STATISTICAL ANALYSIS PLAN
(OPEN-LABEL TREATMENT PERIOD)

A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED EFFICACY AND SAFETY STUDY OF ATALUREN IN PATIENTS WITH NONSENSE MUTATION DUCHENNE MUSCULAR DYSTROPHY AND OPEN-LABEL EXTENSION

PTC124-GD-041-DMD

25 SEPTEMBER 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 ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Specialist Term	Explanation
6MWD	6-minute walk distance
6MWT	6-minute walk test
AE	adverse event
ANCOVA	analysis of covariance
ATC	anatomical therapeutic chemical
CI	confidence interval
CTCAE	Common Terminology Criteria for Adverse Events
DMD	Duchenne muscular dystrophy
DMDSAT	DMD Functional Ability Self-Assessment Tool
ECG	electrocardiogram
eCRF	electrical case report form
EQ-5D	EuroQoL 5-Dimension
ET	early termination
FVC	forced vital capacity
HRQL	health-related quality of life
IRT	interactive response technology
ITT	intention-to-treat
LoA	loss of ambulation
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MEP	maximum expiratory pressure
MI	multiple imputation
MIP	maximum inspiratory pressure
mITT	modified intention-to-treat
MMRM	mixed model for repeated measures
MRI	magnetic resonance imaging
NSAA	North Star Ambulatory Assessment
PK	plasma pharmacokinetic
PROM	Patient-Reported Outcome Measure
PT	preferred term
PUL	Performance of Upper Limb
rNSAA	revised North Star Ambulatory Assessment total score
SAP	statistical analysis plan
SOC	system organ class
TEAE	treatment-emergent adverse event
TFT	timed function tests
UN	unstructured covariance matrix
WHODRUG	World Health Organization Drug Dictionary

1. OVERVIEW

This statistical analysis plan (SAP) details the statistical methods to be used in the analyses and presentation of the data collected during the **entire study including open-label treatment period** in Study PTC124-GD-041-DMD, also referred to as Study 041. A separate SAP was prepared to describe the statistical methods to be used for the data collected during the **double-blind treatment period** only.

This document is prepared on the basis of the final study protocol version 4.0 (dated 21JUL2020). The reader is referred to the study protocol, the electronic case report form (eCRF), general eCRF completion guidelines, and various data collection instruments employed in the study for details of study design, conduct and data collection.

There will be 2 database locks for this study. The first database lock (soft lock) occurred on 11April2022 when all subjects completed the 72-week double-blind treatment period. The final database lock will occur when all subjects have completed 72-week open-label treatment period.

2. STUDY OVERVIEW

2.1. Study Design

Study 041 is a Phase 3, international, multicenter, randomized, double-blind, placebo-controlled, efficacy and safety study of ataluren in subjects with nonsense mutation Duchenne muscular dystrophy and its open-label extension.

This study will enroll subjects ≥5 years old. Subjects who fulfill all inclusion/exclusion criteria will be randomized in a 1:1 ratio to placebo or ataluren.

Treatment will comprise continuous daily administration of double-blind ataluren or placebo for 72 weeks, followed by open-label ataluren for another 72 weeks. Throughout the study, study drug should be taken three times per day – the 1st dose (10 mg/kg) in the morning, the 2nd dose (10 mg/kg) at mid-day, and the 3rd dose (20 mg/kg) in the evening.

Visits will be performed every 12 weeks during the double-blind treatment period and every 24 weeks during the open-label extension. Visit windows are ±7 days. The Baseline and Week 72 visits consist of 2 consecutive days. Subjects that discontinue from the study early will complete an early termination (ET) visit. All subjects will have a follow-up phone visit 4 weeks (±7 days) after the last dose of study drug unless the subject transfers to other ataluren long term follow-up studies.

The primary analysis will be targeted on the subset of subjects who are 7 to 16 years old with 6-minute walk distance (6MWD) ≥300 meters and time to stand from supine ≥5 seconds at baseline. In subjects who are 5 and 6 years old, a treatment benefit may be obscured due to the confounding effect of normal growth and development. Therefore, subjects who are 5 and 6 years old are eligible for the study but will not be included in the primary analysis. Subjects who are ≥17 years old and ambulatory at baseline may be considered to have a milder phenotype and a slower trajectory of disease progression. To limit the variability of study results, these subjects will not be included in the primary analysis population.

2.2. Study Objectives

2.2.1. Primary Objective

The primary objective of this study is to determine the effect of ataluren on ambulation and endurance as assessed by the 6-minute walk test (6MWT).

2.2.2. Secondary Objectives

The secondary objectives of this study are to:

- Determine the effects of ataluren on ambulation and burst activity as assessed by timed function tests (TFT)
- Determine the effects of ataluren on lower-limb muscle function as assessed by the North Star Ambulatory Assessment (NSAA)
- Assess the safety profile of ataluren
- Evaluate the correlation between plasma concentration of ataluren and functional outcomes

- Evaluate the plasma pharmacokinetic (PK) profile of ataluren

2.2.3. Exploratory Objectives

The exploratory objectives of this study are to:

- Determine the effects of ataluren on upper-limb muscle function strength as assessed by the Performance of Upper Limb (PUL) and by the DMD Upper Limb PROM (in subjects ≥7 years old at baseline)
- Determine the effects of ataluren on muscle strength as assessed by myometry (in subjects <7 years old at baseline)
- Determine the effects of ataluren on skeletal muscle integrity as assessed by magnetic resonance imaging (MRI) (at pre-qualified sites only)
- Determine the effects of ataluren on subject- and parent/caregiver-reported health-related quality of life (HRQL) as assessed by at-home questionnaire
- Determine the effects of ataluren on pulmonary function as assessed by forced vital capacity (FVC)

2.3. Study Endpoints for Open-Label Treatment Period

The primary and key secondary efficacy endpoints of this study focus on the comparison of ataluren versus placebo in the double-blind treatment period (baseline to Week 72). These endpoints are described in the SAP for the double-blind treatment period.

The efficacy endpoints below focus on the comparison of subjects receiving ataluren from randomization to Week 144 (hereafter will be referred to as early-start ataluren or ataluren/ataluren) versus subjects receiving placebo from randomization to Week 72 and ataluren from Week 72 to Week 144 (hereafter will be referred to as delayed-start ataluren or placebo/ataluren) across the double-blind and open-label treatment periods.

2.3.1. Primary Endpoint

Not applicable for open-label treatment period

2.3.2. Secondary Endpoints

The following endpoints are considered as secondary endpoints for this study:

- Change from baseline to Week 144 in 6MWD
- Composite of average change in times to run/walk 10 meters, climb 4 stairs, and descend 4 stairs at Week 144
- Change from baseline to Week 144 in time to run/walk 10 meters
- Change from baseline to Week 144 in time to climb 4 stairs
- Change from baseline to Week 144 in time to descend 4 stairs
- Change from baseline to Week 144 in NSAA total score
- Time to loss of ambulation over 144 weeks

- Time to loss of stair-climbing over 144 weeks
- Time to loss of stair-descending over 144 weeks
- Risk of loss of NSAA items over 144 weeks
- Ataluren safety profile characterized by type, frequency, severity, and relationship to study drug of any AEs, or of abnormalities of laboratory tests, vital signs, physical examinations, or ECGs
- Plasma PK at pre-morning dose and 2 hours post-morning dose at Week 144

2.3.3. Exploratory Endpoints

The following endpoints are considered as exploratory endpoints:

- Changes from baseline to Week 144 in PUL total score and domain subscores (in subjects ≥7 years old at baseline)
- Change from baseline to Week 144 in DMD Upper Limb PROM total score (in subjects ≥7 years old at baseline)
- Risk of loss of DMD Upper Limb PROM items over 144 weeks (in subjects ≥7 years old at baseline)
- Change from baseline to Week 144 in myometry parameters (in subjects <7 years old at baseline)
- Change from baseline to Week 144 in muscle fat fraction as assessed by MRI (at pre-qualified sites only)
- Changes from baseline to Week 144 in HRQL as assessed by EQ-5D
- Change from baseline to Week 144 in FVC

2.3.4. Sample Size

Not applicable for open-label treatment period.

2.4. Randomization

Not applicable for open-label treatment period.

2.5. Blinding

Not applicable for open-label treatment period.

2.6. Study Assessments

2.6.1. 6-Minute Walk Test

The 6MWT will be performed at Screening (Visit 1), baseline (Visit 2, two 6MWTs performed), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8, two 6MWTs performed), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

2.6.2. Timed Function Tests

Timed function tests of peak physical capacity, including the times taken to run/walk 10 meters, climb 4 stairs, descend 4 stairs, and stand from supine, will be performed using standardized procedures. These assessments will be performed at Screening (Visit 1), baseline (Visit 2), Week 12 (Visit 3), Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

The time taken to stand from supine will not be used as an endpoint in this study.

2.6.3. North Star Ambulatory Assessment

The NSAA measures change in physical function by using standardized procedures and consists of 17 activities, each scored as 0, 1, or 2. The NSAA will be used to evaluate physical function, using standardized procedures. The NSAA will be performed at Screening (Visit 1), baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

2.6.4. Myometry

Myometry will be performed only in subjects who are <7 years old at baseline.

Upper and lower extremity myometry will be performed using a myometer following standardized procedures. Muscle groups to be evaluated include knee extensors and elbow flexors. Bilateral assessments should be done, and 3 measurements should be recorded from each muscle group on each side if possible. These parameters will be monitored at baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

2.6.5. PUL and DMD Upper Limb PROM

Both PUL (version 2.0) and DMD Upper Limb PROM will be assessed only in subjects who are ≥7 years old at baseline.

PUL is assessed using standard equipment and procedures. The DMD Upper Limb PROM questionnaire will be available in all languages relevant for this study and will be completed by the parent/caregiver. If possible, the same parent/caregiver should complete the questionnaire each time. Both assessments will be performed at baseline (Visit 2), Week 12 (Visit 3), Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

2.6.6. Magnetic Resonance Imaging

MRI and/or magnetic resonance spectroscopy (MRS) will be conducted at a subset of sites that have been qualified by a central imaging vendor to perform this assessment. MRI/MRS data will be analyzed centrally. MRI/MRS will be performed at baseline (Visit 2), Week 24 (Visit 4), Week 48 (Visit 6), and Week 72 (Visit 8), and Week 144 (Visit 11 – End of

Treatment/Premature Discontinuation). The baseline imaging may be performed during the screening period, and post-baseline imaging may be performed ±4 weeks of visit date.

2.6.7. Spirometry

Spirometric evaluation of FVC and other parameters (eg, maximal inspiratory and expiratory pressures, peak expiratory flow) in the sitting position will be performed at screening (Visit 1), Baseline (Visit 2), and Week 72 (Visit 8), and Week 144 (Visit 11 – End of Treatment/Premature Discontinuation). If a patient loses ambulation as determined by inability to perform the 10-meter run/walk test within 30 seconds at a clinic visit, then spirometry will be performed at that visit and at all subsequent visits.

2.6.8. HRQL At-Home Questionnaires

HRQL will be measured via the EQ-5D and patient-reported functional ability will be measured using the DMD Self-Assessment Tool (DMDSAT).

These questionnaires will only be completed where validated versions in the local language are available and will be completed by the subject (where possible) and a parent/caregiver.
If possible, the same parent/caregiver should complete the questionnaire each time. The questionnaires will be provided at Visit 1 and completed at home by the subject and a parent/caregiver between Visit 1 and Visit 2 (pre-treatment) and approximately once per month for the rest of the study. Site personnel will assess subject compliance with the questionnaires will be checked at baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

2.6.9. Pharmacokinetic (PK) Assessments

Ataluren concentrations in PK samples will be determined using a validated LC-MS/MS method. PK samples at pre-morning dose and 2 hours post-morning dose at Week 144 will be collected.

2.6.10. Adverse Events

An Adverse Event (AE) is any untoward medical occurrence associated with the use of a drug in humans, whether or not it is considered related to the drug.

All AEs (both serious and nonserious) that occur in subjects during the AE reporting period must be recorded. All AEs are to be recorded in the source documents and on the eCRF using concise medical terminology; whenever possible terms contained in Medical Dictionary for Regulatory Activities (MedDRA) should be employed.

2.6.11. Laboratory Assessment

Hematology, biochemistry, and urinalysis laboratory assessment will be analyzed by the central laboratory. These parameters will be measured at Screening (Visit 1), baseline (Visit 2), Week 24 (Visit 4), Week 48 (Visit 6), and Week 72 (Visit 8), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

2.6.12. Vital Signs, Height, Weight, and Physical Examination

Vital signs (including systolic and diastolic blood pressure, pulse rate, and body temperature), height, and weight will be monitored at Screening (Visit 1), baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8), Week 96 (Visit 9), Week 120 (Visit 10), and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

A full physical examination (including evaluation of cardiovascular system, chest and lungs, thyroid, abdomen, nervous system, skin and mucosae, musculoskeletal system, eyes, ears, nose, mouth, throat, spine, lymph nodes, extremities, and genitourinary) will be conducted at Screening (Visit 1), Week 72 (Visit 8) and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

A symptom-directed physical examination will be conducted at Baseline (Visit 2), Week 24 (Visit 4), and Week 48 (Visit 6).

2.6.13. 12-Lead Electrocardiogram (ECG)

A 12-lead electrocardiogram (ECG) will be obtained at Screening (Visit 1), Week 24 (Visit 4), Week 48 (Visit 6), Week 72 (Visit 8 and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation). The ECG will be performed and interpreted locally. The findings will be captured with the eCRF.

2.6.14. Echocardiogram

Echocardiogram will be obtained at Screening (Visit 1), Week 72 (Visit 8) and Week 144 (Visit 11 - End of Treatment/Premature Discontinuation).

3. STUDY POPULATIONS

The following study populations will be used to summarize open-label assessments. Refer to double-blind treatment period SAP for other study populations applicable to the double-blind treatment data analyses.

3.1. Intention-to-Treat Population

The intention-to-treat (ITT) population defined in the double-blind treatment period SAP will be used for summary and analysis of open-label efficacy endpoints.

3.2. As-Treated Population for On-Ataluren Analysis

The as-treated population for open-label analysis (hereafter referred to as as-treated-OA) consists of all randomized subjects who receive at least 1 dose of ataluren anytime during the study. This population will be used for summary and analysis of open-label safety endpoints.

4. GENERAL CONSIDERATIONS

4.1. Definitions of Treatment Periods

In general, the open-label data analyses will be based on two treatment periods, the overall treatment period and the on-ataluren treatment period.

Overall treatment period: defined as the period from randomization to the end of study, will be used for analyzing and summarizing efficacy endpoints.

On-ataluren treatment period: defined as the period from the first dose of ataluren, regardless of double-blind or open-label ataluren, to the end of the study, will be used for summarizing safety endpoints.

The corresponding baseline for each treatment period is defined in Section 4.3. The corresponding analysis windows are defined in Section 4.9.

4.2. Tables, Figures, and Listings

Summary statistics for continuous variables will include n (number of subjects), mean, standard deviation, standard error, 95% confidence interval (CI) on the mean, median, minimum, and maximum. Summary statistics for categorical variables will include N (number of subjects in population), n (number of subjects in category), % (percent of subjects in category). Where applicable, the summary data (mean, standard error) will also be provided in graphical presentation.

By-subject data listings will be created for each eCRF domain sorted by treatment, subject, and associated dates, where applicable.

4.3. Baseline and Endpoint Definitions

Two baselines will be used for open-label data analysis, study baseline and on-ataluren baseline.

The study baseline, defined as the last assessment prior to randomization, will be used for analyzing and summarizing efficacy endpoints. Study baseline 6MWD is defined as the maximum measurement of valid Day 1 and Day 2 6MWD values.

The on-ataluren baseline, defined as the last available measurement prior to the first dose of ataluren, will be used for summarizing safety endpoints.

4.4. Baseline Characteristics and Treatment Group Comparability

Baseline characteristics and treatment group comparability will not be summarized in the open-label treatment period.

4.5. Interim Analyses

There is no interim analysis 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 completed the double-blind treatment (Week 72). The final database lock will occur when all subjects have completed open-label treatment period.

4.6. Multicenter Studies

Summary statistics will be provided by geographic region (North and Latin America, Europe, and Asia Pacific) and/or country to examine potential region effects for the primary endpoint as needed. No statistical tests will be conducted.

4.7. Multiplicity Control

No multiplicity adjustment will be applied to open-label analyses.

4.8. Missing Data Handling

The analyses using a mixed model repeated measures (MMRM) model will be performed based on available data assuming the missing assessments are missing at random (MAR). As sensitivity analyses, missing assessments for the primary and key secondary endpoints based on the analysis of covariance (ANCOVA) will be imputed using pattern-mixture multiple imputation (MI) as described in Section 6.

4.9. Analysis Window

Two sets of analysis windows will be assigned in open-label analyses depending on the treatment period.

Analysis visit windows during overall treatment period will be derived as in Table 1 based on days from randomization to the corresponding visit date. The analysis window during the overall treatment period will be used for efficacy analysis or summaries as appropriate.

Table 1: Analysis Window for Overall Treatment Period

Analysis Visit	Scheduled Visit Number (study day)	Analysis Window (study day)
Week 12	Visit 3 (85)	[43, 126]
Week 24	Visit 4 (169)	[127, 210]
Week 36	Visit 5 (253)	[211, 294]
Week 48	Visit 6 (337)	[295, 378]
Week 60	Visit 7 (421)	[379, 462]
Week 72	Visit 8 (505)	[463, 546]
Week 96	Visit 9 (673)	[547, 756]
Week 120	Visit 10 (841)	[757, 924]
Week 144	Visit 11 (1009)	Any visit occurred on or after 925 days

Study day is calculated as visit day – randomization date +1.

Analysis window during on-ataluren treatment period will be derived as in Table 2 based on days from first day of ataluren use to the corresponding visit date. The analysis window during the on-ataluren treatment period will be used for summarizing safety assessments as appropriate.

Table 2: Analysis Window for On-Ataluren Treatment Period

Analysis Visit for Early-start Ataluren	Scheduled Visit Number (On-ataluren Day)	Analysis Window (On-ataluren days)	Analysis Visit for Delayed-start Ataluren	Scheduled Visit Number (On-ataluren Day)	Analysis Window (On-ataluren days)
OA Week 12	Visit 3 (85)	[43, 126]			
OA Week 24	Visit 4 (169)	[127, 210]	OA Week 24	Visit 9 (169)	[2, 253]
OA Week 36	Visit 5 (253)	[211, 294]			
OA Week 48	Visit 6 (337)	[295, 378]	OA Week 48	Visit 10 (337)	[254, 421]
OA Week 60	Visit 7 (421)	[379, 462]			
OA Week 72	Visit 8 (505)	[463, 546]	OA Week 72	Visit 11 (505)	Any On Ataluren visit occurred after 422 OA days
OA Week 96	Visit 9 (673)	[547, 756]			
OA Week 120	Visit 10 (841)	[757, 924]			
OA Week 144	Visit 11 (1009)	Any visit occurred on or after 925 days			

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

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

For 6MWD at Visit 8, the date corresponding to the maximum of Day 1 and Day 2 will be used for deriving the study analysis visit.

4.10. Merging Strata

If one of the strata includes subjects from only one of the two treatment arms in ITT population, the stratum cell will be combined with the closest neighbor stratum cell based on the subjects' baseline assessments within corresponding stratum for the corresponding population.

5. SUBJECT DATA

5.1. Subject Disposition and Study Populations

The number of subjects who entered and discontinued early (along with reasons for discontinuation) will be summarized during overall treatment period for ITT population, and during on-ataluren treatment period for as-treated-OL population. In addition, the number of subjects in as-treated-OL population will also be summarized.

5.2. Duration of Treatment with Study Drug

Duration of treatment with study drug during the overall treatment period will be calculated as:

$$Duration\ (days) = Date\ of\ last\ dose\ of\ study\ drug\ -\ Date\ of\ first\ dose\ of\ study\ drug + 1$$

Duration of treatment with study drug in weeks for the overall treatment period will be summarized continuously and categorically (<24 weeks, 24 weeks to <48 weeks, 48 to <72 weeks, 72 to < 96 weeks, 96 to < 120 weeks, 120 to < 144 weeks, ≥144 weeks) based on ITT population.

Duration of treatment with ataluren during the on-ataluren treatment period will be calculated as:

$$Duration\ (days) = Date\ of\ last\ dose\ of\ ataluren\ -\ Date\ of\ first\ dose\ of\ ataluren\ + 1$$

Duration of treatment with ataluren in weeks for the on-ataluren treatment period will be summarized continuously and categorically (<24 weeks, 24 to <48 weeks, 48 to <72 weeks, 72 to <96 weeks, 96 weeks to <120 weeks, 120 to <144 weeks, ≥144 weeks) based on as-treated-OL population.

5.3. Study Drug Compliance

To evaluate study drug compliance, number of subjects with treatment interruption, and the duration of treatment interruptions will be summarized during overall treatment period for ITT population, and during on-ataluren treatment period for as-treated-OL population.

5.4. Demographic and Baseline Characteristics

Not applicable for open-label treatment period.

5.5. Disease Characteristics

Not applicable for open-label treatment period.

5.6. Medical History

Not applicable for open-label treatment period.

5.7. Concomitant Medications and Non-Drug Treatments for the On-Ataluren Treatment Period

Concomitant medications and non-drug treatments will be coded using the World Health Organization Drug Dictionary (WHODD), version September 2021 or higher.

Concomitant medications and non-drug treatments during the on-ataluren treatment period are defined as those taken any time from the first dose date of Ataluren through the last follow up date within study.

The use of concomitant medications and non-drug treatments will be summarized by Anatomical Therapeutic Chemical Level 3 and preferred term (PT) based on ITT, and as-treated-OL population. In addition, the use of corticosteroids and the changes in regimen of corticosteroids will be also summarized during overall treatment period for ITT population, and during on-ataluren treatment period for as-treated-OL population.

6. EFFICACY EVALUATION

The efficacy evaluations will be performed based on the overall treatment period using study baseline. By-visit summaries will be generated by treatment groups based on analysis visits.

6.1. Primary Analysis

Not applicable.

6.2. Secondary Efficacy Variables

6.2.1. Six-Minutes Walking Distance

The average rate of change from baseline in 6MWD over 144-weeks of treatment in the ITT population is considered the key secondary efficacy endpoint of interest for this study in open-label period. The ITT population was specified in the original SAP as a supportive population for the primary endpoint analysis given its importance in providing information on the broad study population (N=359) encompassing the full range of age and ambulatory function. The MMRM category model with a random intercept, baseline 6MWD as a covariate, and factors of baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), baseline time to stand from supine (<5 seconds, ≥5 seconds), treatment, visit (analysis visit as a categorical variable), interaction of visit and treatment, interaction of visit and baseline 6MWD, and unstructured covariance matrix (UN), is considered the primary analysis model of interest. In case of non-convergence, other types of covariance matrix are to be used in the order of heterogenous Toeplitz, Toeplitz, and compound symmetric. For this analysis, any subject who lose ambulation during the study was assigned a 6MWD of 0 meters for the first visit in which the subjects lose ambulation and set to 0 meters for all subsequent visits while they participate in the study. The average rate of change in 6MWD over 144-weeks of treatment is calculated from the estimated change from baseline in 6MWD at Week 144 from the MMRM categorical model.

The example SAS code is listed below:

```
PROC MIXED data=<dataset> method=reml;
    CLASS Cortico TTSTBLC Subjid Trt C6WMD AVISIT;
    MODEL CHG6WMD = Cortico C6WMD TTSTBLC Trt B6WMD AVISIT AVISIT*B6WMD
                    AVISIT*TRT/S CL;
    RANDOM Int / subject=Subjid;
    REPEATED AVISIT / type= un subject=subjid;
    LSMEANS TRT*AVISIT / pdiff  CL;
RUN;
```

where

CHG6WMD = Change from baseline in 6WMD at each visit (Study baseline to Week 144)
B6WMD = Study baseline 6WMD
C6WMD = Study baseline 6WMD category
Cortico = Study baseline concomitant corticosteroid type
TTSTBLC = Time to stand from supine category
AVISIT = Analysis visits as categorical variable
Trt = Treatment group
Subjid = Subject
Int = Intercept

An ANCOVA model will also be used to evaluate the difference in change in 6MWD from baseline to Week 144 after missing data have been imputed via a pattern mixture model utilizing MI based on ITT population. The estimated treatment difference in change from baseline of 6MWD at Week 144 and the corresponding 95% CI will be provided. The ANCOVA model will include baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), treatment, as well as baseline 6MWD as a covariate. The baseline time to stand from supine (<5 seconds, ≥5 seconds) will be included in the testing model in the ITT analysis. For this analysis, all 6MWD will be assigned to an analysis visit according to the analysis visit window described in Section 4.9. Any subjects who lose ambulation during the study will be assigned a 6MWT result of 0 meters for the visit in which the subjects lose ambulation and for all remaining visits while they participate in the study. For subjects with missing 6MWD due to other reasons, the missing assessments at post-baseline visits will be imputed using the pattern-mixture model MI assuming the missing values are missing not at random (MNAR) utilizing the complete case missing value pattern (CCMV - subjects who completed primary efficacy assessments without missing values). Detailed imputation method is described in Appendix 2.

The example SAS code is listed below:

```
PROC MIXED data=<dataset> method=reml;
  CLASS Cortico TTSTBLC Trt C6MWD;
  MODEL CHG6MWD = Cortico TTSTBLC C6MWD Trt B6MWD /CL;
  LSMEANS TRT / pdiff CL;
RUN;
```

where

CHG6MWD = change from baseline in 6MWD at Week 144
B6MWD = Baseline 6MWD
C6MWD = Baseline 6MWD category
Cortico = Baseline concomitant corticosteroid type
TTSTBLC = Time to stand from supine category
Trt = Treatment group

The 6MWD values and the corresponding changes from study baseline values will be summarized by visit during overall treatment period based on ITT population. For these summaries, any subjects who lose ambulation during the study will be assigned a 6MWT result of 0 meters for the visit in which the subjects lose ambulation and for all remaining visits while they participate in the study.

6.2.1.1. Subgroup Analysis

The change from baseline in 6MWD will be summarized by visit for subgroups below for ITT population:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)

In addition, the change from baseline in 6MWD will be summarized at each post-baseline visit for the following subgroups for ITT population:

- Baseline disease severity

- Baseline 6MWD <300 meters
- Baseline 6MWD ≥300 meters to <400 meters
- Baseline 6MWD ≥400 meters

6.2.1.2. Time to 10% Persistent Worsening in 6MWD

Time to 10% persistent worsening in 6MWD, defined as last time that 6MWD was not 10% worse than baseline, 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 ITT populations. For subjects who do not have 10% 6MWD worsening, time to 10% persistent 6MWD worsening will be censored at the time of the last 6MWT during the overall period. For subjects who have all post-baseline assessments more than 10% worse than the baseline, the event will be assigned to Day 1. Subjects who become non-ambulatory will be considered to have 10% worsening at the time of becoming non-ambulatory. The proportion of 10% persistent worsening by Week 144 will also be summarized.

In addition, time to 30 meter drop in 6MWD will also be analyzed in a similar way to time to 10% persistent worsening in 6MWD.

6.2.2. Timed Function Tests

A composite score of the average in times to run/walk 10 meters, climb 4 stairs, and descend 4 stairs will be calculated. This calculated composite score, the individual timed function tests, and their corresponding change from baselines will be analyzed during the overall treatment period for ITT population using the similar MMRM categorical model as for 6MWD. The MMRM model will include factors of concomitant corticosteroid type (deflazacort, prednisone/prednisolone) at study baseline, 6MWD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters) at study baseline, baseline time to stand from supine (<5 seconds, ≥5 seconds), treatment, visit (analysis visit as a categorical variable), interaction of visit and treatment, interaction of visit and corresponding study baseline TFT score, and corresponding study baseline TFT score.

Subjects who cannot perform a timed function test within 30 seconds, including those who is loss of ambulation (LoA) or the timed function test is above 30 seconds, will be assigned a value of 30 seconds for the respective test for calculating the average change in times. If the average time of the 3 TFT tests is 30 seconds for multiple visits for a subject, the 30 seconds will be kept for the first visit with the assessment of 30 seconds and set to missing for all visits afterwards.

ANCOVA analysis after MI used for 6MWD will also be performed for the composite TFT score and the three individual timed function tests during the overall treatment period for ITT population.

The individual TFT results, and the calculated composite score will be summarized by visit during overall treatment period for ITT population.

The change from baseline in composite average of timed function will be summarized by visit for subgroups below for ITT population:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)

- Baseline disease severity
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

By-visit summaries on the overall treatment period for the ITT population will also be provided to evaluate the 10-meter run/walk results in subjects with a study baseline 6WMD <300 meters.

6.2.3. North Star Ambulatory Assessment

The NSAA consists of 17 activities, each scored as 0, 1, or 2. The sum of 16 scores (except for lifts head) will be used to form a total score [Mayhew 2013a]. Hereafter, this NSAA total score will be referred as revised NSAA (rNSAA) total score. If an activity cannot be performed due to disease progression/loss of ambulation, a score of 0 will be assigned, and will not be considered as missing. If fewer than 13 of the 16 activities are performed, the total score will be considered missing. If from 13 to 16 activities are performed, the total score will be standardized by multiplying the sum of the scores in the x activities by 16/x.

The total score will be analyzed using the similar ANCOVA after MI model as 6MWD during the overall treatment period for ITT population. The estimated difference in the change from study baseline at Week 144 will be provided. The linear transformation of the total score [Mayhew 2013a] will be analyzed using the similar ANCOVA after MI model. The total score and linear score will also be summarized by visit during overall treatment period based on ITT population.

In addition, the number of subjects with loss of function of each NSAA items (score of 1 or 2 at baseline transitioning to a 0 score post-baseline) will also be summarized by treatment groups during overall treatment period for ITT population.

The change from baseline in rNSAA total scores and linear scores will be summarized by visit for subgroups below for ITT population:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)
- Baseline disease severity
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

6.2.4. Loss of Ambulation and Stair Climb/Stair Descend

Time to loss of ambulation (LoA), defined as persistent inability to perform the 10-meter run/walk test within 30 seconds at any post-baseline visit and for all remaining visits, will be evaluated using Kaplan-Meier estimation, Log-rank test, and Cox proportional hazards model (assuming similar covariates as the primary efficacy model) during overall treatment period based on ITT population. Similar analyses will be performed for time to loss of 4-stairs climb and time to loss of 4-stairs descend, defined as persistent inability to perform the 4-stairs climb test or 4-stairs descend, respectively, within 30 seconds at any post-baseline visit and for all remaining visits.

A separate analysis will evaluate time to LoA during the overall treatment period in subjects with study baseline 6MWD <300 meters.

6.2.5. Ataluren Plasma Concentration

Plasma concentration at pre-morning dose and 2 hours post-morning dose collected at the last visit (Week 144) will be summarized with descriptive statistics including n (number of subjects), arithmetic mean, standard deviation, standard error, median, minimum, and maximum, %CV, geometric mean, and geometric %CV. In addition, scatter plots and the Pearson's correlation will also be provided based on the ITT population between the ataluren plasma concentration and the following endpoints:

- change from study baseline in 6MWD
- time to 10% worsening from study baseline in 6MWD
- change from study baseline in timed function tests (10m run/walk, 4-stairs climb, 4-stairs descend)

6.3. Exploratory Variables

6.3.1. PUL and DMD Upper Limb PROM

PUL includes 22 items (with an entry item to define starting functional level), and 21 items subdivided into shoulder level (4 items), elbow level (9 items), and distal level (8 items) dimensions. The DMD Upper Limb PROM includes 32 items categorized in 4 domains of daily life: food (7 items), self-care (8 items), household and environment (6 items), and leisure and communication (11 items). The domain sub scores will be calculated by summing the scores from all items within each domain, and the total score will be calculated as the sum of all domain scores. A domain sub score is considered missing if any of the individual item is missing within the respective domain. The total score is considered missing if any of the individual item is missing from a giving assessment. The PUL and DMD Upper Limb PROM total scores will be analyzed using the similar ANCOVA after MI model as 6MWD during the overall treatment period for ITT population.

The PUL total and domain sub scores will also be summarized by age group (7 to < 8 years old, 8 to <13years old, 13 to <22 years old, and >=22 years old) using descriptive statistics during overall treatment period. The last assessment during the specific age group will be used for a given subject in this summary.

The number of subjects with loss of function for each PROM item by treatment group will also be summarized for ITT population during the overall treatment period.

By-visit summaries of total and domain sub scores for PUL and DMD Upper Limb PROM during overall treatment period will also be provided for subjects with the following subgroups based on ITT population:

- Baseline 6MWD <300 meters
- Baseline 6MWD ≥300 meters to <400 meters
- Baseline 6MWD ≥400 meters

By-visit summaries of total and domain sub scores for PUL and DMD Upper Limb PROM during overall treatment period will also be provided for subjects who lose the ability to stand from supine in <30 seconds during the study.

6.3.2. Myometry Parameters

The myometry parameters will be analyzed using the similar ANCOVA after MI model as 6MWD during overall treatment period for ITT population. The change from study baseline in myometer parameters (knee extension and elbow flexion) will be summarized at each visit during overall treatment period for ITT population.

6.3.3. Muscle Fat Fraction

The change from baseline in muscle fat fraction data (proton density fat fraction and MRS fat fraction) obtained from MRI will be summarized at each visit during overall treatment period for ITT population. The muscle fat fraction will also be summarized for subjects with the following subgroups:

- Baseline 6MWD <300 meters
- Baseline 6MWD ≥300 meters to <400 meters
- Baseline 6MWD ≥400 meters

6.3.4. Spirometry – Pulmonary Function Test

The change from baseline in FVC assessments (FVC, FEV1, and PEF) will be summarized in 12- or 24-week intervals from the time of LoA visit during overall treatment period for ITT population. The baseline of FVC assessment is defined as the FVC assessment taken at the visit when subjects become non-ambulatory in this analysis.

Other spirometry parameters including maximum inspiratory (MIP) and expiratory pressures (MEP) will be presented in subject listings only.

6.3.5. HRQL At-Home Questionnaires

HRQL data obtained from the EQ-5D and DMDSAT will be summarized by treatment group at each visit during overall treatment period based on ITT population.

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, some problems, 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 treatment period for ITT population.

The DMDSAT comprised a total of eight questions covering four domains (arm function, mobility, transfers, and ventilation status). The domain scores will be summed up to form the total scores. The domain scores will be tabulated and the mean and mean change in total score will be summarized by treatment groups at each visit during overall treatment period in ITT population.

7. SAFETY EVALUATION

The safety evaluations will be summarized by treatment groups and overall during on-ataluren treatment period based on as-treated-OL population using on-ataluren baseline. By-visit summaries will be provided based on the on-ataluren analysis windows.

7.1. Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Adverse events will be coded to the MedDRA (Version 24.1 or higher) lower level term closest to the verbatim term. The linked MedDRA PT and primary system organ class (SOC) are also captured in the database. The severity of AEs will be graded using the latest version of the Common Terminology Criteria for Adverse Events (CTCAE), Version 5.0 whenever possible.

A treatment-emergent adverse event (TEAE) during on-ataluren treatment period is defined as an adverse event that occurs or worsens while on ataluren (on or after first dose of ataluren) up to 4 weeks after last dose of ataluren.

Only those AEs that are treatment-emergent will be included in summary tables. All AEs, treatment-emergent or otherwise, will be presented in subject data listings. An overview table of TEAEs, including number of subjects with TEAEs, treatment-emergent serious adverse events (SAEs), deaths, TEAEs classified as CTCAE Grade 3 or higher, study drug related TEAEs, TEAEs leading to study drug withdrawal will be provided for each treatment group. The following summaries will be produced for the TEAEs by treatment group:

- 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 treatment.

In the summary tables subjects may be counted under multiple SOCs and PTs, but for each SOC and PT, subjects are only counted once. If a subject has the same AE on multiple occasions, the highest severity (5=fatal, 4=life-threatening, 3=severe, 2=moderate, 1=mild) recorded for the event will be presented and the highest drug relationship (1 = ‘Unrelated’, 2 = ‘Unlikely to be Related’, 3 = ‘Possibly Related’, 4 = ‘Probably Related’=), reclassified into Related (‘Possibly Related’, ‘Probably Related’) or Not Related (‘Unrelated’), will be presented on the respective tables. Percentages are based on the number of subjects in the safety population.

The following summaries will also be presented for the treatment-emergent SAEs (TESAEs):

- Incidence of TESAEs by SOC and PT
- Incidence of TESAEs by PT in descending order
- Incidence of treatment-related TESAEs by SOC and PT.

In addition, the number and percentage of subjects with TEAEs and treatment-related TEAEs leading to discontinuation from study treatment will also be summarized by MedDRA SOC and PT for each treatment group.

SAE, AE leading to discontinuation from study treatment, and death will be listed.

7.1.1. Exposure-Adjusted TEAE

The exposure-adjusted TEAEs will be summarized by treatment group during on-ataluren treatment period for as-treated-OL population as follows:

- Number of subjects with TEAEs 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 occurrences of an event from the same subject will be counted only once for each subject. For event rate, repeated occurrences 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 ataluren treatment in days divided by 365.25 days.

7.1.2. Adverse Events of Special Interest

The following AEs are considered as the adverse events of special interest (AESI), and will be summarized by AESI Category and PT:

Table 3: Adverse Events of Special Interest (AESI)

AESI Category	Definition
Potentiation of aminoglycoside renal toxicity	• SMQ: Acute renal failure (Narrow and Broad), and; • Concomitant medication using the WHODD ATC code J01G
Long-term cardiovascular effects including changes in lipid profile	• SMQ: Dyslipidaemia (Narrow), or; • SOC: Cardiac disorders
Hypertension with use of concomitant systemic corticosteroids	• SMQ: Hypertension (Narrow and Broad), and; • Concomitant medication using the WHODD ATC code H02A
Renal toxicity	• SMQ: Acute renal failure (Narrow and Broad)
Hepatic toxicity	• SOC: Hepatobiliary disorders, or; • HLGT: Hepatobiliary Investigations
Malignancies in general	• SOC: Neoplasms benign, malignant and unspecified (incl cysts and polyps)

For potentiation of aminoglycoside renal toxicity and hypertension with use of concomitant systemic corticosteroids, the concomitant medications are those started prior to the onset of corresponding AE.
Abbreviation: AESI = adverse event of special interest; SMQ = Standardized MedDRA Query; ATC = Anatomical Therapeutic Chemical; SOC = System Organ Class; HLGT = High Level Group Term; WHODD = World Health Organization Drug Dictionary

7.2. Clinical Laboratory

Mean and mean change from baseline in clinical laboratory data from central laboratory (including parameters for assessment of hepatic and renal monitoring) listed in Table 4 will be summarized by analysis visit, as well as the last assessment during the on-ataluren treatment period. In addition, shift tables for each laboratory parameters from on-ataluren baseline to each of the post-baseline visit will be provided.

Table 4: Clinical Laboratory Parameters

Type	Parameters
Hematology	leukocytes, basophils, basophils/leukocytes, eosinophils, eosinophils/leukocytes, lymphocytes, lymphocytes/leukocytes, monocytes, monocytes/leukocytes, neutrophils, neutrophils/leukocytes, hemoglobin, hematocrit, erythrocytes, platelets
Biochemistry	sodium, potassium, bicarbonate, blood urea nitrogen, creatinine, magnesium, calcium, uric acid, glucose, total protein, bilirubin (direct and indirect), aspartate aminotransferase, alanine Aminotransferase, gamma glutamyl Transferase, creatine kinase, alkaline phosphatase, total cholesterol, lactate dehydrogenase (LDH), high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides, and cystatin C
Urinalysis	pH, glucose, ketones, blood, protein

All post-baseline clinical laboratory results will be graded according to CTCAE severity grade (criteria in Appendix 3), when applicable. For parameters graded according to CTCAE severity, the CTCAE grades increase from baseline to Grade 2 or higher grade will be considered as clinically significant (CS) and the incidence will be summarized for as-treated-OL population.

The incidence of all post-baseline specific renal laboratory parameters provides information on action to be taken for study drug will be summarized.

Table 5: Renal Monitoring Parameters and Actions to Be Taken

Laboratory Parameter	Stop Study Drug Immediately, Confirm Abnormal Value, and Then Start Work-Up	Stop Study Drug After Confirming[a] Abnormal Value, and Then Start Work-Up
Serum cystatin C	>2.00 mg/L	>1.33 – 2.00 mg/L
Serum creatinine	≥ Grade 2 (≥1.5 x ULN for age)	Grade 1 (>ULN – 1.5 x ULN for age)
Serum BUN	≥3.0 x ULN	≥1.5 – 3.0 x ULN

a Laboratory abnormalities may be confirmed immediately or at the next scheduled clinic visit based on investigator judgment.
Abbreviations: BUN = blood urea nitrogen, ULN = upper limit of normal

Laboratory abnormality related to elevated liver function test will also be summarized for all post-baseline assessments according to Table 6.

Table 6: Criteria for Elevated Liver Function Test

Category	Criteria
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

Abbreviations: ALT = alanine aminotransferase; AST = aspartate aminotransferase; AP = alkaline phosphatase; ULN = upper limit of normal.

A listing of all CS abnormal laboratory values, renal laboratory abnormal values, and elevated liver function test will be provided.

7.3. Electrocardiogram/Echocardiogram

Overall interpretation of ECG data, including normal, abnormal and clinically significant, abnormal and not clinically significant will be summarized descriptively by the number and percentages of subjects at each analysis visit, and the last assessment during the on-ataluren period.

A listing of all abnormal values for ECG will be provided.

Echocardiogram (left ventricle ejection fraction and sphericity index) will be summarized descriptively, including change from on-ataluren baseline by analysis visit, and the last assessment during the on-ataluren period. Overall interpretation of echocardiogram (normal, abnormal) will be summarized descriptively by the number and percentage of subjects at each visit, and the last assessment during the on-ataluren period.

7.4. Physical Examination

Physical examination results will be listed in subject listings only.

7.5. Vital Signs

Height, weight, and vital signs data will be summarized descriptively, including change from on-ataluren baseline by analysis visit and last assessment during on-ataluren period. A summary will also be provided by analysis visit and overall of the number of subjects meeting criteria for hypertension based on age, gender, and height-adjusted systolic blood pressure and diastolic blood pressure percentile results [Flynn 2017].

- Normal BP: BP <90th percentile for age, sex, and height; or <120/<80 mm Hg for adolescents ≥13 years old;
- Elevated BP: BP reading ≥90th percentile and <95th percentile for age, sex, and height; or 120 to 129/<80 mm Hg for adolescents ≥13 years old;
- Hypertension: BP ≥95th percentile for age, sex, and height; or ≥130/80 mm Hg for adolescents ≥13 years old.

8. CHANGES FROM THE PROTOCOL

Major changes to the planned analyses from the protocol are summarized as follows:

- NSAA total score will be calculated as the sum of the 16 scores (except for lifts head), instead of 17, based on Mayhew's publication [Mayhew 2013a].
- Slope of change in 6MWD over 144 weeks is removed from the list of secondary endpoints in section 2.3.2.
- Modified Intention-to-Treat (mITT) Population is removed from the list of study populations
- MMRM model was modified from a linear regression model to a MMRM model with visit time as a categorical variable to account for the nonlinearity of study data.

9. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study PTC124-GD-041-DMD is based on protocol version 4.0 (dated 21JUL2020).

SAP Version	Approval Date	Description
1.0	04APR2022	Original version
2.0	25SEP2023	Summary of Changes: To be aligned with the "Changes from Planned Analysis" SAP for the double-blind period: 1) mITT population removed throughout this SAP 2) MMRM slope model replaced with MMRM categorical model as primary analysis method 3) slope of change in 6MWD over 144 weeks removed from the list of secondary endpoints 4) baseline disease severity category updated throughout this SAP Other changes: 5) analyses on time to 10% persistent worsening and 30meter drop from baseline in 6MWD were added as sensitivity analyses

10. BIBLIOGRAPHY

Mayhew AG, Cano SJ, Scott E, Eagle M, Bushby K, Manzur A, Muntoni F; North Star Clinical Network for Neuromuscular Disease. Detecting meaningful change using the North Star Ambulatory Assessment in Duchenne muscular dystrophy. Dev Med Child Neurol. 2013a Nov;55(11):1046-52.

Flynn, JT and Falkner, BE. New Clinical Practice Guideline for the Management of High Blood Pressure in Children and Adolescents. Hypertension 2017;70(4):683-686.

11. TABLE OF CONTENTS FOR TABLES, FIGURES, AND LISTINGS

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

APPENDIX 1. SCHEDULE OF ASSESSMENTS

Study Period		**Screening**	**Double-Blind Treatment**									**Notes**
Study Week		**-2 weeks**	**1**		**12**	**24**	**36**	**48**	**60**	**72**		2-week screening period may be extended by 1 week if necessary to schedule pretreatment MRI
Study Day			**1**		**84**	**168**	**252**	**336**	**420**	**504**		
Visit Window (days)					±7	±7	±7	±7	±7	±7		
Visit Number		**1**	**2**		**3**	**4**	**5**	**6**	**7**	**8**		At Visit 2, complete assessments before first dose
Visit Day			**1**	**2**						**1**	**2**	2-day visits only at Visit 2 and Visit 8
Informed consent		X										May be obtained prior to Visit 1 (eg, by phone)
Enter subject in IRT system		X										Access IRT system to enter subject
Clinical and medication history		X										
Hepatitis screen		X										
Urinalysis sample		X	X			X		X		X		
EQ-5D and DMDSAT		X	X		X	X	X	X	X	X		Provided to subject at Visit 1 for at-home completion between visits and reviewed for compliance in clinic
DMD Upper Limb PROM			X		X	X	X	X	X	X		Performed only in subjects ≥7 years at baseline
Height		X	X		X	X	X	X	X	X		
Weight		X	X		X	X	X	X	X	X		
Physical examination		X	X			X		X		X		Full (Visit 1, 8) or symptom-directed (Visit 2, 4, 6)
Concomitant medications		X	X		X	X	X	X	X	X		
Adverse events		X	X		X	X	X	X	X	X		
Genotyping sample			X									
Hematology sample		X	X			X		X		X		
Biochemistry sample		X	X			X		X		X		To be collected in fasted state (except Visit 1)
Vital signs		X	X		X	X	X	X	X	X		
Myometry		X	X		X	X	X	X	X	X		Performed only in subjects <7 years at baseline
PUL		X	X		X	X	X	X	X	X		Performed only in subjects ≥7 years at baseline
6-minute walk test (1st test)		X	X		X	X	X	X	X	X		
6-minute walk test (2nd test)				X							X	
NSAA		X	X		X	X	X	X	X	X		

Study Period		**Screening**	**Double-Blind Treatment**									**Notes**
Study Week		**-2 weeks**	**1**		**12**	**24**	**36**	**48**	**60**	**72**		2-week screening period may be extended by 1 week if necessary to schedule pretreatment MRI
Study Day			**1**		**84**	**168**	**252**	**336**	**420**	**504**		
Visit Window (days)					±7	±7	±7	±7	±7	±7		
Visit Number		**1**	**2**		**3**	**4**	**5**	**6**	**7**	**8**		At Visit 2, complete assessments before first dose
Visit Day			**1**	**2**						**1**	**2**	2-day visits only at Visit 2 and Visit 8
Timed function tests		X	X		X	X	X	X	X	X		
Study drug dispensed				X	X	X	X	X	X		X	Open-label study drug dispensed at Visit 8
Study drug compliance					X	X	X	X	X	X		
12-lead ECG		X				X		X		X		
Spirometry		X	X							X		Required at all visits after subject loses ambulation
Echocardiogram		X								X		
Randomization				X								Access IRT system to randomize subject
Upper leg imaging			X			X		X		X		Performed only at pre-qualified sites. May be done between Visits 1 and 2 (pre-treatment assessment) or ±4 weeks of visit (post-treatment assessments).

Abbreviations: DMD = Duchenne Muscular Dystrophy, DMDSAT = DMD Functional Ability Self-Assessment Tool, ECG = electrocardiogram, EQ-5D = EuroQoL 5-Dimension, IRT = interactive response technology, NSAA = North Star Ambulatory Assessment, PROM = Patient-Reported Outcome Measure, PUL = Performance of Upper Limb

Study Period	**Open-Label Treatment**			**Follow-Up**	**Notes**
Study Week	**96**	**120**	**144**	**148**	
Visit Window (days)	±7	±7	±7	±7	
Visit Number	**9**	**10**	**11**	**12 (by phone)**	Subjects who discontinue prematurely should complete Visit 11 and Visit 12 (by phone)
Urinalysis	X	X	X		
EQ-5D and DMDSAT	X	X	X		Provided to subject at Visit 1 for at-home completion between visits and reviewed for compliance in clinic
DMD Upper Limb PROM	X	X	X		Performed only in subjects ≥7 years old at baseline
Height	X	X	X		
Weight	X	X	X		
Physical examination			X		Full (Visit 11)
Concomitant medications	X	X	X	X	
Adverse events	X	X	X	X	
Hematology sample	X	X	X		
Biochemistry sample	X	X	X		To be collected in fasted state
PK sample			X		Pre-morning dose and 2 hours post-morning dose
Vital signs	X	X	X		
PUL	X	X	X		Performed only in subjects ≥7 years old at baseline
Myometry	X	X	X		Performed only in subjects <7 years old at baseline
6-minute walk test	X	X	X		
NSAA	X	X	X		
Timed function tests	X	X	X		
Study drug dispensed	X	X			Open-label study drug dispensed at Visit 8
Study drug compliance	X	X	X		
12-lead ECG			X		
Spirometry			X		Required at all visits after subject loses ambulation
Echocardiogram			X		
Upper leg imaging			X		Performed only at pre-qualified sites. May be assessed ±4 weeks of visit date

Abbreviations: DMD = Duchenne muscular dystrophy, DMDSAT = DMD Functional Ability Self-Assessment Tool, ECG = electrocardiogram, EQ-5D = EuroQoL 5-Dimension, NSAA = North Star Ambulatory Assessment, PK = pharmacokinetic; PROM = Patient-Reported Outcome Measure, PUL = Performance of Upper Limb

APPENDIX 2. SAMPLE SAS PROGRAM FOR THE PRIMARY ANALYSIS FOR PRIMARY ENDPOINT USING ANCOVA MODEL

The analysis will employ of the following SAS procedures:

- 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.
- 100 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.
- 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 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

TRTn: Treatment (0=Placebo, 1=Ataluren)
Cortico: Steriod use at randomization
C6mwd: Baseline 6MWD categories
Stand: Baseline stand from supine categories
ITTfl: Identifier for ITT population
*********************************************/

* 1st: Generate 100 datasets to impute missing interim data by treatment
group using MCMC option;
proc sort data=smwt; by trtn usubjid; run;
proc MI data=smwt out=smwt_mono minimum=0 nimpute=100 minmaxiter=5000
seed=93182;
  WHERE ITTfl="Y";
  BY trtn;
  VAR SmwdWKBL SmwdWK12 SmwdWK24 SmwdWK36 SmwdWK48 SmwdWK60 SmwdWK72
SmwdWK96 SmwdWK120 SmwdWK144;
  MCMC chain=multiple impute=monotone;
```

```
run;

* 2nd: Impute missing ending data using monotone statement with regression
option and assuming MNAR using CCMV;

proc sort data=smwt_mono out=smwt_mono; by _imputation_ trtn usubjid; run;
proc MI data=smwt_mono out=smwt_ccmv minimum=0 nimpute=1 minmaxiter=5000
seed=93035;
  WHERE ITTfl="Y";
  BY _imputation_ trtn;
  VAR SmwdWKBL SmwdWK12 SmwdWK24 SmwdWK36 SmwdWK48 SmwdWK60 SmwdWK72 SmwdWK96
SmwdWK120 SmwdWK144;
  MONOTONE reg(/details);
  MNAR Model (SmwdWK12 SmwdWK24 SmwdWK36 SmwdWK48 SmwdWK60 SmwdWK72 SmwdWK96
SmwdWK120 SmwdWK144/Modelobs=CCMV);
run;

* 3rd: Create datasets for ANCOVA;
data smwt_reg;
  set smwt_ccmv;
  AWeek= 0; Avalimp=SmwdWK0;  CFBimp=SmwdWK0-BL6mwd;  output;
  AWeek=12; Avalimp=SmwdWK12; CFBimp=SmwdWK12-BL6mwd; output;
  AWeek=24; Avalimp=SmwdWK24; CFBimp=SmwdWK24-BL6mwd; output;
  AWeek=36; Avalimp=SmwdWK36; CFBimp=SmwdWK36-BL6mwd; output;
  AWeek=48; Avalimp=SmwdWK48; CFBimp=SmwdWK48-BL6mwd; output;
  AWeek=60; Avalimp=SmwdWK60; CFBimp=SmwdWK60-BL6mwd; output;
  AWeek=72; Avalimp=SmwdWK72; CFBimp=SmwdWK72-BL6mwd; output;
  AWeek=96; Avalimp=SmwdWK96; CFBimp=SmwdWK96-BL6mwd; output;
  AWeek=120; Avalimp=SmwdWK120; CFBimp=SmwdWK120-BL6mwd; output;
  AWeek=144; Avalimp=SmwdWK144; CFBimp=SmwdWK144-BL6mwd; output;

run;

…

* 4th: Conduct ANCOVA analysis on imputed datasets;
proc sort data=adsmwdimp out=adsmwdimp; by _imputation_; run;
ods exclude all;
proc mixed data=adsmwdimp method=REML;
  WHERE ITTfl="Y" and Aweek=144;
  BY _imputation;
  CLASS Cortico Trtn C6mwd Stand;
  MODEL CFBimp=BL6MWD Trtn Cortico Stand C6MWD / CL;
  LSMEANS Trtn/ diff=control('0') cl alpha=0.05;
  ods output diffs=diff_mi lsmeans=lsm_mi;
run;
ods exclude none;

* 5th: Use Proc MIAnalyze to obtain final analyiss results;
* LS Means;
proc mianalyze parms(classvar=full)=lsm_mi;
  CLASS Trtn;
  MODELEFFECTS Trtn;
  ods output parameterestimates=lsm_out;
run;

* Treatment differences;
```

```
proc mianalyze parms(classvar=full)=diff_mi;
  CLASS Trtn;
  MODELEFFECTS Trtn;
  ods output parameterestimates=diff_out;
run;
```

The table below lists the random seeds to be used for the primary and secondary analyses and some additional seeds should they be required. The random seeds were created using SAS uniform generator RAND("Uniform") with a seed of 20220316.

Endpoint/Analysis	**Seed for Step 1 (impute missing interim data)**	**seed for Step 2 (impute missing ending data)**
6MWD (mITT, Section 6.2.1)	91137	7262
6MWD (ITT, Section 6.2.1)	1457	59445
10 Meter Run/Walk (mITT, Section 6.2.2)	77829	5735
10 Meter Run/Walk (ITT, Section 6.2.2)	47225	22804
4-stair climb (mITT, Section 6.2.2)	89124	91655
4-stair climb (ITT, Section 6.2.2)	65399	28168
4-stair descend (mITT, Section 6.2.2)	10797	25074
4-stair descend (ITT, Section 6.2.2)	65528	98908
NSAA (mITT, Section 6.2.3)	49813	69116
NSAA (ITT, Section 6.2.3)	70669	12277
PUL (mITT, Section 6.3.1)	88445	32933
PUL (ITT, Section 6.3.1)	44851	95936
DMD Upper Limb PROM (mITT, Section 6.3.1)	66837	18935
DMD Upper Limb PROM (ITT, Section 6.3.1)	66322	26365
Myometry (ITT, Section 6.3.2)	35523	63532
If additional analyses need to be performed, seeds below will be used		
Additional Analysis #1	1371	4851
Additional Analysis #2	10065	77968
Additional Analysis #3	51500	83604

APPENDIX 3. LABORATORY RELATED CTCAE SEVERITY GRADE

AE Term	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
White blood cell decreased	<LLN - 3000/mm^3; <LLN - 3.0 x 10^9/L	<3000 - 2000/mm^3; <3.0 - 2.0 x 10^9/L	<2000 - 1000/mm^3; <2.0 - 1.0 x 10^9 /L	<1000/mm^3; <1.0 x 10^9/L	
White blood cell increased (leukocytosis)	-	-	>100,000 mm^3	Clinical manifestations of leucostasis; urgent intervention indicated	Death
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hgb <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hgb <8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated	Death
Lymphocyte count decreased	<LLN - 800/mm^3; <LLN - 0.8 x10^9 /L	<800 - 500/mm^3; <0.8 - 0.5 x 10^9 /L	<500 - 200/mm^3; <0.5 - 0.2 x 10^9 /L	<200/mm^3; <0.2 x 10^9 /L	
Lymphocyte count increased	-	>4000 - 20,000 mm^3	>20,000 mm^3	-	
Platelet count decreased	<LLN - 75,000/mm^3; <LLN - 75.0 x 10^9/L	<75,000 - 50,000/mm^3; <75.0 - 50.0 x 10^9/L	<50,000 - 25,000/mm^3; <50.0 - 25.0 x 10^9/L	<25,000/mm^3; <25.0 x 10^9/L	
Serum amylase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; >2.0 - 5.0 x ULN and asymptomatic	>2.0 - 5.0 x ULN with signs or symptoms; >5.0 x ULN and asymptomatic	>5.0 x ULN and with signs or symptoms	
Alkaline phosphatase increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	
Alanine aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	
Aspartate aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	

AE Term	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Blood bilirubin increased	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal	
GGT increased	>ULN – 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	
Hemoglobin increased	Increase in >0- 2 gm/dl above ULN	Increase in >2- 4 gm/dl above ULN	Increase in >4 gm/dl above ULN	-	
Lipase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; >2.0 - 5.0 x ULN and asymptomatic	>2.0 - 5.0 x ULN with signs or symptoms; >5.0 x ULN and asymptomatic	>5.0 x ULN and with signs or symptoms	
Hypercalcemia	Corrected serum calcium of >ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L; Ionized calcium >ULN - 1.5 mmol/L	Corrected serum calcium of >11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L; Ionized calcium >1.5 - 1.6 mmol/L; symptomatic	Corrected serum calcium of >12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L; hospitalization indicated	Corrected serum calcium of >13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L; life-threatening consequences	Death
Hypocalcemia	Corrected serum calcium of <LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L; Ionized calcium <LLN - 1.0 mmol/L	Corrected serum calcium of <8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L; Ionized calcium <1.0 - 0.9 mmol/L; symptomatic	Corrected serum calcium of <7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L; Ionized calcium <0.9 - 0.8 mmol/L; hospitalization indicated	Corrected serum calcium of <6.0 mg/dL; <1.5 mmol/L; Ionized calcium <0.8 mmol/L; life-threatening consequences	Death
Creatinine increased	>ULN - 1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 x baseline; >3.0 - 6.0 x ULN	>6.0 x ULN	
Hyperglycemia	Abnormal glucose above baseline with no medical intervention	Change in daily management from baseline for a diabetic; oral antiglycemic agent initiated; workup for diabetes	Insulin therapy initiated; hospitalization indicated	Life-threatening consequences; urgent intervention indicated	Death

AE Term	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Hypoglycemia	<LLN - 55 mg/dL; <LLN - 3.0 mmol/L	<55 - 40 mg/dL; <3.0 - 2.2 mmol/L	<40 - 30 mg/dL; <2.2 - 1.7 mmol/L	<30 mg/dL; <1.7 mmol/L; life-threatening consequences; seizures	Death
Hypophosphatemia	~~<~~Laboratory finding only and intervention not indicated	Oral replacement therapy indicated	Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of existing hospitalization indicated	life-threatening consequences	Death
Hyperkalemia	>ULN - 5.5 mmol/L	>5.5 - 6.0 mmol/L; intervention initiated	>6.0 - 7.0 mmol/L; hospitalization indicated	>7.0 mmol/L; life-threatening consequences	Death
Hypokalemia	<LLN - 3.0 mmol/L	<LLN - 3.0 mmol/L; symptomatic; intervention indicated	<3.0 - 2.5 mmol/L; hospitalization indicated	<2.5 mmol/L; life-threatening consequences	Death
Hyponatremia	<LLN - 130 mmol/L	125-129 mmol/L and asymptomatic	125-129 mmol/L symptomatic; 120-124 mmol/L regardless of symptoms	<120 mmol/L; life-threatening consequences	Death
Hypernatremia	>ULN - 150 mmol/L	>150 - 155 mmol/L; intervention initiated	>155 - 160 mmol/L; hospitalization indicated	>160 mmol/L; life-threatening consequences	Death
Hyperuricemia	>ULN without physiologic consequences	-	>ULN with physiologic consequences	life-threatening consequences	Death

Signature Page for Global PTC124-GD-041-DMD Statistical Analysis Plan (SAP) - V2

Statistics Approval	[redacted] I approve the document(s) 25-Sep-2023 14:03:31 GMT+0000

Statistics Approval	[redacted] I approve the document(s) 25-Sep-2023 14:03:40 GMT+0000

Clinical Approval	[redacted] I approve the document(s) 25-Sep-2023 14:04:46 GMT+0000

Signature Page for VV-CLIN-005685 v3.0

STATISTICAL ANALYSIS PLAN
(DOUBLE-BLIND TREATMENT PERIOD)
for EMA/MHRA

A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED EFFICACY AND SAFETY STUDY OF ATALUREN IN PATIENTS WITH NONSENSE MUTATION DUCHENNE MUSCULAR DYSTROPHY AND OPEN-LABEL EXTENSION

PTC124-GD-041-DMD

04 APRIL 2022
VERSION 1.0

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 ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation or Specialist Term	Explanation
6MWD	6-minute walk distance
6MWT	6-minute walk test
AE	adverse event
AESI	adverse event of special interest
ANCOVA	analysis of covariance
ATC	anatomical therapeutic chemical
BMI	body mass index
CCMV	completed case missing value pattern
CI	confidence interval
CS	clinically significant
CTCAE	Common Terminology Criteria for Adverse Events
COVID19	Coronavirus Disease of 2019
DMD	Duchenne muscular dystrophy
DMDSAT	DMD Functional Ability Self-Assessment Tool
ECG	electrocardiogram
eCRF	electronic case report form
EMA	European Medicines Agency
EQ-5D	EuroQoL 5-Dimension
EQ-VAS	EuroQoL visual analogue scale
ET	early termination
FVC	forced vital capacity
HRQL	health-related quality of life
IRT	interactive response technology
ITT	intention-to-treat
LoA	loss of ambulation
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MEP	maximum expiratory pressure
MHRA	Medicines and Healthcare products Regulatory Agency
MI	multiple imputation
MIP	maximum inspiratory pressure
mITT	modified intention-to-treat
MMRM	mixed model for repeated measures
MNAR	missing not at random
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
NSAA	North Star Ambulatory Assessment
PK	pharmacokinetic
PP	per-protocol
PT	preferred term
PROM	Patient-Reported Outcome Measure
PUL	Performance of Upper Limb
rNSAA	revised North Star Ambulatory Assessment total score
SAE	serious adverse event
SAP	statistical analysis plan
SOC	system organ class
TEAE	treatment-emergent adverse event
TESAE	treatment-emergent serious adverse event
UN	unstructured covariance matrix

Abbreviation or Specialist Term	Explanation
WHODD	World Health Organization Drug Dictionary

1. OVERVIEW

This statistical analysis plan (SAP) details the statistical methods to be used in the analyses and presentation of the data collected during the **double-blind treatment period** in Study PTC124-GD-041-DMD, also referred to as Study 041, for **EMA/MHRA**. The statistical methods to be used for the data collected during the entire study including open-label treatment period will be described in a separate SAP.

This document is prepared on the basis of the final study protocol version 4.0 (dated 21JUL2020). The reader is referred to the study protocol, the electronic case report form (eCRF), general eCRF completion guidelines, and various data collection instruments employed in the study for details of study design, conduct and data collection.

This SAP is to be reviewed and approved prior to the first study database lock. There will be 2 database locks for this study. The first database lock (soft lock) will occur when all subjects completed the 72-week double-blind treatment. The final database lock will occur when all subjects have completed 72-week open-label treatment period.

2. STUDY OVERVIEW

2.1. Study Design

Study 041 is a Phase 3, international, multicenter, randomized, double-blind, parallel, placebo-controlled, efficacy and safety study of ataluren in subjects with nonsense mutation Duchenne muscular dystrophy (DMD) and its open-label extension.

This study will enroll subjects ≥5 years old. Subjects who fulfill all the inclusion/exclusion criteria will be randomized in a 1:1 ratio to placebo or ataluren.

Treatment will comprise continuous daily administration of double-blind ataluren or placebo for 72 weeks, followed by open-label ataluren for another 72 weeks. Throughout the study, study drug should be taken three times per day - the 1st dose (10 mg/kg) in the morning, the 2nd dose (10 mg/kg) at mid-day, and the 3rd dose (20 mg/kg) in the evening.

Visits will be performed every 12 weeks during the double-blind treatment period and every 24 weeks during the open-label extension. Visit windows are ±7 days. The Baseline and Week 72 visits consist of 2 consecutive days. Subjects that discontinue from the study early will complete an early termination (ET) visit. All subjects will have a follow-up phone visit 4 weeks (±7 days) after the last dose of study drug.

The primary analysis will be targeted on the subset of subjects who are 7 to 16 years old with 6-minute walk distance (6MWD) ≥300 meters and time to stand from supine ≥5 seconds at baseline. In subjects who are 5 and 6 years old, a treatment benefit may be obscured due to the confounding effect of normal growth and development. Therefore, subjects who are 5 and 6 years old are eligible for the study but will not be included in the primary analysis. Subjects who are ≥17 years old and ambulatory at baseline may be considered to have a milder phenotype and a slower trajectory of disease progression. To limit the variability of study results, these subjects will not be included in the primary analysis population.

2.2. Study Objectives

2.2.1. Primary Objective

The primary objective of this study is to determine the effect of ataluren on ambulation and endurance as assessed by the 6-minute walk test (6MWT).

2.2.2. Secondary Objectives

The secondary objectives of this study are to:

- Determine the effects of ataluren on ambulation and burst activity as assessed by timed function tests
- Determine the effects of ataluren on lower-limb muscle function as assessed by the North Star Ambulatory Assessment (NSAA)
- Assess the safety profile of ataluren
- Evaluate the correlation between plasma concentration of ataluren and functional outcomes

- Evaluate the plasma pharmacokinetic (PK) profile of ataluren

2.2.3. Exploratory Objectives

The exploratory objectives of this study are to:

- Determine the effects of ataluren on upper-limb muscle function strength as assessed by the Performance of Upper Limb (PUL) and by the DMD Upper Limb Patient-Reported Outcome Measure (PROM) (in subjects ≥7 years old at baseline)
- Determine the effects of ataluren on muscle strength as assessed by myometry (in subjects <7 years old at baseline)
- Determine the effects of ataluren on skeletal muscle integrity as assessed by magnetic resonance imaging (MRI) (at pre-qualified sites only)
- Determine the effects of ataluren on subject- and parent/caregiver-reported health -related quality of life (HRQL) as assessed by at-home questionnaire
- Determine the effects of ataluren on pulmonary function as assessed by forced vital capacity (FVC)

2.3. Study Endpoints for Double-Blind Treatment Period

2.3.1. Primary Endpoint

The primary endpoint is slope of change in 6MWD over 72 weeks.

2.3.2. Secondary Endpoints

The key secondary endpoints are:

- Slope of composite of average change in times to run/walk 10 meters, climb 4 stairs, and descend 4 stairs over 72 weeks
- Change from baseline to Week 72 in NSAA total score

Other secondary endpoints are:

- Change from baseline to Week 72 in 6MWD
- Composite of average change in times to run/walk 10 meters, climb 4 stairs, and descend 4 stairs at Week 72
- Change from baseline to Week 72 in time to run/walk 10 meters
- Change from baseline to Week 72 in time to climb 4 stairs
- Change from baseline to Week 72 in time to descend 4 stairs
- Time to loss of ambulation (LoA) over 72 weeks
- Time to loss of stair-climbing over 72 weeks
- Time to loss of stair-descending over 72 weeks
- Risk of loss of NSAA items over 72 weeks

- Ataluren safety profile characterized by type, frequency, severity, and relationship to study drug of any adverse events (AEs), or of abnormalities of laboratory tests, vital signs, physical examinations, or electrocardiograms (ECGs)

2.3.3. Exploratory Endpoints

Exploratory endpoints include:

- Changes from baseline to Week 72 in PUL total score and domain subscores (in subjects ≥7 years old at baseline)
- Change from baseline to Week 72 in DMD Upper Limb PROM total score and domain subscores (in subjects ≥7 years old at baseline)
- Risk of loss of DMD Upper Limb PROM items over 72 weeks (in subjects ≥7 years old at baseline)
- Change from baseline to Week 72 in myometry parameters (in subjects <7 years old at baseline)
- Change from baseline to Week 72 in muscle fat fraction as assessed by MRI (at pre-qualified sites only)
- Changes from baseline to Week 72 in HRQL as assessed by EuroQoL 5-Dimension (EQ-5D)
- Change from baseline to Week 72 in FVC

2.4. Sample Size

The hypothesis of this study is that the slope of change in 6MWD from baseline to Week 72 (end of double-blind treatment period) will be 1.0 meter/week better in the ataluren arm than in the placebo arm, within a subset of subjects who are ≥7 to 16 years old with baseline 6MWD ≥300 meters and baseline time to stand from supine ≥5 seconds. The baseline 6MWD is defined as the maximum of the 6MWDs from a valid 6MWT on Day 1 and Day 2. Similarly, the Week 72 6MWD is defined as the maximum of the 6MWDs from a valid 6MWT on Day 1 and Day 2 of the Week 72 visit (end of double-blind treatment period). With a 1:1 randomization, 154 subjects will be required (77 subjects in the ataluren treatment arm and 77 subjects in the placebo arm) to detect a difference of 1.0 meter per week between ataluren and placebo with 85% power (two-sided $\alpha = 0.05$), assuming a common standard deviation of 2.058 meters per week. Further assuming a premature discontinuation rate of ~5%, a total of 162 subjects (81 subjects in the ataluren treatment arm and 81 subjects in the placebo arm) will be enrolled for the primary endpoint analysis. The study is expected to enroll ~340 subjects overall, including those outside of the primary analysis population.

2.5. Randomization

Randomization is an accepted means to reduce bias and allows for the highest standard of evidence in documenting a treatment effect. The stratified, block randomization approach will be used. The system will assign the next available randomization number, within the list for the stratum, to the patient. Such a method allows treatment arms to be balanced with respect to the predefined stratification factors as well as for the number of subjects in each arm. The process will be established and performed centrally through an interactive response technology (IRT) system to maximize the integrity and security of the randomization and ensure appropriate access and convenience-of-use by the investigational sites.

In order to balance treatment allocation, the following stratification factors will be used:

- Baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone)
- Baseline 6MWD (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters) where baseline 6MWD is the maximum of the Day 1 and Day 2 6MWDs
- Baseline time to stand from supine (<5 seconds, ≥5 seconds)

2.6. Blinding

Double-blind administration of study treatments during double-blind period is utilized to provide additional and substantial protection against motivational bias on the part of subjects, parents/caregivers, clinic staff, and other study personnel, with respect to assessment of efficacy and safety parameters.

2.7. Study Assessments

2.7.1. 6-Minute Walk Test

The 6MWT will be performed at Screening (Visit 1), baseline (Visit 2, two 6MWTs performed), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8, two 6MWTs performed) in the double-blind treatment period.

2.7.2. Timed Function Tests

Timed function tests of peak physical capacity, including the times taken to run/walk 10 meters, climb 4 stairs, descend 4 stairs, and stand from supine, will be performed using standardized procedures. These assessments will be performed at Screening (Visit 1), baseline (Visit 2), Week 12 (Visit 3), Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8) in the double-blind treatment period.

The time taken to stand from supine will not be used as an endpoint in this study.

2.7.3. North Star Ambulatory Assessment

The NSAA measures change in physical function by using standardized procedures and consists of 17 activities. The NSAA will be used to evaluate physical function, using standardized procedures. The NSAA will be performed at Screening (Visit 1), baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8) in the double-blind treatment period.

2.7.4. Myometry

Myometry will be performed only in subjects who are <7 years old at baseline.

Upper and lower extremity myometry will be performed using a myometer following standardized procedures. Muscle groups to be evaluated include knee extensors and elbow flexors. Bilateral assessments should be done, and 3 measurements should be recorded from each muscle group on each side if possible. These parameters will be monitored at baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8) in the double-blind treatment period.

2.7.5. PUL and DMD Upper Limb PROM

PUL (version 2.0) and DMD Upper Limb PROM will be assessed only in subjects who are ≥7 years old at baseline.

PUL is assessed using standard equipment and procedures at baseline (Visit 2), Week 12 (Visit 3), Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8) in the double-blind treatment period.

The DMD Upper Limb PROM questionnaire will be available in all languages relevant for this study and will be completed by the parent/caregiver. If possible, the same parent/caregiver should complete the questionnaire each time. The DMD Upper Limb PROM questionnaire will be completed at baseline (Visit 2), Week 12 (Visit 3), Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8) in the double-blind treatment period.

2.7.6. Magnetic Resonance Imaging (MRI)

MRI and/or magnetic resonance spectroscopy (MRS) will be conducted at a subset of sites that have been qualified by a central imaging vendor to perform this assessment. MRI/MRS data will be analyzed centrally. MRI/MRS will be performed at baseline (Visit 2), Week 24 (Visit 4), Week 48 (Visit 6), and Week 72 (Visit 8) in the double-blind treatment period. The baseline imaging may be performed during the screening period, and post-baseline imaging may be performed ±4 weeks of visit date.

2.7.7. Spirometry

Spirometric evaluation of FVC and other parameters (eg, maximal inspiratory and expiratory pressures, peak expiratory flow) in the sitting position will be performed at screening (Visit 1), Baseline (Visit 2), and Week 72 (Visit 8) in the double-blind treatment period. If a patient loses ambulation as determined by inability to perform the 10-meter run/walk test within 30 seconds at a clinic visit, then spirometry will be performed at that visit and at all subsequent visits.

2.7.8. HRQL At-Home Questionnaires

HRQL will be measured via the EQ-5D and patient-reported functional ability will be measured using the DMD Self-Assessment Tool (DMDSAT).

These questionnaires will only be completed where validated versions in the local language are available and will be completed by the subject (where possible) and a parent/caregiver. If possible, the same parent/caregiver should complete the questionnaire each time. The questionnaires will be provided at Visit 1 and completed at home by the subject and a parent/caregiver between Visit 1 and Visit 2 (pre-treatment) and approximately once per month for the rest of the study. Site personnel will assess subject compliance with the questionnaires will be checked at baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8) in the double-blind treatment period.

2.7.9. Pharmacokinetic (PK)

No PK sample will be collected during double-blind period.

2.7.10. Adverse Events

An AE is any untoward medical occurrence associated with the use of a drug in humans, whether or not it is considered related to the study drug.

All AEs (both serious and nonserious) that occur in subjects during the AE reporting period must be recorded. All AEs are to be recorded in the source documents and on the eCRF using concise medical terminology; whenever possible terms contained in Medical Dictionary for Regulatory Activities (MedDRA) should be employed.

2.7.11. Laboratory Assessment

Hematology, biochemistry, and urinalysis laboratory assessment will be analyzed by the central laboratory. These parameters will be measured at Screening (Visit 1), baseline (Visit 2), Week 24 (Visit 4), Week 48 (Visit 6), and Week 72 (Visit 8) in the double-blind treatment period.

2.7.12. Vital Signs, Height, Weight, and Physical Examination

Vital signs (including systolic and diastolic blood pressure, pulse rate, and body temperature), height, and weight will be monitored at Screening (Visit 1), baseline (Visit 2), Week 12 (Visit 3) Week 24 (Visit 4), Week 36 (Visit 5), Week 48 (Visit 6), Week 60 (Visit 7), and Week 72 (Visit 8) in the double-blind treatment period.

A full physical examination (including evaluation of cardiovascular system, chest and lungs, thyroid, abdomen, nervous system, skin and mucosae, musculoskeletal system, eyes, ears, nose, mouth, throat, spine, lymph nodes, extremities, and genitourinary) will be conducted at Screening (Visit 1), and Week 72 (Visit 8) in the double-blind treatment period.

A symptom-directed physical examination will be conducted at Baseline (Visit 2), Week 24 (Visit 4), and Week 48 (Visit 6) in the double-blind treatment period.

2.7.13. 12-Lead Electrocardiogram (ECG)

A 12-lead ECG will be obtained at Screening (Visit 1), Week 24 (Visit 4), Week 48 (Visit 6), and Week 72 (Visit 8) in double-blind treatment period. The ECG will be performed and interpreted locally. The findings will be captured on the eCRF.

2.7.14. Echocardiogram

Echocardiogram will be obtained at Screening (Visit 1) and Week 72 (Visit 8) in the double-blind treatment period.

3. STUDY POPULATIONS

3.1. Intention-to-Treat Population

The intention-to-treat (ITT) population will include all subjects who are randomized, with treatment assignments designated according to initial randomization, regardless of whether subjects receive a different study treatment from the one randomized. In addition, subjects in this population must have a valid 6MWT at baseline, and at least one valid post-baseline 6MWT. This population will be used for summary and analysis of efficacy endpoints.

3.2. Modified Intention-to-Treat Population

The modified intention-to-treat (mITT) population will include all subjects in the ITT population who meet the following additional criteria: 7 to 16 years old with 6MWD ≥300 meters and time to stand from supine ≥5 seconds at baseline. This population will be used for the primary and secondary efficacy endpoint analyses, and other supportive efficacy analyses.

3.3. As-Treated Population

The as-treated population consists of all randomized subjects who receive study treatment, with treatment assignments designated according to actual treatment received. This population will be used for summary and analysis of safety endpoints.

3.4. Per-Protocol Population

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

- Received study treatment different from the randomized treatment throughout the double-blind period
- Did not have a valid 6MWT at baseline or a valid Visit 8 6MWT within the Study Analysis Visit window for Week 72
- Non-compliance to study drug administration
- Had significant inclusion or exclusion criteria violations
- Had protocol deviations which may impact effectiveness of study treatment

This population will be used for primary efficacy endpoint as supportive efficacy analysis.

A separate document including detailed PP population exclusion criteria and the list of subjects excluded from PP population will be finalized prior to treatment unblinding.

4. GENERAL CONSIDERATIONS

4.1. Tables, Figures, and Listings

Summaries will be presented by treatment (placebo, ataluren) and in total (where applicable). Summary statistics for continuous variables will include n (number of subjects), mean, standard deviation, median, minimum, and maximum. Summary statistics for categorical variables will include N (number of subjects in population), n (number of subjects in category), and % (percent of subjects in category). Where applicable, the summary data (eg, mean, standard error) will also be provided in graphical presentation. Unless otherwise specified, all inferential analyses will be 2-sided at the alpha (α) of 0.05 level of significance.

By-subject data listings will be created for each eCRF domain sorted by treatment, subject, and associated dates, where applicable.

4.2. Baseline and Endpoint Definitions

Baseline is defined as the last available measurement prior to the first dose of double-blind study drug unless specified otherwise.

Baseline 6MWD is defined as the maximum measurement of valid Day 1 and Day 2 6MWD values.

Similarly, the Week 72 6MWD will be defined as the maximum of the 6MWDs from a valid 6MWT on Day 1 and Day 2 at Week 72 visit.

4.3. Baseline Characteristics and Treatment Group Comparability

Baseline subject characteristics for each study population, will be summarized using frequency tables for categorical variables and descriptive statistics for quantitative variables. The baseline measurements will be included in the analysis model as a covariate for the primary and secondary efficacy endpoints. Details are specified in Section 6.

4.4. Interim Analyses

There are no planned interim analyses.

4.5. Multicenter Study

Summary statistics will be provided by geographic region (North and Latin America, Europe, and Asia Pacific) and/or country and displayed graphically to examine potential region effects for the primary endpoint as needed. No statistical tests in this regard will be conducted.

4.6. Multiplicity Control

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 in the table below.

Table 1: Test Sequences

Primary Endpoint
Slope of change in 6-minute walk distance (6MWD) over 72 weeks (mITT)
Key Secondary Endpoints
(1) Slope of composite of average change in times to run/walk 10 meters, climb 4 stairs, and descend 4 stairs over 72 weeks (mITT)
(2) Change from baseline to Week 72 in North Star Ambulatory (NSAA) total score (mITT)

Abbreviations: 6MWD, 6-minute walk distance; mITT, modified intention-to-treat; NSAA, North Star Ambulatory Assessment

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

4.7. Missing Data Handling

The 12-week interval of patient visits permits inclusion of 6 post-baseline evaluations of efficacy during the randomized study period (Weeks 12, 24, 36, 48, 60, and 72), ensuring that time trends can be adequately assessed, and that sporadically missing data do not have a substantial impact on data analysis.

The analyses using a mixed model repeated measures (MMRM) model will be performed based on available data assuming the missing assessments are missing at random (MAR). As sensitivity analyses, missing assessments for the primary and key secondary endpoints based on the analysis of covariance (ANCOVA) will be imputed using pattern-mixture multiple imputation (MI) as described in Section 6. Detailed imputation method is described in Appendix 2.

Missing data will not be imputed for summaries for all safety endpoints and for by-visit summaries for efficacy endpoints.

4.8. Study Analysis Visit

Study analysis visits during double-blind treatment period will be derived as listed in Table 2 based on 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.

Table 2: Study Analysis Visit

Analysis Visit	Scheduled Visit Number (Study Day)	Analysis Window (Study Day)
DB Week 12	Visit 3 (85)	[43, 126]
DB Week 24	Visit 4 (169)	[127, 210]
DB Week 36	Visit 5 (253)	[211, 294]
DB Week 48	Visit 6 (337)	[295, 378]
DB Week 60	Visit 7 (421)	[379, 462]
DB Week 72	Visit 8 (505)	Any double-blind visit occurred on or after Day 463

Abbreviations: DB, double-blind.
Study day is calculated as visit date - randomization 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 number of days to the scheduled visit date, the later assessment will be used for that given analysis visit.

For 6MWD at the Week 72 Visit (Visit 8), the date corresponding to the maximum of Day 1 and Day 2 will be used for deriving the study analysis visit.

4.9. Merging Strata

If one of the strata includes subjects from only one of the two treatment arms in ITT, mITT, and PP populations, the stratum cell will be combined with the closest neighbor stratum cell based on the patients' baseline assessments within corresponding stratum for the corresponding population.

4.10. Addressing the Impact of Coronavirus Disease of 2019 (COVID19)

Study visits have been impacted by site closures and travel restrictions due to Coronavirus Disease of 2019 (COVID19). Mitigation strategies have been implemented to reduce the number of missed or delayed visits for the study, including allowing subjects to have visits outside of the protocol defined window (±7 days). To minimize the impact from the out-of-window assessments, the actual week, instead of nominal visit, will be used in statistical modelling, where appropriate.

The protocol described MMRM model using actual week at the visit will be considered as primary efficacy analysis method. The protocol proposed the analysis of ANCOVA model analyzing 6MWD at Week 72 after MI assumes that almost all subjects have visits within protocol defined window will be considered as sensitivity analysis.

In addition, to assess the impact of COVID19, a sensitivity analysis for 6MWD using MMRM and ANCOVA after MI after removing all COVID19 impacted visits will be performed for mITT and ITT populations as described in Section 6.1.3.6.

5. SUBJECT DATA

5.1. Subject Disposition and Study Populations

The number of subjects screened, randomized, discontinued early during the double-blind treatment period (along with reasons for discontinuation) will be summarized. The number of subjects randomized will also be summarized by region, country, and site.

The number of subjects in the ITT, mITT, as-treated, and PP populations (along with reasons excluded from the ITT and PP populations) will be summarized.

The number of subjects with major protocol deviation will also be summarized for ITT, mITT, and as-treated populations.

5.2. Duration of Treatment with Study Drug

Duration of treatment with study drug will be calculated as:

$$\begin{aligned} Duration\ (days) \\ &= Date\ of\ last\ dose\ of\ double\text{-}blind\ study\ drug \\ &- Date\ of\ first\ dose\ of\ study\ drug + 1 \end{aligned}$$

Duration of treatment with study drug in weeks will be summarized continuously and categorically (<12 weeks, 12 to <24 weeks, 24 to <36 weeks, 36 to <48 weeks, 48 to <60 weeks, 60 to <72 weeks, ≥72 weeks) based on ITT, mITT, and as-treated populations.

5.3. Study Drug Compliance

To evaluate study drug compliance, number of subjects with treatment interruption, and the duration of treatment interruptions will be summarized for ITT, mITT, and as-treated populations.

5.4. Demographic and Baseline Characteristics

Demographic and baseline characteristics including age, gender, race, ethnicity, region (North and Latin America, Europe, and Asia Pacific), country, height, weight, body mass index (BMI), steroid use at baseline, baseline 6MWD, and baseline time to stand from supine will be summarized for ITT, mITT, PP, and as-treated populations.

5.5. Disease Characteristics

Disease characteristics including age at onset of phenotypic evidence of dystrophinopathy, age at first observation of difficulty with ambulation, age at diagnosis of dystrophinopathy, signs/symptoms/laboratory results used to support diagnosis, prior dystrophin gene sequencing, location of nonsense point mutation, and stop codon type will be summarized for ITT, mITT, and PP populations.

5.6. Medical History

Medical history of dystrophinopathy will be summarized with disease characteristics as described in Section 5.5. Medical history other than dystrophinopathy will be presented in the by-subject listing for as-treated population.

5.7. Concomitant Medications and Non-Drug Treatments

Concomitant medications and non-drug treatments will be coded using the World Health Organization Drug Dictionary (WHODD), version September 2021 or higher.

Prior medications and non-drug treatments are defined as those taken any time prior to Day 1. Concomitant medications and non-drug treatments for double-blind period are defined as those taken any time during the double-blind treatment period (Day 1 to the date prior to the first dose of open-label study treatment, inclusive).

The use of prior and concomitant medications and non-drug treatments will be summarized by Anatomical Therapeutic Chemical (ATC) Level 3 and preferred term (PT) for ITT, mITT, and as-treated populations. In addition, corticosteroids use at baseline and the incidence of changes in regimen of corticosteroids will be summarized for ITT, mITT, and as-treated populations.

6. EFFICACY EVALUATION

Unless otherwise specified, the actual weeks at each visit will be used for all MMRM analyses using all valid assessments. The by-visit summaries will be generated based on the analysis visits described in Section 4.8.

6.1. Primary Efficacy Variable (6MWD)

6.1.1. By-Visit Summary

The mean and mean changes from baseline in 6MWD will be summarized by treatment group at each visit. For these summaries, any subjects who lose ambulation during the study will be assigned a 6MWT result of 0 meters for the visit in which the subjects lose ambulation and for all remaining visits while they participate in the study.

6.1.2. Primary Analysis

The rate of change in 6MWD over 72 weeks (slope) in the mITT population will be considered as the primary efficacy endpoint.

A MMRM model will be employed using a random intercept and a random slope of change per week using observed data assuming the missing assessments are MAR. The slope will be compared between ataluren and placebo, and the estimated treatment difference with corresponding 95% confidence interval (CI) will be provided. The estimated difference in the change at Week 72 will also be provided to facilitate interpretation. The MMRM model will include baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6MWD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), baseline 6MWD, treatment, week (actual week at the visit as a continuous variable), the interaction of week and baseline 6MWD, and the interaction of week and treatment. The unstructured covariance matrix (UN) will be used in the MMRM model. Other types of covariance matrix (heterogenous Toeplitz, Toeplitz, and compound symmetric) will be explored in the case of non-convergence of UN. For this analysis, any subjects who lose ambulation during the study will be assigned a 6MWT result of 0 meters for the first visit in which the subjects lose ambulation and set for missing for all visits afterwards while they participate in the study.

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

The example SAS code is listed below:

```
PROC MIXED data=<dataset> method=reml;
  CLASS Cortico Subjid Trt C6WMD;
  MODEL N6WMD = Cortico C6WMD Trt B6WMD Week Week*B6WMD
        Week*TRT/S CL;
  RANDOM Int Week/type=un subject=Subjid;
  LSMEANS TRT / pdiff at Week=72 CL;
RUN;
```

where

N6WMD = 6WMD at each visit (Baseline to Week 72)

B6WMD = Baseline 6WMD
C6WMD = Baseline 6WMD category
Cortico = Baseline concomitant corticosteroid type
Week = Actual study week at each visit
Trt = Treatment group
Subjid = Subject
Int = Intercept

6.1.3. Supportive/Sensitivity Analyses

6.1.3.1. Analysis of Covariance (ANCOVA) Using Multiple Imputation (MI)

As a sensitivity analysis, an ANCOVA model will be used to evaluate the difference in change in 6MWD from baseline to Week 72 after missing data have been imputed via a pattern mixture model utilizing MI. The estimated treatment difference in change from baseline of 6MWD at Week 72 and the corresponding 95% confidence interval (CI) will be provided. The ANCOVA model will include baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), treatment, as well as baseline 6MWD as a covariate. For this analysis, all 6MWD will be assigned to an analysis visit according to the analysis visit window described in Section 4.8. Any subjects who lose ambulation during the study will be assigned a 6MWT result of 0 meters for the visit in which the subjects lose ambulation and for all remaining visits while they participate in the study. For subjects with missing 6MWD due to other reasons, the missing assessments at post-baseline visits will be imputed using the pattern-mixture model MI assuming the missing values are missing not at random (MNAR) utilizing the complete case missing value pattern (CCMV - subjects who completed primary efficacy assessments without missing values). Detailed imputation method is described in Appendix 2.

The example SAS code is listed below:

```
PROC MIXED data=<dataset> method=reml;
  CLASS Cortico Trt C6MWD;
  MODEL CHG6MWD = Cortico C6MWD Trt B6MWD /CL;
  LSMEANS TRT / pdiff CL;
RUN;
```

where

CHG6MWD = change from baseline in 6MWD at Week 72
B6MWD = Baseline 6MWD
C6MWD = Baseline 6MWD category
Cortico = Baseline concomitant corticosteroid type
Trt = Treatment group

6.1.3.2. ITT and PP Populations

As a supportive analysis, the primary MMRM analysis will be repeated using both the ITT and PP populations. In addition, the ANCOVA model after MI will also be repeated for ITT population. The baseline time to stand from supine (<5 seconds, ≥5 seconds) will be included in the testing model in the ITT analysis.

6.1.3.3. Modification of mITT Criteria (+/- 15%)

As supportive analyses, separate MMRM models similar to the primary analysis will be used to evaluate the change in 6MWD from baseline to Week 72 for all subjects with a baseline 6MWD ≥255 meters (ie, -15% from the 6MWD ≥300 meter lower bound) or baseline 6MWD ≥345 meters (ie, +15% from the 6MWD ≥300 meter lower bound)) and who also meet the age and stand from supine criteria for the mITT population (ie, 7 to 16 years old and time to stand from supine ≥5 seconds at baseline).

6.1.3.4. Checking Missing at Random Assumption

For the primary analysis, an additional sensitivity analysis will be employed to check the MAR assumption underlying the MMRM model where the reason for drop-out is considered. Subjects with following discontinuation reasons will be imputed using MI method using subjects in placebo arm who completed double-blind period of the study without missing values.

- Adverse event
- Withdrew consent
- Investigator decision

This analysis will be performed for mITT population. Detailed method and SAS codes are described in Appendix 3.

6.1.3.5. Using Average of Day 1 and Day 2 6MWDs Rather than Maximum

The primary analysis uses the maximum of valid Day 1 and Day 2 6MWDs for baseline and Week 72. As a sensitivity analysis, the same analysis will be repeated with the average of valid Day 1 and Day 2 6MWDs for baseline and Week 72 for mITT population. For this analysis, the actual week for Baseline and Visit 8 will be calculated based on the first day of the visit.

6.1.3.6. Assess the Impact of COVID19

To assess the impact of COVID19, a sensitivity analysis will be performed similarly to the primary MMRM analysis and ANCOVA analysis after MI by removing all COVID19 impacted visits (as collected on CRF) for mITT and ITT populations. The baseline time to stand from supine (<5 seconds, ≥5 seconds) will be included in the testing model in the ITT analysis.

6.1.3.7. Assess the Impact from Potential Correlation among Siblings

In this study, brothers participating in the study will receive the same treatment group designation to minimize potential compromise of the study drug blinding and to reduce the chance of dosing errors within families. In the primary analysis, siblings will be considered as independent subjects. However, the assessments from two or more brothers may be highly correlated. To assess the impact of potential correlation among siblings, the primary MMRM analysis will be repeated by including only data from the first sibling for mITT and ITT populations. The baseline time to stand from supine (<5 seconds, ≥5 seconds) will be included in the testing model in the ITT analysis.

6.1.3.8. Non-parametric analysis

As a sensitivity analysis, the change from baseline at Week 72 in 6MWD using observed data between two treatment groups will be compared using Wilcoxon rank sum test stratified by baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone) and baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters) for mITT population.

6.1.3.9. Subgroup Analysis

The change from baseline in 6MWD will be summarized by visit for subgroups below for mITT and ITT populations:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)

In addition, the change from baseline in 6MWD will be summarized at each post-baseline visit for the following subgroups based on baseline disease severity for ITT population:

- Baseline disease severity
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and 7 to 16 years old
  - Baseline 6MWD ≥300 meters and time to rise from supine <5 seconds at baseline, or baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and younger than 7 or older than 16 years old

6.2. Secondary Efficacy Variables

6.2.1. Key Secondary Efficacy Variables

6.2.1.1. Timed Function Tests - Composite Scores

The rate of change over 72 weeks (slope) in composite TFT scores will be considered as the first key secondary efficacy endpoint. The composite TFT scores will be defined as the average in times to run/walk 10 meters, climb 4 stairs, and descend 4 stairs. The change in composite TFT scores will be analyzed using the same MMRM model used for the primary analysis with factors of baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), treatment, visit, the average time in the three timed function tests at baseline, interaction of visit and treatment, and the interaction of week and baseline average time. The slope will be compared between ataluren and placebo, and the estimated difference in the change at Week 72 will also be provided to facilitate interpretation.

Subjects who cannot perform a timed function test within 30 seconds, including those who is LoA or the timed function test is above 30 seconds, will be assigned a value of 30 seconds for the respective test for calculating the average change in times. Similarly to the primary analysis, if the average time of the 3 TFT tests is 30 seconds for multiple visits for a subject, the 30 seconds will be kept for the first visit with the assessment of 30 seconds, and set for missing for all visits afterwards.

If the test of the primary endpoint is statistically significant, then this key secondary endpoints will be tested at the 0.05 (two-sided) significance level.

As supportive analyses, the same analysis will be repeated:

- ANCOVA based on mITT population after MI
- MMRM model based on ITT population with additional factor of the baseline time to stand from supine (<5 seconds, ≥5 seconds) in the model
- ANCOVA based on ITT population after MI (with additional factor of the baseline time to stand from supine (<5 seconds, ≥5 seconds) in the model)

The change from baseline in composite average of timed function will be summarized by visit for mITT and ITT populations, and for subgroups below for mITT and ITT populations:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)

In addition, the change from baseline in composite average of timed function will also be summarized at each post-baseline visit for ITT population for the following subgroups:

- Baseline disease severity
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and 7 to 16 years old
  - Baseline 6MWD ≥300 meters and time to rise from supine <5 seconds at baseline, or baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and younger than 7 or older than 16 years old

6.2.1.2. North Star Ambulatory Assessment - Total Scores

The NSAA consists of 17 activities, each scored as 0, 1, or 2. The sum of 16 scores (except for lifts head) will be used to form a total score [Mayhew 2013a]. Hereafter, this NSAA total score will be referred to as revised NSAA (rNSAA) total score. If an activity cannot be performed due to disease progression/LoA, a score of 0 will be assigned, and will not be considered as missing. If fewer than 13 of the 16 activities are performed, the total score will be considered missing. If from 13 to 16 activities are performed, the total score will be standardized by multiplying the sum of the scores in the x activities by 16/x.

The change from baseline in rNSAA total score at Week 72 based on the mITT population will be the second key secondary endpoint. An ANCOVA model will be used to evaluate the difference in change in rNSAA from baseline to Week 72 after missing data have been imputed via a pattern mixture model utilizing MI. The ANCOVA model will include factors of baseline

concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), and treatment, and covariate of baseline rNSAA total score. The estimated difference in the change at Week 72 will be provided.

If both primary efficacy endpoint and the first key secondary endpoint are statistically significant at the 0.05 significance level, the rNSAA will be tested at the 0.05 (two-sided) significance level.

The same ANCOVA model will also be performed on linear transformation of the total score [Mayhew 2013a].

As a sensitivity analysis, rNSAA total score and linear score will be analyzed using a MMRM model with a random intercept. The MMRM model will include baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), treatment, visit (analysis visit as a categorical variable), the baseline rNSAA score as the covariate, as well as interaction of visit and treatment, and the interaction of visit and baseline rNSAA total score. The unstructured covariance matrix (UN) will be used in the MMRM model. Other types of covariance matrix (heterogenous Toeplitz, Toeplitz, and compound symmetric) will be explored in the case of non-convergence of UN.

A sample code is listed as below:

```
PROC MIXED data=<dataset> method=reml;
  CLASS Cortico Subjid Trt C6WMD AVISIT;
  MODEL N6WMD = Cortico C6WMD Trt B6WMD AVISIT AVISIT*B6WMD
        AVISIT*TRT/CL;
  RANDOM Int / subject=Subjid;
  REPEATED AVISIT / type=un subject=Subjid;
  LSMEANS TRT*AVISIT / pdiff CL;
RUN;
```

where

N6WMD = 6WMD at each visit (Baseline to Week 72)
B6WMD = Baseline 6WMD
C6WMD = Baseline 6WMD category
Cortico = Baseline concomitant corticosteroid type
AVISIT = Analysis visits as categorical variable
Trt = Treatment group
Subjid = Subject
Int = Intercept

The same analysis will be repeated for the change from baseline in rNSAA total scores:

- ANCOVA after MI for ITT population (with additional factor of the baseline time to stand from supine (<5 seconds, ≥5 seconds) in the model)
- MMRM model based on ITT population with additional factor of the baseline time to stand from supine (<5 seconds, ≥5 seconds) in the model

The change from baseline in rNSAA total scores and linear scores will be summarized by visit for mITT and ITT populations, and for subgroups below for mITT and ITT populations:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)

In addition, the change from baseline in revised NSAA total scores and linear scores will also be summarized at each post-baseline visit for ITT population for the following subgroups:

- Baseline disease severity:
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and 7 to 16 years old
  - Baseline 6MWD ≥300 meters and time to rise from supine <5 seconds at baseline, or baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and younger than 7 or older than 16 years old

6.2.2. Other Secondary Efficacy Variables

6.2.2.1. Time to 10% Persistent Worsening in 6MWD

Time to 10% persistent worsening in 6MWD, defined as last time that 6MWD was not 10% worse than baseline, 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 populations. For subjects who do not have 10% 6MWD worsening, time to 10% persistent 6MWD worsening will be censored at the time of the last 6MWT during the double-blind period. For subjects who have all post-baseline assessments more than 10% worse than the baseline, the event will be assigned to Day 1. Subjects who become non-ambulatory will be considered to have 10% worsening at the time of becoming non-ambulatory. The proportion of 10% persistent worsening by Week 72 will also be summarized.

6.2.2.2. Timed Function Tests - Individual Test

The 3 timed function tests (10-meter run/walk, 4-stair stair-climb, and 4-stair descend) will be summarized and analyzed separately for each function test using the MMRM model assuming MAR based on both mITT and ITT populations. The MMRM model will include baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), respective baseline timed function test result, treatment, visit, the interaction of visit and treatment, and the interaction of week and corresponding timed function test result at baseline. The baseline time to stand from supine (<5 seconds, ≥5 seconds) will be included in the testing model in the ITT analysis.

In addition, ANCOVA analysis after MI will be performed for the three timed function tests as sensitivity analysis. The individual TFT results will be summarized by visit during the double-blind period for mITT and ITT populations.

A separate analysis will be performed to evaluate the 10-meter run/walk results in subjects with a baseline 6WMD <300 meters in ITT population.

6.2.2.3. North Star Ambulatory Assessment – Individual Items

The frequency of loss of function of each NSAA items (score of 1 or 2 at baseline transitioning to a 0 score post-baseline) will be tabulated by treatment group for both mITT and ITT populations.

6.2.2.4. Loss of Ambulation (LoA) and Stair Climb/Stair Descend

LoA will be defined as persistent inability to perform the 10-meter run/walk test within 30 seconds at any post-baseline visit and for all remaining visits. Loss of stair-climbing will be defined as persistent inability to perform the 4-stair climb test within 30 seconds at any post-baseline visit and for all remaining visits. Loss of stair-descending will be defined as persistent inability to perform the 4-stair descend test within 30 seconds at any post-baseline visit and for all remaining visits. (Note: All subjects will have the ability to perform these tests within 30 seconds at screening and baseline, per the inclusion criteria.)

Time to LoA will be evaluated using Kaplan-Meier estimation, Log-rank test, and Cox proportional hazards model (assuming similar covariates as the primary efficacy model) in both mITT and ITT populations. Similar analyses will be performed for time to loss of stair-climbing and time to loss of stair-descending.

A separate analysis will evaluate time to LoA in subjects with baseline 6MWD <300 meters in ITT population.

6.3. Exploratory Variables

6.3.1. PUL and DMD Upper Limb PROM

PUL includes 22 items (with an entry item to define starting functional level), and 21 items subdivided into shoulder level (4 items), elbow level (9 items), and distal level (8 items) dimensions. The DMD Upper Limb PROM includes 32 items categorized in 4 domains of daily life: food (7 items), self-care (8 items), household and environment (6 items), and leisure and communication (11 items). The domain subscores will be calculated by summing the scores from all items within each domain, and the total score will be calculated as the sum of all domain scores. A domain subscore is considered missing if any of the individual item is missing within the respective domain. The total score is considered missing if any of the individual item is missing from a giving assessment.

The PUL and DMD Upper Limb PROM total scores will be summarized and analyzed separately using the ANCOVA after MI based on both mITT and ITT populations. The ANCOVA model will include baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), and treatment as factors, and respective baseline scores as a covariate. The baseline time to stand from supine (<5 seconds, ≥5 seconds) will be included in the testing model in the ITT analysis.

The frequency of losing each PROM item will also be tabulated by treatment group for mITT and ITT populations.

In addition, the PUL total and domain subscores will also be summarized by age group (7 to <8 years old, 8 to <13years old, 13 to <22 years old, and >=22 years old) using descriptive statistics within each treatment group for mITT and ITT populations. The last assessment during the specific age group will be used for a given subject in this summary.

The PUL total and domain subscores will also be summarized and analyzed for subjects with the following subgroups for ITT population:

- Baseline 6MWD <300 meters
- Baseline 6MWD ≥300 meters and time to rise from supine <5 seconds at baseline, or baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and younger than 7 or older than 16 years old

6.3.2. Myometry Parameters

Myometry parameters (knee extension and elbow flexion) will be analyzed using the ANCOVA after MI as described for the primary endpoint and summarized by treatment groups at each visit based on ITT population.

6.3.3. Muscle Fat Fraction

Muscle fat fraction data obtained (proton density fat fraction and MRS fat fraction) from MRI will be summarized by treatment groups at each visit based on both the mITT and ITT populations. The muscle fat fraction will also be summarized for subjects with the following subgroups:

- Baseline 6MWD <300 meters
- Baseline 6MWD ≥300 meters and time to rise from supine <5 seconds at baseline, or baseline 6MWD ≥300 meters, time to rise from supine ≥5 seconds at baseline, and younger than 7 or older than 16 years old

6.3.4. Spirometry - Pulmonary Function Test

The change from baseline in FVC assessments (FVC, FEV1, and PEF) and %-predicted FVC will be summarized in 12-week intervals from the time of LoA visit based on both mITT and ITT populations. The baseline of FVC assessment is defined as the FVC assessment taken at the visit when subjects become non-ambulatory in this analysis.

Predicted FVC will be calculated as [Gauld 2003]:

Predicted FVC = exp(0.077 * ulna length + 0.041 * actual age – 1.285)

%-predicted FVC will be calculated as:

%-predicted FVC = observed FVC / predicted FVC * 100%

Other spirometry parameters including maximum inspiratory (MIP) and expiratory pressures (MEP) will be presented in subject listings only.

6.3.5. HRQL At-Home Questionnaires

HRQL data obtained from the EQ-5D and DMDSAT will be summarized by treatment group at each visit based on both mITT and ITT populations.

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, some problems, 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 for mITT and ITT populations.

The DMDSAT comprised a total of eight questions covering four domains (arm function, mobility, transfers, and ventilation status). The domain scores will be summed up to form the total scores. The domain scores will be tabulated and the mean and mean change in total score will be summarized by treatment groups at each visit for mITT and ITT populations.

7. SAFETY EVALUATION

Derived analysis visits based on analysis window described in Section 4.8 will be used for safety analysis, unless otherwise specified.

7.1. Adverse Events

AEs will be coded to the MedDRA (Version 24.1 or higher) lower level term closest to the verbatim term. The linked MedDRA PT and primary system organ class (SOC) are also captured in the database. The severity of AEs will be graded using the latest version of the Common Terminology Criteria for Adverse Events (CTCAE), Version 5.0 whenever possible.

For the double-blind treatment period, a treatment-emergent adverse event (TEAE) is defined as an adverse event that:

- Occurs or worsens during the double-blind treatment period (Day 1 to within 4 weeks of last double-blind treatment, and prior to the first open-label treatment) for subjects who complete the double-blind treatment period; and
- Occurs or worsens from the period extending from Day 1 to 4 weeks after the last dose of double-blind study drug for subjects who discontinue from the study early during the double-blind treatment period, or who complete double-blind treatment period but not enroll into open-label treatment period.

Only those AEs that are treatment-emergent will be included in summary tables. All AEs, treatment-emergent or otherwise, will be presented in subject data listings. An overview table of TEAEs, including number of subjects with TEAEs, treatment-emergent serious adverse events (TESAEs), deaths, TEAEs classified as CTCAE Grade 3 or higher, study drug related TEAEs, TEAEs leading to study drug withdrawal will be provided for each treatment group. The following summaries will be produced for the TEAEs by treatment group:

- 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 treatment.

In the summary tables subjects may be counted under multiple SOCs and PTs, but for each SOC and PT, subjects are only counted once. If a subject has the same AE on multiple occasions, the highest severity (5=fatal, 4=life-threatening, 3=severe, 2=moderate, 1=mild) recorded for the event will be presented and the highest drug relationship (1 = ‘Unrelated’, 2 = ‘Unlikely to be Related’, 3 = ‘Possibly Related’, 4 = ‘Probably Related’), reclassified into Related (‘Possibly Related’, ‘Probably Related’) or Not Related (‘Unrelated’, ‘Unlikely to be Related’), will be presented on the respective tables.

The following summaries will also be presented for the TESAEs:

- Incidence of TESAEs by SOC and PT
- Incidence of TESAEs by PT in descending order

- Incidence of treatment-related TESAEs by SOC and PT.

In addition, number and percentage of subjects with TEAEs and treatment-related TEAEs leading to discontinuation from study treatment will also be summarized by MedDRA SOC and PT for each treatmzent group.

SAE, AE leading to discontinuation from study treatment, and death will be listed.

7.1.1. Adverse Events of Special Interest (AESI)

The following TEAEs are considered as the adverse events of special interest (AESI), and will be summarized by AESI Category and PT for each treatment group:

Table 3: Adverse Events of Special Interest (AESI)

AESI Category	Definition
Potentiation of aminoglycoside renal toxicity	• SMQ: Acute renal failure (Narrow and Broad), and; • Concomitant medication using the WHODD ATC code J01G
Long-term cardiovascular effects including changes in lipid profile	• SMQ: Dyslipidaemia (Narrow), or; • SOC: Cardiac disorders
Hypertension with use of concomitant systemic corticosteroids	• SMQ: Hypertension (Narrow and Broad), and; • Concomitant medication using the WHODD ATC code H02A
Renal toxicity	• SMQ: Acute renal failure (Narrow and Broad)
Hepatic toxicity	• SOC: Hepatobiliary disorders, or; • HLGT: Hepatobiliary Investigations
Malignancies in general	• SOC: Neoplasms benign, malignant and unspecified (incl cysts and polyps)

Abbreviations: AESI, adverse event of special interest; SMQ, Standardized MedDRA Query; ATC, Anatomical Therapeutic Chemical; SOC, System Organ Class; HLGT, High Level Group Term; WHODD, World Health Organization Drug Dictionary
For potentiation of aminoglycoside renal toxicity and hypertension with use of concomitant systemic corticosteroids, the concomitant medications are those started prior to the onset of corresponding AE.

7.2. Clinical Laboratory

Mean and mean change from baseline in clinical laboratory data from central laboratory (including parameters for assessment of hepatic and renal monitoring) listed in Table 4 will be summarized by analysis visit, as well as the last assessment during the double-blind period. In addition, shift tables for each laboratory parameters from baseline to each of the post-baseline visit, and to last assessment will be provided.

Table 4: Clinical Laboratory Parameters

Type	Parameters
Hematology	leukocytes, basophils, basophils/leukocytes, eosinophils, eosinophils/leukocytes, lymphocytes, lymphocytes/leukocytes, monocytes, monocytes/leukocytes, neutrophils, neutrophils/leukocytes, hemoglobin, hematocrit, erythrocytes, platelets
Biochemistry	sodium, potassium, bicarbonate, blood urea nitrogen, creatinine, magnesium, calcium, uric acid, glucose, total protein, bilirubin (direct and indirect), aspartate aminotransferase, alanine Aminotransferase, gamma glutamyl Transferase, creatine kinase, alkaline phosphatase, total cholesterol, lactate dehydrogenase (LDH), high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides, and cystatin C
Urinalysis	pH, glucose, ketones, blood, protein

All post-baseline clinical laboratory results will be graded according to CTCAE severity grade (criteria in Appendix 4), when applicable. For parameters graded according to CTCAE severity, the CTCAE grades increase from baseline to Grade 2 or higher grade will be considered as clinically significant (CS) and the incidence will be summarized for as-treated population.

The incidence of all post-baseline specific renal laboratory parameters provides information on action to be taken for study drug will be summarized.

Table 5: Renal Monitoring Parameters and Actions to Be Taken

Laboratory Parameter	Stop Study Drug Immediately, Confirm Abnormal Value, and Then Start Work-Up	Stop Study Drug After Confirming[a] Abnormal Value, and Then Start Work-Up
Serum cystatin C	>2.00 mg/L	>1.33 - 2.00 mg/L
Serum creatinine	≥ Grade 2 (≥1.5 x ULN for age)	Grade 1 (>ULN - 1.5 x ULN for age)
Serum BUN	≥3.0 x ULN	≥1.5 - 3.0 x ULN

Abbreviations: BUN, blood urea nitrogen; ULN, upper limit of normal
a Laboratory abnormalities may be confirmed immediately or at the next scheduled clinic visit based on investigator judgment.

Laboratory abnormality related to elevated liver function test will also be summarized for all post-baseline assessments according to Table 6.

Table 6: Criteria for Elevated Liver Function Test

Category	Criteria
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

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; AP, alkaline phosphatase; ULN, upper limit of normal

A listing of all CS abnormal laboratory values, renal laboratory abnormal values, and elevated liver function test will be provided.

7.3. Electrocardiogram/Echocardiogram

Overall interpretation of ECG data, including normal, abnormal and clinically significant, abnormal and not clinically significant will be summarized descriptively by the number and percentages of subjects at each visit, and the last assessment during the double-blind period.

A listing of all abnormal values for ECG will be provided.

Echocardiogram parameters (left ventricle ejection fraction and sphericity index) will be summarized descriptively, including change from baseline by visit, and the last assessment during the double-blind period. Overall interpretation of echocardiogram (normal, abnormal) will be summarized descriptively by the number and percentages of subjects at each visit, and the last assessment during the double-blind period.

7.4. Physical Examination

Physical examination results will be listed in subject listings only.

7.5. Vital Signs

Height, weight, and vital signs data will be summarized descriptively, including change from baseline at each post-baseline and the last assessment during double-blind period. A summary will also be provided by visit and overall of the number of subjects meeting criteria for hypertension based on age, gender, and height-adjusted systolic blood pressure and diastolic blood pressure percentile results [Flynn 2017].

- Normal BP: BP <90th percentile for age, sex, and height; or <120/<80 mm Hg for adolescents ≥13 years old;
- Elevated BP: BP reading ≥90th percentile and <95th percentile for age, sex, and height; or 120 to 129/<80 mm Hg for adolescents ≥13 years old;
- Hypertension: BP ≥95th percentile for age, sex, and height; or ≥130/80 mm Hg for adolescents ≥13 years old.

8. CHANGES FROM THE PROTOCOL

The major changes to the planned analyses from the protocol are summarized as follows:

- NSAA total score will be calculated as the sum of the 16 scores (except for lifts head), instead of 17, based on Mayhew's publication [Mayhew 2013a].
- Aggregated odds ratio of NSAA will not be calculated because of the correlated information and possible inconsistent baseline within subjects. Due to the same reason, the aggregated odds ratio of PROM will not be calculated.
- Evaluable population defined in protocol is removed. The planned analysis using Evaluable population will be replaced by subjects with sufficient information in ITT, mITT, or as-treated population, where applicable.
- The protocol proposed comparison with external natural history cohorts will not be described in this SAP. This comparison will be described and performed in future if deemed necessary.

9. REFERENCES

Mayhew AG, Cano SJ, Scott E, Eagle M, Bushby K, Manzur A, Muntoni F; North Star Clinical Network for Neuromuscular Disease. Detecting meaningful change using the North Star Ambulatory Assessment in Duchenne muscular dystrophy. Dev Med Child Neurol. 2013a Nov;55(11):1046-52.

Flynn, JT and Falkner, BE. New Clinical Practice Guideline for the Management of High Blood Pressure in Children and Adolescents. Hypertension 2017;70(4):683-686.

NIH National Cancer Institute. Common Toxicity Criteria for Adverse Events (CTCAE) v5.0. November 27, 2017.
https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm#ctc_50

10. TABLES, GRAPHS, AND LISTINGS SHELLS

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

APPENDIX 1. SCHEDULE OF ASSESSMENTS

Study Period	**Screening**	**Double-Blind Treatment**									**Notes**
Study Week	**-2 weeks**	**1**		**12**	**24**	**36**	**48**	**60**	**72**		2-week screening period may be extended by 1 week if necessary to schedule pretreatment MRI
Study Day		**1**		**84**	**168**	**252**	**336**	**420**	**504**		
Visit Window (days)				±7	±7	±7	±7	±7	±7		
Visit Number	**1**	**2**		**3**	**4**	**5**	**6**	**7**	**8**		At Visit 2, complete assessments before first dose
Visit Day		**1**	**2**						**1**	**2**	2-day visits only at Visit 2 and Visit 8
Informed consent	X										May be obtained prior to Visit 1 (eg, by phone)
Enter subject in IRT system	X										Access IRT system to enter subject
Clinical and medication history	X										
Hepatitis screen	X										
Urinalysis sample	X	X			X		X		X		
EQ-5D and DMDSAT	X	X		X	X	X	X	X	X		Provided to subject at Visit 1 for at-home completion between visits and reviewed for compliance in clinic
DMD Upper Limb PROM		X		X	X	X	X	X	X		Performed only in subjects ≥7 years at baseline
Height	X	X		X	X	X	X	X	X		
Weight	X	X		X	X	X	X	X	X		
Physical examination	X	X			X		X		X		Full (Visit 1, 8) or symptom-directed (Visit 2, 4, 6)
Concomitant medications	X	X		X	X	X	X	X	X		
Adverse events	X	X		X	X	X	X	X	X		
Genotyping sample		X									
Hematology sample	X	X			X		X		X		
Biochemistry sample	X	X			X		X		X		To be collected in fasted state (except Visit 1)
Vital signs	X	X		X	X	X	X	X	X		
Myometry	X	X		X	X	X	X	X	X		Performed only in subjects <7 years at baseline
PUL	X	X		X	X	X	X	X	X		Performed only in subjects ≥7 years at baseline
6-minute walk test (1st test)	X	X		X	X	X	X	X	X		
6-minute walk test (2nd test)			X							X	
NSAA	X	X		X	X	X	X	X	X		
Timed function tests	X	X		X	X	X	X	X	X		
Study drug dispensed			X	X	X	X	X	X		X	Open-label study drug dispensed at Visit 8
Study drug compliance				X	X	X	X	X	X		
12-lead ECG	X				X		X		X		
Spirometry	X	X							X		Required at all visits after subject loses ambulation
Echocardiogram	X								X		

Study Period	**Screening**	**Double-Blind Treatment**									**Notes**
Study Week	**-2 weeks**	**1**		**12**	**24**	**36**	**48**	**60**	**72**		2-week screening period may be extended by 1 week if necessary to schedule pretreatment MRI
Study Day		**1**		**84**	**168**	**252**	**336**	**420**	**504**		
Visit Window (days)				±7	±7	±7	±7	±7	±7		
Visit Number	**1**	**2**		**3**	**4**	**5**	**6**	**7**	**8**		At Visit 2, complete assessments before first dose
Visit Day		**1**	**2**						**1**	**2**	2-day visits only at Visit 2 and Visit 8
Randomization			X								Access IRT system to randomize subject
Upper leg imaging		X			X		X		X		Performed only at pre-qualified sites. May be done between Visits 1 and 2 (pre-treatment assessment) or ±4 weeks of visit (post-treatment assessments).

Abbreviations: DMD, Duchenne Muscular Dystrophy; DMDSAT, DMD Functional Ability Self-Assessment Tool; ECG, electrocardiogram; EQ-5D, EuroQoL 5-Dimension; IRT, interactive response technology; NSAA, North Star Ambulatory Assessment; PROM, Patient-Reported Outcome Measure; PUL, Performance of Upper Limb

APPENDIX 2. SAMPLE SAS PROGRAM FOR THE PRIMARY ANALYSIS FOR PRIMARY ENDPOINT USING ANCOVA MODEL

The analysis will employ of the following SAS procedures:

- 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.
- 100 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.
- 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 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

TRTn: Treatment (0=Placebo, 1=Ataluren)
Cortico: Steriod use at randomization
C6mwd: Baseline 6MWD categories
mITTfl: Identifier for mITT population
*********************************************/

* 1st: Generate 100 datasets to impute missing interim data by treatment
group using MCMC option;
proc sort data=smwt; by trtn usubjid; run;
proc MI data=smwt out=smwt_mono minimum=0 nimpute=100 minmaxiter=5000
seed=93182;
  WHERE mITTfl="Y";
  BY trtn;
  VAR SmwdWKBL SmwdWK12 SmwdWK24 SmwdWK36 SmwdWK48 SmwdWK60 SmwdWK72;
  MCMC chain=multiple impute=monotone;
run;
```

```
* 2nd: Impute missing ending data using monotone statement with regression
option and assuming MNAR using CCMV;

proc sort data=smwt_mono out=smwt_mono; by _imputation_ trtn usubjid; run;
proc MI data=smwt_mono out=smwt_ccmv minimum=0 nimpute=1 minmaxiter=5000
seed=93035;
  WHERE mITTfl="Y";
  BY _imputation_ trtn;
  VAR SmwdWKBL SmwdWK12 SmwdWK24 SmwdWK36 SmwdWK48 SmwdWK60 SmwdWK72;
  MONOTONE reg(/details);
  MNAR Model (SmwdWK12 SmwdWK24 SmwdWK36 SmwdWK48 SmwdWK60
SmwdWK72/Modelobs=CCMV);
run;

* 3rd: Create datasets for ANCOVA;
data smwt_reg;
  set smwt_ccmv;
  AWeek= 0; Avalimp=SmwdWK0;  CFBimp=SmwdWK0-BL6mwd;  output;
  AWeek=12; Avalimp=SmwdWK12; CFBimp=SmwdWK12-BL6mwd; output;
  AWeek=24; Avalimp=SmwdWK24; CFBimp=SmwdWK24-BL6mwd; output;
  AWeek=36; Avalimp=SmwdWK36; CFBimp=SmwdWK36-BL6mwd; output;
  AWeek=48; Avalimp=SmwdWK48; CFBimp=SmwdWK48-BL6mwd; output;
  AWeek=60; Avalimp=SmwdWK60; CFBimp=SmwdWK60-BL6mwd; output;
  AWeek=72; Avalimp=SmwdWK72; CFBimp=SmwdWK72-BL6mwd; output;
run;

…

* 4th: Conduct ANCOVA analysis on imputed datasets;
proc sort data=adsmwdimp out=adsmwdimp; by _imputation_; run;
ods exclude all;
proc mixed data=adsmwdimp method=REML;
  WHERE mITTfl="Y" and Aweek=72;
  BY _imputation;
  CLASS Cortico Trtn C6mwd;
  MODEL CFBimp=BL6MWD Trtn Cortico C6MWD / CL;
  LSMEANS Trtn/ diff=control('0') cl alpha=0.05;
  ods output diffs=diff_mi lsmeans=lsm_mi;
run;
ods exclude none;

* 5th: Use Proc MIAnalyze to obtain final analyiss results;
* LS Means;
proc mianalyze parms(classvar=full)=lsm_mi;
  CLASS Trtn;
  MODELEFFECTS Trtn;
  ods output parameterestimates=lsm_out;
run;

* Treatment differences;
proc mianalyze parms(classvar=full)=diff_mi;
  CLASS Trtn;
  MODELEFFECTS Trtn;
  ods output parameterestimates=diff_out;
run;
```

The table below lists the random seeds to be used for the primary and secondary analyses and some additional seeds should they be required. The random seeds were created using SAS uniform generator RAND("Uniform") with a seed of 20210126.

Endpoint/Analysis	**Seed for Step 1 (impute missing interim data)**	**seed for Step 2 (impute missing ending data)**
6MWD (mITT, Section 6.1.2)	93182	93035
6MWD (ITT, Section 6.1.3.2)	31096	27324
6MWD (PP, Section 6.1.3.2)	96919	60747
6MWD (Covid19 impact, mITT, Section 6.1.3.6)	57237	96213
6MWD (Covid19 impact, ITT, Section 6.1.3.6)	81564	23187
NSAA (mITT, Section 6.2.1.2)	53056	44931
NSAA (ITT, Section 6.2.1.2)	78128	55295
10 Meter Run/Walk (mITT, Section 6.2.2.2)	14792	5438
10 Meter Run/Walk (ITT, Section 6.2.2.2)	93543	15827
10 Meter Run/Walk (baseline 6MWD <300 meters, Section 6.2.2.2)	64644	67275
4-stair climb (mITT, Section 6.2.2.2)	40100	14015
4-stair climb (ITT, Section 6.2.2.2)	75269	1604
4-stair descent (mITT, Section 6.2.2.2)	13807	67037
4-stair descent (ITT, Section 6.2.2.2)	69749	60769
PUL (mITT, Section 6.3.1)	20889	68149
PUL (ITT, Section 6.3.1)	53584	22
DMD Upper Limb PROM (mITT, Section 6.3.1)	57087	56786
DMD Upper Limb PROM (ITT, Section 6.3.1)	37430	59741
Myometry (ITT, Section 6.3.2)	75789	20780
If additional analyses need to be performed, seeds below will be used		
Additional Analysis #1	47997	55087
Additional Analysis #2	12688	70229
Additional Analysis #3	91077	10733
Additional Analysis #4	28506	17466
Additional Analysis #5	17445	33103

APPENDIX 3. SAMPLE SAS PROGRAM FOR THE SENSITIVITY ANALYSIS OF CHECKING THE ASSUMPTION OF MISSING AT RANDOM

As sensitivity analysis of checking the assumption of MAR described in Section 6.1.3.4 will use following steps:

1. Keep all placebo subjects who completed the double-blind treatment period, and the subjects who identified prior to database lock
2. Similar to step 1 in Appendix 2, generating 100 datasets to impute interim data without considering the treatment
3. Similar to step 2 in Appendix 2, impute the missing ending data using monotone statement
4. Merge the imputed information for the identified subjects back to original dataset
5. Perform MMRM analysis
6. Use Proc MIAnalyze to obtain final results

Rand seed used for this multiple imputation are shown below.

Endpoint/Analysis	Seed for Step 1 (impute missing interim data)	seed for Step 2 (impute missing ending data)
6MWD for checking MAR assumption (mITT)	49564	50402

APPENDIX 4. LABORATORY RELATED CTCAE SEVERITY GRADE

AE Term	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
White blood cell decreased	<LLN - 3000/mm^3; <LLN - 3.0 x 10^9/L	<3000 - 2000/mm^3; <3.0 - 2.0 x 10^9/L	<2000 - 1000/mm^3; <2.0 - 1.0 x 10^9 /L	<1000/mm^3; <1.0 x 10^9/L	
White blood cell increased (leukocytosis)	-	-	>100,000 mm^3	Clinical manifestations of leucostasis; urgent intervention indicated	Death
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hgb <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hgb <8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated	Life-threatening consequences; urgent intervention indicated	Death
Lymphocyte count decreased	<LLN - 800/mm^3; <LLN - 0.8 x10^9 /L	<800 - 500/mm^3; <0.8 - 0.5 x 10^9 /L	<500 - 200/mm^3; <0.5 - 0.2 x 10^9 /L	<200/mm^3; <0.2 x 10^9 /L	
Lymphocyte count increased	-	>4000 - 20,000 mm^3	>20,000 mm^3	-	
Platelet count decreased	<LLN - 75,000/mm^3; <LLN - 75.0 x 10^9/L	<75,000 - 50,000/mm^3; <75.0 - 50.0 x 10^9/L	<50,000 - 25,000/mm^3; <50.0 - 25.0 x 10^9/L	<25,000/mm^3; <25.0 x 10^9/L	
Serum amylase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; >2.0 5.0 x ULN and asymptomatic	>2.0 - 5.0 x ULN with signs or symptoms; >5.0 x ULN and asymptomatic	>5.0 x ULN and with signs or symptoms	
Alkaline phosphatase increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	
Alanine aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	
Aspartate aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	

AE Term	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Blood bilirubin increased	>ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal	
GGT increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	
Hemoglobin increased	Increase in >0- 2 gm/dl above ULN	Increase in >2- 4 gm/dl above ULN	Increase in >4 gm/dl above ULN	-	
Lipase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; >2.0 - 5.0 x ULN and asymptomatic	>2.0 - 5.0 x ULN with signs or symptoms; >5.0 x ULN and asymptomatic	>5.0 x ULN and with signs or symptoms	
Hypercalcemia	Corrected serum calcium of >ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L; Ionized calcium >ULN - 1.5 mmol/L	Corrected serum calcium of >11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L; Ionized calcium >1.5 - 1.6 mmol/L; symptomatic	Corrected serum calcium of >12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L; hospitalization indicated	Corrected serum calcium of >13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L; life-threatening consequences	Death
Hypocalcemia	Corrected serum calcium of <LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L; Ionized calcium <LLN - 1.0 mmol/L	Corrected serum calcium of <8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L; Ionized calcium <1.0 - 0.9 mmol/L; symptomatic	Corrected serum calcium of <7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L; Ionized calcium <0.9 - 0.8 mmol/L; hospitalization indicated	Corrected serum calcium of <6.0 mg/dL; <1.5 mmol/L; Ionized calcium <0.8 mmol/L; life-threatening consequences	Death
Creatinine increased	>ULN - 1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 x baseline; >3.0 - 6.0 x ULN	>6.0 x ULN	
Hyperglycemia	Abnormal glucose above baseline with no medical intervention	Change in daily management from baseline for a diabetic; oral antiglycemic agent initiated; workup for diabetes	Insulin therapy initiated; hospitalization indicated	Life-threatening consequences; urgent intervention indicated	Death
Hypoglycemia	<LLN - 55 mg/dL; <LLN - 3.0 mmol/L	<55 - 40 mg/dL; <3.0 - 2.2 mmol/L	<40 - 30 mg/dL; <2.2 - 1.7 mmol/L	<30 mg/dL; <1.7 mmol/L; life-threatening consequences; seizures	Death

AE Term	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
Hypophosphatemia	≤Laboratory finding only and intervention not indicated	Oral replacement therapy indicated	Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of existing hospitalization indicated	life-threatening consequences	Death
Hyperkalemia	>ULN - 5.5 mmol/L	>5.5 - 6.0 mmol/L; intervention initiated	>6.0 - 7.0 mmol/L; hospitalization indicated	>7.0 mmol/L; life-threatening consequences	Death
Hypokalemia	<LLN - 3.0 mmol/L	<LLN - 3.0 mmol/L; symptomatic; intervention indicated	<3.0 - 2.5 mmol/L; hospitalization indicated	<2.5 mmol/L; life-threatening consequences	Death
Hyponatremia	<LLN - 130 mmol/L	125-129 mmol/L and asymptomatic	125-129 mmol/L symptomatic; 120-124 mmol/L regardless of symptoms	<120 mmol/L; life-threatening consequences	Death
Hypernatremia	>ULN - 150 mmol/L	>150 - 155 mmol/L; intervention initiated	>155 - 160 mmol/L; hospitalization indicated	>160 mmol/L; life-threatening consequences	Death
Hyperuricemia	>ULN without physiologic consequences	-	>ULN with physiologic consequences	life-threatening consequences	Death

Note: Based on Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0.

Signature Page for EU PTC124-GD-041-DMD Statistical Analysis Plan (SAP) - Double

Statistics Approval	[redacted] I approve the document(s) 04-Apr-2022 21:07:27 GMT+0000

Clinical Approval	[redacted] I approve the document(s) 05-Apr-2022 09:03:34 GMT+0000

Statistics Approval	[redacted] I approve the document(s) 05-Apr-2022 16:01:30 GMT+0000

Signature Page for VV-CLIN-005686 v1.0

CHANGE FROM PLANNED STATISTICAL ANALYSIS (DOUBLE-BLIND TREATMENT PERIOD AND OPEN-LABEL PERIOD) for EMA/MHRA

A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED EFFICACY AND SAFETY STUDY OF ATALUREN IN PATIENTS WITH NONSENSE MUTATION DUCHENNE MUSCULAR DYSTROPHY AND OPEN-LABEL EXTENSION

PTC124-GD-041-DMD

19 SEPTEMBER 2022

VERSION 1.0

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 or Specialist Term	Explanation
6MWD	6-minute walk distance
ANCOVA	analysis of covariance
CI	confidence interval
CRF	case report form
CSR	clinical study report
COVID19	Coronavirus Disease of 2019
DMD	Duchenne muscular dystrophy
EMA	European Medicines Agency
ITT	intention-to-treat
LoA	loss of ambulation
MAR	missing at random
MHRA	Medicines and Healthcare products Regulatory Agency
MI	multiple imputation
mITT	modified intention-to-treat
MMRM	mixed model for repeated measures
nmDMD	Nonsense mutation Duchenne muscular dystrophy
NSAA	North Star Ambulatory Assessment
PROM	Patient-Reported Outcome Measure
PUL	Performance of Upper Limb
rNSAA	revised North Star Ambulatory Assessment total score
SAP	statistical analysis plan
SE	standard error
TFT	timed function test
UN	unstructured covariance matrix

1. OVERVIEW

This document details the changes and the rationales for the changes in the statistical methods in analyzing Study PTC124-GD-041-DMD (hereafter referred to as Study 041) from the planned analyses for **EMA/MHRA**.

This document is prepared on the basis of the final study protocol version 4.0 (dated 21JUL2020) and the final statistical analysis plan (SAP) version 1.0 (dated 04APR2022). The reasons for the changes are described in Section 2 and the detailed list of changes are given in Section 3.

The key changes made from the final SAP are summarized below:

- The assumption of constant (linear) decline in 6MWD over 72-week underpinning the prespecified linear regression MMRM model did not hold. The observed rate of decline was found to vary over time and between visits in the 72-week study period. To accurately account for this non-linearity the 'visit time' variable in the MMRM model was changed from a continuous (regression) variable to a categorical variable. The assessment for linearity and adjustment of the model were conducted in accordance with EMA guidance EMEA/CHMP/295050/2013 Section 7.4.
- Although there was a clear trend in favor of ataluren's treatment effect in the predefined primary analysis subset (modified intention-to-treat, mITT), it did not fulfill its purpose as the most sensitive subset to capture significant benefit over the finite time period of a clinical trial. Therefore, the results for the full intention-to-treat (ITT) analysis set (a pre-specified supportive population for the primary endpoint analysis) will be presented in-text. Additional sensitivity analyses were added to confirm the robustness of the conclusions from the ITT analysis set (for details see Section 3).

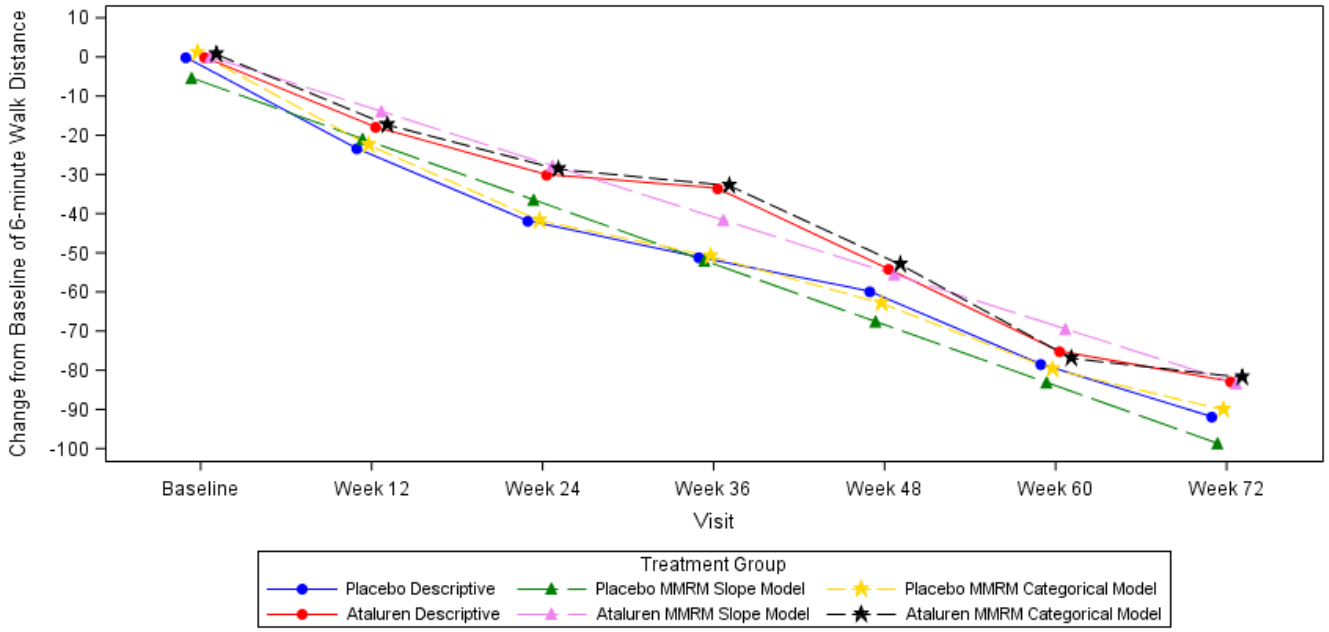
2. REASON FOR CHANGE FROM THE PLANNED ANALYSIS

2.1. Validity of Analysis Model Assumptions

Based on results from previous well controlled studies over 48 weeks the rate of decline in 6MWD was assumed to be constant (linear). Therefore, a linear regression MMRM model was selected as the primary analysis model for this study. As is common and good statistical practice and, in accordance with EMA guidance (EMEA/CHMP/295050/2013 Section 7.4), the validity of the prespecified analysis model assumptions were checked. On assessment of the data following database lock, the linear assumption that underpinned the initial MRMM model did not hold over the 72-weeks study period in study 041.

As illustrated in Figure 1 Change from Baseline in 6MWD, the rate of decline varies over time and in between study visits, see solid lines of descriptive statistics by visit. Also shown are the constant rate of change (pink and green stipulated lines) estimated from the prespecified linear regression MMRM model, these lines do not fit the descriptive statistics very well and appear to overestimate the decline for the placebo group at week 72. In order to account for the nonlinearity of the data, we then modelled the data using an MMRM that treated time on study ("visit time") as a categorical variable instead of as originally planned a continuous variable. The MMRM with study visit as a categorical variable is a standard approach to longitudinal data where change over time is not linear and is the model prespecified in the SAP for the key secondary endpoint NSAA where nonlinearity was assumed a priori. Estimates from the MMRM model with study visit as a categorical variable fit the descriptive statistics well at all visits and appear to provide a more accurate estimate of the decline over 72 weeks. No other changes were made to the originally proposed MMRM model.

Figure 1: Line Chart of Change from Baseline of 6-minute Walk Distance (6MWD): Compare MMRM Models with Descriptive Statistics (mITT Population Double-Blind Period)



Notes: MMRM = mixed-model for repeated measures. MMRM Slope Model refers to the prespecified linear regression MMRM mode. MMRM Categorical Model refers to the MMRM model with study visit as a categorical variable.
Source: Figure 14.2.1.1.1.2

Table 1 shows the estimated mean treatment difference from descriptive statistics (arithmetic mean), the prespecified linear regression MMRM model and ANCOVA (prespecified sensitivity analysis model) as well as the modified MMRM model with study time as a categorical variable.

Table 1: Estimated Treatment Difference for Change from Baseline in 6MWD at Week 72 (mITT Population – Double-blind Period)

	Estimated treatment difference (SE) (Ataluren – Placebo)	
Model	**Change from Baseline at Week 72 (meters)**	**Average rate of change from baseline in 6MWD over 72-week double-blind period (meters/week)**
Descriptive statistics	8.99	0.12
MMRM (linear regression)	9.95 (14.9)	0.14 (0.207)
ANCOVA after MI	9.10 (13.1)	0.13 (0.182)
MMRM (categorical)	8.26 (9.1)	0.11 (0.126)

SE = standard error
Sources: Table 14.2.1.1.1, Table 14.2.1.1.2, Table 14.2.1.1.18, Table 14.2.1.3.1.

Although the prespecified linear regression MMRM model gives a somewhat larger estimate of treatment difference, and the modified categorical MMRM model a somewhat smaller more conservative estimate of treatment effect, the estimates from all models are quite similar and close to the descriptive statistics.

Comparing the precision (SE) of the estimates the linear regression MMRM (which utilizes all seven data points collected for each subject during the double-blind period) does not yield the expected gain in efficiency compared to the simpler ANCOVA model which utilizes only Baseline and Week 72 observations and excludes all intermittent observations. The modified MMRM model with categorical 'visit time' yields the expected gain in efficiency with the standard error of the estimates being reduced with about 30% compared to the ANCOVA model.

In conclusion, the MMRM categorial model with visits being specified as a categorical variable, and otherwise following the exact specification of the planned linear regression MMRM model, fit the data the best and provides the most conservative and precise estimate of treatment. Therefore, the linear regression MMRM model was replaced with the categorical MMRM model for all endpoints where this model was prespecified.

3. CHANGE FROM PLANNED EFFICACY EVALUATION

3.1. Primary Efficacy Variable (6MWD)

There is no change in the primary efficacy endpoint, the mean rate of change from baseline in 6MWD over the 72-weeks of treatment. As explained in Section 2 the MMRM model was modified from a linear regression MMRM to a MMRM model with visit time as a categorical variable to account for the nonlinearity of study data.

3.1.1. Primary Analysis

The average rate of change from baseline in 6MWD over 72-weeks of treatment in the ITT population is considered the primary efficacy endpoint of interest for this study. The ITT population was specified in the original SAP as a supportive population for the primary endpoint analysis given its importance in providing information on the broad study population (N=359) encompassing the full range of age and ambulatory function. The MMRM category model with a random intercept, baseline 6MWD as a covariate, and factors of baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), baseline time to stand from supine (<5 seconds, ≥5 seconds), treatment, visit (analysis visit as a categorical variable), interaction of visit and treatment, interaction of visit and baseline 6MWD, and unstructured covariance matrix (UN), is considered the primary analysis model of interest. In case of non-convergence, other types of covariance matrix are to be used in the order of heterogenous Toeplitz, Toeplitz, and compound symmetric. For this analysis, any subject who lose ambulation during the study was assigned a 6MWD of 0 meters for the first visit in which the subjects lose ambulation and set to 0 meters for all subsequent visits while they participate in the study. The average rate of change in 6MWD over 72-weeks of treatment is calculated from the estimated change from baseline in 6MWD at Week 72 from the MMRM categorical model.

A sample code is listed as below:

```
PROC MIXED data=<dataset> method=reml;
  CLASS Cortico C6MWD TTSTBLC Subjid Trt AVISIT;
  MODEL CHG6WMD = Cortico C6WMD TTSTBLC Trt B6WMD AVISIT
                              AVISIT*B6WMD AVISIT*TRT/CL;
  RANDOM Int / subject=Subjid;
  REPEATED AVISIT / type=un subject=Subjid;
  LSMEANS TRT*AVISIT / pdiff CL;
RUN;
```

where

CHG6WMD = Change from baseline in 6WMD at each visit (Baseline to Week 72)
B6WMD = Baseline 6WMD
C6WMD = Baseline 6WMD category
Cortico = Baseline concomitant corticosteroid type
TTSTBLC = Time to stand from supine category
AVISIT = Analysis visits as categorical variable
Trt = Treatment group
Subjid = Subject

Int = Intercept

3.1.2. Supportive/Sensitivity Analyses

To confirm the robustness of the conclusions from the analyses performed on the ITT analysis set additional sensitivity analyses where added as detailed below.

3.1.2.1. Analysis of Covariance (ANCOVA) Using Multiple Imputation (MI)

The ANCOVA analysis using MI that was prespecified in SAP for the double-blind portion of Study 041 is still considered as a sensitivity analysis based on ITT population.

3.1.2.2. Checking Missing at Random Assumption

An additional sensitivity analysis is planned to check the MAR assumption underlying the MMRM categorical model where the reason for drop-out is considered based on ITT population. Subjects with following discontinuation reasons are to be imputed with MI method using subjects in placebo arm who completed double-blind period of the study without missing values.

- Adverse event
- Withdrew consent
- Investigator decision

3.1.2.3. Using Average of Day 1 and Day 2 6MWDs Rather than Maximum

The primary analysis uses the maximum of valid Day 1 and Day 2 6MWDs for baseline and Week 72 visits. As a sensitivity analysis, the same analysis is to be repeated with the average of valid Day 1 and Day 2 6MWDs for baseline and Week 72 visits for ITT population.

3.1.2.4. Assess the Impact of Coronavirus Disease of 2019 (COVID19)

To assess the impact of coronavirus disease of 2019 (COVID19), a sensitivity analysis is planned similar to the MMRM categorical analysis and ANCOVA analysis after MI by removing all COVID19 impacted visits (as collected on CRF) for ITT population.

3.1.2.5. Assess the Impact from Potential Correlation among Siblings

In this study, brothers participating in the study received the same treatment group designation to minimize potential compromise of the study drug blinding and to reduce the chance of dosing errors within families. In the primary analysis, siblings are considered as independent subjects. However, the assessments from two or more brothers may be highly correlated. To assess the impact of potential correlation among siblings, the MMRM categorical analysis are to be repeated by including only data from the first sibling for ITT population.

3.1.2.6. Non-parametric analysis

As a sensitivity analysis, the change from baseline at Week 72 in 6MWD using observed data between two treatment groups are compared using Wilcoxon rank sum test stratified by baseline concomitant corticosteroid type (deflazacort, prednisone/prednisolone), baseline 6WMD category (<300 meters, 300 to <350 meters, 350 to <400 meters, ≥400 meters), and baseline time to stand from supine (<5 seconds, ≥5 seconds) for ITT population.

3.1.2.7. Subgroup Analysis

The change from baseline in 6MWD are also summarized by visit for subgroups specified below in ITT populations for both the double-blind and overall treatment periods:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)
- Baseline 6MWD
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

3.2. Secondary Efficacy Variables

3.2.1. Key Secondary Efficacy Variables

The key secondary efficacy endpoints remain the same as the prespecified endpoints, TFTs and rNSAA, with the primary population of interest being ITT and the primary analysis model of interest being the MMRM categorial model.

3.2.1.1. Timed Function Tests (TFTs) - Composite Scores

The change in composite timed function tests (TFTs) scores is analyzed using the same MMRM categorical model as used for the primary analysis. Subjects who cannot perform a TFT within 30 seconds, including those who is LoA or the timed function test is above 30 seconds, are assigned a value of 30 seconds for the respective test for calculating the average change in times.

As a supportive analysis, the prespecified ANCOVA model with missing values being imputed using the MI method is also performed.

The change from baseline in composite TFT scores is also summarized by visit for the subgroups specified below in ITT population for both the double-blind and overall treatment periods:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)
- Baseline 6MWD
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

3.2.1.2. North Star Ambulatory Assessment - Total Scores

There is no change to the planned analyses for rNSAA. The prespecified ANCOVA after MI and MMRM categorical models are to be performed for ITT population.

The change from baseline in rNSAA total scores and linear scores are also summarized by visit for the subgroups specified below in ITT population for both the double-blind and overall treatment periods:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)
- Baseline 6MWD:
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

3.2.2. Other Secondary Efficacy Variables

The other secondary efficacy endpoints remain the same as the prespecified endpoints, with the primary population of interest being ITT and the primary analysis model of interest being the MMRM categorial model, where applicable.

3.2.2.1. Time to 10% Persistent Worsening in 6MWD

There is no change to the planned analyses for time to 10% persistent worsening in 6MWD for ITT population.

3.2.2.2. Timed Function Tests (TFTs) - Individual Test

The 3 TFTs (10-meter run/walk, 4-stair stair-climb, and 4-stair descend) are summarized and analyzed separately for each function test using the MMRM categorical model assuming MAR based on ITT population. In addition, the prespecified sensitivity analysis, ANCOVA analysis after MI is also performed for the three timed function tests as sensitivity analysis.

A separate analysis is performed to evaluate the 10-meter run/walk results in subjects with a baseline 6WMD <300 meters in ITT population.

The change from baseline in all timed function tests are also summarized by visit for the subgroups specified below in ITT population for both the double-blind and overall treatment periods:

- Region (North and Latin America, Europe, Asia Pacific)
- Stop codon (UAA, UAG, UGA)
- Baseline 6MWD:
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

3.2.2.3. North Star Ambulatory Assessment – Individual Items

There is no change to the planned analyses for loss of function in individual items of North Star Ambulatory Assessment (NSAA) for ITT population.

3.2.2.4. Loss of Ambulation (LoA) and Stair Climb/Stair Descend

There is no change to the planned analyses for Loss of Ambulation (LoA) and Loss of Stair Climb/Stair Descend for ITT population.

3.3. Exploratory Variables

3.3.1. PUL and DMD Upper Limb PROM

In addition to SAP specified analysis for PUL and PROM for ITT population, PUL and PROM are also summarized by visit for subgroups specified below in ITT population for double-blind period and overall treatment period.

- Baseline 6MWD:
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

3.3.2. Myometry Parameters

There is no change to the planned analyses for myometry parameters for ITT population.

3.3.3. Muscle Fat Fraction

In addition to SAP specified analysis for muscle fat fraction for ITT population, muscle fat fraction is also summarized by visit for subgroups specified below in ITT population for double-blind period and overall treatment period.

- Baseline 6MWD:
  - Baseline 6MWD <300 meters
  - Baseline 6MWD ≥300 meters to <400 meters
  - Baseline 6MWD ≥400 meters

3.3.4. Spirometry – Pulmonary Function Test

There is no change to the planned analyses for spirometry-pulmonary function test for ITT population.

3.3.5. HRQL At-Home Questionnaires

There is no change to the planned analyses for HRQL at-home questionnaires for ITT population.

3.4. Additional Analysis for Subjects with Baseline 6MWD between 300-400 Meters

In order to evaluate the consistency in treatment effects between Study 041 with previous studies (Studies PTC124-GD-007-DMD and PTC124-GD-020-DMD), similar analyses and summaries performed for the ITT population are repeated for the subgroup of subjects with baseline 6WMD ≥300 meters to <400 meters for the primary efficacy variable and selected secondary efficacy

variables (6MWD, 10% persistent worsening in 6MWD, TFTs, rNSAA, LoA, loss of stair climb/stair descend).

3.5. Analyses for Overall Treatment Period

Since this study is still ongoing at the time of the first database lock, more than 40% of the subjects have not completed the open-label period by the data cutoff date. Only summary statistics will be presented for all efficacy endpoints for this Study 041 CSR. Statistical modellings will be performed at the final database lock when all subjects have completed open-label period.

3.6. mITT Analysis

Since the mITT analysis did not meet its primary objective of capturing significant benefit only the results for primary efficacy endpoint, change from baseline in 6MWD over the 72-weeks double-blind period is shown in-text. All other preplanned mITT analyses for other endpoints are presented end-of-text for the purpose of this CSR.

4. REFERENCES

EMA. Guideline on adjustment for baseline covariates in clinical trials. February 26, 2015. Retrieved on 2022-07-31 at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-adjustment-baseline-covariates-clinical-trials_en.pdf

Signature Page for EU/UK PTC124-GD-041-DMD Change of Planned Statistical Analysi

Clinical Approval	[redacted] I approve the document(s) 19-Sep-2022 19:48:35 GMT+0000

Statistics Approval	[redacted] I approve the document(s) 19-Sep-2022 21:44:55 GMT+0000

Statistics Approval	[redacted] I approve the document(s) 20-Sep-2022 17:08:31 GMT+0000

Signature Page for VV-CLIN-007217 v1.0