STATISTICAL ANALYSIS PLAN

A RANDOMIZED, PARALLEL-ARM, DOUBLE-BLIND, PLACEBO CONTROLLED STUDY WITH OPEN-LABEL EXTENSION TO ASSESS THE EFFICACY AND SAFETY OF VATIQUINONE FOR THE TREATMENT OF FRIEDREICH ATAXIA (MOVE-FA)

PTC743-NEU-003-FA

28 MARCH 2023

VERSION 2.0

PTC THERAPEUTICS, INC.
100 CORPORATE COURT
SOUTH PLAINFIELD, NJ 07080 USA

APPROVAL SIGNATURES

[redacted]
PTC Therapeutics, Inc.

Date

[redacted]
PTC Therapeutics, Inc.

Date

[redacted]
PTC Therapeutics, Inc.

Date

The eSignature page is located on the last page.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation	Term
1MWT	1-minute walk test
AE	Adverse Event
ATC	Anatomical Therapeutic Chemical
BMI	Body mass index
CI	confidence interval
CRF	Case report form
DSMB	Data Safety Monitoring Committee
ECG	electrocardiogram
FA	Friedreich Ataxia
FARS-ADL	Friedreich Ataxia Rating Scale Activities of Daily Living
ICE	Intercurrent Event
IRT	interactive response technology
ITT	intention to treat
LoA	loss of ambulation
LOCF	last observation carried forward
MCMC	Monte Carlo Markov Chain
MedDRA	Medical Dictionary for Regulatory Activities
mFARS	modified Friedreich Ataxia Rating Scale
MFIS	Modified Fatigue Impact Scale
MI	multiple imputation
mITT	modified intention to treat
MMRM	mixed model for repeated measures
PK	pharmacokinetic
PO	orally
PP	per-protocol
SAE	serious adverse event
SAP	statistical analysis plan
TEAE	Treatment emergent adverse Event
TFL	table, figure, and listing
tid	ter in die (three times a day)
WHO	World Health Organization

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze data and report results for PTC743-NEU-003-FA (version 5.0 dated 21 May 2021). Table, figure, and listing (TFL) specifications are contained in a separate document.

This SAP is to be reviewed and approved prior to study database unblinding. There will be a database soft lock and a database lock for this study. The database soft lock will occur when all subjects have completed the double-blinded treatment period (72 weeks). The database lock will occur when all subjects have completed the open-label treatment period (96 weeks).

1.1. Study Design

This study will be a parallel-arm, double-blind, placebo-controlled trial during which eligible subjects will be randomized 1:1 to receive either vatiquinone or placebo for 72 weeks using an interactive response technology (IRT) at the Baseline visit (Figure 1). Subjects will be stratified by:

- modified Friedreich Ataxia Rating Scale (mFARS) (<40 versus ≥40),
- age at disease onset (<14 years or ≥14 years), and
- age at Screening (≤21 years or >21 years).

If randomized to vatiquinone, subjects will receive a dose of either 200 mg orally (PO) three times a day (tid) if <12 years of age and weighing <25 kg or 400 mg PO tid if ≥12 years of age and/or ≥25 kg.

Weight will be assessed during every clinic visit, and the dose of study drug will be adjusted to 400 mg or down to 200 mg, as applicable, for any change in weight of ±5 kg from the 25 kg dosing determination threshold. The dose will also be adjusted if the subject's age changes from 11 to 12 years. If randomized to placebo, subjects will receive matching placebo PO tid. Following completion of the randomized placebo-controlled phase, subjects will enter an open-label (24-week) extension phase during which they will receive open-label treatment with vatiquinone at the dose they received in the randomized phase of the study and then a safety follow-up.

For subjects entering the extension phase who initially received vatiquinone, they will continue to receive the same dose of vatiquinone (unless there has been a change in weight or a change in age from 11 to 12 years [see Protocol Section 3.1.2 and 3.1.4]). For subjects entering the extension phase who initially received placebo, the dose of vatiquinone will be determined based on weight as described in Protocol Sections 3.1.2 and 3.1.4. Body weight will continue to be assessed at each visit during the extension phase and dose will be adjusted accordingly, as above. The maximum duration of subject treatment will be 96 weeks (including placebo-controlled phase and extension phase). The overall study duration will be a maximum of approximately 106 weeks per subject. Subjects who complete the study may have the option to enter an open-label expanded access program to continue receiving vatiquinone.

The proposed types and timing of data to be recorded are described in Appendix 1

Figure 1: Study Design

1.2. Study Objectives

1.2.1. Primary Objective

The primary objective of the study is to evaluate the efficacy (using mFARS) and safety of vatiquinone in subjects with Friedreich Ataxia (FA).

1.2.2. Secondary Objectives

Secondary objectives of the study are to:

- Demonstrate the effects of vatiquinone on activities of daily living as assessed by the Friedreich Ataxia Rating Scale Activities of Daily Living (FARS-ADL) scale
- Demonstrate the effects of vatiquinone on ambulation as assessed by the 1-minute walk test (1MWT)
- Demonstrate the effects of vatiquinone on falls as assessed by a fall log

1.2.3. Exploratory Objectives

The exploratory objectives of the study are to:

- Demonstrate the effects of vatiquinone on speech using Redenlab protocol
- Demonstrate the effects of vatiquinone on fatigue as assessed by the Modified Fatigue Impact Scale (MFIS)
- Demonstrate the effects of vatiquinone on health status using the EQ-5D-5L
- Demonstrate the effects of vatiquinone on the Upright Stability subscale of the mFARS

1.3. Endpoints

1.3.1. Primary Endpoint

The primary endpoint will be the change from baseline in the mFARS at Week 72.

1.3.2. Secondary Endpoints

The key secondary efficacy endpoint of the study is change from baseline in the following assessment at Week 72:

- FARS-ADL scale (key secondary endpoint)

Other secondary endpoints include change from baseline in the following assessments at Week 72:

- 1MWT
- Fall log

1.3.3. Exploratory Endpoint

Exploratory endpoints will include change from baseline in the following assessments at Week 72.

- Speech using Redenlab protocol
- MFIS
- EQ-5D-5L
- Upright Stability subscale of the mFARS

1.4. Sample Size Determination

The sample size and power calculation are performed based on the primary efficacy endpoint, change from baseline in mFARS score of subjects between 7 and 21 years old. Based on Phase 2 Study EPI-2010-006 data, vatiquinone 400 mg tid dose demonstrated a change from baseline in mFARS score for 400 mg dose of -1.31 after 18 months of treatment, and natural history showed an estimated increase of 3.26 after 18 months of disease progression (Patel 2016). Assuming a common standard deviation of 6.76, with 1-sided type I error of 2.5% and 90% power, 47 subjects per group is required to detect a treatment difference of -4.57 in mFARS score after 72 weeks of vatiquinone treatment. Assuming a dropout rate of 10%, a total of 106 subjects between 7 and 21 years of age will be randomized to receive vatiquinone or placebo in 1:1 ratio.

In order to explore the treatment efficacy and safety, approximately 20 subjects >21 years of age will also be randomized to receive vatiquinone or placebo in 1:1 ratio for a total of approximately 126 subjects in the study.

2. STUDY ANALYSIS SETS AND PERIODS FOR ANALYSIS

2.1. Study Analysis Sets

2.1.1. Randomized Analysis Set

The Randomized analysis set will include all subjects who are randomized for the study.

2.1.2. Intention-to-Treat Analysis Set

The intent-to-treat (ITT) analysis set will include all subjects who are randomized, have received at least one dose of study drug (vatiquinone or placebo), and have baseline and at least one post-baseline measurement of the primary endpoint. In the ITT analysis, subjects will be analyzed according to the treatment group the subjects are randomized to, regardless of actual treatment received.

Analysis in mFARS change from baseline to Week 72 based on the ITT analysis set will be considered as sensitivity analysis.

2.1.3. Modified Intent-to-Treat Analysis Set

The modified intent-to-treat (mITT) analysis set will include all subjects in the ITT analysis set who are between age of 7 and 21 years, inclusive, at Screening. In the mITT analysis, subjects will be analyzed according to the treatment group the subjects are randomized to, regardless of actual treatment received.

Analysis in mFARS change from baseline to Week 72 performed based on the mITT analysis will be considered as primary efficacy analysis.

2.1.4. Safety Analysis Set

The safety analysis set will include all subjects who receive at least one dose of study drug. This will be the primary analysis set for summarizing safety data. Subjects will be summarized according to the actual treatment received.

The Safety analysis set will be used in the statistical analyses for safety.

2.1.5. Per-Protocol Analysis Set

The per-protocol (PP) analysis set is a subset of the mITT analysis set. Subjects who meet the following criteria will be excluded from the PP analysis set:

- Received study treatment different from the randomized treatment throughout the double-blind period
- Did not have a valid mFARS at baseline or a valid Week 72 mFARS within protocol defined window
- Non-compliance to study drug administration
- Had significant inclusion or exclusion criteria violations
- Had major protocol deviations which may impact effectiveness of study treatment

This analysis set will be used for primary efficacy endpoints as supportive efficacy analyses.

The list of subjects excluded from the PP analysis set will be finalized prior to treatment unblinding.

2.1.6. Pharmacokinetic Analysis Set

The PK analysis set will include all subjects who have any evaluable plasma drug concentration.

2.2. Study Periods

Efficacy analysis will be performed for the double-blind period, and/or overall study period.

Safety analysis will be performed for the overall period and on-vatiquinone period.

2.2.1. Double-blind Period

Double-blind period is defined as the period from randomization to:

- the date prior to the first dose of open-label study drug if the subject continues to the open-label extension period, or
- the last study visit if the subject does not continue to the open-label extension period.

Double-blind period will be used for summarizing efficacy and safety endpoints.

2.2.2. Overall Study Period

Overall study period is defined as the period from the first dose of double-blind study treatment to last study visit.

Overall study period will be used for summarizing efficacy endpoints.

2.2.3. On-vatiquinone Period

On-vatiquinone period is defined as the period from the first dose of vatiquinone, regardless of double-blind or open-label vatiquinone, to the end of the study.

On-vatiquinone period will be used for summarizing safety endpoints.

3. GENERAL CONSIDERATIONS

3.1. Definition of Estimand

Endpoint	Population	Variable	Intercurrent[a]	Summary	Analysis
mFARS	Subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (mITT)	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Treatment difference and corresponding 95% confidence interval (CI) between vatiquinone and placebo in change from baseline to Week 72 in mFARS will be estimated using a mixed model repeated measure (MMRM) after multiple imputation (MI)	Primary analysis for primary efficacy endpoint
mFARS	Subjects with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (ITT)	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Same as primary analysis for primary efficacy endpoint	Sensitivity analysis for primary efficacy endpoint
mFARS	Subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria, with no major protocol violation which may affect treatment effectiveness (PP)	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Same as primary analysis for primary efficacy endpoint	Sensitivity analysis for primary efficacy endpoint

Endpoint	Population	Variable	Intercurrent[a]	Summary	Analysis
mFARS	Completers of subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (mITT completers)	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Same as primary analysis for primary efficacy endpoint	Sensitivity analysis for primary efficacy endpoint
mFARS	Completers of subjects with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (ITT completers)	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Same as primary analysis for primary efficacy endpoint	Sensitivity analysis for primary efficacy endpoint
mFARS	Subjects with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (mITT) excluding subjects with LoA during the double blinded period	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Same as primary analysis for primary efficacy endpoint	Sensitivity analysis for primary efficacy endpoint
mFARS	Subjects with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (ITT) excluding subjects with LoA during the double blinded period	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Same as primary analysis for primary efficacy endpoint	Sensitivity analysis for primary efficacy endpoint

Endpoint	Population	Variable	Intercurrent[a]	Summary	Analysis
mFARS	Subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (mITT)	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Treatment difference and corresponding 95% CI between vatiquinone and placebo in change from baseline to Week 72 in mFARS will be estimated using a mixed model repeated measure (MMRM) assuming missing assessments are missing at random (MAR)	Sensitivity analysis for primary efficacy endpoint
mFARS	Subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (ITT)	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Treatment difference and corresponding 95% CI between vatiquinone and placebo in change from baseline to Week 72 in mFARS will be estimated using a mixed model repeated measure (MMRM) assuming missing assessments are missing at random (MAR)	Sensitivity analysis for primary efficacy endpoint
mFARS	Subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by	Change from baseline in mFARS at Week 72	Early discontinuation: Data will be used as collected.	Change from baseline to Week 72 in mFARS will be compared between vatiquinone and placebo using	Sensitivity analysis for primary efficacy endpoint

Endpoint	Population	Variable	Intercurrent[a]	Summary	Analysis
	the protocol inclusion/exclusion criteria (mITT)			Wilcoxon rank sum test	
mFARS	Subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (mITT)	Change from baseline in mFARS at Week 72	Early discontinuation: Control-based MI will be performed after discontinuation.	Treatment difference and corresponding 95% confidence interval (CI) between vatiquinone and placebo in change from baseline to Week 72 in mFARS will be estimated using a mixed model repeated measure (MMRM), control-based multiple imputation (MI) will be performed for subjects' data collected after early discontinuation.	Sensitivity analysis for primary efficacy endpoint
FARS-ADL	Subjects between 7 and 21 years of age, inclusive with Friedreich ataxia (FA) diagnosed and confirmed by clinical testing as defined by the protocol inclusion/exclusion criteria (mITT)	Change from baseline in FARS-ADL at Week 72	Early discontinuation: Data will be used as collected.	Treatment difference and corresponding 95% CI between vatiquinone and placebo in change from baseline to Week 72 in FARS-ADL will be estimated using a mixed model repeated measure (MMRM) after MI	Analysis of key secondary efficacy endpoint
FARS-ADL	Subjects with Friedreich ataxia (FA) diagnosed and confirmed by clinical	Change from baseline in FARS-	Early discontinuation: Data will be	Same as the primary analysis for secondary endpoint	Analysis of key secondary

Endpoint	Population	Variable	Intercurrent[a]	Summary	Analysis
	testing as defined by the protocol inclusion/exclusion criteria (ITT)	ADL at Week 72	used as collected.		efficacy endpoint

Abbreviations: MI, multiple imputation; mITT, modified intention-to-treat; ITT, intention-to-treat.
a: Early discontinuation=treatment-emergent adverse events leading to discontinuation and death.

3.1.1. Intercurrent Events

Intercurrent events (ICEs) are defined as events that occur after treatment initiation and either preclude observation of the variable of interest or affect its interpretation. The following events will be considered to be ICEs this study:

- Treatment emergent adverse events leading to discontinuation
- Death

Sensitivity analysis performed regarding these ICEs are described in Section 5.1.3.

3.2. General Considerations

For continuous variables, median, mean, standard deviation, minimum, maximum, and number of subjects with non-missing data will be provided for each treatment group, and for overall, where applicable. For categorical variables, the number (percent) of subjects in each category will be provided.

Subjects randomized to vatiquinone will be treated as one group regardless of the dose (200 mg tid or 400 mg tid).

Assessments of change from baseline to post-treatment or the ratio of post-treatment to baseline will include only those subjects with both baseline and post-treatment measurements. The last non-missing value of a variable taken before the first dose of study drug will be used as the study baseline value, unless otherwise specified. In general, missing or dropout data will not be imputed for the purpose of data analysis, unless otherwise specified.

All descriptive summaries will be based on corresponding study analysis visits defined in Section 3.8. All efficacy analysis using mixed model for repeated measure (MMRM) will also use study analysis visit.

In general, statistical inferential statistics will only be performed for double-blinded period using a 2-sided test at the 0.05 significance level.

The descriptive summaries will be presented by treatment groups (placebo, vatiquinone) and in total (where applicable) during double-blind period for both efficacy and safety endpoints. In addition, on-vatiquinone experience from both double-blind and open-label extension period will be combined and summarized for safety endpoints. In this summary, baseline for subjects on placebo in the double-blind period will be the last observation prior to the first vatiquinone dose.

3.3. Data Safety Monitoring Board

A Data Safety Monitoring Board (DSMB) has been set up for this study. The primary responsibility of the DSMB is to protect the safety and welfare of subjects participating in this clinical trial.

Periodic DSMB meetings are planned to monitor subject safety. Unblinded outputs on subject disposition, demographic and baseline characteristics, efficacy, and safety will be reviewed by DSMB members and the unblinded data analyses will be performed by an independent statistical group. Results of efficacy endpoints, if there are any, will be reviewed for safety purpose. To maintain the credibility and integrity of the trial, procedures will be implemented to ensure the DSMB and the independent statistician have sole access to the unblinded safety monitoring data of which PTC Therapeutics personnel will not have any knowledge or access to. Details are provided in the DSMB charter and DSMB SAP.

3.4. Interim Analyses

No formal interim analyses are planned for this study.

However, there will be 2 database locks for this study. The first database lock (soft lock) will occur when all subjects have completed the double-blind treatment (Week 72). The final database lock will occur when all subjects have completed the open-label treatment period.

At the first database lock, all analyses for only the data from double-blind period will be performed as described in this SAP. For the partial data from the open-label period data up to the database lock, only descriptive summary will be provided.

At the second database lock, all analyses based on double-blind period will not be repeated. All analyses including data from open-label period will be performed as described in this SAP using completed data from open-label period.

3.5. Baseline

Baseline for double-blind period will be defined as the last measurement prior to randomization for efficacy endpoints, and prior to first dose for safety endpoints. Baseline (Day 1) is considered as the baseline visit.

For summary based on on-vatiquinone period, the baseline will be defined as the last measurement prior to the first dose of vatiquinone.

3.6. Multiplicity Adjustment

In order to control the family-wise error rate for the primary and key secondary efficacy endpoints, a fixed sequence procedure will be used. Testing orders of the primary endpoint and the key secondary endpoints are listed as below

- Primary endpoint point: change from baseline in the mFARS at Week 72
- The key secondary endpoints:

1. change from baseline in the FARS-ADL scale at Week 72
2. Change from baseline in the 1MWT at Week 72

3. Number of falls from Week 48 to Week 72

The primary endpoint will be tested at the 0.05 significance level (two-sided). If $p<0.05$, then the primary endpoint will be considered statistically significant, and the study will be declared positive. If the primary endpoint is statistically significant, the key secondary endpoints will be tested at the 0.05 significance level (two-sided) in the order specified above. Only if the first key secondary endpoint is statistically significant at the 0.05 significance level (two-sided), the second secondary endpoint will be tested at the 0.05 significance level (two-sided). The third secondary endpoint will also be tested at the same significance level if the second secondary endpoint is statistically significant.

No other efficacy or safety end points will be tested.

3.7. Missing Data Handling

For the primary efficacy endpoint, the missing assessments will be imputed based on pattern-mixture MI prior to fit MMRM model as described in Section 5.1.2.

Unless specified otherwise, missing data will not be imputed for descriptive summaries. mFARS subsection and total score with missing item values will be imputed as described in Section 5.1.1.

In order to minimize missing data in the study, PTC employed the strategies noted in Appendix 3.

3.8. Analysis Visit Windows

Study analysis visits during double-blind period, overall study period, and on-vatiquinone period will be derived as listed in below Table 1, Table 2, and Table 3, respectively, based on the days from randomization date to the corresponding visit date. Study analysis visits will be used in all by-visit summaries for both efficacy and safety assessments. Study analysis visits will also be used in the mixed model for repeated measure (MMRM) model.

Table 1: Analysis Window for Double-blind Period

Study Visit	Scheduled Visit Number (study day)	Analysis Window (study day)
DB Week 12	Week 12 (85)	[2, 126]
DB Week 24	Week 24 (169)	[127, 210]
DB Week 36	Week 36 (253)	[211, 294]
DB Week 48	Week 48 (337)	[295, 378]
DB Week 60	Week 60 (421)	[379, 462]
DB Week 72	Week 72 (505)	Any DB visit that occurred on or after study day 463

DB = double-blind. Study day for efficacy endpoints= visit date – randomization date + 1.
Study day for safety endpoints= visit date – treatment start date + 1.

Table 2: Analysis Window for Overall Study Period

Analysis Visit	Scheduled Visit Number (study day)	Analysis Window (study day)
Week 12	Week 12 (85)	[2, 126]
Week 24	Week 24 (169)	[127, 210]
Week 36	Week 36 (253)	[211, 294]
Week 48	Week 48 (337)	[295, 378]

Analysis Visit	Scheduled Visit Number (study day)	Analysis Window (study day)
Week 60	Week 60 (421)	[379, 462]
Week 72	Week 72 (505)	[463, 546]
Week 84	Week 84 (589)	[547, 630]
Week 96	Week 96 (673)	Any visit that occurred on or after study day 631

Study day for efficacy endpoints= visit date – randomization date + 1.
Study day for safety endpoints= visit date – treatment start date + 1.

Table 3: Study Visits for On-vatiquinone Period

Analysis Visit for Vatiquinone/ Vatiquinone	Scheduled Visit Number (On-vatiquinone day)	Analysis Window (On-vatiquinone days)	Analysis Visit for Placebo/ Vatiquinone	Scheduled Visit Number (On-vatiquinone day)	Analysis Window (On-vatiquinone day)
OV Week 12	Week 12 (85)	[2, 126]	OV Week 12	Week 84 (85)	[2, 126]
OV Week 24	Week 24 (169)	[127, 210]	OV Week 24	Week 96 (169)	Any On vatiquinone visit occurred after 127 OV days
OV Week 36	Week 36 (253)	[211, 294]			
OV Week 48	Week 48 (337)	[295, 378]			
OV Week 60	Week 60 (421)	[379, 462]			
OV Week 72	Week 72 (505)	[463, 546]			
OV Week 84	Week 84 (589)	[547, 630]			
OV Week 96	Week 96 (673)	Any visit occurred on or after 631 days			

OV = on-vatiquinone. On-vatiquinone day is calculated as visit date – first dose of vatiquinone date + 1.

All assessments will be assigned to a study analysis visit based on study days. For a given subject, if multiple assessments are within the same analysis window, the one closest to the scheduled study day will be used for that analysis visit. In case of equal-distance, the later assessment will be used for that given analysis visit.

3.9. Changes of Analysis from Protocol

There are no major changes to the planned analyses from the protocol. LoA was added as a tertiary/exploratory endpoint.

4. SUBJECT DATA

4.1. Participant Dispositions

The disposition of subjects in the double-blind period, including the number of subjects screened, randomized, randomized subjects who received at least 1 dose of study drug, subjects who prematurely discontinue study drug, and the reason for the premature termination will be tabulated for mITT, ITT, Safety, and PP analysis sets.

The similar disposition table will also be produced for open-label extension period for mITT, ITT, and Safety analysis sets.

4.2. Duration of Treatment with Study Drug

Duration of treatment with study drug in double-blinded period will be calculated as number of days from the date of the first dose of double-blinded study drug to the date of the last dose of the double-blinded study drug, inclusive. Duration of treatment with study drug in overall study period will be calculated as number of days from the date of the first dose of study drug to the date of the last dose of the study drug. Total duration of vatiquinone will be calculated as number of the days from the date of the first dose of vatiquinone to the date of the last dose of vatiquinone, inclusive.

Duration of treatment will be summarized by treatment for double-blinded period for mITT, ITT, and Safety analysis sets, by treatment sequence for open-label period for mITT, ITT, and Safety analysis sets, and for total duration for on-vatiquinone for Safety Analysis set.

Duration of treatment will also be summarized by the following categories.

- Double-blind period
  - <23 weeks
  - ≥23 to <47 weeks
  - ≥47 to <71 weeks
  - ≥71 weeks
- On-vatiquinone period/Overall study period
  - <23 weeks
  - ≥23 to <47 weeks
  - ≥47 to <71 weeks
  - ≥71 to <95 weeks
  - ≥95 weeks

4.3. Study Drug Compliance

To evaluate study drug compliance, number of subjects and number of incidences of missed doses and overdoses will be summarized for mITT, ITT, and Safety analysis sets for double-

blind period, for mITT and ITT for open-label period, and for Safety analysis set for on-vatiquinone period.

4.4. Demographic and Baseline Characteristics

Demographics and baseline characteristics including age, sex, race, ethnicity, region, country, height, weight, body mass index (BMI), and stratification factors baseline mFARS score (<40 versus ≥40), age of disease onset (<14 years versus ≥14 years), age at screening (≤21 versus >21 years old), and baseline mFARS, FARS-ADL, 1MWT, and MFIS, will be summarized for mITT, ITT, Safety, and PP analysis sets.

4.5. Disease Characteristics

FA disease characteristics including age of FA diagnosis, age of FA symptom onset, GAA repeat length Allele 1, GAA repeat length Allele 2 will be summarized for mITT, ITT, Safety, and PP analysis sets.

4.6. Medical History

Medical history will be coded with the Medical Dictionary for Regulatory Activities (MedDRA, version 23.0). Medical history will be summarized by System Organ Class (SOC) and preferred terms (PT) by treatment group and overall for mITT, ITT, and Safety analysis sets.

4.7. Concomitant Medications and Non-Drug Treatments

All investigator terms for medications recorded on the case report form (CRF) will be coded using the World Health Organization (WHO) Drug Dictionary (Version MAR2020 or later).

Prior medications will be defined as medications started prior to the first dose of study drug.

Concomitant medications and non-drug treatments during double-blind period are defined as those taken any time during the double-blind treatment period (the date of the first study treatment to the date prior to the first dose of open-label study treatment, inclusive). Concomitant medications and non-drug treatments during open-label period and on-vatiquinone period are defined similarly.

Any medication started before the first dose of study drug and continued at the time of first dose will be considered as both prior and concomitant medications.

Any medication with partial or missing start date in which the prior medication status cannot be determined will be considered as a prior medication. Similarly, any medication with partial or missing end date in which the concomitant medication status for a period cannot be determined will be considered as a concomitant medication for that period.

Prior medication and concomitant medication in double-blind period will be summarized by Anatomical Therapeutic Chemical (ATC) level 2 and PT for mITT, ITT, and Safety analysis sets. In addition, concomitant mediation in on-vatiquinone period will also be summarized by ATC level 2 and PT for Safety analysis set.

5. EFFICACY ANALYSIS

Efficacy analysis will be performed based on the mITT and ITT analysis sets, unless otherwise specified.

5.1. Primary Endpoint(s) Analysis

5.1.1. Definition of Endpoint(s)

The change from baseline in total mFARS at Week 72 in the mITT analysis set will be the primary efficacy endpoint.

The mFARS consists of 25 items with a maximum total score of 93 in 4 subsections, bulbar function (2 items with a maximum score of 5), upper limb coordination (10 items with a maximum score of 36), lower limb coordination (4 items with a maximum score of 16), and upright stability (9 items with at maximum score of 36). The sum of the scores within each subsection will be used for the score for each subsection. The sum of all 25 items will be used to form the mFARS score. Higher score reflects greater Friedreich ataxia disease severity; thus, a negative change from baseline suggests an improvement.

For upright stability subsections, the items e2a, e2b, e3a, e3b, e4, and e5 will be repeated three times; the item-level score will be the average of available assessments.

If a subject has missing score(s) for any subsection at any visit, the following imputation approach will apply:

- Within each subsection, if more than a total of maximum possible score of 20% is missing, that subsection will be considered missing
- If less than a total of maximum possible score of 20% is missing, that subsection score will be imputed.

Subsection score will be imputed proportionally based on the ratio of maximum possible score of non-missing items to total of maximum possible score of that subsection (ie, if the total score is X out of possibly Y from the non-missing items in a subsection, the subsection score will be imputed by X/Y * total maximum possible score in that subsection). Total mFARS score will be the sum of all four subsection scores after imputation. If any one of the subsection scores is missing, the total mFARS score will be considered missing.

Table 4: Measurements of mFARS

Subsection	Examination Code	Examination	Maximum possible scale
Bulbar	A3	Cough	2
	A4	Speech	3
Upper limb coordination	B1a	Finger to finger (Right)	3
	B1b	Finger to finger (Left)	3
	B2a	Nose-finger (Right)	4
	B2b	Nose-finger (Left)	4
	B3a	Dysmetria (Right)	4
	B3b	Dysmetria (Left)	4
	B4a	Rapid movements (Right)	3

Subsection	Examination Code	Examination	Maximum possible scale
	B4b	Rapid movements (Left)	3
	B5a	Finger taps (Right)	4
	B5b	Finger taps (Left)	4
Lower limb coordination	C1a	Heel-shin slide (Right)	4
	C1b	Heel-shin slide (Left)	4
	C2a	Heel-shin tap (Right)	4
	C2b	Heel-shin tap (Left)	4
Upright stability	E1	Sitting position	4
	E2a	Stance feet apart	4
	E2b	Stance feet apart with eyes closed	4
	E3a	Stance feet together	4
	E3b	Stance feet together with eyes closed	4
	E4	Tandem stance	4
	E5	Stance, dom, foot	4
	E6	Tandem walk	3
	E7	Gait	5

5.1.2. Primary Analysis

For primary analysis, a MMRM model with MI described as below will be used. The estimated treatment difference in change from baseline at Week 72 with corresponding 95% CI will be provided.

Missing mFARS score will be imputed using pattern mix model MI assuming missing values are missing not at random (MNAR), and the results will be combined using Rubin's rule (Little 2002) as described in Appendix 2

The change from baseline at each visit will be calculated using the completed data sets after imputation for analysis. The primary efficacy analysis will be performed for subjects between 7 and 21 years of age at screening, and will compare change from baseline in total mFARS to Week 72 between vatiquinone- and placebo-treated subjects using mixed model repeated measure (MMRM) stratified by baseline mFARS score (<40 and ≥40) and age of onset (<14 and ≥14 years) with treatment, region, analysis visit, and analysis visit by treatment interaction as the main effects, and baseline mFARS and age at baseline as covariates using an unstructured variance-covariance matrix structure. If the model does not converge under the unstructured covariance matrix due to overparameterization, less complex alternative variance-covariance structures will be considered in the following order until convergence is reached: Heterogenous Toeplitz, Heterogenous Compound Symmetry, Toeplitz, and Compound Symmetry.

The primary endpoint will be tested at the 0.05 significance level (two-sided). If $p<0.05$, then the primary endpoint will be considered statistically significant, and the study will be declared positive.

The SAS sample code for MMRM model is:

```
PROC MIXED data=<dataset> method=reml;
  CLASS Subjid Trt CmFARS CAgeOS Region Week;
```

```
    MODEL CFBmFARS = Trt CmFARS CAgeos Region BLmFARS Age Week
  Week*TRT/S CL;
    REPEATED Week / TYPE=UN Subject=Subjid;
    LSMEANS TRT*Week / pdiff cl;
  RUN;
```

where Subjid = Subject ID
CFBmFARS = Change from baseline in mFARS at each visit
Trt = Treatment
CmFARS = Baseline mFARS category
CAgeOS = Age of onset category
BLmFARS = Baseline mFARS
Age = Age at baseline

5.1.3. Sensitivity Analysis

The similar analysis as described in the primary analysis will be repeated using the ITT and PP analysis sets as supportive analyses.

A sensitivity analysis will be performed for completers, i.e., all subjects who took medication till the end of double-blind period as specified by the protocol and have mFARS at Baseline and Week 72 visit, for mITT and ITT analysis sets.

Another sensitivity analysis will be performed similarly to the primary efficacy analysis on subjects who did not lose ambulation during double-blind period for mITT and ITT analysis sets. Loss of ambulation (LoA) is defined in Section 5.3.

In addition, another sensitivity analysis will be performed using the similar MMRM model as described for the primary analysis but assuming the missing assessments are missing at random (MAR) ie, without MI for mITT and ITT analysis sets.

As a sensitivity analysis, the change from baseline at Week 72 in mFARS using observed data between two treatment groups will be compared using Wilcoxon rank sum test stratified by baseline mFARS score (<40 and ≥40), age of onset (<14 and ≥14 years), and region for mITT analysis sets.

As another sensitivity analysis, for mITT analysis set, subjects with the ICEs of TEAE leading to discontinuation and death will be imputed using control-based MI method using subjects in the placebo arm who completed double-blind period of the study without missing values as reference.

5.1.4. By-Visit Summary

The mean and mean change from baseline in total mFARS and scores of each subsection (bulbar, upper limb coordination, lower limb coordination, upright stability) will be summarized by treatment group at each visit during overall study period and during the on-vatiquinone period for mITT and ITT analysis sets.

5.1.5. Subgroup Analysis

The mean and mean change from baseline in mFARS score will be summarized for the subgroups below for overall study period, respectively, for mITT and ITT analysis sets:

- Age of FA symptom onset (<14, ≥14 years old)
- Baseline mFARS score (<40, ≥40)
- Sex (female, male)
- Race
- Age at baseline (<18, ≥18 years old))

5.2. Secondary Endpoint(s) Analysis

5.2.1. Key Secondary Endpoint(s)

The key secondary endpoint is change from baseline in FARS-ADL scale to Week 72.

The FARS-ADL consists of 9 items. Each item is scored from 0 (normal) to 4 (most disabled). The sum of 9 items will be used to form the FARS-ADL scale. The maximum scale is 36. If more than 2 items are missing, the total scale will be considered as missing. If 7 or 8 items are not missing, the total scale will be standardized by multiplying the sum of the scale in the X activities by 9/X.

The comparison between two treatment groups will be analyzed using the same model as used in the primary efficacy analysis by using its respective baseline for mITT analysis set.

If the primary endpoint is statistically significant, this key secondary endpoint based on mITT analysis set will also be tested at the 0.05 significance level (two-sided).

The similar analysis will be repeated for ITT analysis set as a supportive analysis.

By-visit and subgroup summaries will also be provided similar to the primary efficacy endpoint.

5.2.2. Supportive Secondary Endpoint(s)

Other secondary efficacy variables include:

- Change from baseline in 1MWT to Week 72
- Number of falls over 72 weeks

The comparison between two treatment groups for change from baseline in 1MWT will be analyzed using the same model as used in the primary efficacy analysis by using its respective baseline for mITT and ITT analysis sets. For analysis of 1MWT, if a subject already become LoA (as defined in Section 5.3) and 1MWT is missing, 0 meter will be assigned to the visit for the subject for 1MWT. By-visit and subgroup summaries will also be provided similar to the primary efficacy endpoint.

If the key secondary endpoint FARS-ADL is statistically significant, change from baseline in 1MWT to Week 72 based on mITT analysis set will be tested at the 0.05 significance level (two-sided).

A summary of the number of falls per 28 days over every 24-week period, as well as over 72 weeks will be presented for mITT and ITT analysis sets. The number of falls per 28 days over last 24 weeks during double-blind period (Week 48 to Week 72) will be compared between the two treatment groups using Wilcoxon rank-sum test for mITT and ITT analysis sets. Number of falls per 28 days during a time interval will be calculated as the number of falls during the period

divided by the number of days during the interval, and multiply by 28. The falls occurred on or after the first LoA visit will be excluded from the analysis.

If the secondary endpoint 1MWT is statistically significant, number of falls over last 24 weeks during double-blind period (Week 48 to Week 72) based on mITT analysis set will also be tested at the 0.05 significance level (two-sided).

5.3. Tertiary/Exploratory Endpoint(s) Analysis

Exploratory efficacy variables include the following:

- Change from baseline in Speech evaluation to Week 72
- Change from baseline in MFIS to Week 72
- Change from baseline in EQ-5D-5L score to Week 72
- Change from baseline in total score of Upright Stability subsection of mFARS to Week 72
- Loss of Ambulation during double-blind period

Handling and comparison between treatment groups for change from baseline in speech will be provided by Redenlab and will be described in a separate document.

The MFIS consists of 21 questions. Each question is scored from 0 (never) to 4 (almost always). The sum of 21 items will be used to form the MFIS scale. The maximum score is 84. If more than 4 items are missing, the total scale will be considered as missing. If there are less than 4 items are missing, the total scale will be standardized by multiplying the sum of the score in the X activities by 21/X.

The comparisons between two treatment groups for change from baseline in MFIS, EQ-5D visual analogue scale (VAS), and Upright Stability subscale of mFARS will be analyzed using the same model as used in the primary efficacy analysis by using its respective baseline for mITT and ITT analysis sets. In addition, by-visit summaries will be provided for each question of MFIS.

The EQ-5D questionnaire has two components, health state description and evaluation. The descriptive part consists of health-related quality of life questionnaires covering 5 dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). Respondents score each dimension, which results in a one-digit number that expresses the level selected (no problems, slight problems, moderate problems, severe problems, and extreme problems) for that dimension. In the evaluation part, the respondents evaluate their overall health status using the visual analogue scale (EQ-VAS). The domain scores will be tabulated, and the mean and mean change in EQ-VAS will be summarized by treatment groups at each visit during overall study period and during the on-vatiquinone period for mITT and ITT populations.

LoA will be defined as attaining a maximum score of 5 on item E7 (unable to walk even with assistance, wheelchair bound) in mFARS assessments for the first time (Rummey 2020). Time to LoA during double-blind period will be evaluated using a Log-rank test and Kaplan-Meier plot, as well as a Cox proportional hazards model (assuming similar covariates as the primary efficacy model) in both mITT and ITT population. For subjects who do not have LoA during the double-blind period will be censored at the last time of mFARS assessment during the double-blind period. Subject who had a score of 5 at baseline will be excluded from this analysis. The proportion of LoA by Week 72 will also be summarized.

6. SAFETY ANALYSES

All safety analyses will be based on the Safety analysis set. Descriptive summary will be based on analysis visits defined in Section 3.8 by treatment group for double-blind period, and by treatment sequence and overall for on-vatiquinone period.

6.1. Adverse Events

The AE verbatim descriptions (investigator terms from the CRF) will be classified into medical terminology using the MedDRA (Version 23.0 or later).

AE summaries will be presented by treatment for double-blind period, and by treatment sequence and overall for on-vatiquinone period, unless specified otherwise.

Only treatment-emergent AEs (TEAEs) will be included in summary tables. All AEs will be presented in subject data listings.

6.1.1. Treatment Emergent Adverse Event

A TEAE will be considered occurring during the double-blind period if the AE had an onset date or date of worsening on or after the date of first dose of double-blind treatment and:

- Prior to the date of the first dose of open-label treatment if subject enrolled into open label period
- Within 30 days of the date of the last dose of double-blind treatment if the subject does not enroll into open-label period

A TEAE will be considered occurring during the on-vatiquinone period if the AE had an onset date or date of worsening on or after the date of the first dose of vatiquinone treatment and within 30 days of the date of the last dose of vatiquinone treatment.

AEs with missing or partial onset date and cannot be determined if it occurred prior to the first dose of study or which treatment period it occurred will be counted as TEAE under that treatment period.

Overview TEAE table, including number and percentage of subjects with TEAEs, treatment-emergent serious AEs (TESAEs), treatment-related TEAEs/TESAEs, TEAEs by CTCAE grade, and TEAEs leading to discontinuation of study treatment will be provided.

The number and percentage of subjects by treatment group for double-blind period and by treatment sequence for on-vatiquinone period will be tabulated for:

- Incidence of TEAEs by SOC and PT
- Incidence of TEAEs by PT in descending order
- Incidence of treatment-related TEAEs by SOC and PT
- Incidence of TEAEs by SOC, PT and maximum CTCAE grade
- Incidence of TEAEs by SOC, PT, and relationship to study drug

6.1.2. Deaths, serious AE, and AE leading to discontinuation

The following summaries will be presented:

- Incidence of TESAEs by SOC and PT
- Incidence of treatment-related TESAEs by SOC and PT
- TEAEs leading to discontinuation of study treatment by SOC and PT

Subjects who die, who have an SAE, and who have an AE leading to discontinuation of study treatment during the study will be listed in three separate listings.

6.1.3. Exposure adjusted TEAE

The exposure-adjusted TEAEs will be summarized by treatment sequences during on-vatiquinone for Safety analysis set as follows:

- Number of subjects with TEAE and incidence rate per patient-years by SOC and PT
- Number of TEAEs and event rate per patient-years by SOC and PT

For incidence rate, repeated occurrence of an event from the same subject will be counted only once for each subject. For event rate, repeated occurrence of an event from the same subject will be counted per number of occurrences regardless of the subjects. Total patient-years of exposure will be calculated as total duration of vatiquinone treatment in days divided by 365.25 days.

6.2. Clinical Laboratory Parameters

Mean and change from baseline in laboratory parameters will be summarized by treatment groups at each visit, and the last assessment during the period for double-blind period, and by treatment sequence at each visit for on-vatiquinone period. Only values collected at scheduled assessment times will be included in the descriptive summary. In addition, shift tables for each laboratory parameter from baseline to each of the post-baseline visits, and to last assessment will be provided. All laboratory assessments will be used for shift table.

Number of subjects with clinically significant value for selected laboratory parameters as defined in Table 5 at any post-baseline visit will be summarized for double-blind period and for on-vatiquinone period.

Table 5: Criteria for Clinically Significant for Selected Laboratory Parameters

Category	Criteria
Liver function tests	
Elevated Aminotransferases	• ALT >3xULN • AST >3xULN • ALT or AST >3xULN
Elevated Bilirubin	• Total Bilirubin >2xULN • AST or ALT >3xULN and Total Bilirubin >2xULN
Elevated Alkaline Phosphatase	• AP >1.5xULN • AST or ALT >3xULN, Total Bilirubin >2xULN, and AP >1.5xULN
Coagulation	
Elevated INR	• INR ≥ Grade 2 • INR > 1.5xULN
Renal	

Category	Criteria
Elevated Creatinine	• Creatinine > 1.5xULN
Neutropenia	• Absolute blood neutrophil count < 1.5 g/L

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; INR, international normalized ratio; ULN, upper limit of normal.

6.3. Vital Signs

Change from baseline in vital signs (temperature, respiratory rate, systolic blood pressure, diastolic blood pressure, heart rate, pulse, oxygen saturation), weight, and BMI will be summarized by treatment groups at each visit for double-blind period, and by treatment sequence at each visit for on-vatiquinone period.

6.4. Electrocardiogram (ECG) / Echocardiogram

Change from baseline in electrocardiogram (ECG) results (heart rate, PR interval, RR interval, QRS interval, QT interval, QTcB, and QTcF) will be summarized by treatment groups for double-blind period, and by treatment sequence for on-vatiquinone period.

The number and percentage of subjects with investigator findings (normal, abnormal non-clinically significant, abnormal clinically significant) as well as shift from baseline to each visit will be summarized by treatment groups for double-blind period, and by treatment sequence for on-vatiquinone period for ECG. Similarly, the number and percentage of subjects with investigator findings will be summarized for echocardiogram.

6.5. Physical examinations

Physical examination findings will be presented in by-subject listing only for Safety analysis set.

6.6. Columbia-Suicide Severity Rating Scale

Suicidality will be assessed using the Columbia-Suicide Severity Rating Scale (C-SSRS) at screening, baseline, each visit during the placebo-controlled phase and the open-label phase. The C-SSRS assesses an individual's degree of suicidality, including both suicidal ideation and suicidal behavior [Nilsson 2013]. At screening, the "Screening/Baseline" version is to be used to determine eligibility; that is, specifying suicidal ideation in the past 3 months and/or suicidal behavior in the last 12 months to determine exclusion. During all subsequent visits, the "Since last visit" version is used to monitor on-study suicidal ideation and behavior after the initial assessment. Analysis will mainly focus on two aspects: suicidal ideation and suicidal behavior and they will be derived for each subject as follows using the "Since Last visit" version as shown below in Table 6.

Table 6: C-SSRS Summary

C-SSRS Item	Derivation
Suicidal Ideation	A "yes" answer to any one of the following five questions from suicidal ideation section on the C-SSRS: 1. Wish to be dead 2. Non-specific active suicidal thoughts 3. Active suicidal ideation with any methods (not plan) without intent to act

C-SSRS Item	Derivation
	4. Active suicidal ideation with some intent to act, without specific plan 5. Active suicidal ideation with specific plan and intent
Suicidal Behavior	A "yes" answer to any one of the following five questions from suicidal behavior section on the C-SSRS: 1. Preparatory acts or behavior 2. Aborted attempt 3. Interrupted attempt 4. Actual attempt 5. Completed suicide
Suicidal ideation or behavior	A "yes" answer to any one of the above ten suicidal ideation and behavior questions on the C-SSRS
Self-injurious behavior without suicidal intent	A "yes" answer to the following question from suicidal behavior section on the C-SSRS: Has subject engaged in Non-Suicidal Self-Injurious Behavior?

The occurrence of suicidal ideation or suicidal behavior at each visit will be summarized for double-blind period, and on-vatiquinone period by treatment group and sequence, respectively, for the Safety analysis set. A subject will be counted once for each of these categories if at least one question is answered positive belonging to the category.

Self-injurious non-suicidal behavior will also be summarized separately for occurrence in a similar fashion to suicidal ideation and behavior.

In addition, number and percentage of subjects' response for each question will be provided at each visit for each treatment group for double-blind period, and for each treatment sequence for on-vatiquinone period for Safety analysis set.

There will be no imputation for missing answers on questions.

All C-SSRS assessments will be listed only for subjects with at least one event of suicidality (suicidal ideation and/or suicidal behavior).

6.7. Other Analyses

6.7.1. Pharmacokinetics analysis

Vatiquinone concentration will be summarized by visit and timepoint for vatiquinone treated subjects. PK concentration data will also be used in Population PK analysis. Population PK analysis plan will be described in a separate document.

7. MOCK TABLES, LISTINGS, AND GRAPHS

Include a sentence that references the current protocol with the date; for SAP version 1 include Table 1 with only the first row. For example:

The tables, listings, and graphs shells for the study will be provided in a separate document.

8. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study PTC743-NEU-003-FA is based on the protocol Version 5.0, dated 21 May 2021.

SAP Version	Approval Date	Change	Rationale
1.0	20 March 2023	Not Applicable	Original version
2.0	28 March 2023	Appendix 3 (Strategies to Minimize Missing Data) was added.	Based on FDA feedback, the SAP was updated to include the strategies used in the study to minimize missing data.

9. REFERENCES

Little R and Rubin D (2002). Statistical Analysis with Missing Data; A John Wiley & Sons, Inc.

Patel M, Isaacs CJ, Seyer L, Brigatti K, Gelbard S, Strawser C, et al. Progression of Friedreich ataxia: quantitative characterization over 5 years. Ann Clin Transl Neurol. 2016;3(9):684-694.

Nilsson M, Suryawanshi S, Gassmann-Mayer C, Dubrava S, McSorley P, and Jiang K (2013). Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide; https://cssrs.columbia.edu/wp-content/uploads/ScoringandDataAnalysisGuide-for-Clinical-Trials-1.pdf

Christian Rummey, Jennifer M. Farmer, David R. Lynch. Predictors of Loss of Ambulation in Friedreich's Ataxia. EClincialMedicine 18 (2020) 100213

APPENDIX 1. SCHEDULE OF EVENTS

Study Period	Screening[1] (Up to 42 days prior to Baseline)	Screening visit (for post-discontinuation mFARS)	Baseline (Day 1)	Placebo-controlled (All visits/contacts ± 1 week)			Open-label extension (All visits/contacts ± 2 weeks)		Follow-up (approximately 30 days [±5 days] after last dose or early termination visit)
Study Week				**Weeks 4, 8, 16, 20, 28, 32, 40, 44, 52, 56, 64, & 68 Telephone calls**	**Weeks 12, 24, 36, 48, 60 In-person visits**	**Week 72 In-person visit**	**Weeks 78 and 90 Telephone calls**	**Weeks 84 and 96[2] In-person visits**	
Informed consent	X								
Eligibility assessment and Inclusion/ exclusion criteria	X		X						
Chart review-previous medical history[3] and demographics	X								
Genetic confirmation of Friedreich ataxia diagnosis[4]	X								
12-lead ECG	X		X		X	X		X	
Echocardiogram[5]	X				X	X		X	
Height, weight, oxygen saturation	X[6]		X		X	X		X	
Physical exam[7] and vital signs (Temp, HR, RR, BP)	X		X		X	X		X	
Pregnancy test[8]	X		X		X	X		X	X
Fasting hematology including CBC with differential and coagulation panel[9]	X		X		X	X		X	
Fasting serum chemistry[9, 10]	X		X		X	X		X	
HbA1c	X				X	X		X	
Urine drug screen (10 panel)	X								
PK samples[11]			X		X	X		X	

Study Period	**Screening[1] (Up to 42 days prior to Baseline)**	**Screening visit (for post-discontinuation mFARS)**	**Baseline (Day 1)**	**Placebo-controlled (All visits/contacts ± 1 week)**			**Open-label extension (All visits/contacts ± 2 weeks)**		**Follow-up (approximately 30 days [±5 days] after last dose or early termination visit)**
Study Week				**Weeks 4, 8, 16, 20, 28, 32, 40, 44, 52, 56, 64, & 68 Telephone calls**	**Weeks 12, 24, 36, 48, 60 In-person visits**	**Week 72 In-person visit**	**Weeks 78 and 90 Telephone calls**	**Weeks 84 and 96[2] In-person visits**	
AE/SAE assessment	SAEs only		X	X	X	X	X	X	X[12]
Concomitant medications	X		X	X	X	X	X	X	X
Randomization			X[13]						
Modified Friedreich Ataxia Rating Scale	X[14]	X[14]	X[14]		X	X		X	
One-minute walk test	X		X		X	X		X	
Modified Fatigue Impact Scale			X		X	X		X	
FARS-ADL			X		X	X		X	
EQ-5D-5L			X		X	X		X	
Fall log review					X	X		X	
Speech assessment	X		X		X	X		X	
C-SSRS[15]	X		X		X	X		X	
Train subject/caregiver on fall log completion			X[16, 17]						
Fall log entry[16, 17]			X—						
Vatiquinone (200 mg or 400 mg) or placebo dose PO tid with 25% fat meal			X[18] —						
Dispense study drug			X[18]		X	X[19]		X	
Drug reconciliation					X	X		X	

Abbreviations: AE, adverse event; BP, blood pressure; CBC, complete blood count; COVID-19, Coronavirus disease 2019; C-SSRS, Columbia-Suicide Severity Rating Scale; ECG, electrocardiogram; FARS-ADL, Friedreich Ataxia Rating Scale Activities of Daily Living; GAA, guanine-adenine-adenine; GGT, gamma-glutamyl transferase; HR, heart rate; ICF, informed consent form; INR, international normalized ratio; LDH, lactate dehydrogenase; mFARS, modified Friedreich Ataxia Rating Scale; PK, pharmacokinetic; PO, by mouth; RR, respiratory rate; SAE, serious adverse event; Temp, temperature; tid, three times daily; wk, week

Note: Clinical lab samples should be collected with pre-dose PK samples to avoid multiple needle sticks. Dosing may occur prior to laboratory results being available.

[1] Screening will be conducted and completed up to 42 days prior to the Baseline Visit for those requiring a washout. Subjects who are unable to adhere to the screening time window specified in the protocol due to unforeseen circumstances (eg, COVID-19) may be rescreened upon Sponsor approval.
[2] The End of Treatment Visit will take place at Week 96.
[3] Fall history will be collected as part of the medical history. A fall log will be maintained as described in Footnote 17.
[4] If previous genetic testing is available, it does not have to be repeated. The size of GAA repeat is not required for eligibility.
[5] Echocardiogram will take place after completion of all other assessments/procedures on the same day, as applicable, for primary and secondary endpoints. Echocardiogram will be performed at Screening and at the Week 36, Week 72, and Week 96 visits. Unscheduled echocardiograms may be performed at the investigator's discretion; any unscheduled echocardiogram is considered relevant at the investigator's discretion.
[6] Height will only be collected at the Screening Visit. All subsequent visits will collect weight and oxygen saturation.
[7] Physical examinations will include a neurologic assessment. Full physical examination will be performed at Screening, and a targeted examination (including neurologic assessment) may be performed at all subsequent visits.
[8] Female subjects of childbearing potential must have a negative serum pregnancy test at screening; urine pregnancy testing may be performed at all subsequent visits. At-home urine pregnancy test kits will be provided; subjects will be required to perform the test at home ~30 days post last dose of study drug and self-report the results via telephone.
[9] Since subjects may sign the ICF then have screening labs obtained immediately after, subjects may obtain screening hematology and serum chemistry laboratory samples under non-fasting conditions. All subsequent laboratory samples should be obtained under fasting conditions (ie, following an 8-hour fast). Screening laboratory samples will be used for inclusion/exclusion; baseline labs do not need to be resulted prior to the first dose of study drug. Local laboratories may be utilized sparingly for screening laboratory values or for INR follow-up for INR AE.
[10] Serum chemistry will include complete metabolic panel, magnesium, phosphorous, LDH, GGT, and lipid panel.
[11] Pharmacokinetic samples will be collected at each clinic visit starting at the Baseline Visit at 0 hour (immediately prior to the morning dose) and >1 hour following the dose for placebo-controlled visits only. Pre-dose PK samples will be collected at each visit during the open-label extension phase.
[12] The follow-up contact will be by telephone and should be completed approximately 30 days (±5 days) after last study drug administration or termination visit, whichever is later. Data to be followed will be limited to AEs that were deemed study drug-related or any SAE and will be followed until resolution or until they have normalized. At the investigator's discretion, an office visit and additional tests may be conducted in order to ensure proper follow-up of an AE/SAE.
[13] Randomization will occur on the same day as the Baseline Visit.
[14] The Screening and the Baseline mFARS must occur at least 24 hours apart. Subjects may have an additional Screening Visit for mFARS assessment after a minimum of 30-days post-discontinuation of interventional CoQ10, vitamin E, or any approved or non-approved medication for FA. If the subject has already been off of these medications for 30 days at the time of the Screening Visit, then an additional Screening Visit for mFARS will not be required.
[15] C-SSRS will be collected at Screening, Baseline, and during every in-person visit while on study treatment.
[16] During the Baseline Visit, the subject and/or caregiver will be trained on how to capture information in the fall log.
[17] The fall log will capture the date and time of each fall. Falls should be recorded daily by the patient or a live-in caregiver. The daily log should capture how often the patient falls during the day, whether the fall resulted in an injury, any other consequences from the fall (eg, went to the hospital, loss of consciousness), what the person was doing prior to the fall, and who reported the fall (ie, patient, caregiver, or another witness who reported it to the caregiver). Outcomes should assess both frequency of falls and fall severity (eg, falls leading to injury). All falls resulting in injury should be reported as an AE.
[18] First dose for the placebo-controlled phase of the study will be taken in the clinic on the same day as the Baseline Visit after all baseline assessments are completed. Study drug will be dispensed during the Baseline Visit and at each subsequent in-person visit (except for the Week 96 visit). Drug may be delivered via delivery service if needed.
[19] The Week 72 visit marks the end of the placebo-controlled phase; open-label study drug is dispensed at the Week 72 visit. Please remember that subjects are to bring back all used/unused study drug to each in person visit for accountability throughout the study.

APPENDIX 2. EXAMPLE PROGRAM FOR MULTIPLE IMPUTATION (MI)

Intermittent missing mFARS total scores will be imputed using Monte Carlo Markov Chain (MCMC) method with adjustment for covariates (for example, but not limited to, treatment group, age, region, gender) to produce 5 monotone missing pattern datasets first. Missing data after dropout will be imputed assuming missing not at random using pattern mixture complete case missing value pattern. Approximately 1000 complete data sets will be produced.

The results of MMRM analysis from each complete imputed dataset will be combined using Rubin's rule (Little 2002). The detailed steps and sample code for MI are as below.

- For each analysis population or identified population for analysis, perform the MI for the population once. The MI results will be used for all analyses planned for the population.
- 5 datasets will be generated where missing data at intermediate visits will be imputed for each treatment group using non-missing data from all subjects within the treatment group by a Monte Carlo Markov Chain (MCMC) imputation model using the MCMC statement in the SAS PROC MI procedure. As a result, each dataset will only have missing ending data, or a monotone missing data pattern.
- For each dataset from the prior step, missing ending data will be imputed by a regression imputation model using the SAS PROC MI procedure with the REGRESSION option in the MONOTONE statement using the CCMV assuming that the missing values are MNAR. The regression imputation model includes an intercept and the slopes of the measurements from all previous visits for the imputation of subsequent visits. Around 200 datasets will be generated for each dataset from prior step.
- The resulting complete datasets from the prior step will be processed via a SAS DATA step to create change from baseline values for each treatment group at each visit.
- The resulting datasets will be analyzed using SAS PROC MIXED to perform the analysis of covariance for the primary efficacy endpoint by study visit.
- The SAS PROC MIANALYZE procedure will be used to combine the LS mean and treatment versus placebo difference estimates to produce the final analysis results.

Sample SAS code corresponding to the above steps appears below:

```
/*********************************************
Multiple imputation assuming MNAR using complete case missing value pattern

TRT: Treatment
aWeek: Analysis visit (0, 12, 24, 36, ... , 72)
mFARSimp: mFARS total score
*********************************************/

* Step 1: transpose data for MI use;
proc sort data=mFARS0 out=MIstep1; by trt usubjid aweek; run;
```

```
proc transpose data=MIstep1 out=MIstep1 prefix=WK_;
  by trt usubjid;
  id aweek;
  var mFARSimp;
run;

* Step 2: generate 5 datasets to impute intermediate missing to produce
monotone missing pattern datasets;
proc MI data=MIstep1 out=MIstep2 nimpute=5 minmaxiter=5000 seed=93182;
  by trt;
  var wk_0 wk_12 wk_24 wk_36 wk_48 wk_60 wk_72;
  mcmc chain=multiple impute=monotone;
run;

* if imputed value out of the range of mFARS score, set to 0 or 93,
respectively;
data MIstep2;
  set MIstep2;
  if .<wk_0 <0 then wk_0 =0; if wk_0 >93 then wk_0 =93;
  if .<wk_12<0 then wk_12=0; if wk_12>93 then wk_12=93;
  if .<wk_24<0 then wk_24=0; if wk_24>93 then wk_24=93;
  if .<wk_36<0 then wk_36=0; if wk_36>93 then wk_36=93;
  if .<wk_48<0 then wk_48=0; if wk_48>93 then wk_48=93;
  if .<wk_60<0 then wk_60=0; if wk_60>93 then wk_60=93;
  if .<wk_72<0 then wk_72=0; if wk_72>93 then wk_72=93;

  rename _imputation_=_imp01_;
run;

* Step 3: impute missing ending data using MNAR with CCMV;
proc sort data=MIstep2 out=MIstep2; by _imp01_ trt usubjid; run;
proc MI data=MIstep2 out=MIstep3 nimpute=200 minmaxiter=5000 seed=93035;
  by  imp01  trt;
  var wk_0 wk_12 wk_24 wk_36 wk_48 wk_60 wk_72;
  monotone reg(/details);
  MNAR Model (wk_12 wk_24 wk_36 wk_48 wk_60 wk_72/Modelobs=CCMV);
run;

* if imputed value out of the range of mFARS score, set to 0 or 93,
respectively;
data MIstep3;
  set MIstep3;
  if .<wk_0 <0 then wk_0 =0; if wk_0 >93 then wk_0 =93;
  if .<wk_12<0 then wk_12=0; if wk_12>93 then wk_12=93;
  if .<wk_24<0 then wk_24=0; if wk_24>93 then wk_24=93;
  if .<wk_36<0 then wk_36=0; if wk_36>93 then wk_36=93;
  if .<wk_48<0 then wk_48=0; if wk_48>93 then wk_48=93;
  if .<wk_60<0 then wk_60=0; if wk_60>93 then wk_60=93;
  if .<wk_72<0 then wk_72=0; if wk_72>93 then wk_72=93;
run;

* Step 4: Create dateset for MMRM;
data MIstep4;
  length BLmFARSC $16.;
  set MIstep3;
  aweek= 0; avalimp=wk_0 ; cfbimp=avalimp-wk_0; output;
  aweek=12; avalimp=wk_12; cfbimp=avalimp-wk_0; output;
```

```
  aweek=24; avalimp=wk_24; cfbimp=avalimp-wk_0; output;
  aweek=36; avalimp=wk_36; cfbimp=avalimp-wk_0; output;
  aweek=48; avalimp=wk_48; cfbimp=avalimp-wk_0; output;
  aweek=60; avalimp=wk_60; cfbimp=avalimp-wk_0; output;
  aweek=72; avalimp=wk_72; cfbimp=avalimp-wk_0; output;
run;

proc sort data=MIstep4 out=MIstep4; by usubjid _imputation_ aWeek; run;
data MIstep4;
  merge MIstep4(in=in1) dmlm;
  by usubjid;
  if in1;
   imputation_=_imputation_+(_imp01_-1)*200;
run;


* Step 5: Perform MMRM and generate output datasets for MIAnalyze;
proc sort data=MIstep4 out=MIstep4; by _imputation_; run;
proc mixed data=MIstep4 method=reml;
  where 0<aweek<=72;
  by _imputation_;
  class usubjid trt AgeFADc BLmFARSC Region aWeek;
  model cfbimp=trt AgeFADc BLmFARSC Region aWeek wk_0 age trt*aweek;
  repeated aWeek / type=UN subject=Usubjid;
  Lsmeans trt*aweek / pdiff cl;
  ods output lsmeans=LSMCFB
             diffs=DiffCFB;
run;

* Step 6: Use Proc MIAnalyze to obtain final analysis results;
data DiffCFB;
  set DiffCFB;
  if aWeek=36 and _aWeek=36;
  effect="trt";
run;

data LSMCFB;
  set LSMCFB;
  if aWeek=36;
  effect="trt";
run;

* LS Means;
proc mianalyze parms(classvar=full)=lsmcfb;
  class trt;
  modeleffects trt;
  ods output parameterestimates=lsmcfb_out;
run;

* Differences;
proc mianalyze parms(classvar=full)=diffcfb;
  class trt;
  modeleffects trt;
  ods output parameterestimates=diffcfb_out;
run;
```

Table 7: Random Seed Used for MI

Analysis	Seed for the First MI	Seed for the Second MI
mFARS total score, mITT, MMRM (Section 5.1.2)	65952	3576
mFARS total score, ITT, MMRM	90400	92626
mFARS total score, PP, MMRM	16874	53962
mFARS total score, mITT completers, MMRM	79094	1881
mFARS total score, ITT completers, MMRM	92422	18900
FARS-ADL, mITT, MMRM	95661	10286
FARS-ADL, ITT, MMRM	42522	91402
1MWT, mITT, MMRM	5175	96503
1MWT, ITT, MMRM	80859	19235
MFIS, mITT, MMRM	63793	66794
MFIS, ITT, MMRM	31044	52789
Upright Stability subsection of mFARS, mITT, MMRM	61393	20552
Upright Stability subsection of mFARS, ITT, MMRM	77595	1741
Additional MI #1	4378	8071
Additional MI #2	27250	52354
Additional MI #3	32577	79518
Additional MI #4	49003	81897
Additional MI #5	2404	57116
Additional MI #6	58749	53423
Additional MI #7	39836	41307
Additional MI #8	11236	62669
Additional MI #9	9364	15280
Additional MI #10	83274	97489

APPENDIX 3. STRATEGIES TO MINIMIZE MISSING DATA

- Good study planning: Careful planning of the study design and data collection protocol helps to minimize missing data. This includes monitoring for missing data and outliers, as well as ensuring that the study design is appropriate for the pediatric population. In addition, it is important to ensure that the data collection process is clearly documented and consistently applied and the questionnaire is administered in a manner that is age-appropriate. The principal investigator and site staff received training on the administration and scoring of study assessments, including mFARS.
- Engaging parents and caregivers: Engaging parents and caregivers can help to minimize missing data by ensuring that they understand the importance of returning for each visit, completing the questionnaire, and providing support to children during the data collection process.
- Child-friendly questionnaires: Using child-friendly questionnaires that are age-appropriate and easy to understand can help to minimize missing data by ensuring that children are able to complete the questionnaire with minimal assistance.
- Use of electronic data capture: Using electronic data capture can help to minimize missing data by providing real-time data entry, automated checks for data completeness, and reminders for missing data. Queries fire automatically if fields are left blank and are required to be entered.
- Weekly data listing report and subject tracker: These contain data management metrics such as the number of missing forms and open queries and are used to minimize and prevent missing data.
- Edit checks: These are performed on the data in order to ensure data quality, completeness, and consistency in data reporting for analysis. These validation checks can include missing values, consistency between fields, and longitudinal consistency checks.