NCT04795622, redacted
version v1.0, 16Jan2023

Statistical Analysis Plan

Title:	Evaluation of the Effectiveness and Safety of the Ulthera® DeepSEE® System for Treating Skin Laxity in the Lower Face and Submentum
Merz Study Number:	M960101056
SAP for	Final Analysis
Sponsor:	Ulthera, Inc. a division of Merz North America, Inc.
Version:	1.0
Date:	30-Sep-2022
Author:	[redacted]

Document References (For Internal Use Only)	
[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.

[redacted]		
Author [redacted] (print name)	Date (dd-mmm-yyyy)	Signature

[redacted]		
Peer Reviewer [redacted] (print name)	Date (dd-mmm-yyyy)	Signature

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.

[redacted]		
[redacted] Merz (print name)	Date (dd-mmm-yyyy)	Signature

[redacted]		
[redacted] Merz (print name)	Date (dd-mmm-yyyy)	Signature

Table of Contents

List of Abbreviations

Abbreviation	Definition
AE	Adverse event
ANOVA	Analysis of variance
ATC	Anatomical Therapeutic Chemical classification system of the World Health Organization
BMI	Body mass index
CI	Confidence interval
CRF	Case report form
CTP	Clinical trial protocol
[redacted]	
eCRF	Electronic case report form
EOT	End of Trial
FAS	Full analysis set
GAIS	Global Aesthetic Improvement Scale
iGAIS	Investigator Global Aesthetic Improvement Scale
IMD	Investigational medical device
MAR	Missing at random
max	maximum
MedDRA	Medical Dictionary for Regulatory Activities
min	minimum
MNAR	Missing not at random
n	Number of values analyzed
NTF	Note to file
PD	Protocol deviation
PEV	Primary endpoint visit
PPS	Per protocol set
[redacted]	
PT	Preferred term
Q1	Lower quartiles
Q3	Upper quartiles
QC	Quality control
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Statistical Analysis Software®
SD	Standard deviation

SES	Safety evaluation set
sGAIS	Subject Global Aesthetic Improvement Scale
SOC	System organ class
SOP	Standard operating procedure
TEAE	Treatment-emergent adverse event
TESAE	Treatment-emergent serious adverse event
USA	United States of America
VAR	SAS variable

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 trial protocol M960101056, version 3.0, dated 08 Jul 2021. CRF, version 2.0, dated 01 Sep 2021.

All programs will be written using SAS[redacted]. A preferred font size of 10 points for the whole document will be used for the tables and figures in Section 14, corresponding to a linesize of 111 digits and a pagesize of 42 lines for an output in A4 landscape format. For listings, a standard font size of 10 points with the linesize and pagesize as defined above 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 files using the Merz internal SAS macro LST2PDF. These PDF files will be generated as needed to populate the subsections of Section 14 and Section 16.2 for the clinical trial/investigational 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 will be applied and adapted to trial specific requirements as laid down in the clinical trial protocol M960101056, version 3.0 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, SDTMs, ADaMs as applicable) will be documented in the external NTF named M960101056_NTF_QC_Statistical Programming_v1.0_20220307.

2 Clinical Trial Design and Objectives

2.1 Clinical Trial Design

This is a prospective, 180-day, randomized, multicenter, evaluator-blind, controlled trial designed to evaluate the effectiveness and safety of the Ulthera DeepSEE (DS) System improving the appearance of skin laxity of the lower face and submentum. Subjects will be enrolled from participating investigative sites in China. A total of 200 subjects will be randomized 1:1 to either the Treatment Group or Control Group stratified by investigative site.

The Treatment Group will receive a single treatment at Day 1 with the Ulthera DeepSEE System to the midface, lower face, submentum, and upper neck with three DS transducers (DS 10-1.5, DS 7-3.0, and DS 7-4.5) and will be followed for 180 days after treatment.

The Control Group will not receive treatment at Day 1 and will remain untreated until Day 90 (delayed treatment). Following completion of all protocol-mandated assessments at Day 90, the Control Group will receive a single treatment with the Ulthera DeepSEE System to the midface, lower face, submentum and upper neck with three DS transducers (DS 10-1.5, DS 7-3.0, and DS 7-4.5) and will be followed for 90 days after treatment.

For both the Treatment Group and Control Group, the starting energy level (EL) will be EL2 for the DS 10-1.5 transducer, and EL4 for the DS 7-3.0 and DS 7-4.5 transducers. [redacted]

All subjects will participate in this trial for 180 days and will be evaluated at screening, Day 1, Day 90, and Day 180. Follow-up telephone calls will occur at 3 days and 14 days after treatment. In case of ongoing transient nerve effects 14 days after treatment, trial personnel will conduct

additional follow-up phone calls with the subject every week until 6 weeks after treatment and document the time of resolution.

After a maximum 14-day screening period, the planned trial duration for individual subjects is 180 days (± 14 days).

[redacted]

2.2 Randomization

Subjects are randomized balanced (1:1) to one of the following groups:

- Treatment Group: Treatment with the Ulthera DeepSEE System at Day 1; or
- Control Group: Untreated control, with delayed treatment at Day 90.

Randomization will be stratified by investigative site and performed in blocks.

2.3 Blinding

In this prospective trial, the treating investigators and subjects will not be blinded to trial treatment. For the primary effectiveness endpoint only (proportion of subjects with improvement in lower face and submental skin laxity), the [redacted] blinded evaluators will be blinded to trial group assignment. [redacted]

2.4 Unblinding

2.4.1 Planned Unblinding

Planned unblinding procedures are not necessary in the current trial design, as the treating investigator, trial coordinator, sponsor staff, and other trial personnel [redacted] are not blinded to trial group assignments.

[redacted]

2.5 Trial Objectives

2.5.1 Effectiveness

Demonstrate superiority of treatment with the Ulthera DeepSEE System compared to untreated control for the improvement of skin laxity of the lower face and submental area.

2.5.2 Safety

Demonstrate the safety of treatment with the Ulthera DeepSEE System for the improvement of skin laxity of the lower face and submental area.

3 Determination of Sample Size

[redacted]

[redacted]

4 Analysis Sets

The following analysis sets will be defined:

- The randomized set is defined as all subjects randomized into the trial;
- The Safety Evaluation Set (SES) is defined as all subjects treated;
- The Full analysis Set (FAS) is defined as all subjects randomized [redacted]. Subjects randomized to the Treatment Group must also finish the treatment;
- The Per Protocol Set (PPS) is the subset of subjects in the FAS who have no major protocol deviations potentially biasing the statistical analysis of the primary endpoint. Final determination of what constitutes major or minor protocol deviations in this sense will be made prior to database close.

5 Endpoints for Analysis

5.1 Effectiveness Endpoints

5.1.1 Primary Effectiveness Endpoint

Proportion of subjects with improvement in lower face and submental skin laxity at Day 90 [redacted].

[redacted]

5.1.2 Secondary Effectiveness Endpoints

5.1.2.1 Displacement of skin in the submentum at Day 90

Displacement of skin (mm) in the submentum at Day 90 [redacted].

5.1.2.2 Proportion of Subjects with Any Improvement on iGAIS at Day 90

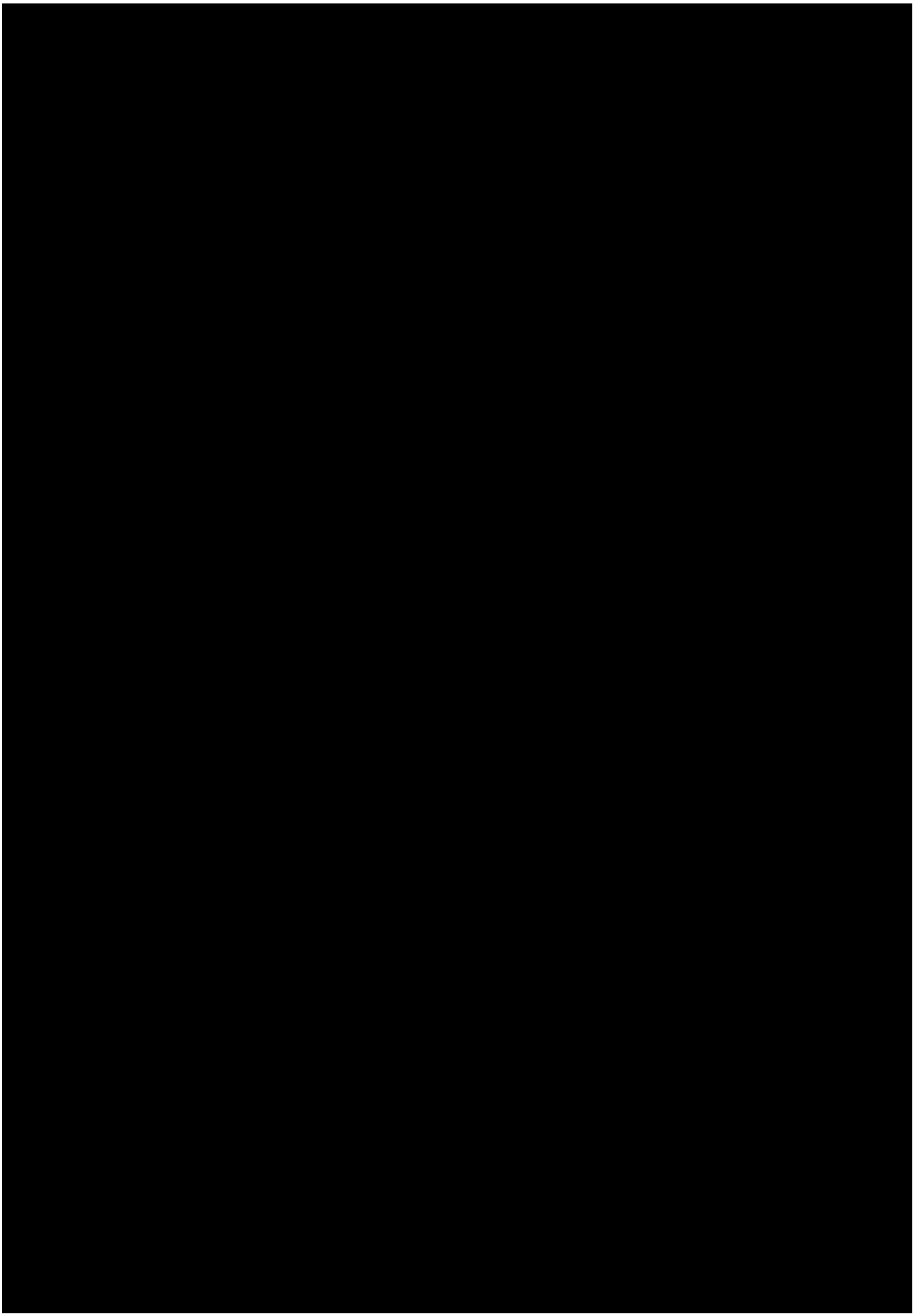
Proportion of subjects with any improvement [redacted] on the Investigator Global Aesthetic Improvement Scale (iGAIS) at Day 90 [redacted], as assessed by the treating investigator in the Treatment Group.

5.1.2.3 Proportion of Subjects with Any Improvement on sGAIS at D90

Proportion of subjects with any improvement [redacted] on the Subject Global Aesthetic Improvement Scale (sGAIS) at Day 90 [redacted], as assessed by the subjects in the Treatment Group.

5.1.2.4 Change from baseline in FACE-Q™ Satisfaction at Day 90

Change from baseline in FACE-Q™ Satisfaction with Lower Face and Jawline score (Rasch-transformed) at Day 90, as assessed by the subjects in the Treatment Group.



5.2 Safety Endpoints

5.2.1 Secondary Safety Endpoint

Incidence of treatment emergent adverse events (TEAEs) related to theUlthera DeepSEE System, as reported by the treating investigator throughout the trial.

5.2.3 Definitions Related to Safety Endpoints

5.2.3.1 TEAE

Treatment-emergent AEs (TEAEs) are defined as AEs with onset at or after the administration of trial treatment. For subjects randomized to treatment, an AE is considered a TEAE if the start date is on or after treatment at Day 1. For subjects randomized to control, an AE is considered a TEAE if the start date is on or after treatment at Day 90.

5.2.3.2 TESAE

TESAEs are defined as TEAEs which meet any seriousness criterion of SAE.

5.2.3.3 Treatment related TEAE

Treatment related AEs are defined as AEs whose causal relationship are related or missing.

6 Statistical Analysis Methods

Metric statistics will be number of values analyzed (n), number of observed values (n obs), number of imputed values (n imp), mean, standard deviation (SD), median, lower quartile (Q1), upper quartile (Q3), minimum (min), and maximum (max). 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 explicitely be given in the Sections below.

Frequency tables for qualitative endpoints/variables will present absolute and percent frequencies (n, %). Frequency tables will include the number of missing values unless otherwise specified. Percentages will be calculated using the denominator of all subjects in a specified population or treatment group. 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 two-sided at type I error rate 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'. P-values above 0.9999 will be presented as '>0.9999'. Confidence intervals (CIs) will be two-sided 95% CIs.

All statistical analyses will be performed using SAS.

6.1 Effectiveness Endpoints

If not otherwise specified, effectiveness analyses will be based on observed cases in the FAS. For the FAS, all subjects will be analyzed as randomized.

6.1.1 Primary Effectiveness Endpoint

6.1.1.1 Primary analysis

Proportion of subjects with improvement at Day 90 = number of subjects with improvement at Day 90 / number of subjects in the analysis set and group * 100, the result will be rounded to 1 decimal place.

If improvement is considered as event, the null and alternative hypotheses to be tested are:

H_0 : odds ratio of Treatment vs. Control = 1 (null hypothesis)

versus H_1 : odds ratio of Treatment vs. Control ≠ 1 (alternative hypothesis).

[redacted] Effectiveness of treatment is confirmed if the p-value of the test is ⩽ 5% and the odds ratio is in favor of Treatment (i.e., > 1).

[redacted]

6.1.1.2 Sensitivity analyses

The following sensitivity analyses will be conducted to support the primary analysis:

1. Observed cases: the same approach as outlined for the primary analysis will be applied to observed cases instead of multiple imputation.

2. PPS: the observed cases analysis will be applied to the PPS.

3. Site effects [redacted].

4. Randomized set: the same approach as outlined for the primary analysis will be applied to all subjects in the randomized set [redacted].

6.1.1.3 Subgroup analyses

No subgroup analysis is planned for primary effectiveness endpoint.

6.1.2 Secondary Effectiveness Endpoints

Displacement of skin (mm) in the submentum in X/Y/Z-direction at Day 90 = the results (mean, median, min and max) of displacement of skin in the submentum in X/Y/Z-direction vectors of all points at Day 90 for the individual subject. Mean and median displacement in skin in the submentum at Day 90 will be analyzed [redacted].

[redacted]

The other secondary effectiveness endpoints will be analyzed by descriptive statistics including CIs. For the iGAIS and sGAIS assessments, the proportion of subjects in the Treatment Group with any improvement [redacted] will be calculated and given as number of subjects with improvement / number of subjects with rating in the corresponding analysis (n/N) and as percentage (%). Specifically, proportion of subjects with any improvement on iGAIS (%) at Day 90 [redacted] on iGAIS at Day 90 / number of subjects with an iGAIS score at this visit in the analysis set and group * 100; proportion of subjects with any improvement on sGAIS (%) at Day 90 [redacted] on sGAIS at Day 90 / number of subjects with a sGAIS score at this visit in the analysis set and group * 100. Results will be rounded to 1 decimal place.

For the FACE-Q satisfaction with lower face and jawline, Rasch-transformed scores, ranging from 0 to 100, will be summarized for the Treatment Group using descriptive statistics. Higher scores reflect a better outcome. If missing data are less than 50% of the scale's items, the mean (round to integer) of the completed items will be inserted. Otherwise, the score will not be imputed and summarized. The following table will be used for conversion of sum scores (range: 5 to 20) to Rasch- transformed scores (range: 0 to 100):

[redacted]

6.2 Safety Endpoints

All safety analyses will be performed on the SES. Subjects will be analyzed as treated. Moreover, safety analyses will be based on the pooled randomized groups if not otherwise specified.

Adverse events will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) version that is in effect at the time the database is closed to allow for analysis by System Organ Class (SOC) and Preferred Term (PT).

Only TEAEs and TESAEs will be summarized. Similarly, only those device deficiencies that are reported to be associated with devices that were used to treat subjects will be summarized.

Frequency tables will give incidence rates by treatment group and overall.

Calculation of time to onset/duration of AEs (days):

- Time to onset of an AE is defined as start date of AE - date of administration of treatment [+ 1 day for TEAEs]
- The duration will be calculated as stop date - onset date + 1 day for AEs.

In case of missing intensity or missing causal relationship of an AE the worst case principle will be applied, i.e., a missing intensity will be set to “severe” and a missing causal relationship will be set to “related”. Missing outcome will be set to “unknown”.

Absolute values and changes from baseline of vital signs will be analyzed by metric statistics, treatment and visit.

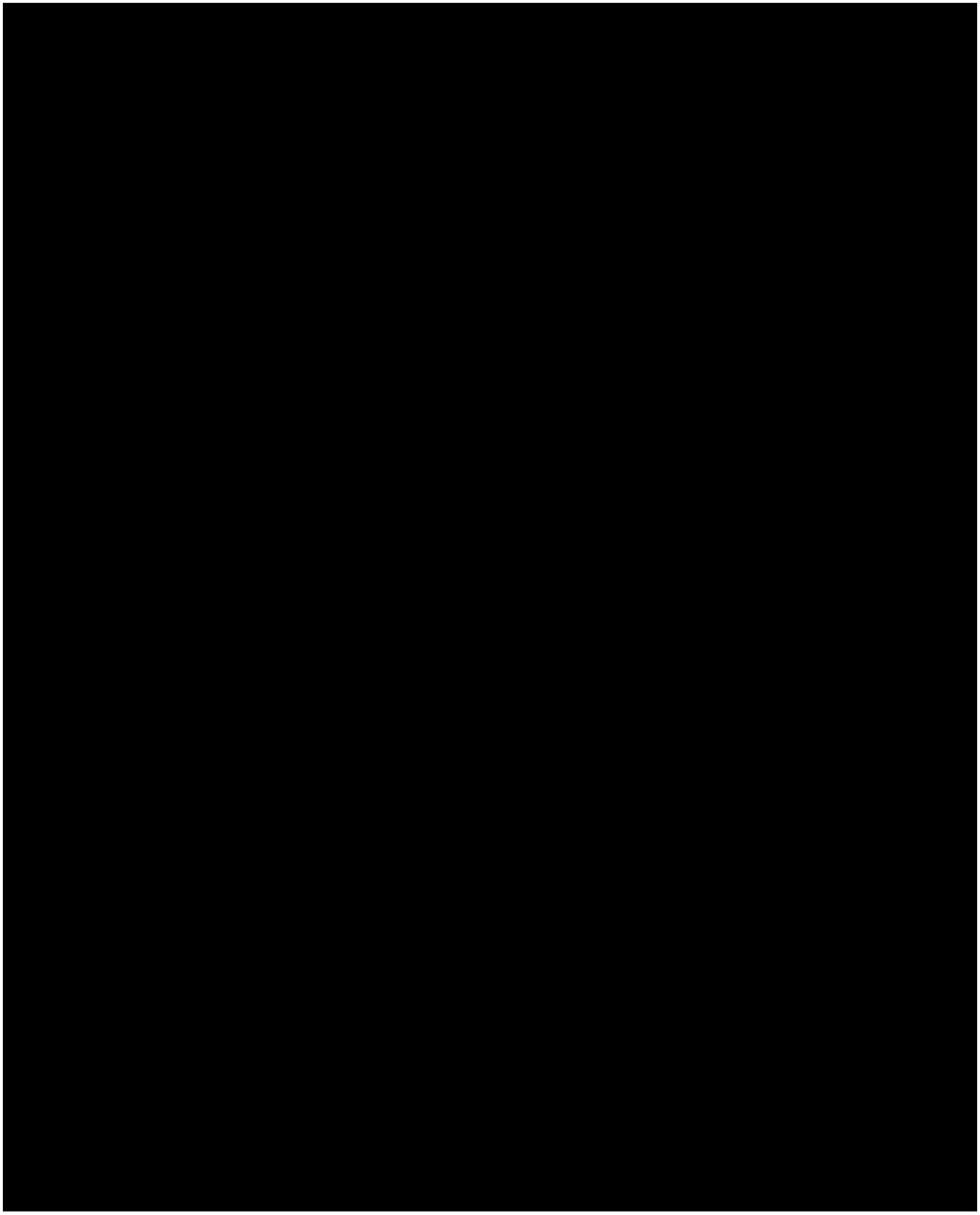
6.2.1 Secondary Safety Endpoints

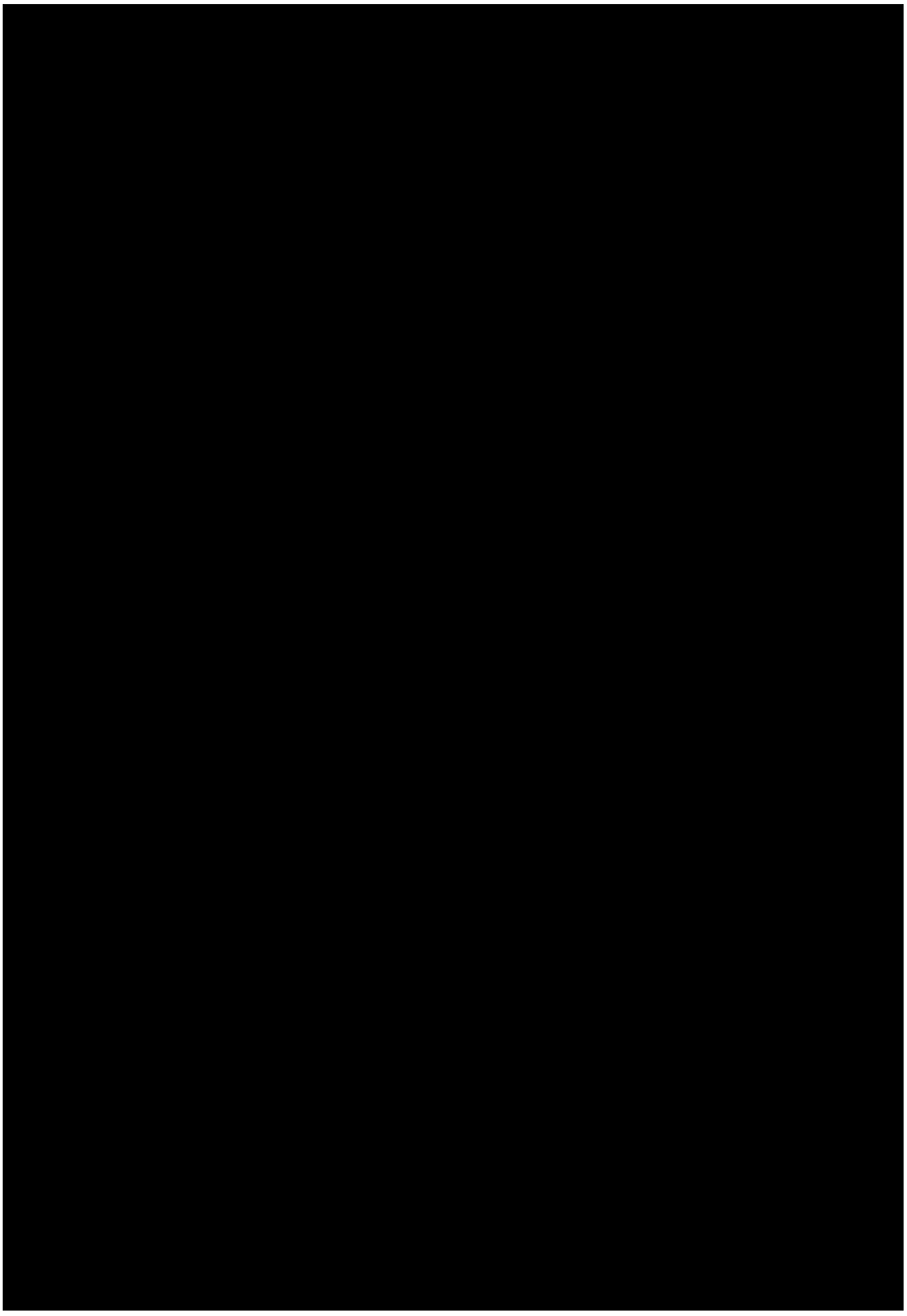
Incidence rates of treatment related TEAEs will be provided by PT:

- Overall for all treated subjects;
- Until 90 days after treatment for all treated subjects;

- Until Day 180 for the Treatment group.

Moreover, incidence rates for treatment related TEAEs will be given by SOC and PT, worst intensity, and worst outcome.TEAEs will be sorted by descending frequencies in SOC and PT in total (overall) column.





6.4 Special Statistical/Analytical Issues

[redacted]

6.4.2 Baseline and Pre-treatment

Baseline is defined for both treatment group and control group as the last observed value or test result before Day 1. Pre-treatment is as same as baseline in treatment group and defined as the last observed value or test result before Day 90 treatment in control group.

6.4.3 Interim Analyses

No interim analysis is planned for this study.

6.4.4 Multiple Comparisons/Multiplicity

No multiplicity adjustments are required. Only one confirmatory analysis of the primary effectiveness endpoint will be performed. [redacted]

[redacted]

8 References

9 Appendix