Statistical Analysis Plan

Title:	Prospective, evaluator-blind, multicenter study to assess the safety and effectiveness of treatment with the Octave System for lifting the eyebrow and improving the appearance of the face and neck
Merz Study Number:	M960101001 / NCT05067504
SAP for	Final Analysis
Sponsor:	Ulthera, Inc., a division of Merz North America, Inc. /Merz Aesthetics
Version:	1.0
Date:	27-SEP-2022
Author:	PPD PPD Biostatistician, PPD PPD

Document References (For Internal Use Only)	
CCI [redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	

SIGNATURE PAGE

I confirm that this Statistical Analysis Plan accurately describes the planned statistical analyses to the best of my knowledge and was finalized before database close.

PPD

Date (dd-mmm-yyyy) Signature

PPD

Date (dd-mmm-yyyy) Signature

PPD

Date (dd-mmm-yyyy) Signature

Table of Contents

SIGNATURE PAGE 2
List of Abbreviations 5
1 General and Technical Aspects 6
2 Clinical Trial Design and Objectives 6
2.1 Clinical Study Design 6
2.2 Clinical Study Objectives 7
3 Determination of Sample Size 7
4 Analysis Sets 7
4.1 Safety Population (SP) 8
4.2 Intent-to-Treat Population (ITT) 8
4.3 Per Protocol Population (PP) 8
5 Endpoints for Analysis 8
5.1 Effectiveness Endpoints 8
5.1.1 Primary Effectiveness Endpoint 8
5.1.2 Secondary Effectiveness Endpoints 8
CCI [redacted]
5.2 Safety Endpoints 9
5.2.1 Secondary Safety Endpoint 9
CCI [redacted]
6 Statistical Analysis Methods 10
6.1 Effectiveness Endpoints 10
6.1.1 Primary Effectiveness Endpoint 10
6.1.2 Secondary Effectiveness Endpoints 13
CCI [redacted]
6.2 Safety Endpoints 16
6.2.1 Primary Safety Endpoint 16
6.2.2 Secondary Safety Endpoints 16
CCI [redacted]
6.4 Special Statistical/Analytical Issues 20
6.4.1 Discontinuations and Missing Data 20
6.4.2 Interim Analyses 22
6.4.3 Multiple Comparisons/Multiplicity 22
6.4.4 Examination of Subgroups 22
6.4.5 Pooling of Sites 22
CCI [redacted]
8 References 23
9 Appendix 24
CCI [redacted]


CCI

List of Abbreviations

AE	Adverse Event
AEH	Average eyebrow height
ATC	Anatomical Therapeutic Chemical classification system of the World Health Organization
BMI	Body Mass Index
CI	Confidence interval
COVID-19	Coronavirus disease 2019
CRO	Contract Research Organization
CCI	[redacted]
DRM	Data Review Meeting
FACE-Q™	Set of subject-reported questionnaire modules
FST	Fitzpatrick skin type
CCI	[redacted]
ITT	Intent-to-Treat population
MedDRA	Medical Dictionary for Regulatory Activities
MEH	Maximum eyebrow height change
PP	Per Protocol Population
PROC MI	SAS multiple imputation procedure
PT	Preferred term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System®
CCI	[redacted]
SOC	System organ class
SP	Safety Population
SSQ	Subject-Satisfaction Questionnaire
TC	Telephone contact
TEAE	Treatment-Emergent Adverse Event
TESAE	Treatment-Emergent Serious Adverse Event
TFL	Tables, Figures, and Listings
VAS	Visual analogue scale
WHO-DD	World Health Organization Drug Dictionary

1 General and Technical Aspects

The objective of this statistical analysis plan (SAP) is to specify the statistical analyses with appropriate detail and precision to serve as a guideline for statistical programming and creation of tables, figures, and listings for clinical study protocol M960101001, dated 16-NOV-2021 , version 6.0 (Amendment 5).

All programs will be written using SAS version 9.4 or higher. A preferred font size of 9 points, minimum font size of 8 points with a unique font size for the whole document required will be used for the tables and figures in Section 14. For listings, a standard font size of 9 points will be used to produce the output in A4 format. Individual SAS programs will be written for all tables, figures, and listings. All outputs will be transferred into PDF. These PDF files will be generated as needed to populate the subsections of Section 14 and Section 16.2 for the clinical study report. Each output file will include the corresponding table of contents, preceding the content of the file.

The Merz standard Tables, Figures, and Listings (TFLs) for medical devices, version 2.0, dated 18-Feb-2020, will be applied and adapted to study specific requirements as laid down in the clinical study protocol and any amendments. These mock TFLs will serve as study-specific output specifications for statistical programming.

Special attention will be paid to planning and performance of quality control measures. Risk scores based on assessments of complexity and impact of errors and quality control measures for statistical programming (including analysis datasets, TFLs) will be documented in the quality control plan for the creation of statistical output.

2 Clinical Trial Design and Objectives

2.1 Clinical Study Design

This is a 180-day, prospective, evaluator-blind, multicenter, staged, confirmatory study designed to evaluate the safety and effectiveness of the Octave System in lifting the brow and improving the appearance of the face and neck. Approximately 50 subjects will be enrolled at up to five investigational sites in the United States.

All Fitzpatrick skin types (FST) will be eligible for study enrollment. It is estimated that approximately 85% of the total sample size will consist of subjects with a FST I, II, or III, and approximately 15% of subjects will be FST IV, V, or VI. Subjects enrolled will be healthy male or female subjects, aged 35 to 65 years at the time of screening, with mild to moderate eyebrow/upper-face laxity and fine lines, wrinkles, and laxity of the mid- to lower face and neck that are amenable to improvement with non-invasive intervention and are not severe enough for surgical intervention.

In this non-comparative, staged, confirmatory study, investigators and subjects will not be blinded to treatment, subjects will not be randomized, and all enrolled subjects will receive a single Octave-Ultherapy treatment of the brow, upper, mid, and lower face, and neck. Subjects in Stage 1 (n = 10) will receive treatment with one transducer [redacted] at energy level 4 and one transducer [redacted] at energy level 2. Remaining subjects in Stage 2 will receive treatment with three transducers [redacted] at energy level 4 and one transducer [redacted] at energy level 2. In Stage 1, an initial cohort of subjects will be enrolled, safety data will be assessed to Day 14, and subjects will participate until study end at Day 90 or Day 180 if reconsented for extended participation. In Stage 2, all enrolled subjects will be followed for 180 days. In addition, to reduce potential bias associated with using only treated-subject photographs for evaluation, photographs of approximately 25 untreated individuals, who meet selected appearance criteria, will be collected at two different time points and subsequently paired for evaluation by the independent, blinded

evaluators. Using this 2:1 ratio of treated subjects to untreated individuals, image sets of these untreated individuals will be randomly distributed among the pre- and post-treatment image sets of the treated subjects and presented to the blinded evaluators. For the primary effectiveness endpoint, the "baseline" photograph will be identified as "Photograph 1", and the independent evaluators will be blinded to the randomly distributed treated versus untreated status of the comparator Photograph 2 (i.e., Day 90 post-treatment among treated subjects or post-baseline for untreated individuals). The same photographic assessment will be used for the secondary effectiveness endpoint with the comparator Photograph 2 collected at Day 180 post-treatment.

Study subjects will have approximately 4 on-site study visits, including screening, baseline/treatment, a Day 90 (± 7 days) post-treatment visit, and a Day 180 (± 14 days) post-treatment visit. Follow-up telephone calls (TCs) will occur at 3 days (± 1 day), 14 days (± 3 days), 28 days (+ 3 days), 120 days (± 7 days), and 150 days (± 7 days) after Octave-Ultherapy treatment. The primary effectiveness endpoint, the proportion of treated subjects with eyebrow lift, will be evaluated using a blinded, independent assessment of photographs at post-treatment Day 90 compared to baseline. Lift is concluded if at least two of three independent, blinded evaluators report "eyebrow lift" in the Day 90 standardized, photographic images when compared to baseline photographic images.

Secondary effectiveness evaluations will include a qualitative assessment of improvement in eyebrow lift at Day 180, as well as Day 90 and Day 180 evaluations of quantitative measurement of eyebrow lift and change from baseline in three FACE-Q modules (i.e., Satisfaction with Forehead and Eyebrows, Satisfaction with Lower Face and Jawline, and Patient-perceived Age visual analogue scale (VAS)).

Standard safety evaluations, including the incidence of adverse events (AEs) and serious adverse events (SAEs), will be assessed.

2.2 Clinical Study Objectives

Primary objective: To demonstrate the effectiveness and safety of the Octave System to lift the eyebrow.

Secondary objective: To demonstrate the effectiveness and safety of the Octave System to improve the appearance of the face and neck.

3 Determination of Sample Size

A responder rate of 72% is assumed for the proportion of subjects with eyebrow lift from baseline to Day 90, as defined by the correct selection of the Day 90 standardized, photographic images by two of three independent, blinded evaluators in the previous Ulthera System eyebrow clearance study (K072505). Monte-Carlo simulations were conducted to obtain sample sizes of evaluable subjects required to achieve an improvement rate (i.e., eyebrow lift reported by two of three independent, blinded evaluators) of ≥ 65% and lower bound of 1-sided 95% confidence interval (CI) above 50%. To achieve this threshold with ≥ 85% power, a sample size of 40 evaluable subjects will be required. Assuming a maximum attrition of 10 subjects, approximately 50 subjects will be enrolled. In Stage 1, 10 subjects will be enrolled, while all other subjects will be enrolled in Stage 2. Because a sample size of 40 subjects who complete the full Ultherapy treatment session is required, study enrollment will continue until the required number of treated subjects is achieved.

4 Analysis Sets

The following analysis sets will be defined for the statistical analysis of this clinical study:

4.1 Safety Population (SP)

The Safety Population (SP) includes all subjects who are enrolled and are treated with the Octave System.

4.2 Intent-to-Treat Population (ITT)

The Intent-to-Treat population (ITT) consists of all subjects who are enrolled.

4.3 Per Protocol Population (PP)

The Per Protocol population (PP) consists of all subjects in the ITT population who complete the full Ultherapy treatment session using the Octave System and do not have a major protocol violation. Major protocol deviations will be defined during the Data Review Meeting (DRM).

5 Endpoints for Analysis

5.1 Effectiveness Endpoints

5.1.1 Primary Effectiveness Endpoint

The primary endpoint of the study is the proportion of treated subjects with eyebrow lift, as demonstrated by comparing photographs taken at Day 90 (post-treatment) with baseline (pre-treatment) photographs. If the subject's screening photograph (Visit 1) is deemed acceptable upon review by the central photography vendor, it will default to the baseline (pre-treatment) photograph and will be used for reference throughout the study.

- Lift is concluded if at least two of three independent, blinded evaluators report "eyebrow lift" in the Day 90 standardized, photographic images when compared to baseline photographic images.
- The success criterion is defined by achieving simultaneously a point estimate ≥ 65% and a lower bound of the one-sided 95% Wilson confidence interval > 50% for the proportion of subjects with eyebrow lift.

5.1.2 Secondary Effectiveness Endpoints

- Proportion of treated subjects with eyebrow lift, as demonstrated by comparing photographs taken at Day 180 (post-treatment) with baseline (pre-treatment) photographs.
- Proportion of treated subjects with eyebrow lift of at least 0.5 mm on frontal 2D photographs at Day 90 and Day 180 compared to baseline.
- Change from baseline in FACE-Q Satisfaction with Forehead and Eyebrows score (Rasch-transformed) at Day 90 and Day 180, as assessed by the subject.
- Change from baseline in FACE-Q Satisfaction with Lower Face and Jawline score (Rasch-transformed) at Day 90 and Day 180, as assessed by the subject.
- Change from baseline in FACE-Q Patient-perceived Age Visual Analogue Scale (VAS) rating at Day 90 and Day 180, as assessed by the subject.

CCI

CCI

[redacted]

5.2 Safety Endpoints

5.2.1 Secondary Safety Endpoint

- Incidence of treatment-emergent adverse events (TEAEs) related to Octave-Ultherapy treatment, as reported throughout the study.

CCI

[redacted]

CCI [redacted]

6 Statistical Analysis Methods

Adequate descriptive statistics will be provided for each effectiveness and safety evaluation. Continuous variables (values and changes from baseline) will be summarized by number of observed values (n obs), number of missing values (n miss), number of imputed values (n imp, where applicable), mean, standard deviation, median, quartiles, minimum and maximum (i.e., metric statistics). Mean, quartiles, and median will be reported to one decimal place more than the data were collected, for the standard deviation two decimal places more will be displayed; for derived data, the number of decimal places will explicitly be given in the sections below.

Categorical variables will be summarized by frequencies and percentages (n, %) per category where the denominator will be chosen according to the adequate analysis population (i.e., frequency statistics). Frequency tables will include the number of missing values. Percentages will be calculated using the denominator of all subjects in a specified population. Percentages will be calculated using the number of subjects with non-missing data at the corresponding visit as denominator (observed cases). The denominator will be specified in a footnote to the tables for clarification if not otherwise obvious. Percentages will be reported to one decimal place.

Ordered categorical data will be summarized by quantitative and frequency statistics.

Shift tables, confidence limits (95%, two-sided), and descriptive p-values will be given, where appropriate.

Due to the COVID-19 pandemic outbreak, onsite visits might not be performed or only performed as virtual contact. In such cases, quantitative evaluations/assessments will not be performed.

CCI [redacted]

6.1 Effectiveness Endpoints

Unless otherwise specified, all effectiveness endpoints will be summarized for the ITT population.

6.1.1 Primary Effectiveness Endpoint

6.1.1.1 Primary analysis

The primary endpoint of the study is the proportion of treated subjects with eyebrow lift by comparing photographs taken at Day 90 (post-treatment) with baseline (pre-treatment) photographs.

Each independent evaluator will assess “Eyebrow Lift”, “No Eyebrow Lift” or “Cannot Evaluate” between Photograph 1 (baseline) and Photograph 2 (post-baseline). Photograph 1 will be identified, and the independent evaluators will be blinded to the random distribution of treated versus untreated status of the comparator Photograph 2 (i.e., Day 90 post-treatment among treated subjects or post-baseline for untreated individuals). Lift will be declared if at least two blinded evaluators indicate “Eyebrow Lift” for the comparator photograph. In case at least two evaluators declare a photograph as not assessable (i.e., “Cannot Evaluate” is ticked), the corresponding lift score over all evaluators will be set to missing for the primary analysis. In case one evaluator rates a photograph as “Cannot Evaluate” and the other two evaluators disagree, the overall lift score will be set to “No Lift”.

The point estimate for the proportion of subjects fulfilling the primary endpoint will be described by frequency statistics for the ITT subset.

Effectiveness is established if:

A. the point estimate is ≥ 65% and

B. the lower bound of the one-sided 95% Wilson CI is > 50%.

In addition, the proportion of untreated individuals fulfilling the primary endpoint will be calculated. It is expected that the point estimate for the proportion of untreated individuals is < 50%.

Missing data, including a missing primary endpoint assessment in subjects with unacceptable (i.e. unassessable) or missing photographs at baseline or Day 90, will be imputed by multiple imputation. In case of complete data, no multiple imputation will be performed and the primary analysis is based on complete data for the ITT.

Imputation is done 50 times using a fully conditional specification logistic regression method, as implemented in SAS PROC MI with statement FCS LOGISTIC, and will use the investigational site and the baseline characteristics age and BMI as independent variables.

In a first step a total of m=50 imputed datasets will be generated, where missing data in the primary endpoint (lift over all evaluators) is modelled via logistic regression based on input variables site, age and BMI at baseline. CCI [redacted] CCI [redacted] CCI [redacted] esponse is coded as ("no lift", "lift"). The lower bound of the one-sided 95% Wilson CI ([WL, WU]) will be calculated using the following steps, as detailed in [1]:

1. The mean sample response $\hat{p}_{Imp}$ and sample variance $\widehat{U}_{Imp}$,

$$\widehat{U}_{Imp} = \frac{\hat{p}_{Imp}\,(1-\hat{p}_{Imp})}{\mathrm{n}},$$

will be calculated for each of the *m=50* imputed datasets with sample size *n*.

2. The overall mean $\bar{p}_m$ and variance $\bar{U}_m$ over all imputed datasets will be calculated:

$$\bar{p}_m = \frac{1}{m}\sum_{Imp=1}^{m} \hat{p}_{Imp},$$

$$\bar{U}_m = \frac{1}{m}\sum_{Imp=1}^{m} \widehat{U}_{Imp}.$$

These will be used to obtain B_m,

$$B_m = \frac{1}{m-1}\sum_{Imp=1}^{m} (\hat{p}_{Imp} - \bar{p}_m)^2,$$

the between-sample variance.

3. The lower boundary of the one-sided 95% Wilson CI (WL and WU) will be calculated as

$$\left[\frac{2\bar{p}_m + \frac{t^2}{n} + \frac{t^2 r_m}{n}}{2\left(1 + \frac{t^2}{n} + \frac{t^2 r_m}{n}\right)} - \sqrt{\frac{\left(2\bar{p}_m + \frac{t^2}{n} + \frac{t^2 r_m}{n}\right)^2}{4\left(1 + \frac{t^2}{n} + \frac{t^2 r_m}{n}\right)^2} - \frac{\bar{p}_m{}^2}{1 + \frac{t^2}{n} + \frac{t^2 r_m}{n}}}\right],$$

where t denotes the 0.95 quantile of the t-distribution with degrees of freedom v,

$$v = (m-1)\left(1+\frac{1}{r_m}\right),$$

and with

$$r_m = \left(1+\frac{1}{m}\right)\frac{B_m}{\overline{U}_m}.$$

The study is considered successful, and consequently effectiveness is established, if the parameter estimate $\bar{p}_m \geq 65\%$ (condition A), and the lower limit of the one-sided 95% Wilson CI exceeds the margin of 50% (condition B).

6.1.1.2 Sensitivity analyses

To assess sensitivity of the results, the primary analyses will be repeated for:

- PP population with observed cases,
- ITT with observed cases, and
- ITT with missing value treated as no eyebrow lift (worst case analysis).

Observed cases are those not missing on primary analysis (section 6.1.1.1).

Additionally, further sensitivity analyses will be performed with regard to image quality as assessed by the central photography vendor. The following analyses will be performed, treating assessments that are based on photographs with not approved image quality as missing:

- PP population with observed cases,
- ITT with observed cases, and
- ITT with missing values imputed by multiple imputation.

For the sensitivity analysis using multiple imputation approach, similar methods as described in section 6.1.1.1 will be performed.

For all other sensitivity analyses, results can be taken directly from the SAS proc freq output:

CCI

, where response_data is the dataset with overall lift data per subject. The point estimate for the proportion of responders and the lower limit of the one-sided 95% Wilson CI will be used to assess sensitivity.

A sensitivity analysis will not be done, if identical to any other analysis.

6.1.1.3 Subgroup analyses

Gain or loss of ≥ 2 BMI units from screening to Day 90 will be considered a major protocol deviation. Subjects with gain or loss of ≥ 2 BMI units from screening to Day 90 will be excluded from PP population. The following subgroups will be defined by the change in body weight from screening to Day 90:

- Subjects in ITT population with gain of ≥ 2 BMI units;
- Subjects in ITT population with loss of ≥ 2 BMI units; and
- Subjects in ITT population with change of < 2 BMI units.

Primary endpoint with observed cases will be analyzed in all subgroups, as defined above, with more than five subjects.

Frequency tables for the proportion of treated subjects with eyebrow lift (primary endpoint) will be provided by age categories (≤ median age, > median age), BMI at baseline in categories (<18.5, 18.5 – 24.9, 25 – 29.9, ≥30.0) and by investigational site.

6.1.2 Secondary Effectiveness Endpoints

Independent Photographic Review for Eyebrow Lift at Day 180

The secondary endpoint will only be assessed among Stage 1 subjects who reconsented for extended study participation and all Stage 2 subjects.

The secondary endpoint of the study is the proportion of treated subjects with eyebrow lift by comparing photographs taken at Day 180 (post-treatment) with baseline (pre-treatment) photographs.

Each independent evaluator will assess "Eyebrow Lift", "No Eyebrow Lift" or "Cannot Evaluate" between Photograph 1 (baseline) and Photograph 2 (post-baseline). Photograph 1 will be identified, and the independent evaluators will be blinded to the random distribution of treated versus untreated status of the comparator Photograph 2 (i.e., Day 180 post-treatment among treated subjects or post-baseline for untreated individuals). Lift will be declared if at least two blinded evaluators indicate "Eyebrow Lift" for the comparator photograph. In case at least two evaluators declare a photograph as not assessable (i.e., "Cannot Evaluate" is ticked), the corresponding lift score over all evaluators will be set to missing for the secondary analysis.

The point estimate for the proportion of subjects fulfilling the qualitative eyebrow lift will be described by frequency statistics for the ITT subset, including the one-sided 95% Wilson CI.

Missing data, including a missing secondary endpoint assessment in subjects with unacceptable or missing photographs at baseline or Day 180, will be imputed by multiple imputation 50 times. Imputation is done using a fully conditional specification logistic regression method, as implemented in SAS PROC MI with statement FCS LOGISTIC, and will use the investigational site and the baseline characteristics age and BMI as informing variables. Point estimate and Wilson CI will be calculated with methods as described in Section 6.1.1.1. For implementation, SAS code similar to the statements for the primary endpoint will be used.

Sensitivity analyses and subgroup analyses will be performed similar to the primary endpoint.

Quantitative Eyebrow Lift

Treatment-related lift in eyebrow position on frontal view will be measured by calculating the differences in eyebrow height using baseline and Day 90 and Day 180 images.

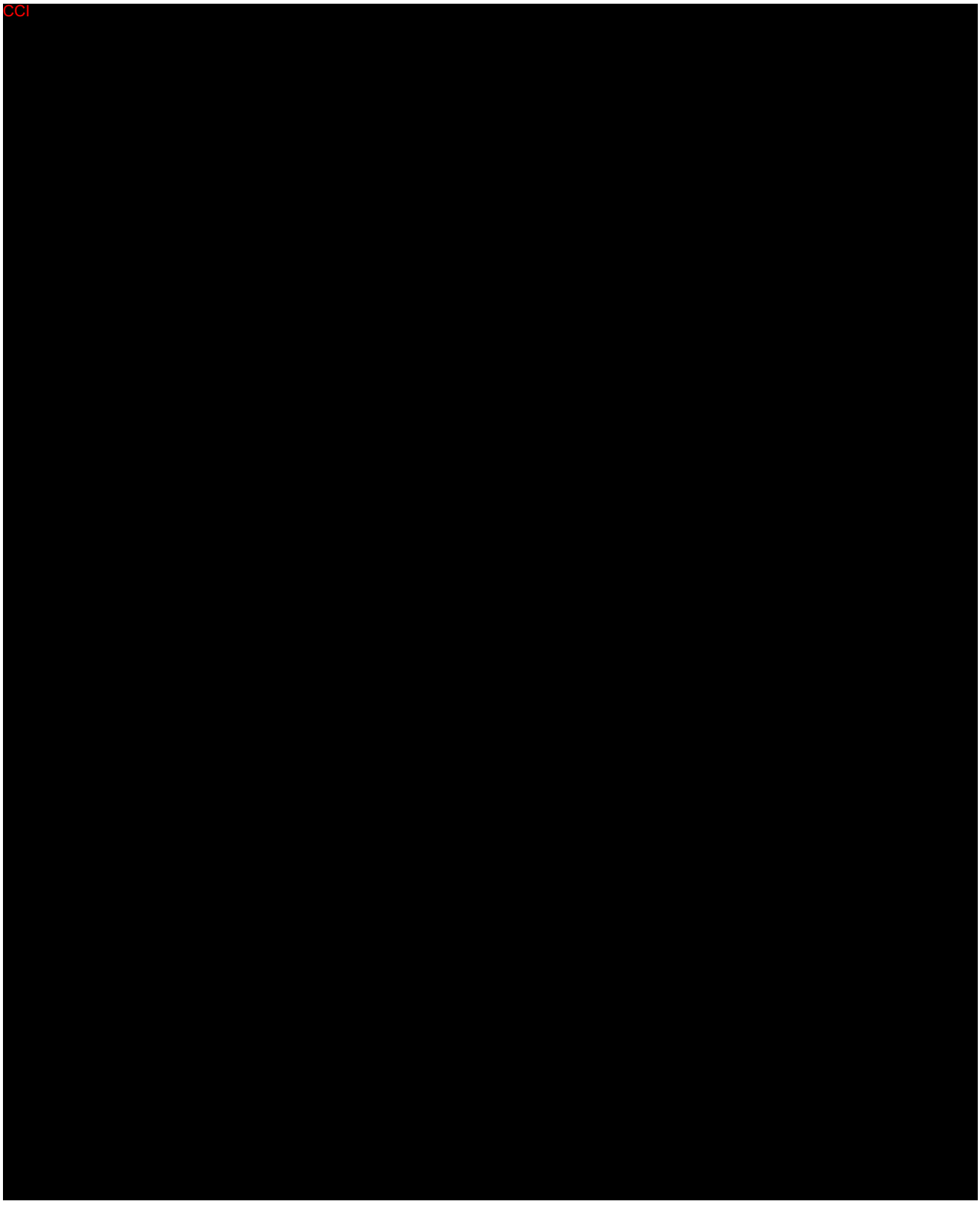
Quantitative results will be calculated using five evenly spaced points on the brow. Height will be calculated at each of the five points and subsequently averaged to generate the average eyebrow height (AEH). Eyebrow lift will be indicated by the calculated change in AEH between baseline and Day 90 and Day 180 height of the brow. For the calculation of change in AEH, only datapoints with non-missing values for both visits (baseline and either Day 90 or Day 180) will be used.

To assess eyebrow lift from baseline to Day 90 and Day 180, the proportion of responders, defined as subjects achieving ≥ 0.5 mm eyebrow lift at Day 90 and Day 180 compared to baseline, will be described by frequency statistics for observed cases in the ITT. Success for the quantitative eyebrow lift will be established if the lower bound of the one-sided 95% Wilson CI for proportion of responders is > 50%.

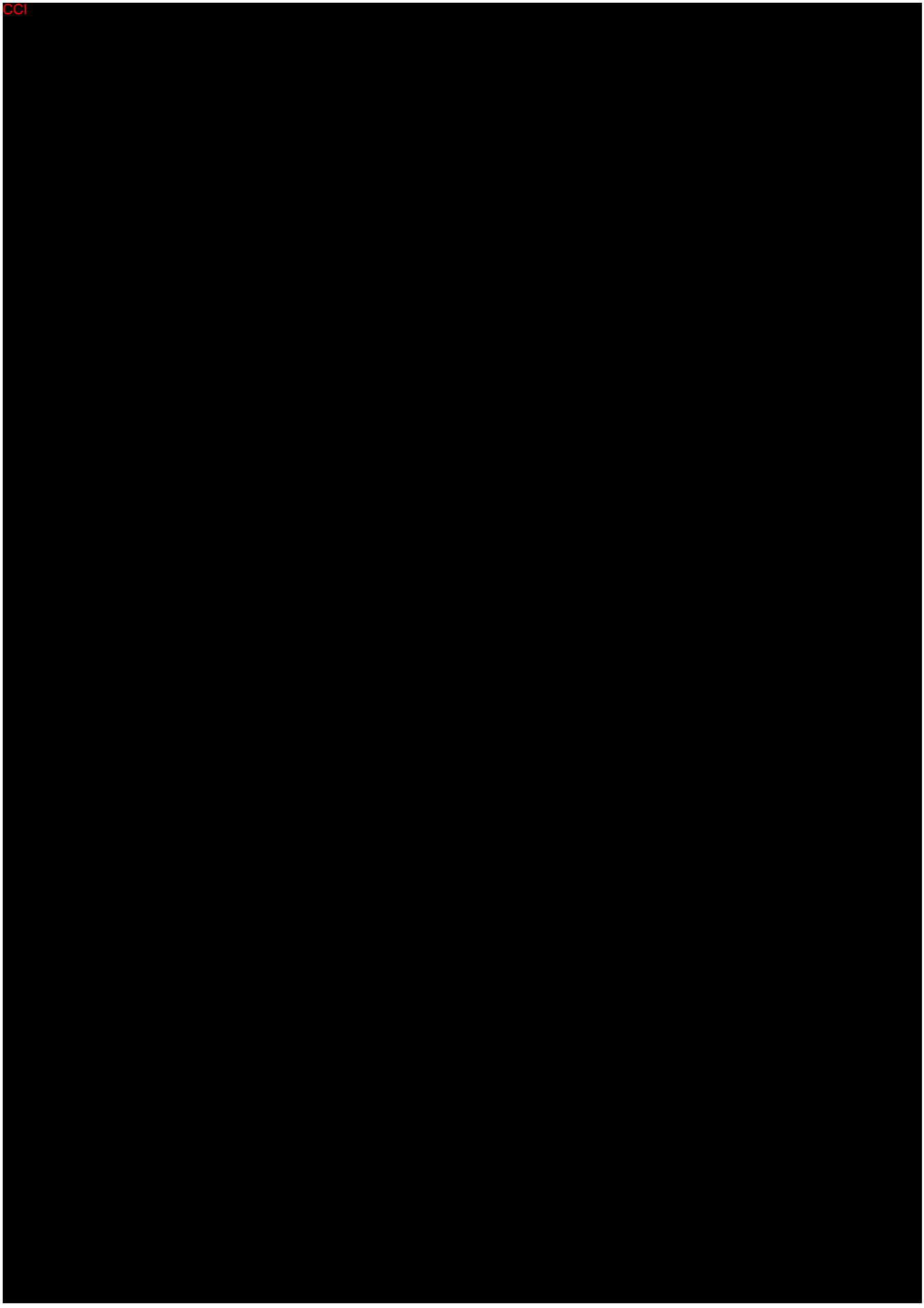
Time course of absolute values and changes from baseline in mm will be described by metric statistics. Additionally, the change in height from baseline at Day 90 and Day 180 will be calculated for each brow point, and maximum eyebrow height change (MEH) will be presented as maximum over all brow points over both sides.

These analyses will be provided by facial side (left/right) and overall. For the overall analysis, a subject will be responder if ≥ 0.5 mm eyebrow lift for both sides is achieved, while metric statistics will be based on the average of the right and left sides on frontal view.

CCI



CCI



CCI

[redacted]

6.2 Safety Endpoints

All safety analyses will be summarized for observed values in the SP.

6.2.1 Primary Safety Endpoint

Not applicable.

6.2.2 Secondary Safety Endpoints

All AEs will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) version in effect at the time the database is closed.

Only treatment-emergent AEs (TEAEs) will be analyzed, which are defined as AEs with onset or worsening at or after the first administration of study treatment. In this regard, an AE, with onset prior to treatment, that worsens at or after first administration of study treatment must be documented as a new TEAE with onset at the time of worsening.

An overall summary of AEs will be provided for the following:

- Any AEs;
- Any non-TEAEs;
- Any TEAEs;
- Any treatment-related TEAEs;
- Any serious TEAEs;
- Any serious treatment-related TEAEs;
- Any TEAEs leading to discontinuation;
- Any treatment-related TEAEs leading to discontinuation;
- Any fatal TEAEs, and
- Any treatment-related fatal TEAEs.

Incidences will be provided for the following classes of AEs:

- TEAEs, subjects with TEAEs and number of TEAEs by MedDRA system organ class (SOC) and preferred term (PT);
- TEAEs, subjects with TEAEs by PT;
- Non-TEAEs, subjects with Non-TEAEs and number of Non-TEAEs by SOC and PT;
- Non-serious TEAEs, subjects with TEAEs and number of TEAEs by SOC and PT;
- TEAEs by maximum intensity (mild, moderate, severe), subjects with TEAEs by PT;
- TEAEs by worst causal relationship (related, not related), subjects with TEAEs by PT;
- TEAEs by worst outcome, subjects with TEAEs by PT;
- TEAEs by affected treatment area, subjects with TEAEs by PT;
- TEAEs by maximum duration (1-3 days, 4-7 days, 8-14 days, ≥15 days), subjects with TEAEs by PT;

- TEAEs by maximum duration and maximum intensity, subjects with TEAEs by PT;
- Related TEAEs, subjects with TEAEs and number of TEAEs by SOC and PT ;
- Related TEAEs, subjects with TEAEs by PT;
- Non-serious related TEAEs, subjects with TEAEs and number of TEAEs by SOC and PT;
- Related TEAEs by maximum intensity, subjects with TEAEs by PT;
- Related TEAEs by worst causal relationship, subjects with TEAEs by PT;
- Related TEAEs by worst outcome, subjects with TEAEs by PT;
- Related TEAEs by affected treatment area, subjects with TEAEs by PT;
- Related TEAEs by maximum duration (1-3 days, 4-7 days, 8-14 days, ≥15 days), subjects with TEAEs by PT;
- Related TEAEs by maximum duration and maximum intensity, subjects with TEAEs by PT;
- Serious TEAEs, subjects by SOC and PT;
- Related serious TEAEs, subjects by SOC and PT;
- Serious TEAEs, listing of subjects (in case of more than 15 serious TEAEs respective table by PT and SOC will be provided); and
- Related serious TEAEs, listing of subjects (in case of more than 15 related serious TEAEs respective table by PT and SOC will be provided).

All tables will be sorted by descending frequencies in Preferred Terms (within SOCs).

The duration of AEs will be calculated by the difference of start and stop day plus 1.

If a subject has more than one outcome within a PT only the worst outcome in the respective time period will be used in the frequency tables. Also on subject level, only the worst outcome category per subject in the respective time period will be counted in the frequency table. The worst outcome is defined in the following order (from best to worst):

- recovered/resolved
- recovered/resolved with sequelae
- recovering/resolving
- not recovered/not resolved
- unknown
- fatal

Moreover, incidences of TEAEs, by PT and SOC, will be provided for the following subgroups:

- Race
- Fitzpatrick Skin Types (I, II, III versus IV, V, VI).

Listings for all AEs, as well as subsets including AEs leading to discontinuation, related SAEs, and deaths, will be provided. AEs with delayed onset will be flagged. A delayed onset AE is defined as an AE with an onset more than 28 days after treatment.

Treatments discontinued due to pain and device deficiencies will also be listed.

CCI

[redacted]

Disposition of subjects

The absolute and relative frequencies for subjects' main reason for premature study discontinuation will be tabulated, based on all treated subjects. Furthermore, the number of subjects screened, the number of subjects treated, and the number of subjects in respective analysis sets will be tabulated.

Visit attendance will be presented in a frequency table for all treated subjects. For 'Subjects expected at visit', the following visits, at a minimum, are defined as expected for all treated subjects: Screening and Day 1 (Baseline). For subjects without reconsent in Stage 1, all following telephone contacts and visits are expected until Day 90 or premature study discontinuation, whatever occurs first. For subjects in Stage 1 with reconsent and all subjects in Stage 2, all following telephone contacts and visits are expected until regular study completion at Day 180 or premature study discontinuation, whatever occurs first. For subjects who discontinue the study, the end of study visit (Day 90 for subjects without reconsent or Day 180 for subjects with reconsent) is expected. Visits after premature discontinuation are generally not expected.

Percentages of 'Subjects expected at visit' will be calculated relative to the total number of treated subjects for Screening visit until Day 90 visit and relative to the number of subjects with reconsent from TC (Day 121) until Day 180 visit.

Percentages relating to 'Visit performed' and 'Visit not performed' will be calculated relative to the number of subjects expected at the respective visit. For study discontinuations, only absolute numbers will be displayed.

Protocol deviations

The absolute and relative frequencies of subjects with major protocol deviations and other reasons for exclusion from analysis sets will be tabulated, based on all treated subjects.

Demographic data and baseline characteristics

Demographic data will be summarized for the SP, the ITT, and the PP. Fitzpatrick skin type will be presented by original value as well as by categories type I-III and type IV-VI. Age will additionally be presented by categories <35 years, 35-40 years, 41-50 years, 51-60 years, 61-65 years and >65 years.

Previous and concomitant medications and non-drug treatments

Separation of previous from concomitant medications will be done according to the start and stop date of the medication in comparison to the date of treatment. Each medication will be allocated unambiguously either to previous medication or to concomitant medication.

- Previous medication: If stop date is before start of treatment.
- Concomitant medication: If the start date is at or after start of treatment or if the stop date is at or after start of treatment or ongoing is ticked.

Previous and concomitant medications will be coded by use of the World Health Organization Drug Dictionary (WHO-DD) whereby the version in effect at the time the database is closed will be used.

Frequencies of previous and concomitant medications will be given on the basis of various Anatomical Therapeutic Chemical classification system of the World Health Organization (ATC) code levels for the SP. Indications for concomitant medications will not be coded and will only be listed.

Non-drug treatments will be coded using MedDRA version which is in effect at the time the database is closed. Non-drug treatments will be reported by SOC and PT levels for the SP.

Medical history and concomitant diseases

Separation of medical history from concomitant diseases will be done according to the stop date of the finding in comparison to the day of treatment. Each finding will be allocated unambiguously either to medical history or to concomitant diseases.

- Medical history: If stop date is before start of treatment.
- Concomitant disease: If the stop date is at or after start of treatment or ongoing is ticked (even if it refers to a cut-off point before start of treatment).

Rules for separation of medical history and concomitant diseases for missing or incomplete dates are given in the Appendix.

Medical history and concomitant diseases will be coded using the MedDRA dictionary and reported by SOC and PT levels for the SP.

Pain pre-treatment medication

Frequency tables of pre-treatment pain medication (initial, additional, and/or administration of nitrous oxide / oxygen mixture) will be given for the SP.

Extent of exposure

A descriptive table for the number of lines by transducer, energy level, and treatment region will be produced for the SP, containing number of observed values (n obs), number of missing values (n miss), mean, standard deviation, median, quartiles, minimum, and maximum.

Frequency tables for premature treatment discontinuation and areas treated will be given for the SP.

Vital Signs

A descriptive summary table will be presented for weight, height, and BMI by visit (i.e., Visits 1, 2, and 3 only; height will be collected at Visit 1 only).

Urine pregnancy test

Results of urine pregnancy tests will be listed.

Device deficiencies

All device deficiencies including device deficiencies for screening failures will be listed.

6.4 Special Statistical/Analytical Issues

6.4.1 Discontinuations and Missing Data

Independent Photographic Review for Eyebrow Lift at Day 90 and Day 180

For analysis of the primary and secondary effectiveness endpoints regarding Independent Photographic Review for Eyebrow lift, the following procedure for incomplete data from the independent evaluators will be applied:

- If data from all three independent evaluators is missing, the outcome of the subject will be set to missing.
- If the result of two out of three evaluators is missing, the outcome of the subject will be set to missing.

- If the result of one evaluator is missing and the other two evaluators agree on the assessment “Eyebrow Lift” then the outcome of the subject is set to “lift” for all analyses.
- If the result of one evaluator is missing and the other two evaluators agree on the assessment “No Eyebrow Lift” then the outcome of the subject is set to “no lift” for all analyses.
- If the result of one evaluator is missing and the other two evaluators disagree, then the outcome of the subject will be set to missing.

For treated subjects with missing outcome, the multiple imputation technique will be employed as described above for the primary and secondary analyses. For the worst case analysis, these subjects will be analyzed as having “no lift”.

All other effectiveness data will be analyzed as observed with no imputation of missing values. The denominator for percentages will be all subjects in the respective analysis set or population (i.e., missing data for binary or categorical variables will be presented as a separate possible category ‘Missing’).

CCI

Adverse events

Missing data of the outcome will be imputed by “unknown”. For any other missing data of adverse events, a worst case strategy will be applied for all analysis tables. The intensity will be imputed

by the worst intensity “severe”, the causal relationship will be imputed by the worst relationship “related”.

6.4.2 Interim Analyses

In Stage 1, an initial cohort of 10 subjects was enrolled and treated with the 3.0 mm and 1.5 mm transducers. An interim database close and interim safety analysis through Day 14 was conducted for Stage 1 subjects. Safety analyses were performed on the SP. Effectiveness data were not included in the interim analysis.

Continuation to Stage 2 was paused during evaluation of the Day 14 safety data from Stage 1 subjects. At the conclusion of the safety assessment for Stage 1, expansion to use the 4.5 mm transducers was considered based on the totality of information available from the Stage 1 safety data, MDR analysis, and the mitigation measures/strategies proposed for this IDE based from the MDR analysis. All Stage 1 subjects participated until study end at 90 days (± 7 days) or 180 days (± 14 days) if reconsented for extended participation.

The interim analysis was only intended to obtain safety data for the first 10 subjects treated before starting recruitment for Stage 2. The interim analysis was not proposed to modify the study design, statistical analyses, or sample size, or for study termination, as this was not an adaptive study design. Since no effectiveness data were analyzed at interim, the interim analysis did not compromise the integrity of the study’s effectiveness data, and it did not bias the final effectiveness analysis.

A separate statistical analysis plan for the interim analysis was written.

6.4.3 Multiple Comparisons/Multiplicity

Primary analysis consists of two conditions which both need to be fulfilled to be able to claim effectiveness, therefore no adjustment of the type-1 error is needed.

6.4.4 Examination of Subgroups

The primary endpoint of observed cases will be additionally analysed by the different subgroups as described in section 6.1.1.3. Similar subgroups will be used for the secondary endpoint Independent Photographic Review for Eyebrow Lift at Day 180.

The safety endpoints for incidences of TEAEs using the SP will be descriptively summarized by Fitzpatrick skin type (I-III versus IV-VI) and race, as described in section 6.2.2.

Subjects with multiple origins (races) will be presented within a separate subgroup category “More than One Race”.

No statistical comparison of subgroups is planned.

6.4.5 Pooling of Sites

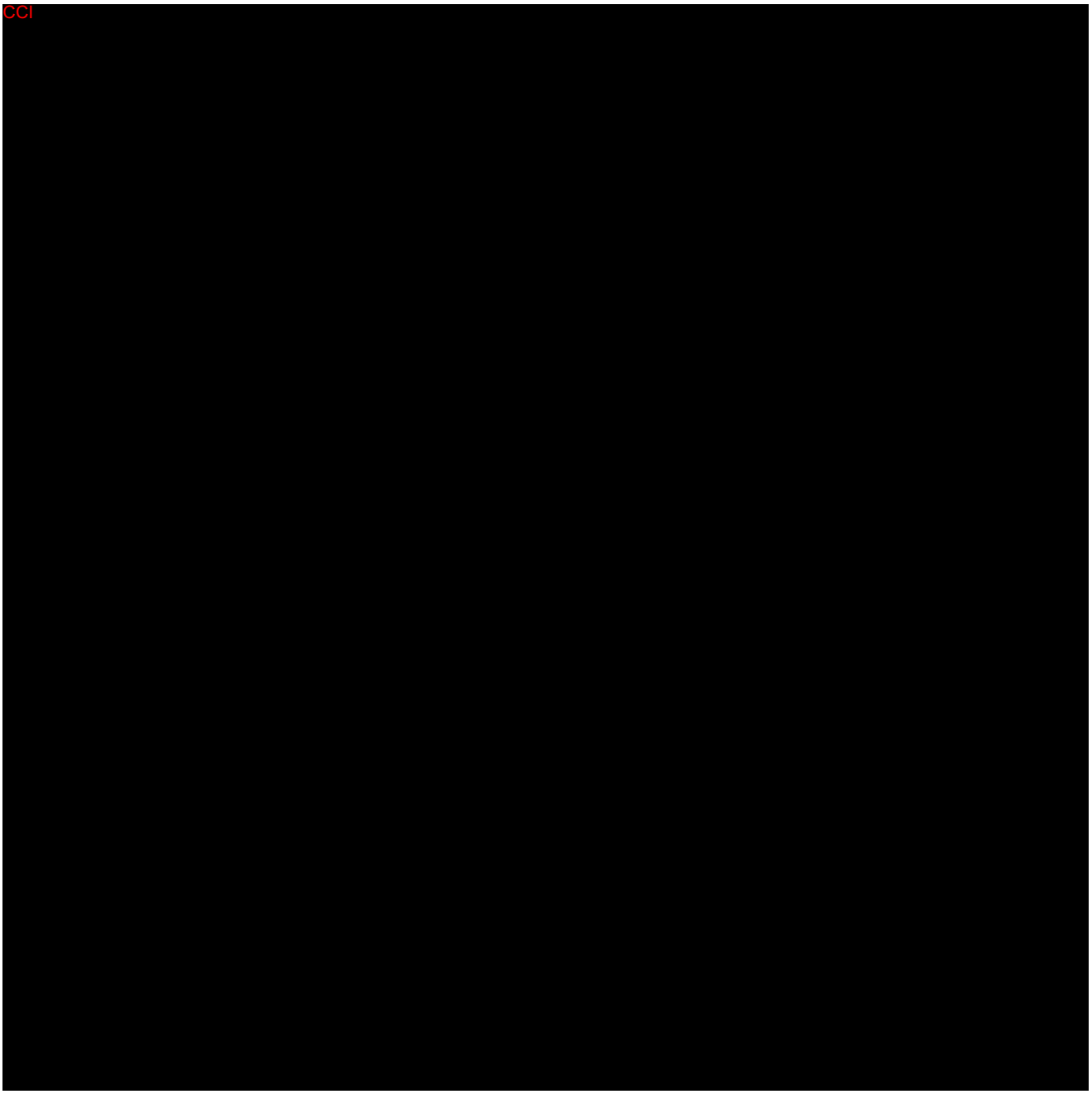
Analysis will be performed on pooled data across all investigational sites.

Treatment site will be used as explanatory variable for the multiple imputation model. In case it turns out that convergence of the model cannot be accomplished, the two smallest sites will be merged and a new model will be calculated. This process will be repeated until convergence is established. The pooled site will be coded as ‘0099999’.

CCI

[redacted]

CCI



8 References

[1] Anne Lott & Jerome P. Reiter (2018): Wilson Confidence Intervals for Binomial Proportions With Multiple Imputation for Missing Data, The American Statistician, DOI: 10.1080/00031305.2018.1473796