Statistical Analysis Plan for Study 2042-201-008

A Multicenter, Double-Blind, Randomized, Placebo-Controlled Parallel-Group Phase 2 Study Evaluating the Safety and Efficacy of OnabotulinumtoxinA X for the Treatment of Moderate to Severe Glabellar Lines

Date: 12 December 2022

Version 2.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for OnabotA X Study 2042-201-008 Glabellar Lines: OnabotulinumtoxinA X in the Treatment of Moderate to Severe Glabellar Lines.

Study 2042-201-008 examines the safety and efficacy of OnabotA X in subjects with moderate to severe glabellar lines.

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

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

2.0 Study Design and Objectives

2.1 Objectives, Hypotheses and Estimands

The primary objectives of the study are to compare the efficacy of OnabotA X and placebo and to evaluate the safety of OnabotA X in subjects with moderate to severe GL.

The null and alternate hypotheses are the following:

H_0: $P_T = P_P$ (The treatment response probability between each OnabotA X treatment group and the placebo group is equal)

H_A: $P_T \neq P_P$ (The treatment response probability between each OnabotA X treatment group and the placebo group is not equal)

Estimand attributes of the primary efficacy endpoint are detailed in Table 1 (see Section 9.3.2). In addition, the attribute of treatment is a single dose of active treatment or placebo.

Differences in proportion of subjects who would achieve ≥ 1-grade improvement from baseline on the investigator-rated AGLSS at maximum contraction at Day 30 had they not

missed AGLSS assessments, for each active treatment group (OnabotA X [redacted]) in comparison with placebo in the ITT population.

Any missing assessments will be assumed to be missing at random and imputed using multiple imputation. Statistical significance of the difference will be tested using Cochran-Mantel-Haenszel (CMH) test stratified by baseline investigator-rated AGLSS at maximum contraction.

[redacted]

The clinical hypothesis is that OnabotA X has an acceptable safety profile when administered to the corrugator and procerus muscles in subjects with moderate to severe GL.

No estimand is defined for the safety evaluations.

2.2 Study Design Overview

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic

2.3 Treatment Assignment and Blinding

Subjects will be randomized in a 2:2:2:1 ratio to receive OnabotA X or placebo. Randomization will be stratified at each investigator site by baseline investigator-rated AGLSS at maximum contraction, using central by block stratified randomization. The same block will not be shared across investigative sites or by baseline GL severity assessment. AbbVie will provide instructions for preparations and treatment administration separately, in the pharmacy manual.

2.4 Sample Size Determination

The sample size was chosen empirically.

[redacted]

3.0 Endpoints

3.1 Primary Endpoint(s)

The primary efficacy endpoint is achievement of ≥ 1-grade improvement from baseline on the Allergan Glabellar Line Severity Scale (AGLSS) according to investigator assessments at maximum contraction at Day 30.

[redacted]

3.2 Secondary Endpoint(s)

- Achievement of None or Mild on the investigator-rated AGLSS assessments at maximum contraction at Day 30;

[redacted]

3.3 Other Efficacy Endpoint(s)

[redacted]

[redacted]

[redacted]

3.4 Safety Endpoint(s)

The safety endpoints for this study include:

- Adverse Events
- Abbreviated Physical Examination (general appearance, head, ears, eyes, nose, throat, and neck)

- Clinical Laboratory Test [redacted]
- Vital Sign Measurements [redacted]
- Electrocardiogram [redacted]
- Neurologic Assessment

3.5 Additional Endpoint(s)

[redacted]

4.0 Analysis Populations

The following population sets will be used for the analyses.

The Intent to Treat (ITT) Population includes all randomized subjects. The ITT Population will be used for all efficacy analyses. [redacted]

The Safety Analysis Set consists of all subjects who received at least 1 dose of study drug. The safety analyses will be based on the safety population. [redacted]

5.0 Subject Disposition

The total number of subjects who were screened, enrolled, and treated will be summarized. Reasons for exclusion, including screen failure, will be summarized.

[redacted]

For end of study participation, the number and percentage of subjects who completed the protocol defined follow-up period (or did not with associated reasons) will be summarized overall and by treatment group.

6.0 Study Drug Duration and Compliance

Duration of exposure will be summarized with descriptive statistics (mean and standard deviation, median, Q1, Q3, minimum, and maximum). The number of subjects exposed for specific period of time [redacted] will also be summarized. Duration of exposure will be calculated as date of study exit minus date of study drug administration +1.

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Demographics, baseline characteristics, medical history, prior and concomitant medications, and concurrent procedure will be summarized overall and by treatment group for the ITT Population. [redacted]

[redacted]

7.1 Demographics and Baseline Characteristics

Continuous demographic variables include age. Categorical demographic variables include sex, ethnicity, race, age group [redacted].

Baseline characteristics include weight, height, body mass index (BMI), Fitzpatrick skin type, investigator-rated and subject-rated AGLSS severity at maximum contraction and at rest, baseline Facial Line Satisfaction Questionnaire (FLSQ©) impact domain score, [redacted]

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). [redacted]

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. [redacted]

[redacted] The number and percentage of subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications.

7.4 Concurrent Procedure

All procedures undergone on or after Study Day 1 visit through the exit visit will be considered concurrent procedures. Concurrent procedure will be summarized by MedDRA high level term and preferred term. [redacted]

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

The primary endpoint and secondary endpoints will be analyzed based on the ITT population, [redacted]

• Subjects who are missing data for any other reason for the endpoint will count as though they hypothetically continued in the study.

9.0 Efficacy Analyses

9.1 General Considerations

All efficacy analyses will be conducted in the ITT Population in each period. [redacted]

[redacted] All efficacy results will be summarized by treatment group.

Unless otherwise specified, any subject who is randomized based on a wrong stratum will be analyzed according to the actual stratum the subject belongs to.

"Baseline" refers to the last non-missing observation before the first administration of study drug or randomization if no study drug is given.

9.2 Handling of Missing Data

Missing data will be imputed using multiple imputation (MI) method for the primary and secondary endpoints.

[redacted]

[redacted]

Each of the 5 imputation data sets will be analyzed individually, then combined to generate the final inferences as described in Section 9.3.2.

9.3 Primary Efficacy Endpoint(s) and Analyses

9.3.1 Primary Efficacy Endpoint(s)

The primary efficacy endpoint is achievement of ≥ 1-grade improvement from baseline on the investigator-based AGLSS at maximum contraction at Day 30.

9.3.2 Main Analysis of Primary Efficacy Endpoint(s)

The attributes of the estimands corresponding to the primary efficacy endpoint are summarized in Table 1.

Table 1. Summary of the Estimand Attributes of the Primary Efficacy Endpoint

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Hypothetical estimand for primary endpoint	OnabotA X or placebo	Achievement of ≥ 1-grade improvement from baseline on the AGLSS according to investigator assessments at maximum contraction at Day 30	ITT (All randomized)	Subjects who discontinue study prior to Day 30 assessments, or who do not have AGLSS assessments will be included in the analysis as a hypothetical scenario in which they had not missed the AGLSS assessments	Difference in response proportions between each active treatment group and placebo, after MI using CMH test stratified by baseline investigator-rated AGLSS at maximum contraction

AGLSS = Allergan Glabellar Line Severity Scale; CMH = Cochran-Mantel-Haenszel; ITT = intent-to-treat; MI = multiple imputation; OnabotA X = onabotulinumtoxinA X

The evaluation of the equality of the proportions of responders will be based on Cochran-Mantel-Haenszel (CMH) test stratified by baseline investigator-rated AGLSS at maximum contraction. Wald confidence intervals for proportions of responders and difference in the proportion of responders will be presented. The Breslow-Day homogeneity of the odds-

ratio test will be performed to test the treatment-by-investigator-rated baseline GL severity at maximum contraction interaction.

To obtain pooled CMH p-value, the Wilson-Hilferty transformation will be used.[1,2]

[redacted]

9.3.3 Sensitivity and Supplementary Analyses of the Primary Efficacy Endpoint(s)

A sensitivity analysis will be performed for the primary efficacy variable using observed data.

9.4 Secondary Efficacy Endpoints and Analyses

9.4.1 Key Secondary Efficacy Endpoints

The key secondary endpoints, which utilize both the AGLSS and FLSQ, include:

- Achievement of None or Mild on the investigator-rated AGLSS assessments at maximum contraction at Day 30;
- Responses of Very satisfied or Mostly satisfied on satisfaction with treatment per the FLSQ follow-up version Item 5 at Day 60;
- Achievement of ≥ 14-point psychosocial impact improvement from baseline at Day 30 per the FLSQ Impact domain, among subjects with baseline scores ≥ 14 points

9.4.2 Main Analyses of Key Secondary Efficacy Endpoints

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

Table 2. Summary of the Estimand Attributes of the Secondary Efficacy Endpoints

	Attributes of the Estimand				
Estimand Label	**Treatment**	**Endpoint**	**Population**	**Handling of Intercurrent Events**	**Statistical Summary**
Hypothetical estimand for secondary endpoint	OnabotA X or placebo	Achievement of None or Mild on the investigator-rated AGLSS assessments at maximum contraction at Day 30	ITT (All randomized)	Subjects who discontinue study prior to Day 30 assessments, or who do not have AGLSS assessments will be included in the analysis as a hypothetical scenario in which they had not missed the AGLSS assessments	Difference in response proportions between each active treatment group and placebo, after MI using CMH test stratified by baseline investigator-rated AGLSS at maximum contraction
		Responses of Very satisfied or Mostly satisfied on satisfaction with treatment per the FLSQ follow-up version Item 5 at Day 60		Subjects who discontinue study prior to Day 60 assessments, or who do not have FLSQ assessments will be included in the analysis as a hypothetical scenario in which they had not missed the FLSQ assessments	
		Achievement of ≥ 14-point psychosocial impact improvement from baseline at Day 30 per the FLSQ Impact domain, among subjects with baseline scores ≥ 14 points		Subjects who discontinue study prior to Day 30 assessments, or who do not have FLSQ assessments will be included in the analysis as a hypothetical scenario in which they had not missed the FLSQ assessments	

AGLSS = Allergan Glabellar Line Severity Scale; CMH = Cochran-Mantel-Haenszel; ITT = intent-to-treat; MI = multiple imputation; OnabotA X = onabotulinumtoxinA X

The key secondary efficacy endpoints are to be analyzed similarly to the primary endpoint using multiple imputation, as specified in Section 9.2. [redacted]

9.4.3 Sensitivity and Supplementary Analyses for Key Secondary Efficacy Endpoints

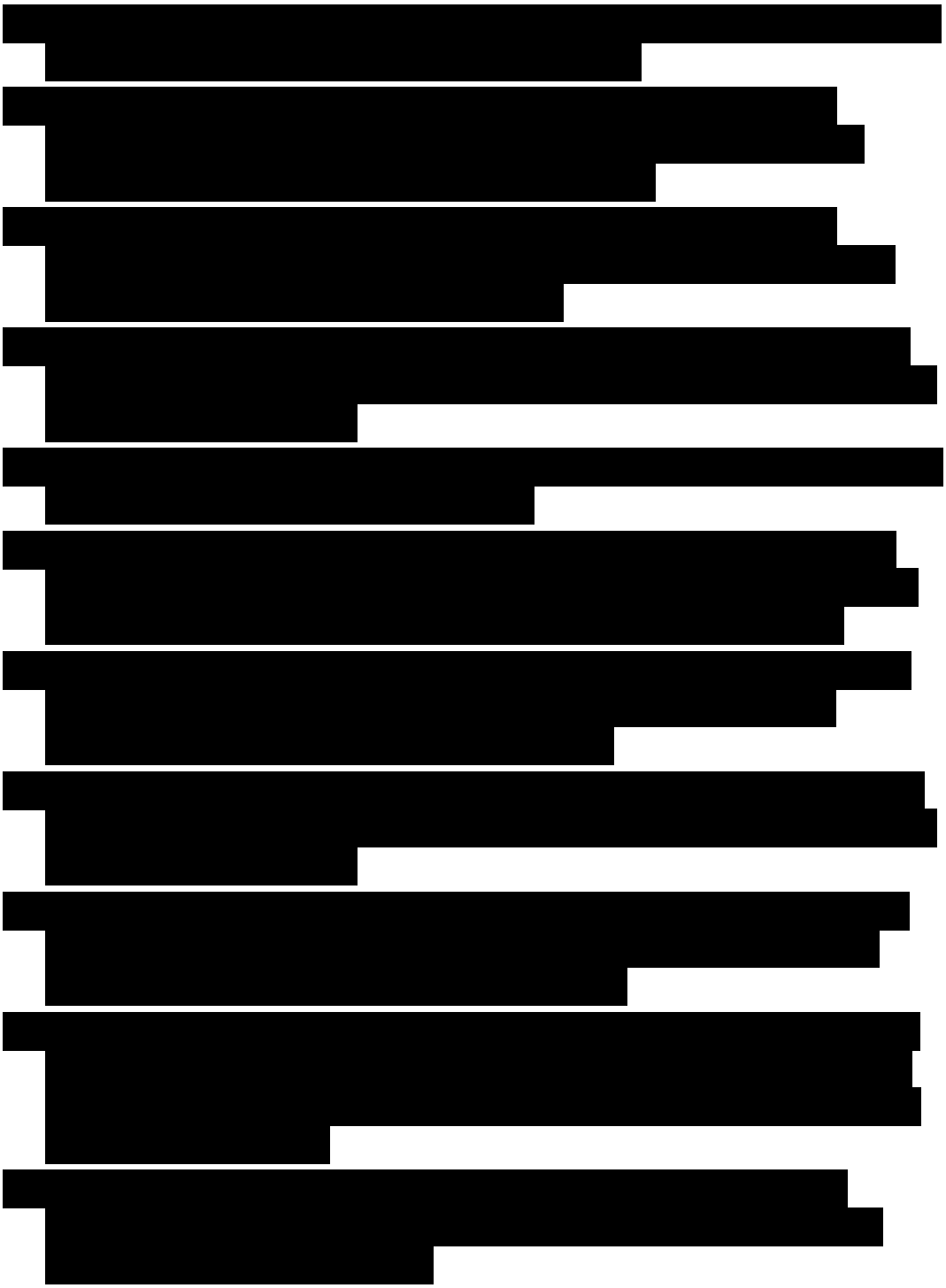
Not applicable.

9.4.4 Supportive Secondary Efficacy Endpoints and Analyses

Not applicable.

9.5 Additional Efficacy Analyses

[redacted]



[redacted]

9.6 Efficacy Subgroup Analyses

Not applicable.

10.0 Safety Analyses

10.1 General Considerations

Safety data will be summarized for the Safety Analysis Set. For the safety analyses, subjects are summarized based on the treatment actually received.

10.2 Adverse Events

Adverse events (AEs) will be summarized overall and by treatment group and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. [redacted]

10.2.1 Treatment-Emergent Adverse Events

[redacted] All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. [redacted]

[redacted]

10.2.2 Adverse Event Overview

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

- Any treatment-emergent AE
- Any treatment-emergent AE related to study treatment according to the investigator

[redacted]

- Any PDSOT TEAEs
- All deaths

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

Treatment-emergent AEs will be summarized by system organ class (SOC) and PT; by maximum severity and SOC and PT; and by subject number and SOC and PT. [redacted]

[redacted]

[redacted]

10.2.4 SAEs (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

SAEs (including deaths) will be summarized by SOC and PT and in listing format.

10.2.5 Potential Distant Spread of Toxin Adverse Events

To assess PDSOT, MedDRA preferred terms that may be associated with botulinum toxin effects have been identified (Appendix B). [redacted]

10.3 Analysis of Laboratory Data

[redacted] The clinical laboratory tests defined in the protocol operations manual (e.g., hematology, clinical chemistry, and urinalysis) will be summarized by treatment group.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

10.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, pulse rate, respiratory rate, and body temperature will be summarized by treatment group.

[redacted]

[redacted]

[redacted]

[redacted]

10.5 Safety Subgroup Analyses

Not applicable.

10.6 Other Safety Analyses

Electrocardiogram (ECG): [redacted]

[redacted]

[redacted]

Neurologic Assessment: [redacted]

11.0 Other Analyses

Immunogenicity: Blood samples for immunogenicity testing will be collected from all subjects [redacted]

Hypersensitivity: Blood samples to be collected within 2 hours of suspected anaphylaxis or systemic hypersensitivity reaction. [redacted]

12.0 Interim Analyses

An interim analysis is planned to occur [redacted] If a dose recommendation cannot be made at this point, then a second interim analysis will be performed [redacted] The interim analyses will be comparative and will be unblinded to a group of AbbVie personnel who are not directly involved in ongoing day-to-day study conduct. [redacted] Further details are provided in DMC Charter.

The primary and secondary efficacy analyses will be performed for each interim analysis, as well as summaries of all safety variables.

12.1 Data Monitoring Committee

An independent DMC will be instituted to review interim safety and efficacy data to provide a dose recommendation for Phase 3. A DMC Charter will be finalized prior to study start. [redacted]

13.0 Overall Type-I Error Control

Analyses will be conducted using a gated hierarchical testing procedure to preserve a familywise Type I error rate of α=0.05 for each OnabotA X group compared to placebo, separately. [redacted]

14.0 Version History

Table 3. SAP Version History Summary

Version	Date of Approval	Summary
1.0	19 August 2021	Initial version
2.0		Section 2.1, added details for estimand Section 4.0, changed the safety analysis set to be analyzed per randomization due to the difficulty of determine the actual treatment Section 8.0, added details for handling potential intercurrent events Section 9.3.2 Table 1, updated column 'Handling of Intercurrent Events' Section 9.4.2 added Table 2, the Summary of the Estimand Attributes of the Secondary Efficacy Endpoints Updated Table B-1 MedDRA preferred terms evaluated for PDSOT, removed the term "Extraocular muscles paresis", added the term "Opthalmoplegia" Updated Table C-1 Criteria for Potentially Clinically Important Clinical Laboratory Values: updated the value corresponding to Cholesterol, Glucose, and Triglycerides. Changed "Inorganic phosphorus" to "Phosphate"

15.0 References

1. EB W, MM H. The distribution of chi-square. Proc Natl Acad Sci U S A. 1931;17(12):684-8.
2. O'Kelly M, Ratitch B. Clinical trials with missing data: a guide for practitioners. First ed: John Wiley & Sons, Ltd; 2014.

Appendix A. Protocol Deviations

The number and percentage of unique subjects reporting significant protocol deviations will be summarized in total and by treatment group for all randomized or treated subjects. The number and percentage of unique subjects with protocol deviation categories below will also be summarized in total and by treatment group.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject received wrong treatment or incorrect dose of study.
- Subject took prohibited concomitant medication or concurrent procedure.

Appendix B. Definition of Possible Distant Spread of Toxin

Table B-1. MedDRA Preferred Terms Evaluated for PDSOT

Accommodation disorder	Hyporeflexia
Aspiration	Hypotonia
Bell's Palsy	Ileus paralytic
Botulism	Muscular weakness
Bradycardia	Opthalmoplegia
Bulbar palsy	Paralysis
Constipation	Paresis cranial nerve
Cranial nerve palsies multiple	Pelvic floor muscle weakness
Cranial nerve paralysis	Peripheral nerve palsy
Diaphragmatic paralysis	Peripheral paralysis
Diplopia	Pneumonia aspiration
Dry mouth	Pupillary reflex impaired
Dysarthria	Respiratory arrest
Dysphagia	Respiratory depression
Dysphonia	Respiratory failure
Dyspnoea	Speech disorder
Eyelid function disorder	Urinary retention
Eyelid ptosis	Vision blurred
Facial paralysis	Vocal cord paralysis
Facial paresis	Vocal cord paresis

Note: Table is based on MedDRA 25.0; the actual list used for analysis will be based on the MedDRA version in use at the time of database lock.

Appendix C. Potentially Clinically Important Criteria for Safety Endpoints