Study Title:
A Multicenter, Safety and Efficacy Study of Taliglucerase Alfa in Subjects With Type 3 Gaucher Disease

NCT Number:
NCT04002830

Statistical Plan Date:
24 September 2020

Protocol WI224302

A MULTICENTER, SAFETY AND EFFICACY STUDY OF TALIGLUCERASE ALFA IN SUBJECTS WITH TYPE 3 GAUCHER DISEASE

Statistical Analysis Plan
(SAP)

Version: **1**

Date: **24 Sep 2020**

TABLE OF CONTENTS

LIST OF TABLES

APPENDICES

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
V1	Amendment 4	Original SAP	No

2. INTRODUCTION

Gaucher disease (GD) is a rare genetic disorder of lipid metabolism and is the most common inherited lysosomal storage disorder, thought to affect around 10,000 people globally. Gaucher disease is classified into 3 major phenotypic subcategories based on the absence (Type 1) or presence and severity (Types 2 and 3) of primary central nervous system involvement. Gaucher disease Type 3 (except Type 3c), the population under study in this current protocol, affects approximately 5% of documented GD cases. Type 3 GD is associated with progressive neurodegenerative disease, with symptom onset usually between infancy and adolescence. Manifestations of Type 3 GD vary among patients, ranging from visceral and bone disease, cytopenia, hearing impairment, supranuclear gaze palsy, developmental delay, and cardiac calcification to myoclonic encephalopathy with seizures and eventual death.

Taliglucerase alfa, a hydrolytic lysosomal glucocerebroside-specific enzyme for intravenous infusion, is a recombinant active form of the lysosomal enzyme, β-GCD (β-glucocerebrosidase), produced using transgenic carrot plant root cells grown in suspension culture with a modified gene sequence encoding human β-GCD. Taliglucerase alfa is currently an approved therapy in the United States and many other countries for adults and children with a confirmed diagnosis of Type 1 GD, and is also approved for use in Type 3 GD in a small number of countries.

The effects of enzyme replacement therapy (ERT) on patients with Type 1 GD have been clearly documented and have a beneficial effect on visceral and hematologic disease parameters. It is known that recombinant enzyme does not pass the blood-brain barrier and has no effect on neurologic involvement. This prospective study aims to objectively evaluate the hematologic and visceral effects of ERT with taliglucerase alfa on a rather clinically and genetically homogenous group of treatment-naïve patients with Type 3 GD. The results of this study are expected to provide a more objective view of the degree of response of this patient type, and potentially create new areas of research.

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study WI224302. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

2.1. Study Objectives, Endpoints, and Estimands

The efficacy objectives of this study are to assess the efficacy of taliglucerase alfa in subjects with Type 3 GD, as measured by percent change from baseline in spleen and liver volume evaluated by magnetic resonance imaging (MRI), hemoglobin, platelet count, and lyso-glucosylsphingosine (Lyso-Gb1); change from baseline in growth and development (height, weight, sexual development by Tanner classification, and bone age by X-ray of left hand and wrist), bone mineral density evaluated by dual-emission X-ray absorptiometry (DEXA), and evaluation of fatigue.

The safety of taliglucerase alfa will be assessed by adverse events, clinical laboratory tests, vital signs, and physical examinations. Anti-taliglucerase alfa antibodies will also be assessed.

2.1.1. Primary Estimand(s)

The primary estimand of this study uses the treatment policy strategy and estimates the effect for all participants are treated and regardless of whether an intercurrent event occurs. It includes the following 4 attributes:

- Population: All patients who suffer from the Type 3 GD as defined by the inclusion and exclusion criteria, and who have received at least one complete dose (or partial dose if adverse events prevent completing infusion dose);
- Variable: Percent change from baseline to Month 12 in spleen volume expressed in MN (Multiples of Normal) and measured by MRI (Magnetic Resonance Imaging);
- Intercurrent event: All data after an intercurrent event (discontinuation of treatment, discontinuation of study, etc), if collected, will be included;
- Population-level summary: median percent change from baseline to Month 12 and p-value from the Wilcoxon signed ranks test.

2.1.2. Secondary Estimand(s)

There are couple of secondary estimands of this study, which estimate the effect for all participants treated and regardless of whether an intercurrent event occurs. Each of them includes the following 4 attributes:

- Population: All patients who suffer from the Type 3 GD as defined by the inclusion and exclusion criteria, and who have received at least one complete dose (or partial dose if adverse events prevent completing infusion dose);
- Variable:
    - Percent change from baseline to Month 12 in liver volume expressed in MN and measured by MRI;

- Percent change from baseline to Month 12 in hemoglobin;
- Percent change from baseline to Month 12 in platelet;
- Percent change from baseline to Month 12 in Lyso-Gb1 (lyso-glucosylsphingosine);
- Percent change from baseline to Month 12 in chitotriosidase.

- Intercurrent event: All data after an intercurrent event (discontinuation of treatment, discontinuation of study, etc), if collected, will be included;
- Population-level summary: median percent change from baseline to Month 12 and p-value from the Wilcoxon signed ranks test.

2.1.3. Additional Estimand(s)

The responders to the primary efficacy endpoint are defined as participants who had the reduction from baseline greater than 20% in spleen volume MN. The responder rate among all patients received at least one complete dose, will be used as the populatin-level summary for this alternative estimand.

2.2. Study Design

This is a multicenter study to assess the safety and efficacy of taliglucerase alfa (60 units/kg) in previously untreated subjects of any age with Type 3 GD. Subjects will receive an infusion of taliglucerase alfa every 2 weeks for 12 months. Subjects who tolerate the infusions well for 3 months will be eligible for home therapy based on approval of the Investigator.

Type 3 GD is a very rare genetic orphan disease. This study will not have a placebo control group because it would be unethical to deny patients the possibility to improve their symptomology with an ERT that has been proved efficacious and safe, not only for Type 1 GD in many countries but also for Type 3 GD in a small number of countries. Thus, all subjects will receive the investigational product in this study.

The parameters chosen as endpoints for this study are the most relevant parameters to GD and allow a significant and relevant evaluation of improvement as a result of treatment for 12 months. Improvement in the clinical manifestations of the disease leads to the achievement of the therapeutic goals of GD treatment in the medical community, which are improved fatigue and decreased risk for morbidities.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

Percent change from baseline will be calculated at the participant level = [(Post baseline value – Baseline)/Baseline] * 100%

3.1. Primary Endpoint

The primary efficacy evaluation will be the percent change in spleen volume (expressed in MN) measured by MRI from baseline to Month 12.

Spleen volume will be measured at the Screening and Month 12 according to the Schedule of Activities in Section 9 of the protocol, measured by MRI by site according to the Appendix 4 of the protocol. The spleen volume (cm^3) will be used for the sponsor to calculate the spleen volume by multiples of normal (MN) data applying the formula below:

$$Spleen\ Volume\ by\ MN = \frac{Spleen\ Volume}{2cm^3 / kg * body\ weight}$$

When body weight (kg) is not available at a visit, the body weight measured in the previous visit will be used to estimate the participant's expected normal volume at the visit.

3.2. Secondary Endpoint(s)

3.2.1. Liver Volume by MN

The secondary efficacy evaluation will be the percent change in liver volume (expressed in MN) measured by MRI from baseline to Month 12.

Liver volume will be measured at the Screening and Month 12 according to the Schedule of Activities in Section 9 of the protocol, measured by MRI by site according to the Appendix 4 of the protocol. The spleen volume (cm^3) will be used for the sponsor to calculate the liver volume by multiples of normal (MN) data applying the formula below:

$$Liver\ Volume\ by\ MN = \frac{Liver\ Volume}{25cm^3 / kg * body\ weight}$$

When body weight (kg) is not available at a visit, the body weight measured in the previous visit will be used to estimate the participant's expected normal volume at the visit.

3.2.2. Hemoglobin (g/dL)

Hemoglobin will be measured at Screening, Baseline, Month 3, Month 6, Month 9 and Month 12 according to the Schedule of Activities in Section 9 of the protocol.

3.2.3. Platelet Count (10^3/uL)

Platelets will be counted at Screening, Baseline, Month 3, Month 6, Month 9 and Month 12 according to the Schedule of Activities in Section 9 of the protocol.

3.2.4. Lyso-Gb1 Biomarker (ng/mL)

Lyso-Gb1 (lyso-glucosylsphinegosine) will be measured at Baseline, Month 3, Month 6, Month 9 and Month 12 according to the Schedule of Activities in Section 9 of the protocol.

3.2.5. Chitotriosidase (nmol/ml/hr)

Chitotriosidase will be measured at Baseline, Month 3, Month 6, Month 9 and Month 12 at the same schedule of the biomarker of Lyso-Gb1 according to the Schedule of Activities in Section 9 of the protocol.

3.3. Other Endpoint(s)

3.3.1. Height SDS and Weight SDS

For change in growth and development, the height and weight will be assigned to a SDS/Z-scores using the World Health Organization (WHO) Child Growth Standards. The Center for Disease Control and Prevention (CDC) reference standards will be used for the age categories where the WHO reference standards are not available.

Height and weight will be collected at Screening, Baseline, Month 3, Month 6, Month 9 and Month 12 according to the Schedule of Activities in Section 9 of the protocol.

Z-scores include height standard deviation score (SDS) by chronologic age, height SDS by bone age, weight SDS by chronologic age, weight SDS by bone age. Z-scores will be calculated by the sponsor using SAS macros following World Health Organization (WHO) reference standards. The input variable of age will be repeatedly applied using chronologic age and bone age in order to derive the above four SDS. Detailed documents on how the physical growth curves and motor milestone windows of achievement were developed as well as application tools to support implementation of the standards is publicly available at WHO website[1] for children ages 0-5 years and for children ages 5-19 years separately. Height SDS and Weight SDS should be calculated using the Statistical Analysis System (SAS) macros available at the same website.[1]

There are no WHO reference standards for weight SDS for pediatric participants >10 years of age. Therefore, as an alternative, weight SDS for all pediatric participants will be also calculated using the CDC reference standards[2] and SAS programs.

In calculation of the percentage change from baseline for SDS, the absolute value of the baseline will be used as the denominator to address the situation where the value is negative.

3.3.2. Tanner Stage

Tanner stage of development will be collected at Baseline and Month 12 according to the Schedule of Activities in Section 9 of the protocol, including pubic hair stage (stage 1 to 5), genitalia stage (stage 1 to 5) and testis volume (mL).

3.3.3. Bone Measurement

Bone age assessed by X-ray of left hand and wrist will be collected at Baseline and Month 12 according to the Schedule of Activities in Section 9 of the protocol.

Bone mineral density (BMD) and Z-score assessed by dual-emission X-ray absorptiometry (DEXA) for L femoral neck, total left hip, R femoral neck, total right hip, lumbar spine, cortical bone and total body will be collected at baseline and Month 12 according to the Schedule of Activities in Section 9 of the protocol. The DEXA will be done in children 9 years of age and above.

3.3.4. Modified Severity Scoring Tool (mSST)

Modified Severity Scoring Tool (mSST) will be collected at Screening, Month 6 and Month 12 according to the Schedule of Activities in Section 9 of the protocol. The total score is a sum of the single scores from gaze palsy, ophthalmology, epilepsy, age at first epilepsy, cognitive ability, ataxia of gait, cerebellar tremor, pyramidal, extrapyramidal, swallowing difficulties/oral bulbar function, speech and spinal alignment. It ranges from 0 to 36, while 0 indicating normal and 36 indicating severe symptom.

3.3.5. Organ Volumes measured by Ultrasound

Spleen volume and liver volume will be also measured by ultrasound at the Screening, Month 6 and Month 12 according to the Schedule of Activities in Section 9 of the protocol. If collected, the spleen volume and liver volume will be calculated as the product of craniocaudal dimension (cm), anteroposterior dimension (cm) and longitudinal dimension (cm) from ultrasound measurement.

The spleen volume by MN will be calculated by sponsor applying the formula in Section 3.1 based on the spleen volume (cm^3) measurement by ultrasound and body weight. The liver volume by MN will be calculated by sponsor applying the formula in Section 3.2.1 based on the liver volume (cm^3) measurement by ultrasound and body weight.

3.4. Baseline Variables

Baseline is defined as Screening visit (Visit 0) for spleen volume defined in Section 3.1, liver volume defined in Section 3.2.1 and mSST defined in Section 3.3.4. Baseline is defined as Day 1 (or Visit 1) for the endpoints specified in Sections 3.2.2 to 3.3, which are corresponding to the secondary efficacy assessment, growth development, biomarker measurement, as well as the adverse events in Section 3.5 for the safety assessment.

3.5. Safety Endpoints

3.5.1. Adverse Events

The adverse event (AE) including time of the onset, severity, action/treatment and outcome of the AE resolution, related to taliglucerase alfa treatment or not, will be collected at each scheduled visit with the visit window specified in the Schedule of Activities in Section 9 of the protocol.

An AE is considered as a treatment emergent adverse event (TEAE), if the event occurring or worsening on or after the first dose of taliglucerase alfa in the study. The AE reporting period for this study will be defined as the duration from the first dose of taliglucerase alfa in the study until 30 calendar days following the final visit dose.

3.5.2. Laboratory Data

Lab data other than hemoglobin and platelet counts will be collected at Screening, baseline, Month 3, Month 6, Month 9 and Month 12 according to the Schedule of Activities in Section 9 of the protocol, including hematology, biochemistry and urinalysis lab tests.

3.5.3. Vital Signs

Vital sign data will be collected during each bi-weekly Taliglucerase alfa infusion. It will be evaluated every 30 minutes if the participant tolerates the previous infusions; otherwise, every 15 minutes for the first hour.

3.5.4. Antibodies

Anti-taliglucerase alfa antibodies will be collected at baseline, Month 6 and Month 12 according to the Schedule of Activities in Section 9 of the protocol, including the status (positive, negative, not done) for IgG, in-vitro neutralizing, cell based neutralizing, and IgE anti-taliglucerase alfa antibodies.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

Data for all participants will be assessed to determine if participants meet the criteria for inclusion in each analysis population prior to releasing the database and classifications will be documented per standard operating procedures.

The analysis populations for reporting are defined in the following table.

Population	Description
Enrolled	All participants who sign the informed consent document.
Evaluable or Full Analysis Set (FAS)	All enrolled participants to the study who take at least 1 complete dose of taliglucerase alfa (or partial dose if adverse events prevent completing infusion dose). Participants will be collected according to the study drug actually received for this non-randomized study.
Safety	All participants who receive at least 1 dose of study drug. Participants will be analyzed according to the study drug they actually received. An enrolled but not treated participant will be excluded from the safety analyses.
Per Protocol set (PP)	All participants who complete the primary efficacy data collection at Month 12 and baseline assessment, and receive at least 1 complete dose of taliglucerase alfa (or partial dose if adverse events prevent completing infusion dose).

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

For the primary efficacy endpoint, a Wilcoxon signed rank test at 2-sided significance level of 0.05 will be applied to test the null hypothesis that the median percent change from baseline to Month 12 in spleen volume MN is equal to -20%. The corresponding p-value will be reported.

The primary null hypothesis is:

$$H_0: \lambda = -20\%$$

and the primary alternative hypotheses are:

$$H_{a1}: \lambda < -20\% \text{ and } H_{a2}: \lambda > -20\%,$$

where λ is the common median of the observed percent difference from the paired data.

Taliglucerase alfa will be considered efficacious with respect to the spleen volume MN, if the median percent change from baseline to Month 12 is significantly less than -20% (or the median precent reduction is greater than 20%) at the 2-sided 0.05 level, ie, p-value from Wilcoxon signed rank test less than 0.05.

For the secondary efficacy endpoints, such as hemoglobin, platelet counts and Lyso-Gb1, the same Wilcoxon signed rank test will be applied for the percent change from baseline to Month 12. No adjustment for multiple comparisons will be made. Though hemoglobin, platelet counts, chitotriosidase and Lyso-Gb1 will be collected at Month 3, 6 and 9 as well, changes from baseline to Months 3, 6 and 9 will be descriptively summarized only.

5.2. General Methods

5.2.1. Analyses for Binary Endpoints

Frequency and percentage of the responders defined in Section 2.1.3, will be provided as a descriptive type of summary. The 95% confidence interval for the responder rate will be provided using exact Clopper-Pearson method via one sample procedure in SAS Proc Freq. The sample code is provided in Appendix 4.

5.2.2. Analyses for Continuous Endpoints

The Wilcoxon signed rank test will be performed for the percent change from baseline to Month 12 in primary and secondary efficacy endpoints due to the possible skewed distribution for the efficacy data.

The rank with positive and negative signs will be assigned to the percent difference (D_i) between baseline and Month 12 values. D_i are a random variable, which are mutually independent from each participant and have the same median. All the zero differences are ignored in the analysis. The Wilcoxon signed rank test inference is based on permutational distribution of its test statistics. This test will be implemented and exact p-values will be produced via SAS Proc Univariate. The Proc Univariate excludes missing values for an analysis variable before calculating statistics. The sample code is provided in Appendix 4.

5.2.3. Analyses for Categorical Endpoints

Frequency and percentage for each category of the variable will be provided as a descriptive summary for each visit. For example, the frequency and percentage of the stage 1 to 5 for tanner stage will be presented at Baseline and Month 12.

5.2.4. Analyses for Time-to-Event Endpoints

Not applicable.

5.3. Methods to Manage Missing Data

The missing data for the primary and secondary efficacy endpoints analyzed by the Wilcoxon signed rank test will be imputed using last observation carried forward (LOCF) method in FAS, ie, the measurement obtained at the early termination visit will be used to impute the Month 12 data.

For imaging-based endpoints, the spleen volume and liver volume measured by the MRI at the early termination visit will be taken as the Month 12 measurement, if the Month 12 visit is missing. If there is no post-baseline MRI measurement, the percentage change from baseline to Month 12 in organ volumes by ultrasound can be used to impute the missing percentage change from baseline in organ volumes by MRI per clinical judgment. The percentage change from baseline needs to be derived within the same imaging test, ie, based on post-baseline organ volume measured by ultrasound and baseline organ volume measured by ultrasound. If there is no post-baseline imaging measurement, the baseline observation carried forward (BOCF) will be applied for missing data, ie, the baseline volume measured by MRI will be used to impute Month 12 MRI data.

Missing safety data will be handled according to the Pfizer Clinical Data Interchange Standards Consortium (CDISC) safety rulebook and data standard rules for imputation, including missing dates (when the date is required for a calculation) and missing severity of adverse events.

6. ANALYSES AND SUMMARIES

6.1. Primary Endpoint

6.1.1. Main Analysis

- Estimand strategy: Treatment policy strategy for percent change from baseline to Month 12 in spleen volume MN measured by MRI.

- Analysis set: FAS (defined in Section 4). All participants assigned to study drug and received at least 1 complete dose of taliglucerase alfa (or partial dose if adverse events prevent completing infusion dose) will be included.

- Analysis methodology: Percent change from baseline to Month 12 will be analyzed using Wilcoxon signed rank test detailed in Section 5.2.2. In order to judge whether this primary efficacy endpoint meets the 20% reduction from baseline, a 20% shift will be added on the percent change from baseline for each participant in the analysis set before the procedure of Proc Univariate (Appendix 4).

- Intercurrent events and missing data: Data after study drug discontinuation will be included. Missing values will be imputed using LOCF and other methods specified in Section 5.3 and be included in calculating percent change from baseline.

- Summary of statistical comparison: The 2-sided p-value from the exact inference for Wilcoxon signed rank statistics for comparing Month 12 to baseline value, as well as the median percent change from baseline, will be presented and used to judge whether efficacious criteria is met. The n, minimum and maximum for the observed percent change from baseline will be presented together as well.

- Descriptive summary including n, mean, SD, median, minimum, and maximum for absolute value, change from baseline and percent change from baseline in spleen volume and spleen volume MN measured by MRI at each visit will be presented.

6.1.2. Sensitivity/Supplementary Analyses

Sensitivity analyses

- To assess the impact of early discontinuation, the same Wilcoxon signed rank test will be repeated in the PP set (Section 4) among those who have the primary efficacy data collected by MRI at Month 12 and Baseline.

- Descriptive summary for absolute value, change from baseline and percent change from baseline in spleen volume and spleen volume MN measured by MRI at each visit will be repeated in the PP set.

Supplementary analysis

- In addition to the main analysis on the primary estimand, the supportive analysis on the alternative estimand of the binary responder rate based on spleen volume MN reduction measured by MRI will be provided in FAS. This includes the frequency and percentage of the responders, as well as the 95% confidence interval (CI) for the responder rate using exact Clopper-Pearson method. The supportive evidence can be obtained, if the lower bound of the 95% CI excludes the responder rate of 50%.

6.2. Key Secondary Endpoints

6.2.1. Main Analysis

- Estimand strategy: Treatment policy strategy for percent change from baseline to Month 12 in liver volume MN measured by MRI, hemoglobin, platelet and Lyso-Gb1 (lyso-glucosylsphingosine), chitotriosidase.

- Analysis set: FAS (defined in Section 4). All participants assigned to study drug and received at least 1 complete dose of taliglucerase alfa (or partial dose if adverse events prevent completing infusion dose) will be included.

- Analysis methodology: Percent change from baseline to Month 12 will be analyzed using Wilcoxon signed rank test detailed in Section 5.2.2.

- Intercurrent events and missing data: Data after study drug discontinuation will be included; Missing values will be imputed using LOCF, and other methods specified in Section 5.3, and included in calculating percent change from baseline.

- Summary of statistical comparison: The 2-sided p-value from the exact inference for Wilcoxon signed rank statistics for comparing Month 12 to baseline value, as well as median percent change from baseline, will be presented. The n, minimum and maximum for observed percent change from baseline will be presented together as well.

- Descriptive summary including n, mean, SD, median, minimum, and maximum for absolute value, change from baseline and percent change from baseline at each visit will be presented. The visit includes baseline and Month 12 for liver volume and liver volume MN measured by MRI, and baseline, Month 3, 6, 9 and 12 for hemoglobin, platelet, chitotriosidase and Lyso-Gb1.

6.2.2. Sensitivity/Supplementary Analysis

There are no sensitivity and supplementary analyses planned for these endpoints.

6.3. Other Secondary Endpoints

- Estimand strategy: Treatment policy strategy for change from baseline in height SDS, weight SDS, tanner stage, bone age, BMD and Z-score for bone assessment, total score from mSST, as well as change from baseline in organ volume measured by ultrasound if available.

- Analysis set: FAS (defined in Section 4). All participants assigned to study drug and who take at least 1 complete dose of taliglucerase alfa (or partial dose if adverse events prevent completing infusion dose).

- Intercurrent events and missing data: Data after study drug discontinuation will be excluded; Missing values won't be included in calculating change from baseline and will not be imputed.

- Descriptive statistics including n, mean, SD, median, minimum, and maximum will be presented for observed continuous variables at each visit. This includes the absolute values, change from baseline and percent change from baseline in height SDS by chronologic age, height SDS by bone age, weight SDS by chronologic age, weight SDS by bone age, bone age, BMD and Z-score for bone assessment, right and left testis volume, total score from mSST, and spleen volume, spleen volume MN, liver volume, liver volume MN measured by ultrasound.

- BDM and Z-score for bone assessment will be descriptively summarized by each bone area, such as L femoral neck, total left hip, total body, etc. that listed in Section 3.3.3.

- Frequency and percentage will be provided for categorical variables at each visit, such as tanner stage.

6.4. Other Endpoint(s)

6.4.1. Immunogenicity (Antibodies)

Frequency and percentage of anti-taliglucerase alfa antibodies in FAS will be reported at baseline, Month 6 and Month 12, if adequate data are collected.

6.5. Subset Analyses

There is no planned subset analysis for this study.

6.6. Baseline and Other Summaries and Analyses

6.6.1. Baseline Summaries

Descriptive summary including n, mean, SD, median, minimum, and maximum will be presented for continuous baseline variables. Frequency and percentage will be provided for categorical baseline variables. The following baseline data and disease characteristics will be summarized, if data permits.

- Demographic data: Age, race, ethnicity, proportion of Ashkenazi Jewish descent (via number of grandparents) will be summarized by gender at birth and 'All Participants' in accordance with the Pfizer reporting standards.

- Primary diagnosis: Type of GD, method of diagnosis and symptoms presented at the diagnosis will be summarized.

- Medical history: Genarl MH will be summarized by SOC and PT according to MedDRA dictionary in the latest available version.

6.6.2. Study Conduct and Participant Disposition

Participant evaluation groups will show end of study participant disposition and will show which participants were analyzed for efficacy and safety (adverse events and laboratory data).

Frequency counts and percentage will be reported for participant discontinuation(s), temporary discontinuation(s) or dose reduction due to adverse events and reasons for discontinuation.

Data will be reported in accordance with the Pfizer reporting standards.

6.6.3. Study Treatment Exposure

Taliglucerase dosing (study drug administration schedule) will be listed for each participant. Dose will be descriptively summarized as n, median, minimum, maximum, mean and SD.

6.6.4. Concomitant Medications and Nondrug Treatments

The concomitant medication(s) and non-drug treatment(s) will be provided in the listings.

6.7. Safety Summaries and Analyses

All the safety data will be summarized descriptively through the appropriate data tablulations in safety population (defined in Section 4) according to Pfizer Safety Rulebook and CDISC standards analyses, to evaluate any potential risk associated with the safety and toleration of administering taliglucerase alfa in Type 3 GD patients.

6.7.1. Adverse Events

Treatment-emergent adverse events (TEAE) will be summarized with counts and percentage for safety population (defined in Section 4) according to Pfizer CDISC standard. All adverse events will be reported in a data listing.

In addition to the system organ class (SOC) and preferred term (PT) defined in the live MedDRA version at the final reporting stage, an additional SOC of hypersensitivity reactions will be summarized using frequency and percentage for each preferred term, including but not limited to:

- Early clinical signs: sensation of warmth and itching, feelings of anxiety or panic;
- Moderate clincal signs: pruritus, flushing, urticaria, chest discomfort, mild hypotension;
- Progressive clinical signs: erythematous or massive urticarial rash;
- Severe clinical signs: hypotension, bronchospasm (wheezing), laryngeal edema (dyspnea, stridor, aphonia, drooling), arrhythmias.

6.7.2. Laboratory Data

Laboratory data will be presented in accordance with the Pfizer CDISC reporting standards.

6.7.3. Vital Signs

All vital sign data will be listed by participant and visit.

Though vital signs will be collected during each infusion visit per standard of care, only vital sign collected at the baseline, Month 3, 6, 9 and 12 will be summarized descriptively. Absolute values, change from baseline and percent change from baseline for vital sign data will be summarized using n, mean, SD, median and range.

6.7.4. Physical Examination

Physical examination and neurologic assessment data will be listed only by participant and visit.

7. INTERIM ANALYSES

There is no planned interim analysis in this study.

8. REFERENCES

1. WHO reference standards and SAS programs for child growth at the WHO website (http://www.who.int/childgrowth/software/en/) for children ages 0-5 years and (http://www.who.int/growthref/en/) for children ages 5-19 years.

2. CDC reference standards for 2000 CDC growth charts and SAS programs at the CDC website (http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm).

9. APPENDICES

Appendix 1. Summary of Efficacy Analyses

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Percent change from baseline to Month 12 in spleen volume MN measured by MRI	Summary	Evaluable (FAS), PP	observed data at each visit	N/A
	Main analysis	Evaluable (FAS)	Month 12 or data from early termination assessment and baseline	Wilcoxon signed rank test (exact p-value), median %change
	Sensitivity analysis	PP	completed both Month 12 and baseline without missing data imputation	Wilcoxon signed rank test (exact p-value), median %change
Responder rate in spleen volume reduction measured by MRI	Supplementary analysis	Evaluable (FAS)	Month 12 or data from early termination assessment and baseline	Clopper-Pearson CI
Percent change from baseline to Month 12 in liver volume MN measured by MRI	Summary	Evaluable (FAS)	observed data at each visit	N/A
	Main analysis	Evaluable (FAS)	Month 12 or data from early termination assessment and baseline	Wilcoxon signed rank test (exact p-value), median %change
Percent change from baseline in hemoglobin	Summary	Evaluable (FAS)	observed data at each visit (baseline, Month 3, 6, 9, 12)	N/A
	Main analysis	Evaluable (FAS)	Month 12 or data from early termination assessment and baseline	Wilcoxon signed rank test (exact p-value), median %change
Percent change from baseline in platelet	Summary	Evaluable (FAS)	observed data at each visit (baseline, Month 3, 6, 9, 12)	N/A
	Main analysis	Evaluable (FAS)	Month 12 or data from early termination assessment and baseline	Wilcoxon signed rank test (exact p-value), median %change
Percent change from baseline in Lyso-Gb1	Summary	Evaluable (FAS)	observed data at each visit (baseline, Month 3, 6, 9, 12)	N/A
	Main analysis	Evaluable (FAS)	Month 12 or data from early termination assessment and baseline	Wilcoxon signed rank test (exact p-value), median %change

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Percent change from baseline in chitotriosidase	Summary	Evaluable (FAS)	observed data at each visit (baseline, Month 3, 6, 9, 12)	N/A
	Main analysis	Evaluable (FAS)	Month 12 or data from early termination assessment and baseline	Wilcoxon signed rank test (exact p-value), median %change
Change from baseline in height SDS	Summary	Evaluable (FAS)	observed data at each visit	N/A
Change from baseline in weight SDS	Summary	Evaluable (FAS)	observed data at each visit	N/A
Change from baseline in bone assessment	Summary	Evaluable (FAS)	observed data at each visit	N/A
Percent change from baseline to Month 12 in spleen volume MN measured by ultrasound	Summary	Evaluable (FAS)	observed data at each visit	N/A
Percent change from baseline to Month 12 in liver volume MN measured by ultrasound	Summary	Evaluable (FAS)	observed data at each visit	N/A
Change from baseline in VAS fatigue score	Summary	Evaluable (FAS)	observed data at each visit	N/A
Tanner stage	Summary	Evaluable (FAS)	observed data at each visit	N/A
Change from baseline in total score from mSST	Summary	Evaluable (FAS)	observed data at each visit	N/A
Anti-taliglucerase antibodies	Summary	Evaluable (FAS)	observed data at each visit	N/A

Appendix 2. Data Derivation Details

Appendix 2.1. Definition and Use of Visit Windows in Reporting

The following table provides the windows categorizing dated assessments under nominal visits for the purpose of reporting.

Nominal Visit	Visit window for clinical-based assessment	Visit window for imaging-based assessment
Baseline	Day 1	date of the first MRI or ultrasound for organ volumes at Screening visit; date of the first assessment around Day 1 (Baseline visit) for other imaging-based measurement.
Month 3	12 weeks -7 days < Visit date[1] – Day 1 +1 <= 12 weeks + 7 days	NA
Month 6	26 weeks -7 days < Visit date – Day 1 +1 <= 26 weeks + 7 days	26 weeks -42 days < Assessment date[2] – Baseline assessment date +1 <= 26 weeks +42 days
Month 9	38 weeks -7 days < Visit date – Day 1 +1 <= 38 weeks + 7 days	NA
Month 12	52 weeks -7 days < Visit date – Day 1 +1 <= 52 weeks + 7 days	52 weeks -42 days < Assessment date[2] – Baseline assessment date +1 <= 52 weeks +42 days

[1] visit date is the date collected on the top of the pages for each visit.
[2] assessment date for organ volume (ie, spleen and liver) and skeletal assessment (eg, skeletal age and bone mineral density) are collected as the "Date of assessment " on the corresponding eCRF modules.

Handling Missing Dates

Missing visit dates will be handled according to Pfizer CDISC data standard for imputation. Missing day in assessment date will be imputed as the first day of each month. If the month is missing for assessment date, the visit date collected at the same visit will be used.

Handling Multiple Observations in the Same Visit Window

If there are multiple observations in one participant falling into one single visit window, the worst assessment will be used for summary. The imaging-based assessment will less likely have multiple observations in the same visit window. The clinical lab-based test, such as hemoglobin and platelet, may have this situation. Since increase from baseline in hemoglobin and platelet indicating clinical improvement, then the least increase from baseline or most decrease from baseline will be considered as the worst assessment in handling multiple observations in the same visit window.

The definition of the worst assessment for some of the biomarkers or measurements may involve other considerations. For example, the worst assessment for bone age estimation needs to be compared to the chronological age. For tanner stage and testis volume, it will be assessed at individual case though lower categorical number or lower volume usually indicates as the worse assessment. In such cases and any parameters not specified above, the definition of the worst assessment will be decided case by case and will be documented in the programming document, such as analysis datasets specification.

Appendix 2.2. Endpoint Derivations

Not applicable.

Appendix 2.3. Definition of Protocol Deviations That Relate to Statistical Analyses/Populations

Not applicable.

Appendix 3. Data Set Descriptions

Not applicable.

Appendix 4. Statistical Methodology Details

1. Sample code for Wilcoxon signed rank test:

```
data one;
      input id $ pre post;
      label pre='Baseline' post='Mon12';
      cards;
      1     8.74  4.22
      2     6.77
      4     1.34  1.95
      5     4.52  3.76
      6     6.72  5.85
      7     2.37  2.74
      8     3.69  2.93
      9     4.55  3.62
      10    17.12 16.84
      11    14.93 18.07
;run;
data two;
      set one;
      pctch=((post-pre)/pre)*100;
      adjpctch=pctch + 20;  ** for primary efficacy endpoint **;
run;

** Wilcoxon signed rank test for primary efficacy endpoint with decision
criteria;
proc univariate data=two;
   var adjpctch;
   output out=rankpct SIGNRANK=statistics PROBS=p median=median;
run;

** Wilcoxon signed rank test for secondary efficacy endpoints;
proc univariate data=two;
   var pctch;
   output out=rankpct SIGNRANK=statistics PROBS=p median=median;
run;
```

2. Sample code for Clopper-Pearson exact test and CI for the responder rate:

```
data three;
      input status $ count;
      cards;
      respond 8
      nonresp 1;
run;
proc freq data=three;
      table status/binomial (exact p=0.5) alpha=0.05;
      weight count;
run;
```

Appendix 5. List of Abbreviations

Abbreviation	Term
Abs	absolute
AE	adverse event
TEAE	Treatment emergent adverse event
BOCF	baseline observation carried forward
BP	blood pressure
CDISC	Clinical Data Interchange Standards Consortium
CDC	Center of Disease Control and Prevention
CI	confidence interval
CRF	case report form
CSR	clinical study report
ECG	electrocardiogram
FAS	full analysis set
FDA	Food and Drug Administration (United States)
ICD	informed consent document
ICH	International Council for Harmonisation
Lyso-Gb1	Lyso-glucosylsphingosine
LOCF	last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
MN	Multiples of Normal
MRI	Magnetic Resonance Imaging
mSST	modified Severity Scoring Tool
N/A	not applicable
PP	per-protocol
PPAS	per-protocol analysis set
PT	preferred term
QTc	corrected QT
QTcF	corrected QT (Fridericia method)
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SOP	standard operating procedure
SUSAR	suspected unexpected serious adverse reaction
TA	therapeutic area
ULN	upper limit of normal
VAS	visual analog scale
WHO	World Health Organization
WHODD	World Health Organization Drug Dictionary