Official Title: A Randomized, Double-Blind, Placebo controlled, Global Phase 3 Study of Edasalonexent in Pediatric Patients with Duchenne Muscular Dystrophy

NCT Number: NCT03703882

Applicant/MAH: Catabasis Pharmaceuticals Inc.

Version Date: 05 OCT 2020

STATISTICAL ANALYSIS PLAN

Protocol CAT-1004-301

A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, GLOBAL PHASE 3 STUDY OF EDASALONEXENT IN PEDIATRIC PATIENTS WITH DUCHENNE MUSCLAR DYSTROPHY

Protocol Number: **(Version Date)**	CAT-1004-301 16 APR 2020
Name of Test Drug:	Edasalonexent (CAT-1004)
Phase:	3
Methodology:	Randomized, Double-blind, Placebo-controlled
Sponsor:	Catabasis Pharmaceuticals, Inc. 100 High Street, 28th Floor Boston, MA 02110 Tel: PI Fax: PI
Sponsor Representative:	PI, M.D., Ph.D. Chief Medical Officer
Sponsor Representative:	PI, MHP Vice President, Clinical Operations
Document Date:	05 OCT 2020
Document Version:	5.0

SIGNATURE PAGE

Protocol Title: A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, GLOBAL PHASE 3 STUDY OF EDASALONEXENT IN PEDIATRIC PATIENTS WITH DUCHENNE MUSCLAR DYSTROPHY

Sponsor:
Catabasis Pharmaceuticals, Inc.
100 High Street, 28th Floor
Boston, MA 02110
Tel: PI
Fax: PI

Protocol Number: CAT-1004-301

Document Date/Version: 05 OCT 2020

Cytel, Inc. Author:
PI
Cytel, Inc.
675 Massachusetts Avenue
Cambridge, MA 02139

Sig **Signature:** PI
Email:
Date: _____

Sponsor Approval

By signing this document, I acknowledge that I have read the document and approve of the planned statistical analyses described herein. I agree that the planned statistical analyses are appropriate for this study, are in accordance with the study objectives, and are consistent with the statistical methodology described in the protocol, clinical development plan, and all applicable regulatory guidance and guidelines.

All changes to the planned analyses will be described in the clinical study report (CSR).

Sponsor Signatory:
PI, M.D., Ph.D.
Chief Medical Officer
Catabasis Pharmaceuticals, Inc.
100 High Street, 28th Floor
Boston, MA 02110
Tel: PI
Fax: PI

Signature: PI
Date: PI 5 Oct 2020

TABLE OF CONTENTS

Section **Page**

Section **Page**

LIST OF IN-TEXT TABLES

Table **Page**

LIST OF IN-TEXT FIGRURES

Figure **Page**

ABBREVIATIONS

Abbreviation	Definition
10MWT	10-meter walk/run test
ACTH	Adrenocorticotrophic hormone
AE	Adverse events
ANCOVA	Analysis of Covariance
ATC	Anatomic Therapeutic Class
BMI	Body mass index
CDC	Centers for disease control and prevention
CFB	Change from Baseline
CIs	Confidence intervals
CK	Creatine kinase
CMH	Cochran-Mantel-Haenszel
CSR	Clinical study report
DMD	Duchenne muscular dystrophy
DXA	Dual-energy X-ray absorptiometry
ECG	Electrocardiogram
FAP	Full Analysis Population
HEENT	Head, ears, eyes, nose, and throat
kg	Kilogram
LS-means	Least-squares means
MAR	Missing at random
MCMC	Markov chain Monte Carlo
MNAR	Missing not at random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed-model repeated-measures
NSAA	North Star Ambulatory Assessment
OC	Observed Case
PD	Pharmacodynamic
PK	Pharmacokinetic
PODCI	Pediatric Outcomes Data Collection Instrument
PP	Per Protocol
PT	Preferred Term
PUL	Performance of Upper Limb
RNA	ribonucleic acid
SAE	Serious Adverse Event
SAP	Statistical analysis plan
SOC	System organ class

Abbreviation	Definition
TEAE	Treatment-Emergent Adverse Event
TFT	Timed function test
WHO	World Health Organization

INTRODUCTION AND OBJECTIVES OF ANALYSIS

1. INTRODUCTION

This statistical analysis plan provides the pre-specification and details for the statistical analyses outlined within protocol CAT-1004-301 entitled "A Randomized, Double-Blind, Placebo-Controlled, Global Phase 3 Study of Edasalonexent in Pediatric Patients with Duchenne Muscular Dystrophy" Protocol Amendment 1 dated 16APR2020.

1.1. Objectives of Statistical Analysis

The primary objective of this Phase 3 study is to assess the efficacy of edasalonexent as measured by change from Baseline (CFB) on the North Star Ambulatory Assessment (NSAA) Total Score in pediatric patients with Duchenne muscular dystrophy (DMD).

Secondary objectives include assessing the safety and tolerability of edasalonexent and assessing the effects of edasalonexent on physical function as measured by the 10-meter walk/run test (10MWT), time to stand from supine, and the 4-stair climb in pediatric patients with DMD.

Additional efficacy assessments will include muscle strength testing, Performance of Upper Limb (PUL) entry item assessments, and the Pediatric Outcomes Data Collection Instrument (PODCI) Transfer and Basic Mobility scale.

This statistical analysis plan (SAP) is designed to outline the methods to be used in the analysis of study data in order to answer these study objectives. Populations for analysis, data handling rules, statistical methods, and formats for data presentation are provided. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results sections of the clinical study report (CSR) for this trial.

This SAP also will outline any differences in the currently planned analytical objectives relative to those planned in the study protocol.

1.2. Summary of Changes

1.2.1. Summary of SAP Version 2 Changes from Version 1

The following is a summary of the changes made in version 2.0 of the SAP (dated 30 JAN 2020) relative to version 1.0 (dated 24 OCT 2019).

- Based on FDA feedback the efficacy endpoint: CFB multivariate rank sum score at Week 52 was changed from a secondary efficacy endpoint to an additional efficacy endpoint. The method for controlling for multiplicity for the primary and secondary efficacy endpoints was updated accordingly.
- Methods for addressing mis-categorization in the randomization strata were included.

1.2.2. Summary of SAP Version 3 Changes from Version 2

The following is a summary of the changes made in version 3.0 of the SAP (dated 31AUG 2020) relative to version 2.0 (dated 30 JAN 2020).

- The algorithms for estimating treatment exposure and compliance were updated and clarified.
- Methods for assessing the impact of COVID-19 were included.
- The method for calculating the total NSAA score when individual NSAA items are missing for reasons unrelated to patient's abilities to complete measures was updated.
- Additional analyses of the NSAA endpoint were included.
- The muscle strength testing algorithm was updated.
- Additional summaries of adverse events were included.
- Additional details were added to the protocol deviations section.
- The CFB multivariate rank sum score exploratory endpoint was modified to only include patients with non-missing Week 52 CFB scores for all 5 of the component endpoints and the analysis was clarified that this is a Week 52 endpoint only. The corresponding analysis was modified to an analysis of covariance (ANCOVA).
- The PODCI analyses was changed from a mixed-model repeated measures (MMRM) to an ANCOVA as it is only assessed at Week 52 post-baseline. A sensitivity analysis for the PODCI endpoint was added.
- Additional minor updates and clarifications were also performed.

1.2.3. Summary of SAP Version 4 Changes from Version 3

The following is a summary of the changes made in version 4.0 of the SAP (dated 28SEP2020) relative to version 3.0 (dated 31AUG 2020).

- Updates to the COVID-19 listings
- Minor updates and edits for clarity.

1.2.4. Summary of SAP Version 5 Changes from Version 4

The following is a summary of the changes made in version 5.0 of the SAP (dated 05OCT2020) relative to version 4.0 (dated 28SEP2020).

- Updates to the Per Protocol Population definition to address study drug compliance assessment issues due to COVID-19 pandemic.

2. STUDY DESIGN

2.1. Synopsis of Study Design

This is a randomized, double-blind, placebo-controlled, Phase 3, multisite study of edasalonexent in pediatric patients with DMD.

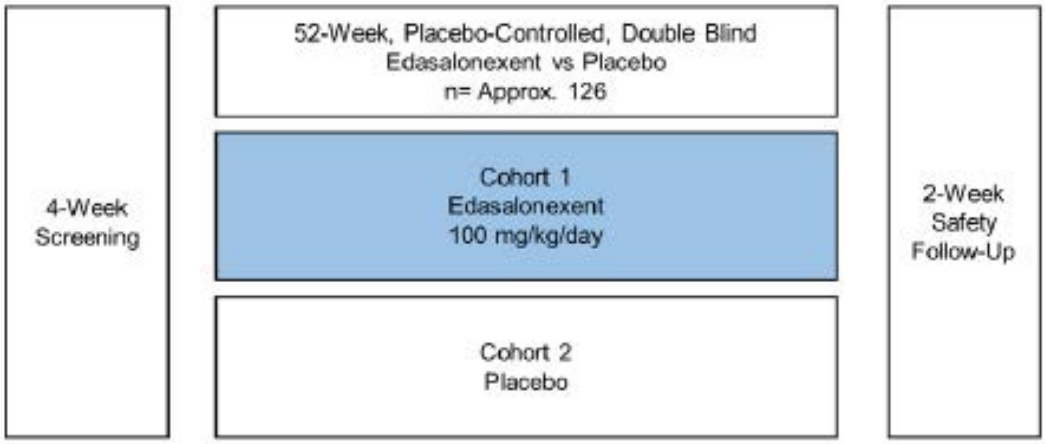
Eligible patients will be ambulatory boys aged >=4.0 to <8.0 years with a confirmed diagnosis of DMD who have not use steroids within 24 weeks prior to treatment start (Day 1). In regions where eteplirsen is approved or there is an established access program, patients who have received at least 24 weeks of treatment with eteplirsen prior to Day 1 will be eligible.

As shown in Figure 1, following a 4-week screening period to determine eligibility, approximately 126 patients will be randomized 2:1 to receive edasalonexent 100 mg/kg/day or placebo in a double-blind fashion daily for 52 weeks. Randomization will be stratified by baseline age (<=6.0 vs. >6.0 years), baseline time to stand from supine (<=5 vs. >5 seconds), treatment with eteplirsen (yes vs. no), and region (North America vs. Europe/Asia/Australia). Note that if siblings are enrolled, they will be assigned to the same treatment group.

Randomized patients (and enrolled siblings who are not randomized) will complete Baseline procedures and begin receiving study drug (edasalonexent or placebo) on Day 1. During the treatment period, patients will have site visits approximately every 13 weeks with site calls in between as specified in the Schedule of Assessments (Table 1). For those patients who do not enter the extension study, a safety follow-up should occur within 2 weeks of the last dose of study drug for collection of adverse events (AEs).

The length of a patient's participation in the study will be from the time of informed consent through the Week 52 visit. A patient will be considered to have completed the study when they have completed the Week 52 assessments. Patients who do not participate in the open-label extension study will be considered to have completed the study after they complete the safety follow-up.

Figure 1 Overview of Study Design



2.2. Randomization Methodology

Patients will be randomized 2:1 to receive edasalonexent or placebo in a double-blind fashion. Randomization will be stratified by Baseline age (<=6.0 vs. >6.0 years), baseline time to stand

from supine (<=5 vs. >5 seconds), treatment with eteplirsen (yes vs. no), and region (North America vs. Europe/Asia/Australia). Note that if siblings are enrolled, they will be assigned to the same treatment group.

Any mis-categorization of a patient's randomization strata (i.e., a patient being assigned to an incorrect randomization strata relative to their actual study data) will be presented in a listing. Analyses by strata (i.e., see Section 4.10.8) will be based on the randomization strata values.

2.3. Unblinding

After completion of the study, patients, monitors, Investigators, study center personnel, and the Sponsor will remain blinded to treatment assignments until results from the study are available.

2.4. Study Procedures

The schedule of assessments, as outlined in the study protocol, is provided in Table 1.

Table 1: Schedule of Assessments

	Screen [a]	BL [b]	Double Blind, Placebo-Controlled Treatment (Weeks 1 - 52)									Safety F/U [c]
Study Week(s)	**-4 - 0**	**0**	**3**	**6**	**13**	**19**	**26**	**32**	**39**	**45**	**52/ET [d]**	
Informed Consent/Assent	X											
Site Visit [e]	X	X			X		X		X		X	
Site Call [e,f]			X	X		X		X		X		X
Med History/Demographics	X											
Confirm Eligibility	X	X										
Randomization [g]		X										
Vital Signs	X	X			X		X		X		X	
Weight	X	X			X		X		X		X	
Height		X					X				X	
Physical Examination	X	X			X		X		X		X	
AEs/Con Meds/Change in Function [h]	X	*Continuous*										
12-Lead ECG [i]	X						X				X	
Cardiac Monitoring [j]		X					X				X	
Hematology/Chemistry [k]	X	X			X		X		X		X	
ACTH/Cortisol [k]		X			X		X		X		X	
Urinalysis	X						X				X	
DNA Sample [l]		X										
Biomarkers, Gene Expression [k]		X			X		X		X		X	
NSAA, Stand from Supine, 10MWT, 4-Stair Climb, Muscle Strength, PUL [l]	X	X			X		X		X		X	
Videotape Functional Assessments [m]	X	X									X	
PODCI [n]		X									X	
Lateral Spinal Radiography, DXA [o]		X									X	
Nutritional Guidance [p]		X										
Drug Dispensation		X			X		X		X			
Drug Accountability			X		X		X		X		X	
Plasma PK [q]		X			X		X				X	

a Per Investigator's discretion, informed consent, medical history/demographic data and lab assessments may be completed remotely by a qualified medical professional.

b The first day of study drug is "Day 1"; unless otherwise specified, all assessments should be completed within a 4 week window prior to Day 1.
c Safety follow-up should occur within 2 weeks after the last dose of study drug. Patients continuing into the open-label extension study are not required to perform the safety follow-up.
d If patient discontinues treatment early, patient will be asked to return for early termination (ET) visit prior to their last dose of study drug. Functional assessments performed within 6 weeks of discontinuing treatment do not need to be repeated at the ET visit. If the ET visit occurs after drug has been discontinued, certain assessments may be omitted per Sponsor discretion.
e Weeks 6-19 should be conducted within the study week with a ±4-day window. Weeks 26-54 have a ±10-day window.
f Per Investigator discretion, site calls may occur by telephone or use of other technology.
g Randomization occurs after Screening assessments have been completed and eligibility is confirmed.
h AEs and concomitant medication/therapy will be assessed at all visits and during site calls.
i ECGs will be performed predose at all timepoints. At Week 26, ECG will per performed predose and 2 hours postdose (+/-15 min).
j Cardiac monitoring will be performed for approximately 48 hours remotely via a wearable cardiac monitoring device if approved for use in the participating country.
k Blood samples will be drawn in the morning when possible.
l A DNA sample may be collected at Baseline for drug metabolizing enzyme and/or transporter analysis as well as pharmacogenomic analysis of potential markers of clinical risk/benefit.
m These assessments should be performed in the order listed and in the morning when possible. Time to stand from supine and 10MWT will be recorded as measured in the NSAA.
n The PUL 6-point entry item scale will be performed at all timepoints. If a patient scores a 6 on the PUL entry item scale, no further PUL testing is required for the visit.
o Videotaping at Baseline is required if Screening video did not pass QC.
p PODCI will be performed when feasible.
q These assessments should be collected within 2 months prior to Day 1. In Germany, bone assessment will not be performed until the respective approvals from the German Federal Office for Radiation Protection (Bundesamt für Strahlenschutz) become available.
r Patient compliance with the nutritional guidance will be confirmed throughout the trial.
s Urine PK collection will be performed at Week 13. Plasma PK collection will be performed predose at Week 13 and Week 52. At Week 26 plasma PK collection will be performed predose, 1 hour, 2 hours (+/-15 min), and 4 hours postdose.

2.5. Efficacy Estimands

The estimands corresponding to the primary and secondary efficacy objectives are provided in Sections 2.5.1 and 2.5.2, respectively.

2.5.1. Primary Objective Estimand

The estimand corresponding to the primary objective for this study is shown in Table 2.

Table 2: Estimand for the Primary Objective

Objective	**Efficacy of edasalonexent - NSAA Total Score**
Estimand	The expected effect in patients if all patients had adhered to the treatment regimen and did not take any corticosteroid therapy or discontinue treatment.
Population	Full Analysis Population (FAP)
Variable	Change from baseline at Week 52 on NSAA Total Score
Intercurrent Events	Treatment discontinuation or initiating corticosteroid therapy or any approved Duchenne therapy: The study is interested in the treatment effect assuming that patients remained on their treatment regimen throughout the study, without initiating corticosteroid therapy or any approved Duchenne therapy.
Population-level Summary	The population summary for the primary estimand is the mean difference in the Week 52 change from baseline on the North Star Ambulatory Assessment Total Score between the edasalonexent and the placebo treatment groups using the model described in Section 4.10.1. This summary is computed using the FAP as defined in Section 3.1

2.5.2. Secondary Objective Estimands

The estimands corresponding to the secondary objectives are shown in Table 3.

Table 3: Estimands for the Secondary Objectives

Objective	**Efficacy of edasalonexent - 10MWT**
Population	FAP
Variable	Change from baseline at Week 52 on the 10MWT speed.
Intercurrent Events	Treatment discontinuation or initiating corticosteroid therapy or any approved Duchenne therapy: The study is interested in the treatment effect assuming that patients remained on their treatment regimen throughout the study, without initiating corticosteroid therapy or any approved Duchenne therapy.

Population-level Summary	The population summary for the estimand is the mean difference in the Week 52 change from baseline on the 10MWT speed between the edasalonexent and the placebo treatment groups. This summary is computed using the FAP as defined in Section 3.1.
Objective	**Efficacy of edasalonexent - Stand from Supine Speed**
Population	FAP
Variable	Change from baseline at Week 52 on the stand from supine speed.
Intercurrent Events	Treatment discontinuation or initiating corticosteroid therapy or any approved Duchenne therapy: The study is interested in the treatment effect assuming that patients remained on their treatment regimen throughout the study, without initiating corticosteroid therapy or any approved Duchenne therapy.
Population-level Summary	The population summary for the estimand is the mean difference in the Week 52 change from baseline on the stand from supine speed between the edasalonexent and the placebo treatment groups. This summary is computed using the FAP as defined in Section 3.1.
Objective	**Efficacy of edasalonexent - 4-Stair Climb**
Population	FAP
Variable	Change from baseline at Week 52 on the 4-Stair climb speed.
Intercurrent Events	Treatment Discontinuation or initiating corticosteroid therapy or any approved Duchenne therapy: The study is interested in the treatment effect assuming that patients remained on their treatment regimen throughout the study, without initiating corticosteroid therapy or any approved Duchenne therapy.
Population-level Summary	The population summary for the estimand is the mean difference in the Week 52 change from baseline on the 4-Stair Climb speed between the edasalonexent and the placebo treatment groups. This summary is computed using the FAP as defined in Section 3.1.

2.6. Efficacy Endpoints

2.6.1. Primary Efficacy Endpoint

The primary efficacy endpoint is the CFB in the NSAA Total Score at Week 52.

The NSAA is a clinician-reported outcome instrument designed to measure ambulatory function in males with DMD. During this assessment, patients are asked to perform 17 different functional activities, including a 10MWT, rising from sit to stand, standing on one leg, climbing and descending a step, stand from supine, lifting the head, standing on heels, and jumping. To optimize standardization and consistency of testing, study procedures must be performed by trained evaluators only, with every effort made to have the same clinical evaluator perform all assessments for a given patient, in the same order, at the same time of day, preferably in the morning. Completion of this test is expected to take approximately 20-30 minutes.

Each function activity will be scored as 0 (unable to achieve independently), 1 (modified method but achieves goal independent of physical assistance from another), or 2 (no obvious modification of activity) or "Not Scored". If the NSAA test was performed and any of the individual items are scored as "not scored" (i.e., for reasons unrelated to the patients physical capabilities), the corresponding total score will be set to missing. The sum of these 17 scores will be used to form an ordinal total score (range 0 – 34).

2.6.2. Secondary Efficacy Endpoints

The 3 secondary efficacy endpoints are:

- The CFB in the 10MWT speed at Week 52
- The CFB in the stand from supine speed at Week 52
- The CFB in the 4-stair climb speed at Week 52

The 10MWT measures time to walk a distance of 10 meters. The stand from supine and 4-stair climb tests measure the amount of time taken to complete those tasks. To optimize standardization and consistency of testing, study procedures must be performed by trained evaluators only, with every effort made to have the same clinical evaluator perform all assessments for a given patient, in the same order, at the same time of day, preferably in the morning. Completion of timed function tests (TFTs) is expected to take approximately 10 minutes.

For all TFTs, both the times to complete (seconds) and the speed of completing the tests (i.e., calculated as the reciprocals of the time to complete; speed=1/time) will be provided. If there are any patients not able to complete the test (reason not done = "subject physically unable to complete test"), a value of 0 will be assigned for the reciprocal test (i.e., speed; in units of task per second). For the "time to complete" endpoints, a value of 12 seconds will be imputed.

Additionally, for the 3 TFTs, the time will be set to 12 seconds and the speed to 0 if the TFT assessment meets the following TFT grading criteria (i.e., each of the TFTs is "graded"):

- 10MWT: a Grade of 1 or 2 (from a 6-point scale)

  1=Unable to walk independently

  2=Unable to walk independently but can walk with knee-ankle foot orthoses or support from a person
- Stand from Supine: a Grade of 1 or 2 (from a 6-point scale)

  1 = Unable to stand from supine, even with use of a chair

  2 = Assisted Gowers – requires furniture for assist in arising from supine to full upright posture (no time to be recorded)
- 4-Stair Climb: a Grade of 1 (from a 6-point scale)

  1 = Unable to climb 4 standard stairs (no time recorded)

2.6.3. Additional Efficacy Endpoints

Additional efficacy endpoints include CFB on the following:

- Muscle strength testing
- PUL entry item assessment

- PODCI Transfer and Basic Mobility Scale
- CFB multivariate rank sum score

2.6.3.1. Muscle Strength Testing

Muscle strength testing will include assessment of knee extension and elbow flexion (in kilograms), and will be performed using a hand-held dynamometer. Knee myometry will be performed while the patient is in a sitting position. Elbow myometry will be performed while the patient is positioned supine.

Muscle testing will be performed by trained evaluators; every effort will be made to have the same evaluator perform all assessments for a given patient at the same time of day. Completion of each muscle test is expected to take approximately 10-15 minutes, for a total of 20-30 minutes.

Multiple assessments are taken on the right and left knee and right and left elbow. The following is the algorithm for determining the patients elbow (knee) scores:

- For both the left and right elbow (separately) the peak score (i.e., maximum value) will be used as the value for the patient.
- The average of the maximum right elbow score and maximum left elbow score will be the elbow score for that visit.
- The same algorithm is used for the knee

As such, at each scheduled timepoint, there will be an elbow flexion and knee flexion endpoint for the muscle testing (units of kilogram (kg)).

2.6.3.2. Performance of Upper Limb (PUL) Entry Item Assessment

The PUL (PUL 2.0; Pane (2018)) was specifically designed to assess upper limb function in ambulant and non-ambulant DMD patients. The PUL is made up of a 6-point entry item assessment to define starting functional level, and 22 items subdivided into shoulder level (6 items), middle level (9 items), and distal level (7 items) dimension. Loss of full overhead reach occurs at median age 9.6 years with some boys losing the ability as early as age 5 or 6.

Patients will perform the PUL entry item assessment; based on the age of patients in this study, most patients are expected to score a 6 on the scale of 0 – 6 and will not be required to perform the additional items.

Patients who score below a 6 on the entry item assessment will perform a subset of the PUL battery including the shoulder level dimension and supination, consisting of 7 items with a maximum score of 14. Completion of the 6-point entry item assessment is expected to take approximately 5 minutes. Completion of the battery of PUL assessments is expected to take approximately 20 minutes. Any missing items on the entry item assessment or the subset of the PUL battery will be considered a zero for the purposes of deriving the overall score.

The following endpoints will be analyzed from the PUL assessments

- Item 1 and Item 2 (Summed): this is the sum of Items 1 and 2:
    - Item 1: Shoulder abduction both arms above head “Raise your arms out to the side and above your head– try and keep straight elbows” a 0 – 2 point scale

  - Item 2: Raise both arms to shoulder height (elbows at shoulder height) "Raise your arms to shoulder level" a 0 – 2 point scale

- PUL Total Score: The PUL total score is a sum of items 1 -7. Each of the 7 items has a score range from 0 – 2 with a higher score indicating better function. The PUL total score ranges from 0 – 14 with a higher score indicating better functioning.

For patients who score a "6" on Item A (and are not required to perform PUL items 1 – 7), the following data imputation will be performed:

- Items 1 and 2 be imputed with a value of 2
- Items 3, 4, 5, 6, 7 will not be imputed, items will remain missing.
- The PUL Total Score will be not be imputed, the score will be missing.

Due to this missing data issue, the analysis of the PUL Total Score will be performed on the subset of FAP patients who did not score a "6" on Item A. The analysis of Items 1 and 2 (i.e., where scores are imputed for patients who score a "6" on Item A) will be performed on the FAP.

2.6.3.3. Pediatric Outcomes Data Collection Instrument (PODCI)

The PODCI is an assessment of health-related functioning in children, and is designed to be completed by the parent/caregiver of children aged 10 or younger (Henricson, 2013). It contains 86 questions, takes approximately 20-30 minutes to complete, and provides information on the following 6 scales:

- Upper Extremity and Physical Function Scale – measures difficulty encountered in performing daily personal care and student activities
- Transfer and Basic Mobility Scale – measures difficulty experienced in performing routine motion and motor activities in daily activities
- Sports and Physical Functioning Scale – measures difficulty or limitations encountered in participating in more active activities or sports
- Pain/Comfort Scale – measures the level of pain experienced during the past week
- Happiness Scale – measures the overall satisfaction with personal looks and sense of similarity to friends and others of own age
- Global Functioning Scale – a general combined scale calculated from the first 4 scales listed above.

For each of the first five scales, the average of the scores in the questions is calculated and standardized on a 0 to 100 scale, where higher numbers represent a positive value. The global functioning scale is an average of the first four subscales (i.e., excluding the "Happiness Scale"). The details of the PODCI scoring are provided in Appendix 3: PODCI Scoring Algorithm.

2.6.3.4. Multivariate Rank Sum Score Endpoint

This multivariate rank sum score endpoint is designed to assess the effect of edasalonexent versus placebo across multiple endpoints using a non-parametric rank-based approach. For Week 52, the endpoint is the sum of the ranks of the following Week 52 CFB scores: NSAA,

10MWT speed, stand from supine speed, 4-stair climb speed, and PODCI transfer and basic mobility scale.

The multivariate rank sum score at Week 52 will be derived using the method described in O'Brien (1984). Specifically, for each patient *i*, the overall rank score is given by:

$$R_{ii} = r_{ij1} + r_{ij2} + r_{ij3} + r_{ij4} + r_{ij5}$$

where:

i = an indicator for a specific patient

j = an indicator for a treatment group

r_{ij1} = patient *i*'s rank of the Week 52 CFB in NSAA score at Week 52

r_{ij2} = patient *i*'s rank of the Week 52 CFB in 10MWT speed score at Week 52

r_{ij3} = patient *i*'s rank of the Week 52 CFB in Stand from Supine Speed score at Week 52

r_{ij4} = patient *i*'s rank of the Week 52 CFB in 4-stair Climb Speed at Week 52

r_{ij5} = patient *i*'s rank of the Week 52 CFB in the PODCI Transfer and Basic Mobility Scale score at Week 52 (described in Section 2.6.3.3)

In determining ranks, if there are ties, the average value of the ranks that would have been assigned to all observations with the same value will be used.

Only patients with non-missing values for each of these 5 Week 52 CFB scores will be included in the analysis.

The patient's baseline rank sum score will be used a covariate for some of the rank-based analyses. This is defined as the sum of the ranks across the 5 endpoints using the baseline value (using the methodology described above).

2.6.4. Safety Endpoints

Safety will be evaluated in terms of all treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs), as well as physical examination, growth parameters (height and weight), vital signs, clinical laboratory parameters (including chemistry, hematology and urinalysis), adrenal function (adrenocorticotrophic hormone (ACTH) and cortisol levels), and electrocardiogram (ECG). Data from ECGs will be locally read and centrally evaluated. Cardiac monitoring and bone evaluations will also be performed.

2.6.5. Additional Endpoints

To access acceptability and palatability of edasalonexent, caregivers will complete a study drug questionnaire, these results will be presented in a listing and the following questions will be summarized by treatment group.

1) The taste and aftertaste (a taste that remains in your mouth after swallowing) of the study drug capsules are acceptable.
2) The smell of the study drug capsules is acceptable.
3) The size and shape of the study drug capsules are acceptable:

These 3 questions are measured on a 1 (strongly agree) to 5 (strongly disagree) point scale. The frequency of each response category will be summarized.

3. PATIENT POPULATIONS

3.1. Population Definitions

The following patient populations will be evaluated and used for presentation and analysis of the data:

- Randomized Population: all patients who are randomized into the study with patients to be analyzed based on the study treatment they were randomized to.
- Safety Population: all patients who receive at least one dose of study drug with patients to be analyzed based on the actual study treatment received.
- Full Analysis Population (FAP): all patients in the Randomized Population who receive at least one dose of study drug and provide at least one non-missing post-Baseline NSAA efficacy assessment.
    - Note that the FAP will exclude patients on eteplirsen at randomization and patients will be analyzed based on the study treatment they were randomized to.
- Per Protocol (PP) Population: all patients in the FAP who complete the 52-week treatment period of the study without any significant protocol deviations with patients to be analyzed based on the study treatment they were randomized to.

  The set of patients in the Per Protocol population may be updated due to study drug compliance information which will not be collected until after database lock and unblinding due to the COVID-19 pandemic, see Section 3.2.

  note: the variable to determine a patient completing the 52-week treatment period is from the End of Study eCRF page “Did subject complete the study?” = Yes.

The FAP will be the primary population for the analysis of efficacy parameters. The Safety Population is the primary population for the analysis of safety endpoints.

Note that if siblings are enrolled, only one sibling will be included in the Randomized Population. This is to avoid having siblings being randomized to different study medications, increasing the risk a subject will inadvertently take incorrect study medication. The second sibling will be assigned to the same treatment group as the randomized sibling. If both siblings meet the inclusion/exclusion criteria, the sibling to be included in the Randomized Population will be the first sibling randomized, both siblings will be included in the Safety Population (see Section 3.1). The sibling not included in the Randomized Population will have their data presented in the listings and will be included in some of the sensitivity analyses.

3.2. Protocol Deviations

At the discretion of the sponsor, significant protocol deviations as determined by a review of the data prior to unblinding of the study may result in the removal of a patient’s data from the Per Protocol Population. The sponsor, or designee, will be responsible for producing the final protocol deviation file (formatted as a Microsoft Excel file), in collaboration with Cytel and the data monitoring group as applicable; this file will include a description of the protocol deviation, and clearly identify whether or not this deviation warrants exclusion from the Per Protocol Population. This file will be finalized prior to database unblinding.

All protocol deviations will be summarized by category and presented in the data listings. The following are the protocol deviation categories:

- Eligibility
- Inclusion/Exclusion criteria
- Informed consent deviation
- Laboratory
- Other
- Outside protocol window
- Procedure or labs not done
- Restricted concomitant medication change
- Study drug
- Study drug dosing
- Study procedures
- Subject non compliance
- Visit not done
- Visit window

Note: Due to the impact of the COVID-19 pandemic on study sites and patients, drug accountability may be incomplete for some patients at the time of database lock and the timeframe for resolution may not be near term or indeed possible. This impacts the assessment of study drug compliance. If, after database lock and unblinding, additional drug accountability information is provided such that a patient's overall study drug compliance is determined to be less than 80%, then the Per Protocol population will be updated to exclude this patient (i.e., if the patient has not previously been excluded from the Per Protocol population). This set of patients (i.e., patients with less than 80% study drug compliance identified after database lock and unblinding) will be provided in the CSR.

4. STATISTICAL METHODS

4.1. Sample Size Justification

Approximately 126 patients will be randomized to either edasalonexent or placebo in a 2:1 ratio. Based on Phase 1/2 data of changes in NSAA in this age group, an effect size of 0.625 is assumed. With this effect size, and assuming a dropout rate of approximately 20%, the study has approximately 80% power to show a difference between the treatment groups at a 2-sided type I error rate of 0.05.

Based on an underlying standard deviation estimate of 4.8 in the change from baseline in NSAA score (derived from the phase 2 study data) this would equate to a treatment difference of 3.0, which would be considered a clinically meaningful change.

4.2. General Statistical Methods and Data Handling

4.2.1. General Methods

Tabulations will be produced for appropriate demographic, Baseline, efficacy, and safety parameters. For categorical variables, summary tabulations of the number and percentage within each category (with a category for missing data) of the parameter will be presented. For continuous variables, the mean, median, standard deviation, minimum and maximum values will be presented. Unless otherwise defined for a particular endpoint, a patient's Baseline value is defined as the last non-missing measurement prior to the initiation of study drug. The first day of study drug is "Day 1".

Formal statistical hypothesis testing will be performed on the primary and secondary efficacy endpoints (Section 2.5.1), with all tests conducted at the 2-sided, 0.05 level of significance. Summary statistics will be presented, as well as confidence intervals (CIs) on selected parameters, as described in the sections below.

4.2.2. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software (Version 9.4), unless otherwise noted. Medical History and adverse events will be coding using the current Medical Dictionary for Regulatory Activities (MedDRA) version. Concomitant medications will be coded using the most current version of the World Health Organization (WHO) Drug Reference List.

4.2.3. Methods of Pooling Data

Not applicable to the present study.

4.2.4. Adjustments for Covariates

Statistical models will adjust for the stratification variables as covariates.

4.2.5. Multiplicity

The following multiplicity adjustment approach will be used to strongly control the overall Type I error rate at 0.05 (two-sided) for the primary endpoint and three secondary endpoints.

A hierarchical testing procedure will be used for the secondary endpoints. At each step the test for treatment effect will be considered statistically significant if the p-value is less than or equal

to 0.05 and all previous tests also meet this level of significance. The testing hierarchy is as follows:

1. Primary endpoint: CFB in the NSAA Total Score at Week 52
2. Secondary endpoint: CFB in the Stand from Supine Speed at Week 52
3. Secondary endpoint: CFB in the 10MWT Speed at Week 52
4. Secondary endpoint: CFB in the 4-Stair Climb Speed at Week 52

This method provides strong control of the family-wise error rate at the 2-sided 0.05 level.

4.2.6. Subgroups

Primary and secondary efficacy analyses will be performed within each of the following subgroups based on the FAP:

- Baseline Age (<= 6.0 years old vs. > 6.0 years old)
- Baseline Time to Stand from Supine (<=5 seconds vs. >5 seconds)
- Region (North America vs. Europe/Asia/Australia)

4.2.7. Withdrawals, Dropouts, Loss to Follow-up

Patients who withdrew from the study will not be replaced.

4.2.8. Missing, Unused, and Spurious Data

For the primary and secondary efficacy endpoint analyses, an MMRM using observed case analysis represents the primary method for handling missing data (as described in Sections 4.10.1 and 4.10.2). Multiple sensitivity analyses on the primary and secondary endpoints which utilize alternative methods of handling missing data are being performed as described in Sections 4.10.4 and 4.10.5.

Additionally, the following missing data conventions are being applied:

- For all TFTs, as described in Section 2.6.2, If there are any patients not able to complete the test (reason not done = “subject physically unable to complete test”), a value of 0 will be assigned for the speed of the TFT (i.e., speed in units of task per second). For the time to complete endpoints, a value of 12 seconds will be imputed.

  Additionally, as described in Section 2.6.2, if the TFT has a grade of 1 or 2 for 10MWT or stand from supine, or a grade of 1 for the 4-stair climb, the speed will be set to 0 and the time set to 12 seconds.
- As described in Section 3.1, if siblings are enrolled, only one sibling will be included in the Randomized Population. Sensitivity analyses for all efficacy variables will be performed with the both siblings’ values to be used (as described in Sections 4.10.1 and 4.10.2).

4.2.9. Visit Windows

Screening takes place in the 4-week interval prior to first day of study drug (Day 1). Week 6 – 19 visits should be conducted with the study week with a ±4-day window. Week 26 – 54 visits have a ±10-day window.

Safety follow-up should occur within 2 weeks after the last dose of study drug.

However, visits which occur outside the windows will still be included in the analysis for that visit.

4.2.10. Early Termination Visit

If a patient has an Early Termination visit, the assessments from that visit will be mapped to a scheduled visit using the algorithm presented in Table 4. Note that if the subject has already had the corresponding scheduled visit, the early termination visit data will not be included in summary statistics, but will be included in listings. The listings will show the visit as "Early Termination", not the corresponding mapped visit.

Table 4: Early Termination Visit Window Algorithm

Study Day of Early Termination Visit	Visit mapping
2 – 136	Week-13 Visit
137 – 227	Week-26 Visit
228 – 318	Week-39 Visit
319 – 364	Week-52 Visit

note: The first day of study drug is "Day 1".

4.2.11. Baseline

A patient's baseline value will be most recent non-missing value occurring prior to initiation of study drug. In most cases this will be the value assessed at the Baseline visit. If the Baseline value is missing, the next most recent non-missing value will be used (i.e., in most cases this will be the Screening visit value).

4.3. Interim Analyses

No interim analyses were planned for this study.

4.4. Patient Disposition

All patients who sign the informed consent will be included in a summary of patient disposition. The number of patients screened, randomized, and dosed in each treatment group will be presented. The number and percentage of patients discontinuing and completing treatment, the number of patients discontinuing before study completion (including reason for discontinuation), and the number of patients completing the study will be presented as well.

The number and percentage of patients in the various analysis populations will be presented separately.

The above information will also be provided in a by-patient listing.

4.5. Demographic and Baseline Characteristics

Demographics, including age at Baseline are considered as a continuous variable, frequency of age by year, age at Baseline (<=6.0 vs. >6.0 years) and Baseline characteristics, including time to stand from supine (both as a continuous and dichotomous variable: <=5 vs. >5 seconds), treatment with eteplirsen (yes vs. no), and region (North America vs. Europe/Asia/Australia) will be summarized by treatment group and overall, and also listed by patient for the Safety

Population. Baseline NSAA score will also be summarized. No inferential statistical comparisons between groups will be performed.

Note that age (in years) will be derived to the tenths place from the date of birth to the date of randomization. For the patients who do not have a date of birth provided (e.g., patients from Germany) the baseline age will be calculated via the years and months old provided at screening plus the time from the screening visit to the date of randomization.

4.6. Concomitant Medications

Concomitant medications will be listed by treatment and coded using the most current version of the WHO Drug Reference List.

4.7. Medical History

Medical History will be listed and coded by patient using the MedDRA version 21.0.

4.8. DMD-Specific Medical History

DMD-specific medical history, including age at symptom onset and diagnosis, will be listed by patient, and summarized by treatment group and overall for the Safety Population.

4.9. Treatment Exposure and Compliance

Throughout the study, patients and their parents/caregivers will complete a diary documenting study drug administration and diet compliance. For the Safety Population, treatment duration, number of administrations, and compliance (summarized by assessing capsule count) will be summarized by treatment group and overall.

Treatment Duration: Treatment Duration will be the number of days from the patient's first dose until their Week 52 visit. If the patient discontinued from the study and did not have a Week 52 visit the subject's Early Termination visit will be used (note: if the subject does not have an early termination visit, the last visit will be used).

The patient's date of first dose is determine by the study drug administration at baseline:

- If for the baseline visit the "first dose of the day administered in clinic" = "No" (i.e., variable EX. EXPERF= "N") then the date of first dose is the date in the field EX.EXSTDAT (i.e., the "Date of study drug administration).
- Otherwise the date of first dose is the date of the baseline visit.

Expected Average Daily Dose: The expected number of capsules that a patient is to take each day is specified at each visit that the patient receives new study medication (e.g., Baseline, Week 13, Week 26, and Week 39, but there may be distribution at unscheduled visits as well). The daily total of capsules is broken up into morning/afternoon/evening administration. The exact capsule count per dose and the capsule dose strength (i.e., 250 mg or 100 mg) will be provided through the IXRS based on the patient weight at each study visit.

The periods are:

- 1st day patient takes study medication to the day before the Week 13 visit
- Week 13 to the day before the Week 26 visit
- Week 26 to the day before the Week 39 visit

- Week 39 to the day of the Week 52 visit.

Please note the following potential modifications to these periods:

- If the patient discontinues the study early, the corresponding period will end at the date of the early termination visit.
- If the patient has an unscheduled visit where study drug is dispensed the periods will be updated to account for this visit. The following is an example of the update to the algorithm where the patient had an unscheduled visit between the Week 26 and Week 39 visits.
    - 1st day patient takes study medication to the day before the Week 13 visit
    - Week 13 to the day before the Week 26 visit
    - Week 26 to the day before the Unscheduled Visit where study drug was dispensed.
    - Unscheduled Visit where study drug was dispensed to the day before the Week 39 visit.
    - Week 39 to the day of the Week 52 visit.

For each study drug administration period, the daily dose (mg/day) is calculated as:

Daily Dose (mg/day) = expected number of capsules per day × capsule strength

(note: Capsule strength is either 250mg or 100mg).

The average daily dose is calculated by taking the average across the patient's different dosing periods weighted by the number of days in the period.

As such the algorithm is:

Average daily dose (mg/day) =

{Daily dose for 1st day patient takes study medication to the day before Week 13 Period × # of days in this period +

Daily dose for Week 13 to the day before the Week 26 visit × # of days in this period +

Daily dose for Week 26 to the day before the Week 39 visit × # of days in this period +

Daily dose for Week 39 to the day before the Week 52 visit × # of days in this period}

/ {Total # of days summed across these periods}.

Please note the following potential modifications to the algorithm:

- If the patient discontinues the study early, the corresponding period will end at the date of the early termination visit.
- If the patient has an unscheduled visit where study drug is dispensed the periods will be updated to account for this visit. The following is an example of the update to the algorithm where the patient had an unscheduled visit between the Week 26 and Week 39 visits.
    - 1st day patient takes study medication to the day before the Week 13 visit

- Week 13 to the day before the Week 26 visit
- Week 26 to the day before the Unscheduled Visit where study drug was dispensed.
- Unscheduled Visit where study drug was dispensed to the day before the Week 39 visit.
- Week 39 to the day of the Week 52 visit.

Actual Number of Capsules: The number of capsules taken will be based on the "number of used units" from the IXRS data which is provided for each study kit. The total number of capsules taken by a patient will be the sum of the "number of used units" across the study.

Note: the following modification to the IXRS data will be applied. If the following holds:

- Is the Kit present at the Site = "No"
- Date returned by subject = "Missing"
- Discrepancy reason is not "Damaged by Subject"

Then the number of used units will be set to 90. (note: in these situations, the convention was for the site pharmacy to set the "Used Units" to 0 and the "Missing Unit" to 90).

Planned Number of Capsules:

Planned Number of Capsules =

{# of Capsules for 1st day patient takes study medication to the day before Week 13 Period × # of days in this period +

# of Capsules for Week 13 to the day before the Week 26 visit × # of days in this period +

# of Capsules for Week 26 to the day before the Week 39 visit × # of days in this period +

# of Capsules for Week 39 to the day before the Week 52 visit × # of days in this period}.

Please note the following potential modifications to the algorithm:

- If the patient discontinues the study early, the corresponding period will end at the date of the early termination visit.
- If the patient has an unscheduled visit where study drug is dispensed the periods will be updated to account for this visit. The following is an example of the update to the algorithm where the patient had an unscheduled visit between the Week 26 and Week 39 visits.
  - 1st day patient takes study medication to the day before the Week 13 visit
  - Week 13 to the day before the Week 26 visit
  - Week 26 to the day before the Unscheduled Visit where study drug was dispensed.
  - Unscheduled Visit where study drug was dispensed to the day before the Week 39 visit.
  - Week 39 to the day of the Week 52 visit.

Percent Compliance: Percent Compliance calculated as:

Percent Compliance = 100× Actual Number of Capsules / Planned Number of Capsules.

4.10. Efficacy Evaluation

The FAP is used as the primary population, and the PP and Randomized Populations will be used for sensitivity analyses as described for each efficacy analysis below.

Efficacy endpoints will be listed by patient, and summarized by treatment group and time point, as applicable. For the timed function tests, both the times to complete (seconds) and their speed (i.e. the reciprocals) will be tabulated. In the case of the 10MWT, units will be 10 meters per second. In the case of the stand from supine and the 4-stair climb, units will be task per second. Descriptive statistics will be calculated for quantitative efficacy data as well as for the CFB, and differences between treatment groups as applicable.

Summary statistics for the primary and secondary efficacy endpoints will be presented by investigative site and overall.

If a patient initiates corticosteroid therapy or discontinues study treatment, all observations from that patient after these events will be censored from the efficacy analyses.

4.10.1. Primary Efficacy Analysis

The primary efficacy endpoint is the CFB in the NSAA Total Score at Week 52. NSAA Total Scores and the CFB values will be summarized by treatment group and time point using descriptive statistics.

If a patient initiates corticosteroid therapy or discontinues study treatment, all observations from that patient after these events will be censored from the primary efficacy analyses.

The primary efficacy endpoint analysis will be an MMRM ANCOVA. The MMRM model will include:

- CFB in the NSAA Total Score as the dependent variable
- Variables in the Model

Visit (a categorical variable)	
Treatment (edasalonexent or placebo)	Treatment × Visit
Region (North America vs. Europe/Asia/Australia)	Region × Visit
Baseline age (as a continuous variable),	Baseline Age × Visit
Baseline time to stand from supine (as a continuous variable)	Baseline time to Stand × Visit
Baseline NSAA score (as a continuous variable)	Baseline NSAA score × Visit

- Patient as a random effect.

The primary efficacy analysis will test the difference between the edasalonexent versus placebo at Week 52 using this MMRM model.

Sample SAS code for the primary efficacy analysis is provided in APPENDIX 1: Sample MMRM Code.

Descriptive summary statistics including number of patients, mean, standard error, least-square means (LS means), LS means differences, along with the p-values and corresponding 95% CIs

will be provided. Specifically, the primary efficacy endpoint hypothesis test of no treatment difference between edasalonexent and placebo is based on the least-squares means (LS-means) difference and corresponding p-value from the Week 52 visit. The results for the treatment difference for the other visits will also be provided to assess the consistency of treatment effect over different visits.

The repeated-measures analysis will be based on the restricted maximum likelihood method assuming an unstructured covariance structure to model the within-patient errors. A Kenward-Roger approximation will be used for the denominator degrees of freedom. If there is a convergence problem due to the unstructured covariance matrix, a compound symmetry covariance structure will be used to model the within-patient variance; if the compound symmetric covariance structure fails to converge, a variance components approach will be applied.

The primary efficacy endpoint analysis is an observed case analysis, with no imputation of missing data.

Descriptive summary statistics will be provided for each of the 17 NSAA items at each of the scheduled assessment timepoints.

4.10.2. Secondary Efficacy Analyses

The TFT secondary efficacy endpoints are:

- CFB in 10MWT Speed at Week 52
- CFB in Stand from Supine Speed at Week 52
- CFB in 4-stair Climb Speed at Week 52

Descriptive statistics will be calculated for the speed and time to complete the TFT secondary efficacy endpoints as well as all CFB values by treatment group and time point.

The secondary efficacy endpoint analyses for these endpoints will be an MMRM ANCOVA, using the methodology described for the primary efficacy endpoint analysis (Section 4.10.1). The difference being that, instead of baseline NSAA as a covariate, the corresponding baseline secondary efficacy endpoint value will be used.

As described in Section 2.6.2, if there are any patients not able to complete the test (reason not done = “subject physically unable to complete test”), a value of 0 will be assigned for the reciprocal test (i.e., speed in units of task per second).

4.10.3. Additional Efficacy Analyses

Additional efficacy endpoints include CFB on the following:

- Muscle strength testing: 2 endpoints: elbow flexion (i.e., average of the left and right elbows), knee extension (average of the left and right knees). If either the left or right value is missing the value will be the non-missing score.

  Note: values which are indicated as not being valid as determined by a response of “No” to “Was test valid? (i.e. representative of child’s true function)” for these assessments will not be included in the analysis (but will be presented in listings).

These additional efficacy endpoints will be analyzed using an MMRM ANCOVA, in a manner similar to that described in Section 4.10.1 with the corresponding baseline muscle strength endpoint replacing baseline NSAA as a covariate.

- PUL endpoints: This includes the PUL Total Score, and the sum of individual items 1 and 2 (note these 2 items are imputed for patients who scored a "6" on Item A as defined in Section 2.6.3.2).

  The sum of individual items 1 and 2, and PUL total score endpoints will be analyzed using an MMRM ANCOVA, in a manner similar to that described in Section 4.10.1 with the corresponding baseline PUL score replacing baseline NSAA as a covariate.

  PODCI Transfer and Basic Mobility Scale: This additional efficacy endpoint will be analyzed using an ANCOVA. The ANCOVA model will include CFB in the PODCI Transfer and Basic Mobility Scale at Week 52 as the dependent variable, treatment and region (North America vs. Europe/Asia/Australia) as factors, and baseline age (as a continuous variable), baseline time to stand from supine (as a continuous variable), and baseline PODCI Transfer and Basic Mobility Scale score (as a continuous variable) as covariates.

  As an additional analysis of the PODCI Transfer and Basic Mobility Scale endpoint, the analysis above will be repeated subset to the subgroup of patients with the same caregiver completing the form at both baseline and Week 52 (i.e., from the PODCI eCRF page "Who completed this form?" with responses of: Caregiver, Father, Mother).

  Note that summary statistics will be provided for the additional POCDI scales at each scheduled timepoint.

  CFB multivariate rank sum score: The analysis for this Week 52 endpoint will be an ANCOVA, The ANCOVA model will include Week 52 CFB multivariate rank sum score as the dependent variable, treatment and region (North America vs. Europe/Asia/Australia) as factors, and baseline age (as a continuous variable), baseline time to stand from supine (as a continuous variable), and the corresponding rank sum of the score at baseline will be used (as described in Section 2.6.2).

4.10.4. Sensitivity Analyses for the Primary Efficacy Endpoint

The following sensitivity analyses of the primary and secondary efficacy endpoints will be performed to assess the robustness of the analyses across endpoints:

a) The primary efficacy endpoint analysis performed on the PP population

The primary efficacy endpoint analysis, as defined in Section 4.10.1 will be performed using the PP Population

b) The primary efficacy endpoint analysis performed on the Randomized population

The primary efficacy endpoint analysis, as defined in Section 4.10.1 will be performed using the Randomized Population (note: this includes randomized patients on eteplirsen)

c) ANCOVA on the primary efficacy endpoint using MAR, MNAR, and OC

The primary efficacy endpoint, the CFB in the NSAA Total Score at Week 52 will be analyzed by an ANCOVA model. The ANCOVA model will include CFB in the NSAA Total Score at Week 52 as the dependent variable, treatment and region (North America vs. Europe/Asia/Australia) as factors, and baseline age (as a continuous variable), baseline time

to stand from supine (as a continuous variable), and baseline NSAA score (as a continuous variable) as covariates.

Three different methods of addressing missing data will be performed:

1) **Missing at Random (MAR)** Missing data will be imputed using the MAR algorithm described in Section 4.10.6.1.
2) **Missing Not at Random (MNAR)** Missing data will be imputed using the MNAR algorithm described in Section 4.10.6.2.
3) **OC** (Observed Case) this ANCOVA will be performed on non-missing data.

These analyses will be performed on the FAP.

d) Observed Week 52 MMRM

The primary efficacy endpoint analysis will be repeated (i.e., as described in Section 4.10.1) subset to those patients who had a non-missing CFB value at Week 52.

e) MMRM with multiple imputation assuming MAR

The primary efficacy endpoint analysis will be repeated (i.e., as described in 4.10.1) with missing data imputed based on a Missing at Random (MAR) pattern as described in 4.10.6.1.

f) MMRM with multiple imputation assuming MNAR

The primary efficacy endpoint analysis will be repeated (i.e., as described in 4.10.1) with missing data imputed based on a Missing Not at Random (MNAR) pattern as described in 4.10.6.2

g) Including siblings

The primary efficacy endpoint analysis will be repeated (i.e., as described in 4.10.1) including the siblings in the FAP.

4.10.5. Sensitivity Analyses for the Secondary Efficacy Endpoints

The sensitivity analyses for the primary efficacy endpoint as defined in Section 4.10.4 will be applied to the secondary endpoints.

4.10.6. Multiple Imputation Methods

As part of the sensitivity analysis, the missing CFB primary and secondary efficacy endpoints will be imputed using multiple imputation (MI) for each of the scheduled assessment visits (i.e., Weeks 13, 26, 39, and 52). Both an MAR and MNAR method will be applied.

The following random number seeds will be used for each endpoint/multiple imputation; these seeds were randomly generated.

Table 5: Random Number Seeds for Multiple Imputation

	Random Number Seed	
Endpoint	**MAR**	**MNAR**
NSAA	29068	24158

10MWT Speed	52987	60804
Stand from Supine Speed	69440	21531
4-stair Climb Speed	36875	55339

Imputation distribution:

The imputation distribution for the missing change from baseline endpoints will be a normal distribution. Although the change from baseline values are integers for the primary change from baseline in NSAA endpoint, imputed values will be rounded to 1 decimal place when used in the analysis. For the change from baseline secondary endpoints (i.e., TFT reciprocal of seconds endpoints) the imputed values will be rounded to 4 decimal places (i.e., the ten-thousandth place).

4.10.6.1. Multiple Imputation Assuming MAR

Imputation algorithm:

Missing post-baseline change from baseline values will be imputed via multiple imputation (MI) assuming MAR for each scheduled assessment visit.

Fifty imputed datasets will be created. Each of the imputed datasets will be analyzed via the specified primary or secondary efficacy endpoint analysis. The results across the multiple imputed data sets will be combined using SAS Proc MIANALZYE (sample SAS code for primary analysis provided in APPENDIX 2: Sample Multiple Imputation Code).

The following are the set of variables to be used in the multiple imputation model:

- Treatment group (edasalonexent or placebo)
- Region (North America vs. Europe/Asia/Australia)
- Baseline age (continuous variable)
- Baseline time to stand from supine (continuous variable)
- Baseline value for the endpoint being imputed (continuous variable)
- Previous change scores (e.g., for the Week 52 CFB imputation, the CFB scores at Weeks 13, 26, and 39 will be included)

The imputations will be performed using monotone linear regression imputation methods which will impute the patients' missing post-baseline change scores at each of the scheduled assessment visits in the study.

If the missing data does not follow a monotonic pattern, a sequential approach to imputing the data via a Markov chain Monte Carlo (MCMC) method to produce a monotone missing data pattern will be applied using the MCMC impute=monotone option in Proc MI (SAS System). Following the method described in Smith (2017) the non-monotonic data imputation will be performed "by" the categorical treatment variable (note: the categorical variable Region will not be included in the non-monotonic imputation). Then the monotone linear regression imputation methods will be applied.

Note: if the multiple imputation model does not converge or produce estimates (i.e., due to over-specification) the set of imputation variables may be modified.

4.10.6.2. MMRM with Multiple Imputation Assuming MNAR

Imputation algorithm:

For this imputation the monotonic missing post-baseline change from baseline scores will be imputed via multiple imputation assuming MNAR for each scheduled assessment visit (i.e., Weeks 13, 26, 39, and 52).

The imputation of monotonic missing scores in the edasalonexent treatment group will be based on the distribution in the placebo group. This method assumes that the trajectory of withdrawals from the edasalonexent treatment arm follows the distribution of the placebo patients. The MNAR option in Proc MI (SAS system) will be used to impute the placebo distribution as described (Yuan, 2014).

Fifty imputed datasets will be created. Each of the imputed datasets will be analyzed via the same model as the primary efficacy endpoint analysis (see Section 4.10.1). The results across the multiple imputed data sets will be combined using SAS Proc MIANALZYE (details provided in APPENDIX 2: Sample Multiple Imputation Code).

The set of variables for this MNAR multiple imputation model is the same set as defined for the MAR multiple imputation (Section 4.10.6.1), with the exception of Treatment group, which is not included in the MNAR imputation (i.e., the placebo distribution is used to impute the data).

The imputations will be performed using monotone linear regression imputation methods which will impute the patients' missing post-baseline change scores at each of the scheduled visits in the study (i.e., Weeks 13, 26, 39, and 52).

If the missing data does not follow a monotonic pattern, a sequential approach to imputing the data via an MCMC method to produce a monotone missing data pattern will be applied using the MCMC impute=monotone option in Proc MI (SAS System). Following the method described in Smith (2017) the non-monotonic data imputation will be performed "by" the categorical treatment variable (note: the categorical variable Region will not be included in the non-monotonic imputation). Then the monotone linear regression imputation methods will be applied.

Note: if the multiple imputation model does not converge or produce estimates (i.e., due to over-specification) the set of imputation variables may be modified.

4.10.7. Supportive Analyses on Primary and Secondary Efficacy Endpoints

The following analyses will be performed to support the primary endpoint analyses. These analyses will be performed on the FAP.

- NSAA Categorical Change from Baseline: The percentage of patients with a change from baseline in NSAA Total Score at Week 52 will be compared between the treatment groups. The following categories for change from baseline in NSAA will be considered
    - Improvement: $\geq 0, \geq 1, \geq 2, \geq 3, \geq 4$
    - Decline: $\leq 0, \leq -1, \leq -2, , \leq -3 , \leq -4, , \leq -5$

The difference between the treatment groups will be compared by a chi-square test, except for the scenarios where, in either group, there are only 5 patients or less in the numerator (or, conversely, N-5 patients or more) in which case a Fisher's exact test will be used.

Additionally, a Kolmogorov–Smirnov test for differences in the change from baseline in NSAA Total Score between treatment groups will be performed.

- Increase in function in Week 52 NSAA Items: For each of the 17 NSAA items, the number and percent of patients with an increase in function will be summarized. A patient is defined as having an increase in function if the baseline items score is <2 and the post-baseline value is greater than the baseline value. Patients with a baseline item score of 2 or a missing Week 52 NSAA item value are excluded from the analysis.

  For each item, the percentages, odds ratio (with 95% CI), and a p-value comparing the treatment group difference in proportions will be presented.

  For the Overall NSAA increase in function, the overall odds ratio, 95% CI, and p-value will be estimated from a generalized estimating equations (GEE) model with treatment and NSAA item as class variables in the model, increase in NSSA item (i.e., a dichotomous variable) as the dependent variable, baseline item score as a covariate, and subject as the repeated term with an exchangeable working correlation matrix.

  The patient count and percent of number of Week 52 NSAA items which increased in function will be presented (i.e., the "count" is the number of NSAA items with a baseline score <2 and the post-baseline value is greater than the baseline value. The percent is the "count" divided by the number of NSAA items with a baseline score <2 and non-missing Week 52 value).

- Preservation of function in Week 52 NSAA Items: For each of the 17 NSAA items, the number and percent of patients with preservation of function will be summarized. A patient is defined as having a preservation of function if the baseline items score is >0 and the post-baseline value is >0. Patients with a baseline item score of 0 or a missing Week 52 NSAA item value are excluded from the analysis.

  For each item, the percentages, odds ratio (with 95% CI), and a p-value comparing the treatment group difference in proportions will be presented.

  For the Overall NSAA increase in function, the overall odds ratio, 95% CI, and p-value will be estimated from a generalized estimating equations (GEE) model with treatment and NSAA item as class variables in the model, NSSA item score >0 (i.e., a dichotomous variable) as the dependent variable, baseline item score as a covariate, and subject as the repeated term with an exchangeable working correlation matrix.

  The patient count and percent of number of Week 52 NSAA items with preserved function will be presented (i.e., the "count" is the number of NSAA items with a baseline score >0 and the post-baseline value is >0. The percent is the "count" divided by the is the number of NSAA items with a baseline score >0 and non-missing Week 52 value).

- Worsening of function in Week 52 NSAA Items: For each of the 17 NSAA items, the number and percent of patients with a worsening of function will be summarized. A patient is defined as having a worsening of function if the baseline items score is >0 and the post-baseline value is less than the baseline value. Patients with a baseline item score of 0 or a missing Week 52 NSAA item value are excluded from the analysis.

  For each item, the percentages, odds ratio (with 95% CI), and a p-value comparing the treatment group difference in proportions will be presented.

  For the Overall NSAA worsening in function, the overall odds ratio, 95% CI, and p-value will be estimated from a generalized estimating equations (GEE) model with treatment and NSAA item as class variables in the model, worsening of function (i.e., a dichotomous variable) as the dependent variable, baseline item score as a covariate, and subject as the repeated term with an exchangeable working correlation matrix.

  The patient count and percent of number of Week 52 NSAA items with worsening function will be presented (i.e., the "count" is the number of NSAA items with a baseline score >0 and the post-baseline value is less than the baseline value. The percent is the "count" divided by the is the number of NSAA items with a baseline score >0 and non-missing Week 52 value).

- Loss in function in Week 52 NSAA Items: For each of the 17 NSAA items, the number and percent of patients with a loss of function will be summarized. A patient is defined as having a loss of function if the baseline items score is >0 and the post-baseline value is 0. Patients with a baseline item score of 0 or a missing Week 52 NSAA item value are excluded from the analysis.

  For each item, the percentages, odds ratio (with 95% CI), and a p-value comparing the treatment group difference in proportions will be presented.

  For the Overall NSAA loss in function, the overall odds ratio, 95% CI, and p-value will be estimated from a generalized estimating equations (GEE) model with treatment and NSAA item as class variables in the model, loss of function (i.e., a dichotomous variable) as the dependent variable, baseline item score as a covariate, and subject as the repeated term with an exchangeable working correlation matrix.

  The patient count and percent of number of Week 52 NSAA items with loss of function will be presented (i.e., the "count" is the number of NSAA items with a baseline score >0 and the post-baseline value is 0. The percent is the "count" divided by the is the number of NSAA items with a baseline score >0 and non-missing Week 52 value).

- Gain of function in Week 52 NSAA Items: For each of the 17 NSAA items, the number and percent of patients with a gain in function will be summarized. A patient is defined as having a gain in function if the baseline items score is 0 and the post-baseline value is > 0 (i.e., a score of 1 or 2). Patients with a baseline item score of 1 or 2 or a missing Week 52 NSAA item value are excluded from the analysis (i.e., this analysis is based on the items with a baseline score of 0).

  For each item, the percentages, odds ratio (with 95% CI), and a p-value comparing the treatment group difference in proportions will be presented.

  For the Overall NSAA gain of function, the overall odds ratio, 95% CI, and p-value will be estimated from a generalized estimating equations (GEE) model with treatment and

NSAA item as class variables in the model, gain of function (i.e., a dichotomous variable) as the dependent variable, baseline item score as a covariate, and subject as the repeated term with an exchangeable working correlation matrix.

The patient count and percent of number of Week 52 NSAA items with a gain in function will be presented (i.e., the "count" is the number of NSAA items with a baseline score of 0 and the post-baseline value is >0. The percent is the "count" divided by the is the number of NSAA items with a baseline score of 0 and non-missing Week 52 value).

- NSAA Items: Change from Baseline: For each NSAA item at each scheduled visit, the change from baseline will be summarized and the two treatment groups will be compared using CMH test on change from baseline score controlling for baseline NSAA item score.

4.10.8. Subgroup Analyses on Primary and Secondary Efficacy Endpoints

Subgroup analyses of the primary and secondary endpoints will be performed in the same manner as the primary analysis within each of the following subgroups based on the FAP:

- Baseline Age (<= 6.0 years old and > 6.0 years old)
- Baseline Time to Stand from Supine (<=5 seconds and >5 seconds)
- Region (North America and Europe/Asia/Australia).

The MMRM model for the previous treatment with eteplirsen subgroup will be the same as the main model. The MMRM model for the other subgroups will be the same as the main model, excluding the subgroup as a factor. If particular subgroups have two few observations (patients) to provide a meaningful MMRM analysis, then descriptive statistics only will be presented for those particular subgroups.

Note that these analyses will be performed on the randomization strata.

4.11. Pharmacokinetic Analyses

Pharmacokinetic (PK) analyses will be described in a separate analysis plan.

4.12. Pharmacodynamic Analyses

Pharmacodynamic (PD) analyses may include evaluation of gene expression as well as circulating protein and micro- ribonucleic acid (RNA) biomarkers that help inform target engagement and efficacy. These assessments and blood samples will be collected at the times specified in Table 1.

Descriptive statistics will be presented by treatment group for applicable PD endpoints and will be included as an appendix to the CSR (note: PD data may not be available at the time of database lock).

4.13. Safety Analyses

Safety analyses will be conducted using the Safety Population. Results of all safety assessments will be listed by patient, and summarized by treatment group and time point as applicable. These safety analyses will be compared qualitatively between treatment groups; however, there will be

no inferential statistical comparisons between groups (except for specific analyses, noted below).

4.13.1. Adverse Events

Adverse events will be coded using the MedDRA dictionary and displayed in tables and listings using System/Organ/Class (SOC) and Preferred Term (PT).

Analyses of adverse events will be performed for those events that are considered TEAEs, where treatment emergent is defined as any adverse event with onset after the administration of study medication through the end of the study (Week 52 visit) or any event that was present at Baseline but worsened in intensity or was subsequently considered drug-related by the investigator through the end of the study.

If the start date of an AE is partially or completely missing, then the date will be compared as far as possible with the date of the randomization. The AE will be assumed to be treatment emergent if it cannot be definitively shown that the AE did not occur or worsen during the post-randomization period (worst case approach).
The following general rules will be used:

- If the start day is missing, but the start month and year are complete, an AE will only be excluded as being on-study if the start month/year is before the month/year of randomization or if the stop date is before randomization.
- If the start day and month are missing, but the start year is complete, an AE will only be excluded as being on-study if start year is before the year of randomization or if the stop date is before randomization.
- If the start date is completely missing, an AE will be considered on-study unless the stop date is before randomization.

TEAE summaries will include the number and percentage of patients in each treatment group and overall. Summaries will include patients with:

- Any TEAEs
- Treatment-related TEAEs (Related, Possibly related)
- Maximum Severity of TEAEs
- Serious TEAEs, including any events resulting in death
- Treatment-related serious TEAEs
- TEAEs (and treatment-related TEAEs) leading to discontinuation of study drug

TEAEs are summarized by patient incidence rates, therefore, in any tabulation, a patient contributes only once to the count for a given adverse event (SOC or preferred term). TEAEs will also be summarized by SOC, PT, and severity and by SOC, PT, and relatedness. The most severe occurrence of a TEAE, as well as the most extreme relationship of the TEAE to the study procedure or treatment, will be used in tabulations of severity and relatedness in cases of multiple occurrences of the same TEAE.

Additionally, the prevalence of common treatment-related TEAEs will be summarized by treatment group and overall across the treatment period stratified into 4-week intervals. A common treatment-related TEAE is one which occurs in over 10% of the patients in either treatment group. The objective of this analysis is to determine the time course of the incidence of these TEAEs. The following are the 13 4-week intervals:

1) Baseline – Week 4 (i.e., Day 1 – Day 28)
2) Week 5 – Week 8 (i.e., Day 29 – Day 56)
3) Week 9 – Week 12 (i.e., Day 57 – Day 84)
4) Week 13 – Week 16 (i.e., Day 85 – Day 112)
5) Week 17 – Week 20 (i.e., Day 113 – Day 140)
6) Week 21 – Week 24 (i.e., Day 141 – Day 168)
7) Week 25 – Week 28 (i.e., Day 169 – Day 196)
8) Week 29 – Week 32 (i.e., Day 197 – Day 224)
9) Week 33 – Week 36 (i.e., Day 225 – Day 252)
10) Week 37 – Week 40 (i.e., Day 253 – Day 280)
11) Week 41 – Week 44 (i.e., Day 281 – Day 308)
12) Week 45 – Week 48 (i.e., Day 309 – Day 336)
13) Week 49 – Week 52 (i.e., Day 337 – Day 364)

For each interval, the percentage of patients with the treatment-related TEAE in the interval will be presented. For each interval, the denominator for the percentage will be the number of patients who were in the study for any part of that period (i.e., patients who discontinued from the study prior to the start of the interval will be excluded). For each related TEAE, a patient will be included in the numerator if any of the following criteria are met:

- The start date of the TEAE is in the interval
- The stop date of the TEAE is in the interval
- The start date of the TEAE is before the interval and the stop date is after the interval
- The start date of the TEAE is before the interval and TEAE is ongoing

Note: For this prevalence TEAE analysis, if the start day and/or month of the AE is missing, a value of 01 for the day and January for the month will be used. If the stop day and/or month of the AE is missing, December will be used for the month and the last day of the month will be used for the day.

All adverse events occurring on study will be listed in patient data listings.

By-patient listings also will be provided for the following: patient deaths; serious adverse events; and adverse events leading to discontinuation of study drug.

4.13.2. Physical Examination and Growth Parameters

Physical examination (general appearance; skin; head, ears, eyes, nose, and throat [HEENT]; etc.) results will be summarized in a shift table from Baseline, and abnormal results will also be summarized.

Growth Parameters (height, weight, body mass index [BMI]) will be summarized by age-normative z-scores. The age-normative height, weight, and BMI scores will be based on a Centers for Disease Control and Prevention (CDC) SAS macro which calculates the percentiles and z-scores for a child's sex and age.

The CDC SAS macro is found here (date accessed: 08Sep2019):

https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm

These summaries will be performed on the Safety Population.

4.13.3. Laboratory Data

Values at all study time points and changes from Baseline to post-Baseline time points in clinical chemistry, hematology and urinalysis results will be summarized descriptively.

Abnormal clinical laboratory values will be identified as outside (above or below) the normal range and will be evaluated for clinically notable abnormalities.

All laboratory data and normal ranges will be provided in data listings. A subset listing will be presented for all abnormal laboratory values (i.e., values which are indicated as being notable or critical low/high).

The following abbreviations for flags in the central lab database.

- L: Reference Low
- NL: Notable Low
- CL: Critical Low
- R: Rechecked Value
- H: Reference High
- NH: Notable High
- CH: Critical High

In addition, to assess decreases in creatine kinase (CK), the change from baseline in CK will be assessed at each scheduled assessment visit using a non-parametric signed rank test (for a within-group comparison) and a Wilcoxon rank sum test for the comparison of edasalonexent versus placebo.

4.13.4. Vital Signs

The actual value and change from Baseline will be summarized for vital signs (temperature, heart rate, respiratory rate, and systolic and diastolic blood pressure) descriptively at Baseline and other study time points.

By-patient listings of vital sign measurements will also be provided.

4.13.5. Electrocardiogram

ECG results will be summarized descriptively, including the number and percent of patients with normal, abnormal, and clinically significant abnormal results at Baseline and each study time point.

In addition, an ANCOVA model including CFB in various ECG parameters at Week 52 (i.e., corrected QT [Bazett and Fridericia], RR interval, PR interval, QRS duration, Heart Rate) as the dependent variable, treatment as a fixed effect, and baseline ECG parameter value, will be used to assess the treatment difference in the ECG parameters at Week 52.

Treatment group differences in heart rate will be compared via an MMRM ANCOVA model with treatment, baseline heart rate, visit, visit×treatment (interaction term), and baseline heart

rate×treatment (interaction term) as fixed effects and patient as a random effect. In this model, visit is a categorical variable, treatment is a dichotomous variable, and baseline heart rate is a continuous variable. This analysis will be performed on the Safety Population.

This repeated-measures analysis will be based on the restricted maximum likelihood method assuming an unstructured covariance structure to model the within-patient errors. A Kenward-Roger approximation will be used for the denominator degrees of freedom. If there is a convergence problem due to the unstructured covariance matrix, a compound symmetry covariance structure will be used to model the within-patient variance; if the compound symmetric covariance structure fails to converge, a variance components approach will be applied.

All ECG data for each patient will be provided in data listings.

4.13.6. Cardiac Monitoring

Measures of cardiac autonomic dysfunction, including heart rate variability, will be measured at Screening, Week 26, and Week 52. Cardiac monitoring will be performed for approximately 48 hours remotely via a wearable cardiac monitoring device if approved for use in the participating country. Cardiac monitoring data analysis methods will be provided in a separate analysis plan.

4.13.7. Bone Evaluation

Lateral thoracolumbar spine radiography scans and Dual-energy X-ray absorptiometry (DXA) scans will be performed at Baseline and Week 52. In Germany, bone assessment will not be performed until the respective approvals from the German Federal Office for Radiation Protection become available. Bone radiography will be locally read and centrally evaluated. Bone evaluation data analysis methods will be provided in a separate analysis plan.

4.13.8. Prior and Concomitant Medications and Therapeutic Procedures

Prior and concomitant medications and therapeutic procedures will be coded using the WHO Drug dictionary. Results will be tabulated by Anatomic Therapeutic Class (ATC; level 4) and preferred term.

Prior medications are those that started and stopped before exposure to study medication (i.e., date of first dose); concomitant medications are all medications taken during the study period, including those started before but on going at first dose. Where a medication start date is partially or fully missing, and it is unclear as to whether the medication is prior or concomitant, it will be assumed that it is concomitant.

The use of concomitant medications will be included in a by-patient data listing and summarized in a table. The use of prior medications and therapeutic procedures will be included in by-patient data listing.

Additionally, a by-patient listing of concomitant medications that started or stopped during the treatment period will be created, specifically indicating the following:

- Treatment Group
- Patient ID:
- Medication Name/Preferred Term/ATC Term Modification:
- Indication

- Start/Stop Date (ongoing)
- Dose (unit)
- Frequency/Route

The determination of a medication starting or stopping with the treatment period will be based on the following:

- Date of First Exposure to Treatment ≤ Date of medication start date ≤ Date of Last Exposure to Treatment

  (i.e., ADCM.TRTSDT ≤ ADCM. ASTDT ≤ ADCM.TRTEDT)

  The following general rules will be used for partially missing medication start dates:
  - If the start day is missing, but the start month and year are complete, the medication will be considered as having started during the treatment period if the month/year is between the treatment start/stop month/year values (inclusive).
  - If the start day and month are missing, but the start year is complete, the medication will be considered as having started during the treatment period if the year is between the treatment start/stop year values (inclusive).
  - If the start date is completely missing, the medication will not be considered as having started in the treatment period.

- Date of First Exposure to Treatment ≤ Date of medication stop date ≤ Date of Last Exposure to Treatment

  (i.e., ADCM.TRTSDT ≤ ADCM. AENDT ≤ ADCM.TRTEDT)

  The following general rules will be used for partially missing medication stop dates:
  - If the stop day is missing, but the stop month and year are complete, the medication will be considered as having stopped during the treatment period if the month/year is between the treatment start/stop month/year values (inclusive).
  - If the stop day and month are missing, but the stop year is complete, the medication will be considered as having stopped during the treatment period if the year is between the treatment start/stop year values (inclusive).
  - If the stop date is completely missing, the medication will not be considered as having stopped in the treatment period.

  Note: Here ADCM refers to the ADaM concomitant medications dataset. The words after "ADCM." indicate variables in the ADaM concomitant medications dataset.

5. COVID-19 ASSESSMENTS AND RELATED ANALYSES

The following sections describe the COVID-19 related assessments collected and additional analyses to address the potential impact of COVID-19.

5.1. COVID-19 Site-Level Assessments

The following site-level information will be listed for each site impacted by COVID-19.

- Site Number
- Site Name
- Trial Activities Impacted by COVID: Yes, No
- Site Closed to Research: Yes/No
- Start/Stop Dates of Closure
- Site Open but Limited Capacity: Yes, No, NA
- Start/Stop Dates at Limited Capacity

This information will be collected by the sponsor and provided in an excel spreadsheet for incorporation into the database. Only sites impacted by COVID-19 will be included.

5.2. COVID-19 Patient-Level Listings

The following data will be listed to provide patient-level information on which assessments were impacted by COVID-19.

- Patient ID
- Visit Week
- Reason for Impact: This information will be collected via discussions with the site (and provided by the sponsor for incorporation into the database) examples of the responses will include:
    - Patient Discretion due to COVID
    - Site closed due to COVID
    - Travel Restrictions due to COVID
    - Reduced Capacity due to COVID
- Out of window visit:
    - Yes, No

  A visit will be considered out of window if the difference between the date of visit and the scheduled date is outside of the ± 10 day window. The scheduled dates for the Week 26, 39, and 52 visits, are 182, 273, and 364 days from the patient's date of randomization, respectively. For this determination the date of NSAA assessment will be

used. If that date is missing, the date of the visit (i.e., from the DOV SAS dataset) will be used.

- Remote Visit:
    - Yes, No

  This will be determined by the "Telehealth Contact" variable in the Date of Visit eCRF page (i.e., Telehealth Contact value of "Yes" = Remote Visit = "Yes"; Telehealth Contact value of "No" = Remote Visit = "No"). Note for Week 26 this information will be provided in the spreadsheet (i.e., the "Remote Visit" Column).

- Re-supply Study Drug:
    - Yes, No

  This will be determined by the "Re-supply Study Drug" variable in the Date of Visit eCRF page. Note for Week 26 this information will be provided in the spreadsheet (i.e., the "Re-supply" column).

- Safety Lab:
    - Yes, No

  This will be determined by the "Safety Labs" variable in the Date of Visit eCRF page. Note for Week 26 this information will be provided in the spreadsheet (i.e., the "Safety Labs" column).

- Local Lab:
    - Yes, No

  This will be determined by the SAS variable LB.LOCALYN (i.e., from the eCRF Central Laboratory page "Were local labs used for this visit?"). Note for Week 26 this information will be provided in the spreadsheet (i.e., the "Local Labs" column).

- Height Collected, Weight Collected:
    - Yes, No

  If the patient's height (weight) is non-missing for this visit, the value will be "Yes", "No" otherwise

- Remote Efficacy:
    - Yes, No

  This will be determined by the "Remote Efficacy" variable in the Date of Visit eCRF page; a value of "Yes" indicates a response of "Remote", a value of "No" indicates "Site". Note there is no remote efficacy at Week 26 (i.e., for Week 26 Remote Efficacy = "No" should be populated in the listing).

- NSAA Performed, 10 MWT Performed, 4SC Performed, TTS Performed, Muscle Strength Performed:
    - Yes, No

These variables will be determined by the "was the test performed variables" listed below:

Assessment	SAS Variable
NSAA Performed	MK_NSAA. MKPERF
10 MWT Performed	MK_TFT.WT_MKPERF
4SC Performed	MK_TFT.SCT_MKPERF
TTS Performed	MK_TFT.RFS_MKPERF
Muscle Strength Performed	MK_MST. MKPERF
PODCI Performed	QS_PODCI.QSPERF

- PE Collected, ECG Collected, PK Collected, Biomarkers Collected, Cardiac Monitoring Performed, DXA Performed, Spine Films Performed:
    - Yes, No

  These variables will be determined by the "was the test performed variables" listed below:

Assessment	SAS Variable
PE Collected	PE.PEPERF
ECG Collected	EG.EGPERF
PK Collected	PC_BL.PCPERF
Biomarkers Collected	PC_PD.PCPERF
Cardiac Monitoring Performed	FA_CM. FAPERF
DXA Performed	PR_DXA. PROCCUR
Spine Films Performed	PR_LSR.PROCCUR

There is also a comment field.

Only assessments impacted by COVID-19 will be included.

5.3. COVID-19 Analyses

The primary and secondary efficacy analysis (Sections 4.10.1 and 4.10.2) will be repeated for the following two sensitivity analyses (i.e., a total of 8 new analyses: 2 sensitivity analyses across the primary and 3 secondary endpoints).

1) Excluding all remote efficacy assessments
2) Excluding all assessments which occurred during the time period the patient's site was closed.

6. CHANGES TO PLANNED ANALYSES

The following are changes to the planned analyses as specified in the 16APR2020 version of the protocol:

- The MMRM model for the primary and secondary endpoint analyses was updated to include baseline covariate by study visit interaction terms and the Baseline age and Baseline time to stand are continuous variables (as specified in Section 4.10.1).
- An additional endpoint was added, the CFB rank sum score at Week 52. This endpoint is designed to assess the effect of edasalonexent versus placebo across multiple endpoints using a non-parametric rank-based approach. The endpoint is the sum of the ranks of the following Week 52 CFB scores: NSAA, 10MWT speed, stand from supine speed, 4-stair climb speed, and PODCI transfer and basic mobility scale.

7. REFERENCES

American Academy of Orthopaedic Surgeons (1997). Pediatric Orthopaedic Society of North America. American Academy of Pediatrics. Shriners Hospitals. Scoring Algorithms for Pediatrics Outcomes Data Collection Instrument. 2.0 ed.

Henricson, E., R. Abresch, et al. (2013). "The 6-minute walk test and person-reported outcomes in boys with duchenne muscular dystrophy and typically developing controls: longitudinal comparisons and clinically-meaningful changes over one year." PLoS Curr **5**.
O'Brien (1984). "Procedures for Comparing Samples with Multiple Endpoints." (Biometrics 40): 1079-1087.
Pane, M., G. Coratti, et al. (2018). "Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal data." PLOS ONE **13**(6): e0199223.
Smith, C., Kosten S. (2017). "Multiple Imputation: A Statistical Programming Story." PharmaSUG: 2017 - Paper SP2001.
Yuan, Y. (2014). "Sensitivity Analysis in Multiple Imputation for Missing Data." SUGI paper: SAS270-2014.

8. CLINICAL STUDY REPORT APPENDICES

The set of tables and listings along with the corresponding shells will be provided in a separate document.