STATISTICAL ANALYSIS PLAN

Protocol Number	CAP-1002-DMD-02
Trial Phase	Phase 2
Protocol Title	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 (HOPE-2)
Protocol Version	Amendment 5.0, 27 September 2019
Product Name or Number	CAP-1002 Allogeneic Cardiosphere-Derived Cells
Indication	Duchenne Muscular Dystrophy
Sponsor	Capricor, Inc. 8840 Wilshire Blvd., 2nd Floor Beverly Hills, CA 90211 USA
Document Type	Statistical Analysis Plan
Prepared by	[redacted]
Version Number (Date)	Final v. 3.0 (20 March 2020)

STATISTICAL ANALYSIS PLAN

Final v. 3.0 (20 March 2020)

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 (HOPE-2)

STATISTICAL ANALYSIS PLAN

Final v. 3.0 (20 March 2020)

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 (HOPE-2)

STATISTICAL ANALYSIS PLAN

Final v. 3.0 (20 March 2020)

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 (HOPE-2)

REVISION HISTORY

Version	Date	Summary of Revision(s)
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Term	Explanation
10MWRT	10-Meter Walk/Run Time
AE	Adverse Event
ALT	Alanine Aminotransferase
AR(1)	Auto-Regressive
AST	Aspartate Aminotransferase
ATC	Anatomic Therapeutic Chemical
BMI	Body Mass Index
BSA	Body Surface Area
BUN	Blood Urea Nitrogen
CBC	Complete Blood Count
CDC	Cardiosphere-Derived Cells
CEC	Clinical Events Committee
CK-MB	Creatine kinase MB Isoenzyme
CK-MM	Creatine kinase MM Isoenzyme
CLIA	Clinical Laboratory Improvement Amendments
cm	Centimeters
CURE	Circumferential Uniformity Ratio Estimate
CXCL10	C-X-C Motif Chemokine 10
DMD	Duchenne Muscular Dystrophy
DMD UL- PROM	DMD Upper Limb Patient-Reported Outcome Measures
DMD-LMS	DMD Lifetime Mobility Scale
DSA	Donor-Specific Antibody
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
FEV_1	Forced Expiratory Volume in 1 Second
FVC	Force Vital Capacity
GGT	Gamma-Glutamyl Transferase
HLA	Human Leukocyte Antigen
IFR	Inspiratory Flow Reserve
IL	Interleukin
IP	Investigational Product
ITT	Intent-to-Treat
IV	Intravenous
IWRS	Interactive Web Response System
kg	Kilograms
K-R	Kenward-Roger
LV	Left Ventricle / Left Ventricular
M cells	Million cells
MedDRA	Medical Dictionary for Regulatory Activities
MEP	Maximum Expiratory Pressure
MFI	Mean Fluorescence Intensity
MIP	Mean Inspiratory Pressure
mITT	Modified Intent-to-Treat
MMRM	Mixed-Model Repeated Measures
MRI	Magnetic Resonance Image
NSAA	North Star Ambulatory Assessment
PCF	Peak Cough Flow
PEF	Peak Expiratory Flow
PFT	Pulmonary Function Test
PI	Primary Investigator
PODCI	Pediatric Outcomes Data Collection Instrument
PP	Per Protocol

Term	Explanation
PROM	Patient-Reported Outcome Measure
PT	Preferred Term
PUL	Performance of the Upper Limb
RBC	Red Blood Cell
SAP	Statistical Analysis Plan
SD	Standard Deviation
SOC	System Organ Class
SOP	Standard Operating Procedures

Term	Explanation
SpO_2	Peripheral Capillary Hemoglobin Oxygen Saturation
SVC	Slow Vital Capacity
TEAE	Treatment-Emergent Adverse Events
TLF	Tables, Listings and Figures
TNF-α	Tumor Necrosis Factor Alpha
WBC	White Blood Cell
WHODrug Global	World Health Organization Drug Global Dictionary

TABLE OF CONTENTS

1 STUDY OVERVIEW

Duchenne muscular dystrophy (DMD) is a severe, X-linked, progressive disease affecting approximately one in 3,600 to 9,200 male births caused by mutations in the dystrophin gene resulting in the absence of or non-functional dystrophin protein.

DMD occurs in early childhood, with degeneration occurring progressively in the skeletal musculature and ultimately in the heart and respiratory muscles, resulting in premature death. Progressive weakness and muscle atrophy caused by degenerating muscle fibers begins in the lower extremities and pelvis before spreading into the upper extremities. Other symptoms include loss of some reflexes, a waddling gait, frequent falls, difficulty when rising from a sitting or lying position or when climbing stairs, changes to overall posture, and impaired breathing. Many children precipitously lose the ability to run or jump. The atrophied muscles, in particular the calf muscles, and less commonly, muscles in the buttocks, shoulders, and arms, may be enlarged by an accumulation of fat and connective tissue, causing them to look larger and healthier than they actually are ("pseudohypertrophy"). Bone thinning and scoliosis are common. Ultimately, a wheelchair becomes necessary, in most cases between 12 to 15 years of age.

No pharmacologic or biological therapies proven to stop or reverse the progression of DMD have been approved in the US. Disease management consists primarily of preventive measures as well as active interventions to address the primary and secondary aspects of the disorder.

CAP-1002 is intended to be used as a therapeutic to improve the morbidities associated with DMD and thus improve or prevent decline in muscle function through its composite immune-modulatory, anti-fibrotic, and regenerative mechanisms of action. [redacted]

2 STUDY OBJECTIVES AND ENDPOINTS

2.1 Study Objectives

2.1.1 Primary

To evaluate the safety and efficacy of intravenous CAP-1002 administered every three months in subjects with DMD and impaired skeletal muscle function.

2.1.2 Secondary

To evaluate the impact of repeated intravenous (IV) administrations of CAP-1002 on exploratory efficacy assessments of skeletal and cardiac muscle function, and quality of life.

2.2 Study Endpoints

2.2.1 Primary Safety Endpoint

The ***primary safety endpoints*** include the incidence of the following from baseline through the 12-month timepoint:

- Acute respiratory decompensation within 2 hours following Investigational Product (IP) administration
- Hypersensitivity reaction
  - Hypersensitivity reaction is defined as a clinical syndrome including, but not limited to, fever, leukocytosis, or rash with onset ≤ 2 hours post infusion and lasting < 24 hours, in the absence of clinical signs of concomitant infection.
- All-cause mortality
- Serious adverse events
- Treatment-emergent adverse events related to IP or administration procedure
- Immune sensitization syndrome
  - Immune sensitization syndrome shall be defined as: (a) clinical signs and symptoms consistent with systemic inflammation (e.g., fever, leukocytosis, rash, or arthralgia) with onset ≥ 24 hours post infusion and the absence of clinical signs of concomitant infection, **AND** (b) elevation of anti-human leukocyte antigen (HLA) antibodies against the donor cells [i.e., donor-specific antibodies (DSAs)], detected ≤ 30 days following onset of syndrome, of (i) ≥ 2000 mean fluorescent intensity (MFI) if baseline MFI ≤ 1000, or (ii) ≥ 2 times baseline otherwise

2.2.2 Secondary Safety Endpoints

The ***secondary safety endpoint*** is the incidence and severity of all adverse events (AEs) through the 12-month timepoint.

2.2.3 Primary Efficacy Endpoint

The ***primary efficacy endpoint*** is change from baseline to the 12-month timepoint in functional capacity as assessed by the mid-level (elbow) dimension of the PUL 1.2.

2.2.4 Secondary Efficacy Endpoints

The ***secondary efficacy endpoints*** include the change from baseline for the following:

- Mid-level (elbow) dimension of the PUL 1.2 at Months 3, 6, and 9 *(Month 12 is the primary endpoint).*
- Regional systolic left ventricular (LV) wall thickening as assessed by cardiac magnetic resonance imaging (MRI) at Months 6 and 12.

2.2.5 Exploratory Efficacy Endpoints

The ***exploratory efficacy endpoints*** include the change from baseline to each assessment timepoint for the following assessments:

- All subjects
  - High-level (shoulder) dimension of the PUL 1.2
  - High-level (shoulder) dimension of the PUL 2.0

  - Mid-level (elbow) dimension of the PUL 2.0
  - Distal-level (wrist and hand) dimension of the PUL 1.2
  - Distal-level (wrist and hand) dimension of the PUL 2.0
  - Grip strength
  - Key and tip-to-tip pinch strength
  - Elbow flexion strength
  - Slow vital capacity (SVC), forced expiratory volume in one second (FEV_1), forced vital capacity (FVC), peak expiratory flow (PEF), maximum inspiratory pressure (MIP), maximum expiratory pressure (MEP), peak cough flow (PCF), and inspiratory flow reserve (IFR)
  - LV structure and function as assessed by cardiac MRI including ejection fraction, end-diastolic volume, end-systolic volume, stroke volume, regional wall thickness, and circumferential strain
  - Biomarkers may include those for muscle inflammation (osteopontin); systemic inflammation (interleukin 1 beta [IL-1ß], IL-2, -6 and -10; tumor necrosis factor alpha [TNF-α] and C-X-C motif chemokine 10 [CXCL10]); muscle damage (troponin I [cardiac, fast skeletal, slow skeletal], troponin T, creatine kinase MM isoenzyme [CK-MM], creatine kinase MB isoenzyme [CK-MB], creatine kinase [total], creatine phosphate, and creatinine); and muscle regeneration (myostatin and follistatin)
  - DMD Upper Limb Patient-Reported Outcome Measures (DMD UL-PROM)
  - Pediatric Outcomes Data Collection Instrument (PODCI)
  - DMD Lifetime Mobility Scale (DMD-LMS)
  - Resource utilization including the incidence of hospitalizations for orthopedic injury or surgery, pulmonary infections, and antibiotic usage

- Ambulatory subjects only
  - 10-Meter Walk/Run Time (10MWRT)
  - Incidence of loss of ambulation (defined by 10MWRT > 30 seconds)
  - North Star Ambulatory Assessment (NSAA)

3 STUDY DESIGN

This Phase 2, multi-center, randomized, double-blind, placebo-controlled trial will assess the safety and efficacy of allogeneic cardiosphere-derived cells (CDCs), or CAP-1002, administered as four IV infusions, one every three months, for the treatment of subjects with DMD. Subjects with a clinical diagnosis of DMD confirmed by genetic testing and evidence of skeletal muscle impairment will undergo a battery of tests and procedures during a 30-day screening period to determine eligibility based on protocol inclusion and exclusion criteria.

Eligible subjects will be prospectively randomized in stratified permuted blocks to CAP-1002 or placebo (1:1 ratio) approximately 7 days prior to the first IV infusion of IP on Day 1. Randomization will be stratified by site and entry item score of the PUL 1.2. Subjects will undergo baseline safety and efficacy assessments (APPENDIX A) prior to the first infusion of IP (CAP-1002 or placebo). Administration of IP (Day 1) should occur within a maximum of 14 days following randomization; if a delay of more than 14 days between randomization and IP administration is unavoidable, a conversation between the Investigator and Medical Monitor

should occur to determine the need for repeat assessments prior to infusion. Subjects that are not randomized may be re-screened at the discretion of the Investigator.

Subjects will complete study assessments at Screening; Baseline (≤ 2 days prior to Day 1); Day 1; Week 4 (±7 days); and Months 3, 6, 9, and 12 (±14 days, each). If the subject can travel to the investigative site, blood samples will be collected at Week 6 (±7 days) and Month 4.5 (±14 days).

All IV infusions will be conducted in an outpatient setting at the investigative site on Day 1 and Months 3, 6, and 9. Prior to each IP infusion, medication(s) will be administered to the subject as determined by the Investigator based on the pre-treatment guidelines and/or institutional protocols to minimize the risk of potential severe allergic reactions such as anaphylaxis. Subjects will be observed in the outpatient setting for at least two hours post infusion and then discharged the same day if medically cleared by the site Investigator.

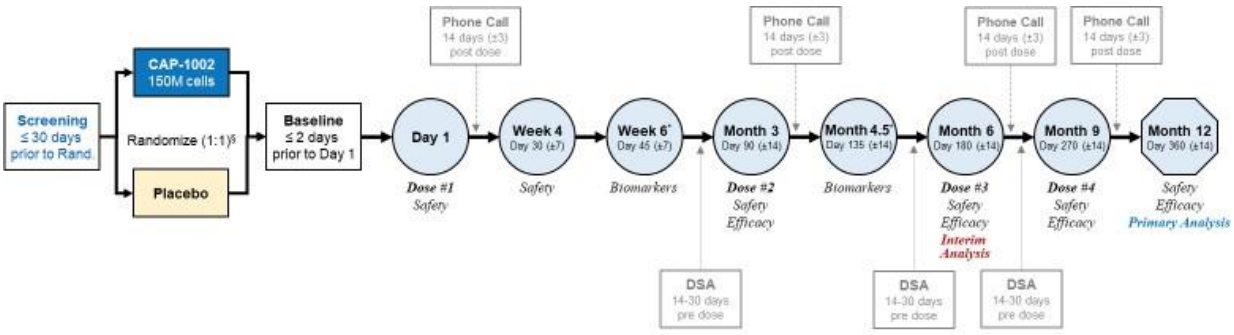
As part of these visits, subjects will complete a safety phone call 14 days (±3 days) after each IP infusion, and if clinically indicated, an unscheduled in-person visit will be performed at the investigative site with targeted assessments based on presentation of signs and symptoms.

Blood samples for DSA testing will be collected at Baseline, Week 4, 14-30 days prior to the IP infusions scheduled for the Month 3, 6 and 9 visits, and Month 12. Blood collections for clinical laboratory and DSA testing at Week 4 and pre-dose DSA testing at Months 3, 6, and 9 will occur either at the investigative site or remotely at a designated central laboratory patient service center. Blood collection for biomarker and exploratory biomarker testing at Week 6 and Month 4.5 will be conducted for those subjects who are able to travel to the investigative site.

A schematic of the study design is displayed in Figure 1.

If trial evidence suggests an appropriate risk/benefit profile of CAP-1002, Capricor, upon the recommendation of the Data Safety Monitoring Board (DSMB), will introduce an open-label extension study to offer CAP-1002 to subjects who were randomized to placebo and completed the 12-month follow-up period of the original trial.

Figure 1 Trial Design



§ Approximately 7 days prior to Day 1
* Only completed for subjects that are able to travel to the investigative site

3.1 Sample Size and Randomization

This trial will be limited to a total of 20 randomized subjects.

This is a double-blind, placebo-controlled trial. Subjects who meet all enrollment criteria will be prospectively randomized in stratified permuted blocks to CAP-1002 or placebo (1:1 ratio) via an interactive web-based response system (IWRS) accessed directly by site personnel who have been delegated this responsibility by the Principal Investigator (PI). Randomization will be stratified by site and entry item score of the PUL 1.2 at Screening (Stratum I: Score of 2 or 3; Stratum II: Score of 4 or 5). Subjects will be randomized approximately 7 days in advance of their first IP infusion (Day 1). The central randomization will be generated in SAS or comparable software using permuted random block sizes held in confidence by the statistician preparing the master randomization list.

3.2 Data Safety Monitoring Board

The DSMB will act in a senior advisory capacity to Capricor regarding data and safety matters throughout the duration of the trial. The objectives, roles and responsibilities of the DSMB, as well as the format and frequency of their meetings, are documented in the DSMB charter.

3.3 Interim Analyses

3.3.1 Interim Analysis

The interim analysis was performed [redacted]

The objective of the interim analysis was to evaluate demographic data for the Intent-to-Treat (ITT) Population, study drug exposure data for the Safety Population, efficacy data for the ITT Population, Modified Intent-to-Treat (mITT) Population and Per Protocol (PP) Population, and safety and adverse event data for the Safety Population. A summary of all listings and tables that were generated to support the analysis of the interim data are outlined in Section 9.2 (APPENDIX B).

The results of the interim analysis were reviewed by the DSMB. Following the interim analysis, enrollment in the trial was discontinued (Protocol Amendment 5 dated September 27, 2019). Subsequently, three additional interim analyses with updated datasets were performed and are described below.

3.3.2 Interim Plus Analysis

The interim plus analysis was performed [redacted]. The objective of the interim plus analysis was to further evaluate unblinded efficacy data for the ITT and PP populations. A summary of all listings and tables that were generated are outlined in Section 9.3 (APPENDIX C).

3.3.3 Interim Plus 2 Analysis

The interim plus analysis was performed on [redacted] The objective of the interim plus 2 analysis was to further evaluate unblinded efficacy for the ITT and PP populations and

unblinded adverse event data for the Safety Population. A summary of all listings and tables that were generated are outlined in Section 9.4 (APPENDIX D).

3.3.4 Interim Plus 3 Analysis

The interim plus analysis was performed on [redacted]. The objective of the interim plus 3 analysis was to further evaluate unblinded efficacy data for the ITT Population and unblinded adverse event data for the Safety Population. A summary of all listings and tables that were generated are outlined in Section 9.5 (APPENDIX E).

4 GENERAL STRATEGIES FOR DATA PRESENTATION

Data with qualifiers (e.g., "<") will be listed with the qualifier but summarized without the qualifier. Subject data listings will include data collected for all subjects. Listings will be sorted by treatment, site, subject number, nominal visit, date and time (as applicable), unless otherwise indicated.

Categorical analyses will be summarized using counts and percentages. Percentages will be based on the number of subjects in the analysis set for whom there are non-missing data, unless otherwise specified. Continuous variables, including change from baseline, will be summarized using descriptive statistics [n, mean, standard deviation (SD), median, minimum, maximum].

All statistical comparisons will be performed using two-sided tests at α=0.05 significance level, unless specifically stated otherwise. P-values will be presented to four decimal places. P-values <0.0001 will be presented as "< 0.0001."

All analyses will be performed using SAS v 9.3 or higher (SAS Institute, Inc, Cary, North Carolina, USA). Validation and quality control of the tables, listings and figures (TLFs) will follow the appropriate [redacted] standard operating procedures (SOPs).

4.1 Treatment Groups

The treatment groups to be presented will be those who receive CAP-1002, those who receive placebo, and all subjects (as applicable).

4.2 Study Day

The day of first IP administration is defined as Day 1. All other study days will be labeled relative to Day 1. Thus, study day for a particular event date on or after Day 1 is calculated as: *(Date of event – Date of first IP administration + 1)*. An event that occurs prior to Day 1 is calculated as: *(Date of event – Date of IP administration)*.

The duration of an event will be calculated as *(Event end date – Event start date + 1)*. Day 0 will not be used.

The baseline value for each subject is the last non-missing value obtained prior to the time of the first IP administration on Day 1.

4.3 Handling of Dropouts or Missing Data

Subject-level listings will present data as reported. Missing or partially missing dates that are required for date-dependent definitions (e.g., treatment-emergent adverse events (TEAEs), concomitant medications) will be assumed to be the most conservative date possible.

For the primary efficacy endpoint, no imputation will be performed and the mixed-model repeated measures (MMRM) analysis will use all available data to estimate the mean treatment effect.

An AE with a completely missing start date will be considered treatment-emergent; similarly, an AE that started the same month and year as IP administration but with missing start day will be considered treatment-emergent.

AEs with missing seriousness will be counted as "serious" in tables and missing in listings; likewise, AEs with missing severity will be counted as "severe" in tables and missing in listings, and AEs with missing relatedness to the IP and/or IP administration procedure will be counted as "possible" and missing in listings.

Medical history with missing stop dates will be considered ongoing. Medications with missing stop dates will be considered in concurrent use during the study and counted in the summary table of concomitant medications unless the start date is after the last IP administration.

4.4 Multiplicity

No multiplicity adjustments will be made to any alpha levels used for statistical testing for efficacy endpoints. P-values < 0.05 at the interim analysis were used as an indication of treatment effect for all efficacy endpoints. P-values for all subsequent analyses performed are considered nominal, without conclusions drawn about statistical significance levels.

4.5 Analysis Populations

The following analysis populations will be defined.

Safety Population: The Safety population will be defined as all subjects who receive IP. Safety endpoints for subjects will be summarized and analyzed according to the treatment actually received. All safety analyses will be based on the Safety population.

Intent-to-Treat (ITT) Population: The ITT population will be defined as all subjects who are randomized. Subjects will be summarized and analyzed in the treatment group to which they are randomized. The analysis of the primary and secondary efficacy endpoints (mid-level PUL 1.2) will be based on the ITT population.

Modified Intent-to-Treat (mITT) Population: For each efficacy parameter, the mITT population included subjects in the ITT population who had at least a baseline observation (i.e., mITT populations are parameter-specific). Subjects were summarized and analyzed in the treatment group to which they were randomized. At the interim analysis, the primary and secondary efficacy (mid-level PUL 1.2) parameter analyses were repeated using the mITT population as

a supportive analysis. It was found that the mITT Population is equivalent to the ITT Population, so the mITT population was not considered for subsequent analyses and will not be considered for the final analysis.

Per Protocol (PP) Population: The PP population will include subjects who receive IP with no protocol deviations/violations that could significantly impact the completeness, accuracy and/or reliability of the trial data. The list of subjects in the per protocol population will be compiled prior to database lock. The primary, secondary and exploratory efficacy endpoint analyses will be evaluated using the PP population.

A summary of the number and percentage of subjects in each analysis population will be provided. For the mITT population, the number and percentage will be based on the primary efficacy endpoint. A listing of subject assignment into each analysis population will also be displayed.

5 STUDY POPULATION PARAMETERS

The study population parameters to be listed and summarized are described below. The ITT population will be used for the listings and summaries. Some of the summaries may also be done for other analysis populations defined in Section 4.5. All summaries will be presented by treatment group.

5.1 Eligibility and Informed Consent

Eligibility and informed consent parameters will be listed and will include date of informed consent/assent, consent to the collection and storage of exploratory biomarkers, participation in HOPE-1 protocol, protocol version, inclusion and/or exclusion criteria that were not met, if applicable, randomization date, and PUL 1.2 entry item score at randomization.

The number of subjects meeting all screening eligibility criteria and reasons for screen failure will be summarized.

5.2 Protocol Deviations

Protocol deviations will be listed and may include, but are not limited to:

- Visit not done or out-of-window
- Assessment not completed or out-of-window
- Inclusion/Exclusion
- Informed Consent/Assent
- IP Administration

The number and percentage of subjects with a protocol deviation will be summarized by protocol deviation category.

5.3 Demographics and Baseline Characteristics

Demographic and baseline characteristics will be listed and summarized. Demographic characteristics consist of age, sex, ethnicity, and race. A subject's age in years is calculated using the date of the informed consent/assent and date of birth using the following formula:

Age = [(Informed Consent/Assent Date – Birth Date + 1)/365.25].

Based on age, patients will be categorized as either pediatric (age <= 16) or adult (age > 16).

Baseline characteristics to be summarized include ulna length, height, weight, body mass index (BMI), screening PUL 1.2 entry item score, screening PUL 1.2 entry item score binary stratification (2, 3 vs. 4, 5) and dystrophin mutation type. In addition, subjects using a wheelchair or scooter for part-time and for full-time will be summarized. Subject height is either measured (i.e., standing height) or calculated from the subject's ulna length.

5.4 Medical Status Questionnaire

Subject medical status will be displayed in subject listings, which will include date of completion, frequency of walker and manual/power wheelchair or scooter use, ventilatory support, supplemental oxygen use, and frequency of falls in the last 30 days.

5.5 General Medical History

General medical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 21.0. Medical history will be listed and summarized. The summary will be presented by MedDRA System Organ Class (SOC) and Preferred Term (PT). At each level of subject summarization, a subject will be counted only once if that subject experienced the event more than once.

5.6 DMD Medical History

The following DMD medical history will be listed and summarized:

- Age at diagnosis (years),
- Age at first chronic glucocorticoid (years),
- Loss of stair climb and age at loss of stair climb (years),
- Loss of stair descend and age at loss of stair descend (years),
- Use of a walker and age at first walker use (years),
- Part-time use of a manual/power wheelchair or scooter and age at first part-time use of a manual/power wheelchair or scooter (years),
- Transition to full-time use of a manual/power wheelchair or scooter and age at first full-time use of a manual/power wheelchair or scooter (years),
- Start of non-invasive ventilatory support and age non-invasive ventilatory support started (years),

- Use of a mechanical cough assist device and age at first mechanical cough assist device use (years),
- History of spine surgery,
- History of symptomatic heart failure and age at first symptoms of heart failure (years), and
- History of medical (non-surgical) hospitalizations.

The following genetic testing results included in the DMD medical history will be listed:

- Date genetic testing sample collected
- Name of testing lab (if Clinical Laboratory Improvement Amendments (CLIA) certified)
- Mutation type
- Exon

5.7 Donor-Specific Antibodies

Donor, donor profiles, allele, HLA, HLA Class and MFI will be listed by subject and visit. Subject HLA profiles will be provided in a separate listing.

DSA data will be analyzed in the context of immune sensitization syndrome should there be an occurrence.

5.8 Physical Examination

All physical examination results will be listed. A shift table will be presented by body system for physical examination findings.

5.9 Prior and Concomitant Medications

Prescription, over-the-counter, and alternative medication use will be coded to drug class, preferred drug name, and generic/trade drug name using the World Health Organization Drug Global dictionary (WHODrug Global) version September 1, 2017. Medications that were stopped before the start of the IP administration procedure will be considered "pre-treatment." All other medications will be considered "concomitant." Medications that were started or ongoing at the time the IP administration procedure was started will be considered "baseline" (a subset of "concomitant"). Pre-treatment, baseline, and concomitant medications will be listed separately.

Frequencies and percentages of subjects reporting or receiving cardiac and steroid medications will be summarized by WHODrug Global Anatomic Therapeutic Chemical (ATC) Level 2 and Preferred Term. Pre-treatment and baseline cardiac and steroid medications will be summarized separately from concomitant cardiac and steroid medications.

At each level of subject summarization, subjects who reported one or more medication within that level are only counted once for that level.

5.10 Planned Medical/Surgical Procedures

All reported elective medical and/or surgical procedures (e.g., wisdom tooth extraction) that were not the result of an adverse event will be displayed in subject listings. For each procedure, the type of procedure, indication, and start/end date will be reported.

Any administered medication(s) related to the planned medical/surgical procedure(s) will be captured as a concomitant medication.

5.11 IP Administration

Details of IP administration will be listed. The number of subjects that completed IP infusion at each visit will be summarized. This is defined as subjects that completed (without interruption or syringe pump interrupted and restarted) both IP syringe administrations. For each visit, the total dose delivered (M cells) from both the first and second IP syringe administration, the number of patients who completed each IP syringe administration, reasons IP syringe was not completed, duration of IP administration (minutes), and total dose of IP syringe delivered (M cells) will be summarized. In addition, for the second IP syringe, the reasons the second IP syringe was not administered will also be summarized.

5.12 Subject Study Progress

A listing of subject study progress will show dates of per protocol study visits and unscheduled visits. The number of subjects who completed each visit will be summarized.

5.13 Subject Disposition

All screen failures will be listed along with the primary reason for screen failure.

Subject disposition will be summarized by treatment group for the ITT population overall, by site, and by the binary stratification of entry item score of PUL 1.2 at screening. The number of subjects who completed the study, the number of subjects who discontinued from the study, and the primary reason for discontinuation will be summarized.

6 EFFICACY ANALYSES

Listings and summaries of the efficacy parameters will be done for the ITT population, unless indicated otherwise. All summaries will be presented by treatment group. Efficacy parameters are described below.

6.1 Primary Efficacy Endpoint

The primary efficacy endpoint is change from baseline to Month 12 in the mid-level (elbow) dimension of the PUL 1.2. The hypothesis set to be tested is:

$$H_o: \mu_c - \mu_p = 0$$

$$H_a: \mu_c - \mu_p \neq 0$$

where μ_c and μ_p are population mean absolute changes from baseline in CAP-1002- and placebo-treated patients, respectively.

A MMRM analysis will be used to analyze the primary efficacy endpoint using the ITT population. The results for Month 3, 6, 9 and 12 visits will be modeled with visit, treatment, and treatment-by-visit interaction as the main effects; baseline measurement, site and binary stratification of entry item score of the PUL 1.2 at screening (2, 3 vs. 4, 5) will be included as covariates in the model. An autoregressive (AR(1)) covariance structure will be the covariance structure to model within-subject errors. The Kenward-Roger's (K-R) approximation will be used to estimate the denominator degrees of freedom. The primary comparison will be the contrast between CAP-1002 and placebo at Month 12.

As a supportive analysis, the analysis may be evaluated in the PP population.

6.2 Secondary Efficacy Endpoints

Supportive secondary efficacy endpoints will be change from baseline in:

- Mid-level (elbow) dimension of the PUL 1.2 at Months 3, 6, and 9;
- Regional systolic LV wall thickening at Months 6 and 12.

For both secondary efficacy endpoints, the same model described in Section 6.1 will be used. For the mid-level (elbow) PUL 1.2 analysis, the comparisons of interest will be the contrasts between CAP-1002 and placebo at Months 3, 6, and 9. For the regional systolic LV wall thickening analysis, the comparisons of interest will be the contrasts between CAP-1002 and placebo at Months 6 and 12.

As supportive analyses, these analyses may be evaluated in the PP population.

6.3 Exploratory Efficacy Endpoints

For an exploratory analysis, the mean change from baseline for the endpoints defined below will be analyzed using the same model described in Section 6.1.

As supportive analyses, these analyses may be evaluated in the PP population.

6.3.1 PUL Scale

All subjects, regardless of ambulatory status, will complete testing in both the 1.2 and 2.0 modules in the same preferred arm. All observed values will be listed. Observed totals and changes from baseline for each dimension [high-level (shoulder), mid-level (elbow), distal-level (wrist and hand)] will also be listed.

Using the model defined in Section 6.1, the comparisons of interest will be the contrasts between CAP-1002 and placebo for change from baseline in the following:

- Combined mid-level (elbow) and distal-level (hand and wrist) dimensions of the PUL 1.2 at Months 3, 6, 9, and 12;
- Combined high-level (shoulder), mid-level (elbow) and distal-level (hand and wrist) dimensions of the PUL 1.2 at Months 3, 6, 9, and 12;
- Mid-level (elbow) dimension of the PUL 2.0 at Months 3, 6, 9, and 12;
- Combined mid-level (elbow) and distal-level (hand and wrist) dimensions of the PUL 2.0 at Months 3, 6, 9, and 12;

- Combined high-level (shoulder), mid-level (elbow) and distal-level (hand and wrist) dimensions of the PUL 2.0 at Months 3, 6, 9, and 12.

6.3.2 Strength Testing

All subjects, regardless of ambulatory status, will complete grip, pinch (tip and key), and elbow flexion strength testing on the same side throughout the duration of the trial. Absolute values will be measured and their corresponding changes from baseline will be listed and summarized.

Using the model defined in Section 6.1, the comparisons of interest will be the contrasts between CAP-1002 and placebo for change from baseline in the following:

- Grip strength at Months 3, 6, 9, and 12;
- Tip-to-tip pinch strength at Months 3, 6, 9, and 12;
- Key pinch strength at Months 3, 6, 9, and 12;
- Elbow flexion strength at Months 3, 6, 9, and 12.

6.3.3 DMD Upper Limb PROM

The subject or parent/caretaker will evaluate the perceived difficulty in performing the activities of daily living classified into four domains including Food/Nutrition, Self Care, Household/Environment, and Leisure and Communication. Observed values and changes from baseline, overall and for each domain separately, will be listed and summarized.

Using the model defined in Section 6.1, the comparisons of interest will be the contrasts between CAP-1002 and placebo at Months 3, 6, 9, and 12 for change from baseline in each domain and overall total score.

6.3.4 Pediatric Outcomes Data Collection Instrument

Subjects and parent/caretaker will complete the PODCI questionnaires independently of one another. Observed values for each of the 6 scales (Upper Extremity and Physical Function, Transfer and Basic Mobility, Sports and Physical Functioning, Pain/Comfort, Happiness and Global Function) will be listed.

6.3.5 DMD Lifetime Mobility Scale

The subject and parents/caretakers will complete the DMD-LMS questionnaires independently of one another. Observed values for each of the 3 sub-domains (Walking and Moving, Changing and Maintaining Body Positions, Moving and Handling Objects) will be listed separately for subject-reported and parent/caretaker-reported values.

Using the model defined in Section 6.1, the comparisons of interest will be the contrasts between CAP-1002 and placebo at Months 3, 6, 9, and 12. The variable of interest is the change from baseline in the total score of a domain, calculated by summing selected items from the respective domain:

- Changing and Maintaining Body Positions Domain – Items 1, 2, 3, 6, 7, 8, 10, 13, 16, 17, 20, 22, 23, 24, 27

- Moving and Handling Objects Domain – Items 1, 2, 3, 4, 5, 7, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 25, 28, 29, 31, 33, 36, 39, 42, 45, 47

6.3.6 North Star Ambulatory Assessment and 10-Meter Walk/Run Time

Ambulatory subjects will complete the NSAA and 10MWRT. All observed values will be listed.

6.3.7 Pulmonary Function Testing

Pulmonary function tests (PFTs) will measure SVC, FVC, FEV_1, PEF, MIP, MEP, PCF, and IFR. All observed values will be displayed in subject listings.

Observed absolute values and changes from baseline will be summarized for FVC, PEF, and IFR. Percent predicted and changes from baseline will be summarized for FVC and PEF.

Using the model defined in Section 6.1, the comparisons of interest will be the contrasts between CAP-1002 and placebo at Months 3, 6, 9, and 12 for change from baseline in the following parameters:

- FVC, absolute value and percent predicted
- PEF, absolute value and percent predicted
- IFR, absolute value

6.3.8 Cardiac MRI

LV structure and function as assessed by cardiac MRI include the following parameters:

- LV (myocardium) mass
- LV end-systolic volume
- LV end-systolic volume index
- LV end-diastolic volume
- LV end-diastolic volume index
- LV stroke volume
- LV ejection fraction
- systolic wall thickening (anterior LV, lateral LV, inferior LV, septal LV)
- regional wall motion [dyssynchrony index, circumferential uniformity ratio estimate (CURE)]
- circumferential strain (global, anterior LV, lateral LV, inferior LV, septal LV)
- LV wall thickness
    - end-systolic wall thickness average (anterior, lateral, inferior, septal)
    - end-diastolic wall thickness average (anterior, lateral, inferior, septal)

In additional, the LV mass index (g/m^2) will be calculated as

$$\text{LV Mass Index } (g/m^2) = \text{LV Mass } (g)/\text{BSA } (m^2)$$

where Body Surface Area (BSA) = $0.007184 * \text{Height (cm)}^{0.725} * \text{Weight (kg)}^{0.425}$. Height and weight measurements closest to the MRI scan will be used. The standing height will be used, if provided; otherwise, the height from ulna length will be used.

All observed values will be displayed in subject listings. Observed values and changes from baseline will be summarized for all parameters except LV wall thickness.

Using the model defined in Section 6.1, the comparisons of interest will be the contrasts between CAP-1002 and placebo at Months 6 and 12 for change from baseline in all parameters except LV wall thickness.

6.3.9 Biomarkers

All observed values will be listed for biomarkers, which may include osteopontin, IL-1ß, IL-2, IL-6, IL-10, TNF-α, CXCL10, troponin I (cardiac, fast skeletal, slow skeletal), troponin T, CK-MM, CK-MB, creatine kinase (total), creatine phosphate, creatinine, myostatin, and follistatin.

6.3.10 Resource Utilization

No endpoint-specific displays will be produced.

7 SAFETY ANALYSES

Listings and summaries of the safety parameters will be done for the Safety population. All summaries will be presented by treatment group. Safety parameters are described below.

7.1 Adverse Events

AEs will be coded using MedDRA coding system (Version 21.0). Potential primary safety endpoints, described in Section 2.2.1, will be adjudicated by an independent Clinical Events Committee (CEC).

The following AEs are defined as treatment-emergent AEs (TEAEs):

- AEs occurring after the initiation of the IV catheter placement for the initial dose of IP; or
- AEs with completely missing start date, or similarly, AEs that started the same month and year as IP administration but with missing start day.

Subject listings of all adverse events as reported will be provided. A separate listing will be done for TEAEs as adjudicated by the CEC. Serious AEs, primary safety endpoints as adjudicated, TEAEs adjudicated as related to IP or administration procedure, and hypersensitivity reactions as adjudicated will be listed separately.

All summary tables will be reported using TEAEs as adjudicated by the CEC. A summary of the primary safety endpoints will be presented. In addition, TEAEs related to IP or administration procedure and hypersensitivity reactions will be summarized.

AEs will be displayed by MedDRA SOC and PT, with subjects who have the same AE counted only once for that event and subjects who have more than one AE within a SOC counted only once in that SOC. The numbers and percentages of subjects reporting an event, as well as the number of events that were reported, will be calculated by treatment group.

7.1.1 Clinical Events Committee

In order to apply some level of consistency and standardization, all potential primary safety endpoints will be adjudicated by a CEC. The CEC can combine AEs reported separately into a

single AE, can change the event description (i.e., the term used to report the AE) and can change the status of the reported event to a non-AE.

Any adjudicated event, including changed event descriptions, will be MedDRA-coded.

7.2 Partner Reported Pregnancy

Data from the Pregnancy Report Form will be listed, if applicable.

7.3 Laboratory Tests (Hematology, Serum Chemistry, and Urinalysis)

Hematological testing will include complete blood count (CBC), white blood cells (WBC) differential (absolute and percentage), hemoglobin, hematocrit, and platelet count.

Serum chemistry testing will include basic metabolic panel (Glucose, Sodium, Potassium, Chloride, Bicarbonate, Blood Urea Nitrogen (BUN), Creatinine, Calcium), comprehensive hepatic panel (Albumin, Alkaline Phosphatase, Total Protein, Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Gamma-Glutamyl Transferase (GGT), Direct Bilirubin, Total Bilirubin), and Creatine Kinase. Creatine Kinase isoenzymes (CK – Total, CK – BB, CK – MB, and CK – MM) will be included in displays as serum chemistry parameters.

Urinalysis analytes that will be tested include appearance, bilirubin, color, nitrite, occult blood, pH, protein, specific gravity, glucose, ketones, red blood cells (RBC) and WBC. If required, microscopic analysis of sediment will be performed and reported per occurrence.

For hematology, serum chemistry, and urinalysis quantitative tests, summary statistics will be presented for the observed values at each visit. Summary statistics will also be presented for the change from baseline values to each post-baseline visit. Scheduled assessments will be summarized, and unscheduled events will only be listed.

For all tests, results will be displayed in subject listings, with those values falling outside the laboratory reference range flagged. Laboratory reference ranges will be provided by the laboratory site(s) and will be included in an appendix of the clinical study report.

7.4 Vital Signs, Height, Weight and BMI

Vital signs will include systolic and diastolic blood pressure, heart rate, respiratory rate, body temperature and pulse oximetry (SpO_2). Pre-infusion observed values and changes from baseline will be summarized at each post-baseline visit. Scheduled assessments will be summarized, and unscheduled events will only be listed.

Listings of vital signs, ulna length, height, weight, and BMI will be provided. Height will either be measured (i.e., standing height) or calculated from the subject's ulna length.

7.5 12-Lead Electrocardiograms

All 12-lead electrocardiogram (ECG) results will be displayed in subject listings.

8 DEVIATIONS FROM STATISTICAL METHODS IN THE PROTOCOL

This final version 3.0 of the statistical analysis plan (SAP) supersedes all previous versions.

There were no changes to the conduct of the study since the last protocol amendment. Analyses planned but not performed and the rationales for the changes are as follows.

The following exploratory efficacy endpoints will not be analyzed:

- High-level (shoulder) dimension of the PUL 1.2;
- Distal-level (wrist and hand) dimension of the PUL 1.2;
- High-level (shoulder) dimension of the PUL 2.0;
- Distal-level (wrist and hand) dimension of the PUL 2.0;
- Grip strength, percent predicted, at Months 3, 6, 9, and 12;
- Tip-to-tip pinch strength, percent predicted, at Months 3, 6, 9, and 12;
- Key pinch strength, percent predicted, at Months 3, 6, 9, and 12;
- Elbow flexion strength, percent predicted, at Months 3, 6, 9, and 12;
- PODCI scales (subject and parent) at Months 3, 6, 9, and 12;
- Parent/caretaker-reported DMD-LMS at Months 3, 6, 9, and 12;
- NSAA at Months 3, 6, 9, and 12;
- 10MWRT at Months 3, 6, 9, and 12;
- Incidence of loss of ambulation (10MWRT > 30 seconds) at Months 3, 6, 9, and 12;
- Biomarkers at Months 3, 6, 9, and 12;
- Resource utilization, including hospitalizations for orthopedic injury or surgery and for pulmonary infections, over the course of the trial.

Originally, it was planned to perform a supportive analysis of the primary efficacy endpoint on the mITT Population. However, it was discovered that the mITT Population is equivalent to the ITT Population, so analyses of efficacy will only be performed on the ITT and PP populations.

Also, originally, it was planned to perform the secondary and exploratory efficacy endpoint analyses on the PP population only. However, it was decided that the same model will be used for the secondary and exploratory endpoints as in the primary efficacy endpoint, so the analyses will be done on the ITT Population as well for all efficacy endpoints except DMD-LMS.

9 APPENDICES

9.1 APPENDIX A: Schedule of Assessments

Procedure / Event [1]	**Screening**	**Baseline**	**Day 1**		**Week 4**	**Week 6**	**Month 3**			**Month 4.5**	**Month 6**			**Month 9**			**Mo 12 / ET**[2]
Trial Day (Visit Window)	*≤ 30 d prior to Rand.*	*≤ 2 d prior to Day 1*	*1 d*	*14 d (±3) post Inf #1*	*30 d (±7)*	*45 d (±7)*	*14-30 days pre Inf #2*	*90 d (±14)*	*14d (±3) post Inf #2*	*135 d (±14)*	*14-30 days pre Inf #3*	*180 d (±14)*	*14d (±3) post Inf #3*	*14-30 days pre Inf #4*	*270 d (±14)*	*14d (±3) post Inf #4*	*360 d (±14)*
Intravenous Infusion #			*#1*					*#2*				*#3*			*#4*		
Informed Consent / Assent	X																
Demographics	X																
Medical History	X																
Eligibility Assessment	X																
Randomization [3]	X																
Medical Status Questionnaire [4]	X	X						X				X			X		X
Prior & Concomitant Medications	X	X	X	X				X	X			X	X		X	X	X
Adverse Events[5]	X	X	X	X				X	X			X	X		X	X	X
IP Intravenous Infusion [6]			X					X				X			X		
Post-Infusion Safety Call [7]				X					X				X			X	
DMD Upper Limb PROM	X	X						X				X			X		X
PODCI	X	X						X				X			X		X
DMD Lifetime Mobility Scale	X	X						X				X			X		X
Vital Signs	X	X	X					X				X			X		X
Height / Ulna Length [8]	X	X						X				X			X		X
Weight	X	X	X					X				X			X		X

Procedure / Event [1]	**Screening**	**Baseline**	**Day 1**		**Week 4**	**Week 6**	**Month 3**			**Month 4.5**	**Month 6**			**Month 9**			**Mo 12 / ET**[2]
Trial Day (Visit Window)	*≤ 30 d prior to Rand.*	*≤ 2 d prior to Day 1*	*1 d*	*14 d (±3) post Inf #1*	*30 d (±7)*	*45 d (±7)*	*14-30 days pre Inf #2*	*90 d (±14)*	*14d (±3) post Inf #2*	*135 d (±14)*	*14-30 days pre Inf #3*	*180 d (±14)*	*14d (±3) post Inf #3*	*14-30 days pre Inf #4*	*270 d (±14)*	*14d (±3) post Inf #4*	*360 d (±14)*
Physical Examination	X											X					X
12-Lead ECG		X						X				X			X		X
Pulmonary Function Testing [9]	X2	X						X				X			X		X
PUL 1.2 and 2.0	X	X						X				X			X		X
Grip Strength	X	X						X				X			X		X
Pinch Strength (Key and Tip)	X	X						X				X			X		X
Elbow Flexion Strength	X	X						X				X			X		X
NSAA *(ambulatory subjects only)*	X	X						X				X			X		X
PSC Lab Collection [10]					X		X				X			X			
Serum Chemistry [11 12]	X	X			X			X				X			X		X
Hematology [11 13]	X	X			X			X				X			X		X
Urinalysis [11 14]	X	X			X			X				X			X		X
DMD Genetic Testing [11 15]	X																
HLA Typing [11]		X															
Donor-Specific Antibodies [11]		X			X		X				X			X			X
Biomarkers [16][17]		X				X		X		X		X			X		X
Exploratory Biomarkers [16 18]		X				X		X		X		X			X		X
Cardiac MRI [19]	X											X					X

Gray-shaded columns indicate visits that may be performed at either the investigative site or remotely at a designated central laboratory patient service center (see footnote #9)

[1] Sites will complete assessments in the following sequence: 1) QOL (DMD UL-PROM, PODCI, and DMD-LMS), 2) safety and other trial assessments, excluding blood and urine collections (e.g., vital signs, height, weight, 12-lead electrocardiogram [ECG], etc.), 3) pulmonary function testing (PFT), 4) PUL 1.2 and 2.0, 5) grip strength, 6) pinch strength (key and tip-to-tip), 7) elbow flexion strength, 8) NSAA (if ambulatory), 9) blood and urine collections, 10) cardiac MRI (if applicable), and 11) IP Infusion (if applicable). All efforts must be made for the same clinical evaluator to complete assessments for the same subject at the same time of day throughout the trial (preferably in the morning).

[2] All attempts must be made to perform the trial assessments indicated for the Month 12 visit (i.e., final comprehensive visit) for subjects that decide to early terminate from the trial before completion and after starting at least one IP infusion.

[3] Subjects will be randomized between approximately 7 days in advance of their first IP infusion on Day 1. This lead time is to ensure adequate IP shipping times (see IP Manual for shipping specifications). Randomization will be stratified by site and entry item score of the PUL 1.2 at screening. Administration of IP on Day 1 should occur within a maximum of 14 days following randomization; if a delay of more than 14 days between randomization and IP administration is unavoidable, a conversation between the Investigator and Medical Monitor should occur to determine the need for repeat assessments prior to infusion. Subjects that are not randomized may be re-screened at the discretion of the Investigator.

[4] Updates regarding frequency of wheelchair use, transition to wheelchair full time, ventilatory support, and frequency of falls.

[5] A blood sample for tryptase should be obtained within 3 hours of the onset of allergic signs or symptoms, and a second red top tube should be drawn within 3 hours of a severe allergic reaction, and frozen at -20°C, for other future analyses.

[6] All IP infusions will be conducted in an outpatient setting at the investigative site. Prior to each IP infusion, medications will be administered to the subject as determined by the Investigator based on the pre-treatment guidelines provided by Capricor and/or institutional protocols to minimize the risk of potential severe allergic reactions such as anaphylaxis. Subjects will be observed for at least 2 hours post infusion, including pulse oximetry monitoring for at least 30 minutes post infusion. Sites will observe local institutional policies related to parenteral infusions and post-infusion monitoring.

[7] A safety phone call will be performed 14 days (±3 days) after each IP infusion, and if clinically indicated, an unscheduled in-person visit will be performed at the investigative site with targeted assessments based on presentation of signs and symptoms.

[8] Ulna length will be measured in all subjects. Standing height will be measured if a subject is capable. If standing height cannot be measured, height will be calculated using a measurement of ulna length per Section 13.1.2 of the protocol.

[9] Paired forced-maneuver testing will be performed on Screening Day 1 and Day 2 (approximately 24 hours between testing sessions) to confirm acceptable FVC variance (exclusion criterion 7). PFT testing sequence: SVC, forced maneuver (FEV_1/FVC/PEF), MIP, MEP, PCF, and IFR.

[10] Blood collections for clinical laboratory and DSA testing at Week 4 and pre-dose DSA testing at Months 3, 6, and 9 will occur either at the investigative site or remotely at a designated central laboratory patient service center.

[11] Blood and urine samples will be collected using trial-specific laboratory kits, and then shipped to and tested at a central laboratory. Required sample collection for a visit may occur on multiple days; if the visit includes an IP infusion, sample collections should occur prior to IP administration. Baseline blood samples may be collected on Day 1 prior to IP administration to reduce number of venipunctures.

[12] Basic metabolic panel (Glucose, Sodium, Potassium, Chloride, Bicarbonate, BUN, Creatinine, Calcium), comprehensive hepatic panel (Albumin, Alkaline Phosphatase, Total Protein, ALT, AST, GGT, Direct Bilirubin, Total Bilirubin).

[13] CBC with WBC differential, hemoglobin, hematocrit and platelet count

[14] Appearance, bilirubin, color, nitrite, occult blood, pH, protein, specific gravity, glucose, ketones, RBC, and WBC. If required, microscopic analysis of sediment will be performed and reported per occurrence

[15] DMD genetic testing will be available at screening for subjects who have not previously completed genetic testing or whose historical genetic results were not performed at a CLIA-certified laboratory as determined by an Investigator. The turn-around time for final results may exceed the 30-day screening window; allow 29-49 days for receipt of final test results.

[16] Week 6 and Month 4.5 blood samples will only be collected at the investigative site.

[17] Biomarkers may include those of muscle inflammation (osteopontin); systemic inflammation (IL-1ß, IL-2, IL-6, IL-10, TNF-α, CXCL10); muscle damage (troponin I [cardiac, fast skeletal, slow skeletal], troponin T, CK-MM, CK-MB, creatine kinase [total], creatine phosphate, and creatinine); and muscle regeneration (myostatin and follistatin).

[18] Collected only if separate informed consent and/or assent provided

[19] Subjects will undergo cardiac MRI if they are physically capable as determined by an Investigator. Subjects that cannot complete a cardiac MRI as part of the trial will complete an echocardiogram at screening using the site's local equipment, acquisition protocol, and evaluation procedures. The reported LVEF from a local interpretation, whether it's measured via cardiac MRI or echocardiogram, will be used to assess eligibility (exclusion criterion 1).

9.2 APPENDIX B: Table of Contents – Interim Analysis Tables and Listings

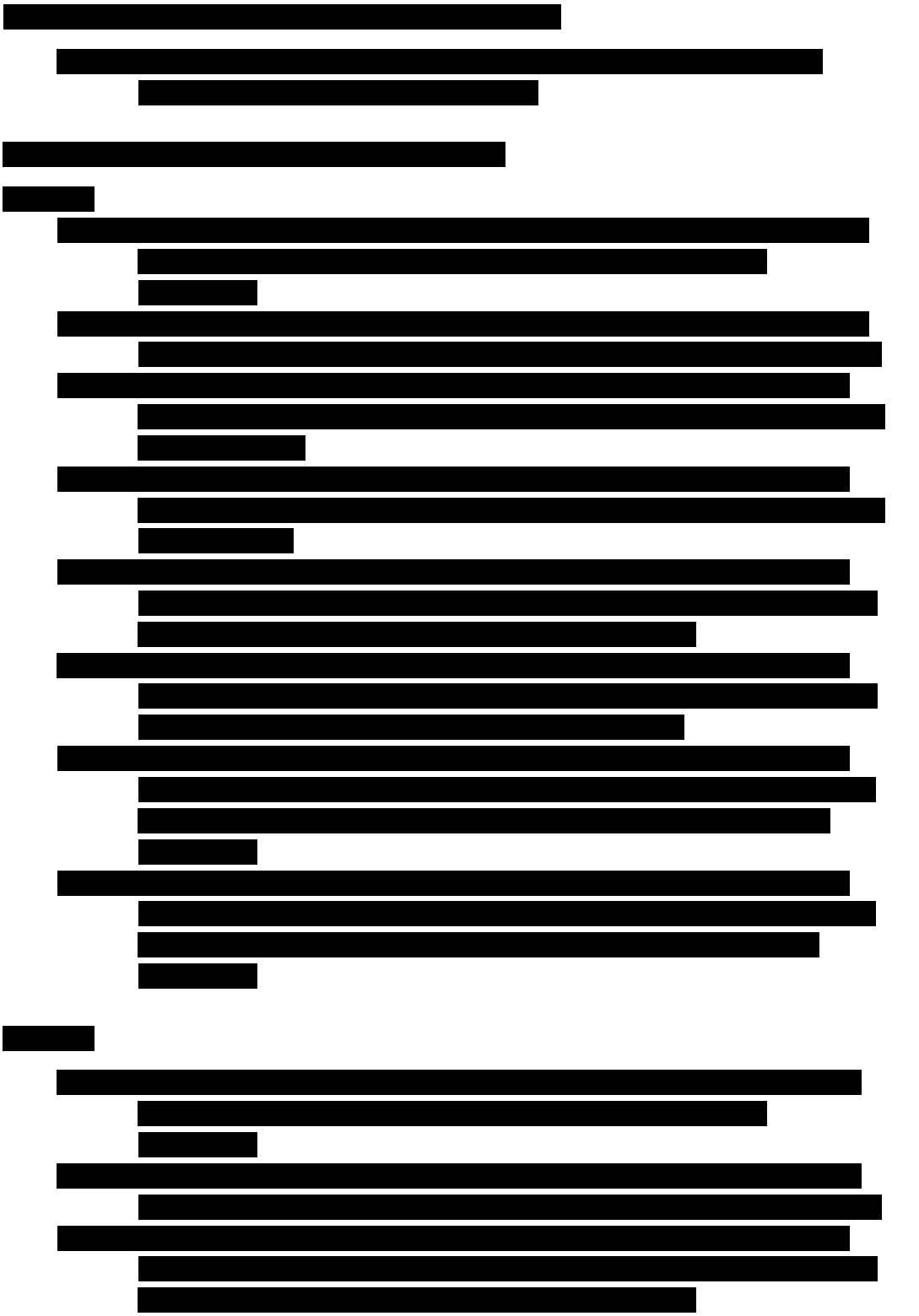
[redacted]

[redacted]

[redacted]

[redacted]

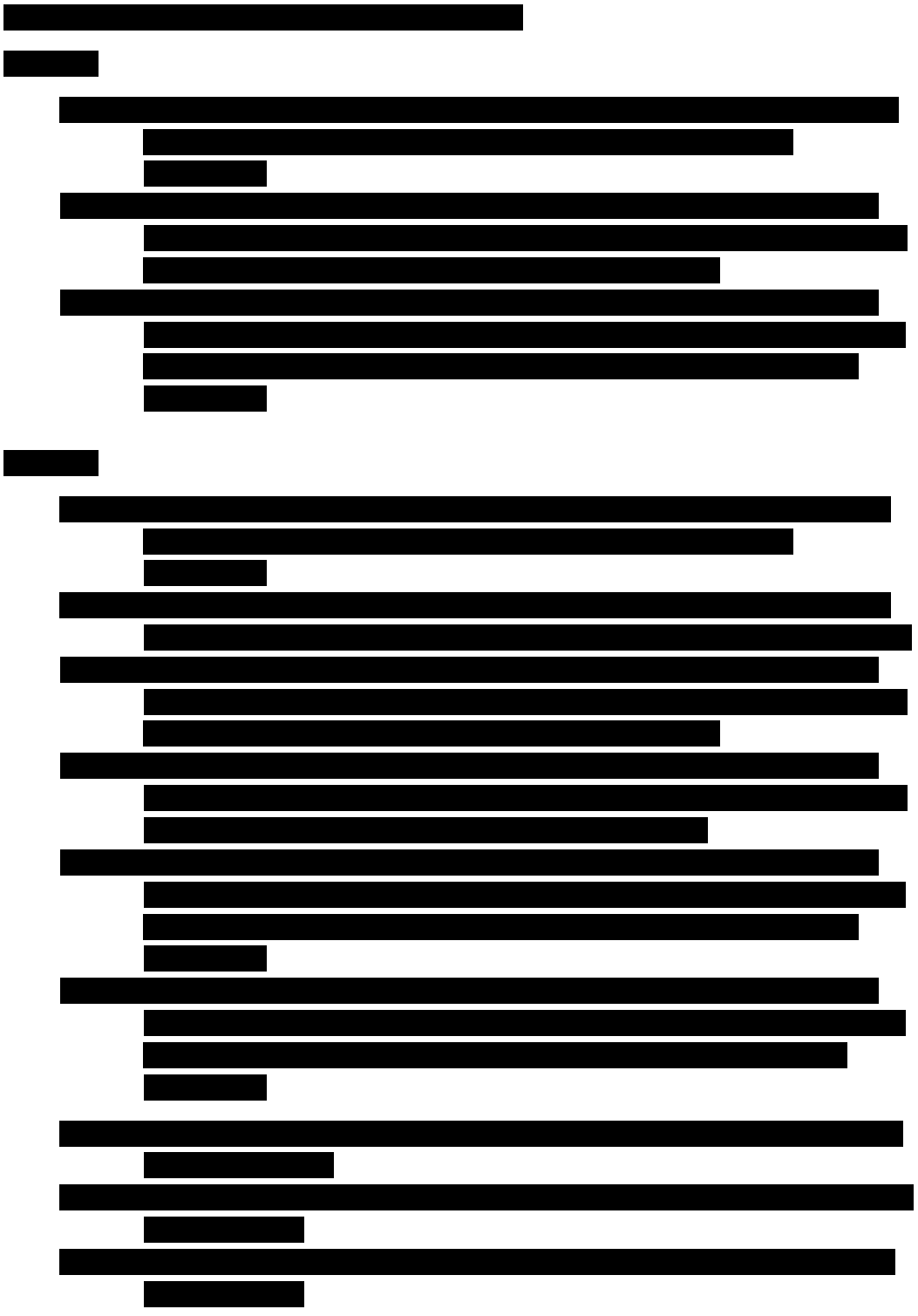
9.3 APPENDIX C: Table of Contents – Interim Plus Analysis Tables and Listings



[redacted]

[redacted]

9.4 APPENDIX D: Table of Contents – Interim Plus 2 Analysis Tables and Listings



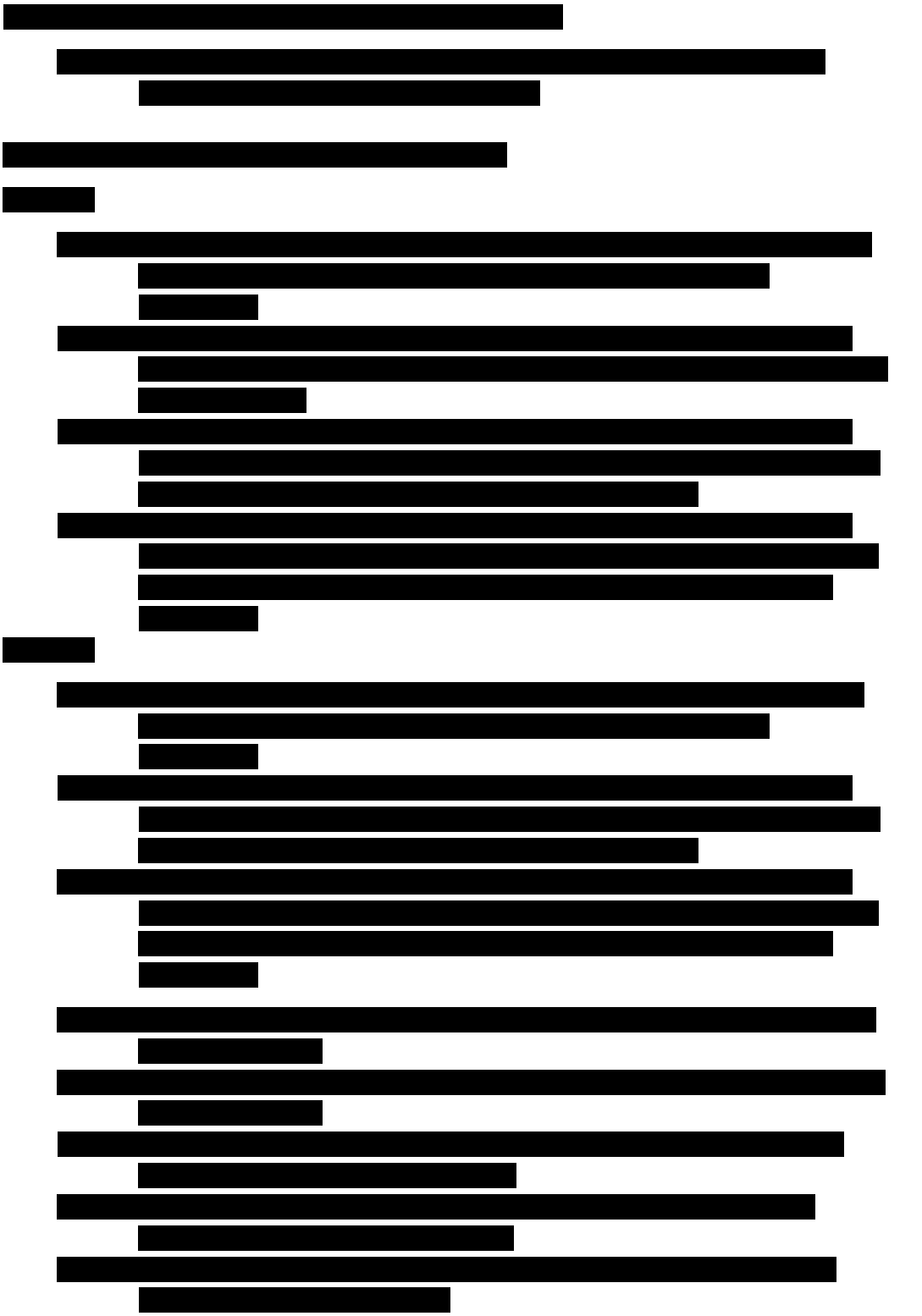
[redacted]

[redacted]

[redacted]

[redacted]

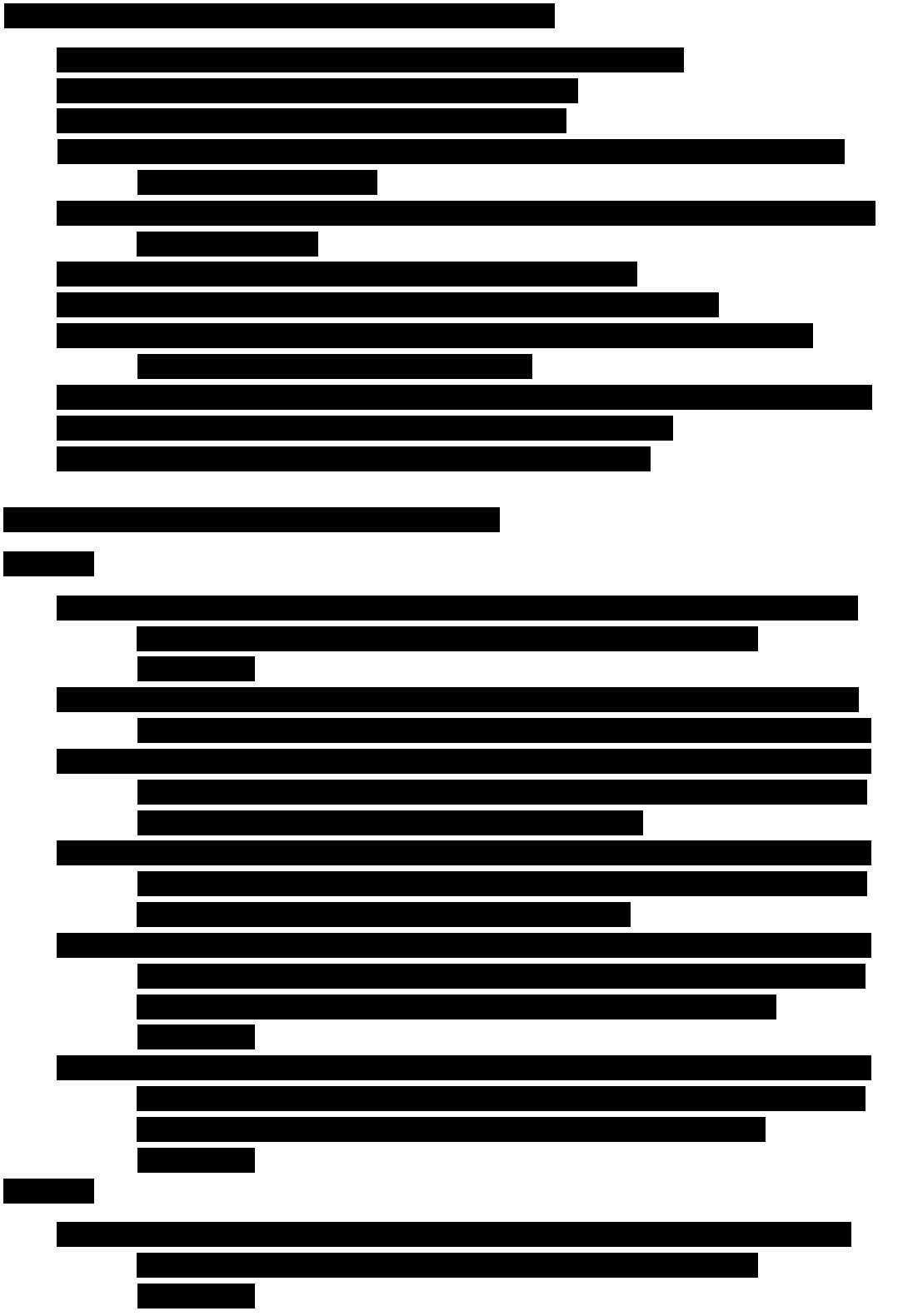
9.5 APPENDIX E: Table of Contents – Interim Plus 3 Analysis Tables and Listings



[redacted]

[redacted]

9.6 APPENDIX F: Table of Contents – Final Analysis Tables and Listings



[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

SAP Addendum – Post Hoc Analysis HOPE-2 Study

1. Introduction

The following analyses are post-hoc. The additional efficacy analyses outlined below were deemed necessary to assess the robustness of inferences from the pre-specified analyses to violations of the normality assumptions in the pre-specified analyses that arose from the smaller-than-planned sample size. For completeness, this Addendum reinstates multiple protocol-specified endpoints previously removed from the final analysis via the Final SAP v3.0. In addition, Adverse Event and cardiac and steroid medication summaries and listings are further detailed in this Addendum.

2. Efficacy Analyses

A parametric linear mixed effects model for repeated measures (MMRM) is the pre-specified analysis for all primary, secondary, and exploratory endpoints of the study. The original study and statistical analysis plan anticipate a total sample size of around 80 subjects. However, the final number of subjects enrolled in the study is 20. Because of the smaller sample size, the primary analysis model is less likely to be robust to departures for certain assumptions[1]. Beyond simply identifying if violations of assumptions exist, a critical aspect of model checking is to understand the consequences violations (if present) had on data interpretations[2].

Specifically, the smaller than intended sample size necessitates increased scrutiny on the assumption of the normality of residuals in the pre-specified primary, parametric MMRM analysis. Concern regarding robustness to departures from normality of the residuals supports the use of additional analyses to supplement the pre-specified model in the HOPE-2 trial.

A common method for dealing with non-normality in a study with less than 30 subjects is to employ a non-parametric test[1]. Non-parametric models generally require fewer assumptions than parametric models and therefore inferences from non-parametric models are less dependent on assumptions[1,2]. However, parametric analyses generally have more power if assumptions are valid, which is also an important consideration when sample size is small. Data transformations can be used to alter the original distribution into a transformed distribution for which the normality assumption is more likely to be valid with small sample sizes while retaining use of a parametric model. Therefore, a data transformation can be used to satisfy normality assumptions without the loss of power inherent to non-parametric methods.

Results from a non-parametric and / or data-transformed analyses can be compared to the corresponding results from a parametric analysis to assess the consequences of departures from assumed conditions. Additional sensitivity analyses can be used to assess the robustness of other aspects of the primary and sensitivity modeling approaches[2].

2.1. Assessment of Departures from Normality Assumptions and their Consequences in the HOPE-2 Trial

The residuals from the pre-specified model of the primary analysis in the HOPE-2 trial will be plotted in a histogram and inspected visually. An inferential assessment for departure from normality will be implemented via the Shapiro-Wilk test.

A data transformation approach will be implemented that retains as many aspects of the primary analysis as possible. This approach will model changes from baseline (and baseline values) after transforming (converting) them to percentile ranks. The primary model will then be fit to the percentile ranks. The estimates from this alternative analysis will be re-transformed back to the original scale to facilitate interpretation on the original scale.

2.2. Details of the Data Transformed Model

The primary efficacy analysis is change from baseline to Month 12 in the mid-level (elbow) dimension of the PUL 1.2 assessed via a parametric linear MMRM model to compare treatment groups in mean change from baseline to Month 12. If non-normality of the residuals is identified via the procedure identified in Section 2, a data transformation will be applied and then the primary MMRM model will be applied to the transformed data. This primary MMRM model on the transformed data tests the same hypotheses.

Ho: $\mu c - \mu p = 0$
Ha: $\mu c - \mu p \neq 0$

where μc and μp are population mean of changes from baseline in CAP-1002- and placebo-treated subjects, respectively, which on the transformed data is also a test of the medians because the mean and the median of a distribution of percentiles are the same.

To implement this model, an MMRM analysis will be used to analyze the percentile ranked primary efficacy endpoint data using the same pre-specified ITT population. Data from the post-baseline visits (Month 3, 6, 9 and 12) will be modelled with visit, treatment, and treatment-by-visit interaction as the main effects, and ranked baseline score, site and binary stratification of entry item score of the PUL 1.2 at screening (2, 3 vs. 4, 5) as covariates. A First Order Autoregressive (AR(1)) covariance structure will be used to model the within-subject errors. The Kenward-Roger's (K-R) approximation will be used to estimate the denominator degrees of freedom. The primary comparison will be the contrast between CAP-1002 and placebo at Month 12.

For a given outcome, this analysis will be implemented by calculating the percentile rank of each change-from-baseline value relative to all observed change-from-baseline values for that outcome (across all subjects and all post-baseline observation times). The percentile ranked change-from-baseline is the dependent variable in the model.
This modeling approach will be applied to primary, secondary, and some exploratory endpoints for the ITT Population only.

In addition, for PUL and Cardiac MRI outcomes, Global Statistical Tests[3] (GST) will be calculated as the averaged percentile ranked change-from-baseline.

2.3. Sensitivity Models

Several sensitivity models will be constructed to assess the robustness of the non-parametric model described above to other potential aspects of model misspecification. These analyses are described below and will only be performed for the primary endpoint in the ITT Population.

2.3.1. Permutation Test

Permutation tests at each analysis Visit (Month 3, 6, 9, and 12) will also be conducted as a non-parametric test for assessing robustness of the primary MMRM result. Permutation tests will be conducted by randomly sampling (without replacement) all possible permutations of the $N_{p,t}$ and $N_{T,t}$ values of the outcome variable, where $N_{p,t}$, $N_{T,t}$ are the number of subjects in placebo and treatment, respectively, at time t. The p-value is the proportion of permuted samples that have a test statistic larger than that of the observed data.

2.3.2. *Study Site*

An additional issue related to the smaller than originally intended sample size is the number of subjects per site. Ideally, there should be at least two subjects per site per treatment at each post baseline visit in order to cleanly estimate the site effect and the treatment effect along with the uncertainty / inconsistency in those effects. If the number of observations falls below this threshold, the treatment and site effects become (partially) confounded and treatment effect estimates may be biased.

In the smaller than planned HOPE-2 study, two sites have only 1 subject, and 3 sites have only 1 subject per arm. Therefore, confounding and the resulting potential bias in the estimated treatment effect is a concern. Two sensitivity models will be performed to assess the effect from this potential confounding. Each change detailed below will be implemented in the non-parametric model described above.

1. Site will be pooled until there were at least 2 subjects per pooled site per treatment arm. This will be achieved by starting with the site(s) with the fewest subjects and pooling with the next largest site(s) until the criterion is achieved.
2. Site will be removed from the model.

2.3.3. *Outlier Removal*

In addition to conducting the non-parametric analyses, outlier (unusual) observations that are driving the non-normality can be removed so that the normality assumptions of the pre-specified parametric model hold. After removal of the outlier subject(s) from the data, the residuals can be reassessed for normality both graphically and inferentially via the Shapiro-Wilk Test.

Although outlier removal can be a useful sensitivity analysis, it is not usually an appropriate primary approach because it violates the intention-to-treat principle that is central to unbiased assessment and is a focus of ICH E9 statistical guidance. Removing just one subject in this small

sample would remove 5% of the data and is therefore deemed less useful than non-parametric analyses, especially given that no a priori cut off for identifying outliers was pre-specified.

2.3.4. *Imputation of Missing Data*

The selected endpoints to be evaluated using imputation of missing data are as follows:

- change from baseline for the mid-level (elbow) dimension of the PUL 1.2 at Months 3, 6, 9, and 12
- change from baseline for the combined high-level (shoulder), mid-level (elbow), and distal-level (hand and wrist) dimensions of the PUL 2.0 at Months 3, 6, 9, and 12
- change from baseline for LVEF as assessed by cardiac MRI at Months 3, 6, 9, and 12

The sensitivity of the analysis of the selected endpoints above to missing data will be evaluated in the ITT population using two different techniques to impute that missing data:

1. Last z-score carried forward
   and
2. Multiple imputation with 500 replicates using control-based pattern imputation for monotone missing data

To implement the first method, z-scores will be calculated by visit using all raw observed values at that visit prior to carrying the last available z-score forward to any missing data points. Once all missing data is imputed with the last available z-score, change from baseline in raw z-score will be calculated and analyzed using the primary MMRM, and percentile ranked change from baseline in raw z-score will be calculated and analyzed using the primary MMRM.

Subjects who do not have data for an intermediate visit but do have data for subsequent visits may reasonably follow the same trajectory as subjects in their respective arm, but subjects who discontinue treatment may follow the same trajectory as placebo-arm subjects. To use this reasoning with multiple imputation, intermediate missing data will first be imputed with random sampling using non-missing data from that subject's respective treatment arm so the resulting missing data pattern will be monotone. The remaining missing data will then be imputed with random sampling from non-missing placebo arm data. Once all missing data is imputed, change from baseline will be calculated and analyzed using the primary MMRM, and percentile ranked change from baseline will be calculated and analyzed using the primary MMRM.

2.3.5. *Additional Transformations*

The sensitivity of the analysis of change from baseline for the mid-level (elbow) dimension of the PUL 1.2 at Months 3, 6, 9, and 12 to data transformations will be evaluated in the ITT population using two additional data transformation methods on the raw baseline and change from baseline values:

1. Ordinal ranking
   and
2. Cube-root transformation

Once the data is transformed using the above methods, the transformed values will be analyzed using the primary MMRM.

2.4. Additional Parameters

In addition to the protocol-specified efficacy endpoints, other meaningful parameters may be analyzed, such as combined PUL dimensions (e.g., combined mid-level (elbow) and distal-level (hand and wrist) dimension) for the ITT Population and PP Populations.

2.5. Data Visualization

Figures depicting model-adjusted least-squares means (LS-Means) by visit broken out by treatment arm will be produced for the primary, secondary, and some exploratory efficacy outcomes. This will apply to both raw and transformed outcome measures and their corresponding model estimates.

A forest plot will be produced for all PUL endpoints showing the magnitude, significance, and errors associated with several measures at Month 12 simultaneously. This plot will also include the GST of all PUL outcomes at Month 12. The same forest plot will be produced for Cardiac MRI outcomes, including the corresponding GST. Forest plots will only be produced for transformed outcome measures.

All figures will be produced for the ITT Population only.

2.6. Protocol-Specified Endpoints Previously Removed from Analysis

The final version 3.0 of the statistical analysis plan removed some protocol-specified exploratory endpoints from the final analysis. This Addendum reinstates the following exploratory efficacy endpoints for analysis:

- High-level (shoulder) dimension of the PUL 1.2;
- Distal-level (wrist and hand) dimension of the PUL 1.2;
- High-level (shoulder) dimension of the PUL 2.0;
- Distal-level (wrist and hand) dimension of the PUL 2.0;
- Grip strength, percent predicted, at Months 3, 6, 9, and 12;
- Tip-to-tip pinch strength, percent predicted, at Months 3, 6, 9, and 12;
- Key pinch strength, percent predicted, at Months 3, 6, 9, and 12;
- Elbow flexion strength, percent predicted, at Months 3, 6, 9, and 12;
- Pediatric Outcomes Data Collection Instrument (PODCI) scales (subject and parent) at Months 3, 6, 9, and 12;
- Parent/caretaker-reported DMD Lifetime Mobility Scale (DMD-LMS) at Months 3, 6, 9, and 12;

- North Star Ambulatory Assessment (NSAA) at Months 3, 6, 9, and 12;
- Biomarkers at Months 3, 6, 9, and 12.

All PUL and strength endpoints above will be analyzed using summary statistics, the primary MMRM, and the percentile ranked model described above for the ITT and PP Populations. All available data listed for the ITT population.

CK Isoenzyme biomarkers will be analyzed similarly but for the safety population, with all available data listed for the safety population within the serum chemistry laboratory listing.

Other biomarkers will be analyzed only using the percentile ranked model for the ITT population, with all available data listed for the ITT population.

PODCI, parent/caretaker-reported DMD-LMS, and NSAA endpoints above will be listed for the ITT Population.

The following exploratory efficacy endpoints will not be reinstated:

- 10-Meter Walk/Run Time (10MWRT) at Months 3, 6, 9, and 12;
- Incidence of loss of ambulation (10MWRT > 30 seconds) at Months 3, 6, 9, and 12;
- Resource utilization, including hospitalizations for orthopedic injury or surgery and for pulmonary infections, over the course of the trial.

3. Safety Analyses

3.1. Adverse Events

This study utilized an independent Clinical Events Committee (CEC) to adjudicate all potential primary safety endpoints. The CEC data review was conducted in a blinded fashion and the results were provided only when the adjudication process yielded a different conclusion on the Treatment-Emergent Adverse Events (TEAE) as compared to the original TEAE assessment (i.e., the raw TEAE data). Since the adjudicated AE data set only contains the subset of adverse events which the adjudication committee reviewed, these data will be listed separately and not summarized due to their sparseness. In order to provide a comprehensive assessment of all TEAEs, the raw TEAE data will be reported in listings and summary tables for the safety population.

3.2. Cardiac and Steroid Medications

Frequencies and percentages of subjects reported or receiving WHODrug Global cardiac medications at any time on or after Day 1 (date of first IP exposure) will be both summarized by WHODrug Global Anatomic Therapeutic Chemical (ATC) Level 2 and Preferred Term and listed.

Frequencies and percentages of subjects reported or receiving WHODrug Global steroid medications at any time on or after Day 1 (date of first IP exposure) will be summarized separately for systemic use and non-systemic use, where steroid medications with reported

indications as DMD, DMD treatment, or SOC for DMD are considered for systemic use, and steroid medications with other indications are considered for non-systemic use. Both summaries will be presented by WHODrug Preferred Term. All steroid medications on study, along with their reported indications, will be listed.

4. Pharmacoeconomic Analysis

An exploratory analysis was planned to explore all-cause hospitalizations, hospitalizations for orthopedic injury/surgery or pulmonary infections, and antibiotic usage. Billing information was not to be collected; rather, standardized, composite or bundled costings of national averages was planned to be used in the analysis. Given the small sample size and minimal observations on trial, this analysis was not completed.

References

1. Siegel S. Nonparametric Statistics. Am Stat 1957;11(3):13–9.

2. Mallinckrodt CH. and Ilya Lipkovich. A practical guide to analyzing longitudinal clinical trial data. CRC Press. Boca Raton. 2017 (Chapter 11).

3. O'Brien PC. Procedures for comparing samples with multiple endpoints. Biometrics 1984; 40: 1079-87.

Signature Approval of SAP Addendum – Post Hoc Analysis HOPE-2 Study

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 (HOPE-2)