Statistical Analysis Plan

Title:	Prospective, multicenter, controlled, evaluator-blind, randomized study to investigate the effectiveness and safety of diluted RADIESSE® for treatment of décolleté wrinkles
Merz Study Number:	M930521003 / NCT05163353
SAP for	Interim and Final Analysis
Sponsor:	Merz North America, Inc.
Version:	1,0
Date:	28-Jul-2023
Author:	PPD Biostatisician, PPD PPD

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

1.1 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 interim analysis.

PPD

Date (dd-mmm-yyyy)	Signature
Date (dd-mmm-yyyy)	Signature
Date (dd-mmm-yyyy)	Signature

Table of Contents

List of Abbreviations

AE	Adverse Event
CI	Confidence Interval
CSP	Clinical Study Protocol
CCI	[redacted]
DRM	Data Review Meeting
CCI	[redacted]
FST	Fitzpatrick skin type
iGAIS	Investigator Global Aesthetic Improvement Scale
IMD	Investigational medical device
CCI	[redacted]
ITT	Intent-to-treat
MAS	Merz Aesthetic Scale
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
PPS	Per Protocol Set
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System®
SES	Safety Evaluation Set
sGAIS	Subject Global Aesthetic Improvement Scale
SOC	System Organ Class
CCI	[redacted]
TEAE	Treatment Emergent Adverse Event
TESAE	Treatment-emergent serious adverse event
TFLs	Table, Figures, and Listings
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 M930521003, dated 04-Aug-2020 and the following amendments, dated 29-Jul-2021, 02-Sep-2021, 26-Oct-2021, and 10-Mar-2022. This SAP describes the interim analysis as well as the final analysis at the end of the study. This SAP will be finalized before interim analysis and it might be amended before the final analysis, if required.

All programs will be written using Statistical Analysis System® (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 or directly created as PDF files. 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.

CCI Table, Figures, and Listings (TFLs) for medical devices, version 2.0, dated 18-Feb-2020, will be applied and adapted to trial specific requirements as laid down in the Clinical Study Protocol (CSP), 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 prospective, multicenter, randomized, evaluator-blinded, parallel-group study with a main period and a safety follow-up period, designed to evaluate the effectiveness and safety of subdermal implantation of diluted Radiesse for the correction of moderate to severe décolleté wrinkles. Healthy female subjects aged 30 to 65 years at the time of screening, CCI [redacted] as assessed live by a blinded evaluator, will be eligible for enrollment.

Approximately 152 subjects will be randomized at up to ten study sites in the US. CCI [redacted] At each study site, the number of subjects randomized is intended to be at least 8 and should not exceed 24 to ensure a reasonable distribution of subjects across all investigational sites. Females of all Fitzpatrick skin types (FST) are eligible for study participation. CCI [redacted]

All eligible subjects will be randomized at a 3:1 ratio to treatment with diluted Radiesse (Group A, from here on labeled "Treatment") or to untreated control followed by delayed treatment with diluted Radiesse (Group B, from here on labeled "Control/Delayed treatment"). For each treatment, a maximum 4.5 mL total volume (1.5 mL of Radiesse diluted 1:2 with 3.0 mL of sterile saline) will be injected using a cannula in the décolleté treatment area. Subjects will be followed

[redacted] to Week 84 for a safety follow-up[redacted]

Subjects in 'Treatment' group will receive three treatments in total at Day 1, Week 6, and Week 12. [redacted]

Subjects in 'Control/Delayed treatment' group will receive three treatments upon completion of all effectiveness assessments at Week 24. [redacted]

[redacted] Untreated-control/delayed-treatment subjects will not be offered an additional optional retreatment.

[redacted]

[redacted] Subjects will have a screening period of up to 10 days. All enrolled subjects will participate in the study for a maximum duration of 84 weeks (± 14 days).

The primary effectiveness endpoint, the proportion of responders on MAS Décolleté Wrinkles-At Rest, will be assessed live by a blinded evaluator at Week 24 and compared to baseline. Response is defined as at least 1-point improvement from baseline. Standard safety parameters will be monitored throughout the study. [redacted]

2.2 Clinical Study Objectives

Effectiveness

- Confirm the effectiveness of treatment with diluted Radiesse for correction of moderate to severe décolleté wrinkles by demonstrating superiority versus untreated-control.

Safety

- Demonstrate the safety of repeat treatment with diluted Radiesse for correction of décolleté wrinkles.

3 Determination of Sample Size

The primary effectiveness analysis considers the proportion of subjects with at least 1-point (≥ 1-point) improvement on the MAS Décolleté Wrinkles-At Rest scale at Week 24 compared to baseline as assessed live by a blinded evaluator. Subjects will be randomized into two treatment groups as detailed in Section 2.1.

[redacted]

CCI [redacted]

- Group ratio: 3:1;

CCI [redacted]

CCI [redacted] approximately 152 subjects, 114 subjects in the 'Treatment' group and 38 subjects in the 'Control/Delayed treatment' group, will be randomized for this study. Regarding the safety evaluation, in a sample size of 152 subjects an adverse event with a true incidence of 1.5% will be observed at least once with a probability of 89.9%.

CCI [redacted]

Sample size calculations were performed using nQuery software (Version 8.5, Statistical Solutions Ltd., 2019).

4 Analysis Sets

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

Safety Evaluation Set (SES)

The Safety Evaluation Set (SES) will consist of all subjects treated at least once. If not otherwise specified, subjects in the SES will be analyzed as treated.

Intent-to-treat (ITT)

The Intent-to-treat (ITT) will include all randomized subjects. Subjects in the ITT analyses will be analyzed as randomized. This will be the primary population used for the effectiveness analyses.

Per Protocol Set (PPS)

The Per Protocol Set (PPS) is a subset of subjects in the ITT without major protocol deviations potentially affecting statistical analysis. Final determination of what constitutes major or minor protocol deviations in this sense will be made prior to database close CCI [redacted]

5 Endpoints for Analysis

5.1 Effectiveness Endpoints

All endpoints will be assessed separately for the diluted Radiesse 'Treatment' group and the 'Control/Delayed treatment' group, if not otherwise specified. Table 1 gives an overview of all effectiveness endpoints, CCI [redacted]

5.1.1 Primary Effectiveness Endpoint

- Proportion of responders at Week 24 on Merz Aesthetic Scale (MAS) Décolleté Wrinkles-At Rest, as assessed live by a blinded evaluator, where response is defined as at least 1-point improvement from baseline.

5.1.2 Secondary Effectiveness Endpoints

- Proportion of responders at Week 24 on MAS Décolleté Wrinkles-Dynamic, as assessed live by a blinded evaluator, where response is defined as at least 1-point improvement from baseline.
- Proportion of subjects with any improvement, defined as a rating of + 1, + 2 or + 3 on Subject Global Aesthetic Improvement Scale (GAIS) at Week 24.
- Proportion of subjects with any improvement, defined as a rating of + 1, + 2 or + 3 on Investigator GAIS at Week 24.

CCI [redacted]

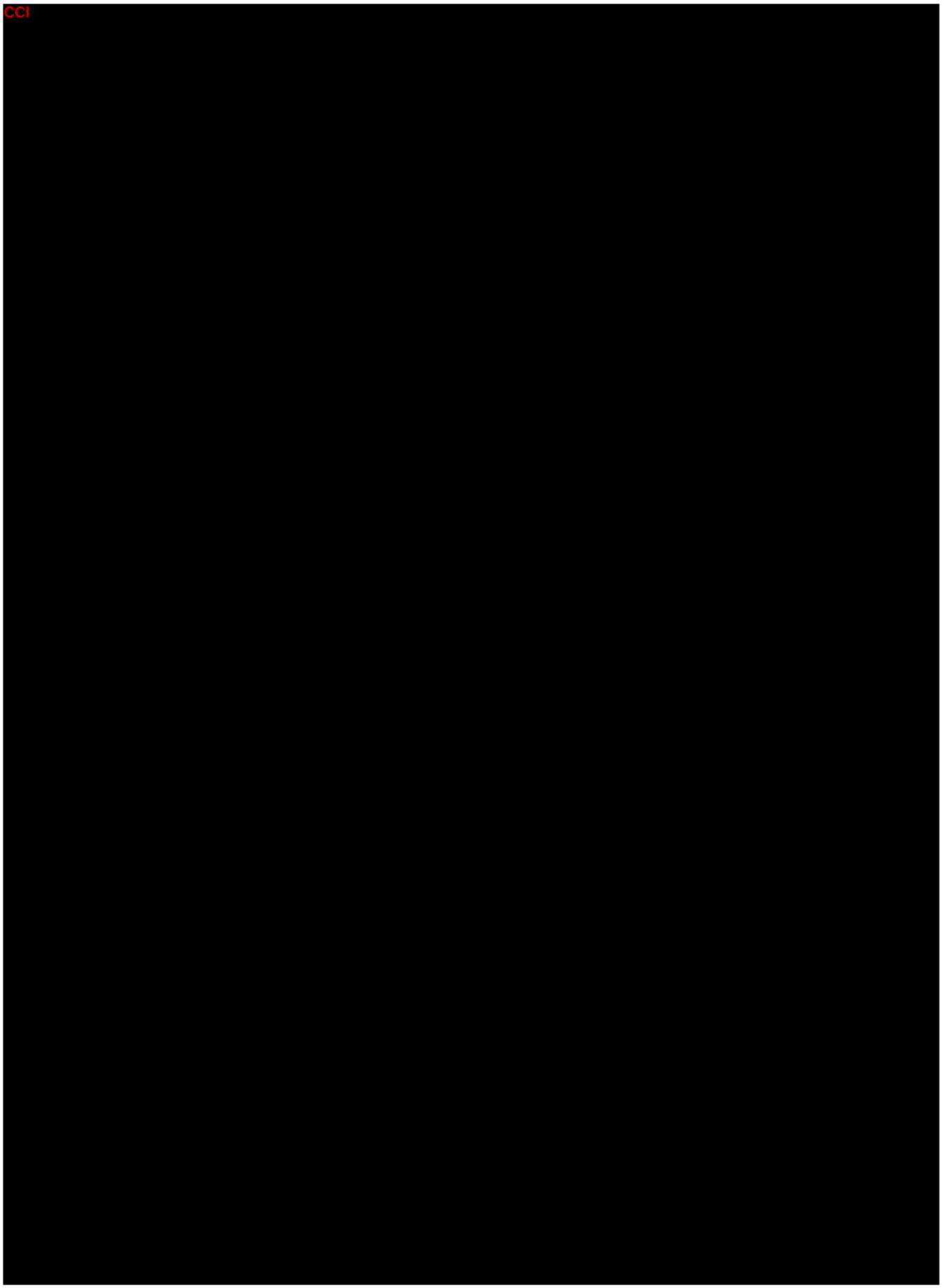

CCI

Table 1: Overview of effectiveness endpoints

Effectiveness endpoint	Description	Analysis time point(s)	Reference time point(s)	Analysis population
Primary	Proportion of responders with at least 1-point improvement from baseline on MAS Décolleté Wrinkles-At Rest	Week 24	Screening	ITT (with missing MAS data replaced by MI procedure), ITT (with observed cases) and PPS
Secondary	Proportion of responders with at least 1-point improvement from baseline on MAS Décolleté Wrinkles-Dynamic	Week 24	Screening	ITT (with missing MAS data replaced by MI procedure), ITT (with observed cases), and PPS
Secondary	Proportion of subjects with a rating of +1, + 2 or + 3 on sGAIS/ iGAIS	Treatment: Week 24	Screening	ITT (with observed cases), and PPS

CCI

[redacted]

CCI

[redacted]

5.2 Safety Endpoints

5.2.1 Primary Safety Endpoint

Not applicable.

5.2.2 Secondary Safety Endpoint

- Incidence of treatment emergent adverse events (TEAEs) related to treatment with diluted Radiesse, as reported by the treating investigator throughout the study.

CCI

[redacted]

CCI [redacted]

5.3 Other Subject Data

Other variables of interest are:

- Subject disposition (including number of subjects enrolled, number of subjects with retreatment, number of discontinuations and reason for discontinuations)
- Incidence and nature of protocol deviations
- Demographic data and baseline characteristics CCI [redacted]

CCI [redacted]

- Previous and concomitant medication
- Previous and concomitant non-drug treatment and procedures
- Medical history and concomitant diseases

CCI [redacted]

6 Statistical Analysis Methods

CCI [redacted]

Adequate descriptive statistics will be provided for each endpoint, by treatment group and overall, where appropriate. Metric descriptive statistics will comprise number of observations, mean, standard deviation, quartiles, minimum, median, and maximum. Frequency tables for qualitative endpoints/variables will display absolute and percent frequencies (n, %) per category where the denominator will be chosen according to the adequate analysis population. Ordered categorical data will be summarized by metric descriptive statistics and frequency tables, where appropriate. All variables will be analyzed as absolute data and as change from baseline assessment, as applicable. Summaries by categorical variables/groups should always include a 'Total' category. 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, an adequate number of decimal places will be chosen. Frequency tables will include the number of missing values. Percentages will be calculated using the denominator of all subjects in a specified population or treatment group. The denominator will be specified in a footnote to the tables for clarification if necessary. Percentages will be reported to one decimal place.

Descriptive confidence limits and descriptive p-values will be given, where appropriate. If not otherwise specified, statistical tests will be conducted two-sided at type I error rate 5% and CIs

will be two-sided with confidence level 95%. All data captured in the eCRF and all image analysis data will be listed. For (sub-) groups with five or less subjects, no statistical test will be performed and confidence intervals will not be displayed due to the low number of subjects.

P-values will be reported to four decimal places. P-values below 0.0001 will be presented as '<0.0001'.

CCI

In the sections below, the term baseline severity refers to the severity of wrinkles on the MAS Décolleté Wrinkles-At Rest as assessed by a blinded live rater.

CCI

6.1 Effectiveness Endpoints

Table 1 gives an overview on the effectiveness endpoints and their corresponding analysis population.

6.1.1 Primary Effectiveness Endpoint

The primary effectiveness endpoint is the proportion of responders (responder rate) at Week 24, defined as subjects with at least 1-point improvement on MAS "Décolleté Wrinkles-At Rest" from baseline, as assessed live by the blinded evaluator. For baseline, the scores assessed at Screening will be used for both treatment groups.

The primary effectiveness endpoint will be analyzed based on the ITT population. The responder rate at Week 24 will be summarized as counts and percentages for the treatment and control groups.

CCI

CCI

[redacted]

6.1.1.1 Primary analysis

The first hypothesis of the primary analysis is to demonstrate that the response rate Π_A at Week 24 is statistically significantly larger than 50% in treated subjects, i.e.,

H_{10}: $\Pi_A \leq 50$ % versus H_{11}: $\Pi_A > 50$ %.

This 50% threshold is commonly used in aesthetic medical device studies to measure a meaningful response.

The second hypothesis is to show statistical superiority of treatment over untreated control. Superiority of treatment means that the response rate in the 'Treatment' group, Π_A, is statistically significantly larger than the response rate in the 'Control/Delayed treatment' group, Π_B, i.e.

H_{20}: $\Pi_A \leq \Pi_B$ versus H_{21}: $\Pi_A > \Pi_B$.

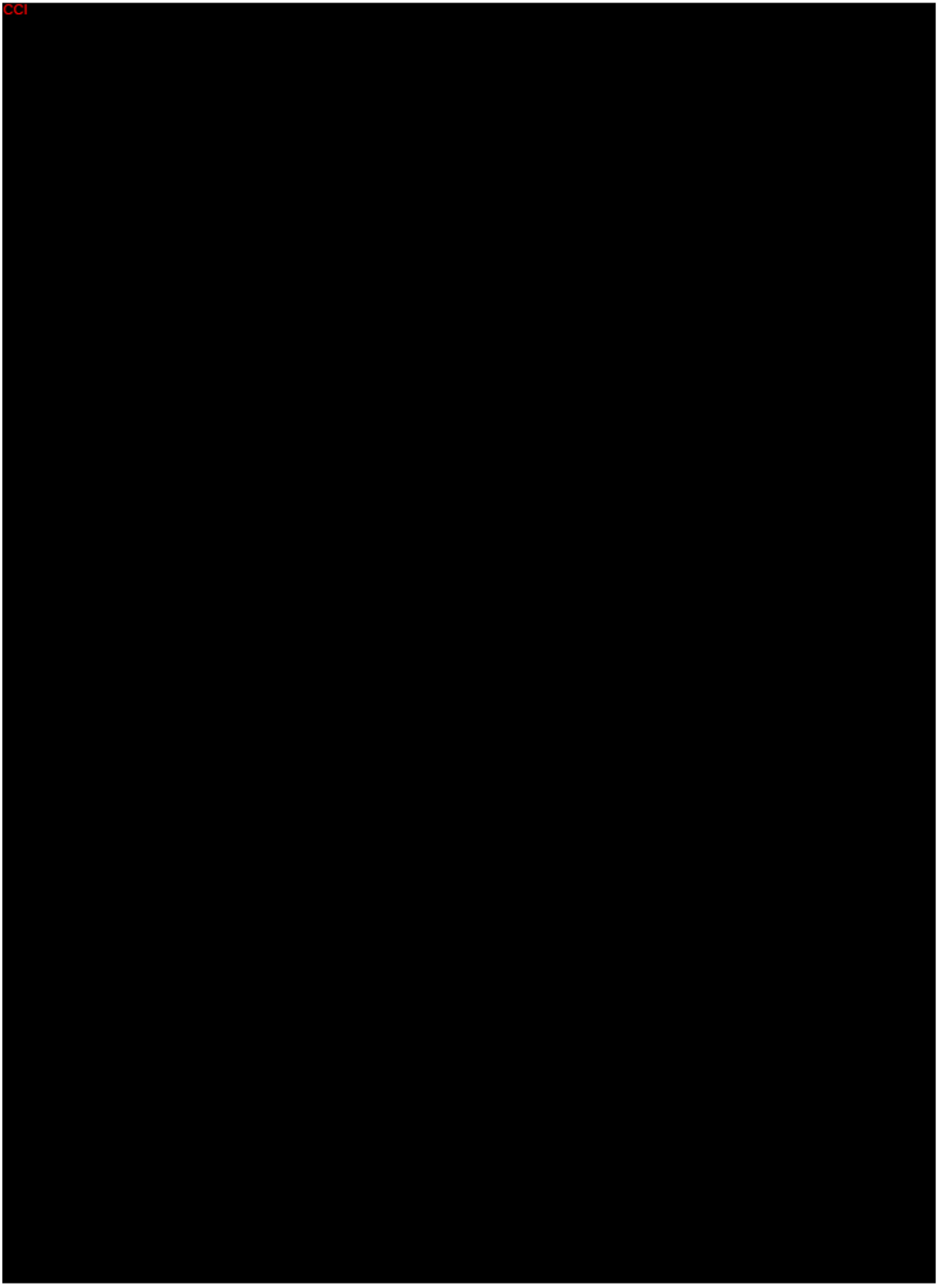
Both hypotheses will be tested using the lower limits of the 95% CIs (two-sided), based on Wilson scores. The lower limit of the 95% Wilson CI must exceed the margin of 50% responder rate to reject the first Null hypothesis H_{10}. The lower limit of the 95% Newcombe CI for the difference between the responder rates in the 'Treatment' group and the 'Control/Delayed treatment' group ($\Pi_A - \Pi_B$) must be greater than zero to reject the second Null hypothesis (H_{20}).

CCI

[redacted]

CCI

[redacted]


CCI

CCI

CCI [redacted]

6.1.2 Secondary Effectiveness Endpoints

MAS Décolleté Wrinkles-Dynamic response

Treatment comparisons for the secondary effectiveness endpoint proportion of responders at Week 24 on MAS Décolleté Wrinkles Dynamic as assessed live by a blinded evaluator will be conducted analogous to the primary effectiveness endpoint. Descriptive two-sided 95% Wilson score-based CIs, CCI [redacted] will be performed CCI [redacted] descriptive summary statistics. CCI [redacted]

CCI [redacted]

The Global Aesthetic Improvement Scale

The iGAIS and sGAIS assessments describing the overall impression of post-treatment aesthetic change specific to the décolleté compared to baseline photograph as well as percentage of subjects with improvement in iGAIS and sGAIS will be evaluated at Week 24 by baseline severity and in total for the 'Treatment' group.

For each GAIS category (i.e., -3 = very much worse, -2 = much worse, -1 = worse, 0 = no change, +1 = improved, +2 = much improved, +3 = very much improved), counts (n) and percentages (%) will be presented.

CCI [redacted]

CCI

[redacted]

The Global Aesthetic Improvement Scale

The iGAIS and sGAIS assessments describing the overall impression of post-treatment aesthetic change specific to the décolleté compared to baseline photograph as well as percentage of subjects with improvement will be evaluated for each post-treatment follow-up visit (except Week 24) and by baseline severity and in total. For subjects receiving an optional retreatment at Week 36 in 'Treatment' group, an additional summary will be used to report iGAIS and sGAIS scores relative to the Week 36 reference photograph.

For each GAIS category (i.e., -3 = very much worse, -2 = much worse, -1 = worse, 0 = no change, +1 = improved, +2 = much improved, +3 = very much improved), counts (n) and percentages (%) will be presented by treatment group.

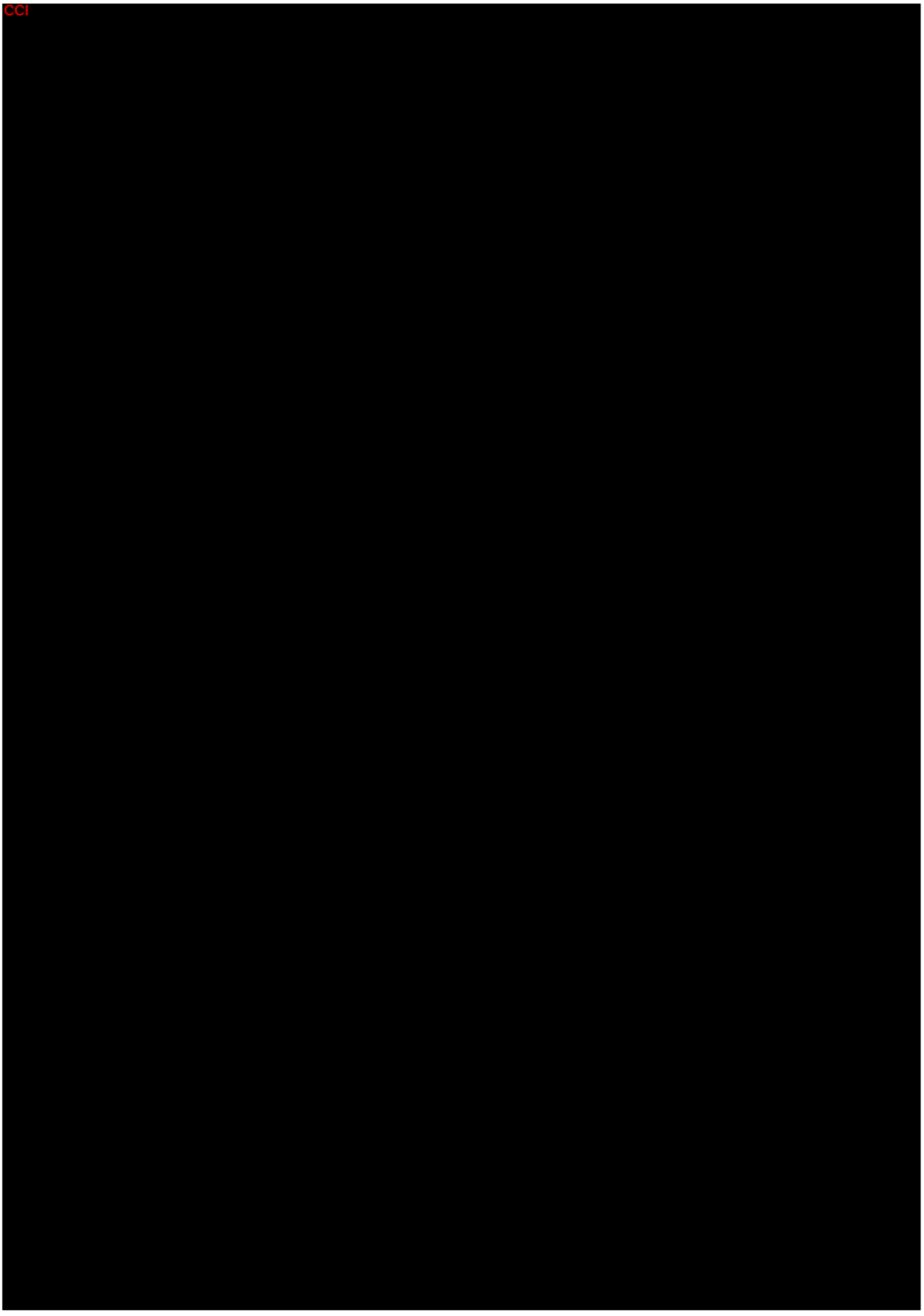
Additionally, iGAIS and sGAIS will be categorized in classes of "improvement" (scores +1 to +3) and "no improvement" (scores -3 to 0) and frequencies for these categories will be provided. Descriptive two-sided 95% Wilson CIs will be calculated for subjects with "improvement".

Moreover, metric descriptive summary statistics will be provided for the iGAIS and sGAIS by treatment group.

CCI

[redacted]

CCI



CCI [redacted]

6.2 Safety Analyses

All safety analyses will be performed on the SES, if not specified otherwise.

For interim and final analysis adverse events (AEs) will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) version in effect at the time the respective database is closed.

Treatment-emergent AEs (TEAEs) 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.

Only TEAEs will be summarized CCI [redacted] using MedDRA preferred terms (PTs) within the system organ classes (SOCs).

CCI [redacted]

CCI

[redacted]

6.2.1 Primary Safety Endpoint

Not applicable.

6.2.2 Secondary Safety Endpoints

Incidences of TEAEs related to the treatment with diluted Radiesse will be summarized by treatment group and total, overall and by treatment period by PT and SOC.

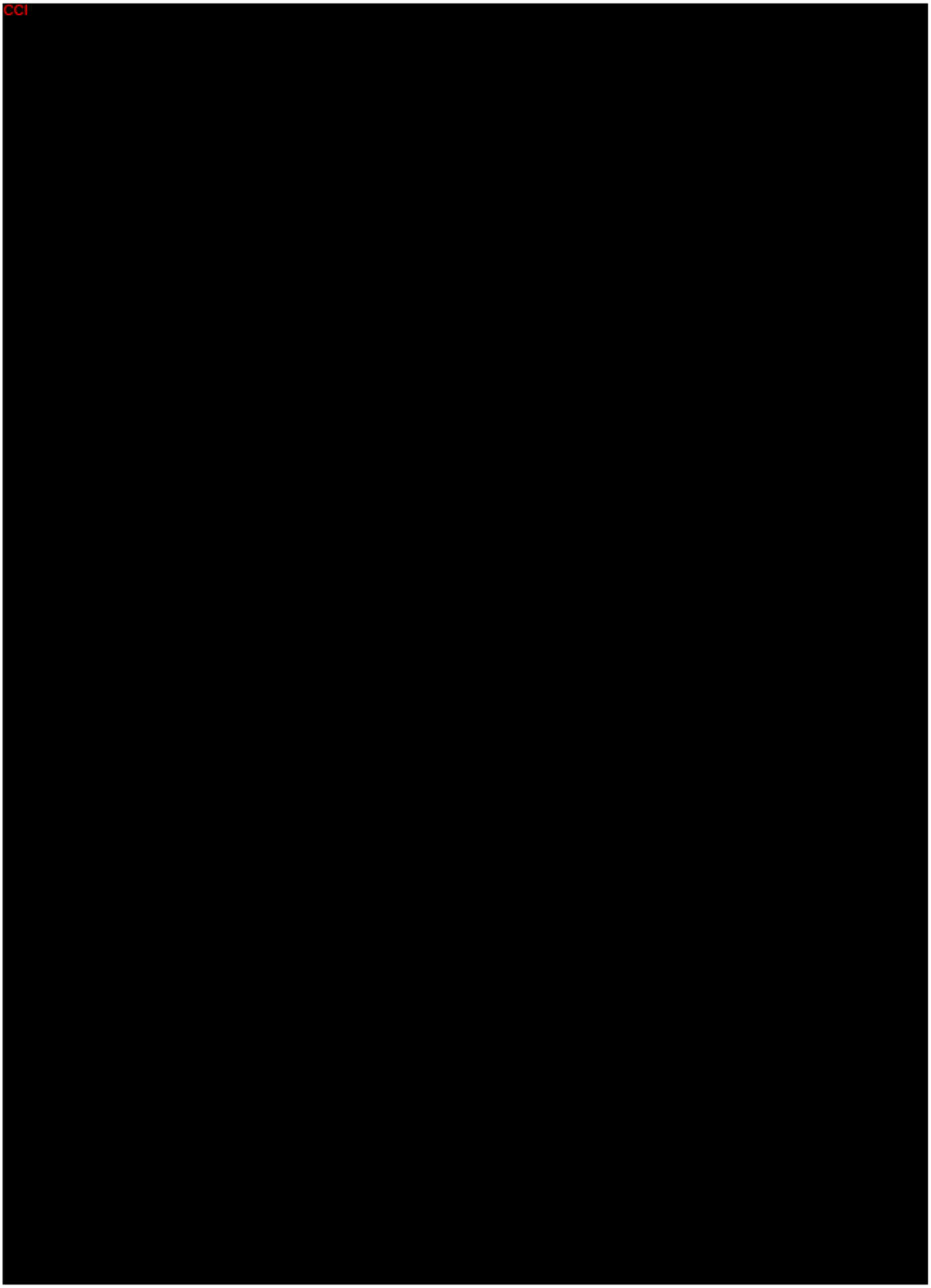
CCI

[redacted]

CCI

[redacted]

CCI



CCI [redacted]

6.3 Other Subject Data

Study subjects will be characterized by descriptive statistics for:

- Disposition of subjects;
- Protocol deviations;
- Demographics and other baseline characteristics;
- Medical history and concomitant diseases;
- Prior and concomitant therapies (medication, non-drug treatments, and procedures); and
- Extent of exposure.

Disposition of subjects, demographic data, and baseline characteristics will be presented using standard descriptive statistics. Demographic data will be summarized for SES, ITT, and PPS; disposition of subjects will be summarized based on all randomized subjects.

Disposition of subjects

The number of subjects screened will be provided. Other subject disposition (i.e., number of subjects randomized, CCI [redacted] the number of subjects in each analysis population) will be summarized by treatment group and overall. CCI [redacted]

The absolute and relative frequencies for subject's main reason of premature discontinuation will be tabulated CCI [redacted]

Visit attendance CCI [redacted] will be presented in a frequency table overall and for both treatment groups.

CCI [redacted]

Protocol deviation

The absolute and relative frequencies of major protocol deviations will be tabulated overall and by treatment group based on all randomized subjects.

All protocol deviations including minor protocol deviations will be listed.

Demographics and other baseline characteristics

Demographic data as well as weight, height, BMI will be summarized overall and by treatment group for the SES, the ITT, and the PPS.

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 start of study 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 (even if it refers to a cut-off point before start of treatment) is ticked.

Medical history and concomitant diseases will be described based on MedDRA SOC and PT levels for the SES.

Previous and concomitant medication

Separation of previous from concomitant therapies will be done according to the start and stop date of the therapy in comparison to the date of first injection. Each therapy will be allocated unambiguously either to previous therapies or to concomitant therapies.

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

CCI [redacted]

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

Frequencies of previous and concomitant medications will be given based on different Anatomical Therapeutic Chemical Classification system code levels of the World Health Organization (ATC) for the SES. Indications are taken into consideration to assign the appropriate ATC code when a drug is coded with WHO-DD, but will not be coded itself with MedDRA and only listed.

CCI [redacted]

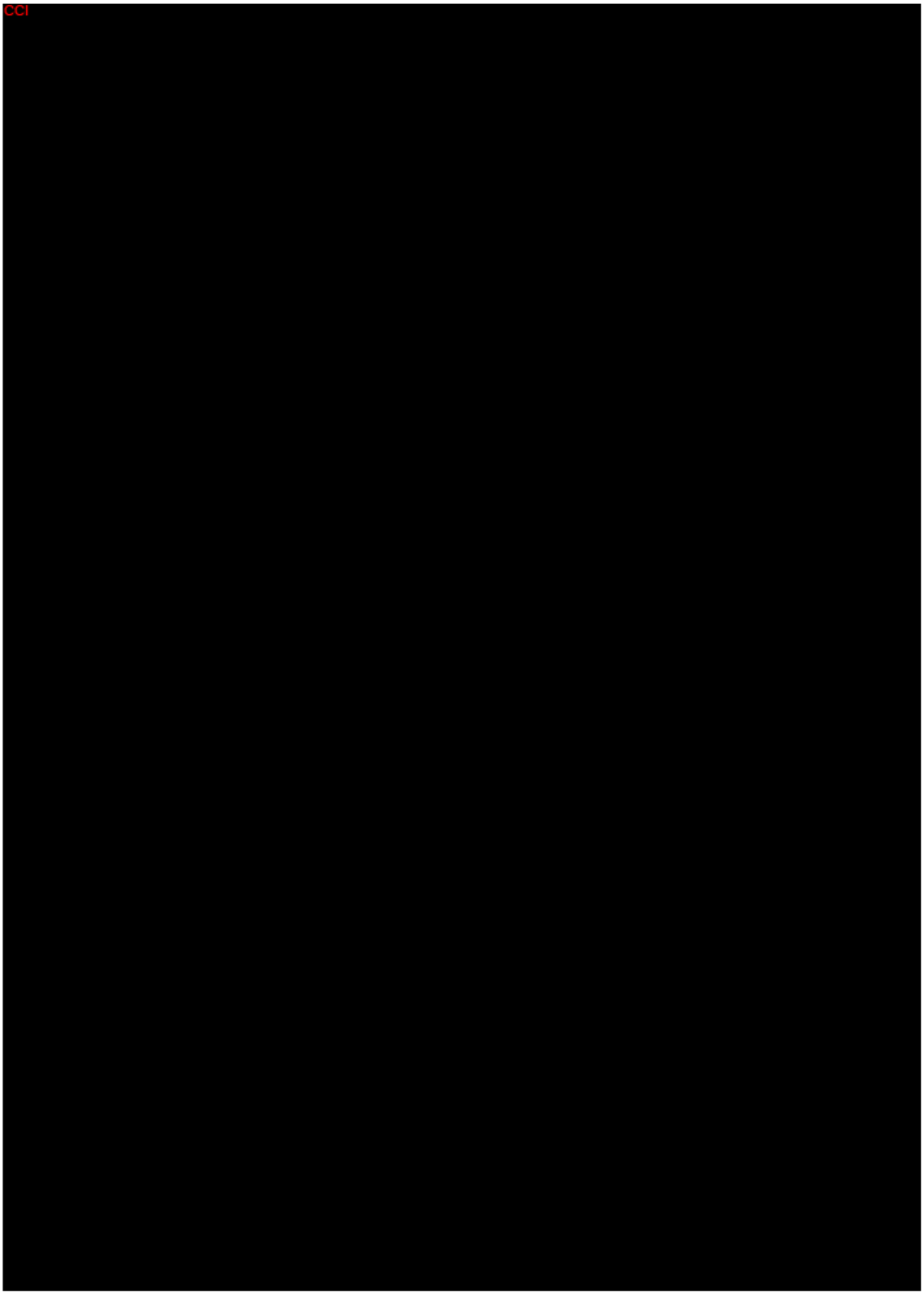
Extent of exposure

Frequencies of subjects receiving treatment CCI [redacted] will be summarized by treatment group for the SES.

CCI [redacted]

6.4 Special Statistical/Analytical Issues

CCI [redacted]


CCI

6.4.2 Interim Analyses

. Safety analyses will be performed on the SES. Effectiveness data will be analyzed on the ITT and PPS where applicable. The interim analysis is not proposed to modify the study design, statistical analyses, or sample size, or for study termination, as this is not an adaptive study design.

After the end of the safety follow-up the database until Week 84/ will be closed and the data until Week will be analyzed. All tables and listings planned for the final analysis will also be done for the interim analysis

6.4.3 Multiple Comparisons/Multiplicity

Based on the primary effectiveness endpoint, two hypothesis tests will be performed. To control for multiplicity, a hierarchical-testing procedure will be performed (see Section 6.1.1).

CCI

[redacted]

8 References

[1] Ian R. White, Rhian Daniel, Patrick Royston (2010): Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables. Comput Stat Data Anal. 2010 Oct 1;54(10):2267-2275. doi: 10.1016/j.csda.2010.04.005. PMID: 24748700; PMCID: PMC3990447.

CCI

[redacted]