COVER PAGE

Official Title:	A Phase 2 Study of the Safety, Efficacy, and Pharmacodynamics of RTA 408 in the Treatment of Friedreich's Ataxia
NCT number:	NCT02255435
Document Date:	Part 1: 15 May 2017 Part 2: 18 Sept 2019 Extension: 12 Jan 2026

STATISTICAL ANALYSIS PLAN
STUDY PART 1

Study Title:	**A Phase 2 Study of the Safety, Efficacy, and Pharmacodynamics of RTA 408 in the Treatment of Friedreich's Ataxia**
Name of Test Drug:	RTA 408 Capsules
Indication:	Friedreich's Ataxia
Sponsor:	Reata Pharmaceuticals, Inc.
Protocol No.:	RTA 408-C-1402
Protocol Version/Date	Version 7.0 (April 13, 2017)
Author Biostatistician:	[redacted], PhD
Analysis Plan Version/Date:	Version 2.0 (Part 1)/15-May-2017

CONFIDENTIAL AND PROPRIETARY INFORMATION

SPONSOR APPROVAL

of the

Statistical Analysis Plan – Study Part 1

A Phase 2 Study of the Safety, Efficacy, and Pharmacodynamics of RTA 408 in the Treatment of Friedreich's Ataxia

[redacted]

[redacted], PhD Date

[redacted]

[redacted] Date

Reata Pharmaceuticals, Inc.

[redacted] MS

Reata Pharmaceuticals, Inc.

TABLE OF CONTENTS

1. INTRODUCTION 7
1.1 Study Objectives 7
1.1.1 Primary Objectives 7
1.1.2 Secondary Objective 7
1.1.3 Exploratory Objectives 7
1.2 Part 1 Study Design 8
1.2.1 Study Visits and Assessments 9
1.2.2 Data Safety Monitoring Board 9
1.2.3 Discontinuation of Treatment 9
1.2.4 Patient Termination 9
1.2.5 Randomization and Unblinding 10
1.3 Sample Size and Power 10
2. PLANNED ANALYSIS OF PART 1 11
2.1 Timing of Analysis 11
2.2 Interim Analysis 11
2.3 Endpoint Measures 11
2.3.1 Efficacy Endpoints 11
2.3.2 Safety Endpoints 11
2.3.3 Pharmacodynamic Endpoints 11
2.3.4 Pharmacokinetics Endpoints 12
2.4 Changes from Protocol-Specified Analysis 12
3. GENERAL CONSIDERATIONS FOR DATA ANALYSIS 13
3.1 Analysis Considerations 13
3.2 Analysis Sets 13
3.2.1 Randomized Analysis Set 13
3.2.2 Intent-to-treat Population (ITT) 13
3.2.3 Safety Population 13
3.2.4 Pharmacokinetic 13
3.3 Strata and Covariates 13
3.4 Examination of Patient Subsets 14
3.5 Multiple Comparisons 14
3.6 Missing Data 14
3.7 Visit Windows 14
3.7.1 Definition of Study Day 14

3.7.2 Definition of Study Baseline 14
3.7.3 Analysis Visits 14
3.7.4 Analysis Windows for Unscheduled Visits 15
3.7.5 Selection of Data in the Event of Multiple Records in a Visit 15
3.8 Rounding 15
4. BASELINE CHARACTERISTICS AND PATIENT DISPOSITION 16
4.1 Disposition of Patients 16
4.2 Demographics and Baseline Characteristics 16
4.3 Study Drug Administration 17
4.4 Inclusion and Exclusion Criteria 17
4.5 Medical History 17
4.6 Prior and Concomitant Medications 18
4.7 Protocol Deviations 18
5. EFFICACY ANALYSIS 19
5.1 Primary Efficacy 19
5.2 Secondary Efficacy 19
5.3 Exploratory Efficacy 20
5.3.2 Other Maximal Exercise Testing Measures (including Oxygen Utilization) 21
5.3.3 25-Foot Timed Walk Test (T25-FW) 22
5.3.4 Low-Contrast Letter Visual Acuity Test 22
5.3.5 9-Hole Peg Test (9-HPT) 22
5.3.6 Fatigue Severity Scale 23
5.3.7 SF-36 Health Survey Update 23
5.3.8 Full Neurological FARS Score 24
[redacted] [redacted] 24
6. SAFETY ANALYSIS 25
6.1 Adverse Events 25
6.1.1 General Considerations for Analysis of Adverse Events 25
6.1.2 Adverse Event Dictionary 25
6.1.3 Treatment-Emergent Adverse Events 25
6.1.4 Adverse Event Severity 26
6.1.5 Relationship of Adverse Events to Study Drug 26
6.1.6 Serious Adverse Events 26
6.2 Clinical Laboratory Evaluations 27
6.2.1 General Considerations for Analyses of Laboratory Data 27

6.2.2 Summaries of Laboratory Results 27
6.2.3 Laboratory Abnormalities 27
6.2.4 Summaries of Laboratory Abnormalities 27
6.3 Vital Signs 28
6.4 Echocardiogram 28
6.5 Electrocardiogram 28
6.6 Physical Exam 29
6.7 Pregnancy 29
7. PHARMACODYNAMIC PARAMETERS 30
[redacted] 30
7.2 Platelet Samples 30
[redacted] 30
8. REFERENCES 31
9. APPENDICES 32

Rationale for Version 2.0

The changes from Version 1.0 to Version 2.0 of the Statistical Analysis Plan were motivated by protocol amendments subsequent to the finalization of Version 1 of the SAP. Version 1 was finalized 13Oct2014 under protocol version 3. The protocol was amended four times since Version 1 of the SAP (Protocol Version 4, 12May2015; Protocol Version 5, 29Sep2015; Protocol Version 6 23May2016; Protocol Version 7 13Apr2017). The current protocol allows for up to approximately 9 cohorts of patients in the dose-ranging portion of the study (part 1). Additional modifications were made to provide clarity to the statistical methods used in the planned analysis.

The Version 2 of the SAP also specifies separate analysis plans for Part 1 and Part 2 of this study. Having separate analysis plans for each study part will allow for information learned in Part 1 to be applied to the analysis of Part 2.

Other notable modifications include:

- Number of patients increased with the addition of more cohorts in Part 1
- Clarify that separate analysis sets will be defined for each Part
- Patients without pes cavus have been identified as a particular subgroup of interest based on emerging information.

1. INTRODUCTION

Study 408-C-1402 (the MOXIE study) is a 2-part, randomized, double blind, placebo-controlled, Phase 2 study designed to compare the efficacy and safety of RTA 408 to placebo in patients with Friedreich's ataxia. This document describes the statistical analysis methods and data presentations that Reata Pharmaceuticals, Inc. (Reata) will use to analyze data from the MOXIE study. [redacted] is MOXIE's blinded study statistical group responsible for conduc[redacted]n this document. Within each study part, [redacted] will remain blinded to study treatment assignments until after the database is locked. Study related documents include the study protocol and case report form. A data safety monitoring board (DSMB) will be reviewing data on a regular basis to monitor the study for safety. The DSMB is managed by an independent unblinded statistical group, [redacted] A separate document contains details regarding the DSMB.

The database will be locked after each study part (*i.e.,* Part 1 and Part 2). This version of the SAP describes the analyses planned for Part 1 prior to database lock of Part 1. Unless otherwise specified, the analyses described in this document will be performed after database lock for Part 1 for inclusion in the discussion of Part 1 data in the final clinical study report (CSR). Any substantive changes made to the statistical analysis plan after database lock will be clearly documented and a justification will be given in the CSR.

This SAP is based on Version 7 of the study protocol dated April 13, 2017. If the protocol is subsequently amended, this SAP may be amended as well. Should the SAP and the protocol be inconsistent with respect to the planned analyses, the language of the SAP is governing. All analyses will be conducted using SAS version 9.2 or higher.

1.1 Study Objectives

In patients with Friedreich's ataxia, comparing those receiving RTA 408 versus those receiving placebo, the objectives of the MOXIE study are as follows:

1.1.1 Primary Objectives

The primary objectives of this study are:

- To evaluate the change in peak work during maximal exercise testing
- To evaluate the safety and tolerability of RTA 408

1.1.2 Secondary Objective

The secondary objective of this study is:

- To evaluate the change in the modified Friedreich's ataxia rating scale (FARS) score

1.1.3 Exploratory Objectives

- To evaluate the change in peak oxygen utilization during maximal exercise testing
- To evaluate the change in performance on a timed 25-foot walk test
- To evaluate the change in performance on a low-contrast letter visual acuity test
- To evaluate the change in performance on a 9-hole peg test (9-HPT)
- To evaluate the change in Fatigue Severity Scale score

- To evaluate the change in SF-36® Health Survey Update (SF-36) score
- To evaluate the change in the full FARS score
- [redacted]
- To evaluate the change in pharmacodynamic (PD) markers of activity in platelet, [redacted] samples [redacted]
- To characterize the pharmacokinetics of RTA 408 and potential metabolites after oral administration of RTA 408 Capsules

1.2 Part 1 Study Design

Part 1 of this study will evaluate the efficacy, safety, PK, and PD of RTA 408 in the treatment of up to approximately 72 patients with Friedreich's ataxia.

The first part of the MOXIE study will be a randomized, placebo-controlled, double-blind, dose-escalation study to evaluate the safety of RTA 408 at 2.5 mg, 5 mg, 10 mg, 20 mg and higher dose levels (not to exceed 300 mg) in up to approximately 9 cohorts of patients with Friedreich's ataxia. A cohort consists of the next eight eligible patients randomized 1:3 to placebo (n=2) or the cohort specific dose of RTA 408 (n=6). Up to 9 cohorts will be enrolled in Part 1 to allow for adequate dose-ranging for selection of RTA 408 dose(s) to be used in Part 2.

Intra-patient dose-escalation will only be utilized in the first cohort to evaluate RTA 408 at the first two dose levels (2.5 mg and 5 mg). Patients enrolling in the first cohort will be randomized to RTA 408 2.5 mg or placebo. After the Week 2 visit, each patient in the first cohort will dose escalate to 5 mg (or remain on placebo) on Day 15 unless a dose-limiting toxicity (DLT) is reported in that patient. After the last patient in the first cohort completes their Week 4 visit (i.e., 2 weeks on 2.5 mg daily [or matching placebo] followed by 2 weeks on 5 mg daily [or matching placebo]), the data safety monitoring board (DSMB) and Sponsor will review all available safety information and make a decision regarding enrollment of the next cohort. Beginning with the second 8-patient cohort, once the eighth patient enrolled completes their Week 2 visit the DSMB will review all available safety information and recommend the dose of RTA 408 for the subsequent cohort. The DSMB dose recommendation for each cohort must not exceed 100% more than the highest dose of RTA 408 evaluated in this study, and the maximum permitted dose of RTA 408 is 300 mg. The dose-level for each new cohort will not exceed the DSMB recommended dose level, and it will be selected by the Sponsor based on review of available safety, efficacy, PK, and PD data. Prior to opening each cohort in Part 1 for enrollment, the Sponsor will evaluate all available data from doses studied in Part 1 to determine if enough information is available to select doses for Part 2 of the study. Once doses are selected for Part 2 by the Sponsor, no additional cohorts will be enrolled in Part 1.

1.2.1 Study Visits and Assessments

Detailed schedules of assessments for the MOXIE study are included in the study Protocol. Patients will self-administer study treatment once daily until they have completed 12 weeks (84 days) of study treatment, are discontinued from study treatment, or have withdrawn consent to participate in the study. A follow-up visit for safety will occur at Week 16 (4 weeks after last dose).

1.2.2 Data Safety Monitoring Board

An independent Data Safety Monitoring Board (DSMB) will be monitoring data from the study on a monthly basis throughout Part 1 to ensure the safety of the patients in the trial.

An independent statistical group, [redacted], will prepare unblinded analyses for the DSMB and will not have a role in the statistical analysis plan (SAP) or day-to-day support of the study.

1.2.3 Discontinuation of Treatment

Discontinuation refers to a patient stopping administration of study drug. Reasons for study drug discontinuation may include the following:

- Occurrence of an adverse event or change in medical status that leads the investigator to be concerned about the patient's welfare
- Protocol violations
- Administrative reasons (e.g., inability to continue)
- Sponsor termination of the study
- Voluntary withdrawal
- Pregnancy during the study
- Investigator unblinding
- Other

Patients who are discontinued from study drug should still complete all study visits and undergo all scheduled study assessments, if possible.

1.2.4 Patient Termination

Termination refers to a patient stopping study drug and all study assessments and visits. Reasons for study termination include the following:

- Administrative reasons (e.g., inability to continue, lost to follow-up)
- Death
- Withdrawal of consent
- Other

Patients who terminate the study for any reason may not re-initiate study drug or study assessments at any time.

1.2.5 Randomization and Unblinding

The first cohort enrolled in Part 1 will be randomized 3:1 to RTA 408 2.5 mg or placebo. Beginning with the second cohort enrolled in Part 1, patients will be randomized 3:1 to a specific dose of RTA 408 or placebo as recommended by the DSMB.

To maintain the study blind, all study drug kits will be packaged with blinded labels. Investigators will distribute the blinded study drug kits by kit number to patients as assigned by the IWRS. All patients, investigators, site personnel, and laboratories with direct involvement in the conduct of the study or their designees will be blinded to treatment assignments and appropriate measures will be taken to ensure the blind is maintained to reduce potential bias. Some Sponsor personnel may have access to treatment assignments during dose escalation (Part 1).

1.3 Sample Size and Power

The sample size for Part 1 (i.e., 8 per cohort for 72 total patients in Part 1) is based on a dose-escalation scheme to evaluate initial safety and PD activity of RTA 408 in this patient population.

2. PLANNED ANALYSIS OF PART 1

This study consists of two parts: a placebo-controlled, dose-ranging phase (Part 1) followed by a second placebo-controlled phase (Part 2) utilizing one dose selected from Part 1. Analyses planned for Study Part 2 will be described in a separate document.

2.1 Timing of Analysis

The primary analysis of Part 1 will be performed after all patients enrolled in Part 1 have completed Part 1 and the Part 1 records are locked in the EDC system. Other analyses of Part 1 data may be performed prior to database lock in order to plan for Part 2 of this study.

2.2 Interim Analysis

A DSMB will be reviewing data on an ongoing basis from the study to ensure patient safety. The DSMB will not recommend stopping the study for efficacy. The DSMB charter describes the interim reviews of safety. No other formal interim analyses are planned for this study.

2.3 Endpoint Measures

Unless otherwise specified, efficacy measures are obtained for all patients.

2.3.1 Efficacy Endpoints

The primary efficacy endpoint is change in peak work. Peak work will be derived by dividing the maximum work load (watts) at each visit by the patient's baseline weight (kg).

The secondary efficacy endpoint is the change in modified FARS (mFARS) score. The mFARS consists of only the neurological portion of the full FARS neurological assessment (i.e., the full FARS minus the peripheral nervous system). Additional details regarding the derivation of the modified FARS score can be found in Section 5.2.

Exploratory endpoints for this trial include:

- Parameters (other than peak work) collected during maximal exercise testing
- 25-Foot Timed Walk Test (T25-FW)
- Low-Contrast visual acuity test at different contrast levels (100%, 2.5%, 1.25%)
- 9-Hole Peg Test (9-HPT)
- Fatigue Severity Scale
- SF-36 Health Survey Update (PCS, MCS, Utility Index Score)
- Full Neurological FARS Score
- [redacted]

2.3.2 Safety Endpoints

Safety parameters include results of echocardiogram, ECG, vital sign measurements, weight, BMI, physical examination, AEs, SAEs, concomitant medications, and laboratory tests (clinical chemistry, hematology, urinalysis, microscopy, and pregnancy tests [as indicated]).

2.3.3 Pharmacodynamic Endpoints

A number of PD markers will be assessed in this study. Frataxin levels will be assessed using [redacted], platelet samples at Sponsor approved sites, [redacted].

[redacted] and platelet data may not be available for analysis at the time of database lock. Should this be the case, [redacted] and platelet data will be summarized in separate reports.

[redacted]

Available pharmacodynamic endpoints will be summarized in a similar manner as clinical laboratory data in Section 6.2 if data are collected from a sufficient number of patients. All PD endpoints will be listed.

2.3.4 Pharmacokinetics Endpoints

Pharmacokinetics will be summarized and reported in a stand-alone PK report performed by either Reata or a vendor other than [redacted].

2.4 Changes from Protocol-Specified Analysis

There are no planned changes to analyses from what is described in the protocol.

3. GENERAL CONSIDERATIONS FOR DATA ANALYSIS

3.1 Analysis Considerations

All individual data will be listed as recorded in the database. All statistical summaries and analyses will be performed using SAS® software (SAS Institute, Cary, North Carolina, USA).

Only patients in the appropriate analysis set will be included in summary statistics. For endpoints discussed in this SAP, all available data for each patient will be included in summaries.

Individual patient data will be presented in data listings. The listings will include all available data for all patients. Unless otherwise noted, all data will be listed.

Continuous data will be summarized by treatment using descriptive statistics (number, mean, standard deviation [SD], minimum, median, quartiles (Q1, Q3), and maximum). Categorical data will be summarized by treatment using frequency tables (number and percentage).

3.2 Analysis Sets

For Part 1 of the study, the following analysis sets will be defined.

3.2.1 Randomized Analysis Set

The Randomized Analysis Set will include all patients who are randomized for the study. The randomized analysis set will be used to assess patient disposition. Patients will be summarized according to the treatment to which they were randomized.

3.2.2 Intent-to-treat Population (ITT)

The intent-to-treat (ITT) population includes all enrolled patients in Part 1, whether or not they received study drug. Patients randomized to 2.5 mg in Part 1, Cohort 1, dose escalate to 5 mg at Week 2. All patients receiving active dose in Cohort 1 will be summarized as the 5 mg dose. All other patients remain on their randomized treatment throughout the study (i.e., no intra-patient dose-escalation) and will be categorized by their randomized treatment group.

3.2.3 Safety Population

The safety population includes all patients who received at least 1 dose of study drug. Patients who receive at least 1 dose of RTA 408 will be classified in the RTA 408 group at the highest dose level received.

3.2.4 Pharmacokinetic

The Pharmacokinetic analysis set will be defined in a separate analysis plan as it will be a stand-alone analysis.

3.3 Strata and Covariates

Randomization in Part 1 of this study is not stratified. Since sample size per dose level will be small, no covariates will be considered for the primary analysis of Part 1.

3.4 Examination of Patient Subsets

Emerging data have identified patients with pes cavus as a particular subgroup of interest. Some safety and efficacy analyses may be performed on subjects with pes cavus.

3.5 Multiple Comparisons

No adjustments for multiplicity will be considered for analysis of Part 1 data.

3.6 Missing Data

Missing data will not be imputed for safety, PK, PD, or exploratory efficacy endpoints. The planned statistical methods use all available data, therefore no imputation is planned for the primary analysis of the primary and secondary efficacy endpoints (i.e., peak work and mFARS). If more than 10% of Week 12 assessments are missing, a sensitivity analysis of the primary and secondary endpoints will be performed with missing data imputed using Jump to Reference (J2R) multiple imputation (Ratitch B, 2013).

3.7 Visit Windows

3.7.1 Definition of Study Day

Study day is the day relative to the first dose of study drug administration. Study day for events on or after the date of the first dose will be defined as the number of days from the date of the first dose of study drug, plus 1 day, so that the date of the first dose will be defined as Day 1.

For events before the date of the first dose, study day will be calculated as the difference in days between the date of the first dose and the date of interest. Thus, the day before the date of the first dose will be defined as Day -1.

3.7.2 Definition of Study Baseline

The average of Screening and Day 1 assessments will be used as baseline for FARS, modified FARS, maximal exercise test parameters, 25FWT, and 9HPT. For all other parameters, study baseline is defined as the last non-missing observation obtained prior to administration of the first dose on Study Day 1. Baseline will be defined as analysis visit 0.

3.7.3 Analysis Visits

The CRF nominal study visits will be used for all safety analyses. When specified, summarizes by visit will be done using analysis visits defined below.

Analysis Visit	Label	Visit Number	Approximate Study Day
0	Baseline	0,1	≤1
4	Week 4	5	28
8	Week 8	7	56
12	Week 12	9	83
12	Week 12	10	84

3.7.4 Analysis Windows for Unscheduled Visits

Unscheduled visits will be reflected in summarization of changes to worst post-baseline measures.

Additionally, unscheduled assessments will be used in the place of missing scheduled efficacy assessments, if the unscheduled assessment was collected within ±7 days of the scheduled assessment. Study day will be calculated for each unscheduled assessment and compared to the protocol defined study day for each visit. In general, analysis windows for unscheduled visits will be defined as follows:

Visit Number	Analysis Visit	Study Day	Analysis Window
1	0	1	Study Day ≤1
5	4	28	21 ≤ Study Day ≤ 35
7	8	56	49 ≤ Study Day ≤ 63
9	12	83	76 ≤ Study Day ≤ 90
10	12	84	77 ≤ Study Day ≤ 91

Records from unscheduled visits that do not fall within an analysis window will be listed, but will not be analyzed.

3.7.5 Selection of Data in the Event of Multiple Records in a Visit

If multiple assessments fall within the same analysis visit, the assessment closest to the target visit day specified in the protocol study procedures will be used. If two assessments are equidistant from a post-baseline target visit day, the earlier assessment will be used.

3.8 Rounding

The method of rounding for data presentation is provided in Appendix 2.

4. BASELINE CHARACTERISTICS AND PATIENT DISPOSITION

4.1 Disposition of Patients

A disposition summary will include:

- Number and percentage of patients randomized in each cohort
- Number and percentage of patients in the ITT
- Number and percentage of patients in the safety analysis set
- Number of patients dose escalated (Cohort 1 only)
- Number and percentage of patients completing treatment at each dose
- Number and percentage of patients discontinuing treatment at each dose
- Number and percentage of patients by reason for discontinuing treatment at each dose
- Number and percentage of patients completing study at each dose
- Number and percentage of patients terminating study at each dose
- Number and percentage of patients by reason for terminating study at each dose

Disposition summaries will be presented by randomized treatment. Percentages will be based on the number of randomized patients.

A listing of patient disposition will include the date of the last dose of study drug, study completion status, and reason for termination (if applicable).

4.2 Demographics and Baseline Characteristics

Demographic and baseline characteristics summaries, for patients in the safety analysis set, will include:

- Age (years at study baseline)
- Sex
- Race
- Ethnicity
- Weight (kg)
- Height (cm)
- Baseline peak work

Due to the number of measures, baseline characteristics will be provided in a separate summary. Continuous statistics will be provided for:

- Age at FA onset
- Duration of diagnosis

- Average number of hours exercised per week
- GAA1 repeat length
- GAA2 repeat length

Frequency tables (number and percentage) will be provided for discrete measures:

- Ambulatory status
- Ambulatory assistive devices
- History of cardiomyopathy
- History of areflexia
- Presence of pes cavus
- Presence of extensor plantar reflex(Babinski reflex)
- Foot surgery
- History of sensory neuropathy
- History of swallowing difficulties
- Visual acuity less than 20/25
- History of scoliosis
- Prior scoliosis surgery

Demographic and baseline characteristics will be summarized by randomized treatment group.

4.3 Study Drug Administration

Study drug compliance will be assessed using the number of pills dispensed and the number returned from each bottle. The number of doses taken is assumed to be the difference between the total dispensed and the total returned. The expected number of doses will be derived based on the interval between date dispensed and date returned. Percent compliance is derived as ((taken/expected)*100).

Number and percentage patients with >80% and ≤80% compliance will be summarized descriptively by treatment and visit. The percentages will be based on the number of patients in the safety analysis set. No treatment comparisons will be made.

A listing will be provided that displays study drug administration data including compliance.

4.4 Inclusion and Exclusion Criteria

A listing of inclusion or exclusion violations for randomized patients as well as a similar listing for screen failures will be provided.

4.5 Medical History

Medical history summary will be presented by treatment. Percentages will be based on the number of patients in the safety analysis set. Medical history will be sorted alphabetically and all data will be listed.

4.6 Prior and Concomitant Medications

Prior and concomitant medications will be coded using the World Health Organization (WHO) drug dictionary (December, 2012) for anatomical therapeutic chemical classification (ATC) and preferred drug name. A patient who used multiple medications will be counted only once for each ATC and preferred drug name. ATC and preferred drug name within each ATC will be sorted alphabetically. Coded concomitant medications will be summarized by treatment. Percentages will be based on the number of patients in the safety analysis set.

Prior medications are defined as those medications disclosed as being taken prior to Study Day 1 on the CRF. Concomitant medications are defined as any medications that was either not taken prior to Study Day 1 (according to the CRF), or was taken prior to Study Day 1 but was ongoing at time of the first dose administration.

4.7 Protocol Deviations

All protocol deviations will be captured on the CRF and reviewed by the Sponsor according to their SOPs. Major protocol deviations will be identified by the Sponsor based on blinded review of the data, and will be flagged in the listing.

5. EFFICACY ANALYSIS

Primary analysis of the efficacy data will be based on the ITT population, which will include all patients randomized in Part 1 of the study. Analyses will be performed for each RTA 408 dose level and also all RTA 408 patients pooled for comparison with all placebo patients pooled.

Summary statistics for observed values, change from baseline, and percent change from baseline (including 95% CI and quartiles) will be presented by randomized treatment group, and will include an additional column to summarize the pooled RTA 408 patients.

5.1 Primary Efficacy

Peak work (w/kg) is calculated as the maximum workload (w) at each visit divided by baseline weight (kg) for each patient. A positive change from baseline suggests an improvement.

Peak work for patients treated with RTA 408 will be compared with placebo after 12 weeks of treatment using mixed models repeated measures (MMRM) analysis, with treatment group, time, and the interaction between treatment and time as fixed factors. Analysis visits 0, 4, and 12 will be used. An unstructured covariance matrix will be assumed. If convergence can't be achieved, a compound symmetry covariance matrix will be evaluated.

The difference between RTA 408 (each dose level and all doses pooled) and placebo in change from baseline of mean peak work will be estimated along with the 95% confidence interval at each protocol scheduled time point. A p-value to test the effect of RTA 408 on peak work will only be provided at Week 12.

Suppose it is assumed that the treatment variable ***trt*** has two levels (1=placebo, 2=active) and values for ***analysis visit*** are (0, 4, 12). The following SAS code serves as sample code to provide the estimated difference, confidence interval, and p-value.

```
Proc Mixed data=x order=data;
  Class patno trt week;
  model resp =  trt week trt*week /s;
  Repeated week / sub = patno type = un;
  lsmeans trt*week / diff cl;
  contrast "Active vs Pbo" trt -1 1  trt*week 0 -1 0 1 ;
  estimate "Active vs Pbo" trt -1 1  trt*week 0 -1 0 1 ;
Run;
```

The pairwise dose group comparisons with placebo will be estimated using a separate repeated measures model than the model that pools the RTA 408 dose levels. If it is assumed the treatment variable ***trt*** now has three levels (1=placebo, 2=low dose, 3=high dose), the contrast and estimate statements to obtain pairwise differences are:

```
  contrast "Low vs Pbo" trt -1 1 0 trt*week 0 -1 0 1 0 0;
  estimate "Low vs Pbo" trt -1 1 0 trt*week 0 -1 0 1 0 0;
  contrast "High vs Pbo" trt -1 0 1 trt*week 0 -1 0 0 0 1;
  estimate "High vs Pbo" trt -1 0 1 trt*week 0 -1 0 0 0 1;
```

This requires a separate execution of Proc Mixed where the active treatments are not pooled.

5.2 Secondary Efficacy

The secondary endpoint of this study is the modified FARS score. The Friedreich's Ataxia Rating Scale is a validated composite measure consisting of scores obtained from a standard neurological examination for Friedreich's ataxia patients. The FARS examination for the

MOXIE study is divided into 5 sections: Bulbar (section A), Upper Limb Coordination (section B), Lower Limb Coordination (section C), Peripheral Nervous System (section D), and Upright Stability (section E).

A higher FARS or mFARS score reflects greater Friedreich's ataxia disease severity thus, a negative change from baseline suggests an improvement.

Assessments in each section are added to get the section score. FARS is the total of sections A through E. Modified FARS is the total of sections A through C, and E (i.e., omit section D from the full FARS score). If the FARS assessment is performed but a measurement or section of the FARS is missing, values should be imputed using LOCF for just the missing measurement or section to avoid an artificial score decrease. If the FARS assessment is not performed the FARS result should remain null for the primary analysis, since missing data are not imputed.

The modified FARS score will be analyzed using repeated measures analysis of variance with the same model used for the primary efficacy endpoint. Analysis visits 0, 4, 8 and 12 will be used. The pairwise dose group comparisons with placebo will be estimated using the difference in adjusted means and 95% CI for the difference in changes from baseline to Week 12.

5.3 Exploratory Efficacy

Summary statistics and 95% confidence intervals for treatment differences in mean change from baseline based on MMRM used in the primary and secondary endpoints will be provided for all exploratory endpoints unless otherwise noted. P-values will be provided for descriptive purposes only.

As with the primary and secondary endpoints, exploratory endpoints will be analyzed using pooled patients at doses of RTA 408 versus placebo pooled. A separate model that does not pool the RTA 408 doses will be used to obtain pairwise differences for each dose level versus pooled placebo.

5.3.2 Other Maximal Exercise Testing Measures (including Oxygen Utilization)

In addition to the primary efficacy endpoint of peak work, the following other parameters are also assessed during maximal exercise testing.

	At Rest (Prior to Test)	At Maximal Work (End of Test)
Heart rate (bpm)	X	X
VO_2[a] **(mL/kg/min)**	X	X
O_2 pulse[b] **(mL/beat)**	X	X
Systolic blood pressure (mmHg)	X	X
Diastolic blood pressure (mmHg)	X	X
Oxygen saturation (SpO_2) (%)	X	X
Forced vital capacity (FVC) (L)	X	
Forced expiration volume within one second (FEV1) (L)	X	
Maximum voluntary ventilation (MVV) (L/min)	X	
%Predicted FVC (%)	X	
VCO_2[a] **(mL/kg/min)**		X
Anaerobic threshold (L/min)		X
Time to reach maximal work (Min:Sec)		X
Maximum rating of perceived exertion (BORG; 0-20 scale)		X
Maximum minute ventilation (VE) BTPS (L/min)		X
Tidal Volume (L)		X
Maximum respiratory rate (breaths/min)		X
Breathing Reserve (%)		X

[a] Analysis of maximal VO2 and resting VO2 will be performed using adjustments for baseline weight:

- Maximal_VO2_adj=Maximal_VO2*1000/weight (kg);
- Resting_VO2_adj=Resting_VO2*1000/weight (kg);

[b] O_2 pulse will be calculated for analysis: O_2 pulse = (Maximal_VO2 (in units of mL/min)*1000)/Maximal_HR)

Observed values and difference between resting and peak performance for these parameters will be summarized at baseline and at each timepoint. For all assessments, each parameter will be summarized at baseline and at each time point along with the change from baseline by treatment. For parameters measured at resting and at the end of the maximal exercise test [heart rate, VO_2

(L/min) and VO_2 (ml/mg/min), diastolic blood pressure, systolic blood pressure, and SPO_2 (%)], the change from resting and the change from baseline should be calculated. The change from resting (delta) for parameters measured at resting and at the end of the maximal exercise test is defined as:

- Delta = Parameter at maximal work - Parameter at resting

The change from baseline is defined for parameters measured at resting and at the end of the maximal exercise test is defined as:

- Change from baseline in Delta= [Parameter at maximal work(At Visit X) - Parameter at resting(At Visit X)] - [Parameter at maximal work(At Baseline) - Parameter at resting(At Baseline)]

Only summary statistics will be provided.

5.3.3 25-Foot Timed Walk Test (T25-FW)

The T25-FW is a quantitative mobility and leg function performance test based on time in seconds to complete a 25- foot walk. Only the completed walk from each visit will be used for analyses. Observed and reciprocal values (i.e., 1/time) will be analyzed at each visit. Changes in reciprocal values are calculated as:

- Change from baseline = 1/(Time to complete 25-foot walk test at Visit X) - 1/(Time to complete 25-foot walk test at Baseline)

Analysis of reciprocals has limited utility but is provided to show information on general trends only. Transforming changes from baseline on the reciprocal scale are not interpretable.

T25-FW will be summarized at baseline and each time point along with the change from baseline by treatment. Analysis visits 0, 4, 8, and 12 will be used. Longer walk times reflect more impairment in mobility thus, a negative change from baseline suggests an improvement.

5.3.4 Low-Contrast Letter Visual Acuity Test

Loss of low contrast visual acuity is a principal ophthalmologic manifestation of Friedreich's ataxia. Low-contrast letter visual acuity is collected at three different contrast levels (100%, 2.5%, or 1.25%). The aggregate score (100%+2.5%+1.25%) will be summarized at baseline and each time point along with the change from baseline by treatment. Only the measure for both eyes will be summarized. Analysis visits 0, 4, 8, and 12 will be used. Lower score indicates greater visual impairment thus, a positive change from baseline suggests an improvement.

5.3.5 9-Hole Peg Test (9-HPT)

The 9-HPT is a brief, standardized, quantitative test of upper extremity function. Both the dominant and non-dominant hands are tested twice. The scores for the dominant and non-dominant hand will be analyzed separately. The two trials for each hand are averaged. The average time for each hand will be summarized at baseline and each time point along with change from baseline by treatment. Observed and reciprocal values (i.e., 1/time) will be analyzed at each visit. Changes in reciprocal values are calculated as:

- Change from baseline = 1/(Time to complete 9-hole peg test at Visit X) - 1/(Time to complete 9-hole peg test at Baseline)

Analysis of reciprocals has limited utility but is provided to show information on general trends only. Transforming changes from baseline on the reciprocal scale are not interpretable.

Analysis visits 0, 4, 8, and 12 will be used. Longer time reflects more impact on one's upper extremity function by the disease thus, a negative change from baseline suggests an improvement.

5.3.6 Fatigue Severity Scale

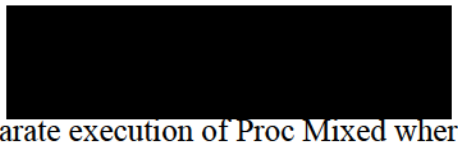
The Fatigue Severity Scale is a self-administered test designed to measure the severity of the fatigue symptoms based on 9 questions (Krupp, 1989). An individual score for each question ranges from 1 to 7 where 1 indicates strong disagreement and 7 indicates strong agreement. The average of the 9 questions for each patient is used for analysis at each visit. If an individual score is missing, the average of available scores will be used. This average score will be summarized at Baseline and Week 12 along with change from baseline by treatment. Larger scores reflect more impact of the disease thus, a negative change from baseline suggests an improvement.

Since there is only one scheduled post-baseline assessment, an analysis of covariance model with treatment group as a fixed factor will be used. The pairwise dose group comparisons with placebo will be estimated using the difference in adjusted means and 95% CI for the difference in changes from baseline to week 12. Analysis visits 0 and 12 will be used.

Suppose it is assumed that the treatment variable ***trt*** has two levels (1=placebo, 2=active). The following SAS code serves as sample code to provide the estimated difference, confidence interval, and p-value.



To obtain estimates of the pairwise differences, the code will need to be updated to account for additional levels of the treatment variable. The following contrast and estimate statements may be used to obtain pairwise differences.



This requires a separate execution of Proc Mixed where the active treatments are not pooled.

5.3.7 SF-36 Health Survey Update

The SF-36 is a multi-purpose, short-form health survey with 36 questions. It yields an eight-scale profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures and a preference-based health utility index. It is a generic measure, as opposed to one that targets a specific age, disease, or treatment group.

SF-36 breaks into 8 domains. Physical Functioning (items 3a-3j), Role-Physical (items 4a-4d), Role-Emotional (items 5a-5c), Social Functioning (items 6, 10), Bodily Pain (items 7, 8), Mental Health (items 9b, 9c, 9d, 9f, 9h), Vitality (items 9a, 9e, 9g, 9i) and General Health Perceptions

(items 1, 11a-11d). Change in Health (item 2) is an overall health status assessment that compares to one year ago. The score is converted to a 0 - 100 scale, with 0 indicating poor health and 100 indicating good health.

The overall composite score (SF-6D Utility Index Score), the MCS (mental component scale) and the PCS (physical component scale) as defined by the SF-36 manual will be summarized at baseline and analysis visit 12 along with the change from baseline by treatment. Lower scores reflect poor quality of life thus, a positive change from baseline suggests an improvement.

Since there is only one scheduled post-baseline assessment, an analysis of covariance model with treatment group as a fixed factor. The pairwise dose group comparisons with placebo will be estimated using the difference in adjusted means and 95% CI for the difference in changes from baseline to week 12. Analysis visits 0 and 12 will be used.

5.3.8 **Full Neurological FARS Score**

Scores will be summarized at Baseline and each time point along with change from baseline by treatment. Analysis visits 0, 4, and 12 will be used. Both the overall Neurological section score (sum of sections A through E) as well as the total score from each subsection will be summarized. Higher scores reflect greater Friedreich's ataxia disease severity thus, a negative change from baseline suggests an improvement.

[redacted]

6. SAFETY ANALYSIS

Safety data (including AEs, laboratory data, vital signs, electrocardiogram [ECG] data, echocardiogram data, and physical examinations) will be listed and summarized for patients in the safety analysis set. All safety data collected on or after the date of the first dose of study drug through completion of the study Part 1 will be summarized by treatment.

The CRF nominal study visits will be used for all safety analyses. Unscheduled visits will be reflected in summarization of changes to worst post-baseline measures where appropriate.

6.1 Adverse Events

AEs will be summarized by treatment as defined by the safety analysis set.

6.1.1 General Considerations for Analysis of Adverse Events

General considerations for AE summaries and calculations are:

- Multiple events by preferred term (PT) and system organ class (SOC) will be counted once only per patient for each treatment.
- For summaries by severity, only the most severe event will be counted per patient for each treatment.
- For summaries by relationship, only the most related event will be counted per patient for each treatment.
- An AE with a missing resolution date or incomplete date that is not identified as continuing will be assumed to be continuing and no duration will be calculated.
- AEs will be summarized by the highest dose received.
- Only treatment-emergent adverse events (TEAEs) will be included in summaries.

6.1.2 Adverse Event Dictionary

AEs will be coded using MedDRA® (Medical Dictionary for Regulatory Activities) version 16.0 or higher. In MedDRA, each verbatim term is mapped to a preferred term and high level term (HLT), which is then mapped to a system organ class. Tables and listings will present data at the SOC and PT level.

6.1.3 Treatment-Emergent Adverse Events

6.1.3.1 Definition of Treatment-Emergent

Treatment-emergent adverse events are events that either:

- Date of onset on or after the date of the date of first dose and not more than 30 days after the date of the last dose of study drug, or
- Had no recorded date of onset with a stop date after the first dose of study drug, or
- Had no recorded date of onset or stop date

6.1.3.2 Incomplete Dates

If the date of onset is incomplete, then the month and year of onset (or year alone if month is not recorded) determines whether the event is treatment-emergent. The event is treatment-emergent if the month and year of onset (or year of onset) of the event is:

- The same as or after the month and year (or year) of the first dose of study drug, and
- The same as or before the month and year (or year) of the date of the last visit.

6.1.4 Adverse Event Severity

The investigator will grade the severity of the AEs as mild, moderate, or severe as defined in the study protocol.

6.1.5 Relationship of Adverse Events to Study Drug

Association or relatedness to the study medication will be graded by the investigator as either probably, possibly, unlikely, or unrelated according to criteria specified in the study protocol.

6.1.6 Serious Adverse Events

As defined in the protocol and captured on the CRF, an SAE is an adverse event that results in any of the following:

- Death;
- A life-threatening adverse drug experience;
- Inpatient hospitalization or prolongation of existing hospitalization;
- Results in persistent or significant incapacity or substantial disruption of the ability to conduct normal life functions;
- Results in a congenital anomaly or birth defect in an offspring of a patient taking study drug;
- Is an important medical event.

6.1.6.1 Summaries of Adverse Events

Treatment-emergent AEs will be summarized by treatment at onset of the AEs. For each treatment, SOC, and PT, the number and percentage of patients reporting an event will be calculated. In summary tables, SOC will be presented alphabetically and events within SOC will be presented by decreasing frequency count.

Summary tables (number and percentage of patients) of AEs (by SOC and PT) will be provided by treatment as follows:

- All treatment-emergent AEs
- All treatment-emergent related AEs (probably or possibly related)
- All treatment-emergent AEs by severity
- All treatment-emergent serious adverse events (including deaths)
- All treatment-emergent adverse events leading to discontinuation of study drug

Listings will be provided showing:

- All AEs
- Serious adverse events (including deaths)
- AEs leading to discontinuation of study drug

6.1.6.2 Additional Analysis of Adverse Events

No additional analysis of AEs is planned.

6.2 Clinical Laboratory Evaluations

Laboratory data will be summarized at baseline and at each time point by treatment. Only values obtained at scheduled assessment times will be included in by-visit summary statistics. When appropriate, change from baseline to worst post-baseline (including unscheduled visits) value will be summarized. Specific assessment times are given in the protocol.

Continuous laboratory data results which are less than the lower limit of quantification or above the upper limit of quantification will be imputed to the value of the lower or upper limit plus or minus one significant digit, respectively (e.g., if the result of a continuous laboratory test is < 20, a value of 19 will be assigned). A baseline value is defined as the last available measurement obtained prior to administration of the first dose. If multiple values for a given lab test exist for a particular day, the last one drawn will be used for summarization.

6.2.1 General Considerations for Analyses of Laboratory Data

Clinical laboratory results will be provided by a central lab. Clinical laboratory test results, including results from hematology, chemistry and urinalysis will be listed by patient.

6.2.2 Summaries of Laboratory Results

Selected laboratory evaluations and change from baseline will be summarized by treatment, laboratory category (hematology, chemistry), test, and study visit using continuous statistics. Laboratory tests to be summarized are provided in Appendix 1. Line graphs of change from baseline will be generated for selected laboratory tests, such as ALT, AST, GGT, Ferritin, and CK. Line graphs will include mean ± SE over time for change from baseline.

Due to the nature of urinalysis parameters, summaries of continuous statistics will not be provided.

6.2.3 Laboratory Abnormalities

Laboratory results are not captured on the CRF. All flags with respect to normal or abnormal results will be identified by the central lab and presented in the data listings.

6.2.4 Summaries of Laboratory Abnormalities

The number and percentage of patients with laboratory normality and abnormality categories (Normal, Low, High) as well as the treatment-emergent shift will be summarized by treatment, laboratory category (hematology, chemistry, and urinalysis), and laboratory test. Baseline, Last post-baseline, and shift to last post-baseline value will be presented. An initial set of parameters of specific interest (ALT, AST, GGT, Ferritin, and CK) will be summarized using shift tables, though additional parameters may be added in an ad hoc manner.

6.2.4.1 Transaminases

To assess the potential for drug induced liver injury, the follow criteria will be summarized:

- Number of patients with ALT more than 3x, 5x, 10x or 20xULN
- Number of patients with AST more than 3x, 5x, 10x or 20xULN
- Number of patients with either ALT or AST more than 3x, 5x, 10x or 20xULN

- Number of patients with ALP >1.5xULN
- Number of patients with Total Bilirubin >2xULN
- Number of patients with AST or ALT >3xULN with an associated Total Bilirubin >1.5xULN
- Number of patients with AST or ALT >3xULN with an associated Total Bilirubin >2xULN

A summary table that includes frequencies and percentages of patients that meet any of the above criteria at the same visit will be provided. A listing of subjects with abnormal ALT, AST, or Total Bilirubin will also be provided.

6.3 Vital Signs

Vital signs assessments include systolic blood pressure (SBP, mmHg) and diastolic blood pressure (DBP, mmHg), oral body temperature (°C), heart rate (HR, bpm), respiration rate (RR, breaths/min), height (m), weight (kg), and BMI (m^2/kg).

Vital signs will be summarized at baseline and at each time point along with the change from baseline by treatment. Only values obtained at scheduled assessment times will be included in by-visit summary statistics. Specific assessment times are given in the protocol. All data will be listed.

6.4 Echocardiogram

Cardiac function [ejection fraction (%), left and right wall thickness (mm), septum thickness (mm) and LVMI (g/m^2)] measurements are captured on Echocardiogram will be summarized at baseline and each time point along with the change from baseline by treatment, only values obtained at scheduled assessment times will be included. Specific assessment times are given in the protocol.

6.5 Electrocardiogram

ECG results will include ventricular rate (bpm), PR interval (ms), QRS duration (ms), QT (ms), and QTc (ms).

ECG results will be summarized at baseline and at each time point by treatment. Only values obtained at scheduled assessment times will be included. Specific assessment times are given in the protocol.

The number and percentage of patients with investigator findings of normal, abnormal (not clinically significant) and abnormal (clinically significant) as well as the shift to abnormal (clinically significant) will be presented by treatment and protocol scheduled assessment times. In addition to by-visit frequencies, the most abnormal as well as the shift to most abnormal will be summarized.

All ECG data will be listed.

6.6 Physical Exam

Physical exam results will be summarized at baseline and at each time point by treatment. Only values obtained at scheduled assessment times will be included in by-visit summary statistics. Specific assessment times are given in the protocol.

The number and percentage of patients with investigator findings of normal, abnormal (not clinically significant) and abnormal (clinically significant) as well as the shift to abnormal (clinically significant) will be presented by treatment and protocol scheduled assessment times. In addition to by-visit frequencies, the most abnormal as well as the shift to most abnormal will be summarized.

All physical exam findings will be listed.

6.7 Pregnancy

A listing will be provided serum and urine pregnancy results of all on-study pregnancies.

7. PHARMACODYNAMIC PARAMETERS

Levels of frataxin and other mitochondrial enzymes will be summarized in a similar manner as clinical laboratory data in Section 6.2 using the Safety Analysis set.

[redacted]

7.2 Platelet Samples

If data are available for at least 10% of patients at Week 12, frataxin levels and changes from baseline from platelet samples will be listed and summarized by treatment.

[redacted]

8. REFERENCES

Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The Fatigue Severity Scale: Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46:1121-1123

Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. Pharm Stat 2013;12:337-47.

9. APPENDICES

Appendix 1. List of Laboratory Tests

Appendix 2. Programming Specifications

Appendix 3. Tables, Listings, and Figures

Appendix 1. List of Laboratory Tests

Blood samples will be collected throughout the study for hematology, chemistry, and urinalysis for clinical laboratory evaluation.

Test panels will include the following:

Hematology	Chemistry	Urinalysis
Hematocrit	Blood urea nitrogen (BUN)	Specific gravity
Hemoglobin	Creatinine	Ketones
HbA1C	Total bilirubin	pH
Red blood cell (RBC) count	Alanine aminotransferase (ALT)	Protein
White blood cell (WBC) count	Aspartate aminotransferase (AST)	Blood
Neutrophils	Alkaline phosphatase (ALP)	Glucose
Bands (if detected)	Ferritin	Urobilinogen
Lymphocytes	Sodium	Bilirubin
Monocytes	Potassium	
Basophils (if detected)	Calcium	
Eosinophils (if detected)	Inorganic phosphorus	
Absolute platelet count	Magnesium	
Mean corpuscular hemoglobin (MCH)	Chloride	
Mean corpuscular volume (MCV)	Bicarbonate	
Mean corpuscular hemoglobin concentration (MCHC)	Uric acid	
Reticulocyte count	Cholesterol	
	Total protein	
	Glucose	
	Triglycerides	
	Albumin	
	Creatine phosphokinase (CPK)	
	Lactate dehydrogenase (LDH)	
	High-density lipoprotein cholesterol (HDL-C)	
	Low-density lipoprotein cholesterol (LDL-C)	
	Very-low-density lipoprotein cholesterol (VLDL-C)	
	Gamma-glutamyl transpeptidase (GGT)	
	Estimated glomerular filtration rate (eGFR) using the MDRD-4 formula	

Appendix 2. Programming Specifications

Continuous data will be listed corresponding to the precision measured or calculated. Measures of central tendency will be presented using one additional decimal place than the precision of the data. Variability summaries will be presented using two additional significant digits relative to the precision of the underlying data. Quartiles, the minimum, and the maximum will be presented using the precision of the data.

All percentages are to be expressed as integers with one decimal place. The convention for rounding percentages is as follows:

- Values greater than or equal to x.x5% are rounded up
- Values between 0 and x.x5% are rounded down
- Values that are between 0 and .5% should be displayed as "<0.5%"

Appendix 3. Tables, Listings, and Figures

	Table of Contents	
Table 14.1.1	Patient Disposition	Randomized
Table 14.1.2	Demographic and Baseline Characteristics	ITT
Table 14.1.3	Baseline Characteristics	ITT
Table 14.1.4	Medical History	ITT
Table 14.1.5	Prior Medications	ITT
Table 14.1.6	Concomitant Medications	ITT
Table 14.2.1	Study Drug Exposure by Visit	ITT
Table 14.2.2.1	Peak Work	ITT
Table 14.2.2.2	Peak Work	ITT Subjects without Pes Cavus
Table 14.2.2.3	Peak Work	ITT Subjects with Pes Cavus
Table 14.2.2.4	Peak Work (Multiple Imputation)	ITT
Table 14.2.3.1	Modified FARS	ITT
Table 14.2.3.2	Modified FARS	ITT Subjects without Pes Cavus
Table 14.2.3.3	Modified FARS	ITT Subjects with Pes Cavus
Table 14.2.3.4	Modified FARS (Multiple Imputation)	ITT
Table 14.2.4.1	Oxygen Utilization	ITT
Table 14.2.4.2	Spirometry	ITT
Table 14.2.4.3	Other Exercise Peak Performance Measures	ITT
Table 14.2.4.4	Peak Exercise Ventilatory Response	ITT
Table 14.2.5	25-Foot Timed Walk Test	ITT
Table 14.2.6	Low Contrast Letter Visual Acuity Test	ITT
Table 14.2.7	9-Hole Peg Test	ITT
Table 14.2.8	Fatigue Severity Scale	ITT
Table 14.2.9	SF-36 Questionnaire - Version 2.0	ITT
Table 14.2.10	Full FARS	ITT
[redacted]	[redacted]	[redacted]

Table 14.3.1.1.1	Overall Summary of Treatment-Emergent Adverse Events	Safety
Table 14.3.1.1.2	Overall Summary of Treatment-Emergent Adverse Events	Safety without Pes Cavus
Table 14.3.1.1.3	Overall Summary of Treatment-Emergent Adverse Events	Safety with Pes Cavus
Table 14.3.1.2.1	Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety
Table 14.3.1.2.2	Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety without Pes Cavus
Table 14.3.1.2.3	Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	Safety with Pes Cavus
Table 14.3.1.3	Treatment-Emergent Adverse Events Possibly or Probably Related to Study Drug	Safety
Table 14.3.1.4	Treatment-Emergent Adverse Events by Maximum Severity	Safety
Table 14.3.1.5	Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term	Safety
Table 14.3.1.6	Treatment-Emergent Adverse Events Leading to Discontinuation of Study Drug	Safety
Table 14.3.2.1	Hematology Laboratory Results	Safety
Table 14.3.2.2	Shift in Hematology Laboratory Parameters	Safety
Table 14.3.2.3	Chemistry Laboratory Results	Safety
Table 14.3.2.4	Shift in Chemistry Laboratory Parameters	Safety
Table 14.3.2.6	Overall Summary of Liver Function Tests	Safety
Table 14.3.3	Vital Signs	Safety
Table 14.3.4	Echocardiogram Results	Safety
Table 14.3.5.1	12-Lead ECG Results	Safety
Table 14.3.5.2	12-Lead ECG Findings	Safety
Table 14.3.6	Physical Exam	Safety
[redacted]	[redacted]	[redacted]
Table 14.4.2	Frataxin Levels from Platelet Samples	Safety
[redacted]	[redacted]	[redacted]
Figure 14.2.2.1	Mean ± SE of Peak Work Over Time	ITT
Figure 14.2.2.2	Mean ± SE of Peak Work Over Time	ITT without pes cavus
Figure 14.2.2.3	Mean ± SE of Peak Work Over Time	ITT with pes cavus

Figure 14.2.3.1	Mean ± SE of Modified FARS Over Time	ITT
Figure 14.2.3.2	Mean ± SE of Modified FARS Over Time	ITT without pes cavus
Figure 14.2.3.3	Mean ± SE of Modified FARS Over Time	ITT with pes cavus
Figure 14.2.4.1	Mean ± SE Chemistry Laboratory Results	Safety
Listing 16.4.1	Patients Disposition	All Patients
Listing 16.4.2	Demographics and Baseline Characteristics	All Patients
Listing 16.4.3.1	Inclusion/Exclusion Criteria	All Randomized Patients
Listing 16.4.3.2	Listing of Screen Failures	Screen Failures
Listing 16.4.4	Protocol Deviations	All Patients
Listing 16.4.5	Pregnancy Test	All Patients
Listing 16.4.6	Medical History	All Patients
Listing 16.4.7	Prior and Concomitant Medications	All Patients
Listing 16.4.8	Study Drug Dosing Compliance	All Patients
Listing 16.4.9	Maximal Exercise	All Patients
Listing 16.4.10	Modified FARS	All Patients
Listing 16.4.11	Other Exercise Parameters	All Patients
Listing 16.4.12	25-Foot Timed Walk Test	All Patients
Listing 16.4.13	Low Contrast Letter Visual Acuity Test	All Patients
Listing 16.4.14	9-Hole Peg Test	All Patients
Listing 16.4.15	Fatigue Severity Scale	All Patients
Listing 16.4.16	SF-36 Questionnaire - Version 2.0	All Patients
Listing 16.4.17	Full FARS	All Patients
[redacted]	[redacted]	[redacted]
Listing 16.4.19.1	Adverse Events	All Patients
Listing 16.4.19.2	Serious Adverse Events	All Patients
Listing 16.4.19.3	Adverse Events Leading to Discontinuation of Treatment	All Patients
Listing 16.4.20.1	Lab - Hematology	All Patients

Listing 16.4.20.2	Lab - Chemistry	All Patients
Listing 16.4.20.3	Lab – Urinalysis	All Patients
Listing 16.4.21	Patients with Abnormal Liver Function Tests	All Patients
Listing 16.4.22	Vital Signs	All Patients
Listing 16.4.23	Echocardiogram	All Patients
Listing 16.4.24	12-Lead ECG Results	All Patients
Listing 16.4.25	Physical Exam	All Patients
[redacted]	[redacted]	[redacted]
Listing 16.4.27	Frataxin Levels from Platelet Samples	All Patients
[redacted]	[redacted]	[redacted]

STATISTICAL ANALYSIS PLAN
STUDY PART 2

VERSION: 2.0

DATE OF PLAN:

18-SEPTEMBER-2019

BASED ON:

Protocol Version 9.0 (August 17, 2017)

STUDY DRUG:

RTA 408, OMAVELOXOLONE

PROTOCOL NUMBER:

408-C-1402, Part 2

STUDY TITLE:

A Phase 2 Study of the Safety, Efficacy, and Pharmacodynamics of RTA 408 in the Treatment of Friedreich's Ataxia

SPONSOR:

Reata Pharmaceuticals, Inc.

2801 Gateway Drive, Suite 150

Irving, Texas 75063

(972) 865-2219

This study is being conducted in compliance with good clinical practice, including the archiving of essential documents.

SIGNATURE PAGE

This document has been prepared and/or reviewed by:

[redacted] MS
[redacted]
Reata Pharmaceuticals, Inc.

[redacted]
Date

[redacted]
Date

[redacted]
Reata Pharmaceuticals, Inc.

This document has been reviewed and accepted by:

[redacted] MD
[redacted]
Reata Pharmaceuticals, Inc.

[redacted]
Date

[redacted]
Date

[redacted]
Reata Pharmaceuticals, Inc.

REVISION HISTORY

Version	Revision Description	Date
2.0	• Specified that patients in the primary efficacy analysis population must have at least one post-baseline measurement • Changed analysis of the two key secondary endpoints to analysis of covariance (ANCOVA) and specified how to handle missing responses • Added tipping point and control-based multiple imputation as sensitivity analyses for the primary and key secondary endpoints to assess the impact of missing data • Updated power calculations to account for dropouts in sample size planning • Added "site" as a covariate for analysis of the primary and key secondary endpoints • Added analysis of efficacy using the "All Randomized" and "Pes Cavus" populations for descriptive purposes. Analyses of the "All Randomized" population includes a covariate for "pes cavus" [redacted] [redacted] • Clarifications throughout	17 September 2019
1.1	Corrected minor typographical errors throughout the document.	04 June 2019
1.0	New document.	16 May 2019

TECHNICAL SUMMARY REPORT (TSR)

Name of Sponsor/Company **Reata Pharmaceuticals**	**Individual Study Table Referring to Part of the Dossier:** **Volume:**	***(For National Authority Use Only)*:**
Name of Finished Product: Omaveloxolone capsules	**Page:**	
Name of Active Ingredient: Omaveloxolone		

Title of Study: A Phase 2 Study of the Safety, Efficacy, and Pharmacodynamics of RTA 408 in the Treatment of Friedreich's Ataxia

Investigators: [redacted] M.D.; [redacted] M.D; [redacted] M.D., PhD.; [redacted] M.D.; [redacted] M.D.; [redacted] M.D.; [redacted] M.D.; [redacted] M.D.; [redacted] M.D.; [redacted] M.D., Ph.D.; [redacted] M.D.
Study Center(s): 11

Studied period (years): 2	**Phase of development:** 2

Objectives:

Primary:

- To evaluate the change in the modified Friedreich's ataxia rating scale (mFARS) score at Week 48
- To evaluate the safety and tolerability of omaveloxolone

Key Secondary:

- To evaluate the Patient Global Impression of Change (PGIC) at Week 48
- To evaluate the Clinical Global Impression of Change (CGIC) at Week 48

Secondary:

- To evaluate the change in nine-hole peg test (9-HPT) at Week 48
- To evaluate the change in timed 25-foot timed walk test (T25-FWT) at Week 48
- To evaluate the frequency of falls
- To evaluate the change in peak work during maximal exercise testing at Week 48
- To evaluate the change in Activities of Daily Living (ADL) at Week 48

Exploratory:

- To evaluate the change in raters' assessments of videos of normal walking at Week 48
- To evaluate the change in SF-36 at Week 48
- To characterize the pharmacokinetics of omaveloxolone and potential metabolites after oral administration of omaveloxolone capsules

Methodology:

Part 2 of the MOXIe Phase 2 study is a randomized, placebo-controlled, double-blind, parallel-group study to evaluate the safety and efficacy of omaveloxolone 150 mg in patients with Friedreich's ataxia. Patients enrolled in Part 2 are randomized 1:1 to receive omaveloxolone 150 mg, or matching placebo. Randomization is stratified by pes cavus status (pes cavus vs. no pes cavus).

Patients with pes cavus have a musculoskeletal foot deformity and may represent a different subtype of FA, having different pathophysiology and clinical phenotype. Analysis of Part 1 data showed that treatment with omaveloxolone did not statistically improve studied endpoints (i.e., mFARS, exercise testing) in patients with pes cavus. Although the study part's small sample size limited its ability to detect a treatment effect, the presence of pes cavus also likely interferes with the ability to perform assessments that require standing or pedaling. Two of the four subsections that comprise the study's primary endpoint, mFARS, include assessments of lower limb coordination and upright stability. As a result, the primary analysis of efficacy is based on the stratum of patients enrolled without pes cavus. Because pes cavus is common in patients with FA, patients with pes cavus are also included in the study but will not comprise more than 20% of patients enrolled in Part 2. Additional secondary endpoints that are less likely to be affected by pes cavus are also assessed in the study to determine whether a therapeutic benefit can be detected in these patients. Efficacy endpoints are summarized descriptively for patients enrolled in the stratum with pes cavus. Safety is assessed for all patients enrolled in both strata.

Selection of the omaveloxolone dose of 150 mg for Part 2 was based on Data Safety Monitoring Board (DSMB) and Sponsor review of available data from Part 1, including safety, efficacy, and pharmacodynamic (PD) data. Following randomization on Day 1, patients self-administer study treatment once daily for 48 weeks. A follow-up visit for safety occurs at Week 52 (4 weeks after the last dose). The DSMB performs quarterly reviews of unblinded data for safety throughout Part 2.

Number of Subjects (planned):

Planned: 100 (with pes cavus n=20; without pes cavus n=80)

Diagnosis and main criteria for inclusion (see protocol section 8.1):

1. Have genetically confirmed Friedreich's ataxia
2. Have a mFARS score ≥ 20 and ≤ 80. The average of the two mFARS values collected at Screening and Day 1 visits must fall within the allowable range, and they must be within 4.5 points of each other
3. Be male or female and ≥ 16 years of age and ≤ 40 years of age
4. Have no changes to their exercise regimen within 30 days prior to Study Day 1 and be willing to remain on the same exercise regimen during the study period
5. Have the ability to complete maximal exercise testing
6. Have adequate kidney function defined as an estimated glomerular filtration rate (eGFR) ≥ 60 mL/min/1.73 m^2 using the Modification of Diet in Renal Disease (MDRD) 4 variable formula
7. Have a left ventricular ejection fraction $\geq 40\%$ (based on echocardiogram performed at Screening Visit or within 90 days prior to Screening Visit)
8. Be able to swallow capsules
9. Be willing and able to cooperate with all aspects of the protocol
10. Be willing to practice medically acceptable methods of birth control
11. Provide written informed consent for study participation, approved by the appropriate Institutional Review Board (IRB)

Test product, dose and mode of administration:

Capsules containing omaveloxolone at the 50-mg strength are used in this study. Treatment kits containing three bottles are provided to patients for self-administration. Patients are instructed to self-administer 3 capsules of study drug orally once daily in the morning on an empty stomach (approximately 1 hour before or 2 hours after eating).

Duration of treatment: 48 weeks

Reference therapy, dose and mode of administration: Matching capsules containing placebo are used in this study. Treatment kits containing three bottles are provided to patients for self-administration. Patients are instructed to self-administer 3 capsules orally once daily in the morning on an empty stomach (approximately 1 hour before or 2 hours after eating).
Criteria for evaluation (see protocol section 11.2, 12.0): *Efficacy*: mFARS; parameters collected during maximal exercise testing (including peak work), 9-HPT, T25-FWT, ADL, SF-36, PGIC, CGIC, fall diary, and videos of normal walking. *Safety*: Results of echocardiogram, electrocardiogram, vital sign measurements, weight, body mass index, physical examinations, adverse events, serious adverse events, concomitant medications, and laboratory test results (clinical chemistry, hematology, urinalysis, microscopy, and pregnancy tests [as indicated]).
Statistical methods: The primary analyses of efficacy in Part 2 of the study are based on the Full Analysis Set (FAS) that includes all patients enrolled without pes cavus. Analyses compare omaveloxolone patients to placebo patients. Mixed models repeated measures (MMRM) analysis is used to compare the difference in the change from baseline in the mFARS and peak work data at Week 48. The distribution of the Patient Global Impression of Change and Clinical Global Impression of Change is evaluated using analysis of covariance (ANCOVA) to compare treatment groups at Week 48. The body of this SAP describes the methods to be used for the other secondary and exploratory outcomes.

TABLE OF CONTENTS

TITLE PAGE 1
1. LIST OF ABBREVIATIONS 14
2. INTRODUCTION 17
3. STUDY OBJECTIVES AND ENDPOINTS 19
3.1. Study Objectives 19
3.1.1. Primary Objectives 19
3.1.2. Key Secondary Objectives 19
3.1.3. Secondary Objectives 19
3.1.4. Exploratory Objectives 19
4. STUDY DESIGN 20
4.1. Summary of Study Design 20
4.2. Definition of Study Drugs 20
4.3. Sample Size Considerations 20
4.3.1. Patients Without Pes Cavus 20
[redacted] 21
4.4. Randomization 21
4.5. Clinical Assessments 21
5. PLANNED ANALYSES 25
5.1. Final Analyses 25
6. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING 26
6.1. General Summary Table and Individual Subject Data Listing Considerations 26
6.2. Data Presentation Conventions 26
6.3. Analysis Populations 26
6.3.1. Full Analysis Set (FAS) 27
6.3.2. All Randomized Population (ARP) 27
6.3.3. Pes Cavus Population (PCP) 27
6.3.4. Safety Population 27
6.3.5. Per-Protocol (PP) Population 27
6.4. Baseline Definition 27
6.4.1. Efficacy Assessments 27
6.4.1.1. mFARS 27

6.4.1.2. Other Efficacy Assessments 28
6.4.2. Safety Assessments 28
6.5. Derived and Transformed Data 28
6.5.1. Baseline Age 28
6.5.2. Study Day 28
6.5.3. Change from Baseline 28
6.5.4. Visit Windows 29
6.5.4.1. Off-Treatment Visit Windows 29
6.5.5. Years Since FA Onset 30
6.5.6. GAA1 and GAA2 Repeat Length 30
6.5.7. Clinical Laboratory Results Outside the Limit of Quantification 30
6.5.8. Estimated Glomerular Filtration Rate 30
6.5.9. mFARS 31
6.5.10. 9-Hole Peg Test (9-HPT) 31
6.5.11. Timed 25-Foot Walk Test (T25-FWT) 32
6.5.12. Electrocardiogram Intervals 32
6.6. Handling of Missing Data 32
6.6.1. Missing Efficacy Endpoint Data 32
6.6.2. Missing Start and Stop Dates for Concomitant Medication 32
6.6.3. Missing Start and Stop Dates for Adverse Events 33
7. STUDY POPULATIONS 34
7.1. Subject Disposition 34
7.2. Screen Failures 34
7.3. Protocol Deviations 34
7.4. Demographic and Baseline Characteristics 34
7.5. Listing of Subject Inclusion and Exclusion Criteria 35
7.6. Medical History 35
8. PRIMARY EFFICACY 36
8.1. General Considerations 36
8.2. Statement of the Null and Alternate Hypotheses 36
8.3. Subgroup Analyses 36
8.4. Multiple Comparisons and Multiplicity 37
8.5. Primary Analysis of the Primary Efficacy Endpoint 37

8.5.1. Primary Efficacy Analysis 37
8.5.2. Sensitivity Analyses of the Primary Efficacy Results 38
8.5.2.1. Tipping Point 38
8.5.2.2. Treatment-Based Multiple Imputation 39
8.5.2.3. Control-Based Multiple Imputation 39
8.5.2.4. Per-Protocol Population Analysis 39
8.6. Primary Analysis of the Two Key Secondary Efficacy Endpoints 39
8.6.1. Patient Global Impression of Change 39
8.6.2. Clinical Global Impression of Change 40
8.6.3. Sensitivity Analyses of the Key Secondary Efficacy Results 40
8.6.3.1. Tipping Point 40
8.6.3.2. Control-Based Multiple Imputation 41
8.6.3.3. Per-Protocol Population Analysis 41
8.7. Primary Analysis of the Secondary Efficacy Endpoints 41
8.7.1. 9-Hole Peg Test (9-HPT) 41
8.7.2. Timed 25-Foot Walk Test (T25-FWT) 41
8.7.3. Frequency of Falls 42
8.7.4. Peak Work 42
8.7.5. Activities of Daily Living 42
8.8. Analysis of the Exploratory Efficacy Endpoints 43
8.8.1. Videos of Normal Walking 43
8.8.2. SF-36 Health Survey Update 43
9. EXPLORATORY EFFICACY 44
9.1. Exploratory Analyses of Efficacy Using the Full Analysis Set 44
9.1.1. mFARS 44
9.1.2. Patient Global Impression of Change 44
9.1.3. Clinical Global Impression of Change 44
9.1.4. 9-Hole Peg Test (9-HPT) 44
9.1.5. Timed 25-Foot Walk Test (T25-FWT) 45
9.1.6. Frequency of Falls 45
9.1.7. Maximal Exercise Test 45
9.1.8. Activities of Daily Living 46
9.2. Exploratory Analyses of Efficacy Using Other Analysis Populations 46

10. SAFETY AND TOLERABILITY 47
10.1. Adverse Event Preferred Term and Body/Organ System Summary Tables 47
10.1.1. Missing and Partial AE Onset Dates 48
10.1.2. Summaries of Adverse Events for Serious Adverse Events (SAE), Adverse Event Dropouts, and Death 48
10.2. Exposure and Compliance 48
10.3. Concomitant and Other Medications 49
10.3.1. Missing and Partial Concomitant and Other Medication Start and Stop Dates 49
10.4. Clinical Laboratory Evaluations 49
10.4.1. Summaries of Laboratory Results 50
10.4.2. Transaminases and Total Bilirubin 50
10.5. Vital Signs 51
10.6. Physical Exam 51
10.7. Echocardiogram 51
10.8. Electrocardiogram 51
10.9. Pregnancy 52
11. PHARMACOKINETICS 53
12. CHANGES FROM PROTOCOL-SPECIFIED ANALYSIS 54
13. REFERENCES 55
14. APPENDIX 56
14.1. Programming Specifications 56
14.2. Rationale for Programming Conventions 56
14.3. Mixed Model Repeated Measures Code 57
14.4. Code for Frequency of Falls 59
14.5. Treatment-based imputation and analysis of the primary endpoint 60
14.6. Control-based imputation and analysis of the primary endpoint 64
14.7. Tipping point analysis of primary endpoint 68
14.8. ANCOVA analysis of the secondary endpoint 77
14.9. Missing data approaches for secondary endpoint analysis 78
14.10. Tipping point analysis of the secondary endpoint 86
14.11. Random seed specification 93
14.12. Other Maximal Exercise Testing Parameters 95
14.13. List of Laboratory Tests 95

14.14. SF-36 Scoring 97

LIST OF TABLES

Table 1: List of Abbreviations 14
Table 2: Schedule of Assessments 22
Table 3: Analysis Visits 29
Table 4: Analysis Visits for Post-Treatment Safety Analysis 30
Table 5: Scoring the FARS Neurologic Exam 31
Table 6: Pre-Specified Threshold Levels for ECG Parameters 51
Table 7: Parameter Names Used in Analysis 56
Table 8: SAS code for MMRM analysis 57
Table 9: SAS code for MMRM analysis: All Randomized Population 57
Table 10: SAS code for treatment-based imputation of the primary endpoint 61
Table 11: SAS code for control-based imputation of the primary endpoint 65
Table 12: SAS code to create datasets satisfying monotone missing condition for continuous measures 69
Table 13: SAS code for tipping point analysis 71
Table 14: SAS code for MMRM analysis 74
Table 15: SAS code to obtain the overall estimate and standard error 75
Table 16: SAS code for PGIC/GCIC ANCOVA 78
Table 17: SAS code for PGIC/GCIC ANCOVA in the All Randomized Population 78
Table 18: SAS code to generate imputation of missing data 79
Table 19: SAS code to generate imputation of missing data in the All Randomized Population 80
Table 20: SAS code for multiple imputation of the secondary endpoint 81
Table 21: SAS code for multiple imputation of the secondary endpoint in the All Randomized Population 84
Table 22: SAS code to generate imputation of missing data in placebo subjects 86
Table 23: SAS Code for Tipping Point Analysis 88
Table 24: SAS code for analysis of the PGIC/CGIC tipping point analysis 91
Table 25: SAS code for random seed generation 93
Table 26: Pre-Specified Seed Values and Analyses 94
Table 27: Summary of Other Maximal Exercise Testing Parameters 95
Table 28: List of Laboratory Tests 96
Table 29: SF-36® final item values 97

Table 30: SF-36® raw scale scores....98
Table 31: SF-36® summary scores....100

1. LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Term
μ_{OMAV}	Omaveloxolone average
$\mu_{PLACEBO}$	Placebo average
T25-FWT	25-Foot Timed Walk Test
9-HPT	9-hole peg test
ADL	Activities of Daily Living
AE	Adverse Event
ALT	alanine aminotransferase
ALP	Alkaline phosphatase
ANCOVA	Analysis of covariance
ARP	All randomized population
AST	aspartate aminotransferase
ATC	Anatomical/Therapeutic/Chemical
BMI	body mass index
BNP	B-type natriuretic peptide
CGIC	Clinical Global Impression of Change
CI	Confidence interval
CK	creatinine kinase
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
cm	Centimeters
CPK	Creatine phosphokinase
CRF	Case Report Form
CSR	Clinical Study Report
ECG	Electrocardiogram
DSMB	Data Safety Monitoring Board
eGFR	estimated glomerular filtration rate
f	truncation fraction
FA	Friedreich's Ataxia
FARS	Friedreich's Ataxia Rating Scale
FAS	Full analysis set
FDA	Food and Drug Administration

GGT	Gamma-Glutamyl Transferase
HDL-C	High-density lipoprotein cholesterol
HIV	Human immunodeficiency virus
HLT	high level term
ICH E9	International Conference on Harmonisation Tripartite Guideline for Good Clinical Practice E9
IRB	Institutional Review Board
IWRS	Interactive Web Response System
kg	Kilogram
LDH	Lactate dehydrogenase
LDL-C	Low-density lipoprotein cholesterol
LLD	Lower Limit of Detection
LS	least squares
LVMI	left ventricular mass index
MAR	missing at random
MCH	Mean corpuscular hemoglobin
MCHC	Mean corpuscular hemoglobin concentration
MCV	Mean corpuscular volume
MedDRA	Medical Dictionary for Regulatory Activities Terminology
MMRM	mixed model repeated measures
MNAR	Missing not at random
mFARS	modified Friedreich's Ataxia Rating Scale
MDRD	Modification of Diet in Renal Disease
NT-pro-BNP	N-terminal prohormone of B-type natriuretic peptide
PCP	Pes Cavus population
PD	Pharmacodynamic
PGIC	Patient Global Impression of Change
pH	potential of hydrogen
PK	Pharmacokinetic
PT	Preferred term
Q1	First quartile
Q3	Third quartile
QTc	corrected QT interval

QTcF	Fridericia corrected QT interval
RBC	Red blood cell
RTA 408	Omaveloxolone
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SD	Standard deviation
SE	Standard error
SF-36	SF-36® Health Survey Update
SOC	System Organ Class
TBL	Total bilirubin
TEAE	treatment emergent adverse events
TLF	Tables, listings, and figures
ULD	Upper Limit of Detection
ULN	Upper Limit of Normal
VLDL-C	Very-low-density lipoprotein cholesterol
WBC	White blood cell
WHO	World Health Organization

2. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the planned analyses and data displays to be included in the Clinical Study Report (CSR) for Protocol 408-C-1402, Part 2.

Protocol Revision Chronology:		
Protocol version 1	23-July-2014	Original
Amendment 1 (Protocol version 2)	11-Sept-2014	• The primary endpoint to evaluate peak workload (watts) was changed to evaluate peak work (watts/kg) • Part 1 study design was modified based on FDA feedback. The design was changed to a randomized, placebo-controlled, double-blind, dose-escalation study to evaluate the safety of omaveloxolone at 2.5 mg, 5 mg, and 10 mg
Amendment 2 (Protocol version 3)	13-Oct-2014	• Per FDA recommendation, PK characterization was added as an objective of the study
Amendment 3 (Protocol version 4)	12-May-2015	• Part 1 of the study was modified to allow for up to 2 additional cohorts of 8 patients each
Amendment 4 (Protocol version 5)	29-Sept-2015	• Part 1 of the study was modified to include 8 cohorts for dose ranging analysis
Amendment 5 (Protocol version 6)	23-May-2016	• Part 1 of the study included 9 cohorts and enrollment was increased to 108 patients. • The maximum permitted dose of omaveloxolone was raised to 300 mg
Amendment 6 (Protocol version 7)	13-Apr-2017	• Part 2 enrollment was increased to 100 patients • Part 2 study design was modified based on FDA feedback
Amendment 7 (Protocol version 8)	29-July-2017	• Patient Global Impression of Change and Clinical Global Impression of Change were added as secondary endpoints
Amendment 8 (Protocol version 9)	17-Aug-2017	• Part 2 study treatment duration was extended to 48 weeks

This SAP was developed in accordance with ICH E9 guidelines. All decisions regarding final analysis, as defined in this SAP document, were made prior to Database Lock (unblinding) of the study data. Further information can be found in the protocol. Changes to protocol specified analyses, including analysis populations, can be found in Section 12.

The statistical analysis plan (SAP) is based on:

- Protocol No. 408-C-1402, Version 9, dated August 17, 2017
- ICH guidelines E4 and E9 (Statistical Principles for Clinical Trials)
- Feedback received from FDA throughout clinical development, including the following communications:
  - FDA Meeting Minutes [redacted]
  - FDA teleconference with Reata on [redacted]
  - FDA Study May Proceed letter [redacted]
  - Email communication from the Division on [redacted]

This SAP describes the study populations, how variables are derived, how missing data are handled, and details concerning the statistical methods to be used to analyze the safety and efficacy data in Part 2 of study 408-C-1402. Should the SAP and the protocol be inconsistent with respect to the planned analyses, the language of the SAP is governing.

The SAP is finalized, approved by the Sponsor, and placed on file before the database is locked. This version of the SAP describes the analyses planned prior to the database lock. Unless otherwise specified, these analyses are summarized in the clinical study report (CSR). Any substantive changes made to the SAP after the database lock are clearly identified, and any analyses in addition to those specified in the SAP prior to the database lock are considered ad hoc. The CSR will describe any deviations from the planned analyses.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1. Study Objectives

In patients with Friedreich's ataxia, the study will compare those receiving omaveloxolone versus those receiving placebo with respect to several objectives.

3.1.1. Primary Objectives

- To evaluate the change in the modified Friedreich's ataxia rating scale (mFARS) score at Week 48
- To evaluate the safety and tolerability of omaveloxolone

3.1.2. Key Secondary Objectives

- To evaluate the Patient Global Impression of Change (PGIC) at Week 48
- To evaluate the Clinical Global Impression of Change (CGIC) at Week 48

3.1.3. Secondary Objectives

- To evaluate the change in performance on a 9-hole peg test (9-HPT) at Week 48
- To evaluate the change in performance on a timed 25-foot timed walk test (T25-FWT) at Week 48
- To evaluate the frequency of falls over 48 weeks
- To evaluate the change in peak work during maximal exercise testing at Week 48
- To evaluate the change in the Activities of Daily Living (ADL) score at Week 48

3.1.4. Exploratory Objectives

- To evaluate the change in raters' assessments of videos of normal walking at Week 48
- To evaluate the change in SF-36® Health Survey Update (SF-36) score at Week 48
- To characterize the pharmacokinetics of omaveloxolone and potential metabolites after oral administration of omaveloxolone capsules

4. STUDY DESIGN

4.1. Summary of Study Design

Part 2 of the 408-C-1402 study is a randomized, placebo-controlled, double-blind, parallel-group study to evaluate the safety and efficacy of omaveloxolone 150 mg in patients with Friedreich's ataxia. Patients enrolled in Part 2 are randomized 1:1 to receive omaveloxolone 150 mg, or placebo. Randomization is stratified by pes cavus status: pes cavus and no pes cavus.

Patients with pes cavus have a musculoskeletal foot deformity and may represent a different subtype of FA, having different pathophysiology and clinical phenotype. Analysis of Part 1 data showed that treatment with omaveloxolone did not statistically improve studied endpoints (i.e., mFARs, exercise testing) in patients with pes cavus. Although the small sample size in Part 1 of the study limited the ability to detect a treatment effect, the presence of pes cavus also likely interferes with the ability to perform assessments that require standing or pedaling. Two of the four subsections that comprise the study's primary endpoint, mFARS, include assessments of lower limb coordination and upright stability. As a result, the primary analysis population for Part 2 efficacy is based on the stratum of patients enrolled without pes cavus (Section 6.3.1). Because pes cavus is common in patients with FA, patients with pes cavus are also included in the study but do not comprise more than 20% of patients enrolled in Part 2, and randomization was stratified by pes cavus status (with pes cavus vs without pes cavus). Secondary endpoints that are less likely to be affected by pes cavus are also assessed in the study to determine whether a therapeutic benefit can be detected in patients with pes cavus. All efficacy endpoints are summarized descriptively for patients enrolled in the stratum with pes cavus. Safety is assessed for all patients enrolled in both strata.

Selection of the omaveloxolone dose of 150 mg for Part 2 was based on Data Safety Monitoring Board (DSMB) and Sponsor review of available data from Part 1, including safety, efficacy, and pharmacodynamic (PD) data. Following randomization on Day 1, patients self-administer study treatment once daily for 48 weeks. A follow-up visit for safety occurs at Week 52 (4 weeks after the last dose). The DSMB performs quarterly reviews of unblinded data for safety throughout Part 2.

4.2. Definition of Study Drugs

Capsules containing omaveloxolone at the 50 mg strength, or the corresponding placebos, are used in this study.

4.3. Sample Size Considerations

4.3.1. Patients Without Pes Cavus

With 80 patients in the stratum without pes cavus, the Part 2 portion of the study has approximately 85% power to test the difference between the two treatment groups with respect to change from baseline in mFARS at Week 48. Th [redacted]

- [redacted]
- Two-sided Type I error rate of 0.05
- A difference of 2.0 points between the change in mFARS in the omaveloxolone and placebo groups
- Standard deviation of change in mFARS of 3.5 points
- [redacted]

[redacted]

4.4. Randomization

Patients enrolled in Part 2 are randomized 1:1 to either omaveloxolone or placebo. Randomization is stratified by pes cavus status (pes cavus vs. no pes cavus). Patients in the pes cavus stratum do not comprise more than 20% of patients enrolled in MOXIe Part 2. Randomization is generated using a centralized IWRS.

4.5. Clinical Assessments

All patients in Part 2 of the study follow the same visit and assessment schedule. Table 2 lists the overall schedule of assessments for the study. Following the first dose of study drug on Day 1, patients are scheduled for in-person visits to be assessed during treatment at Weeks 2, 4, 12, 18, 24, 36, and 48, as well as by telephone contact on Days 7, 56, 210, and 294. Patients are assessed at an in-person follow-up visit for safety at Week 52 (4 weeks after last dose).

Table 2: Schedule of Assessments

Visit Number	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8	Visit 9	Visit 10	Visit 11	Visit 12	Visit 13	Visit 14
Study Day/Week	Screening	Day 1	Week 1 (telephone)	Week 2	Week 4	Week 8 (telephone)	Week 12	Week 18	Week 24	Week 30 (telephone)	Week 36	Week 42 (telephone)	Week 48 End of Treatment	Week 52[a] End of Study/ 4-week follow-up
Day Relative to First Dose	-60 to -1[b]	1[c]	7 (±3 days)	14 (±3 days)	28 (±3 days)	56 (±3 days)	84 (±3 days)	126 (±3 days)	168 (±3 days)	210 (±3 days)	252 (±3 days)	294 (±3 days)	336 (±3 days)	192 (±3 days)
Informed consent	X													
Inclusion/Exclusion criteria assessment	X	X												
Demographics and baseline disease characteristics	X													
Pes cavus assessment	X													
Foot X-ray	X[d]													
Prior and concomitant medication assessment	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Medical history	X													
Height	X													
Echocardiogram	X[e]								X				X	
Electrocardiogram	X[e]			X	X		X	X	X		X		X	X
Vital sign measurements	X	X		X	X		X	X	X		X		X	X
Weight and BMI	X	X		X	X		X	X	X		X		X	X
Physical examination	X								X				X	X
Adverse event collection		X	X	X	X	X	X	X	X	X	X	X	X	X
Clinical chemistry	X[f]	X		X	X		X	X	X		X		X	X
Hematology	X[f]	X		X	X		X	X	X		X		X	X
Urinalysis and microscopy	X[f]	X		X	X		X	X	X		X		X	X
BNP and NT-proBNP[g]	X[f]	X		X	X		X	X	X		X		X	X
Hepatitis B and C and HIV[h]	X[f,h]													
Pregnancy test WOCBP[i]	X[f,i]	X		X	X		X	X	X		X		X	X
Exercise regimen reporting	X	X												
Randomization		X												
Study drug dispensation		X			X		X		X		X			
Study drug return and pill count / diary[j]					X		X		X		X		X	
Study drug administration[k]		←					X						→	

Visit Number	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8	Visit 9	Visit 10	Visit 11	Visit 12	Visit 13	Visit 14
Study Day/Week	Screening	Day 1	Week 1 (telephone)	Week 2	Week 4	Week 8 (telephone)	Week 12	Week 18	Week 24	Week 30 (telephone)	Week 36	Week 42 (telephone)	Week 48 End of Treatment	Week 52[a] End of Study/ 4-week follow-up
Day Relative to First Dose	-60 to -1[b]	1[c]	7 (±3 days)	14 (±3 days)	28 (±3 days)	56 (±3 days)	84 (±3 days)	126 (±3 days)	168 (±3 days)	210 (±3 days)	252 (±3 days)	294 (±3 days)	336 (±3 days)	192 (±3 days)
Maximal exercise test[l]	X	X			X		X	X	X		X		X	
Neurologic FARS (practice)	X[m]													
Neurologic FARS[n]	X[o]	X[o]			X		X	X	X		X		X	
9-hole peg test	X	X							X				X	
25-foot timed walk test	X	X[p]							X[p]				X[p]	
Video Collection of Normal Walking [q]		X							X				X	
SF-36 Health Survey Update		X							X				X	
Activities of Daily Living	X								X		X		X	
Patient Global Impression of Change							X[r]		X[r]		X[r]		X[r]	
Clinical Global Impression of Change							X[r]		X[r]		X[r]		X[r]	
Falls Diary	X[s]	←					X						→	
PK analysis				X[t]			X[u]		X[u]				X[u]	

[a] These procedures should also be performed in the event of early termination.
[b] Study Day -1 is the day prior to first dose of study drug.
[c] All Day 1 procedures should be performed prior to administration of first dose of study drug.
[d] Only 1 x-ray is collected during the study. Only x-rays of the right foot are completed for the foot x-ray assessment. Patients should complete the x-ray foot assessment at Screening; however, if the patient was not able to complete the x-ray foot assessment at Screening, it may be completed during any visit while the patient is participating in the study.
[e] For patients with echocardiograms and electrocardiograms collected within 90 days prior to the Screening Visit, the most recent echocardiogram and electrocardiogram can be used to assess patient eligibility.
[f] A home health nurse visit may be used to collect all lab samples required at the Screening Visit.
[g] Patients must be allowed to rest for a minimum period of 1 hour following maximal exercise test before this blood sample is collected (at the same time as all central lab blood draws). This sample should be taken with the patient in the same position (e.g., sitting or semi-recumbent) at all appropriate visits.
[h] Blood samples should be collected for hepatitis B and C and HIV antibodies only in patients lacking evidence of a negative titer in the past year.
[i] Negative serum pregnancy test results are required at the Screening Visit before study enrollment, and negative urine pregnancy test results are required at all other times indicated for continued participation in the study.
[j] A dosing diary check must be performed at study drug return and pill count.
[k] Study drug should be administered in the presence of study staff in the clinic on Day 1 after all Day 1 assessments have been completed. Study drug should also be administered in the clinic on Day 14 (Visit 4), Day 84 (Visit 7), Day 168 (Visit 9), and Day 336 (Visit 13) after the blood collection for predose PK analysis. All other doses can be administered at home. Study drug should be administered once daily through Day 336 (Visit 13).

[l] On study days where multiple assessments are to be completed, the maximal exercise test is the first functional assessment performed.
[m] The practice FARS assessment is not required for patients having completed a FARS assessment within 90 days prior to Screening. The practice FARS assessment must be performed prior to the screening maximal exercise test.
[n] Other than the practice FARS, all FARS assessments must always be performed after the maximal exercise test.
[o] Both Screening and Day 1 mFARS assessments are required to determine patient eligibility.
[p] Video of the 25-foot timed walk test is optional and is collected only for those patients who provide consent to video.
[q] Patients who consent to video collection of the 25-foot timed walk test also has video collection during normal walking with a 25-foot (un-timed) walk using his/her normal gait. The normal walk must be performed after the 25-foot timed walk test.
[r] Patients and investigators must complete the patient global impression of change and clinical global impression of change following completion of the neurological FARS exam.
[s] Patients are provided the fall diary at Screening and must record fall incidents between Screening and Week 48.
[t] Blood samples for PK analysis at Day 14 (Visit 4) should be collected prior to study drug administration as well as 1, 2, 4, and 8 hours after study drug administration.
[u] A single blood sample for PK analysis at Day 84 (Visit 7), Day 168 (Visit 9), and Day 336 (Visit 13) should be collected prior to study drug administration.

Abbreviations: BMI=body mass index; BNP= B-type natriuretic peptide; FARS=Friedreich's ataxia rating scale; HIV=human immunodeficiency virus; NT-proBNP= N-terminal prohormone of B-type natriuretic peptide; PK=pharmacokinetic; WOCBP=women of childbearing potential.

5. PLANNED ANALYSES

5.1. Final Analyses

The final analyses of efficacy are based on locked data and performed after all enrolled patients have completed all efficacy assessments (i.e., Week 48). The final analyses of safety are based on locked data and performed after all enrolled patients have completed all safety assessments (i.e., Week 52). The database lock plan describes details of the database lock.

6. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING

The efficacy and safety analyses use the analysis sets defined in Section 6.3. Patient listings of all analysis data that support summary tables and figures are provided along with their source data. Measurements from patients excluded from the pre-defined analysis sets or extra measurements (such as unscheduled or repeat assessments) are not included in summary tables unless otherwise specified, but they are included in the patient listings. Missing data are not imputed, unless otherwise specified. In general, patient listings are sorted by patient number and assessment date (time and parameter, as applicable).

6.1. General Summary Table and Individual Subject Data Listing Considerations

Results of statistical analyses are reported using summary tables, listings, and figures (TLFs). All TLFs will use ICH numbering conventions. The following conventions are used:

- Unless otherwise noted, all statistical testing is two-sided and is performed at the 0.05 significance level.
- Tests are declared statistically significant if the calculated p-value is <0.05.

All analyses and summaries are produced using SAS® version 9.3 (or higher).

6.2. Data Presentation Conventions

Unless otherwise specified, descriptive statistics for continuous variables include the number of patients with data (N), mean, standard deviation (SD), quartiles (i.e., median, 25th, and 75th percentiles), minimum, and maximum. The same number of decimal places as in the observed value are presented when reporting minimum and maximum; 1 more decimal place than in the observed value is generally presented when reporting mean and quartiles; and 2 more decimal places than in the observed value are presented when reporting SD. When calculating the minimum and maximum of average baseline values, the same number of decimal places as in the observed value will be presented when reporting minimum and maximum. For imputed values and the average of baseline values, the average values will be rounded to the same number of decimal places as the source value.

Categorical (qualitative) data are presented using frequency counts and percentages. All percentages are rounded to 1 decimal place, unless otherwise specified. Percentages equal to 100 are presented as 100% and no percentages are presented for zero frequencies. Where individual variable values are missing, summaries of categorical data are based on reduced denominators (i.e., only patients with available data are included in the denominators) and the number of missing values is presented. For summaries of AEs and concomitant medications (CM), the percentages are based on the number of patients who received study drug.

6.3. Analysis Populations

Analysis populations defined in this section pertain to patients enrolled in Part 2 of the trial.

6.3.1. Full Analysis Set (FAS)

The primary analysis of efficacy is based on patients without pes cavus. The Full Analysis Set (FAS) includes patients enrolled without pes cavus who have at least one post-baseline measurement, categorized by their randomized treatment group (whether or not they received study drug).

6.3.2. All Randomized Population (ARP)

The All Randomized Population (ARP) includes all patients randomized, categorized by their randomized treatment group (whether or not they received study drug). Only descriptive analyses of efficacy are performed using the ARP.

6.3.3. Pes Cavus Population (PCP)

The Pes Cavus Population (PCP) includes all patients in the pes cavus stratum categorized by their randomized treatment group (whether or not they received study drug). Only descriptive analyses of efficacy will be performed using the PCP.

6.3.4. Safety Population

Safety analyses are based on all enrolled patients. The safety population includes all patients who received at least 1 dose of randomized study drug. The safety population is used for evaluation of safety variables. Patients who receive at least one dose of omaveloxolone are classified in the omaveloxolone group. Patients who receive at least one dose of placebo and no dose of omaveloxolone are classified in the placebo group.

6.3.5. Per-Protocol (PP) Population

A sensitivity analysis exploring the robustness of the primary FAS findings is based on the patients in the per-protocol population. The per-protocol population is defined as patients without pes cavus who:

- Received study drug through Week 48; and
- Had no major protocol deviation that could potentially affect the efficacy assessments.

6.4. Baseline Definition

Baseline values are defined as the last non-missing assessment prior to the first study drug administration, unless otherwise specified below. If the first study drug administration occurs after the date of randomization, the last measurement prior to the first study drug administration is considered the Day 1 measurement for the calculation of baseline.

6.4.1. Efficacy Assessments

6.4.1.1. mFARS

The protocol specifies that the Screening and Day 1 mFARS (Section 6.5.9) must be within 4.5 points to meet inclusion criteria for enrollment. If mFARS scores for Screening and Day 1 exceed the protocol allowed difference of 4.5 points for enrollment by at least two-fold (i.e., at

least 9 points different), then the Screening mFARS assessment is considered invalid, and the mFARS collected closest to randomization (i.e., Day 1) is used as the baseline.

Otherwise, the mean of Screening and Day 1 assessments is used as the baseline mFARS.

6.4.1.2. Other Efficacy Assessments

The mean of Screening and Day 1 assessments is used as baseline for continuous efficacy parameters: peak work, T25-FWT, 9HPT, and ADL.

6.4.2. Safety Assessments

Baseline for continuous safety assessments (i.e., vital sign assessments, weight, BMI, and laboratory measurements) is defined as the average value of measurements collected up through, but prior to, first study drug administration.

6.5. Derived and Transformed Data

6.5.1. Baseline Age

Subject's age in years is defined as the age at consent (Screen A).

- Age (year) = Floor ((date of consent – date of birth)/365.25)

6.5.2. Study Day

Study day is the day relative to the date of randomization. Day 1 is defined as the date of randomization, unless otherwise specified for the calculation of baseline (Section 6.4).

Assessments that occur after randomization but before the first study drug administration are considered to occur on study Day 1. Assessments collected on the same date as the first date of study drug administration will be considered to occur before the first dose of study drug administration.

For visits (or events) after randomization, day is calculated as:

- Study day = visit (or event) date - date of randomization + 1

For visits (or events) before randomization, day is calculated as:

- Study day = visit (or event) date - date of randomization

For listings (such as for adverse events) the quantity 'days since first (or last dose)' is defined as:

- days since first (or last dose) = event date – date of first (or last dose) + 1

For summaries that present distribution of time expressed in weeks and months, weeks will be defined as days divided by seven and months as days divided by 30.4.

6.5.3. Change from Baseline

Change from baseline is calculated using the baseline value (Section 6.4) and the value closest to the target study day, using the rules defined in Section 6.5.4.

6.5.4. Visit Windows

Because clinical visits may occur outside protocol-specified windows, instead of relying on visit labels in the clinical database, analysis visits and their windows are defined using derived study day (Section 6.5.2). Study day is calculated using the actual date for each scheduled and unscheduled assessment and compared to the target study day for each analysis visit. Data analysis and summaries are based on the collection date that is closest to the protocol scheduled target study day (Table 3).

Table 3: Analysis Visits

Analysis Visit	Label	Target Study Day	Analysis Window
2	Week 2	14	$7 \leq$ Study Day ≤ 21
4	Week 4	28	$22 \leq$ Study Day ≤ 55
12	Week 12	84	$56 \leq$ Study Day ≤ 104
18	Week 18	126	$105 \leq$ Study Day ≤ 146
24	Week 24	168	$147 \leq$ Study Day ≤ 210
36	Week 36	252	$211 \leq$ Study Day ≤ 294
48	Week 48	336	$295 \leq$ Study Day ≤ 350

The safety follow-up (Week 52) is based on days since last dose. If more than one assessment exists during the 4-week follow-up after last dose, the one closest to 28 days following the date of the last study drug administration is used for analysis and summary.

If a parameter is assessed or measured more than once within a visit window, the one that is closest to the protocol-scheduled time point (or target study day) is used for the purposes of data analysis and summary. If two assessments are equidistant from a target study day, the earlier assessment is used. If the visit used for analysis includes two assessments on the same day, the average of the two measurements will be used.

Records from visits not closest to the target study day, and therefore not used in analyses, are presented in by-subject data listings.

6.5.4.1. Off-Treatment Visit Windows

Off-treatment values for clinical laboratory evaluations (Section 10.4), vital signs (Section 10.5), and electrocardiogram (Section 10.8) will be summarized based on the date of last dose. 'Off-treatment' measurements are those that occur after the last dose date. The last assessment while on study drug is defined as the on-treatment result with the latest date. Table 4 indicates the windows used to attribute results to post-treatment timepoints.

Table 4: Analysis Visits for Post-Treatment Safety Analysis

Analysis Visit	Label	Target Study Day (Reference Day is Days After Last Dose)	Analysis Window
40	4-weeks – Off Treatment	28	15 ≤ Study Day ≤ 43
80	8-weeks – Off Treatment	56	44 ≤ Study Day ≤ 71
120	12-weeks– Off Treatment	84	72 ≤ Study Day ≤ 99
240	24-weeks– Off Treatment	168	100 ≤ Study Day ≤ 252
480	48-weeks– Off Treatment	336	253 ≤ Study Day ≤ 350

6.5.5. Years Since FA Onset

The years since FA onset are calculated as the difference between Baseline Age (Section 6.5.1) and Age at Onset.

6.5.6. GAA1 and GAA2 Repeat Length

Where GAA1 and GAA2 repeat length are entered as ranges, the value at the lowest end of the range is used. All GAA1 and GAA2 values are listed, including the range of GAA1 and GAA2 values when entered in the CRF.

6.5.7. Clinical Laboratory Results Outside the Limit of Quantification

Laboratory results reported as less than the lower limit of detection (i.e., <LLD) are imputed as LLD/2. Laboratory results reported as greater than the upper limit of detection (i.e., >ULD) are imputed as the ULD.

6.5.8. Estimated Glomerular Filtration Rate

The estimated glomerular filtration rate (eGFR) is calculated using the formula specified below.

For patients consented at age 18 years and older, the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation is used:

- $eGFR\ (mL/min/1.73\ m^2) = 141 \times \min(S_{cr}/\kappa, 1)^{\alpha} \times \max(S_{cr}/\kappa, 1)^{-1.209} \times 0.993^{Age} \times 1.018$ [if female] × 1.159 [if black]

For patients consented at <18 years old, the Bedside Schwartz equation is used:

- $eGFR\ (mL/min/1.73\ m^2) = (0.41 \times \text{Height in cm}) / S_{cr}$

Where S_{cr} is serum creatinine (mg/dL), κ is 0.7 for females or 0.9 for males, and α is -0.329 for females or -0.411 for males. Min indicates the minimum of S_{cr}/κ and 1 and max indicates the maximum of S_{cr}/κ and 1. Age and height at screening are used for eGFR calculations.

6.5.9. mFARS

The Friedreich Ataxia Rating Scale (FARS) is a neurological-exam-based rating scale with five sections: Bulbar (section A), Upper Limb Coordination (section B), Lower Limb Coordination (section C), Peripheral Nervous System (section D), and Upright Stability (section E). Assessments in each section are added according to the logic in Table 5 as specified in the FARS manual (Lynch 2006). mFARS is the sum of sections A, B, C, and E.

Table 5: Scoring the FARS Neurologic Exam

Component Score of Neurologic FARS	Logic	Handling of Missing Data
A	Sum of all of the section A responses	If question is left blank but FARS assessment was completed, use question response from last visit
B	Sum of all of the section B responses	
C	Sum of all of the section C responses	
D	Sum of all of the section D responses, except 5a and 5b. For Question 5, only the results of Questions 5c1 and 5c2 is included in the section D sum.	
E	Average for items 2-5 item in the database; Reports sums for questions 1, 6, and 7 and sums of averages for questions 2-5	If questions 1-5 are left blank but FARS assessment was completed, take average of existing responses (e.g., average of 3a.1 and 3a.2). If no existing responses exist at the current visit for questions 1-5, use the question response from the last visit. For questions 6-7, use the response from the last visit.

6.5.10. 9-Hole Peg Test (9-HPT)

Both the dominant and non-dominant hands are tested twice at each assessment. The times recorded for the two trials for each hand are averaged, and the average value for each hand is used in analyses.

The average time value for each hand is transformed for analysis to allow for a more normal distribution of the data using the following methods:

- The number of pegs per second (pegs/second) for each hand is calculated based on nine pegs placed compared to the time needed to complete the test: 9 / (average time)
- The reciprocal of the average time is calculated as 1 / (average time) (Lynch 2005).

6.5.11. Timed 25-Foot Walk Test (T25-FWT)

Patients are instructed to attempt two T25-FWT trials at each visit. If both walk trials are completed, walk times are averaged for use in analyses. Otherwise, only the completed walk from each visit is used for analyses. The walk time is transformed for analysis to allow for a more normal distribution of the data using the following methods:

- The number of feet per second (feet/second) for each walk is calculated based on the distance (25 feet) and the total time needed to complete the test: 25 / (average time)
- The reciprocal of the walk time is calculated as 1 / (average time) (Lynch 2005).

6.5.12. Electrocardiogram Intervals

Electrocardiogram QTcF intervals are calculated from QTc and RR intervals using the following formula:

$$QTcF = QT/\sqrt[3]{RR}$$

where RR = 60 / (Heart Rate).

6.6. Handling of Missing Data

6.6.1. Missing Efficacy Endpoint Data

The primary analysis of efficacy is based on an assumption of missing at random (MAR). Missing mFARS data are not imputed for the MMRM analysis of the primary endpoint. Missing PGIC and CGIC data are imputed for the primary ANCOVA analysis of the key secondary endpoints using multiple imputation based on an assumption of MAR. A set of sensitivity analyses is included to assess the robustness of conclusions to the MAR assumption (see Section 8.5.2 and Section 8.6.3) for the analysis of the primary and key secondary endpoints.

6.6.2. Missing Start and Stop Dates for Concomitant Medication

Missing start dates for concomitant medications are not imputed.

Concomitant medications with incomplete end dates are considered concomitant medications if:

- Day and month are missing, and the year is equal to or after the year of the first date of study drug administration;
- Day is missing and the year is after the year of the first date of study drug administration;
- Day is missing and the year is equal to the year of the first date of study drug administration and the month is equal to or after the month of the first date of study drug administration; or
- Year is missing.

6.6.3. Missing Start and Stop Dates for Adverse Events

Treatment-emergent adverse events (TEAEs) are events that either:

- Had a date of onset on or after the date of the date of first dose of study drug and not more than 30 days after the date of the last dose of study drug, or
- Had no recorded date of onset with a stop date after the first dose of study drug, or
- Had no recorded date of onset or stop date.

Adverse events with incomplete start dates are considered after the date of first dose if:

- Day and month are missing, and the year is equal to or after the year of the first date of study drug administration;
- Day is missing and the year is after the year of the first date of study drug administration;
- Day is missing and the year is equal to the year of the first date of study drug administration and the month is equal to or after the month of the first date of study drug administration; or
- Year is missing.

Adverse events with incomplete start dates are considered on or within 30 days of last dose, if:

- Day and month are missing, and the year is equal to or before the year of the date of last dose of study drug plus 30 days;
- Day is missing and the year is equal to or before the year of the date of last dose of study drug plus 30 days, and month is equal to or before the month of the date of last dose of study drug plus 30 days;
- Year is missing.

7. STUDY POPULATIONS

Data are summarized for the Full Analysis Set (FAS), the All Randomized Population (ARP), and the Pes Cavus Population (PCP).

7.1. Subject Disposition

A disposition summary includes the number and percentage of patients in the following categories:

- Full Analysis Set (FAS)
- All Randomized Population (ARP)
- Pes Cavus Population (PCP)
- Safety population
- Per-protocol population
- Completed treatment through Week 48
- Discontinued treatment prior to Week 48
    - Reason for discontinuing treatment
- Completed study follow-up through Week 52
    - on treatment
    - discontinued treatment early but completed study visits
- Terminated early from the study
    - Reason for terminating study

A listing of disposition is provided for all enrolled patients.

7.2. Screen Failures

Screen failures are not summarized.

7.3. Protocol Deviations

Prior to database lock, a blinded team will identify all deviations, including major protocol deviations that could potentially affect the efficacy or safety conclusions of the study. Protocol deviations for excluded medications, patients who entered the study even though they did not satisfy the entry criteria, patients who received the wrong treatment or incorrect dose, and major protocol deviations are listed.

7.4. Demographic and Baseline Characteristics

Summaries of demographic and other baseline characteristics summaries are presented by treatment group for all analysis populations.

Demographic and other baseline characteristics include the following:

- Baseline Age, Age category (<18, ≥18)
- Sex, Race, Ethnicity
- Weight (kg), Height (cm), BMI (kg/m^2)
- Diastolic and systolic blood pressure (mmHg), Heart rate (bpm)
- Peak work (watts/kg), mFARS, each FARS subsection (i.e., A, B, C, D, E), ADL
- Age at FA onset, Years since FA onset
- Average number of hours exercised per week
- GAA1 repeat length, GAA2 repeat length, GAA1 repeat length ≥ 675 kbps
- X-ray measurements (calcaneal pitch, naviculocuboid overlap, talo calcaneal, talo first metatarsal, tibia calcaneal, tibia talar)
- Ambulatory status, Ambulatory assistive devices
- History of cardiomyopathy, areflexia, foot surgery, sensory neuropathy, swallowing difficulties, scoliosis, scoliosis surgery.

7.5. Listing of Subject Inclusion and Exclusion Criteria

A listing of enrolled patients who did not meet inclusion or exclusion criteria is generated.

7.6. Medical History

Medical history is summarized by treatment. Medical history is coded using MedDRA (Medical Dictionary for Regulatory Activities) version 14.1. Medical history items are summarized by MedDRA SOC and PT. Patient listings are also provided.

8. PRIMARY EFFICACY

Analyses of efficacy described in this section are the primary analyses of the efficacy endpoints, and pertain to the Full Analysis Set (i.e., patients without pes cavus). Analyses of efficacy described in this section will also be performed using the All Randomized and Pes Cavus Populations as exploratory analyses of efficacy for descriptive purposes. Analyses of efficacy using the All Randomized Population will include pes cavus (yes/no) as an additional covariate.

8.1. General Considerations

Analyses are performed for all omaveloxolone patients in comparison with all placebo patients. Summary statistics for observed values, change from baseline, and percent change from baseline (including 95% CI and quartiles) are presented by randomized treatment group.

8.2. Statement of the Null and Alternate Hypotheses

All efficacy endpoints compare patients randomized to omaveloxolone (μ_{OMAV}) to patients randomized to placebo ($\mu_{placebo}$). The primary and key secondary efficacy objectives will be evaluated according to the following hypotheses:

- H_0: $(\mu_{OMAV}) - (\mu_{placebo}) = 0$
- H_1: $(\mu_{OMAV}) - (\mu_{placebo}) \neq 0$

Because mFARS scores increase with disease progression, a decrease in mFARS score is considered evidence of benefit. Because PGIC and CGIC are assessed using a 7-point scale ranging from 1 (very much improved) to 7 (very much worse), a decrease in PGIC and CGIC is considered evidence of benefit.

8.3. Subgroup Analyses

The primary efficacy endpoint, change in mFARS score at Week 48, will be analyzed for subgroups of interest to assess homogeneity of the effect size for those subgroups that are sufficiently large and have an appropriate number of subjects to warrant these analyses. Key secondary (PGIC and CGIC) and secondary outcomes (9-HPT, T25-FWT, falls, peak work, and ADL) will also be analyzed for these subgroups. Additionally, safety analyses will be performed for these subgroups.

The following subgroup analyses will be tabulated:

- Age: <18; ≥18
- Sex: female; male
- Geographic location: US; Other
- Ethnicity: Non-Hispanic/Latino; Hispanic/Latino
- Race (White, Non-White)
- GAA1 Repeat Length ≥675: Yes, No

8.4. Multiple Comparisons and Multiplicity

This study has a single primary efficacy outcome: change in mFARS score at Week 48. If the study shows statistically significant evidence of benefit for this endpoint (change in mFARS score at Week 48), then the key secondary and secondary endpoints are analyzed using a hierarchical approach to maintain the family-wise overall Type I error rate of 0.05.

Each endpoint is tested at the 0.05 significance level. Formal testing in the hierarchy may proceed so long as statistically significant evidence of benefit continues to be shown. At such point in the hierarchy that one endpoint does not show statistically significant evidence of benefit, formal statistical testing of subsequent endpoints will not occur. Key secondary and secondary endpoints are tested in the following order:

1. PGIC at Week 48 (key secondary)
2. CGIC at Week 48 (key secondary)
3. Change in 9-HPT (reciprocal time measure of the non-dominant hand) at Week 48
4. Change in T25-FWT (reciprocal time measure) at Week 48
5. Frequency of falls over 48 weeks
6. Change in peak work during maximal exercise testing at Week 48
7. Change in ADL score at Week 48

All other objectives listed as "Exploratory" are presented with nominal significance levels for descriptive purposes only, including visits other than Week 48.

8.5. Primary Analysis of the Primary Efficacy Endpoint

8.5.1. Primary Efficacy Analysis

The primary endpoint of this study is the change in mFARS score (Section 6.5.9) at Week 48.

The mFARS scores for patients treated with omaveloxolone are compared with placebo at Week 48 using mixed models repeated measures (MMRM) analysis, with site (SITE) and baseline mFARS (BASE_MFARS) as covariates and the following fixed factors: treatment group (TRT), time (VISITNUM), the interaction between treatment and time (TRT*VISITNUM), the interaction between baseline and time (BASE_MFARS*VISITNUM) (code provided in Section 14.3). The analysis will use analysis visits 4, 12, 18, 24, 36, and 48 (Section 6.5.4). An unstructured covariance matrix is assumed.

In the event the MMRM model with an unstructured covariance structure does not converge, the following covariance structures are substituted, in the order listed. Each subsequent covariance structure is used only if each previous covariance structure is used and no previous model converged.

1. Heterogeneous Toeplitz covariance structure (assuming different variances at each time point and that measurements taken closer together in time are more highly correlated than those taken farther apart).

2. Toeplitz covariance structure (assuming measurements taken closer together in time are more highly correlated than those taken farther apart).
3. First order auto-regressive [AR(1)] covariance structure (assuming measurements taken closer together in time are more highly correlated than those taken farther apart, but more constrained than the Toeplitz structure).
4. Compound symmetry covariance structure (assuming equal correlation for measurements from a patient, regardless of how far apart in time when they were taken). [redacted]

The difference between omaveloxolone and placebo in change from baseline of mFARS will be estimated along with the 95% confidence interval at Week 48 for the primary analysis. The primary analysis of omaveloxolone vs placebo on mFARS score is at Week 48.

8.5.2. Sensitivity Analyses of the Primary Efficacy Results

Sensitivity analyses are included to assess the robustness of conclusions to the primary analysis of the primary efficacy endpoints (i.e., change in mFARS at Week 48), and to assess the assumption of MAR. The following sets of sensitivity analyses are summarized in tables and forest plots: tipping point with multiple imputation, treatment-based multiple imputation, control-based multiple imputation, and the MMRM model fit to the per-protocol population. Additional sensitivity analyses may be performed as appropriate. Sensitivity analyses are only performed in the FAS population (Section 6.3.1).

8.5.2.1. Tipping Point

The following tipping point sensitivity analysis will be performed to assess how severe departures from MAR must be in order to overturn conclusions from the primary analysis:

1. The PROC MI procedure in SAS will be used to generate 100 datasets satisfying the assumption of monotone missingness, imputing any missing mFARS values using multiple imputation. Imputation will be based on the non-missing observations within each treatment group.
2. The PROC MI procedure in SAS will be applied to each of the 100 datasets and all missing values in the placebo group will be imputed. These 100 datasets will then be used for the tipping point analysis.
3. For the tipping point analysis with a shift parameter, patients in the omaveloxolone group with a missing value will be assigned a shift parameter in the imputation procedure for progressively worse (higher) scores to find the point at which statistical significance is lost. Specifically, if the outcome of the hypothesis test favors omaveloxolone over the placebo group (P-value <0.05), then one point will be added as a shift parameter in the imputation for the omaveloxolone group, whilst the data for the placebo group will not be changed.
4. A mixed model repeated measures (MMRM) model (see Section 8.5.1 and Section 14.3) will be used to analyze each of the 100 completed data sets and PROC MIANALYZE will be used to combine these results to obtain the final estimates for the given value of the shift parameter.

5. This "mFARS shifting", as outlined in steps 3 and 4, will be repeated with one shift point at a time until the hypothesis test no longer rejects the null hypothesis in favor of the omaveloxolone group over the placebo group (i.e. when the P-value becomes greater than 0.05). Additional increments may be used to locate the tipping point. If the result is marginally significant, progressively lower scores will be added to find the point at which statistical significance is attained.

Further details of the tipping point analysis and associated code are provided in the Appendix (see Section 14.5).

8.5.2.2. Treatment-Based Multiple Imputation

As a sensitivity analysis, all missing Week 48 mFARS values are imputed with multiple imputation using the Week 48 data based on the randomized treatment group. The treatment-based sensitivity analysis is performed using the MMRM statistical model defined in Section 8.5.1.

8.5.2.3. Control-Based Multiple Imputation

As a sensitivity analysis, all missing Week 48 mFARS values are imputed with multiple imputation using the Week 48 data from the placebo group. The control-based sensitivity analysis is performed using the MMRM statistical model defined in Section 8.5.1.

8.5.2.4. Per-Protocol Population Analysis

The per-protocol population analysis will be performed in the per-protocol population using the MMRM statistical model defined in Section 8.5.1.

8.6. Primary Analysis of the Two Key Secondary Efficacy Endpoints

The following sections describe the primary analysis of the key secondary endpoints. Sensitivity analyses of key secondary endpoints may be performed as appropriate.

8.6.1. Patient Global Impression of Change

The PGIC is a 7-point scale that requires the patient to assess how much the patient's illness has improved or worsened relative to a baseline state at the beginning of an intervention (Guy 1976). The PGIC is assessed by completing the following statement "since I began trial treatment, my overall status is: very much improved (1), much improved (2), minimally improved (3), no change (4), minimally worse (5), much worse (6), very much worse (7)." All PGIC values collected at each visit using the appropriate analysis windows (Section 6.5.4) will be used; the number of patients in each category is summarized at each analysis visit by treatment group.

The PGIC responses for patients treated with omaveloxolone are compared to placebo at Week 48 using analysis of covariance (ANCOVA), with treatment group and site as covariates (Section 14.8) and Week 48 value of PGIC as the outcome. Missing Week 48 data are imputed as defined in Section 14.9 using multiple imputation based on the treatment group to which the subject is assigned.

8.6.2. Clinical Global Impression of Change

The CGIC is a 7-point scale that requires the clinician to assess how much the patient's illness has improved or worsened relative to a baseline state at the beginning of an intervention (Guy 1976). The CGIC is assessed by completing the following statement "Compared to the patient's condition at the start of the trial, this patient's overall status is: very much improved (1), much improved (2), minimally improved (3), no change (4), minimally worse (5), much worse (6), very much worse (7)." The number of patients in each category are summed at each analysis visit by treatment group.

CGIC is analyzed using the same methods specified for PGIC in Section 8.6.1 with the primary analysis at Week 48.

8.6.3. Sensitivity Analyses of the Key Secondary Efficacy Results

Sensitivity analyses are included to assess the robustness of conclusions to the primary analysis of the key secondary efficacy endpoints (i.e., PGIC and CGIC at Week 48), and to assess the assumption of MAR. The following sets of sensitivity analyses are summarized in tables and forest plots: tipping point with multiple imputation, control-based multiple imputation, and the ANCOVA model fit to the per-protocol population. Additional sensitivity analyses may be performed as appropriate. Sensitivity analyses are only performed in the FAS population (Section 6.3.1).

8.6.3.1. Tipping Point

The following tipping point sensitivity analysis will be performed to assess how severe missing data results must be in order to overturn conclusions from the key secondary analysis:

1. The PROC MI procedure in SAS will be used to generate 100 datasets imputing any missing values using multiple imputation
2. For the tipping point analysis with a shift parameter in the omaveloxolone group, patients in the omaveloxolone group with a missing value will be assigned a shift parameter in the imputation procedure for progressively worse (higher) scores to find the point at which statistical significance is lost. Specifically, if the outcome of the hypothesis test favors omaveloxolone over the placebo group (P-value <0.05), then one point will be added as a shift parameter in the imputation for the omaveloxolone group, whilst the data for the placebo group will not be changed.
3. A analysis of covariance (ANCOVA) model (see Section 14.8) will be used to analyze each of the 100 completed data sets and PROC MIANALYZE will be used to combine these results to obtain the final estimates for the given value of the shift parameter.
4. This "shifting", as outlined in steps 2 and 3, will be repeated with one shift point at a time until the hypothesis test no longer rejects the null hypothesis in favor of the omaveloxolone group over the placebo group (i.e. when the P-value becomes greater than 0.05). Additional increments may be used to locate the tipping point. If the result is marginally significant, progressively lower scores will be added to find the point at which statistical significance is attained.

Further details of the tipping point analysis and associated code are provided in the Appendix (see Section 14.10).

8.6.3.2. Control-Based Multiple Imputation

As a sensitivity analysis, all missing Week 48 PGIC and CGIC values are imputed with multiple imputation using the Week 48 data from the placebo group. The control-based sensitivity analysis is performed using the ANCOVA statistical model defined in Section 8.6.1.

8.6.3.3. Per-Protocol Population Analysis

The per-protocol population analysis will be performed in the per-protocol population using the ANCOVA statistical model defined in Section 8.6.1.

8.7. Primary Analysis of the Secondary Efficacy Endpoints

The following sections describe the primary analyses of the secondary endpoints. Sensitivity analyses of secondary endpoints may be performed as appropriate.

8.7.1. 9-Hole Peg Test (9-HPT)

The 9-HPT is a brief, standardized, quantitative test of upper extremity function. Two trials are attempted using each hand, and the time for the dominant and the non-dominant hand are analyzed separately. The average observed time for each hand is summarized descriptively at baseline, Week 24, and Week 48. Change from baseline by treatment is summarized for each hand at Weeks 24 and 48. Longer test times reflect more impairment on the patient's upper extremity function, thus a decrease from baseline suggests an improvement. Reciprocal time values and pegs/second values (Section 6.5.10) are also summarized at each time point along with change from baseline in these parameters. The primary analysis compares the change from baseline in the reciprocal time measure of the non-dominant hand for patients treated with omaveloxolone to patients treated with placebo at Week 48 with MMRM analysis using the statistical model defined in Section 8.5.1. The comparison of omaveloxolone with placebo is estimated using the difference in adjusted means and 95% CI for the difference in changes from baseline to Week 48. All other analyses of 9-HPT are considered exploratory and the nominal p-value will be reported for these analyses.

8.7.2. Timed 25-Foot Walk Test (T25-FWT)

The T25-FWT is a quantitative mobility and leg function performance test based on time in seconds to complete a 25-foot walk. The average observed time is summarized descriptively at baseline, Week 24, and Week 48. Change from baseline by treatment is summarized at Weeks 24 and 48. Longer test times reflect more impairment on the patient's ability to walk, thus a decrease from baseline suggests an improvement. Reciprocal time values and feet/second values (Section 6.5.11) are also summarized at each time point along with change from baseline in these parameters. The primary analysis is the change from baseline in the reciprocal time measure for patients treated with omaveloxolone compared to patients treated with placebo at Week 48 using MMRM analysis using the statistical model defined in Section 8.5.1. All other analyses of T25-FWT are considered exploratory and the nominal p-value will be reported for these analyses.

8.7.3. Frequency of Falls

Throughout Part 2 of the study, patients are instructed to record any instances of falls in a fall diary. A fall is defined as "the patient unintentionally coming to rest on the ground or at a lower level." Items included on the falls diary may include, but are not limited to, the date and time of each fall, the location of the fall, the preceding activity prior to the fall, the perceived cause of the fall, and if an injury was sustained after the fall. Patients are provided the fall diary at Screening and must record fall incidents between Screening and Week 48. The total number of falls before and after study drug administration is reported. The mean, quartiles (i.e., median, 25th, and 75th percentiles), and confidence interval of the total number of falls per patient recorded are summarized descriptively by treatment group. The number of falls by 12-week exposure interval (i.e., Day 1 to Week 12, Week 12 to Week 24, Week 24 to Week 36, and Week 36 to Week 48) overall and normalized by the number of patients per group, will also be summarized descriptively.

The primary analysis is the comparison of the total number of falls after randomization, summarized by computing the incidence rate of falls using a Poisson model fit with PROC GENMOD in SAS (Section 14.4). The comparison with placebo will be estimated from the Poisson model with the natural logarithm of time on study (days) included as an offset term; 95% confidence intervals of the incidence rates and the difference between those rates will be computed.

8.7.4. Peak Work

The change in peak work in (W/kg) is analyzed using an MMRM approach analogous to the method outlined for the primary efficacy endpoint (baseline peak work will be substituted for baseline mFARS as covariate and change from baseline in peak work will be the outcome) as described in Section 8.5.1. The comparison of omaveloxolone with placebo is estimated using the difference in adjusted means and 95% CI for the changes from baseline to Week 48.

8.7.5. Activities of Daily Living

The Activities of Daily Living scale examines the ability of patients to complete everyday tasks (e.g., holding a fork, dressing). The Activities of Daily Living is a 9-question assessment, with the total score being the sum of 9 questions. A score of 0 is considered normal, therefore, a decrease from baseline would be considered an improvement. If a question result is unanswered, then the total Activities of Daily Living score is considered missing. The number of patients with an unanswered ADL question will be presented. In addition, summaries of each Activities of Daily Living question are generated.

A mixed model (MMRM) will be fit with the change from baseline in ADL as the outcome and baseline ADL, treatment group, visit, treatment group by visit interaction, and baseline ADL by visit interaction as covariates. The model uses analysis visits 24, 36, and 48 weeks. The SAS code for this analysis can be found in Section 14.3. The comparison of omaveloxolone with placebo is estimated using the difference in adjusted means and 95% CI for the difference in changes from baseline to Week 48.

8.8. Analysis of the Exploratory Efficacy Endpoints

Summary statistics and 95% confidence intervals for treatment differences are provided for all exploratory endpoints as described below. P-values are provided for descriptive purposes only. Sensitivity analyses of exploratory endpoints may be performed as appropriate.

8.8.1. Videos of Normal Walking

Patients have the option to consent to have videos of normal walking recorded. Patients who consent to video collection will have video collection during normal walking with a 25-foot (un-timed) walk using his/her normal gait. The normal walk is not part of the T25-FWT, and must be performed after the 25-foot timed walk test.

Paired videos of normal walking at baseline to Week 24, and baseline to Week 48 (i.e., end-of-treatment) will be assessed by a movement specialist using the 7-point Likert scale to determine the clinical global impression of change (CGIC-w). CGIC-w values from the video analysis will be analyzed using the same methods specified in Section 8.6.1.

8.8.2. SF-36 Health Survey Update

The SF-36 is a multi-purpose, short-form health survey with 36 questions. It yields a profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures and a preference-based health utility index. The SF-36 is a generic measure, as opposed to one that targets a specific age, disease, or treatment group.

There are eight domains of the SF-36 as follows: Physical Functioning (items 3a-3j), Role-Physical (items 4a-4d), Role-Emotional (items 5a-5c), Social Functioning (items 6, 10), Bodily Pain (items 7, 8), Mental Health (items 9b, 9c, 9d, 9f, 9h), Vitality (items 9a, 9e, 9g, 9i) and General Health Perceptions (items 1, 11a-11d). Change in Health (item 2) is an overall health status assessment that compares to one year ago. The score is converted to a 0 - 100 scale, with 0 indicating poor health and 100 indicating good health (Section 14.14).

The MCS (mental component scale) and the PCS (physical component scale) as defined by the SF-36 manual are summarized at baseline and Week 48, along with the change from baseline by treatment. Lower scores reflect poor quality of life thus, a positive change from baseline suggests an improvement.

There are two assessments of SF-36 post-baseline, one at Week 24 and one at Week 48. A MMRM will be fit with change from baseline as the outcome and baseline MCS/PCS, treatment group, visit, treatment group by visit interaction, and baseline MCS/PCS by visit interaction as covariates. The SAS code for this analysis is similar to that provided in Section 14.3. The comparison of omaveloxolone with placebo is estimated using the difference in adjusted means and 95% CI for the difference in changes from baseline to Week 24 and baseline to Week 48.

9. EXPLORATORY EFFICACY

9.1. Exploratory Analyses of Efficacy Using the Full Analysis Set

Analyses of efficacy described in this section are exploratory analyses of the efficacy endpoints and are examined in all populations.

9.1.1. mFARS

Changes in mFARS for the omaveloxolone group versus the placebo group will be evaluated at Weeks 4, 12, 18, 24, and 36 using the statistical model defined in Section 8.5.1.

Changes in each FARS subsection score (Section 6.5.9) for the omaveloxolone group versus the placebo group will be evaluated at Weeks 4, 12, 18, 24, 36, and 48 using the statistical model defined in Section 8.5.1.

The proportion of patients at Week 48 with an mFARS improvement from baseline at or less than -1 at 1-point increments (e.g., ≤-1, ≤-2, ≤-3, ≤-4, ≤-5, ≤-6, ≤-7, ≤-8) at Week 48 are summarized by treatment group. Similarly, the proportion of patients at Week 48 with an mFARS worsening from baseline at or above 1 at 1-point increments (e.g., ≥1, ≥2, ≥3, ≥4, ≥5, ≥6, ≥7, ≥8) at Week 48 are summarized by treatment group.

A plot of the cumulative distribution of mFARS change from baseline at Week 48 will be generated. A Kruskal-Wallis rank sum test by treatment group of mFARS change from baseline at Week 48 will be used to assess for a difference in distributions of Week 48 mFARS changes between treatment groups.

Waterfall plots showing individual mFARS change from baseline, labeled by treatment group, at Week 48 will be generated.

9.1.2. Patient Global Impression of Change

[redacted]

To describe the PGIC over time, the proportion of patients having a PGIC response of "very much improved" or "much improved" (combined categories) are summarized at each visit. The treatments are compared descriptively by calculating the differences in proportions along with the 95% confidence intervals for those differences.

The correlation of PGIC response and change from baseline in mFARS at Week 48 will be calculated to describe the relationship between PGIC and the primary endpoint.

9.1.3. Clinical Global Impression of Change

Exploratory analyses specified for PGIC (Section 9.1.2) will be performed for CGIC.

9.1.4. 9-Hole Peg Test (9-HPT)

Changes in reciprocal 9-HPT time measure for the omaveloxolone group versus the placebo group will be evaluated for the non-dominant hand at Week 24 using the statistical model defined in Section 8.5.1.

[redacted]

9.1.5. Timed 25-Foot Walk Test (T25-FWT)

Changes in the reciprocal T25-FWT time measure for the omaveloxolone group versus the placebo group will be evaluated at Week 24 using the statistical model defined in Section 8.5.1.

[redacted]

9.1.6. Frequency of Falls

No exploratory statistical comparisons are planned for frequency of falls.

9.1.7. Maximal Exercise Test

Change from baseline in peak work (W/kg) for the omaveloxolone group versus the placebo group will be evaluated at Weeks 4, 12, 18, 24, and 36 using the statistical model defined in Section 8.5.1.

Change from baseline in other parameters collected during the maximal exercise test (Section 14.12) for the omaveloxolone group versus the placebo group will be summarized descriptively. For parameters measured at resting and at the end of the maximal exercise test (Table 27), the change from resting is also analyzed at each visit and for change from baseline.

[redacted]

9.1.8. Activities of Daily Living

Change from baseline in ADL for the omaveloxolone group versus the placebo group will be evaluated at Weeks 24 and 36 using the statistical model defined in Section 8.7.5.

[redacted]

9.2. Exploratory Analyses of Efficacy Using Other Analysis Populations

The primary analysis of efficacy is based on the full analysis set, which does not include patients with pes cavus. Because pes cavus is common in patients with FA, patients with pes cavus are also included in the study but do not comprise more than 20% of patients enrolled in Part 2. Secondary endpoints that are less likely to be affected by pes cavus are also assessed in the study to determine whether a therapeutic benefit can be detected in patients with pes cavus.

As indicated in Section 8, all primary, key secondary, secondary, and exploratory efficacy endpoints listed in Section 3.1 are summarized descriptively at each visit and change from baseline using the all randomized population (ARP) and pes cavus population (PCP). General considerations for data analysis (Section 6) are followed. Inferential testing performed using the ARP or PCP will be for descriptive purposes only.

10. SAFETY AND TOLERABILITY

Safety data (including AEs, laboratory data, vital signs, electrocardiogram [ECG] data, echocardiogram data, and physical examinations) are listed and summarized for patients in the safety population (Section 6.3.2). Safety data are summarized by treatment both overall and by pes cavus stratum.

The analysis visit windows (Section 6.5.4) are used for all safety analyses. Unscheduled visits that are outside the visit windows are listed and reflected in summaries of change to worst post-baseline measures where appropriate.

10.1. Adverse Event Preferred Term and Body/Organ System Summary Tables

AEs are summarized by treatment as defined by the safety analysis set. General considerations for AE summaries and calculations are:

- Multiple events by preferred term (PT) and system organ class (SOC) are counted once only per patient for each treatment.
- For summaries by severity, only the most severe event is counted per patient for each treatment.
- For summaries by relationship, for frequency counts by patient, only the most related event is counted per patient for each treatment.
- An AE with a missing resolution date or incomplete date that is not identified as continuing is assumed to be continuing and no duration is calculated.
- Only treatment-emergent adverse events (TEAEs) are included in summaries.

AEs are coded using MedDRA (Medical Dictionary for Regulatory Activities) version 14.1. In MedDRA, each verbatim term is mapped to a preferred term and high level term (HLT), which is then mapped to a system organ class. Tables and listings present data at the SOC and PT level.

Treatment-emergent adverse events (TEAEs) are events that either:

- Had a date of onset on or after the date of the date of first dose and not more than 30 days after the date of the last dose of study drug, or
- Had no recorded date of onset with a stop date after the first dose of study drug, or
- Had no recorded date of onset or stop date.

In addition, adverse events which occurred >30 days after the date of last dose of study drug are summarized as late-onset AEs (LOAEs). These are adverse events that had a date of onset more than 30 days after the date of the last dose of study drug.

The investigator grades the severity of the AEs as mild, moderate, or severe as defined in the study protocol.

Association or relatedness to the study medication is graded by the investigator according to criteria specified in Section 12.4 of the study protocol.

As defined in the protocol and captured on the CRF, a serious adverse event (SAE) is an adverse event that results in any of the following:

- Death;
- A life-threatening adverse drug experience;
- Inpatient hospitalization or prolongation of existing hospitalization;
- Results in persistent or significant incapacity or substantial disruption of the ability to conduct normal life functions;
- Results in a congenital anomaly or birth defect in an offspring of a patient taking study drug;
- Is an important medical event.

10.1.1. Missing and Partial AE Onset Dates

Rules for handling missing partial AE Onset Dates are included in Section 6.6.3.

10.1.2. Summaries of Adverse Events for Serious Adverse Events (SAE), Adverse Event Dropouts, and Death

Treatment-emergent AEs are summarized by treatment at onset of the AEs. For each treatment, SOC, and PT, the number and percentage of patients reporting an event are calculated. In summary tables, SOC are presented alphabetically and events within SOC are presented by decreasing frequency count.

Summary tables (number and percentage of patients) of AEs (by SOC and PT) are provided by treatment as follows:

- All treatment-emergent AEs
- All treatment-emergent related AEs (definitely, probably, or possibly related)
- All treatment-emergent AEs by severity
- All treatment-emergent serious adverse events (including deaths)
- All treatment-emergent adverse events leading to discontinuation of study drug.

Listings are provided showing:

- All AEs
- Serious adverse events (including deaths)
- AEs leading to discontinuation of study drug.

10.2. Exposure and Compliance

The duration of study drug exposure is defined as the number of days on treatment from the first dose of study drug until the last dose of study drug (last dose – first dose + 1). Study drug exposure is summarized by descriptive statistics. Summaries include the total dose (mg) received (based on the number of pills returned), study drug compliance, the number and

percentage of patients receiving study drug by scheduled dispense visit, and duration (days) of exposure during the study treatment period.

Total number of doses dispensed and total dose (mg) dispensed are calculated from total number of kits (bottles) recorded on the Study Drug Dispensation eCRF. Total number of doses received is calculated from information on the eCRF of Study Drug Return and Study Drug Dispensation, as the total number of doses dispensed – total number of doses returned. Study drug compliance (%) is calculated as 100 × (total number of doses received) / (total number of doses dispensed). The proportion of patients who have ≥80% compliance are summarized.

In addition, the following exposure categories are summarized by number and percentage of patients:

- ≤16 Weeks
- >16 Weeks
- >24 Weeks

10.3. Concomitant and Other Medications

Concomitant medications are coded using the World Health Organization (WHO) drug dictionary (Enhanced version, September 2011, B2 format) for anatomical therapeutic chemical classification (ATC) and preferred drug name. A patient who used multiple medications is counted only once for each ATC and preferred drug name. ATC and preferred drug name within each ATC are sorted alphabetically. Coded concomitant medications are summarized by treatment. Percentages are based on the number of patients in the safety analysis set.

A concomitant medication is any medication taken at the time of first study treatment or a medication that was started after the start of study drug dosing. Specifically, concomitant medications are medications

- that are continued from screening and continued after the first study drug dosing, or
- that have start dates or stop dates within the treatment period.

Concomitant medications include those with an end date after the first study drug administration as well as medications without an end date. Medications with an end date on the date of first study drug administration are not to be considered concomitant medications.

In addition, patients who take excluded medications (defined in the Protocol Section 9.2.1) during the study are listed.

10.3.1. Missing and Partial Concomitant and Other Medication Start and Stop Dates

Missing and partial concomitant medication start and stop dates are detailed in Section 6.6.2.

10.4. Clinical Laboratory Evaluations

Laboratory data are summarized at baseline and at each time point by treatment. Only values obtained within analysis study windows (Section 6.5.4 and Section 6.5.4.1) are included in by-visit summary statistics. On-treatment values will be summarized according to the analysis study

windows in Section 6.5.4 and off-treatment values will be summarized according to the analysis study windows in Section 6.5.4.1.

10.4.1. Summaries of Laboratory Results

Selected laboratory evaluations and change from baseline are summarized by treatment, laboratory category (hematology, chemistry), test, and study visit using continuous statistics. Laboratory tests to be summarized are provided in Section 14.13. The estimated glomerular filtration rate (eGFR) results are calculated using formulas described in Section 6.5.8.

Line graphs of change from baseline are generated for selected laboratory tests, such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase (GGT), ferritin, and creatine kinase (CK). Line graphs include mean ± SE over time for change from baseline.

Due to the nature of urinalysis parameters, summaries of continuous statistics are not provided. Urinalysis results (other than pH and specific gravity) of ketones, protein, blood, glucose, clarity, color, leukocytes, nitrite, bilirubin, and microscopic examinations (if indicated based on laboratory results), urine microscopic findings, and pregnancy test results are not summarized.

The number and percentage of patients with laboratory normality and abnormality categories at any time during the study (Normal, Low, High) as well as the treatment-emergent shift are summarized by treatment, laboratory category (hematology, chemistry, and urinalysis), and laboratory test. Shift tables summarizing Baseline, Last post-baseline, and shift to last post-baseline value are presented for lab parameters as appropriate.

10.4.2. Transaminases and Total Bilirubin

To assess the potential for drug induced liver injury, the number of patients meeting the follow criteria is summarized:

- ALT more than 3x, 5x, 10x or 20x ULN
- AST more than 3x, 5x, 10x or 20x ULN
- either ALT or AST more than 3x, 5x, 10x or 20x ULN
- either ALT or AST more than 5x ULN for more than two consecutive weeks
- alkaline phosphatase (ALP) >1.5x ULN
- TBL >2x ULN
- AST or ALT >3x ULN with an associated TBL>1.5x ULN
- AST or ALT >3xULN with an associated TBL>2x ULN.

A summary table that includes frequencies and percentages of patients that meet any of the above criteria at any time during the study are provided. A listing of subjects with abnormal ALT, AST, or TBL is also provided.

Moreover, the following shift tables of ALT, AST, and TBL are generated:

- Baseline to highest on-treatment;

- Baseline to follow-up (i.e., off-treatment); and
- Highest on-treatment to follow-up.

In addition, evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plots are generated for ALT and AST versus total bilirubin (TBL).

10.5. Vital Signs

Vital signs assessments include systolic blood pressure (SBP, mmHg) and diastolic blood pressure (DBP, mmHg), body temperature (°C), heart rate (HR, bpm), height (cm), weight (kg), and BMI (kg/m^2). Only values obtained within analysis study windows (Section 6.5.4 and Section 6.5.4.1) are included in by-visit summary statistics. On-treatment values will be summarized according to the analysis study windows in Section 6.5.4 and off-treatment values will be summarized according to the analysis study windows in Section 6.5.4.1.

Vital signs are summarized at baseline and at each time point along with the change from baseline by treatment. All data are listed.

10.6. Physical Exam

Physical exam results are listed.

10.7. Echocardiogram

Cardiac function [ejection fraction (%), left and right wall thickness (mm), septum thickness (mm) and left ventricular mass index (LVMI) (g/m^2)] measurements captured on echocardiogram are summarized at baseline and each time point along with the change from baseline.

10.8. Electrocardiogram

Electrocardiogram (ECG) data, such as clinical interpretation of ECGs, ventricular rate and interval assessments of PR, QRS, and QT are collected on the eCRF. QTcF is calculated (Section 6.5.12). Descriptive statistics for observed values and change from baseline at each time point are presented for these 12-lead ECG interval assessments. Only values obtained within analysis study windows (Section 6.5.4 and Section 6.5.4.1) are included in by-visit summary statistics. On-treatment values will be summarized according to the analysis study windows in Section 6.5.4 and off-treatment values will be summarized according to the analysis study windows in Section 6.5.4.1.

In addition, number and percentage of patients with any abnormal values at any time during the study (i.e., above a pre-specified threshold) are summarized by time point and overall while on study drug. The pre-specified levels of ECG QTc thresholds are consistent with FDA guidance.

Table 6: Pre-Specified Threshold Levels for ECG Parameters

ECG Parameter	Pre-Specified Level
PR	>200 msec

ECG Parameter	Pre-Specified Level
QTcF	>450, >480 or >500 msec, >30 or >60 msec increase from baseline
Heart rate	<40, >100 beats/min

10.9. Pregnancy

A listing is provided for serum and urine pregnancy results of all on-study pregnancies.

11. PHARMACOKINETICS

A separate document contains details regarding the planned pharmacokinetic analyses.

12. CHANGES FROM PROTOCOL-SPECIFIED ANALYSIS

This analysis plan specifies (Section 6.3) that the primary analysis of efficacy is based on the Full Analysis Set, which includes only patients without pes cavus. The protocol (Section 14.3.1) indicates that all enrolled patients are included in the ITT population for analysis of efficacy. The rationale for this change is based on insights from Part 1 analyses. Analysis of Part 1 data showed that treatment with omaveloxolone did not statistically improve studied endpoints (i.e., mFARs, exercise testing) in patients with pes cavus. Although the small sample size in Part 1 of the study limited the ability to detect a treatment effect, the presence of pes cavus also likely interferes with the ability to perform assessments that require standing or pedaling. Two of the four subsections that comprise the study's primary endpoint, mFARS, include assessments of lower limb coordination and upright stability. As a result, the primary analysis population for Part 2 efficacy is based on the stratum of patients enrolled without pes cavus (Section 6.3.1). Because pes cavus is common in patients with FA, patients with pes cavus are also included in the study but do not comprise more than 20% of patients enrolled in Part 2, and randomization was stratified by pes cavus status (with pes cavus vs without pes cavus). Secondary endpoints that are less likely to be affected by pes cavus are also assessed in the study to determine whether a therapeutic benefit can be detected in patients with pes cavus. This analysis plan specifies that patients having pes cavus are assessed descriptively for efficacy in the Pes Cavus Stratum.

The protocol sample size calculations (Section 14.1) were based on 100 patients (total enrollment) [redacted]. With the primary analysis of efficacy specified as being based on only those patients without pes cavus, this analysis plan includes power calculations for the 80 patients in the primary ITT population (Section 4.3.1). [redacted]

Additionally, this SAP changes the defined order of some efficacy endpoints. This analysis plan elevated PGIC and CGIC from secondary endpoints in the protocol to key secondary endpoints. Select protocol-specified exploratory endpoints (9-HPT, T25-FWT, frequency of falls, and ADL) are elevated to secondary endpoints in this analysis plan. The protocol did not specify change in raters' assessments of videos of normal walking as an endpoint, but this analysis plan added it as an exploratory endpoint.

The protocol includes only a high-level description of the analysis of the primary endpoint. No change is being made to that (a mixed model for repeated measurements), but this document sets out all the important details of how that analysis is done, including issues around handling missing data and planned sensitivity analyses.

This SAP specifies CKD-EPI and Bedside Schwartz as the eGFR formulas to be used for analyses of eGFR (Section 6.5.8). The protocol specified the Modification of Diet in Renal Disease (MDRD) 4 variable formula to calculate eGFR. This change aligns eGFR calculations across the sponsor's development programs.

13. REFERENCES

Guy, W. ECDEU Assessment Manual For Psychopharmacology, DHEW Publication No. ADM 76–338, US Government Printing Office, Washington, DC, USA, 1976.

Lynch DR, Farmer JM, Wilson RL, Balcer LJ.Performance measures in Friedreich ataxia: potential utility as clinical outcome tools. Mov Disord. 2005 Jul;20(7):777-82.

Lynch DR, Farmer JM, Tsou AY, Perlman S, Subramony SH, Gomez CM, Ashizawa T, Wilmot GR, Wilson RB, Balcer LJ. Measuring Friedreich ataxia: complementary features of examination and performance measures. Neurology. 2006 Jun 13;66(11):1711-6.

14. APPENDIX

14.1. Programming Specifications

The example code for programming specifications represent best approximations based on existing data structures and content. However, the specific SAS coding conventions may change depending on the structure and content of the data.

Continuous data are listed corresponding to the precision measured or calculated. Measures of central tendency are presented using one more decimal place than the precision of the data. Summaries of variability are presented using two significant digits more than the precision of the underlying data. Quartiles, the minimum, and the maximum are presented using the precision of the data. Rounding for derived values should be one additional decimal place to which the raw data were entered. For instance, if the raw data are rounded to the 0.001 decimal place, the derived values must be rounded to the 0.0001 decimal place.

All percentages are to be expressed as integers with one decimal place. The convention for rounding percentages is as follows:

- Values greater than or equal to x.x5% are rounded up
- Values between 0 and x.x5% are rounded down

Table 7: Parameter Names Used in Analysis

Description	Parameter
Baseline mFARS	BASE_MFARS
Change from Baseline mFARS	DELT_MFARS
Total Number of Falls	FALLS
Patient Number	PATNO
Pes Cavus (Yes/No)	PES_CAVUS
PGIC/CGIC Values	PGIC
Site	SITE
Treatment Group	TRT
Visit	VISIT
Visit Number	VISITNUM

14.2. Rationale for Programming Conventions

Assumption of No Monotone Missing for ANCOVA Analysis of PGIC/CGIC at Week 48

Since PGIC/CGIC results by visit week are not highly correlated within a patient, and because past subjective PGIC/CGIC results are not indicative of future results, no monotone missing datasets are generated to compute the Week 48 PGIC/CGIC results. In essence, past PGIC/CGIC results have no influence on the imputation of Week 48 PGIC/CGIC results, since PGIC/CGIC is a highly subjective measure.

Specification of A Single Seed Per Analysis (Table 26)

In order to both aid in validation and pre-define the seed values for multiple imputation analyses, a single seed is defined for each analysis. For tipping point analyses, only a single seed is used.

14.3. Mixed Model Repeated Measures Code

Sample SAS code for the primary analysis will be as follows:

Table 8: SAS code for MMRM analysis

```
proc mixed data=OUT2 method=reml alpha=0.05 covtest;
   class  VISITNUM(ref=LAST)
          TRT PATNO SITE;
   model CHG_MFARS =  BASE_MFARS TRT VISITNUM
         TRT*VISITNUM BASE_MFARS*VISITNUM
          SITE / solution ddfm=kr;
   repeated VISITNUM / subject=PATNO type=un r;
   lsmeans TRT*VISITNUM / diff om cl e;
   estimate 'estimate at Week 48' TRT 1 -1 VISITNUM*TRTGRP 0 0 0
    0 0 0 0 0 0 0 1 -1
        / CL;
   ods output estimates=SI_ESTIMATES;
run;
```

The pes cavus covariate is only utilized when analyzing results for the all randomized population (Section 6.3.2), as shown in Table 9.

[redacted]

The primary result will be obtained from the estimate statement and will be checked against the lsmeans statement to ensure that the correct result is obtained. This code will also be used for additional MMRM analyses such as that for peak work, SF-36, and ADL. The estimate statement will be adjusted to obtain results for Week 24 in models where the outcome is measured at baseline, Week 24 and Week 48.

14.4. Code for Frequency of Falls

The following SAS code serves as sample code to provide the estimated differences, confidence intervals, and p-value. A Poisson model is fit to the data with the natural logarithm of time (days) on study included as an offset term and an option to account for potential over dispersion (dscale). The test of the statistical significance of the coefficient associated with treatment group membership will test the null hypothesis of no differences in the number of falls between the two groups.



The pes cavus covariate is only utilized when analyzing results for the all randomized population (Section 6.3.2).



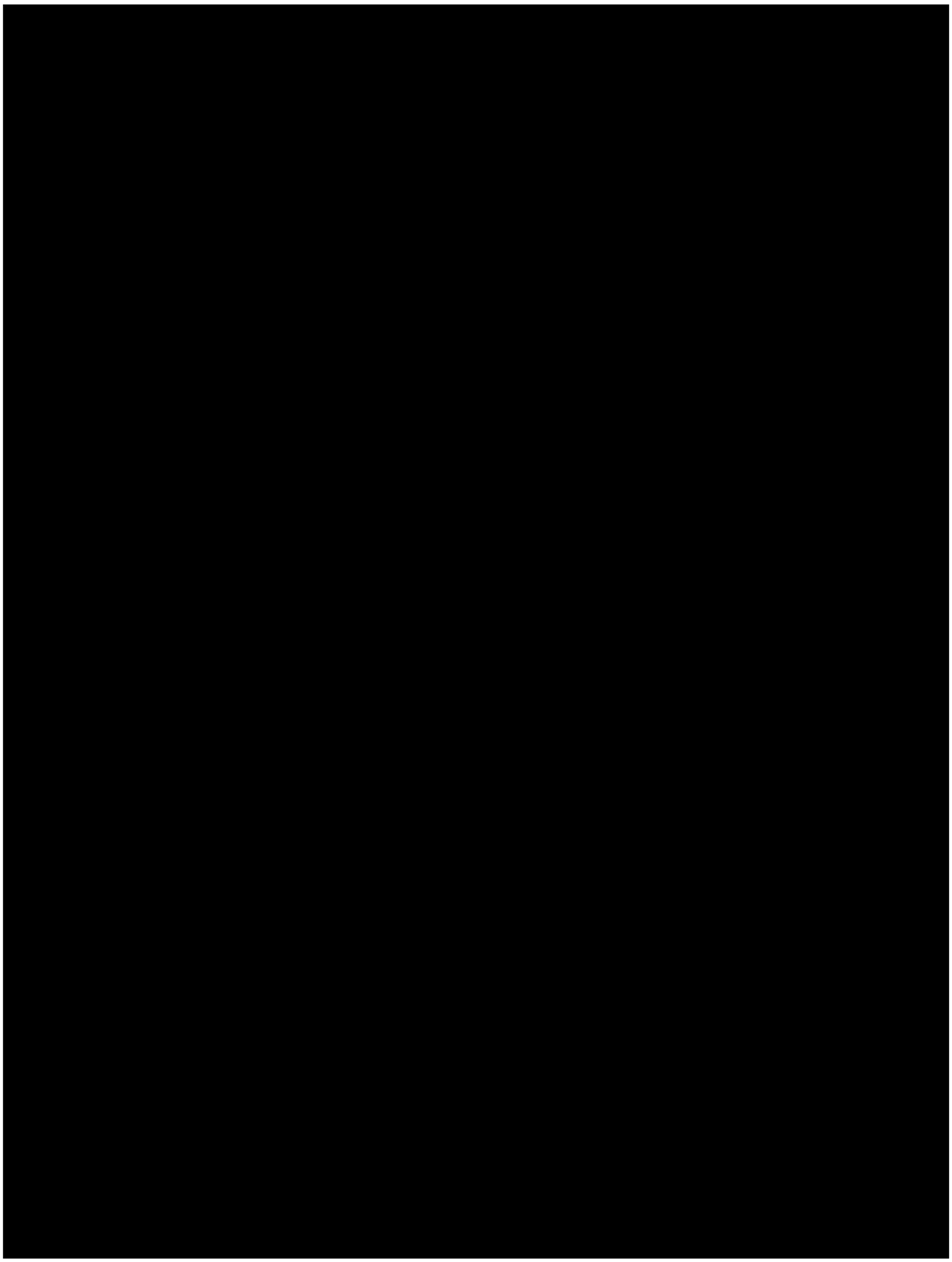
14.5. Treatment-based imputation and analysis of the primary endpoint

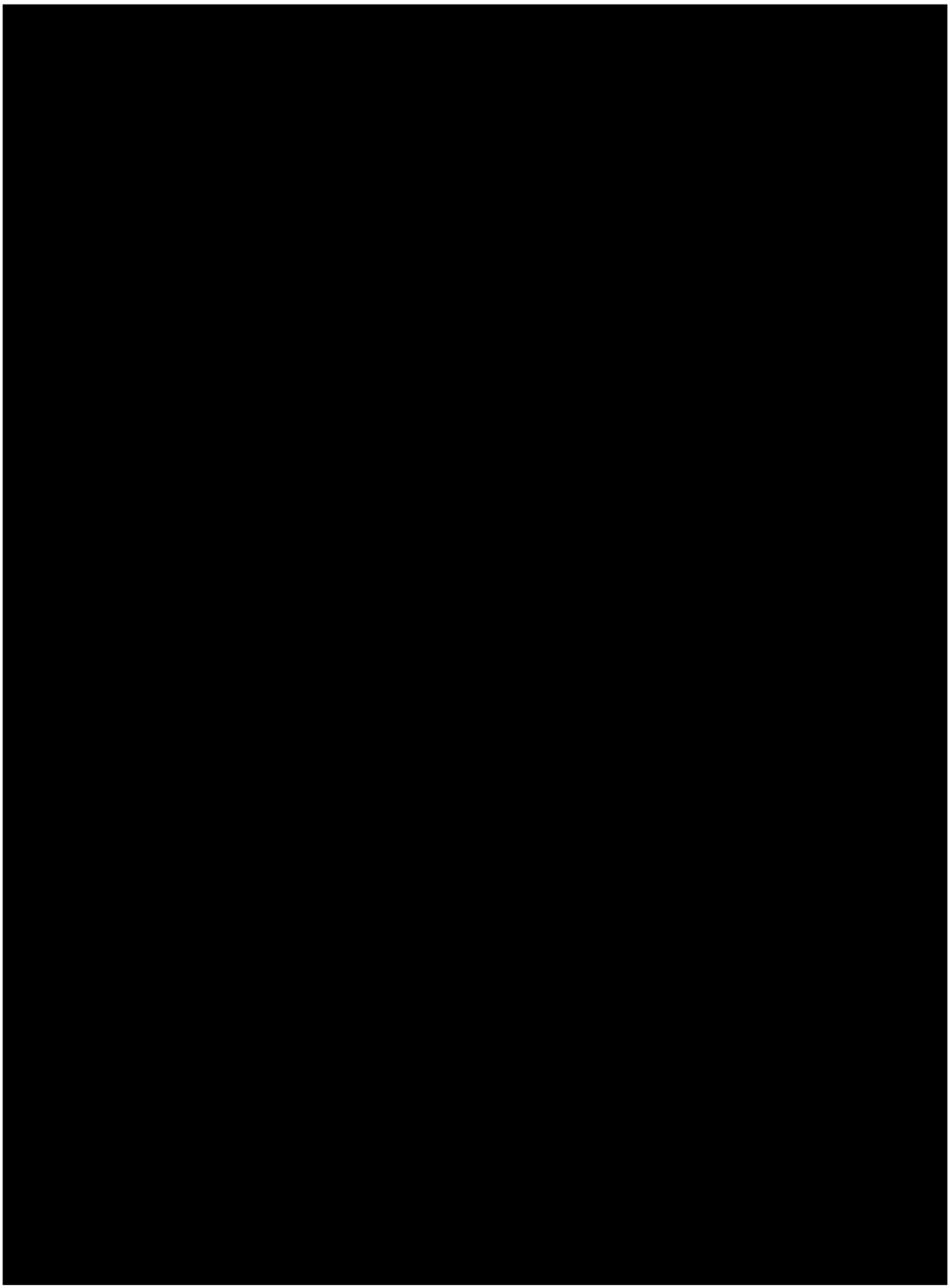
This section provides details and sample SAS programming code for the treatment-based imputation of the primary endpoint analysis outline in Section 8.5.2.

The treatment-based imputation of the primary endpoint will use a single seed from Section 14.11.

The dataset DATA_IMP is the transposed input dataset, with columns for changes in mFARS at Weeks 1-48. The OUT_IMP_CONT dataset and the code described in Table 10 creates a monotone missing dataset with columns for change from baseline by analysis study week.

Table 10: SAS code for treatment-based imputation of the primary endpoint







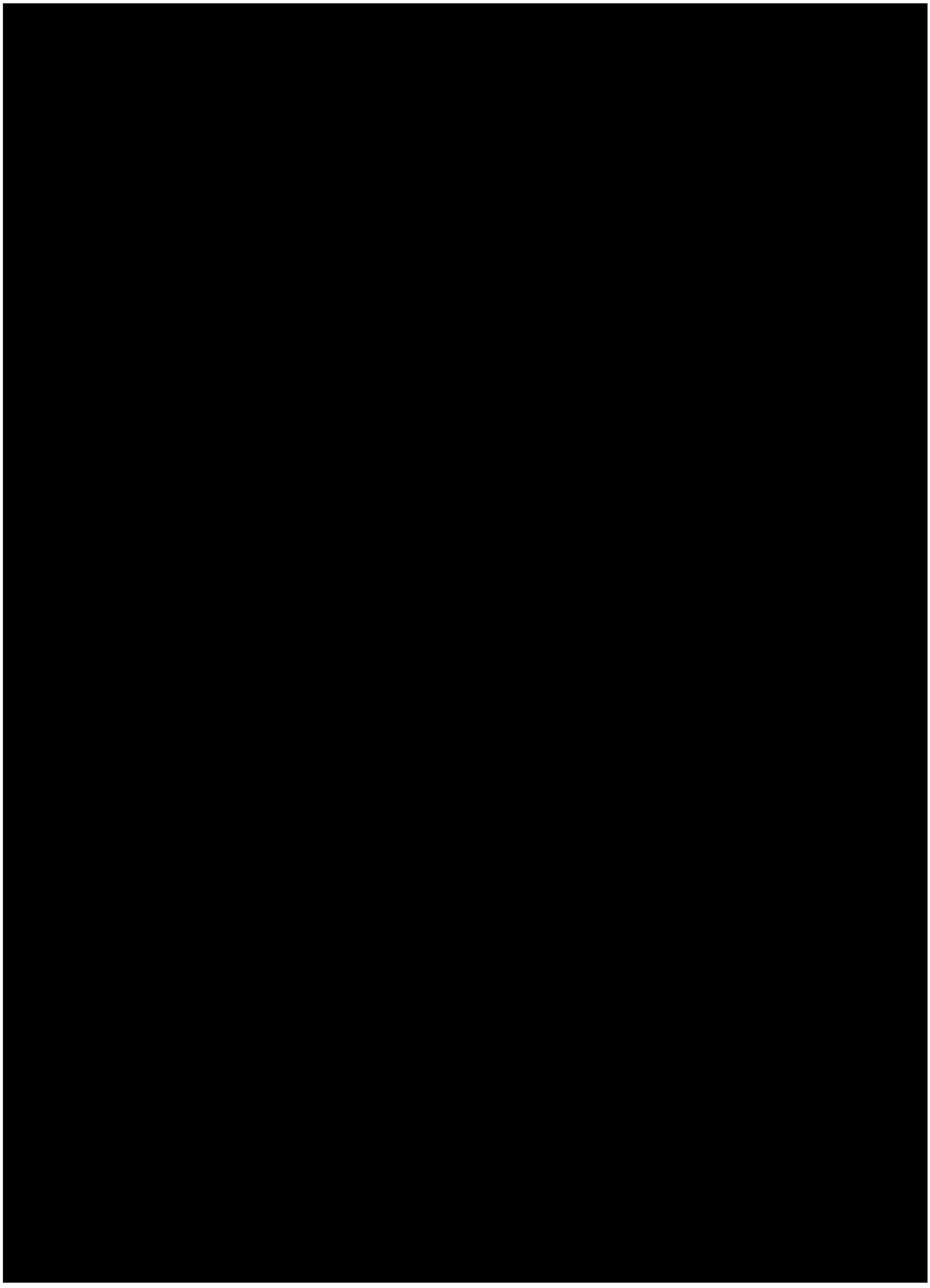
14.6. Control-based imputation and analysis of the primary endpoint

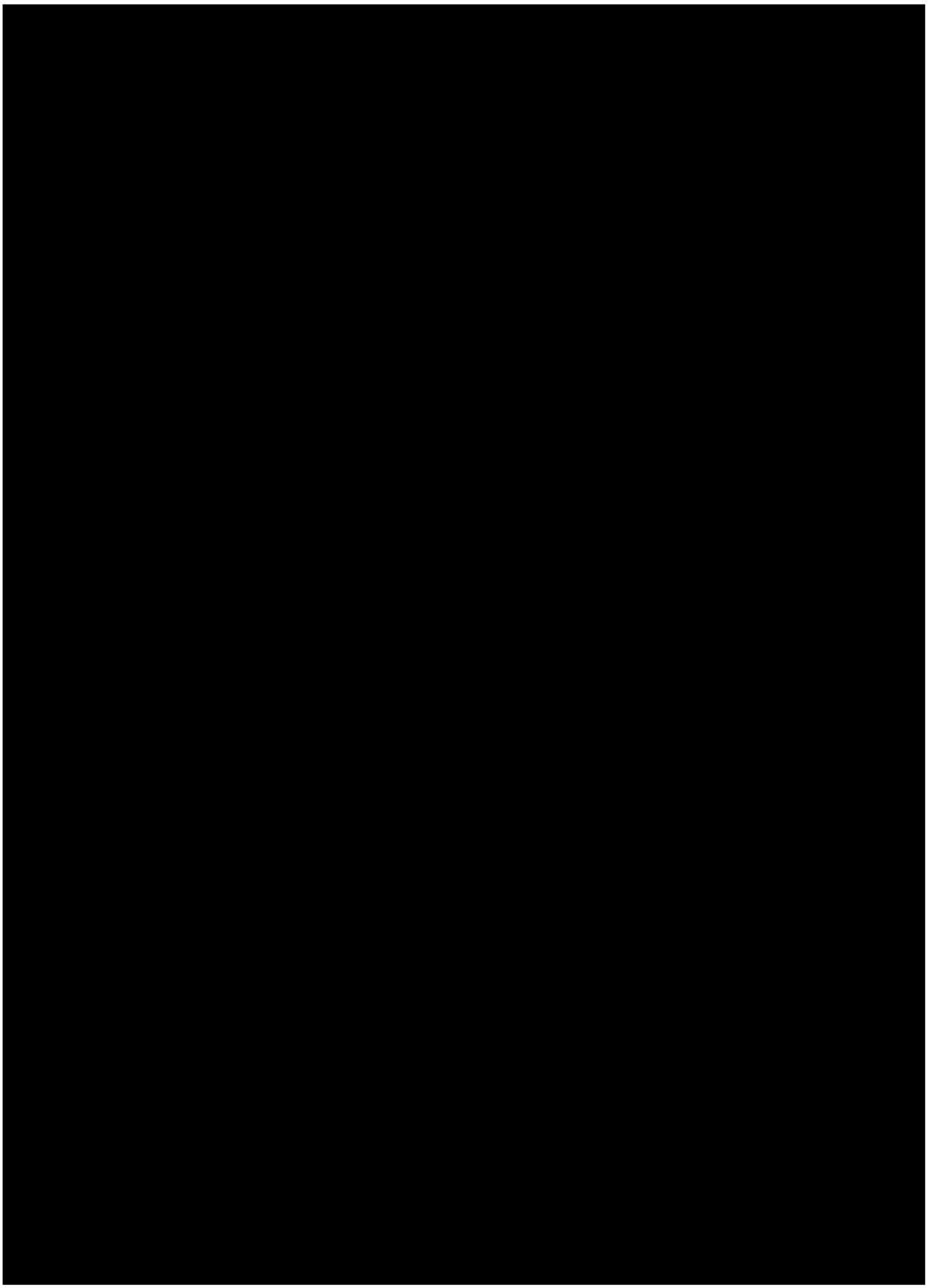
This section provides details and sample SAS programming code for the control-based imputation of the primary endpoint analysis outline in Section 8.5.2.

The control-based imputation of the primary endpoint will use a single seed from Section 14.11.

The dataset DATA_IMP is the transposed input dataset, with columns for changes in mFARS at Weeks 1-48. The OUT_IMP_CONT dataset and the code described in Table 11 creates a monotone missing dataset with columns for change from baseline by analysis study week.

Table 11: SAS code for control-based imputation of the primary endpoint







14.7. Tipping point analysis of primary endpoint

This section provides details and sample SAS programming code for the tipping point analysis outline in Section 8.5.2.1 of the SAP.

The tipping point analysis assesses the effect of potential deviations from the assumption of MAR and explores the consequences of assuming data in the omaveloxolone arm are missing not at random (MNAR) (i.e., subjects in the omaveloxolone arm with missing outcome are assumed to have lower mFARS values than subjects in the placebo arm). This analysis is based on the assumptions of monotone missingness using the regression approach and that the trajectory of mFARS changes once the subject stops use of omaveloxolone. The specific steps are as follows:

1. The missing information on mFARS prior to subject discontinuation will be imputed using the monotone method with missing data imputed based on their assigned treatment group. This step will create 100 data sets that satisfy the assumption of monotone missingness. The code for this step is provided in Table 12.

2. The missing information on mFARS will be imputed in the placebo group only for all remaining values. This is the data set that will be used for Step 3 in all analyses.
3. The remaining missing data for mFARS in the omaveloxolone arm will be multiply imputed using a shift parameter S, where S can change in both directions progressively to find the point at which statistical significance is lost or achieved. The sample SAS code for performing multiple imputation with a shift S in the omaveloxolone arm is presented in Table 13. Missing data in the placebo arm are imputed assuming MAR, using the code in Table 13 subset to the placebo subjects.
4. The multiply-imputed data will be analysed using Rubin's rule to combine the results and to obtain an estimated treatment effect and its associated significance level.
5. Steps 3 and 4 will then be repeated. The shift parameter S starts from 0, which corresponds to the primary efficacy analysis with no shift effect, and the effect potentially increased in both a negative and positive direction by a certain amount in each step until the analysis reaches the "tipping point", the point at which the effect of omaveloxolone is no longer superior to that of placebo. The more the tipping point diverges from the observed data, the more robust the conclusion based on primary efficacy analysis.

The tipping point analysis will use a single seed from Section 14.11.

The dataset DATA_IMP is the transposed input dataset, with columns for changes in mFARS at Weeks 1-48. The OUT_IMP_CONT dataset and the code described in Table 12 creates a monotone missing dataset with columns for change from baseline by analysis study week.

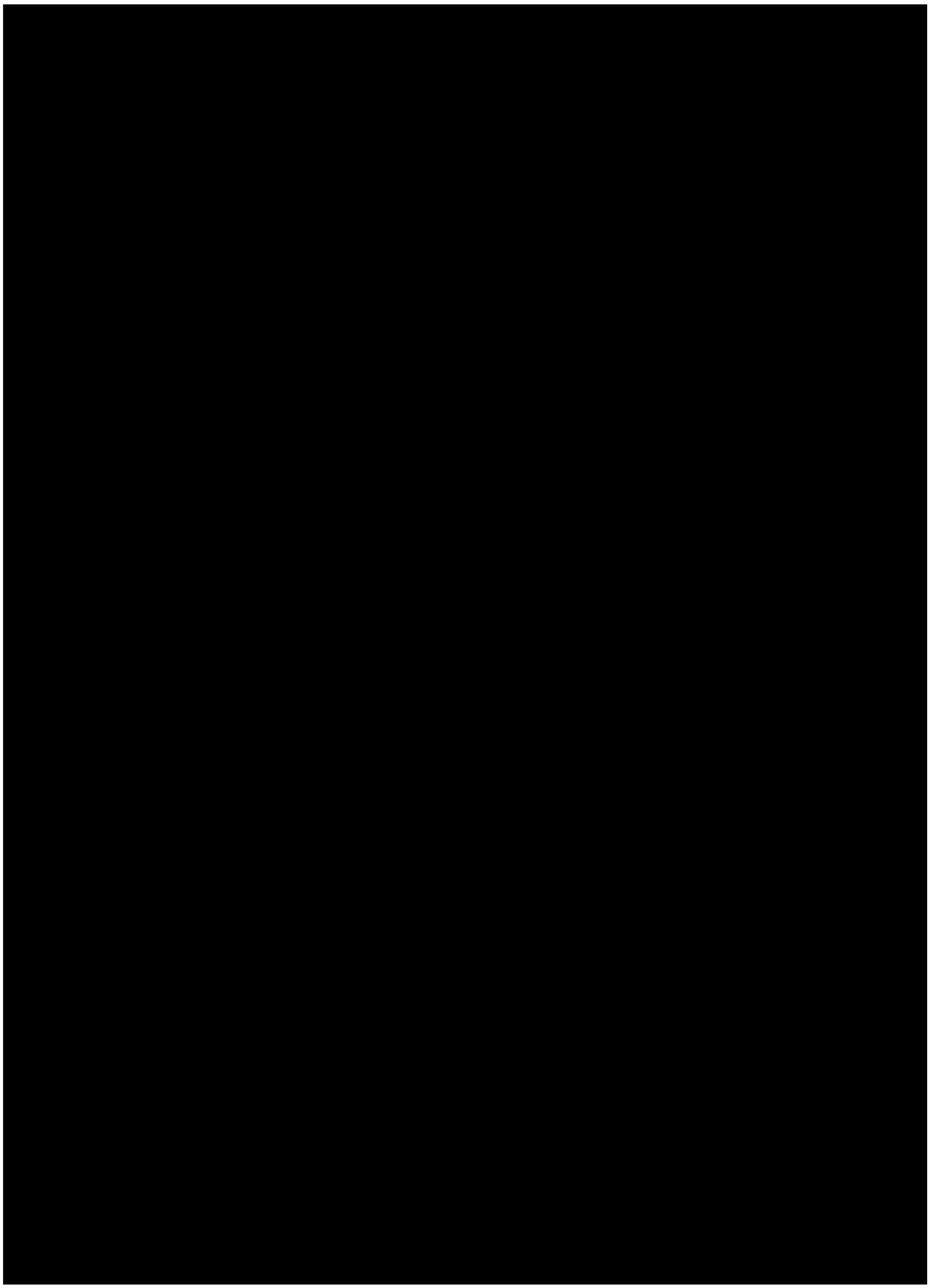
Table 12: SAS code to create datasets satisfying monotone missing condition for continuous measures

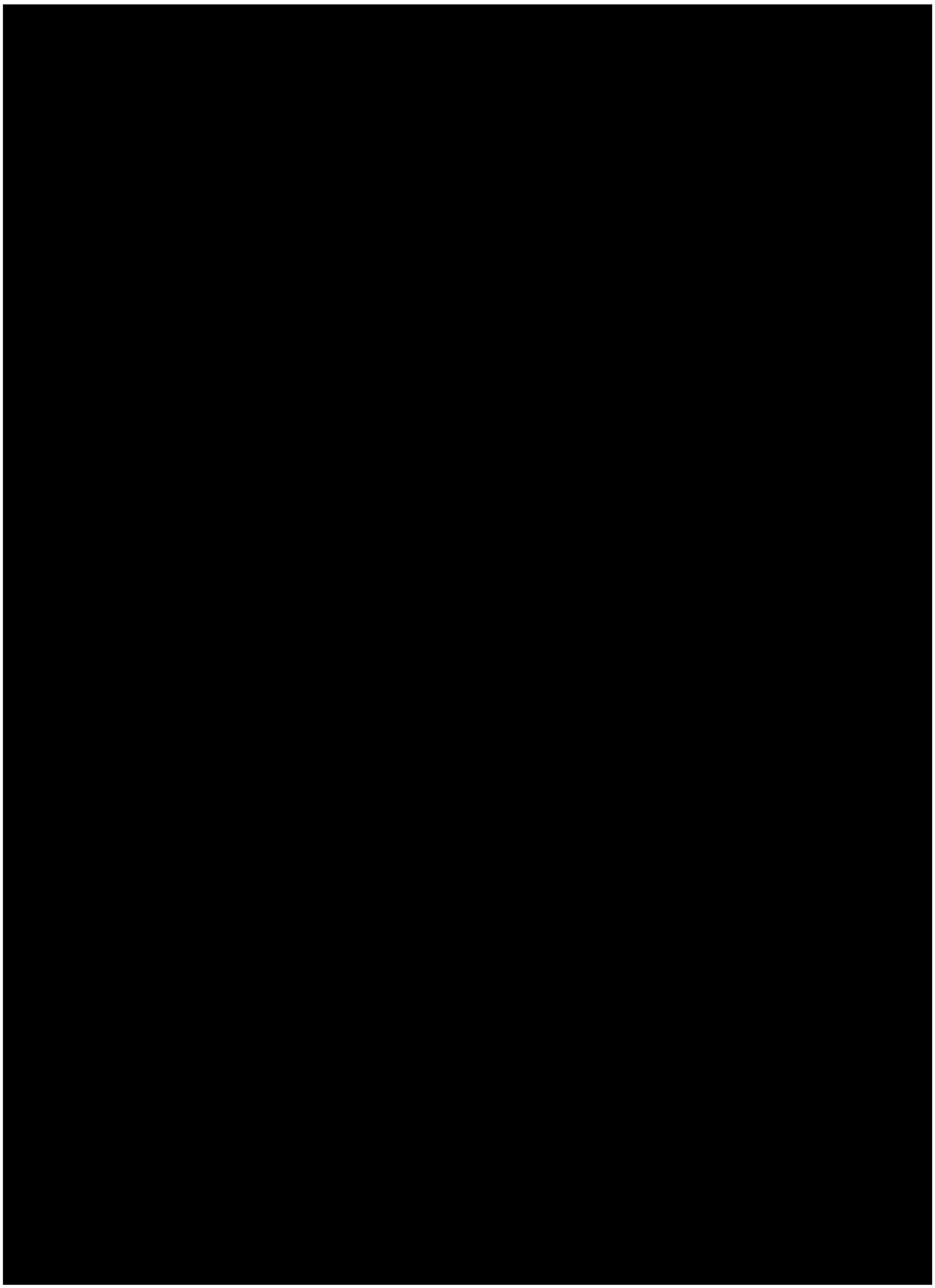


The code to impute the data sets for the tipping point analysis is in Table 13 below. Note that a single group of 100 datasets is created from this analysis. This set of 100 datasets of imputed values for the placebo data is used for the full tipping point analysis. Note that the code used to impute the placebo data and create the dataset IMPDAT_PBO only needs to be executed once. After each shift dataset is created, each of these sets of imputed datasets are then analysed using PROC MIXED and combined using PROC MIANALYZE.

For each shift value, the code in Table 14 and Table 15 is also executed within the %do loop and macro shown inTable 13. For simplicity, Table 14 and Table 15 show the mixed model code and analysis of results for a single shift parameter. Once the code from Table 15 is executed within the %do loop in Table 13 for a single shift parameter, the results for that parameter are output using the &outparms variable for a particular shift value.

Table 13: SAS code for tipping point analysis

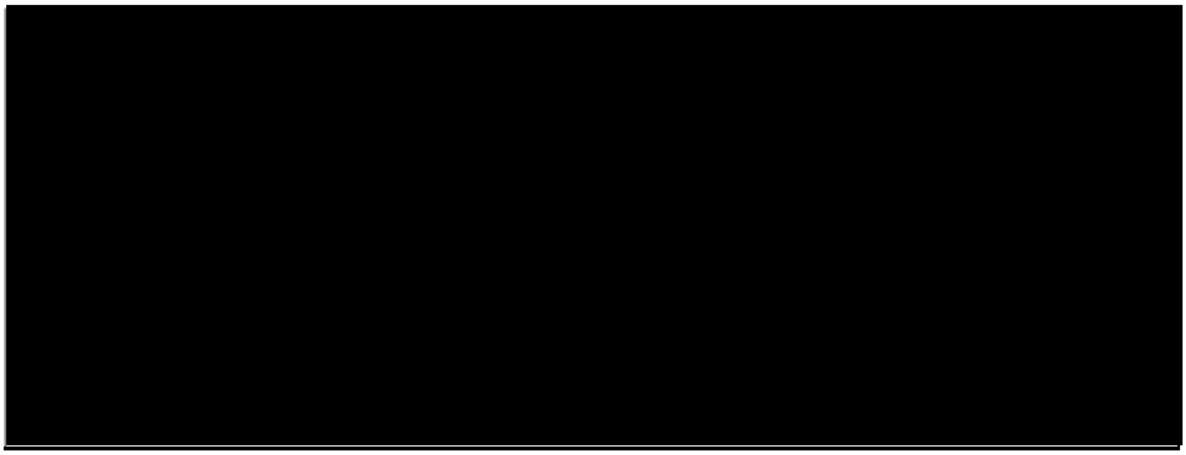






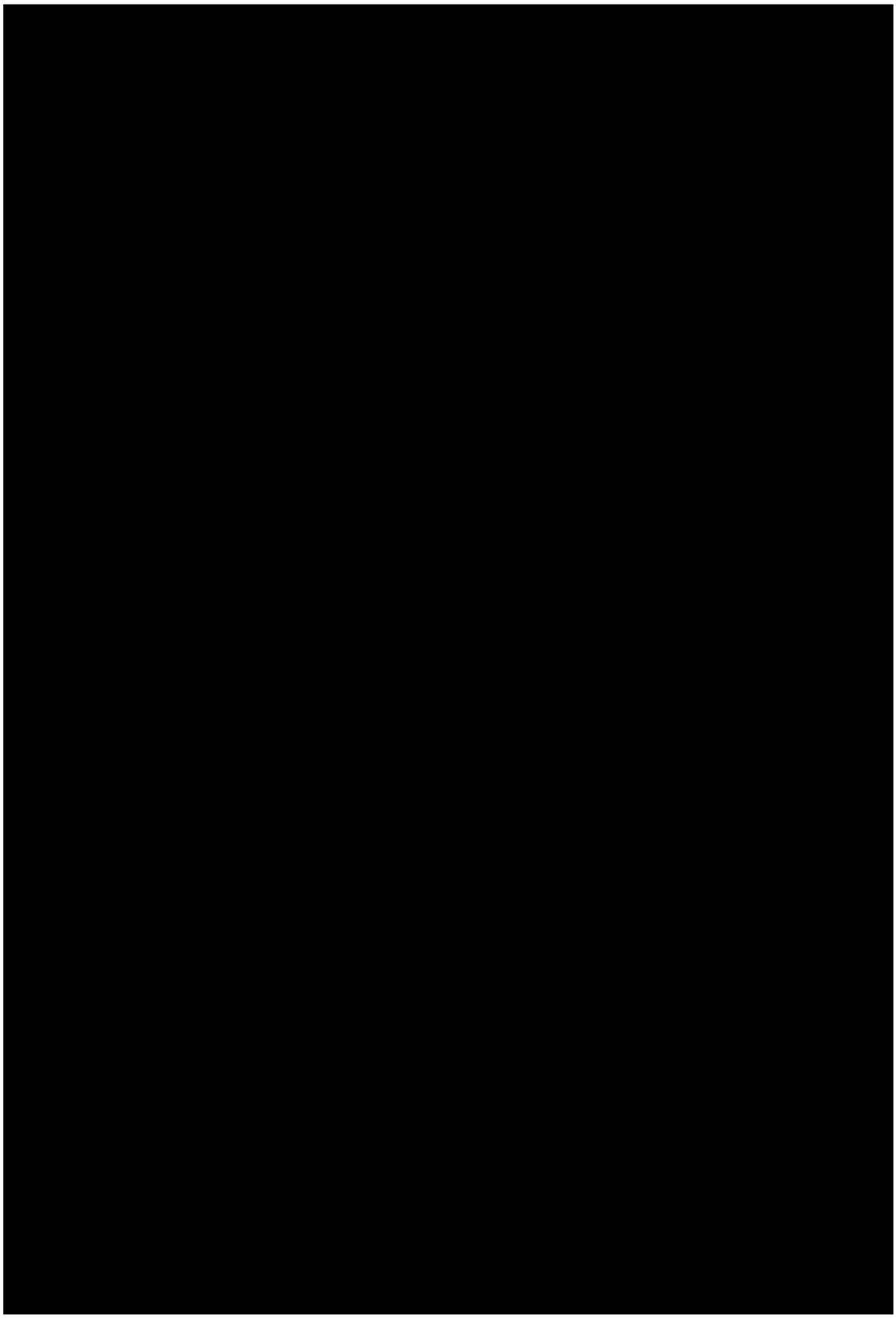
The example code below fits a mixed model to each of the sets of 100 imputed datasets created from the code in Table 13. Note that the number of datasets to be analyzed depends on the number of sensitivity parameters used.

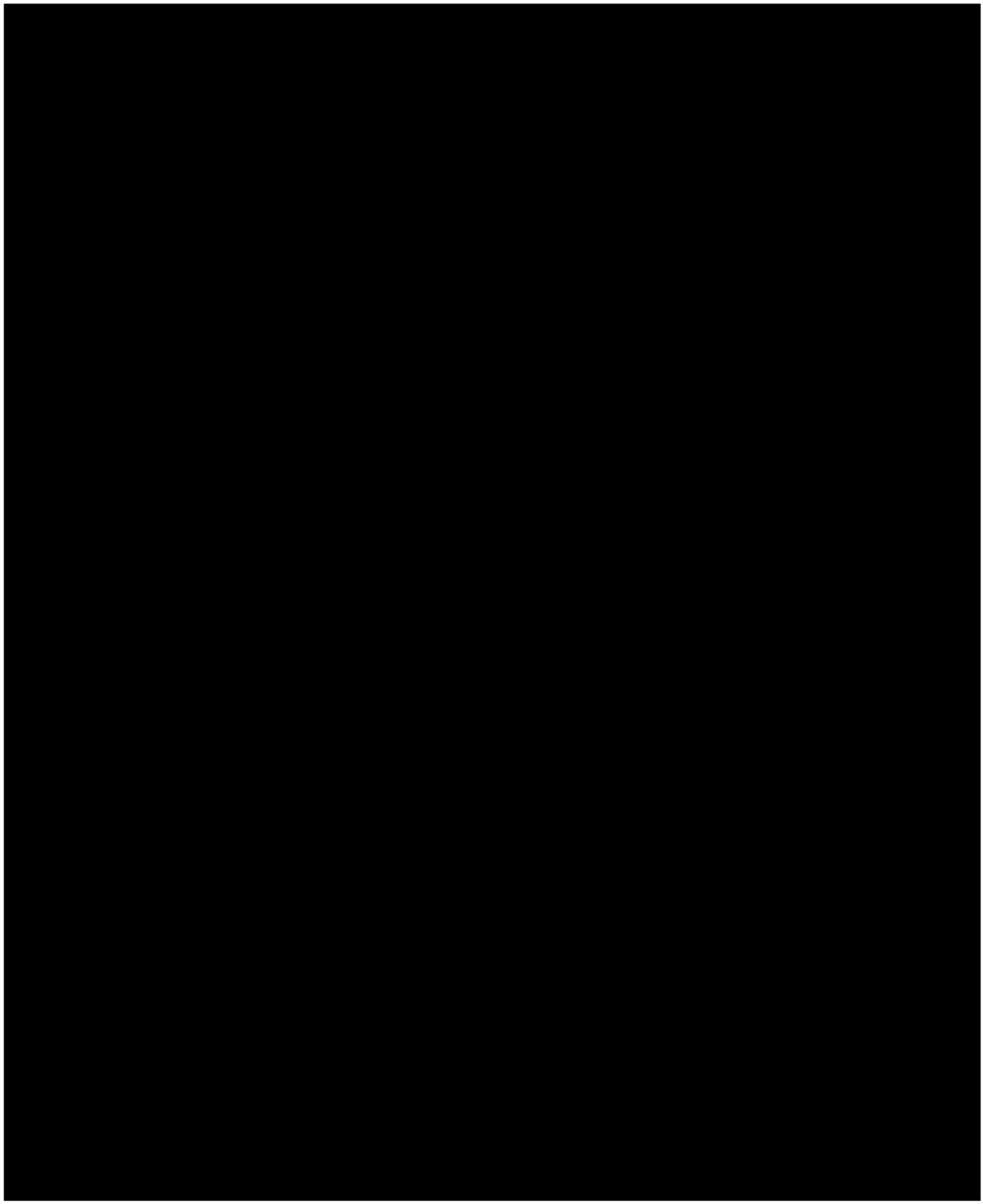
Table 14: SAS code for MMRM analysis



The input data set "COMBINE"" includes the imputed data for the treatment and placebo groups from Table 13. This dataset is then input into PROC MIANALYZE to obtain the final result for each sensitivity parameter.

Table 15: SAS code to obtain the overall estimate and standard error





14.8. ANCOVA analysis of the secondary endpoint

The key secondary endpoint of change from baseline in PGIC/CGIC at Week 48 is based on an ANCOVA model. The code for this model is provided in: Table 16.

The test of the null hypothesis of no difference between groups in the change from baseline at Week 48 is obtained from the lsmeans statement.

Table 16: SAS code for PGIC/GCIC ANCOVA



The pes cavus covariate is only utilized when analyzing results for the all randomized population (Section 6.3.2), as shown in Table 17.

Table 17: SAS code for PGIC/GCIC ANCOVA in the All Randomized Population



14.9. Missing data approaches for secondary endpoint analysis

The ANCOVA model presented in Table 16 includes the Week 48 timepoints only. All missing values are imputed based on the treatment group to which the patients were randomized. The code for where missing values are imputed based on the treatment group the subject was randomized to, is provided in Table 18.

The OUT_IMP_CONT is the transposed input dataset, with columns for PGIC/CGIC at Week 48.

Table 18: SAS code to generate imputation of missing data



The pes cavus covariate is only utilized when analyzing results for the all randomized population (Section 6.3.2), as shown in Table 19.

Table 19: SAS code to generate imputation of missing data in the All Randomized Population



The code for the remainder of the multiple imputation process is outlined in Table 20 below, and in Table 21 for the All Randomized Population. The final result is obtained from the output of PROC MIANALYZE.

Table 20: SAS code for multiple imputation of the secondary endpoint

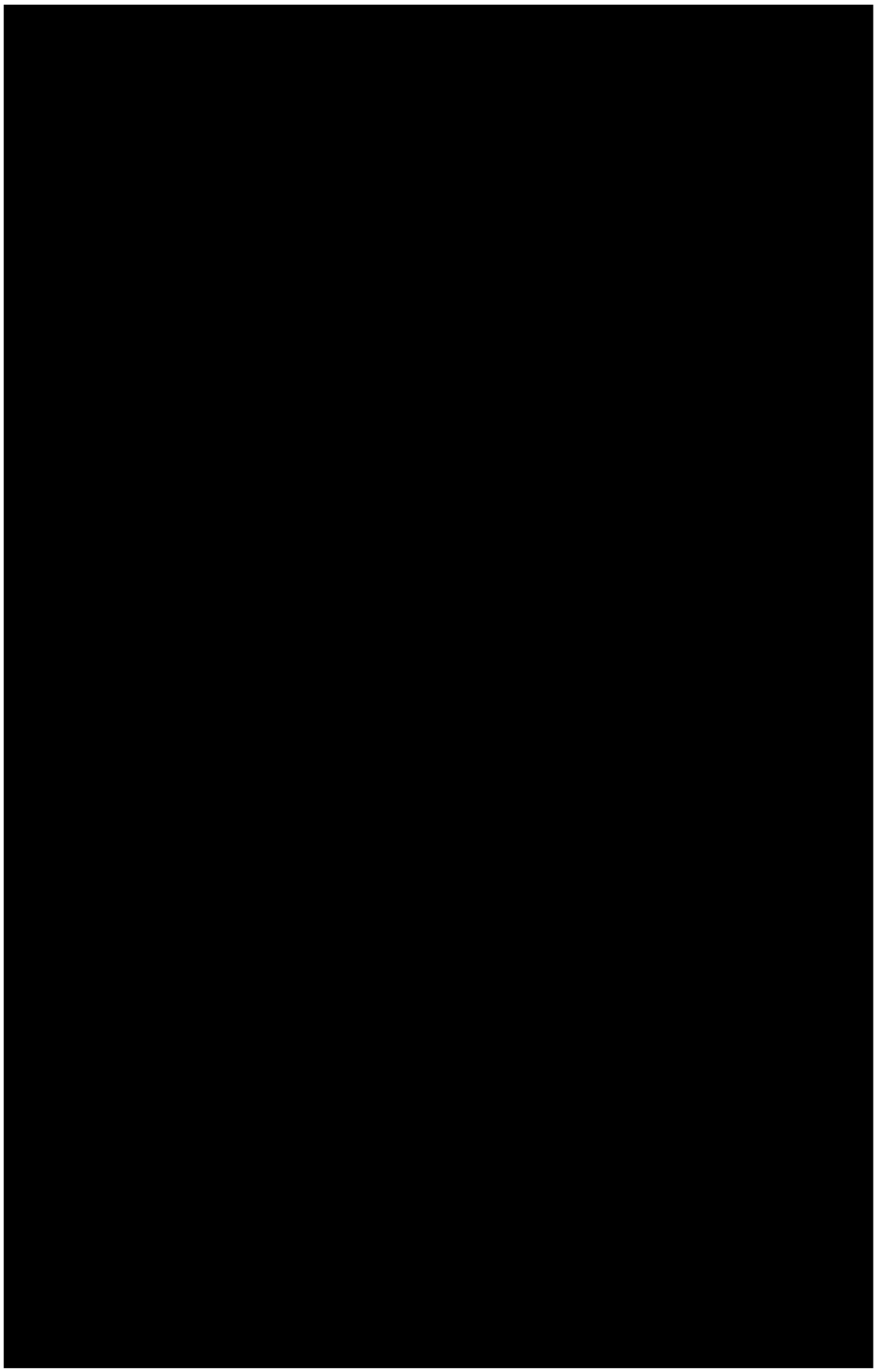
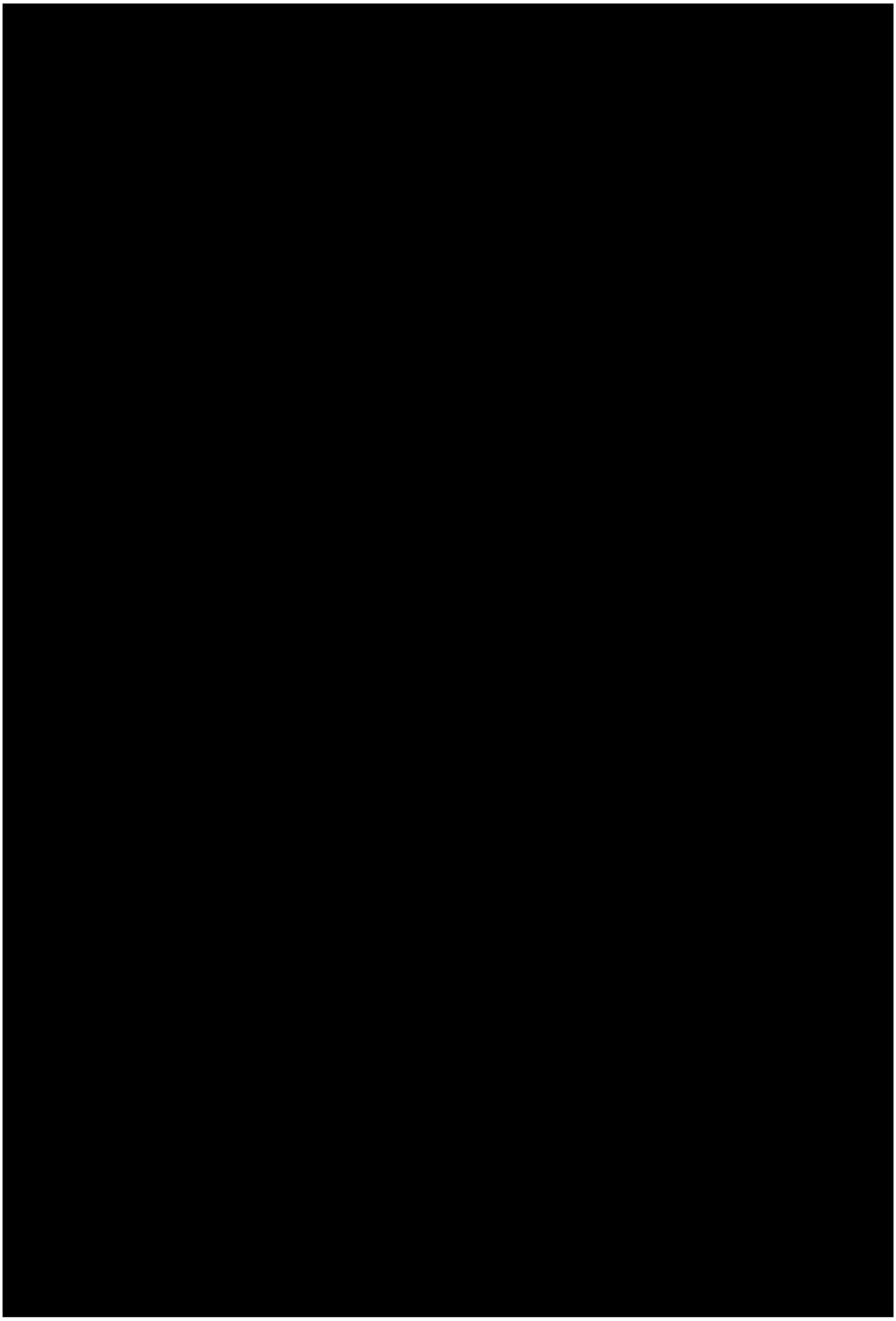




Table 21: SAS code for multiple imputation of the secondary endpoint in the All Randomized Population





Control-based imputation assumes that all data are imputed based on the placebo data. The data for the imputation is subsetted to all placebo subjects and any treated subjects with missing data at Week 48. The code for this analysis is provided in Table 22.

Table 22: SAS code to generate imputation of missing data in placebo subjects



Analysis of the results follows the code in Table 20.

14.10. Tipping point analysis of the secondary endpoint

The tipping point analysis is only performed for the Full Analysis Set (Section 6.3.1). The specific steps are as follows:

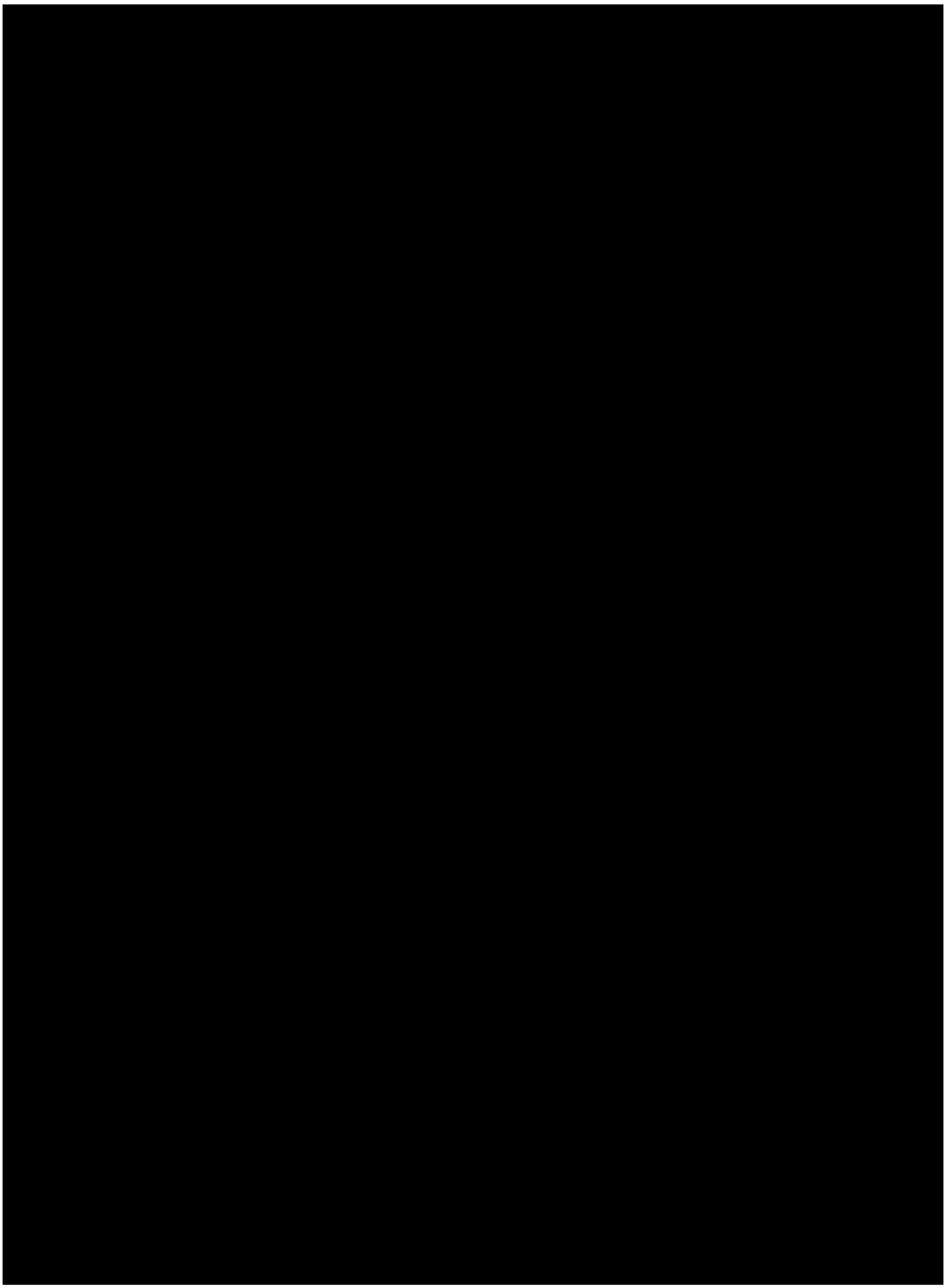
1. The missing information on PGIC/CGIC will be imputed in the placebo group only for all remaining values. This is the data set that will be used for Step 2 in all analyses.
2. The remaining missing data for PGIC/CGIC in the omaveloxolone arm will be imputed using a shift parameter S in the omaveloxolone group, where S can change in both directions progressively to find the point at which statistical significance is lost or achieved. The sample SAS code for performing multiple imputation with a shift S in the omaveloxolone arm is presented in Table 23. Missing data in the placebo arm are imputed assuming MAR, using the code in Table 23 subset to the placebo subjects.
3. The -imputed data will be analysed using standard multiple imputation combining rules to obtain an estimated treatment effect and its associated significance level.

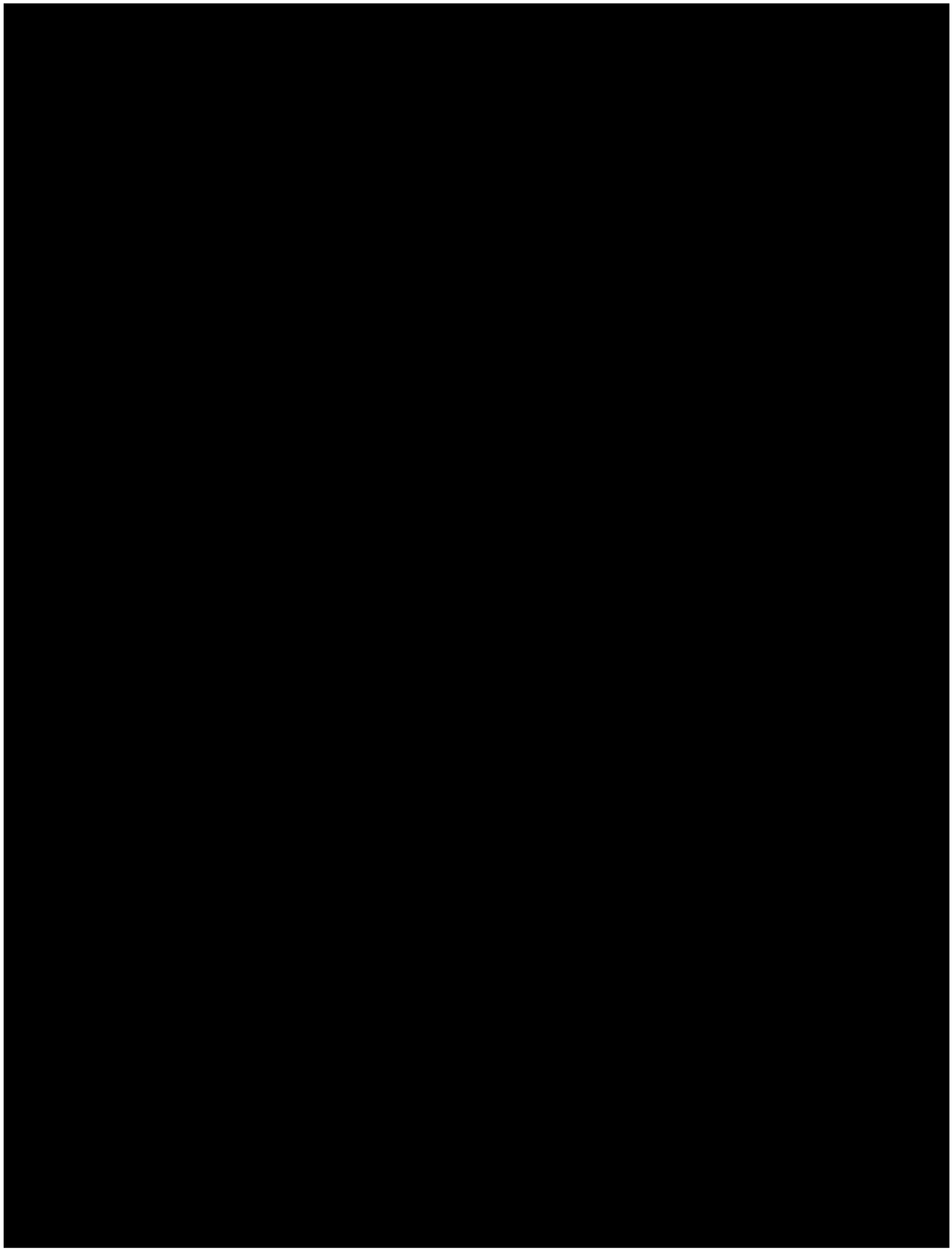
4. Steps 3 and 4 will then be repeated. The shift parameter S starts from 0, which corresponds to the primary efficacy analysis with no shift effect, and the effect potentially increased in both a negative and positive direction by a certain amount in each step until the analysis reaches the "tipping point", the point at which the effect of omaveloxolone is no longer superior to that of placebo. The more the tipping point diverges from the observed data, the more robust the conclusion based on primary efficacy analysis.

The tipping point analysis will use a single seed from Section 14.11.

The PGIC is a transposed dataset with columns for the Week 48 PGIC values.

Table 23: SAS Code for Tipping Point Analysis



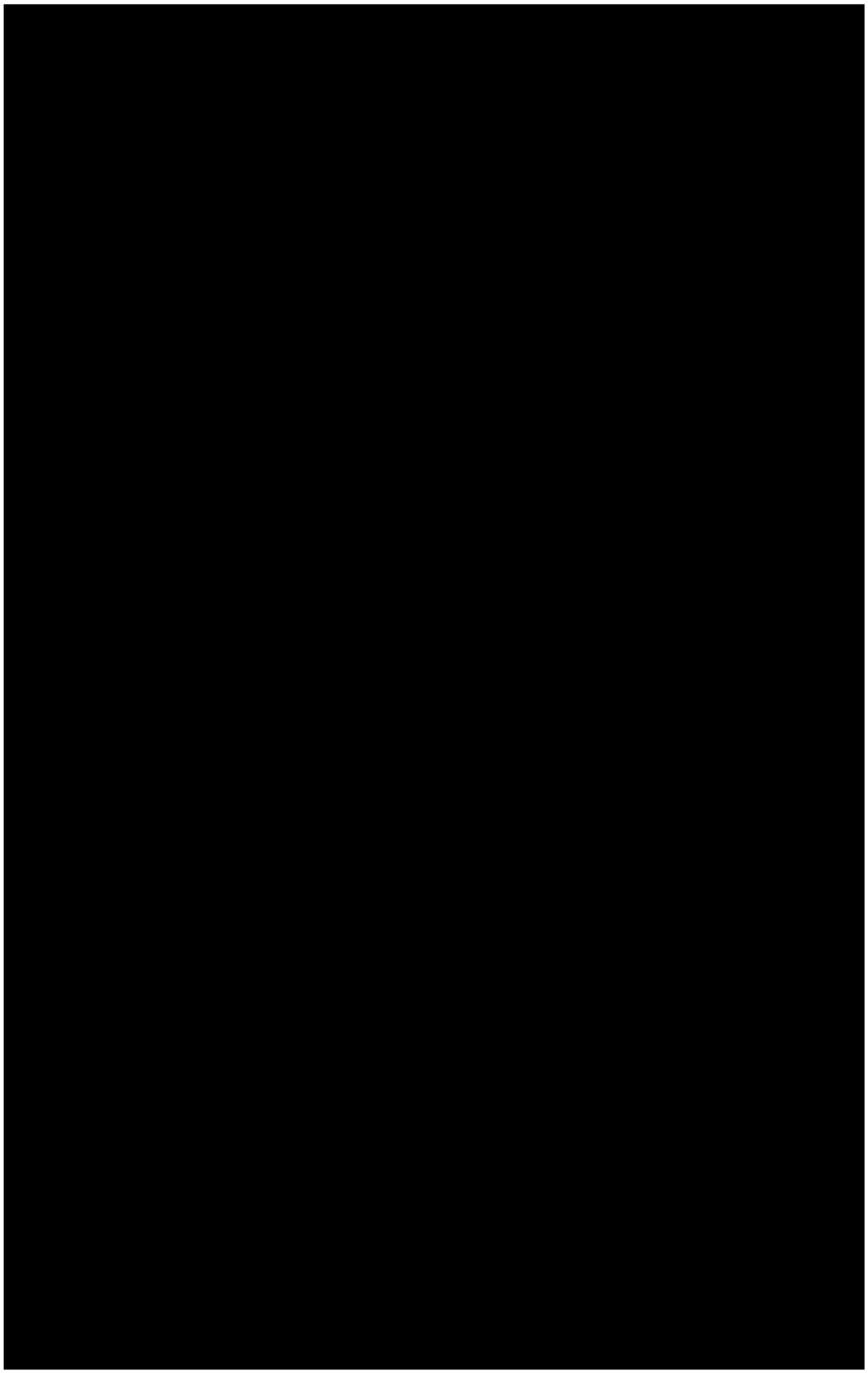


The code for the remainder of the multiple imputation process is outlined in Table 24 below. The final result is obtained from the output of PROC MIANALYZE.

For each shift value, the code in Table 24 and is also executed within the %do loop and macro shown in Table 16. For simplicity, Table 24 displays the model code and analysis of results for a single shift parameter. Once the code from Table 24 is executed within the %do loop in Table 23 for a single shift parameter, the results for that parameter are output using the &outparms variable for a particular shift value.

Table 24: SAS code for analysis of the PGIC/CGIC tipping point analysis







14.11. Random seed specification

The example code in Table 25 will be used to generate the random seeds for the analyses outlined above. The MASTER seed will be 406191150. The NUMSEED will be 13, and the seeds will be used in the order in which they are generated for the analyses.

Table 25: SAS code for random seed generation



Table 26: Pre-Specified Seed Values and Analyses

Number	Seed	Description
1	1583283832	Tipping point analysis of the primary endpoint: FAS population
2	1495344738	Treatment-based imputation of the primary endpoint: FAS population
3	188497609	Control-based imputation of the primary endpoint: FAS population
4	16733020	ANCOVA analysis of the secondary endpoint: PGIC in the FAS population
5	1664555052	Tipping point analysis of the secondary endpoint: PGIC in the FAS population
6	294012164	Control-based imputation of the secondary endpoint: PGIC in the FAS population
7	284685512	ANCOVA analysis of the secondary endpoint: CGIC in the FAS population
8	987560035	Tipping point analysis of the secondary endpoint: CGIC in the FAS population
9	1955031032	Control-based imputation of the secondary endpoint: CGIC in the FAS population
10	656058139	ANCOVA analysis of the secondary endpoint: PGIC in the PCP population
11	960440053	ANCOVA analysis of the secondary endpoint: PGIC in the ARP population
12	2120233908	ANCOVA analysis of the secondary endpoint: CGIC in the PCP population
13	1207777846	ANCOVA analysis of the secondary endpoint: CGIC in the ARP population

14.12. Other Maximal Exercise Testing Parameters

Table 27: Summary of Other Maximal Exercise Testing Parameters

	At Rest (Prior to Test)	At Maximal Work (End of Test)	Change from Resting[c]
Heart rate (bpm)	X	X	X
VO_2[a] (mL/kg/min)	X	X	X
O_2 pulse[b] (mL/beat)	X	X	X
Systolic blood pressure (mmHg)	X	X	X
Diastolic blood pressure (mmHg)	X	X	X
Oxygen saturation (SpO_2) (%)	X	X	X
Forced vital capacity (FVC) (L)	X		
Forced expiration volume within one second (FEV1) (L)	X		
Maximum voluntary ventilation (MVV) (L/min)	X		
%Predicted FVC (%)	X		
VCO_2[a] (mL/kg/min)		X	
Anaerobic threshold (L/min)		X	
Time to reach maximal work (Min:Sec)		X	
Maximum rating of perceived exertion (BORG; 0-20 scale)		X	
Maximum minute ventilation (VE) BTPS (L/min)		X	
Tidal Volume (L)		X	
Maximum respiratory rate (breaths/min)		X	
Breathing Reserve (%)		X	

[a] Analysis of maximal VO2, VCO2, and resting VO2 are performed using adjustments for baseline weight:
- Maximal_VO_2_adj=Maximal_VO_2*1000/Baseline weight (kg);
- Maximal_VCO2_adj=Maximal_VCO2*1000/Baseline weight (kg);
- Resting_VO_2_adj=Resting_VO_2*1000/Baseline weight (kg);

[b] O_2 pulse are calculated for analysis: O_2 pulse = (Maximal_VO_2 (in units of mL/min)*1000)/Maximal_HR

[c] The change from resting (delta) for parameters measured at resting and at the end of the maximal exercise test is defined as:
- Delta = Parameter at maximal work - Parameter at resting

The change from baseline is defined for parameters measured at resting and at the end of the maximal exercise test is defined as:

- Change from baseline in Delta =
  [Parameter at maximal work(At Visit X) - Parameter at resting(At Visit X)] -
  [Parameter at maximal work(At Baseline) - Parameter at resting(At Baseline)]

14.13. List of Laboratory Tests

Blood samples are collected throughout the study for hematology, chemistry, and urinalysis for clinical laboratory evaluation. Test panels include the following:

Table 28: List of Laboratory Tests

Hematology	Chemistry	Urinalysis
Hematocrit	Blood urea nitrogen (BUN)	Specific gravity
Hemoglobin	Creatinine	Ketones
HbA1C	Total bilirubin	pH
Red blood cell (RBC) count	Alanine aminotransferase (ALT)	Protein
White blood cell (WBC) count	Aspartate aminotransferase (AST)	Blood
Neutrophils	Alkaline phosphatase (ALP)	Glucose
Bands (if detected)	Ferritin	Urobilinogen
Lymphocytes	Sodium	Bilirubin
Monocytes	Potassium	
Basophils (if detected)	Calcium	
Eosinophils (if detected)	Inorganic phosphorus	
Absolute platelet count	Magnesium	
Mean corpuscular hemoglobin (MCH)	Chloride	
Mean corpuscular volume (MCV)	Bicarbonate	
Mean corpuscular hemoglobin concentration (MCHC)	Uric acid	
Reticulocyte count	Cholesterol	
	Total protein	
	Glucose	
	Triglycerides	
	Albumin	
	Creatine phosphokinase (CPK)	
	Lactate dehydrogenase (LDH)	
	High-density lipoprotein cholesterol (HDL-C)	
	Low-density lipoprotein cholesterol (LDL-C)	
	Very-low-density lipoprotein cholesterol (VLDL-C)	
	Gama-glutamyl transpeptidase (GGT)	
	Estimated glomerular filtration rate (eGFR) using the CKD-EPI / Schwartz Bedside Equation	

14.14. SF-36 Scoring

This study also uses the SF-36®, a 36-item health survey to evaluate the patient's physical, social and mental well-being. It comprises 35 items that provide the generic core and an overall heath rating item.

For the SF-36®, the raw scale scores will be calculated for each of the eight dimensions (Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional, Mental Health) from 0 to 100 with 0 representing poor health and 100 good health as described in the SF-36® Health Survey Manual.

According to the SF-36® Health Survey Manual, the pre-coded item values of the SF-36® should first be recoded as follows to generate the final item value:

Table 29: SF-36® final item values

Recoding of item			Final item value
3a, 3b, 3c, 3d, 3e, 3f, 3g, 3h, 3i, 3j, 4a, 4b, 4c, 4d, 5a, 5b, 5c, 9b, 9c, 9f, 9g, 9i, 10, 11a, 11c			= precoded item value
6, 11b, 11d			= 6 – precoded item value
9a, 9d, 9e, 9h			= 7 – precoded item value
1:	precoded item value =	1	= 5.0
		2	= 4.4
		3	= 3.4
		4	= 2.0
		5	= 1.0
7:	precoded item value =	1	= 6.0
		2	= 5.4
		3	= 4.2
		4	= 3.1
		5	= 2.2
		6	= 1.0
8 if Item 7 is answered:		1	= 6.0 (if Item 7 = 1)
	precoded item value =	1	= 5.0 (if Item 7 ≠ 1)
		2	= 4.0
		3	= 3.0
		4	= 2.0
		5	= 1.0
8 if Item 7 is not answered:		1	= 6.0
	precoded item value =	2	= 4.75
		3	= 3.5
		4	= 2.25
		5	= 1.0

The raw scale scores will then be calculated for each of the eight dimensions as the sum of the corresponding final item values as follows:

Table 30: SF-36® raw scale scores

Dimension	Sum of final item values	Possible raw score		
		Lowest	Highest	Range
Physical Functioning (PF)	3a, 3b, 3c, 3d, 3e, 3f, 3g, 3h, 3i, 3j	10	30	20
Role-Physical (RP)	4a, 4b, 4c, 4d	4	20	16
Bodily Pain (BP)	7, 8	2	12	10
General Health (GH)	1, 11a, 11b, 11c, 11d	5	25	20
Vitality (VT)	9a, 9e, 9g, 9i	6	22	16
Social Functioning (SF)	6, 10	2	10	8
Role-Emotional (RE)	5a, 5b, 5c	3	15	12
Mental Health (MH)	9b, 9c, 9d, 9f, 9h	7	27	20

A dimension score will only be calculated if at least half of the items in the scale are not missing; otherwise the raw scale score will be set to missing. If fewer than half the items are missing, the missing items will be imputed as the mean across completed items in the same dimension scale.

The following formula will be used to convert thee raw scale scores into transformed scale scores:

$$\text{Transformed scale} = \frac{\text{(Actual raw score – lowest possible score)}}{\text{Possible raw score range}} \times 100$$

Additionally, two summary scale scores, based on weighted combinations of the eight dimension scores, will be computed: the Physical Component Summary (PCS) and the Mental Component Summary (MCS). The algorithm for calculating PCS and MCS, which consists of three steps, is based on Table 31.

Table 31: SF-36® summary scores

	US population characteristics		Factor score coefficients	
SF-36® dimension	Mean	Standard Deviation	Coefficients for the aggregate physical score	Coefficients for the aggregate mental score
PF	84.52404	22.89490	0.42402	-0.22999
RP	81.19907	33.79729	0.35119	-0.12329
BP	75.49196	23.55879	0.31754	-0.09731
GH	72.21316	20.16964	0.24954	-0.01571
VT	61.05453	20.86942	0.02877	0.23534
SF	83.59753	22.37642	-0.00753	0.26876
RE	81.29467	33.02717	-0.19206	0.43407
MH	74.84212	18.01189	-0.22069	0.48581

Step 1: Calculate the z-score standardizations of SF-36® subscales in this study by subtracting the US population mean and dividing by the US population SD for each dimension. For instance, the z-score for PF is given by Z = (PF-84.52404)/22.89490).

Step 2: Calculate the aggregate physical and mental component scores (AGG_PHYS and AGG_MENT) by summing up all the z-scores weighted with the factor score coefficients corresponding to PCS and MCS respectively. The PCS and MCS will only be calculated if at least half of the dimension scores are not missing.

Aggregate Physical Component Score = AGG_PHYS. The score, a linear combination of standardized SF-36 subscale scores, is given in the following formula:

$$\begin{aligned}AGG_PHYS = {} & 0.42402 \times (PF - 84.52404)/22.89490 \\ & + 0.35119 \times (RP - 81.19907)/33.79729 \\ & + 0.31754 \times (BP - 75.49196)/23.55879 \\ & + 0.24954 \times (GH - 72.21316)/20.16964 \\ & + 0.02877 \times (VT - 61.05453)/20.86942 \\ & - 0.00753 \times (SF - 83.59753)/22.37642 \\ & - 0.19206 \times (RE - 81.29467)/33.02717 \\ & - 0.22069 \times (MH - 74.84212)/18.01189\end{aligned}$$

Aggregate Mental Component Score = AGG_MENT. The score, a linear combination of standardized SF-36 subscale scores, is given in the following formula:

$$\begin{aligned}AGG_MENT = {} & -0.22999 \times (PF - 84.52404)/22.89490 \\ & - 0.12329 \times (RP - 81.19907)/33.79729 \\ & - 0.09731 \times (BP - 75.49196)/23.55879 \\ & - 0.01571 \times (GH - 72.21316)/20.16964 \\ & + 0.23534 \times (VT - 61.05453)/20.86942 \\ & + 0.26876 \times (SF - 83.59753)/22.37642 \\ & + 0.43407 \times (RE - 81.29467)/33.02717 \\ & + 0.48581 \times (MH - 74.84212)/18.01189\end{aligned}$$

Step 3: Calculate the T-score transformation of the aggregate component scores:

$$PCS = 50 + (AGG_PHYS \times 10)$$

$$MCS = 50 + (AGG_MENT \times 10)$$

STATISTICAL ANALYSIS PLAN EXTENSION

VERSION: 6.0

DATE OF PLAN:

12 JAN 2026

BASED ON:

Protocol Version 14.0 (21 APR 2025)

STUDY DRUG:

RTA 408, OMAVELOXOLONE

PROTOCOL NUMBER:

408-C-1402, Extension

STUDY TITLE:

A Phase 2 Study of the Safety, Efficacy, and Pharmacodynamics of RTA 408 in the Treatment of Friedreich's Ataxia

SPONSOR:

Reata Pharmaceuticals, Inc.

5320 Legacy Drive

Plano, Texas 75024

(972) 865-2219

This study is being conducted in compliance with good clinical practice, including the archiving of essential documents.

SIGNATURE PAGE

This document has been reviewed and approved by:		
[redacted]	DocuSigned by: [redacted] Signature	[redacted] Date
[redacted]	DocuSigned by: [redacted] Signature	[redacted] Date
[redacted] SMT Medical Lead	Signed by: [redacted] Signature	[redacted] Date

TECHNICAL SUMMARY REPORT (TSR)

Name of Sponsor/Company **Reata Pharmaceuticals**	**Individual Study Table Referring to Part of the Dossier:** **Volume:**	***(For National Authority Use Only)*:**
Name of Finished Product: Omaveloxolone capsules	**Page:**	
Name of Active Ingredient: Omaveloxolone		
Title Of Study: A Phase 2 Study of the Safety, Efficacy, and Pharmacodynamics of RTA 408 in the Treatment of Friedreich's Ataxia		
Studied period (years): Until the drug is available through commercial channels or alternate post-trial access mechanisms	**Phase of development:** 2	
Objectives: To provide continuing open-label treatment with RTA 408 while collecting ongoing safety and tolerability data with RTA 408		
Methodology: The extension examines the long-term safety and tolerability of omaveloxolone in qualified patients with Friedreich's ataxia following completion of Part 1 or Part 2. Patients are not unblinded to study treatment in Part 1 or Part 2 upon entering the extension study. Patients receive open-label omaveloxolone (150 mg) once daily until the drug is available through commercial channels, or until patient withdrawal, whichever is sooner. All patients in the extension follow the same visit and assessment schedule except United Kingdom (UK) patients. Extension Day 1 is defined as the first day treatment is dispensed to the patient for the extension, following completion of Part 1 or Part 2. All other extension visits are relative to extension Day 1. Patients are scheduled for in-person assessments during treatment in the extension at Day 1, Week 4, Week 24, and every 24 weeks thereafter. Patients are also assessed by telephone contact on Day 14. Patients in the UK are scheduled for in-person assessments during treatment in the extension at Day 1, Weeks 4, 10, 16, and 24, and every 24 weeks thereafter. Patients in the UK are assessed by telephone contact on Day 14. Since Part 2 patients will be followed continuously within the study, eligibility for the extension is assessed at the Week 52 visit, and a separate screening visit is not required. After confirming eligibility for Part 2 patients, study treatment for extension Day 1 should be dispensed following the Week 52 assessments. Part 1 patients will not have been followed continuously within this study and therefore must complete a screening visit to confirm eligibility. The conduct of the extension phase, according to protocol specifications, is impacted by the COVID-19 pandemic. As a result, and as of Version 10 and 10.1 of the protocol, modifications intended to address access to and administration of investigational product, and adherence to protocol-specified visits and laboratory assessments are implemented. These modifications are described in Protocol Appendix 13 (COVID-19 Mitigations).		
Number of Subjects (planned and analyzed): Up to 172 (Part 1 [$n \leq 69$]; Part 2 [$n \leq 103$])		

Diagnosis and main criteria for inclusion (see Protocol Appendix 1):
Patients must complete 12 weeks of treatment in Part 1, or 48 weeks of treatment in Part 2, and have no major protocol deviations.

1. Have adequate kidney function defined as an estimated glomerular filtration rate ≥ 60 mL/min/1.73 m^2 using the Modification of Diet in Renal Disease 4 variable formula
2. Have a left ventricular ejection fraction ≥ 40% (based on echocardiogram performed at Screening Visit or within 90 days prior to Screening Visit)
3. Be able to swallow capsules
4. Be willing and able to cooperate with all aspects of the extension
5. Be willing to practice medically acceptable methods of birth control
6. Provide written informed consent for study participation, approved by the appropriate Institutional Review Board
7. Have been enrolled in Part 1 or Part 2 and completed assessments through the follow-up visit with no major protocol deviations that would, in the opinion of the investigator, deem the patient unsuitable for enrollment in the extension phase of the study

Test product, dose and mode of administration: Capsules containing omaveloxolone at the 50-mg strength are used in this study. Treatment kits containing 3 bottles (each containing 30 count of active capsules), or 1 bottle (containing 90 count of active capsules) are provided to patients for self-administration. Patients are instructed to self-administer 3 capsules of study drug orally once daily in the morning on an empty stomach (approximately 1 hour before or 2 hours after eating).

Duration of treatment: RTA 408 administered orally once daily and with extended access until the drug is available through commercial channels or alternate post-trial access mechanisms, or until patient withdrawal, whichever is sooner.

Reference therapy, dose and mode of administration: No reference therapy

Criteria for evaluation: Results of physical examinations, laboratory results, vital sign measurements, weight, AEs, SAEs, concomitant medications, pregnancy tests (as indicated), FARS, ADL, 9-HPT, 25-foot timed walk test.

Statistical methods: The sample size for the extension is limited to the number of patients having completed Part 1 or Part 2. As the extension is of an open-label design with no comparator group, all statistical analyses will be descriptive. The summary tables will be presented for the overall group of patients, and also split by previous treatment groups (i.e., RTA 408 or placebo).

Benefit-Risk Assessment:
Treatment with RTA 408 (omaveloxolone) provides the possibility of benefit based on the efficacy of omaveloxolone previously demonstrated in the present Study 408-C-1402, involving adult and adolescent participants with Friedreich's ataxia aged 16 years and older.
The risks of participation are primarily those associated with adverse reactions to the study treatment, although there may also be some discomfort from collection of blood samples and other study procedures. More information about the known and expected benefits, risks, and reasonably anticipated AEs associated with omaveloxolone may be found in the Investigator's Brochure.
The potential risks related to participation in this study are justified by the anticipated benefit to participants. The risk to participants in this study may be minimized by compliance with the eligibility criteria and study assessments and procedures.
All safety findings will be closely monitored and reviewed by a DSMB in Part 1 and Part 2, and by the Sponsor during the extension phase of the study.

TABLE OF CONTENTS

TITLE PAGE 1
SIGNATURE PAGE 2
TECHNICAL SUMMARY REPORT (TSR) 3
1. LIST OF ABBREVIATIONS 9
2. INTRODUCTION 11
3. STUDY OBJECTIVES AND ENDPOINTS 11
3.1. Study Objectives 11
4. STUDY DESIGN 12
4.1. Summary of Study Design 12
4.2. Definition of Study Drugs 12
4.3. Sample Size Considerations 12
4.4. Randomization 13
4.5. Treatment Analysis Group Assignment 13
4.6. Clinical Assessments 13
5. PLANNED ANALYSES 16
5.1. Planned Interim Analyses 16
5.2. Final Analyses and Reporting 16
6. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING 17
6.1. General Summary Table and Individual Subject Data Listing Considerations 17
6.2. Analysis Populations 18
6.2.1. Safety Population 18
6.2.2. Subgroups 18
6.3. Baseline Definition 19
6.3.1. Efficacy Assessments 19
6.3.1.1. Modified Friedreich's Ataxia Rating Scale without Prior Exercise Testing 19
6.3.1.2. Other Efficacy Assessments 19
6.3.2. Safety Assessments 20
6.4. Derived and Transformed Data 20

6.4.1. Demographic and Baseline Characteristics 20
6.4.2. Baseline Age 21
6.4.3. Years Since FA Onset 21
6.4.4. Pes Cavus (yes/no) 21
6.4.5. Study Day 21
6.4.6. Change from Baseline 22
6.4.7. Visit Windows 22
6.4.8. Clinical Laboratory Results Outside the Limit of Quantification 23
6.4.9. Estimated Glomerular Filtration Rate 23
6.4.10. mFARS without Prior Exercise Testing 25
6.4.11. 9-Hole Peg Test 26
6.4.12. Timed 25-Foot Walk Test 26
6.4.13. Electrocardiogram Intervals 26
6.5. Handling of Missing Data 27
6.5.1. Missing Efficacy Data 27
6.5.2. Missing Start and Stop Dates for Concomitant Medication 27
6.5.3. Missing Start and Stop Dates for Adverse Events 27
7. STUDY POPULATIONS 29
7.1. Subject Disposition 29
7.2. Screen Failures 29
7.3. Protocol Deviation 30
7.4. Demographic and Baseline Characteristics 30
7.5. Listing of Subject Inclusion and Exclusion Criteria 31
7.6. Medical History 31
8. EFFICACY 31
8.1. Modified FARS without Prior Exercise Testing 31
8.2. 9-Hole Peg Test 32
8.3. Timed 25-Foot Walk Test 32
8.4. Patient Global Impression of Change 32
8.5. Clinical Global Impression of Change 32

8.6. Activities of Daily Living 33
9. SAFETY AND TOLERABILITY 33
9.1. Adverse Event Preferred Term and Body/Organ System Summary Tables 33
9.1.1. Missing and Partial AE Onset Dates 35
9.1.2. Summaries of AEs for SAEs, AE Dropouts, and Death 35
9.2. Exposure and Compliance 36
9.3. Concomitant and Other Medications 36
9.3.1. Missing and Partial Concomitant and Other Medication Start and Stop Dates 37
9.4. Clinical Laboratory Evaluations 37
9.4.1. Continuous Summaries of Laboratory Results 38
9.4.2. Categorical Summary of Laboratory Results 38
9.4.3. Transaminases and Total Bilirubin 39
9.5. Vital Signs 40
9.6. Echocardiogram 40
9.7. Electrocardiogram 41
9.8. Pregnancy 42
10. PHARMACOKINETICS 42
11. COVID-19 IMPACT 43
11.1. Operational Impact 43
11.2. Efficacy 44
11.3. Safety 44
11.4. Visit Window 45
12. REFERENCES 46
13. APPENDIX 47
13.1. SAP Change Summary 47
13.2. Programming Specifications 49
13.3. List of Laboratory Tests 49

LIST OF TABLES

Table 1: List of Abbreviations 9
Table 2: Extension Schedule of Assessments 15
Table 3: Analysis Visits 22
Table 4: Scoring of the Neurologic FARS 24
Table 5: Prespecified Threshold Levels for ECG Parameters 37
Table 6: List of Laboratory Tests 45

1. LIST OF ABBREVIATIONS

The list is sorted alphabetically.

Table 1: List of Abbreviations

Abbreviation	Term
9-HPT	9-hole peg test
ADL	activities of daily living
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical (drug classification system)
BMI	body mass index
BNP	B-type natriuretic peptide
CGIC	Clinical Global Impression of Change
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CSR	clinical study report
DBL	database lock
ECG	electrocardiogram
ECHO	echocardiogram
eCRF	electronic case report form
eGFR	estimated glomerular filtration rate
FA	Friedreich's ataxia
FARS	Friedreich's ataxia rating scale
FDA	Food and Drug Administration
GAA	trinucleotide composed of one guanine and two adenines

HDL	high-density lipoprotein
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IP	investigational product
LDL	low-density lipoprotein
LLD	lower limit of detection
MedDRA	Medical Dictionary for Regulatory Activities
mFARS	modified Friedreich's ataxia rating scale
NT-proBNP	N-terminal prohormone B-type natriuretic peptide
PGIC	Patient Global Impression of Change
PT	preferred term
QTcB	corrected QT interval using Bazett's formula
QTcF	corrected QT interval using Fridericia's formula
RTA 408	omaveloxolone
SAE	serious adverse event
SAP	Statistical Analysis Plan
SOC	System Organ Class
T25FW	timed 25-foot walk
TBL	total bilirubin
TEAE	treatment-emergent adverse events
UK	United Kingdom
ULD	upper limit of detection
ULN	upper limit of normal
WHO	World Health Organization

WOCBP	women of child-bearing potential

2. INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to describe the planned analyses and data displays to be included in the clinical study report (CSR) for Protocol 408-C-1402 Extension. This SAP was developed in accordance with International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) E9 guidelines. Further information can be found in the protocol. The SAP is based on:

- Protocol No. 408-C-1402, Version 14.0, dated 21-Apr-2025
- ICH guidelines E4 and E9 (Statistical Principles for Clinical Trials)
- Feedback received from Food and Drug Administration (FDA) throughout clinical development, including the following communications:
  - FDA Meeting Minutes [redacted]
  - FDA teleconference with Reata on [redacted]
  - FDA Study May Proceed letter [redacted]
  - Email communication from the Division on [redacted]

This SAP describes the study populations, how variables are derived, how missing data are handled, and details concerning the statistical methods to be used to analyze the safety and efficacy data in the 408-C-1402 extension study. Should the SAP and the protocol be inconsistent with respect to the planned analyses, the language of the SAP is governing.

The SAP version 1.0 dated 26-Aug-2019 was finalized, approved by the Sponsor, and placed on file before the first database lock (DBL) on 13-Dec-2019. Unless otherwise specified, these analyses are summarized in the CSR associated with the DBL. The CSR will describe any deviations from the planned analyses.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1. Study Objectives

The objective of the extension is to provide continuing treatment with omaveloxolone as part of this extended access program while collecting ongoing safety and tolerability data with omaveloxolone.

4. STUDY DESIGN

4.1. Summary of Study Design

The extension examines the long-term safety and tolerability of omaveloxolone in qualified patients with Friedreich's ataxia (FA) following completion of Part 1 or Part 2. Patients are not unblinded to study treatment in Part 1 or Part 2 upon entering the extension study.

Patients receive open-label omaveloxolone (150 mg) once daily until the drug is available through commercial channels, or until patient withdrawal, whichever is sooner. All patients in the extension follow the same visit and assessment schedule except United Kingdom (UK) patients. Extension Day 1 is defined as the first day treatment is dispensed to the patient for the extension, following completion of Part 1 or Part 2. All other extension visits are relative to extension Day 1. Patients are scheduled for in-person assessments during treatment in the extension at Day 1, Week 4, Week 24, and every 24 weeks thereafter. Patients are also assessed by telephone contact on Day 14. Patients in the UK are scheduled for in-person assessments during treatment in the extension at Day 1, Weeks 4, 10, 16, and 24, and every 24 weeks thereafter. Patients in the UK are assessed by telephone contact on Day 14.

Since Part 2 patients will be followed continuously within the study (ie, seamless rollover), eligibility for the extension is assessed at the Week 52 visit, and a separate screening visit is not required. After confirming eligibility for Part 2 patients, study treatment for extension Day 1 should be dispensed following the Week 52 assessments. Part 1 patients will not have been followed continuously within this study (ie, non-seamless rollover), and therefore must complete a screening visit to confirm eligibility.

The conduct of the extension phase, according to protocol specifications, is impacted by the COVID-19 pandemic. As a result, and as of Version 10 and 10.1 of

the protocol, modifications intended to address access to and administration of investigational product, and adherence to protocol-specified visits and laboratory assessments are implemented. These modifications are described in Protocol Appendix 13 (COVID-19 Mitigations).

The COVID-19-related analyses are specified in Section 11 of this SAP.

Safety data of the open-label extension study is overseen by Reata's internal safety management team.

4.2. Definition of Study Drugs

Capsules containing omaveloxolone at the 50 mg strength are used in this study.

4.3. Sample Size Considerations

The sample size for the extension is limited to the number of patients having completed 408-C-1402 Part 1 or Part 2. The study is not powered to show efficacy or safety effects, and there is no placebo comparator in this open-label extension trial. A maximum of 172 patients will be enrolled in this study, with 69 from Part 1 and 103 from Part 2.

4.4. Randomization

This is the non-randomized, open-label extension portion of the study that includes patients who previously completed Part 1 or Part 2. No randomization or blinding is performed, and patients enrolled in the extension are assigned to omaveloxolone 150 mg.

4.5. Treatment Analysis Group Assignment

The treatment analysis groups consist of:

1. Patients initially randomized to placebo in Part 2 or patients participating in Part 1 of the study, and
2. Patients initially randomized to omaveloxolone in Part 2.

4.6. Clinical Assessments

All patients in the extension follow the same visit and assessment schedule. Extension Day 1 is defined as the first day treatment is dispensed to the patient for the extension, following completion of Part 1 or Part 2. All other extension visits are relative to extension Day 1. Patients are scheduled for in-person assessments during treatment in the extension at Day 1, Week 4, Week 24, and every 24 weeks thereafter. Patients are also assessed by telephone contact on Day 14. Patients at the UK site have additional Week 10 and Week 16 in-person visits, which are included in the listings.

Table 2: Extension Schedule of Assessments

Study Day/Week	Screening[a]	Day 1[b]	Week 2 (telephone)	Week 4	Week 24 + Every 24 Weeks[k]	End of Study[k]
Day Relative to First Dose in Extension	**-30 to -1**	**1**	**14 (±3 days)**	**28 (±3 days)**	**168, 336, 504 etc. (±14 days)**	**Within 30 days after last dose**
Informed consent	X					
Inclusion/Exclusion	X	X				
Pes cavus assessment	X					
Medical history	X					
Adverse event collection		X[c]	X	X	X	X
Prior and Concomitant medications	X	X	X	X	X	X
Weight and BMI	X	X[d]		X	X	X
Vital sign measurements	X	X[d]		X	X	X
Physical examination	X	X[d]		X	X	X
Echocardiogram[e]	X					
Height	X			X[j]		
Electrocardiogram[e]	X					
Clinical Chemistry	X			X	X	X
Hematology	X			X	X	X
Urinalysis and microscopy	X			X	X	X
BNP and NT-proBNP[f]	X			X	X	X
Hepatitis B and C and HIV[g]	X					
Pregnancy test WOCBP[h]	X	X		X	X	X
Dispense study drug		X		X	X	
Study drug return and pill count				X	X	X
Neurologic FARS		X			X	X
Activities of Daily Living		X			X	X
Patient Global Impression of Change					X[i]	

Clinical Global Impression of Change					X[i]	
25-foot Timed Walk Test		X			X	X
9-Hole Peg Test		X			X	X

Abbreviations: BMI=body mass index; BNP= B-type natriuretic peptide; EC=ethics committee; eCRF=electronic case report form; FARS=Friedreich's ataxia rating scale; IRB=Institutional Review Board; NT-proBNP= N-terminal prohormone B-type natriuretic peptide; WOCBP=women of child-bearing potential.

[a] The screening visit is only required for Part 1 patients. Assessments collected at Part 2 - Week 48 or Part 2 - Week 52 should be used to assess eligibility for Part 2 patients.

[b] On Day 1, all required procedures should be performed and eligibility confirmed before study drug is dispensed in the extension.

[c] Adverse event assessments on Day 1 should be performed following study drug administration.

[d] For patients enrolled in Part 2 of the study, visit procedures required at Week 52 and also Extension Day 1 should only be performed once since Part 2 Week 52 and Extension Day 1 should occur on the same day.

[e] For patients with echocardiograms and electrocardiograms collected within 90 days prior to the Screening Visit, the most recent echocardiogram and electrocardiogram can be used to assess patient eligibility.

[f] Patients must be allowed to rest for a minimum period of 1 hour following maximal exercise test before this blood sample is collected (at the same time as all central lab blood draws). This sample should be taken with the patient in the same position (eg, sitting or semi-recumbent) at all appropriate visits.

[g] Blood samples should be collected for hepatitis B and C and HIV antibodies only in patients lacking evidence of a negative titer in the past year.

[h] A serum pregnancy test will be performed at any point in time if a pregnancy is suspected. All other pregnancy assessments will be urine pregnancy tests. Additional pregnancy assessments will be performed more frequently if required by local law or requested by local health authorities or IRBs/ECs.

[i] Patients and investigators must complete the patient global impression of change and clinical global impression of change following completion of the neurological FARS exam

[j] Height will be performed at Week 4 for Part 2 patients enrolled in the Extension.

[k] At week 24+ visits and end of study visit, patients should be asked whether they had a recent electrocardiogram and/or echocardiogram done. If the answer is yes, the site is to request that data and, if provided, enter that data in the appropriate eCRF.

5. PLANNED ANALYSES

5.1. Planned Interim Analyses

Interim analyses are based on locked data and performed as appropriate to support commercial approval of omaveloxolone. A DBL plan will describe the details of each DBL. All outputs identified in this SAP and planned for the final analysis will be prepared for each interim analysis, and an interim CSR will be prepared.

At the time of each planned interim analysis, no individual patient's data are considered fully locked; therefore, interim results are interpreted as preliminary.

5.2. Final Analyses and Reporting

All final, planned analyses identified in this SAP will be performed after the last patient has completed the study and the database has been locked. A final CSR will be prepared based on the final analyses.

6. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING

The half-life of omaveloxolone 150 mg was determined to be approximately 26 ± 13 h in Part 1 of the MOXIe study (Investigator Brochure, Section 5.1.3); patients who remain off-treatment for > 8 weeks should have no lingering omaveloxolone drug effects at > 8 weeks after the date of last dose. For this reason, patients participating in Part 1 of the study are treated as placebo-omaveloxolone patients in the extension study analysis.

Results will be summarized by the treatment analysis groups defined in Section 4.5.

Patient listings of all analysis data that support summary tables and figures are provided. Measurements from patients excluded from the pre-defined analysis sets or extra measurements (such as unscheduled or repeat assessments) are not included in summary tables unless specified otherwise, but they are included in the patient listings. Missing data are not imputed, unless otherwise specified. In general, patient listings are sorted by patient number and assessment date (time and parameter, as applicable).

6.1. General Summary Table and Individual Subject Data Listing Considerations

Results of statistical analyses are reported using summary tables, listings, and figures (TLFs). All TLFs will use ICH numbering conventions.

Data from Part 2 Week 52 was entered only into the Part 2 database, but also serves as the Extension Day 1 data. For patients in Part 2, the Day 1(Week 52) efficacy and safety results will be pulled from the Part 2 analysis datasets

All analyses and summaries are produced using Statistical Analysis Software (SAS)® version 9.3 (or higher).

Unless otherwise specified, descriptive statistics for continuous variables include the number of patients with data (N), mean, SD, median, minimum, and maximum. The same number of decimal places as in the observed value are presented when reporting minimum and maximum; 1 more decimal place than in the observed value is generally presented when reporting mean; and 2 more decimal places than in the observed value are presented when reporting SD. When calculating the minimum and maximum of average baseline values, the same number of decimal places as in the observed value will be presented when reporting minimum and maximum. For

imputed values and the average of baseline values, the average values will be rounded to the same number of decimal places as the source value.

Categorical (qualitative) data are presented using frequency counts and percentages. All percentages are rounded to 1 decimal place, unless otherwise specified. Percentages equal to 100 are presented as 100% and no percentages are presented for zero frequencies. Where individual variable values are missing, summaries of categorical data are based on reduced denominators (ie, only patients with available data are included in the denominators). For summaries of adverse events (AEs) and concomitant medications, the percentages are based on the number of patients who received study drug within each summary group.

6.2. Analysis Populations

Analysis populations defined in this section pertain to patients enrolled in the extension portion of the trial.

6.2.1. Safety Population

The safety population includes all patients who received at least 1 dose of open-label omaveloxolone. The safety population is used for evaluation of both efficacy and safety variables.

6.2.2. Subgroups

Efficacy endpoints will be summarized separately using the following subgroups:

- Age: < 18; ≥ 18
- Sex: female; male
- Geographic location: United States; Other
- Pes Cavus: (yes/no)
- Ethnicity: Non-Hispanic/Latino; Hispanic/Latino
- Race: (White, Non-White)
- Trinucleotide composed of one guanine and two adenines (GAA)1 repeat length $\geq$ 675: Yes, No
- COVID-19 subgroup: Yes, No (patients with vs without treatment emergent adverse event preferred term of coronavirus and start date $\geq$ 01 Dec 2019)

- COVID-19 investigational product (IP) interruption: (Yes; No)

Safety endpoints will be summarized using the following subgroups:

- Pes cavus: (yes/no)
- Age: < 18; ≥ 18
- COVID-19 IP interruption: (Yes; No)

6.3. Baseline Definition

For patients initially randomized to placebo in Part 2 or patients participating in Part 1 of the study, baseline values are defined as the last non-missing assessment prior to the first study drug administration in the extension study, unless otherwise specified below. If the first study drug administration in the extension study occurs after the date of drug dispense in the extension, the last measurement prior to the first study drug administration is considered the Day 1 measurement for the calculation of baseline. For patients in Part 2, the Extension Day 1(Week 52) efficacy and safety results are pulled from the Part 2 analysis datasets if patients are not rescreened before extension study.

6.3.1. Efficacy Assessments

6.3.1.1. Modified Friedreich's Ataxia Rating Scale without Prior Exercise Testing

For patients participating in Part 1 of the study, the Day 1 assessment is used as the baseline modified Friedreich's ataxia rating scale (mFARS).

For patients in Part 2, baseline mFARS values are the Week 52 / Extension Day 1 mFARS values as defined in the 408-C-1402 Part 2 protocol (Section 7.8).

6.3.1.2. Other Efficacy Assessments

For patients participating in Part 1 of the study, the Day 1 assessment is used as baseline for continuous efficacy parameters: timed 25-foot walk (T25FW), 9-hole peg test (9HPT), and activities of daily living (ADL).

For patients in Part 2, baseline values are the Week 52 / Extension Day 1 values as defined in the 408-C-1402 Part 2 protocol (Section 7.8).

6.3.2. Safety Assessments

Baseline for visit-based, continuous safety assessments (ie, vital sign assessments, weight, body mass index [BMI], and laboratory measurements) is defined as the average value of measurements collected up through, but prior to, first study drug administration in the extension study unless otherwise specified. For patients participating in Part 2, Week 48 values or any on-treatment values in Part 2 will not be incorporated in the baseline safety assessments unless otherwise specified.

6.4. Derived and Transformed Data

6.4.1. Demographic and Baseline Characteristics

Unless otherwise specified in Section 6.3 or Section 6.4.2, demographic and baseline characteristics for patients participating in Part 1 are the values used in the Part 1 analysis and defined in the 408-C-1402 Part 1 SAP (Section 4.2). Within the safety Part 1 safety population, these include the following:

- Sex
- Race
- Ethnicity
- Age at FA onset
- GAA1 repeat length
- GAA2 repeat length
- Ambulatory status, ambulatory assistive devices
- Pes Cavus (yes/no) pulled from extension Day 1
- History of cardiomyopathy, areflexia, foot surgery, sensory neuropathy, swallowing difficulties, scoliosis, scoliosis surgery

Unless otherwise specified in Section 6.3 or Section 6.4.2, demographics and baseline characteristics for patients participating in Part 2 are the values used in the Part 2 analysis and defined in the 408-C-1402 Part 2 SAP (Section 7.4). Within the safety Part 2 safety population, these include the items listed above.

6.4.2. Baseline Age

For patients participating in Part 1 of the study, subject's age in years is defined as the age at extension study consent (Screening for Part 1 patients, Day 1 Extension /Week 52 for Part 2 patients). For Part 2 patients, if no date of extension study consent is present, the age at the Day 1 Extension/Week 52 visit is used.

- Age (year) = Floor ([date of extension study consent – date of birth]/365.25)

6.4.3. Years Since FA Onset

The years since FA onset are calculated as the difference between Baseline Age (Section 6.4.2) and Age at Onset (Section 6.4.1).

6.4.4. Pes Cavus (yes/no)

The assessment of pes cavus was inconsistent between Part 1 and Part 2 of the 408-C-1402 study. In Part 1, the assessment of pes cavus was solely based on the investigator's clinical judgement. In Part 2, the clinical diagnosis of pes cavus is defined as having a loss of lateral support. Loss of lateral support is determined by having a patient stand barefoot and bear weight. If light from a flashlight can be seen under the patient's arch, then the patient has pes cavus per the study's definition. Since Part 1 utilizes different test methodology for pes cavus, patients tested as pes cavus 'yes' in Part 1 and tested as pes cavus 'no' in extension Day 1 are listed.

For patients participating in Part 2 of the study, baseline pes cavus values are the baseline values used in the Part 2 analysis and defined in the 408-C-1402 Part 2 SAP.

For patients participating in Part 1 of the study, baseline pes cavus values are the Extension Day 1 values.

6.4.5. Study Day

Study day is the day relative to the date of drug dispensed in the extension study. Day 1 is defined as the date of study drug dispensation, unless otherwise specified for the calculation of baseline (Section 6.3).

Assessments that occur after extension study drug dispensing but before the first study drug administration are considered to occur on study Day 1. Assessments collected on the same date

as the first date of study drug administration will be considered to occur before the first dose of study drug administration.

For visits (or events) after the date of drug dispensing, day is calculated as:

- Study day = visit (or event) date - date of drug dispense + 1

For visits (or events) before the date of drug dispensing, day is calculated as:

- Study day = visit (or event) date - date of drug dispense

For listings (such as for adverse events) the quantity 'days since first (or last dose)' is defined as:

- days since first (or last dose) = event date – date of first (or last dose) + 1

For summaries that present distribution of time expressed in weeks and months, weeks will be defined as days divided by seven and months as days divided by 30.4.

6.4.6. Change from Baseline

Change from baseline is calculated using the baseline value (Section 6.3) and the value closest to the target study day, using the rules defined in Section 6.4.5.

6.4.7. Visit Windows

Because clinical visits may occur outside protocol-specified windows, instead of relying on visit labels in the clinical database, analysis visits and their windows are defined using derived study day (Section 6.4.3). Study day is calculated using the actual date for each scheduled and unscheduled assessment and compared to the target study day for each analysis visit. Data analysis and summaries are based on the collection date that is closest to the protocol scheduled target study day. Efficacy and safety summaries are based on Table 3.

Table 3: Analysis Visits

Analysis Visit	Label	Target Study Day	Analysis Window
4	Week 4	28	$21 \leq$ Study Day ≤ 55
24	Week 24	168	$112 \leq$ Study Day ≤ 224
48	Week 48	336	$280 \leq$ Study Day ≤ 392

72	Week 72	504	448≤ Study Day ≤ 560
96	Week 96	672	616 ≤ Study Day ≤ 728
120, 144, etc.	Week 120, Week 144, Every 24 Weeks Thereafter	Analysis Visit × 7	(Analysis Visit × 7 - 56) ≤ Study Day ≤ (Analysis Visit × 7 + 56)
FU	Follow-up	28 days after last dose	14 ≤ Days after last dose

Abbreviations: FU=follow-up

The efficacy and safety follow-up is based on days since last dose. If more than one assessment exists during the follow-up after last dose, the one closest to 28 days following the date of the last study drug administration is used for analysis and summary.

If a parameter is assessed or measured more than once within a visit window, the one that is closest to the protocol-scheduled time point (or target study day) is used for the purposes of data analysis and summary. If two assessments are equidistant from a target study day, the earlier assessment is used.

Records from visits not closest to the target study day, and therefore not used in analyses, are presented in by-subject data listings.

6.4.8. Clinical Laboratory Results Outside the Limit of Quantification

Laboratory results reported as less than the lower limit of detection (ie, <LLD) are imputed as LLD/2. Laboratory results reported as greater than the upper limit of detection (ie, >ULD) are imputed as the ULD.

6.4.9. Estimated Glomerular Filtration Rate

The estimated glomerular filtration rate (eGFR) is calculated using the formula specified below.

For patients participating in Part 1 of the study who consented to the extension study at age 18 years and older, the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation is used:

- eGFR (mL/min/1.73 m^2) = 141 × min(S_{cr}/κ, 1)$^{\alpha}$ × max(S_{cr}/κ, 1)$^{-1.209}$ × 0.993^{Age} × 1.018 [if female] × 1.159 [if black]

For patients participating in Part 2 of the study who consented to Part 2 of the study at age 18 years and older, the CKD-EPI equation is used.

For patients participating in Part 1 of the study who consented to the extension study at < 18 years old, the Bedside Schwartz equation is used:

- eGFR (mL/min/1.73 m^2) = (0.41 × Height in cm) / S_{cr}

Where S_{cr} is serum creatinine (mg/dL), κ is 0.7 for females or 0.9 for males, and α is -0.329 for females or -0.411 for males. Min indicates the minimum of S_{cr}/κ and 1 and max indicates the maximum of S_{cr}/κ and 1. Age and height at screening are used for eGFR calculations.

For patients participating in Part 2 of the study who consented to Part 2 of the study at < 18 years old, the Bedside Schwartz equation is used.

6.4.10. mFARS without Prior Exercise Testing

The Friedreich Ataxia Rating Scale (FARS) is a neurological-exam-based rating scale with five sections: Bulbar (section A), Upper Limb Coordination (section B), Lower Limb Coordination (section C), Peripheral Nervous System (section D), and Upright Stability (section E). Assessments in each section are added according to the logic in Table 4 as specified in the FARS manual (Lynch, 2006). The mFARS is the sum of sections A, B, C, and E.

Table 4: Scoring of the Neurologic FARS

Component Score of Neurologic FARS	Logic	Handling of Missing Data
A	Sum of all of the section A responses	If question is left blank but FARS assessment was completed, use question response from last visit.
B	Sum of all of the section B responses	
C	Sum of all of the section C responses	
D	Sum of all of the section D responses, except 5a and 5b. For Question 5, only the results of Questions 5c1 and 5c2 is included in the section D sum.	
E	Average of items 2-5 item in the database; reports sums for questions 1, 6, and 7 and sums of averages for questions 2-5.	If questions 1-5 are left blank but FARS assessment was completed, take average of existing responses (eg, average of 3a.1 and 3a.2). If no existing responses exist at the current visit for questions 1-5, use the question response from the last visit. For questions 6-7, use the response from the last visit.

Abbreviations: FARS=Friedreich's ataxia rating scale

6.4.11. 9-Hole Peg Test

Both the dominant and non-dominant hands are tested twice at each assessment. The times recorded for the two trials for each hand are averaged, and the average value for each hand is used in analyses.

The average time value for each hand is transformed for analysis to allow for a more normal distribution of the data using the following methods:

- The number of pegs per second (pegs/second) for each hand is calculated based on nine pegs placed compared to the time needed to complete the test: 9 / (average time)
- The reciprocal of the average time is calculated as 1 / (average time) (Lynch, 2005).

6.4.12. Timed 25-Foot Walk Test

Patients are instructed to attempt two T25FW trials at each visit. If both walk trials are completed, walk times are averaged for use in analyses. Otherwise, only the completed walk from each visit is used for analyses. The walk time is transformed for analysis to allow for a more normal distribution of the data using the following methods:

- The number of feet per second (feet/second) for each walk is calculated based on the distance (25 feet) and the total time needed to complete the test: 25 / (average time)
- The reciprocal of the walk time is calculated as 1 / (average time) (Lynch, 2005)

6.4.13. Electrocardiogram Intervals

Electrocardiogram intervals are calculated from QT and RR intervals using the following formulas:

- Corrected QT interval using Fridericia's formula (QTcF): $QTcF = QT/\sqrt[3]{RR}$
- Corrected QT interval using Bazett's formula (QTcB): $QTcB = QT/\sqrt[2]{RR}$ where RR = 60 / (Heart Rate).

6.5. Handling of Missing Data

6.5.1. Missing Efficacy Data

Missing efficacy data are not imputed unless specified otherwise.

6.5.2. Missing Start and Stop Dates for Concomitant Medication

Missing start dates for concomitant medications are not imputed.

Concomitant medications with incomplete end dates are considered concomitant medications if:

- Day and month are missing, and the year is equal to or after the year of the first date of study drug administration;
- Day is missing and the year is after the year of the first date of study drug administration;
- Day is missing and the year is equal to the year of the first date of study drug administration and the month is equal to or after the month of the first date of study drug administration; or
- Year is missing.

6.5.3. Missing Start and Stop Dates for Adverse Events

Treatment-emergent adverse events (TEAEs) are events that either:

- Had a date of onset on or after the date of first dose of study drug and not more than 30 days after the date of the last dose of study drug, or
- Had no recorded date of onset with a stop date after the first dose of study drug, or
- Had no recorded date of onset or stop date.

Adverse events with incomplete start dates are considered after the date of first dose if:

- Day and month are missing, and the year is equal to or after the year of the first date of study drug administration;
- Day is missing and the year is after the year of the first date of study drug administration;

- Day is missing and the year is equal to the year of the first date of study drug administration and the month is equal to or after the month of the first date of study drug administration; or
- Year is missing.

Adverse events with incomplete start dates are considered on or within 30 days of last dose, if:

- Day and month are missing, and the year is equal to or before the year of the date of last dose of study drug plus 30 days;
- Day is missing and the year is equal to or before the year of the date of last dose of study drug plus 30 days, and month is equal to or before the month of the date of last dose of study drug plus 30 days;
- Year is missing.

7. STUDY POPULATIONS

Data are summarized for the safety population (Section 6.2.1).

7.1. Subject Disposition

A disposition summary includes the number and percentage of patients in the following categories, are summarized by treatment analysis group (Section 4.5):

- Safety Patients in 408-C-1402 Part 1
  - Safety Patients in 408-C-1402 Part 1 and randomized to placebo
  - Safety Patients in 408-C-1402 Part 1 and randomized to omaveloxolone
- Safety Patients in 408-C-1402 Part 2
  - Safety Patients in 408-C-1402 Part 2 and randomized to placebo
  - Safety Patients in 408-C-1402 Part 2 and randomized to omaveloxolone
- Safety population
  - From 408-C-1402 Part 1
  - From 408-C-1402 Part 2
- Continuing (or completed) treatment • Terminated early from the treatment
  - Reason for terminating treatment
  - Terminated early from the treatment due to COVID-19
- Terminated early from the study
  - Reason for terminating study
  - Terminated early from the study related to COVID-19

Results will be summarized by analysis population (Section 6.2). A listing of disposition is provided for all enrolled patients.

7.2. Screen Failures

Screen failures are not summarized.

7.3. Protocol Deviation

Protocol deviations for patients who entered the study even though they did not satisfy all entry criteria, patients who received an incorrect dose, major protocol deviations and protocol deviations related to COVID-19 are listed.

7.4. Demographic and Baseline Characteristics

Summaries of demographic and other baseline characteristics summaries are presented by treatment analysis group for all analysis populations. Derivation of demographic and other baseline characteristics are described in Section 6.4.1 and Section 6.4.2. For variables not described in Section 6.4.1 or Section 6.4.2, baseline values are used in demographic and baseline characteristic summaries according to Section 6.3.

Demographic and other baseline characteristics include the following:

- Baseline Age
- Age category ($< 18, \geq 18$)
- Sex: female; male
- Race: (White, Non-White)
- Ethnicity: Non-Hispanic/Latino; Hispanic/Latino
- Weight (kg), BMI (kg/m^2)
- Diastolic and systolic blood pressure (mmHg), Heart rate (bpm)
- mFARS, each FARS subsection (ie, A, B, C, D, E), ADL
- Pes Cavus: (yes/no)
- Age at FA onset
- Years since FA onset
- GAA1 repeat length, GAA2 repeat length, GAA1 repeat length ≥ 675
- Ambulatory status, Ambulatory assistive devices
- History of cardiomyopathy, areflexia, foot surgery, sensory neuropathy, swallowing difficulties, scoliosis, scoliosis surgery.

7.5. Listing of Subject Inclusion and Exclusion Criteria

A listing of enrolled patients who did not meet inclusion or exclusion criteria is generated.

7.6. Medical History

Medical history is summarized by treatment analysis group (Section 4.5, as well as all patients combined). For patients participating in Part 1 of the study, medical history values are the values recorded at the extension study Screening visit. For patients participating in Part 2 of the study, medical history values are the values recorded at the Part 2 Screening visit. Medical history is coded using MedDRA (Medical Dictionary for Regulatory Activities) version 28.1. Medical history items are summarized by MedDRA System Organ Class (SOC) and preferred term (PT). Patient listings are also provided.

8. EFFICACY

Analyses of efficacy (mFARS, ADL, 9-HPT, T25FW) include only summary statistics by visit (Section 6.4.7). Summary statistics for observed values and change from baseline are presented. Analyses of Patient Global Impression of Change (PGIC) and Clinical Global Impression of Change (CGIC) results will include frequency counts and summary statistics by visit. Results will be summarized by treatment analysis group (Section 4.5, as well as all patients combined).

Analyses of efficacy described in this section are the primary analyses of the efficacy endpoints and pertain only to the Full Analysis Set (ie, patients without pes cavus).

8.1. Modified FARS without Prior Exercise Testing

The mFARS is not performed after maximal exercise testing, as it was in 408-C-1402 Part 1 and Part 2. Because of the operational challenges in completing the maximal exercise test and the extension study objective of examining the long-term safety of omaveloxolone treatment at 150 mg, maximal exercise testing is not performed in the extension study. Therefore, the mFARS values may differ from those recorded in Part 1 and Part 2 of the study and should not be directly compared, since in the extension study patients were not sufficiently fatigued by the maximal exercise test before completing the FARS assessments.

Summary statistics for observed values and change from baseline are presented. In addition, summaries of all subsections of the FARS (A, B, C, D, and E) are generated.

8.2. 9-Hole Peg Test

The 9-HPT is a brief, standardized, quantitative test of upper extremity function. Two trials are attempted using each hand, and the time for the dominant and non-dominant hand are analyzed separately. The average observed time for each hand is summarized descriptively at each timepoint by treatment analysis group (placebo-omaveloxolone, omaveloxolone all). Change from baseline treatment is summarized for each hand. Longer test times reflect more impairment on the patient's upper extremity function, thus a decrease from baseline suggests an improvement. Reciprocal time values and pegs/second values (Section 6.4.11) are also summarized at each time point along with change from baseline in these parameters.

8.3. Timed 25-Foot Walk Test

The T25FW is a quantitative mobility and leg function performance test based on time in seconds to complete a 25-foot walk. The average observed time is summarized descriptively at all visits by treatment analysis group (placebo-omaveloxolone, omaveloxolone all). Change from baseline by treatment is summarized at all visits. Longer test times reflect more impairment on the patient's ability to walk, thus a decrease from baseline suggests an improvement. Reciprocal time values and feet/second values (Section 6.4.12) are also summarized at each time point along with change from baseline in these parameters.

8.4. Patient Global Impression of Change

The PGIC is a 7-point scale that requires the patient to assess how much the patient's illness has improved or worsened relative to a baseline state at the beginning of an intervention (Guy, 1976). The PGIC is assessed by completing the following statement "since I began trial treatment, my overall status is: very much improved (1), much improved (2), minimally improved (3), no change (4), minimally worse (5), much worse (6), very much worse (7)." All PGIC values collected at each visit using the appropriate analysis windows (Section 6.4.7) will be used; the number of patients in each category is summarized at each analysis visit by treatment analysis group (Section 4.5).

8.5. Clinical Global Impression of Change

The CGIC is a 7-point scale that requires the clinician to assess how much the patient's illness has improved or worsened relative to a baseline state at the beginning of an intervention (Guy,

1976). The CGIC is assessed by completing the following statement "Compared to the patient's condition at the start of the trial, this patient's overall status is: very much improved (1), much improved (2), minimally improved (3), no change (4), minimally worse (5), much worse (6), very much worse (7)." The number of patients in each category are summed at each analysis visit by treatment analysis group (Section 4.5). There are several patients who transferred to different study sites during participation in the extension study. CGIC results collected after patient transfers will be excluded from the summary and listed separately.

8.6. Activities of Daily Living

The ADL scale examines the ability of patients to complete everyday tasks (eg, holding a fork, dressing). The ADL is a 9-question assessment, with the total score being the sum of 9 questions. If a question result is unanswered, then the total ADL score is considered missing. The number of patients with an unanswered ADL question will be presented. In addition, summaries of each ADL question are generated by treatment analysis group (Section 4.5).

9. SAFETY AND TOLERABILITY

Safety data (including AEs, laboratory data, vital signs) are listed and summarized for patients in the safety population (Section 6.2.1). Results will be summarized by 1) patients initially randomized to placebo in Part 2 or patients participating in Part 1 of the study (placebo-omaveloxolone), 2) patients initially randomized to omaveloxolone in Part 2 (omaveloxolone), 3) all patients (all).

The analysis visit windows (Section 6.4.7) are used for all safety analyses. Visits that are outside the visit windows are listed where appropriate.

Echocardiogram (ECHO) and electrocardiogram (ECG) (Section 9.6 and Section 9.7) data will be summarized descriptively for visits where a sufficient number of patients have data collected.

9.1. Adverse Event Preferred Term and Body/Organ System Summary Tables

The AEs are summarized by treatment as defined by the safety analysis set. General considerations for AE summaries and calculations are:

- Multiple events by PT and SOC are counted once only per patient for each treatment.

- For summaries by severity, only the most severe event is counted per patient for each treatment.
- For summaries by relationship, for frequency counts by patient, only the most related event is counted per patient for each treatment.
- An AE with a missing resolution date or incomplete date that is not identified as continuing is assumed to be continuing and no duration is calculated.
- Only TEAEs are included in summaries.

All AEs are coded using MedDRA version 28.1. In MedDRA, each verbatim term is mapped to a PT and high-level term, which is then mapped to an SOC. Tables and listings present data at the SOC and PT level.

Treatment-emergent AEs are events that either:

- Had a date of onset on or after the date of the date of first dose in the extension study and not more than 30 days after the date of the last dose of study drug in the extension study, or
- Had no recorded date of onset with a stop date after the first dose of study drug, or
- Had no recorded date of onset or stop date.

Adverse events with a date of onset in Part 2 that continued during the extension study are not considered TEAEs.

In addition, AEs that occurred > 30 days after the date of last dose of study drug are summarized as late-onset AEs. These are AEs that had a date of onset > 30 days after the date of the last dose of study drug.

The investigator grades the severity of the AEs as mild, moderate, or severe as defined in the study protocol.

Association or relatedness to the study medication are graded by the investigator according to criteria specified in protocol Section 12.4.

As defined in the protocol and captured on the eCRF, an SAE is an AE that results in any of the following:

- Death;
- A life-threatening adverse drug experience;

- Inpatient hospitalization or prolongation of existing hospitalization;
- Results in persistent or significant incapacity or substantial disruption of the ability to conduct normal life functions;
- Results in a congenital anomaly or birth defect in an offspring of a patient taking study drug;
- Is an important medical event.

9.1.1. Missing and Partial AE Onset Dates

Rules for handling missing partial AE Onset Dates are included in Section 6.5.3.

9.1.2. Summaries of AEs for SAEs, AE Dropouts, and Death

Treatment-emergent AEs are summarized by treatment at onset of the AEs. For each treatment, SOC, and PT, the number and percentage of patients reporting an event are calculated. In summary tables, SOCs are presented alphabetically and events within SOCs are presented by decreasing frequency count.

Summary tables (number and percentage of patients) of AEs (by SOC and PT) are provided by treatment as follows:

- All TEAEs
- All treatment-emergent related AEs (definitely, probably, or possibly related)
- All TEAEs by severity
- All treatment-emergent SAEs (including deaths)
- All related treatment-emergent SAEs (including deaths)
- All TEAEs leading to permanent discontinuation of study drug.
- All treatment-emergent non-serious AEs

Listings are provided showing:

- All AEs
- SAEs (including deaths)
- AEs leading to discontinuation of study drug.

9.2. Exposure and Compliance

The duration of study drug exposure is defined as the number of days on treatment from the first dose of study drug until the last dose of study drug (last dose – first dose + 1). Study drug exposure is summarized by descriptive statistics. Summaries include the total dose (mg) received (based on the number of pills returned), study drug compliance, the number and percentage of patients receiving study drug by dispense visit (Section 6.4.7), and duration (days) of exposure during the study treatment period.

In addition, a summary of the number and percentage of patients on omaveloxolone by visit is generated.

The following exposure categories are summarized by number and percentage of patients:

- ≤ 24 Weeks
- > 24 Weeks
- ≤ 48 Weeks
- > 48 Weeks
- ≤ 72 Weeks
- > 72 Weeks
- ≤ 96 Weeks
- > 96 Weeks

Total number of doses dispensed and total dose (mg) dispensed are calculated from total number of kits (bottles) recorded on the Study Drug Dispensation eCRF. Total number of doses received are calculated from information on the eCRF of Study Drug Return and Study Drug Dispensation, as the total number of doses dispensed – total number of doses returned. Study drug compliance (%) is calculated as 100 × total number of doses received / total number of doses dispensed. For patients who remain on study treatment at the time of an interim or final data cut, the last kit dispensed date is considered the date of last dose.

9.3. Concomitant and Other Medications

Concomitant medications are coded using the World Health Organization (WHO) drug dictionary Global B3 September 2025 for anatomical therapeutic chemical (drug classification system) (ATC) and preferred drug name. A patient who used multiple medications are counted only once

for each ATC and preferred drug name. The ATC and preferred drug name within each ATC are sorted alphabetically. Coded concomitant medications are summarized by treatment analysis group (Section 4.5). Percentages are based on the number of patients in the safety analysis set.

A concomitant medication is any medication taken at the time of first study treatment during the extension study or a medication that was started after the start of study drug dosing. Specifically, concomitant medications are medications

- That are continued from screening or Part 2 and continued after the first study drug dosing, or
- That have start dates or stop dates within the treatment period.

Concomitant medications are summarized for each treatment group by WHO ATC class and preferred name. Patients may have more than one medication per ATC class and preferred name. At each level of patient summarization, a patient is counted once if he/she reported one or more medications at that level. Each summary is ordered by descending order of incidence of ATC class and preferred name within each ATC class.

Concomitant medications include those with an end date after the first study drug administration as well as medications without an end date. Medications with an end date on the date of first study drug administration are not considered concomitant medications.

In addition, patients who take excluded or prohibited medications (defined in Protocol Section 9.2.1 or Protocol Section 9.2.2, respectively) during the study are listed.

9.3.1. Missing and Partial Concomitant and Other Medication Start and Stop Dates

Missing and partial concomitant medication start and stop dates are detailed in Section 6.5.2.

9.4. Clinical Laboratory Evaluations

Laboratory data are summarized at baseline and at each time point by treatment analysis group (Section 4.5). Only values obtained within analysis study windows (Section 6.4.7) are included in by-visit summary statistics. For patients diagnosed with COVID-19 after first dose in 408-C-1402 extension study, laboratory assessment values after the onset date of severe/moderate AE or SAE PT of coronavirus are not used in summary tables but will be presented in by-patient listings. Post severe/moderate or serious COVID-19 laboratory assessment values will be summarized separately.

9.4.1. Continuous Summaries of Laboratory Results

Selected laboratory evaluations and change from baseline are summarized by treatment analysis group (Section 4.5), laboratory category (hematology, chemistry), test, and study visit using continuous statistics. Laboratory tests to be summarized are provided in Section 13.3. The eGFR results are calculated using formulas described in Section 6.4.9.

Line graphs of change from baseline are generated for laboratory tests (hematology and chemistry). Line graphs include mean ± SE over time for change from baseline.

Due to the nature of urinalysis parameters, summaries of continuous statistics are not provided. Urinalysis results (other than potential of hydrogen and specific gravity) of ketones, protein, blood, glucose, clarity, color, leukocytes, nitrite, bilirubin, and microscopic examinations (if indicated based on laboratory results), urine microscopic findings, and pregnancy test results are not summarized.

In addition, the number and percentage of patients with any abnormal values while on study drug are listed.

9.4.2. Categorical Summary of Laboratory Results

The number and percentage of patients by laboratory normality and abnormality categories at any time during the study (Normal, Low, High) are summarized by treatment analysis group (Section 4.5), laboratory category (hematology, chemistry, and urinalysis), and laboratory test. The worst abnormality values are defined as the maximum values while on study treatment, except for magnesium and hemoglobin. The worst abnormal values for magnesium and hemoglobin are defined as the minimum values while on study treatment.

Low-density lipoprotein (LDL) and high-density lipoprotein (HDL) are summarized using normality and abnormality categories (Normal, Low, High) shift tables from baseline to end of treatment, baseline to worst on-treatment and baseline to follow-up.

An initial set of parameters of specific interest (alanine aminotransferase [ALT], aspartate aminotransferase [AST]), are summarized using shift tables, though additional parameters may be added as appropriate. Shift tables summarizing (1) baseline to end of study treatment, and (2) baseline to worst on-treatment, (3) baseline to follow-up, and (4) worst on-treatment to follow-up are presented for lab parameters as appropriate.

In addition, a summary table is provided by treatment analysis group (Section 4.5) for the number and percentage of patients meeting a following prespecified threshold level at any time during the study:

- Magnesium < 1.3 mEq/L (0.65 mmol/L)
- BNP > 200 pg/mL
- NT-proBNP > 1000 pg/mL

9.4.3. Transaminases and Total Bilirubin

To assess the potential for drug-induced liver injury, the number of patients meeting the follow criteria are summarized:

- ALT more than 3×, 5×, 10×, or 20× upper limit of normal (ULN)
- AST more than 3×, 5×, 10×, or 20× ULN
- Either ALT or AST more than 3×, 5×, 10×, or 20× ULN
- Either ALT or AST more than 5× ULN for more than two consecutive weeks
- Alkaline phosphatase (ALP) > 1.5× ULN
- Total bilirubin (TBL) >2× ULN
- AST or ALT > 3× ULN with an associated TBL > 1.5× ULN
- AST or ALT > 3×ULN with an associated TBL > 2× ULN
- AST or ALT > 3× ULN with associated eosinophils (> 5%)

A summary table that includes frequencies and percentages of patients that meet any of the above criteria at any time during the study are provided. A listing of subjects with abnormal ALT, AST, or TBL is also provided.

In addition, evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plots are generated for ALT and AST versus TBL. Graphs of ALT and AST relative to ULN and the appearance of selected AEs (eg fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash) will be generated for subjects with ALT or AST >3xULN.

9.5. Vital Signs

Vital signs assessments include systolic blood pressure (SBP, mmHg) and diastolic blood pressure (DBP, mmHg), body temperature (°C), heart rate (HR, bpm), weight (kg), and BMI (kg/m^2). Only values obtained within analysis study windows (Section 6.4.7) are included in by-visit summary statistics. Vital signs are summarized at baseline and at each time point along with the change from baseline by treatment analysis group (Section 4.5, and all patients combined). All data are listed.

For patients diagnosed with COVID-19, vital signs assessment values after the start date of a moderate/severe AE or SAE PT of coronavirus are not used in summary tables but will be presented in by-patient listings. Post moderate/severe or serious COVID-19 vital signs assessment values will be summarized separately.

Body mass index is summarized through a shift table from baseline to lowest on-treatment based on the following categories: < 18.5, 18.5 to < 25, 25 to < 30, ≥30. Patients with lowest on-treatment BMI values < 18.5 will be separately listed.

Percentage of change from baseline of body weight will be summarized and listed.

In addition, a summary table is provided for the number and percentage of patients meeting the following prespecified post baseline weight change categories:

- Body Weight (kg):
  - Maximum increase from baseline: 5 to 7%; > 7 to 10%; > 10 to 20%; > 20%
  - Maximum decrease from baseline: 5 to 7%; > 7 to 10%; > 10 to 20%; > 20%

9.6. Echocardiogram

At week 24+ visits and end of study visits, patients are asked whether they had a recent ECHO done. If the answer is yes, the site is to request that data and, if provided, enter that data in the appropriate eCRF.

Cardiac function (ejection fraction [%], left and right wall thickness [cm], septum thickness [cm], and left ventricular mass index [g/m^2]) measurements captured via the ECHO are summarized at baseline and at each available time point along with the change from baseline.

Baseline for ECHO is the last assessment prior to first study drug administration. For patients participating in Part 2 of the study, Week 48 ECHO is the last assessment collected in Part 2 and used in the extension screening, which will be incorporated in the baseline.

For patients diagnosed with COVID-19, ECHO assessment values after the start date of a moderate/severe AE or SAE PT of coronavirus will not be used in summary tables but will be presented in by-patient listings.

9.7. Electrocardiogram

At week 24+ visits and end of study visits, patients are asked whether they had a recent ECG done. If the answer is yes, the site is to request that data and, if provided, enter that data in the appropriate eCRF.

All ECG data, such as clinical interpretation of ECGs, ventricular rate, and interval assessments of PR, QRS, and QT are collected on the eCRF.

The QTcF and QTcB are calculated (Section 6.4.13). Descriptive statistics for observed values and change from baseline at each available time point are presented for these 12-lead ECG interval assessments.

Baseline for ECG is the last assessment prior to first study drug administration.

In addition, the number and percentage of patients with any abnormal values at any time during the study (ie, above a prespecified threshold) are summarized by the available time points and overall while on study drug. The prespecified levels of ECG QTcF thresholds (Table 5) are consistent with FDA guidance.

For patients diagnosed with COVID-19, ECG assessment values after the start date of a moderate/severe AE or SAE PT of coronavirus will not be used in summary tables but will be presented in by-patient listings.

Table 5: Prespecified Threshold Levels for ECG Parameters

ECG Parameter	Prespecified Level
PR	> 200 msec
QTcF	> 450, > 480, or > 500 msec, > 30 or > 60 msec increase from baseline
Heart rate	< 40, > 100 bpm

Abbreviations: ECG=electrocardiogram; QTcF=corrected QT interval using Fridericia's formula

9.8. Pregnancy

A listing is provided for serum and urine pregnancy results of all on-study pregnancies.

10. PHARMACOKINETICS

There are no pharmacokinetic samples collected in the MOXIe study.

11. COVID-19 IMPACT

11.1. Operational Impact

The COVID-19 pandemic has impacted the conduct of the study per protocol.

The mitigation strategies are adopted to protect the health of participants in the extension study, while maintaining compliance with good clinical practice and minimizing the risk to trial integrity (Protocol Appendix 13). For any visits after 01-Mar-2020 sites should record if visits, data collection, drug dispensation, and drug administration are impacted by or related to COVID-19.

The number and percentage of patients impacted by COVID-19 are summarized by visit and treatment analysis group for the following categories:

- Visit impact
  - Visit not done
  - Visit conducted out of window
  - One or more procedures could not be performed
- Data collection methods
  - No change (occurred as expected per protocol)
  - Telephone visit
  - Video conference visit
  - Home health care visit
  - Local lab
  - Other
- Drug dispensation methods
  - No change (office visit)
  - Mailed to patient
  - Home health care provided

  - Other

The number and percentage of patients impacted by COVID-19 are summarized by treatment analysis group for the following categories:

- Drug Interruption due to COVID-19
  - Tested positive for COVID-19 (yes, no, not tested)
  - Primary reason for interruption (medical monitor/sponsor recommendation, site mandate, patient choice, insufficient IP supply, other)

Patient narratives are provided for all patients who test positive for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of COVID-19.

Early discontinuations of treatment or study due to COVID-19 are summarized and listed (specified in Section 7.1).

All deviations due to the impacts of COVID-19 (after 01 Mar 2020) are identified and documented accordingly by the site and the Sponsor. The failure to complete a protocol visit is not considered as a reason for study discontinuation or a major deviation. All COVID-19 related deviations will be identified in the protocol deviation listing (Section 7.3).

11.2. Efficacy

Subgroup analyses are performed as appropriate to assess the potential impact of the COVID-19 pandemic on the treatment effect (Section 6.2.2). Missing data due to COVID-19 are not imputed. Patients affected by COVID-19 with visits not done, visits outside the window, one or more efficacy procedures not performed, and other impact to study assessments or procedures are listed.

11.3. Safety

For patients who are unable to complete an in-clinic visit, safety assessments can be conducted via telemedicine (eg, phone or video). At a minimum, telemedicine visits for all sites complete the following assessments:

- AE assessment.
- Prior and concomitant medication assessment.

- Pregnancy testing for women of childbearing potential (WOCBP).

Where in-person clinic visits are not possible, in-home visits may be provided in sites in Australia, Austria, Italy, and United States by home health services to perform the following assessments:

- Urine pregnancy tests for WOCBP (where applicable).
- Vital signs.
- Safety labs (not required, may be performed upon site request).

In the UK, every effort is made to obtain safety laboratory assessments every 24 weeks using either the central laboratory, local laboratories, or home health services (if available). In cases when laboratory assessments cannot be performed as a result of COVID-19 related restrictions, alternative proportionate mechanisms of oversight should be followed to ensure patient safety and well-being as specified in Protocol Appendix 13.

Patients with changes in safety data collection methods due to COVID-19 including telephone, video conference, home health, local lab and other methods, are listed.

Subgroup analyses are performed to assess the potential impact of the COVID-19 pandemic on the treatment effect on safety outcomes (Section 6.2.2).

11.4. Visit Window

When the investigator deems that the patients recovered from COVID-19 can safely return to the site for an in-person visit, the investigator should conduct an unscheduled visit and perform all assessments as in Week 4. If there are no changes in clinical status that would preclude resuming IP, IP may be resumed. A follow-up telephone visit should be conducted approximately 2 weeks after this unscheduled visit to collect AEs and prior and concomitant medications. Approximately 4 weeks after the unscheduled visit, the patient should return for repeat assessments as in the Week 4 visit. The usual visit schedule may resume thereafter. For patients recovered from COVID-19 infection, the week 4 unscheduled visit and scheduled Week 4 visit are mapped into the analysis visit window (Section 6.4.7) based on the study day accordingly.

12. REFERENCES

Guy, W. ECDEU Assessment Manual For Psychopharmacology, DHEW Publication No. ADM 76–338, US Government Printing Office, Washington, DC, USA, 1976.

Lynch DR, Farmer JM, Wilson RL, Balcer LJ. Performance measures in Friedreich ataxia: potential utility as clinical outcome tools. Mov Disord. 2005 Jul;20(7):777-82.

Lynch DR, Farmer JM, Tsou AY, Perlman S, Subramony SH, Gomez CM, Ashizawa T, Wilmot GR, Wilson RB, Balcer LJ. Measuring Friedreich ataxia: complementary features of examination and performance measures. Neurology. 2006 Jun 13;66(11):1711-6.

13. APPENDIX

13.1. SAP Change Summary

SAP Revision Chronology:			
SAP version 1.0	26-Aug-2019	Original	Main Rationale*
SAP version 2.0 (Amendment 1)	10-Jul-2020	• Added COVID-19 impact section • Added COVID-19 subgroup for safety and efficacy analyses • Updated patient disposition section to include early termination due to COVID-19 and remove enrolled population • Added protocol deviation section • Added summary statistics for PGIC and CGIC besides frequency counts • Removed physical exam, electrocardiogram, echocardiogram and some laboratory parameters from listing and summary since data is not collected	Protocol Amendment
SAP version 3.0 (Amendment 2)	12-Aug-2021	• Updated schedule of assessments to align with amended protocol • Expanded visit window for additional ± 28 days • Added categorical summaries for clinical laboratory assessment since additional lab values are available • Specified that patients with lab and vital signs assessment values after the start date of a moderate/severe or serious adverse event preferred term of coronavirus are summarized separately • Removed COVID-19 Yes/No subgroup from safety summary • Added summary tables treatment-emergent nonserious AEs	Protocol Amendment

SAP version 4.0 (Amendment 3)	13-Feb-2023	• Updated Schedule of Assessments to align with amended protocol • Added ECHO and ECG summary due to the availability of post baseline assessments • Added description of Reata's internal safety management team oversight of the open-label extension study	Protocol Amendment
SAP version 5.0 (Amendment 4)	12-Aug-2025	• Update MedDRA version of Adverse Event and Medical history from 14.1 to 28.0 due to database up-versioning. • Update concomitant medications WHO drug dictionary version from September 2011 B2 format to Global B3 March 2025 due to database up-versioning. • Update the TSR section based on protocol v14.	Protocol Amendment and MedDRA/WHO drug up-versioning
SAP version 6.0 (Amendment 5)	12-Jan-2026	• Update MedDRA version of Adverse Event and Medical history from 28.0 to 28.1 due to database up-versioning. • Update concomitant medications WHO drug dictionary version from Global B3 March 2025 to Global B3 September 2025 due to database up-versioning.	MedDRA/WHO drug up-versioning

* Detailed rationale for each change item is included in a separate SAP amendment change document. (Excluding amendments #4 and #5 due to minimum changes).

13.2. Programming Specifications

Continuous data are listed corresponding to the precision measured or calculated. Measures of central tendency are presented using one more decimal place than the precision of the data. Summaries of variability are presented using two significant digits more than the precision of the underlying data. Quartiles, the minimum, and the maximum are presented using the precision of the data.

All percentages are to be expressed as integers with one decimal place. The convention for rounding percentages is as follows:

- Values greater than or equal to x.x5% are rounded up
- Values between 0 and x.x5% are rounded down

13.3. List of Laboratory Tests

Blood samples are collected throughout the study for hematology, chemistry, and urinalysis for clinical laboratory evaluation. Specific clinical laboratory evaluations are shown in Table 6.

Table 6: List of Laboratory Tests

Hematology	Chemistry	Urinalysis
Hematocrit	Blood urea nitrogen (BUN)	Specific gravity Ketones pH
Hemoglobin	B-type natriuretic peptide (BNP)	Protein
HbA1C	N-terminal prohormone of B-type natriuretic peptide (NT-pro BNP)	Blood
Red blood cell (RBC) count	Creatinine	Glucose
White blood cell (WBC) count	Total bilirubin	Urobilinogen
Neutrophils	Alanine aminotransferase (ALT)	Bilirubin
Neutrophils/Leukocytes	Aspartate aminotransferase (AST)	
Lymphocytes	Alkaline phosphatase (ALP)	
Lymphocytes /Leukocytes	Ferritin	
Monocytes	Sodium	
Monocytes/Leukocytes Basophils (if detected)	Potassium	
Basophils/Leukocytes Eosinophils (if detected)	Calcium	
Eosinophils/Leukocytes Absolute platelet count	Inorganic phosphorus	
Mean corpuscular hemoglobin (MCH)	Magnesium	
Mean corpuscular volume (MCV)	Chloride	
Mean corpuscular hemoglobin concentration (MCHC)	Bicarbonate	
Reticulocytes/Erythrocytes	Uric acid	
	Cholesterol	
	Total protein	
	Glucose	
	Triglycerides	
	Albumin	
	Creatine phosphokinase (CPK)	
	Lactate dehydrogenase (LDH)	
	High-density lipoprotein cholesterol (HDL-C)	
	Low-density lipoprotein cholesterol (LDL-C)	
	Very-low-density lipoprotein cholesterol (VLDL-C)	
	Gama-glutamyl transpeptidase (GGT)	
	Estimated glomerular filtration rate (eGFR) using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) / Schwartz Bedside Equation	

docusign

Certificate Of Completion

Envelope Id: [redacted]
Status: Completed
Subject: Complete with Docusign: 408-C-1402 Extension Statistical Analysis Plan (SAP) v6.0 Final.pdf
Vendor Name:
Source Envelope:
Document Pages: 50
Signatures: 3
Certificate Pages: 2
Initials: 0
AutoNav: Enabled
EnvelopeId Stamping: Enabled
Time Zone: (UTC-05:00) Eastern Time (US & Canada)

Envelope Originator:
[redacted]
225 Binney St
Cambridge, MA 02142
[redacted]
IP Address: [redacted]

Record Tracking

Status: Original
12 January 2026 | 12:26

Holder: [redacted]

Location: DocuSign

Signer Events	Signature	Timestamp
[redacted] Study Medical Director, [redacted] Biogen Security Level: Email, Account Authentication (None)	Signed by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 12 January 2026 | 12:28 Viewed: 12 January 2026 | 12:52 Signed: 12 January 2026 | 12:53
Electronic Record and Signature Disclosure: Not Offered via Docusign		
[redacted] Sr. Principal Biostatistician Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 12 January 2026 | 12:28 Viewed: 12 January 2026 | 12:28 Signed: 12 January 2026 | 12:28
Electronic Record and Signature Disclosure: Not Offered via Docusign		
[redacted] Biogen Security Level: Email, Account Authentication (None)	DocuSigned by: [redacted] Signature Adoption: Pre-selected Style Using IP Address: [redacted]	Sent: 12 January 2026 | 12:28 Viewed: 12 January 2026 | 14:22 Signed: 12 January 2026 | 14:22
Electronic Record and Signature Disclosure: Not Offered via Docusign		

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	12 January 2026 | 12:28
Certified Delivered	Security Checked	12 January 2026 | 14:22
Signing Complete	Security Checked	12 January 2026 | 14:22
Completed	Security Checked	12 January 2026 | 14:22

Payment Events	Status	Timestamps