NCT03583359, redacted
version v1.0, 02May2022

Merz North America, Inc.

Statistical Analysis Plan

Evaluation of Effectiveness and Safety of Radiesse (+) to Improve the Contour of Jawline by Adding Volume to the Jawline

Pivotal Study

M900391004

Version 2.0

Date: 15-JUNE-2020

Author(s):

[redacted], Merz North America, Inc.
[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 (print name)	Date (dd-MMM-yyyy)	Signature

[redacted]

Peer Reviewer (print name)	Date (dd-MMM-yyyy)	Signature

[redacted]
Merz North America, Inc

Peer Reviewer/Contributor (print name)	Date (dd-MMM-yyyy)	Signature

TABLE OF CONTENTS

SIGNATURE PAGE 2
TABLE OF CONTENTS 3
1 LIST OF ABBREVIATIONS 5
2 GENERAL AND TECHNICAL ASPECTS 6
3 Clinical Study Design and Objectives 7
3.1 Clinical Study Design 7
3.2 Clinical Study Objectives 8
4 Determination of Sample Size 9
5 Analysis Sets 9
6 Endpoints for Analysis 11
6.1 Effectiveness Endpoints 11
6.1.1 Primary effectiveness endpoint 11
6.1.2 Secondary effectiveness endpoints 11
6.1.3 [redacted]
6.2 Safety Endpoints 14
6.2.1 Primary safety endpoints 14
6.2.2 [redacted]
7 Statistical Analysis Methods 15
7.1 Effectiveness Endpoints 15
7.1.1 Primary Effectiveness Endpoint 15
7.1.2 Secondary Effectiveness Endpoint 17
7.1.3 [redacted]
7.2 Safety Endpoints 23
7.3 Specificity for Other Summaries of Data and Variables 26
7.3.1 Subject Disposition 26
7.3.2 Subject Populations 26
7.3.3 Medical History 26
7.3.4 Previous Medication and Therapy 27
7.3.5 Demographic Data 27
7.3.6 Exposure Data 27
7.4 Special Statistical/Analytical Issues 28
7.4.1 [redacted]
7.4.2 Interim Analyses 31
7.4.3 [redacted]
8 [redacted]
9 REFERENCES 32

10 Appendices 33

1 LIST OF ABBREVIATIONS

AE	Adverse event
ATC	Anatomical Therapeutic Chemical
BMI	Body mass index
CSP	Clinical study protocol
CSR	Clinical study report
[redacted]	
DMP	Data management plan
DRM	Data review meeting
eCRF	Electronic case report form
FDA	Food and Drug Administration
GAIS	Global aesthetic improvement scale
[redacted]	
ITT	Intent-to-treat
MAR	Missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MJAS	Merz jawline assessment scale
MNAR	Missing not at random
MVTF	Missing value treated as failure
n	Number of non-missing observations
OC	Observed cases
PDF	Portable document format
PP	Per protocol population
PT	Preferred term
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS®	Statistical Analysis System
SD	Standard deviation
SOC	System organ class
TFLs	Tables, figures, and listings
TEAE	Treatment emergent adverse event
TOC	Table of content
[redacted]	

2 GENERAL AND TECHNICAL ASPECTS

The objective of this statistical analysis plan (SAP) is to specify the statistical analyses in greater detail than stated in the clinical study protocol so as to be precise enough to serve as a guideline for statistical programming and creation of tables, figures, and listings (TFLs).

This statistical analysis plan is based on the clinical study protocol (version 4.0), dated 20-July-2018.

All programs will be written using SAS version 9.4 or higher. Individual SAS programs will be written for all tables, figures, and listings. All outputs will be transferred into PDF files. The PDF files will be generated separately for the tables and figures of section 14 and the listings of section 16 of the appendix of the clinical study report (CSR). Each PDF file will include the corresponding table of contents, preceding the content of the file.

[redacted] TFLs will be applied and adapted to study specific requirements as specified in the clinical study protocol (CSP) and any amendments, if applicable.

3 CLINICAL STUDY DESIGN AND OBJECTIVES

3.1 Clinical Study Design

This is a 60-week, prospective, multicenter, randomized, controlled, [redacted] pivotal clinical study to investigate the effectiveness and safety of Radiesse (+) to improve the contour of jawline by adding to the jawline. Approximately 180 subjects with a grade of 2 or 3 on the MJAS will be enrolled into the study. [redacted] Subjects will be enrolled at up to 15 sites in the United States. [redacted]

At the baseline visit, enrolled subjects will be randomized (2:1 allocation ratio) to either treatment with Radiesse (+) or to untreated control. Those allocated to the control group at baseline will remain untreated until assessment of the primary effectiveness endpoint at Week 12; these subjects will then be eligible for delayed treatment with Radiesse (+) in the jawline at Week 12 and will be followed for 48 weeks post-treatment. [redacted] Subjects who achieve a ≥1-point improvement on the MJAS [redacted] may have a touch-up [redacted]. The primary effectiveness endpoint will be assessed 12 weeks post-treatment. Subjects randomized to the treatment group at baseline will have the option for re-treatment with Radiesse (+), upon agreement between the subject and the investigator, at Week 48 and will then be followed for an additional 12 weeks, for a total study duration of 60 weeks. Subjects who do not receive a re-treatment at Week 48 will also be followed until Week 60.

Subjects randomized to the control group at baseline will remain untreated until completion of the primary endpoint assessment at Week 12. After all primary endpoint assessments have been completed these control subjects will be treated with Radiesse (+) (i.e., delayed treatment) and will then be followed 48 weeks post-treatment. These delayed-treatment subjects will be required to have a touch-up in [redacted] if a subject does not achieve a ≥1-point improvement on the MJAS [redacted]. Subjects who achieve a ≥1-point improvement on the MJAS [redacted]

[redacted]
may have a touch-up [redacted] for further optimal correction at the discretion of the treating investigator. Control subjects will not be offered re-treatment.

[redacted]

3.2 Clinical Study Objectives

Effectiveness Objective:

The primary objective of the study is to demonstrate the effectiveness of Radiesse (+) following deep (subdermal and/or supraperiosteal) injection to improve the contour of jawline by adding volume to the jawline.

Safety Objective:

The safety objectives include the identification and description of adverse events (AEs), and serious adverse events (SAEs) during the course of the study. [redacted]

4 DETERMINATION OF SAMPLE SIZE

[redacted]

The statistical power calculation software, nQuery Advisor Version 7.0, was used for sample size calculations.

5 ANALYSIS SETS

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

Intent-to-Treat (ITT)

The Intent-to-Treat (ITT) population will consist of all randomized subjects. This will be the primary population used for the effectiveness analyses. All effectiveness endpoints will be analyzed as randomized

Safety Evaluation Set (SES)

The SES is the subset of all subjects who were exposed to study medication at least once. All safety endpoints and exposure data will be analyzed as per actual treatment received.

Per Protocol Population (PP)

The Per Protocol (PP) population is a subset of subjects in the ITT population without major protocol deviations. Final determination of what constitutes major or minor protocol deviations will be made prior to database lock. [redacted]

[redacted]

Major protocol deviations will be finalized during the Data Review Meeting that will be carried out prior to database lock.

Populations for Primary and Secondary Effectiveness Endpoint Analyses

The primary and secondary effectiveness endpoints will be summarized using the ITT population, and additionally, for sensitivity purposes, on the PP. [redacted] Additional details of the use of analysis populations are provided below in the Statistical Analysis Methods section. All safety endpoints will be summarized using the SES.

6 ENDPOINTS FOR ANALYSIS

6.1 Effectiveness Endpoints

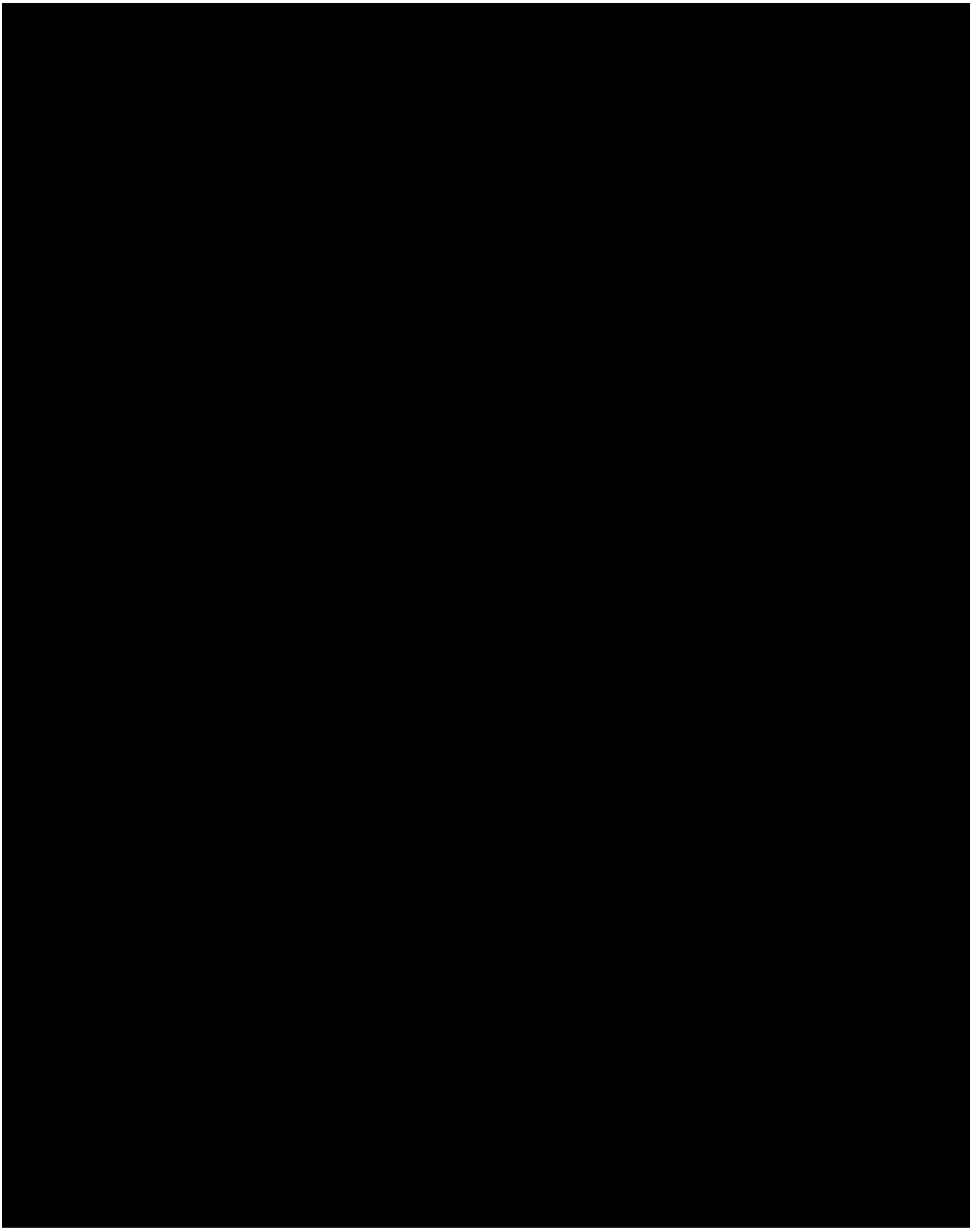
6.1.1 Primary effectiveness endpoint

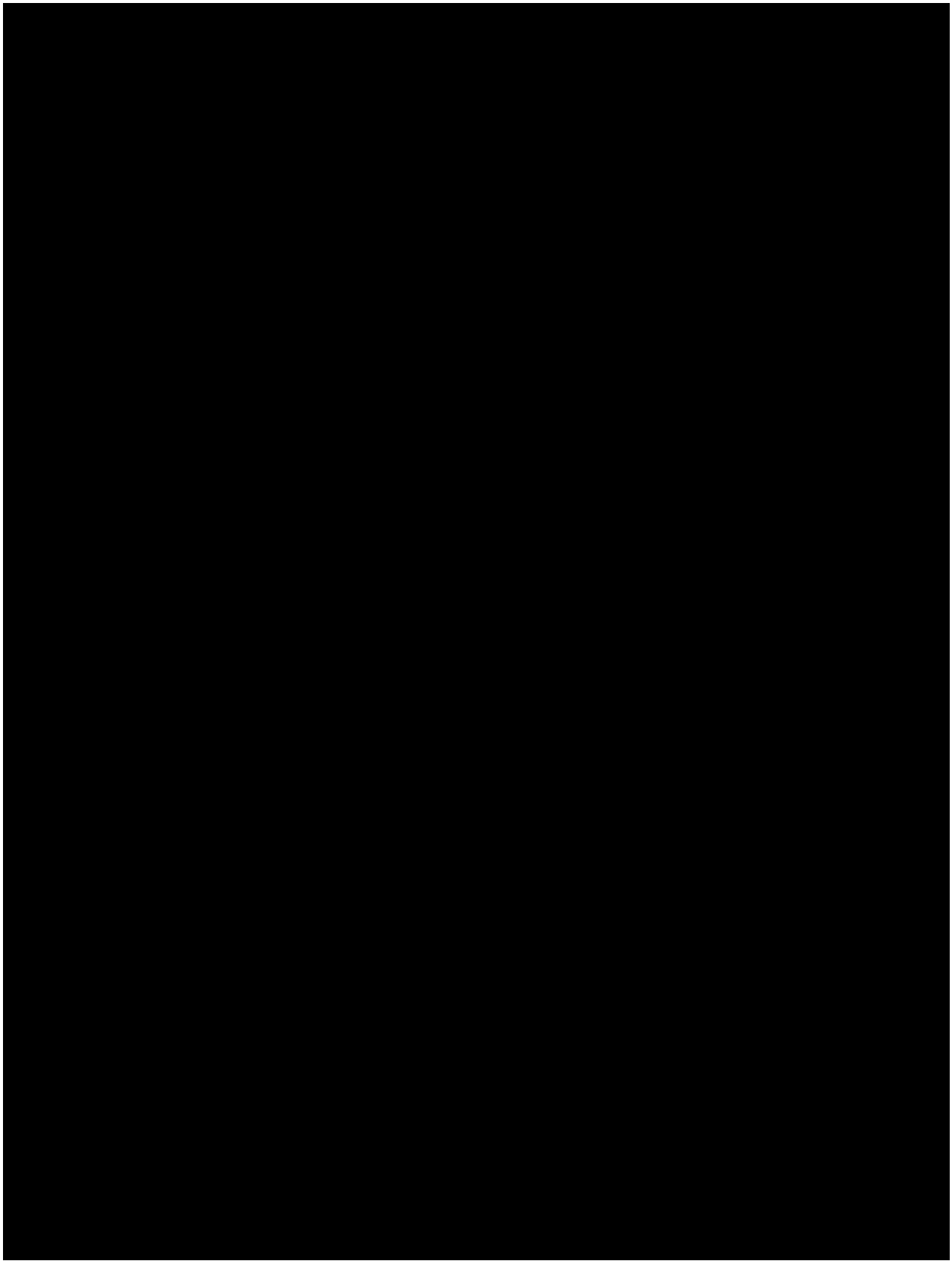
Comparison of the responder rate between the treatment group and the untreated control group at Week 12, according to the Merz Jawline Assessment Scale (MJAS) [redacted]. Treatment response is defined as ≥ 1-point improvement [redacted]. The study device will be clinically effective if at least 50% of treated subjects are responders.

6.1.2 Secondary effectiveness endpoints

- [redacted] FACE-Q satisfaction of the lower face and jawline for treated subjects [redacted] as average percent change from baseline to Week 12. [redacted]
- [redacted] Global Aesthetic Improvement Scale (GAIS) scores for treated subjects at Week 12, as completed by the treating investigator. [redacted]
- [redacted] GAIS scores for treated subjects at Week 12, as completed by the subject. [redacted]
- [redacted] responder rates in the treatment group and the untreated control group at Week 12, according to the MJAS [redacted] Treatment response is defined as ≥ 1-point improvement [redacted].

6.1.3 [redacted]





[redacted]

6.2 Safety Endpoints

6.2.1 Primary safety endpoints

- Incidence and nature of device- and/or injection-related AEs and SAEs observed during the study.

[redacted]

6.2.2 [redacted]

[redacted]

7 STATISTICAL ANALYSIS METHODS

7.1 Effectiveness Endpoints

The primary effectiveness analysis and the analysis of the secondary effectiveness endpoints will be based primarily on the ITT and additionally, for sensitivity purposes, on the PP. [redacted] Statistical tests on the primary effectiveness endpoint will be conducted as a one-sided hypothesis to determine evidence of between-treatment differences.

For summary purposes, continuous variables (values and changes from baseline) will be summarized by n, mean, standard deviation (SD), median, minimum, and maximum. For qualitative variables, absolute and percent frequencies (n, %) and, if applicable, shift tables will be displayed. For all test statistics used for analyzing the effectiveness endpoints, p-values will be used either for inferential or for descriptive purposes, as appropriate.

Change from baseline is defined as post-baseline value minus the corresponding baseline value; percent change from baseline is defined as change from baseline divided by the corresponding baseline value, expressed in percentage.

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'.

7.1.1 Primary Effectiveness Endpoint

The primary effectiveness endpoint, the proportion of responders, subjects with ≥ 1-grade improvement [redacted] on the MJAS from baseline to Week 12 will be summarized as counts and percentages for the treatment and control groups.

Two hypothesis tests will be performed for the primary endpoint, in a sequential order. The first hypothesis is to demonstrate at least 50% of treated subjects are responders and the second hypothesis is to compare the treatment group with the control group. Each hypothesis test is a one-sided test at a significance level of 0.025.

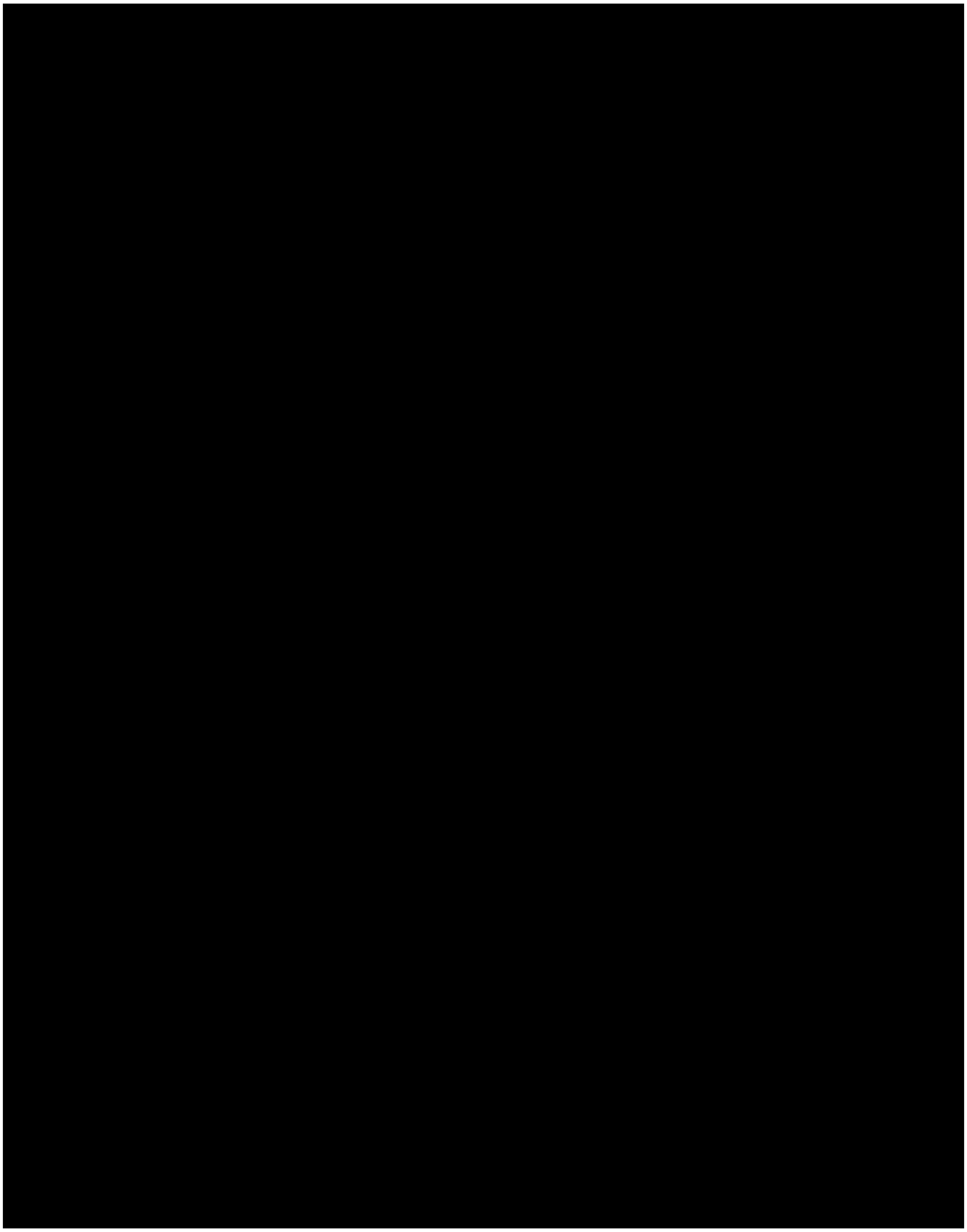
Hypothesis test 1:

$$H_{01}: P_{treatment} \leq 50\% \qquad H_{11}: P_{treatment} > 50\%$$

Hypothesis test 2:

$$H_{02}: P_{treatment} \leq P_{control} \qquad H_{12}: P_{treatment} > P_{control}$$

For primary hypothesis testing of at least 50% of treated subjects are responders (H_{01} versus H_{11}), the binomial test will be utilized. For primary testing of statistical superiority of treatment over control (H_{02} versus H_{12}), the Fisher's exact test will be used.



[redacted]

7.1.2 Secondary Effectiveness Endpoint

The secondary effectiveness endpoints analyses will be conducted using the ITT population using OC, i.e. without imputation of missing values. Additionally, summary of the secondary effectiveness endpoints will also be conducted using the PP population, with OC.

- For the FACE-Q satisfaction with the lower face and jawline assessment, among treated subjects, sum scores and the equivalent Rasch-transformed scores will be summarized using descriptive statistics [redacted].
- The investigator GAIS for treated subjects will be descriptively summarized at Week 12 using counts (n) and percentages (%) for each GAIS category [redacted].
- The subject-rated GAIS for treated subjects will be descriptively summarized at Week 12 using counts (n) and percentages (%) for each GAIS category [redacted].

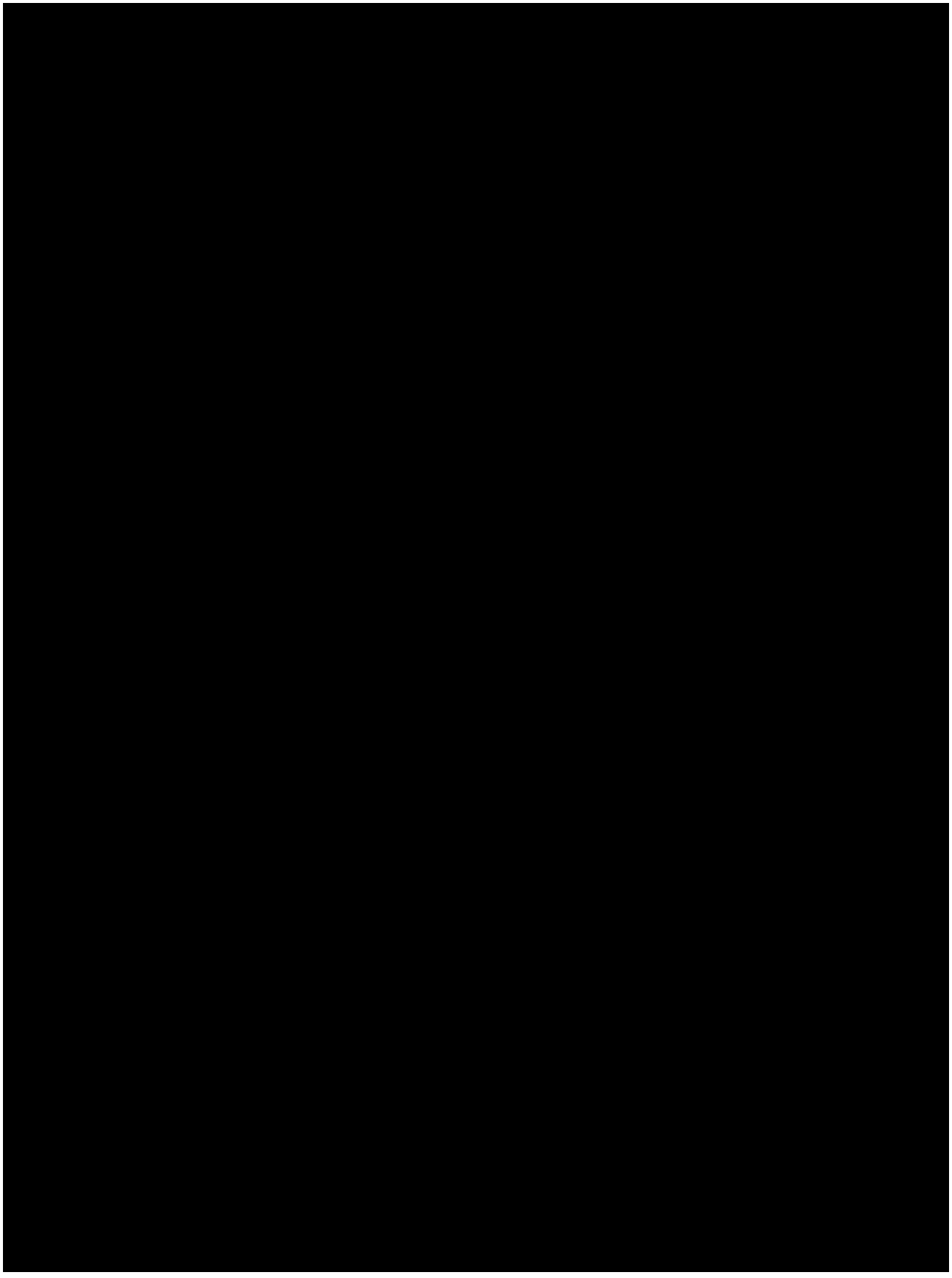
[redacted]

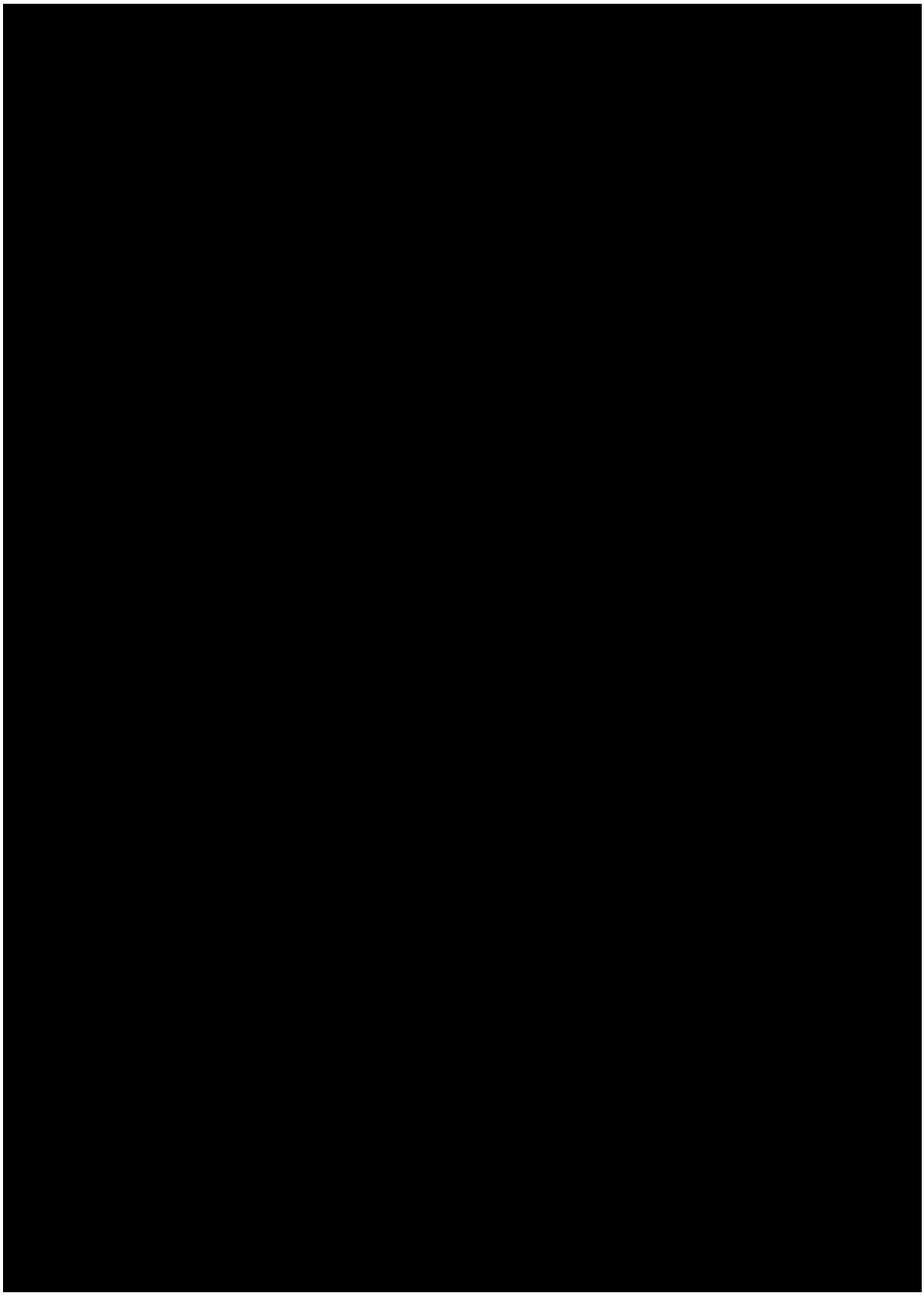
- The proportion of responders in the treatment group and the untreated control group at Week 12, according to the MJAS, [redacted]

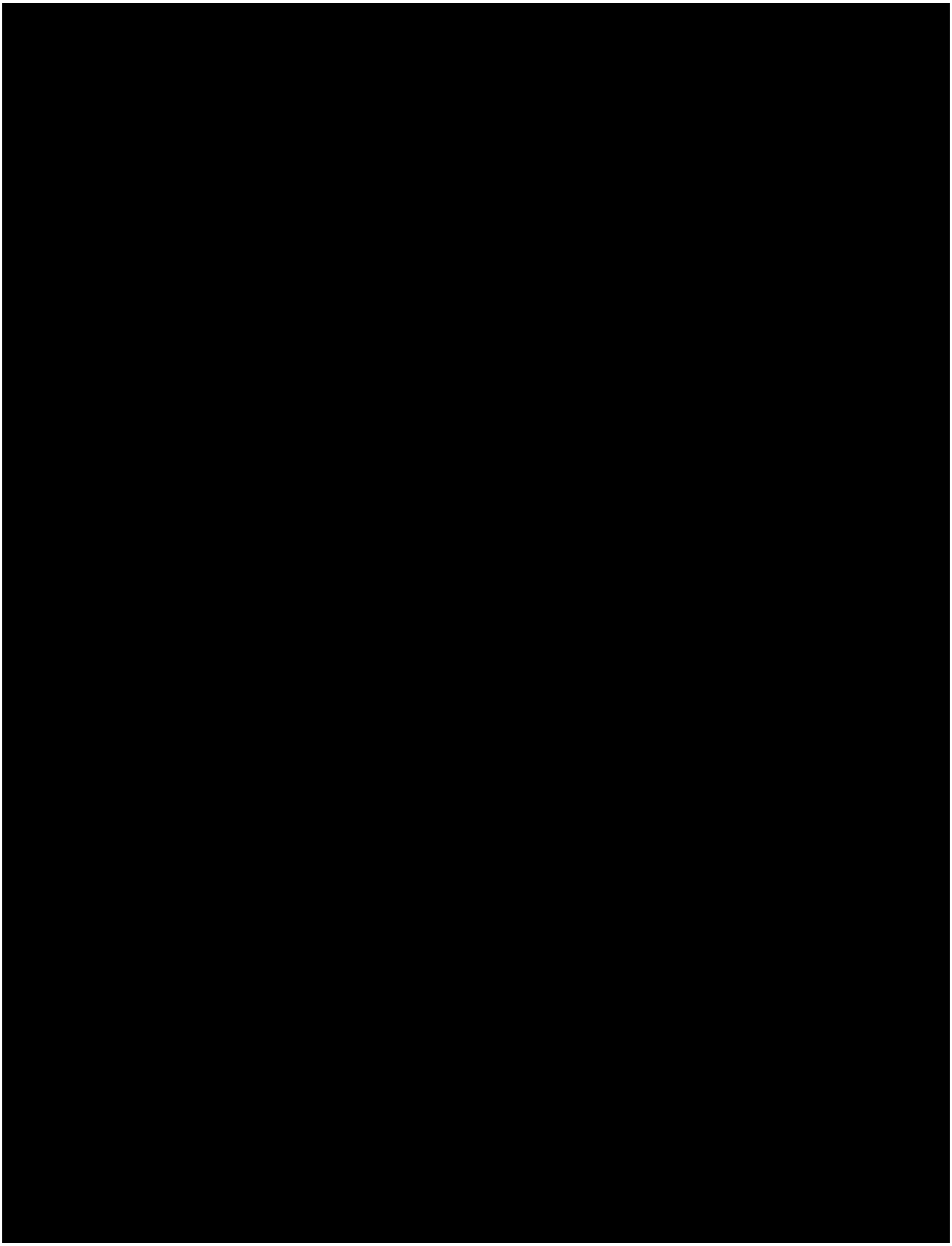
[redacted] will be summarized using counts and percentages. Treatment response is defined as ≥ 1-point improvement [redacted]

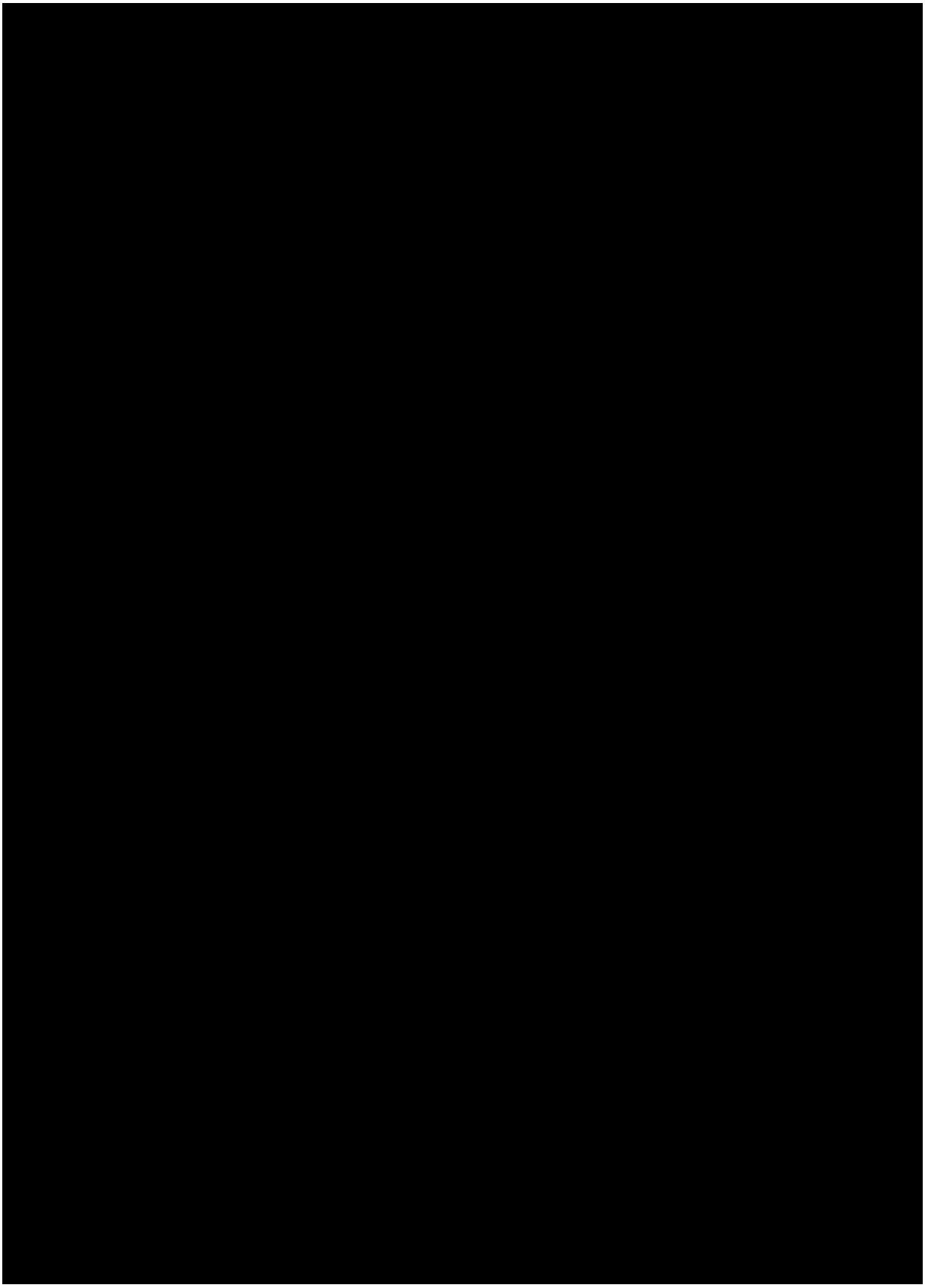
7.1.3 [redacted]

[redacted]









[redacted]

7.2 Safety Endpoints

All safety analyses will be performed on the SES. AEs will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA). The version of MedDRA used will be documented in the data management plan (DMP).

Only treatment emergent adverse events (TEAEs) will be analyzed. [redacted]

[redacted]

[redacted]

An overall summary of AEs will be provided by treatment group and overall, for the following:

- Any AEs
- Any non-TEAEs
- Any TEAEs
- Any treatment-related TEAEs [redacted]
- Any serious TEAEs
- Any TEAEs leading to discontinuation

[redacted]

[redacted]

Listings for all AEs as well as subsets such as TEAEs leading to discontinuation, serious TEAEs, and deaths will be provided.

[redacted]

[redacted]

7.3 Specificity for Other Summaries of Data and Variables

7.3.1 Subject Disposition

Subject disposition will be presented using standard descriptive statistics (i.e., absolute and percent frequencies (n, %)).

The number of subjects who are screened, randomized, attend each visit, and complete the study will be presented and summarized by treatment group and overall. Additionally, the number of subjects attending each visit, including optional visits and telephone contacts (i.e., 72 post-treatment safety phone calls), and reasons for early withdrawal from the study, where applicable, will be listed and summarized by treatment group and overall.

7.3.2 Subject Populations

The number of subjects in each of the ITT, SES, and PP dataset populations will be listed and summarized by treatment group and overall. All randomized subjects will be included in ITT population. Subjects will be excluded from SES only if they were not exposed. The reasons for exclusion from PP population and major protocol violations will be listed and summarized by treatment group and overall. Minor protocol violations will be listed only. [redacted]

7.3.3 Medical History

Medical history will be categorized according to the start and stop dates of the condition as:

- Conditions/diseases started and stopped prior to the study start (baseline visit) will be regarded as Medical History.

- Conditions/diseases prior to study start (baseline visit) and stopped or ongoing during the study will be treated as Concomitant Diseases.

Medical history and concomitant diseases will be summarized based on MedDRA system organ class and preferred term levels for the ITT population. Where there is insufficient information available to confirm whether a disease/condition should be classified as previous or concomitant, it will be treated as previous (i.e., Medical History).

7.3.4 Previous Medication and Therapy

Previous medications/therapies will be categorized according to the start and stop dates of the medication as:

- Medication/therapies started and stopped prior to the baseline visit will be treated as Previous Medication.
- Medication/therapies started prior to the baseline visit and stopped or ongoing during the study period will be treated as Concomitant Medications.
- Medications/therapies started after the baseline visit will be treated as Concomitant Medication.

Frequencies of concomitant medications will be given based on different ATC code levels as well as by WHO preferred name for the ITT population. Indications for concomitant procedures will be coded and only listed. Where there is insufficient data available to confirm whether a medication/therapy should be classified as previous or concomitant, it will be treated as previous.

7.3.5 Demographic Data

Demographic characteristics [redacted] and baseline characteristics [redacted] will be presented using standard descriptive statistics for continuous variables (n, mean, SD, median, minimum, and maximum) and for qualitative variables (absolute and percent frequencies (n, %)). Demographic and baseline characteristics will be summarized for the ITT, SES and the PP populations.

The body mass index will be calculated as follows:

BMI [kg/m^2] = body weight[kg] / (height[cm]/100)^2

Post-baseline weight will only be listed.

7.3.6 Exposure Data

Subjects' exposure to study product will be summarized for the SES (as exposed to product if different from as randomized for any reason) according to the following:

(1) Frequencies of subjects receiving initial treatment and touch-up will be summarized by treatment group [redacted]

(2) Frequency of subjects receiving the optional re-treatment (Week 48) will also be summarized for the baseline treatment group;

[redacted]

[redacted]

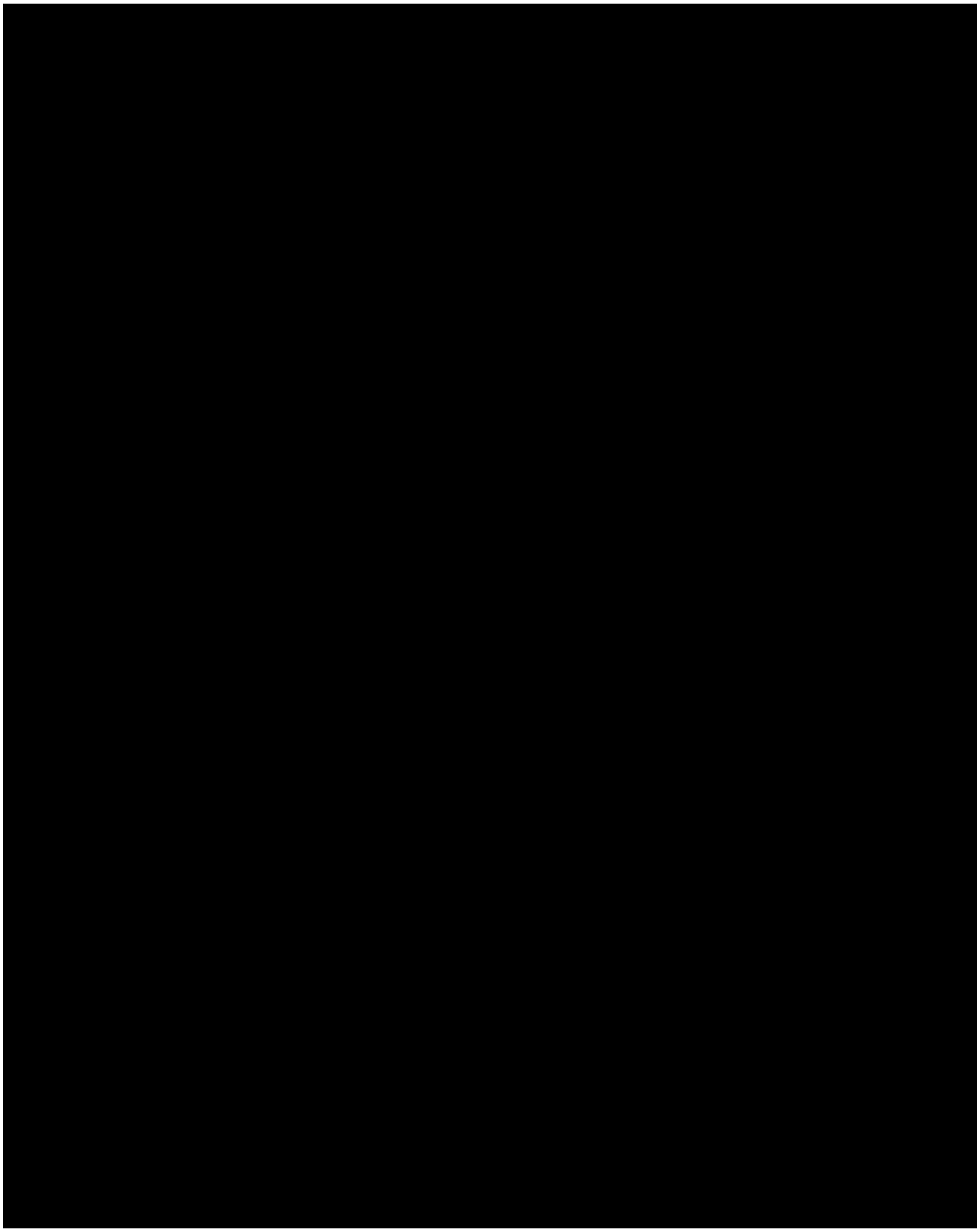
Frequencies of pre- and post-treatment skin applications will be given based on different ATC code levels.

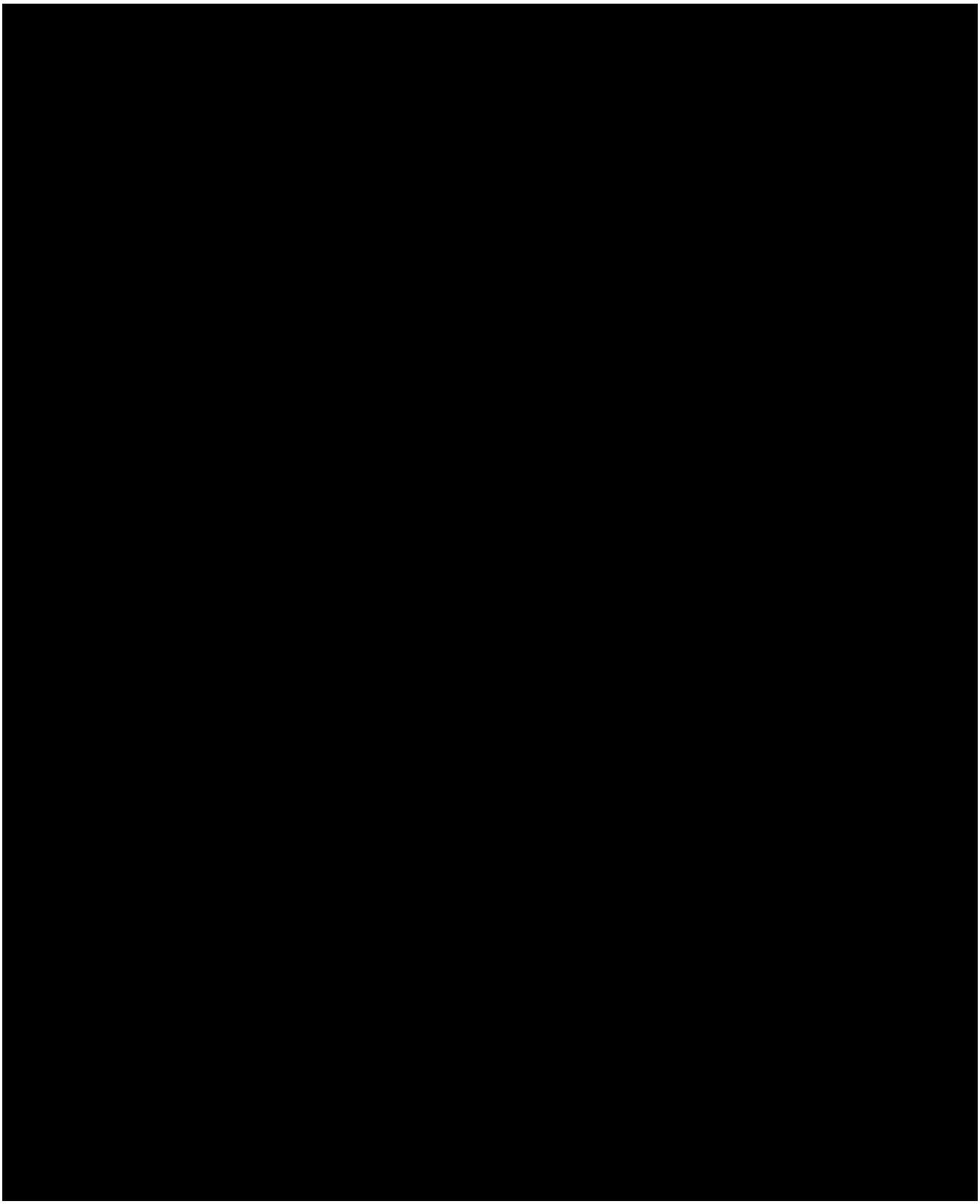
Data regarding device malfunction/technical complaint associated with the treatment injections will be listed only.

7.4 Special Statistical/Analytical Issues

7.4.1 [redacted]

[redacted]





7.4.2 Interim Analyses

No formal interim analysis is planned. [redacted] The annual reports generally may require enrollment updates (i.e., number of subjects screened and enrolled (overall and by site), numbers randomized, etc.), subject disposition, and safety listing.

7.4.3 [redacted]

[redacted]

8 [redacted]

[redacted]

9 REFERENCES

[1] The American Society for Aesthetic Plastic Surgery. 2016 Cosmetic Surgery National Data Bank Statistics. Available at: https://www.surgery.org/sites/default/files/ASAPS-Stats2016.pdf [accessed April 4, 2017].

[2] Carpenter, J. R. and Kenward, M. G. (2013), Multiple Imputation and Its Application, New York: John Wiley &Sons.

[3] Ratitch, B., Lipkovich, I., O'Kelly, M. (2013) Combining Analysis Results from Multiply Imputed Categorical Data. Available at https://www.pharmasug.org /proceedings/2013/SP/PharmaSUG-2013-SP03.pdf [accessed June 1, 2020].

[4] Newcombe, R. (2012) Confidence Intervals for Proportions and Related Measures of Effect Size, CRC Press.

10 APPENDICES