NCT01917006

Study ID: 191622-133

Title: An Exploratory Study of the Safety and Efficacy of BOTOX® for the Treatment of Premature Ejaculation

Statistical Analysis Plan Date: 08 Nov 2016

Allergan
Biostatistics
Analysis Plan – Clinical Study Report

Study ID: 191622-133

Study Title: An Exploratory Study of the Safety and Efficacy of BOTOX® for the Treatment of Premature Ejaculation

Table of Contents

1. Introduction 4
1.1 Primary Study Objectives and Design 4
1.2 Secondary and Other Objectives 5
1.3 Sample Size 5
1.4 Experimental Unit and Analysis Unit 5
2. Analysis Populations and Data Conventions 5
2.1 Analysis Populations 5
2.1.1 Handling Mis-randomization and Mis-stratification 6
2.2 Analysis Visit Windows 6
2.3 Data Conventions 7
3. Disposition and Exit Status 8
3.1 Screening Log Data 8
3.2 Disposition and Exit Status 8
3.3 Study Duration 9
3.4 Protocol Deviations 9
4. Demographics and Other Baseline Characteristics 9
4.1 Demographics 9
4.2 Disease Characteristics 10
4.3 Prior Medications 10
4.4 Concomitant Medications/Procedures 10
4.5 Past Medical History 10
5. Efficacy Analyses 11
5.1 Collection of Primary Efficacy Measurement(s) and Derivation of Primary Efficacy Variable(s) 11
5.2 Primary Efficacy Analyses 12
5.2.1 Primary Analyses of Primary Efficacy Variable(s) 12
5.2.2 Other Analyses of Primary Efficacy Variable(s) 12
5.3 Secondary Efficacy Analyses 13
5.3.1 Secondary Efficacy Analysis 13
5.3.2 Multiplicity Consideration among Secondary Efficacy Analyses 14
[redacted]

[redacted]
5.6 Subgroup Analyses for Efficacy Variables 16
6. Safety Analyses 16
6.1 Study Treatment – Exposure and Administration 17
6.1.1 Exposure to Study Treatment(s) 17
6.1.2 Administration of Study Treatment(s) 17
6.2 Pre-Treatment Adverse Events 17
6.3 Pre-Treatment Serious Adverse Events 18
6.4 Adverse Events 18
6.4.1 Potential Distant Spread of Toxin Adverse Events (applicable to BOTOX® studies only) 19
6.5 Serious Adverse Events 20
6.6 Clinical Laboratory Evaluations 20
6.7 Vital Signs 21
[redacted]
6.9 Subgroup Analyses for Safety Variables 22
7. Pharmacokinetic, Biomarker, Genomic, or Immunogenicity Data Analyses 22
[redacted]
9. Interim Analyses 24
10. Analysis for US FDA 24
11. Data Collected but not Analyzed 24
12. Deviations from Protocol 25
13. References 25
14. Amendment(s) 25

1. Introduction

This document describes the statistical analysis plan for study 191622-133, an exploratory study of the safety and efficacy of BOTOX® for the treatment of premature ejaculation (PE). Patients will receive only 1 treatment of BOTOX® or placebo during the study and are followed up to 24 weeks for safety and efficacy assessments. The final statistical analysis will be conducted after database lock, when all patients have completed or exited the study. No interim analysis is planned for this study. An independent DRC will review cumulative, un-blinded safety and efficacy data. Details of analyses conducted for the DRC review can be found in the DRC statistical analysis plan.

1.1 Primary Study Objectives and Design

The primary study objectives are to explore the safety and efficacy of a range of doses of BOTOX® for the treatment of PE in male patients. This is a phase 2a, multicenter, randomized, double-blind, placebo-controlled, single treatment, pilot study, followed by an open-label observation period, to assess a range of BOTOX® doses for the treatment of male patients with PE. Patients will attend a minimum of 6 or 7 clinic visits and also have 1 or 2 telephone visits in the randomization period: screening, day 1 (treatment), day 2 (telephone visit), week 1 (phone/clinic visit), and weeks 2, 4, 8 and 12/study exit. For the open-label period, patients will have 2 clinic visits (weeks 4 and 12/study exit), along with 3 telephone visits (day 2, weeks 2, and 8).

For Cohorts 1-5, patients will be enrolled in cohorts of 10 patients. Within each cohort, 8 patients will receive BOTOX® and 2 patients will receive placebo. Patients will receive a single treatment of the study medication, with the total BOTOX® dose ranging from [redacted] The possible doses are [redacted], as well as any interim doses that are recommended by the DRC. Sentinel dosing of the first 2 patients (1 BOTOX®, 1 placebo) will occur during cohort 1 and also for any subsequent cohort that escalates the dose. Up to 2 additional sentinel patients might be enrolled to be treated with BOTOX®, prior to enrolling more patients in the dose level. These additional sentinel patients are not considered as randomized patients. For Cohorts 1-5, DRC will review cumulative unblinded safety and efficacy data when at least 4 weeks of data are available for all patients remaining in a cohort. The DRC will make recommendations for dose escalations, dose de-escalations, assessing alternative doses from the planned dose escalation scheme, repeating cohorts, and

stopping the study.
Based on DRC recommendations, Cohort 6 is added, where patients will be enrolled into the randomized period in a 1:1 ratio, with 12 patients receiving a single treatment of BOTOX® [redacted] and 12 patients receiving placebo. Cohort 6 patients may choose to participate in the open-label period where they will receive BOTOX® [redacted] For Cohort 6, no data review will be conducted by the DRC as this will be the final cohort of the study.

1.2 Secondary and Other Objectives

None.

1.3 Sample Size

Sample size calculation and power consideration is discussed in Section 7.5 of the study protocol.

1.4 Experimental Unit and Analysis Unit

Both the experimental unit and the analysis unit are the study subject.

2. Analysis Populations and Data Conventions

2.1 Analysis Populations

There are 3 analysis populations defined in this study: modified intent-to-treat (mITT), per protocol (PP), and safety:

- The mITT population will be used to analyze all efficacy variables and summarize baseline characteristics. The mITT population is defined as all randomized patients that have received study medication and have both baseline and postbaseline IELT data available. Analyses on the mITT population will be based on the dose actually received by the patient.
- The PP population includes patients in the mITT population who have had no protocol deviations that affect the primary efficacy variable. The PP population will be used for analyzing the primary efficacy variable and summarizing baseline characteristics based on the dose received by the patient.
- The safety population includes all treated patients and will be used in the analysis of all safety data. Analyses on the safety population will be based on the dose received by the patient.

2.1.1 Handling Mis-randomization and Mis-stratification

There is no stratification factor in the randomization of this study. If a subject is mis-randomized, i.e., taking the wrong medication as randomized, the subject will be analyzed as treated (using the treatment received) for the mITT and Safety populations, while excluded from analysis using the PP population.

2.2 Analysis Visit Windows

All days are referenced to day 1 (day of study medication treatment). Unless otherwise defined, baseline value is the last non-missing observation prior to treatment on day 1. Unless otherwise specified, visit windows are defined as below for mITT/Safety population for up to the end of the randomized period of the study, using study days relative to the day of study medication treatment for the randomized period of the study:

For the open-label period of the study, the visit windows are defined as in the table below. All the study days are relative to the day of study medication treatment for the open-label period of the study.

Unless otherwise stated, when multiple observations are captured in the same visit window, the observation closest to the target date of visit will be used for the window. If 2

observations are equidistant from the target date, then the latter observation will be used for the window. The only exception is for clinical laboratory variables, where the last non-missing observation will be used. Because re-runs of laboratory variables may only involve one or a few variables, leaving the rest missing, this rule will be applied separately for each variable for non-missing data only.
A separate window for the PP population would not be generated as the population is only used to analyze IELT data, which is not window-based.

2.3 Data Conventions

- A one-sided, alpha=0.1 significance level will be used for this study for treatment comparisons, i.e., treatment differences are considered significant if the 1-sided p-value of the test used in the comparison is less than or equal to 0.1.
- Data will be pooled across centers in all analyses. Placebo patients from different cohorts will also be pooled to form the placebo treatment group in all analyses.
- Day 1 is defined as the day of the first study medication administration. Study day is calculated as: Study day=visit date-day 1 date +1 for visits occurred after day 1 and study=visit date -day 1 date for visits occurred before day 1.
- The Medical Dictionary for Regulatory Activities (MedDRA) will be used to code all adverse events and medical history.
- Medications will be coded by the World Health Organization Drug Dictionary Enhanced (WHO DDE) preferred name for the medication's name and MedDRA for the medication's indication.
- All adverse events/medications with missing onset/stop dates will be identified and the missing dates will be imputed as follows:
  a) For adverse event onset date: if day and month are missing but year is available then the imputed date would be January 1st of that year, or day 1 if they are in the same year. If day is missing but month and year are available then the imputed day will be the first day of the month or day 1 if they are in the same month and year.
  b) For adverse event stop date: if day and month are missing but year is available then the imputed day will be December 31st of that year, or the study exit date if they are in the same year. If only day is missing but moth and year are available then the imputed day will be the last day of the month of the year, or the study exit date if they are in the same month and year.

- Data from unscheduled/re-test visits, excluding lab data, will also be treated the same way as scheduled visits, i.e., put in the visit window and decided by the visit day's distance to target day as which one to be used in the analyses.
- Weight and height will be presented with the metric system units and laboratory data will be presented with the Standard International (SI) units.
- Summary statistics for continuous/ordinal data include sample size (N), mean, standard deviation, minimum, median, and maximum and that for categorical data include sample size (N), frequency count and percentage.
- Change from baseline is calculated as follow-up minus baseline, and treatment difference will be calculated as BOTOX® minus placebo.
- The type III sums of squares will be used for all analysis of covariance (ANCOVA) and analysis of variance (ANOVA) models.
- Unless otherwise specified, comparisons on categorical data will be performed using Pearson's Chi-square test. When 25% or more of the cells have expected counts of less 5, Fisher's exact test will be used.
- Except for AE and medication onset/stop dates, missing data will not be imputed.

3. Disposition and Exit Status

3.1 Screening Log Data

The number and percent of patients who are screen failures as well as who qualified, along with the reasons for failure, will be tabulated. The demographic information (age, race and sex) of the failed and qualified patients will be summarized. A by-site listing of the screen failures with failure reasons will be also provided.

3.2 Disposition and Exit Status

For each analysis population (safety, mITT and PP), patient disposition (enrolled, completed, discontinued, and reasons for discontinuation) will be presented by treatment group. A cumulative frequency table showing patient disposition (continuing, completed, discontinued and reasons for discontinuation) will be provided for the mITT population by visit.
A by-patient data listing presenting exclusions from the PP analyses and the primary reason(s) for exclusion will be provided.
The exit status of the patients is classified as completed or discontinued. The reason for discontinuation include adverse event, lack of efficacy, lost to follow-up, personal reasons, protocol violation and other. A listing of discontinued patients with exit reasons will be provided.

Enrollment by site will be tabulated with the number and percent of patients screened, randomized, treated, completed and discontinued.
The summaries will be provided for both the randomized period and for the open-label period of the study, where applicable.

3.3 Study Duration

The length of study duration for each subject is defined as the duration between the baseline and exit dates. Summary statistics of study duration by treatment will be provided for the randomized period, the open-label period and the whole study period (including the randomized period and the open-label period). Duration for the open-label period starts from the treatment day (day 1) of the open-label period.

3.4 Protocol Deviations

Significant protocol deviations will be summarized with incidence by deviation type broken further by whether deviation leads to data exclusion from per protocol analysis for both the randomized period and the open-label period.

4. Demographics and Other Baseline Characteristics

Unless otherwise specified, summaries planned in this section will be presented using the mITT population, for patents who participated in the randomized period and for those who participated in the open-label period. If a subject participates in the open-label period, the demographics and baseline characteristics for the open-label period use the same values as those for the randomized period. For summaries for the open-label period, comparisons will not be performed.

4.1 Demographics

Age, race, sex are collected in the demographics Case Report Form (CRF) and height and weight are collected as part of the pre-treatment vital signs. All will be summarized using frequency table and/or summary statistics, both by treatment group and overall. Age will be further categorized into <35 and ≥ 35 years of age groups. Race will be summarized as Caucasian, Black, Asian, Hispanic and other, as well as Caucasian vs. non-Caucasian (Black, Asian, Hispanic and other combined).
Patients' age, weight and height will be analyzed using a one-way ANOVA to evaluate any baseline differences among the treatment groups. The distributions of race (Caucasian and

non-Caucasian) will be compared among the treatment groups using Pearson's Chi-square test or Fisher's exact test. A patient level listing will be provided for all demographic data, including height and weight.

4.2 Disease Characteristics

Baseline disease characteristics include baseline IELT, [redacted] Summary statistics will be provided for each variable by treatment group as well as overall. An one-way ANOVA will be used to compare these characteristics among treatment groups. For IELT, as the eligibility is based on the average of the latest four eligible events collected at during screening period, arithmetic mean will be used.

4.3 Prior Medications

Prior medications are medications captured in the CRF that are taken prior to the day 1. A medication is a prior medication if the (imputed) onset date is before day 1.
The number and percent of patients who had taken prior medications will be presented for each treatment group and overall by the WHO DDE preferred name and the MedDRA preferred term (PT) of the medication's indication.

4.4 Concomitant Medications/Procedures

Concomitant medications are medications taken on or after the day of study medication administration. A medication is a concomitant medication if the (imputed) onset date is on or after day 1, or if the onset date is before day 1 and the stop date is after day 1 (discontinued after day 1 or ongoing during the study).
The number and percent of patients who had taken concomitant medications will be presented for each treatment group and overall by the WHO DDE preferred name and the MedDRA PT of the medication's indication.

4.5 Past Medical History

Medical history with onset dates prior to the informed consent date will be tabulated and presented for each treatment group and overall. Medical history will be coded using MedDRA primary system organ class (SOC). All medical history will be presented by

treatment group and overall. No statistical comparisons will be performed. A patient-level listing of medical and surgical histories will also be provided for each treatment group.

Patients are also asked to fill out a sexual history CRF which will record their PE therapy or techniques together with start/stop date information. The un-coded therapy/technique will be tabulated with number and percent of patients using each therapy/technique by treatment group and overall. A patient-level listing of the information collected will also be presented.

5. Efficacy Analyses

5.1 Collection of Primary Efficacy Measurement(s) and Derivation of Primary Efficacy Variable(s)

The primary efficacy measurement is the patient's IELT (second) collected from the Sex Intercourse Diary (SID). The SID will be completed by the patient and/or their female partner after any vaginal sexual activity during the screening visit and for 12 or 24 weeks following study treatment. Only IELTs from evaluable events recorded in the SID will be used for analyses.

An evaluable event is defined as:

- An event in which the patient and his partner had vaginal intercourse, and the couple have recorded that the ejaculation occurred "intravaginally" and duration of the time from penetration to ejaculation has been recorded

OR

- An event in which the patient and his partner attempted to have vaginal intercourse. However, the couple have recorded that ejaculation occurred "prior to vaginal penetration". In this scenario, the duration of time will be recorded as zero.

Any other scenarios will not be considered as evaluable and will not be included in deriving the primary efficacy variable, although time to withdrawal can be used in exploratory analyses. Number of evaluable events will be summarized by treatment group at baseline, weeks 2, 4, 8 and 12 for the randomized period of the study.

The primary efficacy variable is the change from baseline in logarithm value of the individual patient's geometric mean IELT for the entire 12 weeks of the randomized period of the study.

For baseline, the geometric mean IELT is calculated from the most 4 recent evaluable events prior to study medication administration. The geometric mean of the IELTs from all postbaseline evaluable events up to week 12 is then calculated and the difference in the logarithm of the two geometric mean IELTs (postbaseline - baseline) will be calculated.

For IELTs that are zero (events in which ejaculation occurred prior to vaginal penetration), 0.5 (second) will be added for the calculation of geometric means.

5.2 Primary Efficacy Analyses

The analyses on the primary efficacy variables will be performed on both the mITT and the PP populations. Comparisons will be made between each of the BOTOX® dose levels and placebo without multiplicity adjustment.

5.2.1 Primary Analyses of Primary Efficacy Variable(s)

The primary null hypothesis is that at week 12 the difference in mean change from baseline in logarithm value of the geometric mean IELT between BOTOX® and placebo is at zero, and there is no dose response shown in the change from baseline in logarithm value of the geometric mean IELT.

To test the primary null hypothesis, the primary efficacy variable in relation to dose levels will be analyzed using an ANCOVA model with treatment as the fixed effect and baseline logarithm value in geometric mean IELT as the covariate to assess the relationship of response across the BOTOX® dose levels. Comparisons of the LSMEANS from the ANCOVA model will be made for each of the active BOTOX® dose levels vs. placebo. Coefficient of variation (CV) will be presented for the geometric IELT, and the LSMEAN and its 95% confidence interval (CI) as well as treatment difference in LSMEAN and its 95% CI for each active BOTOX® dose level vs. placebo will be back-transformed to original unit (seconds) to facilitate review.

The same ANCOVA model using the PP population will be performed.

To assess the dose-response of the active treatment, a linear regression model with baseline logarithm value in geometric mean IELT and dose level (numeric value) as the independent variables and the primary efficacy variable as the dependent variable in the model will be conducted at week 12. The same analysis will also be carried out for the PP population.

5.2.2 Other Analyses of Primary Efficacy Variable(s)

As a sensitivity analysis to the log-normal assumption of the geometric mean IELT, Wilcoxon rank-sum test will be performed on the primary efficacy variable for the treatment difference for each of the active BOTOX® dose level vs. placebo.

In addition, responder analyses for the primary efficacy variable will also be performed. As treatment responder in PE are defined as 3-8 fold changes in geometric mean IELT, corresponding responders will be defined in the logarithmic value of a 3, 4, 5, 6, 7 and 8 fold

change (1.10, 1.39, 1.61, 1.79, 1.95, and 2.08, respectively) for the primary efficacy variable and the Pearson's chi-squared test will be used to compare the proportion of responders between each of the active BOTOX® dose group vs. placebo.

5.3 Secondary Efficacy Analyses

Secondary efficacy variables include the following:

- Geometric mean IELT up to weeks 2, 4, 6, 8 and 10 of the randomized period, respectively.
- Average (arithmetic mean) IELT up to weeks 2, 4, 6, 8, 10 and 12 of the randomized period, respectively

Interval-based geometric mean IELT and average IELT for week 4, 8 and 12 (ie, events between baseline and week 4, between week 4 and week 8, and between week 8 and week 12) will also be calculated from the above two as differences between periods.

All analyses on the secondary efficacy variables will not be window-based as indicated in Section 2.2. The target day for each time point, except for week 12, will be used to determine to which interval or period a certain event belongs to. For example, all events on or before week 10 target day, day 70, will be included for analyses with IELTs up to week 10, while all events between day 29 (inclusive) and day 56 (inclusive) will be included for analyses with IELTs between week 4 and week 8. Any event data collected on or after the week 12 target date of day 84 of the randomized period, and before the day 1 (treatment day) of the open-label period, if applicable, will be included in analyses related to week 12.

Comparisons will be made for each active BOTOX® dose level vs. placebo without multiplicity adjustment.

5.3.1 Secondary Efficacy Analysis

Summary statistics will be provided for each of the secondary efficacy variables. Change from baseline in log-transformed geometric mean IELTs and change from baseline in average mean IELTs will also be calculated and summarized.

Same ANCOVA model used in Section 5.2.1 will be used to the change from baseline in the log-transformed geometric mean IELT up to weeks 2, 4, 6, 8, and 10, respectively as well as change from baseline in interval-based log-transformed geometric mean IELT for weeks 4, 8

and 12, with treatment as the fixed effect, and log-transformed baseline geometric mean IELT as the covariate.

Summary statistics will be provided for average IELT as well as change from baseline in average IELT by timepoint. A similar ANCOVA model will be used on change from baseline in average IELT as well as the interval-based average IELTs: the model will have treatment as the fixed effect and the baseline average IELT as the covariate. Change from baseline in average IELTs up to weeks 2, 4, 6, 8, 10 and 12 and the interval-based change from baseline in average IELT for weeks 4, 8 and 12 will not be log-transformed prior to analyses.

5.3.2 Multiplicity Consideration among Secondary Efficacy Analyses

Not applicable

5.4 Other Analyses of Secondary Efficacy Variables

Nonparametric analysis using the Wilcoxon rank-sum test will be performed on the geometric mean IELT up to weeks 2, 4, 6, 8 and 10 as well as interval-based geometric mean IELT for weeks 4, 8 and 12, to compare each active BOTOX® dose level vs. placebo. Responder analyses described in 5.2.2 will also be carried out for geometric mean IELTs up to weeks 2, 4, 6, 8 and 10.

Treatment comparisons of each active BOTOX® dose level vs. placebo Geometric mean IELT up to weeks 4, 8 and 12 will also be performed using methods described in Zhou et al, 1997:

1. The likelihood-based Z-score test: the comparison between an active BOTOX® dose level *i* and placebo is evaluated by the statistic Z_i:

$$Z_i = \frac{\hat{\mu}_i - \hat{\mu}_0 + \frac{1}{2}(S_i^2 - S_0^2)}{\sqrt{\frac{S_i^2}{n_i} + \frac{S_0^2}{n_0} + \frac{1}{2}\left(\frac{S_i^4}{n_i - 1} + \frac{S_0^4}{n_0 - 1}\right)}}$$

Where $\hat{\mu}_i$ and $\hat{\mu}_0$ are the sample mean of the log-transformed geometric mean IELT of dose level *i* and placebo, S_i^2 and S_0^2 are the sample variance of the log-transformed geometric mean IELT of dose level *i* and placebo, and n_i and n_0 are the sample size of dose level *i* and placebo, respectively.

Assuming that for each of the active dose level *i* and placebo, the geometric mean IELT are independently log-normally distributed with means M_i and M_0, respectively, then Z_i tests the null hypothesis of M_i=M_0, and under the null hypothesis Z_i is approximately standard normal.

2. Bootstrap approach:

   Let $\overline{T}_i$, $\overline{T}_0$ and $\overline{T}^{(i,0)}$ be the sample mean of the geometric mean IELT of the dose level i, placebo and the two groups combined, respectively. $T_k^{(i)} = T_{ik} - \overline{T}_i + \overline{T}^{(i,0)}$ and $T_j^{(0)} = T_{0j} - \overline{T}_0 + \overline{T}^{(i,0)}$ form a sample with common mean $\overline{T}^{(i,0)}$, where T_{ik} and T_{0j} are the geometric mean IELT for the k[th] patient in dose level i and the j[th] patient in the placebo group, for k=1, …, n_i and j=1,…, n_0. Resample from them with replacement 2000 times, $Z_{i1}^{*(b)},\ldots, Z_{in_i}^{*(b)}$ and $Z_{01}^{*(b)},\ldots, Z_{0n_0}^{*(b)}$ for each b=1, …, 2000. Compute

$$t_{b,i}^{*} = \frac{\overline{Z}_i^{*(b)} - \overline{Z}_0^{*(b)}}{\sqrt{\frac{\tau_i^{*(b)2}}{n_i} + \frac{\tau_0^{*(b)2}}{n_0}}}, \text{ for b=1,…, 2000.}$$

   Where $\overline{Z}_i^{*(b)}, \tau_i^{*(b)2}$ $\overline{Z}_0^{*(b)}$, and $\tau_0^{*(b)2}$ are the sample mean and variance for the active does level i and placebo for the b[th] resample, respectively.

   The test statistic,

$$t_{obs,i} = \frac{\overline{T}_i - \overline{T}_0}{\sqrt{\frac{\hat{\tau}_i^2}{n_i} + \frac{\hat{\tau}_0^2}{n_0}}}$$

   Where $\hat{\tau}_i^2$ and $\hat{\tau}_0^2$ are the sample variances of the observed geometric mean IELT of active dose level i and placebo, respectively.

   Then the achieved significance level (ASL) for testing H_0: $M_i=M_0$ vs. H_a: $M_i \neq M_0$, for the dose level i, where M_i and M_0 are the means of the IELTs of dose level i and the placebo under the log-normal distribution, respectively.

$$\hat{ASL}_i = \frac{\#\ of\ times\ |t_{b,i}^{*}| \geq |t_{obx,i}|}{2000}$$

[redacted]

[redacted]

5.6 Subgroup Analyses for Efficacy Variables

No subgroup analysis is planned.

[redacted]

[redacted]

6.1 Study Treatment – Exposure and Administration

6.1.1 Exposure to Study Treatment(s)

Patient exposure to the study medication will be characterized by study days. Study day is defined as the number of days from day 1 to exit; if the date of exit is missing, the date of the last visit will be used. Data will be summarized with descriptive statistics for each treatment group using safety population.

6.1.2 Administration of Study Treatment(s)

Patients will receive a single treatment administration during the study, consisting of 2 injections (each containing half of the total dose) of the assigned study medication to be delivered bilaterally (1 injection per side) to the bulbospongiosus muscle. Injection of the study medication will be performed using ultrasound guidance.
The following options are permitted either alone or in combination before treatment administration:

- No anesthesia
- Topical local anesthesia
    - Sterile local anesthetic topical preparation (eg, lidocaine cream) may be applied to the skin around the injection site
- Local anesthetic injections to skin
    - Sterile local anesthetic solution (eg, lidocaine solution) may be injected to the skin overlaying the injection site
- Sedation
    - If it is clinically indicated, may be administered to the patient at the investigator's discretion

Information on injection of study medication, including date, time of injection, number of injection sites, total volume injected and injection location will be provided in a data listing.

6.2 Pre-Treatment Adverse Events

Adverse Events with an onset date that is before day 1 of the randomized period of the study will be counted as pretreatment adverse events.

Adverse events will be coded from the verbatim text into MedDRA primary SOC and PT.
Pretreatment AEs will be presented in descending order of incidence rate by the treatment group patient was randomized to for all randomized patients.
No statistical comparison will be performed.
A patient listing will be generated for pretreatment adverse events.

6.3 Pre-Treatment Serious Adverse Events

Pre-treatment serious adverse event will be collected and coded the same way as mentioned in Section 6.2 and a listing will be generated for pretreatment serious adverse events.

6.4 Adverse Events

AEs will be coded from the verbatim text into PT and the primary SOC using MedDRA.

AEs are collected both for the screening/baseline period pretreatment (which are referred to as pretreatment AEs) and for the follow-up period after treatment is initiated (which are referred to as posttreatment AEs). A treatment-emergent adverse event (TEAE) is an adverse event with onset on or 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. For the randomized period of the study, TEAEs are those with onset on or after the first study treatment, and prior to the study treatment for the open-label period, if applicable; for the open-label period, TEAEs are those with onset on or after the study treatment for the open-label period.

Three incidence rate tables are standard for summarizing all TEAEs:

- by descending order of incidence rate
- by primary SOC and PT
- by primary SOC, PT and severity

Number and percent of patients with TEAEs will be tabulated by relationship to study drug and study procedures. All treatment-related TEAEs, either study drug-related or study procedure-related will also be presented by descending incidence and by primary SOC and PT for each treatment group. Separate tables will be generated for study drug-related and study procedure-related TEAEs. TEAEs leading to study discontinuation will also be tabulated by primary SOC and PT for each treatment group.

No statistical comparisons will be made for any adverse event table.

6.4.1 Potential Distant Spread of Toxin Adverse Events (applicable to BOTOX® studies only)

To assess possible distant spread of toxin (PDSOT), 40 MedDRA (version 17.0) preferred terms that may be associated with botulinum toxin effects have been identified. All AEs associated with PDSOT will be tabulated by SOC and treatment group; in addition, 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
- 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
- VIIth nerve paralysis
- Facial paresis
- Hyporeflexia
- Hypotonia
- Paralysis
- Paralysis flaccid
- Paresis cranial nerve
- Peripheral nerve palsy
- Peripheral paralysis
- Speech disorder
- Vocal cord paralysis
- Vocal cord paresis

Renal and Urinary Disorders
- 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: 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 list may need to be updated before database lock due to version change of the MedDRA.

6.5 Serious Adverse Events

The number and percent of patients with serious TEAEs will be tabulated by treatment group for all serious TEAEs and treatment (study drug- or study procedure-, separately) related serious TEAEs by primary SOC and PT. Serious TEAEs will leading to study discontinuation will also be summarized by Primary SOC.
All serious adverse events and deaths will also be presented in a listing for each treatment group.

6.6 Clinical Laboratory Evaluations

Hematology and non-fasting clinical chemistry blood samples and urine sample are collected and processed by a central laboratory at screening, week 4 and week 12 for both the randomized period and the open-label period of the study. Measures collected are:

- Hematology: Hematocrit, hemoglobin, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, mean corpuscular volume (MCV), platelets, red blood cell (RBC) count, RBC morphology, total white blood cell (WBC) count and differential (eg, neutrophils, bands, lymphocytes, monocytes, basophils, and eosinophils).
- Nonfasting serum chemistry: albumin, alkaline phosphatase, alanine transminase, aspartae transminase, bicarbonate, calcium, chloride, creatinine, creatine kinase, direct bilirubin, nonfasting glucose, indirect bilirubin, magnesium, phosphorous, potassium, sodium, total bilirubin, total cholesterol, total protein, urea nitrogen, and uric acid.
- Urine samples: clarity, color, bilirubin, blood, glucose, ketones, leukocyte esterase, nitrite, pH, protein, specific gravity, urobilinogen, microscopic sediment (WBCs, RBCs, casts, bacteria, crystals, and epithelial cells)

An urine culture and sensitivity will be analyzed by the central lab when the central lab urine analysis results are suggestive of a urinary tract infection (eg. Positive leukocyte esterase, nitrites, blood, or microscopic sediments such as WBCs, RBCs and/or bacteria)
All clinical laboratory data analyses and presentations are to be expressed in the SI (Standard International) units.
Hematology, blood chemistry and urine analysis values will be summarized by visit with descriptive statistics. Change from study baseline (defined as the last assessment prior to study medication administration) will be summarized with descriptive statistics. For the

randomized period, baseline is the last assessment prior to the study medication of the randomized period of the study. For the open-label period, baseline is the last assessment prior to the study medication of the open-label period of the study. Within-group change will be evaluated by Wilcoxin Signed-rank test.

Laboratory values will be categorized as low, normal, and high according to the reference normal range. For each variable, cross-tabulations (shift tables) of low/normal/high values at study baseline will be compared to those at each post-treatment visit.

Listing of all laboratory data, as well as a patient-level listing of all abnormal (high/low) laboratory values will be provided.

6.7 Vital Signs

Vital signs include pulse rate (bpm) systolic/diastolic blood pressure (mmHg) and body temperature (°C) are collected at screening, day 1 (both pre- and post-treatment), week 1 (sentinel patients only), weeks 2 ,4,8 and 12. Summary statistics for the raw value as well as change from baseline will be provided by treatment group at each time point the data are collected. No statistical comparison will be made for vital signs. A patient-level listing will also be provided. Baseline for the randomized period and for the open-label period are defined similarly to those for the lab assessments.

[redacted]

6.9 Subgroup Analyses for Safety Variables

No subgroup analysis for safety variables is planned.

7. Pharmacokinetic, Biomarker, Genomic, or Immunogenicity Data Analyses

Not applicable for this study.

[redacted]

[redacted]

9. Interim Analyses

No interim analysis is planned for this study.

10. Analysis for US FDA

Not applicable for this study.

11. Data Collected but not Analyzed

Not applicable for this study.

12. Deviations from Protocol

None.

13. References

Zhou XH, Gao S, Hui SL. Method for comparing the means of two independent log-normal samples. Biometrics. 1997; 53(3): 1129-1135.

14. Amendment(s)

- Section 1.1 amended to include the addition of Cohort 6 based on DRC recommendations and also the changes in doses to Cohort 1- 5.
- Section 2.2 amended to add the Analysis Visit Windows for the open-label period of the study.
- Section 3.2 – amended to add analyses on disposition and exit status for the open-label period of the study.
- Section 3.3 – amended to add analyses on study duration for the open-label period of the study.
- Section 3.4 – amended to add analyses on protocol deviations for the open-label period of the study.
- Section 4 – amended to add analyses on demographics and baseline characteristics for the open-label period of the study.
- Section 5.1 and 5.3 – amended to limit the efficacy analyses to the randomized period only.
- Sections 5.5.1 and 5.5.2 – amended to limit the analyses on BSM measurements to Cohort 5 only, and to add exploratory analyses for the open-label period.
- Sections 6.2, 6.4, 6.5, 6.6, and 6.7 – amended to add analyses on safety for the open-label period of the study.
- Sections 8.1, 8.2, 8.3, and 8.4 – amended to add exploratory analyses and safety analyses for the open-label period of the study.

- Section 9 – Deleted the statements that DRC will perform ongoing unblinded review of both safety and efficacy, to be consistent with the protocol.

ALLERGAN

Analysis Plan - 191622-133 Amendment 2

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