Allergan
Biostatistics
Analysis Plan

Study ID: 191622-137

Study Title: BOTOX® in the Treatment of Urinary Incontinence Due to Overactive Bladder in Patients 12 to 17 Years of Age

Table of Contents

List of Tables

1. Introduction

This document details the planned analyses for Study 191622-137, a multicenter, randomized, double-blind, parallel group, multi dose study to evaluate the safety and efficacy of BOTOX for the treatment of urinary incontinence due to overactive bladder (OAB) in patients 12 to 17 years of age who have not been adequately managed with anticholinergic therapy.

There will only be one analysis for this study. When all patients in Study 191622-137 exit the study, there will be a database lock (DBL), and a clinical study report (CSR) will be generated. This analysis plan outlines the outputs to be included in the CSR.

1.1 Primary Study Objectives and Design

The primary objective is to evaluate the safety and efficacy for the treatment of urinary incontinence due to OAB in patients 12 to 17 years of age who have not been adequately managed with anticholinergic therapy.

Patients may participate in the study for 96 weeks following entry into the study and should have at least 12 weeks follow-up since the last treatment prior to exiting the study. Eligible patients are randomized to 1 of 3 treatment groups in a 1:1:1 ratio:

- 25 units (U) BOTOX (not to exceed 6 U/kg)
- 50 U BOTOX (not to exceed 6 U/kg)
- 100 U BOTOX (not to exceed 6 U/kg)

In order to ensure that an upper dosing limit of 6 U/kg is not exceeded, the actual dose administered will be adjusted based on the patient's weight if necessary.

Patients will be centrally randomized. In order to ensure balance across treatment groups, patients will be stratified by baseline daytime urinary urgency incontinence episodes (a total of ≤ 6 episodes or > 6 episodes over the 2-day diary collection period).

Multiple treatments may be administered in this study. Doses received at subsequent retreatments will be determined by the investigator and based on the patient's response to the preceding blinded treatment. The investigator can elect to keep the dose the same or increase the dose compared with the preceding treatment. If it is deemed that a decrease in dose is warranted, the patient will be exited from the study. The first treatment will be administered on day 1.

All patients will be evaluated at scheduled clinic visits at weeks 2, 6, and 12 post-treatment, and thereafter at alternating telephone and clinic visits every 6 weeks until the patient qualifies for further retreatment or exits the study. Patients can request retreatment from week 12 since the preceding treatment. Once qualified for retreatment, the same visit schedule will be followed.

The primary efficacy measure is daytime urinary incontinence episodes, and the primary time point is week 12 after treatment 1.

1.2 Secondary and Other Objectives

This section is not applicable.

1.3 Sample Size

The sample size calculation for this study is determined empirically. A total of approximately 108 patients are to be enrolled for the study (32 per treatment group plus 10% attrition).

The following table displays the estimated differences between 100 U of BOTOX (maximum 6 U/kg) and 25 U of BOTOX (maximum 6 U/kg) in mean daily average frequency of daytime episodes of urinary incontinence change from baseline at week 12 for a range of common standard deviations (SDs). All calculations assume 32 patients per treatment group and a 2-sided test significance level of 0.05. The calculation was performed using the commercial software nQuery Advisor (procedure MTC0-1), version 6.01 (Statistical Solutions Ltd., Cork, Ireland), and the CI for the difference of 2 means based on normal distribution.

Standard Deviation	Distance From Mean to the Limit
2.0	0.980
2.5	1.225
3.0	1.470
3.5	1.715
4.0	1.960

Note: The ranges of the SDs are determined based on the values from Allergan's 2 pivotal phase 3 OAB studies, 191622-095 and 191622-520. The estimate of the SD for the 100-U BOTOX treatment group is 3.33 for Study 191622-095, 3.58 for Study 191622-520, and 3.46 for the 2 studies combined. The distance from mean to limit indicates how much the CI for the mean extends from the observed difference in means in either direction.

1.4 Experimental Unit and Analysis Unit

The experimental unit is study subject. The primary efficacy assessment is daytime urinary incontinence episodes measured at the patient level.

2. Analysis Populations and Data Conventions

2.1 Analysis Populations

Patients may receive multiple treatments in this study. Eligible patients are initially randomized to one of three treatment groups, namely, 25, 50, or 100 U BOTOX treatment groups. The investigator can elect to administer subsequent retreatments in a blinded fashion by keeping the dose the same or increase the dose compared with the preceding treatment. The dose that is actually received by a patient can vary due to the dose limit of 6 U/kg (i.e., the dose is adjusted based on the patient's weight, see Section 12.1 of the study protocol for details).

In order to maintain blinding of the BOTOX dose received, the drug is reconstituted by an independent drug reconstitutor (IDR) based on treatment assigned by IVRS/IWRS (which also includes the patient's weight as provided in IVRS/IWRS). Study drug accountability and reconstitution records are then reviewed and confirmed by independent drug monitors (IDM). If an incorrect dose was administered, the IDM will complete the "Incorrect Unit Dose Administered" form and the "incorrect unit dose administered" will be entered into the electronic case report form (eCRF).

The actual dose unit received by a patient will be derived as follows:

1. If the "incorrect unit dose administered" is populated in the eCRF, this dose will represent the actual dose unit received.
2. If the "incorrect unit dose administered" is not populated in the eCRF, then the IVRS/IWRS assigned dose (ie, 25U, 50U, or 100U) and the patient's weight recorded in IVRS/IWRS at the treatment visit will be utilized to determine the actual dose unit administered as described in Section 12.1 of the study protocol.

One analysis population will be used in the statistical analysis of this study: the BOTOX-Treated Population. The BOTOX-Treated Population will include all patients enrolled into the study who received at least 1 BOTOX treatment and will be used for all efficacy and safety analyses. The BOTOX treatment group used for analysis in each treatment cycle will be derived as stated in Table 1 below based on the dose actually received in the respective treatment cycle.

Table 1 BOTOX Treatment Groups Based on Actual Dose Received

Actual BOTOX Dose Unit Received	BOTOX Treatment Group used for Analysis
< 37.5 U	25 U BOTOX
37.5 U ≥ and < 75 U	50 U BOTOX
≥ 75 U	100 U BOTOX

2.1.1 Handling Mis-stratification

For any analysis that are by stratum or for analysis that adjusts for stratum effect, the actual strata will be used instead of the strata assigned by the site's entry into the IWRS at randomization.

2.2 Visit Windows

The target day and visit window for each post-treatment visit will be referenced to days after treatment. The visit window except for the first three and the last (Week 96) visit windows is defined as:

[(Target day of current visit + Target day of previous visit)/2+1] to [(Target day of current visit + Target day of next visit)/2]

The visit window for each study treatment cycle is given in the following table. If there are values from multiple visits in a given window, the value collected from the visit closest to the target day will be used to represent the window. If, instead of a single closest visit, 2 visits with values are equidistant from the target day, the value from the first of the 2 visits chronologically will be used to represent the window. For clinical laboratory variables, the last non-missing observation for that visit window will be used. This rule will be applied separately for each variable for non-missing data only. (The exception is because re-runs of laboratory variables may only involve 1 or a few variables and would thus include missing data for many variables).

Table 2 Analysis Visit Windows

Visit	Target Day of the Visit[a]	Study Visit Window[a]
Screening		≤ Day -1
Treatment 1 Day 1	1	Day 1
Treatment 1 Week 2	15	Day 2 – Day 29
Treatment 1 Week 6	43	Day 30 – Day 64
Treatment 1 Week 12	85	Day 65 – Day 127
Treatment 1 Week 24	169	Day 128 – Day 211
Treatment 1 Week 36	253	Day 212 – Day 295
Treatment 1 Week 48	337	Day 296 – Day 379
Treatment 1 Week 60	421	Day 380 – Day 463
Treatment 1 Week 72	505	Day 464 – Day 547
Treatment 1 Week 84	589	Day 548 – Day 631
Treatment 1 Week 96	673	Day 632 – Day 715/Exit
Treatment X Day 1	1	Day 1
Treatment X Week Y	Repeat same as treatment 1	Repeat same as treatment 1

[a] Study Day is relative to the most recent treatment day.

2.3 Data Conventions

The following data conventions will be applied to all analyses:

- Only data with exceptions will be listed for conciseness. For example, the following would not be listed: normal findings for medical history and physical examination, pregnancy tests results for males or females when not applicable, etc.
- The level of significance used for all statistical tests will be 0.05, 2-sided, unless stated otherwise. P-values ≤0.050 will be considered as statistically significant.
- The type III sums of squares will be used for all analysis of covariance models (ANCOVA) and analysis of variance models (ANOVA).
- The variance for Kaplan-Meier estimates will be calculated using Greenwood's formula.
- Descriptive statistics for continuous variables will include the sample size (n), mean, standard deviation (SD), median, minimum (min), first quartile (Q1), third quartile (Q3), and maximum (max).
- Summary statistics for categorical variables will include the sample size (n), frequency count and percent.

- The Medical Dictionary for Regulatory Activities (MedDRA) dictionary will be used to code all adverse events (AEs) and medical history.
- World Health Organization Drug Dictionary Enhanced (WHO DDE) preferred name and MedDRA will be used to code all medications.
- Study duration from randomization/day 1 will be calculated as:

  Study days = visit date – randomization/day 1 visit date + 1.
- The Cycle duration for treatment X will be calculated as:

  Study days = visit date – treatment X date + 1.
- The algorithm presented in Table 3, together with the rules below will be used to impute incomplete or missing dates for adverse events and medications as follows:

  (a) AE start dates will be imputed as the minimum of the following:
    - initial imputed date, where target date = first study drug injection date.
    - complete end date

  (b) Medication start dates will be imputed as the minimum of the following:
    - initial imputed date, where target date = first study drug injection date – 1.
    - complete end date

  (c) AE and Medication end dates will be imputed as the minimum of the following:
    - initial imputed date, where target date = study exit date + 30.
    - Death date
- For each treatment cycle, AEs will be counted with onset date between the day when the study treatment is received and the day before the next treatment is received or exit day.
- Concurrent medication will be classified using the following convention: 1) If the start date of taking medication is after or on the first BOTOX injection date, then it will be counted as concomitant medication; 2) If the start date of taking medication is prior to the first study drug injection date and stop date is on or after first study drug injection date, then it will be counted as both prior medication and concomitant medication.
- All partial dates (including AE, concurrent and prior medication) will be listed “as is” in the data listings.
- All statistical analyses will be performed using SAS® version 9 or higher.

Table 3 Initial Imputed Date Algorithm

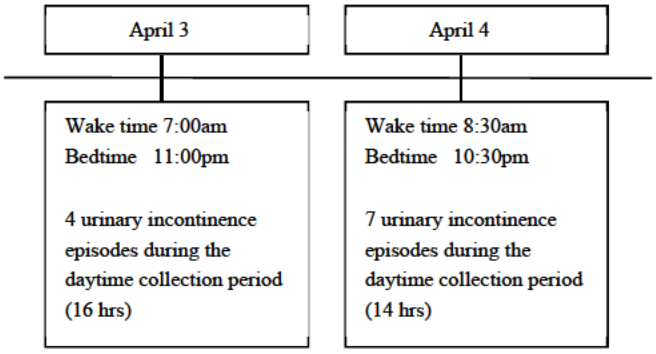
Available Year (YYYY)	Available Month (MM)			
	Missing	**< Target Month**	**= Target Month**	**> Target Month**
Missing	Target Date	—		
< Target Year	YYYY-12-31	YYYY-MM-LD		
= Target Year	Target Date	YYYY-MM-LD	Target Date	YYYY-MM-01
> Target Year	YYYY-01-01	YYYY-MM-01		

2.3.1 Diary Data Conventions

- For baseline and post-treatment visits, analyses will be based on the diary data collected over 2 consecutive days. For each day, the bladder diary will be used to collect wake time (time patients woke up to start the day), bedtime (time patients go to bed to sleep at night), urinary episodes (incontinence, urgency, catheterization, or voluntary void) during the daytime collection period, volume of each void over 1 daytime period, and presence/absence of nighttime urinary incontinence.
- A valid diary day is defined as a day where there are 1 or more urinary episodes of any type (incontinence, catheterization, or voluntary void) during the daytime collection period. In the case of patients who partially complete their diary, providing that at least 1 valid diary day out of 2 is available, the 2-day frequency of daytime urinary incontinence will be prorated with the value from the valid diary day, i.e. it is the average of the daytime urinary incontinence in the valid diary day multiplied by 2.
- If less than 1 diary day out of 2 has been completed, the frequency of daytime urinary incontinence will be set to missing.
- Missing daily average frequency of episodes of daytime urinary incontinence at baseline will be imputed using the median of all non-missing values at baseline within the same treatment dose group received for the first treatment (grouped to the nearest dose group based on the dose actually received). For scheduled visits up to week 12 after the first study treatment, missing daily average frequency of episodes of daytime urinary incontinence will be replaced using the last observation carried forward (LOCF) approach. Baseline values will not be carried forward. No imputation will be done for missing values on visits after the post-treatment 1 week 12 visit.

- For normalization of frequency of daytime urinary incontinence episodes in a given day to a 12-hour daytime period, the following algorithm will be used:

  1) Apply visit window defined in Section 2.2 that uses days from treatment as reference.
  2) Identify wake time and bedtime within the visit window. Daytime period (in hours) is determined by time between wake time and bedtime for that day.
  3) Using the example below, 4 urinary incontinence episodes are recorded during the first daytime period of 16 hours (7:00am to 11:00pm), and 7 urinary incontinence episodes are recorded during the second daytime period of 14 hours (8:30am to 10:30pm).



  On a given day, the number of daytime urinary incontinence episodes normalized to a 12-hour daytime period will be calculated by

$$\frac{12}{\text{daytime period}} \times \text{number of daytime urinary incontinence episodes},$$

  which, in the example, adjusts to 3 (= 12/16 × 4) and 6 (=12/14 × 7) normalized urinary incontinence episodes, respectively. The daily average frequency of daytime urinary incontinence episodes by the normalized daytime period in a given week visit will be 4.5 (= (3 + 6)/2) in the example.

  4) If wake time is missing, the missing wake time will be imputed by either the first urinary episode time of any type recorded on that day or 7:00 am, whichever comes earlier. If bed time is missing, the missing time will be imputed by either the last urinary episode time of any type recorded on that day or 10:00 pm, whichever comes later.

3. Disposition and Exit Status

3.1 Screening Log Data

A listing of patients that failed screening will be generated for patients that signed informed consent. The listing will include information on demographic characteristic (sex, race and age) and reason for screen failure.

3.2 Disposition and Exit Status

Patient disposition (enrolled, completed, discontinued, and reason for discontinuation) will be presented according to the dose (grouped to the nearest dose group, ie, 25, 50, or 100 U of BOTOX) received at the first treatment cycle, as well as an overall BOTOX group (ie, regardless of dose). Data will also be presented by BOTOX treatment cycle.

A listing will be generated for discontinued patients with the corresponding reason(s) for early withdrawal from the study. A data listing for the patients discontinued due to AEs will be presented.

A summary table will be generated for the number of patients who switched from one dose to the next higher dose (remaining on the same dose or a dose escalation).

3.3 Study Duration

Patients will participate in the study for at least 96 weeks following entry into the study and should have at least 12 weeks follow-up since the last treatment prior to exiting the study. The minimum duration is therefore 96 weeks, and the maximum duration is approximately 108 weeks (for patients who received their last treatment at week 96 with 12 weeks posttreatment follow-up). Thus, patients could receive a maximum of 8 BOTOX treatment cycles. For details on how to analyze duration of exposure to study treatment see Section 6.1.1.

3.4 Protocol Deviations

Significant and non-significant deviations are defined in the Protocol Deviations Specification Document and will be determined prior to database lock. A protocol deviation table and patient listing will be produced for the significant protocol deviations. The deviation type and detail will be presented.

4. Demographics and Other Baseline Characteristics

The procedures scheduled for baseline are to occur prior to receiving any study treatment.

Demographic and baseline characteristics will be summarized on the BOTOX-treated population according to the dose (grouped to the nearest dose group, ie, 25, 50, or 100 U of BOTOX) received at the first treatment cycle, as well as for the total of all BOTOX groups (ie, regardless of dose).

4.1 Demographics

Age, race, sex, weight (kg), and height (cm) will be summarized by treatment group and overall. Race will be summarized as Caucasian, Black, Asian, Hispanic, and other, as well as just Caucasian or non-Caucasian (Black, Asian, Hispanic, and other).

4.2 Disease Characteristics

Baseline characteristics, including the baseline daily average frequencies of urinary incontinence, urinary urgency, and urinary urgency incontinence, will be summarized according to treatment groups, as well as for the total of all BOTOX groups. Similarly, volume voided per micturition will be summarized according to treatment groups, as well as for the total of all BOTOX groups.

The following other baseline parameters will also be summarized by treatment group and overall: duration of OAB history and post-void residual (PVR) urine volume.

4.3 Prior Medications

A prior medication is one taken prior to the first study treatment injection in the study. Summary tables for prior anticholinergic medications for OAB and prior non-anticholinergic medications for OAB will be provided according to WHO DDE preferred drug name and treatment groups, as well as for the total of all BOTOX groups.

4.4 Concomitant Medications/Procedures

Concomitant medications are defined as those taken on or after the first study treatment, regardless of when the medications started. The number and percentage of patients taking each medication will be presented by WHO DDE preferred drug name for all BOTOX group.

All prophylactic antibiotic medication associated with study drug treatment will be summarized by WHO Drug Class and WHO DDE preferred drug name.

Concurrent procedures are defined as those undertaken on or after the first study treatment. A list of concurrent procedures will be provided for all BOTOX group.

4.5 Medical History and Surgical History

Data for medical history with onset date prior to the first study treatment will be tabulated and presented for treatment groups as well as for the total of all BOTOX groups.

A list of surgical history will also be provided.

5. Efficacy Analyses

All efficacy analyses will be performed on the BOTOX-treated population. The primary timepoint will be week 12 after treatment 1. Efficacy data will be presented by study visit in each BOTOX treatment cycle.

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

The following efficacy measures will be collected and summarized by BOTOX treatment cycle, BOTOX dose group, and all BOTOX group on the BOTOX-treated population:

- number of daytime urinary incontinence episodes
- number of daytime micturition episodes
- number of daytime urgency episodes
- presence or absence of nighttime urinary incontinence
- volume voided per micturition
- time to patient's request and time to patient's qualification for retreatment
- pediatric incontinence questionnaire (PinQ) score
- modified treatment benefit scale (TBS)

Change from study baseline analyses will be performed. Study baseline is defined as the data collected prior to the first treatment.

5.2 Primary Efficacy Analyses

For each of the BOTOX doses of 100 and 50 U, the null hypothesis to test is that there is no treatment difference between these dose groups and the 25 U BOTOX dose group in mean change from study baseline in daily average frequency of daytime urinary incontinence episodes at week 12 post-treatment 1. The alternative hypothesis is that there is a treatment difference in mean change from study baseline in daily average frequency of daytime urinary incontinence episodes at week 12 post-treatment 1.

5.2.1 Primary Analyses of Primary Efficacy Variables

The primary efficacy variable is the change from study baseline post treatment 1 in the daily average frequency of daytime urinary incontinence episodes (daytime is defined as the time between waking up to start the day and going to bed to sleep for the night).

Each daytime period recorded in the bladder diary will be normalized to represent a 12-hour period. The daily average frequency of daytime urinary incontinence episodes will be adjusted by the normalized daytime period.

For BOTOX treatment cycle 1, the treatment difference (least-squares [LS] mean difference) and 95% CI of the treatment difference in mean change from study baseline for 100 versus 25 U of BOTOX and 50 versus 25 U of BOTOX will be calculated. Pairwise comparisons of 100 versus 25 U of BOTOX and 50 versus 25 U of BOTOX up to week 12 posttreatment 1 will be evaluated using an ANCOVA model with study baseline value as covariate, and treatment group as factor. A hierarchical analysis strategy (Lubsen and Kirwan, 2002) to adjust for multiplicity will be used. In the order of (1) 100 versus 25 U of BOTOX, and (2) 50 versus 25 U of BOTOX, treatment group differences at week 12 will be tested at the 0.05 significance level in a fixed sequence fashion. Results of hypothesis testing for 50 versus 25 U of BOTOX will be considered for statistical significance only if the treatment difference for 100 versus 25 U of BOTOX is shown to be statistically significant. A similar analysis, without a hierarchal testing strategy, will be applied to other timepoints.

For each BOTOX treatment cycle, the mean raw value, mean change from study baseline, and 95% CIs for the arithmetic mean change from the study baseline will be summarized by study visit.

5.2.2 Other Analyses of Primary Efficacy Variables

As a sensitivity analysis of the LOCF imputation method, mixed-effect model repeated measures analysis (MMRM) (Mallinckrodt et al, 2001a and 2001b) will be conducted on the

primary efficacy variable with the observed data. The data will include all the observed cases up to and including the primary timepoint (week 12 after the first treatment). The unstructured variance-covariance structure will be used in the model. In contrast to LOCF that imputes missing values by carrying the last observed values forward, sensitivity analysis using MMRM allows for estimation of patient's missing response by using all the observed data and nonconstant correlations among the timepoints.

5.3 Secondary Efficacy Analyses

5.3.1 Proportion of Responders

Responder analyses will be performed using the daily average frequency of normalized daytime urinary incontinence episodes. The following definitions of responder will be used for daytime urinary incontinence at week 12 post-treatment 1.

1. A responder is a patient who has at least a 50% reduction from study baseline in daily average frequency of normalized daytime urinary incontinence episodes at week 12 post-treatment 1.

2. In addition, the proportions of patients with 100% reduction from study baseline in daily average frequency of normalized daytime urinary incontinence episodes at week 12 post-treatment 1 will be evaluated.

For week 12 posttreatment 1, the difference of the proportions of patients responding between the treatment groups (100 versus 25 U of BOTOX, and 50 versus 25 U of BOTOX) will be analyzed using the Cochran-Mantel-Haenszel (CMH) method controlling for baseline daytime urinary urgency incontinence episodes (a total of ≤ 6 episodes or > 6 episodes over the 2-day diary collection period).
For each treatment cycle, the proportion of patients with ≥50%, and 100% reduction from study baseline in normalized daytime urinary incontinence episodes will be presented by treatment group for each post-treatment visit.

5.4 Efficacy Analyses for the Secondary Efficacy Variables

No imputation will be done for the missing values of the secondary efficacy variables. Efficacy analyses for the secondary efficacy variables will also be conducted with normalization to a 12-hour daytime period unless specified otherwise. The hierarchical analysis strategy to adjust for multiplicity will not be implemented.

5.4.1 Number of Daytime Micturition Episodes

For each treatment cycle, the mean raw value, mean change from study baseline, and 95% CI for the mean change from study baseline in daily average frequency of daytime micturition episodes (arithmetic mean) over the 2-day diary period will be summarized by treatment group at post-treatment visits.

For BOTOX treatment cycle 1, the treatment difference (LS mean difference) and 95% CI of the treatment difference in mean change from study baseline for 100 versus 25 U of BOTOX and 50 versus 25 U of BOTOX will be calculated. Pairwise comparisons of 100 versus 25 U of BOTOX and 50 versus 25 U of BOTOX up to week 12 posttreatment 1 will be evaluated using an ANCOVA model with study baseline value as covariate, and treatment group and stratification (baseline daytime urinary urgency incontinence episodes [a total of ≤ 6 episodes or > 6 episodes over the 2-day diary collection period]) as factor. A hierarchical analysis strategy to adjust for multiplicity will not be implemented.

5.4.2 Number of Daytime Urgency Episodes

For each treatment cycle, the mean raw value, mean change from study baseline and 95% CI for the mean change from study baseline in daytime urgency episodes (arithmetic mean) over the 2-day diary period will be summarized by treatment group at post-treatment visits.

For BOTOX treatment cycle 1, the same ANCOVA model described in Section 5.4.1 will be used.

5.4.3 Volume (mL) per Micturition

Volume per micturition will be derived from the total volume voided over 1 daytime period during the 2-day bladder diary collection period divided by the number of voids in the same daytime period.

For each treatment cycle, the mean raw value, mean change from study baseline, and 95% CI for the mean change from study baseline in volume per micturition (arithmetic mean) over the 2-day diary period will be summarized by treatment group at post-treatment visits.

For BOTOX treatment cycle 1, the same ANCOVA model described in Section 5.4.1 will be used.

5.4.4 Night Time Urinary Incontinence Episodes

For each treatment cycle, a summary table for the number and proportion of patients who experienced nighttime urinary incontinence on 0, 1, or 2 nights will be presented by treatment group for study baseline and each post-treatment visit.

5.4.5 PinQ Total Score

The PinQ is a 20-item questionnaire completed prior to treatment at day1 weeks 6 and 12 post-treatment and then at the qualification for retreatment visit(s) if the patient qualifies.

For each treatment cycle, the mean raw value, mean change from study baseline, and 95% CI for the mean change from study baseline in the PinQ total score (arithmetic mean) will be summarized by treatment group for each post-treatment visit.

For BOTOX treatment cycle 1, the same ANCOVA model described in Section 5.4.1 will be used.

The 3 prespecified item scores (worry about smell [question 3], feel bad about myself [question 10], and being with friends [question 20]), these will be analyzed in the same manner as described above for the PinQ total score.

5.4.6 Modified TBS

The TBS is single-item scale completed at weeks 6 & 12 posttreatment, at the qualification for retreatment visit(s) if the patient qualifies and at the study exit visit.

For each treatment cycle, the numbers and proportions of patients with a positive treatment response in the modified TBS, defined as either "greatly improved" or "improved", will be presented and the 95% CI of the proportion of patients with a positive treatment response will also be provided by treatment group for each post-treatment visit.

For week 12 post-treatment 1, the difference of the proportions of patients responding between the treatment groups (100 versus 25 U of BOTOX, and 50 versus 25 U of BOTOX) will be analyzed using the Cochran-Mantel-Haenszel (CMH) method controlling for baseline daytime urinary urgency incontinence episodes (a total of ≤ 6 episodes or > 6 episodes over the 2-day diary collection period).

5.5 Other Efficacy Analyses

5.5.1 Duration of Effect

Duration of treatment effect will be evaluated by BOTOX treatment group for each treatment cycle in 2 different ways:

- Time between a treatment and patient's first request for subsequent treatment
- Time between a treatment and qualification for subsequent treatment

Patients who did not request retreatment will be treated as censored at the time of their last study visit or study exit.

The time from the day of each BOTOX treatment to the request for the subsequent treatment will be estimated using a Kaplan-Meier survival method for each treatment group for each treatment cycle. The median duration in days (and in weeks) will be presented.

Analysis for time to qualification will be performed and presented in the same manner as described for time to request for retreatment.

5.6 Subgroup Analyses for Primary Efficacy Variables

For each treatment cycle, the mean raw value, mean change from study baseline, and 95% CIs for the daily average frequency of normalized daytime urinary incontinence episodes (arithmetic mean) over the 2-day diary period will be presented by treatment group at study visit by

- baseline daytime urinary urgency incontinence episodes (a total of ≤ 6 versus > 6)
- race (Caucasian or non-Caucasian)
- sex (male or female)
- Region (North America versus Other Regions). North America = U.S.A and Canada; Other Regions = Australia, Belgium, Czech Republic, Great Britain, Italy, and Poland.

6. Safety Analyses

All analyses will be performed on the BOTOX-treated population. Adverse events and all other safety analyses will be presented by BOTOX treatment cycle in this population and grouped according to the actual dose received at that treatment cycle (25, 50, or 100 U of BOTOX as well as an 'All BOTOX' group).

6.1 Study Treatment – Exposure and Administration

6.1.1 Exposure to Study Treatments

The cumulative duration of treatment exposure will be summarized for the overall study and by each BOTOX treatment cycle. For each patient, the total duration of treatment exposure is defined as the number of days from the day of the first BOTOX treatment to the exit day or the day of data cut, as applicable. For each patient, the number of follow-up days for a given BOTOX treatment cycle is defined as the number of days from the day of treatment to the day prior to the subsequent treatment or exit day or the data cut day, as applicable.

Data will be presented in the following intervals: weeks ≥ 2 weeks, ≥ 6 weeks, ≥12 weeks, ≥18 weeks, ≥24 weeks, ≥30 weeks, ≥36 weeks, ≥42 weeks, ≥48 weeks, ≥ 54 weeks, ≥ 60 weeks, ≥ 66 weeks, ≥ 72 weeks, ≥ 78 weeks, ≥84 weeks, ≥90 weeks and ≥ 96 weeks. Duration of exposure is based on calendar weeks and descriptive summary statistics of the treatment duration in weeks will be presented.

6.1.2 Administration of Study Treatments

For the BOTOX injection, a patient listing will be produced. The listing will present the time and date of injection and whether the patient was injected per protocol. If the patient was not injected per protocol, the number of injection sites, location of injection sites and volume injected will also be included.

6.2 Adverse Events

AEs will be coded from the verbatim text into preferred term and the primary SOC by using the MedDRA dictionary. Events will be summarized by BOTOX treatment cycle.

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

For a given treatment cycle, a TEAE is an AE with onset after the initiation of cycle treatment or an AE with onset prior to cycle treatment that worsened in severity or became serious after the initiation of cycle treatment.

An AE that stops during treatment cycle 1 and recurs in treatment cycle 2 is considered as a TEAE in treatment cycle 2 if it fulfills the TEAE definition above with respect to initiation of cycle 2 treatment.

The incidence of TEAEs will be summarized and presented for all events and will be tabulated by treatment group as follows:

i. By relationship to treatment. Classified for all TEAEs and treatment-related TEAEs.
ii. By descending order of frequency.
iii. By primary SOC and preferred term. At each level of summarization (overall, primary SOC, and preferred term) a patient will be counted once if he/she reports 1 or more experiences at that level.
iv. By primary SOC, preferred term and severity. For a given adverse event of a patient, if more than 1 severity grade was reported, the worst severity grade will be included in the tabulation. Statistical hypothesis tests will not be performed for this summary.

All treatment-related TEAEs (study drug-related, and injection procedure-related adverse events) will be presented. In addition, study drug-related and injection procedure-related AEs will be presented separately.

Treatment-related TEAEs will be analyzed by the same procedure outlined above (ii, iii, and iv). In addition, the number and percent of patients discontinued due to treatment-related TEAEs will also be tabulated by treatment group.

All the summaries discussed above will be based on all AEs observed during the entire treatment cycle. Since the design of this study allows patients to request/receive retreatment from week 12 (day 84) onward, an additional summary showing the number and percentage of patients with TEAEs by preferred term and treatment group in descending order of incidence will be presented for the TEAE observed prior to or on day 84 after each injection cycle. This will allow for a comparison across dose groups for the same duration of time (ie, since patients can request retreatment from week 12 onwards, patients remain in a cycle for varying time periods). Therefore this analysis provides a more direct comparison across dose groups.

In addition, the number and percentage of patients who have a treatment-emergent UTI during the first 2 weeks (14 days), the first 12 weeks (84 days), and during the overall treatment cycle will be presented by the treatment received in each treatment cycle.

A patient listing will be generated for TEAEs, sorted by primary SOC, preferred term, relationship, and severity. In addition, the number and percent of patients discontinued due to TEAEs will also be tabulated by BOTOX treatment group for each BOTOX treatment cycle. A listing will also be generated for patients who discontinued due to TEAEs.

6.2.1 Potential Distant Spread of Toxin Adverse Events (Applicable to BOTOX Studies Only)

To assess possible distant spread of toxin (PDSOT), 40 MedDRA 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
- Bell's palsy
- 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*

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

*: 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 patient. In accordance with this, since BOTOX is injected into the urinary bladder for this indication, and since urinary retention is considered an expected localized effect, the preferred term 'urinary retention' will not be considered a PDSOT event.

6.3 Serious Adverse Events

A treatment emergent SAE is an TEAE as defined above that further meets one or more SAE criteria at any time during the study treatment period.

The number and percent of patients with serious adverse events will be tabulated by treatment group for all serious adverse events and for treatment-related serious adverse events by SOC for each BOTOX treatment cycle. A listing will be generated of patients with serious adverse events.

6.4 Clinical Laboratory Evaluations

The clinical laboratory data (hematology, blood chemistry, and urinalysis) will be summarized by BOTOX treatment group for each BOTOX treatment cycle. Descriptive statistics will be generated for baseline values and change from study baseline by visit. Study baseline is defined as the last assessment prior to treatment 1.

Clinical laboratory test values will be considered potentially clinically significant (PCS) if they meet either the lower-limit or higher-limit PCS criteria. PCS criteria are specified on Table 4. The number and percentage of patients who have PCS postbaseline clinical laboratory values will be tabulated for each cycle by treatment group. The percentages will be calculated relative to the number of patients with available non-PCS study baseline values and at least 1 postbaseline assessment during the treatment cycle of interest. The numerator will be the total number of patients with available non-PCS baseline values and at least 1 PCS postbaseline value during the treatment cycle of interest. A supportive listing of patients with PCS postbaseline values will be provided, including the patient number, baseline and all postbaseline (including non-PCS) values.

Table 4 Clinical Laboratory PCS Criteria

Laboratory Group	Parameter	SI Unit	PCS Low Limit	PCS High Limit
Hematology	Basophils, absolute cell count	10^9/L	—	> 2.0 × ULN
	Eosinophils absolute cell count	10^9/L	—	> 2 × ULN
	Hematocrit	%	< 0.9 × LLN	> 1.1 × ULN
	Hemoglobin	g/L	< 0.9 × LLN	> 1.1 × ULN
	Lymphocytes absolute cell count	10^9/L	< 0.7 × LLN	> 1.3 × ULN
	Monocytes, absolute cell count	10^9/L	< 0.5 × LLN	> 2.0 × ULN
	Neutrophils absolute cell count	10^9/L	< 0.7 × LLN	> 1.5 × ULN
	Platelet count (thrombocytes)	10^9/L	< 0.5 × LLN	> 1.5 × ULN
	Red blood cell count	10^{12}/L	< 0.9 × LLN	> 1.1 × ULN
	White blood cell count	10^9/L	< 0.9 × LLN	> 1.5 × ULN
Chemistry	Alanine aminotransferase	U/L	—	> 3 × ULN
	Alkaline phosphatase	U/L	—	> 1.2 × ULN: 5-11 years old > 3 x ULN: ≥ 12 years old
	Aspartate aminotransferase	U/L	—	> 3 × ULN
	Bicarbonate	mmol/L	< 0.9 × LLN	> 1.1 × ULN
	Bilirubin, total	µmol/L	—	> 1.5 × ULN
	Blood urea nitrogen	mmol/L	—	> 1.5 × ULN
	Calcium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
	Chloride	mmol/L	< 0.9 × LLN	> 1.1 × ULN
	Creatinine	µmol/L	—	> 1.5 × ULN
	Glucose, nonfasting	mmol/L	< 0.8 × LLN	> 1.5 × ULN
	Phosphorus	mmol/L	< 0.9 × LLN	> 1.1 × ULN
	Potassium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
	Magnesium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
	Protein, total	g/L	< 0.9 × LLN	> 1.1 × ULN
	Sodium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
	Uric acid (urate)	µmol/L	—	> 1.2 × ULN
Urinalysis	Glucose	µmol/L	—	Positive
	pH	—	< 0.9 × LLN	> 1.1 × ULN
	Protein	g/L	—	Positive
	Specific gravity	—	—	> 1.1 × ULN

LLN = lower limit of normal (value provided by the laboratory); SI = *Le Système International d'Unités* (International System of Units); ULN = upper limit of normal (value provided by the laboratory).

6.5 Vital Signs Measurements

Vital signs including pulse rate (beats/min), blood pressure (mm Hg), respiration rate (breaths per minute), and temperature (°C), will be summarized as descriptive statistics for baseline value and change from study baseline by BOTOX treatment group for each BOTOX treatment cycle.

Vital sign values will be considered potentially clinically significant (PCS) if they meet both the observed-value criteria and the change-from-baseline criteria. PCS criteria are specified on Table 5. The number and percentage of patients with PCS postbaseline values will be tabulated for each cycle by treatment group. The percentages will be calculated relative to the number of patients with available non-PCS study baseline values and at least 1 postbaseline assessment in the cycle of interest. The numerator will be the total number of patients with available non-PCS study baseline values and at least 1 PCS postbaseline value in the cycle of interest. A supportive listing of patients with PCS postbaseline values will be provided, including the patient number, baseline and all postbaseline (including non-PCS) values.

Table 5 Vital Sign PCS Criteria

Parameter	***Flag***	***Criteria***[a]	
		Observed Value	*Change From Baseline*
Sitting systolic blood pressure, mm Hg	High	≥ 155	Increase of ≥ 20
	Low	≤ 90	Decrease of ≥ 20
Sitting diastolic blood pressure, mm Hg	High	≥105	Increase of ≥ 15
	Low	≤ 45	Decrease of ≥ 15
Sitting pulse rate, bpm	High	≥ 120	Increase of ≥ 15
	Low	≤ 40	Decrease of ≥ 15
Respiratory rate (breaths per minute)	High	> 30	—
	Low	< 12	—
Temperature (°C)	High	> 38	—
	Low	< 36	—
Weight, kg	High	—	Increase of ≥ 5%
	Low	—	Decrease of ≥ 5%

a A postbaseline value is considered potentially clinically significant if it meets both the observed-value and the change-from-baseline criteria.

bpm = beats per minute.

6.6 Other Safety Analyses

6.6.1 Physical Examinations

Physical examinations will be performed at screening and study exit visit. Abnormal findings will be summarized by BOTOX treatment group. No statistical tests will be performed.

6.6.2 Pregnancy Test Results

Pregnancy test results for female patients will be provided in a data listing.

6.6.3 Post-void Residual (PVR) Urine Volume

Summary statistics for mean values at each visit, baseline and mean change from study baseline, and 95% CI for the mean change from baseline at each applicable visit will be tabulated by treatment group.

As per the protocol guidance, patients with elevated PVR or patients who initiated clean intermittent catheterization (CIC) may have further specific follow-up visits where PVR and the need for CIC are evaluated. Therefore, these patients may have multiple post-baseline PVR values recorded within 1 visit window, and in those cases, the highest PVR value will be presented in the summary table.

The proportion of patients with total PVR urine volume of ≤100, > 100 to < 200, ≥ 200 to < and ≥350mL will be summarized by treatment group and visit for each BOTOX treatment cycle.

6.6.4 Use of Clean Intermittent Catheterization Post-treatment

A listing of patients will be produced for patients using CIC containing all data from the eCRF information page.

6.6.5 Bladder and Kidney Ultrasound

A listing of all bladder and kidney ultrasound results will be provided for patients with at least 1 positive finding.

6.6.6 Positive Urine Culture Events

A listing of patients' urine culture results will be presented by treatment group.

6.6.7 Renal Function Testing

Renal function will be monitored from an estimate of the glomerular filtration rate (eGFR) determined from the serum creatinine levels obtained on the day of each treatment (prior to treatment), week 12 after each treatment, and at study exit. The calculation of eGFR will be based on eGFR = 0.41 × height (cm)/serum creatinine (mg/dL). For the derivation of estimated GFR, the height closest to the serum creatinine assessment will be used.

Summary statistics for mean values at each visit, study baseline and mean change from study baseline will be tabulated by treatment group. For study baseline eGFR, the latest value prior to the first treatment will be presented in the summary table.

6.6.8 Type of Anesthesia Used During Administration of The Study Drug

One or more of the following was used for each participant during administration of the study drug:

- Intraurethral local anesthetic gel
- Instillation of local anesthetic solution into the bladder
- Sedative
- General anesthesia

Within each treatment cycle, the number and percentage of participants taking each type of anesthesia will be presented by treatment group.

In addition, because participants can receive one or more of the above anesthesia options, participants will be summarized based on whether they received general anesthesia or no general anesthesia.

6.7 Subgroup Analyses for Safety Variables

The incidences of TEAEs occurring in treatment cycle 1 will be summarized by treatment group for each of the following subgroups: gender (male, female), race (White, Non-White), geographic region (North America vs. Other Regions), and type of anesthesia received (general anesthesia vs. no general anesthesia). The number and percentage of patients reporting TEAE in each treatment group will be tabulated by descending percentage of the preferred terms for each of the subgroups. Since the study design allows patients to request/receive retreatment from week 12 onward, only TEAEs observed prior to or on day 84 after the first study injection will be presented.

7. Pharmacokinetic Data Analyses

No pharmacokinetic samples will be collected.

8. Immunogenicity Data Analyses

Two methods will be used to analyze immunogenicity samples: the enzyme linked immunosorbent assay (ELISA) method to detect the toxin-binding antibodies and the mouse protection assay (MPA) to detect the toxin-neutralizing antibodies to BOTOX. In the ELISA step, samples will be analyzed using a toxin-binding assay to determine the presence of any toxin-binding antibodies. Only samples from patients who are positive for the presence of toxin-binding assay will be subsequently analyzed using the MPA to determine the presence of toxin-neutralizing antibodies. For the toxin-binding assay, immunogenicity data will be reported as negative, positive or inconclusive. The percent of patients who have a serum sample that is positive for binding antibodies, as assessed by the screening and immunodepletion steps, will be reported by BOTOX treatment group at baseline, week 12 and study exit. The titer values for all samples will also be summarized.

The toxin-neutralizing data will be reported as protected, not protected, or inconclusive for the mouse protection assay (MPA). Samples considered protected are those that contain neutralizing antibodies towards BOTOX and thus protect the mouse from the lethal effects of BOTOX. Samples considered not protected do not protect the mouse from the lethal effects of BOTOX. The final result will depend on the number of mice found dead or alive. If 3 out of 4 mice or more are found dead then the sample will be considered negative (not protected) for neutralizing antibodies; if the number of dead mice is less than or equal to 1, then the sample will be considered positive (protected) for neutralizing antibodies; if 2 mice are found dead, or if the assay results or sample integrity are in question, then the sample will be considered as inconclusive. Antibody test results will be summarized and tabulated by BOTOX treatment group and by visit. The percent of patients considered protected, not protected, or inconclusive for toxin-neutralizing assay will be reported.

A listing of antibody assay results will be provided for patients who had a positive toxin-neutralizing antibody result.

9. Health Outcomes Data Analyses

All analyses of health outcome variables are described in Sections 5.4.5 and 5.4.6.

10. Interim Analyses

No interim analysis is planned for this study.

11. Analysis for US FDA

Not applicable.

12. Data Collected But Not Analyzed

Not applicable.

13. Deviations From Protocol

There are no deviations from the protocol.

14. References

Lubsen J and Kirwan BA. Combined endpoints: can we use them? Statistics in Medicine. 2002;21(19):2959-2970.

Mallinckrodt CH, Clark WS, David SR. Type I error rates from mixed-effects model repeated measures versus fixed effects analysis of variance with missing values imputed via last observation carried forward. Drug Information J. 2001a; 35:1215-1225.

Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using mixed-effects models. J Biopharmaceutical Statistics. 2001b;11: 9-21.

15. Amendments

Following is a summary of content-oriented changes that were made to each section of the SAP, and a brief rationale for these changes. Minor editorial and document formatting revisions have not been summarized.

Sections	Revision	Rationale
Section 1. Introduction	Revised to indicate that there will be no primary analysis, that only one analysis will be performed in this study	FDA agreed with Sponsor's request to stop further recruitment in this study, hence there was no longer a need for a primary analysis
Section 2.1: Analysis Populations	Deleted Per-protocol population. All analyses on per-protocol population were also deleted in other sections (e.g., Section 5.2.2)	To reflect current company practices on use of per-protocol population analyses
Section 2.1: Analysis Populations	Additional text included to clarify how participants are assigned to actual BOTOX treatment group based on actual BOTOX dose unit received	For added clarity
Section 2.1.1: Handling Mis-stratification	This section was added to discuss how mis-stratified patients will be handled in analysis	For added clarity
Section 2.2: Visit Windows	This section was modified by removing telephone visits	To reflect removal of telephone visits since minimal data is collected, and no data for which primary/secondary endpoints analyzed
Section 2.3: Data Conventions	Modified imputation rules for completely missing or partial dates for adverse events and prior/concomitant medications	To reflect current programming standard conventions
Sections 4.1 and 4.2: Demographics and Disease Characteristics	Deleted Summary of Pretreatment Adverse Events	To reflect current company standards
Section 5.2.1: Primary Analyses of Primary Efficacy Variables	Removed last paragraph describing how treatment by investigator interaction will be assessed	Due to the large number of sites and small number of patients per site, considered not feasible to assess treatment by investigator interaction
Section 5.4.6: Modified TBS	Added CMH test	To present between group-comparisons
Section 5.5: Dose response using AUC of primary efficacy measure	Section deleted	No longer considered an analysis of interest
Section 5.6: Subgroup Analyses for Primary Efficacy Variables	Replaced "investigator site" with "Region" as a subgroup	Due to the high number of sites and low sample size per site. Added "Region" as extrinsic factor.

Sections	Revision	Rationale
Section 6.1.1 Exposure to Study Treatments	Slightly modified the intervals for which duration of follow-up is summarized	Same information is summarized in an easier to interpret manner
Section 6.2 and 6.3: Pretreatments AEs and SAEs	Pretreatments AEs and SAEs are not summarized	Pretreatments AEs and SAEs are presented in Listings
Section 6.2: Adverse Events	Added summary of AEs observed up to Week 12 (Day 84)	Since patients remain in a treatment cycle for varying time periods, this provides a more direct comparison across dose groups.
Section 6.2.1: PDSOT adverse events	Paralysis flaccid was removed from the list of preferred terms. The preferred term “VII nerve paralysis” was replaced by “Facial paralysis”. Bell’s palsy was added to the list	Due to update in MedDRA version
Section 6.4. Clinical Laboratory Evaluations and Section 6.5. Vital Signs	Definition of potentially clinically significant (PCS) was included	To reflect current company standards