Galderma Research and Development, LLC

STATISTICAL ANALYSIS PLAN

Clinical Trial Number (CTN) 43USCH2208

Protocol V4.0 (24MAY2023)

A randomized, evaluator-blinded, parallel group, comparator-controlled, multicenter study to evaluate the safety and effectiveness of Restylane Lyft with Lidocaine for augmentation of the chin region to improve the chin profile

SPONSOR

Q-Med AB
Seminariegatan 21
SE-752 28 Uppsala, Sweden

Version Final 1.0
28 May 2024

Confidential

APPROVAL SIGNATURE PAGE

Protocol Title: *A randomized, evaluator-blinded, parallel group, comparator-controlled, multicenter study to evaluate the safety and effectiveness of Restylane Lyft with Lidocaine for augmentation of the chin region to improve the chin profile*

Sponsor: Q-Med AB
Seminariegatan 21
SE-752 28 Uppsala, Sweden
Telephone: +46 18 474 90 00

Protocol Number: 43USCH2208

PPD

By signing this document, I acknowledge that I have read the document and approve of the planned statistical analyses described herein. I agree that the planned statistical analyses are appropriate for this study, are in accordance with the study objectives, and are consistent with the statistical methodology described in the protocol, clinical development plan, and all applicable regulatory guidances and guidelines.

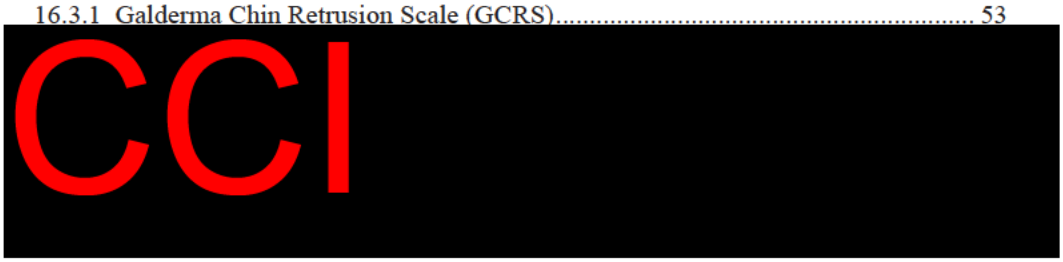
I have discussed any questions I have regarding the contents of this document with the biostatistical author.

I also understand that any subsequent changes to the planned statistical analyses, as described herein, may have a regulatory impact and/or result in timeline adjustments. All changes to the planned analyses will be described in the clinical study report.

PPD

Table of Contents

[redacted]



Tables

1 List of Abbreviations and Definition of Terms

Abbreviation	Description
AE	Adverse Event
AESI	Adverse Event of Special Interest
AIC	Akaike Information Criterion
ATC	Anatomical Therapeutic Chemical
ATC-3	Anatomical Therapeutic Chemical 3rd level
ATC-4	Anatomical Therapeutic Chemical 4th level
BE	Blinded Evaluator
BOCF	Baseline Observation Carried Forward
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence Interval
CTN	Clinical Trial Number
EOS	End of Study
eCRF	Electronic Case Report Form
FSFV	First Subject First Visit
FST	Fitzpatrick Skin Type
CCI	[redacted]
GCRS	Galderma Chin Retrusion Scale
ICF	Informed Consent Form
ICH	International Conference on Harmonization
CCI	[redacted]
IRB	Institutional Review Board
CCI	[redacted]
ITT	Intention-to-Treat
KM	Kaplan-Meier
KR	Kenward-Rodger
LS	Least Squares
LSLV	Last Subject Last Visit
MAR	Missing at Random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
ML	Maximum Likelihood
NPS	Numeric Pain Scale
OC	Observed Cases
PRO	Patient-Reported Outcome
PP	Per-Protocol
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SOC	System Organ Class
CCI	[redacted]

Abbreviation	Description
TI	Treating Investigator
UADE	Unanticipated Adverse Device Effect
VA	Visual Acuity
WHO	World Health Organization

2 Introduction

This statistical analysis plan (SAP) describes the effectiveness and safety summaries and analyses that will be performed for Clinical Trial Number (CTN) 43USCH2208, *A randomized, evaluator-blinded, parallel group, comparator-controlled, multicenter study to evaluate the safety and effectiveness of Restylane Lyft with Lidocaine for augmentation of the chin region to improve the chin profile* and is based on the study protocol Version 4.0 dated 24MAY2023 and electronic Case Report Form (eCRF) Version 4.03 dated 14FEB2024.

3 Study Objectives and Overall Design

3.1 Study Objectives and Endpoints

3.1.1 Primary Effectiveness Objective and Endpoint

The primary objective of the study is to demonstrate non-inferiority of Restylane Lyft with Lidocaine versus Juvéderm® Voluma™ XC for augmentation of the chin region to improve the chin profile by comparing change from baseline in the Blinded Evaluator live assessment of the Galderma Chin Retrusion Scale (GCRS) at 3 months after last treatment.

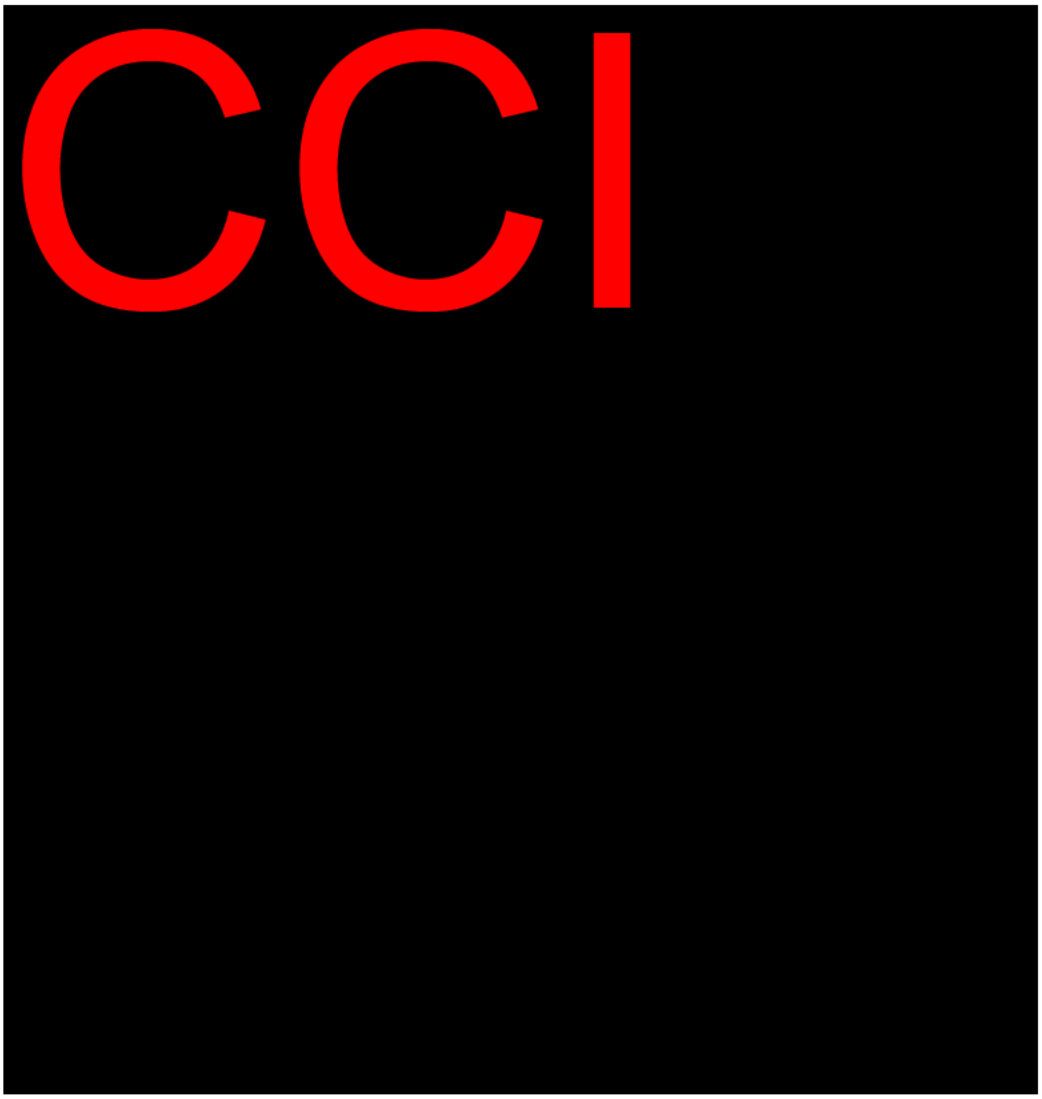
3.1.2 Secondary Effectiveness Objective and Endpoints

The secondary objective is to evaluate the effectiveness of Restylane Lyft with Lidocaine versus a comparator-control for augmentation of the chin region to improve the chin profile.

This objective will be based on the following endpoints:

1. Responder rate based on the Blinded Evaluators' (BE) live assessment using the GCRS at 3, 6, 9 and 12 months after last treatment.

   A responder is defined as a subject with at least 1 grade improvement from Baseline on the GCRS.



3.2 Overall Design

This is a prospective, randomized, evaluator-blinded, parallel group, comparator-controlled, multicenter study in the US to evaluate the safety and effectiveness of Restylane Lyft with Lidocaine for augmentation of the chin region to improve the chin profile. The chin region is comprised of the anatomical areas of the pogonion, menton, and pre-jowl sulcus. Before the study starts, a randomization list will be prepared under the supervision of a designated biostatistician from the Sponsor.

Injection technique is at the Treating Investigator's discretion with needle only or a combination of needle and cannula. Restylane Lyft with Lidocaine and the comparator-

control will be administered by needle in the pogonion. The mentum and pre-jowl sulci may be injected by needle or cannula.

At least 35 subjects will be Fitzpatrick skin type (FST) IV through VI, this includes at least 18 subjects with FST V – VI, where at least 9 subjects will be FST V and at least 9 subjects will be FST VI.

After providing informed consent, eligible subjects randomized to receive either Restylane Lyft with Lidocaine or a comparator-control will be injected by the Treating Investigator at Baseline (Day 1). A follow-up telephone call should be made 72 hours after treatment and a follow-up visit should be scheduled after 14 days.

Optional touch-up treatment may be administered 1 month (4 weeks) after initial treatment, if deemed necessary by the Treating Investigator and the subject to obtain optimal aesthetic improvement. If optional touch-up is performed, a 72-hour follow-up telephone call, a follow-up visit at 14 days, and a follow-up visit at 1 month should be scheduled.

Appropriate injection volume for the chin area will be determined by the Treating Investigator but is not to exceed a maximum total volume of 4.0 mL for initial and touch-up treatments combined. Injection in the pogonion is only permitted with a needle. Both needle and cannula are permitted for injection in the other treatment areas.

Investigator blinding will be accomplished by using a Treating Investigator to administer the treatments and a Blinded Evaluator, to whom randomization and treatment are concealed, to conduct the blinded assessments. Safety assessments will be performed by non-blinded personnel.

Effectiveness and safety data will be collected for up to 12 months (48 weeks) after last treatment including physical follow-up visits at 1, 3, 6, 9 and 12 months as well as 14 days after each treatment. A subject will be involved in the study for up to 14 months, including a 21-day screening period.

3.2.1 Randomization Methodology

Approximately 174 subjects will be randomized (2:1) to treatment with either Restylane Lyft with Lidocaine (Treatment Group) or Juvéderm® Voluma™ XC (Control Group).

Before starting the study, a computer-generated randomization list will be prepared under the supervision of a designated biostatistician from the Sponsor. The randomization will be stratified by FST (I-III, IV or V-VI). Subjects in the FST I-III stratum will be further stratified by study site; subjects in the FST IV, or FST V-VI strata will not be further stratified by study site due to the smaller sample size in these groups.

Randomization will be performed using an Interactive Response System by assigning each subject to treatment with Restylane Lyft with Lidocaine or Juvéderm Voluma XC

according to the randomization list. Randomization numbers will be allocated in ascending sequential order to each subject.

3.2.2 Study Subjects

A total of 174 adult male and female subjects over the age of 21 years with the intent to undergo treatment for augmentation of the chin region, is planned to be included in this study.

To minimize potential bias in the study results, 35 subjects will be the maximum allowed to be enrolled at any one study site.

3.2.3 Sample Size Justification

Approximately 174 subjects will be treated in the study, randomized in a 2:1 ratio (116 with Restylane Lyft with Lidocaine: 58 with Juvéderm Voluma XC).

For sample size calculation of 174, it was assumed that the true effectiveness of Restylane Lyft with Lidocaine compared to control treatment assessed by the GCRS is approximately 0.1 point lower. A common standard deviation of 0.7 point was assumed for the treatments. Based on the non-inferiority margin of 0.5, a total sample size of 147 evaluable subjects is then required to show non-inferiority of Restylane Lyft with Lidocaine to the control treatment with a power of 90%, using a two-sided 95% CI for the difference between treatment means.

The GCRS scale is designed so that a complete scale step considers a meaningful clinical difference. A change of up to 0.5 points is not considered clinically important, thus, the non-inferiority margin of 0.5 is chosen. We have estimated the standard deviation from previous studies where the scale GCRS has been used.

Accounting for approximately 15% dropouts and non-evaluable subjects due to protocol deviations at 3 months, a total of approximately 174 subjects will be randomized in the study.

3.2.4 Effectiveness Assessments

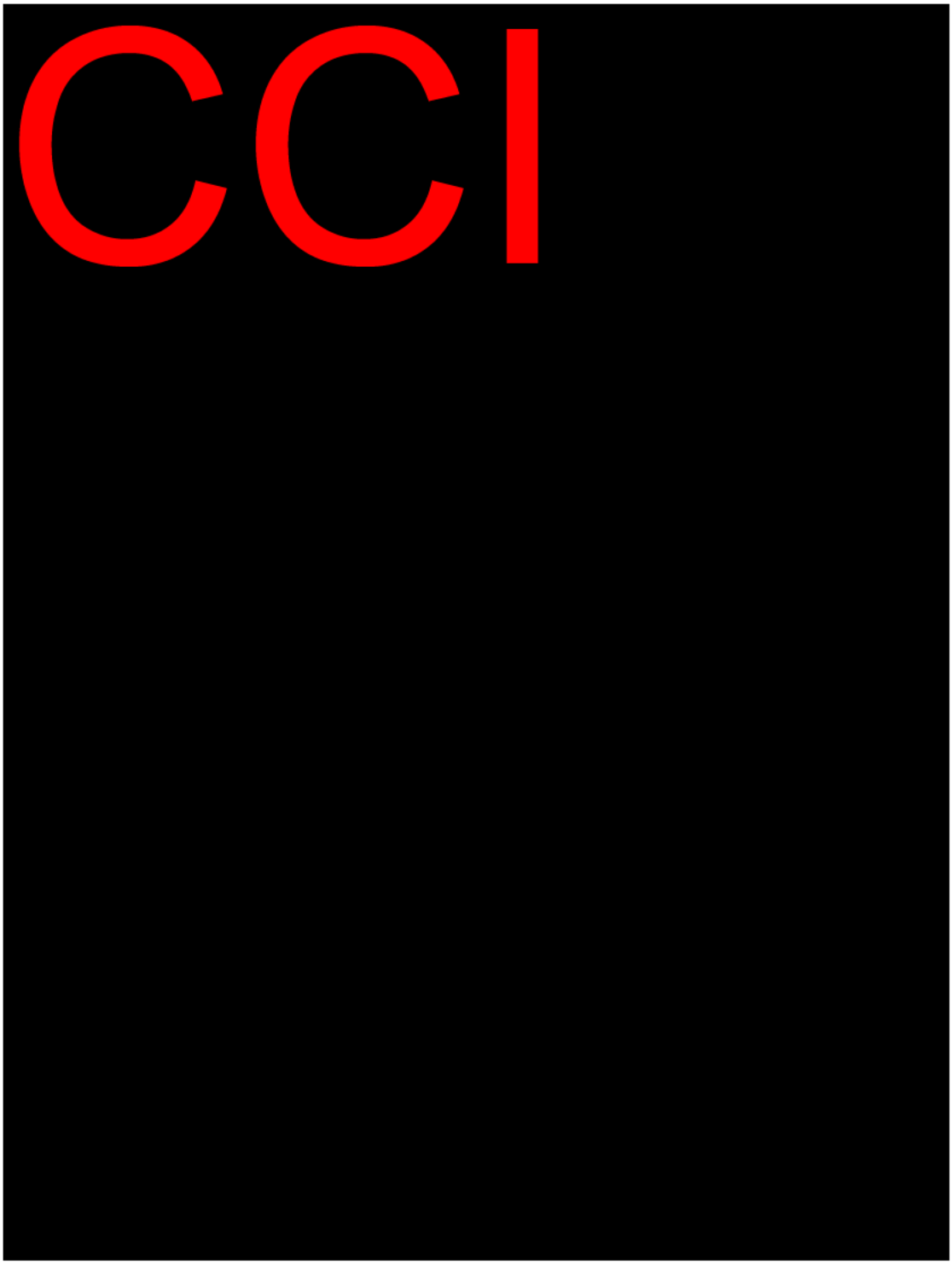
3.2.4.1 Galderma Chin Retrusion Scale (GCRS)

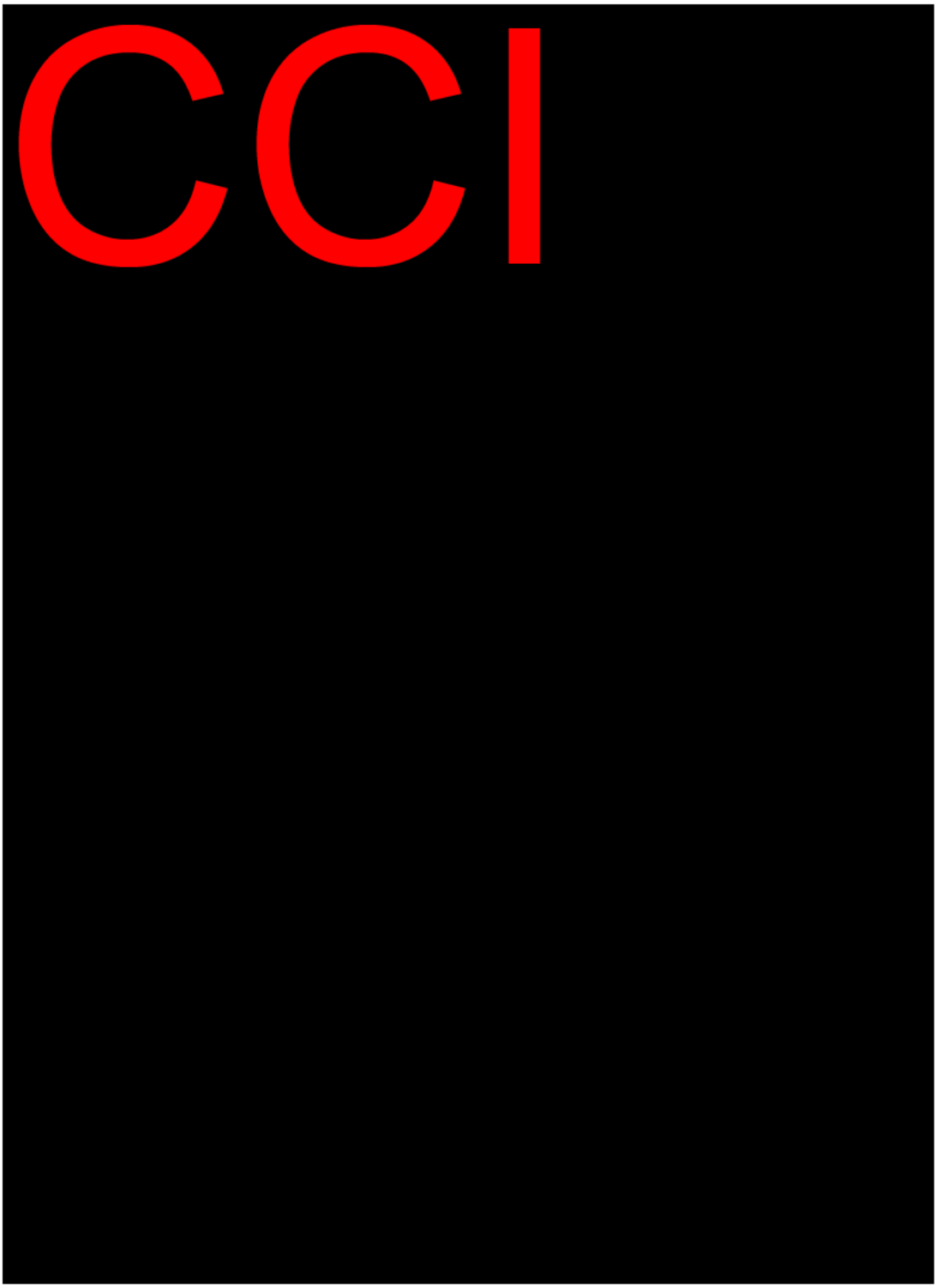
The GCRS is a validated 4 grade scale for assessment of Chin retrusion as described in Table 1 and detailed in Section 16.3.1. Each score in the GCRS is exemplified by photographic images of the scale. The Blinded Evaluator and Treating Investigator will perform a live assessment of the subject's chin separately. The Blinded Evaluator will assess the GCRS at Screening, Baseline and 3, 6, 9 and 12 months after last treatment. The Treating Investigator will assess the GCRS at Screening, Baseline and 1 month after initial treatment. On visits where a treatment is performed, assessments will be done prior to any planned treatment.

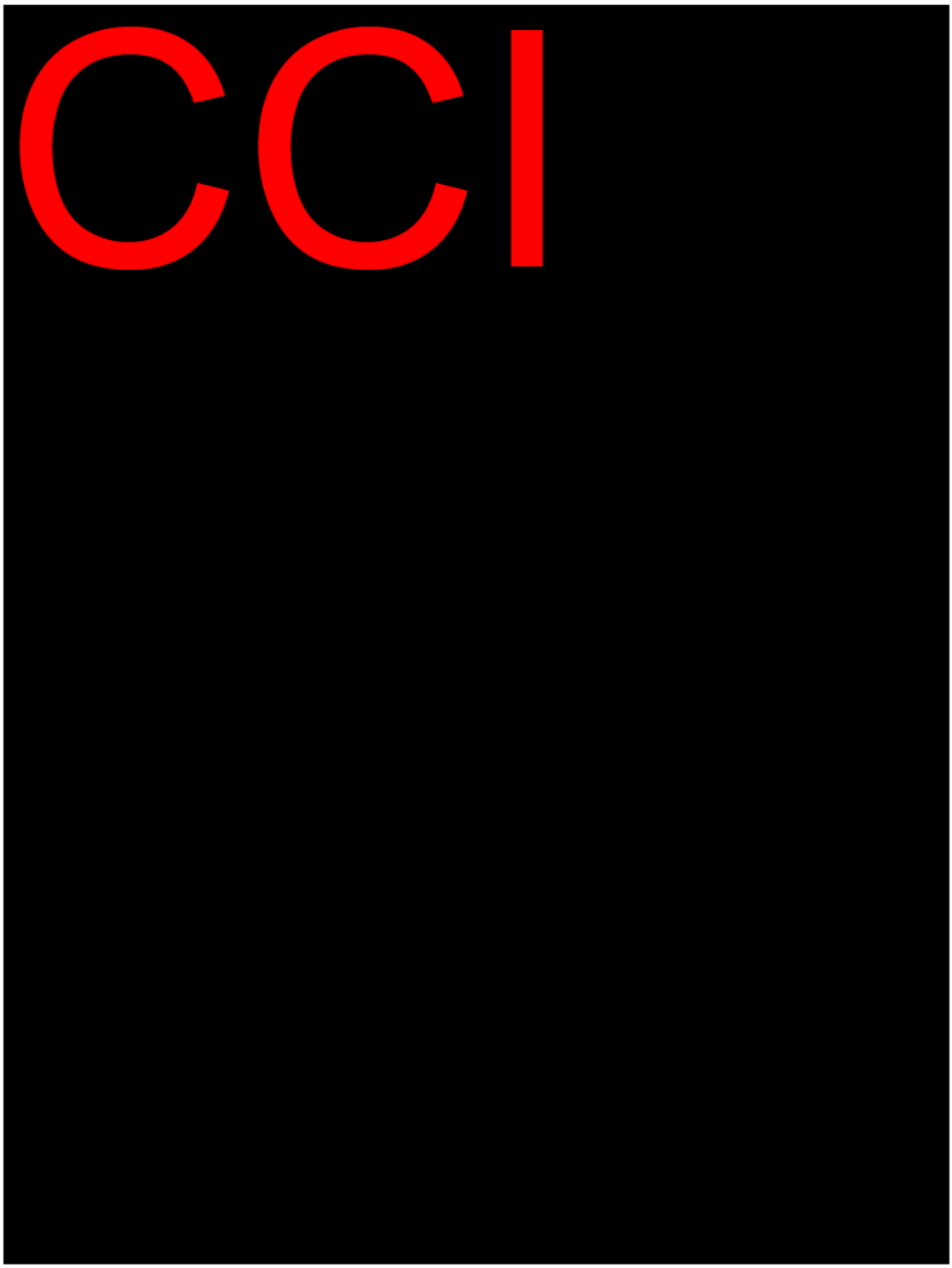
Table 1: Galderma Chin Retrusion Scale (GCRS)

Score	Description
0	No Retrusion: The most anterior portion of the chin is at or near a vertical line drawn from the vermillion border of the lower lip.
1	Mild Retrusion: The most anterior portion of the chin is clearly recessed, but less than midway, between vertical lines drawn from the vermillion border of the lower lip and the oral commissure.
2	Moderate Retrusion: The most anterior portion of the chin is recessed approximately midway between vertical lines drawn from the vermillion border of the lower lip and the oral commissure.
3	Severe Retrusion: The most anterior portion of the chin is clearly posterior to the midway point between vertical lines drawn from the vermillion border of the lower lip and the oral commissure.

A responder indicator will be created to identify subjects that achieved at least 1 grade improvement from Baseline, on the Blinded Evaluator GCRS at a given post-Baseline visit.







3.2.4.9 Returning to Social Engagement

A follow-up telephone call should be made 72 hours after each treatment, where the treated subject will be asked when (in hours after treatment) s/he felt comfortable returning to social engagement, such as returning to the office or other public workplace, having dinner in a public restaurant, attending a social event/gathering such as a dinner party, etc. CCI [redacted]

For each subject, time, in hours, until the subject feels comfortable returning to social engagement will be computed. A separate flag for each treatment will be created to determine the time in hours at which the subject responded they felt comfortable returning to social engagement. Time, in hours, until the subject feels comfortable returning to social engagement will be calculated as follows:

Number of hours until subject feels comfortable returning to social engagement = Date and Hour of Response Flag – Date and Hour of Treatment

Subjects who never feel comfortable returning to social engagement will be censored at 23:59 on the date of their telephone call visit (Visit 2a for initial treatment, Visit 3a for optional touch-up treatment). Subjects who felt comfortable returning to social engagement, but who provided a date, and no time will be censored at 23:59 on the date provided. Subjects who felt comfortable returning to social engagement, but who provided a time, and no date will be censored at the provided time on the date of their telephone call visit. Subjects who felt comfortable returning to social engagement, but who did not provide a date and time will be censored at 23:59 on the date of their telephone call visit. Subjects who never respond to the 72-hour phone call will be censored at 23:59 on the target date of their telephone call visit. Subjects dropping out before feeling comfortable returning to social engagement will also be censored at 23:59 on the target date of their telephone call visit or date of withdrawal, whichever comes first. The target date can be calculated by taking the date of treatment + 72 hours.

3.2.4.10 3D Imaging of the Chin region

The change in topography of the chin region from Baseline will be assessed by 3D imaging to 3 and 12 months after last treatment. Assessments may include, but will not be limited to, Pogonion Projection, Total Volume Change, Positive Volume Change, and Negative Volume Change.

CCI [redacted]

3.2.5.2 Sensory tests

The sensory tests will be performed at screening to confirm eligibility. A baseline value of sensation, evaluating the chin and lower lip, will be obtained and new assessments will be done at all physical visits thereafter. A study staff member who is qualified by training and experience to perform safety assessments will score sensation as "Normal" or "Abnormal."

Chin and lower lip sensation will be tested using two methods: the monofilament test and the cotton wisp test.

3.2.5.3 Functionality tests

The functionality tests will be performed at screening to confirm eligibility. A baseline value of lip, and lip and chin movement will be obtained, and new assessments will be done at all physical visits thereafter. A study staff member who is qualified by training and experience to perform safety assessments will score lip movement and lip and chin function as "Normal" or "Abnormal." After receiving treatment, any subject with abnormal ratings should be assessed by the Investigator for confirmation of any potential AE.

3.2.5.4 Assessment of palpability

Palpability will be assessed at screening, baseline and at each physical follow-up visit, prior to treatment. A study staff member who is qualified, by training and experience, to perform safety assessments will palpate the chin for detection of any unexpected feeling such as lumpiness, mass, non-uniform density, or unexpected aggregation of product. The study products are generally palpable (i.e., can be felt under the skin) and have uniform density. Any unexpected feel of the study product upon palpation of the chin will be assessed as "Abnormal". Any abnormality should be assessed by the investigator for confirmation of anu potential AE, including a description of the localization, feel and approximate size.

3.2.5.5 Changes in hair growth

The Treating Investigator will ask if the subject has noticed any changes in hair growth (e.g., loss or growth) in the treated area at physical follow-up visits after baseline. Clinically significant changes in hair growth, as determined by the Treating Investigator, will be reported as an AE.

3.2.5.6 Visual Function Assessment

Three visual function assessments will be performed at screening, baseline and at all physical follow-up visits. At the treatment visits, the visual function assessment tests will be performed prior to and approximately 30 minutes post injection of the study product.

The sections below provide descriptions of the visual function assessments that will be performed.

3.2.5.6.1 Snellen Visual Acuity

A Snellen eye chart will be used to objectively assess visual acuity for distance vision. Visual acuity will be conducted using the subject's best distance correction (e.g., contacts or eyeglasses) at a distance of 6 feet from the chart. Each eye will be measured separately by either using an occluder or having the subject cover each eye. The subject will be asked to start reading the letters at the top of the chart working their way to the bottom. The smallest row of letters that the subject can read will indicate their visual acuity listed on the chart. Any subject requiring vision acuity correction must be assessed using the prescribed correction (e.g., contacts or eyeglasses).

Unacceptable vision for a subject to be eligible for enrollment would be an absolute score on the Snellen eye chart 20/50 or worse in one or both eyes.

A 1+ worsening line change in Snellen Visual Acuity (VA) is when a subject moves up at least one line of letters on the Snellen Eye Chart (Figure 1). For example, if a subject's VA at baseline is 20/40 for the left and right eyes and then at Visit 1d is 20/70 for the left eye and 20/50 in the right eye, the subject experienced a 2-line worsening for the left eye and a 1-line worsening in the right eye.

Figure 1: Snellen Eye Chart

Given that VA is assessed at every study visit (pre-and-post treatment for applicable visits), there are many ways to assess a 1+ worsening line change:

- Acute worsening – A 1+ worsening line change between the pre-treatment and post-treatment VA, for visits where an injection is given.
- Compared to Previous visit – A 1+ worsening line change between the current visit VA and the previous visit VA (e.g., Visit 4 VA compared to Visit 3 VA). If the 'previous' visit has both pre-and-post treatment VA values, the post-treatment VA will be used; if the 'current' visit has both pre-and-post treatment VA values, both VAs will be compared to the 'previous' visit VA.
- Change from Baseline – A 1+ worsening line change between a post-baseline visit VA and Baseline (pre-treatment) VA. For post-baseline visits with both pre-and-post treatment VA, both will be compared to the Baseline (pre-treatment) VA.

The Snellen chart provided to sites for use in this study is 6ft. If a subject uses a different distance Snellen chart and their scores from the non-6ft Snellen chart match values found on the 6ft Snellen chart, then the 6ft Snellen chart equivalent will be used, and the subject will continue the study using the 6ft Snellen chart. For example, if a subject using a non-6ft Snellen chart scores values of 20/20, 20/25, 20/30, 20/40, 20/50, 20/70, 20/100, or 20/200 (all of which are found on the 6ft Snellen chart), then the subject's Snellen scores will remain the same, but the subject will proceed with the study using the 6ft Snellen chart.

On the other hand, if a subject uses a different distance Snellen chart and their scores from the non-6ft Snellen chart *do not consistently* match values found on the 6ft Snellen chart, then the subject will continue using the non-6ft Snellen chart for the duration of the study for safety purposes, and their observed Snellen values will remain the same in the data and data listings. In outputs presenting Snellen line changes, the line change calculation using the subject's non-6ft Snellen chart will be performed to determine the number of line changes separately for each eye (left and right). For example, if at any point a subject using a non-6ft Snellen chart scores values of 20/16, 20/60, 20/80, etc., then the data and data listings will present these non-6ft values. In the Snellen line change outputs, the type of Snellen chart will be considered when calculating line changes.

At the time this SAP was drafted, one site (Dr. Rivkin's site 8651) was identified as having used a non-6ft Snellen chart. For subjects at site 8651 with Snellen chart scores that do not match the scores on the 6-ft Snellen chart (i.e. any score other than 20/20, 20/25, 20/30, 20/40, 20/50, 20/70, 20/100, or 20/200), the following Dr. Rivkin 10-ft Snellen chart will be considered when calculating line changes.

Figure 2: Dr. Rivkin Site 8651 10ft Snellen Chart

3.2.5.6.2 Extraocular Muscle Function

Extraocular muscle function testing examines the function of the eye muscles. This test observes the movement of the eyes in six specific directions to evaluate weakness or other problems in the extraocular muscles. The subject will be asked to sit or stand with his/her head up and looking straight ahead. The assessor will hold a pen or other object approximately 16 inches in front of the subject's face. The assessor will then move the object in several directions and ask the subject to follow it with their eyes, without moving their head. The result shall be reported as "normal" or "abnormal" for each eye.

3.2.5.6.3 Confrontation Visual Field

The confrontation visual field test will be used to assess the subject's peripheral vision. The examiner will sit facing the subject, about 3 to 4 feet away from the subject and ask the subject to fix their gaze on the examiner's eye. The examiner will hold their arms straight out to the side and bring their hands into the subject's visual field from the sides in each quadrant. The subject will signal as soon as the hand is seen. Each eye will be recorded as normal or abnormal.

CCI

4 General Analysis Conventions

Data for all subjects in the clinical database will be included in the data listings and sorted by treatment and subject. Calculated (derived) variables will be included in listings when appropriate. All summaries, statistical analyses, and individual subject data listings will be completed using Version 9.4 or later of SAS software (SAS Institute, Inc., Cary, North Carolina). A list of proposed statistical tables and data listings will be provided in a separate document. Any changes from the SAP will be detailed in the Clinical Study Report.

Unless otherwise specified, summary tables will be presented with different columns for each treatment (Restylane Lyft with Lidocaine or Juvéderm Voluma). An additional column for overall summaries will be presented for non-efficacy endpoints. Continuous endpoints will be summarized using the number of subjects (n), mean, median, SD, minimum and maximum values. Categorical endpoints will be presented in frequency tables with number and percentage of observations for each level.

Unless otherwise specified, confidence intervals (CIs) will be two-tailed and constructed at a confidence level of 95%, statistical tests will be performed at a significance level of 5%, and p-values will be two-sided.

Study days will be calculated relative to the first injection of study drug. Day 1 will be the first day of study drug administration in the study, and the day prior to the first dose of study drug administration will be Day -1. There will be no Day 0.

Adverse events, medical history, prior dermatological procedures and implant history, and concomitant procedures/treatments data will be coded using Medical Dictionary for Regulatory Activities (MedDRA) Version 25.1. Concomitant medications, procedural anesthetics and vaccines data will be coded using the World Health Organization (WHO) Drug Dictionary (Global B3/September 2022).

4.1 Study Visits

The study visits are illustrated in Section 16.1 Study Flow Chart and Section 16.2 Schedule of Events.

4.2 Visit Windows

Study visits are expected to occur according to the protocol schedule. 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 of all dates on or after Baseline 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, unless used for Baseline determination. Unscheduled visits, if any, will be presented in listings only.

Analysis visit values will be generated using the raw VISIT values for use in tables. See Table below:

Table 6: AVISIT and VISIT Values

Visit (VISIT)	Presentation in Tables (AVISIT)	Presentation in Listings
Visit 1a Screening Day -21 to Day 1	Visit 1a (Screening) {or Baseline}	Visit 1a
Visit 1b Baseline/Tx Day 1	Visit 1b (Baseline) {or Baseline}	Visit 1b
Visit 2 (TC) 72 hrs after Baseline (+/- 24 hrs)	Visit 2 (Hour 72 after Treatment)	Visit 2
Visit 2a 14 days after Baseline (+ 7 days)	Visit 2a (Day 14 after Treatment)	Visit 2a
Visit 3 1 month after Baseline (+ 7 days)	Visit 3 (Month 1 after Treatment)	Visit 3
Visit 3a (TC) 72 hrs after Optional Touch Up (+/- 24 hrs)	Visit 3a (Hour 72 after Optional Touch-up)	Visit 3a
Visit 3b 14 days after Optional Touch Up (+ 7 days)	Visit 3b (Day 14 after Optional Touch-up)	Visit 3b
Visit 3c 1 month after Optional Touch Up (+ 7 days)	Visit 3c (Month 1 after Optional Touch-up)	Visit 3c
Visit 4 Follow-up 3 Months after Last Treatment (+/- 14 days)	Visit 4 (Month 3 after Last Treatment)	Visit 4
Visit 5 Follow-up 6 Months after Last Treatment (+/- 14 days)	Visit 5 (Month 6 after Last Treatment)	Visit 5
Visit 6 Follow-up 9 Months after Last Treatment (+/- 14 days)	Visit 6 (Month 9 after Last Treatment)	Visit 6
Visit 7 Follow-up 12 Months after Last Treatment (+/- 14 days)	Visit 7 (Month 12 after Last Treatment)	Visit 7

4.3 Study Duration

First subject first visit (FSFV) to last subject last visit (LSLV): estimated to approximately 18 months including 4 months enrollment period. Duration for each Subject will be up to 14 months, including:

- Up to 21 days screening.
- Treatment at Baseline (Day 1) and optional touch-up at 1 month.
- Data collection and safety follow-up visits at 1, 3, 6, 9 and 12 months after last treatment as well as 14 days after each treatment.

One month is defined as 4 weeks in the study and 4 weeks is defined as 28 days.

4.4 Baseline Definition

Baseline will be the last assessment prior to the injection of study drug unless otherwise indicated. The Screening Visit 1 (Day -21 to Day 1) will be considered the visit prior to injection of study drug. Because the Screening visit and Baseline visit (Day 1) may be performed on the same day, the Screening visit can also be Day 1.

For all analyses, Baseline will be defined as the most recent measurement prior to the first administration of study product, including unscheduled visits conducted prior to study treatment. The change from Baseline (Δ) will be defined as

$$\Delta = Visit\ X\ Assessment\ Value - Baseline\ Assessment\ Value$$

and will be summarized using the same descriptive statistics as other continuous endpoints.

In change from baseline analyses that treat the response as ordinal, the Δ representing change from baseline will be treated as a categorical variable.

4.5 Injection Timepoint Definition

4.5.1 Injection Timepoint for PPD Study Drug Exposure

PPD

PPD Timepoint will refer to the two following periods of the study:

- Initial Treatment – At or in the 28 days following Baseline injection, prior to touch-up treatment.
- Touch-up Treatment – If subject received a touch-up treatment at Month 1, at or in the 28 days following optional Month 1 touch-up injection.

4.5.2 Injection Timepoint for Adverse Events

For adverse events summaries, Injection timepoint will refer to the following periods of the study:

- Initial Treatment: at or after Baseline injection until optional touch-up treatment or end of study (EOS), whichever comes first. All adverse events for subjects who did not receive optional touch-up treatment will be included in this injection timepoint.

- Initial Treatment + Touch-up: at or after optional touch-up treatment until end of study (EOS). Subjects who did not receive optional touch-up treatment will not have adverse events included in this injection timepoint.

5 Analysis Populations

5.1 Intention-to-Treat Population

The Intention-to-Treat (ITT) Population will include all subjects who are randomized and will be analyzed according to the randomization scheme. Subjects will be analyzed according to the treatment they are planned to receive according to the randomization schedule. The ITT Population is the primary population for the analysis of effectiveness parameters.

5.2 Per-Protocol Population

The Per-Protocol (PP) Population will include all subjects in the ITT Population who complete the primary endpoint assessment at 3 months after baseline or last treatment without any deviations considered to have substantial impact on the primary effectiveness. Subjects will be analyzed according to the treatment they are planned to receive according to the randomization schedule. A sensitivity analysis of the primary effectiveness endpoint will be performed based on the PP Population.

5.3 Safety Population

The Safety Population will include all subjects who were treated with Restylane Lyft with Lidocaine or Juvéderm Voluma XC, and will be analyzed according to the as-treated principle. Subjects will be analyzed according to the treatment they have received according to the actual treatment. Safety analysis will be performed based on the Safety Population set.

6 Subject Disposition

Subject disposition will be summarized in a table. The number of screened, randomized, treated, completed, and withdrawn subjects will be presented, as well as number and percentages of subjects in each analysis population (ITT, Per-Protocol and Safety Populations). Percentages for the number of screen failures and randomized subjects are based on the number of subjects screened. Percentages for the number of subjects treated, the number of subjects who completed the study, the number of subjects who did not complete the study, and the number of subjects in the PP Population are based on the number of subjects in the ITT Population. The numbers and percentages of subjects associated with withdrawal reasons will be presented, with percentages based on the number of withdrawn subjects per treatment. A subject disposition table by study site will also be presented, with the same parameters reported.

Subject accountability will be summarized by visit with the following:

- Number of subjects expected at each visit.
- Number of subjects who performed each visit.
- Number of subjects who missed each visit.
- Number of subjects withdrawn at each visit.

Subjects who were excluded from the study will be presented in a data listing, along with reasons for exclusion. All subjects' completion and discontinuation information will be presented in a data listing that will report the last visit performed with date, the date of completion or discontinuation and the reason for discontinuation.

A study visit listing will include the reason for missing the study visit and how the remote visit was conducted when it applies. For unscheduled visits, changes since last visit concerning adverse events, concomitant medications or procedures will also be displayed.

An informed consent and randomization data listing will include the date and version of the initial consent and information on re-consent if it applies, as well as the date and time of randomization and the assigned treatment.

7 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 the PP Population will be used as a sensitivity analysis for the primary analysis at 3 months after last treatment only, the focus will be on deviations occurring before and on the 3 months after last treatment visit day, as they might compromise the primary endpoint. For this study, the protocol deviations that will exclude subjects from the PP Population are identified (but not limited to) in Table 6 below.

Table 7. Protocol deviations

Deviation
Any subject that conducts their 3 months after last treatment visit out of window by greater than 21 days or earlier than 7 days
Any subject not treated according to the assigned treatment.
Any subject that does not have a Blinded Evaluator GCRS onsite live assessment at the primary endpoint visit (3 months after last treatment).
Any subject that does not have an available Baseline (or screening if combined) Blinded Evaluator GCRS onsite live assessment.
Any subject that does not have grade 1 or 2 on either the Blinded Evaluator or Treating Investigator GCRS onsite live assessment at screening and Baseline.

Deviation
Any subject with prohibited concomitant treatments/procedures prior to the 3 months after last treatment 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 the 3 months after last treatment visit considered to have a substantial impact on the primary efficacy outcome.

Protocol deviations will be presented descriptively. The number and percentage of subjects who have protocol deviations, the total number of deviations and the type of protocol deviation will be summarized using the ITT Population. The above summary of protocol deviations will be repeated by site as well. Protocol deviation listings will be reviewed, and each subject will be classified as belonging to the per-protocol set or not. Exclusions from the PP Population will be identified and documented prior to database lock.

Protocol deviations will be presented in a data listing that will display the visit where the deviation was reported, the date where the deviation occurred, the deviation type and description, the action taken, if it was reported to the Institutional Review Board (IRB) and if the deviation was due to COVID-19. Subjects who are excluded from the PP Population will also be presented in a data listing, along with their reason(s) for exclusion.

7.1 Out of Window Visit Duration

When a subject performs a planned study visit outside of the protocol-specified visit windows (Section 16.2), the subject has an out of window study visit, which is considered a type of protocol deviation for this study. Knowing a subject's Screening visit (Visit 1a) and Baseline visit (Visit 1b) and using the protocol-specified visit windowing, one can calculate a subject's target study visit window date range for each planned study visit. Then, by comparing a subject's actual study visit date to the target study visit window date range, one can calculate the days over/under the actual study visit in relation to the target study visit date, corresponding to the study visit out of window duration.

The duration, in days, a study visit is out of window will be summarized by treatment group, overall, and by visit. Out of window visit durations will also be presented by site. Duration will be split into the following below categories. Duration categories will include both days before and days after the target planned visit date.

- +/- 1-2 days.
- +/- 3-7 days.
- +/- 8-20 days.
- +/- 21 or more days.

The number and percentage of subjects within each out of window duration category will be presented descriptively. For the overall analysis, percentages will be calculated using the number of subjects with an out of window visit at the respective visit as the denominator; for the by site analysis, percentages will be calculated using the number of

subjects at the respective site with an out of window visit at the respective visit as the denominator.

8 Demographic and Baseline Characteristics

8.1 Demographic Characteristics

Demographic assessments for this study include:

- Age (years);
- Age Category (≤ ITT Population median age, > ITT Population median age);
- Sex at Birth (Male, Female);
- Gender identity (Female, Male, Transgender male-to-female, Transgender female-to-male, Gender nonconforming (neither exclusively male nor female), Other, Decline to answer);
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino);
- Race (White, Black or African American, Asian (Chinese, Asian Indian, Filipino, Japanese, Korean, Vietnamese, Other Asian), American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other);
- Height (cm);
- Weight (kg);
- BMI (kg/m^2);
- Fitzpatrick Skin Type (FST) Score (I, II, III, IV, V, VI);
- Fitzpatrick Skin Type (FST) Cohort at Randomization (I-III, IV, V-VI);
- Baseline GCRS by Blinded Evaluator (0, 1, 2, 3);
- Baseline GCRS by Treating Investigator (0, 1, 2, 3).

Demographic and Baseline characteristics will be summarized using the ITT Population. No formal statistical comparisons between treatment groups will be performed. Continuous variables to be summarized include age (in years), height (in cm), weight (in kg), BMI (in kg/m^2). Categorical variables include age category, gender, ethnicity, FST, Baseline GCRS by Blinded Evaluator and Baseline GCRS by Treating Investigator.

FST Cohort at Randomization will be used in tables. FST Cohort at Randomization and FST Score will both be included in listings.

Demographic and Baseline characteristics, as well as the childbearing potential and breastfeeding status for female subjects will also be reported in a data listing.

8.2 Medical History

Subjects' medical history will be collected at Screening and at Baseline visits. An event that occurs after the subject signs the Informed Consent Form (ICF) but before enrollment will be recorded in the subject's medical history.

Medical history summaries will be reported using the ITT Population. The number and percentage of subjects reporting medical history and the number of events will be summarized by MedDRA System Organ Class (SOC) and Preferred Term (PT). System Organ Classes and PTs will be sorted in descending frequency first (by Restylane Lyft with Lidocaine group) and then alphabetically if there are ties. Each subject will only be counted once within each SOC and SOC/PT combination. The MedDRA version used will be noted as a footnote in the tables and listings.

Medical history information will also be reported in a data listing presenting the SOC, the PT, the description, and the onset and stop date (or if ongoing).

8.3 Dermatological Procedures and Implant History

Subjects' dermatological procedures and implant history will be collected at Screening and at Baseline visits and summarized using the ITT Population in the same way as for subject's medical history (Section 8.2).

Dermatological procedures and implant history listing will include the SOC, the PT, the procedure or product name, the location and the date.

8.4 Concomitant Procedures/Non-Pharmacological Treatments

Concomitant procedures/non-pharmacological treatments are defined as follows:

- Any existing procedures ongoing at the time of the Screening visit.
- Any changes to existing procedures (such as changes in dose or formulation) during the study.
- Any new procedures received by the subject since the Screening visit.

Subjects' concomitant procedures will be collected at all study visits.

Concomitant procedures will be summarized using the ITT Population in the same way as for the medical history. The subset of concomitant procedures administered for treatment-related adverse events (see Section 10.2.3) will be summarized in the same way in an additional table.

Concomitant procedures listing will include the SOC, the PT, the procedure name, the start and stop date, the reason for procedure, and the related medical history and/or adverse event when it applies.

8.5 Prior and Concomitant Medications

Prior medications are defined as medications that have been used within 30 days preceding the Screening visit or within the timelines specified in the Inclusion/Exclusion criteria, and then stopped prior to the Screening visit. Concomitant medications are any existing medications ongoing at the time of the Screening visit, any changes to existing medications during the study or any new medications received since the Screening visit. Subjects' prior medications will be collected at Screening and concomitant medications will be collected at all study visits.

Prior and concomitant medication summaries will be reported using the ITT Population in two different tables. The number and percentage of subjects reporting prior/concomitant medications and the number of events will be summarized by the WHO Drug Dictionary Anatomical Therapeutic Chemical 4th level (ATC-4) and the Preferred Term (PT). If the 4th level term is not available, the next available level (e.g., Anatomical Therapeutic Chemical 3rd level (ATC-3)) will be used. ATC-4 and PT will be presented by overall descending frequency first, then alphabetically if there are ties. Anatomical Therapeutic Chemical 4th level and PTs will be sorted in descending frequency first (by Restylane Lyft with Lidocaine group) and then alphabetically if there are ties. Each subject will only be counted once within each Anatomical Therapeutic Chemical (ATC) and ATC/PT combination. The subset of concomitant medication administered for treatment-related adverse events (see Section 10.2.3) will be summarized in the same way in an additional table.

All prior and concomitant medication data for each subject will be included in a data listing that will present the name of the medication, the ATC and PT, the indication, the total daily dose, the route, the start and end date, the reason of administration, and the related medical history and/or adverse event when it applies.

9 Effectiveness Analysis

Unless otherwise stated, all effectiveness analyses will be conducted primarily using the ITT Population. The analyses may be reproduced using the Per-Protocol Population if the PP Population contains less than 90% of the subjects in the ITT Population.

9.1 Primary Effectiveness Endpoint

The primary endpoint is defined as the change from Baseline based on the Blinded Evaluator's live assessment at 3 months after last treatment using the GCRS. The change from Baseline to 3 months after last treatment GCRS Blinded Evaluator will be calculated using the equation specified in Section 4.4.

9.1.1 Primary Effectiveness Estimand

The primary estimand is the treatment difference of Restylane Lyft with Lidocaine versus a comparator-control for augmentation of the chin region to improve the chin profile by comparing change from baseline in the Blinded Evaluator live assessment of the GCRS at 3 months after last treatment, in all randomized subjects (ITT Population). To address the impact of intercurrent events, sensitivity analyses will be performed separately from the primary endpoint analysis as outlined in Section 9.1.3.

Primary Endpoint	**Estimand**	
Change from baseline in the Blinded Evaluators' live assessment using the GCRS at 3 months after last treatment	**Population:** all randomized subjects (ITT Population)	
	Endpoint: The subject's GCRS based on Blinded Evaluator's live assessment at Month 3 after last treatment	
	Intercurrent Events: 1. Use of prohibited treatment in or near the treatment area prior to 3 Months after last treatment. 2. Use of Hyaluronidase prior to 3 Months after last treatment. 3. Missed visit, early termination, or missing Blinded Evaluator's live assessment of the GCRS leading to missing 3 Months after last treatment observation. 4. Administered treatment not according to randomization. 5. Subjects who were Lost to Follow Up prior to Visit 4 and re-	**Handling of Intercurrent Events:** 1. Use observed response. 2. Use observed response. 3. Use multiple imputation (MI). 4. Use observed response. 5. Assume "No Change from Baseline".

	enrolled after discontinuation.	
	Summary measure: Treatment Difference of Restylane Lyft with Lidocaine versus Juvéderm Voluma XC (control) using the GCRS in average change from Baseline.	

Note: Intercurrent events are defined as events occurring after treatment initiation that affect either the interpretation or the existence of the measurement associated with the clinical question.

9.1.2 Primary Effectiveness Analysis

The primary effectiveness analysis is a test to demonstrate non-inferiority of Restylane Lyft with Lidocaine to a comparator-control Juvéderm Voluma XC in the ITT and PP Population analysis sets, using the confidence interval (CI) approach.

The change from baseline in GCRS, as assessed by the Blinded Evaluator, will be analyzed using a mixed effect model with treatment and Baseline GCRS value as fixed effects. The change from baseline at Month 3 after last treatment, will be the response variable, as specified below:

$$Y(\Delta) = \beta_0 + \beta_1 Treatment\ Group + \beta_2\ Baseline\ GCRS\ value + \varepsilon_i$$

Where:

- $Y(\Delta)$ is the change from Baseline to Month 3 after last treatment in GCRS value;
- β_k is the slope of a given covariate (i.e., fixed effect), k=1,2;
- ε_i is the error for subject i, $\varepsilon_i \sim MVN(0, \sigma^2)$;
- β_0 is the fixed intercept.

The maximum likelihood (ML) will be used for the model estimation method.

Missing Month 3 after last treatment BE GCRS values will be imputed using Multiple Imputation (MI). The MI methods are detailed in Section 11.6.

The two-sided 95% CI around the Least Square (LS) mean estimate of the Restylane Lyft with Lidocaine – Control difference will be obtained from the model and used for assessing non-inferiority in the following way: non-inferiority will be established if the upper limit of the CI for the mean difference between treatment groups is below the non-inferiority margin of 0.5 units. Superiority will be declared if the upper limit of the CI for the mean difference between treatment groups is below 0. Non-inferiority must be shown for both ITT and PP Populations. If non-inferiority is shown in the ITT and PP Population, superiority will be declared if the upper limit is below 0 in the ITT Population. This decision criterion is consistent with a gate-keeping strategy (testing for superiority following the test for non-inferiority) to maintain the overall type I error rate at 2.5%.

The null and alternative hypotheses are stated below.

Δ= the LS mean estimate of (Restylane Lyft with Lidocaine - Juvéderm Voluma XC), the statistical hypotheses to be tested are:

H_0: $\Delta \geq 0.5$
H_1: $\Delta < 0.5$

Results from the mixed model will be presented with the LS mean estimates of each treatment group with their corresponding CIs and the LS mean estimate of the treatment difference with its standard error and CI.

The primary variable is the Month 3 after last treatment GCRS value minus the Baseline GCRS value. This change from Baseline measure indicates a favorable result for a subject if the value is negative. From the model a 95% confidence interval for the LS mean estimate of the treatment difference (Restylane – Control) is calculated. A negative value for the mean treatment difference is a favorable result for Restylane, and a positive value for the mean treatment difference is a favorable result for the Comparator.

During the study, subjects 8546-002, 8546-006, 8546-008, 8546-009, and 8546-013 were identified as Lost to Follow Up. Later in the study, the subjects were re-enrolled using the same subject ID numbers. Subjects were re-enrolled at the visit with a target date closest to the date at which they reentered the study. For subjects in this situation who missed the primary endpoint visit due to being Lost to Follow Up, the subjects will be handled assuming No Change from Baseline in the primary endpoint analysis and any sensitivity or subgroup analyses using multiple imputation (see Section 9.1.1). No Change from Baseline will be used in lieu of multiple imputation because the missing data may be informative and it is reasonable to consider the missing data pattern missing not at random. For subjects in this situation who did not miss the primary endpoint visit, the primary endpoint analysis will be conducted using the data as collected.

9.1.3 Sensitivity Analysis of Primary Endpoint

To assess the robustness of the primary efficacy results, the following additional sensitivity analyses will be conducted:

- To evaluate the impact of missing data on the primary endpoint, the primary analysis specified in Section 9.1.2 will be repeated using the ITT Population with Observed Cases (OC), that is, no imputation will be done.
- To evaluate the impact of missing data on the primary endpoint, the primary analysis specified in Section 9.1.2 will be repeated using the ITT Population with baseline observation carried forward (BOCF) method.
- To evaluate the impact of missing data on the primary endpoint, the primary analysis specified in Section 9.1.2 will be repeated using the ITT Population with imputation of the worst case. For the GCRS, the worst case is defined as 3 – Severe Retrusion.

- To evaluate the impact of intercurrent events, the primary analysis specified in Section 9.1.1 will be repeated using the ITT Population with MI but excluding subjects that had a prohibited treatment in or near the treatment area prior to 3 months after last treatment. Subjects with prohibited treatments are defined as subjects with a protocol deviation type of "Prohibited medication or procedure" prior to or on their primary endpoint visit.
- To evaluate the impact of intercurrent events, the primary analysis specified in Section 9.1.1 will be repeated using the ITT Population with MI but excluding subjects who were administered treatment not according to randomization. Subjects who were administered treatment not according to randomization are defined as subjects who were randomized but not treated, subjects whose actual treatment at initial treatment or optional touch-up treatment does not match their planned treatment according to the randomization schedule, or subjects who were treated but not randomized.
- To evaluate the impact of intercurrent events, the primary analysis specified in Section 9.1.1 will be repeated using the ITT Population with MI but excluding subjects that used Hyaluronidase prior to 3 months after last treatment.
- A sensitivity analysis of the primary effectiveness endpoint specified in Section 9.1.2 will be performed based on the PP Population with using observed cases.
- The change from baseline (i.e., number of subjects scored as -3, -2, -1, 0, 1, 2, or 3) will be analyzed as an ordinary outcome, using an Ordinal Logistic Regression. A similar model described for the primary endpoint in Section 9.1.2 will be applied using ordinal logistic regression models. Treatment group (Restylane Lyft with Lidocaine, Juvederm Voluma XC) and Baseline GCRS value (1, 2) will be included in the model as predictors. The reference category for Treatment group will be Restylane Lyft with Lidocaine. The reference category for Baseline GCRS value will be 2. PROC LOGISTIC will be used to generate Odds Ratio point estimates and Wald confidence intervals for Treatment Group and Baseline GCRS value. If the proportional odds assumption is not met, a partial proportional odds model will be used in lieu of a proportional odds model. This partial proportional odds model will be generated using PROC LOGISTIC with an UNEQUALSLOPES option.
  - A separate table will present the predicted probabilities of each change from baseline value by treatment. The same PROC LOGISTIC used above, with the addition of the PREDPROBS(I) option will be used to output the predicted probabilities. Using the predicted probabilities, the expected value of the change from baseline will be calculated as:

$$\sum [Predicted\ Probability_i \times Change\ from\ Baseline_i]$$

where i = *Change from Baseline value*. The expected value will be presented in the output by treatment.

9.2 Secondary Effectiveness Analysis

Secondary effectiveness analysis will be evaluated based on the OC in ITT Population, unless otherwise indicated.

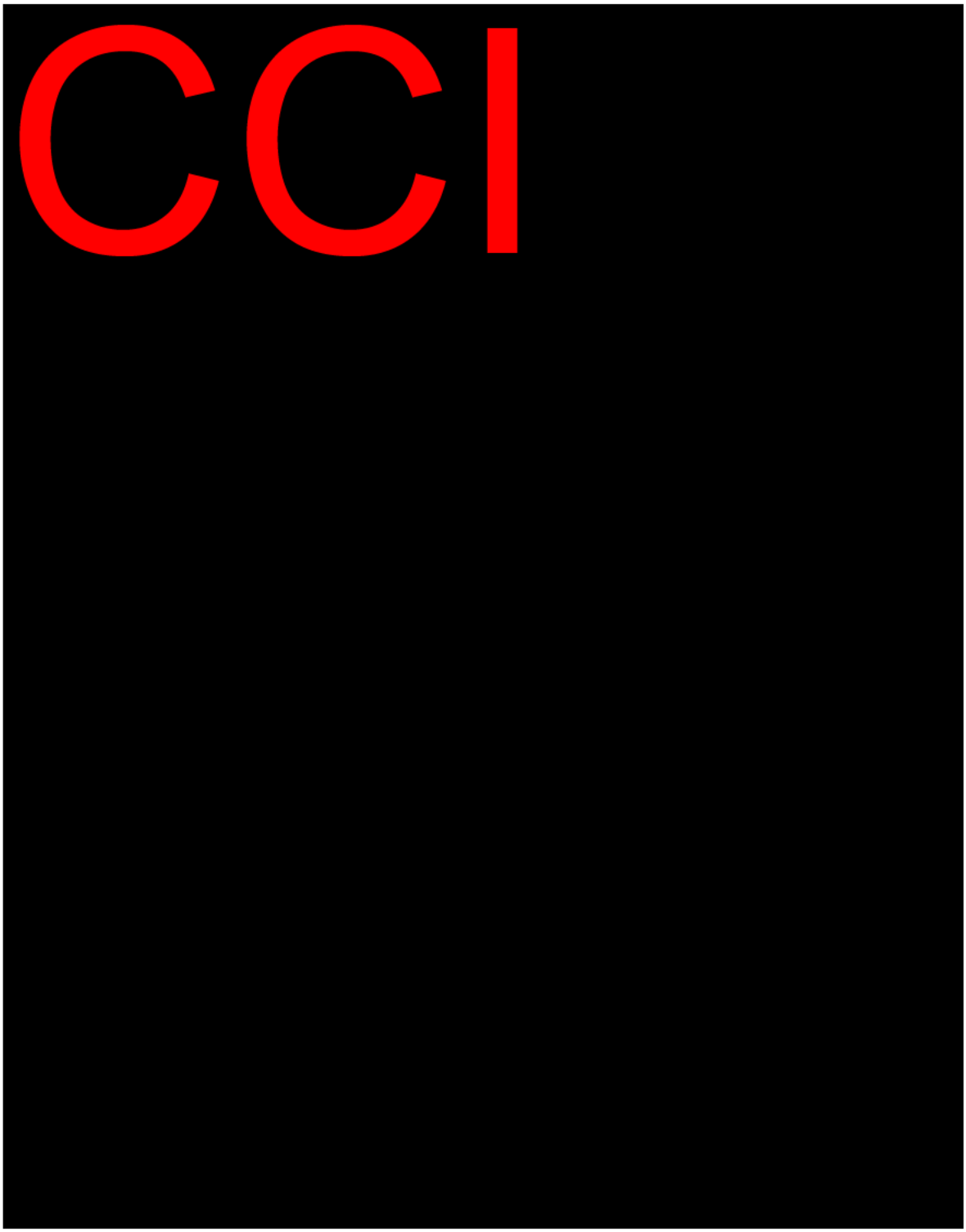
P-values obtained for any endpoint other than the primary will only be provided for ease of interpretation of results. All other secondary effectiveness analyses will be performed descriptively only.

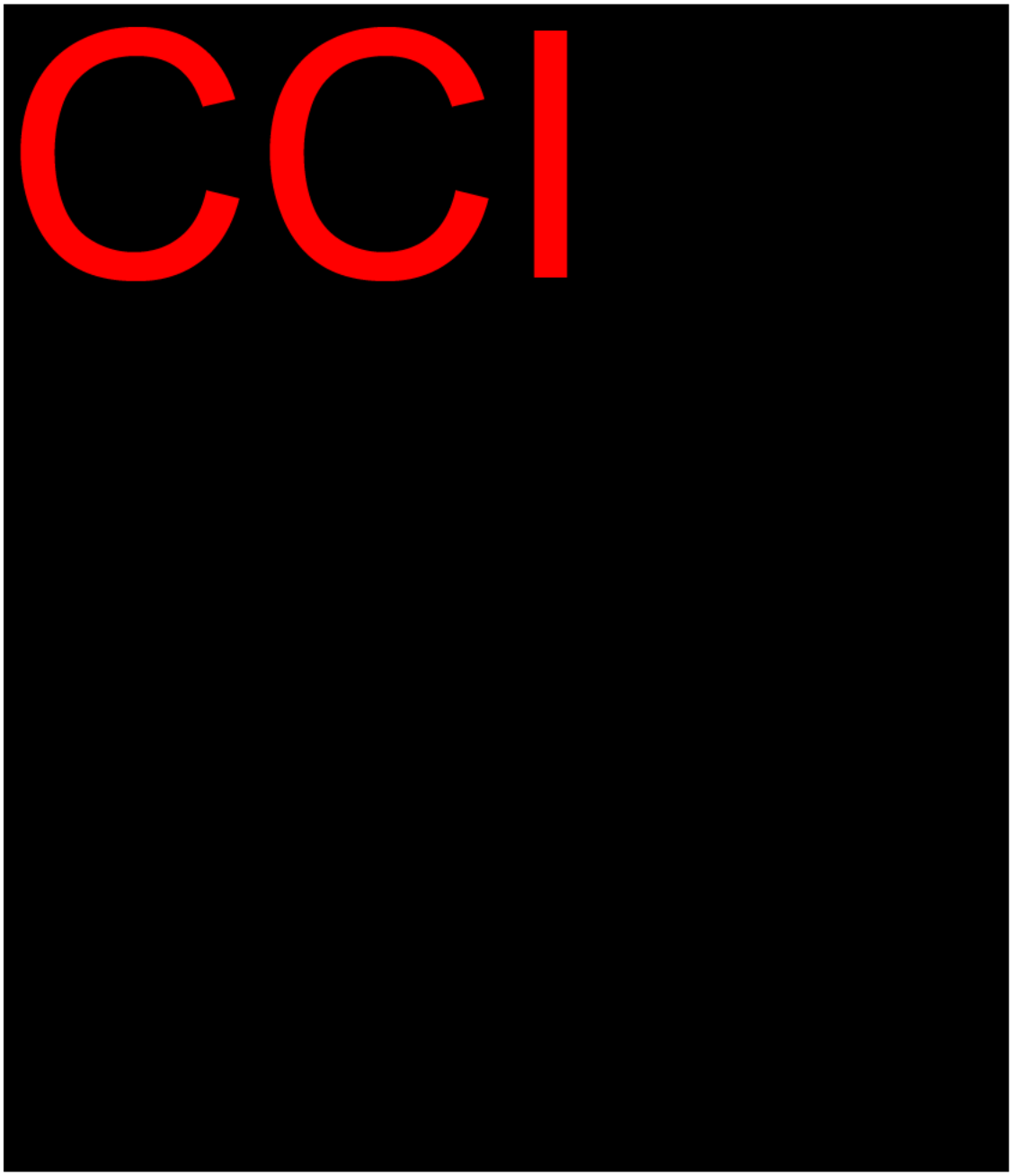
9.2.1 Galderma Chin Retrusion Scale (GCRS) Responders Rate

The responder rates based on GCRS as assessed by the BE (defined in Section 3.2.2.1) at Months 3, 6, 9 and 12 after last treatment will be presented by treatment group with their 95% CIs using the Clopper-Pearson method, and the two-sided p-value of the comparison of responder rates between the treatment groups at each visit using the Fisher's exact test. Results will include estimates of the responder rates, difference in responder rates, and corresponding 95% CIs, as well as a graph showing the responder rates for the two treatment groups over time. The difference in responder rates corresponding 95% CI will be generated using continuity-corrected confidence limits by specifying RISKDIFFC.

The number and percentage of subjects in each response category, from No Retrusion (0) to Severe Retrusion (3) of the BE's GCRS, will be summarized by visit for each treatment group. The number and percentage of subjects in each change from baseline bin (-3, -2, -1, 0, 1, 2, 3) will be summarized by visit for each treatment group.

The Treating Investigator's and the Blinded Evaluator's GCRS results, including the actual GCRS ratings, the change from Baseline and the BE responder status (yes/no), will be presented in a data listing.




CCI

10.1 Study Drug Exposure

Data on extent of exposure and treatment procedure will be summarized by treatment group and by Injection Timepoint as defined in section 4.5.2.

10.1.1 Injection Characteristics

Injection characteristics include injection administration, procedural anesthetics, and post-injection procedures used. No overall column will be included. The following parameters will be summarized by Injection Timepoint:

- Total number of subjects treated.
- Treatment Area (Sublabial Crease, Pogonium, Menton, Prejowl sulcus area – left hand side, Prejowl sulcus area – right hand side).
- Total number of syringes used.
- Injection tool(s) used by treatment area.
- Injection method(s) used.
- Depth(s) of injection (Deep, Supraperiosteal, other).
- Use of procedural anesthetics (yes/no).
- Type of anesthetics (Topical or local injection).
- Post-injection procedures (none, massage, ice pack, other).

10.1.2 Injection Volume Administered

The injection volume (mL) administered throughout the study will be presented descriptively by Injection Timepoints. No overall column will be included. Additionally, a cumulative sum will be presented adding Initial treatment and Touch-up.

All injection volume, injection characteristics, and post-injection procedures data will be presented by treatment and subject in a data listing for each visit with treatment. Procedural anesthetics will also be reported, specifying the anesthetic name, concentration, volume and type.

For the optional touch-up, the eligibility, and the reason for not performing the injection will also be presented by treatment and subject in a data listing.

10.2 Adverse Events

Adverse Event reporting on each subject will start once a subject is enrolled in the study. All other events that occur after the subject signs the ICF but before enrollment will be recorded in the subject's medical history. The reporting shall continue during each follow-up visit (including telephone contacts and extra visits between planned visits) until the last scheduled visit in the study.

All AEs reported will be listed, documenting the verbatim term, MedDRA Preferred Term (PT) and System Organ Class (SOC), affected area, event start and end date (or ongoing), seriousness (yes/no), intensity, relationship to study product (yes/no), relationship to injection procedure (yes/no), UADE status, AESI status, action taken, outcome, time to onset, duration and Injection Timepoint. All AE endpoints will be presented Overall. The Overall Summary of Adverse Events table and the Adverse Events by System Organ Class and Preferred Term table will be presented Overall and by Injection Timepoint in which they occurred (see definition in Section 4.5.2). Additional presentations of AE endpoints may be presented by Injection Timepoint on an as needed basis.

Only AEs occurring after treatment will be included in analysis. AEs occurring before initial treatment, if any, will only be presented in data listings. Missing dates, intensity, relationship, and seriousness will be imputed as described in Section 11.1 and Section 11.2.

10.2.1 Summary of Adverse Events

A summary of AEs will be provided, which will include the number and percentage of subjects and number of events in the following categories:

- Any AE.
- Any AE related to study product or injection procedure.
- Any AE unrelated to study product or injection procedure.
- Any serious AE.
- Any serious AE related to study product or injection procedure.

- Any serious AE unrelated to study product or injection procedure.
- Any AE leading to study discontinuation.
- Any AE ongoing at the end of the study.
- Any AESI
- Any UADE
- Subjects who did not have any AE.

The number and percentage of subjects with at least one AE and the number of events will be summarized by SOC and PT. System Organ Classes and PTs will be sorted in descending frequency first (by Restylane Lyft with Lidocaine group) and then alphabetically if there are ties. This sorting method will apply to all other AE tables described below.

10.2.2 Intensity of the Adverse Events

The intensity of the AEs (mild, moderate, or severe) will be recorded. If the intensity changes within one day, the maximum intensity of the AE during that day will be recorded. If the intensity assessment is missing, the intensity of 'Severe' will be assumed. AEs will be summarized by SOC, PT and intensity. If a subject has multiple occurrences of the same MedDRA SOC or PT, then only the most severe event will be summarized in the tables for that SOC and PT.

10.2.3 Relationship to Study Product of Injection Procedure

Each AE will be assessed by the Investigator for a reasonable possibility of a causal relationship with the study product and with its use (the injection procedure), both on a two-point scale (Yes or No). If the response to any of the questions "*Do you consider that there is a reasonable possibility that the event may have been caused by the study product?*" or "*Do you consider that there is a reasonable possibility that the event may have been cause by the study product injection procedure?*" is "Yes", then an AE is considered related for the analysis. Missing relationships will be considered related.

Related AEs will be summarized by SOC, PT and maximum intensity (defined in Section 10.2.2). Unrelated AEs, as well as serious related AEs, will also each be summarized by SOC, PT and maximum intensity. If a subject has multiple occurrences of the same MedDRA SOC or PT, then only the most severe event will be summarized in the tables for that SOC and PT.

Finally, the number of subjects with related AEs will be summarized by SOC, PT, and action taken (none, medication treatment, non-pharmacological treatment, other procedures/tests, subject withdrawn). Subjects will be only counted in 'None' category if no other action was taken.

10.2.4 Time to Onset and Duration of AEs

Time to onset and duration of related AEs will be recorded. Time to onset of an AE will be derived as the AE start date minus the date of most recent prior treatment. If the AE start

date is partially/completely missing, it will be imputed using the imputation rules specified in Section 11.2.

Duration of an AE will be derived as the AE stop date minus the AE start date + 1. If the AE start date is partially/completely missing, it will be imputed using the imputation rules specified in Section 11.2. A completely missing AE stop date will not be imputed and therefore no duration will be calculated in these cases.

The number of days to onset and the duration of each respective related AEs will be summarized by SOC and PT using the mean, SD, min, max and median. These summaries are at the event level so that subjects will be counted once for each multiple related AE they experience. Related AEs with onset >21 days after the most recent treatment will be listed.

10.2.5 Adverse Events of Special Interest

For this study, all instances of visual disturbances, regardless of relationship to study product or seriousness are considered AESIs and are outlined in the protocol.

In addition, the following events are considered AESIs:

- All abnormal findings in functionality tests, assessed by the Investigator, regardless of relationship to the study product or seriousness
- Stroke
- Skin necrosis

AESIs, as identified in the EDC, will be summarized the same way as AEs, by SOC and PT, and will be presented in a listing.

10.2.6 Death, Serious Adverse Events, Adverse Events Leading to Discontinuation of Study

Serious AEs, AEs resulting in death on study and AEs leading to discontinuation of study will be reported in separate listings. These listings will include the same variables as the listing for all AEs, and the Serious Adverse Events (SAEs) listing will additionally include the SAE criteria.

10.3 Subject pain assessment

The Numeric Pain Scale (NPS) will be treated as a continuous variable and will be summarized descriptively by visit, timepoint (pre- and post-injection). The change from pre- to post-injection pain will also be presented.

Subject pain assessment will also be presented in a data listing that will include the reason why the assessment was not done, if applicable.

10.4 Sensory tests

Sensory tests, consisting of the Monofilament Test and the Cotton Wisp Test, will be summarized with counts and percentages of subjects with normal/abnormal findings by visit.

The sensory test will be presented in a data listing that will include the assessments of sensation at all locations for both tests, including the final test scores. When an abnormality is identified, the description of this abnormality will appear in the listing.

10.5 Functionality tests

Lip Movement - Word Pronunciation, will be described with counts and percentages for each word and the word score will be presented using mean, SD, minimum, maximum, and median statistics by visit. Functionality tests, consisting of the Lip Function - Drinking from Straw and Chin Function - Open and Close Mouth, will be summarized with counts and percentages of subjects with normal/abnormal findings by visit.

The functionality test will be presented in a data listing that will include the assessments of lip movement and the lip and chin function. When an abnormality is identified, the description of this abnormality will appear in the listing.

10.6 Assessment of palpability

Palpation data will be summarized with counts and percentages of subjects in each category (Yes – Expected Feel (Normal), Yes – Unexpected Feel (Abnormal)) by visit.

The palpation data will be presented in a data listing that will include the description of abnormality when it applies and the reason why the assessment was not done when it applies.

10.7 Changes in hair growth

Changes in hair growth data will be summarized with counts and percentages of subjects in each category (Yes, No) by visit.

The changes in hair growth data will be presented in a data listing that will include the description of abnormality when it applies.

10.8 Visual Function Assessments

10.8.1 Snellen Visual Acuity

The number and percentage of subjects that experience an Acute, Compared to Previous Visit, and/or Change from Baseline worsening of 1+ line change in at least one eye during the study will be summarized by eye, visit and timepoint (pre- or post-injection, if

applicable). The number of subjects with a non-missing Snellen assessment at a respective visit/timepoint will be used as the denominator.

Snellen VA Line Changes will be presented by eye in an (a, b) format, where a represents the individual change in the left eye and b represents the individual change in the right eye (e.g., (-2, -1) indicates a 2-line worsening in the left eye and a 1-line worsening in the right eye).

Snellen Visual Acuity test data will be presented in a listing.

10.8.2 Extraocular Muscle Function

The number and percentage of subjects reporting normal and abnormal test results on the Extraocular Muscle Function test will be presented by eye, visit and timepoint (pre- or post-injection, if applicable).

Extraocular Muscle Function test data will be presented in a listing.

10.8.3 Confrontation Visual Field

The number and percentage of subjects reporting normal and abnormal test results on the Confrontation Visual Field test will be presented by eye, visit and timepoint (pre- or post-injection, if applicable).

Confrontation Visual Field test data will be presented in a listing.

10.10 Height and Weight

Subjects' height and weight will be collected at Baseline and at the Month 12 visit (weight only) and recorded in a listing.

10.11 Urine Pregnancy Test

A urine pregnancy test will be performed at Screening and Baseline for all women of childbearing potential. At Month 1 after Baseline, a test is done only if the touch-up treatment is performed. During visits with a treatment, the test is done prior to treatment administration.

Results will be reported in a data listing that will include the date of the pregnancy test, the test result and the reason it was not done where it applies. Pregnancies reported during the study will also be presented in this data listing.

10.12 Vaccines

Vaccine data will be reported in a data listing that will include the type of vaccine, the name, the manufacturer, the route, the location and the date and time of administration.

10.13 Device Deficiencies

A device deficiency is defined as an inadequacy of a medical device with respect to its identity, quality, durability, reliability, safety, or performance. Device deficiencies will be evaluated at Baseline and if the touch-up treatment is performed.

Any reported product complaint related to device deficiency will be presented in a data listing.

11 Statistical/Analytical Issues

11.1 Handling of Missing Data

The number of missing values will be summarized and reported as appropriate in all outputs. Calculations for all questionnaires will follow the methods specified in the specific sections of this SAP for the handling of the data.

For ITT analysis of the change from Baseline based on the Blinded Evaluator's live assessment of GCRS at Month 3 after last treatment (primary endpoint), the analysis will be performed using MI as the primary imputation method. Multiple imputation methods are specified in Section 11.6.

Sensitivity analyses will be performed to evaluate the impact of missing data on the primary endpoint using BOCF, worst case, and observed cases as described in Section 9.1.3.

All other effectiveness endpoints will be evaluated based on the OC in the ITT analysis set. Descriptive statistics of all safety data will be performed on OC in the Safety Population.

In the event the intensity, relationship, or seriousness of an AE is missing, the most severe value will be assigned. Intensity will be assigned to "Severe", relationship will be assigned to "Related to Study Product and Study Procedure", and seriousness will be imputed to SAE.

11.2 Handling of Missing/Partial Dates

While every effort will be made to obtain full, complete information on all data collected, Table 8 outlines the imputation rules to be followed for any missing dates. Dates will be presented as collected in the listings.

Table 8. Imputation Criteria for Missing/Partial Dates

Variable	Missing Day	Missing Month or Day and Month	Missing Year, Month and Year, or Day and Year	Missing Day, Month, Year
Adverse Event/Medication Start Date	Assign to the first of the month (i.e., UNK-JAN-2023 becomes 01-JAN-2023)	Assign to the month of 'June' (i.e., 01-UNK-2023 becomes 01-JUN-2023), provided the imputed date is on or after the subject's Baseline study treatment date; otherwise, the subsequent month after study treatment will be used.	Assign to year of Baseline study treatment (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 date of Baseline study treatment will be used.	Assign to Baseline study treatment date
Adverse Event/Medication End Date	Assign the last day of the month (i.e., UNK-JAN-2023 becomes 31-JAN-2023).	Assign to the subsequent month after the start date.	Assign to the year of Baseline study treatment (i.e., 01-JAN-UNK becomes 01-JAN-2023), provided the imputed date is after the start date; otherwise, the subsequent year after start date will be used.	Assumed AE/ medication is ongoing; no imputation

11.3 Pooling of Sites

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 site-level analyses, the number of subjects at each site may need to be considered for reliable and accurate results. If any site enrolls 5 or fewer subjects, the site will be removed from analyses. Sites enrolling 5 or fewer subjects will still be presented in descriptive statistics analyses.

If no sites enroll 5 or fewer subjects, then the efficacy site-level analyses will be presented for all sites.

11.4 Multiple Comparisons/Multiplicity

No adjustments for multiple comparisons or multiplicity will be made.

11.5 Examination of Subgroups

11.5.1 Effectiveness

To examine consistency of the results of the analyses of the primary endpoint across different subgroups, the following factors will be used for definition of subgroups, respectively:

- Study site;
- Race (White vs. Black vs. Other); If a subject indicates more than one race, they will be considered 'Other'.
- Ethnicity (Hispanic or Latino vs. Not Hispanic or Latino);
- Sex at Birth (Male vs. Female);
- Age category (≤ ITT median age vs. > ITT median age).
- FST (I-III vs. IV-VI); FST Cohort at Randomization will be used to generate FST subgroups.
- Injection volume (≤median injection volume vs. >median injection volume)
- Injection tool (Needle, Cannula, Needle and Cannula)

Subgroup analysis will be performed using the same mixed model as for the primary effectiveness analysis (specified in Section 9.1.2) but it will also include subgroup factor-by-treatment group interaction term. These analyses will be done using the ITT Population and MI as the imputation method.

Results from the above mixed models will be presented in tables that will display the LS mean estimates with their CIs for each treatment*factor interaction and the LS mean estimate of the treatment difference within each factor with its standard error and CI.

The number and percentage of subjects in each category of the BE's GCRS, from No Retrusion (0) to Severe Retrusion (3), will also be summarized by visit and subgroup category for each treatment group. The number and percentage of subjects in each change from baseline bin (-3, -2, -1, 0, 1, 2, 3) will be summarized by visit and subgroup category for each treatment group.

11.5.2 Safety

The consistency of AE data across different subgroups will be evaluated. The following subgroup factors will be used:

- Study site;

- Race (White vs. Black vs. Other); If a subject indicates more than race, they will be considered 'Other'.
- Ethnicity (Hispanic or Latino vs. Not Hispanic or Latino);
- Sex at Birth (Male vs. Female);
- Age category (≤ Safety Population median age vs. > Safety Population median age).
- FST (I-III vs. IV-VI); FST Cohort at Randomization will be used to generate FST subgroups.
- Injection volume (≤median injection volume vs. >median injection volume)
- Injection tool (Needle, Cannula, Needle and Cannula)

For the injection volume subgroup analysis, the median injection volume will be computed for the Injection Timepoint defined in Section 4.5.2.

Related AEs will be summarized within each subgroup by SOC, PT and maximum intensity.

11.6 Multiple Imputation Method

The MI procedure will be performed for the primary analysis using the following methods:

Missing data for 3 months after last treatment Blinded Evaluator GCRS assessment will be imputed using MI under the Missing at Random (MAR) assumption. Change from Baseline in Blinded Evaluator 3 months after last treatment GCRS will be derived from the corresponding imputed values. The following steps will be followed to create the imputations:

1. Regardless of the actual pattern of missing data, the Markov Chain Monte Carlo (MCMC) method of the SAS PROC MI procedure will first be used to make it monotone. The single chain method will be used. The minimum values for imputed variables will be set to 0, in order to force PROC MI to redraw another value for imputation when an intended imputed value is less than the 0, the maximum value for imputed variables will be set to 3, in order to force PROC MI to redraw another value for imputation when an intended imputed value is greater than the 3. Imputed values will be rounded to the nearest integer. The seed number will be set to 2208 and 5 imputations will be created.
2. The SAS PROC MI will be used for imputing missing values of data with monotone missing pattern. One imputation will be made using each of the 5 MCMC-imputed datasets (Step 1). A linear regression model will be used with covariates for treatment, site, race group, ethnicity, sex, Fitzpatrick Skin Type, age, total injection volume, and non-missing GCRS values from earlier scheduled time points including Baseline. The minimum values for imputed variables will be set to 0, in order to force PROC MI to redraw another value for imputation when an intended imputed value is less than the 0. Imputed values will be rounded to the nearest integer. The seed number will be set to

2208. Change from Baseline in 3 months after baseline GCRS will be derived from the corresponding imputed values. The Lost to Follow Up subjects who were reenrolled described in Section 9.1.1 and 9.1.2 will be handled assuming No Change from Baseline after the missing values have been imputed with multiple imputation. Protocol deviations will be used to identify subjects who missed their primary endpoint visit due to being Lost to Follow Up and later reenrolled.

3. The imputed datasets will be analyzed as specified in Section 9.1.2.
4. The results from the mixed model analysis on the imputed datasets will then be combined using the SAS PROC MIANALYZE to produce a single set of statistics.

12 Interim Analysis and Data Monitoring

No interim analysis is planned for this study.

13 Quality Control

All data displays and analyses will adhere to the International Conference on Harmonization (ICH) *Harmonized Tripartite Guideline: Structure and Content of Clinical Study Reports (ICH Topic E3).*

All analyses will be performed using SAS® Version 9.4 (or later). CCI will follow its standard operating procedures in the creation and quality control of all tables, listings, figures, and analyses. The Sponsor or its designee will review all tables, listings, and figures prior to final database lock. All final SAS programs and associated output files will be transferred to the Sponsor in agreed-upon format at project completion.

14 Tables and Listings Conventions

Mock-ups for statistical tables and listings will be provided. Final formats for the statistical tables and listings may deviate from these mock-ups upon agreement with the Sponsor. Footnotes will be used as needed to clarify the information that is presented in the tables and listings. Unless otherwise requested by the Sponsor, the term 'subject' will be used in all tables and listings, in accordance with Clinical Data Interchange Standards Consortium (CDISC) standards.

The table and listing mock shells, along with their programming convention, will be provided in a separate document.

15 Changes to Planned Analyses

15.1 Changes to Planned Analyses

This is the first version of the SAP for the study.

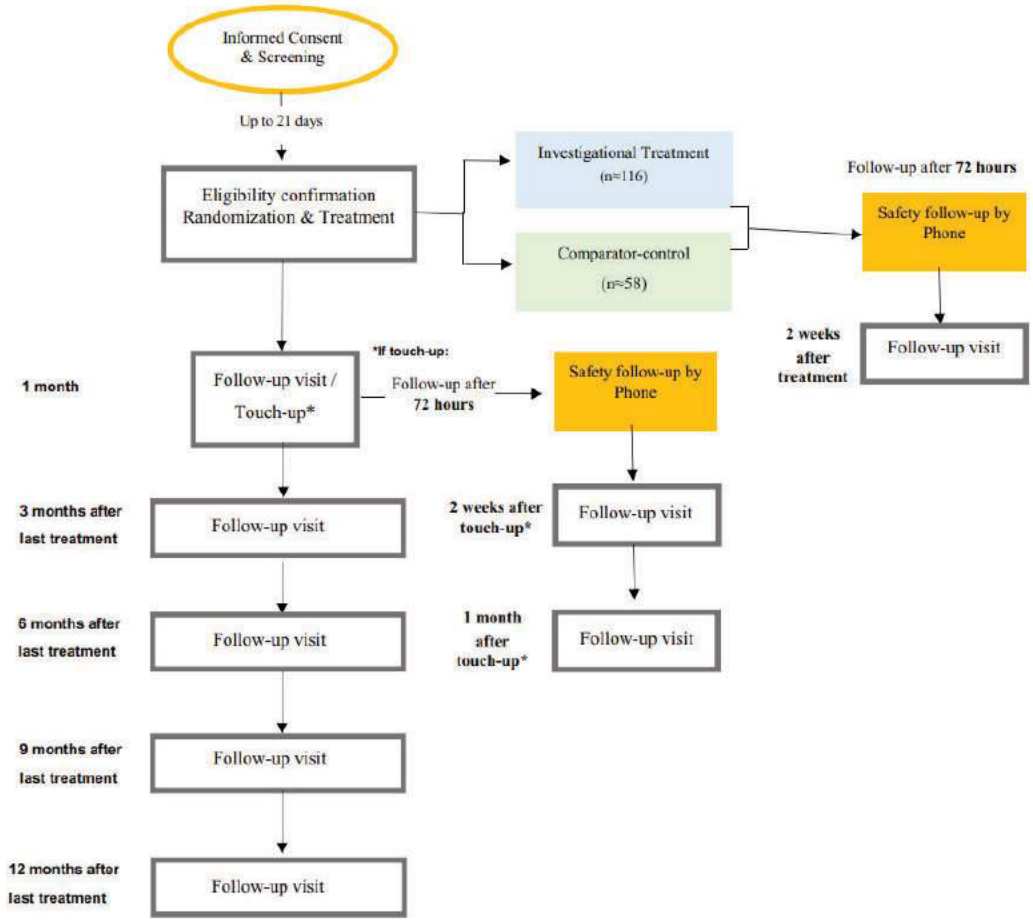
15.2 Changes in Analysis Planned in Protocol

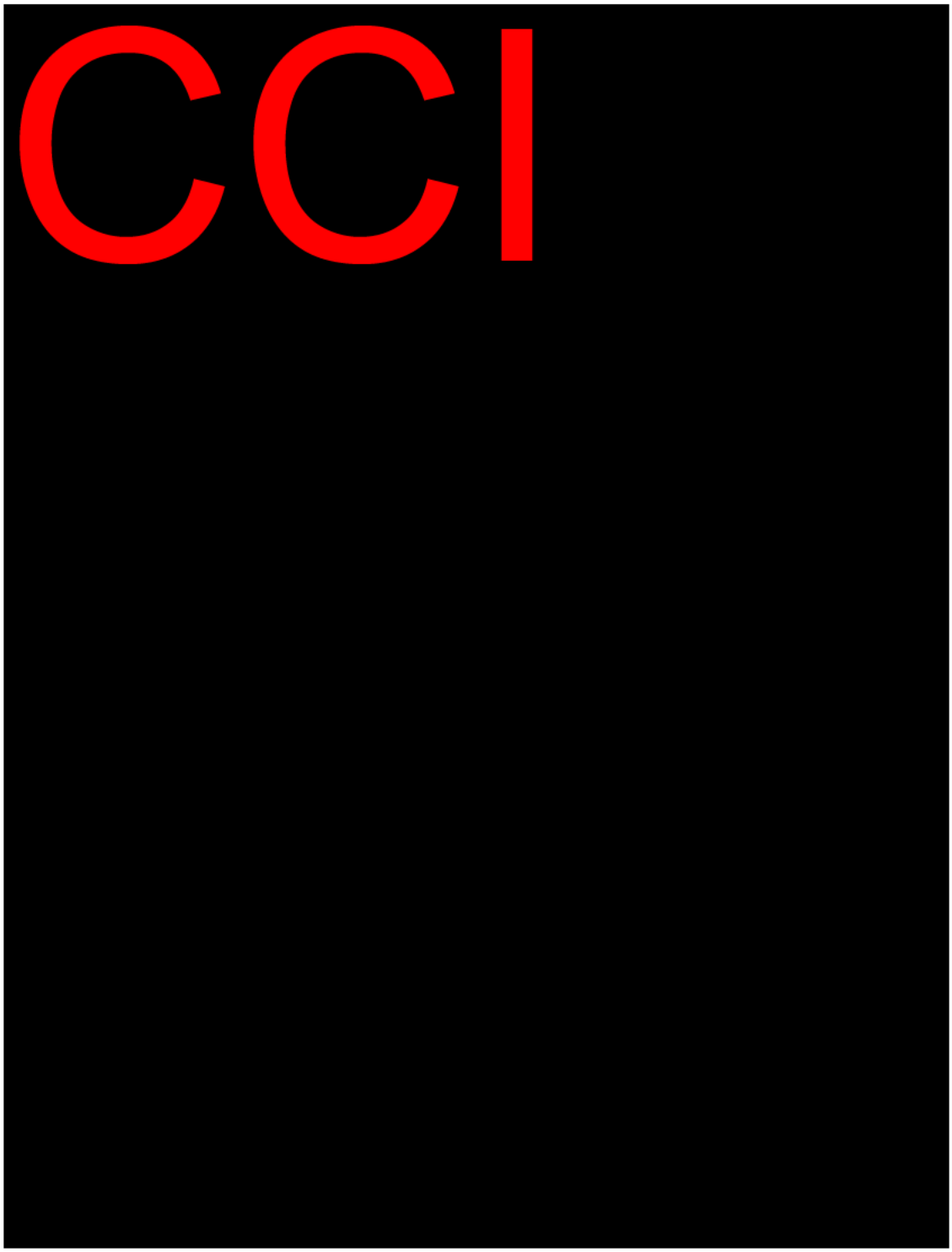
The Investigator Satisfaction Questionnaire is collected at the subject level instead of the investigator level. Due to the nature of the questionnaire, the analysis in the SAP differs from the protocol definition for this endpoint. Sections 3.2.4.8 and 9.2.8 have been updated to reflect these changes.

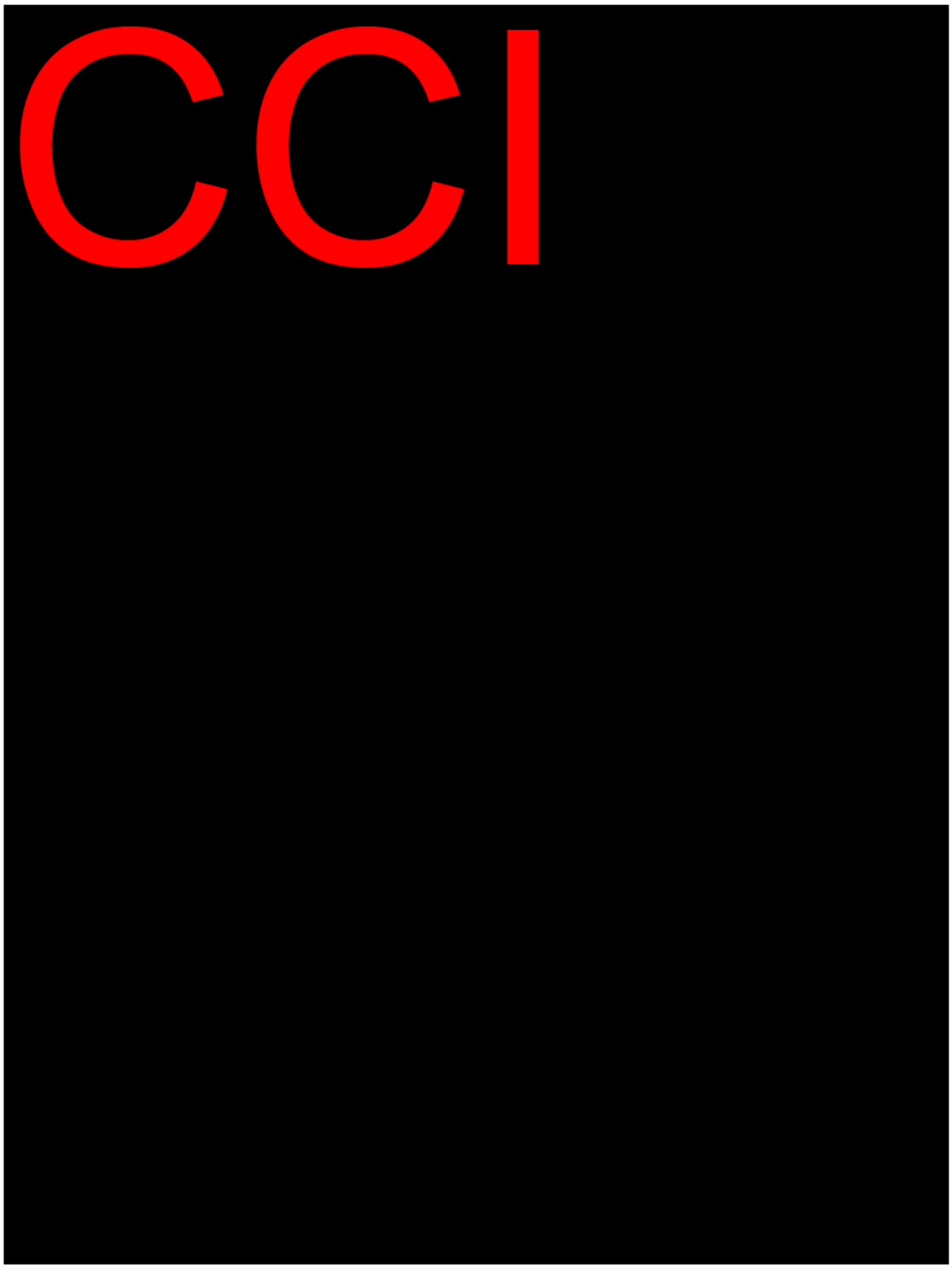
The protocol reads that "If the PP Population contains less than 90% of the subjects in the ITT, a sensitivity analysis of the primary effectiveness endpoint will be performed based on the PP Population. Regardless of the amount of overlap between the PP Population and the ITT Population, a sensitivity analysis will be performed using the PP Population. Sections 5.2 and 9.1.3 have been updated to reflect these changes.

16 Appendices

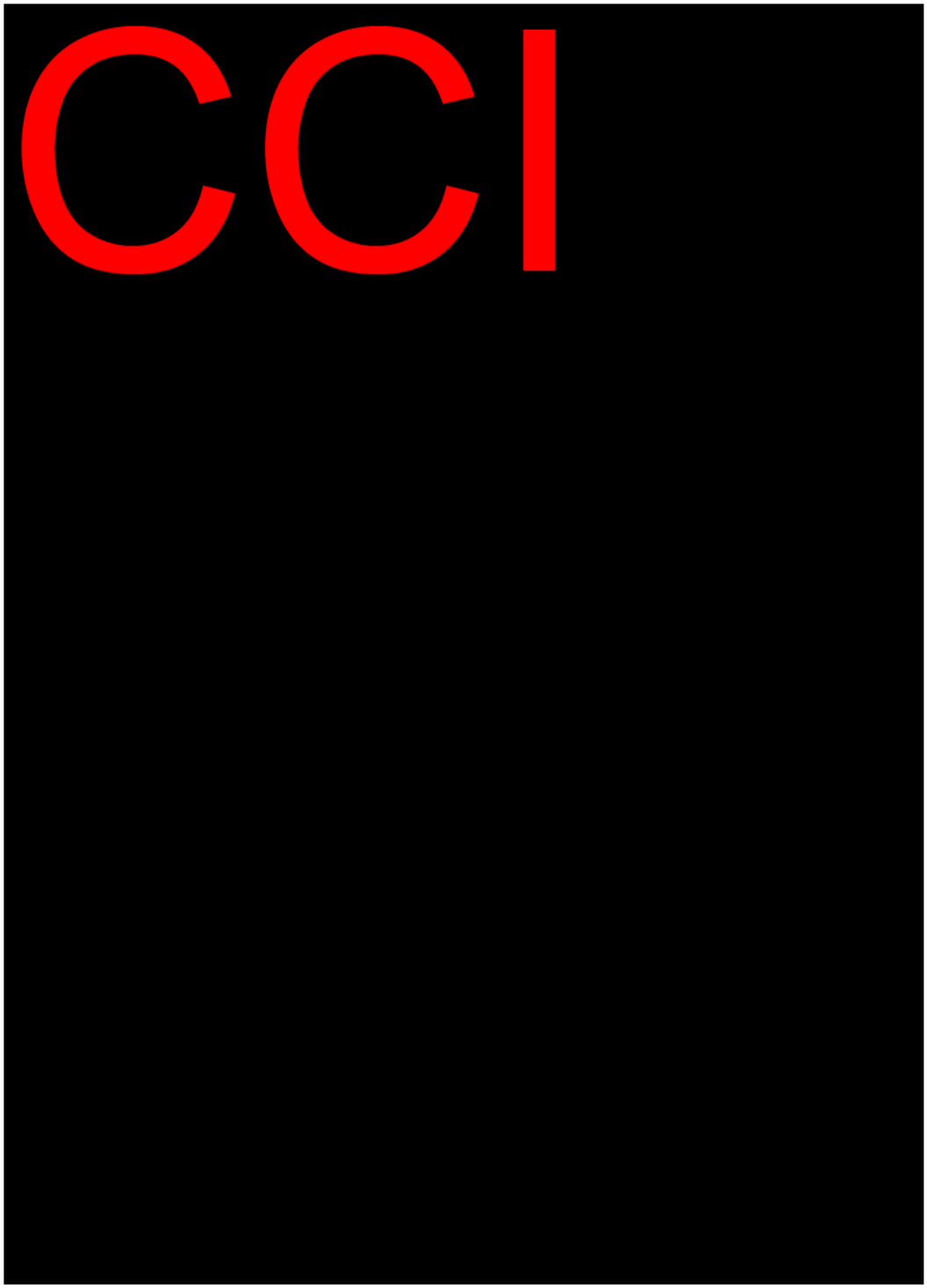
16.1 Study Flow Chart

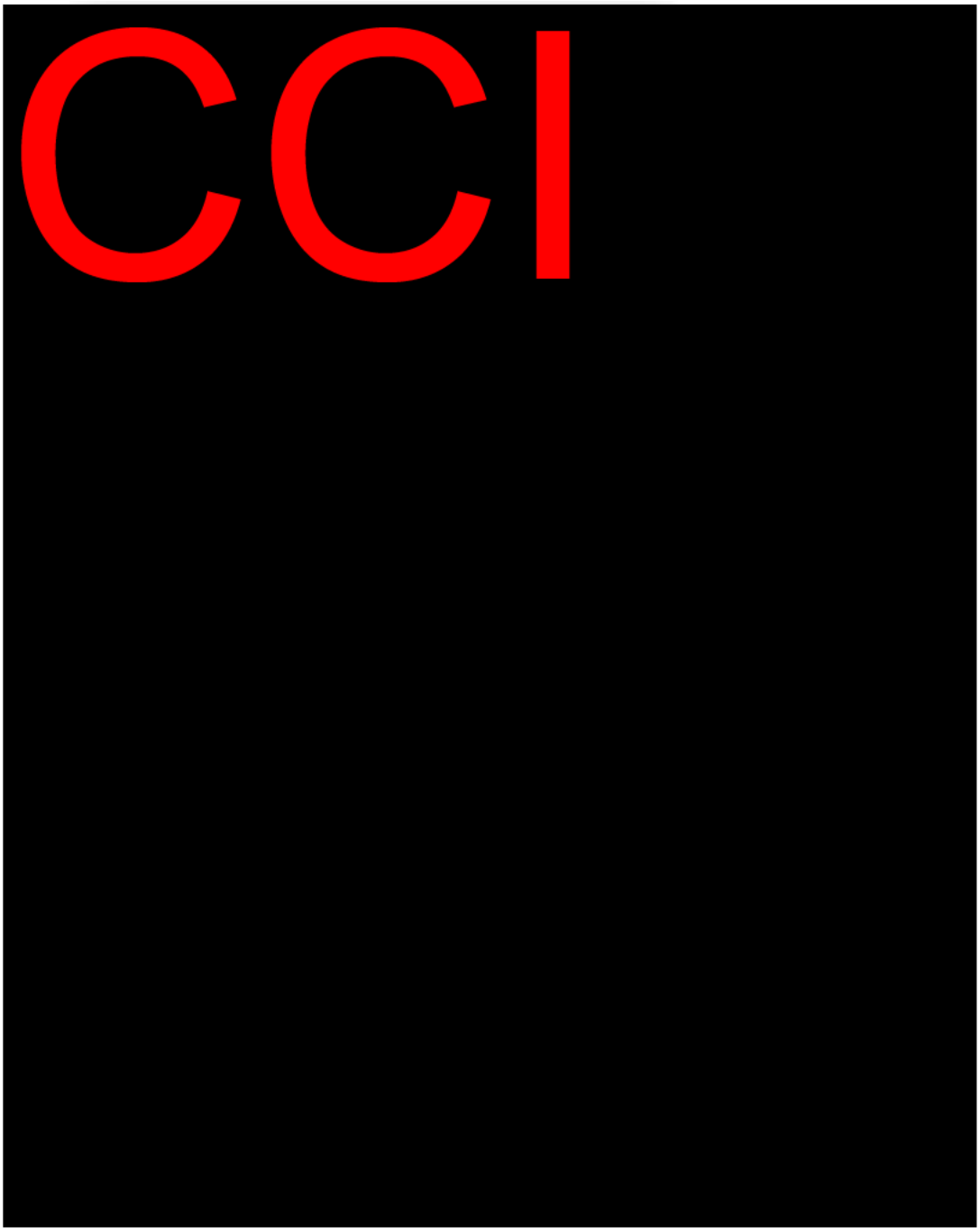


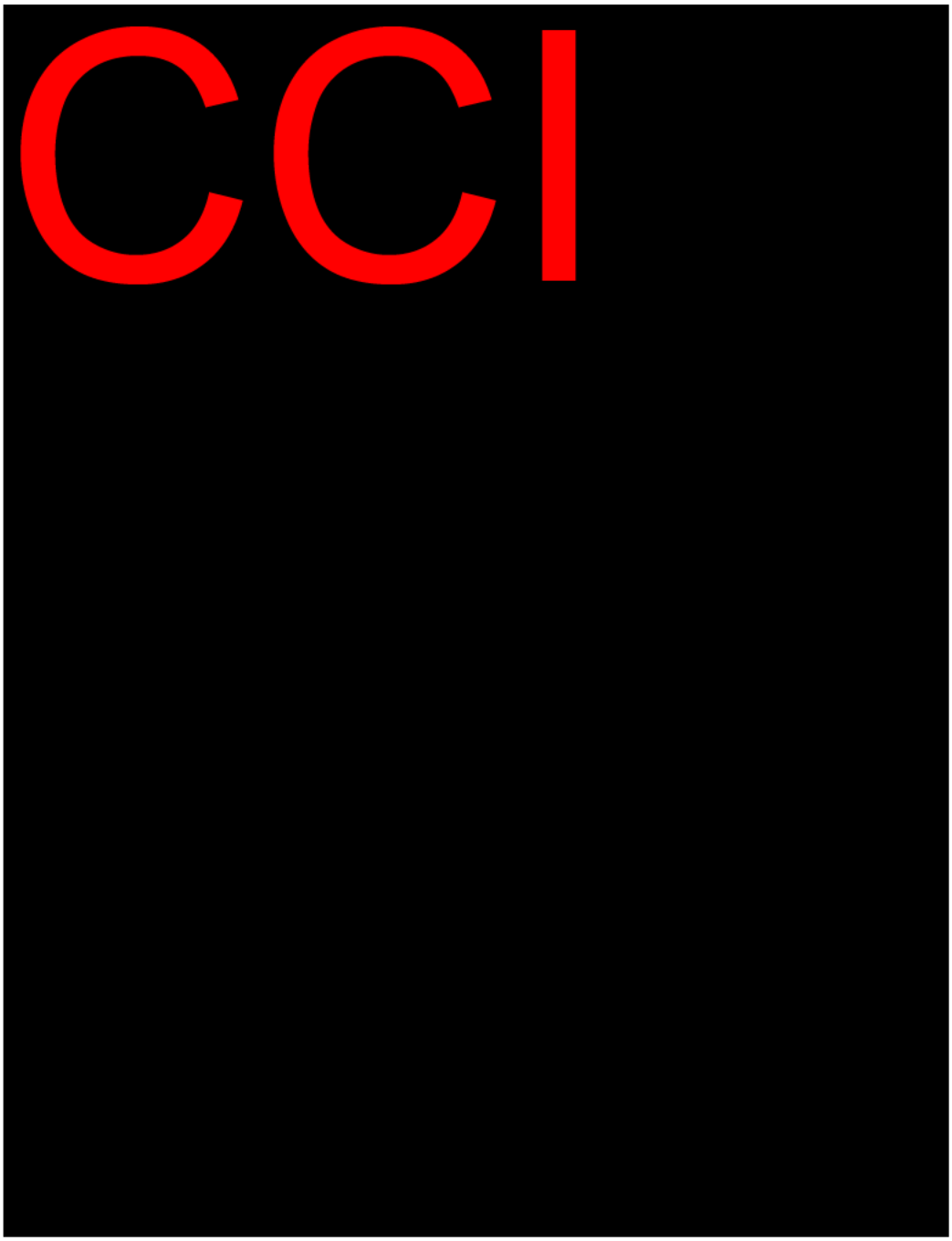





CCI







GALDERMA	Title **43USCH2208 Statistical Analysis Plan (SAP)**	Doc id **MA-60980**

SIGNATURES PAGE

PPD