Statistical Analysis Plan

Clinical Trial Number: 43QM1602

Table of Contents

1 Study Information

This statistical analysis plan (SAP) describes the efficacy and safety summaries and analyses that will be performed for Clinical Trial Number (CTN) 43QM1602, *A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of QM1114-DP for the Treatment of Moderate to Severe Glabellar Lines (READY - 1)* and is based on the study protocol Version 4 dated 10JUN2020.

1.1 Background

1.1.1 Study design

This is a phase 3, multi-center, randomized, double blind, placebo-controlled study to assess the efficacy and safety of 50 units of QM1114-DP in the treatment of moderate to severe glabellar lines (GL).

Each subject will receive a single treatment of QM1114-DP 50 units or placebo divided equally among 5 injection points (0.1 mL per injection point) in the glabellar region. Following treatment at baseline, subjects will be monitored for safety and efficacy over a period of approximately 6.5 months.

1.1.2 Number of subjects and randomization

The study will screen approximately 330 male and female adults, 18 years of age and older, with moderate to severe glabellar lines at maximum frown in up to 15 study sites. Following the screening process, eligible subjects will be randomized at the baseline visit (Day 0) in a 3:1 ratio to QM1114-DP or placebo, stratified by site.

1.2 Study Objectives

The objective of the study is to evaluate the efficacy and safety of a single dose of 50 units of QM1114-DP compared to placebo in the treatment of moderate to severe GL.

1.2.1 Primary efficacy objective

The primary efficacy objective of the study as assessed using the endpoints in Section 1.4.1 is to evaluate a single dose of 50 units of QM1114-DP compared to placebo for the treatment of moderate to severe GL using the Validated 4-point Photographic Scale of Glabellar Line Severity: Investigator Live Assessment (GL-ILA) (described in Section 1.3.1) and Static 4-point Categorical Scale of Glabellar Line Severity: Subject Live Assessment (GL-SLA) (described in Section 1.3.2) at maximum frown.

1.2.2 Secondary efficacy objectives

The secondary efficacy objective of the study as assessed using the endpoints in Section 1.4.2 is to evaluate the efficacy of a single dose of 50 units of QM1114-DP compared to placebo for the treatment of moderate to severe GL using the GL-ILA.

CCI

1.2.4 Safety objective

The safety objective of the study is to evaluate the safety of a single dose of 50 units of QM1114-DP compared to placebo for the treatment of moderate to severe GL.

1.3 Efficacy assessment

For all assessments, baseline will be defined as the observation that is closest to but prior to study injection on Day 0. Likewise, in general change from baseline (Δ) will be calculated as the value at a given time point, *X*, minus the baseline value:

$$\Delta = X\ Value - Baseline\ (Day\ 0)\ Value$$

1.3.1 Validated 4-point Photographic Scale of Glabellar Line Severity: Investigator Live Assessment (GL-ILA)

The GL-ILA is a 4-point validated scale for assessment of glabellar lines. The validated 4-point Photographic Scale of Glabellar Line Severity (Appendix A) includes two grading systems: one for investigator live assessments at maximum frown, CCI . The scale represents the severity of GL from none (grade 0), mild (grade 1), moderate (grade 2) to severe GL (grade 3) as shown in Table 1.

Table 1. The static 4-point categorical scale.

Grade	Severity of Glabellar Lines	Description
0	No wrinkles	Smooth skin
1	Mild wrinkles	Fairly smooth skin
2	Moderate wrinkles	Frown lines
3	Severe wrinkles	Deep frown lines

The Investigators will use the GL-ILA for direct, live comparison with the subject's face at CCI at maximum frown. The Investigator will perform the GL-ILA at:

- Screening/Baseline (prior to treatment) visit(s)
- all post-treatment visits.

Multiple responder indicators will need to be created as follows:

- Primary endpoint GL-ILA indicator: subjects that achieved a score of 0 or 1, and had at least 2 grade improvement from baseline at maximum frown at Month 1 will have the value

'1'. This indicator will be used in combination with the primary endpoint GL-SLA indicator to define the primary endpoint composite responder.

- Subjects who achieve a score of 0 or 1 in the GL-ILA at maximum frown at each respective post-baseline visit will be considered a GL-ILA at maximum frown responder for that visit

CCI

[redacted]

1.3.2 Static 4-point Categorical Scale of Glabellar Line Severity: Subject Live Assessment (GL-SLA)

Similar to the GL-ILA, subjects will also assess their GL severity using the GL-SLA. The GL-SLA consists of 4 scaled options, presented in Table 1. This assessment will be done independently of the Investigator's assessment. Subjects will be asked to evaluate their GL at maximum frown at:

- Screening/Baseline (prior to treatment) visit(s)
- all post-treatment visits

Multiple responder indicators will need to be created as follows:

- Primary endpoint GL-SLA indicator: subjects that achieved a score of 0 or 1, and had at least 2 grade improvement from baseline at maximum frown at Month 1 will have the value '1'. This indicator will be used in combination with the primary endpoint GL-ILA indicator to define the primary endpoint composite responder.
- Subjects who achieve a score of 0 or 1 in the GL-SLA at maximum frown at each respective post-baseline visit will be considered a GL-SLA at maximum frown responder for that visit

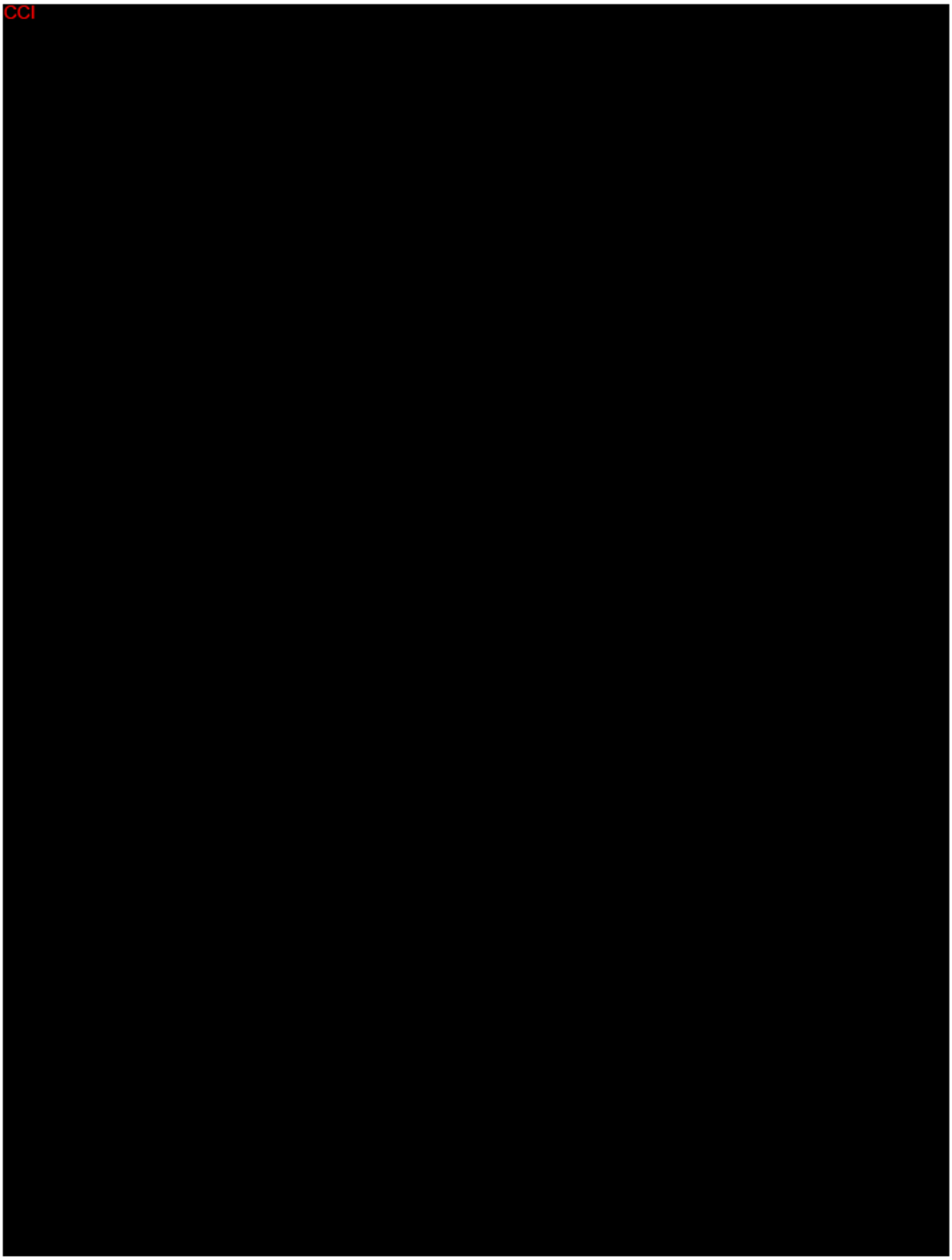
CCI

[redacted]

CCI

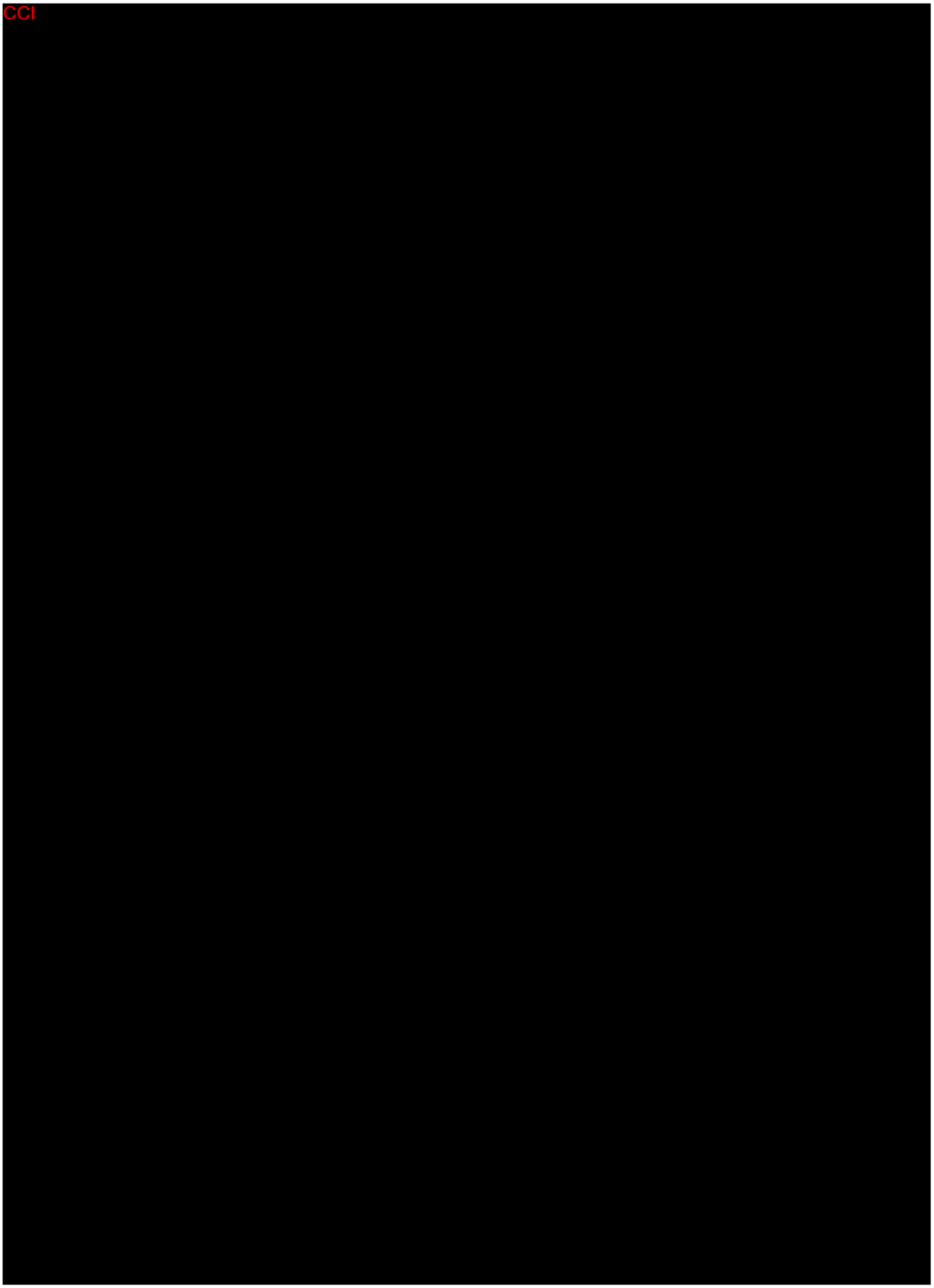
[redacted]

CCI



GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

CCI



CCI

[redacted]

1.4 Efficacy endpoints

1.4.1 Primary efficacy endpoint

The primary efficacy endpoint is the composite responder rate at Month 1 using GL-ILA and GL-SLA at maximum frown.

A composite responder is defined as a subject who achieves a score of 0 or 1 in glabellar line severity, and at least 2-grades improvement from baseline on both the GL-ILA and GL-SLA scales, concurrently, at Month 1.

1.4.2 Secondary efficacy endpoints

Secondary endpoints include:

(i) Percentage of subjects who achieve a score of 0 or 1 at each post-treatment visit using the GL-ILA at maximum frown

CCI

[redacted]

CCI

1.5 Safety assessments and endpoints

For details regarding the safety assessments, please refer to the CSP Section 7.2.

1.5.1 Adverse events

Adverse events (AEs) are to be monitored throughout the course of the study. All AEs reported will be coded using the Medical Dictionary for Regulatory Activities (MedDRA, version in force at the time of database freeze) and classified by MedDRA preferred term (PT) and system organ class (SOC). AEs will be defined as treatment-emergent adverse events (TEAEs) if the AE had an onset time greater than or equal to the time of study treatment. The study period for the purpose of AE collection is defined as the period from the signing of a study specific informed consent to study exit.

A two-point scale ("Yes" or "No" response) will be used for the causality assessments. The Treating Investigator should be asked to indicate a response to each of the following questions in the electronic Case Report Form (eCRF:):

- "Do you consider that there is a reasonable possibility that the event may have been caused by the study product?"
- "Do you consider that there is a reasonable possibility that the event may have been caused by the study product injection procedure?"

If either of these questions is answered with a 'Yes', the AE will be considered related.

AE endpoints include incidence and severity of TEAEs.

1.5.2 Laboratory safety tests

Hematology and blood chemistry laboratory tests will be performed at baseline (before treatment) and Month 6. Laboratory safety test endpoints include:

- Values collected at each visit
- Changes from baseline

1.5.3 Focused physical examination

Physical examination will be done at screening/baseline (before treatment), Day 7, Day 14, Month 1 and Month 6. Normal, abnormal and clinically significant findings will be assessed.

1.5.4 Vital signs

Vital signs will be assessed from baseline (before and after treatment), Day 7, Day 14, Month 1, and Month 6. Vital signs endpoints include:

- Values collected at each visit
- Changes from baseline

1.5.5 Neutralizing antibody testing

Blood samples will be taken for measurement of serum neutralizing antibody testing against QM1114-DP at baseline (before treatment), Month 1, and Month 6. Analysis of the blood samples include in vitro screening and confirmatory (if positive screening result) ELISA assays to test for the presence of binding antibodies, and in vivo mouse protection assay (MPA) to test for the presence of neutralizing antibodies. The MPA will only be conducted if the subject has a positive confirmatory result

1.5.6 Electrocardiogram (ECG)

ECG recordings will be done at baseline (before treatment), Month 1 and Month 6. The following items will be measured:

- RR interval
- PR interval
- QRS interval
- QT interval
- Heart rate (HR)
- QTcB interval
- QTcF interval

During each visit where ECG is conducted, the ECG will be recorded in triplicates. In addition to the individual measurements the above items will be aggregated into a mean value per visit (mean of all beats). ECG endpoints include:

- Aggregated values collected at each visit
- Changes from baseline
- QTcB and QTcF prolongation criteria
- Overall assessment (Normal, Abnormal Not Clinically Significant, Abnormal Clinically Significant) of the ECG

2 Statistical Methods

2.1 General methods

All tables, listings, and figures will be programmed using SAS Version 9.4 or higher. Data collected in this study will be documented using summary tables and subject data listings created by using the SAS® system. Confidence intervals (CI) and p-values will be 2-sided and performed at a significance level of 5%, unless otherwise specified. Data for all subjects in the clinical database will be included in the data listings. Calculated (derived) variables will be listed as appropriate. Any changes from the SAP will be detailed in the clinical study report.

All efficacy, safety and baseline characteristics variables will be presented using descriptive statistics within each treatment group, and graphs as appropriate. Continuous variables will be summarized using descriptive statistics (number of observations, mean, standard deviation [SD], median, minimum, and maximum). Categorical variables will be presented in frequency tables with number and percentage of observations for each level. Missing counts for all variables will be presented for informational purposes only and will not be included in percentage calculations.

Study days will be calculated relative to the injection of study drug. Day 0 will be the day of study drug administration. Baseline will be the last assessment prior to the injection of study drug unless otherwise indicated. The Screening Visit 1 (Day -14 to Day 0) will be considered the visit prior to injection of study drug. Because the Screening visit and Baseline visit (Day 0) may be performed on the same day, the Screening visit can also be Day 0.

Adverse events, cosmetic/aesthetic procedures and implant history events, medical history events, and concomitant treatments/procedures will be coded using MedDRA, Version 23.0. Prior/concomitant medications and procedural anesthetics will be coded using the World Health Organization (WHO) Drug Dictionary Global, 1 Sep 2019 B3 or higher.

In general, efficacy, safety, CCI [redacted] analyses will be performed and summarized by treatment group (QM114-DP, placebo), unless otherwise stated.

2.1.1 Visit windows

Study visits are expected to occur according to the protocol schedule in Appendix E. All data will be tabulated per the evaluation visit as recorded on the eCRF even if the assessment is outside of the visit window. In data listings, the relative study day (in relation to date of study drug administration) of all dates will be presented. There will not be any windowing for unscheduled visits in the analysis, and unscheduled visits will not be included in any analyses. Unscheduled visits, if any, will be presented in listings only.

2.1.2 Pooling of Centers

As this is a multi-center study, it is possible that some sites may only enroll a small number of subjects. Since this study plans to conduct some site-level analyses, pooling of sites may need to be considered for reliable and accurate results. If any site enrolls less than 5 subjects, then sites will be pooled by the following geographic regions:

- US sites
  - East
  - North
  - South

  - West
- Canadian sites
  - If only one site exists in this country then it's region will be pooled with the respective US region above (i.e. Canada – East will be pooled with US – East); otherwise sites will be pooled into the following geographic regions:
  - East
  - North
  - South
  - West

This region-level site grouping will then be used for all site-level analyses. This will be done prior to unblinding.

2.2 Analysis Populations

The statistical analyses will be performed based on the following four subject populations:

- Modified Intent-to-Treat efficacy population – The modified Intention-to-treat (mITT) population includes all subjects who are randomized and dispensed the investigational product, and will be analyzed according to the randomization scheme; subjects with a photographic and categorical scale Month 1 assessment via a remote visit will be excluded from the mITT.
- Intention-to-treat efficacy population - The Intention-to-treat (ITT) population includes all subjects who are randomized and dispensed the investigational product and will be analyzed according to the randomization scheme.
- Per-protocol efficacy population - The Per-protocol (PP) population is a subset of the mITT subjects who have no protocol deviations that are considered to have a substantial impact on the primary efficacy outcome.
- Safety population - The safety population includes all subjects who were administered the study product and will be analyzed according to as-treated principle.

2.3 Study subjects

2.3.1 Subject disposition

Subject disposition will be presented by treatment group and overall. The number of subjects in each study population (i.e. mITT, ITT, PP and Safety) will be summarized. Study population variables will also be presented in a data listing. Study completion, as well as early discontinuation, will be described for all subjects as well as by visit and by center.

Reasons for early discontinuation will be summarized and listed. All withdrawn subjects will be listed individually, including at least subject number, date and reason for withdrawal, and last visit performed, along with the relevant comments recorded on the eCRF (i.e., the Exit Form).

2.3.2 Protocol deviations

A protocol deviation occurs when a subject deviates from the protocol procedures. Depending on the seriousness of the deviation, the subject might be excluded from the PP analysis. Since PP will be used for the primary analysis at Month 1 only, the focus will be on deviations occurring before and on Month 1 visit day, as they might compromise the primary endpoint.

For this study, the protocol deviations that will exclude subjects from PP are identified (but not limited to) in Table 5 below.

Table 5. Protocol deviations

Deviation
Any subject that conducts their Month 1 visit out of window by greater than 14 days or earlier than 7 days
Any subject not treated according to the assigned treatment
Any subject that does not complete the GL-ILA at maximum frown at the primary endpoint visit (Month 1)
Any subject that does not complete the GL-SLA at maximum frown at the primary endpoint visit (Month 1)
Any subject that does not have an available screening or baseline GL-ILA at maximum frown assessment
Any subject that does not have an available screening or baseline GL-SLA at maximum frown assessment
Any subject that does not have grade 2 or 3 at maximum frown on the GL-ILA at baseline
Any subject that do not have grade 2 or 3 at maximum frown on the GL-SLA at baseline
Any subject with prohibited concomitant treatments/procedures prior to Month 1 visit considered to have a substantial impact on the primary efficacy outcome.
Any subject with a prohibited medical history, unstable medical history condition, or medical history condition that worsens prior to Month 1 visit considered to have a substantial impact on the primary efficacy outcome

Before unblinding the subject data, and as a part of preparations for the database lock, the study team will review all protocol deviations. All protocol deviations will be presented in a data listing.

Handling of Protocol Deviations During COVID-19

Due to the public health emergency related to the COVID-19 pandemic during 2020, steps have been taken to ensure patient and practitioner safety in alignment with FDA Guidance dated May 11, 2020 (Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency). Most notably, in partnership with clinical sites and the Institutional Review Board (IRB), optional remote assessment procedures for efficacy and safety endpoints has been implemented to ensure safety and respect localized and elective restrictions.

Protocol deviations will presented descriptively, overall and by treatment group. The total number of deviations, the type of protocol deviation, and if the deviation was reportable to IRB will be summarized. These summaries will be stratified by relatedness to COVID-19 (related, not related). The above summary of protocol deviations will be repeated by site as well.

A listing of all protocol deviations reported throughout the study, including their relatedness to COVID-19, be provided.

2.3.3 Demographic characteristics

Demographic assessments for this study include:

- Age (years)
- Height (in)
- Weight (lbs.)

For this study, the protocol deviations that will exclude subjects from PP are identified (but not limited to) in Table 5 below.

Table 5. Protocol deviations

Deviation
Any subject that conducts their Month 1 visit out of window by greater than 14 days or earlier than 7 days
Any subject not treated according to the assigned treatment
Any subject that does not complete the GL-ILA at maximum frown at the primary endpoint visit (Month 1)
Any subject that does not complete the GL-SLA at maximum frown at the primary endpoint visit (Month 1)
Any subject that does not have an available screening or baseline GL-ILA at maximum frown assessment
Any subject that does not have an available screening or baseline GL-SLA at maximum frown assessment
Any subject that does not have grade 2 or 3 at maximum frown on the GL-ILA at baseline
Any subject that do not have grade 2 or 3 at maximum frown on the GL-SLA at baseline
Any subject with prohibited concomitant treatments/procedures prior to Month 1 visit considered to have a substantial impact on the primary efficacy outcome.
Any subject with a prohibited medical history, unstable medical history condition, or medical history condition that worsens prior to Month 1 visit considered to have a substantial impact on the primary efficacy outcome

Before unblinding the subject data, and as a part of preparations for the database lock, the study team will review all protocol deviations. All protocol deviations will be presented in a data listing.

Handling of Protocol Deviations During COVID-19

Due to the public health emergency related to the COVID-19 pandemic during 2020, steps have been taken to ensure patient and practitioner safety in alignment with FDA Guidance dated May 11, 2020 (Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency). Most notably, in partnership with clinical sites and the Institutional Review Board (IRB), optional remote assessment procedures for efficacy and safety endpoints has been implemented to ensure safety and respect localized and elective restrictions.

Protocol deviations will presented descriptively, overall and by treatment group. The total number of deviations, the type of protocol deviation, and if the deviation was reportable to IRB will be summarized. These summaries will be stratified by relatedness to COVID-19 (related, not related). The above summary of protocol deviations will be repeated by site as well.

A listing of all protocol deviations reported throughout the study, including their relatedness to COVID-19, be provided.

2.3.3 Demographic characteristics

Demographic assessments for this study include:

- Age (years)
- Height (in)
- Weight (lbs.)

- BMI (kg/m^2)
- Gender (Male, Female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other)
- FST score (I, II, III, IV, V, VI)
- Prior use of botulinum toxin (Naive, Non-Naive)

Subject demographic data will be summarized for the ITT population by treatment group and overall. Age, height, weight, and BMI will be analyzed as continuous variables. Gender, race, ethnicity, Fitzpatrick skin type, and prior botulinum toxin use status will be analyzed as categorical variables.

Demographics and baseline characteristics will be presented by subject in a data listing.

2.3.4 Medical history and previous/concomitant medication (including drugs and medical and surgical procedures)

All summaries will be done by treatment group based on the ITT population. History of relevant or clinically significant surgical events and medical conditions, including any prior cosmetic/aesthetic procedures or implants, will be collected. Medical History will be coded according to MedDRA; the version used will be noted as a footnote in the tables and listings.

The number and percentage of subjects reporting medical history will be summarized by system organ class (SOC) and preferred term (PT). System organ class and PTs will be presented in descending frequency first based on the QM1114-DP group, and then alphabetically if there are ties. Each subject will contribute at most one count per summarization category. In other words, if a subject has more than one medical history event with same PT, the subject will be counted only once for that PT. Similarly, if a subject has more than one medical history event for a SOC, the subject will be counted only once in that SOC.

Cosmetic/aesthetic procedures and/or implant history will also be presented and will follow the same methods specified above.

Medical history, and prior cosmetic/aesthetic procedures or implants will be provided in the subject data listing. Medical history will be further presented in a separate listing by past and ongoing GL medical history.

Concomitant medications for this study are defined as any ongoing medications with a start date prior to the date of injection, any changes to existing medications (such as dose or formulation) during the course of the study, or any new medications received by the subject since the date of injection. Prior medications are medications with stop dates prior to study treatment and used within 4 weeks preceding screening visit. Medications will be coded using the World Health Organization (WHO) Drug Dictionary; concomitant procedures will be coded according to MedDRA. The versions used for the coding will be noted as a footnote in the tables and listings.

The number and percentage of subjects who receive prior and concomitant medications will be summarized by the WHO Drug Dictionary Anatomical Therapeutic Chemical 3rd level (ATC-3) and the preferred name. If the 3rd level term is not available, the next available level (e.g., ATC-2)

will be used. In addition, the number and percentage of subjects reporting a concomitant medication/therapy will be summarized by reason (medical history, adverse event, concomitant procedure, contraception, or other). Therapies and procedures that started due to an AE will be summarized separately from those who did not start due to an AE.

ATC-3 and preferred name will be presented in descending frequency first based on the QM1114-DP group, and then alphabetically if there are ties. Each subject will contribute at most one count per summarization category. In other words, if a subject has more than one medication with same preferred name, the subject will be counted only once for that preferred name. Similarly, if a subject has more than one medication for an ATC-3 level, the subject will be counted only once in that ATC-3 level and preferred name. Concomitant procedures will also be presented and will follow the same methods specified for medical history.

Prior and concomitant medications/procedures will be presented by subject in a data listing.

Handling of Missing/Partial Dates

While every effort will be made to obtain full, complete information on every reported medication, the following imputation rules will be followed for any respective missing medication data:

For the purpose of determining whether a medication is considered prior or concomitant, the following date imputation rules will used. Dates will be presented as collected in the listings.

- Start Date
  - If the start date is completely missing, it will be assumed that the medication started on the study treatment date.
  - If the start date is missing the day, the first of the month will be used (i.e. UNK-JAN-2019 becomes 01-JAN-2019), provided the imputed date is on or after the subject's study treatment date; otherwise, the day of study treatment will be used.
  - If the start date is missing the month, the month of 'June' will be used (i.e. 01-UNK-2019 becomes 01-JUN-2019), provided the imputed date is on or after the subject's study treatment date; otherwise, the subsequent month after study treatment will be used.
  - If the start date is missing the year, the year of study treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is on or after the subject's study treatment date; otherwise, the subsequent year after treatment will be used.
- End Date
  - If the end date is completely missing, it will be assumed that the medication is still ongoing and will not be imputed.
  - If the end date is missing the day, the last day of the month will be used (i.e. UNK-JAN-2019 becomes 31-JAN-2019).
  - If the end date is missing the month, the subsequent month after the start date will be used.
  - If the end date is missing the year, the year of study treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is after the start date; otherwise, the subsequent year after start date will be used.

2.4 Efficacy analysis

2.4.1 Datasets analyzed

The primary efficacy analysis will be analyzed based on the mITT population. All secondary efficacy CCI [redacted] variables will be analyzed based on the ITT population, unless otherwise specified below.

2.4.2 Handling of missing data

In general, the number of subjects with missing values will be summarized and reported as appropriate in all outputs. The primary analysis will be performed using multiple imputation (MI) as the primary imputation method and repeated using baseline observation carried forward (BOCF) as a sensitivity analysis for missing values. With the BOCF imputation method, if a subject is missing their Month 1 GL-ILA score or GL-SLA Month 1 score, their GL-ILA/GL-SLA score at baseline will be used, respectively.

The imputation using MI will assume the Missing Completely at Random (MCAR) missing data assumption. Regardless of the actual pattern of missing data, the Markov Chain Monte Carlo (MCMC) method of the MI procedure from the SAS® system will first be used to create a monotonic pattern of missing data. Then, a second MI procedure will be used to generate five sets of data with missing values imputed from observed data. Linear regressions will be employed to model the missing GL-ILA and GL-SLA scores, separately, with the following covariates included in each imputation model: treatment group and non-missing GL-ILA or GL-SLA data, respectively, from earlier timepoints (baseline, Day 7, and Day 14). The imputed datasets will be used to create the composite responder variable, which will then be analyzed using the methodology described for the primary analysis of responder rates at Month 1 (Section 2.4.3). The results from the analysis of the multiple imputed datasets will be combined by the MIANALYZE procedure of the SAS® system. The seed number to be used will be the number 110196.

The secondary efficacy analysis (described in Section 2.4.4) will be performed using the Observed Cases (OC), that is, no imputation will be done. CCI [redacted]
CCI [redacted]

2.4.3 Primary analysis

The primary efficacy endpoint will be the composite responder rate based on the GL-ILA and GL-SLA of GL severity at maximum frown at Month 1 (described in Section 1.3.1 and Section 1.3.2, respectively). To evaluate the effectiveness of QM1114-DP versus placebo in the treatment of moderate to severe GL, the responder rates of the QM1114-DP and placebo will be compared using the Cochran-Mantel-Haenszel (CMH) test stratified by site at the 5% significance level (2-sided).

The null hypothesis of no relationship between treatment and responder rate (i.e. the responder rates are the same in both groups) will be tested against the alternative hypothesis that there is a relationship between treatment and responder rate (i.e. the responder rates are different in the two groups). For a significant result, the two-sided p-value of the comparison of the composite GL-ILA/GL-SLA responder rates between the treated and untreated subjects at Month 1 using the CMH test needs to be smaller than 0.05. For consistency across strata (site) the Breslow-Day test will be used to assess the homogeneity of the odds ratios across all sites.

The estimates of the composite GL-ILA/GL-SLA responder rates in each treatment group will be presented as well as the difference in responder rates (QM1114-DP responder rate – placebo responder rate). Corresponding 95% CI for the treatment group composite GL-ILA/GL-SLA responder rates and the difference in responder rates along with the p-value for the difference will also be presented. The normal approximation (Wald) method will be used to calculate both the 95% CI for the individual treatment group composite GL-ILA/GL-SLA responder rates and the 95% CI for the difference in responder rates. The above responder rates will be presented in figures by visit and treatment group.

2.4.3.1 Sensitivity analysis

To evaluate the impact of missing data on the primary endpoint, sensitivity analyses will be performed. The primary analysis specified in Section 2.4.3 will be repeated using the BOCF method (detailed in Section 2.4.2), and also using the ITT OC.

In addition, a sensitivity analysis of the primary efficacy endpoint will be performed based on the PP population, and also on the Safety population to account for potential errors in randomization. The analysis method described above in Section 2.4.3 will be repeated but using the PP population and Safety Population.

2.4.3.2 Subgroup analysis

Additionally, to evaluate the consistency of the results of the primary analysis across different subgroups of interest, the primary analysis specified above (Section 2.4.3) will be repeated, stratifying for each of the following subgroups specified below:

- Age (less than 65 years old, 65+ years old)
- Gender (Male, Female; if enough males are recruited)
- Baseline Severity score of the ILA at maximum frown (Grade 2 (Moderate), Grade 3 (Severe))
- Prior botulinum toxin use (Yes, No)
- Fitzpatrick skin type (I-III, IV-VI)
- Site
- Type of Month 1 visit (Onsite, Remote)

For each subgroup, the responder rates, difference in responder rates between the treatment groups, and the corresponding 95% CI will be presented. The Breslow-Day test will be used to assess the homogeneity of the odds ratios across all the subjects within each subgroup. The subgroup analysis for type of Month 1 visit will be performed based on the ITT population.

2.4.4 Secondary analysis

To evaluate the effectiveness of QM1114-DP versus placebo in the treatment of moderate to severe GL, the proportion of subjects who achieve 0 or 1 on the GL-ILA at maximum frown at all post-treatment visits will be compared using the CMH test stratified by site. To control the type I error rate, the fixed sequence testing procedure will be used, which requires no adjustment to the level of significance. However, the comparisons must be made in a pre-specified order following the procedure outlined below:

1. Primary analysis must be done first, prior to any other analysis

Secondary analysis:

2. Month 1 responder rate
3. Day 14 responder rate
4. Month 2 responder rate
5. Month 3 responder rate
6. Month 4 responder rate
7. Month 5 responder rate
8. Month 6 responder rate
9. Day 7 responder rate

If one of the tests in the sequence is not significant ($p > 0.05$), no confirmatory claims can be made based on tests following (and including) that one.

The secondary analysis in the sequence will be based on the ITT population with the remote visit assessments excluded. A sensitivity analysis will be run using the ITT population including the remote visit assessments.

Results will include estimates of the percentages, difference in percentages, 95% CIs, and CMH p-value as well as a graph showing the percentages for the two treatment groups over time. For consistency across strata, the Breslow-Day test will be used.

CCI

2.5 Safety analysis

All safety data will be summarized descriptively based on the safety population. All of the safety analyses will be done using OC. If deemed necessary, any analyses may be repeated using OC, BOCF, or MI as appropriate.

2.5.1 Extent of exposure

There will only be limited exposure to QM1114-DP since it is injected only once. The number of subjects receiving a single dose of QM1114-DP and placebo, respectively, will be presented.

2.5.2 Adverse events

All AE data will be summarized by treatment group. Missing dates will be imputed as described below. AEs will be summarized by SOC and PT. AEs occurring before treatment will be presented in listings only. The MedDRA version used for the coding will be noted as a footnote in the tables and listings.

A summary of all AEs will be provided, which will include:

- number (%) of subjects who did not have an AE
- number (%) of subjects with at least one TEAE and number of events
- number (%) of subjects with at least one TEAE related to study product or injection procedure
- number (%) of subjects with at least one TEAE not related to study product or injection procedure
- number (%) of subjects with at least one TEAE leading to discontinuation
- number (%) of subjects with at least one serious TEAE.

Summaries of TEAEs (including the total number of events, number and percentage of subjects) will be displayed by treatment group according to the following:

- All TEAEs by SOC and PT
- Treatment emergent SAE by SOC, PT, maximum intensity (mild, moderate, severe), and causality
- Related TEAEs by SOC and PT, and maximum intensity (mild, moderate, severe)
- Unrelated TEAEs by SOC and PT, and maximum intensity (mild, moderate, severe)
- TEAEs leading to discontinuation by SOC and PT, and maximum intensity (mild, moderate, severe)
- TEAEs by SOC, PT, and action taken (none, medical treatment, non-pharmacological treatment, subject withdrawn)

For the subject level analyses, the number and percentage of subjects who experienced at least one of the events listed above will be summarized overall and for each SOC and each PT. System organ class and PTs will be presented in descending frequency first (by QM1114-DP group), and then alphabetically if there are ties. Each subject will contribute at most one count per summarization category. In other words, if a subject has more than one TEAE with same PT, the subject will be counted only once for that PT. Similarly, if a subject has more than one TEAE for a SOC, the subject will be counted only once in that SOC and PT. For the event level analyses, the counts of each respective event will be presented. In general, percentages will be calculated using the number of subjects in the safety population for the denominator. For the subgroup analyses (described in Section 2.5.2.1) percentages will be calculated using the number of subjects in the safety population for each respective sub-category will be used for the denominator.

For the "action taken" summary specifically, subjects will be only counted in 'None' category if no other action was taken. The onset/duration summaries will be presented at the event level (i.e. will include multiple AEs within the same SOC and PT). Number of days to onset and duration of event will be summarized by SOC and PT, using mean, SD, minimum, maximum, and median statistics. Time to onset will be calculated as the first day with the AE *minus* the Day 0. Duration will be calculated as the last day with the AE *minus* the first day with the AE *plus* one. The text below lists the imputation rules for partial/completely missing dates. Missing stop date will not be imputed and therefore no duration will be calculated in these cases. Instead, the number of AEs that were ongoing at the end of the study will be given.

Handling of Missing/Partial Dates

While every effort will be made to obtain full, complete information on every reported AE, the following imputation rules will be followed for any respective missing AE data:

For the purpose of calculating treatment emergence, onset time, and duration, the following date imputation rules will used. Dates will be presented as is in the listings.

- Start Date
  - If start date is completely missing, it will be assumed that the AE started on the treatment date.
  - If the start date is missing the day, the first of the month will be used (i.e. UNK-JAN-2019 becomes 01-JAN-2019), provided the imputed date is on or after the subject's treatment date; otherwise, the day of treatment will be used.
  - If the start date is missing the month, the month of 'June' will be used (i.e. 01-UNK-2019 becomes 01-JUN-2019), provided the imputed date is on or after the subject's treatment date; otherwise, the subsequent month after treatment will be used.
  - If the start date is missing the year, the year of treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is on or after the subject's treatment date; otherwise, the subsequent year after treatment will be used.
- End Date
  - If end date is completely missing, it will be assumed that the AE is still ongoing and will not be imputed.
  - If the end date is missing the day, the last of the month will be used (i.e. UNK-JAN-2019 becomes 31-JAN-2019)
  - If the end date is missing the month, the subsequent month after the start date will be used.
  - If the end date is missing the year, the year of treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is after the start date; otherwise, the subsequent year after start date will be used.

2.5.2.1 Subgroup analysis

To evaluate the consistency of the AE data, subgroup analyses will be performed. All TEAEs related to study product or injection procedure by SOC, PT, and maximum intensity will be repeated by the following subgroups. The same methods specified above (Section 2.5.2) will be followed

- Age (less than 65 years old, 65+ years old)
- Gender (Male, Female)
- Prior use of botulinum toxin (Naive, Non-Naive)
- Race (White, Black, Other)
- Fitzpatrick skin type (I-III, IV-VI)
- Site

2.5.2.2 Assessment of local/remote toxin spread and hypersensitivity

Any potential or suspected toxin spread or toxin hypersensitivity events will be evaluated separately. The same methods specified in Section 2.5.2 will be followed. Suspected toxin spread events and suspected hypersensitivity events will be summarized in separate tables. The investigator will conduct the investigation and evaluate if an AE is also a suspected toxin spread or hypersensitivity event; however, Appendix 6 in the CSP lists AEs potentially suggestive of spread of toxin.

2.5.3 Laboratory assessments

The laboratory data (hematology and clinical chemistry) at baseline Day 0 and Month 6 and the changes from baseline will be summarized by descriptive statistics. All clinically significant out-of-range laboratory values for blood samples collected at baseline will be recorded in the subject's medical history and all clinically significant out of range laboratory values for blood samples collected after baseline are to be reported as an AE if this abnormality was not present at the baseline visit or is assessed as having worsened since the baseline visit.

Throughout the study it is possible that certain laboratory parameter results are below the limit of quantification (BLQ). Any BLQ laboratory values will be summarized by the imputed value of half the limit of quantification for that parameter.

All laboratory data will be presented in data listings as collected, without any imputations for BLQ values.

2.5.4 Neutralizing antibody testing

Sampling for antibody testing will be conducted at baseline Day 0, Month 1, and Month 6. For the antibody testing results, due to the timing and availability of the data, results will be presented in a separate report.

2.5.5 Physical examinations and vital signs

The number and percentage of subjects with normal/abnormal results in physical examination will be presented by visit and treatment. A shift table will be created to present any change from baseline in normal/abnormal results in physical examination across the study visits for each treatment group.

Vital signs at baseline, Day 7, Day 14, Month 1, and Month 6, and the changes from baseline will be summarized by treatment and visit using descriptive statistics.

2.5.6 ECG

The aggregated value of each ECG measurement (described in Section 1.5.6) at baseline, Month 1, Month 6, and the changes from baseline will be summarized descriptively by treatment group at each visit. In addition, a shift table will be created to show any change from baseline to Month 6 in the overall ECG assessment.

QTcF and QTcB interval prolongation will be presented separately as well. The number and percentage of subjects with a QTcF and/or QTcB interval of >450 msec, >480 msec, and >500 msec will be summarized by treatment and visit, along with the number and percentage of subjects with a change from baseline in QTcF and/or QTcB interval of 30 - < 60 msec and ≥60 msec.

2.5.7 Urine pregnancy test

A data listing will be provided to summarize the results of the urine pregnancy tests. All pregnancy SAEs will be flagged within the data listing.

2.6 Interim Analysis

Not applicable.

2.7 Determination of Sample Size

The sample size is determined by the number of subjects exposed to QM1114-DP in the GL, and amount of long-term safety data. Further, due to the public health emergency related to the COVID-19 pandemic during 2020, steps have been taken to ensure patient and practitioner safety in alignment with FDA Guidance dated May 11, 2020 (Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency). Most notably, in partnership with clinical sites and the IRB, optional remote assessment procedures for efficacy and safety endpoints has been implemented to ensure safety and respect localized and elective restrictions. However, the number of subjects in the mITT population used in the assessment of the primary efficacy endpoint, should still satisfy the sample size requirements for the primary efficacy analysis.

CCI [redacted]

CCI [redacted]

2.7.3 Sample size calculation

The study is planned to include approximately 300 subjects, of whom 225 will be treated with QM1114-DP and 75 subjects will receive placebo. CCI [redacted]

2.8 Changes in Analysis Planned in the Protocol

The table below outlines all rationale and changes in the analyses specified in the SAP that differ from the analyses specified in the protocol.

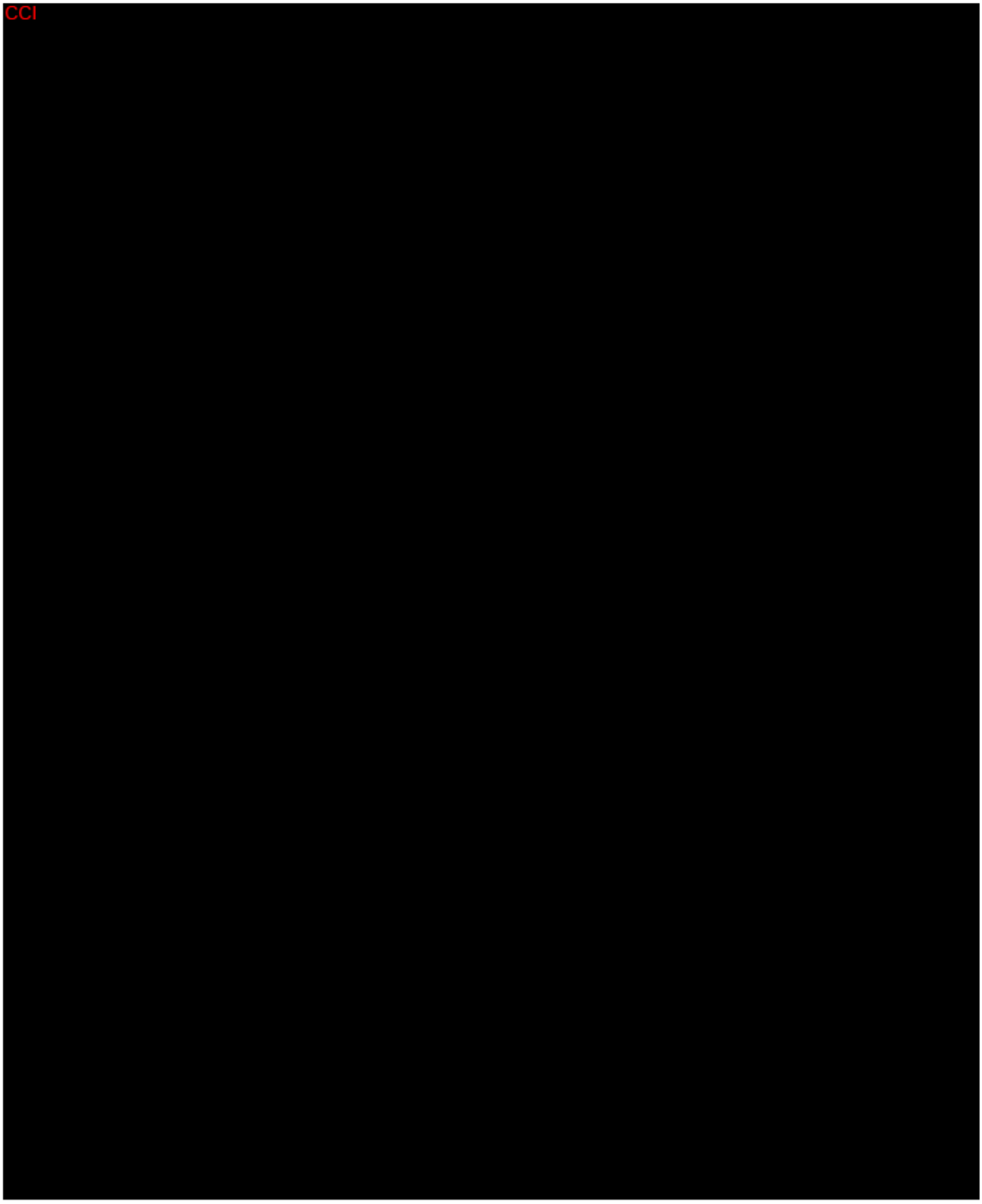
Table 6 Changes in Analysis Planned in the Protocol

CSP Section	SAP section	Description/Rationale of Change
9.1.2.3 Per-Protocol (PP) Efficacy Population	2.4.3.1 Sensitivity Analysis	Regardless of if the PP population contains less than 90% of the mITT population, sensitivity analyses for the primary efficacy endpoint using the PP population will be conducted.

3 Reference List

4 Appendix A GL-ILA Scales

CCI



Maximal Frown

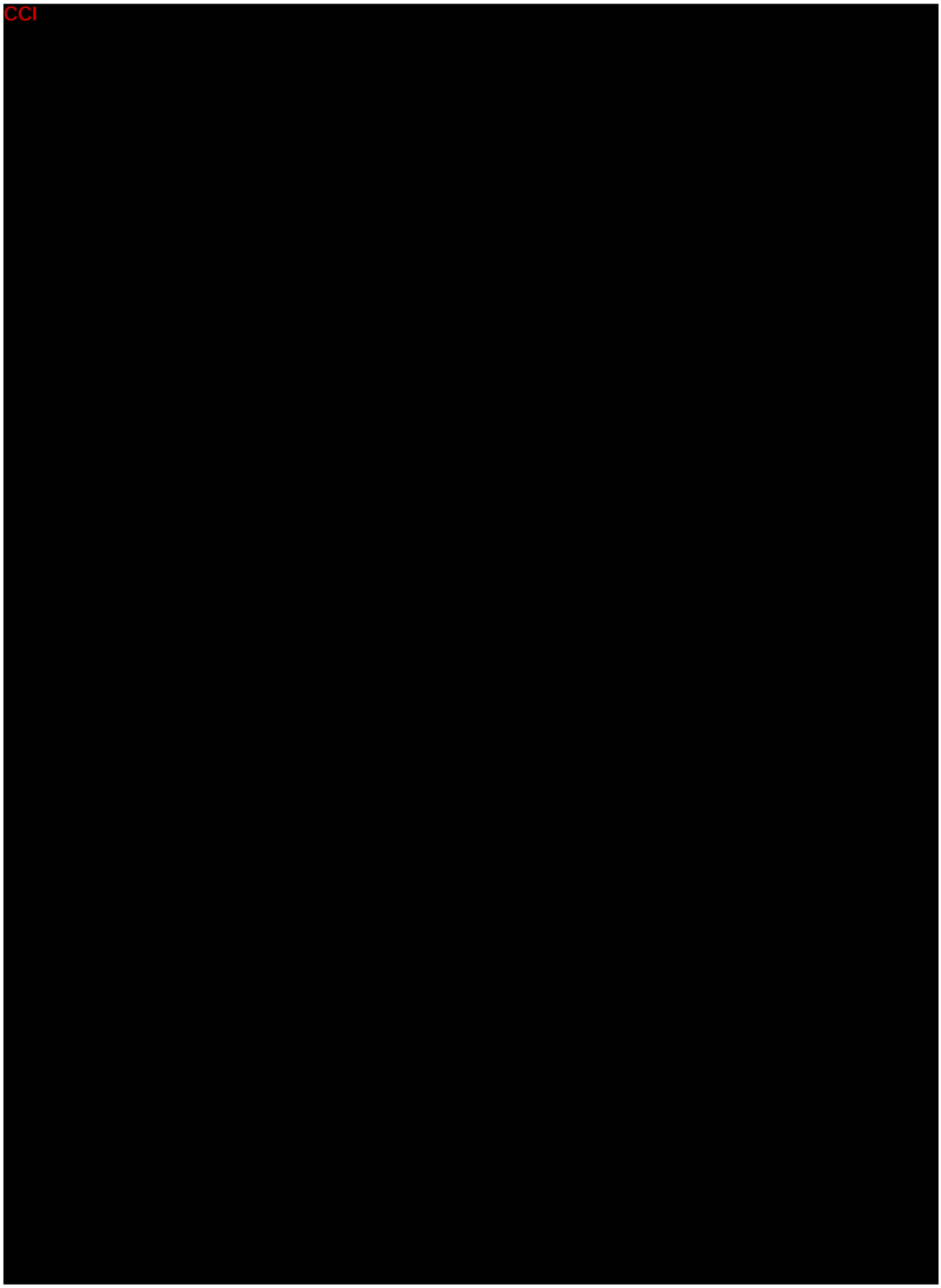
CCI

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

CCI

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

CCI



8 Appendix E Schedule of Assessments

1 month = 4 weeks/28 days All visit windows are calculated from Baseline/Day 0	Visit 1 Screening[1]	Visit 2 Baseline/Day 0[1] (within 2 weeks after screening)	Visit 3 Day 7 (±1 day)	Visit 4 Day 14 (± 3 days)	Visit 5 Month 1 (±5 days)	Visit 6 Month 2 (±5 days)	Visit 7 Month 3 (±5 days)	Visit 8 Month 4 (±5 days)	Visit 9 Month 5 (±5 days)	Visit 10 Month 6/ET[2] (±5 days)
Informed Consent	X									
Demographic Data[3] including, Fitzpatrick skin type, medical history & concurrent diseases, previous facial treatments/procedures (toxin naïve/non-toxin naïve)	X									
Inclusion /Exclusion Criteria	X	X[4]								
Concomitant Therapies/ Procedures	X	X[5]	X	X	X	X	X	X	X	X
Adverse Events	X	X[5]	X	X	X	X	X	X	X	X
Urine Pregnancy Test[6]	X	X[4]								X
Vital Signs (blood pressure, heart rate, and respiratory rate)[7]		X[5]	X	X	X					X
ECG		X[4]			X					X
Blood sample clinical chemistry and hematology		X[4]								X
Blood sample for serum antibody testing		X[4]			X					X
ILA of GL Severity	X	X[4]	X	X	X	X	X	X	X	X
CCI										
Photography		X[4]	X	X	X	X	X	X	X	X
Randomization		X[4]								
Treatment		X[9]								
CCI										
Subject Assessments										
SLA of GL Severity[10]	X	X[4]	X	X	X	X	X	X	X	X
CCI										

1. *Screening and baseline visits may be on the same day. If completed on the same day, only perform study assessments once (i.e., PE, UPT, SLA, ILA, AE, concomitant therapies/procedures, inclusion/exclusion review).*
2. *If the subject withdraws before the final visit the assessments at Month 6/ET should be completed, if possible.*
3. *Includes date of birth, gender, race, ethnicity, height, weight.*
4. *To be performed before treatment.*
5. *Performed before treatment and post-treatment.*
6. *Females of childbearing potential*
7. *Vital signs are taken seated after 10 minutes rest. Vital signs are taken prior to any blood draw (excluding post-*

CCI

9. *Following treatment administration, subjects will be monitored at the study center for 30 minutes.*
10. *Subjects will make their GL assessments independently of the Investigator's assessment.*

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

PPD

Statistical Analysis Plan Approval Form

Client:	Galderma	
Identifier (e.g., Protocol Name and Number, ISS, ISE, SCS, or CSE):	43QM1602	
Version (Protocol Only):	☐ Original, Date:	☒ Amendment, Number 4.0; Date: 10JUN2020
Title:	A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of QM1114-DP for the Treatment of Moderate to Severe Glabellar Lines (READY - 1	
Statistical Analysis Plan Version:	☒ Original, Date: 27OCT2020	☐ Amendment, Number ; Date:

AC Statistical Analysis Plan Author:

PPD
Senior Biostatistician
Printed Name & Title

Handwritten Signature **OR** *Electronic Signature*

Signature: _____

Date: _____
(dd/mmm/yyyy)

PPD

DocuSign

AC Statistical Analysis Plan Reviewer:

PPD
Senior Biostatistician
Printed Name & Title

Handwritten Signature **OR** *Electronic Signature*

Signature: _____

Date: _____
(dd/mmm/yyyy)

PPD

DocuSign

First Client Statistical Analysis Plan Approval:

PPD
Printed Name & Title

PPD *Handwritten Signature*

Signature: _____

Date: PPD
(dd/mmm/yyyy)

Second Client Statistical Analysis Plan Approval ☐ Not Applicable

PPD
Printed Name & Title

PPD *Handwritten Signature*

Signature: _____

Date: PPD
(dd/mmm/yyyy)

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

SIGNATURES PAGE

Date	Signed by
2020-11-27 15:58	PPD
Justification	Approved by Technical Expert

2020-11-30 08:14	PPD
Justification	Approved by Owner

2021-02-05 13:17	PPD
Justification	Approved by

- BMI (kg/m^2)
- Gender (Male, Female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other)
- FST score (I, II, III, IV, V, VI)
- Prior use of botulinum toxin (Naive, Non-Naive)

Subject demographic data will be summarized for the ITT population by treatment group and overall. Age, height, weight, and BMI will be analyzed as continuous variables. Gender, race, ethnicity, Fitzpatrick skin type, and prior botulinum toxin use status will be analyzed as categorical variables.

Demographics and baseline characteristics will be presented by subject in a data listing.

2.3.4 Medical history and previous/concomitant medication (including drugs and medical and surgical procedures)

All summaries will be done by treatment group based on the ITT population. History of relevant or clinically significant surgical events and medical conditions, including any prior cosmetic/aesthetic procedures or implants, will be collected. Medical History will be coded according to MedDRA; the version used will be noted as a footnote in the tables and listings.

The number and percentage of subjects reporting medical history will be summarized by system organ class (SOC) and preferred term (PT). System organ class and PTs will be presented in descending frequency first based on the QM1114-DP group, and then alphabetically if there are ties. Each subject will contribute at most one count per summarization category. In other words, if a subject has more than one medical history event with same PT, the subject will be counted only once for that PT. Similarly, if a subject has more than one medical history event for a SOC, the subject will be counted only once in that SOC.

Cosmetic/aesthetic procedures and/or implant history will also be presented and will follow the same methods specified above.

Medical history, and prior cosmetic/aesthetic procedures or implants will be provided in the subject data listing. Medical history will be further presented in a separate listing by past and ongoing GL medical history.

Concomitant medications for this study are defined as any ongoing medications with a start date prior to the date of injection, any changes to existing medications (such as dose or formulation) during the course of the study, or any new medications received by the subject since the date of injection. Prior medications are medications with stop dates prior to study treatment and used within 4 weeks preceding screening visit. Medications will be coded using the World Health Organization (WHO) Drug Dictionary; concomitant procedures will be coded according to MedDRA. The versions used for the coding will be noted as a footnote in the tables and listings.

The number and percentage of subjects who receive prior and concomitant medications will be summarized by the WHO Drug Dictionary Anatomical Therapeutic Chemical 3rd level (ATC-3) and the preferred name. If the 3rd level term is not available, the next available level (e.g., ATC-2)

will be used. In addition, the number and percentage of subjects reporting a concomitant medication/therapy will be summarized by reason (medical history, adverse event, concomitant procedure, contraception, or other). Therapies and procedures that started due to an AE will be summarized separately from those who did not start due to an AE.

ATC-3 and preferred name will be presented in descending frequency first based on the QM1114-DP group, and then alphabetically if there are ties. Each subject will contribute at most one count per summarization category. In other words, if a subject has more than one medication with same preferred name, the subject will be counted only once for that preferred name. Similarly, if a subject has more than one medication for an ATC-3 level, the subject will be counted only once in that ATC-3 level and preferred name. Concomitant procedures will also be presented and will follow the same methods specified for medical history.

Prior and concomitant medications/procedures will be presented by subject in a data listing.

Handling of Missing/Partial Dates

While every effort will be made to obtain full, complete information on every reported medication, the following imputation rules will be followed for any respective missing medication data:

For the purpose of determining whether a medication is considered prior or concomitant, the following date imputation rules will used. Dates will be presented as collected in the listings.

- Start Date
  - If the start date is completely missing, it will be assumed that the medication started on the study treatment date.
  - If the start date is missing the day, the first of the month will be used (i.e. UNK-JAN-2019 becomes 01-JAN-2019), provided the imputed date is on or after the subject's study treatment date; otherwise, the day of study treatment will be used.
  - If the start date is missing the month, the month of 'June' will be used (i.e. 01-UNK-2019 becomes 01-JUN-2019), provided the imputed date is on or after the subject's study treatment date; otherwise, the subsequent month after study treatment will be used.
  - If the start date is missing the year, the year of study treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is on or after the subject's study treatment date; otherwise, the subsequent year after treatment will be used.
- End Date
  - If the end date is completely missing, it will be assumed that the medication is still ongoing and will not be imputed.
  - If the end date is missing the day, the last day of the month will be used (i.e. UNK-JAN-2019 becomes 31-JAN-2019).
  - If the end date is missing the month, the subsequent month after the start date will be used.
  - If the end date is missing the year, the year of study treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is after the start date; otherwise, the subsequent year after start date will be used.

2.4 Efficacy analysis

2.4.1 Datasets analyzed

The primary efficacy analysis will be analyzed based on the mITT population. All secondary efficacy CCI [redacted] variables will be analyzed based on the ITT population, unless otherwise specified below.

2.4.2 Handling of missing data

In general, the number of subjects with missing values will be summarized and reported as appropriate in all outputs. The primary analysis will be performed using multiple imputation (MI) as the primary imputation method and repeated using baseline observation carried forward (BOCF) as a sensitivity analysis for missing values. With the BOCF imputation method, if a subject is missing their Month 1 GL-ILA score or GL-SLA Month 1 score, their GL-ILA/GL-SLA score at baseline will be used, respectively.

The imputation using MI will assume the Missing Completely at Random (MCAR) missing data assumption. Regardless of the actual pattern of missing data, the Markov Chain Monte Carlo (MCMC) method of the MI procedure from the SAS® system will first be used to create a monotonic pattern of missing data. Then, a second MI procedure will be used to generate five sets of data with missing values imputed from observed data. Linear regressions will be employed to model the missing GL-ILA and GL-SLA scores, separately, with the following covariates included in each imputation model: treatment group and non-missing GL-ILA or GL-SLA data, respectively, from earlier timepoints (baseline, Day 7, and Day 14). The imputed datasets will be used to create the composite responder variable, which will then be analyzed using the methodology described for the primary analysis of responder rates at Month 1 (Section 2.4.3). The results from the analysis of the multiple imputed datasets will be combined by the MIANALYZE procedure of the SAS® system. The seed number to be used will be the number 110196.

The secondary efficacy analysis (described in Section 2.4.4) will be performed using the Observed Cases (OC), that is, no imputation will be done. CCI [redacted]
CCI [redacted]

2.4.3 Primary analysis

The primary efficacy endpoint will be the composite responder rate based on the GL-ILA and GL-SLA of GL severity at maximum frown at Month 1 (described in Section 1.3.1 and Section 1.3.2, respectively). To evaluate the effectiveness of QM1114-DP versus placebo in the treatment of moderate to severe GL, the responder rates of the QM1114-DP and placebo will be compared using the Cochran-Mantel-Haenszel (CMH) test stratified by site at the 5% significance level (2-sided).

The null hypothesis of no relationship between treatment and responder rate (i.e. the responder rates are the same in both groups) will be tested against the alternative hypothesis that there is a relationship between treatment and responder rate (i.e. the responder rates are different in the two groups). For a significant result, the two-sided p-value of the comparison of the composite GL-ILA/GL-SLA responder rates between the treated and untreated subjects at Month 1 using the CMH test needs to be smaller than 0.05. For consistency across strata (site) the Breslow-Day test will be used to assess the homogeneity of the odds ratios across all sites.

The estimates of the composite GL-ILA/GL-SLA responder rates in each treatment group will be presented as well as the difference in responder rates (QM1114-DP responder rate – placebo responder rate). Corresponding 95% CI for the treatment group composite GL-ILA/GL-SLA responder rates and the difference in responder rates along with the p-value for the difference will also be presented. The normal approximation (Wald) method will be used to calculate both the 95% CI for the individual treatment group composite GL-ILA/GL-SLA responder rates and the 95% CI for the difference in responder rates. The above responder rates will be presented in figures by visit and treatment group.

2.4.3.1 Sensitivity analysis

To evaluate the impact of missing data on the primary endpoint, sensitivity analyses will be performed. The primary analysis specified in Section 2.4.3 will be repeated using the BOCF method (detailed in Section 2.4.2), and also using the ITT OC.

In addition, a sensitivity analysis of the primary efficacy endpoint will be performed based on the PP population, and also on the Safety population to account for potential errors in randomization. The analysis method described above in Section 2.4.3 will be repeated but using the PP population and Safety Population.

2.4.3.2 Subgroup analysis

Additionally, to evaluate the consistency of the results of the primary analysis across different subgroups of interest, the primary analysis specified above (Section 2.4.3) will be repeated, stratifying for each of the following subgroups specified below:

- Age (less than 65 years old, 65+ years old)
- Gender (Male, Female; if enough males are recruited)
- Baseline Severity score of the ILA at maximum frown (Grade 2 (Moderate), Grade 3 (Severe))
- Prior botulinum toxin use (Yes, No)
- Fitzpatrick skin type (I-III, IV-VI)
- Site
- Type of Month 1 visit (Onsite, Remote)

For each subgroup, the responder rates, difference in responder rates between the treatment groups, and the corresponding 95% CI will be presented. The Breslow-Day test will be used to assess the homogeneity of the odds ratios across all the subjects within each subgroup. The subgroup analysis for type of Month 1 visit will be performed based on the ITT population.

2.4.4 Secondary analysis

To evaluate the effectiveness of QM1114-DP versus placebo in the treatment of moderate to severe GL, the proportion of subjects who achieve 0 or 1 on the GL-ILA at maximum frown at all post-treatment visits will be compared using the CMH test stratified by site. To control the type I error rate, the fixed sequence testing procedure will be used, which requires no adjustment to the level of significance. However, the comparisons must be made in a pre-specified order following the procedure outlined below:

1. Primary analysis must be done first, prior to any other analysis

Secondary analysis:

2. Month 1 responder rate
3. Day 14 responder rate
4. Month 2 responder rate
5. Month 3 responder rate
6. Month 4 responder rate
7. Month 5 responder rate
8. Month 6 responder rate
9. Day 7 responder rate

If one of the tests in the sequence is not significant ($p > 0.05$), no confirmatory claims can be made based on tests following (and including) that one.

The secondary analysis in the sequence will be based on the ITT population with the remote visit assessments excluded. A sensitivity analysis will be run using the ITT population including the remote visit assessments.

Results will include estimates of the percentages, difference in percentages, 95% CIs, and CMH p-value as well as a graph showing the percentages for the two treatment groups over time. For consistency across strata, the Breslow-Day test will be used.

CCI

2.5 Safety analysis

All safety data will be summarized descriptively based on the safety population. All of the safety analyses will be done using OC. If deemed necessary, any analyses may be repeated using OC, BOCF, or MI as appropriate.

2.5.1 Extent of exposure

There will only be limited exposure to QM1114-DP since it is injected only once. The number of subjects receiving a single dose of QM1114-DP and placebo, respectively, will be presented.

2.5.2 Adverse events

All AE data will be summarized by treatment group. Missing dates will be imputed as described below. AEs will be summarized by SOC and PT. AEs occurring before treatment will be presented in listings only. The MedDRA version used for the coding will be noted as a footnote in the tables and listings.

A summary of all AEs will be provided, which will include:

- number (%) of subjects who did not have an AE
- number (%) of subjects with at least one TEAE and number of events
- number (%) of subjects with at least one TEAE related to study product or injection procedure
- number (%) of subjects with at least one TEAE not related to study product or injection procedure
- number (%) of subjects with at least one TEAE leading to discontinuation
- number (%) of subjects with at least one serious TEAE.

Summaries of TEAEs (including the total number of events, number and percentage of subjects) will be displayed by treatment group according to the following:

- All TEAEs by SOC and PT
- Treatment emergent SAE by SOC, PT, maximum intensity (mild, moderate, severe), and causality
- Related TEAEs by SOC and PT, and maximum intensity (mild, moderate, severe)
- Unrelated TEAEs by SOC and PT, and maximum intensity (mild, moderate, severe)
- TEAEs leading to discontinuation by SOC and PT, and maximum intensity (mild, moderate, severe)
- TEAEs by SOC, PT, and action taken (none, medical treatment, non-pharmacological treatment, subject withdrawn)

For the subject level analyses, the number and percentage of subjects who experienced at least one of the events listed above will be summarized overall and for each SOC and each PT. System organ class and PTs will be presented in descending frequency first (by QM1114-DP group), and then alphabetically if there are ties. Each subject will contribute at most one count per summarization category. In other words, if a subject has more than one TEAE with same PT, the subject will be counted only once for that PT. Similarly, if a subject has more than one TEAE for a SOC, the subject will be counted only once in that SOC and PT. For the event level analyses, the counts of each respective event will be presented. In general, percentages will be calculated using the number of subjects in the safety population for the denominator. For the subgroup analyses (described in Section 2.5.2.1) percentages will be calculated using the number of subjects in the safety population for each respective sub-category will be used for the denominator.

For the "action taken" summary specifically, subjects will be only counted in 'None' category if no other action was taken. The onset/duration summaries will be presented at the event level (i.e. will include multiple AEs within the same SOC and PT). Number of days to onset and duration of event will be summarized by SOC and PT, using mean, SD, minimum, maximum, and median statistics. Time to onset will be calculated as the first day with the AE *minus* the Day 0. Duration will be calculated as the last day with the AE *minus* the first day with the AE *plus* one. The text below lists the imputation rules for partial/completely missing dates. Missing stop date will not be imputed and therefore no duration will be calculated in these cases. Instead, the number of AEs that were ongoing at the end of the study will be given.

Handling of Missing/Partial Dates

While every effort will be made to obtain full, complete information on every reported AE, the following imputation rules will be followed for any respective missing AE data:

For the purpose of calculating treatment emergence, onset time, and duration, the following date imputation rules will used. Dates will be presented as is in the listings.

- Start Date
  - If start date is completely missing, it will be assumed that the AE started on the treatment date.
  - If the start date is missing the day, the first of the month will be used (i.e. UNK-JAN-2019 becomes 01-JAN-2019), provided the imputed date is on or after the subject's treatment date; otherwise, the day of treatment will be used.
  - If the start date is missing the month, the month of 'June' will be used (i.e. 01-UNK-2019 becomes 01-JUN-2019), provided the imputed date is on or after the subject's treatment date; otherwise, the subsequent month after treatment will be used.
  - If the start date is missing the year, the year of treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is on or after the subject's treatment date; otherwise, the subsequent year after treatment will be used.
- End Date
  - If end date is completely missing, it will be assumed that the AE is still ongoing and will not be imputed.
  - If the end date is missing the day, the last of the month will be used (i.e. UNK-JAN-2019 becomes 31-JAN-2019)
  - If the end date is missing the month, the subsequent month after the start date will be used.
  - If the end date is missing the year, the year of treatment will be used (i.e. 01-JAN-UNK becomes 01-JAN-2019), provided the imputed date is after the start date; otherwise, the subsequent year after start date will be used.

2.5.2.1 Subgroup analysis

To evaluate the consistency of the AE data, subgroup analyses will be performed. All TEAEs related to study product or injection procedure by SOC, PT, and maximum intensity will be repeated by the following subgroups. The same methods specified above (Section 2.5.2) will be followed

- Age (less than 65 years old, 65+ years old)
- Gender (Male, Female)
- Prior use of botulinum toxin (Naive, Non-Naive)
- Race (White, Black, Other)
- Fitzpatrick skin type (I-III, IV-VI)
- Site

2.5.2.2 Assessment of local/remote toxin spread and hypersensitivity

Any potential or suspected toxin spread or toxin hypersensitivity events will be evaluated separately. The same methods specified in Section 2.5.2 will be followed. Suspected toxin spread events and suspected hypersensitivity events will be summarized in separate tables. The investigator will conduct the investigation and evaluate if an AE is also a suspected toxin spread or hypersensitivity event; however, Appendix 6 in the CSP lists AEs potentially suggestive of spread of toxin.

2.5.3 Laboratory assessments

The laboratory data (hematology and clinical chemistry) at baseline Day 0 and Month 6 and the changes from baseline will be summarized by descriptive statistics. All clinically significant out-of-range laboratory values for blood samples collected at baseline will be recorded in the subject's medical history and all clinically significant out of range laboratory values for blood samples collected after baseline are to be reported as an AE if this abnormality was not present at the baseline visit or is assessed as having worsened since the baseline visit.

Throughout the study it is possible that certain laboratory parameter results are below the limit of quantification (BLQ). Any BLQ laboratory values will be summarized by the imputed value of half the limit of quantification for that parameter.

All laboratory data will be presented in data listings as collected, without any imputations for BLQ values.

2.5.4 Neutralizing antibody testing

Sampling for antibody testing will be conducted at baseline Day 0, Month 1, and Month 6. For the antibody testing results, due to the timing and availability of the data, results will be presented in a separate report.

2.5.5 Physical examinations and vital signs

The number and percentage of subjects with normal/abnormal results in physical examination will be presented by visit and treatment. A shift table will be created to present any change from baseline in normal/abnormal results in physical examination across the study visits for each treatment group.

Vital signs at baseline, Day 7, Day 14, Month 1, and Month 6, and the changes from baseline will be summarized by treatment and visit using descriptive statistics.

2.5.6 ECG

The aggregated value of each ECG measurement (described in Section 1.5.6) at baseline, Month 1, Month 6, and the changes from baseline will be summarized descriptively by treatment group at each visit. In addition, a shift table will be created to show any change from baseline to Month 6 in the overall ECG assessment.

QTcF and QTcB interval prolongation will be presented separately as well. The number and percentage of subjects with a QTcF and/or QTcB interval of >450 msec, >480 msec, and >500 msec will be summarized by treatment and visit, along with the number and percentage of subjects with a change from baseline in QTcF and/or QTcB interval of 30 - < 60 msec and ≥60 msec.

2.5.7 Urine pregnancy test

A data listing will be provided to summarize the results of the urine pregnancy tests. All pregnancy SAEs will be flagged within the data listing.

2.6 Interim Analysis

Not applicable.

2.7 Determination of Sample Size

The sample size is determined by the number of subjects exposed to QM1114-DP in the GL, and amount of long-term safety data. Further, due to the public health emergency related to the COVID-19 pandemic during 2020, steps have been taken to ensure patient and practitioner safety in alignment with FDA Guidance dated May 11, 2020 (Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency). Most notably, in partnership with clinical sites and the IRB, optional remote assessment procedures for efficacy and safety endpoints has been implemented to ensure safety and respect localized and elective restrictions. However, the number of subjects in the mITT population used in the assessment of the primary efficacy endpoint, should still satisfy the sample size requirements for the primary efficacy analysis.

CCI [redacted]

CCI [redacted]

2.7.3 Sample size calculation

The study is planned to include approximately 300 subjects, of whom 225 will be treated with QM1114-DP and 75 subjects will receive placebo. CCI [redacted]

2.8 Changes in Analysis Planned in the Protocol

The table below outlines all rationale and changes in the analyses specified in the SAP that differ from the analyses specified in the protocol.

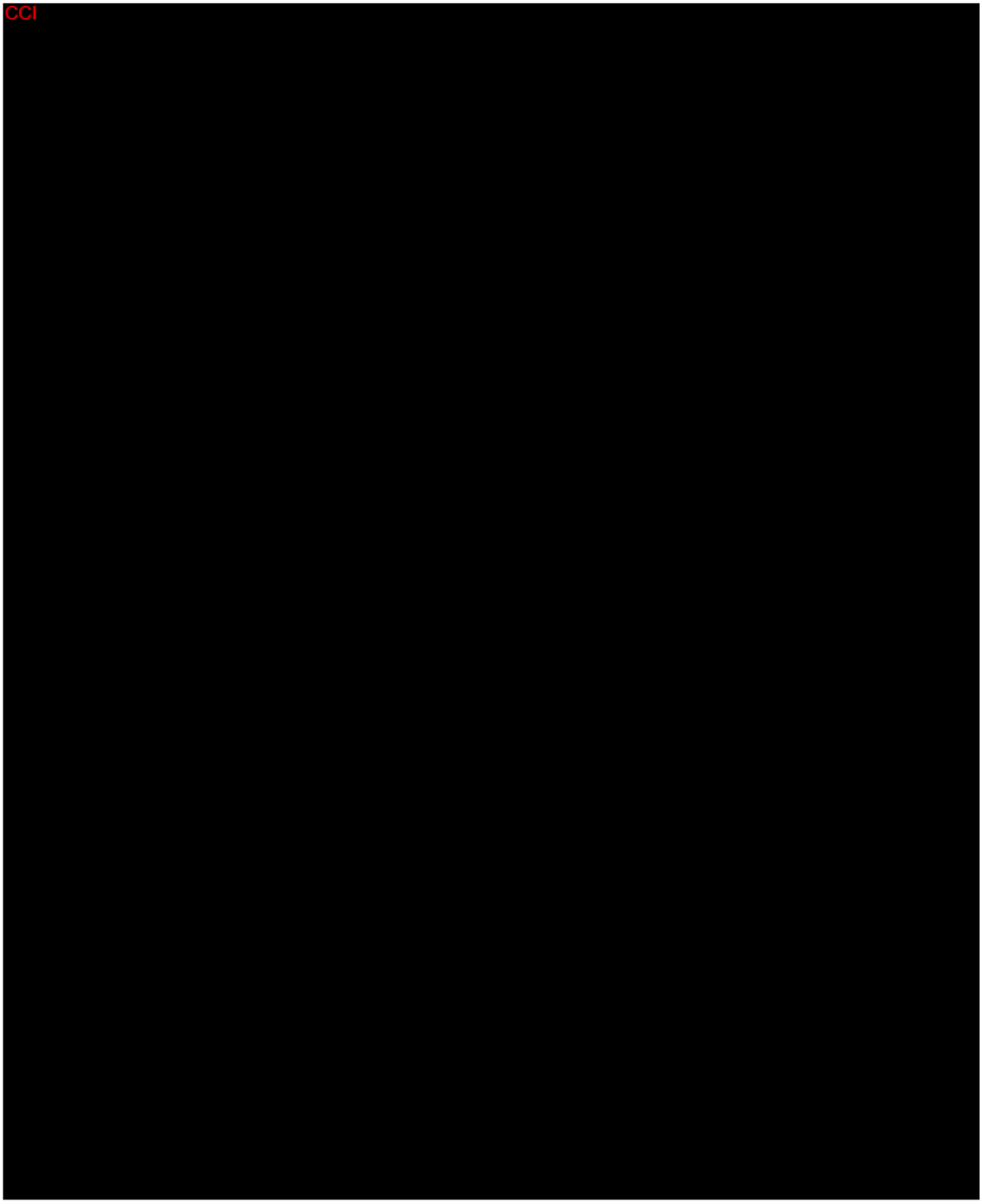
Table 6 Changes in Analysis Planned in the Protocol

CSP Section	SAP section	Description/Rationale of Change
9.1.2.3 Per-Protocol (PP) Efficacy Population	2.4.3.1 Sensitivity Analysis	Regardless of if the PP population contains less than 90% of the mITT population, sensitivity analyses for the primary efficacy endpoint using the PP population will be conducted.

3 Reference List

4 Appendix A GL-ILA Scales

CCI



Maximal Frown

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

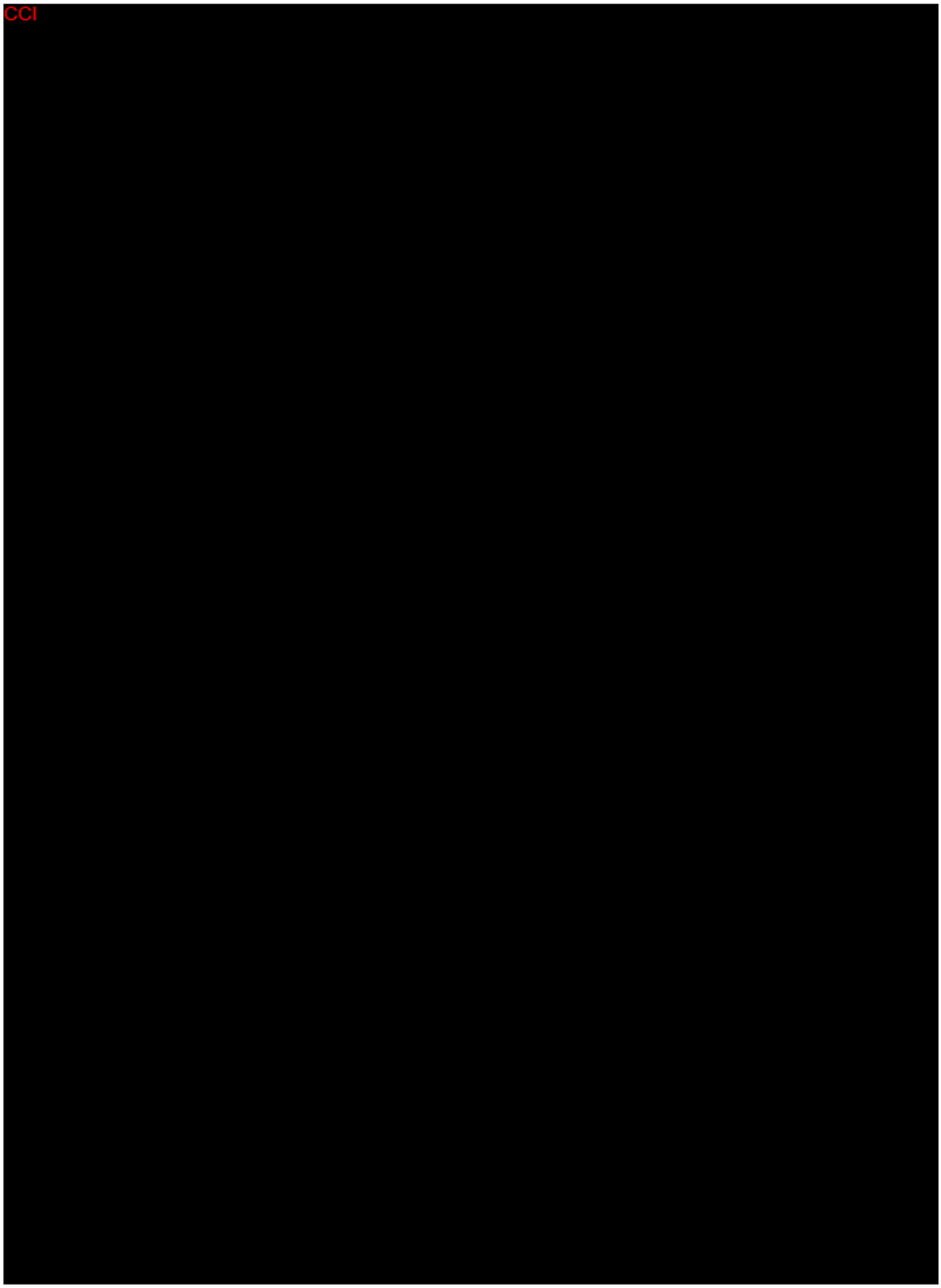
CCI

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

CCI

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

CCI



GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

CCI

8 Appendix E Schedule of Assessments

1 month = 4 weeks/28 days All visit windows are calculated from Baseline/Day 0	Visit 1 Screening[1]	Visit 2 Baseline/Day 0[1] (within 2 weeks after screening)	Visit 3 Day 7 (±1 day)	Visit 4 Day 14 (± 3 days)	Visit 5 Month 1 (±5 days)	Visit 6 Month 2 (±5 days)	Visit 7 Month 3 (±5 days)	Visit 8 Month 4 (±5 days)	Visit 9 Month 5 (±5 days)	Visit 10 Month 6/ET[2] (±5 days)
Informed Consent	X									
Demographic Data[3] including, Fitzpatrick skin type, medical history & concurrent diseases, previous facial treatments/procedures (toxin naïve/non-toxin naïve)	X									
Inclusion /Exclusion Criteria	X	X[4]								
Concomitant Therapies/ Procedures	X	X[5]	X	X	X	X	X	X	X	X
Adverse Events	X	X[5]	X	X	X	X	X	X	X	X
Urine Pregnancy Test[6]	X	X[4]								X
Vital Signs (blood pressure, heart rate, and respiratory rate)[7]		X[5]	X	X	X					X
ECG		X[4]			X					X
Blood sample clinical chemistry and hematology		X[4]								X
Blood sample for serum antibody testing		X[4]			X					X
ILA of GL Severity	X	X[4]	X	X	X	X	X	X	X	X
CCI										
Photography		X[4]	X	X	X	X	X	X	X	X
Randomization		X[4]								
Treatment		X[9]								
CCI										
Subject Assessments										
SLA of GL Severity[10]	X	X[4]	X	X	X	X	X	X	X	X
CCI										

1. *Screening and baseline visits may be on the same day. If completed on the same day, only perform study assessments once (i.e., PE, UPT, SLA, ILA, AE, concomitant therapies/procedures, inclusion/exclusion review).*
2. *If the subject withdraws before the final visit the assessments at Month 6/ET should be completed, if possible.*
3. *Includes date of birth, gender, race, ethnicity, height, weight.*
4. *To be performed before treatment.*
5. *Performed before treatment and post-treatment.*
6. *Females of childbearing potential*
7. *Vital signs are taken seated after 10 minutes rest. Vital signs are taken prior to any blood draw (excluding post-*

CCI

9. *Following treatment administration, subjects will be monitored at the study center for 30 minutes.*
10. *Subjects will make their GL assessments independently of the Investigator's assessment.*

PPD

Statistical Analysis Plan Approval Form

Client:	Galderma	
Identifier (e.g., Protocol Name and Number, ISS, ISE, SCS, or CSE):	43QM1602	
Version (Protocol Only):	☐ Original, Date:	☒ Amendment, Number 4.0; Date: 10JUN2020
Title:	A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of QM1114-DP for the Treatment of Moderate to Severe Glabellar Lines (READY - 1	
Statistical Analysis Plan Version:	☒ Original, Date: 27OCT2020	☐ Amendment, Number ; Date:

AC Statistical Analysis Plan Author:

PPD
Senior Biostatistician
Printed Name & Title

Handwritten Signature **OR** *Electronic Signature*

Signature: _____

Date: _____
(dd/mmm/yyyy)

PPD

DocuSign

AC Statistical Analysis Plan Reviewer:

PPD
Senior Biostatistician
Printed Name & Title

Handwritten Signature **OR** *Electronic Signature*

Signature: _____

Date: _____
(dd/mmm/yyyy)

PPD

DocuSign

First Client Statistical Analysis Plan Approval:

PPD
Printed Name & Title

PPD
Signature: _____ *Handwritten Signature*

Date: PPD
(dd/mmm/yyyy)

Second Client Statistical Analysis Plan Approval ☐ Not Applicable

PPD
Printed Name & Title

PPD
Signature: _____ *Handwritten Signature*

Date: PPD
(dd/mmm/yyyy)

GALDERMA	Title 43QM1602 Statistical Analysis Plan US - QM1114-DP - GL	Doc id MA-30369

SIGNATURES PAGE

Date	Signed by
2020-11-27 15:58	PPD
Justification	Approved by Technical Expert

2020-11-30 08:14	PPD
Justification	Approved by Owner

2021-02-05 13:17	PPD
Justification	Approved by