NCT02010775

Study ID: 191622-130

Title: BOTOX® TREATMENT OF MASSETER MUSCLE HYPERTROPHY

Statistical Analysis Plan Amendment 2 Date: 12-Jun-2018

Allergan

Biostatistics

Analysis Plan

Study ID: 191622-130

Study Title: BOTOX® TREATMENT OF MASSETER MUSCLE HYPERTROPHY

Table of Contents

List of Tables

1. Introduction

This analysis plan provides an expanded and detailed description of the statistical methods presented in the Study 191622-130 protocol.

This analysis plan is based on Protocol Amendment 1, approved on 28 February 2014.

There are 4 cohorts in total to be enrolled into the study. For each cohort, there will be a database snapshot with key data through the day 90 visit, for the purpose of the Data Review Committee (DRC). There will be a final database lock with all study data after all subjects have completed the study (final cohort day 360). Subjects' treatment assignments will remain blinded for all site personnel, study subjects, and Allergan personnel directly involved with ongoing operational activities of the clinical study, until completion of the final database lock.

1.1 Primary Study Objectives and Design

The objectives of this study are to evaluate the safety and efficacy of a range of doses of BOTOX for the treatment of subjects with bilateral masseter muscle hypertrophy (MMH).

This study is a multicenter, randomized, double-blind, placebo-controlled, up to 2-treatment, dose escalation study. BOTOX or placebo will be administered intramuscularly to the bilateral masseter muscles (6 total injections, 3 injections/masseter). The initial BOTOX total dose will be 24 U (12 U/masseter) and the maximum BOTOX total dose will be 96 U (48 U/masseter). Stratification will be by baseline (day 1) Masseter Muscle Prominence Scale (MMPS) grade at baseline (day 1). The planned dose escalation is as follows:

Cohort	Total Dose	Sample Size	Total Subjects/Cohort
1	24 U	N=40	50
	placebo	N=10	
2	48 U	N=40	50
	placebo	N=10	
3	72 U	N=40	50
	placebo	N=10	
4	96 U	N=40	50
	placebo	N=10	

The second treatment is only allowed at the day 180 visit. Subjects will receive the same treatment as they received at the day 1 visit.

Subjects who meet the following retreatment criteria at the day 180 visit will receive a second treatment:

- Allergan written notification allowing retreatment of the current cohort

- subject has marked (grade 4) or very marked (grade 5) bilateral masseter muscle hypertrophy, as assessed by the investigator using the MMPS
- females of childbearing potential must have a negative urine pregnancy test prior to treatment

The primary efficacy variable, change from baseline in lower facial volume (cm^3, note that $1cm^3$=1mL), is calculated from VECTRA M3 digital photography system 3D image models [redacted]

The primary safety analysis is the calculation of incidence of treatment emergent adverse events (TEAE) for the entire study period, as well as for each treatment cycle.

1.2 Secondary and Other Objectives

The secondary efficacy variable is the investigator's assessment of MMH using the MMPS (1 = minimal, 2 = mild, 3 = moderate, 4 = marked, 5 = very marked).

[redacted]

2. Analysis Populations and Data Conventions

2.1 Analysis Populations

Two different analysis populations will be used in the analysis of this study.

All efficacy analyses will be performed on the modified intent-to-treat (mITT) population, consisting of all randomized subjects who received study treatment and had at least 1 follow-up visit. For the mITT population, analyses will be performed using subjects' planned treatment.

The safety analyses will be based on the safety population, which will consist of all subjects who receive at least 1 injection of study treatment. All safety analyses will be performed with subjects analyzed by their actual treatment received.

Due to missing source documents, subject [redacted] will be excluded from all analyses but still included in the listings.

[redacted]

2.3 Data Conventions

Certain data conventions and definitions are stated in this section and apply to all analyses.

2.3.1 Treatment Group and Treatment Cycles

A treatment cycle is considered to start on the day of treatment with study medication and end on the day prior to the next study treatment or on the study exit day if there are no subsequent treatments.

Between group comparisons will only be performed for each BOTOX group against the placebo group, and for each BOTOX group against the next higher dose BOTOX group.

2.3.2 Baseline, Changes from Baseline and Direction of Differences

The latest measurement taken on or before day 1 prior to the first injection will be the baseline measurement. In addition, within each treatment cycle, "cycle baseline" is defined as the most recent evaluation prior to the current cycle's treatment. Changes from baseline will be based on the day 1 baseline (rather than cycle baseline) unless otherwise stated.

Change from baseline will be calculated as followup minus baseline. The interpretation of a positive or negative change from baseline is specific to the outcome measure and will be discussed as needed. Treatment differences will be reported as BOTOX minus placebo and as BOTOX (higher total dose) minus BOTOX (lower total dose).

2.3.3 Time Variables

Study days will be calculated based on day 1 as follows:

- Study day = visit date minus date of day 1 plus 1.

Treatment cycle days will be calculated as follows:

- Treatment cycle day = visit date minus date of treatment cycle injection plus 1.

In general, there will be no substitution for 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 pre-study 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.

If, for some reason, a partial visit date exists, it will be imputed on a case-by-case basis as necessary.

Partial adverse event onset date will be imputed as follows: 1) if day is missing but month is not, impute the date as the first day of the month; 2) if both day and month are missing, impute the date as 01 Jan; 3) if imputed partial onset date is before the first treatment for adverse event with onset after treatment, impute the date as the first injection date; 4) if imputed partial onset date is before period 2 for adverse event with onset during period 2, impute the date as the day 180 visit date. Imputed partial adverse event onset date will only be used to determine the adverse event onset cycle.

Partial adverse event stop date will be imputed as follows: 1) if day is missing but month is not, impute the date as the last day of the month; 2) if both day and month are missing, impute the date as 31 Dec.

The date of first increase in severity to moderate, if a partial date, will be imputed as follows: 1) if day is missing but month is not, impute the date as the 15th of the month; 2) if both day and month are missing, impute the date as 15 June; 3) if imputed partial date is before the first treatment while the maximum severity during pretreatment period is mild, impute the date as the first injection date; 4) if imputed partial date is during study period 1 (on or after first treatment date but before day 180 visit date) while the maximum severity during study period 1 is mild, impute the date as the day 180 visit date.

The date of first increase in severity to severe, if a partial date, will be imputed as follows: 1) if day is missing but month is not, impute the date as the 15th of the month; 2) if both day and month are missing, impute the date as 15 June; 3) if imputed partial date is before the first treatment while the maximum severity during pretreatment period is mild or moderate, impute the date as the first injection date; 4) if imputed partial date is during study period 1 (on or after first treatment date but before day 180 visit date) while the maximum severity during study period 1 is mild or moderate, impute the date as the day 180 visit date.

Imputed partial adverse event dates will only be used to determine TEAE. All partial dates will be listed "as is" in the data listings.

2.3.4 Stratification

Subject randomization is stratified by baseline MMPS (4 or 5) at baseline (day 1). In case of mis-stratification, the subjects' actual baseline MMPS will be used for analyses.

2.3.5 Handling Drop-out or Missing Data

For efficacy analyses of mITT population of the change from baseline in lower facial volume using VECTRA 3D images, MMPS, lower facial width, mandibular facial angle, CT masseter muscle volume, LFSC, and LFSQ, last observation carried forward (LOCF) will be used to impute missing values for each scheduled visit. Only missing data for visits up to day 180 of each treatment cycle will be imputed. Missing data for visits after day 180 of each cycle will not be imputed. If a subject exited the study before day 180 of a treatment cycle, visits after the study exit will not be imputed.

Missing baseline values for lower facial width, mandibular facial angle, CT masseter muscle volume, LFSC, and LFSQ will be imputed using the average of subjects with non-missing values. Missing values for change from baseline in lower facial volume using VECTRA 3D images at day 30 will be imputed using the average of subjects with non-missing values.

For angle measurements, the average degree from both sides of the face will be determined first; if the degree of angle is missing for 1 side of the face, then the angle for the subject will be based on the angle of the other side of the face. If a subject's average is a half-degree value, it will be rounded up to an integer (eg, 30.5 degree will be rounded up to 31 degrees).

For rating an individual's MMPS score, the average of the scores from both sides of the face will be determined first; if a score is missing for 1 side of the face, then the score for the subject will be based on the score of the other side of the face. If a subject's average is a half-grade value, it will be rounded up to an integer to remain consistent with the original scale (eg, 1.50 will be rounded up to 2).

When implemented, imputation of missing values will be performed only for subjects who receive a treatment in the pertinent treatment cycle. For example, only subjects receiving the second treatment will have values imputed in treatment cycle 2; subjects who discontinued from the study during the first treatment will not have any data imputed in treatment cycle 2.

2.3.6 Descriptive Statistics

Data will be summarized with descriptive statistics and/or response frequencies. Descriptive statistics for continuous variables include the sample size (N), mean (Mean), standard deviation (SD), median (Median), minimum (Min), and maximum (Max). Descriptive statistics for categorical variables include the sample size (N) and frequency (counts and percentages).

2.3.7 Significance Level

The level of significance used for all statistical tests will be 0.05, 2-sided, unless stated otherwise.

2.3.8 Confidence Intervals

For dichotomous variables, 2-sided 95% confidence intervals for proportions will be calculated based on the normal distribution approximation, unless otherwise stated.

2.3.9 Dictionaries

The Medical Dictionary for Regulatory Activities (MedDRA) will be used to code adverse events, medical history, and medications. MedDRA primary system organ class (SOC), MedDRA high level term (HLT), together with the drug record number and base preferred drug name from the enhanced version of World Health Organization Drug Dictionary (WHO DDE) will be used to classify all medications (ie, prior or concomitant) recorded in the study. Base preferred name is the WHO DDE drug preferred name assigned to the 11-digit code of drug record number+01+001.

2.3.10 Units

Metric system units will be used; degrees Celsius for temperature, kilograms (kg) for body weight, and centimeters (cm) for height.

3. Disposition and Exit Status

3.1 Screening Log Data

The subject counts and percentage for the reasons for screening failure will be displayed for the following categories: inclusion criteria, exclusion criteria, and total. In addition, subjects who fail to pass screening will be compared to randomized subjects with descriptive statistics for age, sex, and race. No statistical tests are planned to compare groups.

3.2 Disposition and Exit Status

The analysis will be performed using the mITT, and Safety population for the entire study and by treatment cycle.

The disposition of study subjects will consist of a summary of the number of subjects enrolled and available in the mITT, and Safety analysis populations by treatment regimen and overall.

The summary will be further enhanced by tabulating the number of subjects remaining in the respective treatment groups of each treatment cycle.

Tabulation of the number and percentage of subjects in each exit status category (ie, adverse event, pregnancy, lost to follow-up, personal reasons, protocol violations and other) will be provided for each treatment regimen of each treatment cycle. Discontinued subjects will be listed along with the corresponding reason(s) for early withdrawal from the study.

4. Demographics and Other Baseline Characteristics

The analysis and summary of all demographic and other baseline characteristics will be performed using the mITT population unless otherwise specified.

4.1 Demographics

Demographic data will be collected at the screening visit, including age, gender, race, weight, height, and body mass index (BMI). Age, weight, height and BMI will be analyzed with descriptive statistics (sample size, mean, SD, median, Min, and Max). Age group (18-35 years versus 36-50 years), gender, and race (Asian versus non-Asian) will be analyzed with frequency (counts and percentages).

Similar analyses will be done for the Safety population.

4.2 Disease Characteristics

For the mITT population, the distribution of the baseline MMPS will be summarized.

The distribution of the baseline lower facial width, mandibular facial angle, CT masseter muscle volume (total and defined region), investigator and subject LFSC will be summarized by treatment regimen for the mITT population.

4.3 Prior and Concomitant Medications

Prior medications include all medications taken prior to day 1 (randomization/treatment visit), whether or not the medication is continuing beyond day 1. Concomitant medications encompass all medicinal products that the subject was taking prior to the day 1 visit that are ongoing at the visit, in addition to all medications that have a start date on or after the day 1 visit date.

Prior and concomitant medication will be tabulated separately. The frequency (number and percentages) of subjects who have taken each medication will be tabulated for each treatment

regimen by MedDRA primary SOC, MedDRA HLT, and WHO DDE base preferred term. MedDRA SOC, HLT within each SOC, and WHO DDE base preferred term within each HLT will appear alphabetically.

If start or stop dates for medications are only partially reported but can be classified as prior to day 1, then the medications will be included in the summary of prior medications. If start or stop dates for medications are only partially reported and cannot be definitively classified as having stopped prior to day 1, then the medications will be included in the summary of concomitant medications.

4.4 Concurrent Procedures

All procedures undergone on or after study day 1 visit through the exit visit will be considered concurrent procedures. The frequency (number and percentage) of subjects who have undergone each procedure will be tabulated for treatment regimen by MedDRA HLT and preferred term. MedDRA HLT will appear alphabetically, and preferred term within each HLT will be sorted by descending incidence.

4.5 Medical History

Medical history includes all medical conditions that the subject has had in the past or currently has prior to injection on the day 1 visit. The frequency (numbers and percentage) of subjects reporting each medical history will be tabulated for the mITT population by MedDRA primary SOC in descending order.

4.6 Dental History and Habits

Dental history will be coded using the MedDRA dictionary. Dental history data will be summarized by preferred term using frequency distributions.

Descriptive statistics (sample size, count, percentage) will be presented for each dental habit.

Dental history and habits will be presented in the listing.

4.7 Pretreatment Adverse Events

Pretreatment adverse events (PTAEs) are defined as any adverse events with onset date prior to the first dosing of study medication. Pretreatment adverse events are discussed in Section 6.2.

5. Efficacy Analyses

Pairwise comparisons will be done between each BOTOX group and the placebo group. A pairwise comparison will also be done between a BOTOX group and the next higher dose BOTOX group.

All subjects who received placebo will be pooled together as placebo group for analyses.

5.1 Collection and Derivation of Primary Efficacy Assessments

The primary efficacy variable, change from baseline in lower facial volume (cm^3), is calculated from VECTRA M3 digital photography system 3D image models [redacted]

Visit windows for all time points are specified in Section 2.2. The primary time point is day 90 of the first treatment cycle.

Between group comparisons will only be performed for each BOTOX group against the placebo group, and for each BOTOX group against the next higher dose BOTOX group.

For analysis of the mITT population, the last observation carried forward (LOCF) method will be used to impute missing values.

5.2 Primary Efficacy Analyses

Descriptive statistics for change from baseline in lower facial volume for each study visit (sample size, mean, standard deviation, median, minimum, and maximum) will be provided by treatment regimen for all scheduled visits.

The primary efficacy analysis will be based on the mITT population. Day 90 is the primary timepoint.

The following set of hypotheses will be used to compare the BOTOX groups with placebo:

- Null hypothesis: BOTOX and placebo are equally effective in changing lower facial volume at day 90.
- Alternative hypothesis: BOTOX and placebo are not equally effective in changing lower facial volume at day 90.

The change from baseline in lower facial volume at each visit will be analyzed using an ANOVA model with treatment, and baseline MMPS grade as factors. In addition, 2-sided 95% confidence intervals for the treatment differences will be provided.

Analysis will also be performed based on observed data (ie, strictly data that falls within the visit windows) for the mITT population.

Among-group comparisons for the placebo groups will be analyzed using an ANOVA model with treatment as factor to test whether they are combinable. If the difference among the placebo groups are significant, sensitivity analysis will also be done for the primary variable using an ANOVA model with treatment, cohort number (as numeric variable) and baseline MMPS grade as factors.

5.3 Secondary Efficacy Analyses

The secondary efficacy variable is the investigator's assessment of MMH using the MMPS (1 = minimal, 2 = mild, 3 = moderate, 4 = marked, 5 = very marked).

The proportion of responders will be analyzed with a responder defined as subjects who achieve an MMPS grade of ≤ 3. The proportion of responders will be analyzed using Cochran-Mantel-Haenszel (CMH) tests stratified by baseline MMPS grade. In addition, 2-sided 95% confidence intervals for the treatment differences in response rates will be provided. Between group comparisons will only be performed for each BOTOX group against the placebo group, and for each BOTOX group against the next higher dose BOTOX group.

[redacted]

[redacted]

[redacted]

6.1 Exposure to Study Treatment

Subjects' exposure to BOTOX will be summarized by number of treatments received.

6.2 Adverse Events

Adverse events will be coded from the verbatim text into preferred term (PT) and the primary system organ class (SOC) by using the MedDRA dictionary.

Adverse events are collected both for the screening/baseline period pretreatment (which are referred to as pretreatment AE) and for the followup period after treatment is initiated (which are referred to as postbaseline AE).

A treatment emergent adverse event (TEAE)is an adverse event with onset after the initiation of study treatment, or an adverse event with onset prior to study treatment that worsened in severity, or became serious after the initiation of study treatment.

A TEAE for a specific treatment cycle is an adverse event with onset after the initiation of study treatment for the specific cycle, or an adverse event with onset prior to the cycle that worsened in severity during the cycle, or became serious after the initiation of study treatment for the specific cycle.

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

The MedDRA nomenclature will be used to code all adverse events. Three summary tables as shown below will be presented for TEAEs, and for a subset of TEAEs that were considered by investigator to be treatment related (ie, treatment related TEAEs).

The 3 summary tables to be generated are:

1. by preferred term in descending order of incidence rate (summarized by treatment group)

2. by preferred term within primary SOC in alphabetical order (the incidence rates of each preferred term within a primary SOC as well as overall primary SOC incidence rates will be summarized by treatment group. Primary SOCs will appear alphabetically and preferred terms within each primary SOC will be sorted by descending incidence of adverse events. If a subject has multiple reports coded to the same preferred term within a primary SOC, that subject will be counted only once for that preferred term within that primary SOC).

3. by maximum severity of preferred term within primary SOC in alphabetical order (the incidence rates of the maximum severity of each preferred term within a primary SOC will be summarized by treatment group. Primary SOCs will appear alphabetically and preferred terms within a primary SOC will be sorted in descending incidence of adverse events. If a subject has multiple reports coding to the same preferred term within a primary SOC, the maximum severity for that subject will be used and that subject will be counted only once for that preferred term within that primary SOC.)

TEAEs and treatment-related TEAEs will be summarized for each treatment regimen by treatment cycle and over the entire study.

All adverse events leading to study discontinuation, treatment-related adverse events leading to study discontinuation, and serious adverse events leading to study discontinuation will be summarized by preferred term within primary SOC. A listing of all adverse events leading to study discontinuation will be provided.

A subject listing will be generated for all adverse events, including subject age, sex, and race, primary SOC, preferred term, onset and stop date, onset day relative to the most recent dose, relationship, and severity.

A listing of all pretreatment AEs (PTAEs) will be provided.

6.3 Serious Adverse Events

A serious adverse event is defined as any adverse event occurring at any dose that results in any of the following outcomes: death, a life-threatening adverse event, inpatient hospitalization or prolongation of existing hospitalization, a persistent or significant disability/incapacity, or a congenital anomaly/birth defect. Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered a serious adverse event when, based upon appropriate medical judgment, they may jeopardize the patient or subject and may require medical or surgical intervention to prevent 1 of the outcomes listed in this definition. All cancer adverse events are considered serious adverse events by Allergan. Allergan also considers all abortions, spontaneous, or elective as serious adverse events.

Serious TEAEs will be summarized similar to TEAEs for entire study, and by treatment cycle. A subject listing will be generated with serious adverse events including subjects' age, gender, race, onset day, duration (number of days the adverse event lasted), relationship to study treatment, and serious criteria (congenital anomaly, death, hospitalization, life-threatening, significant disability, and other).

6.4 Adverse Events Possibly Associated with Effects Remote to Site of Injection (Possible Spread of Toxin, PSOTs) and Local Effects

Possible distant spread of toxin is defined as a possible pharmacologic effect of botulinum toxin at sites noncontiguous and distant from the site of injection. Possible spread of toxin (PSOT) is defined as those events classified as possible distant spread of toxin and locally-occurring adverse events that could be due to a possible pharmacologic effect of botulinum toxin. In order to retrieve cases potentially meeting this definition, the steps outlined below are employed. This analytic methodology is approved by the Allergan Reference Member State that reviews the PSURs, is the current standard utilized in PSURs, and was agreed with the US FDA.

The index search MedDRA preferred terms identified for the possible distant spread of toxin analysis were originally based upon individual reported symptoms of the naturally occurring condition of clinical botulism (Merck Manual, Beers, 2006). It was acknowledged that possible local effects may also occur anatomically proximal to the injections site, such as ipsilateral ptosis

following injection for strabismus, or dysphagia following injection for cervical dystonia. The preferred terms used in this analysis were purposely chosen to be broad, knowing that cases would subsequently be further characterized. These 39 preferred terms are listed in Table 2 below.

Table 2 MedDRA Version 20.1 Preferred Terms Evaluated for Possible Spread of Toxin

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

Step 1. Adverse event listings will be searched for cases (regardless of serious, listedness, or medical confirmation) mapping to the 39 MedDRA preferred terms from Table 2.

Step 2. In order to focus on the most informative cases, reports meeting the criteria in Step 1 together with abnormal neurologic findings recorded on a NEF will then be medically reviewed and placed into 1 of the following categories. Where cases containing elements of more than 1 category, medical judgment will be applied to determine the most significant factor in the case, and placed into 1 category accordingly. These categories are:

a. *Insufficient Information:* Cases considered insufficiently documented for medical assessment or discussion of possible spread of toxin will not be further reviewed in this analysis, but additional information will be requested in follow-up as appropriate.

b. *Inconsistent with Pharmacology of Botulinum Toxin Effects:* Cases that reported a clinical course of events inconsistent with the pharmacology of botulinum toxin (eg, rapidly fluctuating signs and symptoms, increasing spasticity) will not be reviewed further in this

analysis. Specific note is made of the constitutional symptoms (Peters 2005) that have been reported as soon as within hours of treatment. These include complaints of generalized fatigue, flu-like symptoms, dizziness, asthenia, nausea, etc. These constitutional symptoms are not accepted as due to possible spread of toxin in this analysis of neurologic symptoms. Hypersensitivity events, are immunologic rather than pharmacologic and are not consistent with the definition of PSOT.

c. *Localized Neuromuscular Weakness:* Cases reporting only locally occurring events relative to the site of injection will be captured separately for this analysis.

d. *Confounded Cases:* Cases reporting events that are more likely attributed to other medications and/or underlying medical conditions will not be reviewed further in this analysis.

e. *Remaining Cases:* Cases not meeting the criteria above will be analyzed further. These remaining cases will then be subjected to an aggregate reassessment according to Step 3.

Step 3. Allergan will apply additional medical context to the evaluation of the remaining cases to increase precision of the analysis. This will be done by taking into consideration the clinical presentation of neurological signs and symptoms in individual case reports, with forethought to the known clinical presentation of foodborne and wound botulism. Cases not considered to be clinically related to PSOT include presentations with no neurologic signs or symptoms distant from the site of injection and, therefore, will be candidates for exclusion.

Any remaining cases subsequent to the Step 3 analysis will be further described, including case narratives and concluding observations. In addition, cases with localized neuromuscular weakness (local spread of toxin) will have been identified in Step 2c. The incidence of PSOT TEAE identified in Step 1 will be summarized by preferred term in descending order of incidence rate for the entire study, and for each treatment cycle.

6.5 CT Scan Safety Measures

Using a standardized protocol, CT scans for each subject will be captured by a CT technologist at screening (baseline), treatment cycle 1 day 90, and the exit visit, and the 9 measures listed below will be assessed by a central reader.

For each of the quantitative measures, summary statistics will be presented for baseline, mean change from baseline, and percent change from baseline (sample size, mean, standard deviation, median, minimum, and maximum) by treatment regimen for all scheduled visits.

For each of the qualitative measures, the incidence of change from baseline will be summarized on a per-subject basis. Interpretation of change from baseline will be provided by the central reader in comments.

Quantitative Safety Measures

1. bigonial width (mm)
2. cortical thickness under the insertion point of the masseter muscle at the mandibular ramus (mm) (left and right)
3. mandibular flare angle (degrees; left and right)
4. gonial angle (degrees; left and right)

Qualitative Safety Measures

1. mandibular condylar head shape (A-L)
2. mandibular condylar cortical surface (Yes or No)
3. mandibular condylar head (Regular, Thin, Thick)
4. mandibular condylar position (A-J)
5. glenoid fossa and articular eminence (Yes or No)

CT scan is only done at screening (baseline), treatment cycle 1 day 90 and the exit visit. For the CT safety measurements, the exit visit will be displayed as treatment cycle 1 exit visit and treatment cycle 2 exit visit. Only subjects who received the second injection will be included in the treatment cycle 2 exit visit. Subjects who received only 1 treatment will be included in the treatment cycle 1 exit visit.

6.6 Dental Examination

Dental history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) dictionary. A subject listing will be generated for the dental history.

For each dental measurement, the incidence of normal, abnormal not clinically relevant, and abnormal clinically relevant will be summarized by visit. Interpretation of any measurement noted as abnormal not clinically relevant will be provided by the dentist in comments.

Shift table will also be provided for post baseline dental examination vs baseline examination.

6.7 Body Weight

Summary statistics will be presented for body weight for baseline and changes from baseline (sample size, mean, standard deviation, median, minimum, and maximum) by treatment regimen for all scheduled visits.

6.8 Clinical Laboratory Evaluations

No laboratory data will be collected for this study.

6.9 Pregnancy Tests

Urine pregnancy tests are performed prior to each treatment, day 90 of treatment cycle 1, and at the exit visit for females of childbearing potential. Subjects with positive pregnancy test results will be listed by treatment regimen, including urine sample collection date, days since day 1 treatment and days since most recent treatment.

6.10 Subgroup Analyses for Safety Variables

Key safety analyses, or data summaries as appropriate, will be provided by subgroups of race (Asian and non-Asian), gender, and age (18 to 35 years and 36 to 50 years).

7. Interim Analyses

For each cohort, there will be a database snapshot with key data through the day 90 visit, for the purpose of the DRC. An interim database lock may be performed to support a regulatory filing, if needed. The treatment assignments will be unblinded to selected individuals for interim analyses. Unblinded data will be disseminated only to Allergan personnel who are directly involved in submission-related activities. These data will remain blinded for all Allergan personnel who are directly involved with the ongoing operational activities of the clinical study, all subjects, and all site personnel until the final database lock.

8. Pharmacokinetic Data Analyses

There are no pharmacokinetic data to be analyzed in this study.

[redacted]

10. Analysis for US FDA HFD-550

Not applicable.

11. Deviations from Protocol Section 7

Not applicable.

12. References

Beers MH. Anaerobic Bacteria. In: Beers MH, editor. The Merck Manual of Diagnosis and Therapy: Infectious Diseases. 18th ed. Whitehouse Station, NJ: Merck Research Laboratories; 2006:1498-1499.

Peters CJ. Infections Caused by Arthropod- and Rodent-Borne Viruses. In: Kasper DL, Braunwald E, Fauci AS, et al (eds) Harrison's Principles Of Internal Medicine; 16th Ed.. Mc-Graw Hill Companies, 2005:1161-1175.

13. Amendment(s)

Amendment 1

Date of Amendment: 30 Oct 2015

Summary

- Removed pvalue calculations for baseline and other baseline characteristics (Section 4), and Safety Analysis (Section 6).
- Removed Per Protocol (PP) population (Section 2.1) and corresponding analysis.
- Removed efficacy subgroup analyses (Section 5.5).
- Added shift table for dental examination (Section 6.6).
- Removed safety subgroup analysis by baseline MMPS grade (Section 6.10).

Amendment 2

Date of Amendment: 02 Feb 2018

Summary

- [redacted] from all analyses due to missing source documents (Section 2.1).
- Updated MedDRA version from 17.0 to 20.1 and updated preferred terms to 39 terms (Section 6.4).
- Clarified that planned treatment is used for mITT analyses (Section 2.1).

ALLERGAN

191622-130 Statistical Analysis Plan Amendment 2

Date (DD/MMM/YYYY)/Time (PT)	Signed by:	Justification
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]