days

Days will only be directly captured as headache-free days or headache days if they are reported on the calendar day or within a 1-day missing-recall window, as defined in Section 9.3.2. Outside such 1-day missing-recall windows, days will only be counted indirectly when pro-rating and/or imputation applies, as indicated in Section 9.3.2.5 and Section 9.3.2.7.

9.3.2.3 Definition of Headache Day, Headache-Free Day and [redacted]

A headache day will be defined as a calendar day for which a subject reports at least [redacted] hours of headache. The [redacted]-hour constraint is removed if an acute headache medication is taken that day. This definition will be altered for sensitivity analyses (to [redacted] hours and [redacted] minutes of headache).

A migraine day will be defined as a calendar day for which a subject reports at least [redacted] hours of migraine. A "probable migraine" day will be defined as a calendar day for which a subject reports at least [redacted] hours of a probable migraine. The [redacted]-hour constraint is removed if an acute migraine medication is taken that day. A "migraine/probable migraine" day will be defined as a day that is a migraine and/or probable migraine day.

A single headache day will be counted regardless of the number of headache attacks during that day. This applies separately to all types of headache days (e.g., headache days, migraine days, probable migraine days and migraine/probable migraine days).

For headache days of 1 or more total hours, a headache-free day is defined as a calendar day with total hours of headache recorded in the headache diary as 0 or less than 1 total hour without headache medication use, or reported as a day without headache. For headache days of any duration, a headache-free day is defined as a calendar day with total hours of headache recorded in the headache diary as 0 or reported as a day without headache. For headache days of [redacted] or more total hours, a headache-free day is defined as a calendar day with total hours of headache recorded in the headache diary as 0 or less than [redacted] total hours without headache medication use, or reported as a day without headache.

A [redacted]. Calendar days begin at midnight and last until 11:59 PM (23:59).

9.3.2.4 Missing-Recall Data

During a diary entry, subjects will be queried about total duration of headaches that day and any associated migraine characteristics and symptoms and headache medication. First though, they will be similarly queried about headaches that occurred during the immediately preceding calendar day, which is the missing-recall window.

Suppose that a subject completes e-Diary on Sunday, 12/12/21 and reports a headache duration [redacted] hours with associated characteristics or symptoms. Then the subject doesn't complete the e-Diary again until Wednesday, 12/15/21, and retrospectively reports no headache occurred on Tuesday, 12/14/21 or Wednesday, 12/15/21 prior to completing the e-Diary. In that case, 12/12/21 would be reported a headache day (as the duration of the headache was [redacted] hours or longer on that day), while 12/13/21 would be captured as a missing diary-data day, because it was not in the 1-day missing-recall window. Since no headache involving 12/14/21 was reported when completing the e-Diary, it would be counted as a headache-free day. Then, the subject completes the e-Diary on Thursday, 12/16/21, and recalls a headache started on 12/15/21 and lasted 6 hours and a headache didn't occur on 12/16/21. Thus, 12/15/21 would be counted as a reported headache day. That is, just as the missing-recall day can be captured as headache days, they can also be captured as headache-free days if there is no data that the subject had a headache day of sufficient duration. Similarly, if the subject does not complete the Diary again until Sunday, 12/19/21 and reports that a headache duration lasting 5 hours occurred on Saturday, 12/18/21, in this case, 12/18/21 will be counted as a reported headache day. Thus, the seven-day period from 12/12 through 12/18 includes three headache days, one of which is a directly-reported headache day (12/12) and two of which are missing-recall headache days (12/15 and 12/18). The seven-day period also includes two headache-free days, one of which is a missing-recall headache-free day (12/14) and one of which is a directly-reported headache-free day (12/16). The seven-day period also includes two days with missing data (12/13 and 12/17). The pro-rated and rounded count of headache days for the subject would be (3/5)*7 4 headache days per week and (3/5)*28 17 headache days standardized to a 28-day period, as described in the prorating section below.

9.3.2.5 Prorating

For the monthly counts, if there are at least 14 days for which the subject has reported headache data (for either headache days or headache-free days), but less than 28 days in a diary period, the data will be prorated accordingly and rounded to the nearest whole number. The prorating will be based on the number of days with reported data in that 28-day period.

For example, if there are 24 days with reported data in the third diary period, the day counts will be multiplied by 28/24 and rounded to a whole number. Then, if the subject's counts for the 28-day period included 14 headache days and 10 headache-free days, after prorating by 28/24, the subject's counts for the 28-day period would be 16 headache days and 12 headache-free days.

For the weekly counts, if there are at least 4 days for which the subject has reported headache data (for either headache days or headache-free days), but less than 7 days in a diary period, the day counts will be prorated accordingly and rounded to the nearest whole number. The prorating will be based on the number of days with reported data in that 7-day period.

For example, if there are 5 days with reported data in a given 7-day diary period, the day counts will be multiplied by 28/5 and rounded to a whole number. Then, if the subject's counts for the 7-day period included 2 headache days and 3 headache-free days, after prorating by 28/5, the subject's counts for the 7-day period would be 11 headache days and 17 headache-free days.

9.3.2.6 Truncation

If the baseline period starting with the screening visit and preceding the first injection exceeds 28 days, then the baseline period will only include the last 28 days. For 28-day periods that occur at the end of an injection cycle, the 28 days ending with Day 84 will be counted, unless truncated by an early injection visit; any days beyond Day 84 and prior to re-injection will be in an overflow period that will not be included in efficacy analyses.

9.3.2.7 Imputation Rule

In this study, the imputation methods are only intended for the diary data collected in the e-diary and secondary variables (i.e., MSQ v2.1 RFR domain score, AIM-D Physical Impairment domain score, HIT-6 score). No imputation will occur for other variables.

If a subject reports any diary data for less than 14 days of a 28-day period, the subject's scores (such as headache day count) for that period will be set to missing for summary tables of observed data. For tables of imputed data, such a subject's score for that period will be imputed by a modified last observation carried forward (mLOCF) rate-change method and rounded to the nearest whole number. Specifically, the substitution will be the subject's previous 28-day period score multiplied by the ratio of the mean within the same treatment regimen in the 28-day period of interest to the mean within the same treatment regimen in the previous 28-day period.

If a subject reports diary data for at least 14 days of a 28-day period, the scores for the period will be determined by the simple prorating method (discussed in Section 9.3.2.5).

If a subject reports any diary data for less than 4 days of a 7-day period, the subject's scores (such as headache day count) for that period will be set to missing for summary tables of observed data. For tables of imputed data, such a subject's score for that period will be imputed by a modified last observation carried forward (mLOCF) rate-change method and rounded to the nearest whole number. Specifically, the substitution will be the subject's previous 7-day period score multiplied by the ratio of the mean within the same treatment regimen in the 7-day period of interest to the mean within the same treatment regimen in the previous 7-day period.

If a subject reports diary data for at least 4 days of a 7-day period, the scores for the period will be determined by the simple prorating method (discussed in Section 9.3.2.5).

For example, for imputation of the 28-day period, the substitution will be iterative by 28-day period, in that imputation of a given 28-day period (e.g., Week 8) will follow

imputation of the preceding 28-day period (e.g., Week 4). The following is an example of how the mLOCF algorithm works:

Observed data:

Subject #	Treatment	Headache Day Count				
		Baseline	Week 4	Week 8	Week 12	Week 16
100001	TRT-A	14	12	13	---	---
100002	TRT-A	10	9	9	8	6
100003	TRT-A	12	13	10	8	5
100004	TRT-A	6	6	5	6	4
Mean	TRT-A	10.5	10	9.25	7.33	5

The Subject 100001 missing headache day count at Week 12 will be estimated by 13 * (7.33/9.25) 10.3, rounded to 10. So, the data set becomes:

Modified LOCF (mean-change adjusted LOCF) imputation:

Subject #	Treatment	Headache Day Count				
		Baseline	Week 4	Week 8	Week 12	Week 16
100001	TRT-A	14	12	13	***10***	---
100002	TRT-A	10	9	9	8	6
100003	TRT-A	12	13	10	8	5
100004	TRT-A	6	6	5	6	4
Mean	TRT-A	10.5	10	9.25	8	5

To estimate the headache day count of Week 16 for Subject 100001, this imputed value (10) will be used to update the overall mean for Week 12, which becomes (10+8+8+6)/4 8. Then the headache day count for Subject 100001 of Week 16 will be estimated by 10* (5/8) 6.25, rounded to 6. So, the data set becomes:

Modified LOCF (mean-change adjusted LOCF) imputation:

Subject #	Treatment	Headache Day Count				
		Baseline	Week 4	Week 8	Week 12	Week 16
100001	TRT-A	14	12	13	10	***6***
100002	TRT-A	10	9	9	8	6
100003	TRT-A	12	13	10	8	5
100004	TRT-A	6	6	5	6	4
Mean	TRT-A	10.5	10	9.25	8	5.25

9.3.3 Main Analysis of Primary Efficacy Endpoint

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

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

Estimand Label	Attributes of the Estimand				
	Treatment	Endpoint	Population	Handling of Intercurrent Events	Statistical Summary
Primary	Arm A: Placebo Arm B: BOTOX 155 U Arm C: BOTOX 195 U	Change from baseline in the frequency of monthly migraine days across Months 5 and 6	The Intent-to-Treat (ITT) Population	IE1: Missing data due to study discontinuations will not be imputed and will be handled as MAR with MMRM. IE2: Regardless of whether rescue medications are taken or not, collected data will be included in the analysis. IE3: Subjects who discontinue study treatment due to any reason will have their data collected after discontinuation of study treatment, and collected data will be included in the analysis.	The difference in the change from baseline in the frequency of monthly migraine days across Months 5 and 6 between each BOTOX dose group and placebo

A mixed model for repeated measures (MMRM) analysis will be used for the primary analysis, using data for baseline and each month of the double-blind, placebo-controlled phase (i.e., months ending with Weeks 4, 8, 12, 16, 20, and 24). The statistical model will include treatment (BOTOX 195 U, BOTOX 155 U, and placebo), month, country, strata of [redacted], and treatment group-by-month interaction as fixed effects, with the baseline number of monthly migraine days as a covariate, included as a continuous variable rather than as the binomial stratification variable. Subject and residual errors will be random effects. Model parameters will be estimated using a restricted maximum likelihood method incorporating all observed data during the double-blind phase. Unstructured covariance structure will be used. Toeplitz or compound symmetry may be considered if convergency issues arise for the unstructured covariance structure. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom. Contrasts will be constructed to obtain the average treatment effects across Months 5 and 6 to compare each BOTOX dose (195 U and 155 U) to placebo group. The ITT population will be used for primary and sensitivity analyses. A few subjects (less than 1%) who did not have post-baseline scores will not be included in the primary MMRM due to having missing values for the dependent variable. These subjects will be included in sensitivity analyses and have post-baseline scores imputed as described in Section 9.2. The primary comparisons of interest will be the pairwise comparisons of each BOTOX dose (195 U and 155 U) versus placebo across Months 5 and 6.

Each dose comparison will be analyzed across primary and secondary endpoints in the order described in Section 13.0, using a ranked-order serial gatekeeping approach to control the overall type I (false positive) error rate. For each BOTOX dose versus placebo, each variable/dose pair will be examined for significance only if significance was determined for all variable/dose pairs ranked ahead of it.

A 2-sided test with a p-value ≤ 0.05 will be considered as statistically significant for the treatment group comparisons of each BOTOX dose (BOTOX 195 U and BOTOX 155 U) versus placebo.

Interaction effects will be examined at the 0.10 level by including treatment group-by-country interaction in a separate expansion of the primary MMRM model. If there is a significant interaction ($p \leq 0.10$) in the expanded MMRM model, the relevant reduced model will be examined for interpretation (i.e., by country).

9.3.4 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoint

A sensitivity analysis will use the same MMRM model after imputation of missing data by the mLOCF rate-change method, which adjusts the most recent observation by the average change since that observation within the same treatment regimen. This method has been chosen for consistency with methods used for the Phase 3 CM studies of BOTOX.

An additional sensitivity analysis will use the same MMRM, after imputation of missing values by the mLOCF rate-change method, but with imputation instead by reversion toward baseline for post-baseline data (including missing and non-missing observed data) after the date of intercurrent events (e.g., took rescue medications, skipped Week 12 injection but staying in the study, exit due to lack of efficacy, due to adverse events or due to lost to follow-up) that could affect the primary endpoint.

In order to examine the sensitivity of the assumption, which relies on a missing at random/missing completely at random missing-data mechanism (MAR/MCAR), a sensitivity analysis will instead impute a subject's missing scores as the average non-missing score across all previous 28-day periods, including baseline, for the subject. This is defined as "reversion toward baseline" imputation. The reason for this adjustment is that such missing scores might be non-ignorable due to a missing not at random mechanism (MNAR) and the mLOCF assumption of continued effect could be misleading for such subjects, especially regarding the placebo component of any improvement from baseline. This sensitivity imputation will likely dampen down the mean improvement from baseline within treatment groups, but the interest will be in how it affects the treatment comparisons. If conclusions about treatment differences using observed data

and the sensitivity imputations are not qualitatively consistent, the reasons will be explored.

Another sensitivity analysis will use the same MMRM model after imputation of missing data by the multiple imputation (MI) method. The missing values of monthly migraine days will be imputed on the 0 to 28 scale by treatment group with 30 imputed datasets, using the fully conditional specification (FCS) method for ordinal data (via SAS® procedure MI with FCS logistic statement). Changes from Baseline will be calculated after imputation of the monthly migraine days.

Thus, the sensitivity analyses for comparison to the primary analysis using observed data are planned to include the following analyses of the primary variable.

1. MMRM after mLOCF imputation within treatments
2. Reversion toward baseline
3. MI for monthly migraine days

As a supplementary analysis of the primary endpoint, secondary analyses of covariance (ANCOVAs) of the primary efficacy variable (monthly migraine days) will be made for each month of the double-blind, placebo-controlled phase (i.e., months ending with Weeks 4, 8, 12, 16, 20, and 24), using the same model as for the primary MMRM analysis, but without repeat measures (i.e., omitting fixed effect of month and treatment group-by-month interaction).

9.4 Secondary Efficacy Endpoints and Analyses

9.4.1 Secondary Efficacy Endpoints

The secondary endpoints are listed below:

1. Change from baseline in the frequency of monthly headache days across Months 5 and 6

2. Responder status of 50% reduction from baseline in the frequency of monthly migraine days across Months 5 and 6
3. Change from baseline in the frequency of monthly acute headache medication days across Months 5 and 6
4. Change from baseline in Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQ v2.1) Role Function Restrictive (RFR) domain score at Month 6
5. Change from baseline in the Activity Impairment in Migraine Diary (AIM-D) Physical Impairment domain score across Months 5 and 6 (total 6-item Headache Impact Test [HIT-6] score for EU)

9.4.2 Main Analyses of Secondary Efficacy Endpoints

The attributes of the estimands corresponding to the secondary efficacy objectives are summarized in Table 2.

Table 2. Summary of the Estimand Attributes Corresponding to the Secondary Efficacy Objectives

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Secondary 1	Arm A: Placebo Arm B: BOTOX 155 U Arm C: BOTOX 195 U	Change from baseline in the frequency of monthly headache days across Months 5 and 6	ITT Population	IE1: Missing data due to study discontinuations, will not be imputed and will be handled as MAR with MMRM. IE2: Regardless of whether rescue medications are taken or not, collected data will be included in the analysis. IE3: Subjects who discontinue study treatment due to any reason will have their data collected after discontinuation of study treatment, and collected data will be included in the analysis.	The difference in the change from baseline in the frequency of monthly headache days across Months 5 and 6 between each BOTOX dose group and placebo

Estimand Label	Attributes of the Estimand				
	Treatment	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Secondary 2	Arm A: Placebo Arm B: BOTOX 155 U Arm C: BOTOX 195 U	Responder status of 50% reduction from baseline in the frequency of monthly migraine days across Months 5 and 6	ITT population	IE1: Missing data due to study discontinuations, will not be imputed and will be handled as MAR with MMRM. IE2: Regardless of whether rescue medications are taken or not, collected data will be included in the analysis. IE3: Subjects who discontinue study treatment due to any reason will have their data collected after discontinuation of study treatment, and collected data will be included in the analysis.	The odds ratio in participants responder status of 50% reduction from baseline in the frequency of monthly migraine days across Months 5 and 6 between each BOTOX dose group and placebo

Estimand Label	Attributes of the Estimand				
	Treatment	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Secondary 3	Arm A: Placebo Arm B: BOTOX 155 U Arm C: BOTOX 195 U	Change from baseline in the frequency of monthly acute headache medication days across Months 5 and 6	ITT Population	IE1: Missing data due to study discontinuations, will not be imputed and will be handled as MAR with MMRM. IE2: Regardless of whether rescue medications are taken or not, collected data will be included in the analysis. IE3: Subjects who discontinue study treatment due to any reason will have their data collected after discontinuation of study treatment, and collected data will be included in the analysis.	The difference in the change from baseline in the frequency of monthly acute headache medication days across Months 5 and 6 between each BOTOX dose group and placebo

Estimand Label	Attributes of the Estimand				
	Treatment	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Secondary 4	Arm A: Placebo Arm B: BOTOX 155 U Arm C: BOTOX 195 U	Change from baseline in Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQ v2.1) Role Function – Restrictive (RFR) domain score at Month 6	ITT population	IE1: Missing data due to study discontinuations, will not be imputed and will be handled as MAR with MMRM. IE2: Regardless of whether rescue medications are taken or not, collected data will be included in the analysis. IE3: Subjects who discontinue study treatment due to any reason will have their data collected after discontinuation of study treatment, and collected data will be included in the analysis.	The difference in the change from baseline in Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQ v2.1) Role Function – Restrictive (RFR) domain score at Month 6 between each BOTOX dose group and placebo

Estimand Label	Attributes of the Estimand				
	Treatment	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Secondary 5a	Arm A: Placebo Arm B: BOTOX 155 U Arm C: BOTOX 195 U	Change from baseline in the Activity Impairment in Migraine – Diary (AIM-D) Physical Impairment domain score across Months 5 and 6 in US	ITT Population	IE1: Missing data due to study discontinuations, will not be imputed and will be handled as MAR with MMRM. IE2: Regardless of whether rescue medications are taken or not, collected data will be included in the analysis. IE3: Subjects who discontinue study treatment due to any reason will have their data collected after discontinuation of study treatment, and collected data will be included in the analysis.	The difference in the change from baseline in the Activity Impairment in Migraine – Diary (AIM-D) Physical Impairment domain score across Months 5 and 6 between each BOTOX dose group and placebo

Estimand Label	Attributes of the Estimand				
	Treatment	Endpoint	Population	Handling of Intercurrent Events	Statistical Summary
Secondary 5b	Arm A: Placebo Arm B: BOTOX 155 U Arm C: BOTOX 195 U	Change from baseline in the total 6-item Headache Impact Test [HIT-6] score across Months 5 and 6 in Europe	ITT population	IE1: Missing data due to study discontinuations, will not be imputed and will be handled as MAR with MMRM. IE2: Regardless of whether rescue medications are taken or not, collected data will be included in the analysis. IE3: Subjects who discontinue study treatment due to any reason will have their data collected after discontinuation of study treatment, and collected data will be included in the analysis.	The difference in the change from baseline in the total 6-item Headache Impact Test [HIT-6] score across Months 5 and 6 between each BOTOX dose group and placebo

Each dose comparison will be analyzed across primary and secondary endpoints in the order described in Section 13.0, using a ranked-order serial gatekeeping approach to control the overall type I (false-positive) error rate. For each BOTOX dose versus placebo, each variable/dose pair will be examined for significance only if significance was determined for all variable/dose pairs ranked ahead of it.

The comparisons of interest for each of these secondary endpoints will be the pairwise comparisons of each BOTOX dose (195 U and 155 U) versus placebo, with no statistical comparison between the BOTOX doses.

The analysis results of the change from baseline in monthly headache days, monthly acute headache medication days, AIM-D Physical Impairment domain score (or HIT-6 score

[for Europe]), and MSQ v2.1 RFR domain score at Month 6 will come from the same type of MMRM analysis of observed data used for the primary efficacy endpoint, except that the baseline covariates will be monthly headache days, monthly acute headache medication days, AIM-D Physical Impairment domain score (or HIT-6 score [for Europe]), and MSQ v2.1 RFR domain score, respectively.

Logistic regression will be used to analyze responder variables using the same ANCOVA analysis model described for the primary variable (i.e., the MMRM model used for the primary analysis, but without repeat measures). The proportion of subjects with at least 50% reduction in the change from baseline in monthly migraine days across Months 5 and 6 will be summarized, along with p-values from the logistic regression. The primary assessment will be the pairwise comparison of each BOTOX dose versus placebo.

9.4.3 Sensitivity and Supplementary Analyses for Secondary Efficacy Endpoints

Sensitivity analyses are not planned for secondary variables. If they are done, they will be done in the same manner as for the primary variable. That is, the secondary endpoint of responder status of 50% reduction from baseline in the frequency of monthly migraine days across Months 5 and 6 is derived from the primary variable monthly migraine days, a sensitivity analysis will use the same logistic model, as described in Section 9.4.2, after imputation of missing data for the underlying primary variable by the mLOCF rate-change method.

The same supplementary analysis of ANCOVA, described in Section 4.3.5 and Section 9.3.4, will be performed for secondary endpoints, except that the corresponding baseline covariate will be for the variable being analyzed.

9.5 Additional Efficacy Endpoints and Analyses

The other endpoints are listed in Section 3.3. Analysis of these variables will be performed at the nominal significance level, without adjusting for multiplicity. Hence,

results will be considered descriptive and potentially supportive of the results for the primary and secondary endpoints.

The comparisons of interest for each of the other endpoints will be the pairwise comparisons of each BOTOX dose (195 U and 155 U) versus placebo, with no statistical comparison between the BOTOX doses.

In general, though, for continuous and ordinal variables (e.g., headache days, acute headache medication days, [redacted], scale scores), the change from baseline analysis will be done with the same type of MMRM or ANCOVA analyses described in Section 9.3.3 and Section 9.3.4 for the primary efficacy endpoint, except that the baseline covariate will be for the variable being analyzed. For [redacted] which is recorded as change scores, baseline for [redacted] score will be used as a surrogate baseline covariate, to adjust for variation in response that is due to baseline differences between treatment groups.

In general, for responder variables (e.g., responder status of 75% or 100% reduction from baseline in the frequency of monthly migraine days , responder status of at least 5-grade improvement in total HIT-6 score), analysis of the responder status will be done with the same methods described in Section 9.4.2 for the secondary variable of 50% responder status.

For monthly migraine days across weeks, the change from baseline analysis will be done with the same type of MMRM or ANCOVA analyses described in Section 9.3.3 and Section 9.3.4 for the primary efficacy endpoint, except that the model will use weeks instead of months. The counts for a given week will be prorated to 28-day counts before calculating the change from baseline.

The decile responder status of t% (i.e., t is from 0 to 100 by tens) reduction from baseline in the frequency of monthly migraine days across Months 5 and 6, defined as a subject with at least a t% reduction from baseline in the 2-month average of monthly migraine days over Months 5 and 6, will be assessed for each individual.

Continuous responder curve for responder status in frequency of monthly migraine days across Months 5 and 6 will be presented by treatment group across Months 5 and 6, respectively. The plots will display cumulative proportion of subjects with decile t% reduction (y-axis), with error bars for 95% CI, versus responder status of reduction from 0% to 100% (x-axis).

The Kaplan Meier curves of time to event data (e.g., first day of 50% improvement during the preceding week) will be provided. The difference between treatments in the medians for time to event will be the main interest.

Plots of least squares mean of change from baseline and their standard errors for monthly migraine days, monthly headache days and monthly acute headache medication days from the MMRM will be presented by treatment group and 28-day period and primary time point across Months 5 and 6.

Plots of visit mean and standard errors for monthly migraine days, monthly headache days and monthly acute headache medication days will be presented by treatment group and 28-day period and primary time point across Months 5 and 6.

The plot of achievement of ≥ 50% improvement (decrease) in monthly migraine days will be presented by treatment group and 28-day period and primary time point across Months 5 and 6.

9.5.1 [redacted]

The [redacted] is to be reported at Day 1, Week 4, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, Week 36, Week 40, Week 44 and Week 48 visits. The [redacted] is a single item questionnaire used to measure the subject's [redacted] in relation to migraine symptoms overall at the time of administration of the measure. The [redacted] uses a 5-point rating scale with responses ranging from [redacted] The [redacted] score ranges from 1 to 5, where lower score indicates a better condition.

The main summary will be a change from baseline score, potentially ranging from -4 to +4, best to worst. The scores will be summarized as an ordinal variable, using observed data, with the pairwise comparisons of each BOTOX dose (195 U and 155 U) versus placebo examined by the same type of ANCOVA analysis described in Section 9.3.4. The summary statistics will include response frequencies as well as distribution parameters.

9.5.2 [redacted]

The [redacted] is to be reported at Week 4, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, Week 36, Week 40, Week 44 and Week 48 visits. The [redacted] is a single item used to measure the subject's [redacted] in migraine since the first dose of study intervention. The measure uses a 7-point rating scale with responses ranging from [redacted] The [redacted] score ranges from 1 to 7, where 7 should indicate worsening and 1 should indicate improvement.

The main summary is that the [redacted] scores will be summarized as an ordinal variable, using observed data, with the pairwise comparisons of each BOTOX dose (195 U and 155 U) versus placebo examined by the same type of ANCOVA analysis described in Section 9.5. The summary statistics will include response frequencies as well as distribution parameters.

The second summary will collapse scores to a dichotomous responder status of participants with [redacted] score of "much better" or "very much better" across Months 5 and 6. The analysis of the responder status will be done with the same methods described for the secondary variable of 50% responder status (Details described in Section 9.4.2).

9.5.3 [redacted]

The [redacted] is to be reported at screening, Day 1, Week 12, Week 24, Week 36, and Week 48 visits. The [redacted] focuses on the 9 diagnostic criteria, with each question on a 0 (not at all) to 3 (nearly every day) scale. Subjects with a total score of 15 or greater at Visit 1 (or potentially confounding [redacted] per the investigator's discretion), are to be

excluded from study entry. The [redacted] can range from 0 to 27, the higher the score the more symptoms of [redacted] a subject experiences, and the more severe their [redacted] Typical thresholds in the [redacted] scoring are as follows: 0 to 4 (none); 5 to 9 (mild); 10 to 14 (moderate); 15 to 19 (moderately severe); and ≥ 20 (severe).

The response is a change from baseline score. The scores will be summarized as an ordinal variable, using observed data, with the pairwise comparisons of each BOTOX dose (195 U and 155 U) versus placebo examined by the same type of ANCOVA analysis described in Section 9.3.4. The summary statistics will include response frequencies as well as distribution parameters.

9.5.4 [redacted]

The [redacted] is to be reported at Day 1, Week 12, Week 24, Week 36, and Week 48 visits. The [redacted] is a 5-item questionnaire designed to quantify [redacted] over a 3-month period. In this study, a modified version of the [redacted] will be utilized to shorten the recall period to 1 month. The [redacted] score is the sum of [redacted] due to headaches and in the month prior to the assessment.

The baseline and change from baseline for the total [redacted] score will be summarized as a continuous variable, using observed data, with the pairwise comparisons of each BOTOX dose (195 U and 155 U) versus placebo examined with the same type of ANCOVA analysis described in Section 9.3.4.

9.5.5 [redacted]

The [redacted] is to be reported at Day 1, Week 4, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, Week 36, Week 40, Week 44 and Week 48 visits. It will be used to assess [redacted]. The measure uses a 1-week recall and contains 6 questions related to [redacted]. The [redacted] measures

both [redacted]. The measure yields 4 scores, expressed as percentage of impairment, each ranging from 0% to 100%: [redacted].

For the [redacted] scores, each with ordered response categories, comparisons of the changes from baseline between treatment groups will be done with the same type of ANCOVA analysis described in Section 9.3.4.

9.5.6 [redacted]

There are 2 variables derived from [redacted], the [redacted] descriptive system and the [redacted] The [redacted] is to be reported at Day 1, Week 12, Week 24, Week 36, and Week 48 visits.

- [redacted] The health state is a 5 digit number, with each digit being the problem level: Level 1 no problems, Level 2 slight problems, Level 3 moderate problems, Level 4 severe problems, and Level 5 extreme problems) for the 5 health state dimensions in order (digit 1 mobility, digit 2 self-care, digit 3 usual activities, digit 4 pain/discomfort, and digit 5 anxiety/ depression). For example, health state 11111 indicates no problem for any of the five dimensions, while health state 11235 indicates no problems with mobility or self-care, slight problems with performing usual activities, moderate pain or discomfort, and extreme anxiety and depression. The health state score will be calculated from the 5digit health state using the country-specific cross-walk value sets. For a country without a value set developed, another country-specific cross-walk set of values that most closely approximates that country will be used.
- [redacted] to indicate the subject's own health state, for baseline and changes from baseline.

The [redacted] will each be summarized for baseline and change from baseline. Summary statistics for each treatment for each visit window will be as described in Section 4.3.4. For changes from baseline, positive scores will indicate improvement for [redacted]. There is no plan to analyze the 5 dimensions of the [redacted] separately.

For the [redacted], treatment comparisons of baseline and the changes from baseline will be done with the same type of ANCOVA analysis described in Section 9.3.4.

9.5.7 Responder Analysis for Clinical Outcome Assessment Scores

Empirical cumulative distribution functions (eCDFs) are recommended as supportive analyses to inform the interpretation of meaningful treatment benefit for change from baseline endpoints based on clinical outcome assessments (COAs) included as primary and/or key secondary endpoints intended for approval and labeling.

To aid in the interpretation and contextualization of secondary endpoints constructed from COAs specifically, the MSQ v2.1 RFR domain score, and AIM-D PI domain score eCDFs of the change from baseline at Month 6 for the MSQ v2.1 RFR endpoint, and across Months 5 and 6 for the AIM-D PI endpoint will be plotted by treatment group. The eCDF curves for the MSQ v2.1 RFR and AIM-D PI will include all participants who meet all ITT population definition requirements and have an evaluable value at baseline and the specified outcome period. Change from baseline scores will be plotted on the horizontal axis and the cumulative percentage of subjects experiencing that specific level of change on the vertical axis.

eCDFs provide a visual display of the cumulative percentage of subjects within each group that achieve within-patient change across the entire distribution of change, thereby illustrating the percentage of responders across the range of potentially meaningful change threshold values. The following meaningful within-patient change (MWPC) responder thresholds will be utilized to develop the following responder endpoints:

1. MSQ v2.1 RFR responder at Month 6 using cutoff of [redacted]: Individual-level score change of ≥ [redacted] points can be considered clinically meaningful; (i.e., a participant whose MSQ v2.1 RFR domain score improves (increases) by ≥ [redacted] points from baseline at Month 6 can be considered a responder).[3]

2. MSQ v2.1 RFR responder at Month 6 using cutoff of [redacted]: Individual-level score change of ≥ [redacted] points can be considered clinically meaningful; (i.e., a participant whose MSQ v2.1 RFR domain score improves (increases) by ≥ [redacted] points from baseline at Month 6 can be considered a responder).[4]

3. AIM-D PI responder across Months 5 and 6 using the cutoff of [redacted]: Individual-level score change of ≤ [redacted] points can be considered clinically meaningful; (i.e., a participant whose PI domain score of the AIM-D improves (decreases) by ≥ [redacted] points from baseline across Months 5 and 6 can be considered a responder).[3]This MWPC threshold for the AIM-D PI was corroborated through psychometric evaluation and score interpretation analyses using the Phase 3 data (M21307/PRECLUDE).

For the MSQ v2.1 RFR and AIM-D PI responder endpoints defined above, a raw summary of the proportion of responders will be provided and comparisons of each BOTOX dose (195 U and 155 U) versus placebo will be done using logistic regression using the same ANCOVA analysis model described for the primary variable (i.e., the MMRM model used for the primary analysis, but without repeat measures). The analysis model will include treatment (BOTOX 195 U, BOTOX 155 U, and placebo), country and strata of [redacted] as fixed effects, with the corresponding baseline as a covariate, included as a continuous variable. The analysis will be performed based on postbaseline values at the specific timepoints. The treatment difference in terms of odds ratio between each BOTOX dose group and placebo will be estimated and tested from these models.

9.6 [redacted]

At the end of the double-blind phase of study (Week 24), in order to evaluate treatment [redacted], the participant will be asked to indicate which [redacted]. This assessment will include 2 questions: (1) The [redacted], and (2) [redacted]

The [redacted] assessment asks the participant to indicate the response using the scale below that best describes the [redacted] assigned:

[redacted]

For data summary, responses to the Question (1) above will be summarized on a 5-point scale: [redacted] The response frequencies will be summarized with descriptive statistics, but inferential statistics will not be generated. Interpretation of the results should be done with awareness that [redacted] is confounded with efficacy and safety.

The second question asks the participant to indicate the reason(s) why they chose their response to Question (1):

[redacted]

Response option selected on above question will be populated with the response option selected by the participant on Question (1).

The subject [redacted] will be calculated and presented for each treatment group.

9.7 Efficacy Subgroup Analyses

The primary effectiveness analysis (MMRM of change in monthly migraine days without the fixed effect of the corresponding subgroup variable) will be repeated for subgroups defined by each randomization stratification subgroup (monthly migraine days at baseline ≥ [redacted] [yes/no] and [redacted] [yes/no]), for medication overuse subgroup [yes/no] (see Section 8.2 in the Operations Manual), for each gender, for each country, age [< 40, ≥ 40], for each race and for prior prophylactic treatment failure subgroup [failed fewer than [redacted] preventive treatments or failed at least [redacted] preventive treatments].

10.0 Safety Analyses

10.1 General Considerations

In general, adverse events (AEs) data will be analyzed and presented for

1. The double-blind phase (i.e., AEs with start date or increase in severity on or after Day 1 through the last day preceding the Week-24 injection).
2. The open-label phase (i.e., AEs with start date or increase in severity on or after the injection day of the Week-24 treatment through Week 48 visit). Subjects who complete the double-blind, placebo-controlled phase to enter the open-label phase will be included in this analysis.
3. The entire post-injection study duration (i.e., AEs with start date or increase in severity on or after Day 1 through Week 48 visit).
4. Each treatment cycle (i.e., AEs with start date or increase in severity on or after day of injection through last day preceding subsequent injection of study treatment).

In each of cases 1-4 above, a specific AE will only count once per subject, 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 4 cases.

Safety data will be summarized for the Safety Analysis Set. Safety summaries will be presented by treatment group. For the safety analysis, participants are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized.

The safety parameters will include adverse events (AEs), clinical laboratory test results (hematology, chemistry and urinalysis), vital signs, C-SSRS, and concomitant medications and procedures. For clinical laboratory and vital signs, the last non-missing safety assessment before the first dose 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, Q1, Q3, minimum, and maximum values. Categorical variables will be summarized by number and percentage of participants.

10.2 Adverse Events

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

10.2.1 Treatment-Emergent Adverse Events

Adverse events will be coded by system organ class and preferred term using the Medical Dictionary for Regulatory Activities (MedDRA), version 23.0 or newer.

Any AEs that may occur before randomization will be listed but not summarized. Treatment-emergent adverse events (TEAEs) are defined as any AE began or worsened (increased in severity or became serious) on or after the first dose of study treatment and no later than the last dose of study treatment plus 30 days, or exit date, or death date within 30 days after exit visit, whichever occurs last. Events where the onset date is the same as the study treatment start date are assumed to be treatment-emergent. All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in descending percentage within each SOC.

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

10.2.2 Adverse Event Overview

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

- Any treatment-emergent AE
- Any treatment-emergent AE related to study treatment according to the investigator
- Any severe treatment-emergent AE
- Any serious treatment-emergent AE
- Any treatment-emergent AE related to distant spread of toxin
- Any treatment-emergent AE leading to discontinuation of study treatment
- Any treatment-emergent AE leading to death
- Any treatment-emergent AE related to COVID-19
- All deaths

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

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study treatment as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum severity and SOC and PT; and by maximum duration and SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency across treatment groups.

The number and percentage of subjects with treatment-related TEAEs in the treatment and placebo groups will similarly be tabulated by descending percentage, by SOC and PT and, separately, by SOC, PT, and severity.

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

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

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

10.2.5 Possible Distant Spread of Toxin (PDSOT)

Possible distant spread of toxin (PDSOT) is defined as a possible pharmacologic effect of botulinum toxin at sites noncontiguous and distant from the site of injection. The PDSOT will be summarized by SOC and PT according to the version of the MedDRA coding dictionary used for the study at the time of database lock.

To assess possible distant spread of toxin (PDSOT), 40 MedDRA preferred terms that may be associated with botulinum toxin effects have been identified. All PDSOT AEs will be listed by subject. The 40 terms are listed below.

MedDRA Preferred Terms Evaluated for Possible Distant Spread of Toxin	
Cardiac Disorders	**Nervous System Disorders**
Bradycardia	Bell's palsy
Eye Disorders	Bulbar palsy
Accommodation disorder	Cranial nerve palsies multiple
Diplopia	Cranial nerve paralysis
Eyelid function disorder	Dysarthria
Eyelid ptosis	Facial paralysis
Ophthalmoplegia	Facial paresis
Pupillary reflex impaired	Hyporeflexia
Vision blurred	Hypotonia
Gastrointestinal Disorders	Paralysis
Constipation	Paresis cranial nerve
Dry mouth	Peripheral nerve palsy
Dysphagia	Peripheral paralysis
Ileus paralytic	Speech disorder
Infections and Infestations	Vocal cord paralysis
Botulism	Vocal cord paresis
Musculoskeletal and Connective Tissue Disorders	**Renal and Urinary Disorders**
Muscular weakness	Urinary retention
	Respiratory, Thoracic and Mediastinal Disorders
	Aspiration
	Diaphragmatic paralysis
	Dysphonia
	Dyspnoea
	Pneumonia aspiration
	Respiratory arrest
	Respiratory depression
	Respiratory failure
	Reproductive System and Breast Disorders
	Pelvic floor muscle weakness

Note: Table is based on MedDRA 27.0. The evaluation of events mapping to these terms will take into consideration the known mechanism of action of BOTOX, the temporal relationship of the event (time to onset of the AE), the duration of the event, any re-challenge information if applicable, confounding factors that may include co-morbidities,

past medical history, concomitant medications and other non-specific constitutional symptoms of a subject.

The analyses described above for treatment versus placebo groups will be for AEs occurring during the 24-week double-blind, placebo-controlled phase. The analyses will be repeated to summarize AE results for treated subjects over the open-label phase for subjects who receive open-label BOTOX.

10.3 Analysis of Laboratory Data

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

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

Changes in laboratory parameters will be tabulated using shift tables categorized as low, normal, or high based on the normal ranges of the laboratory used for each sample. A shift table will be provided to summarize shifts from baseline to post-baseline values.

Laboratory abnormalities will be evaluated based on Potentially Clinically Important (PCI) criteria (Appendix B). For each laboratory PCI criterion, the number and percentage of subjects who have a laboratory value meeting the criteria will be summarized. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCI criteria. Hematology variables include hematocrit, hemoglobin, white blood cell count, neutrophils, lymphocytes, monocytes, platelet count (estimate not acceptable), blasts (if present).

Urinalysis variables include specific gravity, ketones, pH, protein, blood, glucose.

Chemistry variables include blood urea nitrogen, creatinine, total bilirubin, direct and in direct bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase, sodium, potassium, calcium, inorganic phosphorus, total protein, glucose, hemoglobin A1c, albumin, chloride, bicarbonate, creatinine clearance (Cockcroft-Gault calculation).

In addition, ALT/SGPT, AST/SGOT, alkaline phosphatase, and total bilirubin will be categorized as follows:

- $> 1.5 \times$ ULN to $\leq 3.0 \times$ ULN
- $> 3.0 \times$ ULN to $\leq 5.0 \times$ ULN
- $> 5.0 \times$ ULN to $\leq 10.0 \times$ ULN
- $> 10.0 \times$ ULN to $\leq 20.0 \times$ ULN
- $> 20.0 \times$ ULN

A listing of subjects with ALT or AST $> 3 \times$ ULN and total bilirubin $> 2 \times$ ULN, not necessarily concurrent, will be provided to evaluate potential Drug Induced Serious Hepatotoxicity.

A listing of urine pregnancy test results will be provided for female participants of child-bearing potential with at least one positive result.

10.4 Analysis of Vital Signs

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

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

Vital sign variables will be evaluated based on potentially clinically important (PCI) criteria (Appendix B). For each vital sign PCI criterion, the number and percentage of subjects who have a vital sign value meeting the criteria will be summarized. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCI criteria.

10.5 Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS responses will be recorded at each regular study visit: screening, Day 1, Week 4, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, Week 36, Week 40, Week 44 and Week 48 visits. At Visit 1 (screening), the C-SSRS will be completed for the subject's lifetime history and for the 6 months prior to screening. At all other visits the C-SSRS will be completed for ideation and behavior since the previous visit.

Among the questions for each assessment time period (lifetime prior to screening, 6 months prior to screening, and since previous visit) are a set of 10 questions intended to elicit a yes or no response from the subject (as deduced by the investigator site's interview of the subject). Among those questions are 5 regarding suicidal ideation, arranged in order of increasing severity (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 and 5 active suicidal ideation with specific plan and intent). Also included are 5 questions in the suicidal behavior section of the C-SSRS (0 no suicidal behavior, 1 preparatory acts or behavior, 2 aborted attempt, 3 interrupted attempt, 4 actual attempt).

C-SSRS data will be summarized separately for the following sets of responses: (1) lifetime prior to screening, (2) 6 months prior to screening, (3) baseline as reported under "since screening" at the randomization visit, (4) by visit for post-treatment visits (Week 4, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, Week 36, Week 40, Week 44 and Week 48 visits), reported as "since last visit."

In each analysis table, subjects in the safety population will be grouped by treatment for analysis and will also be combined across treatment groups. Summary statistics will include the count and percentage of subjects for each of the 10 ideation and behavior categories separately and also for any ideation, for any behavior and for any ideation or behavior.

10.6 Other Safety Analyses

10.6.1 Pregnancy Test

The serum pregnancy test is performed at screening. The urine pregnancy test will be performed at Day 1, Week 12, Week 24, and Week 36. Pregnancy test results (any pregnancies will be reported in the clinical study report) will be displayed in data listings.

10.7 Safety Subgroup Analyses

The analysis of treatment-related TEAEs (by descending order of PT incidence) will be repeated for subgroups defined by each randomization stratification group (monthly migraine days at baseline ≥ [redacted] [yes/no] and [redacted] [yes/no]), for medication overuse subgroup [yes/no] (see Section 8.2 in the Operations Manual), for each gender, for each country, for age [< 40, ≥ 40], and for each race.

11.0 COVID-19 Related Analysis

This section specified analyses related to COVID-19 pandemic from the following aspects:

- Study visits (Type of visit, Primary reason for impact to visit)
- Protocol deviation
- Treatment interruption due to COVID-19
- TEAEs related with COVID-19 and supplemental signs and symptoms
- COVID-19 status (COVID-19 testing results or contact with a COVID-19 positive person), and test type (Molecular, Antigen, Unknown, or Other)

- COVID-19 vaccine details (Vaccine name, Date, AEs and SAEs related to the vaccination, etc.)

The Safety Population will be used for the planned analyses described above. The number of participants with study visits impacted by COVID-19 will be summarized by the classification of study visits. Furthermore, the number of participants who missed at least one entire visit due to COVID-19 will be summarized.

The number of participants with a significant protocol deviation due to COVID-19 will be provided. The number of participants with study drug disruption due to COVID-19 will be provided as well. The number of participants with TEAEs related to COVID-19 will be tabulated by preferred terms, and related supplemental signs and symptoms will be listed. COVID-19 status, i.e., testing results or contact with a COVID-19 positive person, will be summarized.

The number and percentage of participants who received a COVID-19 vaccine on or after the first dose date of study treatment will be tabulated by Anatomical Therapeutic Chemical (ATC) 4 class and preferred term (PT). The number and percentage of participants with TEAEs and serious TEAEs related to COVID-19 vaccine will be summarized by PT and overall.

Information collected on subjects impacted by the ongoing COVID-19 pandemic will be presented in data listings.

12.0 Interim Analyses

There is no interim analysis planned for this study. However, to assess the validity of the assumptions for powering this Phase 3 trial, a blinded sample size re-estimation was planned to occur when approximately 33% of the subjects have reached the Week 24 visit (i.e., with diary data available for the primary endpoint of migraine days during Months 5 and 6). The observed standard deviation (SD) for monthly migraine days across Months 5 and 6 is less than specified SD (SD [redacted]) in protocol. Therefore, there is no need to increase sample size.

12.1 Data Monitoring Committee

No data monitoring committee is planned for this study.

13.0 Overall Type-I Error Control

All statistical tests will be 2-sided hypothesis tests performed at the 5% level of significance for main effects. All confidence intervals (CIs) will be 2-sided 95% CIs, unless stated otherwise.

To control the overall type I error rate at the 0.05 level, multiple comparisons are accommodated for the 2 BOTOX dose comparisons to placebo across each of the primary efficacy endpoint and the 5 secondary variables in the multiplicity control by serial gatekeeping, as illustrated in Figure 2. Specifically, the hypotheses will be tested in a stepwise manner, with endpoints analyzed in the order illustrated in the figure. The testing process will be terminated for subsequent hypotheses whenever a statistical test indicates that a higher ranked hypothesis is not significant at the 0.05 level (i.e., all subsequent tests of any of the dose/variable combinations ranked in the remaining steps will be considered not significant). Results from the analyses that are not included in this gatekeeping algorithm will be considered supportive and exploratory.

Figure 2. Schematic of the Multiplicity Method to Control Overall Type I Error Rate

Key
- Primary - Change from baseline in the frequency of monthly migraine days across Months 5 and 6
- S1 - Change from baseline in the frequency of monthly headache days across Months 5 and 6
- S2 - Responder status of 50% reduction from baseline in the frequency of monthly migraine days across Months 5 and 6
- S3 - Change from baseline in the frequency of monthly acute headache medication days across Months 5 and 6
- S4 - Change from baseline in MSQ v2.1 RFR domain score at Month 6
- S5 - Change from baseline in AIM-D physical impairment score across Months 5 and 6 *[Europe: Change from baseline in HIT-6 score across Months 5 and 6]*

14.0 Version History

Table 3. SAP Version History Summary

Version	Date	Section	Description
1.0	14 September 2022		Initial version
2.0	23 December 2023	1.0	Modified 4-week screening phase to 4-week screening/baseline phase; Removed "under both L-AGN Linux and L-ABV Unix operating systems."
		2.1	Updated the wording of the objective to align with protocol.
		2.3	Added wording for randomization within country; Updated wording of enrollment targets to allow flexibility.
		2.4	Added "Due to enrollment challenges a maximum number of 24 subjects (around 3% of 777 subjects) randomized per site was allowed."

Version	Date	Section	Description
		3.3	Added the statement of "The primary and secondary efficacy measures will also be evaluated at monthly timepoints"; Modified analyses time points from "monthly time points" to "Weeks 12, 24, 36, and 48" for [redacted] descriptive system index score and [redacted] score, [redacted] Added an additional efficacy endpoint for [redacted] Removed the efficacy endpoint analysis of monthly acute headache medication days by medication class due to low counts of subjects taking individual medication classes at baseline, the results are not informative.
		3.4	Modified AESIs to PDSOT; Removed "and serum" and add "urinalysis" for clinical laboratory testing; Modified safety analysis baseline cutoff date from "randomization" to "the date of first study treatment administration."
		4.2	Updated the data dealing rule for laboratory variables with multiple values within same analysis window.
		4.3.1	Modified safety analysis baseline cutoff date from "randomization" to "the date of first study treatment administration."
		4.3.5	Removed analyses using Wilcoxon rank-sum test; Removed comparisons between treatment groups at baseline.
		4.3.6	Modified the analysis methods for Nominal Variables.
		4.3.7	Added "the average across Months 5 and 6 will be rounded to tenth."
		5.0	Modified "reason" to "primary reason"; Added a paragraph for summary of subjects who prematurely discontinued from study; Added comparing inconsistency of randomized strata from IRT versus actual strata derived from eDiary, eCRF.
		6.0	Removed the analysis for the percentage of injection received and keep the summary of subjects treated in each treatment cycle.

Version	Date	Section	Description
		7.1	Updated race categories, add summary for Ethnicity, Region 1 (US, Ex-US) and Region 2 (North America, Europe, Other).
		7.1.1	Removed comparisons between treatment groups at baseline for [redacted]
		7.1.2	Removed comparisons between treatment groups at baseline for [redacted].
		7.2.1	Modified PT sorting order from "alphabetical" to "decreasing percentage"; Removed comparisons for incidences of medical conditions.
		7.2.2	Added "Prior and concomitant medications will be summarized separately."
		7.2.2.1	Removed comparisons for prior medications, prior procedures.
		7.2.2.2	Added a section title for study medication, updated section number from 7.2.2.2 to 7.2.2.3 for concomitant medications and concurrent procedures; Removed comparisons for study medication.
		7.2.2.3	Added "injection" after "Day 1"; Removed comparisons for concomitant medications, concurrent procedures.
		9.1	Removed comparison between combined BOTOX doses versus Placebo; Modified safety analysis baseline cutoff date from "randomization" to "the date of first study treatment administration." Removed laboratory evaluations, vital signs. Updated the wording of analysis related to actual stratum.
		9.2 9.3.2	Original "15 days" criteria for using prorated approach were modified to "14 days."
		9.3.2.1 9.3.2.3	Modified sensitivity analysis using [redacted] mins headache duration instead of [redacted] hours.
		9.3.2.5 9.3.2.7	Original "15 days" criteria for using prorated approach were modified to "14 days."
		9.3.3	Updated intercurrent events definitions and handling methods for primary endpoints; Removed comparison between combined BOTOX doses versus Placebo.

Version	Date	Section	Description
		9.4.2	Updated intercurrent events definitions and handling methods for secondary endpoints.
		9.5	Updated baseline for the [redacted] score will be used as a surrogate baseline covariate for [redacted] which is not collected at baseline.
		9.5.1	Modified Wilcoxon rank-sum test to ANCOVA for [redacted].
		9.5.2	Removed redundant statements for [redacted] Updated the wording for [redacted] analyses; Modified Wilcoxon rank-sum test to ANCOVA for [redacted]
		9.5.3	Updated "> 20" to "≥ 20"; Modified Wilcoxon rank-sum test to ANCOVA for [redacted].
		9.5.4	Modified Wilcoxon rank-sum test to ANCOVA for [redacted]
		9.5.5	Modified Wilcoxon rank-sum test to ANCOVA for [redacted]
		9.5.6	Modified Wilcoxon rank-sum test to ANCOVA for [redacted]
		9.7	Added efficacy subgroup analyses for prior prophylactic treatment failure subgroup, age, race.
		10.1	Removed "and serum" and add "urinalysis" for clinical laboratory testing. Modified safety analysis baseline cutoff date from "randomization" to "the date of first study treatment administration."
		10.2.1	Updated the upper bound for AEs to be considered as TEAEs. Modified the sorting order of AEs by PTs from "alphabetical order" to "descending percentage."
		10.2.3	Removed AEs by subject number and SOC and PT, added AEs by maximum duration and SOC and PT; Modified the sorting order of AEs from "for total active group" to "across treatment groups." Added analyses for treatment-related TEAEs.
		10.2.5	Updated the section AESIs to the section of PDSOT; PDSOT preferred terms were updated according to MedDRA 26.0.

Version	Date	Section	Description
		10.3	Updated analyses regarding Lab shift tables.
		10.4	Updated vital sign related analyses.
		10.5	Removed C-SSRS analyses by treatment period; Removed C-SSRS comparisons between treatment groups.
		10.7	Added safety subgroup analyses for age and race.
		12.0	Added planned blinded sample size re-estimation result.
		Appendix A	Updated list of SAP signatories.
		Appendix B	Added PCI criteria for urinalysis values.
3.0	11 April 2024	3.2	Moved the last secondary endpoint of [redacted] to Section 3.3 Additional Efficacy Endpoints to align with protocol amendment.
		3.3	Added three endpoints for responder analyses of MSQ v2.1 RFR and AIM-D PI.
		5.0	Added summary of subject accountability by country; Updated enrollment failure analysis.
		9.3.3	Added details regarding primary analysis model regarding covariance structure and method to estimate denominator degrees of freedom; Added the statement "The ITT population will be used for primary and sensitivity analyses. A few subjects (less than 1%) who did not have post-baseline scores will not be included in the primary MMRM due to having missing value for the dependent variable. These subjects will be included in sensitivity analyses and have post-baseline scores imputed as described in Section 9.2."
		9.3.4	Removed reversion to baseline for the imputation of data after intercurrent events, instead use reversion toward baseline for the imputation of data after specified intercurrents events.
		9.4.1	Removed the last secondary endpoint of [redacted] to align with protocol amendment.

Version	Date	Section	Description
		9.4.2	Removed estimand table for the last secondary endpoint of [redacted] and associated analysis from this section and was added to Section 9.5.
		9.4.3	Added the statement of supplementary analysis for secondary endpoints.
		9.5.7	Added the section of responder analysis for MSQ v2.1 RFR and AIM-D PI scores.
		10.2.1	Updated the upper bound for TEAEs to include TEAEs leading to death in study periods.
		13.0	Updated 6 secondary variables to 5 secondary variables; Updated Figure 2 by removing the last ranked secondary endpoint of [redacted]
		Appendix A	Updated list of SAP signatories.
4.0	03 June 2024		Updated "patient" to "subject" as necessary in the document.
		6.0	Added the summary of treatment exposure duration by 12-week intervals and the summary of subjects treated in at least 1, 2, 3, and 4 injection cycles
		7.3	Updated protocol deviation categories.
		9.2	Added the statement "A few subjects (less than 1%) who did not have post-baseline scores will not be included in the primary MMRM due to having missing value for the dependent variable" to be consistent with Section 9.3.3.
		9.3.3	Added details regarding primary analysis model, i.e., REML estimation and contrasts.
		9.4.2	Added "at least" prior to 50% reduction in the change from baseline in monthly migraine days.
		9.4.3	Added a sensitivity analysis for the secondary endpoint of responder status of 50% reduction from baseline in the frequency of monthly migraine days.
		9.5	Added the statement for planned figures.
		9.5.7	Added the statement regarding AIM-D PI MWPC threshold derived using the Phase 3 data (M21307/PRECLUDE).
		10.2.5	Updated MedDRA version for PDSOT to v27.0.

Version	Date	Section	Description
		Appendix A	Updated list of SAP signatories.

14.1 Changes to Planned Analyses in the Protocol

Three new endpoints were added to SAP Section 3.3 listed as below and corresponding analyses were added to Section 9.5.7.

- Responder status of at least [redacted]-point improvement (increase) in MSQ v2.1 RFR score at Month 6
- Responder status of at least [redacted]-point improvement (increase) in MSQ v2.1 RFR score at Month 6
- Responder status of at least [redacted]-point improvement (decrease) in AIM-D PI score across Months 5 and 6

The efficacy subgroup analyses for age [< 40, ≥ 40], for each race, and for prophylactic treatment failure subgroup [failed fewer than [redacted] preventive treatments or failed at least [redacted] preventive treatments] for the primary efficacy endpoint were added in SAP Section 9.7.

The safety subgroup analyses for age [< 40, ≥ 40] and for each race for treatment-related TEAEs were added in SAP Section 10.7.

15.0 References

1. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. Cephalalgia. 2018;38(1):1-211.
2. Katsarava Z, Schneeweiss S, Kurth T, et al. Incidence and predictors for chronicity of headache in patients with episodic migraine. Neurology. 2004;62(5):788-90.
3. Center for Drug Evaluation and Research. Qulipta (Atogepant) for prevention of episodic migraine in adults. 2021; Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2021/215206Orig1s000IntegratedR.pdf. Accessed January 30, 2024.

4. Center for Drug Evaluation and Research. Emgality (galcanezumab) for prevention of migraines. 2018; Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2018/761063Orig1s000MedR.pdf. Accessed January 30, 2024.

Appendix A. List of SAP Signatories

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

Appendix B. Potentially Clinically Important Criteria for Safety Endpoints

The criteria for Potentially Clinically Important (PCI) laboratory findings are described in Table B-1, Table B-2 and Table B-3 and the PCI criteria for vital sign findings are described in Table B-4.

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

		Criteria
Hematology Variables	**Flag**	**Observed Value**
Hemoglobin (g/L)	High	Increase in > 2 g/dL above ULN or > baseline if baseline > ULN
	Low	< 100
WBC Count (10^9/L)	High	-
	Low	< 3
Neutrophils (10^9/L)	High	-
	Low	< 1.5
Lymphocytes (10^9/L)	High	> 4
	Low	< 0.8
Platelets (10^9/L)	High	-
	Low	< 75.0

ULN Upper Limit of Normal

Note: A post baseline value must be more extreme than the baseline value to be considered a potentially clinically important finding.

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

		Criteria[a]
Chemistry Variables	**Flag**	**Observed Value**
Albumin (g/L)	High	-
	Low	< 30
Alkaline Phosphatase (U/L)	High	> 2.5 × ULN
	Low	-
Alanine Aminotransferase (U/L)	High	> 3.0 × ULN
	Low	-
Aspartate Aminotransferase (U/L)	High	> 3.0 × ULN
	Low	-
Total Bilirubin (umol/L)	High	> 1.5 × ULN
	Low	-
Calcium (mmol/L)	High	Corrected serum calcium > 2.9@; Ionized calcium > 1.5
	Low	Corrected serum calcium < 2.0@; Ionized calcium < 1.0
Creatinine (umol/L)	High	> 1.5 × ULN
	Low	-
Glucose (mmol/L)	High	> 13.9
	Low	< 2.2
Potassium (mmol/L)	High	> 5.5
	Low	< LLN
Sodium (mmol/L)	High	> 150
	Low	< 130

ULN Upper Limit of Normal; LLN Lower Limit of Normal

@ Calcium corrected for Albumin (can be used if it is confirmed with the lab that the result has been corrected for Albumin and not a different protein).

a. A post baseline value must be more extreme than the baseline value to be considered a potentially clinically important finding.

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

		Criteria
Hematology Variables	**Flag**	**Observed Value**
pH	High	> 1.1 × ULN
	Low	< 0.9 × LLN
Glucose	High	At least 1 +
	Low	-
Protein	High	At least 1 +
	Low	-
Specific gravity	High	> 1.1 × ULN
	Low	-

ULN Upper Limit of Normal; LLN lower limit of normal value

Note: A post baseline value must be more extreme than the baseline value to be considered a potentially clinically important finding.

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

		Criteria	
Vital Signs Variables	**Criterion**	**Observed Value**	**Change From Baseline**
Sitting Systolic Blood Pressure (mmHg)	High	≥ 160	Increase of ≥ 20
	Low	≤ 90	Decrease of ≥ 20
Sitting Diastolic Blood Pressure (mmHg)	High	≥ 100	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Sitting Pulse rate (bpm)	High	≥ 110	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15

bpm beats per minute

Note: A post baseline value is considered potentially clinically important if it meets both the observed value and the change from baseline criteria.

Document Approval

Study M21307 - Statistical Analysis Plan Version 4 - 03Jun2024 (E3 16.1.9)

Version: 1.0 **Date:** 05-Jun-2024 **Company ID:** 20240605-0900f9f687b7e2a9-1.0-en

Signed by:	Date:	Meaning of Signature:
[redacted]	05-Jun-2024 00:15 UTC	Approver - Statistics
	04-Jun-2024 19:40 UTC	Approver - Statistics
	04-Jun-2024 17:36 UTC	Approver
	04-Jun-2024 17:24 UTC	Approver
	03-Jun-2024 23:25 UTC	Author