Protocol B5161004

A MULTICENTER, OPEN-LABEL EXTENSION STUDY TO EVALUATE THE LONG TERM SAFETY OF PF-06252616 IN BOYS WITH DUCHENNE MUSCULAR DYSTROPHY

Statistical Analysis Plan (SAP)

Version: 4.0

Author: PPD Cambridge MA

Date: 12-Feb-2018

TABLE OF CONTENTS

LIST OF TABLES

1. VERSION HISTORY

Version 4 of the Statistical Analysis Plan (SAP) for study B5161004 clarified analyses to integrate the efficacy and safety results across B5161002 and B5161004. It added new analyses to compare outcomes with a historical control group. It also made a few clarifications and added additional summaries for the safety, efficacy and PD endpoints.

Version 3 of the Statistical Analysis Plan (SAP) for study B5161004 clarified that summaries of change from overall baseline will involve combination of B5161002 and B5161004 data. It also clarified the definition of loss of ambulation and removed analysis of loss of ability to perform 6 minute walk test, 4 stair climb, rise from floor or run/walk 10 meters.

Version 2 of the Statistical Analysis Plan (SAP) for study B5161004 reflected the changes in the revised B5161002 protocol dated 15 AUG2016 (Amendment 2). The SAP clarified the subgroup category of baseline 4SC.

The initial version of the Statistical Analysis Plan (SAP) for study B5161004 is based on the protocol dated 08Apr2016.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
1	Not Applicable	Not Applicable
2	Update the author's affiliation Section 6.2.1: Update subgroup category	To comply with the current organizational structure To comply with the parent study B5161002 SAP amendment 3.
3	Section 6.2.1: Added clarification that summaries of change from overall baseline will combine B5161002 and B5161004 data. Removed summary of number and percentage of subjects who lose the ability to perform 6 minute walk test, 4 stair	To investigate the long term effect of PF-06252616.

	climb, rise from floor or run/walk 10 meters. Modified definition of loss of ambulation in B5161002.	Analysis not needed. To get more accurate definition of loss of ambulation.
4	In Section 3.2.1 and Section 3.3.2.	Clarified endpoint details including the pre-specified subsets, the timed function tests within the NSAA endpoint, and the PODCI score.
4	In Section 3.5.1	Clarified the AE Tier-2 event definition to align with the "Rule of 4" given the trial will have less than 400 or fewer subjects per treatment group, and clarified the AE Tier-3 event analysis.
4	In Section 6.3.2.1	Added descriptions about the historical control group due to the B5161002 protocol amendment 2 (dated 15Aug2016).
4	In Section 6.2.1	Added a summary plot of the cumulative percent of patients with a percent change from baseline on 4SC time less than a given set of thresholds as an additional sensitivity analysis.
4	In Section 5.1, 5.2 and 6.3.2.2	Added comparisons of sequence 1 data to the historical control group due to the B5161002 protocol amendment 2 (dated 15Aug2016).
4	In Section 6.4.3	Added descriptive summary of the study drug exposure to facilitate data review.
4	In Section 6.5.5	Added shift table analysis to further characterize the cardiac safety.
4	In Section 6.5.7	Clarified that the safety data will be either listed and/or summarized. Added details about analysis of the bone mineral density data.

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study B5161004. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment. In this document any text taken directly from the protocol is ***italicized***.

2.1. Study Objectives

This study is an open-label extension (OLE) to Protocol B5161002 and will provide an assessment of the long term safety, efficacy, pharmacodynamics (PD) and pharmacokinetics (PK) of intravenous (IV) dosing of PF-06252616 in boys with Duchenne muscular dystrophy (DMD).

2.1.1. Primary Objective

To evaluate the long-term safety of IV dosing of PF-06252616 in boys with DMD.

2.1.2. Secondary Objectives

- *To evaluate the long-term efficacy of PF-06252616 using functional assessments and strength.*
- *To assess the PK and immunogenicity of PF-06252616.*

2.1.3. Exploratory Objectives

- *To evaluate PD markers that may be informative in demonstrating the pharmacologic effect of PF-06252616.*
- *To evaluate the long-term functional health effects of PF-06252616 on Parent-report and Adolescent self-report scores on the Pediatric Outcomes Data Collection Instrument (PODCI).*
- *To assess the long-term effects of PF-06252616 on Health-related quality of life (HRQL) and healthcare resource utilization (HRU) in boys with DMD.*
- *To evaluate the long-term effects of PF-06252616 on caregiver burden, HRQL and work productivity and activity impairment.*
- *To collect exploratory biomarker samples for bio-banking.*

2.2. Study Design

Approximately 105 eligible subjects will be assigned to receive an individualized maximum tolerated dose based on their tolerability profile/data from B5161002. No placebo comparator will be assessed. Consenting subjects who complete the B5161002 study will be invited to transition directly to the OLE study. Subjects' results from the B5161002 study may be used as screening data for the current study.

The individual dose level selected for each subject will be based on the maximum tolerated dose from the parent Study B5161002. Subjects will be dosed with one of three PF-06252616 dose levels: 5 mg/kg, 20 mg/kg or 40 mg/kg.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS ANDCONVENTIONS

3.1. Primary Endpoint(s)

3.1.1. Safety

- *Incidence and/or rate of intolerability or dose limiting treatment related adverse events (AEs) following up to 4 years of treatment.*

- *Incidence and/or rate, severity and causal relationship of treatment emergent AEs (TEAEs) and withdrawals due to TEAEs following up to 4 years of treatment.*

- *Incidence and magnitude of abnormal laboratory findings (clinical laboratory tests: hematology, chemistry, gamma-glutamyl transferase (GGT), glutamate dehydrogenase (GLDH), prothrombin time (PT), activated partial thromboplastin time (aPTT), creatine kinase, amylase, serum ferritin, serum iron, total iron binding capacity (TIBC), % transferrin saturation. Hormones: luteinizing hormone [LH], follicle stimulating hormone [FSH], thyroxine [T4], thyroid stimulating hormone [TSH]. Fecal occult blood, cardiac Troponin I and urinalysis) following up to 4 years of treatment.*

- *Abnormal and clinically relevant changes in liver magnetic resonance imaging (MRI) and physical examinations (including nose and throat mucosal exam and Tanner stage and Testicular Volume), weight, vitals, electrocardiogram (ECG), left ventricular ejection fraction (LVEF) measured by cardiac MRI (or echocardiogram), bone mineral density (BMD) by dual energy x-ray absorptiometry (DXA), x-ray (hand and wrist for bone age evaluation) and Columbia Suicide Severity Rating Scale (C-SSRS).*

3.2. Secondary Endpoint(s)

3.2.1. Efficacy

- *Mean change from baseline following up to 4 years of treatment in the following functional assessment tests: pulmonary function tests (PFTs) (to include forced vital capacity [FVC], forced expiratory volume in one second [FEV_{1}] and peak expiratory flow rate [PEFR]), four stair climb (4SC), Northstar Ambulatory Assessment (NSAA), Performance of Upper Limb (PUL), range of motion (ROM) and 6 minute walk distance (6MWD).*

- *Mean change from baseline in muscle strength measured by myometry following up to 4 years of treatment.*

The time function tests including time-to-stand and 10 meter walk/run from the Northstar ambulatory assessment will be analyzed separately for summary tabulation along with the total NSAA score.

For all the efficacy endpoints, the baseline is defined as the last pre-dose assessment prior to the first day of dosing in study B5161002. This baseline value is denoted as "overall baseline". To evaluate the changes in the efficacy endpoints during the long term follow-up study B5161004, a second baseline will be defined as the last assessment prior to dosing on Day 1 in B5161004. This baseline value is denoted as "B5161004 baseline". The baseline value for B5161004 will be the screening visit which may in some cases, be the same data as the Week 97 visit for study B5161002. PFTs will be repeated up to six times at the site.

For the pulmonary function testing, testing will be conducted to produce 3 technically adequate results for each measure (with no more than 6 attempts). Only the single best (maximum value) measure will be reported on the CRF. The highest PEFR will be recorded in the CRF in L/min. The percent predicted FVC and FEV_1 will be calculated based on the best (maximum) FVC and FEV_1 measurement. The best measurement for FVC and FEV_1 may occur on different efforts.

3.2.2. PK and Immunogenicity

- *Trough serum PF-06252616 concentrations for all subjects receiving active drug.*
- *Incidence of anti-drug antibody (ADA) and neutralizing antibody (Nab).*

3.3. Other Endpoints

3.3.1. Pharmacodynamic Endpoints

- *Change from baseline of Lean Body Mass (LBM) determined via whole body DXA after up to 4 years of treatment.*

Two baselines will be considered for the purposes of the analyses. The first baseline will be the last acceptable pre-dose measurement before the first day of dosing in study B5161002. This baseline value is denoted as "overall baseline". To evaluate the changes in LBM during the long term follow-up study B5161004, a second baseline will be defined as the last assessment prior to dosing on Day 1 in B5161004. This baseline value is denoted as "B5161004 baseline". The baseline value for B5161004 will be the screening visit which may in some cases, be the same data as the Week 97 visit for study B5161002 .

3.3.2. Clinical Outcome Assessments

- *Changes from baseline in the Pediatric Outcomes Data Collection Instrument (PODCI) Parent-report and Adolescent self-report scores.*
- *Change from baseline in the EuroQol 5 Dimensions 3 Levels (EQ-5D-3L) Health Questionnaire.*

- *Change from baseline in the EuroQol 5 Dimensions - Youth (EQ-5D-Y) Health Questionnaire.*

- *Change from baseline in Healthcare Resource Utilization (HRU) questionnaire.*

- *Change from baseline in Zarit Burden Interview (ZBI).*

- *Change from baseline in the Work Productivity and Activity Impairment. Questionnaire adapted for caregiving (WPAI:CG) in the percent work time missed, percent impairment while working, percent overall work impairment, and percent activity impairment due to a child's Duchenne muscular dystrophy.*

- *Pooled or exploratory analyses, if conducted, utilizing the biobanked exploratory genomic and biomarker samples will be documented in a separate statistical analysis plan.*

Baseline for PODCI parent-report score will be the last acceptable pre-dose measurement taken before the first day of dosing in study B5161002. This baseline value is denoted as "overall baseline".

To evaluate the changes in PODCI parent-report score during the long term follow-up study B5161004, a second baseline will be defined as the last assessment prior to dosing on Day 1 in B5161004. This baseline value is denoted as "B5161004 baseline". The baseline value for B5161004 will be the screening visit which may in some cases, be the same data as the Week 97 visit for study B5161002 .

PODCI Adolescent self-reported version will be completed in study B5161004 once the subjects are 11 years or older. Baseline for PODCI Adolescent self-report is not applicable because subjects start self-reporting at different times in the study.

The PODCI results will be presented by the Global Function score and the five domains: Upper Extremity and Physical Function; Transfer and Basic Mobility; Sports and Physical Functioning; Pain/Comfort and Happiness.

For other outcome assessments, since they are not collected in B5161002, baseline will be the screening measurement in study B5161004.

3.4. Baseline Variables

Baseline variables including PFTs, 4SC, 6MWD, or demographics (age at enrolment, weight [continuous and/or Pfizer standard categorical cut-offs]), may be explored as covariates in the modeling of the efficacy endpoints.

Covariates will be summarized for the Safety Analysis Set and Per Protocol Analysis Set (PPAS) and by treatment sequence to which subjects had been randomized for study B5161002 (abbreviated as "treatment sequence" in the following paragraphs). Continuous baseline covariates will be summarized by: n, mean, median, standard deviation, min and max. Binary and categorical covariates will be summarized by percent and counts.

In study B5161002, subjects are stratified at the time of randomization into two groups based on their ability to complete the 4 stair climb (4SC) in ≤8 seconds or >8 second at baseline. This stratification factor may be used as a covariate in the statistical model for study B5161004 when modeling data from tests other than the 4SC. For the 4SC analysis, the actual baseline value instead of the stratification factor will be included in the model. The baseline value may also be included as a covariate in the model for the other functional, strength, or imaging assessments.

Additional covariates including the qualitative assessment for the 4SC will be summarized over time and may be explored in the modeling of the 4SC statistical analysis.

3.5. Safety Endpoints

3.5.1. Adverse Events

If an AE starts prior to dosing on Day 1 in study B5161004 and is ongoing at the time of dosing on Day 1, it will not be considered as a TEAE in B5161004. This will prevent double counting an AE in both the parent study B5161002 and follow-up study B5161004. A 3-tier approach will be used to summarize AEs. Under this approach, AEs are classified into 1 of 3 tiers. Different analyses will be performed for different tiers (See Section 6.5.1).

Tier-1 events: These are pre-specified events of clinical importance and are maintained in a list in the product's Safety Review Plan (GDMS Location: /Compounds/PF-06/PF-06252616/Product Strategy and Administration/Product Strategy/Safety Review Plan or see link for current version: http://gdms.pfizer.com/gdms/drl/objectId/090177e18692171a)

Tier-2 events: These are events that are not tier-1 but are "common". A MedDRA PT is defined as a tier-2 event if there are 4 or more subjects with the same PT in any one treatment group, given the trial will have less than 400 or fewer subjects per treatment group ("Rule of 4").

Tier-3 events: These are events that are neither tier-1 nor tier-2 events. Pfizer standard safety output where all AEs will be included (ie, no new outputs)

Severity is the key baseline information for adverse events. To judge an increase in adverse event severity after dosing, the post-dosing adverse event severity will be compared to the adverse event severity reported prior to the first day of study drug administration.

3.5.2. Laboratory Data

Using the current Pfizer data standards, the last pre-dose value before the first day of dosing in study B5161002 is used as the baseline for all laboratory parameters. This baseline value is denoted as "overall baseline".

To evaluate the changes in lab parameters during the long term follow-up study B5161004, a second baseline will be defined as the last assessment prior to dosing on Day 1 in B5161004. This baseline value is denoted as "B5161004 baseline". The baseline value for B5161004 will be the screening visit which may in some cases, be the same data as the Week 97 visit for study B5161002.

3.5.3. Vitals and ECGs

Baseline values for vital signs and ECG will be the last acceptable pre-dose measurement taken either on the day of dosing or at the baseline or screening visit in study B5161002. This baseline value is denoted as "overall baseline".

To evaluate the changes in vital sign and ECG during the long term follow-up study B5161004, a second baseline will be defined as the last assessment prior to dosing on Day 1 in B5161004. This baseline value is denoted as "B5161004 baseline". The baseline value for B5161004 will be the screening visit which may in some cases, be the same data as the Week 97 visit for study B5161002 .

If not supplied, QTcF will be derived using Fridericia's heart rate correction formula: QTcF = QT / (RR)^(1/3), where RR = 60/HR (if RR is not provided).

3.5.4. Other Safety Endpoints

For all other safety endpoints including liver MRI and physical examinations (including nose and throat mucosal exam and Tanner stage and Testicular Volume), LVEF measured by cardiac MRI (or echocardiogram), bone mineral density by DXA, x-ray (hand and wrist for bone age evaluation) and C-SSRS, baseline will be the the last acceptable pre-dose measurement before the first day of dosing in study B5161002. This baseline value is denoted as "overall baseline".

To evaluate the changes in these safety endpoints during the long term follow-up study B5161004, a second baseline will be defined as the last assessment prior to dosing on Day 1 in B5161004. This baseline value is denoted as "B5161004 baseline". The baseline value for B5161004 will be the screening visit which may in some cases, be the same data as the Week 97 visit for study B5161002 .

4. ANALYSIS SETS

4.1. Full Analysis Set

The Full Analysis Set (FAS) is defined as all subjects who have received at least one dose of treatment in B5161004. It is the same as the Safety Analysis Set.

4.2. Per Protocol Analysis Set

The Per Protocol Analysis Set (PPAS) will be a subset of the FAS dataset. See Section 4.6 for the methods used to define this subset.

Exclusions from the PPAS will be reviewed prior to database release at study completion. Reasons for excluding subjects from the PPAS and subject numbers excluded will be documented and forwarded to programming.

4.3. Safety Analysis Set

The Safety Analysis Set is defined as all subjects who have received at least 1 dose of study medication in B5161004.

4.4. Other Analysis Sets

4.4.1. PK Concentration Analysis Set

The PK Concentration Analysis Set is defined as all subjects who have received at least 1 dose of study medication in B5161004 and have at least 1 PF-06252616 concentration measured.

4.5. Treatment Misallocations

If a subject was:

- Randomized but not treated, then they will be reported under their randomized treatment group for efficacy analyses. However, they are by definition excluded from the safety analyses as actual treatment is missing;
- Treated but not randomized, then by definition they will be excluded from the efficacy analyses since randomized treatment is missing, but will be reported under the treatment they actually received for all safety analyses;
- Randomized but took incorrect treatment, then they will be reported under their randomized treatment group for all efficacy analyses, but will be reported under the treatment they actually received for all safety analyses.

4.6. Protocol Deviations

The method used to identify the PPAS is defined in the table below. The clinical triad consisting of clinician, clinical pharmacologist and statistician, will review listings of source-verified data from the database prior to database release, to assess subject eligibility.

Following assessment of the data for all subjects, the project statistician will ensure that the list of subjects and the reason for exclusion is clearly documented. This list should be approved by the clinician and forwarded by the statistician to the clinical programmer for inclusion in the final analyses and Clinical Study Report (CSR) tables.

Table 2. Potential Exclusions from the PPAS

Reasons for Subject Exclusion	Responsible	Action/Source	Required Listings from Programming
Eligibility Criteria	Programming/Clinical	Programming review list of inclusion/exclusion criteria not met Clinical reviews before excluding subjects from population	Inclusion and Exclusion
Dosing Interruptions	Programming/Clinical	Programming checks for interruptions to the dosing schedule and sends to clinical for review	Dosing Information
Other Possible Exclusions	Clinical	Review protocol deviations and site reports for lack of protocol adherence	

Subjects may be excluded from the PPAS if they did not meet all eligibility criteria, experienced dosing interruptions, or other protocol deviations as supported by adequate justification and documentation. Subjects who are not in PPAS in study B5161002 will also be excluded from the PPAS in B5161004.

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

Statistical hypothesis testing against the historical control group will be performed in an exploratory manner. No statistical decision rules will apply.

5.2. General Methods

Continuous variables will be summarized by the N, mean, median, standard deviation, minimum and maximum. Categorical variables will be summarized by percent and counts. Efficacy data will be listed, tabulated and graphically represented, as appropriate.

Subjects were randomized in a 1:1:1 ratio to three treatment sequences in study B5161002:

Sequence 1:

Period 1: Active treatment (PF-06252616) within subject dose escalation (5, 20 and 40mg/kg).

Period 2: Active treatment (PF-06252616) at the maximum tolerated dose in Period 1.

Sequence 2:

Period 1: Active treatment (PF-06252616) within subject dose escalation (5, 20 and 40mg/kg).

Period 2: Placebo.

Sequence 3:

Period 1: Placebo.

Period 2: Active treatment (PF-06252616) within subject dose escalation (5, 20 and 40mg/kg).

In order to compare the effects of different treatment dosage, duration and sequence which subjects experienced during study B5161002 on the safety and efficacy endpoints in study B5161004, all summaries will be displayed for the entire population and by treatment sequences.

All plots produced may be used as in-text plots in the CSR and therefore need to be optimized for inclusion, ie, legends, axis labels and titles should be clearly legible when inserted into one half of a page within the CSR.

For the comparison with historical control group of change from baseline in functional assessments, a mixed-effect model for repeated measures(MMRM) will be utilized. In this analysis, the treated group will consist of B5161004 data combined with subjects who were randomized to Sequence 1 in B5161002, and will be compared to the historical control group (see Section 6.3.2 for the construction of the historical control group). Baseline, treatment, time and treatment-by-time interaction will be included as fixed effects in the model. The stratification factor of time to complete the 4 stair climb at baseline will be included in the model, unless the baseline time to climb 4 stairs is included in the model. Subject will be included as random effect and the model will be fitted with an unstructured covariance for the repeated measures. If there are model convergence issues when using an unstructured covariance, then compound symmetry and autoregressive covariance structures will be explored to potentially address convergence issues and model fit.

Transformations (including log transformation) will be evaluated to ensure the normality assumption is met. Contrasts of the treated group against the historical control group will be created to estimate the differences in change from baseline at the end of each long term follow up visit (Week 13, 25, 49, 73, 97, 121, 145, 169 and 193 in B5161004; refer to Section 6.2.1 for visit mapping when combining data with B5161002). Additionally, non Gaussian models, such as Wilcoxon rank- sum test or Wei- Lachin test, may be explored in the case of non normal data.

To explore the data further and to assess the goodness-of-fit of all statistical models, separate SAS datasets will be provided by the clinical programmer to the statistician. Covariates for data exploration may include baseline and/or demographic covariates such as gender, age, weight (continuous and/or Pfizer standard categorical cut-offs), site, country, steroid usage, 4SC at baseline, etc.

5.3. Methods to Manage Missing Data

In general, missing data on the efficacy endpoints will not be imputed. Time to complete a functional assessment will be transformed to velocity (the reciprocal of time to complete the functional test), so that subjects who lose the ability to complete a functional assessment and/or ambulate will be assumed to have a velocity of zero.

Missing data for the PODCI outcome will be handled according to the guidance of the developers of the PODCI questionnaire. For other questionnaires, missing data will also be handled according to the guidance of the developers of the questionnaire.

Values below the limit of quantification (BLQ) will be analyzed at the limit of quantification for that parameter except for PK data. For laboratory listings, the <BLQ will be used with the actual limit of quantification used in place of BLQ.

6. ANALYSES AND SUMMARIES

6.1. Primary Safety Endpoints

Safety analyses will be based on the Safety Analysis Set. See Section 6.5 for details.

6.2. Secondary Endpoints

6.2.1. Efficacy and PD Endpoints

Baseline values, change from baseline to last visit in PFTs, 4SC, NSAA, PUL, 6MWD, muscle strength by myometry and lean body mass by DXA will be described based on summary statistics including minimum, median, mean, maximum, and standard deviation. Change from baseline may also be assessed for any or all intermediate visits. The endpoints will be tabulated overall and by treatment sequence and study visit. For all the summaries in this section, two sets of tables will be generated, summarizing change from overall baseline and B5161004 baseline, respectively. For the summaries of changing from the overall baseline, B5161004 data will be combined with B5161002 data to show the long term change of the efficacy and PD endpoints. When summarizing change from the overall baseline, change from baseline in the historical control group will also be summarized together at each applicable study visit, if the endpoint is available in the historical control dataset. Mapping of B5161004 visits to a new visit that is continuous with B5161002 is provided in the table below:

B5161004 Visit Number	B5161004 Week Number	Mapped Week Number for combined B5161002 and B5161004
1	1	98
4	13	110
7	25	122
10	37	134
13	49	146
14	53	150
16	61	158
19	73	170
22	85	182
25	97	194
26	101	198
28	109	206
31	121	218
34	133	230
37	145	242
40	157	254
43	169	266
46	181	278
49	193	290

These summaries will be based on the FAS and PPAS. Additionally, a sensitivity analysis based on velocity may also be performed as described in Section 5.3 for timed function tests based on FAS.

Plots of mean change from baseline against nominal time will be generated, with one line representing one treatment sequence. Standard error of the mean will be presented together with the mean. A separate figure will be generated for each efficacy endpoint. Two sets of figures will be generated, presenting change from overall baseline and B5161004 baseline, respectively. For the plots of changing from the overall baseline, B5161004 data will be combined with B5161002 data to show the long term change of the efficacy and PD endpoints.

In the pre-specified subset of subjects who have B5161002 baseline 4SC <3.5 seconds, ≥3.5 seconds and ≤8 seconds, or >8 seconds, change from baseline will be summarized by treatment sequence. These summaries will be based on the FAS. Other exploratory analysis may be performed, such as summary of change from baseline in the efficacy and PD endpoints by subgroups defined by their B5161002 baseline function on FVC testing or on the 6MWT, or demographics (age, weight [continuous and/or Pfizer standard categorical cut-offs], site, country, steroid usage, 4SC at baseline, etc). Additionally, summaries based on velocity may also be performed as described in Section 5.3 for the efficacy endpoints based on FAS.

A summary plot of the cumulative percent of patients with a percent change from baseline on 4SC time less than a given set of thresholds (from <0% to <100%) will also be produced at each study visit.

6.2.2. PK Analyses

6.2.2.1. Concentrations Below the Limit of Quantification

In all data presentations (except listings), concentrations below the limit of quantification (BLQ) will be set to zero. (In listings BLQ values will be reported as "<LLQ", where LLQ will be replaced with the value for the lower limit of quantification).

6.2.2.2. Deviations, Missing Concentrations and Anomalous Values

In summary tables at each time point, statistics will be calculated having set concentrations to missing if 1 of the following cases is true:

1. A concentration has been collected as ND (ie not done) or NS (ie no sample);

2. A deviation in sampling time is of sufficient concern or a concentration has been flagged anomalous by the pharmacokineticist.

Note that summary statistics will not be presented at a particular time point if more than 50% of the data are missing.

Summary statistics will be provided for trough serum concentrations by dose and timepoint.

6.2.3. Immunogenicity

Immunogenicity analysis will be based on the Safety Analysis Set and will include anti-drug antibody and neutralizing antibody development from the start of B5161002 until the last visit in B5161004. Both continuous endpoints (titer) and categorical endpoints (ie, positive, negative and inconclusive) will be reported for the anti-drug antibody and neutralizing antibody assays overall and by treatment sequence and sampling time points. Data permitting, the incidence of anti-drug antibody and neutralizing antibody may be summarized by time points and treatment sequence. The impact of anti-drug antibody and neutralizing antibody on PK and PD parameters and profiles, efficacy and safety may be also evaluated but no statistical inference will be drawn.

6.3. Other Endpoint(s)

6.3.1. Clinical Outcomes

Baseline and change from baseline of the following health outcome endpoints will be summarized descriptively overall and by treatment sequence and time point as applicable.

- Two sets of tables will be provided for the PODCI parent-reported score (global function score and subcores for the five domains), summarizing change from overall baseline and B5161004 baseline, respectively. The PODCI Adolescent self-reported version will be completed in study B5161004 once the subjects are ≥11 years old.

The raw score of PODCI (both parent-reported and self-reported) will be summarized overall and by treatment sequence and study visit, and broken down by completers of the questionnaire within a specific study visit, if appropriate.

- EuroQoL 5 Dimension 3 Levels (EQ-5D-3L) Health Questionnaire.
- EuroQoL 5 Dimensions – Youth.
- Healthcare Resource Utilization (HRU) questionnaire.
- Zarit Burden Interview (ZBI).
- Work Productivity and Activity Impairment Questionnaire adapted for Caregiving (WPAI:CG) the percent work time missed, percent impairment while working, percent overall work impairment, and percent activity impairment due to a child's Duchenne muscular dystrophy.

6.3.2. Additional Analyses on Efficacy Endpoints

6.3.2.1. Historical Control

In keeping with global guidance (eg, Food and Drug Administration's Guidance from 2015: Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment and Eurpoean Medicinal Agency[EMA]/Committee for Medicinal Product for Human Use [CHMP] Guideline on the Clinical Investigation of Medicinal Products for the Treatment of Duchenne and Becker muscular dystrophy, 2016) and related advice received from regulators, the longterm effects on functional assessments following treatment with PF06252616 will be characterized compared to a historical control group. This historical control group will be established by filtering available natural history datasets (eg, CINRG dataset) to match covariates including but not limited to baseline 4SC time range, baseline age interval, and the baseline glucocorticosteroids requirements defined by the B5161002 protocol (here the baseline refers to the B5161002 baseline). To ensure appropriate comparison with the B5161002 and B5161004 study data, visits in the historical control group will be rounded to the nearest study visit as if the subjects are on the same visit schedule as in B5161002 or B5161004.

6.3.2.2. Statistical Analysis

The functional tests endpoints mentioned above, will be analyzed based on the FAS using the mixed effects model outlined in Section 5.2. These analyses will be performed based on the cumulative data from subjects in sequence 1 B5161002 and B5161004 compared to the historical control group. Covariates for this analysis will include the time to climb 4 stairs at the overall baseline, treatment, visit and treatment-by-visit interaction. Model estimates and change from historical control estimates, 95% two-sided confidence intervals, and p-values will be displayed for study interpretation purposes. Model estimates for each study visit available with confidence intervals will also be displayed graphically. Graphical summaries of the raw data, including boxplots, will be created to highlight outliers and other influential

observations. Analyses using PPAS may be explored to assess the sensitivity of analysis results.

A sensitivity analysis of the efficacy endpoints will be based on the velocity (defined as the reciprocal of time). This analysis will use the same MMRM model, but subjects who can no longer complete the assessment or ambulate will be analyzed with a velocity of 0 instead of a missing value at that time point. Transformation of the velocity data will be used to meet model assumptions, if the velocity data are not normally distributed based on plots of the residuals.

In the pre-specified subset of subjects who have B5161002 baseline 4SC <3.5 seconds, ≥3.5 seconds and ≤8 seconds, or >8 seconds, changes from baseline as compared to the historical control group may be analyzed using the same statistical methods as described in the previous paragraph, if data permits. Similarly, a sensitivity analysis based on velocity may be carried out as described above.

A time-to-event analysis will be performed for loss of ambulation. Loss of ambulation is defined as the inability to walk unassisted and without braces for at least 10m, as assessed and reported by the investigator at each study visit, and confirmed by the inability to walk/run 10m (as one component of the NSAA) evaluated at the next visit at which timed function tests are performed. The time-to-loss of ambulation will be calculated as (Date of investigator reported lost ambulation - Day 1 in study B5161002+1). If the subject is not able to walk independently (defined as an AE preferred term "Abasia") during B5161002, the date of loss of ambulation will be defined as the start date of abasia. Subjects who are still ambulatory at the end of the study/at early withdrawal visit will be censored at their last study visit. A Kaplan-Meier plot of time-to-loss of ambulation will be generated, with one curve for each treatment sequence. The time-to-loss of amublation curve for the historical control group will be presented on the same plot as well. Loss of ambulation is defined as daily wheelchair use for the historical control group. Time-to-loss of ambulation for the historical control group will be calculated as (age at loss of ambulation – age at baseline visit)*365.25.

The median time to loss of ambulation and 95%CI, obtained from the Kaplan-Meier curves, will be reported for each treatment sequence, as well as for the historical control group. If data permits, a log-rank test may be performed to treated group to the historical control group. The time to loss of ambulation analysis will be based on FAS.

6.4. Baseline and Other Summaries and Analyses

6.4.1. Baseline Summaries

A breakdown of demographic data will be provided for age, race, weight, body mass index, and height. Each will be summarized overall and by treatment sequence in accordance with the sponsor reporting standards.

6.4.2. Study Conduct and Subject Disposition

Subject evaluation groups will show end of study subject disposition and will show which subjects were analyzed for efficacy (FAS or PPAS), pharmacokinetics (PK), and safety (adverse events and laboratory data). Frequency counts will be supplied for subject discontinuation(s) overall and by treatment sequence. Data will be reported in accordance with the sponsor reporting standards.

6.4.3. Study Treatment Exposure

Study drug administration will be provided in a listing in accordance with the sponsor reporting standards. In addition, duration of study drug exposure will be summarized descriptively overall population and by treatment sequence.

6.4.4. Concomitant Medications and Non-Drug Treatments

All concomitant medication(s) as well as non-drug treatment(s) will be provided in the listings.

6.5. Safety Summaries and Analyses

6.5.1. Adverse Events

All adverse event summaries will be presented for the open label extension study B5161004 overall population and by treatment sequence according to the sponsor's reporting standard. Additionally, summary tables of the most frequent adverse events will be presented by 12-month dosing intervals. Based on the review of the frequency and timing of adverse events, additional summaries may be requested.

6.5.2. Laboratory Data

Laboratory data will be listed and tabulated overall population and by treatment sequence and study visit according to the sponsor's data reporting standard.

Two sets of tables will be provided, summarizing change from overall baseline and B5161004 baseline, respectively.

6.5.3. Vital Signs

Vital signs will be listed and tabulated overall and by treatment sequence and study visit with descriptive statistics. Change from baseline will also be summarized using the same descriptive statistics overall and by treatment sequence and study visit.

Two sets of tables will be provided, summarizing change from overall baseline and B5161004 baseline, respectively.

6.5.4. Electrocardiogram

The following ECG data will be listed: QT, QTc (Fridericia's), heart rate, QRS duration, PR and RR interval.

Baseline and change from baseline for QT, QTcF, heart rate, QRS, PR and RR will be summarized using descriptive statistics overall and by treatment sequence and study visit. The baseline for ECG parameters will be the average of the triplicate pre-dose measurements at Week 0 in B5161002. Any triplicate measurements will be averaged prior to the calculation of summary statistics. For QTcF a classification of absolute values and increase from baseline will be used. Two sets of tables will be provided, summarizing change from overall baseline and B5161004 baseline, respectively.

The number of subjects with average of the triplicate QTcF <450 ms, 450 ms ≤QTcF <480 ms, 480 ms ≤ QTcF <500 ms and QTcF values ≥500 ms will be tabulated overall and by treatment sequence and study visit. The number of subjects with maximum increase from baseline QTcF <30 ms, 30 ms ≤ QTcF <60 ms and QTcF ≥60 ms will be tabulated overall and by treatment sequence and study visit. Two sets of tables will be provided, summarizing maximum increase from overall baseline and B5161004 baseline, respectively. In addition, the number of subjects with uncorrected QT values ≥500 ms will be summarized.

6.5.5. Cardiac MRI/Echocardiogram

The mean absolute and percent change from baseline in LVEF will be summarized descriptively overall and by treatment sequence, imaging modality and study visit. LVEF shift tables will be prepared for overall population and by treatment sequence. Left ventricular wall thickness, ventricular wall strain and cardiac fibrosis will also be summarized where data from cardiac MRI are available.

Two sets of tables will be provided, summarizing change from overall baseline and B5161004 baseline, respectively.

6.5.6. Liver MRI

The mean absolute and percent change from baseline in the R2* value will be tabulated for each subject. Separate summaries by magnet field strength will be presented. Summaries of the categorical assessment of iron overload (normal, above normal, etc) will be presented.

Two sets of tables will be provided, summarizing change from overall baseline and B5161004 baseline overall and by treatment sequence, respectively.

6.5.7. Other Safety Data

Tanner stage, testicular volume, C-SSRs, ratio of bone age to chronologic age (hand and wrist x-ray), bone mineral density (whole body and spine DXA), prior medication(s), non-drug treatment(s), medical history and physical examination will be summarized and/or listed in accordance with the sponsor reporting standards but not subjected to formal statistical analysis.

Annual percent change in AP spine total L1 to L4 bone mineral density, as well as categorical summary of BMD changes exceeding 4% threshold will be summarized overall and by treatment sequence and study visit. Two sets of tables will be provided, summarizing change from overall baseline and B5161004 baseline, respectively.Any other screening data that is captured on the study database, will be listed.

7. INTERIM ANALYSES

There is no interim analysis planned for this study. The safety, PK, PD and efficacy data will be routinely reviewed by the study team for internal decisions regarding future study planning.

The E-DMC will periodically review safety data emerging from the ongoing study. The frequency and timing of safety reviews, as well as additional information concerning the composition and function of the E-DMC are contained in the E-DMC charter.