Statistical Analysis Plan

TITLE PAGE

PROTOCOL TITLE:	A Phase 3, Randomized, Double-Blind, Trial of Pamrevlumab (FG-3019) or Placebo in Combination with Systemic Corticosteroids in Subjects with Non-ambulatory Duchenne Muscular Dystrophy (DMD)
PROTOCOL NUMBER:	Protocol FGCL-3019-093 Amendment 6: 01 November 2022
STUDY SPONSOR:	FibroGen, Inc. 409 Illinois Street San Francisco, California 94158 USA
STUDY DRUG:	Pamrevlumab
INDICATION:	Duchenne Muscular Dystrophy (DMD)
SAP VERSION	Final 1.0
RELEASE DATE	May 23, 2023

SIGNATURES AND APPROVALS

Approvals

I have reviewed and accepted the information in this document to be a true and accurate representation of the Statistical Analysis Plan.

Signature: see appended final page for 21CFR Part 11 compliant approval

Initiator:

[redacted], FibroGen, Inc.

Reviewed by:

[redacted] FibroGen, Inc.
[redacted] FibroGen, Inc.
[redacted] FibroGen, Inc.
[redacted] FibroGen, Inc.
[redacted] FibroGen, Inc.
[redacted] FibroGen, Inc.
[redacted] FibroGen, Inc.
[redacted] FibroGen, Inc.

Signature Significance

The following significance is lent to the signatures on the Approvals page of this document.

Signatory	Significance
Initiator	By signing, the author is attesting that the content of the document is complete and accurate.
Reviewer	By signing, the reviewer(s) are attesting that the document's approach and contents are compliant with the study protocol, all appropriate, regulatory requirements, and other significant guidelines. This individual(s) has reviewed the document for accuracy and completeness.

CHANGE HISTORY

Version	Date	Description
1.0	23May2023	Final approved version

LIST OF TABLES

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Definition
ADaM	Analysis Data Model
AE	Adverse Event
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
BP	Blood Pressure
BPM	Beats Per Minute
BSA	Body Surface Area
BTR	Best-Test Read
CBC	Complete Blood Count
CDE	Center for Drug Evaluation (China)
CRF	Case Report Form
CS	Clinically Significant
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DBL	Database Lock
DCM	Dilated Cardiomyopathy
DMD	Duchenne Muscular Dystrophy
DMC	Data Monitoring Committee
ECG	Electrocardiogram
EE	Efficacy Evaluable
EoS	End of Study
EoT	End of Treatment
FEV_1	Forced Expiratory Volume in 1 Second
FG-3019	FibroGen-3019 (Recombinant Human Monoclonal Antibody), Pamrevlumab
FVC	Forced Vital Capacity
GCS	Global Circumferential Strain
HAHA	Human Anti-Human Antibody

Abbreviation	Definition
HHM	Handheld Myometry
ICF	Informed Consent Form
IP	Investigational Product
ITT	Intent-to-Treat
IV	Intravenous
kg	Kilogram
LGE	Late Gadolinium Enhancement
LLN	Lower Limit of Normal
LTFU	Lost to Follow-up
LVEF	Ventricular Ejection Fraction
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
MMRM	Mixed Model Repeated Measures
MRI	Magnetic Resonance Imaging
NCI	National Cancer Institute
NCS	Not Clinically Significant
OLE	Open Label Extension
PD	Pharmacodynamics
PEF	Peak Expiratory Flow
PFT	Pulmonary Function Test
PK	Pharmacokinetic
ppFVC	percent predicted FVC
PRO	Patient Reported Outcome
PT	Preferred Term
PTT	Partial Prothrombin Time
PUL	Performance of Upper Limb
QoL	Quality of Life
ROM	Range of Motion
RCM	Random Coefficient Model
SAF	Safety Analysis Set
SAE	Serious Adverse Event

Abbreviation	Definition
SAP	Statistical Analysis Plan
SD	Standard Deviation
SDTM	Study Data Tabulation Model
SE	Standard Error
SoA	Schedule of Assessments
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TESAE	Treatment Emergent Serous Adverse Event
TFLs	Tables, Figures and Listings
ULN	Upper Limit of Normal
WHODD	World Health Organization Drug Dictionary

1. INTRODUCTION

This statistical analysis plan (SAP) provides a more technical and detailed elaboration of the statistical analyses of study data as specified in the Study Protocol FGCL-3019-093 Amendment 6, 01 November 2022. Specifications of tables, figures, and listings (TFL) are contained in a separate document. 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 study. The SAP will be finalized and signed off prior to database lock (DBL). Any major modification of this SAP after the signoff will be documented in an SAP amendment or the CSR.

This SAP is only for the main study period. The Open Label Extension (OLE) will have a separate SAP. A population PK analysis, as well as an exposure-response analysis, will be defined in a separate PK analysis plan.

Based on regional regulatory filing needs, regional specific population will be analyzed following this same SAP. Refer to Appendix 6 if regulatory submission in China is pursued.

2. STUDY OBJECTIVES

The overall objective of this trial is to evaluate the efficacy and safety of pamrevlumab versus placebo in combination with systemic corticosteroids in subjects with non-ambulatory Duchenne Muscular Dystrophy (DMD) (age 12 years and older).

3. STUDY DESIGN

3.1. Overview

Study FGCL-3019-093 is a Phase 3, global, randomized, double-blind, placebo-controlled multi-center study. The goals of the study are to evaluate the efficacy and safety of pamrevlumab compared to placebo in subjects with non-ambulatory Duchenne Muscular Dystrophy over a 52-week period.

3.2. Study Population

Approximately 92 male subjects will be enrolled in this trial, globally.

3.3. Sample Size Determination

This section is updated based on latest publications and the study enrollment.

This study is planned to enroll a total of approximately 92 subjects with a 1:1 randomization ratio of assigning to pamrevlumab or placebo, respectively. With Performance of Upper Limb (PUL) version 2.0, the sample size of 92 subjects is needed to detect a treatment difference of 2.0 between treatment arms in change from baseline in the total score of PUL at Week 52, assuming a standard deviation of 2.9 with a 90% power, based on a two-sided two-sample t-test at the alpha level of 0.05. A sample size of 82 subjects will provide the study with at least 85% power to detect the same assumed treatment difference.

3.4. Randomization and Treatment Assignment

Subjects will be randomized with a 1:1 ratio to one of the two study treatment arms as follows:

- Arm A: pamrevlumab 35 mg/kg IV infusion Q2 weeks + systemic deflazacort or equivalent potency of corticosteroids administered orally.
- Arm B: matching placebo IV infusion Q2 weeks + systemic deflazacort or equivalent potency of corticosteroids administered orally.

3.5. Study Periods

This study consists of the following study periods (Figure 1):

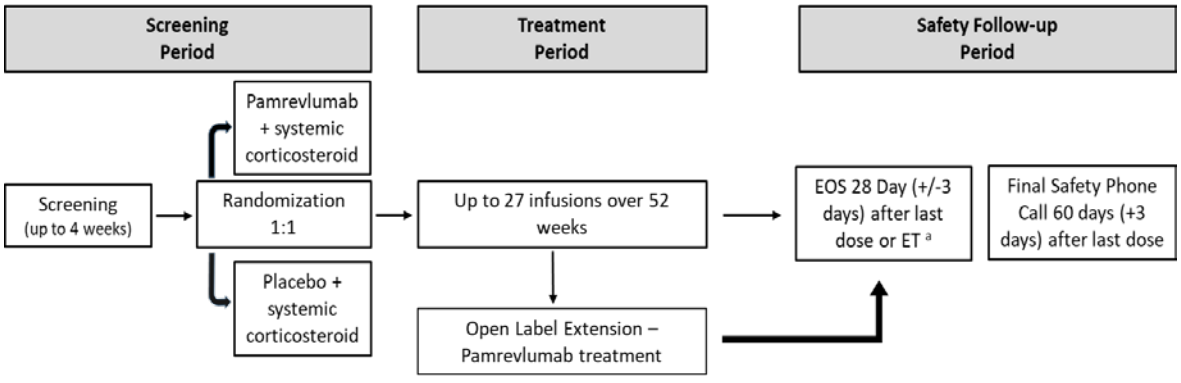
- Main (double-blind, placebo-controlled) study period:
  - Screening period: Up to 4 weeks
  - Treatment period: 52 weeks
- Optional, open label extension (OLE) period:

  Subjects who complete week 52 of the main study period (regardless of the number of study drug infusions received) will be eligible to participate in the open-label extension (OLE) where all subjects will be treated with pamrevlumab. If a subject discontinues the main study period early for any reason, they will not be eligible for inclusion in the OLE.
- Follow-up period/final safety assessment:
  - 28 days after last dose: scheduled visit
  - 60 days after last dose: follow-up phone call, for a final safety assessment

Subjects who discontinue study treatment for any reason should be encouraged to return to the investigative site to complete final safety and efficacy assessments.

A schematic overview of the study is provided in Figure 1.

Figure 1: FGCL-3019-093 Study Schema



Abbreviations: EOS= End of Study, ET=Early Termination
a) Subjects who discontinue the study early, for any reason, will complete an early termination visit 28 days (+/- 3 Days) and a final safety follow-up phone call 60 days (+ 3 days) after the last infusion.

3.6. Study Assessments

The intent of this study is to evaluate the efficacy and safety profile of pamrevlumab in combination with systemic corticosteroids in subjects with non-ambulatory Duchenne muscular dystrophy (age 12 years and older).

The following assessments will be assessed by the schedules defined in the protocol (Schedule of Assessments (SoA) in the Appendix of the protocol):

- Performance of Upper Limb (PUL) V2.0
- Pulmonary Function Test (PFT) including
    - Forced Vital Capacity (FVC)
    - Peak Expiratory Flow (PEF)
    - Forced Expiratory Volume in 1 second (FEV1)
- Grip strength of the hands by Handheld Myometry (HHM)
- Left Ventricular Ejection Fraction percentage (LVEF %), assessed by magnetic resonance imaging (MRI)
- Duchenne Video Assessment (DVA)
- Dilated Cardiomyopathy (DCM) by genetic analysis
- Fibrosis score of the biceps brachii, assessed by MRI
- Cardiac fibrosis score, assessed by Late Gadolinium Enhancement (LGE)
- Myocardial Circumferential Strain [Global Circumferential Strain (GCS)], assessed by cardiac MRI

Safety assessments will be assessed throughout the study and in accordance with the SoA in the protocol, including but not limited to, adverse events (AEs), concomitant medications, laboratory tests, vital signs, physical exams, etc.

The following assessments will be assessed centrally by independent external vendors or study committees: PFTs, MRI, PUL V2.0. and DVA.

In addition, an independent Data Monitoring Committee (DMC) will review safety and other clinical data (with the authority to unblind such data as needed) on a periodic basis to monitor overall subject safety.

4. STUDY ENDPOINTS

4.1. Primary Endpoint

4.1.1. Function assessment:

- Change in the total score of Performance of Upper Limb [(PUL) version 2.0], from baseline to Week 52

4.2. Secondary Endpoints

- Change in percent predicted forced vital capacity (ppFVC) from baseline to Week 52, assessed by spirometry
- Change in the total score of Mid-level of PUL from baseline to Week 52
- Change in the Grip strength of the hands from baseline to Week 52, assessed by Handheld Myometry (HHM)
- Change in Left Ventricular Ejection Fraction percentage (LVEF %) from baseline to Week 52, assessed by MRI
- **Composite of endpoints**

  Mean observed z score for each of the 3 composite endpoints (PUL, ppFVC), (PUL, ppFVC, Grip Strength), (PUL, ppFVC, Grip Strength, LVEF%).

4.3. Safety Assessments

- All treatment emergent adverse events (TEAEs), treatment emergent serious adverse events (TESAEs), clinically significant laboratory test abnormalities, discontinuation of treatment due to treatment-related AEs, hypersensitivity/anaphylactic reactions, and infusion reactions
- Number and percentage of subjects with hospitalizations due to any serious adverse events with pulmonary and/or cardiac cause(s)
- Number and percentage of subjects with bone fractures
- Ulna length measurements for indirect measure of growth velocity (cm/year) for subjects under the age of 18

4.4. Exploratory Endpoints

- Change in the subscores of the regional dimensions (High-level (Shoulder), Distal-level (Wrist and Hand)) of PUL, from baseline to Week 52
- Change in Duchenne Video Assessment (DVA) severity percentage from baseline to Week 52
- Change in percent predicted Forced Expiratory Volume at 1 second (ppFEV1) from baseline to Week 52

- Change in percent predicted peak expiratory flow (ppPEF) from baseline to Week 52, assessed by spirometry
- Progression of Dilated Cardiomyopathy (DCM) by genetic analysis at Week 52

Fibrosis/MRI assessments:

- Changes in fibrosis score of the biceps brachii, from baseline to Week 52, assessed by MRI
- Changes in cardiac fibrosis score from baseline to Week 52, assessed by Late Gadolinium Enhancement (LGE)
- Change in Myocardial Circumferential Strain [Global Circumferential Strain (GCS)] percentage from baseline to Week 52, assessed by cardiac MRI

4.5. Pharmacokinetics/pharmacodynamics (PK/PD) assessment

- Population PK/PD analysis

5. GENERAL STATISTICAL CONSIDERATIONS

5.1. General Conventions

All analyses will be performed using SAS® Version 9.4 or higher. All data collected will be presented by data listings for review and substantiation of summary tables.

5.1.1. Analysis for Categorical Endpoints

Frequency distributions (number and percentage of subjects) will be presented for categorical variables.

For categorical variables, the confidence interval of proportions for each treatment group will be calculated using the Clopper-Pearson method as appropriate. Bar graphs of proportions for categorical variables will be plotted as appropriate.

5.1.2. Analysis for Continuous Endpoints

For continuous variables, descriptive statistics, including n, mean, standard deviation (SD), median, minimum, and maximum, will be presented. Unless otherwise stated, all confidence intervals will be two-sided 95% confidence intervals.

For selected continuous variables, line graphs of group mean (and standard error (SE)) and mean change from baseline (and SE) values will be plotted over visits.

Baseline characteristics, safety, efficacy, and other data will be summarized by treatment arm based on available data in the mITT/ITT or Safety Analysis Set, except for the parameters specifically indicated otherwise.

Efficacy and safety analyses for treatment comparisons will be based on data observed during the main study period (52 weeks).

The statistical comparisons will be done between treatment groups as deemed appropriate. The treatment groups will be presented as follows:

- Pamrevlumab
- Placebo

5.2. Analysis Sets

The following analysis sets will be used for statistical analysis:

5.2.1. Intent-to-Treat Set (ITT)

The ITT set will include all randomized subjects. Subjects will be analyzed according to their randomized treatment arm.

5.2.2. Modified ITT Population (mITT)

The mITT population enriched for expected change will include subjects from the ITT set with a PUL Entry score >= 2 (excluding PUL entry scores of 1) at baseline.

5.2.3. Per-Protocol Set (PPS)

The per-protocol set will be defined as all randomized subjects who completed at least 24 doses of treatment, with baseline and at least one post-baseline PUL assessments, did not early terminated from treatment or study, and no major protocol deviation(s) that significantly impact efficacy analyses.

5.2.4. Safety Analysis Set (SAF)

The safety analysis set will include all subjects who receive any dose of study medication. All safety data will be analyzed using the SAF. Subjects will be analyzed according to the treatment actually received.

5.3. General Data Handling Rules and Presentation Specifications

The following general guidelines will apply to all statistical analyses and data presentations:

5.3.1. Analysis Period

5.3.1.1. On-Study Period

Unless otherwise specified, the efficacy analysis will be based on the on-study period, which is defined as from randomization to the day of last efficacy assessment for the main study period prior to first study drug infusion in the OLE period for subjects who enrolled in the OLE period; or from randomization to the last day of any efficacy assessment for subjects who didn't enroll in the OLE period.

For the safety analysis, all safety assessments beyond 60 days of last dose will be excluded from summary tables or figures.

5.3.1.2. On-Treatment Period

The on-treatment analysis includes only the assessments that are observed during On-Study Period as defined above (Section 5.3.1.1), or within 4 weeks after last study drug infusion in the main study period. On-treatment analysis as one of the sensitivity analyses will only be conducted for the primary efficacy endpoint.

5.3.2. Baseline and Change from Baseline Definitions

Baseline is defined as the last available value obtained prior to the first dose of study drug (or randomization date if no infusion received) if not specified otherwise (Section 5.5). The acceptable value on Day 1 will be used to define the baseline values. Unscheduled visits will be considered for baseline.

Change from baseline will be calculated by subtracting the baseline value from the post-dose assessment for each subject.

5.3.3. Study Day Calculation

The day when a subject receives the first dose of study drug after randomization will be considered as Day 1 for all analysis.

Study day of an assessment/procedure is calculated as follows.

- For events on or after first dose:

  Study day = Event date – Day 1 date + 1.

- For events earlier than first dose:

  Study day = Event date – Day 1 date.

Note: Day 1 date will be the first dose date if the subject receives study treatment. Otherwise, it will be the randomization date if the subject does not receive the study treatment after randomization.

5.3.4. General Instructions on TFLs

For continuous variables that are recorded as "<X" or ">X", the value of "X" will be used in the calculation of summary statistics.

All percentages will be rounded to one decimal place. The percentage will be suppressed when the count is zero. All durations of time will have 1 decimal place.

All tables and listings will have a header showing "FibroGen, Inc.", protocol number (study nickname), date of data cutoff, and Page x of y. A footer will show the program file path/name, date of data extraction, run date and run time.

More details are available in Appendix 5.

5.3.5. Handling Dropouts and Missing Data

All assessments collected will be considered for analyses regardless of whether such data are collected during treatment or after a subject discontinued treatment. All analyses assume the

missing data are missing at random (MAR), unless stated otherwise. Detailed missing data handling are described in the analysis of specific endpoints.

5.4. Interim Analysis and Data Monitoring Committee

In addition to routine safety monitoring, an independent DMC is established to review safety data on an ongoing basis. A DMC charter will establish the procedures, meeting frequency, and scope of responsibilities of the committee.

This study has no planned or pre-specified interim analysis for either efficacy or futility.

5.5. Analysis Visit Windows

In accordance with FDA/EMA Guidance: During the COVID-19 pandemic, the visit modality (e.g.: in-person versus remotely) and scheduling may be adjusted and conducted at the discretion of the Investigator, in accordance with the site's rules and recommendations, and using all necessary precautions. If a visit must be done remotely, it can be conducted with any technology available to the site and study subjects, such as via tele-health visits, phone calls, etc.

Analysis visits, instead of the nominal visits from case report form (CRF), derived from visit dates and visit time windows will be used in the by-visit analyses. Unscheduled visits within a visit window (defined in Appendix 2) will be grouped into the closest scheduled visits based on the visit date. For subjects who have more than one measurement at a certain analysis visit, the measurement closest to the target date or the later assessment if there is a tie will be used, with the following exceptions:

- Liver function tests, such as Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Gamma-Glutamyl transferase (GGT), Alkaline Phosphatase (ALP), and total bilirubin, in which the maximum value will be used.
- Vital signs: the values on the infusion days will be used if applicable.
- Spirometry data: the Best Test Read (BTR) values will be used for analysis.

Efficacy parameters will be summarized by analysis visit defined by the following assessment windows (Appendix 2). The date of the first dose will be considered as the date of Day 1 for all analyses. Subjects will receive study drug every 2 weeks. The visit window for these visits is ±3 days. Duchenne Video Assessment is to be performed in the subject's home up to 14 days prior to the scheduled study visit.

6. STATISTICAL ANALYSES

6.1. Subject Enrollment and Disposition

6.1.1. Eligibility

Eligibility will be summarized for all screened subjects. The data will be summarized with respect to:

- number of subjects screened

- number (%) of subjects screen-failed
- number (%) of subjects for each failed inclusion/exclusion criterion

Subject level inclusion criteria not met/exclusion criteria met listings will be provided.

6.1.2. Subject Accountability and Disposition

The number (%) of subjects randomized (ITT), dosed (SAF), mITT, Per-Protocol Set (PPS), completed the main study, entered the OLE period, and discontinued prematurely from treatment and study in the main study will be presented for each treatment group and for all subjects pooled.

Reasons for premature discontinuation of treatment in the main study will be summarized by treatment group for the ITT population. A listing for the subjects who discontinue treatment prematurely in the main study will be presented.

6.2. Important Protocol Deviations

Important protocol deviations of interest may include, but are not limited to, the following:

- Entry Deviation: Subjects who entered the study, but did not meet inclusion/exclusion criteria;
- Withdrawal Deviation: Subject met withdrawal criteria during the study but was not withdrawn;
- Dosing Deviation: Subject received the wrong treatment or incorrect dose; including incorrect timing of a dose; or subjects who missed more than 10% doses of prescribed study medication during the overall treatment period;
- Prohibited Medication Deviation: Subject received disallowed concomitant medications or non-drug therapy.

All above protocol deviations will go through the medical review process.

A subset of pre-specified important protocol deviations will be defined as major protocol deviations and subjects with major protocol deviations will be excluded from PPS analysis. The number and percentage of subjects with important protocol deviations will be categorized and tabulated as appropriate for the ITT population. COVID-19 related important protocol deviations will be summarized separately. All protocol deviations will be finalized prior to database lock and unblinding.

6.3. Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized descriptively for subjects in mITT, ITT and SAF. Each parameter will be presented in data listings.

Baseline values for efficacy assessments will be presented in baseline tables as appropriate.

6.3.1. Demographics and Baseline Characteristics

Demographic parameters and other important baseline and disease characteristics will be summarized by treatment group. These include but may not be limited to age, age group (≤16

years, >16 years), ethnicity, race, weight, body-mass index (BMI), side of dominant hand, and body surface area (BSA), PUL 2.0 total score and subscores.

Computation formula:

BSA = [Weight $^{0.425}$ (kg) * Height $^{0.725}$ (cm)] x 0.007184
Age = Year of informed consent date – Year of birth
Height (cm) = 4.605 * Ulna length (cm) +1.308 * Age at PFT assessment+28.003, where age < 18 years old during PFT visits; For subjects that enter the study at age 18 or turn 18 during the study, the last recorded ulna length/height will be used for the remainder of the study.

Descriptive statistics (n, mean, standard deviation, median, interquartile (Q1, Q3), range (=minimum, maximum)) will be presented for continuous variables. Frequency distributions (number and percentage of subjects) will be presented for categorical variables.

6.3.2. Baseline DMD Disease Characteristics

Baseline DMD disease characteristics include, but not limited to,:

1. Age in years when DMD was diagnosed = Year of date of diagnosis - Year of birth
2. Age when subject became non-ambulant = Year of date of when subject became non-ambulant - Year of birth
3. Years since subject became non-ambulant = Current age – Age when subject became non-ambulant
4. Age when subject began ventilation support = Year of date of subject began ventilation support- Year of birth
5. Ventilation support: yes or no. If yes, years since subject began ventilation support = Current age– Age when subject began ventilation support.
6. Corticosteroids use: yes or no. If yes, age when subject began corticosteroids (= Year of date of subject began corticosteroids- Year of birth), and years since subject began corticosteroids = Current age – Age when subject began corticosteroids.
7. Spine surgery: yes or no. If yes, age when spine surgery (= Year of date of subject began spine surgery- Year of birth) was done, and years since the most recent spine surgery = Year of Day 1 Visit Date - Year of spine surgery,
8. Genetic characteristics (check all that apply)
   - Exonic deletion
   - Duplication
   - Point mutation
   - None of above
   - Unknown

Additionally, the following baseline disease characteristics, but are not limited to, will be summarized.

9. Bone parameters
    - Bone fractures (Section 6.8.8)
    - Scoliosis
    - Osteoporosis
    - Spine surgery
10. Cardiac parameters
    - Cardiomyopathy
    - LVEF (%)
11. Respiratory parameters
    - Ventilation support
    - Years since ventilation support
    - Sleep apnea
    - ppFVC
12. Metabolic parameters
    - Insulin resistance
    - Obesity
    - Vit D deficiency
    - Delayed puberty
13. Corticosteroids (CS) use
    - Age at CS initiation
    - Prednisone/Prednisolone
    - Deflazacort
    - Regimen
        - Daily
        - Other

6.3.3. Medical History

Medical History of interest including allergies and surgeries, coded in system organ class (SOC) and preferred term (PT) using MedDRA (version 26.0 or higher), will be summarized for the ITT Analysis Set. Subject medical history listings will be provided.

6.4. Prior and Concomitant Medications

The World Health Organization Drug Dictionary (WHODD) Version (WHO Drug March 1, 2020, or later version) will be used to classify prior and concomitant medications by therapeutic class and generic name based on Anatomical Therapeutic Chemical (ATC) code level 3.

Prior medication is defined as any medication taken and stopped prior to the first infusion of the study medication. Concomitant medication is defined as any medication used concomitantly with the study drug that were not stopped before the first infusion or used concomitantly after the first infusion, ending 60 days after the last treatment. Partially or incomplete missing prior/concomitant medication start or stop date will be imputed (Appendix 1).

Both prior and concomitant medication usage will be summarized by the number (%) of subjects receiving the drug within each therapeutic class and ATC code level 3 and preferred term for the SAF. Multiple usage of the same drug by a patient will be counted only once.

In addition, a listing of all medications (prior or concomitant) taken during the course of the study (from screening through the end of study) captured in Concomitant Medication CRF, as well as Non-Drug Therapies CRF will be provided.

Data on concomitant DMD medications administered during the study are collected and analyzed accordingly. Summary of concomitant DMD medications such as types of corticosteroids and duration of corticosteroid will be provided.

To evaluate the use of co-administered medications that are narrow therapeutic CYP substrates, the clinical database of concomitant medication will be searched for presence of narrow therapeutic index cytochrome P450 substrates and summarized. The list of related medications will be finalized before database lock.

6.5. Study Drug Exposure and Treatment Compliance

6.5.1. Study Drug Exposure

The number and percentage of subjects who receive study medication will be summarized by treatment group for the SAF.

Duration of weekly exposure/treatment is calculated as:

Weeks in treatment = (last dose date - first dose date + 1)/7 (keep 1 decimal place).

Duration of weekly exposure will be tabulated by treatment group for the Safety population.

The number of infusions and average infusion dose amount in mg and mg/kg, any interruption during infusion (Y/N), and reason for missed dose or interruption will be summarized for the SAF analysis set.

6.5.2. Treatment Compliance

Dosing for pamrevlumab (or placebo) will be based on subject's body weight, which will be measured at Screening and every 12 weeks thereafter to determine dose for the subsequent 12-week interval. The total dose of pamrevlumab (or placebo) is not to exceed 4.1g per infusion or

to align with the protocol. Subjects weighing more than 117 kg will receive the maximum allowable dose of 4.1g.

Compliance will be calculated as the total amount of dosage (mg) the subject received divided by the total amount of dosage (mg) the subject is scheduled to receive during the participation in treatment.

$$Compliance\ (\%) = \frac{\text{actual total dosage received (mg)}}{\text{total planned dosage (mg) while actively in treatment}} \times 100$$

Descriptive statistics for study medication compliance will be presented by treatment group for the SAF. Treatment compliance will be summarized as a continuous variable and as a categorical variable (<70%, 70% –<80%, 80% –<90%, and 90% - 100%, >100%).

6.6. Efficacy Analyses

The primary and secondary endpoints will be tested using a fixed sequence procedure to preserve the study-wide error rate of 5%. Under the sequential analysis, the primary and secondary efficacy endpoints will be tested in a predefined sequence according to the order listed in Table 1 each at the usual alpha= 0.05 level of statistical significance. The testing will cease when a failure occurs in the pre-determined sequential hypothesis testing and all p-values for the subsequent testing will be considered nominal.

All analyses for efficacy endpoints will be performed for the mITT during the On-Study period (Section 5.3.1.1), unless noted otherwise. Sensitivity analysis on ITT will also be performed on primary and secondary endpoints as supportive evidence.

Table 1: Testing Sequence of Primary and Secondary Endpoints

Testing Sequence	Endpoint
1 (primary endpoint)	Change in the total score of Performance of Upper Limb [(PUL) version 2.0], from baseline to Week 52.
2	Change in percent predicted Forced Vital Capacity (ppFVC) from baseline to Week 52, assessed by spirometry.
3	Change in the total score of Mid-level of PUL from baseline to Week 52
4	Change in the grip strength of the hands from baseline to Week 52, assessed by Hand-Held Myometry (HHM).
5	Change in Left Ventricular Ejection Fraction percentage (LVEF %) from baseline to Week 52, assessed by MRI.

6.6.1. Primary Endpoint Estimands and Analyses

The primary efficacy endpoint in this study is defined as Change from Baseline in the total score of Performance of Upper Limb [(PUL) version 2.0] at Week 52.

PUL 2.0 scale:

- The PUL 2.0 scale, with maximum total score = 42

- High Level: Shoulder Dimension, 6 items: Max score = 12
- Mid-level: Elbow Dimension, 9 items: Max score = 17
- Distal level: Wrist and Hand Dimension, 7 items: Max score = 13

Only 1 item in each domain is allowed missing. The score of the missing item will be imputed by the mean of the non-missing items in the domain. If the total score of any domain is missing (the high-level domain only applies to subjects with Item A score ≥ 3), then the total score will be missing. The total score of PUL is the sum of the total scores of the three domains (if Item A score ≥ 3) or two domains (Mid-level and Distal level, if Item A score <3).

The last measurement prior to first dose of study drug from Muscle Function Tests (MFTs) at Day 1 will be used to establish baseline.

The primary efficacy endpoint will be analyzed using mITT as the primary analysis set.

The hypothesis to be tested for the primary efficacy analysis is:

H_0: Change from Baseline in the total score of PUL at Week 52 for the pamrevlumab arm = Change from Baseline in the total score of PUL at Week 52 for the placebo group

Versus:

H_1: Change from Baseline in the total score of PUL at Week 52 for the pamrevlumab arm ≠ Change from Baseline in the total score of PUL at Week 52 for the placebo group

H_0 will be tested at the two-sided alpha = 0.05 level of significance and will be rejected if the p < 0.05 from the test.

6.6.1.1. Primary Analysis with the Random Coefficient Model (RCM)

6.6.1.1.1. Estimand Strategy

The primary estimand is intended to provide a population level estimate of the treatment effect of the pamrevlumab on a continuous endpoint, regardless of participant compliance with the IP dosing. Treatment policy strategy will be implemented, which means all observed measurements after randomization will be used in the primary analysis.

6.6.1.1.2. Population of Interest

The mITT population will used as defined in Section 5.2.2 .

6.6.1.1.3. Intercurrent Event Handling

Treatment discontinuation (such as due to adverse event, lost to follow-up, withdrawal by subject, physician decision, protocol deviations, etc.): Missing data will be implicitly imputed under missing at random assumption.

Death: The worst postbaseline score of all patients in the mITT population would be used to impute all missing values due to death. Missing study day will be imputed as target day of the corresponding analysis visit.

6.6.1.1.4. Analysis Variable

Change from baseline in the total score of all PUL assessments in the randomization treatment period, including scheduled, unscheduled, and available assessments after treatment discontinuation, up to the targeted week 52 during the On-Study period (Section 5.3.1.1), will be included in the analysis.

6.6.1.1.5. Population Summary for Treatment Comparison

Treatment difference of Least-square mean (LSMean) and SE at week 52 and corresponding 95% CI will be presented.

The random coefficient linear regression model (RCM), where within-patient errors follow a random coefficient regression model will be performed. The RCM model will include fixed effects of time, treatment, treatment-by-time interaction, with baseline ordinal PUL entry score as covariate. The time in the RCM model will be calculated as the elapsed weeks (considered as a continuous variable with 1 decimal place) from first dose date (or randomization date for subjects who drop out prior to any study medication) to the assessment date. All valid values including from unscheduled visits will be included.
The residual plot from the PROC MIXED output will be examined. If the residual plot shows a systematic pattern rather than a random scatter, then the linearity assumption may not be appropriate. In case of non-linearity, MMRM will be used as the primary analysis method, which is currently listed as one of the sensitivity analyses in next section.

Covariance Structure Strategy

The unstructured covariance structure for the within-patient errors in the model will be applied first. The by-treatment-group option is included to the covariance pattern to improve the model efficiency.

If the algorithm for unstructured covariance pattern does not converge, the following covariance structures will be tested in sequence until the model converges: heterogeneous Toeplitz, homogeneous Toeplitz, first-order autoregressive, compound symmetry, and variance component. The sandwich estimator will be used if there is convergence issue. When the random effects model does not converge or does not have a positive definite Hessian, the fixed effects instead of random effects will be performed instead. The Least Square Mean (LSMean) (SE) and 95% CI estimated changes from baseline to Week 52 will be presented.

Sample SAS code:

```
PROC MIXED DATA = XXX;
  CLASS TRTPN SUBJID ItemAgrp;
  MODEL CHG = TRTPN AWK TRTPN*AWK ItemAgrp / RESIDUAL CL COVB DDFM=KR;
  RANDOM AWK / SUBJECT =SUBJID TYPE=UN GROUP= TRTPN;
  LSMEANS TRTPN / at AWK = 52 PDIFF CL;
  ODS OUTPUT LSMEANS=MIXLSMNS DIFFS=MISDIFS;
RUN;
```

(Note: ItemAgrp denotes the categorical Item A (or PUL entry) scores at baseline)

6.6.1.2. Sensitivity Analyses of the Primary Endpoint

6.6.1.2.1. MMRM Analysis for Primary Efficacy Endpoint

6.6.1.2.1.1. Estimand Strategy

Same as Section 6.6.1.1.1.

6.6.1.2.1.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.1.2.1.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3.

6.6.1.2.1.4. Analysis Variable

Same as Section 6.6.1.1.4.

6.6.1.2.1.5. Population Summary for Treatment Comparison

All post-baseline visits during the double-blind period will be included in the analysis. Change from baseline for the specified endpoints (excluding baseline visits) will be analyzed using a Mixed Model for Repeated Measures (MMRM) approach with fixed effects for treatment, visit (as a factor), treatment-by-visit interaction, and covariates (baseline values). The covariance structure strategy is the same as described in Section 6.6.1.1.5.

Sample SAS code:

```
PROC MIXED DATA = XXX;
  CLASS TRTPN SUBJID AVISITN ItemAgrp;
  MODEL CHG = TRTPN AVISITN TRTPN* AVISITN ItemAgrp / DDFM=KR;
  REPEATED AVISITN / SUBJECT =SUBJID TYPE=UN GROUP= TRTPN;
  LSMEANS TRTPN / at AVISITN = 52 PDIFF CL;
  ODS OUTPUT LSMEANS=MIXLSMNS DIFFS=MISDIFS;
RUN;
```

6.6.1.2.2. ITT for Primary Efficacy Endpoint

6.6.1.2.2.1. Estimand Strategy

Same as Section 6.6.1.1.1.

6.6.1.2.2.2. Population of Interest

ITT population will be used as defined in Section 5.2.1.

6.6.1.2.2.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3.

6.6.1.2.2.4. Analysis Variable

Same as Section 6.6.1.1.4.

6.6.1.2.2.5. Population Summary for Treatment Comparison

Same as Section 6.6.1.1.5.

6.6.1.2.3. MI ANCOVA

Multiple Imputation Analysis of Covariance model (MI ANCOVA) will be performed to evaluate the robustness of the primary analysis as a sensitivity analysis.

6.6.1.2.3.1. Estimand Strategy

Same as Section 6.6.1.1.1

6.6.1.2.3.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.1.2.3.3. Intercurrent Event Handling

Under the MAR assumption, subjects who discontinued from the treatment early alive are assumed to have the same data pattern as subjects who remain in the study for the same treatment arm. Multiple imputation will be carried out to impute missing data for visits up to Week 52.

Death: Same as Section 6.6.1.1.3.

6.6.1.2.3.4. Analysis Variable

Same as Section 6.6.1.1.4 (Missing value imputation rules are defined below).

6.6.1.2.3.5. Population Summary for Treatment Comparison

The combined treatment difference of LSMean (and SE) and the corresponding 95% CI for the estimated change from baseline in PUL total score at Week 52 will be presented.

Analysis will be performed in 3 steps.

Step 1a - First, the intermittent missing PUL data will be imputed for each treatment relying on non-missing data from all subjects within each treatment group using the Markov Chain Monte Carlo (MCMC) imputation model with baseline PUL and other baseline factors, and the available non-missing PUL to generate a monotone missing data pattern that only has missing data at end for planned visits through Week 52. The reported (observed) value will be used for multiple imputation. Change from baseline values will be derived using the difference between the imputed or observed post-baseline values and baseline values.

Step 1b - Then, for each of the 200 datasets with monotone missing data generated above, missing data at end through Week 52 will be imputed to derive 200 imputed datasets with non-missing data.

Sample SAS code:

```
/*examine missing patterns*/
```

```
PROC MI DATA = INDAT NIMPUTE = 0;
        BY TRTPN;
    VAR BASE WEEK16 WEEK28 WEEK40 WEEK52;
RUN;
```

Suppose the output from PROC MI above indicates that the missing pattern is non-monotone, it is necessary to perform Step 1a, which is Partial imputation (just enough to get the monotone missing pattern).

```
/*partial imputation to get monotone missing pattern*/
PROC MI DATA= INDAT OUT=xx1 SEED=9978991 NIMPUTE=200 OUT = MONO;
        BY TRTPN;
        VAR BASE WEEK16 WEEK28 WEEK40 WEEK52;
        MCMC CHAIN=multiple IMPUTE=monotone;
RUN;

/*The above procedure will output MONO dataset with a monotone missing data pattern. */

PROC SORT DATA= MONO OUT =XX1;
        BY _IMPUTATION_;
RUN;

PROC MI DATA=XX1 OUT=XX SEED=9978 NIMPUTE=1;
        BY _IMPUTATION_;
        VAR TRTPN BASE WEEK16 WEEK28 WEEK40 WEEK52;
        CLASS TRTPN;
        MONOTONE REGRESSION (/details);
RUN;

DATA MIANCOVA;
        SET XX;
        CHG_WK52= WEEK52 - BASE;
RUN;
```

Step 2 - The 200 multiple-imputation datasets with imputed and observed PUL data at Week 52 will be analyzed separately for each imputation using the ANCOVA method. The ANCOVA model will contain terms for treatment, baseline PUL measurements. The LSMean and corresponding SE for the change from baseline in PUL total score at Week 52 will be estimated.

Sample SAS code:

```
PROC MIXED data= MIANCOVA;
        BY _IMPUTATION_
        CLASS TRTPN ItemAgrp;
        MODEL CHG_WK52 = TRTPN ItemAgrp COVARIATES/SOLUTION DDFM=KR;
        LSMEANS TRTPN / PDIFF CL;
        ODS OUTPUT DIFFS=LSDIFS LSMEANS=LSM SOLUTIONF=PARMS;
RUN;
```

Step 3 - The SAS PROC MIANALYZE will be used to derive the final estimates and test statistics summarizing the 200 analysis results.

Sample SAS code:

```
PROC SORT DATA=LSM;
        BY TRTPN;
RUN;

PROC MIANALYZE DATA=LSM;
        BY TRTPN;
        MODELEFFECTS ESTIMATE;
        STDERR;
        ODS OUTPUT ParameterEstimates=LSM2;
RUN;

PROC MIANALYZE DATA= LSDIFS;
        MODELEFFECTS ESTIMATE;
        STDERR;
        ODS OUTPUT ParameterEstimates=lsdiffs2;
RUN;
```

6.6.1.2.4. Pattern Mixture Models under MNAR for Missing Data: Jump-to-Control Analysis

The goal of the following analyses is to address the possibility of data being missing not at random (MNAR).

6.6.1.2.4.1. Estimand Strategy

Same as Section 6.6.1.1.1

6.6.1.2.4.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.1.2.4.3. Intercurrent Event Handling

Treatment discontinuation (such as due to AEs, lost to follow-up, withdrawal by subject, physician decision, protocol deviations, etc.): The missing data pattern for the pamrevlumab subjects after withdrawal from the study can be assumed to switch to the same data pattern as subjects on the placebo treatment. Subjects that discontinued from the placebo arm are assumed to have the same data pattern as placebo subjects that remain in the study. This is often called the jump-to-control approach.

Death: same as Section 6.6.1.1.3.

6.6.1.2.4.4. Analysis Variable

Change from baseline in the total score of all PUL assessments at Week 52 (Missing value imputation rules are defined below).

6.6.1.2.4.5. Population summary for treatment comparison

Under the jump-to-control assumption, the analysis will be carried out in 3 steps.

Step 1 - the missing PUL data will be imputed to derive 200 imputed datasets with non-missing data according to the jump-to-control data pattern.

Sample SAS code:

```
PROC MI DATA=aval OUT=xx SEED=8767892 NIMPUTE=200;
        CLASS TRTP;
        VAR BASE WEEK16 WEEK28 WEEK40 WEEK52;
        MONOTONE REG(/details);
        MNAR MODEL (WEEK52/MODELOBS=(TRTP='Placebo'));
RUN;
```

Step 2 - The 200 multiple imputation datasets with imputed and observed data will be analyzed separately using the ANCOVA method separately as described in Section 6.6.1.2.3.5. The LSMean and corresponding SE for the change from baseline in PUL total score at Week 52 will be estimated.

Step 3 - The SAS PROC MIANALYZE will be used to derive the final estimates and test statistics summarizing the 200 dataset results.

The combined treatment difference of LSMean, corresponding SE and 95% CI for the estimated change from baseline in PUL total score at week 52 will be presented.

6.6.1.2.5. Pattern Mixture Models under MNAR for Missing Data: Delta-Adjusting (Tipping Point) Analysis

6.6.1.2.5.1. Estimand Strategy

Same as Section 6.6.1.1.1

6.6.1.2.5.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.1.2.5.3. Intercurrent Event Handling

Treatment discontinuation (such as due to AEs, lost to follow-up, withdrawal by subject, physician decision, protocol deviations, etc.): An alternative assumption is that the missing data for the pamrevlumab treated subjects who discontinue early have a lower expected value than the pamrevlumab subjects remaining in the study, while subjects who discontinue from the placebo arm are assumed to have the same data pattern as placebo subjects remaining in the study. This is often called the delta-adjusting (or tipping point) approach.

Death: same as Section 6.6.1.1.3.

6.6.1.2.5.4. Analysis Variable

Change from baseline in the total score of all PUL assessments at Week 52 (Missing value imputation rules are defined below).

6.6.1.2.5.5. Population summary for treatment comparison

The integrated treatment difference of LSMean, corresponding SE and 95% CI for the change from baseline in PUL total score at week 52 will be presented.

The multiple imputation analysis will be performed as follows.

Step 1 - the missing PUL data will be imputed to derive 200 imputed datasets with non-missing data according to the delta-adjusting data approach.

Sample SAS code:

```
PROC MI DATA=aval OUT=zz SEED=9876783 NIMPUTE=200;
        CLASS TRTP;
        VAR TRTP BASE WEEK16 WEEK28 WEEK40 WEEK52;
        MONOTONE REG;
        MNAR ADJUST (WEEK16 /SHIFT=-#.# ADJUSTOBS =(TRTP='pamrevlumab'));
        MNAR ADJUST (WEEK28 /SHIFT=-#.# ADJUSTOBS =(TRTP='pamrevlumab'));
        MNAR ADJUST (WEEK40 /SHIFT=-#.# ADJUSTOBS =(TRTP='pamrevlumab'));
        MNAR ADJUST (WEEK52 /SHIFT=-#.# ADJUSTOBS =(TRTP='pamrevlumab'));
RUN;
```

Step 2- The observed and imputed PUL data will be analyzed using the ANCOVA method separately as described in Section 6.6.1.2.3.5 for each of the 200 imputed datasets. The LSMean and corresponding SE for the change from baseline in PUL at week 52 will be estimated.

Step 3- The SAS PROC MIANALYZE will be used to summarize the final estimates and test statistics from the 200 dataset results.

6.6.1.2.6. On-treatment Analysis for Primary Efficacy Endpoint

A sensitivity analysis will be performed for PUL assessed during on-treatment period as defined in Section 5.3.1.2.

6.6.1.2.6.1. Estimand Strategy

Same as Section 6.6.1.1.1

6.6.1.2.6.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.1.2.6.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3

6.6.1.2.6.4. Analysis Variable

Only on-treatment PUL assessments will be included in the analysis

6.6.1.2.6.5. Population Summary for Treatment Comparison

If linear assumption was met, the RCM linear regression model as described in Section 6.6.1.1.5 will be performed; otherwise, ANCOVA model as described in Section 6.6.1.2.3.5 will be performed.

6.6.1.2.7. Per-Protocol Set Analysis for Primary Efficacy Endpoint

6.6.1.2.7.1. Estimand Strategy

Same as Section 6.6.1.1.1.

6.6.1.2.7.2. Population of Interest

PPS as defined in Section 5.2.3.

6.6.1.2.7.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3.

6.6.1.2.7.4. Analysis Variable

Same as Section 6.6.1.1.4.

6.6.1.2.7.5. Population Summary for Treatment Comparison

If the linear assumption was met, the RCM linear regression model as described in Section 6.6.1.1.5 will be performed; otherwise, the ANCOVA model as described in Section 6.6.1.2.3.5 will be performed.

6.6.1.2.8. Cumulative Loss of Function Analysis for Primary Efficacy Endpoint

This sensitivity analysis will use data representing cumulative failure to perform 22 items of PUL 2.0 in patients at multiple time points over 52 weeks.

6.6.1.2.8.1. Estimand Strategy

Failure to perform an item is defined as a score transition from 2 or 1 to 0 at evaluation. Only post-baseline failures will be considered in this analysis. For each treatment group, the mean of all such patients' individual curves will be constructed, reflecting the average cumulative number of failures over time. The higher the curve, the worse the study treatment effect. To quantify the group difference, we assume that the ratio of two corresponding underlying group-wise curves is constant over time, and use the Lin, Wei, Yang and Ying (LWYY) [17] analytic method, which is the same as the Anderson-Gill method with the robust variance estimate, to estimate this constant ratio. This summary measure can also be interpreted as the ratio of the intensities of occurrences of failures over time. If the assumption of constant ratio is not plausible, the resulting estimate reflects an averaged ratio between the two curves over time. The lower the ratio (Pamrevlumab over Placebo), the greater the treatment effect.

6.6.1.2.8.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.1.2.8.3. Intercurrent Event Handling Strategy

If a patient discontinued, the cumulative failure number will be censored upon discontinuation.

6.6.1.2.8.4. Analysis Variable

Cumulative counts of failure to perform 22 items of PUL 2.0 at each protocol specified timepoint.

6.6.1.2.8.5. Population Summary for Treatment Comparison

All post-baseline visits during the double-blind period will be included in the analysis. Cumulative counts of failure to perform 22 items of PUL 2.0 at post-baseline visits will be analyzed using the LWYY analytic method to estimate the ratio between treatment arms. Rate ratio between 2 treatment arms including 95% CI and p-value will be reported. Sample SAS code is provided below.

```
PROC PHREG DATA=FINAL COVS(AGGREGATE);
  MODEL (TSTART, TSTOP) * STATUS(1) = TRT NUMBER;
  ID SUBJID;
  WHERE TSTART < TSTOP;
RUN;
```

Note: 1= censored. For more information, please refer to SAS online help.

6.6.2. Secondary Endpoint Estimands and Analyses

The secondary endpoints during the on-study period (as defined in Section 5.3.1.1) will be analyzed in the order specified in Table 1.

6.6.2.1. Continuous Secondary Endpoints with RCM Model

6.6.2.1.1.1. Estimand Strategy

Same as Section 6.6.1.1.1

6.6.2.1.1.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.2.1.1.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3

6.6.2.1.2. Analysis Variables

6.6.2.1.2.1. Change in Percent Predicted Forced Vital Capacity (ppFVC) from baseline to Week 52, assessed by spirometry

6.6.2.1.3. Change in the total score of Mid-level of PUL from baseline to Week 52Population Summary for Treatment Comparison

Same as Section 6.6.1.1.5.

6.6.2.2. Continuous Secondary Endpoints with MMRM Model

6.6.2.2.1.1. Estimand Strategy

Same as Section 6.6.1.1.1

6.6.2.2.1.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.2.2.1.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3

6.6.2.2.2. Analysis Variables

6.6.2.2.2.1. Absolute Change from Baseline in Grip Strength at Week 52

The absolute change from baseline in the grip strength of the hands is derived with the formula below:

$$grip\ strength\ of\ the\ hands\ value\ at\ Post\ baseline\ visit - grip\ strength\ of\ the\ hands\ value\ at\ Baseline$$

The grip strength should be reported by dominant and non-dominant hand separately.

6.6.2.2.3. Covariance Structure Strategy

The unstructured covariance structure for the within-patient errors in the model will be applied first. The by-treatment-group option may will be added to the covariance pattern to improve the model fit as appropriate.

If the algorithm for unstructured covariance pattern does not converge, the following covariance structures will be tested in sequence until the model converges: heterogeneous Toeplitz, homogeneous Toeplitz, first-order autoregressive, compound symmetry, and variance component, until the model converges. The sandwich estimator will be used if there is convergence issue. If the model doesn't converge for all covariance structures listed above, some least significant factors or interaction terms (p>0.05) can be excluded from the model to achieve convergence. The revised model with fewer factors or interaction terms will be tested using the same sequence as specified above.

Sample SAS code:

```
/* excluding baseline visit */
  PROC MIXED DATA=MMRM (WHERE=(AVISITN>0));
        CLASS TRTP AVISIT;
        MODEL CHG = TRTP AVISIT TRTP*AVISIT BASE /DDFM=KR;
        REPEATED AVISIT /SUBJECT=SUBJID TYPE=UN GROUP=TRTP;
        LSMEANS TRTP /PDIFF CL;
        LSMEANS TRTP*AVISIT /PDIFF CL;
        ODS OUTPUT LSMEANS=MIXLSMNS DIFFS=MIXDIFS;
  RUN;
```

6.6.2.3. Continuous Secondary Endpoints with ANCOVA Model

6.6.2.3.1.1. Estimand Strategy

Same as Section 6.6.1.1.1

6.6.2.3.1.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.2.3.1.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3

6.6.2.3.2. Analysis Variables

6.6.2.3.2.1. Change in Left Ventricular Ejection Fraction Percentage (LVEF %) from Baseline to Week 52, Assessed by MRI

6.6.2.3.3. Population Summary for Treatment Comparison

The change in the LVEF% from baseline to Week 52 will be analyzed using an ANCOVA model, which contains terms for treatment, baseline LVEF measurements. The LSMean and corresponding SE for the change from baseline in LVEF at Week 52 will be estimated.

Sample SAS code:

```
/* excluding baseline visit */
  PROC MIXED DATA=ANCOVA METHOD=ML;
        CLASS TRTP;
        MODEL CHG = TRTP BASE / DDFM = KR;
        LSMEANS TRTP /PDIFF CL;
  RUN;
```

6.6.2.4. Analysis of A Composite of Endpoints

6.6.2.4.1.1. Estimand Strategy

To evaluate the totality of the treatment effect in multiple muscle groups (i.e., skeletal, respiratory, and cardiac muscle), a composite endpoint will be used to combine the results from multiple outcomes including primary endpoint and 3 key secondary endpoints (PUL, ppFVC, Grip Strength, LVEF%) as a secondary analysis following the method introduced by Li, et al in 2020.

6.6.2.4.1.2. Population of Interest

Same as Section 6.6.1.1.2

6.6.2.4.1.3. Intercurrent Event Handling Strategy

Same as Section 6.6.1.1.3

6.6.2.4.2. Analysis Variables

The z-score is defined as LSmean of the treatment difference divided by its corresponding SE at week 52 from the primary analysis model. The mean z-score will be calculated sequentially across 4 endpoints (PUL, ppFVC, Grip Strength, LVEF%), namely, the mean z-scores for PUL + ppFVC, PUL + ppFVC + Grip Strength, PUL + ppFVC + Grip Strength + LVEF will be reported.

6.6.2.4.3. Population Summary for Treatment Comparison

The mean z-score will be calculated sequentially across 4 endpoints (PUL, ppFVC, Grip Strength, LVEF%) and p-value will be derived by a permutation method to rerandomize patient vectors of the observed values from the 4 endpoints in 1000 trials. A distribution of the mean z-score will be derived from these 1000 trials. A p-value will be computed from the z-score from this clinical trial against the backdrop of the random distribution of the 1000-trial z-scores. p-values corresponding to the z-scores for PUL + ppFVC, PUL + ppFVC + Grip Strength, PUL + ppFVC + Grip Strength + LVEF will be reported. If the p-value is < 0.025 (1-sided), the composite endpoint is deemed to be statistically significant.

The method was introduced by Li, et al in 2020 [16].

6.6.2.5. Assessments by ITT populations of secondary endpoints:

Assessments by ITT populations of secondary endpoints will be considered as sensitivity analysis for these secondary endpoints. Summary of the following endpoints will also be conducted on the ITT population. The change from baseline to Week 52 in these endpoints will be summarized by analysis visit and analyzed using the RCM or MMRM or ANCOVA model as described in corresponding sections in Section 6.6.2 .

- Change in percent predicted forced vital capacity (ppFVC) from baseline to Week 52, assessed by spirometry
- Change in the total score of Mid-level of PUL from baseline to Week 52
- Change in the Grip strength of the hands from baseline to Week 52, assessed by Handheld Myometry (HHM)
- Change in Left Ventricular Ejection Fraction percentage (LVEF %) from baseline to Week 52, assessed by MRI
- Sequential composites of endpoints - PUL + ppFVC, PUL + ppFVC + Grip Strength, PUL + ppFVC + Grip Strength + LVEF

6.6.3. Exploratory Analyses

6.6.3.1. Change in the Subscores of Regional Dimensions (High-level (Shoulder), Distal-level (Wrist and Hand)) of PUL, from Baseline to Week 52

Summary of the listings and analyses of subscores of regional dimensions (High-level (shoulder), Distal-level (wrist and hand)) of PUL will be summarized for subjects in the ITT population.

The change from baseline to Week 52 in subscores of regional dimensions (High-level (shoulder), Distal-level (wrist and hand)) of PUL, will be analyzed using the same analysis method as the one for the primary endpoint.

6.6.3.2. Change in Duchenne Video Assessment Severity Percentage from Baseline to Week 52

The Duchenne Video Assessment (DVA) tool provides a standardized way to document and assess patient quality of movement. Caregivers will video record subjects doing specific movement tasks at home using a secure mobile application within 2 weeks prior to the clinic visits specified in the Schedule of Assessments. Trained physical therapists will score the videos in a secure scoring dashboard using scorecards with pre-specified compensatory movement criteria. Casimir, the organization that developed the DVA, will manage DVA data collection, quality, and scoring.

Summary of the listings and analyses Duchenne Video Assessment severity percentage and CGI-C scores will be based on the ITT population. The change from baseline to Week 52 in Duchenne Video Assessment severity percentage will be summarized by analysis visit and analyzed using the MMRM model as described above.

The CGI-C will be analyzed using the proportional odds model, a special case of generalized linear mixed models (GLMM). The GLMM model will include treatment as fixed effect. The model will be fitted using GLMM with a multinomial distribution and a cumulative logit link function. The proportional odds ratio estimate between Pamrevlumab and placebo, the corresponding 95% CI and P-value will be presented. If the proportional odds assumption is contradicted ($p < 0.05$ in Score Test for the Proportional Odds Assumption) by the observed data, the GLMM method will be replaced by the nonparametric van Elteren test to test whether there is a difference in the location of distribution of the CGI-C at 52 weeks between Pamrevlumab and placebo.

6.6.3.3. Change in Percent Predicted Forced Expiratory Volume at 1 Second (ppFEV1) from Baseline to Week 52

Standardized PFTs including FVC, PEF, FEV1 will be collected for consistency during the course of the study. The PFT vendor will provide the SpiroSphere®. Within a PFT session, eight attempts may be made. Only sessions with Best Test Read (BTR) grade of borderline or acceptable are considered valid. The session may be repeated up to three times during the screening window. The Day 1 PFT session may also be repeated up to three times. If sessions are repeated, the last session conducted for screening, and the last session conducted for Day 1 supersede all others and are used to determine the average for baseline.

Summary of the listings and analyses percent predicted forced expiratory volume at 1 second (ppFEV1) will be based on the ITT population. The change from baseline to Week 52 in ppFEV1 will be summarized by analysis visit and analyzed using the RCM model as described above.

6.6.3.4. Change in percent predicted peak expiratory flow (ppPEF) from baseline to Week 52, assessed by spirometry

The change from baseline to Week 52 in ppPEF between two treatment arms will be summarized by analysis visit and analyzed using the RCM model as described above.

6.6.3.5. Analysis of Progression of Dilated Cardiomyopathy (DCM) by genetic analysis

A whole blood sample will be collected in a PAXgene blood DNA tube for DNA analysis on Day 1 before dosing for all subjects who agree to DNA analysis (to include but not limited to LTBP4 & SPP1). This testing is optional, requires specific consent by participating subjects and subjects may refuse DNA testing. In the event that a subject later agrees to genetic testing, a blood sample may be collected as an unscheduled lab draw. Only tests for genetic loci associated with dilated cardiomyopathy (DCM) will be performed and all samples will be destroyed after testing is completed.

Summary of the progression of dilated cardiomyopathy (DCM) by genetic analysis will be based on the ITT population. Proportion of patients with progression of dilated cardiomyopathy (DCM) will be summarized.

6.6.3.6. Fibrosis/MRI assessments:

6.6.3.6.1. Changes in Fibrosis Score of the Biceps Brachii from Baseline to Week 52, Assessed by MRI

Summary of the listings and analyses of upper arm (bicep) muscle fibrosis assessed by MRI, will be summarized based on ITT population.

The change from baseline to Week 52 in fibrosis score of the biceps brachii muscle assessed by MRI between two treatment arms will be compared by ANCOVA. The ANCOVA model will contain terms for treatment and baseline measurements.

6.6.3.6.2. Changes in Cardiac Fibrosis Score from Baseline to Week 52, Assessed by Late Gadolinium Enhancement (LGE).

Summary of the listings and analyses of changes in cardiac fibrosis (LGE) will be based on the ITT population.

The change from baseline to Week 52 in cardiac fibrosis score assessed by MRI and Mass of Late Gadolinium Enhancement (Scar Mass) between two treatment arms, will be compared by ANCOVA. The ANCOVA model will contain terms for treatment and baseline measurements.

In addition, dichotomous responder analysis (proportion of patients with change from baseline in cardiac fibrosis score ≥ 0 at Week 52) will be summarized and analyzed. The logistic regression model will be used to compare the proportion of subjects with change from baseline in cardiac fibrosis >=0. The model will include treatment and baseline cardiac fibrosis score.

6.6.3.7. Change in Myocardial Circumferential Strain [Global Circumferential Strain (GCS)%] percentage from Baseline to Week 52, Assessed by cardiac MRI.

Summary of the listings and analyses of annual changes from baseline for the Global Myocardial Circumferential Strain (GCS) will be summarized for subjects in the ITT population,

The change from baseline to Week 52 in GCS score between two treatment arms will be compared by ANCOVA. The ANCOVA model will contain terms for treatment and baseline measurements.

In addition, dichotomous responder analysis (proportion of patients with change from baseline in GSC score ≥ 0 at Week 52) will be summarized and analyzed. The logistic regression model will

be used to compare the proportion of subjects with change from baseline in GCS score >=0. The model will include treatment and baseline cardiac fibrosis score.

6.6.4. Subgroup Analyses of Specified Endpoints

The subgroups may include but are not limited to,

- Age group (<=16 years, >16 years),
- Race (White, Others),
- Corticosteroids use (Deflazacort, Prednisone/Prednisolone).
- Baseline PUL level (<= Median, > Median)
- Baseline PUL entry Item A score (=1, >1)
- Baseline PUL entry Item A score (2-5, 1 or 6)
- Regional subgroups:
  - US, Non-US
  - China, Non-China

The primary endpoint PUL will be repeated for relevant and appropriate subgroups. If linear assumption was met, the RCM linear regression model as described in Section 6.6.1.1 will be performed for subgroup analysis; otherwise, MMRM model as described in Section 6.6.1.2.1 will be performed. Secondary endpoints ppFVC, Mid-level of PUL, grip strength, and LVEF will also be analyzed by subgroups.

If subjects who took the prohibited medications constitute to more than 10% of the population, then the subgroup analysis – with or without prohibited medications will be performed.

The LSMean of treatment difference and corresponding 95% CI will be presented in a forest plot.

6.7. PK Analyses

A summary of drug concentrations by visit as well as a listing will be provided. A population PK analysis, as well as an exposure-response analysis, will be defined in a separate PK analysis plan.

6.8. Safety Analyses

6.8.1. Adverse Events

Adverse events will be coded using MedDRA version 26.0 or higher.

A new or worsening AE occurring on or after the first dose of study medication and within 60 days after the last dose of study drug is defined as a treatment emergent adverse event (TEAE). Partially or incomplete missing AE start/stop date/time will be imputed (Appendix 1).

If more than one event occurs with the same system organ class (SOC) and preferred term (PT) for the same patient, the patient will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by common terminology criteria for adverse events (CTCAE) severity grade and by relationship to the study medication. Relationship to study drug will be imputed as “Related” for any TEAE with missing value for relationship.

The following summary AE tables including number (%) of subjects will be produced:

- Summary of all AEs
- TEAEs by PT
- TEAEs by SOC and PT
- TEAEs by SOC, PT, and maximum severity
- TEAEs with severity grade ≥3 by SOC and PT
- TEAEs related to study medication determined by investigator by SOC and PT
- Common TEAEs (≥ 5% of subjects in either treatment arm) by SOC and PT
- Common non-serious TEAEs (≥ 5% of subjects in either treatment arm) by SOC and PT (Note: this is required for ClinicalTrials.gov)
- TEAEs leading to discontinuation of study medication by SOC and PT
- TEAEs leading to interruption of study medication by SOC and PT
- Treatment-emergent serious AEs (TESAEs) by SOC and PT
- TESAEs by SOC, PT, and maximum severity
- TESAEs related to study medication determined by investigator by SOC and PT
- Fatal TESAEs (i.e., adverse events that has an outcome of death) by SOC and PT
- All cause deaths

Listings of serious adverse events (SAEs), adverse events leading to study drug discontinuation, hypersensitivity, and infusion reactions, and all cause deaths will be provided.

Additionally, the following safety analyses will be conducted:

- Number and percentage subjects with bone fractures
- Ulna length measurements for indirect measure of growth velocity (cm/year) for subjects under the age of 18 (Refer to Section 6.8.7)

Refer to Section 6.8.6 for special safety events such as hypersensitivity reactions and related topics.

6.8.2. Clinical Laboratory Assessments

Blood samples are drawn for the following analyses showed in Table 2.

Table 2: Laboratory Tests

CBC:	**Chemistry Panel:**
Absolute neutrophil count (ANC)	BUN
Eosinophils	Creatine Kinase
RBCs (Erythrocyte count)	Creatinine
Hematocrit %	Chloride
Hemoglobin	Magnesium

WBCs (Leukocyte count)	ALP
Lymphocytes	ALT
Monocytes	AST
Neutrophils	Bilirubin, total
Platelets	Albumin
CRP	Phosphorous
Basophils	Potassium
	Sodium
	GGT
	Calcium
	GFR
	Cystatin-C

Laboratory test results and change from baseline are summarized by analysis visit and by treatment arm.

CTCAE grade 3 or higher lab test results will be considered potentially clinically significant. These results are summarized and presented in a data listing.

Shift tables to summarize changes from baseline to each visit in CTCAE v5.0 categories are tabulated. Shift from baseline to most severe CTCAE category during the study is also summarized.

Due to central lab kit shortages caused by Covid-19, lab data may be collected using local labs in lieu of the central lab. The local lab data will be integrated with central lab data, when appropriate and feasible. Sensitivity analyses may be performed for key analyses of lab parameters with and without local lab values (i.e., set to missing).

An eDISH (evaluation of Drug Induce Severe Hepatotoxicity) analytical graph, which is a scatter plot of maximum observed total bilirubin versus maximum observed ALT or AST, will be generated to identify cases in Hy's law range.

6.8.3. Vital Signs

Pulse (beats/min), diastolic and systolic blood pressure (mmHg), respiration (breaths/min), and temperature (°C) will be descriptively summarized by treatment at selected visits.

Vital sign values are potentially clinically significant (PCS) if they meet both the observed value criteria and the change from baseline criteria listed in Table 3 below. The number and percentage of subjects with post-baseline PCS values will be tabulated by treatment group. The percentages are to be calculated relative to the number of subjects with baseline and at least one post-baseline assessment. The numerator is the total number of subjects with at least one post-baseline PCS vital sign value. A supportive listing of subjects with post-baseline PCS values will be provided including the treatment, subject ID, study center ID, baseline, and post-baseline values. A listing of all AEs for subjects with PCS vital signs will also be provided.

Table 3: Criteria for Potentially Clinically Significant Vital Signs

Vital Sign Parameter	Flag	Criteria*	
		Observed Value	Change from Baseline
Systolic Blood Pressure (mmHg)	High	≥170	Increase of ≥ 20
	Low	≤90	Decrease of ≥ 20
Diastolic Blood Pressure (mmHg)	High	≥110	Increase of ≥15
	Low	≤45	Decrease of ≥ 15
Pulse Rate (bpm)	High	≥120	Increase of ≥20
	Low	≤50	Decrease of ≥ 20
Weight (kg)	High	-	Increase of ≥20%
	Low	-	Decrease of ≥ 20%

* Except for body weight, a post-baseline value is considered as a PCS value if it meets both criteria for observed change from baseline.

6.8.4. Physical Examination

Clinically significant changes in PE results from baseline will be summarized by treatment over visit. A shift table will be provided if data appropriate. A listing of clinically significant changes in PE results from baseline will be provided.

6.8.5. Electrocardiogram

ECG results will be provided in a listing only.

6.8.6. Special Safety Events

Treatment-emergent special safety events including:

1. Hypersensitivity (any time)
2. Infusion reactions (24 hours: on day of infusion or 1 day post any study drug infusion)
3. Anaphylactic reactions (24 hours: on day of infusion or 1 day post any study drug infusion)

Items 1 and 2 include both hypersensitivity and angioedema events. Both items will be listed and summarized similarly to TEAEs:

- Events by event type, SOC, PT, and maximum severity,
- Events by event type, SOC, PT, and subgroups (same as TEAEs).

Items 3 will be listed and summarized by treatment and PT only.

The preferred term list for these special safety events will be finalized prior to database lock.

6.8.7. Growth Velocity (cm/year)

Ulna length measurements is used for indirect measure of annualized growth velocity (cm/year) for subjects under the age of 18. For subjects that enter the study at age 18 or turn 18 during the study, the last recorded ulna length/height will be used for the remainder of the study.

Annualized growth velocity (cm/year) is defined to be the ratio of change in height (cm) between the baseline and post-baseline to the duration of the observed time (years), where height is estimated by the formula described in Section 6.3.1.

6.8.8. Bone Fractures

Bone fractures are summarized, but not limited to, by types of bone fractures, namely, long bone fractures and vertebral compression fractures.

- Long bones fractures
  - Lower extremities: femur (thigh); tibia; fibula (leg), metatarsus, phalanges (foot). Joint fractures: hip, knee, ankle;
  - Upper extremities: humerus (arm); radius; ulna (forearm); metacarpus, phalanges (hand, palm); Joint fractures: shoulder, elbow, wrist.
- Vertebral compression fractures

  Spine compression fracture, vertebral compression fracture, spine fracture, or vertebral fracture indicate the same type of vertebral compression fractures. The level of compression fracture is indicated by vertebrae: C1–C7 (cervical spine); T1–T12 (thoracic spine); L1–L5 (lumbar spine); S1-S5 (sacrum, fused), coccyx (3-5, fused).
- Other

  Hip bone fracture; foot fracture (other than long bones); hand fracture (other than long bones); other.

6.8.9. Subgroup Analyses of Safety Endpoints

The subgroup analysis will be conducted to safety endpoints such as TEAE and TESAE. The subgroups are specified in Section 6.6.4, not including subgroup defined by Baseline PUL level.

6.9. Biomarker Endpoint Analysis

A sample for Tryptase at time of immunogenic reaction will be collected. A listing will be provided for the data.

6.10. Immunogenicity Analysis

A listing will include all available Human Anti-Human Antibody (HAHA, ADA) samples will be provided. ADA analysis will be conducted in the OLE phase of the study.

7. CHANGES FROM PROTOCOL

- **Changes for the primary analysis**:
  The mITT population is defined in Section 5.2.2 and is used for the analysis of primary and secondary efficacy endpoints. In addition, sensitivity analyses will be conducted for the primary and secondary efficacy endpoints in the ITT population.

- **Rationale for the changes:**
  - Recently published observational data suggest that patients with PUL 2.0 entry score 1 have minimal PUL 2.0 changes over a one-year treatment period [19]. Consequently, there is no meaningful opportunity to demonstrate the benefit of a therapeutic drug in the PUL 2.0 entry score 1 subpopulation. This information was published in April, 2023, after the study was already designed.
  - This is the largest observational study of PUL 2.0 scores to date in 311 patients followed longitudinally over 24 months with 808 paired assessments [19]. In this study, the decline of PUL 2.0 scores for patients with a PUL 2.0 entry score of 1 was very low at 12 and 24 months compared to patients with higher PUL 2.0 entry scores. For example:
    - At 12 months, the mean PUL 2.0 total score change on all the paired assessments for patients with PUL 2.0 entry scores of 1 was -0.84 compared to -2.05 for PUL 2.0 entry score of 3. Changes were even greater for patients with PUL entry scores of 4, 5, or 6.
    - At 24 months, the mean PUL 2.0 total score change on all the paired assessment for patients with PUL 2.0 entry scores of 1 was -1.78 compared to -3.82 for patients with PUL 2.0 entry score of 3. Changes were even greater for patients with PUL 2.0 entry scores of 4, 5, or 6.
  - Based on these results, the authors concluded that differences in PUL 2.0 entry scores should be taken into account when designing and analyzing clinical studies [19] and this supports the proposed change to the analysis population to exclude patients with a PUL baseline score of 1.
  - Based on these results, the authors concluded that differences in PUL 2.0 entry scores should be taken into account when designing and analyzing clinical studies [19].
  - In an early study by Pane from a cohort of 187 patients with DMD, which included 90 non-ambulatory patients, patients with the lowest PUL entry scores of 0 or 1 had minimal further progression during a 12-month follow-up period.
  - Expert guidance from key opinion leaders for PUL 2.0 [redacted] strongly support the approach of removal of patients with PUL 2.0 entry scores of 1 because these patients do not demonstrate significant decline over a 1-year period and therefore impact meaningful opportunity to demonstrate efficacy.

- Based on changes in the external environment for DMD studies, the Sponsor conduced an analysis of current trials from clinicaltrials.gov and concluded that other clinical trials in the non-ambulatory DMD patient population using the PUL 2.0 as a primary endpoint measure have excluded patients with a PUL 2.0 entry score of 1 at baseline. For example, Capricor Inc./NCT05126758 A Phase 3, Randomized, Double-Blind, Placebo-Controlled Trial Evaluating the Efficacy and Safety of Human Allogeneic Cardiosphere-Derived Cells for the Treatment of Duchenne Muscular Dystrophy A Study of CAP-1002 in Ambulatory and Non-Ambulatory Patients with Duchenne Muscular Dystrophy (HOPE-3).
- Non-ambulatory DMD patients have limited treatment options and have been excluded from clinical trials in emerging treatment classes. Modification of the analysis will result in benefit to non-ambulatory patients as well as increased treatment options.
- The Sponsor has determined that removal of patients with a baseline PUL 2.0 entry score of 1.0 will result in approximately 15 patients removed from the ITT population. The sample size and power calculations are based on the assumptions described in the protocol, and with 15 fewer subjects out of 97 dosed subjects, the trial will have 87% power with 82 subjects in the mITT population.

- Sample size calculation was updated per the latest literature publications and study enrollment. The testing sequence of secondary endpoints was changed from the protocol.
- Changed the analysis method for subscores of regional dimensions (High-level (Shoulder), Distal-level (Wrist and Hand)) of PUL from MMRM to RCM, the same one used in the primary analysis.
- The analysis methods for change from baseline at week 52 in LVEF, Fibrosis Score of the Biceps Brachii, Cardiac Fibrosis Score, Myocardial Circumferential Strain [Global Circumferential Strain (GCS)%] percentage are changed from MMRM to ANCOVA respectively, because there are only two assessments performed during the study for these exploratory endpoints – one at baseline and the other at week 52 and it is not suitable for mixed models of repeated measures.
- MMRM (Section 6.6.1.2.1), Cumulative Loss of Function Analysis (Section 6.6.1.2.8) are added as sensitivity analyses for the primary efficacy endpoint.
- Added the composite of endpoints - mean observed z scores sequentially across 4 endpoints (PUL, ppFVC, Grip Strength, LVEF%) at week 52, namely, the mean z-scores for PUL + ppFVC, PUL + ppFVC + Grip Strength, PUL + ppFVC + Grip Strength + LVEF% , to the secondary endpoint. Analysis with the Composite of Endpoints (Section 6.6.2.4) is added as additional analysis for PUL.
- Change in the total score of Mid-level of PUL from baseline to Week 52 is added as a secondary endpoint.

8. REFERENCES

1. Annexstad EJ, Lund-Petersen I, Rasmussen M. (2014) Duchenne muscular dystrophy. Tidsskr Nor Laegeforen. 134(14):1361-1364.
2. Pane M, Mazzone ES, Fanelli L, et al. Reliability of the Performance of Upper Limb assessment in Duchenne muscular dystrophy. NeuromuscDisord 2014;24:201-206.
3. Mayer OH, Finkel RS, Rummey C et al. Characterization of Pulmonary Function in Duchenne Muscular Dystrophy. PediatrPulmonol. 2015; 50:487–494
4. J. Finder et al. Pulmonary Endpoints in Duchenne Muscular Dystrophy A Workshop Summary Am J Respir Crit Care Med Vol 196, Iss 4, pp 512–519, Aug 15, 2017
5. S. C. Hagenbuch et al. Detection of Progressive Cardiac Dysfunction by Serial Evaluation of Circumferential Strain in Patients With Duchenne Muscular Dystrophy Am J Cardiology, 2010;105:1451–1455
6. V. Rocotti, et al. Respiratory and upper limb function as outcome measures in ambulant and non-ambulant subjects with Duchenne muscular dystrophy: A prospective multicentre study, Neuromuscular Disorders (2019) 1-8.
7. Seferian, et al. Upper Limb Strength and Function Changes during a One-Year Follow-Up in Non-Ambulant Patients with Duchenne Muscular Dystrophy: An Observational Multicenter Trial.PLoS ONE 10(2):1-18.
8. McDonald et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study.Lancet2018; 391: 451–61.
9. Supplement to: McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. *Lancet* 2017
10. Seferian AM, Moraux A, Annoussamy M, et al. (2015) Upper limb strength and function changes during a one-year follow-up in non-ambulant patients with Duchenne Muscular Dystrophy: an observational multicenter trial. PLoS One. 10(2):e0113999.
11. Sun G, Haginoya K, Wu Y, et al. (2008) Connective tissue growth factor is overexpressed in muscles of human muscular dystrophy. J Neurol Sci 267: (1-2):48-56
12. Szigyarto CA, Spitali P. (2018) Biomarkers of Duchenne muscular dystrophy: current findings. Degener Neurol Neuromuscul Dis. 8:1-13
13. Yiu EM, Kornberg AJ. (2015) Duchenne muscular dystrophy. J Pediatric Child Health. 51(8):759-764.
14. Zhou L, Lu H. (2010) Targeting fibrosis in Duchenne muscular dystrophy. J Neuropathol Exp Neurol. 69(8):771-776.
15. McDonald Craig, et al. (2021) Evaluating longitudinal therapy effects via the North Star Ambulatory Assessment. Muscle & Nerve. DOI: 10.1002/mus.27396

16. Daniel Li, et al. (2020) Assessment of Treatment Effect With Multiple Outcomes in 2 Clinical Trials of Patients With Duchenne Muscular Dystrophy. JAMA Network Open. 2020;3(2):e1921306. DOI:10.1001/jamanetworkopen.2019.21306

17. D. Y. Lin, et al. (1999) Semiparametric regression for the mean and rate functions of recurrent events, J. R. Statist. Soc. B (2000) 62, Part 4, pp. 711-730

18. China National Medical Products Administration, Technical Guidance for Accepting Overseas Clinical Trial Data ([2018] No.52), 2018

19. Pane, Marika et al. 'Longitudinal Analysis of PUL 2.0 Domains in Ambulant and Non-Ambulant Duchenne Muscular Dystrophy Patients: How Do They Change in Relation to Functional Ability?' Journal of Neuromuscular Diseases, vol. Pre-press, no. Pre-press, pp. 1-8, 2023.

20. Astellas Pharma Inc. A Randomized, Double Blind, Placebo-Controlled Phase 1b Study With Open-Label Extension to Assess the Safety, Tolerability and Preliminary Efficacy of ASP0367 (MA-0211) in Pediatric Male Patients With Duchenne Muscular Dystrophy (DMD). ClinicalTrials.gov identifier NCT04184882.

21. Capricor. A Study of CAP-1002 in Ambulatory and Non-Ambulatory Patients With Duchenne Muscular Dystrophy. ClinicalTrials.gov identifier NCT05126758. Updated May 3, 2023. Accessed May 8, 2023. https://clinicaltrials.gov/ct2/show/NCT05126758.

22. Capricor. A Phase 2, Randomized, Double-Blind, Placebo-Controlled Trial Evaluating the Safety and Efficacy of Intravenous Delivery of Allogeneic Cardiosphere-Derived Cells in Subjects with Duchenne Muscular Dystrophy. ClinicalTrials.gov identifier NCT03406780.

23. Dyne Therapeutics. A Randomized, Double-Blind, Placebo-Controlled, Multiple Ascending Dose Study Assessing Safety, Tolerability, Pharmacodynamics, Efficacy, and Pharmacokinetics of DYNE-251 Administered to Participants With Duchenne Muscular Dystrophy Amenable to Exon 51 Skipping. ClinicalTrials.gov identifier NCT05524883.

24. McDonald et al (2022). Repeated intravenous cardiosphere-derived cell therapy in late-stage Duchenne muscular dystrophy (HOPE-2): a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. The Lancet, Volume 399, Issue 10329,2022,Pages 1049-1058,ISSN 0140-6736,https://doi.org/10.1016/S0140-6736(22)00012-5.

25. McDonald CM, Gordish-Dressman H, Henricson EK, Duong T, et al. Longitudinal pulmonary function testing outcome measures in Duchenne muscular dystrophy: Long-term natural history with and without glucocorticoids. Neuromuscular Disorders, Volume 28, Issue 11,2018. Pages 897-909. ISSN 0960-8966. https://doi.org/10.1016/j.nmd.2018.07.004.

26. Pane M, Coratti G, Brogna C, Mazzone ES, Mayhew A, et al. (2018) Upper limb function in Duchenne muscular dystrophy: 24 month longitudinal data. PLOS ONE 13(6): e0199223. https://doi.org/10.1371/journal.pone.0199223

27. Pfizer. A Phase 1b Multicenter, Open-Label, Single Ascending Dose Study to Evaluate the Safety and Tolerability of Pf-06939926 in Ambulatory and Non-Ambulatory Subjects with Duchenne Muscular Dystrophy. ClinicalTrials.gov identifier NCT03362502.

9. APPENDICES

APPENDIX 1. HANDLING MISSING/INCOMPLETE DATES

A.1 Missing/Incomplete AE Onset Date

The following imputation rules apply to the case where the start date is incomplete (i.e., partially missing) for adverse events.

A.1.1 Missing start time

AEs with missing start times and which occur on a study-drug-dosing day will be considered as occurred after the study drug administration on that day, that is, it will be considered as TEAE. No imputation on other missing times.

A.1.2 Missing day and month

If the year is same as the year of first day on double-blind study medication, then the day and month of the start date of double-blind study medication will be assigned to the missing fields.

If the year is not the same as the year of first day on double-blind study medication, then January 1 will be assigned to the missing fields.

A.1.3 Missing month only

Treat day as missing and replace both month and day according to the above procedure.

A.1.4 Missing day only

If the month and year are same as the year and month of first day on double-blind study medication, then the start date of double-blind study medication will be assigned to the missing day.

If the month and year are not the same as the year and month of first day on double-blind study medication, then the first day of the month will be assigned to the missing day.

Table 4: Analysis Date Derivation Rules for Missing/Incomplete AE Onset Date

Reported Date	Date of First Drug Intake	Analysis Date (Derived)
--/MM/YYYY	DD/MM/YYYY	
--/02/2008	14/02/2008	14/02/2008*
--/02/2008	14/02/2007	01/02/2008
--/02/2008	14/02/2009	01/02/2008
--/--/YYYY	DD/MM/YYYY	
--/--/2008	14/02/2008	14/02/2008
--/--/2008	14/02/2007	01/01/2008
--/--/2008	14/02/2009	01/01/2008
DD/--/----		No imputation
--/MM/----		
--/--/----		

*Death date has to be taken into consideration when calculating this.

If the stop date is complete and the imputed start date as above is after the stop date, the start date will be imputed by the stop date.

A.2 Missing/Incomplete AE Stop Date

If needed, the following imputation rules apply to the case where the end date is incomplete (i.e., partially missing) for adverse events. Other partial end dates will not be imputed.

A.2.1 Missing day and month, or Missing month only

December 31 will be assigned to the missing fields.

A.2.2 Missing day only

The last day of the month will be assigned to the missing day.

Table 5: Analysis Date Derivation Rules for Missing/Incomplete AE Stop Date

Reported Date	Analysis Date (Derived) *
--/MM/YYYY	Set as last day of the month 31/MM/YYYY or
	30/MM/YYYY or
	29/MM/YYYY or
	28/MM/YYYY
--/--/YYYY	31/12/YYYY
DD/--/----	No imputation
--/MM/----	
--/--/----	

*Death date has to be taken into consideration when deriving this.

A.3 Missing/Incomplete Prior or Concomitant Medication Start Date

For prior or concomitant medications, incomplete (i.e., partially missing) start date is imputed the same way as for the AE described above. When the start date and the stop date are both incomplete for a patient, impute the start date first.

The following rules will be applied to impute the missing start date. If the stop date is complete

and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

A.3.1 Missing day and month

If the year of the incomplete start date is the same as the year of the first dose date of double-blind study medication, then the day and month of the first dose date will be assigned to the missing fields.

If the year of the incomplete start date is not the same as the first dose date of double-blind study medication, then January 1 will be assigned to the missing fields.

A.3.2 Missing month only

Treat day as missing and replace both month and day according to the above procedure.

A.3.3 Missing day only

If the month and year of the incomplete start date are the same as the month and year of the first dose date of double-blind study medication, then the day of the first dose date will be assigned to the missing day.

If the month and year of the incomplete start date are the same as the first dose date of double-blind study medication, then the first day of the month will be assigned to the missing day.

A.4 Missing/Incomplete Prior or Concomitant Medication Stop Date

The following rules will be applied to impute the missing stop date, if needed. If the last dose

date of double-blind study medication is missing, impute it with the last visit date. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be replaced with the start date.

A.4.1 Missing day and month

If the year of the incomplete stop date is the same as the year of the last dose date of double- blind study medication, then the day and month of the last dose date will be assigned to the missing fields.

If the year of the incomplete stop date is not the same as the year of the last dose date of double-blind study medication, then December 31 will be assigned to the missing fields.

A.4.2 Missing month only

Treat day as missing and replace both month and day according to the above procedure.

A.4.3 Missing day only

If the month and year of the incomplete stop date are the same as the month and year of the last dose date of double-blind study medication, then the day of the last dose date will be assigned to the missing day.

If the month and year of the incomplete stop date are not the same as the month and year of the last dose date of double-blind study medication, then the last day of the month will be assigned to the missing day.

A.5 Missing Date Imputation for Last Dose Date

Imputed last dose date = earliest date of (last drug dispense date + number of days of drug dispensed, date of death, date of EOT/EOS visit, and other dates as appropriate).

APPENDIX 2. ANALYSIS VISIT WINDOWS

A.2. Analysis Windows

Analysis visits are defined by the windows that will have the widths of the corresponding assessments centered at the scheduled time. Unscheduled visits within a visit window defined below will be grouped into the closest scheduled visits based on the visit date.

Table 6: Analysis Visit Window for Vital Signs

Analysis Visit	Target Day	Start Day	End Day
Baseline	Last value before the first study drug infusion (or randomization date if no infusion received)		
Day 1 post-dose	1	1	1
Week 2	15	2	22
Week 4	29	23	36
Week 6	43	37	50
Week 8	57	51	64
Week 10	71	65	78
Week 12	85	79	92
Week 14	99	93	106
Week 16	113	107	120
Week 18	127	121	134
Week 20	141	135	148
Week 22	155	149	162
Week 24	169	163	176
Week 26	183	177	190
Week 28	197	191	204
Week 30	211	205	218
Week 32	225	219	232
Week 34	239	233	246
Week 36	253	247	260
Week 38	267	261	274
Week 40	281	275	288
Week 42	295	289	302
Week 44	309	303	316
Week 46	323	317	330
Week 48	337	331	344
Week 50	351	345	358
Week 52	365	359	**

** before the first study drug infusion in the OLE period or end of randomization treatment period
Note: The data on a study day with infusions will be used as a priority for summary tables whenever possible.
Note: All assessments beyond 60 days post last infusion will be excluded from summary tables or figures.

Table 7: Analysis Visit Window for Physical Examination

Analysis Visit	Target Day	Start Day	End Day
Baseline	Last value before the first study drug infusion (or randomization date if no infusion received)		
Week 12	85	2	120
Week 22	155	121	190
Week 32	225	191	260
Week 42	295	261	330
Week 52	365	331	**

** before the first study drug infusion in the OLE period or end of randomization treatment period

Note: All assessments beyond 60 days post last infusion will be excluded from summary tables for labs.

Table 8: Analysis Visit Window for Pulmonary Function Tests (PFT)

Analysis Visit	Target Day	Start Day	End Day
Baseline	See Section 5.3.2		
Week 12	85	2	120
Week 22	155	121	190
Week 32	225	191	260
Week 42	295	261	330
Week 52	365	331	**

** before the first study drug infusion in the OLE period or end of randomization treatment period
Note: All assessments beyond 60 days post last infusion will be excluded from summary tables for labs.

Table 9: Analysis Visit Window for Muscle Function Tests (MFT)

Analysis Visit	Target Day	Start Day	End Day
Baseline	Last value before the first study drug infusion (or randomization date if no infusion received)		
Week 16	113	2	155
Week 28	197	156	239
Week 40	281	240	323
Week 52	365	324	**

** before the first study drug infusion in the OLE period or end of randomization treatment period
Note: All assessments beyond 60 days post last infusion will be excluded from summary tables for labs.

Table 10: Analysis Visit Window for Laboratory Tests

Analysis Visit	Target Day	Start Day	End Day
Baseline	Last value before the first study drug infusion (or randomization date if no infusion received)		
Week 8	57	2	85
Week 16	113	86	141
Week 24	169	142	197
Week 32	225	198	253
Week 40	281	254	309
Week 48	337	310	365
Week 56	393	366	**

** before the first study drug infusion in the OLE period or end of randomization treatment period
Note: All assessments beyond 60 days post last infusion will be excluded from summary tables for labs.

APPENDIX 3. PERFORMANCE OF THE UPPER LIMB (PUL) MODULE

Performance of the Upper Limb Module for DMD 2.0 (PUL for DMD)				
Dominant arm (used for all tests): ☐ Right ☐ Left Elbow extension ROM full = 0°: Right: Left: e.g. 10° contracture = -10° **Supination ROM: Right:** ☐ Full ☐ ¾ ☐½ ☐¼ **Left:** ☐ Full ☐ ¾ ☐½ ☐¼				
Entry item A. – start with A to identify starting point for subsequent tests. Circle score for each item. DO NOT INCLUDE IN TOTAL SCORE				
Item	Score	Description		
A	0	No useful function of hands		
	1	Can use hands to hold pen or pick up a coin or drive a powered chair		
	2	Can raise 1 or 2 hands to mouth but cannot raise a cup with a 200g weight in it to mouth		
	3	Can raise plastic cup with 200g weight in it to mouth using 1 or 2 hands		
	4	Can raise both arms (to shoulder height with or without compensation), i.e. elbow bent or in extension		
	5	Can raise both arms simultaneously above head only by flexing the elbow (shortening circumference of the movement /using accessory muscles)		
	6	Can abduct both arms simultaneously elbows in extension in a full circle until they touch above the head.		
For item A: A score of 3, 4, 5, 6 on item A, start with item 1. A score of, 1, 2 start with item 7				
High level shoulder Dimension: Administer only if subject scored 3, 4, 5, 6 on item A				
Item	Description	0	1	2
1 Score from Entry item above	Shoulder abduction both arms above head "Raise your arms above your head out to the side – try and keep straight elbows"	Unable	Can raise both arms simultaneously **above head** only by flexing the elbow - **with** compensation	Can abduct both arms simultaneously elbows in extension in a full circle until they touch above the head

2	Raise both arms to shoulder height (elbows at shoulder height) "Raise your arms to shoulder level"	Unable	Can raise both arms to shoulder height either one at a time or with elbows flexed (**with compensation**)	Can raise both elbows to shoulder height **without** compensation e.g. simultaneously with elbows straight
3	Shoulder flexion to shoulder height (no weights) "Reach out and touch my hand" – elbow to eye level	Unable	Able **with** compensation	Able **without** compensation
4	Shoulder flexion to shoulder height with 500g weight "Reach out and touch my hand" – elbow to eye level	Unable	Able to lift 500g weight **with compensation**	Able to lift 500g weight **without compensation**
5	Shoulder flexion above shoulder height with 500 g weight Hand on lap – "give me the weight"	Unable	Able to lift 500g weight **with compensation**	Able to lift 500g weight **without compensation**
6	Shoulder flexion above shoulder with 1 kg weight Hand on lap – "give me the weight"	Unable	Able to lift 1 kg weight **with compensation**	Able to lift 1 kg weight **without compensation**

Mid-level elbow Dimension				
Do these tests on all individuals				
Item	Description	0	1	2
7	Hand(s) to mouth "Bring the cup to your mouth with one hand"	Unable	Able to bring 200g in cup with any compensation to mouth (can use more than one hand and / or bring head to hands)	Able to bring 200g in cup to mouth with one hand no elbow support (without compensation)
8	Hands to table from lap "Bring both hands from lap to table"	Unable	Able to bring two hands completely (to wrist crease) to table but **NOT**	Two hands completely on table simultaneously

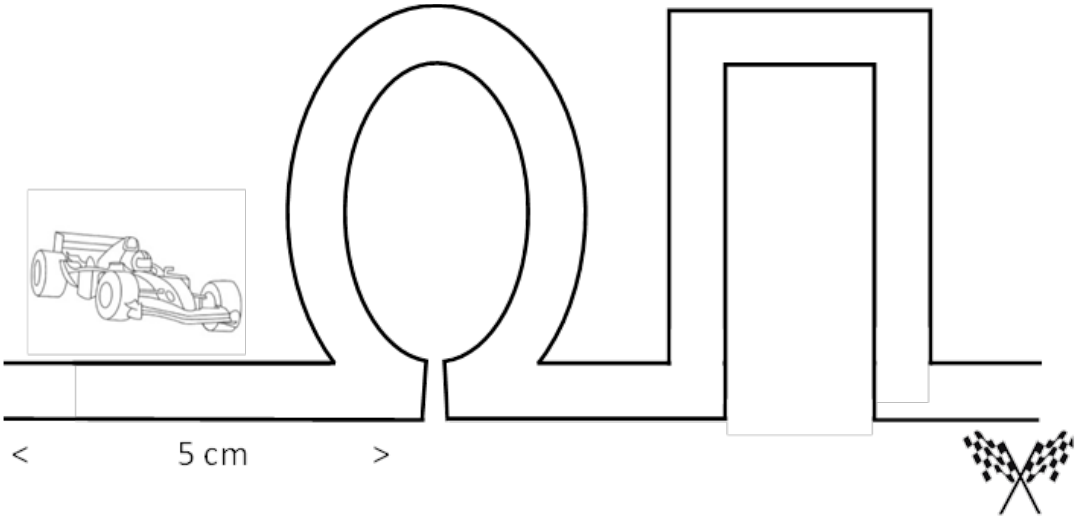
			simultaneously or in one action	
9	Move weight on table 100g "Move the weight from outside circle to centre circle"	Unable	Can move **100g** weight from outer to centre circle using compensation (slide forearm or elbow make contact with table)	Can lift **100g** weight from outer to centre circle without compensation
10	Move weight on table 500g "Move the weight from outside circle to centre circle"	Unable	Can move **500g** weight from outer to centre circle using compensation (slide forearm or elbow make contact with table)	Can lift **500g** weight from outer to centre circle without compensation
11	Move weight on table 1kg "Move the weight from outside circle to centre circle"	Unable	Can move **1kg** weight from outer to centre circle using compensation (slide forearm or elbow make contact with table)	Can lift **1kg** weight from outer to centre circle without compensation
12	Lift heavy can diagonally "Lift can from this circle nearest your hand to this circle furthest away and across your body"	Unable	Can move heavy can from nearest circle across body with compensation (slide forearm or elbow make contact with table)	Can lift heavy can from nearest circle across body without compensation
13	Stack of three cans "Stack these two cans, one at a time on the middle can using one hand"	Unable to stackthirdcan even with compensation	Able to stack third can with compensation	Able to stack third can without compensation
14	Stack of five cans "Stack these two additional cans, one at a time on top of this can using one hand"	Unable to stackfifthcan even with compensation	Able to stack fifth can with compensation	Able to stack fifth can without compensation

15	Remove lid from container **“**Use your hands to open this container”	Unable	Opens completely	

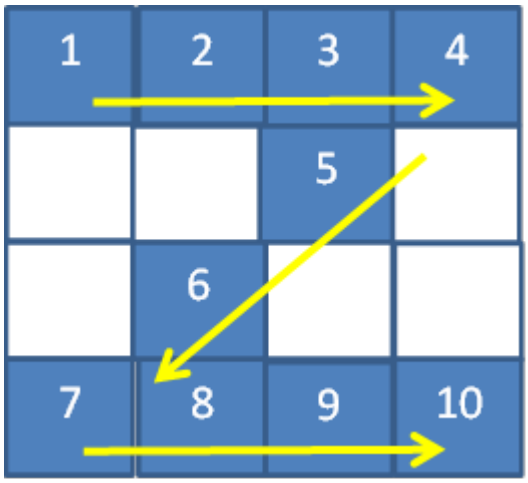
Distal wrist and hand Dimension: Do these tests on all individuals				
Item	Description	0	1	2
16	Tearing paper “Tear the sheet of paper beginning from here”	Unable	Tears the sheet of paper folded in half from the folded edge	Tears the sheet of paper folded in 4, beginning from the folded edge
17	Tracing path “Use your pencil to complete the path in one smooth movement”	Unable	Completes the path with compensation - needs to raise pencil from paper or pivot arm	Able to complete the path **without stops** or raising hand from paper
18	Push on light “Push on the light with the fingers of one hand”	Unable	Able to turn the light on momentarily with fingers of one hand	Able to turn the light on permanently with fingers of one hand
19	Supination “Pick up the light and turn your hand over”	Unable	Picks up the light but either turns hands over incompletely or uses compensation to turn it over	Picks up the light, and turns the hand over completely with no compensatory movements
20	Picking up coins **“**Using one hand, Pick up 6 coins, one at a time”	Cannot pick up one coin	Can pick up **one** coin/ token	Can pick up six coins in one hand
21	Placing finger on number diagram (precision not essential) **“**Using one finger to touch each number on the diagram”	Cannot raise the finger or slide it on the diagram	Able to place finger (slide or lift) between at least two squares	Able to place finger successively on the numbers of the diagram (with or without compensation)
22	Pick up 10g weight finger pinch “Pick up this small weight like this (by body of weight)”	Unable	Able to grip and lift weight off surface	

Total Score PUL	

Additional Material
Item 17: Tracing a path



Item 21: Placing finger on number diagram
Instruction: Starting on the yellow number 1 point to the numbers 1 to 10 in turn following the arrow



Were all tests valid?
(i.e., representative of child's true function) Yes ☐ No ☐

If no, which tests were invalid?		And why?
Hand Held Myometry	□	☐ Testing Environment (i.e. equip. malfunction, fire alarm, disruption) ☐ General health (i.e. stomachache, cold, flu) ☐ Travel delays (i.e. lack of sleep; testing not performed at regular time) ☐ Inappropriate clothing ☐ Behavior ☐ Musculoskeletal issue (i.e. injury, muscle cramping, tendinitis) ☐ Other (please explain in comment section of worksheet)
PUL	□	☐ Testing Environment (i.e. equip. malfunction, fire alarm, disruption) ☐ General health (i.e. stomachache, cold, flu) ☐ Travel delays (i.e. lack of sleep; testing not performed at regular time) ☐ Inappropriate clothing o ☐ Behavior ☐ Musculoskeletal issue (i.e. injury, muscle cramping, tendinitis) ☐ Other (please explain in comment section of worksheet)

APPENDIX 4. LABORATORY TEST CTCAE CRITERIA

The following table is extracted from NCI Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0

Chemistry

		Grade 1	Grade 2	Grade 3	Grade 4
Bicarbonate	Decreased	<LLN and no intervention initiated	None	None	None
Creatinine		ULN – 1.5 x ULN	>1.5 – 3.0 x baseline[2] >1.5 – 3.0 x ULN	>3.0 x baseline[2] >3.0 – 6.0 x ULN	>6.0 x ULN
Albumin	Decreased	3 g/dL – LLN	2 - <3 g/dL	<2 g/dL	
Alkaline phosphatase (ALP)		ULN – 2.5 x ULN[4]; 2.0 - 2.5 x baseline[5]	>2.5 – 5.0 x ULN[4]; >2.5 - 5.0 x baseline[5]	>5.0 – 20.0 x ULN[4]; >5.0 - 20.0 x baseline[5]	> 20.0 x ULN[4]; >20.0 x baseline[5]
ALT		ULN – 3.0 x ULN[4]; 1.5 - 3.0 x baseline[5]	>3.0 – 5.0 x ULN[4]; >3.0 - 5.0 x baseline[5]	>5.0 – 20.0 x ULN[4]; >5.0 - 20.0 x baseline[5]	>20.0 x ULN[4]; >20.0 x baseline[5]
AST		ULN – 3.0 x ULN[4]; 1.5 - 3.0 x baseline[5]	>3.0 – 5.0 x ULN[4]; >3.0 - 5.0 x baseline[5]	>5.0 – 20.0 x ULN[4]; >5.0 - 20.0 x baseline[5]	>20.0 x ULN[4]; >20.0 x baseline[5]
Total bilirubin		ULN – 1.5 x ULN[4]; > 1.0 - 1.5 x baseline[5]	>1.5 – 3.0 x ULN[4]; >1.5 - 3.0 x baseline[5]	>3.0 – 10.0 x ULN[4]; >3.0 - 10.0 x baseline[5]	>10.0 x ULN[4]; >10.0 x baseline[5]
Calcium (Corrected)	Decreased	8.0 mg/dL – LLN	7.0 - <8.0 mg/dL	6.0 - <7.0 mg/dL	<6.0 mg/dL
		ULN – 11.5 mg/dL	>11.5 – 12.5 mg/dL	>12.5 – 13.5 mg/dL	>13.5 mg/dL
Glucose (Random)	Decreased	55 mg/dL – LLN	40 - <55 mg/dL	30 - <40 mg/dL	<30 mg/dL
Potassium	Decreased	3.0 mmol/L – LLN	3.0 mmol/L – LLN[1]	2.5 - <3.0 mmol/L	<2.5 mmol/L
		ULN – 5.5 mmol/L	>5.5 – 6.0 mmol/L[3]	>6.0 – 7.0 mmol/L	>7.0 mmol/L
Sodium	Decreased	130 mmol/L – LLN	125-129 mmol/L and asymptomatic	125-129 mmol/L symptomatic; 120-124 mmol/L regardless of symptoms	<120 mmol/L
		ULN – 150 mmol/L	>150 – 155 mmol/L[3]	>155 – 160 mmol/L	>160 mmol/L
Magnesium	Decreased	1.2 mg/dL – LLN	0.9 - <1.2 mg/dL	0.7 - <0.9 mg/dL	<0.7 mg/dL

		ULN – 3.0 mg/dL	None	>3.0 – 8.0 mg/dL	>8.0 mg/dL

Serum Hematology

		Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin	Decreased	10.0 g/dL – LLN	8.0 - <10.0 g/dL	<8.0 g/dL	
		>0 – 2 g/dL	>2 – 4 g/dL	>4 g/dL	
Platelet	Decreased	75,000 /mm^3 – LLN	50,000 – <75,000 /mm^3	25,000 - <50,000 /mm^3	<25,000 /mm^3
WBC	Decreased	3,000 /mm^3 – LLN	2,000 - <3,000 /mm^3	1,000 - <2,000 /mm^3	<1,000 /mm^3
		None	None	>100,000 /mm^3	
aPTT		ULN – 1.5 x ULN	>1.5 – 2.5 x ULN	>2.5 x ULN	
Lymphocytes	Decreased	800 /mm3 – LLN	500 - <800 /mm^3	200 - <500 /mm^3	<200 /mm^3
		None	>4,000 – 20,000 /mm^3	>20,000 /mm^3	
Neutrophils	Decreased	1500 /mm3 - LLN	1000 - <1500 /mm3	500 - <1000/mm3	<500/mm3
Eosinophils		>ULN and >Baseline	None	Steroids initiated	None

Decreased: below LLN; Otherwise, above ULN;
[1] Symptomatic, Intervention indicated
[2] Baseline is used if it is above ULN
[3] Intervention indicated
[4] ULN is used if Baseline was normal
[5] Baseline is used if Baseline was abnormal

APPENDIX 5. GENERAL SPECIFICATIONS FOR TABLES, LISTINGS, FIGURES

1. Software Used

All programming of tables, listings and figures (TFLs) will be performed using the statistical software package SASÒ version 9.4 or greater.

2. General

All TFLs are based on SDTM and/or ADaM datasets. By default, data listings reflect the actual values captured in SDTM and ADaM datasets, including date/time variables and missing values. Except for concatenation of some variables for compact display purpose, data are presented directly with minimum manipulation. In general, the character standard result variables is presented in data listings.

For continuous variables that are recorded as "<X" or ">X", the value of "X" will be used in the calculation of summary statistics. The value "X" is also captured in the numeric variable in the SDTM datasets as well as in the ADaM datasets for consistency, although SDTMIG recommends capturing missing values in the numeric variables.

3. Page Layout

All column headers (consisting of one or several words) will start with uppercase and thereafter only lowercase characters, except for acronyms and abbreviations. In case values from the database will be displayed in column headers, they may be displayed as in the database. Pages will be numbered as 'Page x of y', where 'y' is the total number of pages of the corresponding table or listing. The page specifications are presented in Table 3.

Table 11: Specifications for Page Layout

Paper Size	Letter
Orientation	Landscape
Alignment	Center
Font size	9
Font type	Courier New (default)
Margins:	
Top	1.0"
Bottom	0.4"
Left	1.0"
Right	0.4"

The margin sizes and font size for listings may be flexible to provide sufficient information on a single page to facilitate review and comparison.
When created using SAS, tables and listings will be created using ODS, and output files will be produced in RTF. When RTF files are produced, titles and footnotes will appear as document headers/footers.

4. Titles and Footnotes

All tables and listings will have a header showing “FibroGen, Inc.”, the protocol number, database cutoff date or ‘Final Database’, and Page x of y. A footer will show the program file path/name, output file path/name, run date and time.

All titles are written in title format, with uppercase at the beginning of each word; articles, prepositions, and conjunctions, which are of three characters length or less will start with lowercase letters (Mixed Case). Footnotes are in regular text format.

Titles

In total there are up to 10 titles available, defined as following:

first title “FibroGen, Inc.” (left aligned) and “Date of Data Cutoff : ddMMMyyyy” (right aligned)

second title protocol number “Protocol: FGCL-3019-093 (LELANTOS(-1))” (left aligned) and “Page x of y” (right aligned)

third title blank

fourth title: table/listing/figure number

fifth title: table/listing/figure title

sixth title: population names if provided in SAP, or brief definition of specific analysis set

Footnotes

Up to 10 footnote lines are available for tables, listings and figures. Footnotes 1, 9 and 10 are standard. Footnotes 2 to 8 (left aligned) might be used as needed. They are to be specified in the Shell.

first footnote is a separating horizontal line.

second – eighth are free text which can be used for explanations. Footnotes will be referenced using numbers in square brackets, starting with [1], followed by [2] etc.

ninth footnote left blank; in case needed may also be used as for explanations.

tenth footnote the program name and the date of data extraction (left aligned); the date and time in the format ddMMMyyyy hh:mm when the output was created; the version (e.g. draft or final); and the word “Confidential”.

Footnotes are denoted by [1], [2], and so on.

5. Table, Figure, Listing Metadata

The table, figure, and listing (TFL) metadata will include the TFL numbers, titles, analysis populations, program names, input dataset names. For tables and figures, PARAMCD, PARAM, and other conditions will be specified. TFL numbers, titles, and footnotes will be imported from this master spreadsheet. In addition, this spreadsheet will record the names of the original programmer and the validator/reviewer and the date of validation approval.

6. Significant Digits of Summary Statistics

All percentages will be rounded to one decimal place and lined up by the decimal place. The percentage will be suppressed when the count is zero.

Any p-values will be rounded to four decimal places and will be presented as '<.0001' if they are less than 0.0001 after rounding.

For variables of direct measurements, summary statistics are displayed with the following specifications of decimal places in Table 12 .

Table 12: Significant Digits of Summary Statistics

Description	Characteristic	Number of decimal places
Count	N	0
Mean	Mean	As in source + 1
Standard deviation	SD	As in source + 2
Standard error of the mean	SE	As in source + 2
Confidence Interval	CI	As in source + 1
Minimum	Min	As in source
Median	Median	As in source + 1
Maximum	Max	As in source
Q1 / Q3	Q1/Q3	As in source + 1
Percentage	%	All percentages will be rounded to one decimal place and lined up by the decimal place. The percentage will be suppressed when the count is zero
Coefficient of variation	CV (%)	1
p-value	p-value	p-values will be rounded to four decimal places and will be presented as '<.0001' if they are less than 0.0001 after rounding

N=number; Std=Standard deviation; CI=Confidence Interval; Min=minimum; Max=maximum; CV=Coefficient of variation

As a general guideline for derived parameters, three significant digits may be displayed for a parameter with an overall mean less than 100; otherwise, one decimal place may be used. If a derived parameter is in the same scale as some related measured parameters, such as MAP, QTc, the same display format may be used as the measured parameters.

Summary Statistics are to be displayed in the following order: Count, Mean, Standard Deviation, <Coefficient of Variation, Standard Error of the Mean, Confidence Interval>, Minimum, <Q1>, Median, <Q3>, Maximum.

For categorical variables the categories will be displayed in the TFLs in the same order they appear in the CRF.

7. Figure Specifications

In general, figures should include annotation of key summary statistics: n, mean, SE, median for continuous variables; n and percent for categorical variables; number of subjects at risk and

cumulative number of events as well as median and 95% CI for time-to-event data. Other statistics such as quartiles, ranges may be included depending on need and space.

P-values should be presented if comparisons are of interest.

For scatter plots, linear or non-linear trend lines should be included if the association of the two variables is of interest. Correlation coefficient or regression coefficients as well as corresponding p-values should be presented.

For box plots, 'BOXSTYLE=SCHEMATIC' should be used. The whiskers are drawn to the most extreme points in the group that lie within the fences. The upper fence is defined as the third quartile (represented by the upper edge of the box) plus 1.5 times the interquartile range. The lower fence is defined as the first quartile (represented by the lower edge of the box) minus 1.5 times the interquartile range. Observations outside the fences are identified with a special symbol.

APPENDIX 6. CHINA CDE REQUIREMENTS

Efficacy data, mainly including overseas key clinical trial data and clinical trial data conducted in China, should not only confirm the efficacy of the study drug as a whole, but also analyze the consistency between Chinese subgroups and the overall population.

The point estimate of treatment effect in China subgroup divided by its counterpart in the overall population will be used for assessing efficacy consistency.

Safety data, including all domestic and foreign data used for safety evaluation, should be analyzed not only for overall safety, but also for consistency between Chinese subgroup and overall population.

All TFLs in China may be provided per China regulatory requirements if regulatory submission in China is pursued. [18]