Statistical Analysis Plan

Title:	Evaluation of Effectiveness and Safety of Belotero Balance® (+) Lidocaine for Volume Augmentation of the Infraorbital Hollow
Merz Study Number:	M930121002 / NCT04594239
SAP for	Final Analysis
Sponsor:	Merz North America, Inc.
Version:	1.0
Date:	22-Aug-2022
Author: (Name, Function, and CRO)	PPD, Biostatistician, PPD PPD

Document References (For Internal Use Only)

CCI [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 Author (print name), PPD PPD	Date (dd-mmm-yyyy)	Signature
PPD Biostatistician, Freelance (print name)	Date (dd-mmm-yyyy)	Signature
PPD Medical Expert Merz (print name)	Date (dd-mmm-yyyy)	Signature

Table of Contents

List of Abbreviations

AE	Adverse event
BBL	Belotero Balance® (+) Lidocaine
BMI	Body mass index
CAN	Pooled groups TC and DTC of subjects treated with cannula
CDTC	Untreated control until Week 8/ Delayed-treatment BBL with cannula (starting at Week 8)
CDTN	Untreated control until Week 8/ Delayed-treatment BBL with needle (starting at Week 8)
CI	Confidence interval
COVID-19	Coronavirus disease 2019
CRO	Contract Research Organization
CSP	Clinical Study Protocol
CTL	Untreated control subjects until Week 8
CCI	[redacted]
DTC	Randomized control subjects with delayed treatment BBL using a cannula (untreated control until Week 8)
DTN	Randomized control subjects with delayed treatment BBL using a needle (untreated control until Week 8)
DRM	Data Review Meeting
DTRT	Delayed-treatment groups DTN and DTC pooled
eCRF	Electronic case report form
eDiary	Electronic Diary
FACE-Q™	Set of subject-reported questionnaire modules
CCI	[redacted]
IOH	Infraorbital Hollow
IPR	Independent panel review
ITT	Intent to treat population
MedDRA	Medical Dictionary for Regulatory Activities
MIHAS	Merz Infraorbital Hollow Assessment Scale
n	Number of values analyzed
NED	Pooled groups TN and DTN of subjects treated with needle
OC	Observed cases
PP	Per protocol population
PT	Preferred term
SAS	Statistical Analysis System®
SOC	System organ class
SP	Safety population
TC	Randomized treatment group, subjects treated with BBL using a cannula
TEAE	Treatment Emergent Adverse Event
TESAE	Treatment-emergent serious adverse event
TN	Randomized treatment group, subjects treated with BBL using a needle
TOTT	Pooled groups NED and CAN (or TRT and DTRT)
TRT	Pooled treatment groups TN and TC
US	United States of America
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 investigational plan M930121002, dated 12-Jun-2020 and the following amendments, dated 14-Jul-2020, 20-Jul-2020, 27-Jan-2021 and 04-Nov-2021.

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 created directly 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.

The Merz standard 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 TFLs, SDTMs, ADaMs) 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, comparative, evaluator-blinded pivotal study designed to evaluate the safety and effectiveness of Belotero Balance® (+) Lidocaine (BBL) for the correction of volume loss in the Infraorbital Hollow (IOH) area. Study subjects will have a screening period of up to 10 days and participate for a maximum duration of 76 weeks (± 14 days).

A total of 150 subjects will be randomized at 9 investigational sites in the United States (US). At each study site, the number of subjects randomized should not exceed 26. A minimum of 20% of enrolled subjects will be Fitzpatrick Skin Type IV, V, or VI. Subjects from the Fitzpatrick Skin Type IV, V, or VI group will be distributed as follows: at least 16 subjects of Type IV; 7 subjects of Type V; and 7 subjects of Type VI. At least 10% of the subjects enrolled will be male.

Subjects eligible for study enrollment will have symmetrical right and left IOHs with the same Merz Infraorbital Hollow Assessment Scale (MIHAS) score of 2 or 3 (moderate or severe), as assessed live by a blinded evaluator. All blinded evaluators will be qualified healthcare practitioners, delegated by the treating investigator and trained by the sponsor. At screening, eligible subjects will be randomized to 4 groups using a 2:2:1:1 ratio as follows: BBL with needle (TN), BBL with cannula (TC), control/delayed-treatment BBL with needle (CDTN), and control/delayed-treatment BBL with cannula (CDTC). The needle or cannula assignment for a subject cannot be interchanged during the study; that is, needle and cannula cannot be used in the same subject.

Subjects randomized to treatment (TN or TC) will receive a BBL injection in both IOHs at Day 1. Treated subjects who do not achieve optimal aesthetic correction or at least 1-point improvement on the MIHAS, as assessed by the blinded evaluator 4 weeks post treatment compared to baseline, may have a touch-up injection in one or both IOH(s). If a touch-up is necessary, it will only be administered if there are no medical contraindications or existing adverse events (AEs) of concern,

as determined by the treating investigator. The treating investigator will be responsible for reviewing whether the threshold for optimal aesthetic correction or at least 1-point MIHAS improvement is met, as assessed by the blinded evaluator, since the blinded evaluator will not have access to the subject's study records.

Subjects who achieve at least 1-point improvement on the MIHAS 4 weeks post treatment when compared to baseline may have a touch-up injection in one or both IOH(s) to achieve optimal correction, at the discretion of the treating investigator and the subject.

Subjects randomized to treatment (TN or TC) at Day 1, will have the option for retreatment, upon agreement between the subject and the treating investigator, at 48 weeks post last injection (i.e., baseline injection or touch-up, if applicable) and will then be followed for an additional 24 weeks, for a total study duration of up to 76 weeks (72 weeks if no touch-up was performed).

Subjects randomized to the control/delayed-treatment group (CDTN or CDTC) will remain untreated until Week 8. After all applicable effectiveness endpoint assessments have been completed at their Week 8 visit, control subjects will receive BBL IOH injections (i.e., delayed treatment) and will be followed for 48 weeks. Subjects who do not achieve optimal aesthetic correction or at least 1-point improvement on the MIHAS as assessed by the blinded evaluator 4 weeks post treatment compared to the delayed-treatment baseline (Week 8), may have a touch-up injection in one or both IOH(s). If a touch-up is necessary, it will only be administered if there are no medical contraindications or existing AEs of concern, as determined by the treating investigator. The treating investigator will be responsible for reviewing whether the threshold for optimal aesthetic correction or at least 1-point MIHAS improvement is met, as assessed by the blinded evaluator, since the blinded evaluator will not have access to the subject's study records.

Subjects from the control/delayed treatment group who achieve at least 1-point improvement on the MIHAS 4 weeks post treatment when compared to the delayed-treatment baseline (Week 8) may have a touch-up injection in one or both IOH(s) to achieve optimal correction, at the discretion of the treating investigator and the subject.

Control/delayed-treatment subjects will not be offered retreatment.

Any time a subject is treated (i.e., initial injection, touch-up, and/or retreatment), he/she will have a 72-hour post-treatment phone call to evaluate safety. If a subject reports a safety concern during any phone call, an unscheduled visit may be necessary.

Effectiveness assessments include blinded evaluator MIHAS, treating investigator and subject Global Aesthetic Improvement Scale (GAIS), subject-reported FACE-Q Satisfaction with Eyes, FACE-Q Appraisal of Lower Eyelids, FACE-Q Patient-Perceived Age visual analogue scale (VAS), a Patient-Perceived Pain VAS, and a likelihood of future treatment assessment. These assessments will be performed throughout the study. Additionally, three blinded, board-certified independent panel review (IPR) experts who are not part of the study sites will assess subject IOHs using the MIHAS and subject photographs.

Standard safety parameters will be monitored throughout the study. Additionally, visual assessments (i.e., visual acuity, confrontation visual field test, ocular motility, and an undilated central retinal exam using non-mydriatic retinal cameras and read by a central reader) will be conducted. A subject electronic diary (eDiary) will be used to record predefined common treatment responses (CTRs) for 28 days after each injection.

2.2 Clinical Study Objectives

Effectiveness

Confirm the effectiveness of BBL injection for the correction of volume loss in the IOH area by demonstrating superiority to untreated control.

Safety

Confirm the safety of BBL injection for the correction of volume loss in the IOH area.

3 Determination of Sample Size

The primary effectiveness analysis considers the proportion of subjects with at least 1-point improvement on the MIHAS for both IOHs, as assessed by the blinded evaluator at Week 8 compared to baseline. For this analysis, subjects from TN and TC groups will be pooled into the Treatment Group TRT, and subjects from CDTN and CDTC groups (data before any treatment) will be pooled into the Control Group (CTL).

A two-step hierarchical-testing procedure is foreseen for the primary analysis:

1. Treatment Group (TRT) shows a response rate of more than 50%, and
2. Treatment Group (TRT) is superior to the Control Group (CTL).

This hierarchical-testing procedure will keep the overall one-sided error rate of $\alpha = 0.025$, and no adjustment for multiplicity is necessary.

Sample-size estimation was based on methods implemented for the binomial and Fisher's exact test. In this context, it should be noted that the use of two-sided 95% Wilson confidence intervals (CI) and Newcombe CIs, instead of binomial and Fisher's exact tests as referenced in the next paragraph, will deliver more efficient results (i.e. smaller CIs). The CI-based methods have the advantage that they can be used for multiple imputation of missing data. Hence, sample size calculations were based on slightly more conservative test procedures, so that the determined power will also suffice for tests based on CIs.

Sample-size estimation for test Step 1 was performed with a one-sided binomial test at an error level of 2.5% to evaluate whether the response rate was significantly larger than 50%. It was aimed to ensure a power of 90%. The binomial test will be conducted on the pooled data (TRT) from both treatment groups (TN and TC). Due to the allocation ratio of 2:2:1:1, the two treatment groups (TN and TC) will be approximately equally represented. Calculations for test Step 2 were performed with a one-sided Fisher's exact test for unequal n's, at an error level of 2.5%, to detect a statistically significant difference between a response rate in Group TRT of 70% compared to 30% in Group CTL. A power of at least 90% was assumed.

Based on data from a pilot study in the same indication (M930121001) showing responder rates of above 80% for observed data and above 70% with missing data treated as no change from baseline (no response), a true responder rate of 70% for this pooled Group TRT sample was assumed. Under these assumptions, a total of 80 subjects are needed to reach a power of at least 94% in test Step 1.

For test Step 2 and maintaining the assumption to have 70% response rate in Group TRT and 30% in Group CTL, a sample size of 80 subjects in Group TRT and 40 subjects in Group CTL reaches a power of at least 98% to demonstrate statistically different response rates between the groups.

To account for up to 20% drop-outs, 150 subjects will be randomized in this study (i.e., 50 subjects into each treatment group [TN and TC] and 25 subjects to each control group [CDTN and CDTC]). The conservative approach for this hierarchical-test procedure has an overall power of above 92%.

If 150 subjects are treated, the sample size is sufficient to observe, with a probability of 80%, at least one AE (adverse event) with actual event probability of 1.1%.

Sample size calculations were performed using nQuery software (Version 8, Statistical Solutions Ltd., 2017).

4 Analysis Sets

In addition to the randomized treatment groups (TN, TC, CDTN and CDTC; see Section 2.1), the following treatment groups will be defined for randomized subjects as follows:

- Treatment group (TRT): All subjects from TN and TC groups;
- Control group (CDTRT): All subjects from CDTN and CDTC;
- CTL: All subjects from CDTN and CDTC groups (only data until first treatment will be considered for analysis.);
- DTN: Subset of subjects within CDTN (only data from first (delayed) treatment onward will be considered for analysis);
- DTC: Subset of subjects within CDTC (only data from first (delayed) treatment onward will be considered for analysis);
- DTRT: All subjects from DTN and DTC (only data from first (delayed) treatment onward will be considered for analysis);
- NED: Subjects that receive treatment with needle-injection technique will be pooled from TN and DTN groups;
- CAN: Subjects that receive treatment with cannula-injection technique will be pooled from TC and DTC groups; and
- TOTT: Subjects that receive treatment at any time during the study (TRT and DTRT) will be pooled from NED and CAN groups.

The following analysis sets will be defined for statistical analysis:

Randomized Population

The randomized population will consist all subjects randomized.

Safety Population (SP)

The SP will consist of all subjects treated.

Intent-to-Treat (ITT) Population

The ITT population will include all randomized subjects of the control group (CTL) and all randomized and treated subjects of the treatment group (TRT).

Per Protocol (PP) Population

The PP population is a subset of subjects in the ITT population without major protocol deviations (PDs). Final determination of what constitutes a major protocol deviation leading to exclusion from PP population will be made in a data review meeting (DRM) prior to database close.

Analyses based on ITT and PP populations and effectiveness analyses based on SP will include subjects as randomized. All other analyses on the SP will include subjects as treated.

5 Endpoints for Analysis

5.1 Effectiveness Endpoints

5.1.1 Primary Effectiveness Endpoint

Proportion of responders at Week 8 on the MIHAS, as assessed live by a blinded evaluator.

Note: Responder is defined as a subject with at least 1-point improvement from baseline on MIHAS of both IOHs.

Note: For subjects randomized to treatment, if no touch-up is performed, the primary effectiveness MIHAS assessment will occur 8 weeks post baseline injection. If a touch-up is performed, the primary effectiveness MIHAS assessment will be completed 8 weeks post touch-up. For subjects randomized to control/delayed-treatment, the primary effectiveness assessment will occur 8 weeks from the screening visit. I.e. MIHAS data from Visit 3 will be used for subjects randomized to treatment and MIHAS data from Visit C1-a will be used for subjects randomized to control/delayed-treatment.

5.1.2 Secondary Effectiveness Endpoints

Note: For all secondary endpoints, Week 8 corresponds to the time from last injection (either baseline or touch-up injection) for subjects randomized to treatment only (TN and TC groups). I.e. data from Visit 3 will be used.

- GAIS score at Week 8, as assessed by the treating investigator.
- GAIS score at Week 8, as assessed by the subject.
- FACE-Q™ Satisfaction with Eyes scores and changes from baseline to Week 8, as assessed by the subject.

  Note: FACE-Q evaluates both eyes concurrently.

CCI

[redacted]

CCI

[redacted]

5.2 Safety Endpoints

5.2.1 Primary Safety Endpoints

Not applicable.

5.2.2 Secondary Safety Endpoints

- Incidence of treatment-emergent adverse events (TEAEs) related to BBL.
- Incidence of treatment-emergent serious adverse events (TESAEs) related to BBL.

CCI

[redacted]

CCI

[redacted]

6 Statistical Analysis Methods

Metric statistics will be number of values analyzed (n), mean, standard deviation, median, quartiles, minimum, and maximum. If available, number of missing (Nmiss) or number of imputed values will be presented. 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.

Frequency tables for qualitative endpoints will be absolute and percent frequencies (n, %). In case of missing values, 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. For visit-based data and if not otherwise specified, 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.

If not otherwise specified, statistical tests will be conducted one-sided at type I error rate 2.5%. P-values will be reported to four decimal places (e.g., p=0.0375). P-values below 0.0001 will be presented as '<0.0001'. CIs will be two-sided 95% CIs.

The following definition of treatment cycles will be used in several analyses.

For subjects of treatment groups TN and TC who receive retreatment, Cycle 1 is the period from initial injection to the retreatment injection (at 48 weeks after the last injection [either initial or touch-up injection]) and Cycle 2 is the period from retreatment injection until study end (i.e. date of last contact). Analyses for Cycle 2 will therefore present results at visit V8 as 8 weeks post-injection in Cycle 2, rather than as 56 weeks post last injection in Cycle 1. Similarly, results at visit V9 will be presented as 24 weeks post-injection in Cycle 2.

For subjects of treatment groups TN and TC who do not receive retreatment and for subjects of treatment groups DTN and DTC, Cycle 1 is defined as the period from initial injection until study end. Results at visit V8 and V9 for TN and TC subjects will be presented as 56 and 72 weeks post last injection (in Cycle 1), respectively.

Subjects who never received treatment do not have a cycle 1 or cycle 2.

Section 6.4.1 describes methods for handling of data collected at the end of study visit (either visit V9 for subjects in TRT or visit C7 for subjects in CDTRT) among subjects prematurely discontinuing the study.

For the primary and secondary effectiveness analyses, if not stated otherwise, "Week 8" refers to visit V3 (8 weeks post last injection in Cycle 1) for subjects in TRT, TN, and TC and visit C1-a (8 weeks post screening) for subjects in CTL.

All data will be listed. If not stated otherwise, listings will be presented based on all subjects randomized.

6.1 Effectiveness Endpoints

6.1.1 Primary Effectiveness Endpoint

The primary effectiveness endpoint is the proportion of responders at Week 8, as assessed live by a blinded evaluator. A responder is defined as a subject with at least 1-point improvement from baseline on MIHAS of both IOHs. For baseline, the MIHAS scores as assessed at Screening will be used. Further details on assessments and visits used for analysis are described in detail in Section 5.1.1.

The primary effectiveness endpoint will be analyzed based on the ITT population. The proportion of responders at Week 8 will be summarized as counts and percentages for the pooled TRT and the pooled CTL groups. For analyses based on multiple imputation as described below, the average number and percentage of responders, over all imputations, will be presented instead. Four hypothesis tests will be performed and will be evaluated in sequential order. The formal hypotheses of each test are given in CCI To control for multiplicity, a hierarchical-testing procedure will be performed. Each test will hereby be performed at a one-sided α level of 0.025.

The first two tests for the primary endpoint of treatment versus control will be confirmatory tests as outlined in CCI and as described in detail in Section 6.1.1.1; if both null hypotheses are rejected, two additional tests will be performed, as confirmative subgroup analyses, to separately investigate the effectiveness of the two injection techniques (i.e., needle and cannula). For the subgroup analyses, the first test is TN vs. CTL and only if this hypothesis can be rejected, the test of TC vs. CTL will be performed. Details on this subgroup analysis are given in Section 6.1.1.3.

CCI

6.1.1.1 Primary analysis

The first hypothesis test of the primary analysis evaluates whether the response rate at Week 8 (Visit 3) in treated subjects is significantly larger than 50%. The analysis will be based on all subjects of the ITT population in the TRT treatment group.

Hypothesis test 1:

H_{10} : $\pi_{TRT} \leq 50\%$

H_{11} : $\pi_{TRT} > 50\%$,

where π_{TRT} denotes the responder rate in percent at Week 8 of treated subjects.

The second hypothesis test evaluates whether the response rate in treated subjects at Week 8 (Visit 3) is higher than the response rate in untreated subjects at Week 8 (Visit C1-a): The analysis will be based on all subjects from the ITT population in the TRT and CTL groups.

Hypothesis test 2:

H_{20} : $\pi_{TRT} \leq \pi_{CTL}$

H_{21} : $\pi_{TRT} > \pi_{CTL}$,

where π_{TRT} and π_{CTL} denote the responder rates in percent at Week 8 of the treated and untreated subjects, respectively.

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 for the responder rate in the TRT group must exceed the margin of 50% responder rate to reject Hypothesis 1 (H_{10}). The lower limit of the 95% Newcombe CI for the difference between the responder rates in the TRT and the CTL groups (TRT-CTL) must be greater than zero to reject Hypothesis 2 (H_{20}).

Using the Wilson and Newcombe CIs in these tests will allow for imputation of missing values of the primary effectiveness endpoint with a multiple-imputation approach. This procedure will replace missing MIHAS values at Week 8 for the ITT population.

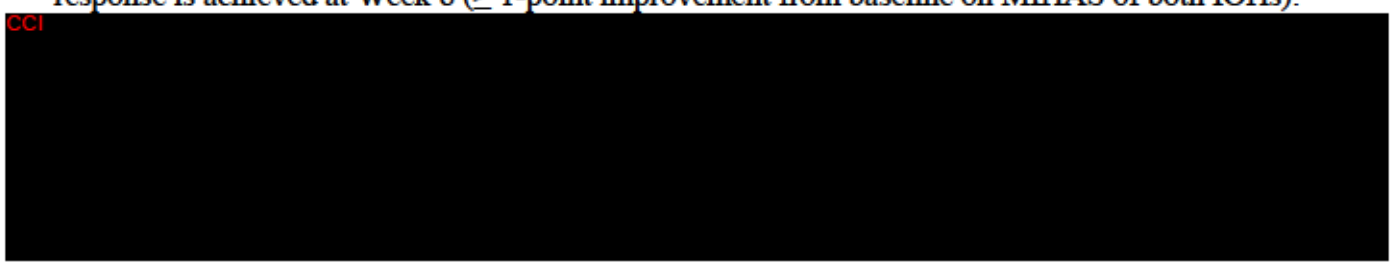
SAS PROC MI will be used to create 100 sets with imputed MIHAS scores per treatment group (TRT and CTL) and per IOH (left and right). For each treatment group and IOH, a different seed will be used for implementation. For treatment group TRT, a seed of 2793 and a seed of 8210 will be used for left and right IOH, respectively. Regarding the control group CTL, a seed of 3913 and a seed of 7956 will be used for left and right IOH, respectively. Baseline MIHAS, Week 4 (Visit 2-a) MIHAS, (pooled) site, and touch-up (yes/no) will be included in the multiple-imputation model for subjects in the TRT group. Details on pooling of sites are described in Section 6.4.5. If Week 4 MIHAS scores are missing, then these will be imputed using the same covariables available up to the respective time point, before imputing Week 8 MIHAS scores. Baseline MIHAS and (pooled) site will be included in the multiple-imputation model for subjects in the CTL group. The default number of burn-in iterations before each imputation (20) will be used for all imputation models. In order to ensure reproducibility, the data will be sorted by subject identifier (USUBJID) prior to imputation. SAS code similar to the following will be used:

CCI

CCI

Subsequently, it will be determined for each subject within each of the imputed datasets, whether response is achieved at Week 8 (≥ 1-point improvement from baseline on MIHAS of both IOHs).

CCI



This code assumes that 'response' is coded with '1' and 'no response' is coded with '0'. Moreover, midata_8w_trt_left, midata_8w_trt_right, midata_8w_ctl_left and midata_8w_ctl_right are the subsets of midata_trt_left, midata_trt_right, midata_ctl_left and midata_ctl_left, respectively, and will include only the MIHAS scores at baseline and at Week 8.
For the Hypothesis Test 1, the Wilson CI for the response rate in the TRT group ([WL_{TRT}, WU_{TRT}]) 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=100* imputed datasets with sample size *n*.

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

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

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

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 two boundaries of the Wilson CI (WL_{TRT} and WU_{TRT}) will be calculated as

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

where t denotes the 0.975 quantile of the t-distribution with v degrees of freedom,

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

and with

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

4. In order to reject Hypothesis 1 (H_{10}), the lower limit of the 95% Wilson CI as given above, which results from the complete multiple imputation procedure must exceed the margin of 50%.

For the test of Hypothesis 2, the Newcombe CI will be calculated using the following procedure (as described in [2]), assuming ignorable missingness:

1. The 95% Wilson CIs will be calculated for groups TRT and CTL as described above, denoted as [WL_{TRT}, WU_{TRT}] and [WL_{CTL}, WU_{CTL}] with upper and lower bounds $WL_{TRT}, WU_{TRT}, WL_{CTL}, WU_{CTL}$.
2. The lower and upper bound of the Newcombe CI will then be computed using the following formulas

$$NL = \bar{p}_{m,TRT} - \bar{p}_{m,CTL} - \sqrt{\left(\bar{p}_{m,TRT} - WL_{TRT}\right)^2 + \left(WU_{CTL} - \bar{p}_{m,CTL}\right)^2},$$

$$NU = \bar{p}_{m,TRT} - \bar{p}_{m,CTL} + \sqrt{\left(WU_{TRT} - \bar{p}_{m,TRT}\right)^2 + \left(\bar{p}_{m,CTL} - WL_{CTL}\right)^2},$$

where $\bar{p}_{m,TRT}$ and $\bar{p}_{m,CTL}$ are the overall means for TRT and CTL groups, respectively (see step 2 in the calculation of the Wilson CI).

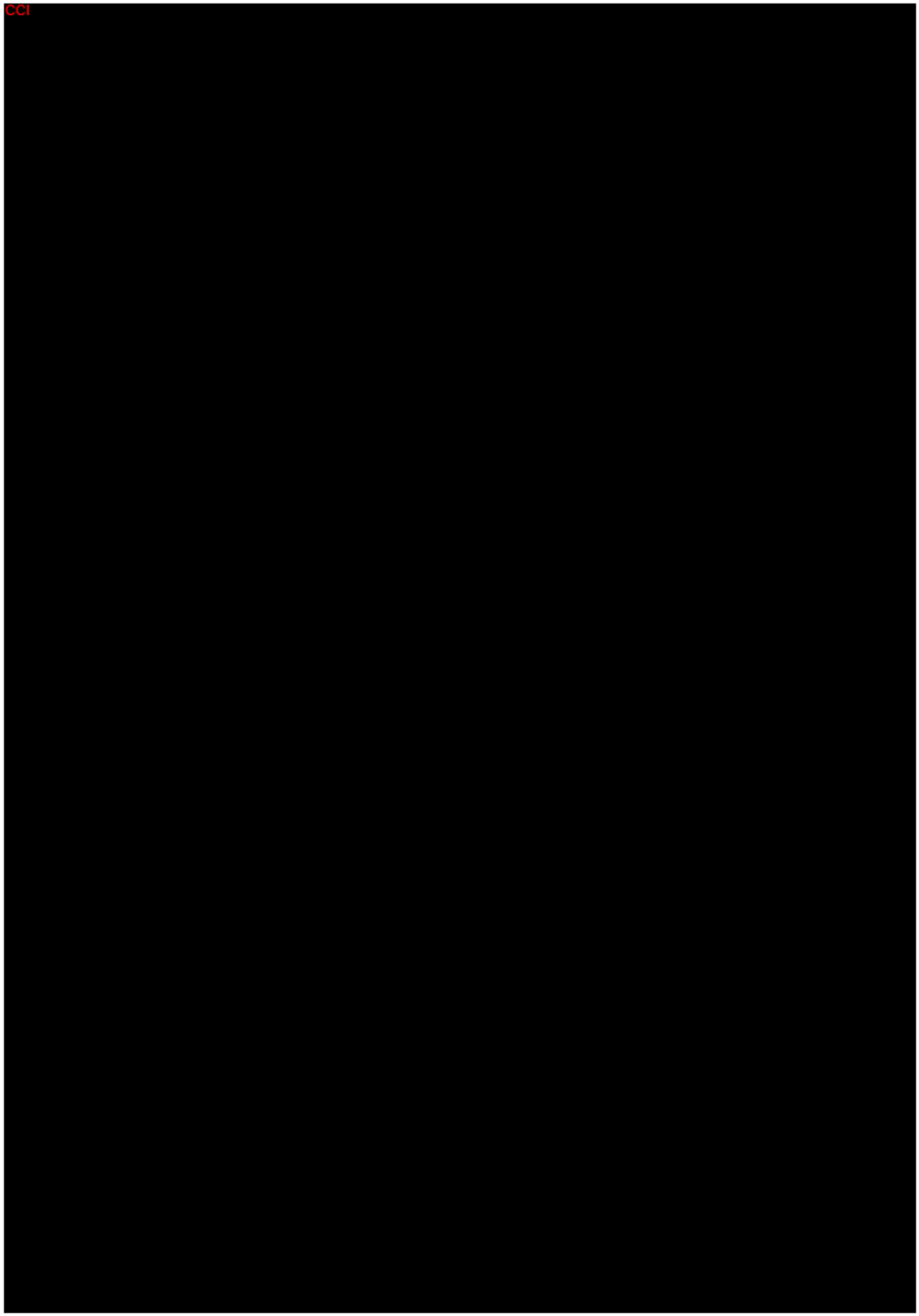
3. The lower limit NL of the calculated two-sided 95% Newcombe CI must be greater than zero to reject Hypothesis 2 (H_{20}).

If both hypotheses H_{10} and H_{20} are rejected, two additional tests will be performed as subgroup analyses (see Section 6.1.1.3).

CCI

[redacted]

CCI



CCI

[redacted]

6.1.1.3 Subgroup analyses

If both Null hypotheses in the primary analysis are rejected, additional hypotheses testing will be performed to evaluate whether treatment with needle-injection technique and treatment with cannula-injection techniques are superior over control as displayed in CCI [redacted] test Steps 3 and 4. Only if H_{30} can be rejected will H_{40} be tested.

A similar procedure as for the test of Hypothesis 2 in Section 6.1.1.1 will be used to separately compare the responder rates of the TN and TC groups with the CTL group. The hypothesis tests will be performed based on the ITT population.

Hypothesis test 3:

$H_{30} : \pi_{TN} \leq \pi_{CTL}$

$H_{31} : \pi_{TN} > \pi_{CTL}$.

Hypothesis test 4:

$H_{40} : \pi_{TC} \leq \pi_{CTL}$

$H_{41} : \pi_{TC} > \pi_{CTL}$.

In the hypotheses above, π_{TN}, π_{TC} and π_{CTL} denote the responder rates in percent of subjects treated with needle or cannula and of the untreated subjects, respectively.

For testing both Hypothesis 3 and 4, missing MIHAS values as replaced by multiple imputation for the primary analysis of the primary endpoint will be used.

Similar to Hypothesis test 2, the Newcombe CI for the differences between responder rates of TN and CTL will be calculated to test Hypothesis 3. The lower limit of the calculated two-sided 95% Newcombe CI for the difference between responder rates of TN and CTL must be greater than zero to reject H_{30}.

If Hypothesis 3 is rejected, Hypothesis 4 will be tested by calculating the two-sided 95% Newcombe CI for the difference between responder rates of TC and CTL. Similar to above, the lower limit of the calculated CI must be greater than zero to reject H_{40}.

These subgroup analyses will be performed based on the ITT population, using multiple imputation and will be repeated based on observed cases for the ITT and PP populations.

Furthermore, response rates for TRT and CTL, as well as the difference of response rates between TRT and CTL, including the corresponding 95% CIs will be provided for the following subgroups:

- Baseline MIHAS (2 vs. 3),
- Site,
- Sex (male vs. female),
- Race (white vs. other),
- Ethnicity (Hispanic or Latino vs. not Hispanic or Latino),
- Age (< 45 years vs. ≥ 45 years), and
- Fitzpatrick Skin Type (I-III vs. IV-VI).

These subgroup analyses will be performed based on observed cases in the ITT population.

6.1.1.4 Supportive analyses

The MIHAS scores of left and right IOH at baseline and Week 8 as assessed by a blinded evaluator will be presented by summary statistics for the TN, TC, TRT and CTL groups. Furthermore, the absolute change from baseline to Week 8 will be summarized descriptively. In this instance, baseline refers to the screening visit for all groups.

Additionally, MIHAS scores (0 = none to minimal, 1 = mild, 2 = moderate, 3 = severe, 4 = extreme) of left and right IOH at baseline and Week 8 are described in frequency tables via counts (n) and percentages (%) for the same groups as for the summary statistics.

Supportive analyses will be based on observed cases in the ITT and the PP population.

6.1.2 Secondary Effectiveness Endpoints

If not otherwise specified, the analyses of secondary effectiveness endpoints will be conducted based on observed cases for the ITT and PP population for the TN, TC and TRT groups only. Subgroup analyses will be performed only based on the ITT population, using observed cases.

GAIS

CCI [redacted]

FACE-Q

CCI [redacted]

CCI

Subgroup analysis

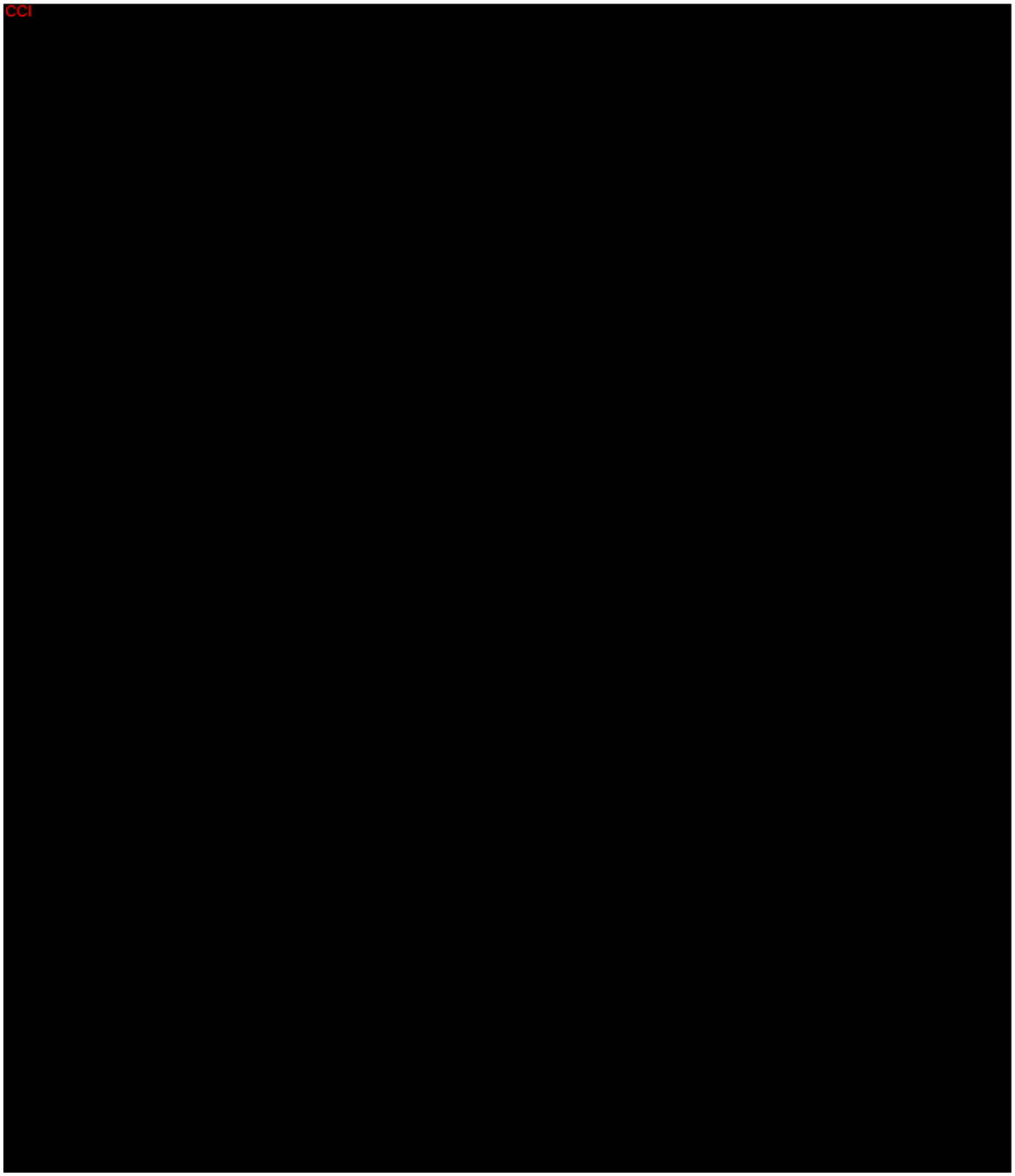
Investigator and subject GAIS, as well as FACE-Q Satisfaction with Eyes Scores, will be analyzed for subgroups TN and TC, as part of the analyses described above.

In addition to the overall summary statistics defined above, summary statistics for investigator and subject GAIS and FACE-Q Satisfaction with Eyes Rasch-Transformed Scores will also be

provided for TN, TC and TRT groups, based on observed cases in the ITT population, for the following subgroups:

- Baseline MIHAS (2 vs. 3),
- Site,
- Sex (male vs. female),
- Fitzpatrick Skin Type (I-III vs. IV-VI).

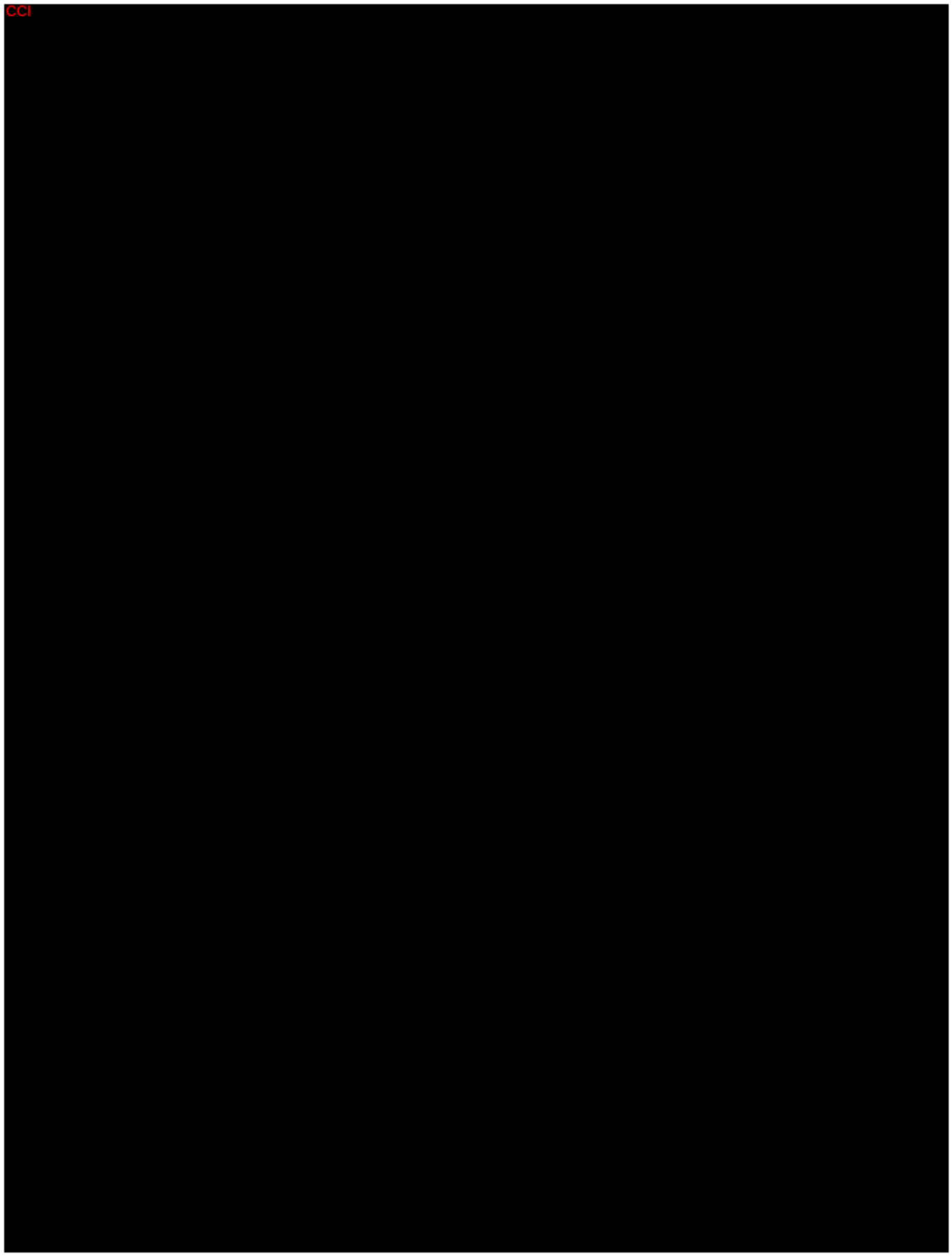
CCI



CCI

[redacted]

CCI



6.2 Safety Endpoints

All safety endpoints will be summarized for observed values in the SP by treatment group (TRT and DTRT) and overall (TOTT) and by injection technique (needle/cannula) (groups TN, DTN and NED as well as TC, DTC and CAN), if not otherwise stated. Analyses by treatment cycle will be based on the subset of subjects in the SP with treatment in the respective cycle.

6.2.1 Adverse Events

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

In the electronic case report form (eCRF), AEs will be reported with their onset and information on study discontinuation (tick box "AE leading to Discontinuation of the Study" = "Yes") and treatment discontinuation (tick box "Action taken with Medical Device" = "Treatment suspended"). The start date of the AE will be the date of the record reporting the onset of the AE and the stop date will be the stop date or ongoing, if not yet resolved. For seriousness, intensity, causal relationship, and outcome the worst attributes of each AE will be documented. 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.

TEAEs are defined as AEs with onset at or after the first administration of study treatment. The handling of missing start and stop dates are described in CCI TEAEs will be assigned to a treatment cycle (see treatment cycles in Section 6) based on the following rules:

- For subjects of treatment groups TN or TC (and TRT), TEAEs will be grouped into two treatment cycles according to the AE onset date, as follows:
    - <u>For subjects without retreatment:</u>

An AE is considered treatment emergent for **Cycle 1** if the start date of the AE is on or after the date of the initial treatment injection.

- For subjects with retreatment:

  An AE is considered treatment emergent for **Cycle 1** if the start date of the AE is on or after the date of the initial injection and before the date of retreatment.

  An AE is considered treatment emergent for **Cycle 2** if the start date of AE is on or after the date of the retreatment. In case of any doubts because of exact day or time of onset information missing, an AE will be rather be assigned to Cycle 2 than to Cycle 1.

- For subjects of groups DTN or DTC (and DTRT), an AE is considered treatment emergent for **Cycle 1** if the start date is on or after the date of delayed treatment at Week 8.

All other AEs are considered to be non-TEAEs (i.e., all AEs with start date prior to initial injection).

For TEAEs, an overall summary table will be provided, displaying the following content by treatment cycle and overall:

- Any TEAEs;
- Any related TEAEs;
- Any TESAEs;
- Any related TESAEs;
- Any TEAEs leading to discontinuation from study;
- Any related TEAEs leading to discontinuation from study;
- Any TEAEs leading to discontinuation of study treatment;
- Any related TEAEs leading to discontinuation of study treatment;
- Any fatal TEAEs and
- Any fatal TEAEs related to treatment.

For TEAEs, related TEAEs, TESAEs, non-serious TEAEs, TEAEs leading to discontinuation from study and related TESAEs, number of subjects with TEAEs and number of TEAEs will be summarized by SOC and PT for each treatment cycle and overall. TESAEs, TEAEs leading to discontinuation from study and related TESAEs will only be summarized if more than 5 subjects (over all treatment groups) experienced such a TEAE.

In addition, the following TEAE incidences will be shown overall and by treatment cycle:

- Number of subjects with TEAEs by PT;
- Number of subjects with TEAEs by worst intensity, by SOC and PT;
- Number of subjects with TEAEs by worst causal relationship, by SOC and PT;
- Number of subjects with TEAEs by worst outcome, by SOC and PT; and
- Number of subjects with TEAEs by maximum duration, by SOC and PT (duration categories: 1 to 3 days, 4 to 7 days, 8 to 14 days, 15 to 28 days, and > 28 days);
- Number of subjects with related TEAEs by worst intensity;

- Number of subjects with related TEAEs by maximum duration (duration categories: 1 to 3 days, 4 to 7 days, 8 to 14 days, 15 to 28 days, and > 28 days);

Moreover, a summary of all related TEAEs will be provided, summarizing the number of events by severity and by duration.

The duration of TEAEs will be calculated by the difference of start and stop date plus 1 day. Missing start and stop days will be imputed as described in CCI If a TEAE is ongoing at study end, TEAE duration will be calculated as the difference between TEAE start date and last contact date + 1.

Moreover, the TEAE overview tables as well as incidences and number of related TEAEs (by PT and SOC) will be provided for the following subgroups, overall and by treatment cycle:

- Baseline MIHAS (2 vs. 3);
- Race (white vs. other);
- Ethnicity (Hispanic or Latino vs. not Hispanic or Latino);
- Age (< 45 years vs. ≥ 45 years);
- Touch-up (yes/no) (applies to Cycle 1 only);
- Sex; and
- Fitzpatrick skin type (I, II, III vs. IV, V, VI).

For all randomized subjects, a summary table presenting the number and percentage of subjects who experienced any AE at any time during study and the number and percentage of subjects with any non-TEAEs will be displayed by randomized treatment groups (TRT, CDTRT) and overall. This analysis will use subjects as randomized and will be based on all randomized subjects.

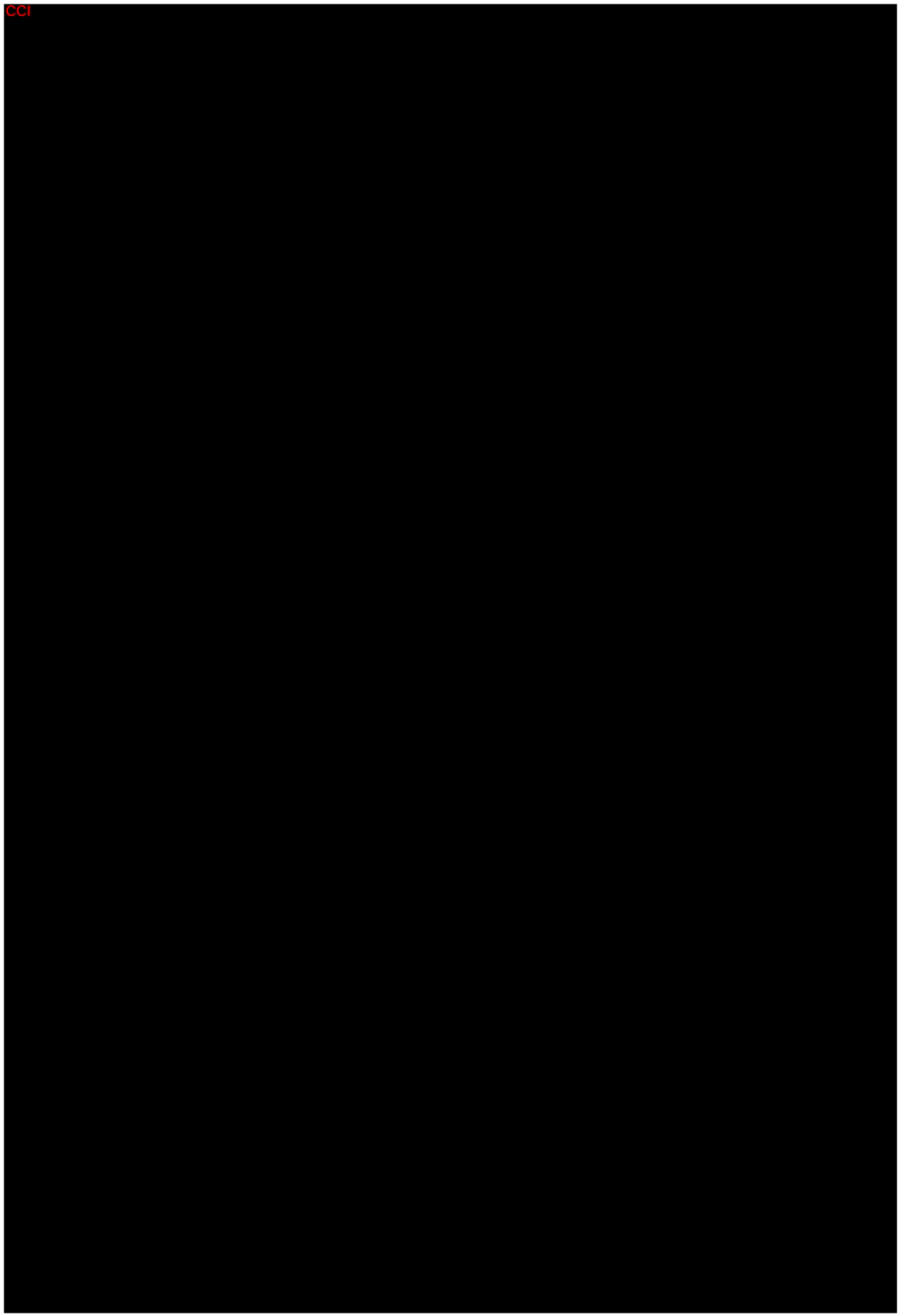
Additionally, AEs reported for subjects in the CTL group from date of randomization until Week 8 (pre-injection) will be summarized separately showing number of subjects with AEs and number of AEs by system organ class (SOC) and preferred term (PT). For subjects who received treatment at Week 8, AEs that started on the day of injection will not be considered in this analysis. For subjects who never received treatment, all AEs until end of study will be considered.

Listings for related TEAEs, TEAEs leading to study discontinuation, serious TEAEs, and serious TEAEs leading to death will be provided in the corresponding table section. For related TEAEs, AEs with delayed onset will be flagged. A delayed onset AE is defined as an AE with an onset more than 28 days after the most recent injection (i.e., the latest injection prior to the AE onset).

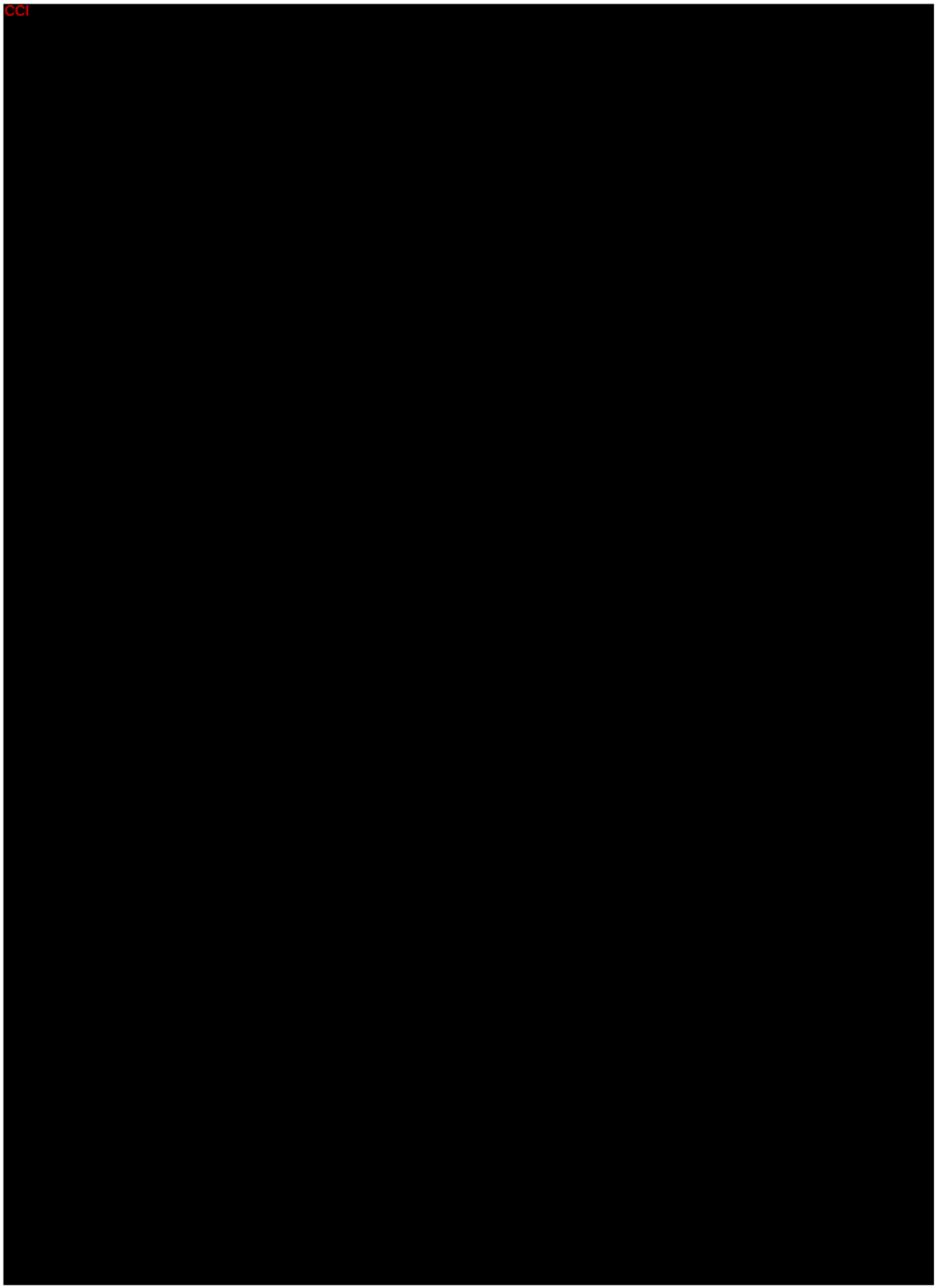
All AEs including non-TEAEs, as well as all reported device deficiencies, will be listed with all details recorded in the listings section.

CCI

CCI



CCI



Disposition of subjects

The number of subjects screened will be provided. Other subject disposition information (i.e., number of subjects randomized, treated and re-treated, as well as the number of subjects in each analysis population) will be summarized by treatment group and overall (TRT, CDTRT and TOTT) and additionally by treatment group and injection technique (needle/cannula) and overall (TN, TC, CDTN, CDTC, and TOTT). Similar descriptive analyses will be performed by site in a separate table.

The absolute and relative frequencies for subject's main reason for premature study discontinuation will be tabulated overall and separately for discontinuations before/at Week 8 and after Week 8 for various treatment groups (TN, TC, TRT, CDTN, CDTC, CDTRT and TOTT), based on all randomized subjects.

The time in weeks from randomization until premature discontinuation from study up to Week 8 will be displayed using a Kaplan-Meier curve. This curve will be provided for subjects in TRT and CTL groups based on all subjects randomized. Subjects who do not discontinue the study within the specified time window will be presented as censored at the date of their Week 8 visit. Subjects in TRT group that missed their Week 8 visit but did not discontinue will be censored at the target day of the Week 8 visit (i.e. Day 57 after last injection in treatment Cycle 1). For subjects who attended only visits up to Week 8 with date of last contact after Week 8, the date of discontinuation will be set to the date of the Week 8 visit.

The time in months from first treatment until premature discontinuation from study will be described by a second Kaplan-Meier curve for the TRT and DTRT groups in SP. Subjects who complete the study will be presented as censored at end of study.

Visit attendance (subjects expected at visit, subjects with visit performed on-site, subjects with visit performed virtually, subjects with missed visit due to COVID-19 pandemic or other reasons, subjects discontinued) will be presented in a frequency table for various treatment groups (i.e., TN, TC, TRT, CDTN, CDTC, CDTRT) based on all randomized subjects.

For 'Study discontinued', the cumulative number of subjects who discontinued the study before the respective visit will be displayed. For 'Subjects expected at visit', the following visits are defined as expected for all randomized subjects: Screening visit, initial treatment visit V1-a, V1-b, V2-a, V3 – V7-a, V8, V9 for TRT subjects; and Screening visit, visits C1-a, C1-b, C2-a, C3 and all following visits for CDTRT subjects, as defined in the visit schedules in the CSP [redacted] for TRT subjects and [redacted] for CDTRT subjects). Visits V2-b/C2-b and V2-c/C2-c are

only expected for subjects receiving an optional touch-up at visit V2-a/C2-a. In case of retreatment at visit V7-a for TRT subjects, follow-up visits V7-b and V7-c are also expected. For subjects who discontinue the study, the end of study visit V9/C7 is expected. Visits after premature discontinuation are generally not expected. The number of 'Subjects expected at visit' is thus calculated as number of randomized subjects – the number of subjects discontinued before the respective visit.

Percentages of 'Visit performed' (total) will be calculated relative to the number of subjects expected at the respective visit.

For columns relating to 'Subjects expected at visit', 'Study discontinued', 'Visit performed' (regularly on-site, virtually due to Coronavirus disease 2019 (COVID-19) pandemic) and 'Visit not performed' (total, due to COVID-19 pandemic, due to other reasons), only absolute numbers will be displayed. All performed and all not performed visits until regular study completion or premature discontinuation (as applicable for the respective subject) will be listed. Flags will also be included for visits performed virtually, visits not performed due to COVID-19 pandemic and visits not performed due to other reasons.

Protocol Deviation

The absolute and relative frequencies of major protocol deviations and other reasons for exclusion from the PP population will be tabulated, based on all randomized subjects for groups TN, TC, TRT, CTL and overall.

Additionally, major protocol deviations and other reasons for exclusion from the PP population related to COVID-19 pandemic will be presented in a separate table. For this purpose, each protocol deviation will be categorized whether it was related to the COVID-19 pandemic or not.

Minor protocol deviations will be listed only.

A listing of all documented impacts of COVID-19 pandemic on study conduct (incl. disruptions and modifications not constituting protocol deviations) will be provided. This listing will present investigational site identifier, randomized treatment group and information on impact of COVID-19 pandemic for all randomized subjects. In case the study conduct was impacted by COVID-19 pandemic, classified disruption or modification (e.g. visit performed virtually, time schedule deviation) and details on disruption or modification will be presented.

Demographic data and baseline characteristics

Demographic data (age, age category, sex, ethnicity, race and Fitzpatrick skin type) as well as weight, height and BMI and MIHAS baseline score as assessed by the blinded evaluator will be summarized by various treatment groups (i.e., TN, TC, TRT, CDTN, CDTC, CDTRT) and overall for the ITT, SP, PP populations and for all randomized subjects. Fitzpatrick skin type will additionally be categorized into type I-III and type IV-VI. Age will additionally be presented in categories of "< 45 years" and "≥ 45 years".

Previous and concomitant medication and non-drug treatment

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 database is closed will be used for analysis.

Rules for separation of previous and concomitant medication are given in CCI also accounting for missing or incomplete start and stop dates.

Frequencies of previous and concomitant treatments will be given based on Anatomical Therapeutic Chemical (ATC) Classification System code levels 2 and 3 for the SP by treatment

groups TN, TC, TRT, CDTN, CDTC, CDTRT, and overall. Indications for concomitant therapies will not be coded and will only be listed.

Non-drug treatments will be coded using the MedDRA version which is in effect at the time the database is closed. Non-drug treatments will be displayed by SOC and PT levels for the SP by groups TN, TC, TRT, CDTN, CDTC, CDTRT, and overall.

Medical history and concomitant diseases

Medical history and concomitant diseases will be coded using the MedDRA dictionary and reported by SOC and PT levels for the SP by groups TN, TC, TRT, CDTN, CDTC, CDTRT and overall.

Rules for separation of medical history and concomitant diseases for missing or incomplete dates are given in CCI [redacted] also accounting for missing or incomplete stop dates.

Extent of exposure

Frequencies of subjects receiving initial treatment, touch-up and retreatment, if applicable, will be summarized by various groups (i.e., TN, TC, TRT, DTN, DTC, DTRT, NED, CAN, and TOTT) for SP.

Data on injection volume, technique, number of injection points, injection plane, and if ice was used (pre- and post-injection) will be analyzed for each treatment of the right IOH and left IOH separately. Injection volume will also be summarized for both IOHs combined. Number of injection points will be presented via summary statistics and also as frequencies using categories of 1, 2, 3, 4-9 and 10-18 injection points. These analyses will be displayed by various groups (i.e., TN, TC, TRT, DTN, DTC, DTRT, NED, CAN, and TOTT) for SP.

Further, injection-technique information (cannula/needle) will be summarized for TRT, DTRT and TOTT groups based on SP for each treatment of the right IOH and left IOH separately.

Additionally, frequencies of reasons for performed and not performed touch-up and retreatment will be summarized by treatment groups TN, TC and TRT for SP.

Listings on extend of exposure will be presented based on all subjects treated.

6.4 Special Statistical/Analytical Issues

6.4.1 Discontinuations and Missing Data

Subjects discontinuing the study prematurely should perform and record end-of-study assessments at V9 (subjects in TRT) or at C7 (subjects in CDTRT), unless they are lost to follow-up. Effectiveness data of discontinued subjects collected at these visits will therefore be analyzed under the study visit which is closest to the recorded V9/C7 date. This will be based on the assessment date relative to last treatment and treatment cycle, using the following visit windows:

Case 1: Subject without touch-up

Timepoint	Subject Status	Window Start	Target Day of Visit	Window End
Cycle 1: V1-b/C1-b		Day 2 (after initial treatment)	Day 15 (after initial treatment)	Day 22 (after initial treatment)
Cycle 1: V2-a/C2-a		Day 23 (after initial treatment)	Day 29 (after initial treatment)	Day 43 (after initial treatment)
Cycle 1: V3/C3		Day 44 (after initial treatment)	Day 57 (after initial treatment)	Day 71 (after initial treatment)

Timepoint	Subject Status	Window Start	Target Day of Visit	Window End
Cycle 1: V4/C4		Day 72 (after initial treatment)	Day 85 (after initial treatment)	Day 127 (after initial treatment)
Cycle 1: V5/C5		Day 128 (after initial treatment)	Day 169 (after initial treatment)	Day 211 (after initial treatment)
Cycle 1: V6/C6		Day 212 (after initial treatment)	Day 253 (after initial treatment)	Day 295 (after initial treatment)
Cycle 1: C7		Day 296 (after initial treatment)	Day 337 (after initial treatment)	Open end
Cycle 1: V7-a	No retreatment at V7-a	Day 296 (after initial treatment)	Day 337 (after initial treatment)	Day 365 (after initial treatment)
Cycle 1: V8	No retreatment at V7-a	Day 366 (after initial treatment)	Day 393 (after initial treatment)	Day 449 (after initial treatment)
Cycle 1: V9	No retreatment at V7-a	Day 450 (after initial treatment)	Day 505 (after initial treatment)	Open end
Cycle 2: V7-a	Retreatment at V7-a	No end of study data will be assigned to this visit as subject received retreatment at this visit.		
Cycle 2: V7-b	Retreatment at V7-a	Day 2 (after retreatment)	Day 15 (after retreatment)	Day 22 (after retreatment)
Cycle 2: V7-c	Retreatment at V7-a	Day 23 (after retreatment)	Day 29 (after retreatment)	Day 43 (after retreatment)
Cycle 2: V8	Retreatment at V7-a	Day 44 (after retreatment)	Day 57 (after retreatment)	Day 113 (after retreatment)
Cycle 2: V9	Retreatment at V7-a	Day 114 (after retreatment)	Day 169 (after retreatment)	Open end

Case 2: Subject with touch-up

Timepoint	Subject Status	Window Start	Target Day of visit	Window End
Cycle 1: V1-b/C1-b		No end of study data will be assigned to this visit as subject received touch-up after this visit.		
Cycle 1: V2-a/C2-a		No end of study data will be assigned to this visit as subject received touch-up at this visit.		
Cycle 1: V2-b/C2-b		Day 2 (after touch-up)	Day 15 (after touch-up)	Day 22 (after touch-up)
Cycle 1: V2-c/C2-c		Day 23 (after touch-up)	Day 29 (after touch-up)	Day 43 (after touch-up)
Cycle 1: V3/C3		Day 44 (after touch-up)	Day 57 (after touch-up)	Day 71 (after touch-up)
Cycle 1: V4/C4		Day 72 (after touch-up)	Day 85 (after touch-up)	Day 127 (after touch-up)
Cycle 1:		Day 128 (after touch-up)	Day 169 (after touch-up)	Day 211 (after touch-up)

Timepoint	Subject Status	Window Start	Target Day of visit	Window End
V5/C5				
Cycle 1: V6/C6		Day 212 (after touch-up)	Day 253 (after touch-up)	Day 295 (after touch-up)
Cycle 1: C7		Day 296 (after touch-up)	Day 337 (after touch-up)	Open end
Cycle 1: V7-a	No retreatment at V7-a	Day 296 (after touch-up)	Day 337 (after touch-up)	Day 365 (after touch-up)
Cycle 1: V8	No retreatment at V7-a	Day 366 (after touch-up)	Day 393 (after touch-up)	Day 449 (after touch-up)
Cycle 1: V9	No retreatment at V7-a	Day 450 (after touch-up)	Day 505 (after touch-up)	Open end
Cycle 2: V7-a	Retreatment at V7-a	No end of study data will be assigned to this visit as subject received re-treatment at this visit.		
Cycle 2: V7-b	Retreatment at V7-a	Day 2 (after re-treatment	Day 15 (after retreatment)	Day 22 (after retreatment)
Cycle 2: V7-c	Retreatment at V7-a	Day 23 (after re-treatment	Day 29 (after retreatment)	Day 43 (after retreatment)
Cycle 2: V8	Retreatment at V7-a	Day 44 (after re-treatment	Day 57 (after retreatment)	Day 113 (after re-treatment
Cycle 2: V9	Retreatment at V7-a	Day 114 (after re-treatment	Day 169 (after retreatment)	Open end

If based on the above windows, a V9/C7 end of study assessment would be assigned to a timepoint for which a regular visit was performed, then the data from the respective visit will be used for analysis, and the V9/C7 data will only be listed. Also, no data recorded under V9/C7 (end of study) visit will be assigned to a visit with initial, touch-up or retreatment performed.

If V8/C6 was performed and the subject later prematurely discontinued the study, V9/C7 visit will be analyzed as documented.

Effectiveness data

Methods for handling of missing primary effectiveness endpoint data are described in Section 6.1.1.

Handling of missing scores of FACE-Q Satisfaction with Eyes scores is described in Section 6.1.2.
Handling of missing scores of FACE-Q Appraisal of Lower Eyelids is described in Section 6.1.3.

For remaining endpoints, observed cases will be analyzed, if not otherwise specified.

Safety data

For the imputed analysis start date of AEs (variable ASTDT in the ADaM domain ADAE), imputation rules related to any cycle are described in Section 6.

Missing data of the worst outcome will be imputed by “unknown”. For any other missing data of AEs, a worst-case strategy will be applied for all analysis tables. The intensity will be imputed by worst intensity “severe” and the causal relationship will be imputed by the worst causal relationship “related”.

Rules for handling missing eDiary entries are described in Section 6.2.

6.4.2 Interim Analyses

No interim analyses are planned.

6.4.3 Multiple Comparisons/Multiplicity

Based on the primary effectiveness endpoint, four hypothesis tests will be performed. To control for multiplicity, a hierarchical-testing procedure will be performed (see Sections 6.1.1.1 and 6.1.1.3).

6.4.4 Examination of Subgroups

Subgroup analyses are specified in Section 6.1.1.3 for the primary effectiveness endpoint, for the secondary effectiveness endpoints in Section 6.1.2, and for IPR endpoints in Section 6.1.1.3

Subgroup analyses for safety endpoints are described in Section 6.2.1 for TEAE and in Section 6.2.2 for CTR analyses.

6.4.5 Pooling of Sites

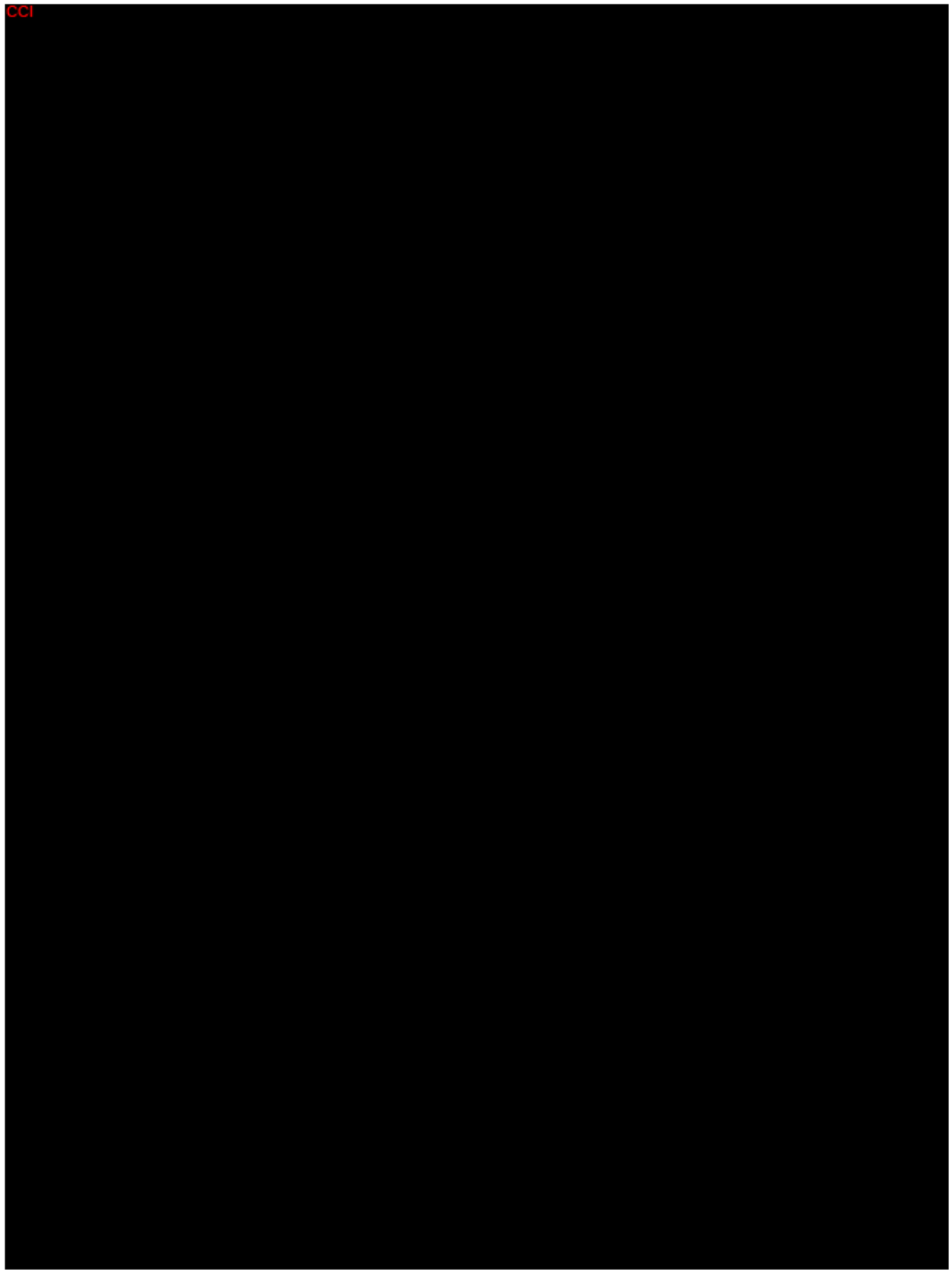
Pooling of investigative sites will be performed according to the algorithm below and only for the usage as a factor in the Multiple Imputation model and in the logistic regression model. For the Multiple Imputation model, pooling will be performed for TRT and CTL groups separately. If convergence problems are observed in either of the two models, the following steps will be performed.

1. The site with the lowest number of randomized subjects will be pooled with the site with second lowest number of subjects. Regarding the logistic regression model, if two or more sites with the same number of subjects are suited for pooling, the site with the lowest number of subjects in the control group will be pooled. If two or more sites with the same number of subjects and same number of subjects in the control group (only applicable for logistic regression model) are suited for pooling, then the site to be pooled will be chosen randomly among all suited sites.
2. Logistic regression/Multiple Imputation will be re-run with pooled site as factor.
3. If convergence problems are resolved, no further pooling of investigative sites will be performed. Otherwise, steps 1) and 2) above will be repeated until convergence problems resolve. If pooling of sites does not resolve convergence problems (unless all sites are pooled), site will be completely removed from the model.

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

[2] Yulia Sidi & Ofer Harel (2021): Difference Between Binomial Proportions Using Newcombe's Method With Multiple Imputation for Incomplete Data, The American Statistician, DOI: 10.1080/00031305.2021.1898468