SANOFI GENZYME

STATISTICAL ANALYSIS PLAN

A phase 1/2, multi-center, open-label, ascending dose study to evaluate the safety, tolerability, pharmacokinetics, pharmacodynamics and exploratory efficacy of olipudase alfa in pediatric patients aged <18 years with acid sphingomyelinase deficiency

GZ402665-DFI13803

STUDY BIOSTATISTICIAN: [redacted]

DATE OF ISSUE: 05-Nov-2019

NCT02292654

Total number of pages: 51

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

ABAS-3:	Adaptive Behavior Assessment System, Third Edition
ACE:	angiotensin-converting enzyme
ADA:	anti-drug antibody
AE:	adverse event
AESI:	adverse events of special interest
ALT:	alanine aminotransferase
ANCOVA:	analysis of covariance
AP:	alkaline phosphatase
APR:	acute phase reaction
ASM:	acid sphingomyelinase
ASMD:	acid sphingomyelinase deficiency
AST:	aspartate aminotransferase
AUC:	area under the plasma concentration curve
$AUC_{0\text{-}inf}$:	area under the concentration versus time curve from time 0 textraoplated to infinity
BMI:	body mass index
BUN:	blood urea nitrogen
CCL18:	chemokine (CC-motif) ligand 18
C_{eio}:	concentration at end of infusion
CHITO:	chitotriosidase
CL:	total body clearance
C_{max}:	maximum plasma concentration
CRS:	cytokine release syndrome
CSR:	clinical study report
CV:	coefficient of variation
DBP:	diastolic blood pressure
DBS:	dried blood spots
DL_{CO}:	diffusing capacity of the lung for carbon monoxide
DP-3:	Developmental Profile 3
ECG:	electrocardiogram
ECHO:	echocardiogram
e-CRF:	electronic case report form
FEV_1:	forced expiratory volume in the 1^{st} second of the FVC maneuver
FVC:	forced vital capacity
Hb:	hemoglobin
HDL:	high density lipoprotein
HIV:	human immunodeficiency virus
HR:	heart rate
HRCT:	high resolution computed tomography
hsCRP:	high sensitivity C-reactive protein
IAR:	infusion-associated reaction

IgE:	immunoglobulin E
IgG:	immunoglobulin G
IL:	interleukin
INR:	international normalized ratio
ITT:	Intent-to treat
IV:	intravenous
LDL:	low density lipoprotein
LLOQ:	Lower Limit of Quantification
MedDRA:	Medical Dictionary for Regulatory Activities
mITT:	Modified intent-to-treat
MN:	multiples of normal
MRI:	magnetic resonance imaging
PCSA:	potentially clinically significant abnormality
PD:	pharmacodynamic(s)
PedsQL:	Pediatric Quality of Life
PFT:	pulmonary function testing
PK:	pharmacokinetc(s)
PKDM:	pharmacokinetics, dynamics and metabolism
Protime:	prothrombin time
PT:	preferred term
PTT:	partial thromboplastin time
QOL:	Quality of Life
QTc:	corrected QT-interval
QTcB:	QTc interval corrected by Bazett
QTcF:	QTc interval corrected by Fridericia
RBC:	red blood cell
rhASM:	recombinant human acid sphingomyelinase
SAE:	serious adverse event
SAP:	Statistical Analysis Plan
SAS:	Statistical Analysis System
SBP:	systolic blood pressure
SD:	standard deviation
SEM:	standard error of the mean
SI:	international system of units, international system of units
SMPD1:	acid sphingomyelinase gene, acid sphingomyelinase gene
SOC:	system organ class
$t_{1/2z}$:	terminal half-life
TEAE:	treatment-emergent adverse event
TLC:	total lung capacity
t_{max}:	time to reach Cmax, time to reach Cmax
ULN:	upper limit of normal
VAS:	visual analogue scale
VLDL:	very low density lipoprotein
V_{SS}:	volume of distribution at steady state.
WBC:	white blood cell
WHO:	World Health Organization

WHO-DRUG:	World Health Organization Drug dictionary
WMW:	Wilcoxon-Mann-Whitney
β-HCG:	beta-human chorionic gonadotropin

1 OVERVIEW AND INVESTIGATIONAL PLAN

This statistical analysis plan (SAP) provides a comprehensive and detailed description of strategy and statistical techniques to be used to analyze data from DFI13803, a study of GZ402665/olipudase alfa, recombinant human acid sphingomyelinase [rhASM] in pediatric patients; it is based on Amended Clinical Trial Protocol 4, dated 21-Aug-2017. The purpose of the SAP is to ensure the credibility of the study findings by pre-specifying the statistical approaches to the analysis of study data prior to database lock. This version replaces the previous versions of the SAP.

1.1 STUDY DESIGN AND RANDOMIZATION

This is a Phase 1/2, multi-center, open-label, ascending dose study to evaluate the safety, tolerability, pharmacokinetics (PK), pharmacodynamics (PD) and exploratory efficacy of olipudase alfa in pediatric patients less than 18 years of age with non-neuronopathic acid sphingomyelinase deficiency (ASMD). At least 12 patients were initially planned to be enrolled in 3 age cohorts with a minimum of 3 patients aged from 12 to <18 years (adolescent cohort), a minimum of 3 patients aged from 6 to <12 years (child cohort), and at least 2 patients aged from birth to <6 years (infant/early child cohort). At least 8 additional pediatric patients <12 years will be enrolled following protocol amendment 04 (dated 21-Aug-2017). There will be at least 20 patients enrolled in the study in total.

For each patient, study participation will consist of a screening/baseline period (up to 60 days) and a 64-week treatment period with intravenous (IV) infusions of olipudase alfa administered every 2 weeks for 64 weeks. Patients will go through a dose escalation period with target dose of 3.0mg/kg; for further details see protocol Section 1.1.

Eligible patients may enroll in a long-term study (LTS13632) following completion of the 64-week treatment period to continue receiving olipudase alfa. Patients not enrolling in the long-term study will undergo a final safety follow-up assessment by telephone 30 to 37 days after the last infusion in this study; patients with a treatment gap between the end of the 64-week treatment period of DFI13803 and enrollment in the long-term study, depending on the length of the gap, may also undergo a final safety follow-up assessment by telephone 30 to 37 days after the last infusion in this study.

1.2 OBJECTIVES

1.2.1 Primary objectives

The primary objective of this study is to evaluate the safety and tolerability of olipudase alfa administered intravenously in pediatric patients every 2 weeks for 64 weeks.

1.2.2 Secondary objectives

The secondary objectives of this study are to characterize the pharmacokinetic profile and evaluate the pharmacodynamics and exploratory efficacy of olipudase alfa administered intravenously in pediatric patients every 2 weeks for up to 64 weeks.

1.3 DETERMINATION OF SAMPLE SIZE

A sample size power calculation was not performed for this study. The study sample size is based upon empirical considerations.

1.4 STUDY PLAN

Refer to Protocol Section 1.2 for the study plan.

1.5 MODIFICATIONS FROM THE STATISTICAL SECTION OF THE PROTOCOL

In the section of statistical considerations of the protocol, the baseline for white blood cell count, platelet count, and hemoglobin was defined as the mean of 2 samples collected greater than 24 hours apart during the screening period. The baseline definition is modified in the SAP as the average of all measurements taken before the first infusion.

The statistical considerations of the protocol mentions that Infusion Associated Reactions (IAR), Acute phase reaction (APR) and cytokine release syndrome (CRS) will be listed as part of adverse events of special interest (AESI). However, per protocol, AESIs are reported as original signs and symptoms and only IARs were specifically identified in the eCRF. Consequently, APR and CRS cannot specifically be listed, but the details of these AESIs will be provided in patient narratives, and discussed in the clinical study report (CSR) as necessary.

In the section of statistical considerations of the protocol, the post-treatment period was defined as the time after the EOS visit. However, no data are expected to be collected after the EOS visit. Therefore, the post-treatment period has been removed from this SAP.

1.6 STATISTICAL MODIFICATIONS MADE IN THE STATISTICAL ANALYSIS PLAN

The main modifications made in this version of SAP compared to the previously approved version (Version 2, Dated 17-Aug-2016) are listed as follows.

- The modified intent-to-treat population was added in Section 2.2.2 to be consistent with the protocol.
- The per protocol population and completer population were removed from Section 2.2 as they are not pre-defined in the protocol and also are not needed for this study.

- Derivations for classifications of the different SMPD1, CHIT1 and UGT1A1 genotypes were added in Section 2.3.1 in order to provide summaries on number of patients and percentages.
- Protocol-defined potential dose limiting toxicities that are considered AESIs have been added in Section 2.6.1.3 to be consistent with the protocol.
- Specific analyses on infections and bleeding have been updated in Section 2.6.1.4 following interactions with health authorities.
- [redacted]
- The paired t-test has been updated to analysis of covariance with baseline included as covariate for exploratory efficacy variables in order to adjust for baseline, since baseline is highly correlated with change from baseline values. See Section 2.8.
- Platelet count and hemoglobin were added as exploratory efficacy endpoints in order to be consistent with other ASMD studies for compound GZ402665/olipudase alfa. See Section 2.8.4.

2 STATISTICAL AND ANALYTICAL PROCEDURE

Data collected in this study will be presented using summary tables, figures, and by-patient data listings. Descriptive statistics for categorical variables will be provided using frequencies (n) and percentages (%). Descriptive statistics for continuous variables will be provided using number of observations (n), mean, standard deviation (SD), median, minimum, and maximum.
If appropriate, p-values and 95% confidence intervals (CI) from analysis of covariance with baseline as covariate may be provided for change from baseline values. All data used in tabular summaries or figures will be supported by listings of underlying data – raw and/or derived, and all figures will be supported by relevant summary information.

2.1 SUBJECT DESCRIPTION

2.1.1 Disposition of subjects

Screened patients are defined as any patients who signed the informed consent or whose parent or guardian signed informed consent.

Enrolled patients are defined as those who meet the inclusion criteria and none of the exclusion criteria.

The number of patients screened, enrolled, treated (receiving at least one study infusion), completing the study, and not completing study along with primary reason for discontinuation will be presented.

2.1.2 Protocol deviations

During the review of the database, compliance with the protocol will be examined with regard to inclusion and exclusion criteria, treatment, prohibited therapies, and timing and availability of planned assessments. Protocol deviations will be identified by the study team before database lock and classified as minor, major or critical deviations. A listing of major or critical protocol deviations for all treated patients will be produced.

In addition to the above, the two conditions listed below will be evaluated:

- Noncompliance: Missing 2 or more consecutive infusions (excluding for cause missed doses, such as when the missed infusions are required by the protocol due to a dose limiting toxicity) or a total of 4 or more infusions for the 64-week treatment period. Whether a missed infusion is for-cause or not will be reviewed and determined by the clinical team; the decision will be used in the determination of noncompliance.
- Prohibited medications intake: A list of medications which potentially modify the effect of olipudase alfa are identified and provided by the clinical team. This list is updated regularly with change of World Health Organization (WHO) Drug dictionary (WHO-DRUG) versions. Any use of such medications in the clinical database will be reviewed by

the clinical team before data base lock with special attention to start and/or stop date of these medications compared to start/stop date of study treatment. The clinical team will review the prohibited medication intake and decide whether it is minor, major or critical protocol deviation.

2.2 ANALYSIS POPULATIONS

2.2.1 Safety population

All subjects who were exposed to the study treatment, olipudase alfa, ie, who received at least one infusion of study treatment (total or partial), will be included in the safety population.

2.2.2 Modified intent-to-treat population

The modified intent-to-treat (mITT) population is the same as the safety population and is used as the primary population for the efficacy analysis.

2.2.3 Pharmacokinetic population

The pharmacokinetic population includes all patients who receive at least 1 infusion of olipudase alfa and have evaluable pharmacokinetic data available post-baseline.

2.2.4 Pharmacodynamic population

Pharmacodynamic population includes all patients who are treated and have at least one evaluable pharmacodynamics measurement available post-baseline.

2.3 DEMOGRAPHIC AND BASELINE CHARACTERISTICS

2.3.1 Subject demographic characteristics, medical history, and diagnoses

Demographic, disease characteristics and medical/surgical history information will be summarized by age group cohorts and overall using the safety population. No formal hypothesis testing will be performed to compare differences between age group cohorts.

Demographic and baseline characteristic parameters include the following: gender, race, ethnicity, ancestry, age (in years and months) at Week 0/Day 1, human immunodeficiency virus (HIV) antibody testing, and Hepatitis tests (B surface antigen and C antibody tests).

Baseline disease characteristic parameters include the following: age at ASMD symptom onset (years and months if available), age at ASMD diagnosis (years and months if available), spleen status, ASMD symptoms present at disease onset, ASMD family history, genotyping of acid sphingomyelinase gene (SMPD1), CHIT1, UGT1A1, severe splenomegaly and severe % predicted DLco adjusted for hemoglobin. Genotype of acid sphingomyelinase gene (SMPD1), CHIT1 and UGT1A1 as well as acid sphingomyelinase (ASM) activity in peripheral leukocytes and CRIM

testing from screening evaluations will be included in data listings. Severe splenomegaly is defined as the baseline spleen volume (MN) > 15 MN; severe % predicted DLco adjusted for hemoglobin is defined as the baseline value < 40%.

The SMPD1 variant type (i.e., missense, frameshift, deletion, nonsense, etc.) will be derived, if possible, from the SMPD1 amino-acid change.

For the classification of CHIT1 genotype, the following derivation will be applied:

- N= Wildtype (WT)/WT = Normal / 2 functional alleles
- H= WT/Duplication 24 (DUP24) = Heterozygous mutation /1 functional allele
- M = DUP24/DUP24 = homozygous mutation / 2 non-functional alleles (no chito activity)

For the classification of UGT1A1 genotype, the following derivation will be applied:

- UGT1A1*1 = Fully functional allele = Not Gilbert variant
- UGT1A1*36 = Functional allele = Not Gilbert variant
- UGT1A1*28 = Gilbert syndrome allele = Gilbert variant
- UGT1A1*37 = Crigler-Najar syndrome type 2 allele (impaired UGT1A1) = Gilbert variant

Patients with 0 or 1 Gilbert variant alleles will be classified as “no Gilbert Syndrome”; patients with 2 Gilbert variants alleles will be classified as “Gilbert Syndrome”.

Medical/surgical history will be collected by body system as Yes/No/Unknown, and if Yes, the current status will be collected. These data will be summarized by body system, by age group cohorts and overall. Additional medical/surgical history regarding specific disease of interest will be collected and summarized.

Smoking and alcohol drinking history data will be provided in data listings.

2.3.2 Baseline safety parameters

Baseline for safety parameters (weight, height, body mass index (BMI), physical exam, vital signs, echocardiograms, immune response assessments, and safety biomarkers) in general is defined as the last available non-missing value for the parameter before the first infusion. Most of the baseline values will thus be Day 1/Week 0 pre-infusion assessments with the exception of hematology parameters and possibly electrocardiogram (ECG). Hematology samples are collected twice at screening and twice on first infusion visit, once before and once after the first infusion; the average of all measurements taken before the first infusion will serve as the baseline. Another exception may be the ECG parameters – Day 1/Week 0 (the first infusion visit) is to collect ECG in triplicate before the infusion; if 2 or more than 2 of these triplicate collections of ECG are before the first infusion, the average of these ECG parameter results will be considered as the baseline value; otherwise, the baseline value of ECG parameters will be taken from the latest observation before the first infusion.

Baseline safety data will be presented along with subsequent values assessed during or after olipudase alfa dosing using safety population.

2.3.3 Baseline pharmacodynamics parameters

Baseline values for sphingomyelin and sphingomyelin metabolites, including, but not limited to, ceramide in dried blood spots (DBS), lyso-sphingomyelin and sphingosine-1-phosphate in plasma and/or DBS will be the last available value before first infusion; typically these will be the screening values.

Baseline pharmacodynamic data will be presented along with subsequent pharmacodynamics values assessed during or after olipudase alfa dosing using pharmacodynamic population.

2.4 EXTENT OF STUDY TREATMENT EXPOSURE AND COMPLIANCE

The extent of study treatment exposure will be assessed and will be summarized using safety population. The duration of exposure is defined as the number of weeks between first infusion and disposition date (ie, the end of study date) in the study: (disposition date of the study – first infusion date + 1)/7. The number of infusions received and the duration of exposure will be summarized by age cohorts and overall.

The distribution of weeks on study treatment (<12 weeks, >=12 weeks to <26 weeks, >=26 weeks to <38 weeks, >=38 weeks to <52 weeks, >=52 weeks to <64 weeks, and >=64 weeks) will be provided using number and percentage of patients by age group cohorts and overall.

Compliance to the treatment regimen will be monitored in terms of the percentage of scheduled infusions the patient receives during the treatment period. A patient is considered noncompliant if the patient missed 2 or more consecutive infusions (not-for-cause, such as when the missed infusions are required by the protocol due to a dose limiting toxicity) or a total of 4 or more infusions for the 64-week treatment period. Whether a missed infusion is for-cause or not will be reviewed and determined by the clinical team; the decision will be used in the analysis of compliance.

Overall Compliance = 100% x Total number of infusions received / Total number of expected infusions

where total number of expected infusions refers to the number of infusions expected between the first infusion and the date of withdrawal or date of completion. It is calculated as the integer part of (Date of withdrawal – Date of the first infusion)/14. For a patient who completes the 64-week treatment period, the total number of expected infusions is 33.

Compliance excluding for-cause = 100% x Total number of infusions received / Total number of expected infusions (excluding for-cause missing infusions)

where total number of not-for-cause expected infusions refers to the number of infusions expected between the first infusion and the date of withdrawal or date of completion, excluding the

for-cause missed dose. It is calculated as the integer part of [(Date of withdrawal – Date of the first infusion)/14] – number of for-cause missed infusions.

Overall compliance and compliance excluding for-cause will be summarized descriptively as quantitative variable by age group cohorts and overall using safety population. The numbers and percentages of patients whose compliance are <80%, >=80% to <90%, >=90% to <100% and >=100% will be provided. Number and percentage of patients in the following two categories will be presented:

- Missed 2 or more not-for-cause consecutive infusions.
- Missed 4 or more scheduled infusions.

2.5 PRIOR AND CONCOMITANT MEDICATION/THERAPY

Medications and therapeutic procedures received by the patient in the 30 days prior to their first infusion until the final visit will be recorded on the Prior and Concomitant Medication the electronic case report form (e-CRF). Medications, including over-the-counter and herbal products as well as supplements, and therapeutic procedures will be collected. Prior and concomitant medications will be coded using the World Health Organization Drug dictionary version dated March of 2019.

- Prior medications will be defined as medications that are taken prior to the first infusion of olipudase alfa in this study.
- Concomitant medications will be defined as medications that are taken on or after the first infusion of olipudase alfa in this study.

A medication that was started before the first infusion but continued beyond the first infusion will be categorized as both prior and concomitant. In case the start/stop date of the medication is missing or incomplete preventing the unambiguous determination of the medication status, the medication will be categorized as both prior and concomitant.

Medication summaries will be presented by WHO-DRUG Anatomic Class and by WHO-DRUG Therapeutic Class for each age group cohort and overall using safety population. Tables will show all medications used during the study, sorted by most prevalent Anatomic Class and then by descending frequency of Therapeutic Class within Anatomic Class based on the prevalence within the overall column.

Prohibited medications include those that are taken during the study and may decrease olipudase alfa activity (eg, fluoxetine, chlorpromazine; tricyclic antidepressants [eg, imipramine, or desipramine]). The medications list with potential prohibited medications will be provided by the data management and will be merged with the concomitant medication data in the database to identify any use of it. Whether the use of such medication constitute a protocol deviation that interferes with assessment of effect of olipudase alfa will be determined by clinical review before the database lock of all prohibited medication intake during the study.

2.6 ANALYSIS OF SAFETY DATA

Safety endpoints include data pertaining to the safety and tolerability of olipudase alfa: assessment of adverse events (AEs)/treatment-emergent adverse events (TEAEs), including AESIs (eg, infusion-associated reactions (IARs), symptomatic overdose, etc.); physical examinations; neurological examinations; clinical laboratory evaluations; vital sign measurements; ECGs; Doppler echocardiography; liver ultrasound Doppler; safety biomarkers; and immune response assessments. For laboratory endpoints, the summaries will use both the international system of units (SI) and US conventional units; figures will use the international system of units (SI); data listings will provide the results in original units as well as in standard units, where applicable.

All the safety analyses will be performed using the safety population. Patients will be summarized and displayed by age group cohorts and overall. Safety data will be reported up to and including the time of treatment withdrawal, and follow-up for AEs, where applicable. The baseline value is the last available value before the first infusion of olipudase alfa, except for white blood cell count, platelet count, hemoglobin and possibly ECG parameters. For complete definition of baseline, refer to Section 2.11.1.

For all safety data, the observation period will be divided into 2 segments:

- The **pretreatment period** is defined as the time between when the patient gives informed consent and the start of the first infusion (excluded).
- The **on-treatment period** is defined as the time from the start of the first infusion (included) till the end of the study.

The following definitions regarding potentially clinically significant abnormality (PCSA) will be applied to laboratory parameters, vital signs and ECG:

The PCSA values are defined as abnormal values considered medically important by the sponsor according to predefined criteria/thresholds based on literature review and defined by the sponsor for clinical laboratory tests, vital signs, and ECG. These values are provided in Section 2.11.5.

- PCSA criteria will determine which patients had at least 1 PCSA taking into account all evaluations performed, including unscheduled or repeated evaluations.
- Before determining a value as PCSA, the value will be rounded to the precision that the PCSA criterion is defined.

2.6.1 Adverse events

2.6.1.1 Definitions

AEs are defined in protocol Section 10.6.1. AEs will be coded according to the version of Medical Dictionary for Regulatory Activities (MedDRA) in use at the time of database lock. Adverse event summaries will include the number of events (N), number (n) and percentage (%) of patients experiencing an adverse event. The denominator for computation of percentages is the safety population for the entire study and for each age group cohort. Sorting order will follow the

internationally agreed MedDRA system organ class (SOC) order, then by decreasing number of patients overall in preferred terms (PTs) within SOCs, and further alphabetically within the same number of patients over different PTs. When more than one PT has same number of events within the same SOC, the order of presentation will be alphabetical in PTs.

The AEs will be classified into different periods of occurrence as follows:

- Pre-treatment adverse events: AEs that started before the start of the first infusion.
- TEAEs: AEs that started during the on-treatment period.

If an AE date/time of onset is incomplete, an imputation algorithm will be used to classify the adverse event within the appropriate adverse event period. The algorithm for imputing date/time of onset will be conservative and will classify an adverse event as treatment emergent unless there is definitive information to determine otherwise. Details on classification of AEs with missing or partial onset dates are provided in Section 2.11.2. If an AE worsens in severity, only the worst severity is captured on the AE log page in the e-CRF.

The AE log page in the e-CRF captures the date when the event become serious. If and when this date is non-missing, it indicates that the AE started as non-serious and became serious after few days. To account for non-serious and serious parts of the AE collected in one AE log page, this AE will be split into two AEs – one non-serious, and one serious AE – with the same AE number, with the following modifications in data for these two AEs:

- Non-serious AE: AE end date, AE outcome, date of death or hospitalization, autopsy information, criterion for serious adverse event (SAE) will be blank.
- Serious AE: AE start date will use the date the event became serious.

Action taken towards study drug and action taken (toward AE) will be considered in relation to the date the event became serious, the last infusion date and dosing information, and a determination will be made, in consultation with safety monitor of the trial, as to whether the recorded fields belong to non-serious AE or serious AE; then the decision will be implemented in analysis. Any such decision will be rationalized in the CSR.

The primary focus of adverse event reporting will be on treatment-emergent adverse events. Pre-treatment adverse events will be listed separately.

Start day of a pre-treatment AE = Start date of AE – first infusion date.

Start day of treatment emergent AE relative to first infusion = Start date of AE – first infusion date + 1.

Start day of treatment emergent AE relative to last infusion = Start date of AE – last infusion date + 1.

2.6.1.2 Treatment-emergent adverse events

An overview of TEAEs will be provided by age group cohorts and overall to include (potentially related is defined as either 'related' or 'possibly related' as assessed by the investigator):

- TEAEs.
- TEAEs potentially related to study treatment.
- TEAEs by the worst severity.
- TEAEs leading to treatment discontinuation.
- TEAEs leading to study withdrawal.
- TEAEs leading to dose reduction.
- TEAEs leading to study treatment interruption.
- Protocol-defined IARs.
- Algorithm-defined IARs.
- Treatment-emergent serious adverse events.
- Treatment-emergent serious adverse events potentially related to study treatment.
- TEAEs with fatal outcome (Deaths).
- Treatment-emergent pregnancies.
- TEAEs which are considered symptomatic overdose.
- Dose-limiting toxicities (collected in the adverse event log page in e-CRF, see protocol Section 8.3.4.2.1 for details).

The following summaries will be produced by primary SOC and PT, by age group cohorts and overall:

- TEAEs.
- TEAEs by the worst severity (ie, mild, moderate, or severe).
- TEAEs potentially related to study treatment.

The most common TEAEs are defined as those with percentage of events >= 2% and number of patients >=2 in all age cohorts combined. The most common TEAEs by SOC and PT will be provided in descending order of overall incidence by age group cohorts and overall.

TEAEs by the latest infusion dose immediately preceding the start of AE will be provided by SOC and PT, by age group cohorts and overall.

2.6.1.3 Adverse events of special interest

The following are AEs of special interest:

Protocol-defined IAR: Any AE that is considered related or possibly related to study treatment as judged by the investigator or the sponsor is defined as a protocol-defined IAR. Sponsor identification will reconcile with investigator identification of IAR before the database lock.

The following summaries will be produced by primary SOC and PT, by age group cohorts and overall:

- Protocol-defined IARs.
- Protocol-defined IARs by the worst severity.
- Protocol-defined IARs started during infusion.
- Protocol-defined IARs started between 0 hours (exclusive) and 3 hours (exclusive) post end of infusion.
- Protocol-defined IARs started between 3 hours (inclusive) and 24 hours (inclusive) post end of infusion.
- Protocol-defined IARs started between 24 hours (exclusive) and 72 hours (inclusive) post end of infusion.
- Protocol-defined IARs started after 72 hours post end of infusion.

Algorithm-defined IAR: All AEs that start between the start of infusion and the end of infusion plus 24 hours are defined as algorithm-defined IARs, irrespective of the perceived relation with study treatment.

The following summaries will be produced by primary SOC and PT, by age group cohorts and overall:

- Algorithm-defined IARs (including any TEAEs with missing time but happened on the same day or the day after the infusion)

A subset of protocol-defined IARs will also be defined programmatically by searching for preferred terms within Hypersensitivity SMQs (identified using SMQ code 20000214 Broad and Narrow scope). The summaries of Hypersensitivity IARs will be provided by primary SOC and PT.

Pregnancy: Pregnancy of a female patient entered in the study as well as pregnancy occurring in female partner of a male patient enrolled in this study is captured as an adverse event. In addition, e-CRF provides urine pregnancy test results as positive or negative. Labs provide serum pregnancy test results. Pregnancies will be provided in the overview of AE summary and the pregnancy test results will be provided as part of patient narrative.

Symptomatic overdose: This is captured as an adverse event in this study. For details, please see Section 10.6.4.1 in the protocol. The symptomatic overdose data will be provided as part of patient narrative.

Asymptomatic overdose: This is captured as an adverse event with verbatim term as 'Asymptomatic overdose'. Asymptomatic overdose data will be provided as part of patient narrative.

Laboratory values: The following protocol-defined potential dose limiting toxicities are also considered AESIs:

- Any increase in AST, ALT, total bilirubin, or AP >3x baseline (before olipudase alfa therapy) and > ULN;
- Any increase in total bilirubin or AP >1.5x baseline, in the presence of AST or ALT > 2x ULN;
- Any increase in ALT or AST >3x ULN combined with an increase in ALT or AST >2x baseline (prior to olipudase alfa therapy), with symptoms of fatigue, nausea, vomiting, fever, rash, or eosinophilia (> ULN).

2.6.1.4 Adverse events of other interest

Infections: TEAEs of infections are defined under the system organ class of Infections and Infestations. These TEAEs will be included in the summary of all TEAEs by primary SOC and PT.

Bleeding: TEAEs of bleeding are defined under the standard MedDRA queries (SMQs) of Hemorrhages (identified using SMQ code 20000038 Broad and Narrow scope). Bleeding AE data will be summarized by SOC and PT for each age cohort and overall.

2.6.1.5 Serious adverse events

The following summaries will be provided by primary SOC and PT, by age group cohorts and overall:

- Treatment-emergent serious adverse events.

2.6.1.6 Adverse events leading to permanent treatment discontinuation

In this study, if the treatment is discontinued permanently, the patient is withdrawn from the study. A temporary dose interruption is not considered permanent discontinuation of study treatment. AEs leading to permanent treatment discontinuation will be provided as part of the patient narrative.

2.6.1.7 Deaths

Death is reported as a fatal outcome on the AE log page in the e-CRF. All deaths reported in the trial will be provided as part of the patient narrative.

2.6.1.8 Potential Quantifiable Dose Limiting Toxicities

Potential quantifiable Dose Limiting Toxicities will be provided in a by patient data listing.

2.6.1.9 [redacted]

[redacted]

2.6.2 Physical examinations

Complete physical examinations data will be collected at screening, Day 1/Week 0, Week 12, 26, 38 and 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal). These examinations will include the evaluation of normal versus abnormal in following body systems: general appearance, examination of the skin; head, eyes, ears, nose, and throat; lymph nodes; heart, lungs, and abdomen; extremities and joints. Complete physical examination (except overall neurological exam) will be classified as normal or abnormal. Frequency shift tables will be provided.

An abbreviated physical examination will assess only the general appearance with evaluation of normal versus abnormal of the patient and is only required prior to each infusion when a complete exam is not indicated. A listing of the abbreviated physical examination data will be provided.

Tanner stage for physical gender-related development will be assessed at Screening, Week 12, Week 26, Week 38 and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal). A data listing will be provided.

2.6.3 Neurological examinations

Neurological examinations assessments will be collected at screening, Day 1/Week 0, Week 12, 26, 38, and 52 (or at withdrawal if withdrawn prematurely). They will include, but are not limited to the patient's mental status, posture, cranial nerves, motor system including muscle atrophy, tone and power, reflexes, sensory system, coordination, gait, and age-appropriate responses.

Neurological examinations will be classified as normal/abnormal findings. Frequency shift table will be provided.

2.6.4 Weight, height, BMI

The weight, height (standing height will be used for patients >2 years of age and supine height will be used for patients <=2 years of age) and BMI will be collected at Screening and within

24 hours prior to every infusion visit from Day 1/Week 0 to Week 64 (or withdrawal if withdrawn prematurely).

BMI = weight in kilograms / (height in meters)^2

For these three parameters, change from baseline, in addition to observed values, will be calculated and summarized for all scheduled visits by age group cohorts and overall.

Age-appropriate growth in height is a matter of efficacy; please see Section 2.8.2 for analysis of height z-scores.

2.6.5 Clinical laboratory evaluations

Clinical laboratory data consists of blood analysis, including hematology, clinical chemistry, and urinalysis. Analysis of all safety laboratory parameters will be conducted by a central laboraty.

The following clinical laboratory evaluations will be assessed by the site as specified in Section 1.2 to 1.4 of the study protocol. The laboratory parameters will be classified as follows:

- Hematology: Complete blood count including white blood cell (WBC, leukocytes) count, platelet count, hemoglobin, hematocrit, red blood cell (RBC , erythrocytes) count, WBC differentials (absolute and percentage of lymphocytes, monocytes, neutrophils, basophils, eosinophils, and (if applicable) abnormal cells). Platelet count and hemoglobin will also be analyzed as exploratory efficacy endpoints. See details in Section 2.8.4.
- Clinical Chemistry: sodium, potassium, chloride, blood urea nitrogen (BUN), glucose, uric acid, calcium, phosphorus, magnesium, albumin, total protein, creatinine.
- Liver function tests: aspartate aminotransferase (AST), alanine aminotransferase (ALT), total bilirubin, direct bilirubin and alkaline phosphatase (AP).
- Coagulation: prothrombin time (Protime), partial thromboplastin time (PTT), international normalized ratio (INR) and D-dimer.
- Urinalysis: dipstick for pH, ketones, glucose, bilirubin, protein and blood.
- If female of child bearing potential, β-human chorionic gonadotropin (β-HCG) serum or urine test as appropriate per Protocol Sections 1.2 and 10.3.

Observed values, change from baseline and percent change from baseline to scheduled study time points will be summarized by age group cohorts and overall. By patient plots of hemoglobin, platelets, WBC, ALT, AST, total bilirubin will be provided. Summary plots over time will also be provided as needed.

Laboratory values will be classified as normal, above normal, or below normal based on normal ranges when provided by the laboratory. Shifts from baseline to scheduled time points, any time during the on-treatment period and the last available value, in the abnormality of hematology, liver function tests and clinical chemistry will be provided by age group cohorts and overall.

PCSA values for hematology, liver function tests and clinical chemistry (provided in Section 2.11.5) during the study (from first infusion start to last available observation) will be summarized by age group cohorts and overall. PCSA values for hematology, liver function test and clinical chemistry parameters will be provided in separate listings.

2.6.6 Vital signs

Vital signs are to be measured at every visit before, halfway through, and after infusion (±10 minutes). Post-infusion vital signs are measured every 30 minutes for the first 2 hours, hourly for 2 to 6 hours post-infusion, and then once every 8 hours until the patient is discharged (±10 minutes). Refer to Section 2.11.1 for the baseline definition.

Vital signs will include heart rate (HR, beats/minute), systolic and diastolic blood pressure (SBP/DBP, mmHg), and body temperature (°F or °C). Each vital sign parameter will be summarized as observed parameter value and change from baseline per-infusion, 3 hours post-infusion, 24 hours post-infusion, 48 hours post-infusion, 72 hours post-infusion and end of infusion at each scheduled visit by age group cohorts and overall.

PCSA values during the study (from first infusion start to last available observation) will be summarized by age group cohorts and overall.

2.6.7 Electrocardiogram

Standard 12-lead ECGs will be recorded at the visits and time points specified in Protocol Sections 1.2 and 1.3, and prior to any blood collections. For Day 1/Week 0, triplicate ECGs will be recorded within 5 minutes with at least 1 minute between 2 replicates prior to the study treatment administration.

Heart rate (in beats per minute, PR (in milliseconds), QRS (in milliseconds), QT-intervals (in milliseconds) and corrected QT-intervals (QTc) will be analyzed as observed parameter values and change from baseline (for HR, PR, QT, QTc). The corrected QT-interval (QTc) data will be analyzed using the Bazett (QTcB) and Fredericia (QTcF) methods.

For all parameters, the observed data and change from baseline will be summarized for each time point by age group cohorts and overall.

PCSA values will be summarized during the study (from first infusion start to last available observation) by age group cohorts and overall.

2.6.8 Doppler echocardiogram

A standard 2-dimensional and M-mode Echocardiogram (ECHO) with Doppler will be conducted at Screening, with the first infusion at 3.0 mg/kg or highest tolerated dose during Week 14 to Week 24 or whenever dose escalation is reached, and at Week 52. Examination will include but not be limited to: left ventricular cavity size, left ventricular mass, valve characterization, ejection fraction, ventricular wall thickness, regional wall motion, systolic and diastolic functions,

pericardium characterization and congenital abnormalities. Pulmonary artery pressure will be determined by Doppler ultrasound.

Shift tables will be created for valve characterization (including aortic regurgitation, mitral regurgitation, pulmonic regurgitation, tricuspid regurgitation). Summary of change from baseline in ejection fraction (%) and pulmonary artery systolic pressure will be provided. By-patient plots of left ventricular calculated ejection fraction (%) will be provided.

2.6.9 Liver ultrasound Doppler

Liver ultrasound Doppler will be performed at Screening and Week 52 to document hepatic blood flow characteristics, principally portal vein pressure and blood flow direction. The structures to be examined include hepatic portal vein, the main hepatic artery and the main hepatic vein. Additional structures that may be examined include the network of intrahepatic portal veins, the main and intrahepatic arteries, the hepatic veins, the main and intrahepatic portal veins, the intrahepatic portion of the inferior vena cava, collateral venous pathways, and transjugular intrahepatic portosystemic shunts. Liver ultrasound Doppler will be performed using methods that are compatible with the standard institutional procedures of the investigational site.

Liver ultrasound Doppler data will be provided in listings.

2.6.10 Safety biomarkers

Biomarkers for monitoring the safety of olipudase alfa include, but are not limited to, high sensitivity C reactive protein (hsCRP), plasma ceramide, iron, ferritin, cardiac-specific troponin-1, calcitonin, interleukin (IL)-6 and IL-8. Samples will be collected at the visits specified in Protocol Sections 1.2 and 1.3.

Observed values, change from baseline and percent change from baseline will be summarized by age group cohorts and overall at each time point. Summary plots over time will also be presented for observed values as appropriate.

2.6.11 Immune response assessments

An immune reaction against an exogenously administered recombinant protein plays a critical role in the safety of such compounds. Therefore, safety assessments will also include blood samples for anti-olipudase alfa immunoglobulin G (IgG) antibodies and neutralizing antibody formation in patients testing positive for ADA when appropriate.

In the event that a patient experiences a moderate, severe, or recurrent IAR suggestive of hypersensitivity reactions, additional blood samples will be collected for the evaluation of:

- Circulating immune complex detection; and,
- Immunoglobulin E (IgE) ADA, serum tryptase, and complement activation.

Additionally, skin testing may be performed, as appropriate, in patients who experience an IAR that meets the following criteria:

- IAR is suggestive of IgE ADA – mediated hypersensitivity reaction, with persistent symptoms of bronchospasm, hypotension and/or urticaria requiring intervention,
- Or any other signs or symptoms at the discretion of the investigator or the sponsor.

For ADA summaries, the baseline is the last non-missing ADA assessment prior to the first IMP.

Data listings will be provided to display results of anti-olipudase IgG antibody, neutralizing antibody, anti-olipudase alfa IgE antibody, serum tryptase activity, complement activation, and skin testing performed. If IgG ADA titer value is reported as <xx, the numeric value will be imputed as xx, the minimum required dilution (MRD), and used for summary purposes. Listing will provide results as reported as well as results as imputed.

A patient with at least one anti-drug antibody (ADA) response available in the database is considered evaluable.

A patient whose ADA status is positive at baseline is considered to have pre-existing ADA.

A patient whose ADA status is positive anytime post-baseline and is negative or missing at baseline is considered to have treatment-induced ADA.

A patient whose ADA status is positive at baseline (pre-existing ADA) and the ADA titer level anytime post-baseline is significantly higher than that at baseline is considered to have treatment-boosted ADA. A difference in titer values between 2 samples representing greater than or equal to twice the dilution level is considered significant. In other words, if the post-baseline titer value divide by the baseline titer value is greater than or equal to 4, then the patient is considered to have treatment boosted ADA. For example, if baseline titer value is 50, the post-baseline titer value has to be >=200.

For patients with treatment-induced ADA, the following 3 ADA responses are defined:

Transient ADA response:

- Treatment-induced ADA detected only at one sampling time point post-baseline (excluding the last sampling time point).
- Treatment-induced ADA detected at two or more sampling time points post-baseline, where the first and last ADA-positive samples (irrespective of any negative samples in between) are separated by a period less than 16 weeks, and the subject's last sampling time point is ADA-negative.

Persistent ADA response:

- Treatment-induced ADA detected at two or more sampling time points post-baseline, where the first and last ADA-positive on-treatment sample (irrespective of any negative samples in between) are separated by at least 16 weeks.
- Treatment-induced ADA detected in the last two sampling time points, irrespective of the time period in between.

The following sub-classifications for persistent ADA response will be considered as well.

- Low response – if a patient's peak titer ≤400 and positive at final assessment. This represents the first titer that is greater than a 4-fold increase from the assay minimum required dilution (MRD). Titers within this range would be considered as Low response.
- Intermediate ADA response - if a patient had persistent ADA response but peak titer is >400 and positive at final assessment.
- Tolerized – if a patient had persistent ADA response, but negative at the final assessment.

Indeterminate ADA response:

- Only the last sampling time point is positive and all previous samples are negative.

Pre-existing ADA and ADA Incidence rate are defined as follows:

Pre-existing ADA = 100*(Number of patients pre-existing ADA at baseline)/number of evaluable patients.

ADA Incidence = 100*(treatment-boosted + treatment-induced ADA positive patients)/number of evaluable patients.

The following information will be provided by age group cohorts and overall:

- Number of ADA evaluable patients
- Number (%) of patients with ADA positive/negative at baseline
- Number (%) of patients that never develop ADA at any time (i.e, ADA always negative during baseline and post-baseline)
- Number (%) of patients with treatment-boosted ADA
- Number (%) of patients with treatment-induced ADA
- Treatment emergent ADA: Number (%) of patients with treatment-boosted/treatment-induced ADA
- Duration of ADA response: Number (%) of patients with transient/persistent/indeterminate ADA response (each category separately)
- Number (%) of patients within persistent ADA response subclassifications: low response, intermediate response, tolerized (each category separately)
- Patients with treatment induced ADA will be further characterized
  - Peak ADA titer: Median and 25th/75th quantiles (min, max) of the ADA titer for peak ADA titer of treatment induced ADA positive patients
  - Last ADA titer (final assessment): Median and 25th/75th quantiles (min, max) of the ADA titer of treatment induced ADA positive patients
- Number (%) of patients with positive neutralizing antibody (NAb)

- NAb inhibition of catalytic activity
- NAb inhibition of cellular uptake

[redacted]

2.7 ANALYSIS OF PHARMACOKINETIC DATA

All the pharmacokinetic analyses will be performed using the pharmacokinetic population. Pharmacokinetic sampling will occur with the first infusion at 0.3, 1.0 and 3.0 mg/kg olipudase alfa and at Week 52. Specific time points are detailed in protocol Section 1.5 for each age cohort. To minimize the amount of blood drawn, post-dose time points have been selected for sparse sampling of pharmacokinetic parameters and microsampling methods (volumes ≤0.5 mL) will be used. If a patient receives only a partial infusion (eg, due to a safety concern), pharmacokinetic sampling will be repeated the next time the patient receives that dose.

2.7.1 Pharmacokinetic parameters

Plasma concentration-time data will be analyzed by noncompartmental methods, or nonlinear mixed effects modeling, based upon patient age and data suitability. Pharmacokinetic parameters if calculated using noncompartmental methods from plasma olipudase alfa concentrations may include, but are not limited to, C_{eio}, C_{max}, t_{max}, AUC_{last}, $AUC_{0-\tau}$, $t_{1/2z}$, CL and V_{SS}. Pharmacokinetic parameters if derived from population PK model may include C_{max}, t_{max}, $AUC_{0-\tau}$, CL and V.

2.7.2 Statistical analysis

Plasma concentration-time data will be analyzed by non-compartmental methods, or nonlinear mixed effects modeling, based upon patient age and data suitability. Values will be reported for individual patient and summarized by age group cohorts, dose level, [redacted] and study week as appropriate.

For ease of presentation, mean values will be arithmetic mean unless specified. Concentration values below the plasma assay limit will be treated as zero in calculating mean values. Mean values below Lower Limit of Quantification (LLOQ) will be reported as LLOQ in the tables and not plotted in the figures if after C_{max}. Mean calculations and their associated statistics will be generated from unrounded numbers and may differ slightly from those values that would be determined using the rounded numbers displayed in the tables. Values expressed in all tables will be for ease of presentation and will not be meant to imply accuracy to more than 3 significant figures.

Pharmacokinetic parameters of olipudase alfa may be summarized using descriptive statistics such as arithmetic mean, geometric mean, SD, standard error of the mean (SEM), coefficient of variation (CV [%]), minimum, median, maximum, and number of observations for each study

visit, dose level and age group cohort under the responsibility of Pharmacokinetics, Dynamics and Metabolism (PKDM) using Phoenix WinNonlin Professional, Version 5.2 or above.

Associations between IgG ADA variables and PK of olipudase alfa may be explored. If applicable, plot of olipudase alfa PK (e.g., AUC) versus PK visit may be provided by ADA classifications.

2.8 ANALYSIS OF EXPLORATORY EFFICACY VARIABLES

All exploratory efficacy analyses to assess changes in organomegaly (spleen and liver), infiltrative lung disease, linear patient growth, pulmonary function, bone age, cycle ergometry endpoints, physician's global assessment, efficacy biomarkers, lipid profile, bone biomarkers, health outcome questionnaires, cognitive functioning and adaptive functioning will be performed using the mITT population. For each exploratory efficacy assessment, observed measures and changes from baseline, as appropriate, will be listed and summarized, including the p-value from analysis of covariance with baseline as covariate and 95% CIs as appropriate, separately by time point, age group cohorts and overall. [redacted]

Baseline for exploratory efficacy variables will be defined as the last non-missing value before the first infusion.

2.8.1 Spleen and liver volumes

Spleen and liver volumes will be measured by abdominal magnetic resonance imaging (MRI) to quantitate the degree of splenomegaly and hepatomegaly at Screening, Week 26 and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal). MRIs will be collected and read centrally by a third party blinded to patient number and study visit.

Spleen volume >=5 multiples of normal (MN) is an inclusion criterion. For screening purpose, the external vendor will provide the spleen volume in the unit of MN to the site. The computation will use the same formula as below. However, the report from the external vendor will be part of the source document of the site and will not be included in the clinical data base. Spleen and liver volumes collected in cm^3 unit (including those from the screening visit) will be converted to multiples of normal (MN) unit for analyses using the following formula:

Spleen Volume (MN) = Spleen Volume (cm^3) / [2*Weight (kg)].

Liver Volume (MN) = Liver Volume (cm^3) / [25*Weight (kg)].

Change from baseline and percent change from baseline for both cm^3 and MN units, in addition to observed values, will be calculated and summarized for all scheduled visits by age group cohorts and overall. The analysis of covariance (ANCOVA) with baseline as the covariate and Wilcoxon-Mann-Whitney (WMW) test p-values for the change from baseline and percent change from baseline will be provided in the summary tables for spleen/liver volumes in MN unit. Box plots of observed values over time for spleen and liver volumes in MN unit will be provided.

Summary plots for the mean spleen/liver volumes (MN) over time will be provided. By-patient plots of the observed spleen/liver volumes (MN) over time will be provided.

Additional sensitivity analyses to assess the effect of age group cohorts: Analysis of covariance (ANCOVA) will be used for the mITT population for spleen/liver volumes in MN unit with the following model:

Percent change from baseline at Week 26 or 52 volume (MN) = Age group cohorts + Baseline value

The observed values, change from baseline and percent change from baseline on spleen volume (MN) will also be summarized by the following disease severity:

- Spleen volume by baseline spleen volume severity (severe vs not severe). Severity is defined as baseline spleen volume > 15 MN.

In addition, the summary of spleen/liver volumes (MN) over time will be provided to exclude patients who satisfy the noncompliance condition in Section 2.1.2 as a sensitivity analysis if applicable.

2.8.2 Height z-score

As stated in Section 2.6.4, patient height will be collected at screening and within 24 hours prior to every infusion visit from Day 1/Week 0 to Week 64 (or withdrawal if withdrawn prematurely). Z-score for height will be evaluated as an exploratory efficacy parameter. This parameter reflects failure to thrive.

Height z-score will be calculated and categorized using the number and percentage of patients in <= -4, > -4 to <= -3.5, > -3.5 to <= -3, > -3 to <= -2.5, > -2.5 to <= -2, > -2 to <= -1.5, > -1.5 to <= -1 and > -1. The categories will be summarized for baseline, Week 12, Week 26, Week 38 and Week 52 by age group cohorts and overall. Change from baseline at Week 52 (as improved, same, and worsened) will also be summarized using the number and percentage of patients. A bar chart for the percentage of patients in each category at baseline and Week 52 will be provided. By-patient plot on height z-score over time will be provided.

Observed values and change from baseline values for height z-score will be summarized at baseline, Week 12, Week 26, Week 38 and Week 52 by age group cohorts and overall. In addition, change from baseline in height z-score will be analyzed with analysis of covariance (ANCOVA) model with baseline height z-score as covariate.

2.8.3 Pulmonary function testing

Pulmonary function testing (PFT) will be performed at Screening, Week 26 and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal) in patients ≥5 years old at Day 1/Week 0. PFT will include, but may not be limited to the assessment of forced vital capacity (FVC), forced expiratory volume in the 1st second of the FVC maneuver (FEV_1), total lung capacity (TLC), and diffusing capacity of carbon monoxide (DL_{CO}). Completion of the assessment may depend upon patient cooperation. The tests are done

in local labs and hence do not use a standardized approach for calculation of percent predicted values. For analyses, we use the following standard formula (height and hemoglobin values are from available ones closest to the PFT evaluation date/time, age is calculated at the PFT evaluation date, and sex=0 for female and 1 for male) to calculate the percent predicted values. The percent predicted values captured in the e-CRF will remain in the database, but will not be used for analysis.

Percent predicted values of FEV1 and FVC will use the Global Lung Initiative (GLI) 2012 reference equations (1).

Predicted TLC (Liters) will use the following formula (Koopman et al. corrected version [2]):

Predicted TLC = exp(24.67 - 4.91*ln(Height[cm]) – 7.16*ln(Age[months]) + 0.059*Sex + 1.47*ln[Height]*ln[Age])

Percent predicted TLC = 100*Observed TLC / Predicted TLC

Predicted DLco will be calculated in the unit of mmol/min/kPa as follows (Koopman et al. corrected version [2]):

Predicted DLco (mmol/min/kPa) = exp(34.80 - 6.89*ln(Height[cm]) – 8.66*ln(Age[months]) + 0.10*Sex + 1.79*ln[Height]*ln[Age])

Unit conversion (Macintyre et al. [3]): mL CO/min/mmHg = mmol/min/kPa*2.987

Hemoglobin-adjustment factor (Macintyre et al. [3]) (to be applied to DLco in unit of mL CO/min/mmHg)

= 1.7*Hemoglobin (g/dL) / (10.22+Hemoglobin) for patients >=15 years of age.

= 1.7*Hemoglobin (g/dL) / (9.38+Hemoglobin) for patients <15 years of age.

Patients' age at screening will be used to decide which formula to use for the hemoglobin-adjusted factor.

Adjusted DLco = Observed DLco (mL CO/min/mmHg) / Hemoglobin-adjusted factor

Percent predicted Hemoglobin-adjusted DLco is calculated as follows:

Percent predicted Hemoglobin-adjusted DLco = 100 * Adjusted DLco / Predicted DLco in unit of mL CO/min/mmHg.

For all 4 tests, FVC (L), FEV_1 (L), TLC (L) and DL_{CO} (mL/min), percent (%) predicted values will be summarized for observed values, change from baseline and percent change from baseline using the mITT population by age group cohorts and overall at scheduled study visit. The analysis of covariance (ANCOVA) with baseline as covariate and WMW test p-values for the change from baseline and percent change from baseline will be provided in the summary tables for all 4 tests. Box plots of observed values over time for these 4 tests will be provided for the mITT population. By-patient plots for all 4 tests over time will be provided.

Additional sensitivity analyses to assess the effect of age group cohorts: ANCOVA will be used for the mITT population for the 4 tests with the following model:

Percent change from baseline at Week 26 or 52 = Age group cohorts + Baseline value.

The summary of percent predicted DLco adjusted for hemoglobin over time will be provided to exclude patients who satisfy the noncompliance condition in Section 2.1.2 as a sensitivity analysis if applicable. In addition, sensitivity analysis to account for missing data of percent predicted DLco adjusted for hemoglobin at baseline or post-baseline visits may be conducted after appropriate imputations.

2.8.4 Platelet count and hemoglobin

Observed values, change from baseline and percent change from baseline for platelet count and hemoglobin at pre-infusion time point will be summarized by age group cohorts and overall. The ANCOVA with baseline as covariate will be used to assess the change from baseline. Additionally, the ANCOVA with age group cohorts and baseline value as covariates will be used to assess the age effect.

2.8.5 Pulmonary imaging by high resolution computed tomography

Pulmonary imaging of the chest using high resolution computed tomography (HRCT) will be assessed at Screening and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal) to quantitate the degree of possible infiltrative lung disease.

The bilateral lung boundaries are determined from 4 pre-defined levels of HRCT images: Level 1 includes the aortic arch, Level 2 includes the carina, Level 3 is midway between the carina and 1 cm above the hemidiaphragm, and Level 4 is 1 cm above the hemidiaphragm. A qualitative assessment will be made for the following features: interstitial lung disease (reticular pattern), ground glass appearance, reticulo-nodular density (nodular densities) and pleural thickening of the right and left lungs based on the following criteria: 0 = No lung disease, 1 = Mild (affecting 1 to 25% of the lung volume), 2 = Moderate (affecting 26 to 50% of the lung volume), 3 = Severe (affecting 51 to 100% of the lung volume), and NE = Not evaluable.

For all 4 features, the mean and standard deviation will be provided for observed values and change from baseline by each side at each level. The mean scores at each scheduled visit are calculated as follows:

Mean score across 4 levels for left (or right) lung = Sum of scores for 4 levels/4.

Mean score across 4 levels and both lungs = (Mean score across 4 levels for left lung + Mean score across 4 levels for right lung)/2.

Summary of the above two types of mean scores and change from baseline for each feature will be provided by age group cohorts and overall. A box plot for the mean scores across 4 levels and both lungs of each feature will be provided. The mean pixel density will be provided in a data listing.

2.8.6 Chest X-ray

A chest X-ray (posterior-anterior and lateral) will be performed at Screening and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal) at selected sites. The lung features (interstitial (overall ILD)), reticular, nodular, consolidation and pleura) will be scored subjectively for the degree of diffuse lung disease (infiltrative lung disease): 0 = No infiltrative lung disease, 1 = Mild (affecting 1 to 25% of the lung volume), 2 = Moderate (affecting 26 to 50% of the lung volume), or 3 = Severe (affecting 51 to 100% of the lung volume).

For chest x-ray imaging, interstitial provides an overall assessment of interstitial lung disease with reticular and nodular may provide more supportive information on type of interstitial involvement. Consolidation provides a measure of air space disease and is assessed in addition to interstitial involvement. As HRCT provides a more robust characterization of the diffuse lung disease, only interstitial (overall ILD) of the chest X-ray will be analyzed; however data for all features will be provided in data listings.

Summary of mean scores and change from baseline for interstitial will be provided by age group cohorts and overall.

2.8.7 Bone age by hand X-ray

Patients will have a hand X-ray performed on their left hand, fingers and wrist at Screening and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal). Bone age will be provided by the central vendor – [redacted]; they will use calculations based on the Greulich & Pyle Atlas (4). A listing of these data will be provided.

For each patient at each scheduled baseline and post-baseline visits, the difference between the bone age and actual age at that visit will be calculated. Then the change from baseline in the bone age and actual age difference will be analyzed with analysis of covariance (ANCOVA) model with baseline bone age and actual age difference as covariate.

2.8.8 Cycle ergometry

Cycle ergometry will be performed at Screening and within 1 week after infusion at Week 26 and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal). This assessment is not required in patients that are ≤6 years of age or <120 cm in height on Day 1/Week 0. Completion of the assessment may depend upon patient cooperation.

Cycle ergometry assessments will include maximum workload (watts), percent predicted maximum workload (%), working time (min), maximum heart rate (breaths/min), maximum percent predicted heart rate (%), maximum O_2 saturation (%), maximum respiratory rate (breaths/min), maximum ventilation (L/min), maximum O_2 uptake (mL/min), maximum percent predicted O_2 uptake (%), maximum CO_2 output (mL/min), and maximum respiratory exchange ratio.

Assessments will be summarized at each time point along with change from baseline by age group cohorts and overall.

2.8.9 Physician's global assessment of change

The physician's global assessment of the patient's progress will be evaluated prior to infusion at Screening, Week 26 and Week 52 (or at withdrawal if withdrawn prematurely), as mild, moderate and severe. At Week 26 and Week 52 the investigator will also evaluate the patient's current clinical status compared with screening (baseline) by marking 1 of 7 categories: "marked improvement of daily activities", "moderate improvement of daily activities", "mild improvement of daily activities", "no change", "mild worsening of daily activities", "moderate worsening of daily activities" or "marked worsening of daily activities".

Frequency shift table for physician's global assessment of patient progress (mild, moderate, severe) from the baseline to Week 26 and Week 52 will be provided by age group cohorts and overall. Frequencies and percentages of patients at each clinical status at Week 26 and Week 52 will be provided by age group cohorts and overall.

7 categories of Physician's global assessment of patient's current clinical status compared with screening (baseline) will be converted to numeric values as follows:

- "Marked improvement of daily activities" = 3.
- "Moderate improvement of daily activities" = 2.
- "Mild improvement of daily activities" = 1.
- "No change" = 0.
- "Mild worsening of daily activities" = -1.
- "Moderate worsening of daily activities" = -2.
- "Marked worsening of daily activities" = -3.

Observed numeric values at each post-baseline visits will be analyzed using analysis of covariance (ANCOVA) model with baseline as covariate. The baseline value will be Physician's global assessment of severity at screening with value of 1=mild, 2=moderate, 3=severe.

2.8.10 Efficacy biomarkers

Efficacy biomarkers of chitotriosidase activity ([CHITO]nmol/hr/mL), chemokine (CC-motif) ligand 18 (CCL18) (ug/L), and angiotensin-converting enzyme (ACE) (uKat/L) will be evaluated at Day 1/Week 0, Week 26, Week 52, and Week 64 (or at withdrawal if withdrawn prematurely). Observed values, change from baseline and percent change from baseline will be summarized by age group cohorts and overall at each time point. Summary plots of observed values and percent change from baseline over time for each efficacy biomarker will be provided.

Chitotriosidase values will be normalized prior to study analyses based on the patient's chitotriosidase genotype:

Chitotriosidase Genotype	Normalization of Chitotriosidase Activity
Normal	None (reported value was used)
Heterozygous	Value multiplied by 2
Homozygous mutation	Value set to “no value”

2.8.11 Lipid profile

A lipid profile will be assessed at Screening, Day 1/Week 0 (not required in patients <=2 years of age), Week 12 (not required in patients <=2 years of age), Week 26, Week 38, Week 52 and Week 64 (or at withdrawal if withdrawn prematurely). Blood will be collected prior to infusion for lipids including total cholesterol (mmol/L), low density lipoprotein (LDL)(mmol/L), high density lipoprotein (HDL)(mmol/L), very low density lipoprotein (VLDL)(mmol/L), triglycerides (mmol/L), apolipoprotein B (g/L), apolipoprotein A1 (g/L), and lipoprotein-[a] (g/L).

Results will be analyzed and presented for observed values and change from baseline values by age group cohorts and overall at each scheduled time point. Summary plots of observed values over time of lipid profile parameters will be provided.

2.8.12 Bone biomarkers

Bone biomarkers of serum bone-specific alkaline phosphatase (ug/L) and C-telopeptide (ug/L) will be evaluated at Day 1/Week 0, Week 26, Week 52 and Week 64 (or at withdrawal if withdrawn prematurely) for patients >2 years of age.

Observed values and change from baseline will be summarized by age group cohorts and overall at each scheduled time point. Summary plots of observed values over time for each bone biomarkers will be provided.

2.8.13 Health outcome questionnaires

Health outcome measures were evaluated using questionnaires designed to measure general quality of life (QOL) with the Pediatric Quality of Life (PedsQL) scale, fatigue with the PedsQL Multidimensional Fatigue Scale, and pain with the PedsQL Pediatric Pain Questionnaire. These will be evaluated at Screening, Week 26 and Week 52 (or at withdrawal if withdrawn prematurely

and not done within 12 weeks prior to withdrawal), and will be analyzed in accordance to the tool manuals.

PedsQL Generic Core Scales

The PedsQL includes a child self-report for patients age 5 to 18 years and a report for parents of patients age birth to 18 years. For patients age 5 to 18 years, the 23-item PedsQL Generic Core Scales for both patient and parent reports include four Scales, Physical Functioning (8 items), Emotional Functioning (5 items), Social Functioning (5 items), and School Functioning (5 items). For patients age 2 to 4 years, the 21-item PedsQL Generic Core Scales for the parent reports include 8 items in Physical Functioning Scale, 5 items in Emotional Functioning Scale, 5 items in Social Functioning Scale, and 3 items in School Functioning Scale.

The PedsQL Infant Scales include two age-appropriate versions for ages 1-12 months and 13-24 months, and assess parents' perceptions of their infant's generic Health-Related Quality of Life. The 36-item PedsQL Infant Scales 1-12 months Version encompasses 5 scales: (1) Physical Functioning (6 items), (2) Physical Symptoms (10 items), (3) Emotional Functioning (12 items), (4) Social Functioning (4 items), and (5) Cognitive Functioning (4 items). The 45-item PedsQL Infant Scales 13-24 months Version contains the same 5 scales and the same 36 items as the 1-12 months Version with 9 additional age-appropriate items: (1) Physical Functioning (9 items), (2) Physical Symptoms (10 items), (3) Emotional Functioning (12 items), (4) Social Functioning (5 items), and (5) Cognitive Functioning (9 items). The format, instructions, Likert response scale, and scoring method for the PedsQL Infant Scales are identical to the PedsQL Generic Core Scales for ages 2-18 years.

Each item in every Scale in both parent report and patient report in the above age ranges uses the 4-point rating scale (0=never to 4=almost always). The items of the four Scales are grouped together on the actual questionnaire. The algorithm of Scale Scores (5) is as follows.

1. On the PedsQL Generic Core Scales, for ease of interpretability, items are reversed scored and linearly transformed to a 0-100 scale, so that higher scores indicate better Health-Related Quality of Life.
2. To reverse score, transform the 0-4 scale items to 0-100 as follows:

   0=100, 1=75, 2=50, 3=25, 4=0.
3. To create Scale Scores, if more than 50% of the items in the scale are missing, the Scale Score will not be computed. If 50% or more of the items in the scale are completed, compute the Scale Score using the sum of non-missing items scores divided by the number of items answered.
4. The PedsQL Generic Core Scales for ages 2-18 years: to create the Psychosocial Health Summary Score, the mean is computed as the sum of the items scores over the number of items answered in the Emotional, Social, and School Functioning Scales; the Physical Health Summary Score is the same as the Physical Functioning Scale Score.
5. The PedsQL Infant Scales 1-12 months and 13-24 months: the Physical Health Summary Score is computed as the sum of the items scores over the number of items answered in the Physical Functioning and Physical Symptoms Scales; the Psychosocial Health Summary

Score is computed as the sum of the items scores over the number of items answered in the Emotional, Social, and Cognitive Functioning Scales.

6. To create the Total Scale Score, the mean is computed as the sum of all the items scores over the number of items answered on all the Scales.

Response Choices	Never	Almost Never	Sometimes	Often	Almost Always
Raw Scores	0	1	2	3	4
0-100 Scale Scores	100	75	50	25	0

Scale Scores, Psychosocial Health Summary Score, Physical Health Summary Score and the Total Scale Score will be calculated based on the 0-100 scale scores. Their computed values and changes from baseline will be summarized by age group cohorts at each scheduled time point.

PedsQL Multidimensional Fatigue Scale

The PedsQL Multidimensional Fatigue Scale is an 18-item instrument encompassing three scales: (1) General Fatigue (6 items), (2) Sleep/Rest Fatigue (6 items), and (3) Cognitive Fatigue (6 items). It includes a child self-report for patients ages 5-7, 8-12 and 13-18 years and a report for parents of patients ages 2-4, 5-7, 8-12 and 13-18 years.

For parent reports of patients ages 2-4, 5-7, 8-12 and 13-18 years and child self-reports of patients ages 2-4, 8-12 and 13-18, a 5-point Likert-type response scale is utilized (0 = never a problem; 1 = almost never a problem; 2 = sometimes a problem; 3 = often a problem; 4 = almost always a problem). The child self-reports of patients ages 5 to 7 years was simplified to a 3-point scale to increase the ease of use (0 = not at all a problem; 2 = sometimes a problem; 4 = a lot of a problem).

The format, instructions, response scale, and scoring method are identical to the PedsQL Generic Core Scales (6). Items in all child self-reports and parent reports are reverse-scored and linearly transformed to a 0-100 scale (0 = 100, 1 = 75, 2 = 50, 3 = 25, 4 = 0), so that higher scores indicate better Health-Related Quality of Life. Scale Scores are computed as the sum of item scores divided by the number of answered items. If more than 50% of the items in the scale are missing, the Scale Score is not computed. If less than 50% of the items in the scale are missing, use the same method for the PedsQL Generic Core Scales to impute the Scale Score. The Total Scale Score is computed as the sum of all the item scores divided by the number of items answered on all the Scales.

Scale Scores and the Total Scale Score will be calculated based on the 0-100 scale scores. Their computed values and changes from baseline will be summarized by age group cohorts at each scheduled time point.

PedsQL Pediatric Pain Questionnaire

The PedsQL Pediatric Pain Questionnaire consists of 3 items and includes a child self-report for patients age 5 to 18 years and a proxy report for parents of patients age 5 to 18 years.

The first item is the degree of present pain, the second item is the degree of the worst pain, and the third item refers to the localization of pain and is not scored. The first two items will be scaled using visual analogue scales (VAS) from 0 (Not hurting / No discomfort / No pain) to 100 mm (Hurting a whole lot / Very uncomfortable / Severe pain). The present pain and worst pain are scored separately. The scores are based on line length to the nearest 5mm. There is no imputation process since these are single item VAS scales.

For the first two items, the observed scores based on the 100 mm VAS and the change from baseline will be summarized by age group cohorts at each scheduled time point.

2.8.14 Cognitive and adaptive function testing

Developmental Profile 3 (DP-3):

Cognitive function will be evaluated using the Developmental Profile 3 (DP-3) (7) assessment tool at Screening, Week 26 and Week 52 (or at withdrawal if withdrawn prematurely and not done within 12 weeks prior to withdrawal) for patients <6 years of age at Day 1/Week 0. The number of items to be answered on each content area of the DP-3 depends on the patient's age and functional level. For each content area (Physical, Adaptive Behavior, Social-Emotional, Cognitive, Communication), the raw score indicates the number of developmental items the subject is credited for being able to complete.

The DP-3 assessments are centrally scored to generate the General Development Score. For each content area and the General Development Score, standard scores, confidence intervals, descriptive categories, and percentile ranks are generated. For each content score, age-equivalence is generated.

Adaptive Behavior Assessment System, Third Edition (ABAS-3):

Adaptive function will be evaluated at the same time points as cognitive function for patients <6 years of age at Day 1/Week 0 using the Adaptive Behavior Assessment System, Third Edition (ABAS-3) (8). The ABAS-3 consists of several age-specific rating forms; for this study, the Parent/Primary Caregiver Form (Ages 0-5) and Parent Form (Ages 5-21) will be used. The ABAS-3 measures adaptive behavior at three different levels. At the highest level is the General Adaptive Composite, which is composed of all measured skill areas. The next level are three adaptive domains, each comprising multiple individual skill areas; these domains are Conceptual, Social, and Practical. At the third level are individual adaptive skill areas. For each adaptive skill area, the raw score is the sum of item behavior frequency ratings for all items.

The ABAS-3 assessments are centrally scored to generate scaled scores for each adaptive skill area. The scaled scores are utilized to generate a standard score, percentile rank, and confidence

interval for the General Adaptive Composite and three adaptive domains (Conceptual, Social, and Practical).

Analysis:

A listing of DP-3 and ABAS-3 raw and generated scores will be provided.

For the DP-3, by-patient bar graphs will be generated to demonstrate subject's performance over time using percentile rank and standard score for each content area (Physical, Adaptive Behavior, Social-Emotional, Cognitive, Communication) and General Development Score at the 3 study time points (baseline, Week 26, and Week 52).

For the ABAS-3, by-patient bar graphs will be generated to demonstrate subject's performance over time using percentile rank and standard scores for the General Adaptive Composite and three adaptive domains (Conceptual, Social, and Practical) at the 3 study time points (baseline, Week 26, and Week 52).

In addition, to illustrate patient's performance on the DP-3 and ABAS-3 collectively, group graphs by patient age will be developed showing tool standard scores of patients over time. One graph for the ABAS-3 General Adaptive Composite Standard Score and one graph for the DP-3 General Development Score Standard Score.

2.9 ANALYSIS OF PHARMACODYNAMIC VARIABLES

All the pharmacodynamic analyses will be performed using the pharmacodynamic population.

2.9.1 Description of pharmacodynamic variables

Sphingomyelin and sphingomyelin metabolites, including, but not limited to the following:

- Ceramide in DBS.
- Lyso-sphingomyelin and sphingosine-1-phosphate in plasma and/or DBS.

Refer to protocol Section 1.2 and 1.3 for more details.

2.9.2 Statistical analysis

Concentration-time data for sphingomyelin and metabolite levels (eg, ceramide, lyso-sphingomyelin and sphingosine-1-phosphate), observed, change from baseline and percent change from baseline, will be summarized by age group cohorts and overall. Summary plots over time will be provided.

2.10 PATIENT NARRATIVES

Patient narratives criteria for this study are as follows:

- Deaths
- Serious adverse event
- Adverse Events leading to permanent treatment discontinuation
- Adverse events of special interest including the followings:
  - Protocol defined IARs (as specified in Section 2.6.1.3)
  - Pregnancy of female patient or partner of a male patient
  - Symptomatic overdose of study drug
  - Any increase in AST, ALT, total bilirubin, or AP >3x baseline (prior to olipudase alfa therapy) and > ULN
  - Any increase in total bilirubin or AP >1.5x baseline in the presence of AST or ALT > 2x ULN
  - Any increase in ALT or AST >3x the ULN combined with an increase in ALT or AST >2x baseline (prior to olipudase alfa therapy) with symptoms of fatigue, nausea, vomiting, fever, rash, or eosinophilia (> ULN)
  - asymptomatic overdose of study drug

Patients who meet these criteria will be listed and provided to the medical writer.

2.11 DATA HANDLING CONVENTIONS

2.11.1 General conventions

In general, the baseline value is defined as latest value prior to the start of first infusion, except for hematology parameters and possibly ECG parameters. For hematology parameters, the baseline value is defined as the average of all available values before the start of first infusion. Other exception may be the ECG parameters – Day 1/Week 0 (the first infusion visit) collects ECG in triplicate before the infusion; if 2 or more than 2 of these triplicate collections of ECG are before the first infusion, the average of these ECG parameter results will be considered as the baseline value; otherwise, the baseline value of ECG parameters will be taken from the latest observation before the first infusion.

The pre-infusion hematology value at Week 52 will be the average of the two samples collected within 24 hours.

Age at ASMD diagnosis = number of years between the date of birth and the date of ASMD diagnosis.

If the date of birth is missing or incomplete, then the patient's age collected in year and month at screening will be compared to the screening date to derive which year and month the patient was born, and then the corresponding year and the 1st day of the month will be assigned as the birth date.

2.11.2 Missing data

In general, no imputation is planned for missing data. The following approaches are default methods for missing data handling. Some exploratory analyses can be planned with different strategies for treating missing outcomes.

- Categorical data at baseline will be summarized for each age group cohort using counts (n) and percentages (%). The number of patients with missing data may be mentioned, but will not be included in the denominator for the calculation of percentages unless otherwise specified. When relevant, the number of patients with missing data is presented.
- Continuous data: The analyses and summaries for variables with continuous scales will be based on observed data only.

Handling of computation of treatment duration if investigational medicinal product end of treatment date is missing

For the calculation of the treatment duration, the date of the last dose of study treatment is reported on the last infusion page in e-CRF. If this date is missing, the exposure duration will be left as missing.

Handling of medication missing/partial dates

No imputation of medication start/end dates or times will be performed. If a medication date or time is missing or partially missing and it cannot be determined whether it was taken prior to the start of first infusion or concomitantly with study, it will be considered a prior as well as a concomitant medication.

Handling of adverse events with missing data

- If the AE is missing start date and/or time, the treatment-emergent adverse event will be determined by the following conservative principle: An adverse event will be considered a treatment-emergent adverse event if it cannot be confirmed that the event is not a Treatment-emergent adverse event due to missing data.
- If the timing of AE with respect to infusion start and/or end time could not be determined unambiguously, the AE will be assumed to be an algorithm-defined IAR.
- If the assessment of the relationship is missing, then the relationship of possibly related is assumed for the summary table, but no imputation will be done at the data level.
- If the worst severity grade is missing, a missing category will be added in the summary table.

2.11.3 Windows for time points

Statistical analysis will be based on the nominal visit which should follow schedule defined in protocol Section 1.2. Visits will not be windowed for analysis purposes, except for unscheduled visits in some cases as described in Section 2.11.4.

2.11.4 Unscheduled visits

In general, unscheduled visit measurements will not be included in the by-visit summaries, but will be used for computation of baseline and/or summaries of “any time” or “end of study” time points or PCSA. If a scheduled post-baseline visit measurement of an exploratory efficacy parameter is missing, but an unscheduled visit value for the same parameter is available in close proximity of time, the windowing of unscheduled visit to substitute the missing scheduled visit value is defined in the following table. Laboratory, vital signs, ECG and PD parameters are collected at several times during a visit with respect to the infusion time, therefore a window for an unscheduled visit to substitute missing schedule values will not be considered.

Parameter	Scheduled Visit/Target Days	Analysis Visit Window, in Study Days
Spleen volume, liver volume, pulmonary imaging with HRCT, PFT, cycle ergometry, physician's global assessment, health outcome questionnaires	Week 26/183 days	92 days to 274 days
	Week 52/365 days	275 days to 456 days
Chest X-ray, bone age by hand X-ray	Week 52/365 days	275 days to 456 days
Efficacy biomarker, serum bone-specific ALP and C-telopeptide	Week 26/183 days	92 days to 274 days
	Week 52/365 days	275 days to 407 days
	Week 64/449 days	408 days to 463 days
Lipid profile	Week 12/85 days	43 days to 134 days
	Week 26/183 days	135 days to 225 days
	Week 38/267 days	226 days to 316 days
	Week 52/365 days	317 days to 407 days
	Week 64/449 days	408 days to 463 days

2.11.5 Potentially clinically significant abnormalities (PCSA) criteria

In the safety data analysis, patients included in the PCSA analysis for a given parameter will be defined as subjects:

- Exposed to study treatment,
- With at least one parameter result after study treatment administration, and,
- With available laboratory limits when required for analysis.

Liver function tests

ALT	> 3 ULN By distribution analysis: > 3 ULN > 5 ULN > 10 ULN > 20 ULN
AST	> 3 ULN By distribution analysis: > 3 ULN > 5 ULN > 10 ULN > 20 ULN
Alkaline phosphatase	> 1.5 ULN
Total bilirubin	>= 1.3 ULN
ALT and total bilirubin	ALT > 3 ULN and total bilirubin > 2 ULN

Hematology

WBC	Birth/0 to 27 days old (Neonates)	<4.0 GIGA/L >25.0 GIGA/L
	28 days/1 month to 23 months old (Infants)	<4.0 GIGA/L >20.0 GIGA/L
	24 months/2 years to <6 years old (Children)	<3.0 GIGA/L >16.0 GIA/L
	6 to <12 years old (Children)	<5.0 GIGA/L >17.0 GIGA/L
	12 to 16/18 years old (Adolescents)	<4.5 GIGA/L >13.5 GIGA/L
Lymphocytes	Birth/0 to 27 days old (Neonates)	<1.2 GIGA/L >17.0 GIGA/L
	28 days/1 month to 23 months old (Infants)	<2.0 GIGA/L >13.5 GIGA/L
	24 months/2 years to <6 years old (Children)	<1.0 GIGA/L >9.5 GIGA/L
	6 to <12 years old (Children)	<1.0 GIGA/L >8.0 GIGA/L
	12 to 16/18 years old (Adolescents)	<0.6 GIGA/L >6.0 GIGA/L

Neutrophils	Birth/0 to 27 days old (Neonates)	<4.0 GIGA/L (1 day old) <1.5 GIGA/L (2 – 7 days old) <1.25 GIGA/L (>7 day – 1 month old) >1 ULN
	28 days/1 month to 23 months old (Infants)	<1.0 GIGA/L (1 – 3 months) <1.2 GIGA/L (3 – 24 months) >1 ULN
	24 months/2 years to <6 years old (Children)	<1.2 GIGA/L >1 ULN
	6 to <12 years old (Children)	<1.2 GIGA/L >1 ULN
	12 to 16/18 years old (Adolescents)	<1.2 GIGA/L >1 ULN
Eosinophils	All age ranges	>0.5 GIGA/L Or >ULN if ULN ≥0.5 GIGA/L
Hemoglobin	Birth/0 to 27 days old (Neonates)	<12.0 g/dL and 20% decrease from baseline, or decrease from baseline >=2 g/dL
	28 days/1 month to 23 months old (Infants)	<9.0 g/dL and 20% decrease from baseline, or decrease from baseline >=2 g/dL
	24 months/2 years to <16/18 years old (Children, Adolescents)	<10.0 g/dL and 20% decrease from baseline, or decrease from baseline >=2 g/dL
Hematocrit	Birth/0 to 27 days old (Neonates)	<0.39 l/l or 40% >0.61 l/l or 47%
	28 days/1 month to 23 months old (Infants)	<0.29 l/l or 29% >0.42 l/l or 42%
	24 months/2 years to <16/18 years old (Children, Adolescents)	<0.32 l/l or 32% >0.47 l/l or 47%
Platelets	All age ranges	<100 GIGA/L and 20% decrease from baseline >=700 GIGA/L

ECG parameters

HR	Birth/0 to 27 days old (Neonates)	<=90 bpm and decrease from baseline >=20 bpm >=190 bpm and increase from baseline >=20 bpm
	28 days/1 month to 23 months old (Infants)	<=80 bpm and decrease form baseline >=20 bpm >=175 bpm and increase from baseline >=20 bpm
	24 months/2 years to <6 years old (Children)	<=75 bpm and decrease from baseline >=20 bpm >=140 bpm and increase from baseline >=20 bpm
	6 to <12 years old (Children)	<=50 bpm and decrease from baseline >=20 bpm >=120 bpm and increase from baseline >=20 bpm
	12 to 16/18 years old (Adolescents)	<=50 bpm and decrease from baseline >=20 bpm >=120 bpm and increase from baseline >=20 bpm
PR	Birth/0 to 27 days old (Neonates)	>=120 ms
	28 days/1 month to 23 months old (Infants)	>=140 ms
	24 months/2 years to <6 years old (Children)	>=160 ms
	6 to <12 years old (Children)	>=170 ms
	12 to 16/18 years old (Adolescents)	>=180 ms
QRS	Birth/0 to 27 days old (Neonates)	>=85 ms
	28 days/1 month to 23 months old (Infants)	>=85 ms
	24 months/2 years to <6 years old (Children)	>=95 ms
	6 to <12 years old (Children)	>=100 ms
	12 to 16/18 years old (Adolescents)	>=110 ms
QTc	Birth/0 to <12 years old (Neonates, Infants, Children)	<u>Absolute values (ms)</u> Borderline: 431 – 450 ms Prolonged: >450 ms Additional: >=500 ms <u>Increase from baseline</u> Borderline: Increase from baseline 30 – 60 ms Prolonged: Increase from baseline >60 ms

	12 to 16/18 years old (Adolescents)	Absolute values (ms) Borderline: 431 – 450 ms (Boys); 451 – 470 ms (Girls) Prolonged: >450 ms (Boys); >470 ms (Girls) Additional: >=500 ms Increase from baseline Borderline: Increase from baseline 30 – 60 ms Prolonged: Increase from baseline >60 ms

Clinical chemistry

Creatinine	Birth/0 to <6 years old (Neonates, Infants, Children)	>53 µmol/L or 0.6 mg/dL
	6 years to <12 years old (Children)	>=90 µmol/L or 1.1 mg/dL
	12 years to 16/18 years old (Adolescents)	>=132 µmol/L or 1.5 mg/dL
Blood urea nitrogen (BUN)	Birth/0 to 27 days old (Neonates)	>=4.3 mmol/L or 12 mg/dl
	28 days/1 month to 16/18 years old (Infants, Children, Adolescents)	>=6.4 mmol/L or 18 mg/dl
Chloride	All age ranges	<80 mmol/L >115 mmol/L
Sodium	All age ranges	<=129 mmol/L >=150 mmol/L
Potassium	Birth/0 to 27 days old (Neonates)	<=3.0 mmol/L >=7.0 mmol/L
	28 days/1 month to 23 months old (Infants)	<=3.5 mmol/L >=6.0 mmol/L
	24 months/2 years to 16/18 years old (Children, Adolescents)	<=3.5 mmol/L >=5.5 mmol/L
Glucose	All age ranges	Hypoglycaemia <2.7 mmol/L Hyperglycaemia >=7 mmol/L (fasted after >12 hours of fast); >=10.0 mmol/L (unfasted)

Vital signs

Heart rate	Birth/0 to 27 days old (Neonates)	<=90 bpm and decrease from baseline >=20 bpm >=190 bpm and increase from baseline >=20 bpm
	28 days/1 month to 23 months old (Infants)	<= 80 bpm and decrease form baseline >=20 bpm >= 175 bpm and increase from baseline >= 20 bpm
	24 months/2 years to <6 years old (Children)	<= 75 bpm and decrease from baseline >=20 bpm >= 140 bpm and increase from baseline >= 20 bpm
	6 to <12 years old (Children)	<=50 bpm and decrease from baseline >=20 bpm >=120 bpm and increase from baseline >=20 bpm
	12 to 16/18 years old (Adolescents)	<=50 bpm and decrease from baseline >=20 bpm >=120 bpm and increase from baseline >=20 bpm
Systolic BP	Birth/0 to 27 days old (Neonates)	<=60 mmHg and decrease from baseline >=20 mmHg >=85 mmHg and increase from baseline >=20 mmHg
	28 days/1 month to 23 months old (Infants)	<=70 mmHg and decrease from baseline >=20 mmHg >=98 mmHg and increase from baseline >=20 mmHg
	24 months/2 years to <6 years old (Children)	<=70 mmHg and decrease from baseline >=20 mmHg >=101 mmHg and increase from baseline >=20 mmHg
	6 to <12 years old (Children)	<=80 mmHg and decrease from baseline >=20 mmHg >=108 mmHg and increase from baseline >=20 mmHg
	12 to 16/18 years old (Adolescents)	<=90 mmHg and decrease from baseline >=20 mmHg >=119 mmHg and increase from baseline >=20 mmHg
Diastolic BP	Birth/0 to 27 days old (Neonates)	<=34 mmHg and decrease from baseline >=10 mmHg >=50 mmHg and increase from baseline >=10 mmHg
	28 days/1 month to 23 months old (Infants)	<=34 mmHg and decrease from baseline >=10 mmHg >=54 mmHg and increase from baseline >=10 mmHg
	24 months/2 years to <6 years old (Children)	<=34 mmHg and decrease from baseline >=10 mmHg >=59 mmHg and increase from baseline >=10 mmHg
	6 to <12 years old (Children)	<=48 mmHg and decrease from baseline >=10 mmHg >=72 mmHg and increase from baseline >=10 mmHg
	12 to 16/18 years old (Adolescents)	<=54 mmHg and decrease from baseline >=10 mmHg >=78 mmHg and increase from baseline >=10 mmHg

3 INTERIM ANALYSIS

Analysis of data, including summary tables, figures and listings, may be conducted intermittently during the study and upon completion of a cohort's key milestones, eg, week 26 or 52 efficacy assessments. However, due to the nature of the open label study, these potential analyses will have no impact on primary and secondary objectives of this study.

4 SOFTWARE DOCUMENTATION

The analysis of clinical data will be performed using Statistical Analysis System (SAS®) (version 9.4 or higher, SAS Institute, NC USA). Statistical analysis of pharmacokinetic parameters will be performed using Phoenix WinNonlin Professional, Version 5.2 or above.

5 REFERENCES

1. Global Lung Initiative (GLI) reference [Online]. [cited 2019 Nov 05]. Available from: URL: http://www.ers-education.org/guidelines/global-lung-function-initiative.aspx

2. Koopman, M, Zanen P, Kruitwagen CL, van der Ent CK, Arets HG. Corrigendum to "Reference values for pediatric pulmonary function testing: The Utrecht dataset". Respiratory Medicine. 2011;105:1970-1.

3. Macintyre, N, Crapo RO, Vieqi G, Johnson DC, van der Grinten CP, Brusasco V, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. Eur Respir J. 2005;26:720-35.

4. Greulich, WW, Pyle SI. Radiographic atlas of skeletal development of the hand and wrist. 2nd ed. Standford, Calif.: Standford University Press; 1959.

5. PedsQL Generic Core Scales scoring algorithm [Online]. [cited 2019 Nov 05]. Available from: URL: http://www.pedsql.org/

6. Varni JW, Beaujean AA, Limbers CA. Factorial invariance of pediatric patient self-reported fatigue across age and gender: a multigroup confirmatory factor analysis approach utilizing the PedsQL™ Multidimensional Fatigue Scale. Qual Life Res. 2013;22:2581-94.

7. Aplern, GD. Developmental Profile 3 (DP-3) [Manual]. Torrance, CA: Western Psychological Services. [Online]. 2007 [cited 2019 Nov 05]. Available from: URL:https://www.wpspublish.com/store/p/2743/dp-3-developmental-profile-3

8. Harrison PL, Oakland T. Adaptive Behavior Assessment System, Third Edition [Manual]. Torrance, CA: Western Psychological Services. [Online]. 2015 [cited 2019 Nov 05]. Available from: URL: http://www.wpspublish.com/store/p/3234/abas-3-adaptive-behavior-assessment-system-third-edition

6 APPENDICES

Adjustment of DLco measurements

1 Introduction

Adjustments of DLco measurements are described in detail the 2005 ERS/ATS statement. In the paper, however, the adjustment is described as a change of the DLco predicted values. After several discussions with opinion leaders, we have to conclude that it is more common to compute a DLadj value that is computed from the measured DLco value. This DLadj value is then compared to the (unchanged) predicted values. The following table describes this:

Comment	Parameter	Predicted	% Predicted
	$DLco_{meas}$	$DLco_{pred}$	$DLco_{pred}/DLco_{meas}$
ERS/ATS statement	$DLco_{meas}$	$DLco_{pred}*\alpha$	$DLco_{pred}*\alpha/DLco_{meas}$
EasyOne pro	$DLadj = DLco_{meas}*\beta$	$DLco_{pred}$	$DLco_{pred}/(DLco_{meas}*\beta)$

As can be seen from the table above, the % Predicted is the same if $\beta = 1/\alpha$. This means that the equations of the ERS/ATS statement have to be inverted (ie, the factor is multiplied instead of divided). If compared to the original formulas in the ERS/ATS statement, the following formulas are therefore 'inverted' (multiplication instead of division and vice versa).

2 Adjustment for hemoglobin

The following formula is applied to correct for hemoglobin (Hb):

Male adults (age >=15): DLadj = DLco/(1.7*Hb/(10.22+Hb)) Hb in g/dL

Female and children (age<15): DLadj = DLco/(1.7*Hb/(9.38+Hb)) Hb in g/dL

Unit conversion: Hb [g/dL] = Hb [mmol/L] /0.616

(according to other sources the factor is 0.6206)

Allowed range for Hb: 0 to 100 g/dL

Default value for male adults (age >=15): 14.6 g/dL (9.00 mmol/L)

Default value for female and children (age <15): 13.4 g/dL (8.26 mmol/L)

 AppNote-DLCO-Adjustment-V01R.doc

Signature Page for VV-CLIN-0195653 v3.0
dfi13803-16-1-9-sap

Approve & eSign	[redacted] Clinical [redacted]

Approve & eSign	[redacted] [redacted]