Statistical Analysis Plan for Interventional Studies

Sponsor Name: Revance Therapeutics, Inc.

Protocol Number: 1820203
Protocol Title: A Phase 2, Randomized, Double-Blind, Placebo-Controlled, Parallel Group, Dose-Ranging, Multi-Center Trial to Evaluate the Efficacy and Safety of DaxibotulinumtoxinA for Injection for the Treatment of Upper Limb Spasticity in Adults After Stroke or Traumatic Brain Injury (JUNIPER)

Protocol Version and Date:

Amendment 3	**29-June-2020**
Amendment 2	**04-November-2019**
Amendment 1	**29-November-2018**
Original Protocol v3.0	**14-September-2018**

[redacted]

Notice of Confidential and Proprietary Information:

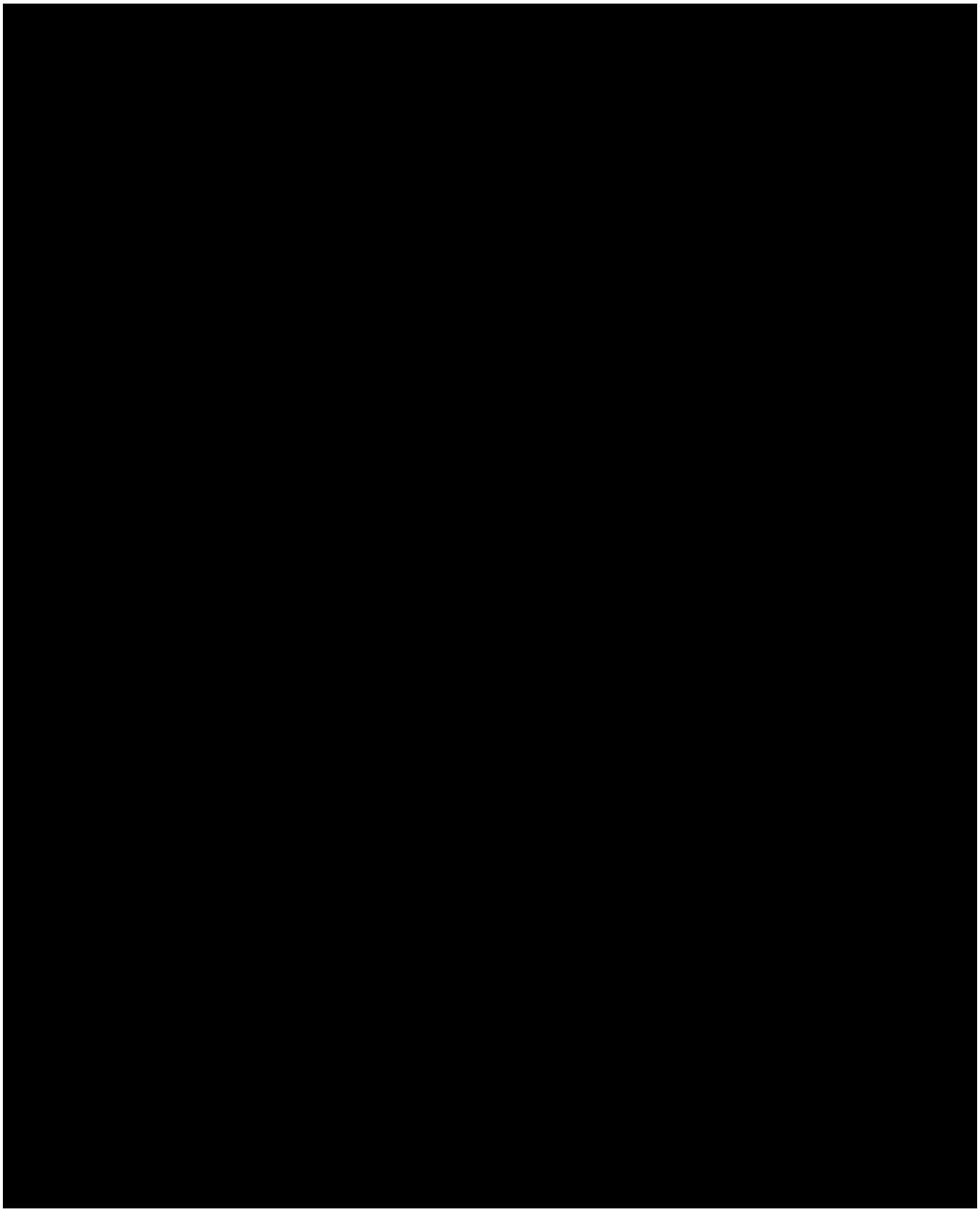


Table of Contents

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

1. Glossary of Abbreviations

Abbreviation	Description
AE	Adverse Event
ANCOVA	Analysis of Covariance
AP	Aggregate Pattern
AROM	Active Range of Motion
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass index
BoNT	botulinum neurotoxin
BoNTA	botulinum neurotoxin type A
CI	Confidence Interval
cm	Centimeters
CMH	Cochran-Mantel-Haenszel
C-SSRS	Columbia-Suicide Severity Rating Scale
DAS	Disability Assessment Scale
DAXI	DaxibotulinumtoxinA
DMC	Data Monitoring Committee
ECG	Electrocardiogram
EQ-5D-5L	EuroQol Five Dimension Scale
EOS	End of Study
FDA	Food and Drug Administration
FEV_1	First Second of Exhalation
FMUE	Fugl-Meyer Upper Extremity
FVC	Forced Vital Capacity
ICH	International Conference on Harmonization
INR	International Normalized Ratio
IXRS	Interactive Web Randomization System
kg	Kilograms
LS	Least Squares
m	Meters
MAR	Missing at Random
MAS	Modified Ashworth Scale

Abbreviation	Description
MedDRA	Medical Dictionary for Regulatory Activities
mL	milliliters
MMRM	Mixed Model Repeated Measures
P/CGIC	Patient/Caregiver Global Impression of Change
PGIC	Physician Global Impression of Change
PP	Per Protocol
PTT	Principal Treatment Target
PT	Preferred Term
QC	Quality Control
QoL	Quality of Life
QQ	Quantile-Quantile
QTcF	Corrected QT Interval using Fridericia's correction formula
ROM	Range of Motion
SAP	Statistical Analysis Plan
SD	Standard Deviation
SF-36	Short Form-36 Survey
SI	Système International
SMG	Suprahypertonic Muscle Group
SOC	System Organ Class
SOP	Standard Operating Procedure
TBI	Traumatic Brain Injury
TEAE	Treatment Emergent Adverse Event
TLF	Table, Listing and Figure
TS	Tardieu Scale
ULS	Upper Limb Spasticity
VAS	Visual Analog Scale
WHO	World Health Organization
WHODD	World Health Organization Drug Dictionary
WOCBP	Women of Child Bearing Potential

2. Purpose

The purpose of this statistical analysis plan (SAP) is to ensure that the data listings, summary tables and figures that will be produced, and the statistical methodologies that will be used, are complete and appropriate to allow valid conclusions regarding the study objectives.

2.1. Responsibilities

Syneos Health will perform the statistical analyses and are responsible for the production and quality control of all tables, figures and listings.

2.2. Timings of Analyses

An independent Data Monitoring Committee (DMC) will review descriptive summaries of accumulating safety data during the study and will monitor quality and completeness of the safety data, as well as signals and outcomes of the safety data. Further description of the DMC analyses and timing of the analyses, including a sentinel safety cohort, can be found in the DMC charter. An unblinded team from Syneos Health within Biostatistics will perform the analyses as to maintain the blinding of the study. Those who are performing the unblinded analysis will not be involved in the study outside of the DMC analyses or generating randomization code.

The complete analysis of safety and efficacy is planned after all subjects complete the final study visit or terminate early from the study, and the clinical study database is locked.

3. Study Objectives

3.1. Primary Objective

The primary objective of this study is to compare the safety and efficacy of a single treatment of DaxibotulinumtoxinA (DAXI) for injection at three dose levels (250 U, 375 U, 500 U) versus placebo in reducing muscle tone of adult subjects with upper limb spasticity (ULS) due to stroke or traumatic brain injury (TBI).

[redacted]

3.4. Brief Description

This is a Phase 2, randomized, double-blind, placebo-controlled, parallel group, dose ranging, multi-center trial to evaluate the efficacy and safety of DAXI for injection in the treatment of ULS in adults after stroke or TBI. The trial will be conducted at approximately 30 sites across the United States.

Approximately 128 subjects with ULS will be randomly assigned [redacted]

[redacted]

The co-primary efficacy endpoints are the mean change from baseline in muscle tone measured using the Modified Ashworth Scale (MAS) in the suprahypertonic muscle group (SMG) of the elbow, wrist, or finger flexors at Week 6 and the mean score on the Physician Global Impression of Change (PGIC) at Week 6.

- For subjects enrolled before COVID-19:
    - The end of study visit for a given subject can occur at Week 12 if the change in MAS score in the SMG is < 1-point reduction from baseline AND the PGIC score is ≤ 0, OR
    - The end of study visit will occur after Week 12 when there is loss of muscle tone improvement in the SMG (i.e., a reduction from baseline in MAS score of < 1-point) AND the PGIC score is ≤ 0. OR
    - Week 36 will be the end of study for all subjects, even those who maintain muscle tone improvement in the SMG (i.e., a reduction from baseline in MAS score of ≥ 1 point) OR PGIC score is ≥ 1.
- For subjects enrolled after the COVID-19 20 March 2020 screening and randomization hold the end of study will be at Week 12.

3.5. Subject Selection

3.5.1. Inclusion Criteria

All participants must meet the following inclusion criteria:

1. Adults, 18 – 75 years of age.
2. Written informed consent including authorization to release health information.
3. Focal ULS after a stroke (as defined by World Health Organization [WHO] criteria) or TBI, last stroke or TBI > 24 weeks prior to Screening.
4. Upper limb spasticity with the primary aggregate pattern (AP): flexed elbow + flexed wrist + clenched fist (flexed fingers) and ≥ 1 other clinical pattern: adducted and internally rotated shoulder, pronated forearm, thumb-in-palm and/or intrinsic plus hand.
5. Moderate to severe ULS with a MAS score ≥ 2 at the elbow, wrist, and finger flexors; with the exception that the MAS score is ≥ 3 at the SMG for subjects with previous injections of botulinum neurotoxin type A (BoNTA) in the paretic limb.
6. Moderate to severe functional disability (Disability Assessment Scale [DAS] score ≥ 2) on the principal target of treatment (1 of 4 functional domains: hygiene, dressing, malposition of the arm/wrist/fingers, and pain).
7. Has sufficient cognitive and communication ability to be able to give informed consent including authorization to release health information.

3.5.2. Exclusion Criteria

Subjects will not be enrolled if they meet any of the following exclusion criteria:

1. ULS attributable to an etiology other than stroke or TBI.

2. Bilateral upper limb paresis or quadriplegia.
3. Initiated in physiotherapy of the upper extremities ≤ 30 days prior to Screening or planned to start physiotherapy of the upper extremities during the course of the study.
4. Previous treatment with BoNT product for any condition ≤ 16 weeks prior to Screening.
5. Botulinum neurotoxin treatment-experienced subjects who have historically required < 200 U of Botox/Xeomin or its equivalent to effectively treat the ULS.
6. Botulinum neurotoxin treatment-experienced subjects who had suboptimal or no treatment response to the most recent BoNTA injection for spasticity, as determined by the investigator; or history of primary or secondary non-response to BoNTA injections, known to have neutralizing antibodies to BoNTA; or have a history of botulinum toxin type B (rimabotulinumtoxinB [Myobloc/Neurobloc]) injection for spasticity due to non-response or suboptimal response to BoNTA.
7. Change in oral medications for spasticity including dosage and dosing frequency ≤ 30 days prior to Screening.
8. Previous or planned treatment of the spastic upper limb with phenol, alcohol injection or surgery.
9. Profound muscular atrophy or fixed contracture (spasticity angle, per the Tardieu Scale, < 10 degrees in the most hypertonic muscle group [elbow, wrist or finger flexors]) of the spastic limb leading to marked limitation on passive range of motion or any other known conditions of the upper limb that could confound muscle tone or functional assessment.
10. Prior treatment with intrathecal baclofen.

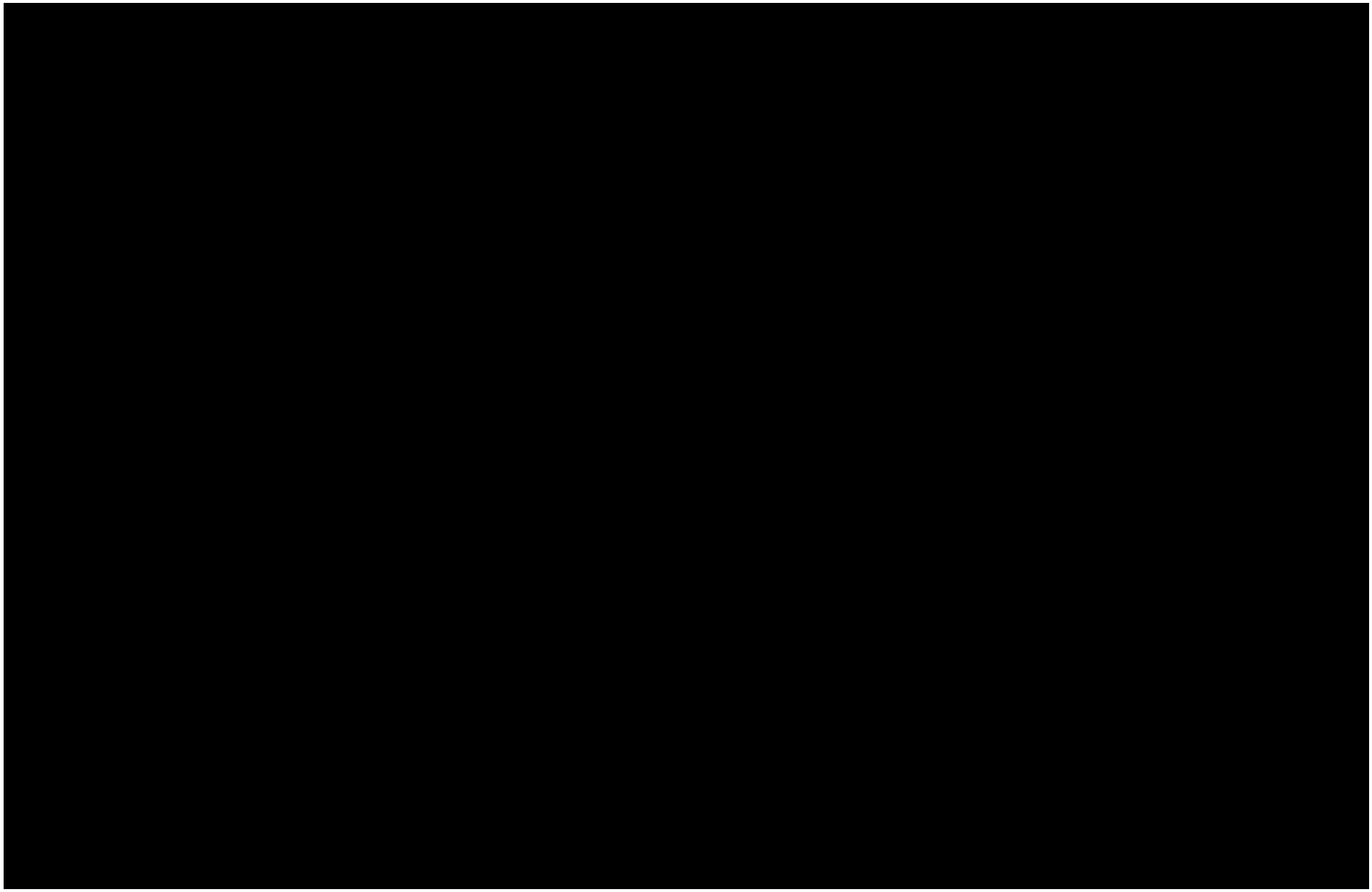
[redacted]

3.6. Determination of Sample Size

The sample size calculation is based on the effect size that was observed for the primary endpoint (change from baseline in MAS in the SMG) from the registration study of Dysport (Gracies J, 2015) for the treatment of ULS. With a mean change from baseline in MAS in the SMG difference between an active treatment and the placebo arm of 0.9 to 1.1 and a common standard deviation of 1.1, the minimum effect size (Cohen's d) targeted is 0.8 to 1.0.

[redacted]

[redacted]

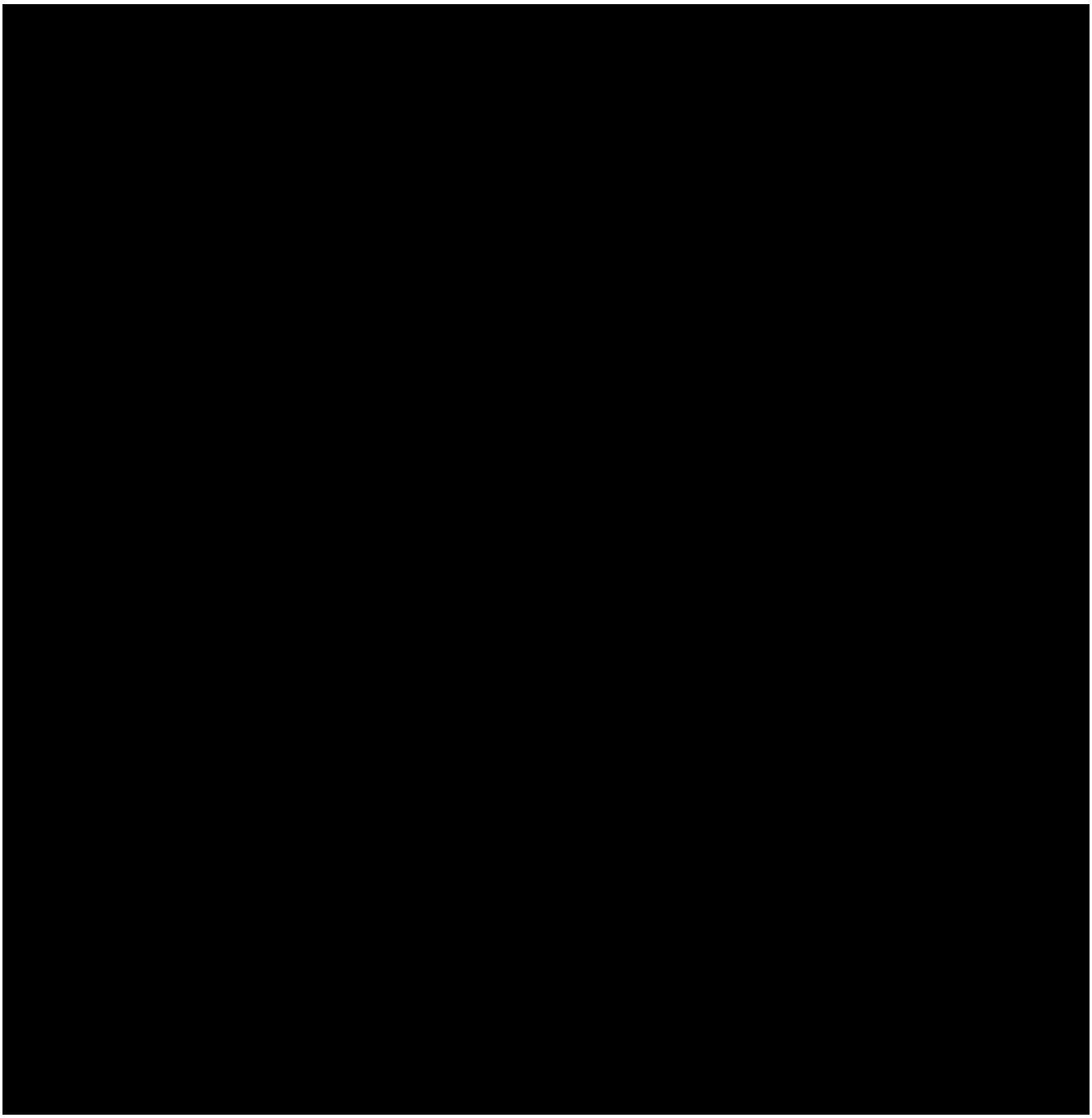


4. Endpoints

4.1. Co-Primary Endpoints

The co-primary endpoints of this study are:

- Mean change from baseline in muscle tone measure with the MAS in the SMG of the elbow, wrist, OR finger flexors at Week 6
- Mean score of the PGIC at Week 6



[redacted]

5. Analysis Populations

5.1. Screened Population

The Screened Population will include all subjects screened for entry into the study. This population will be used for disposition summaries and listings.

5.2. Randomized Population

The Randomized Population will include all subjects randomized into the study. This population will be used for protocol deviation summaries and all listings.

5.3. Safety Population

The safety population is defined as all randomized subjects who received Investigational product. The safety population will be analyzed by treatment arm as treated (i.e., based on the treatment the subject actually received rather than the treatment to which the subject was randomized). The safety population will be used to for all safety analyses.

5.4. Intent-to-Treat Population

The Intent-to-Treat (ITT) population is defined as all subjects who were randomized, received study injections in this study, and have a MAS score in the SMG at baseline and at least one post-baseline assessment. This population will be used for all efficacy analyses. The ITT population will be classified by treatment arm as randomized (i.e., treatment arm will be based on randomization assignment).

5.6. Pre-COVID-19 ITT Population

The Pre-COVID-19 ITT population will be a subset of the ITT population and includes subjects who were enrolled before the COVID-19 enrollment hold on 20-March-2020. This population will be classified by treatment arm as randomized.

5.7. Pre-COVID-19 PP Population

The Pre-COVID-19 PP population will be a subset of the PP population and includes subjects who were enrolled before the COVID-19 enrollment hold on 20-March-2020. This population will be classified by treatment arm as randomized.

5.8. Protocol Deviations

Protocol deviations will be summarized for all randomized subjects by treatment arm and total. A listing will be provided for all protocol deviations. A separate summary table and listing will be provided for any protocol

deviations related to COVID-19. The summary of Covid-19 related PDs will include number of subjects impacted, as well as missed visits, missed assessments, remote visits, and remote assessments.

6. General Aspects for Statistical Analysis

6.1. General Methods

- Unless otherwise specified, summaries will be presented for each treatment group.
- Continuous variables will be summarized using the number of observations (n), mean, standard deviation (SD), median, minimum, and maximum. Categorical variables will be summarized using number of observations (n), frequency and percentages of subjects.
- All relevant subject data will be included in listings. All subjects entered into the database will be included in subject data listings.
- Unless otherwise specified, if a subject has more than one observation at a given time point, the value closest to the target day for that visit will be used. If two observations are equal distance from the target day, the later observation will be used.
- All analyses and outputs will be generated using SAS® or SAS/IML® version 9.4 or higher.

6.2. Key Definitions

Study day will be assigned as follows:

- Before the study medication dose date, study day = visit date – dose date.
- After the study medication dose date, study day = visit date – dose date + 1.

The day the study medication is given will be Day 1, and the day before the study medication is given will be considered Day -1.

Unless otherwise specified, baseline is defined as the last non-missing value before the dose of study medication. Change from baseline is defined as the post-baseline value minus the baseline value (post-baseline – baseline). Percent change from baseline is defined as 100 times (post-baseline value minus baseline value) divided by the baseline value (100*(post-baseline-baseline)/baseline).

[redacted]

6.4. Visit Windows and Remapping of End of Study Visit

Analyses will be based on the visit designation. Visits are required to be within specified time windows relative to the Day 1 or Baseline visit. [redacted]

6.5. Pooling of Centers

Treatment centers with less than 4 subjects will be considered for pooling. Pooling will be based on geographic proximity. No pooled center will contain more than 20 randomized subjects. Final pooling of centers will be determined prior to database lock.

[redacted]

7. Demographic, Other Baseline Characteristics and Medication

7.1. Subject Disposition and Withdrawals

The number and percentage of subjects screened, randomized, completed, and discontinued will be summarized along with the reasons for discontinuation. The number of subjects in each study population will also be presented. Listings will be provided for disposition data, study populations and reasons for exclusion from the study population, and inclusion/exclusion criteria not met.

A summary of the duration of the subject participation in the study will be produced, including the n, mean, SD, median, minimum, and maximum duration in weeks, as well as the number and percentage of subjects in the following categories of duration: <4 weeks, 4 to < 6 weeks, 6 to <12 weeks, 12 to <24 weeks, and 24 to 36 weeks. If a subject is lost to follow-up, their participation in the study will be calculated based on their date of last visit.

7.2. Demographic and Other Baseline Characteristics

Demographic and baseline characteristics will be summarized for the ITT, Pre-COVID-19 ITT, PP, Pre-COVID-19 PP, and safety populations. [redacted]

7.3. Medical History and Concomitant Diseases

Medical history and concomitant diseases will be coded according to Medical Dictionary for Regulatory Activities (MedDRA) version 22.0 or higher. A summary of medical history and concomitant diseases for the safety population by system organ class and preferred term. Terms in the table will be sorted alphabetically. A listing will also be provided that includes the system organ class, preferred term, and verbatim term for each medical history and concomitant disease. The listing will be sorted by treatment group, subject, system organ class, preferred term, and verbatim term.

7.4. Other Baseline Characteristics

Summaries will also be provided for the ITT, Pre-COVID-19 ITT, PP, Pre-COVID-19 PP, and safety populations for ULS history. [redacted]

7.5. Medication

Prior and concomitant medication will be coded using World Health Organization Drug Dictionary (WHODD) version March 2018 or higher. Anatomic Therapeutic Chemical (ATC) level 2 and preferred names will be included in all summaries and listings. Listings will also include verbatim terms. Summary tables will be presented in alphabetical order, and listings will be presented for each subject by earliest start date and then alphabetical order. The listing will indicate if the medication was prior or concomitant. Prior medications for ULS will be summarized and listed separately. Summaries will be provided for the safety population.

7.5.1. Prior Medication

Medications that were stopped before the study medication dose date will be classified as prior. A summary table will be provided for prior medications.

7.5.2. Concomitant Medication

Medications that were taken on or after the study medication dose date will be classified as concomitant. This classification includes medications that were started after the study medication dose date or medications that began before and continued to be taken after the study medication dose date.

8. Efficacy

All efficacy analyses will be performed for the ITT and PP populations, with the ITT population serving as the primary analysis. Efficacy analyses with PP populations will be supportive in nature. Sensitivity analyses will also be performed on the primary and key secondary efficacy endpoints, analyses of duration of effect, and selected exploratory efficacy endpoints using the Pre-COVID-19 ITT and Pre-COVID-19 PP populations. All statistical tests will be 2-sided at the 0.05 significance level, unless otherwise indicated. All efficacy endpoints will be summarized with descriptive statistics by treatment arm and visit.

8.1. Co-Primary Efficacy Endpoints and Analyses

The co-primary efficacy endpoints are:

- Mean change from baseline in muscle tone measured using the MAS in the SMG of the elbow, wrist, OR finger flexors at Week 6
- Mean score on the Physician Global Impression of Change (PGIC) at Week 6.

The MAS measures resistance during passive soft-tissue stretching. It is a quick and easy measure that can help assess the efficacy of treatment. The following conventions prevail:

- The MAS is performed in the supine position (this will garner the most accurate and the lowest score as tension anywhere in the body will increase spasticity)
- Because spasticity is “velocity dependent” (the faster the limb is moved, the more spasticity is encountered), the MAS is performed while moving the limb at the “speed of gravity”; this is defined as the same speed at which a non-spastic limb would naturally drop (fairly fast)
- The test is performed a maximum of three times for each joint; if more than three times, the short-term effect of a stretch can influence the score
- The MAS is performed prior to goniometric testing; goniometric testing provides a stretch, and the short-term effect of a stretch can influence the score

Original MAS Scoring

0 = Normal tone, no increase in tone
1 = Slight increase in muscle tone, manifested by a catch and release or minimal resistance at the end of the range of motion (ROM) when the affected part(s) is moved in flexion or extension
1+ = Slight increase in muscle tone, manifested by a catch, followed by minimal resistance throughout the remainder (less than half) of the ROM
2 = More marked increase in muscle tone through most of the ROM, but affected part(s) easily moved
3 = Considerable increase in muscle tone, passive movement difficult
4 = Affected part(s) rigid in flexion or extension

Because the original MAS score is a categorical variable, the re-coding will be applied to be able to treat it as a continuous variable in the mean change from baseline statistical analysis. Recoding is listed in Table 4. All analyses will use the recoded MAS scores.

Table 4: MAS Re-coding

Original MAS score	Derived MAS score
0	0
1	1
1+	2
2	3
3	4
4	5

The PGIC is a questionnaire that captures the physician's overall impression of the subject's response to study treatment. The physician's selected response maps to a 9-point scale: -4 (Markedly worse), 0 (about the same), to +4 (Markedly improved). The PGIC scale is presented in Table 5 below.

Table 5: PGIC Questionnaire

Please choose the response that best describes the overall change in your subject's upper limb spasticity since the start of the study medication	
☐-4	Markedly worse
☐-3	Very much worse
☐-2	Moderately worse
☐-1	A little worse
☐0	About the same
☐+1	A little improved
☐+2	Moderately improved
☐+3	Very much improved

☐+4	Markedly improved

8.1.1. Primary Analyses of the Co-Primary Endpoints

The primary analysis will be based on the ITT population.

The mean change from baseline in muscle tone measured using the MAS in the SMG at Week 6 between treatment arms will be compared using analysis of covariance (ANCOVA) model with baseline MAS score as covariate and terms for treatment, and the two randomization stratification factors (pooled study center, history of prior treatment with BoNT products). The 3 pairwise comparisons considered for this co-primary endpoint are:

- H111: DAXI 500 U versus placebo at week 6
- H211: DAXI 375 U versus placebo at week 6
- H311: DAXI 250 U versus placebo at week 6

The mean score on the co-primary endpoint PGIC at Week 6 will be compared between treatment arms using the analysis of covariance (ANCOVA) model with terms for treatment, and the two randomization stratification factors (pooled study center, history of prior treatment with BoNT products). The 3 pairwise comparisons considered for this co-primary endpoint are:

- H121: DAXI 500 U versus placebo at week 6
- H221: DAXI 375 U versus placebo at week 6
- H321: DAXI 250 U versus placebo at week 6

[redacted]

8.1.2. Sensitivity Analysis of the Co-Primary Endpoints

The primary analyses will be repeated using the PP population, the Pre-COVID-19 ITT population, and the Pre-COVID-19 PP population as supportive analyses.

The sensitivity of the primary endpoint results to missing data will be evaluated in the following additional analyses, a permutation test, and a tipping point analysis. A permutation test will be done that will fit 1000 datasets using the same ANCOVA model as the primary analysis, but with randomly assigned

pseudo-treatment group designations (Stedman, 2009). The empirical p-value will be obtained from the permutation test. This non-parametric based permutation test provides a conservative way to assess the co-primary efficacy endpoints.

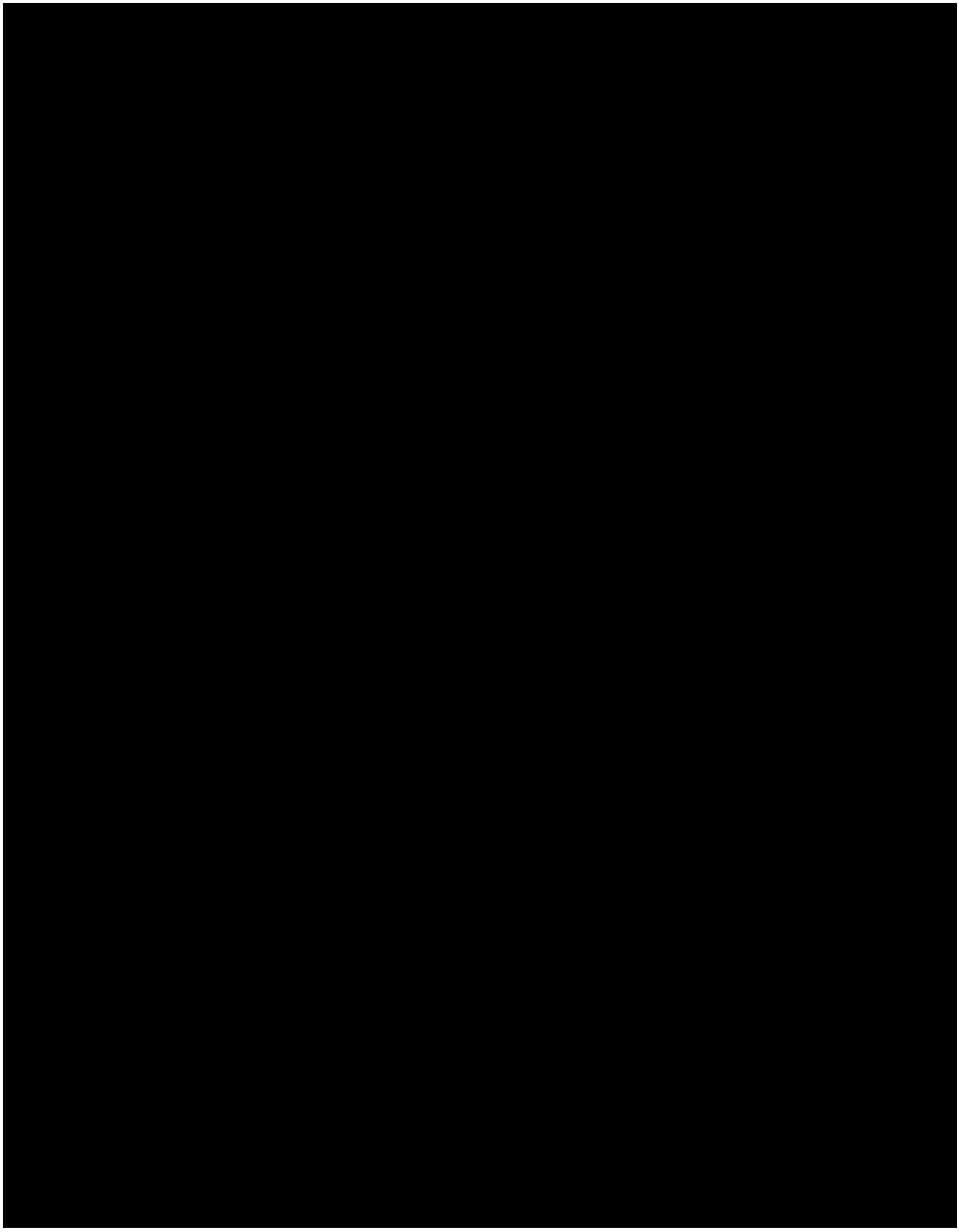
A sensitivity analysis with a tipping point analysis approach will be performed to assess how severe departures from the assumption of data missing at random (MAR) must be to overturn conclusions from the primary analysis. This analysis will be performed for both co-primary efficacy endpoints. Implementing the tipping point approach will include the following steps with the first 3 steps being the standard multiple imputation steps:

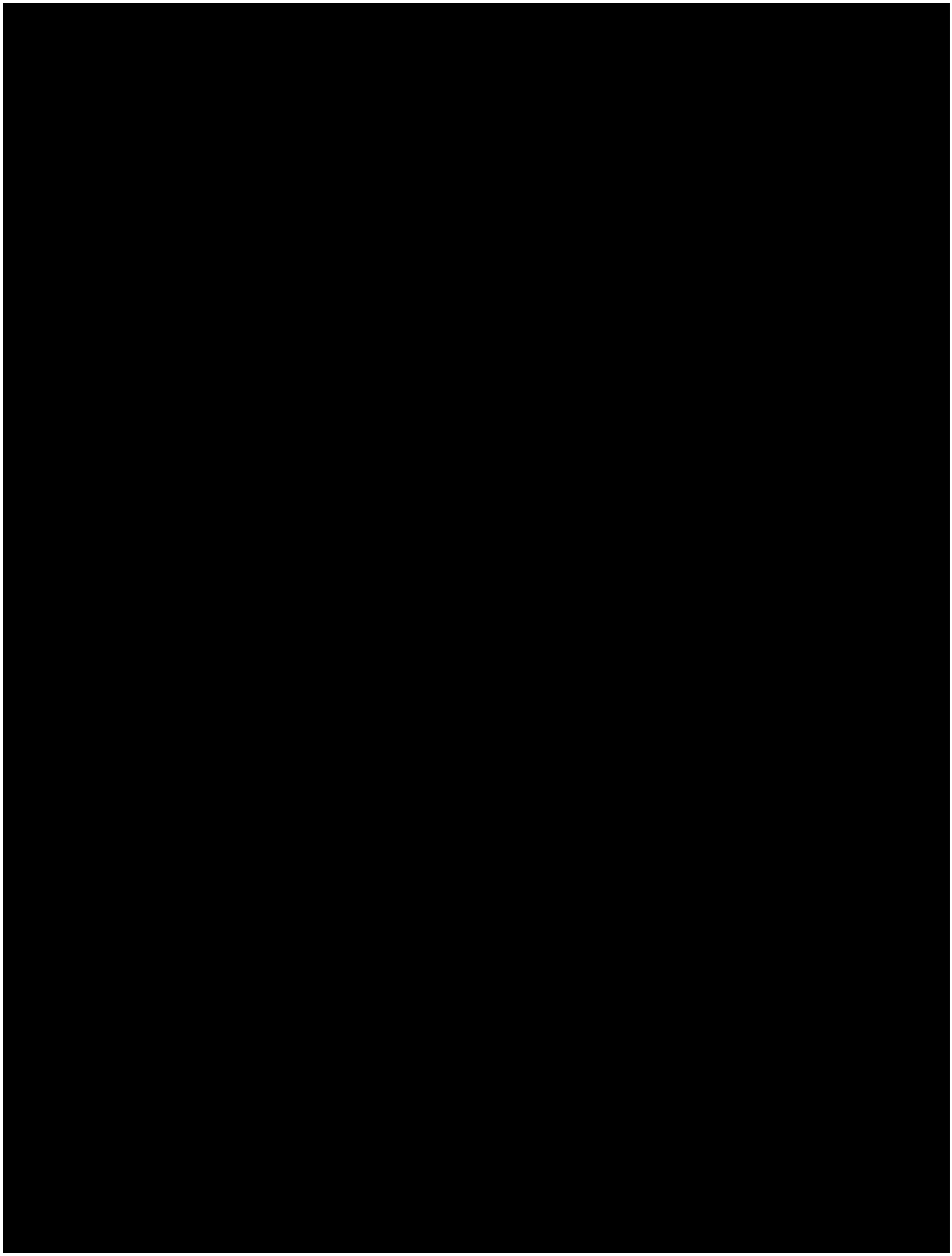
1. Missing data are filled in using multiple imputation to generate 100 complete data sets as described in steps 1 and 2 in Section 19.
2. The 100 complete datasets are analyzed by using the same ANCOVA model as the primary analysis.
3. The results from the 100 complete datasets are combined for inference.
4. Step 1 is repeated to generate multiple imputation datasets with a specified shift in the imputed values for observations in the treated group. No shift will be made for the placebo group.
5. Repeat step 2 for the imputed data sets with the shift parameter applied.
6. Repeat step 3 to assess statistical significance.
7. Repeat steps 4-6 with more stringent shifts until loss of statistical significance.
8. Assess whether the degree of departure from MAR needed to change the substantive results is clinically plausible. If the treatment effect is qualitatively maintained for the range of shift parameters that are considered to be clinically plausible, then the findings are considered to be robust.

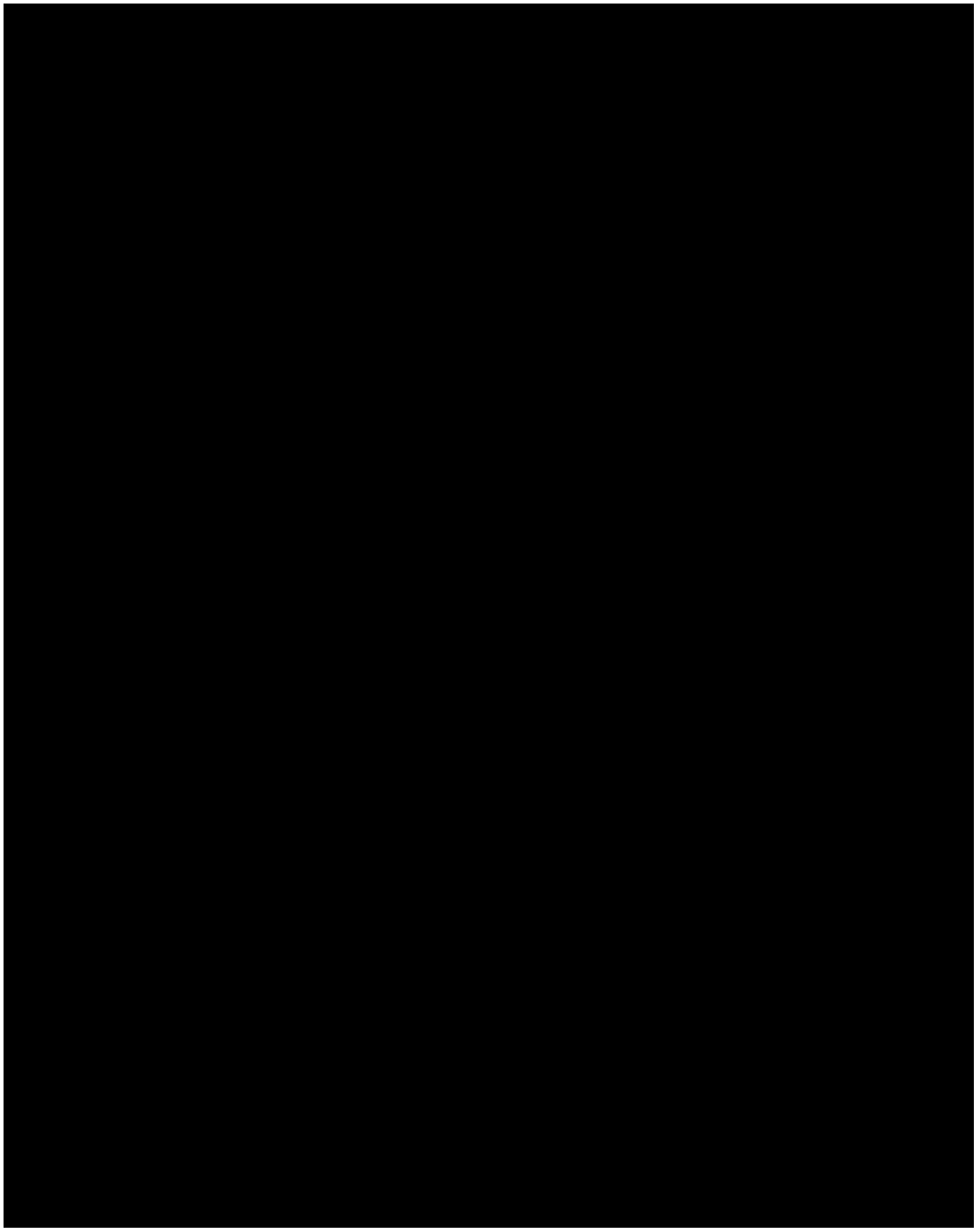
The change from baseline of the MAS in SMG at all post-treatment visits will be analyzed using MMRM with fixed factors of pooled center, history of prior treatment with BoNT products, treatment arm, visit, and treatment arm by visit interaction, and the baseline MAS score as a covariate. The PGIC will be analyzed using MMRM with fixed factors of pooled center, history of prior treatment with BoNT products, treatment arm, visit, and treatment arm by visit interaction. An unstructured variance-covariance (UN) structure will be used in the model. If the model with UN structure does not converge, the autoregressive (AR1) structure will be used with the assumption that measurements that are close to each other in time are going to be more closely correlated and as measurements get farther apart in time they are less correlated.

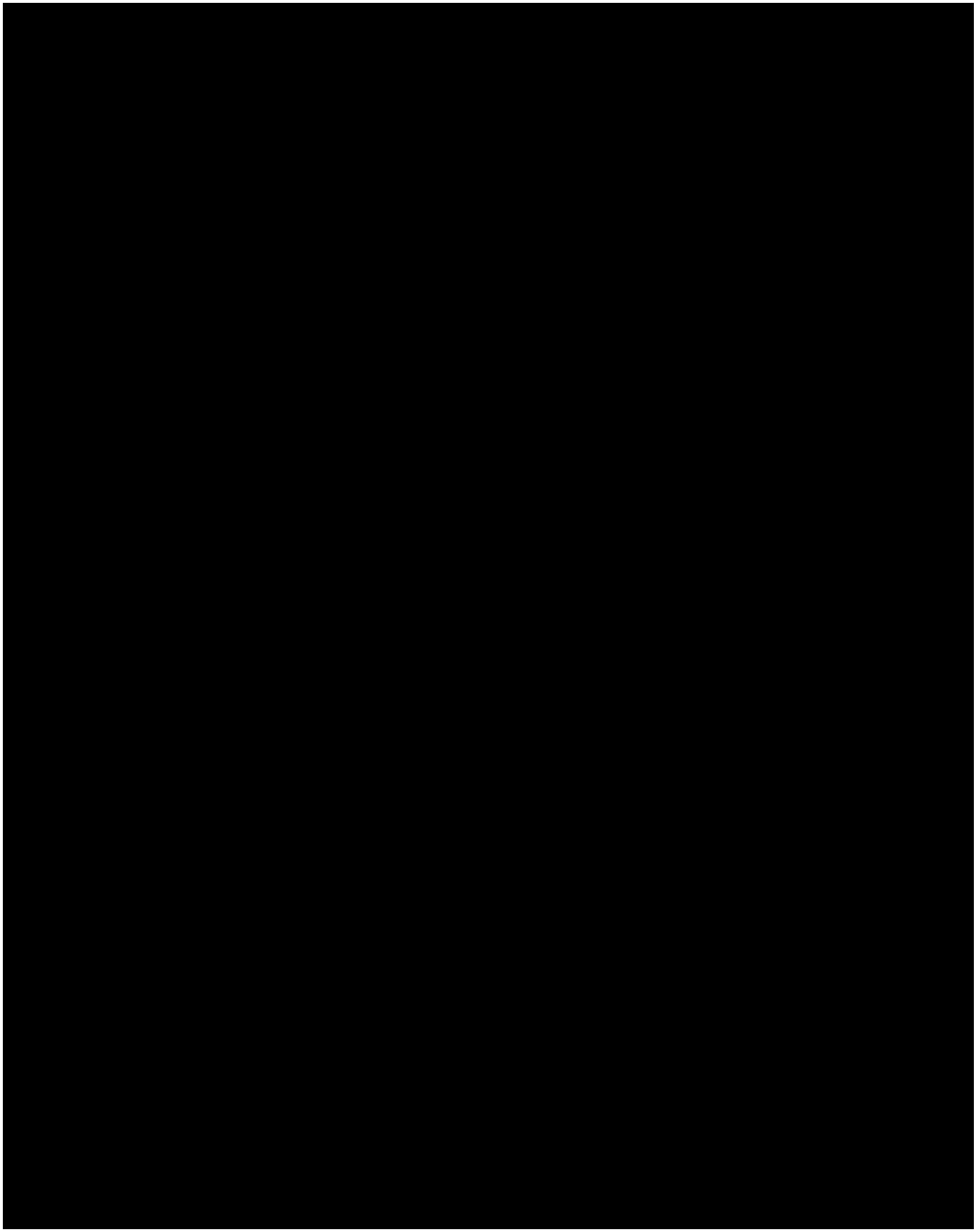
8.1.3. Subgroup Analyses of the Co-Primary Endpoints

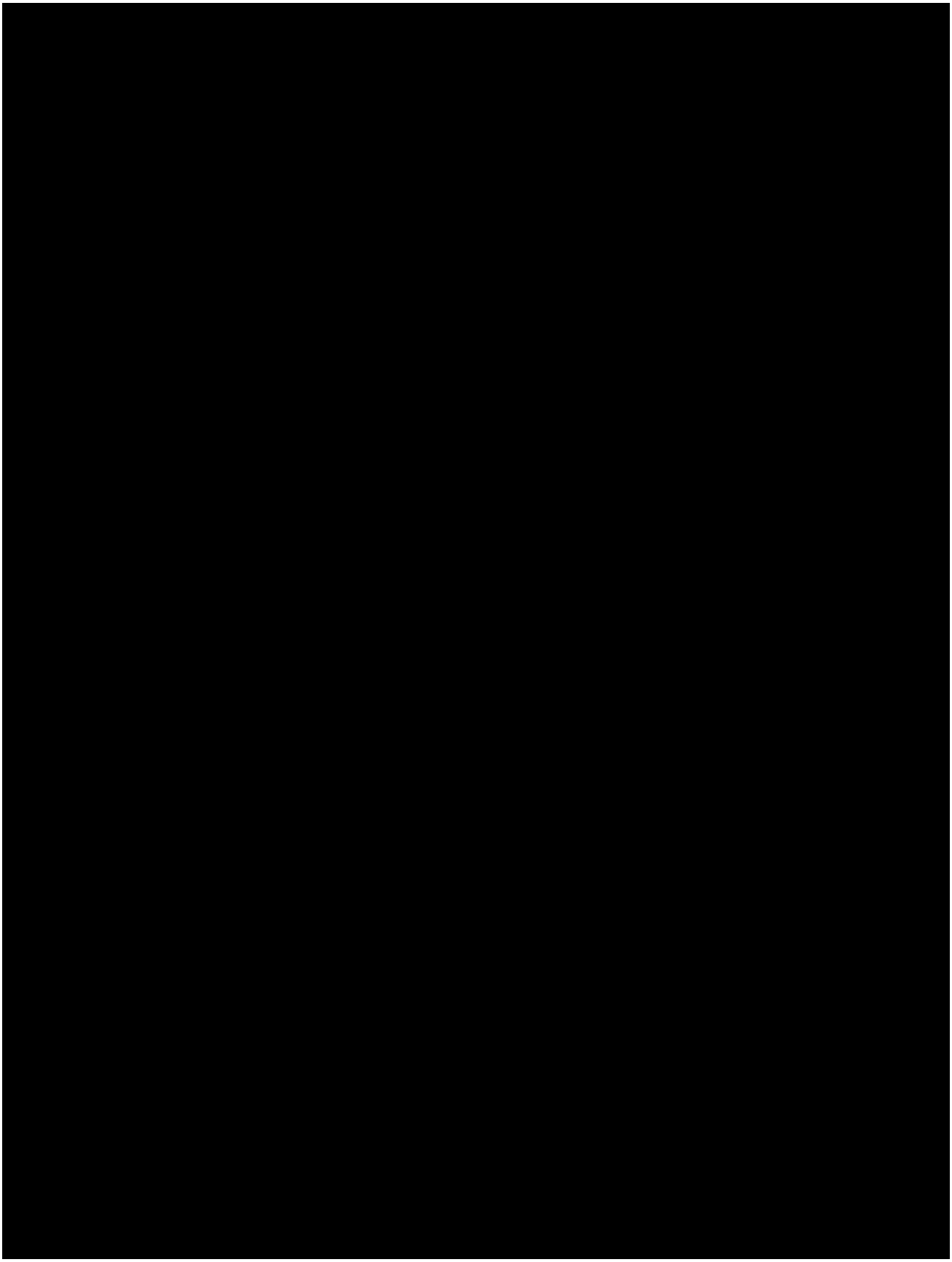
Subgroup analyses by prior BoNT treatment experience, history of ULS due to stroke versus TBI, gender, and age will be performed for both co-primary efficacy endpoint with adjustments in the ANCOVA model for the ITT and populations. Descriptive statistics and the ANCOVA model will be run separately for each subgroup.

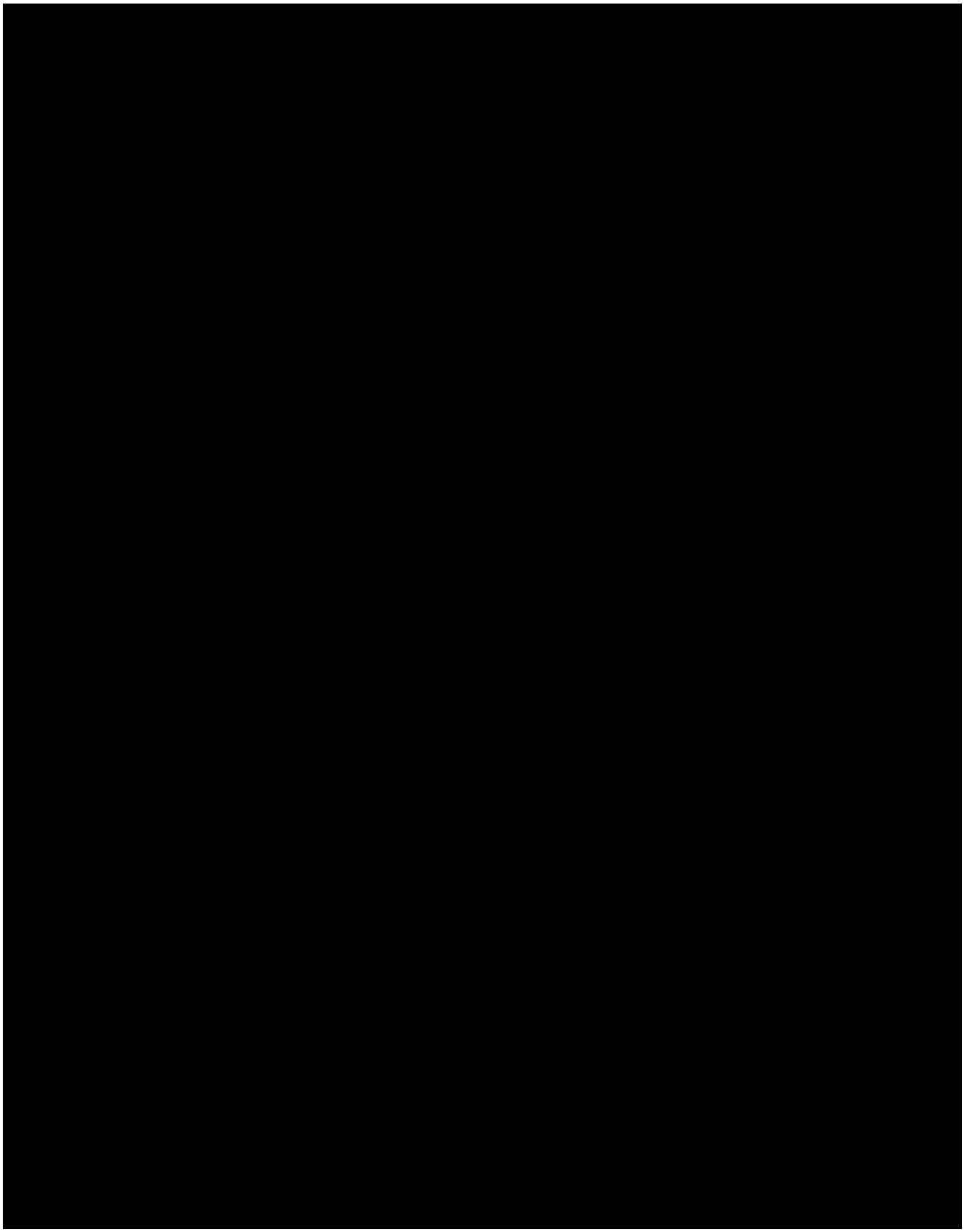












9. Safety

The safety population will be used for all safety analyses. Safety will be assessed on the basis of AEs, clinical laboratory values (hematology, INR, chemistry, and urinalysis), pregnancy test results, vital signs, [redacted] spirometry, and 12-lead electrocardiogram (ECG) results.

9.1. Injection Site Evaluation

The number and percentage of subjects with an injection site reaction will be summarized, along with the type of reaction. All injection site evaluation information will be provided in a listing.

9.2. Muscle Injection Record

The number and percentage of subjects injecting into each muscle will be summarized. All muscle injection record data will be listed.

9.3. Adverse Events

AEs will be coded using MedDRA version 22 or later. An AE will be considered treatment emergent if the event begins on or after the study medication dose date or if the event begins before and worsens in severity after the study medication dose date. An adverse event with a relationship to study medication of definite, probable, or possible is also considered to be treatment-emergent.

AEs will be summarized by system organ class (SOC) and preferred term (PT). An AE is considered related to study treatment if the relationship is considered as definite, probable, or possible. If the relationship to study treatment is missing, the AE will be counted as definite in summary tables. If the severity is missing, then the AE will be counted as severe in summary tables. Actual missing values will be presented in the listings.

Summary tables will be provided for treatment-emergent AEs (TEAEs) for the following:

- Overall summary of TEAEs
- TEAEs by SOC and PT
- TEAEs by SOC and PT occurring in at least 5% of subjects
- TEAEs by SOC, PT, and maximum severity
- TEAEs by SOC, PT, and strongest study medication relationship
- Treatment-related TEAEs by SOC and PT
- Serious TEAEs by SOC and PT
- TEAEs resulting in study medication discontinuation by SOC and PT [redacted]

The overall summary of TEAEs will include the number and percentage of subjects with a TEAE, with treatment-related TEAEs, with serious TEAEs, [redacted], and with TEAEs leading to death. For TEAE summaries, the number and percentage of subjects will be presented by SOC and PT by treatment. Subjects will be counted only once for each SOC and for each PT within each SOC. For summaries by maximum severity and strongest relationship, if a subject has the same event more than once, the subject will be included only once at the worst severity or strongest relationship. Summary tables by SOC and PT will be sorted in descending frequency by SOC and then by PT within each SOC by the DAXI 500 U column. All AEs will be included in the listings. Listings will be provided for all AEs, serious AEs, and AEs resulting in study medication discontinuation.

[redacted]

9.5. Laboratory Evaluations

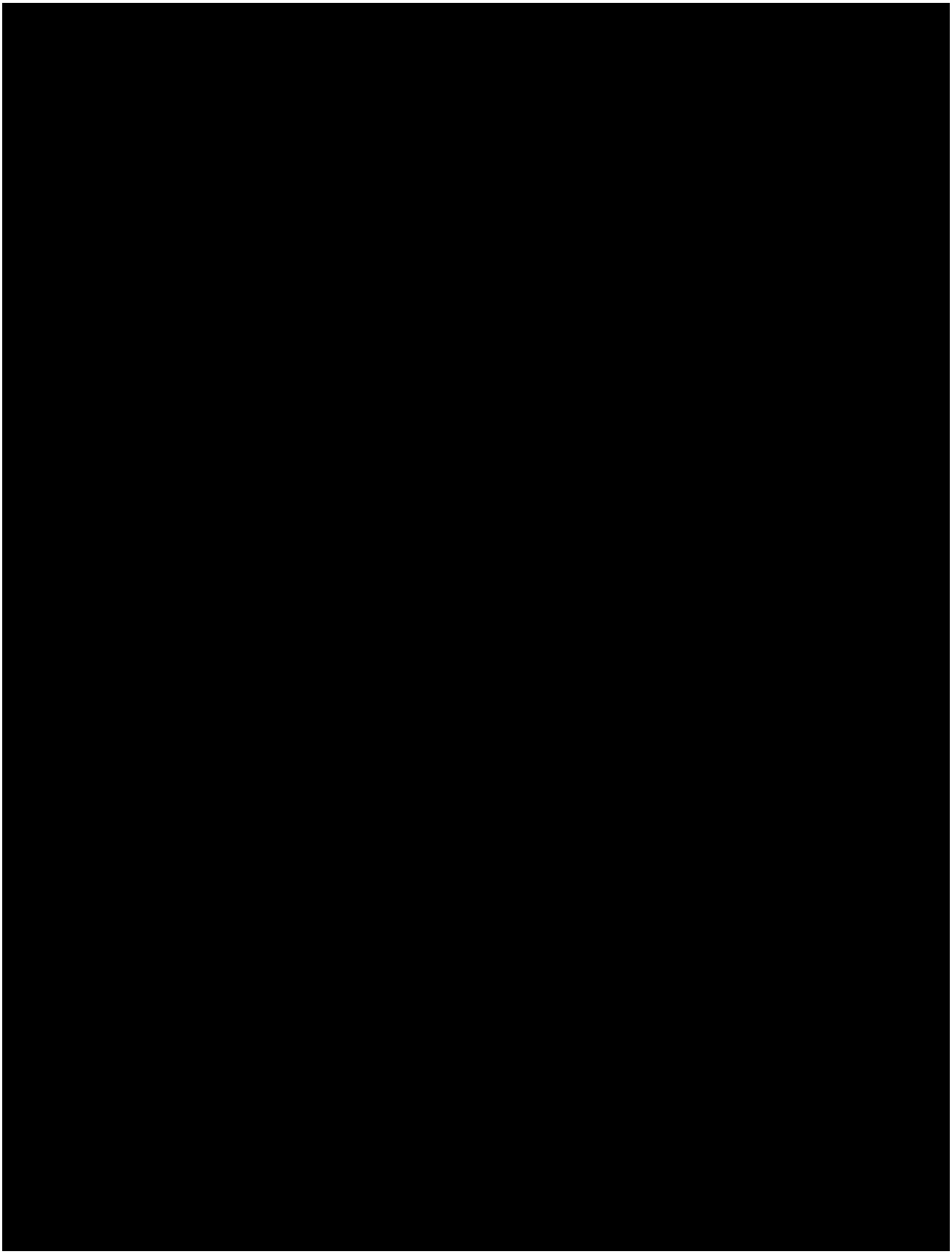
Non-fasting samples will be collected for serum chemistry, hematology, coagulation (INR), and urinalysis at Screening, Week 12, and Week 36/End of Study. In addition, for women of child bearing potential, a pregnancy test will be evaluated at Screening, Baseline, Week 12, and Week 36/End of Study. Blood samples for antibody testing are collected at Screening, Week 6, Week 12, and Week 36/End of Study. The hematology tests that will be collected are: Hemoglobin, Hematocrit, Leukocyte count (total), Leukocyte count (differential), Red blood cell count, and Platelet count. Serum chemistry tests that will be collected are: Sodium, Potassium, Chloride, Carbon dioxide (Bicarbonate), Calcium, Albumin, Glucose, Total bilirubin, Alanine aminotransferase, Aspartate aminotransferase, Alkaline phosphatase, Blood urea nitrogen, and Creatinine. Urinalysis tests that will be collected are: Specific gravity, pH, Glucose, Protein, Blood, Bilirubin, and Ketones. Additional tests that may be collected are: Prothrombin time, urine pregnancy for women of childbearing potential, serum pregnancy test at end of study if urine pregnancy test is positive, and [redacted] Summary tables and listings will report laboratory results using Système International (SI) units.

Summary statistics for observed values and changes from baseline will be provided for chemistry, hematology, coagulation, [redacted]at each visit by treatment. Summary statistics for observed urinalysis values will be provided at each visit by treatment.

Shift tables will be provided for shift from Baseline to Week 12 and shift from Baseline to Week 36/End of Study. Shift tables will be based on normal ranges and results reported as low, normal, or high. For urinalysis values, results will be reported as normal or abnormal.

A listing will be provided for all laboratory values at all visits. A listing of all out-of-range or clinically significant laboratory test results at any evaluation will be provided. Determination of clinical significance for all out-of-range laboratory values were to be made by each investigator and will be included in the listing.

[redacted]



[redacted]

9.7. Vital Signs

Vital signs include weight, body temperature, respiration rate, sitting radial pulse rate, and sitting systolic and diastolic blood pressure and are collected at Screening, Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, and Week 36/End of Study. Height is collected at the Screening visit only. Body Mass Index (BMI) will be computed using the weight at each visit and height at Screening: BMI=weight (kg) / height $(m)^2$.

Vital signs will be summarized by treatment and visit. All vital signs will be included in a listing.

9.8. ECG

A standard 12-lead ECG will be performed after the subject has rested quietly in the supine position at Screening, Week 12, and Week 36/End of Study. ECG data will be submitted to a central reader for review.

A summary of the overall ECG interpretation will be provided by treatment and visit. Summary statistics for ECG interval data will be provided by treatment and visit. The number and percentage of subjects in each of the following QTcF categories will be provided by visit and any post-treatment visit:

- QTcF > 450 msec for males or > 470 msec for females
- QTcF > 480 msec
- QTcF > 500 msec
- Increase from baseline QTcF ≥ 30 msec
- Increase from baseline QTcF ≥ 60 msec

If the ECG is repeated more than once at any timepoint, average values will be used for the summary. All ECG results will be included in a listing

9.9. Physical [redacted]

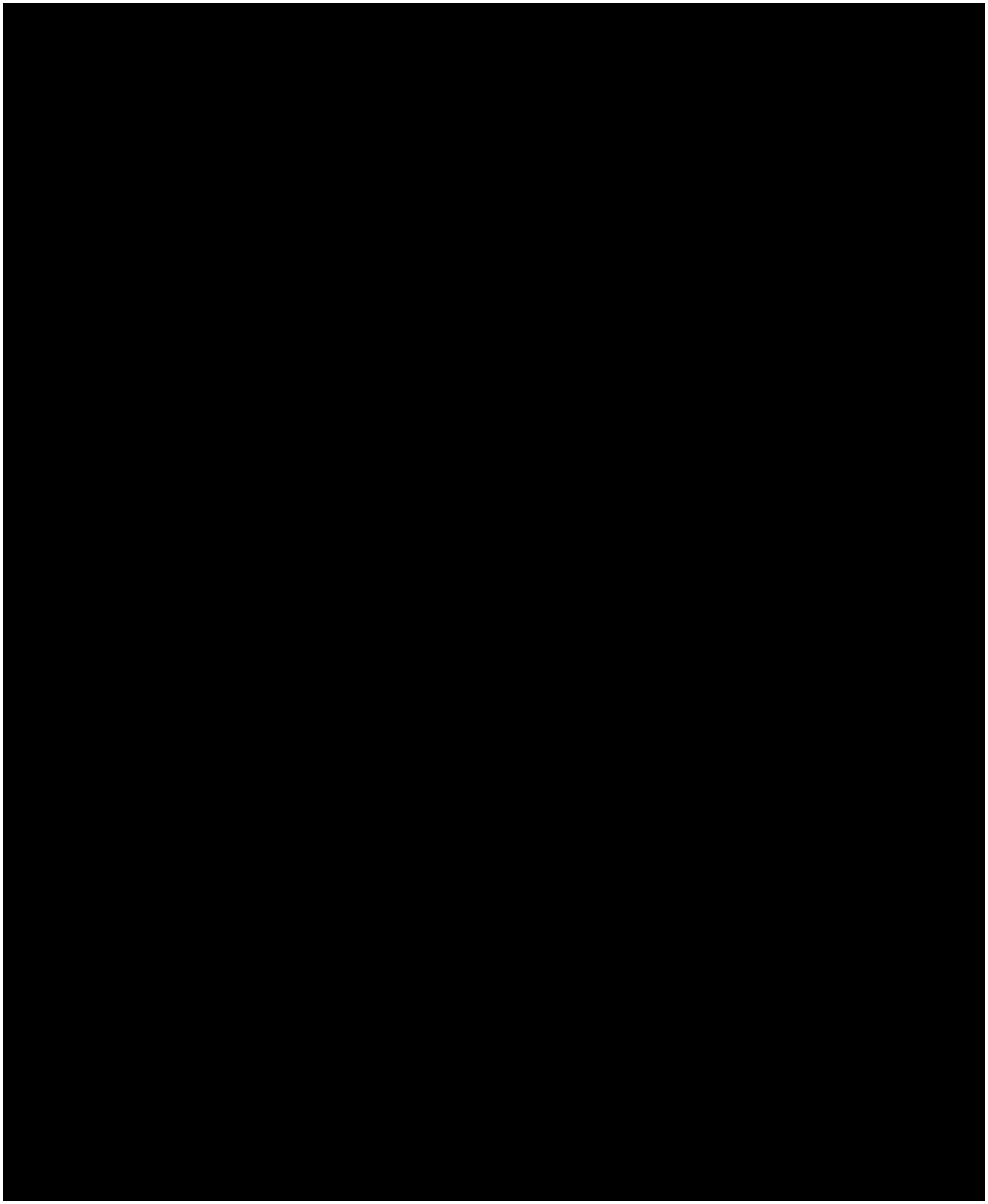
Physical [redacted] results are collected at Screening, Baseline, Week 2, Week 4, Week 6, Week 8, Week 12, Week 16, Week 20, Week 24, Week 28, Week 32, and Week 36/End of Study. A complete physical examination is performed at Screening and Baseline. At post-treatment visits, the physical examination may be abbreviated. Results will be summarized by treatment and visit. All data will be provided in a listing.

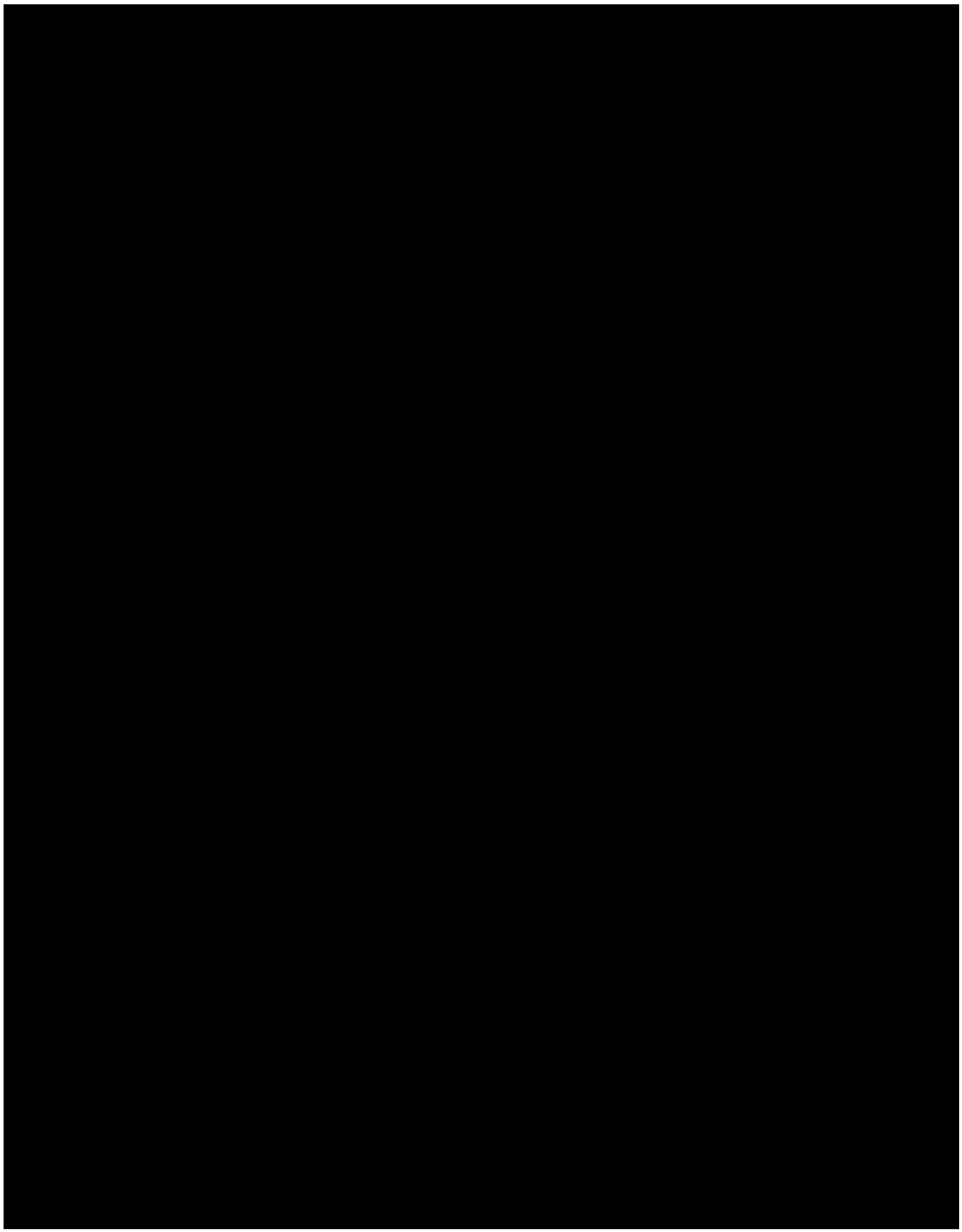
9.10. Pulmonary Function Test

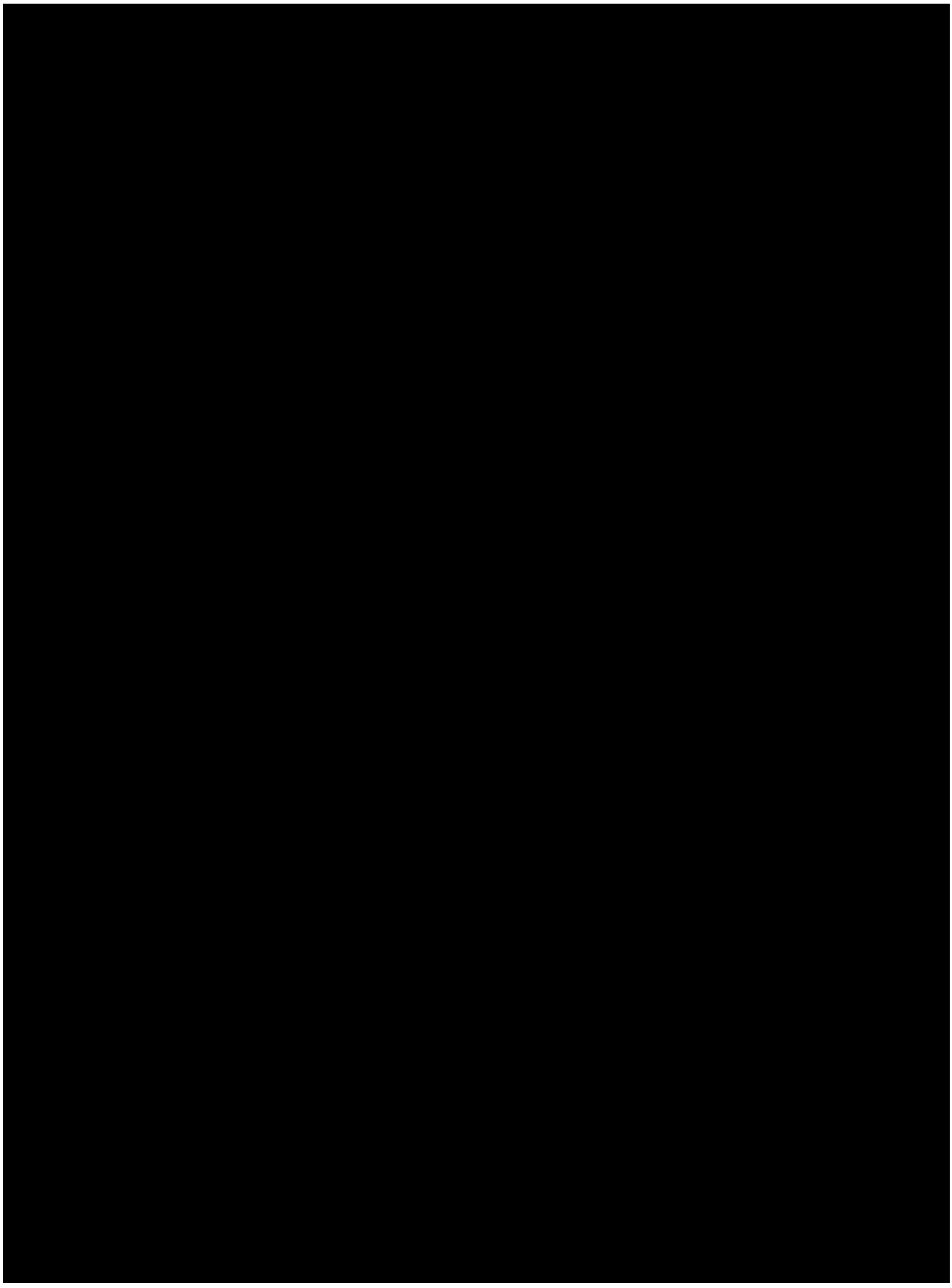
Pulmonary function will be recorded using spirometry. Parameters measured are the forced vital capacity (FVC), which is the total volume of air exhaled during forced expiration, and the forced expiratory volume in the first second of exhalation (FEV_1). Spirometry results are collected at Baseline, Week 6, Week 12, and Week 36/End of Study, however, if there is a post-treatment decline in FVC ≥ 20%, continue to monitor FEV_1 and FVC at all subsequent study visits. The FCV and FEV_1 results and changes from baseline will be summarized by treatment and visit. All data will be provided in a listing. Note, spirometery was discontinued 20 March 2020 to reduce the risk of possible spread of invection due to COVID-19.

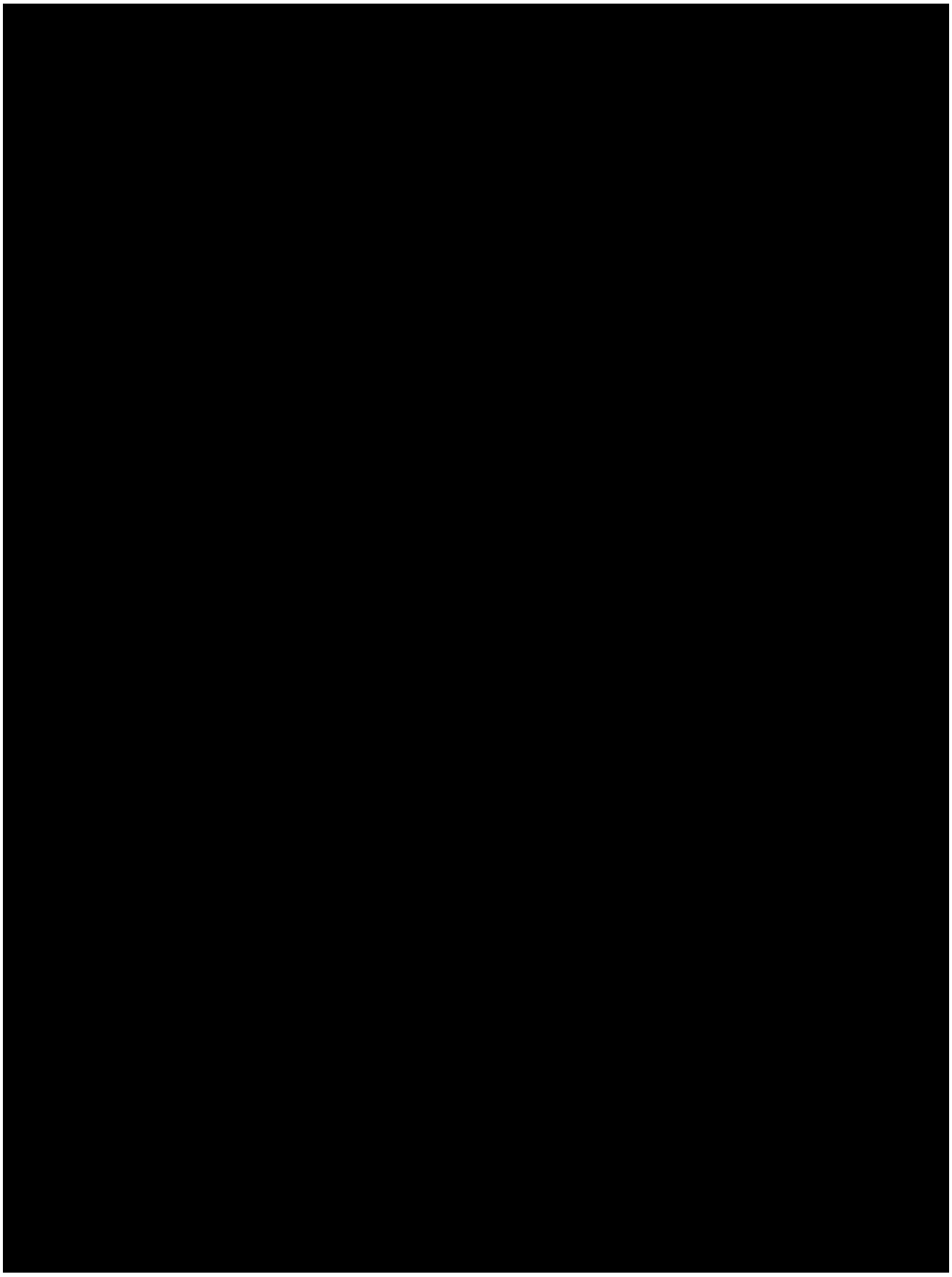
9.11. Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS will be collected at screening and all visits. The number and percentage of subjects in each category of the suicidal ideation items will be summarized by visit and treatment. All C-SSRS information will be provided in a listing.





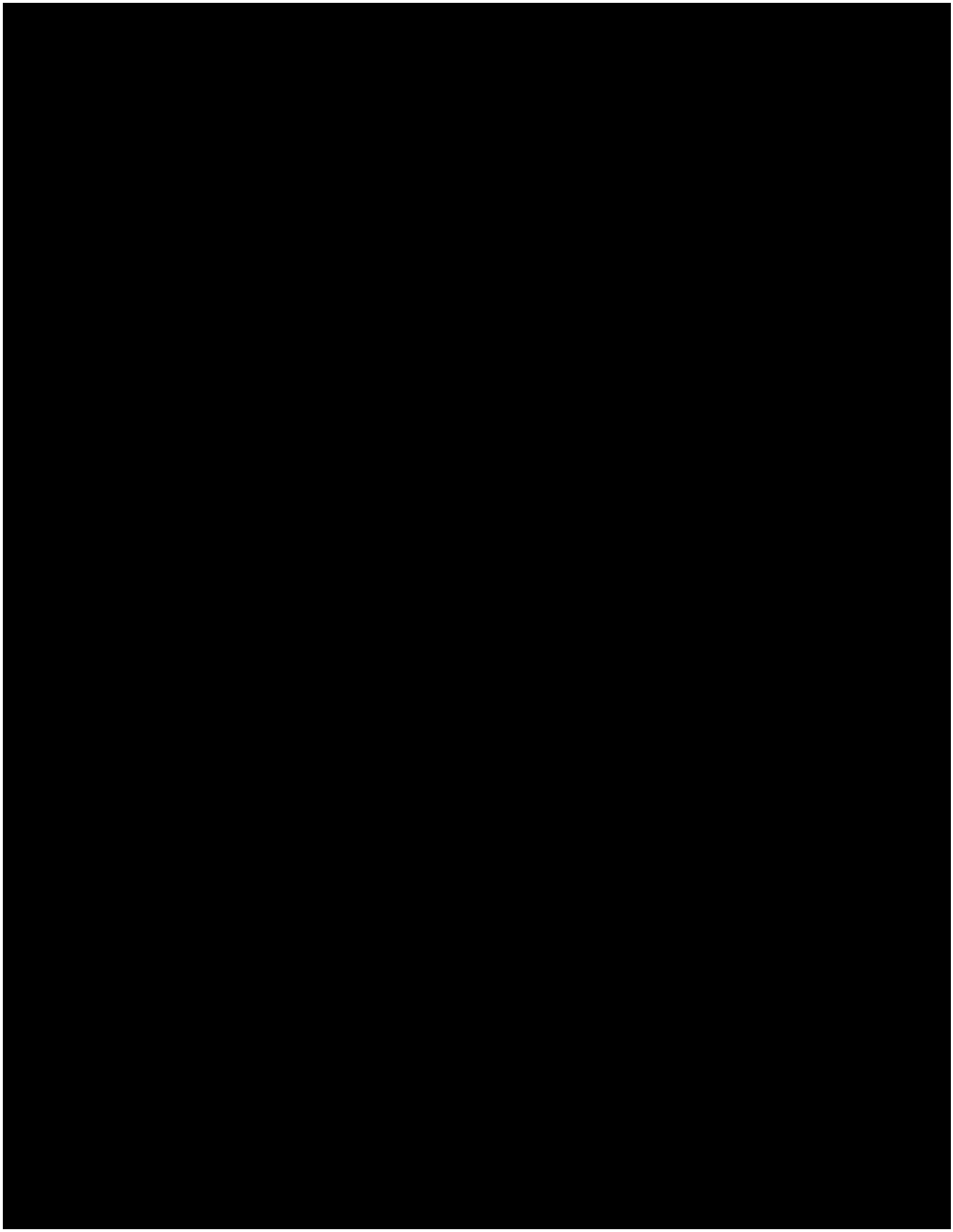




14. Quality Control

SAS programs are developed to produce output such as analysis data sets, summary tables, data listings, figures or statistical analyses. An overview of the development of programs is detailed in Syneos Health Standard Operating Procedure (SOP) Developing Statistical Programs (3907).

Syneos Health SOPs Developing Statistical Programs (3907) and Quality Deliveries (SDTM, ADaM, TLF) (3908) describes the quality control procedures that are performed for all SAS programs and output. Quality control is defined here as the operational techniques and activities undertaken to verify that the SAS programs produce the output by checking for their logic, efficiency and commenting and by review of the produced output."



[redacted]