STATISTICAL ANALYSIS PLAN

Protocol Title: A Phase 2b Double-Blind, Placebo-Controlled, Dose-Ranging Study to Evaluate the Safety and Efficacy of AGN-151586 in Participants With Moderate to Severe Glabellar Lines

Protocol Number: 2034-201-008

Compound Number: AGN-151586 (Botulinum Neurotoxin Serotype E, BoNT/E) for injection

Short Title: AGN-151586 Dose-Ranging Study for Treatment of Glabellar Lines

Sponsor Name: Allergan Sales LLC

Legal Registered Address: Allergan Sales LLC
5 Giralda Farms
Madison, NJ 07940
USA

Regulatory Agency Identifier Number(s) : IND Number 130880

Author: [redacted]

Version 1.0: **15 NOVEMBER 2019**

Amendment 1: **13 OCTOBER 2020**

Table of Contents

List of Tables

SAP Version History

This Statistical Analysis Plan (SAP) Amendment 1 for study 2034-201-008 is based on the Protocol Amendment 1 (dated 05 June 2020).

In addition to minor typographical corrections, other updates from the previous SAP version 1.0 (dated 15 November 2019) include the following:

- Section 1.1: Objectives and Endpoints table added based on objectives and endpoints stated in previous SAP. Estimand framework table updated [redacted].
- Section 5.1: SAS version and MedDRA version updated to current versions. Added information about COVID-19-related reasons for modified visits to be reported as a listing.
- Section 5.3.2: Clarification provided for 95% confidence intervals that will be computed using the normal approximation (ie, Wald). In addition, p-values will be based on the stratified CMH tests, but the rate differences are not based on stratification.
- [redacted]
- Section 5.5: Added that study period (prior to, during, after study intervention administration) will be presented for safety data listings.
- Section 5.5.2: Added a breakdown for treatment-related AEs to consist of study drug-related AEs and study procedure-related AEs. In addition, PDSOTs will be summarized as part of the overall AE table. Removed from summary the TESAEs leading to study discontinuation. Neurologic assessments will be summarized using descriptive statistics, and a listing will be provided. Treatment-emergent neurologic assessment-related AEs will also be presented by SOC and PT.
- Section 5.6.1: Updated immunogenicity analysis planned summary and stated that the results would be summarized if ready by time of the immunogenicity database lock.
- Sections 5.7 & 5.8: Stated that interim analyses are not applicable to the study, and the DMC analyses were moved to a separate section.
- Section 6.2: Added that a change to the protocol-specified analyses would be the summarization of TEAEs, not just AEs, per the primary safety endpoint.

1. Introduction

This statistical analysis plan (SAP) provides a more technical and detailed elaboration of the statistical analyses of the efficacy and safety data as outlined and specified in study 2034-201-008 Protocol Amendment 1 (Approved version dated 05 June 2020). Specifications of tables, figures, and data listings are contained in a separate document.

1.1. Objectives and Endpoints

Objectives	Endpoints
Primary	
• Efficacy: To compare the efficacy between placebo and a range of doses of AGN-151586 for the treatment of GL in participants with moderate to severe GL	• Achievement of a ≥ 2-grade improvement from baseline on the FWS according to investigator assessments of GL severity at maximum frown at any postintervention timepoint through Day 7
• Safety: To compare the safety between placebo and a range of doses of AGN-151586 for the treatment of GL in participants with moderate to severe GL	• Incidence of AEs; change from baseline in hematology/chemistry/urinalysis laboratory, vital sign, and ECG parameters; and presence of antidrug antibodies
Secondary	
• To compare the efficacy between placebo and a range of doses of AGN-151586 for the treatment of GL in participants with moderate to severe GL, as measured by the FWS (participant-assessed, investigator-assessed)	• Achievement of a ≥ 2-grade improvement from baseline on the FWS according to participant assessments of GL severity at maximum frown at any postintervention timepoint through Day 7 • Achievement of a composite ≥2-grade improvement from baseline on the FWS at maximum frown as rated concurrently by the investigator and participant by timepoint • Achievement of a ≥2-grade improvement from baseline on the FWS at maximum frown as rated by the investigator by timepoint • Achievement of a ≥1-grade improvement from baseline on the FWS at maximum frown as rated by the investigator by timepoint • Achievement of a ≥2-grade improvement from baseline on the FWS at maximum frown as rated by the participant by timepoint • Achievement of a ≥1-grade improvement from baseline on the FWS at maximum frown as rated by the participant by timepoint • None or mild rating in FWS at maximum frown as rated by investigator by timepoint • None or mild rating in FWS at maximum frown as rated by participant by timepoint • Responders (defined by primary endpoint) achieving none or mild FWS at maximum frown as rated by the investigator by timepoint

Exploratory	
[redacted]	[redacted]

Estimand Framework:

Objective Clinical Category	Statistical Category	Estimand/Variable
Primary Objectives: • Efficacy: To compare the efficacy between placebo and a range of doses of AGN-151586 for the treatment of glabellar lines (GL) in participants with moderate to severe GL • Safety: To compare the safety between placebo and a range of doses of AGN-151586 for the treatment of GL in participants with moderate to severe GL		
Effectiveness[1]	Primary	• Variable: Achievement of a ≥ 2 grade improvement from baseline on the Facial Wrinkle Scale (FWS) With Photonumeric Guide according to investigator assessments of GL severity at maximum frown at any postintervention timepoint through Day 7 • Population: The target population is the randomized population that has met the inclusion and exclusion criteria; the analysis population is the modified intention-to-treat (mITT) population (randomized participants who have received study intervention and have at least one postintervention measurement of investigator-rated FWS at maximum frown) • IES: Participants with no valid assessments at any postintervention timepoint through Day 7 will be excluded from the analysis • PLS: Response rates between each dose group of AGN-151586 and placebo • Analysis: Cochran-Mantel-Haenzsel (CMH) test stratified by baseline GL severity at maximum frown (investigator assessment), using observed data without imputation or adjustment for multiplicity. Between group comparisons will only be performed for each cohort's AGN-151586 group against the cohort's placebo group. Nominal p-values will be presented for each statistical test.
	Secondary	• Variables: Achievement of a none or mild grade on the FWS according to investigator assessment of GL severity at maximum frown at any postintervention timepoint through Day 7 • Population: The target population is the randomized population that has met the inclusion and exclusion criteria; the analysis population is the mITT population • IES: Participants with no valid assessments at any postintervention timepoint through Day 7 will be excluded from the analyses • PLS: Response rates between each cohort dose group of AGN-151586 and placebo • Analysis: CMH test stratified by baseline GL severity at maximum frown using observed data without imputation or adjustment for multiplicity

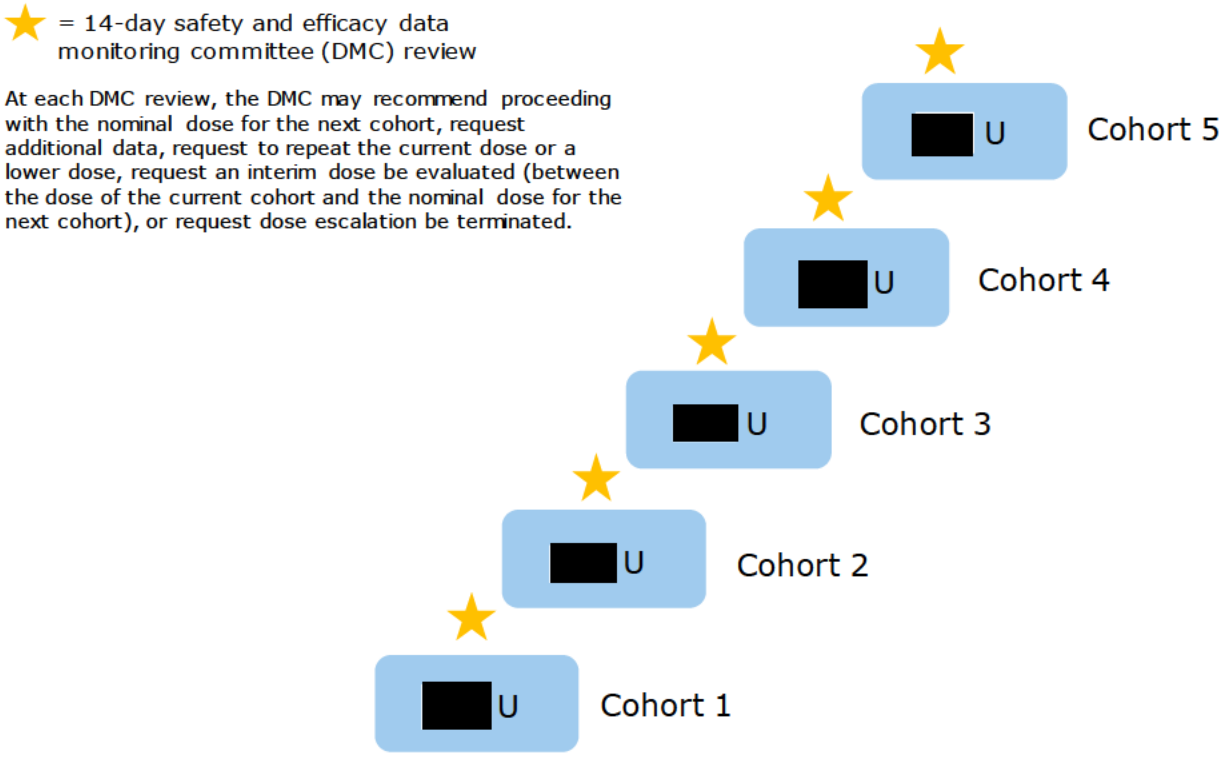
Objective Clinical Category	Statistical Category	Estimand/Variable
[redacted]	[redacted]	[redacted]
Safety	Primary	• Variable: Incidence of treatment-emergent adverse events (TEAEs); change from baseline in hematology/chemistry/urinalysis laboratory, vital signs, and electrocardiogram (ECG) parameters; and presence of antidrug antibodies • Population: Safety Population (participants who have received study intervention) • Analysis: Descriptive statistics

IES = Intercurrent event(s) strategy; PLS = Population-level summary.
[1] All estimand attributes explicitly identified for primary/secondary and select key estimands only. Other secondary and exploratory estimands are similar.

1.2. Study Design

- This is a multicenter, double-blind, randomized, placebo-controlled, dose-escalation study to evaluate the safety and efficacy of a single study intervention cycle of AGN-151586 in participants with moderate to severe GL.
- Participants are adults 18 to 65 years old, with moderate to severe GL at maximum frown (as assessed by the investigator and participant using the FWS).
- Up to 5 doses of AGN-151586 will be evaluated and compared with placebo in a sequential cohort design, starting with the lowest dose.
- Each cohort will enroll a balanced distribution of participants with moderate or severe GL severity at maximum frown per investigator assessment (at least 40% and no more than 60% of either).
- The primary efficacy measure is the investigator's assessments of GL at maximum frown using the FWS.
- Participants will undergo a Screening Visit up to 30 days before randomization on Day 1. All screening data must be available to the investigator prior to randomization on Day 1.
- On Day 1, participants will be randomly assigned in a 3:1 ratio to receive AGN-151586 or placebo. Enrollment stratification will be by baseline GL severity at maximum frown, assessed by the clinician (principal investigator or physician subinvestigator) using the FWS.
- Study intervention on Day 1 will be administered as 5 injections in the glabellar complex (1 in the procerus and 2 in each corrugator).
- Participants will remain in the clinic for the first 12 hours after administration of study intervention for serial assessments at Hours 4, 8 and 12. Thereafter, participants return to the clinic on Day 2 (Hour 24 and Hour 36), Day 3 (Hour 48 and Hour 60), Day 7, Day 14, Day 21, Day 28, Day 35, and Day 42.
- For each dose group, an internal data monitoring committee (DMC), independent from the study team, will review select safety and efficacy data from a subset of participants ($n \geq 16$) with data through 14 days postintervention (in addition to cumulative data from the prior cohort[s]).
- The DMC may recommend proceeding with the nominal dose for the next cohort, request additional data (eg, data from additional participants in the cohort, data from visits after Day 14), request to repeat the current dose or a lower dose, request an intermediate dose be evaluated (between the dose of the current cohort and the nominal dose for the next cohort), or request dose escalation be terminated.

The study scheme is as follows:



The Schedule of Visits and Procedures during the study are presented in Table 1–1.

Table 1–1 Schedule of Activities

Visit Number	1	2			3	4	5	6	7	8	9	10	11	12
Study Period	Screening (Days −30 to −1)	Randomization and Study Intervention (Day 1)[a]	Day 1[b]			Day 2		Day 3		Follow-up Day 7	Day 14	Day 21	Day 28	Day 35	End of Study (Day 42 or Early Exit)
Time after administration of study intervention	--	--	4 h	8 h	12 h	24 h	36 h	48 h	60 h						
Visit windows	--	--	± 1 h	± 1 h	± 1 h	± 2 h	± 2 h	± 2 h	± 2 h	± 1 day	± 1 day	± 1 day	± 1 day	± 1 day	± 1 day
Consent/ authorization	X														
Demographics	X														
Height and weight	X														
Inclusion/ exclusion criteria	X	X													
Medical/surgical history	X														
[redacted]		X		X		X		X		X	X	X	X	X	X
[redacted]		X													
[redacted]				X		X		X		X	X	X	X	X	X
[redacted]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
[redacted]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Standardized facial photography		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Physical examination[d]	X	X	X			X		X		X	X	X	X	X	X

Visit Number	**1**	**2**			**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Study Period	**Screening (Days −30 to −1)**	**Randomization and Study Intervention (Day 1)[a]**	**Day 1[b]**			**Day 2**		**Day 3**		**Follow-up Day 7**	**Day 14**	**Day 21**	**Day 28**	**Day 35**	**End of Study (Day 42 or Early Exit)**
Time after administration of study intervention	**--**	**--**	**4 h**	**8 h**	**12 h**	**24 h**	**36 h**	**48 h**	**60 h**						
Visit windows	**--**	**--**	**± 1 h**	**± 1 h**	**± 1 h**	**± 2 h**	**± 2 h**	**± 2 h**	**± 2 h**	**± 1 day**	**± 1 day**	**± 1 day**	**± 1 day**	**± 1 day**	**± 1 day**
Vital sign measurements[e]	X	X	X			X		X		X	X	X	X	X	X
Neurologic assessments[f]	X	X	X			X		X		X	X	X	X	X	X
Urine pregnancy test (WOCBP)[g]	X	X													X
12-lead ECG[h]	X					X		X		X					X
Collection of blood and urine samples for hematology, chemistry, and urinalysis laboratory testing[i]	X							X		X					X
Collection of blood samples for immunogenicity testing	X									X			X		X
Adverse events	X	X[j]	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concurrent procedures	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Visit Number	1	2			3	4	5	6	7	8	9	10	11	12
Study Period	Screening (Days −30 to −1)	Randomization and Study Intervention (Day 1)[a]	Day 1[b]			Follow-up: Day 2		Follow-up: Day 3		Day 7	Day 14	Day 21	Day 28	Day 35	End of Study (Day 42 or Early Exit)
Time after administration of study intervention	--	--	4 h	8 h	12 h	24 h	36 h	48 h	60 h						
Visit windows	--	--	± 1 h	± 1 h	± 1 h	± 2 h	± 2 h	± 2 h	± 2 h	± 1 day	± 1 day	± 1 day	± 1 day	± 1 day	± 1 day
Study intervention administration		X (0 h)													

ECG = electrocardiogram; [redacted] GL = glabellar lines; WOCBP = women of childbearing potential

[a] All makeup must be removed and hair not obscuring the upper face prior to completing any assessment. All assessments must be completed before administration of study intervention at Hour 0.

[b] Participants must remain at the investigational site on Day 1 (ie, from baseline procedures through Hour 12 assessment).

[c] The investigator and participant must grade the severity of GL using the FWS at the same time, first at rest, then at maximum frown. The investigator and participant must not discuss their results with each other.

[d] A complete physical examination is to be done at the Screening Visit and EOS/Early Exit Visit. An *abbreviated physical examination* is be done at other visits. The *abbreviated physical examination* is limited to an examination of general appearance, heart, lungs, and abdomen.

[e] Vital sign measurements includes pulse rate, respiration rate, blood pressure, and body temperature. Participants are to be seated for at least 2 minutes before measurements are collected.

[f] The neurologic assessment includes the Focused Symptom Questionnaire and Focused Neurologic Examination.

[g] Women of childbearing potential (WOCBP) must have a negative pregnancy test at the Screening Visit and on Day 1 prior to randomization. The investigator may ask the participant to perform a urine pregnancy test at any time, which must be conducted at the investigational site. At each visit, site personnel must discuss with WOCBP and males with partners who are WOCBP compliance with contraceptive use.

[h] A 12-lead ECG will be taken with the participant in a semirecumbent position for ≥ 10 minutes before starting the tracing.

[i] Nonfasting

[j] Participant must be observed for at least 30 minutes after study intervention for adverse events.

2. Statistical Hypotheses

The following hypotheses will be used to compare each dose group of AGN-151586 with placebo within a cohort:

- Null hypothesis: AGN-151586 and placebo are equally effective in the probability of reducing GL severity at maximum frown as measured by a ≥ 2-grade improvement from baseline based on the investigator-rated FWS at any postintervention timepoint through Day 7.
- Alternative hypothesis: AGN-151586 and placebo are not equally effective in the probability of reducing GL severity at maximum frown as measured by a ≥ 2-grade improvement from baseline based on the investigator-rated FWS at any postintervention timepoint through Day 7.

The null and alternative hypotheses are as follows:

$$H_0: P_{agn} = P_{pbo}$$

$$H_a: P_{agn} \neq P_{pbo}$$

where P_{agn} and P_{pbo} are proportions of the participants with reduction in GL severity at maximum frown as measured by a ≥ 2-grade improvement from baseline based on the investigator-rated FWS at any postintervention timepoint through Day 7 for AGN-151586 group and placebo group, respectively.

The CMH test with stratification of baseline investigator-rated GL severity at maximum frown will be used to test the above hypothesis.

3. Sample Size Determination

The primary efficacy analysis will be conducted on the mITT Population, which consists of randomized participants who have received study intervention and have at least one postintervention measurement of investigator-rated FWS at maximum frown. A sample size of 40 participants per cohort will have approximately 88% power to detect a difference between an AGN-151586 dose group and the placebo group, assuming a 5% dropout rate and responder rates of 70% for the AGN-151586 study intervention group and 10% for the placebo group. The calculation is based on a 2-sided Chi-squared test (Fisher's exact test) with a 0.05 significance level using nQuery Advisor version 7.0.

For each cohort subset data reviewed by the DMC, beginning with the lowest dose before progressing to higher doses, the study will have increasing power to detect AE incidence among the participants treated with AGN-151586. For Cohort 1 DMC review, for example, with a subset of 12 participants treated with AGN-151586, there will be approximately 72% chance to detect an AE at a 10% incidence. For subsequent cohorts with increasing doses, for example, with subsets of 21 participants treated with AGN-151586, there will be approximately 90% chance to detect at AE at 10% incidence. A table is included below to show the range of probabilities of detecting AEs with varying incidence rates.

TRUE EVENT RATE (%)	12 on Active Treatment	21 on Active Treatment	**30 on Active Treatment**	60 on Active Treatment
0.5%	0.06 (0, 0.19)	0.10 (0, 0.23)	0.14 (0.02, 0.26)	0.26 (0.15, 0.37)
1%	0.11 (0, 0.29)	0.19 (0.02, 0.36)	0.26 (0.10, 0.42)	0.45 (0.33, 0.58)
3%	0.31 (0.05, 0.57)	0.47 (0.26, 0.69)	0.60 (0.42, 0.77)	0.84 (0.75, 0.93)
5%	0.46 (0.18, 0.74)	0.66 (0.46, 0.86)	0.79 (0.64, 0.93)	0.95 (0.90, 1)
10%	0.72 (0.46, 0.97)	0.89 (0.76, 1)	0.96 (0.89, 1)	>0.99 (0.99, 1)
20%	0.93 (0.79, 1)	0.99 (0.95, 1)	>0.99 (0.99, 1)	>0.99 (0.99, 1)
30%	0.99 (0.92, 1)	>0.99 (0.99, 1)	>0.99 (0.99, 1)	>0.99 (0.99, 1)
50%	>0.99 (0.99, 1)	>0.99 (0.99, 1)	>0.99 (0.99, 1)	>0.99 (0.99, 1)

4. Populations for Analysis

The analysis populations will consist of participants as defined below:

Table 4–1 Analysis Populations

Population	Definition	Study Intervention
mITT Population	All randomized participants who receive study intervention and who have at least one postintervention investigator-rated FWS measurement at maximum frown	As treated
Safety Population	All participants who receive study intervention	As treated

5. Statistical Analyses

5.1. General Considerations

- The primary analyses will be performed after the database is locked and randomization schedule is released.
- The effectiveness endpoints will be analyzed based on the mITT Population. Safety analyses will be performed using the Safety Population. The analyses will be based on the actual study intervention that the participants receive.
- For each cohort, the effectiveness and safety endpoints will be analyzed/summarized by study intervention group, i.e., AGN-151586 group and placebo group, respectively. In addition, placebo groups across cohorts will be pooled to compare with each dose of AGN-151586 group as a sensitivity analysis.
- For applicable effectiveness and safety endpoints, baseline is defined as the last non-missing assessment prior to study intervention.
- Descriptive statistics for continuous variables include the sample size (N), mean, standard deviation, median, 1st and 3rd quartiles, minimum, and maximum.Summary statistics for categorical variables include the sample size (N), frequency count, and percentage.
- The primary analysis of the effectiveness endpoint, ie, the response rate of the participants with achievement of a $\geq$ 2 grade improvement from baseline on the FWS according to investigator assessments of GL severity at maximum frown at any postintervention timepoint through Day 7, will be analyzed using CMH tests stratified by baseline GL severity at maximum frown, to compare the between-group difference. The observed data without imputation will be used. Nominal p-values will be provided for each statistical testing. No multiplicity adjustment will be performed to control overall type I error.
- Summary of TEAEs, medical history, and concomitant medication will be summarized by study intervention group for each cohort. Also, the AGN-151586 groups across cohorts and the placebo groups across cohorts will be pooled.
- Partial dates will be treated as missing in computations, and will be listed in the data listings as they appear on the electronic case report form (eCRFs). No imputation of missing values will be performed, unless otherwise specified.
- The level of significance used for statistical testing of the effectiveness endpoints will be 0.05, 2-sided.

- The change from baseline values will be computed as the value for the postbaseline visit minus the baseline value.
- All statistical analysis will be performed using SAS version 9.4 or higher.
- Medical dictionary for regulatory activities (MedDRA) version 23.0 or higher will be used to code AEs and medical history.
- World Health Organization (WHO) Drug Dictionary will be used to code medications.
- Due to the COVID-19 pandemic, randomization into the study was curtailed for Cohort 3 only (38 of 40 participants randomized into Cohort 3), and visit assessments were modified for Cohorts 3 & 4 only. For Cohort 5, participants were expected to complete the visits as per the protocol; hence, no modifications were made to the visits for Cohort 5. These changes will be reported as a listing for modified visits due to COVID-19-related reasons.

5.2. Participant Disposition

The number of participants in the study populations (mITT and Safety) will be summarized by study intervention group; the number of participants screened will be summarized overall.

Summary of study disposition will be provided by study intervention group for the following:

- Number of randomized participants
- Number of participants who receive study intervention
- Number of participants who discontinue the study
- Reasons for discontinuation of the study

5.3. Primary Effectiveness Endpoint Analysis

5.3.1. Definition of Endpoint

The assessments of the severity of GL at maximum frown using the validated FWS, is based on the following scale:

- 0 = None
- 1 = Mild
- 2 = Moderate
- 3 = Severe

If a participant is determined to have at least a 2-grade improvement from baseline by the investigator-rated FWS at maximum frown at any postintervention assessment through Day 7, he/she will be considered as a responder. The primary effectiveness endpoint is the response rate of the participants with achievement of a ≥ 2-grade improvement from baseline on the FWS according to investigator assessments of GL severity at maximum frown at any postintervention timepoint through Day 7.

5.3.2. Main Analytical Approach

The proportion of responders for the primary endpoint will be analyzed based on mITT Population using CMH tests stratified by investigator assessment of baseline GL severity at maximum frown to compare each AGN-151586 study intervention group versus placebo for each cohort. In addition, placebo groups across cohorts will be pooled. Each AGN-151586 group will be compared against the pooled placebo groups as a sensitivity analysis. The observed data without imputation will be used.

The response rates of the participants with achievement of a ≥ 2-grade improvement from baseline on the FWS according to investigator assessments of GL severity at maximum frown at any postintervention timepoint through Day 7 will be calculated for each study intervention group. The corresponding 95% confidence intervals (CIs) using normal approximation (ie, Wald) will be provided.

The proportion difference, the corresponding Wald 95% CIs and nominal p-values will be provided for the comparison of response rates in the primary endpoint between different study intervention groups. P-values will be calculated based on the stratified CMH tests, but the rate differences are presented without stratification.

The p-values presented for each statistical test are deemed nominal. No multiplicity adjustment will be performed to control the overall Type I error.

5.3.3. Sensitivity Analysis

Sensitivity analyses of the primary efficacy variable will be based on the non-responder imputation (NRI) method. If a participant does not have the valid investigator-rated FWS assessment at maximum frown at any postintervention assessment through Day 7, he/she will be considered as a non-responder. CMH tests stratified by baseline investigator assessment of GL severity at maximum frown will be used for the analysis based on the imputed data. In addition, the pooled placebo group will be compared to each AGN study intervention group for proportion differences for the primary analysis.

5.3.4. Supplemental Approach

Categorical descriptive statistics including the sample size (N), frequency count, and percentage for each response score and change from baseline will be provided by study intervention group and by visit. Change from baseline values will be computed as the value for the postbaseline visit minus the baseline value.

5.4. Secondary and Exploratory Effectiveness Endpoint Analyses

5.4.1. Secondary and Exploratory Effectiveness Endpoints

Secondary effectiveness endpoints include:

- Achievement of a ≥ 2-grade improvement from baseline on the FWS according to **participant assessments** of GL severity at **maximum frown** at any postintervention timepoint through Day 7
- Achievement of a composite ≥2-grade improvement from baseline on the FWS at maximum frown as rated **concurrently by the investigator and participant** by timepoint
- Achievement of a ≥2-grade improvement from baseline on the FWS at maximum frown as rated by the **investigator** by timepoint
- Achievement of a ≥1-grade improvement from baseline on the FWS at maximum frown as rated by the **investigator** by timepoint
- Achievement of a ≥2-grade improvement from baseline on the FWS at maximum frown as rated by the **participant** by timepoint
- Achievement of a ≥1-grade improvement from baseline on the FWS at maximum frown as rated by the **participant** by timepoint
- None or mild rating in FWS at maximum frown as rated by **investigator** by timepoint
- None or mild rating in FWS at maximum frown as rated by **participant** by timepoint
- Responders (defined by primary endpoint) achieving none or mild FWS at maximum frown as rated by the **investigator** by timepoint

Exploratory effectiveness endpoints include:

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

5.4.2. Main Analytical Approach

The responder analyses will be analyzed in a similar way as for the primary effectiveness endpoint. The nominal p-values based on secondary effectiveness endpoints will be considered as supportive.

Descriptive statistics including the sample size (N), mean, standard deviation, median, 1st and 3rd quartiles, minimum, and maximum will be summarized by study intervention group and timepoint using the mITT Population for the FLO-11 item scores and total score, FLSQ item scores and domain scores.

5.5. Safety Analyses

Safety endpoints will be summarized by AGN-151586 dose group and placebo group in each cohort, as well as by total cohort groups and pooled placebo group and pooled AGN group using the Safety Population.

Study period in which events are experienced will be presented for safety data listings:

1 = screening period (prior to study intervention administration)
2 = treatment period (at time of study intervention administration)
3 = safety follow-up period (post-intervention through the end of study)

5.5.1. Treatment Administration

The treatment administration-related variables including the total dose will summarized by study intervention group using descriptive statistics.

Descriptive statistics will include the sample size (N), frequency count, and percentage for categorical endpoints, and sample size (N), mean, standard deviation, median, 1st and 3rd quartiles, minimum, and maximum for continuous endpoints.

5.5.2. Adverse Events

An adverse event (AE) will be considered a treatment-emergent adverse event (TEAE) if:

- The AE began on or after the date and time of the study intervention; or
- The AE was present before the date and time of the study intervention, but increased in severity or became serious on or after the date and time of the study intervention.

A TEAE will be considered a treatment-emergent serious adverse event (TESAE) if it is a TEAE that is also a serious adverse event (SAE).

A TEAE will be considered a treatment-related TEAE if it is considered related to the study intervention or study procedures by the principal investigator.

The number and percentage of participants with TEAEs as well as the events in the following AE categories will be summarized by study intervention group, as well as for all participants who received AGN-151586, as described in Table 5-1.

If more than 1 AE is coded to the same preferred term for the same participant, the participant will be counted only once for that preferred term using the most severe and most related occurrence for the summarizations by severity and by relationship to study intervention.

Listings of all TEAEs, TESAEs, TEAEs leading to study discontinuation by participant, and death will be presented.

Table 5-1 TEAE Tabular Summaries

Endpoint	Description	Methodology
Overall summary	Overall summary for the following categories: • TEAEs • Treatment-related TEAEs (including study drug-related TEAEs, study procedure-related TEAEs) • TESAEs • Treatment-related TESAEs • Possible distant spread of toxin (PDSOT) TEAEs • TEAEs leading to discontinuation from study • Deaths	Categorical descriptive summary statistics
TEAEs	• Overall summary and by system organ class (SOC) and preferred term (PT) • Overall summary and by SOC, PT, and maximum severity • Overall summary and by PT in descending order	Categorical descriptive summary statistics
Treatment-related TEAEs • Study drug-related TEAEs • Study procedure-related TEAEs	• Overall summary and by SOC and PT • Overall summary and by SOC, PT, and maximum severity • Overall summary by maximum severity, duration, time to onset, resolution, and treatment required	Categorical descriptive summary statistics, event descriptives
TESAEs	Overall summary and by SOC and PT if 5 or more participants overall reported such events	Categorical descriptive summary statistics
TEAEs leading to study discontinuation	Overall summary and by SOC and PT if 5 or more participants overall reported such events	Categorical descriptive summary statistics

In addition, neurologic assessments (Focused Symptoms Questionnaire and Focused Neurologic Examination) will be reported using descriptive statistics for baseline, follow-up visits, and for

the entire study period. Listings will also be provided. Treatment-emergent neurologic assessment-related AEs will be identified based on the abnormal findings and reported by SOC and PT.

To assess possible distant spread of toxin (PDSOT), MedDRA preferred terms that may be associated with botulinum toxin effects have been identified. The number and percentage of participants reporting TEAEs that are included in the PDSOT term list will be tabulated. Summary tables will be provided for participants with TESAEs and participants with TEAEs leading to discontinuation if 5 or more participants reported such events. In addition, all PDSOT AEs will be listed by subject. The terms are listed below.

[redacted]

5.5.3. Clinical Laboratory Assessments

Descriptive statistics for clinical laboratory values (in International System of Units [SI] units) and changes from the baseline values at each assessment visit will be presented by study intervention group for the following laboratory parameters:

Hematology: Absolute and differential white blood cell count (neutrophil count, lymphocyte count, monocyte count, basophil count, eosinophil count), hemoglobin, hematocrit, platelet count, red blood cell count, and red blood cell indices (mean corpuscular volume, mean

corpuscular hemoglobin, and mean corpuscular hemoglobin concentration, and distribution width).

Chemistry: Sodium, potassium, calcium, chloride, bicarbonate, magnesium, inorganic phosphate, glucose, blood urea nitrogen, creatinine, total protein, alkaline phosphatase (ALP), albumin, total bilirubin (TBL), alanine aminotransferase (ALT), aspartate aminotransferase (AST), total cholesterol, γglutamyl- transferase (GGT), lactate dehydrogenase (LDH), Uric acid

Urinalysis: Appearance, color, pH, specific gravity, glucose, protein, ketone, blood, urine microscopy

Clinical laboratory test values will be considered potentially clinically significant (PCS) if they meet either the lower-limit or higher-limit PCS criteria listed in Table 6-3. The number and percentage of participants who have PCS postbaseline clinical laboratory values will be tabulated by study intervention group. The percentages will be calculated relative to the number of participants with available non-PCS baseline values and at least 1 postbaseline assessment. The numerator will be the total number of participants with available non-PCS baseline values and at least 1 PCS postbaseline value. A supportive tabular display of participants with PCS postbaseline values will be provided, including the PID number, baseline and all postbaseline (including non-PCS) values. In this listing, any participant with PCS value (if any) will also be included.

In addition, a tabular display showing all AEs that occurred in participants who had PCS postbaseline clinical laboratory values will be provided.

5.5.4. Vital Signs

Descriptive statistics for vital signs (ie, temperature, respiratory rate, pulse rate, systolic and diastolic blood pressure) and changes from baseline values at each assessment visit will be presented by study intervention group.

Vital sign values will be considered as PCS if they meet both the observed-value criteria and the change-from-baseline criteria listed in Table 6-4. The number and percentage of participants with PCS postbaseline values will be tabulated by study intervention group. The percentages will be calculated relative to the number of participants with available non-PCS baseline values and at least 1 postbaseline assessment. The numerator will be the total number of participants with available non-PCS baseline values and at least 1 PCS postbaseline value. A supportive tabular display of participants with PCS postbaseline values will be provided, including the PID number, baseline and all postbaseline (including non-PCS) values.

In addition, a listing showing all AEs that occurred in participants who had PCS postbaseline vital sign values will be provided.

5.5.5. Electrocardiogram

Electrocardiogram (ECG) evaluations by investigators are captured in the eCRF and are classified into 3 categories: normal, abnormal, or not evaluable. These data will be summarized by study intervention group as change from baseline for each parameter, as well as the number and percent of participants in each category for PCS. Data listings which include ECG basic parameters and any ECG abnormalities will be produced.

5.5.6. Pregnancy Test Analyses

Urine pregnancy tests will be performed at the study center at screening, randomization/study intervention, and End of Study (Day 42 or Early Exit). Participants with a positive result will be presented in a listing for the Safety Population.

5.6. Other Analyses

5.6.1. Immunogenicity Analyses

Immunogenicity results, manifested as the presence of binding antibodies (positive), and neutralizing antibodies (positive) after a validated assay becomes available, to AGN-151586, will be summarized in a table for baseline and postbaseline, if available by immunogenicity database lock. Immunogenicity findings (positive) will be tabulated with the number and percentage of participants at each visit separately for each AGN-151586 treatment group. Percentages will be based on the number of treated participants with interpretable antibody assays in each treatment group at the specified visit.

5.7. Interim Analyses

Not applicable.

5.8. Data Monitoring Committee (DMC) Analyses

An internal DMC, independent from the study team, has been instituted to provide oversight in this study. Starting with Cohort 1, the DMC will convene to review select safety and efficacy data after the first subset of participants in Cohort 1 have completed the Day 14 visit. Following this review, the DMC will determine whether the study will proceed to Cohort 2 of dose escalation. After Cohort 2 proceeds, a similar review process will occur to determine whether Cohort 3 will proceed. This pattern repeats prior to the start of each new cohort. At each DMC review, data through Day 14 from a subset of participants ($n \geq 16$) for any given cohort will be

reviewed. Additionally, cumulative data from participants from the prior cohort(s) will also be included in the DMC review.

At each DMC review, the DMC may recommend proceeding with the nominal dose for the next cohort, request additional data (eg, data from additional participants in the cohort, data from visits after Day 14), request to repeat the current dose or a lower dose, request an intermediate dose be evaluated (between the dose of the current cohort and the nominal dose for the next cohort), or request dose escalation be terminated.

The dose-escalation design of this study allows for careful review of select safety and efficacy data by the DMC to perform risk/benefit analyses prior to escalation to the next higher dose. Details of the activities and composition of the DMC are further described in the DMC charter.

6. Supporting Documentation

6.1. Appendix 1 List of Abbreviations

AE	adverse event
ATC	Anatomical Therapeutic Chemical
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
DMC	data monitoring committee
ECG	electrocardiogram
eCRF	electronic case report form
[redacted]	[redacted]
[redacted]	[redacted]
FWS	Facial Wrinkle Scale With Photonumeric Guide
IES	Intercurrent event(s) strategy
MedDRA	medical dictionary for regulatory activities
mITT	modified intention-to-treat
PCS	potentially clinically significant
PDSOT	possible distant spread of toxin
PLS	Population-level summary
SAE	serious adverse event
SAP	statistical analysis plan
SI	L Système International d'Unités (International System of Units)
SOC	system organ class
TEAE	treatment-emergent adverse event
TESAE	treatment-emergent serious adverse event
WHO	World Health Organization

6.2. Appendix 2: Changes to Protocol-Planned Analyses

There is one change to the protocol-planned analyses. For the primary safety endpoints, TEAEs will be summarized, not AEs.

6.3. Appendix 3: Supporting Study Information

6.3.1. Demographics

Demographic parameters (age [years]; sex; race; ethnicity) will be summarized descriptively by study intervention group for the mITT and Safety Populations. Age (years) will be calculated relative to informed consent date.

6.3.2. Baseline and Disease Characteristics

Baseline characteristics including weight, height, body mass index, and disease characteristics including baseline investigators' and participants' assessments of the severity of GL at maximum frown and at rest using the validated FWS, baseline FLO-11 and FLSQ scores will be summarized descriptively by study intervention group for the mITT Population.

6.3.3. Protocol Deviations

Protocol deviations will be defined in a separate document. Unique participants reporting significant protocol deviations will be summarized in total and by treatment group for the mITT Population as described in Table 6-1.

Listing of significant protocol deviations will be provided.

Table 6-1 Protocol Deviation Summary

Endpoint	Description	Methodology
Significant protocol deviations	Number (%) of participants with significant protocol deviation will be summarized.	Categorical descriptive summary statistics

6.3.4. Medical History

Abnormalities in participants' medical, surgical, cosmetic and dental history, encompassing abnormalities, surgeries, and procedures reported as occurring before the Screening Visit, will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 23.0 or newer. The number and percentage of participants with abnormalities in medical and surgical histories in each system organ class and preferred term will be summarized by study intervention group for the mITT Population.

Listings will be provided for the mITT population.

6.3.5. Prior/Concomitant Medications

Medications will be coded using the WHO Drug Dictionary, version March 2017 or newer. Unique participants who reported medications (Anatomical Therapeutic Chemical [ATC] 4 class

and PT) will be summarized by study intervention group and by study period using the mITT Population as described in Table 6-2.

Table 6-2 Prior and Concomitant Medications

Endpoint	Description	Methodology
Prior medications	Medications taken before the date and time of study intervention	Categorical descriptive summary statistics
Concomitant medications	Medications on or after the date and time of study intervention to End-of-Study Visit (Day 42)	Categorical descriptive summary statistics

6.3.6. Criteria for Potentially Clinically Significant Laboratory Results

Criteria for potentially clinically significant laboratory results are listing in Table 6-3.

Table 6-3 Criteria for Potentially Clinically Significant Laboratory Results

Parameter	*SI Unit*	*Lower Limit*	*Higher Limit*
CHEMISTRY			
Albumin	g/L	< 0.9 × LLN	> 1.1 × ULN
Alanine aminotransferase (ALT)	U/L	—	≥ 3 × ULN
Alkaline phosphatase	U/L	—	≥ 3 × ULN
Aspartate aminotransferase (AST)	U/L	—	≥ 3 × ULN
Bilirubin, total	µmol/L	—	> 1.5 × ULN
Calcium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Chloride	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Cholesterol, Total	mmol/L	—	> 1.6× ULN
Creatinine	µmol/L	—	> 1.3 × ULN
Glucose, fasting, serum	mmol/L	< 0.8 × LLN	> 1.4 × ULN
Glucose, random, serum	mmol/L	< 0.8 × LLN	> 1.4 × ULN
Potassium	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
Urea Nitrogen (BUN)	mmol/L	—	> 1.2 × ULN
HEMATOLOGY			
Basophils, absolute cell count	10^9/L	—	> 3 × ULN
Neutrophils, absolute cell count	10^9/L	< 0.8 × LLN	> 1.5 × ULN
Hematocrit	Ratio	< 0.9 × LLN	> 1.1 × ULN
Hemoglobin	g/L	< 0.9 × LLN	> 1.1 × ULN
Platelet count	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	< 0.7 × LLN	> 1.5 × ULN
URINALYSIS			
pH	—	< 0.9 × LLN	> 1.1 × ULN
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.

6.3.7. Criteria for Potentially Clinically Significant Vital Sign Results

Criteria for potentially clinically significant vital sign results are listed in Table 6-4.

Table 6-4 Criteria for Potentially Clinically Significant Vital Sign Results

Parameter	***Flag***	***Criteria[a]***	
		Observed Value	*Change From Baseline*
Sitting systolic blood pressure, mm Hg	High	≥ 160	Increase of ≥ 20
	Low	≤ 90	Decrease of ≥ 20
Sitting diastolic blood pressure, mm Hg	High	≥ 105	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Sitting pulse rate, bpm	High	≥ 110	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Weight, kg	High	—	Increase of ≥ 7%
	Low	—	Decrease of ≥ 7%

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.

6.4. Data Handling Convention

[redacted]

[redacted]

[redacted]

[redacted]

6.4.2.1. Missing/Incomplete AE Start Date

Imputation of dates with missing day and/or month is only applied to TEAEs. If adequate information is available, no imputation is needed. TEAE start dates with missing day or month will be imputed as following:

- If day and month are missing but year is the same as year of study intervention, then the imputed day and month will be the date of study intervention.
- If day is missing but the month and year are available, then the imputed day will be the first day of the month, or the date of study intervention, if they have the same month and year, whichever is later.

6.4.2.2. Missing/Incomplete AE/Medication End Date

Imputation of dates with missing day and/or month is only applied to TEAEs when AE duration is calculated. If adequate information is available, no imputation is needed. TEAE end dates with missing day or month will be imputed as following:

- If day and month are missing but year is available, then the imputed day and month will be 31 Dec or the study exit date if they have the same year, whichever is earlier

If day is missing but the month and year are available, then the imputed day will be the last day of the month or the study exit date if they have the same month and year, whichever is earlier.

7. References

N/A