Effects of N-Acetyl-L-Leucine on GM2 Gangliosidosis (Tay-Sachs and Sandhoff Disease): A multinational, multicenter, open-label, rater-blinded Phase II study

NCT Number: NCT03759665

Version and date of SAP: Final SAP 4.0, 08 Mar 2023

STATISTICAL ANALYSIS PLAN

(Short) study title:	Effects of N-Acetyl-L-Leucine on GM2 Gangliosidosis (Tay-Sachs and Sandhoff Disease): A multinational, multicenter, open-label, rater-blinded Phase II study.
Name of the sponsor:	IntraBio Ltd.
Protocol identification:	IB1001-202, version 7.0 (30Nov2022)
Version and date of SAP:	Final SAP 4.0, 08 Mar 2023

VERSION HISTORY

Version	Date	History list
1.0	24 July 2019	Final version.
1.1	02 July 2020	Updated version for sponsor review: new protocol version, including Extension phase, updated CRF, addition of ITT analysis population, textual adaptations and clarifications, Covid-19 implications.
1.2	05 November 2020	Updated version for sponsor review: new protocol version with additional Extension phase II. Replacement of Bland-Altman with ICC.
1.3	25 November 2020	Adjustments after review comments
2.0	03 Feb 2021	Final version for signature.
3.0	19 May 2021	Final version: Correction of an inconsistency in the statistical analysis of CI-S.
4.0	08 Mar 2023	Updated version because of amended protocol for extension phase. Removal of ASIS analysis for extension phase.

APPROVAL PAGE

I hereby declare that I have read and reviewed this document. To the best of my knowledge, the content accurately states the intended analyses and output to be provided. This document is intended for an agreement on analysis and reporting details between the sponsor and [redacted]

(Lead) Statistician:

[redacted]

[redacted]

21 March 2023
Date

Sponsor representative:

[redacted] Director
IntraBio Ltd.

[redacted]

Signature

20 March 2023
Date

Sponsor Biostatistician:

[redacted]

Signature

20 March 2023
Date

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

_avg	Average
_recipr	Reciprocal
8MWT	8-Meter Walking Test
9HPT-D	9 Hole Peg Test of the Dominant Hand
9HPT-ND	9 Hole Peg Test of the Non-Dominant Hand
ADaM	Analysis Data Model
AE	Adverse Event
ALT	Alanine Aminotransferase
AST	Aspartate Transaminase
ATC	Anatomic Therapeutic Chemical
BDRM	Blinded Data Review Meeting
BL	Baseline
CFB	Change from Baseline
CGI-C	Clinical Global Impression of Change
CGI-I	Clinical Global Impression of Improvement
CGI-S	Clinical Global Impression of Severity
CI	Confidence Interval
CI-CS	Clinical Impression of Change in Severity
CI-S	Clinical Impression of Severity
COVID-19	Coronavirus Disease 2019
CRF	Case Report Form
EMA	European Medicines Agency
EU	European Union
FDA	Food and Drug Administration
FSH	Follicle Stimulating Hormone
GCP	Good Clinical Practice
ICC	Intraclass Correlation Coefficient
ICH	International Council on Harmonisation
LoA	Limits of Agreement
LOCF	Last Observation Carried Forward
M(C)AR	Missing (Completely) at Random
MedDRA	Medical Dictionary for Regulatory Activities
mDRS	Modified Disability Rating Scale
mITT	Modified Intention-to-Treat

PI	Principal Investigator
PK	Pharmacokinetics
PM	Project Manager
PPS	Per Protocol Set
PT	Preferred Term
RBC	Red Blood Cell Count
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SARA	Scale for the Assessment and Rating of Ataxia
SCAFI	Spinocerebellar Ataxia Functional Index
SD	Standard Deviation
SDTM	Study Data Tabulation Model
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
US	United States
VAS	Visual Analogue Scale
WBC	White Blood Cell Count
WHO	World Health Organization

1 GENERAL

This Statistical Analysis Plan (SAP) describes in detail the methods and presentation of the data analyses which will be conducted by [redacted] for study IB1001-202. This plan is written in agreement with protocol version 7.0, dated 30 November 2023 and annotated Case Report Form (CRF), version 3.4, dated 11 June 2021, and the relevant Good Clinical Practice International Council on Harmonisation (GCP-ICH) guidelines. Furthermore, sponsor requirements for reporting will be considered. Additional changes or updates of those documents or requirements may result in a new version of the reporting/statistical analysis plan. This plan is to be finalized preferably prior to enrolment of the first patient in the study, but at least before first programming/data analysis.

2 STUDY INFORMATION

2.1 Study Objectives

2.1.1 Parent study

Primary Objective

The primary objective is to evaluate the efficacy of N-Acetyl-L-Leucine based on blinded raters' clinical impression of change in severity (CI-CS) in the treatment of GM2 Gangliosidosis.

Secondary Objectives

The secondary objectives are:

- To assess the clinical efficacy of N-Acetyl-L-Leucine on symptoms of ataxia, functioning, and quality of life for patients with GM2 Gangliosidosis.
- To evaluate the safety and tolerability of N-Acetyl-L-Leucine at 4 g/day in patients with GM2 Gangliosidosis, including patients aged ≥18 years in the United States and patients aged ≥13 years in Europe, and weight-tiered doses in patients 6 to 12 years of age in Europe.

Exploratory Objectives

The exploratory objective is:

- To characterize the pharmacokinetics (PK) of N-Acetyl-L-Leucine in patients with GM2 Gangliosidosis.

2.1.2 Extension phase

Primary Objective

The primary objective in the Extension Phase is to evaluate the efficacy of N-Acetyl-L-Leucine based on the Modified Disability Rating Scale (mDRS).

Secondary Objectives

The secondary objectives are:

- To evaluate the long-term safety and tolerability of N-Acetyl-L-Leucine at 4 g/day in patients aged ≥13 years, and weight-tiered doses in patients 6 to 12 years of age, with GM2 Gangliosidosis
- To characterize the pharmacokinetics (PK) of N-Acetyl-L-Leucine in patients with GM2 Gangliosidosis.

Exploratory Objectives

The exploratory objectives are:

- To assess the clinical efficacy of long-term treatment with N-Acetyl-L-Leucine on symptoms, functioning, and quality of life for patients with GM2 Gangliosidosis
- To evaluate the effects of a 42-day (+14 days) washout from N-Acetyl-L-Leucine after one-year treatment based on blinded raters' Clinical Impression of Change in Severity (CI-CS)
- To assess the effects of a 42-day (+14 days) washout from N-Acetyl-L-Leucine after one-year treatment on symptoms, functioning, and quality of life for patients with GM2 Gangliosidosis.
- To evaluate the long-term clinical efficacy, safety and tolerability of N-Acetyl-L-leucine for patients with GM2 Gangliosidosis

2.2 Design of the Study

2.2.1 Parent study

This is a multinational, multicenter, open-label, rater-blinded Phase II study investigating the efficacy and safety of N-Acetyl-L-Leucine for the treatment of GM2 Gangliosidosis. Approximately 39 male and female patients aged ≥6 years in Europe or ≥18 years in the United States with a confirmed diagnosis of GM2 Gangliosidosis are planned to be enrolled.

The visit schedule differs between "naïve" (patients not previously exposed to prohibited medications within the 6 weeks (42 days) before the initial screening visit) and "non-naïve" patients: see figures 1 and 2 below.

Figure 1: Study Scheme for Naïve Patients

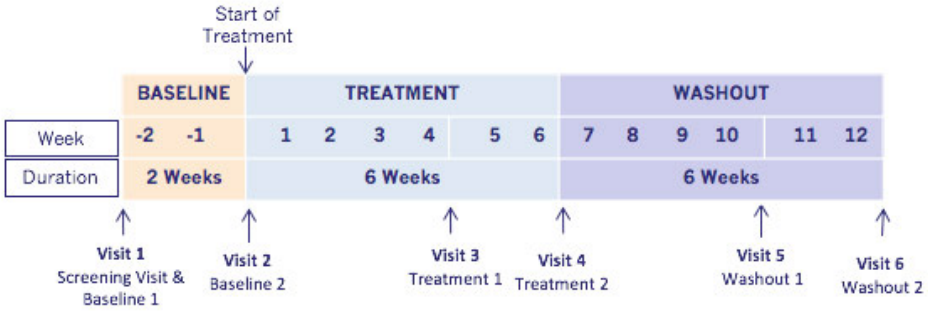
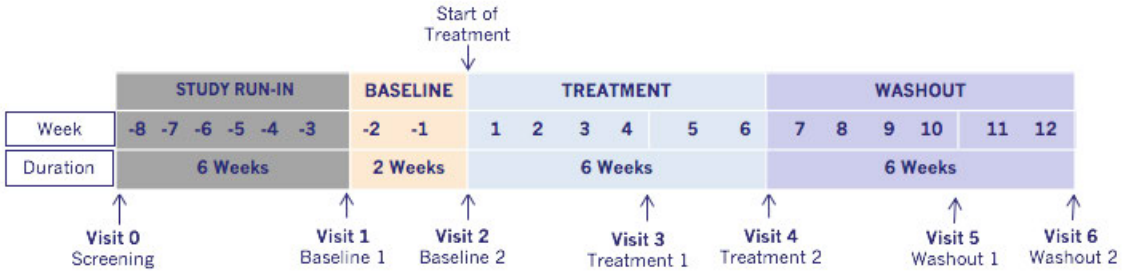


Figure 2: Study Scheme for Non-naïve Patients



2.2.2 Extension phase

The (full) Extension phase is considered the study period starting with the Extension phase baseline (Visit 7) until Visit 12. In some instances, however, it may be beneficial to make the following split:

- Extension phase I is considered the study period starting with the Extension phase Visit 7 until Visit 10.
- Extension phase II is considered the study period starting with the Extension phase Visit 10 until Visit 12.

The Extension phase will enroll patients who have completed Visit 6 of the Parent study phase of IB1001-202.

During the Extension phase, patients will be assessed 6 times over a 25,5-month period.

Baseline Visit of the Extension Phase

Visit 7 (Part A and B) is the baseline of the extension phase**.** The assessments (Part A and Part B) comprising Visit 7 may be conducted over a two-day period ("Part A" and "Part B"). If Visit 7B is scheduled 1 day (+6 days) after Visit 6 (Washout 2 of the Parent study), the assessments conducted at Visit 6 should count as Visit 7A (i.e. the assessments do not need to be repeated).

If Visit 7B cannot be scheduled 1 day (+6 days) after Visit 6, the assessments included in Visit 7A and Visit 7B should be performed over two consecutive days, i.e. Visit 7B occurs +1 day after Visit 7A.

In general Visit 7A will be the same as Visit 6 in the Parent phase, or within one week following Visit 6. But a non-seamless transition of several weeks or months is possible. The same subject numbers will be used throughout the Parent and the Extension phase.

Extension Phase Treatment Period I

From Visit 7B onward**,** all patients will receive treatment with N-Acetyl-L-Leucine for 365 days (+/-14 days).

Visit 8 will occur at approximately Day 180 (+/-14 days).

Visit 9 (Part A and B) will occur at approximately Day 365 (+/-14 days).

The assessments comprising Visit 9 may be conducted over a two-day period ("Part A" and "Part B").

Extension Phase Washout Period

After **Visit 9B**, all patients will enter a 42-day (+14 days) washout phase.

Visit 10 will occur at Day 407 (+ 14 days).

Extension Phase Treatment Period II

From **Visit 10** onward, all patients will receive treatment with N-Acetyl-L-Leucine for 365 days (+/-14 days).

Visit 11 will occur at Day 587 (+/-14 days).

Visit 12 will occur at Day 767 (+/-14 days).

2.3 Study medication

2.3.1 Parent study

All patients are to receive N-Acetyl-L-leucine (IB1001) in this single-arm study. The study drug will be taken during a 6-week (42 days + 7 days) treatment period. The patient's dose for the Parent phase will be determined based on their age and weight at Visit 1. During the treatment phase of this study, the dosing of the study drug is as follows:

- Patients aged ≥13 years in Europe and aged ≥18 years in the United States will take 4 g per day: 2 g in the morning, 1 g in the afternoon, and 1 g in the evening.
- Patients aged 6-12 years weighing 15 to <25 kg will take 2 g per day: 1 g in the morning and 1 g in the evening.
- Patients aged 6-12 years weighing 25 to <35 kg will take 3 g per day: 1 g in the morning, 1 g in the afternoon, and 1 g in the evening.
- Patients aged 6-12 years weighing ≥35 kg will take 4 g per day: 2 g in the morning, 1g in the afternoon and 1 g in the evening (as per patients aged ≥13)

2.3.2 Extension phase

All patients are to receive N-Acetyl-L-leucine (IB1001) during the Extension phase. The study drug will be taken for two 365-day (+/- 14 days) treatment periods (Extension phases I and II). The patient's dose for the Extension phase I will be determined based on their age and weight at Visit 7, following the same tiered doses as for the Parent study (see Section 2.3.1). The patient's dose for the Extension phase II will be determined based on their age and weight at Visit 10, following the same tiered doses as for the Parent study (see Section 2.3.1).

In the event a patient turns 13 years old or a patient aged 6-12 years old changes weight categories between Visit 7 and 8 or Visit 10 and 11 of the Extension Phase, their daily dose will be adjusted accordingly at Visit 8 or Visit 11 (respectively).

2.4 Sample size

2.4.1 Parent study

Under the original protocol, it was postulated that N-Acetyl-L-Leucine will show effectiveness in 30% of patients and this success rate is viewed as being clinically important. Assuming that this group of patients will have scores that are evenly distributed across the values 1 and 2 for the primary endpoint and further that the remaining 70% of patients will have scores that are evenly distributed between the values -1, 0, and 1, the resulting mean score is 0.45 and the standard deviation for the primary endpoint is 1.02. With 30 patients reporting on the primary endpoint, the study will have 76% power to detect a treatment benefit in a 5% one-sided one-sample t-test.

Recruitment will continue until approximately 30 patients complete dosing. The following numbers of patients are foreseen:

- To be enrolled: Approximately 39 patients
- To be analyzed according to the modified Intention to Treat analysis (mITT): Approximately 36 patients
- To be analyzed according to the Per Protocol Set (PPS): Approximately 30 patients

2.4.2 Extension phase

The sample size is determined by the number of patients that move into this Extension phase from the Parent study. The expectation is that between 60% and 70% of patients will rollover into the extension period. With 20 out of 32 patients (62.5%) rolling over, the study will have 80% power to detect an improvement in the success rate to 31% compared to the null hypothesis value of 10%.

Patients who completed Visit 10 of the Extension phase and are willing to continue for an additional year will continue in the Extension Phase until Visit 12.

2.5 Study flow chart

2.5.1 Parent study

Schedule of Events for "Naïve" Patients

Period	Baseline Period		Treatment Period		Wash-Out Period		Early Term.
Duration of the whole period	1 Day	2 Weeks	6 Weeks		6 Weeks		1 Day
Visit number	Visit 1[1]	Visit 2[2]	Visit 3	Visit 4	Visit 5	Visit 6 / EOS	ET
Name of the Visit	Screening/Bsl 1	Baseline 2	Treatment 1	Treatment 2	Washout 1	Washout 2	ET
Timeline (Days)	*Day -14*	*Day 1, Start IMP*	*Day 28*	*Day 42*	*Day 70*	*Day 84*	*XX*
Visit Window allowed	*na*	*+7 days*	*+7 days*	*+7 days*	*+7 days*	*+ 7days*	*na*

Patient information and informed consent process	X					X[19]	
Inclusion / exclusion criteria	X	X				X[19]	
Patient weight and height measurements	X						
Confirmation prohibited medications have not been used in the past 42 days	X	X[3,4]					
Classify patient as "Naïve" or "Non-naïve"	X						
Patient demographics (in accordance with local regulations)	X						
Relevant medical history	X						
60-Day drug history	X						
Documentation of therapy	X	X	X	X	X	X	X
Vital signs	X	X	X	X	X	X	X
12-lead electrocardiogram (ECG)[5]	X		X		X		X
Urine test for N-Acetyl-D-Leucine[6]	X[7]	X			X	X	X
Blood safety laboratory tests[8]	X	X	X	X	X	X	X
Follicle stimulating hormone serum [8, 9]	X						
Urinalysis [8]	X	X	X	X	X	X	X
Urine by dipstick for pregnancy test[10]		X		X		X	X

Blood sample for sparse PK [6]	X	X	X	X	X	X	X
Quality of Life EQ-5D-5L for patients aged ≥18; EQ-5D-Y[11] for children aged <18 years	X	X	X	X	X	X	X
Scale for Ataxia Rating (SARA)	X	X	X	X	X	X	X
Modified Disabling Rating Score (mDRS)	X	X	X	X	X	X	X
Scale for Spinocerebellar Ataxia Functional Index (SCAFI)[12]	X	X	X	X	X	X	X
Cognitive assessment according to standard procedures of the clinical site	X						
Determination of CI-CS Primary Anchor Test (9HPT-D or 8MWT)	X						
Clinical Global Impression of Severity (CGI-S) by Physician	X	X	X	X	X	X	X
Clinical Global Impression of Severity (CGI-S) by Caregiver	X	X	X	X	X	X	X
Clinical Global Impression of Severity (CGI-S) by Patient[13]	X	X	X	X	X	X	X
Clinical Global Impression of Change (CGI-C) by Physician				X		X	X
Clinical Global Impression of Change (CGI-C) by Caregiver				X		X	X
Clinical Global Impression of Change (CGI-C) by Patient[13]				X		X	X
Medication History: Confirm if N-Acetyl-Leucine ever used for "naïve" patients						X	X

Documentation of concomitant medication[14]	X	X	X	X	X	X	X
Documentation of AEs	X	X	X	X	X	X	X
Dispensing of study drug		X	X[15]				
Intake of study drug at site		X[16]					
Return of study drug			X	X			X
Study drug compliance check			X[17]	X[18]			X[18]

[1] If the patient's eligibility is confirmed and their urine screen for N-Acetyl-D-Leucine is below the permitted threshold. The next visit, **Visit 2**, should be planned 14 days (+7 days) from **Visit 1**.

[2] All assessments must be done pre-dose

[3] At **Visit 2**, confirm no prohibited medications since **Visit 1**

[4] If the patient (or caregiver) states the patient has been using prohibited medication at **Visit 2**, they will be classified as non-compliant and withdrawn from the study

[5] If feasible, repeat assessments performed at treating physician's discretion if clinically significant results at **Visit 3**

[6] To be analyzed at PK lab

[7] If the patient's urine sample unexpectedly detects levels of N-Acetyl-D-Leucine above the permitted threshold, they will (provided eligible) switch to "non-naïve"

[8] To be analyzed at the central lab

[9] Only for post-menopausal women of non-child bearing potential with amenorrhea for at least 1 year prior to the first dose (and have not undergone sterilization procedures at least 6 months prior to the first dose)

[10] Only for women of childbearing potential; done at site

[11] In Europe only

[12] Two subtests, the 9-hole Peg Test of the Dominant Hand (9HPT-D) and 8-meter Walk Test (8MWT) will be videoed in a standardized format at every visit (except Visits 0).

[13] If the patient is able to provide the CGI-S/C

[14] Any concomitant medication needs to be recorded, used or not used for (symptoms of) GM2 Gangliosidosis since the last 60 days prior to date of informed consent, up to End of Study / ET

[15] If needed

[16] Patient should not have used any of the prohibited concomitant medication 42 days prior to first dose of study drug, and have had a urine screen for N-Acetyl-D-Leucine below the permitted threshold prior to first dose of study drug.

[17] If feasible

[18] Or after the Visit if IMP is returned via courier

[19] If Investigator determines continued treatment with IB1001 in patient's best interest, and patient would like to participate in the optional Extension Phase, task performed for continuation into the Extension Phase

General note: due to COVID-19, deviations from the schedule of visits and assessments may be necessary, i.e. visits may be skipped, postponed, or performed remotely.

Schedule of Events for "Non-naïve" Patients

	Study Run-In		Baseline Period		Treatment Period		Wash-Out Period		Early Term.
	Screening visit	Pre-treatment Washout							
Duration of the whole period	**1 Day**	**6 Weeks**	**2 Weeks**		**6 Weeks**		**6 Weeks**		**1 Day**
Visit number	**Visit 0**		**Visit 1[1]**	**Visit 2[2]**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6 / EOS**	**ET**
Name of the Visit	**Screening**		**Baseline 1**	**Baseline 2/ Start IMP**	**Treatment 1**	**Treatment 2**	**Washout 1**	**Washout 2**	**ET**
Timeline (Days)	*Day -56*		*Day -14*	*Day 1*	*Day 28*	*Day 42*	*Day 70*	*Day 84*	*na*
Visit Window allowed	*na*		*na*	*+7 days*	*+7 days*	*+7 days*	*+7 days*	*+7 days*	*na*

Patient information and informed consent process	X								
Inclusion / exclusion criteria	X		X	X				X[19]	
Patient weight and height measurements	X							X[19]	
Confirmation prohibited medications have not been used in the past 42 days at Visit 1/since Visit 1 for Visit 2			X	X[3]					
Confirmation prohibited medications have been used within past 42 days[4]	X								
Classify patient as "Naïve" or "Non-naïve"	X								
Patient demographics (in accordance with local regulations)	X								
Relevant medical history	X								
60-Day drug history	X								
Documentation of therapy	X		X	X	X	X	X	X	X
Vital signs	X		X	X	X	X	X	X	X
12-lead electrocardiogram (ECG)[5]			X		X		X		X

Urine Test for N-Acetyl-D-Leucine[6]			X[7]	X			X	X	X
Blood safety laboratory tests[8]			X	X	X	X	X	X	X
Follicle stimulating hormone serum[9]			X						
Urinalysis[8]			X	X	X	X	X	X	X
Urine by dipstick for pregnancy test[10]				X		X		X	X

Blood sample for sparse PK[6]			X	X	X	X	X	X	X
Quality of Life EQ-5D-5L for patients aged ≥18 years; EQ-5D-Y[11] for children aged <18 years			X	X	X	X	X	X	X
Scale for Ataxia Rating (SARA)	X		X	X	X	X	X	X	X
Modified Disabling Rating Score (mDRS)			X	X	X	X	X	X	X
Scale for Spinocerebellar Ataxia Functional Index (SCAFI)[12]	X		X	X	X	X	X	X	X
Cognitive assessment according to standard procedures of the clinical site			X						
Determination of CI-CS Primary Anchor Test (9HPT-D or 8MWT)			X						

Clinical Global Impression of Severity (CGI-S) by Physician			X	X	X	X	X	X	X
Clinical Global Impression of Severity (CGI-S) by Caregiver			X	X	X	X	X	X	X
Clinical Global Impression of Severity (CGI-S) by Patient[13]			X	X	X	X	X	X	X
Clinical Global Impression of Change (CGI-C) by Physician						X		X	X
Clinical Global Impression of Change (CGI-C) by Caregiver						X		X	X
Clinical Global Impression of Change (CGI-C) by Patient[13]						X		X	X
Documentation of concomitant medication[14]	X		X	X	X	X	X	X	X
Documentation of AEs	X		X	X	X	X	X	X	X
Dispensing of study drug				X	X[15]				
Intake of study drug at site				X[16]					
Return of study drug					X	X			X
Study drug compliance check					X[17]	X[18]			X[18]

[1] If the patient's eligibility is confirmed and their urine screen for N-Acetyl-D-Leucine is below the permitted threshold. The next visit, **Visit 2**, should be planned 14 days (+7) days from Visit 1.

[2] All assessments must be done pre-dose

[3] At **Visit 2**, confirm no prohibited medications since **Visit 1**

[4] If the patient (or caregiver) states the patient has been using prohibited medication at **Visit 1** or **Visit 2**, they will be classified as non-compliant and withdrawn from the study

[5] If feasible, repeat assessments performed at treating physician's discretion if clinically significant results at **Visit 3**

[6] Analyzed at PK lab

[7] Patients whose Visit 1 urine sample detects limits of N-Acetyl-D-Leucine above the permitted threshold are classified as non-compliant and withdrawn from the study

[8] Analyzed at central lab

[9] Only for post-menopausal women of non-child bearing potential with amenorrhea for at least 1 year prior to the first dose (and have not undergone sterilization procedures at least 6 months prior to the first dose)

[10] Only for women of childbearing potential; done at site

[11] In Europe only

[12] At Visit 1 through Visit 6 (*not* Visit 0), two subtests, the 9-hole Peg Test of the Dominant Hand (9HPT-D) and 8-meter Walk Test (8MWT) will be videoed in a standardized format at every visit. From **Visit 2**, the selected anchor test (9HPT-D or 8MWT) is to be performed first, before the remaining SCAFI subsets.

[13] If the patient is able to provide the CGI-S/C

[14] Any concomitant medication needs to be recorded, used or not used for (symptoms of) GM2 Gangliosidosis since the last 60 days prior screening, up to End of Study / ET

[15] If needed

[16] Patient should not have used any of the prohibited concomitant medication 42 days prior to first dose of study drug, and have had a urine screen for N-Acetyl-D-Leucine below the permitted threshold prior to first dose of study drug.

[17] If feasible

[18] Or after the Visit if IMP is returned via courier

[19] If Investigator determines continued treatment with IB1001 in patient's best interest, and patient would like to participate in the optional Extension Phase, task performed for continuation into the Extension Phase

General note: due to COVID-19, deviations from the schedule of visits and assessments may be necessary, i.e. visits may be skipped, postponed, or performed remotely.

2.5.2 Extension phase

Period in the Extension Phase (EP)	EP Baseline Period		EP Treatment Period I			EP Wash-Out Period	EP Treatment Period II		EP Early Term.
Duration of the whole period	1 Day	1 Day	1 Year			6 Weeks	1 Year		1 Day
Visit number	Visit 7A[i]	Visit 7B	Visit 8	Visit 9A	Visit 9B	Visit 10	Visit 11	Visit 12 / EOS	ET
Name of the Visit	EP Screening/Bsl	EP Baseline	EP Treatment 1	EP Treatment 2	EP Treatment 2	EP Washout 1	EP Treatment 3	EP Treatment 4	EP ET
Timeline (Days)	*Day -1*	*Day 1, Start IMP*	*Day 180*	*Day 365*	*Day 366*	*Day 407*	*Day 587*	*Day 767*	*XX*
Visit Window allowed	*na*	*+6 day[ii]*	*+/- 14 days*	*+/- 14 days*	*na*	*+14 days*	*+/- 14 days*	*+/- 14 days*	*na*
Patient information and informed consent process	X					X			
Inclusion / exclusion criteria	X								
Patient Weight		X	X		X	X	X	X	X
Physical Examination		X	X		X	X	X	X	X
Documentation of concomitant medication	X		X		X	X	X	X	X
Documentation of frequency of therapy (hours per week)	X		X		X	X	X	X	X
Vital signs	X		X		X	X	X	X	X
12-lead electrocardiogram (ECG)				X					X
Blood safety laboratory tests[iii]	X		X		X	X	X	X	X
Blood samples for research purposes	X				X			X	X
PK Blood Sampling[iv]		X			X				
Urinalysis	X		X	X		X	X	X	X
Urine by dipstick for pregnancy test[v]	X		X	X		X	X	X	X
Urine test for N-Acetyl-D-Leucine	X					X			X
Quality of Life EQ-5D-5L for patients aged ≥18; EQ-5D-Y for children aged <18 years[vi]	X		X	X		X	X	X	X
Scale for Ataxia Rating (SARA)	X		X	X		X	X	X	X

Scale for Spinocerebellar Ataxia Functional Index (SCAFI)	X		X	X		X	X	X	X
Video-Recording Primary Anchor Tests (8MWT + 9HPT-D)	X[vii]		X	X		X	X	X	X
Modified Disability Rating Scale (mDRS)	X	X	X	X		X	X	X	X
Clinical Global Impression of Severity (CGI-S) by Physician	X	X	X	X		X	X	X	X
Clinical Global Impression of Severity (CGI-S) by Caregiver	X		X	X		X	X	X	X
Clinical Global Impression of Severity (CGI-S) by Patient	X		X	X		X	X	X	X
Clinical Global Impression of Change (CGI-C) by Physician			X	X		X	X	X	X
Clinical Global Impression of Change (CGI-C) by Caregiver			X	X		X	X	X	X
Clinical Global Impression of Change (CGI-C) by Patient			X	X		X	X	X	X
Documentation of AEs	X	X	X	X	X	X	X	X	X
Dispensing of study drug		X	X			X	X		
Intake of study drug at site		X			X				
Return of study drug[viii]			X		X		X	X	X
Study drug compliance check			X		X		X	X	X

Abbreviations: EP = Extension Phase; na = not applicable.

[i] Visit 7A may be taken from Visit 6, provided Visit 7B is scheduled within 1 day (+6 days) of Visit 6

[ii] if Visit 7B cannot be scheduled 1 day (+6 days) after Visit 6, Visit 7A and Visit 7B should be conducted over a two-day period on two consecutive days, i.e. Visit 7B occurs +1 day after Visit 7A

[iii] Analyzed at central lab

[iv] Analyzed at PK lab

[v] Only for women of childbearing potential; done at site

[vi] If a patient turns 18 over the course of the study, they should continue to use the EQ-5D-Y

[vii] If Visit 7A is based on Visit 6 assessments, the video recordings from Visit 6 should be used

[viii] Or after visit if IMP is returned via courier

[ix] If applicable

General note: due to COVID-19, deviations from the schedule of visits and assessments may be necessary. Any changes to the schedule of visits/assessments should be discussed between the PI and

3 SUBJECTS FOR ANALYSIS

3.1 Parent study: Analysis populations

3.1.1 Intention-to-Treat (ITT) population

The ITT population will consist of all patients who receive at least one dose of study drug (N-Acetyl-L-Leucine) and with one video recording at either Visit 1 or Visit 2 (or both).

3.1.2 Modified Intention-to-Treat (mITT) population

The mITT population will consist of all patients who receive at least one dose of study drug (N-Acetyl-L-Leucine) and with one video recording at either Visit 1 or Visit 2 (or both) and one video recording at either Visit 3 or Visit 4 (or both).

3.1.3 Per Protocol Set (PPS)

The PPS will consist of all patients with at least one video recording at baseline (either Visit 1 or Visit 2), end of treatment (either Visit 3 or Visit 4), and end of washout (either Visit 5 or Visit 6) and without any major protocol deviations that can influence the validity of the data for the primary efficacy variable.

3.1.4 Safety Analysis Set (SAF)

The Safety Analysis Set (SAF) will consist of all patients who receive at least one dose of study drug (N-Acetyl-L-Leucine).

3.2 Extension phase: Analysis populations

The Extension phase will only include patients who completed Visit 6 of the Parent study and fulfill the inclusion criteria of Extension phase. Patients who completed Visit 10 of the Extension phase and are willing to continue for an additional year will continue in the Extension Phase until Visit 12.

3.2.1 Modified Intention-to-Treat Extension phase (mITTe) population

The mITTe population will consist of all patients who receive at least one dose of study drug (N-Acetyl-L-Leucine) in Extension phase and with a mDRS score at Visit 7 and at either Visit 8 or Visit 9 (or both). This analysis population will be used for efficacy analyses throughout the Extension phase until Visit 12.

3.2.2 Safety Analysis Set Extension phase (SAFe)

The SAFe will consist of all patients who receive at least one dose of study drug (N-Acetyl-L-Leucine) in the Extension phase. This analysis population will be used for all safety analyses throughout the Extension phase until Visit 12. Only reported data will be used in the analysis.

3.3 Protocol deviations

Major protocol deviators will be identified to the extent possible by individuals responsible for data collection/compliance, and its analysis and interpretation. Any subjects or data values excluded from the PPS/PPSe analysis will be identified, along with their reason for exclusion.

3.4 Blinded Data Review Meeting

There will be a BDRM meeting held for the Parent study. No BDRM will be necessary for the Extension phase since no Per Protocol analysis population will be defined then.

The patients will be classified during the BDRM to the analysis populations based upon the protocol deviations. Invitees to this meeting are the Sponsor's Medical Expert, the Sponsor's representative, the Sponsor's Project Manager (PM), the Medical Monitor, the Sponsor statistician, and the [redacted] statistician, but more roles can be invited if considered necessary. Input to this meeting will be supplied by the involved Data Management provider or the responsible PM at least one week in advance of the meeting: a blinded list of all protocol deviations (including missing and likely erroneous data), with specific detailing and description regarding the deviation, preferably in Excel.

Possible compliance issues and unforeseen or reclassification of categories for missingness of the primary endpoint data will be part of the discussions held at the meeting.

Where possible, this meeting will be held in a blinded manner, that is blind to study outcome data (also for open studies). The goal of this meeting is to reach consensus on minor and major protocol deviations. In case of a major deviation impacting the primary or secondary endpoints, the specific patient will be excluded from the PPS population, either completely or from a specific time point onwards. The meeting must be held prior to database lock.

The decisions taken during this meeting and the reasons for those decisions will be documented by the PM (or a delegate as agreed) and sent for review to all parties involved as soon as possible after the meeting, but before database lock. If all parties involved agree, then the document is finalized, signed and stored in the TMF before database lock.

4 STUDY ENDPOINTS

4.1 Parent study

4.1.1 Primary endpoint

The primary efficacy endpoint is based on the blinded raters' Clinical Impression of Change in Severity (CI-CS) score on either the 9 Hole Peg Test of the Dominant Hand (9HPT-D) or the 8 Meter Walk Test (8MWT). The primary endpoint is defined as the CI-CS comparing the end of treatment (Visit 4) with baseline (Visit 2) **minus** the CI-CS comparing the end of washout (Visit 6) with the end of treatment (Visit 4).

4.1.2 Key secondary efficacy endpoints

Supportive secondary endpoints that directly supplement the analysis of the primary endpoint will be evaluated as follows:

- The individual components of the primary endpoint, that is the CI-CS from Visit 2 to Visit 4 and the CI-CS from Visit 4 to Visit 6.
- Difference in the blinded rater's Clinical Impression of Severity (CI-S) values from baseline (average of Visit 1 and Visit 2) to end of treatment (average of Visit 3 and Visit 4) and from end of treatment (average of Visit 3 and Visit 4) to end of washout (average of Visit 5 and Visit 6).

- Measurement of change in performance of the CI-CS score of either the 9HPT-D or the 8MWT reclassified to a 3-point scale.
- An evaluation of the CI-CS for the test (9HPT-D or 8MWT) that was not selected as the primary anchor test.

4.1.3 Additional secondary efficacy endpoints

For all additional secondary endpoints, the change from Visit 2 to Visit 4 and the change from Visit 4 to Visit 6 will be evaluated.

Measurement of Ataxia and Functioning

- Spinocerebellar Ataxia Functional Index (SCAFI).
  The SCAFI is composed of the 8MWT, the 9HPT, and the PATA rate, a measure of speech performance.
- Scale for Assessment and Rating of Ataxia (SARA) score.
  SARA has 8 items that are related to gait, stance, sitting, speech, finger-chase test, nose-finger test, fast alternating movements and heel-shin test.

Measurement of Health-related Quality of Life

- Quality of Life EQ-5D-5L for patients aged ≥18; EQ-5D-Y for children aged 6-17 years (two parts: the EQ visual analogue scale (EQ-VAS) and the EQ descriptive system).

Measurement of Overall Neurological status

- Modified Disability Rating Scale (mDRS)
  The scores of the six components (ambulation, language, manipulation, swallowing, seizures, and ocular movement) of this scale add up to a composite value.

Measurement of Global Impression

- Clinical Global Impression of Severity by the treating physician (CGI-S-physician) at every visit.
- Clinical Global Impression of Change by the treating physician (CGI-C-physician) comparing end of treatment (Visit 4) to baseline (Visit 2), and end of washout (Visit 6) to end of treatment (Visit 4).
- Clinical Global Impression of Severity by the caregiver (CGI-S-caregiver) at every visit.
- Clinical Global Impression of Change by the caregiver (CGI-C-caregiver) comparing Visit 4 to Visit 2 and Visit 6 to Visit 4.
- Clinical Global Impression of Severity by the patient (CGI-S-patient) at every visit *if they are able.*
- Clinical Global Impression of Change by the patient (CGI-C-patient) comparing Visit 4 to Visit 2 and Visit 6 to Visit 4 *if they are able.*

The CGI-S scales are on a 7-point scale ranging from ‘normal, not ill at all’ to ‘among the most ill patients’. The CGI-C scales are also on a 7-point scale ranging from ‘very much improved’ to ‘very much worse’.

4.1.4 Exploratory endpoints

Sparse PK sampling will be collected for biochemical analysis so to characterize the pharmacokinetics of N-Acetyl-L-Leucine in patients with GM2 Gangliosidosis.

(Note that analysis and reporting of pharmacokinetics is outside the scope of this analysis plan. PK data handling, analysis and reporting will be done by an external company that specializes in PK analysis and reporting. The PK report will be added to Appendix 16.1 of the CSR as a separate document.)

4.1.5 Safety endpoints

The safety parameters to be evaluated are:

- Adverse Events
- Laboratory safety measurements: haematology, clinical chemistry, urinalysis
- Vital signs
- Electrocardiography (ECG)

4.2 Extension phase

With respect to the Parent study, all endpoints for the Extension phases are considered secondary to the endpoints of the Parent study phase.

4.2.1 Primary endpoint

The primary endpoint in the Extension phase is "success" based on the mDRS score, with success defined as no change or improvement (decrease) in the mDRS score from Visit 7 to Visit 9.

4.2.2 Secondary efficacy endpoints

Supportive secondary endpoints for the Extension phase that directly supplement the analysis of the primary endpoint for the Extension phase will be evaluated as follows:

- Full PK sampling at Visit 7 and Visit 9 to characterize the pharmacokinetics of N-Acetyl-L-Leucine in patients with GM2 Gangliosidosis.
  (Note that analysis and reporting of pharmacokinetics is outside the scope of this analysis plan. PK data handling, analysis and reporting will be done by an external company that specializes in PK analysis and reporting. The PK report will be added to Appendix 16.1 of the CSR as a separate document.)

4.2.3 Exploratory Endpoints

Exploratory endpoints will investigate other measures of symptoms and quality of life from Visit 7 to Visit 9 for the following measures:

<u>Measurement of Ataxia and Functioning</u>

- Spinocerebellar Ataxia Functional Index (SCAFI).
  The SCAFI is composed of the 8MWT, the 9HPT, and the PATA rate, a measure of speech performance.
- Scale for Assessment and Rating of Ataxia (SARA) score.
  SARA has 8 items that are related to gait, stance, sitting, speech, finger-chase test, nose-finger test, fast alternating movements and heel-shin test.

<u>Measurement of Health-related Quality of Life</u>

- Quality of Life EQ-5D-5L for patients aged ≥18; EQ-5D-Y for children aged 6-17 years (two parts: the EQ visual analogue scale (EQ-VAS) and the EQ descriptive system).

Measurement of Global Impression

- Clinical Global Impression of Severity by the treating physician (CGI-S-physician).
- Clinical Global Impression of Change by the treating physician (CGI-C-physician).
- Clinical Global Impression of Severity by the caregiver (CGI-S-caregiver).
- Clinical Global Impression of Change by the caregiver (CGI-C-caregiver).
- Clinical Global Impression of Severity by the patient (CGI-S-patient) *if they are able.*
- Clinical Global Impression of Change by the patient (CGI-C-patient) *if they are able.*

The CGI-S scales are on a 7-point scale ranging from 'normal, not ill at all' to 'among the most ill patients'. The CGI-C scales are also on a 7-point scale ranging from 'very much improved' to 'very much worse'.

Videos of the primary and non-primary anchor tests are collected throughout the extension phase for the CI-CS. These videos may be analyzed in order to inform the development and validation of the CI-CS.

4.2.4 Additional exploratory endpoints

For all additional exploratory endpoints, the change from Visit 7 to Visit 10, Visit 9 to Visit 10 and Visit 10 to Visit 12 may also be evaluated. In addition, the change from Visit 1 to Visit 12 may also be investigated.

- SCAFI
- 9-HPT of dominant hand
- 8MWT
- SARA
- Health-related Quality of Life (EQ-5D-5L/Y)
- CGI-S by physician
- CGI-S by caregiver
- CGI-S by patient *if able*
- CGI-C by physician
- CGI-C by caregiver
- CGI-C by patient *if able*
- mDRS

4.2.5 Safety endpoints

The safety parameters to be evaluated are:

- Adverse Events
- Laboratory safety measurements: haematology, clinical chemistry, urinalysis
- Vital signs
- ECGs
- Physical Examination

5 PARENT STUDY: STATISTICAL ANALYSIS

5.1 General considerations

The Parent study is considered the study period between screening (Visit 0 or Visit 1) through Visit 6.

Raw data (in listings) will be presented in the same precision as received. Appropriate rounding will be performed for the following summary statistics, where applicable: mean, standard deviation (SD) and two-sided 90% confidence limits will be presented with at least one more decimal than the original data; minimum and maximum values will be presented with the same precision as the original data. In frequency tables, percentages will be presented with 1 decimal unless otherwise stated.

P-values will be presented with 3 decimals, and those smaller than 0.001 will be replaced by <0.001. One-sided p-values smaller than 0.05 will be considered statistically significant for the primary endpoint and indicative for other endpoints. No adjustment for multiple comparisons will be applied. Conclusions regarding treatment efficacy will not solely rely on detecting statistical significance but equal emphasis will be placed on the magnitude and clinical relevance of treatment differences as judged by the point estimates. All evaluations other than the primary endpoint will be considered exploratory in nature.

Descriptive statistics presented in general summary tables will be the number of non-missing observations (n), mean, SD, minimum, median and maximum for quantitative data. For categorical data, frequency counts and percentages will be determined.

In general, baseline is defined as the value measured prior to first study treatment administration (at Visit 2), unless otherwise specified. In case no value at Visit 2 is available, the value measured at Visit 1 will be used. A footnote may be added to those tables/listings presenting baseline value, with an explanation how this value was assessed.

If available in the database, data for screening failures will not be presented in summary tables, except for disposition, medical history and end-of-study displays. Data for screening failures will be listed as available.

A treatment emergent adverse event (TEAE) is defined as an adverse event that appears during or after study treatment and was absent before.

5.2 Adjudication of endpoint data

Each individual video (CI-S) or video pairing (CI-CS) will be read by two independent reviewers. For the CI-S assessments of the primary and non-primary anchor test, the independent raters will be given 6 videos of the patient's performance taken at Visit 1 to Visit 6. For the CI-CS assessments of the primary and non-primary anchor test, the raters will be given three pairs of videos and asked to score the CI-CS for each video pair. The video pairs will be labeled Video A to Video B, Video B to Video C, and Video C to Video A[1]. The videos will be presented in a random order, and the independent raters will be blinded to the timepoint corresponding to each video. The appropriate Likert scale score will be provided to each of the videos and pairs of videos.

[redacted] (2019). Video Review Charter. IntraBio Ltd. Internal Document.

For the CI-CS assessment, pairs of videos will be evaluated with the Likert scale defined as:

- Significantly Improved (Score = +3)
- Much Improved (Score = +2)
- Minimally Improved (Score = +1)
- No Observable Change (Score = 0)
- Minimally Worse (Score = -1)
- Much Worse (Score = -2)
- Significantly Worse (Score = -3)

If there is a scoring difference between the two blinded raters for a CI-CS assessment, the following rules will apply:

- If the difference between the blinded raters is 1, the two scores will be averaged, and the unrounded average value will be used for further analysis.
- If there is a scoring difference of >1 between the two blinded raters for a CI-CS assessment, then a third blinded rater will review the scores and will determine which score is more accurate (adjudication by consensus). The adjudicator's decision will be the final score for that video assessment and used for further analysis.

For the CI-S assessment, individual videos will be evaluated with the Likert scale defined as:

- Normal, not at all ill (Score = +3)
- Borderline ill (Score = +2)
- Mildly ill (Score = +1)
- Moderately ill (Score = 0)
- Markedly ill (Score = -1)
- Severely ill (Score = -2)
- Among the most extremely ill patients (Score = -3)

If there is any scoring difference between the two blinded raters for a CI-S assessment, the two scores will be averaged, and the unrounded average value will be used for further analysis.

Note that in averaging situations for the CI-CS and CI-S assessments, the scores may not fit the Likert scales as mentioned above. In case of average scoring present in the database, presentations will be on score only, without the additional description.

5.3 Missing data

Patients that withdraw from the study are replaced at the discretion of the sponsor. Data from withdrawal patients will be included in the analysis until their last assessment.

For handling missing data of the statistical analyses applied to the primary and key secondary endpoints, refer to the respective analysis sections.

For other endpoints being presented with descriptive statistics only, such as baseline/screening data, safety presentations, additional secondary endpoints and (additional) exploratory endpoints, no imputation will be performed, unless otherwise specified (in the respective analysis section or in Section 10.1). All these analyses will be performed on data available at the visit considered. In summary tables, the number of patients without missing data will be presented (per visit, if

applicable) unless otherwise specified. In calculations of percentages, subjects with missing data will not be considered in numerator or denominator unless otherwise specified.

Using the most conservative approach, missing/incomplete information related to AEs will be handled as listed below, to indicate an AE as being treatment emergent or presented having a certain intensity/causality. Note that the imputed dates or intensity/causality will not appear in data listings, but only in descriptive statistics tables (where applicable).

- In case of (partially) missing onset dates, the AE will be handled as follows:
  - If full start date is available, and on or after first dosing date, the AE is considered treatment emergent.
  - In case full stop date is available and prior to first dosing date, the AE is considered prior.
  - If the day part of the start date is missing:
    - The AE is considered treatment emergent if the month and year of the start date are the same or after the month and the year of the first dosing date.
  - If the day and month part of the start date are missing:
    - The AE is considered treatment emergent if the year of the start date is the same or after the year of the first dosing date.
  - In case the start date is completely missing:
    - If stop date is fully available and on or after the first dosing date, the AE is considered treatment emergent.
    - If the stop date is partially missing, but the month and year (or year alone in case of missing month) are after the month and year (or year alone) of the first dosing date, the AE is considered treatment emergent.
  - In case full start date and full stop date are missing, the AE is considered treatment emergent.

- In case intensity is missing for a certain TEAE, this will be regarded as severe.
- In case causality is missing for a certain TEAE, this will be regarded as related.
- In case seriousness is missing for a certain AE, this is discussed and addressed prior to database lock and unblinding in agreement with the sponsor and data management provider.

Regarding prior/concomitant medication, a similar approach will be followed for partially missing dates, as will be done for AEs as described above. If a comedication is started and stopped prior to the first date of study treatment (i.e. first dosing date), this is considered prior.

Based on the BDRM, certain values can be decided to be excluded from analyses. These values will be listed, but not included in descriptive statistics, plots or statistical analyses. This may be outlier data that are a result of unambiguous measurement errors.

5.4 Interim analysis

Per-patient data listings and summary tables/plots will be provided for the several DSMB meetings and as specified in the DSMB Charter.

There will be an analysis of the primary, secondary, and exploratory endpoints relating to the Parent phase once Visit 6 has been completed, and an interim CSR will be written.

5.5 Baseline characteristics

5.5.1 Inclusion/exclusion criteria

An individual patient listing of the deviations from inclusion and exclusion criteria will be presented.

5.5.2 Demographics

Demographic data at baseline will be summarized for all patients and by the 4 age/weight/dosing groups in tabulations.

Appropriate descriptive statistics for age, height, weight, BMI, ethnic group and sex will be given. The summary will be created for each analysis population separately. Additionally, demographic data will be listed.

5.5.3 Baseline characteristics

The following relevant patient baseline characteristics will be listed and summarized overall (frequency and percentage) and are considered key subgroups. The subgroups differentiated on a median value will use the median value as determined for the specific analysis population.

- Naïve versus non-naïve as determined at screening
- Age (paediatric versus adult)
- Gender (male versus female)
- Age/weight/dosing group
- Region (US versus EU)
- Selected Primary Anchor Test (9HPT-D or 8MWT)
- Tay-Sachs versus Sandhoff Patients
- SARA Subtest Gait: classification based on below (≤)/above the median value at Visit 1.
- Composite of SARA Subtests 1- 4 (Gait, Stance, Sitting, Speech): classification based on below (≤)/above the median value at Visit 1.
- Intra-Patient variability between SARA score at Visit 1 (Baseline 1) vs Visit 2 (Baseline 2): classification based on below (≤)/above the median value of the difference between the two visits.
- Intra-Patient variability between CI-S score (for primary anchor only) Visit 1 (Baseline 1) vs Visit 2 (Baseline 2): classification based on below (≤)/above the median value of the difference between the two visits.
- Age of diagnosis: Infantile (<2 years), Juvenile (2 to <15 years), Adolescent/Late Onset (≥ 15 years)
- Disease severity: classification based on below (≤)/ above the median SARA score at Visit 1.

5.5.4 Disposition

A summary table will be created, stating the number of patients per site, including site number, location/country, and investigator.

To present disposition, a summary of all included, randomised, treated and completed patients will be created. Number/percentage screened will be included as well (if available in the database). Additionally, in the same summary, the number/percentage of patients in the SAF, mITT and PPS will be presented. Furthermore, the frequency and percentage per reason for non-completion will be added.

Additionally, a listing displaying all disposition information (including reasons for withdrawal and reasons for exclusion from the analysis population, if available electronically) on a per patient level will be created. This listing will include follow-up time for each patient, calculated from date of first dose until end of study for the Parent phase.

5.5.5 Medical history

Medical history will be tabulated per System Organ Class (SOC) and Preferred term (PT) according to the Medical Dictionary for Regulatory Activities (MedDRA) terminology list and presented as number/percentage of patients in each SOC and PT for the SAF. SOC and PT will be presented in descending order of frequency of occurrence based on SOC and PT. These data will also be listed on a per patient level.

GM2 Gangliosidosis specific history will be summarized as appropriate, presenting confirmed diagnosis (yes/no) and time since diagnosis. For calculation of time since diagnosis, see Section 10.1.

GM2 Gangliosidosis specific history data as collected will also be listed on a per patient level.

5.5.6 Pregnancy data

Pregnancy test result data will be listed for females of childbearing potential only, on a per visit basis.

5.5.7 Other screening data

Other data collected at screening and/or baseline only, if available, will be listed. These screening/baseline data will be presented using all subjects included in the study, if applicable.

5.6 Statistical analysis efficacy endpoints

The primary analysis population is the mITT for the primary and all secondary endpoints. Analyses of these endpoints based on the PPS will also be undertaken to provide supplementary information on efficacy. The ITT population will be used only for baseline summary tables.

Analyses of the primary endpoint based on the mITT population will utilise a last observation carried forward (LOCF) approach for missing/unreadable videos at Visit 5 and Visit 6. For the primary endpoint CI-CS, this implies that the CI-CS value for Visit 4 to Visit 6 will be assigned the value 0 (stable) if both videos at Visit 5 and Visit 6 are unavailable.

Also for the key secondary endpoint, Change in Severity Score based on Averaged Clinical Impression of Severity (CI-S), it is assumed that if both videos are missing/unreadable for the baseline (Videos from Visit 1 or 2), treatment (Videos from Visit 3 or 4), or washout (Videos from Visit 5 or 6) period, the change from one period to another will be assigned the value 0 (stable).

A sensitivity analysis for missing data will be performed for the statistical analyses to compare patient's outcome on the primary outcome measure and their compliance throughout the study. The robustness of the primary analyses in the mITT analysis set will be assessed as follows:

- If either video for Visit 3 or Visit 4 is available but both videos for Visit 5 and Visit 6 are not available, then the value -1 (worsening) will be used instead of 0 for the CI-CS for Visit 4 to Visit 6 unless the missingness is due to a technical reason (MCAR, e.g. due to machine failure).

For each of the primary and secondary endpoints there will be separate descriptive statistical displays within the key subgroups as listed in Section 5.5.3 without any statistical analysis. For the subgroup presentations, only the mITT population will be used. Spaghetti plots will not be created for the subgroup analyses; these will be presented in forest plots (one for each endpoint, excluding the SARA -, SAFI - and mDRS subscores and excluding the EQ-5D-5L/Y), presenting the mean or (pseudo-)median with a 90% two-sided CI, using a stem at mean or median difference of 0, where applicable. Upon discretion of the programmer, the forest plots may be presented as two plots: one for Visit 4 versus Visit 2 and one for Visit 6 versus Visit 4.

Considering the number of subgroups, the forest plots may become unclear and indecipherable. When that occurs, a split will be made into two forest plots for the endpoints: one for the first 7 subgroups, and one for the remaining.

All individual primary and secondary endpoint data will be listed as well.

5.6.1 Primary endpoint

The primary efficacy endpoint is based on the blinded raters' Clinical Impression of Change in Severity (CI-CS) score over 6 weeks determined by comparing videos showing the patient's performance on a pre-defined anchor clinical symptom scale: either the 9 Hole Peg Test of the Dominant Hand (9HPT-D) or the 8 Meter Walk Test (8MWT).

For each patient, there will be two CI-CS scores. These two scores indicate the level of change in severity observed in Visit 4 (end of treatment) compared to Visit 2 (baseline) and the level of change in severity observed in Visit 6 (end of washout) compared to Visit 4 (end of treatment). Both CI-CS scores are measured on a 7-point Likert-point scale (-3 = *significantly worse* to +3 = *significantly improved*, see Section 5.2 for the precise definition) and the primary endpoint is defined as the difference between both CI-CS scores.

The primary endpoint will be summarized using descriptive statistics, including a frequency table, and frequency percentages will be presented graphically using a bar chart.
The primary statistical analysis will be performed using a one-sample t-test comparing the CI-CS difference score to zero. Statistical output will present the number of patients included in the analysis, both CI-CS scores, the mean of the difference, the corresponding two-sided 90% confidence interval, and a one-sided p-value.

The assumption of normality will be visually checked and if the assumption does not hold, a Wilcoxon signed-rank test will be performed instead of the t-test using the difference score. Statistical output will then present the number of patients included in the analysis, both CI-CS scores, the (pseudo-) median of the difference using the Hodges-Lehmann estimator, the corresponding two-sided 90% confidence interval, and a one-sided p-value.

5.6.2 Key secondary endpoints

All key secondary endpoints will be presented using the mITT and PPS populations.

Clinical Impression of Change in Severity

The CI-CS scores on the pre-defined anchor test for Visit 4 versus Visit 2, and the CI-CS score on the pre-defined anchor test for Visit 6 versus Visit 4 will be evaluated separately. These key secondary endpoints will be summarized using descriptive statistics and frequency tables, and percentages will be shown graphically using a bar chart. No sensitivity analysis will be applied.

Change in severity score based on Averaged Clinical Impression of Severity (CI-S)

CI-S scores are measured on a 7-point Likert-point scale (-3 = *among the most extremely ill patients* to +3 = *normal, not at all ill*; see Section 5.2 for the precise definition) and the primary endpoint is defined as the difference between both CI-CS scores.
Improvement in the pre-defined anchor test measure will be evaluated based on the change in the blinded raters' CI-S scores between the baseline period (average for Visit 1 and Visit 2) and end of treatment period (average for Visit 3 and Visit 4) minus the change in CI-S between end of treatment period (average for Visit 3 and Visit 4) and end of washout period (average for Visit 5 and Visit 6). Note that if a CI-S score is missing at one visit for a period, but available at the other visit in the same period, that one value will be used instead of the average. If both videos in a period are missing, then we cannot determine the average and the period will be assigned a value of 0 (stable). The presentations and statistical analyses will be similar to the primary endpoint, excluding the sensitivity analysis which is only to be used for the primary endpoint. Data received for non-anchor CI-S scores will only be listed.

Clinical Impression of Change in Severity reclassified on a 3-point scale

Using the CI-CS outcome on the pre-defined anchor test based on the data for Visits 2 and Visit 4, any patient given a score <0 on the CI-CS will be classified as worsened (-1). Any patient classified as 0 on the CI-CS will be classified no change (0). Any patient given a score >0 on the CI-CS will be classified as improved (+1). Similarly, the CI-CS comparing Visit 4 and Visit 6 will be reclassified and the differences between these two scores on a 3-point scale will be calculated.
The descriptive table and graphical presentations will be similar to the primary endpoint. No statistical analysis will be performed, and no sensitivity analysis will be applied.

Clinical Impression of Change in Severity (CI-CS) score non-primary anchor test

Descriptive statistics and frequency tables of the CI-CS scores for the test (9HPT-D or 8MWT) that was not selected as the primary anchor test will be presented for Visit 4 versus Visit 2, and separately, for Visit 6 versus Visit 4. No plots will be created, and no sensitivity analysis will be applied.

5.6.3 Additional secondary endpoints

All additional secondary endpoints will be presented using the mITT and PPS populations, and all individual data will be listed as well. No sensitivity analysis will be applied.

Measurement of Ataxia and Functioning

- SARA total score
- SCAFI total score
    - 9HPT-D

- 9HPT-ND
- 8MWT
- PATA test

Descriptive statistics will be used to summarize the endpoints above at each visit.

The subscores of the SARA score (gait, stance, sitting, speech disturbance, finger chase, nose-finger test, fast alternating hand movements, heel-shin slide) will also be summarized for each visit using descriptive statistics. Stacked bar charts will be provided per visit, based on the subscores of the total SARA score.

Spaghetti plots of individual patient SARA total scores and the SCAFI total scores versus time will be made as well as a mean or median display versus time, depending upon the distribution, and corresponding two-sided 90% CI at each visit.

The average SCAFI subscores (8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg) as determined according to Section 10.1, following the scoring manual, will also be summarized with descriptive statistics.

The changes in SARA total score, SCAFI total score, the SCAFI subscores (8MW_avg, 9HPT-D_avg, and PATA_avg) between Visit 2 and Visit 4, as well as the changes between Visit 4 and Visit 6 will be analysed using a one-sample t-test or a Wilcoxon Signed Rank test depending on the distribution of the data. The SCAFI subscore 9HPT-ND_avg will only be descriptively summarized in tabular format.

Measurement of Health-related Quality of Life

The results of EQ-5D-5L as well as the EQ-5D-Y will be combined into a 5-digit number presenting the health status of the subject. Frequency tables will be presented per visit for the 5 domains, as well as for the 5-digit number of the EQ-5D-5L and the EQ-5D-Y separately.
In addition, EQ-5D-5L and EQ-5D-Y domain percentages will be presented in separate bar charts per visit, with all visits combined into one plot (if possible).

The EQ-VAS score will be summarized by visit using descriptive statistics for patients aged < and ≥18 years combined. Spaghetti plots of individual patient EQ-VAS scores versus time will be made as well as a mean or median display versus time, depending upon the distribution, and corresponding two-sided 90% CI at each visit.

Measurement of Overall Neurological Status

The mDRS composite score and individual components (ambulation, language, manipulation, swallowing, seizures, and ocular movement) will be summarized using descriptive statistics at each visit. As graphical display, a stacked bar chart of the 6 mDRS subscores at each visit will be created. Spaghetti plots of individual patient mDRS composite scores versus time will be made as well as a mean or median display versus time, depending upon the distribution, and corresponding two-sided 90% CI at each visit.

Furthermore, changes in the mDRS composite score between Visit 2 and Visit 4 as well as between Visit 4 and Visit 6 will be descriptively presented in a summary table.

Measurement of Global Impression

The three CGI-S scores (scored by treating physician, caregiver, and patient) will be summarized using descriptive statistics and frequency tables for each visit.
Bar charts will be presented for the three CGI-S scores, spaghetti plots of individual CGI-S scores versus time will be made as well as a mean or median display versus time, depending upon the distribution, and corresponding two-sided 90% CI at each visit.

Descriptive statistics and frequency tables will also be used to summarize the three CGI-C scores (treating physician, caregiver and patient) between Visit 2 and Visit 4, and between Visit 4 and Visit 6, and corresponding bar charts will be created.

The CGI-C scores for Visit 4 versus Visit 2, and for Visit 6 versus Visit 4 will be evaluated using descriptive statistics. Similarly, the CGI-C from Visit 2 to Visit 4 minus the CGI-C from Visit 4 to Visit 6 will be presented descriptively. In addition, the CGI-C measures for treating physician will be analyzed similar to the primary endpoint: the CGI-C from Visit 2 to Visit 4 minus the CGI-C from Visit 4 to Visit 6 will be compared to a difference of zero using a one-sample t-test or Wilcoxon signed-rank test.

Note that CGI-C is referred to as CGI-I in the eCRF.

It is important to realize that the CGI-S and CGI-C scores per visit for each patient as provided by the caregiver may be based on varying caregivers and hence additional variation is introduced. No correction for this variation can be done and a footnote will be added to the table to describe this situation.

5.6.4 Other additional statistical analyses

The following analysis will provide an estimate of intra-patient variability and provides information on the stability of the outcome evaluation. Note that, considering the high degree of variability within and between these patients, this is only a simple approach as results may be influenced by e.g. rater variability, disease progression or treatment effect.

The analysis will be done for the mITT population only, and only for the CI-S outcome of the videos. No sensitivity analysis will be applied, and no separate analysis will be performed for subgroups.

The Intraclass Correlation Coefficient (ICC) will be used as an index of reliability to measure agreement between pairs of observations for Visit 1 and 2, Visit 3 and 4, and Visit 5 and 6 respectively, presenting the proportion of the total variance in the observations that is due to the differences between pairs. The ICC takes on values between 0 (meaning no agreement) to 1 (perfect agreement), with and ICC >0.80 representing high reliability. No imputation for missing visits will be done, and the ICC can only be determined for subjects having data for both visits in the pairs. The ICC value and the corresponding 95% confidence interval will be presented in a tabular format, and for additional clarity the number of available pairs will be added as well.

5.7 Safety and tolerability evaluation

5.7.1 Adverse events

A treatment emergent adverse event (TEAE) is defined as an adverse event that appears during or after study treatment and was absent before.

An AE overview table will be created displaying the number and percentage of subjects experiencing a treatment-emergent adverse event (TEAE) and the number of TEAEs for: any TEAE, any mild/moderate/severe TEAE, any related/unrelated TEAE, any serious AE, and any TEAE leading to study discontinuation.

Furthermore, all TEAEs will be tabulated by System Organ Class (SOC) and Preferred Terms (PTs) within each SOC according to the MedDRA terminology list, using frequency counts (number of subjects with at least one event, and number of events) and percentage of subjects with the event.

TEAEs will also be tabulated by severity (mild/moderate/severe) and by relationship to study medication (related/unrelated). Similar tables will be created for TEAEs leading to premature discontinuation, SAEs and deaths, if applicable. These summary tables will be presented by decreasing frequency of occurrence based on SOC and PT.

In addition, similar TEAE tables (overall, by severity and by relationship, premature discontinuation, SAEs, and deaths) by SOC and PT will be presented per treatment cycle, and frequencies and percentages will be determined for the active treatment cycle and the washout cycle. The treatment cycle is the period between start of first treatment at Visit 2 and 24 hours after last dose at Visit 4. The washout cycle starts 24 hours after last dose at Visit 4 until and including Visit 6.

The summary tables will be accompanied by individual subject listings of *all* AEs including information on AE number, actual AE description, date/time of start and end of AE (or ongoing), PT (MedDRA), SOC (MedDRA), severity, relationship, pre-dose (yes/no), seriousness, action taken, outcome and other information collected in the CRF for adverse events. Pre-dose AEs are not considered to be treatment-emergent, except in case of worsening during/after study treatment (to be collected as separate AE in the database). AEs starting prior to administration of the study drug will only be listed. In this listing, a clear distinction will be made between prior and treatment emergent events.

To address the study delays due to Covid-19 in reference to AE occurrences, AE incidence rates will also be determined per study cycle (treatment cycle and washout cycle). Incidence rate of adverse events is defined as the number of adverse event cases per patient-year(s). The calculation of the incidence rate is the number of cases (i.e. patient with an AE) divided by the sum of the follow-up time of all patient at risk (which will be all subjects receiving treatment). This calculation is accomplished as follows:

- Denominator: $\sum p_i \times t_i$, where p_i = patient i and t_i= treatment duration of patient i in years (total # days divided by 365.25), where treatment duration is determined based on first and last (dose) date.
- Numerator: the total number of cases (note: a single patient can be counted more than once if the AE occurs on separate occasions)
- Incidence rate=numerator/denominator

The use of this measure implies the assumption that the incidence rate is constant over different periods of time. Since the calculation is done per cycle, the denominator (patient-years) is also determined per cycle, based on the patients available within that cycle.

Separate listings will be created for SAEs and deaths, if applicable. The SAF population is used for the AE presentations.

5.7.2 Clinical laboratory

The following laboratory safety data are collected for this study:

Haematology	Chemistry	Urinalysis
Hemoglobin	Sodium	Leukocytes
Erythrocytes	LDH(Lactate dehydrogenase)	Nitrite
Hematocrit	Potassium	Urobilinogen
Thrombocytes	Creatinine	Protein
Leukocytes	Serum bilirubin level	pH
	AST (Aspartate aminotransferase)	Occult blood (erythrocytes, leucocytes)
	ALT (Alanine aminotransferase)	Specific gravity
	Urea	Ketones
	ALP (Alkaline phosphatase)	Bilirubin
	FSH (Follicle-stimulating hormone) (for postmenopausal women only)	Glucose

Laboratory safety data for haematology, biochemistry and urinalysis will be summarized using descriptive statistics and listed per visit, using protocol visits. Change from baseline will be calculated and presented as well for quantitative data, using the same summary statistics. If applicable, laboratory safety data collected as additional/unscheduled assessments (i.e. apart from those per protocol) will only be listed and will not be used in summary statistics. Baseline is defined as the Visit 2 measurement. In case the Visit 2 measurement is missing, the Visit 1 measurement will be used.

For laboratory safety data, all recorded and determined laboratory safety data will be listed, including information on the reference ranges, if available. In addition, information regarding age at screening and sex will be added to the listing.

All safety laboratory parameters will be presented in the tables and listings in the same standard units as supplied, which will be SI units.

Lastly, for clinical laboratory parameters, a listing will be created presenting all data that are out of reference range on a per-patient level, including any available unscheduled measurements. Note that clinical significance for out-of-range values is not collected per eCRF and can therefore not be presented. Information regarding age at screening and sex will be added to this listing. The SAF population is used for the presentations.

5.7.3 Vital Signs

Vital sign data consist of measurements for pulse rate and systolic/diastolic blood pressure (sitting). Vital signs will be summarized and listed per visit (protocol visits). Change from baseline will be calculated and presented as well, using the same summary statistics. Visit 2 is considered the baseline visit for vital signs; in case Visit 2 is missing, the Visit 1 measurement will be used as baseline instead. If applicable, vital sign measurements collected as additional/unscheduled assessments (i.e. apart from those per protocol) will only be listed and will not be used in summary statistics. The SAF population is used for the presentations.

5.7.4 Electrocardiography (ECG)

ECG outcome results (corrected QT interval and normal/abnormal result) will be presented descriptively per protocol visit. If applicable, ECG measurements collected as additional/unscheduled assessments (i.e. apart from those per protocol) will only be listed and will not be used in summary statistics. The SAF population is used for the presentations.

5.7.5 Prior and concomitant medication/therapies

The use of prior and concomitant medication/therapies will be listed for all subjects: included will be the preferred term, World Health Organization (WHO) coding information and the Anatomic Therapeutic Chemical (ATC) class code), dose, route of administration, start and stop date, frequency and reason for administration, as well as information if given for an AE. Differentiation will be made between prior and concomitant medication/therapies, by creating two separate listings. A frequency table per ATC class code and WHO drug code will be created if considered relevant, separately for prior and concomitant medication/therapies. The SAF population is used for the presentation.

5.8 Scheduled visits, dosing, and treatment compliance

5.8.1 Visit dates

A listing with actual visit dates (and times, if applicable) per patient will be presented.

5.8.2 Dosing and treatment compliance

Relevant dosing information (first dosing date, last dosing date), scheduled and actual dosing dates/times and treatment compliance information will be determined and listed for each patient. Summary statistics for the number of N-Acetyl-L-Leucine doses taken, based on bottles of dispensed and returned unused bottle counts, will be calculated for Visit 3 and Visit 4 if this information is collected on a per-visit basis. Compliance is defined as 100% * (total number of IB1001 bottles and/or sachets dispensed – total number of IB1001 unused bottles and/or sachets returned), divided by the number of bottles and/or sachets which should have been used. Note that the number of bottles the patients should have used depends on the dosage they should take (see Section 2.3 for reference) and the duration of the treatment period. Considering the fact that the patient's total daily dose may be reduced by up to half of their assigned dose at the discretion of the investigator, the actual dose will be used to determine the correct compliance calculation, if collected in the eCRF. The proportion of patients who take at least 80% of the prescribed medication will also be shown. Total dosing period is defined as the duration (in days) between first dose date and last dose date.

6 EXTENSION PHASE: STATISTICAL ANALYSIS

6.1 General considerations

The Extension phase is considered the study period starting with the Extension phase baseline (Visit 7) until Visit 12. In principle, data presentations will be created for the full Extension phase. Only where applicable and relevant, a separation will be made as follows:

- Extension phase I: Extension phase Visit 7 until Visit 10.
- Extension phase II: Extension phase Visit 10 until Visit 12.

All evaluations will be exploratory in nature, and all endpoints for the complete Extension phase are considered secondary to the endpoints of the Parent study. Statistical analysis will only be applied for some of the endpoints as defined for Extension phase I, no statistical analyses will be done for the endpoints in Extension phase II.

Summary presentations and listings will be presented and only reported data will be presented (i.e. no imputation or LOCF will be done) for data collected during Extension phase II. If not otherwise defined, data will we presented as missing where applicable.

Raw data (in listings) will be presented in the same precision as received. Appropriate rounding will be performed for the following summary statistics, where applicable: mean, standard deviation (SD) and two-sided 90% confidence limits will be presented with at least one more decimal than the original data; minimum and maximum values will be presented with the same precision as the original data. In frequency tables, percentages will be presented with 1 decimal unless otherwise stated.

P-values will be presented with 3 decimals, and those smaller than 0.001 will be replaced by <0.001. One-sided p-values smaller than 0.05 will be considered statistically significant for the primary endpoint and indicative for other endpoints. No adjustment for multiple comparisons will be applied. Conclusions regarding treatment efficacy will not solely rely on detecting statistical significance but equal emphasis will be placed on the magnitude and clinical relevance of treatment differences as judged by the point estimates.

Descriptive statistics presented in general summary tables will be the number of non-missing observations (n), mean, SD, minimum, median and maximum for quantitative data. For categorical data, frequency counts and percentages will be determined.

For the full Extension phase, the measurement at Visit 7 is considered the baseline. In case of missing Visit 7A data for safety measurements (e.g. for safety lab), and Visit 7B has occurred within 1 (+6) days of Visit 6, the Visit 6 data from the Parent study should be used instead (data are not being re-entered in the database): see also Section 2.2.2 for clarification.

No screen failures will be present in the data.

For the full Extension phase, a treatment emergent adverse event (TEAE) is defined as an adverse event that appears during or after the start of study treatment in the Extension phase at Visit 7B and was absent before. As a consequence, adverse events occurring during the washout phase between V9 and V10 will also be considered treatment emergent.

6.2 Adjudication of endpoint data

Videos of the primary and non-primary anchor tests are collected throughout the extension phase for the CI-CS. These videos may be analyzed in order to inform the development and validation of the CI-CS but will not be reported in the CSR (and hence not listed or tabulated). Data received will be included as-is in the CDISC files. In case of analysis, the following process applies: For the CI-CS assessments of both the primary and non-primary anchor tests, after a patient completes Visit 10 of the Extension phase, the independent raters will be given 3 video pairs of from Visit 7 to Visit 9, Visit 9 to Visit 10, and Visit 7 to Visit 10. After a patient completes Visit 12 of the Extension phase, the independent raters will be given a video pair from Visit 10 to Visit 12 (the raters may also assess additional video pairs, i.e. Visit 1 to Visit 12, on an exploratory basis). After a patient completes Visit 12 of the Extension phase, the independent

raters will be given a video pair from Visit 10 to Visit 12 (the raters may also assess additional video pairs, i.e. Visit 1 to Visit 12, on an exploratory basis). The videos will be presented in a random order, and the independent raters will be blinded to the timepoint corresponding to each video. The appropriate Likert scale score will be provided to each of the videos and pairs of videos. For further details regarding scoring and adjudication, which will be similar to that of the Parent study, see Section 5.2.

6.3 Missing data

Data from withdrawal patients will be included in the analysis until their last assessment.

For handling missing data in Extension phase, similar approaches as described in Section 5.3 will be followed, with the alteration that "first dose date" is the first dosing during Extension phase I or II, respectively. For AE and concomitant medication this indicates that events started during the Parent study and continuing during Extension phase will be considered prior for the Extension phase.

6.4 Interim analysis

Per patient data listings and summary tables/plots will be provided for several DSMB meetings as specified in the DSMB Charter created for the Parent study.

The final CSR will be written upon the completion of the full Extension phase, and no interim analysis is performed during the Extension phase.

6.5 Baseline characteristics

6.5.1 Inclusion/exclusion criteria

An individual patient listing of the deviations from inclusion criteria will be presented. No exclusion criteria are collected for the Extension phase, and no separate inclusion is done for Extension phase II.

6.5.2 Demographics

Demographic data at baseline will be summarized for all patients and by the 4 age/weight/dosing groups (using Visit 7 information) in tabulations, even if there is a change in dose between Visit 7 and 8, or between Visit 8 and 10, or Visit 10 and 11 (see Section 2.3.2).

Appropriate descriptive statistics for age, height, weight, BMI, ethnic group and sex will be given. The summary will be created for each analysis population separately. Additionally, demographic data will be listed.

6.5.3 Baseline characteristics

The following relevant patient baseline characteristics will be listed and summarized overall (frequency and percentage) and are considered key subgroups, and will only be used for presentations of the primary endpoint in Extension phase I analyses:

- Naïve versus non-naïve, as determined at screening during the Parent study phase
- Age (paediatric versus adult) at Visit 7
- Gender (male versus female)

- Age/weight/dosing group as applicable at Visit 7
- Selected Primary Anchor Test (9HPT-D or 8MWT) as defined during the Parent study phase
- Tay-Sachs versus Sandhoff Patients
- Composite of SARA Subtests 1- 4 (Gait, Stance, Sitting, Speech): classification based on below(≤)/above the median value at Visit 7
- Age of diagnosis: Infantile (<2 years), Juvenile (2 to <15 years), Adolescent/Late Onset (≥ 15 years)
- Disease severity: classification based on below (≤)/ above the median SARA total score at Visit 1.

6.5.4 Disposition

A summary table will be created, stating the number of patients per site, including site number, location/country, and investigator.

To present disposition, a summary of all included, treated, and completed patients will be created for the Extension phase I and II separately, and for the full Extension phase . Additionally, in the same summary, the number/percentage of patients in the SAFe and mITTe will be presented. Furthermore, the frequency and percentage per reason for non-completion will be added.

A listing displaying all disposition information (including reasons for withdrawal and reasons for exclusion from the analysis population, if available electronically) on a per patient level will be created. This listing will include follow-up time for each patient, calculated from date of first dose during the Extension phase until end of Extension phase or withdrawal date.

6.5.5 Medical history

Note that only for patients not having a seamless transition from the Parent study to the Extension phase, additional medical history data may be collected during the intermediate time. Medical history will be tabulated per System Organ Class (SOC) and Preferred term (PT) according to the Medical Dictionary for Regulatory Activities (MedDRA) terminology list and presented as number/percentage of patients in each SOC and PT for the SAFe. SOC and PT will be presented in descending order of frequency of occurrence based on SOC and PT. These data will also be listed on a per patient level.

6.5.6 Pregnancy data

Pregnancy test result data will be listed for females of childbearing potential only, on a per visit basis.

6.5.7 Other screening data

Other data collected at screening and/or baseline only, if available, will be listed. These screening/baseline data will be presented using all subjects included in the Extension phase, if applicable.

6.6 Statistical analysis efficacy endpoints

The analysis for all efficacy endpoints in the Extension phase is the mITTe.

Analyses of the primary endpoint based on the mITTe population will utilise a last observation carried forward (LOCF) approach for missing mDRS at Visit 9. For the primary mDRS endpoint this implies that the measurement of Visit 8 will be used instead. Other endpoints will not be imputed, and missing data will be reported as is. No sensitivity analyses will be applied.

For the primary endpoint for the Extension phase there will be separate descriptive statistical displays within the key subgroups as listed in Section 6.5.3, without any statistical analysis. For the subgroup presentations, only the mITTe population will be used. No plots will be created. All individual primary endpoint data will be listed as well.

All efficacy endpoints for Extension phase Visit 10 to Visit 12 are considered exploratory, and therefore described under section 6.6.4.

6.6.1 Primary endpoint Extension phase

The primary efficacy endpoint is based on the Modified Disability Rating Scale (mDRS). The mDRS measures neurological status via clinical signs and symptoms in 6 domains (ambulation, language, manipulation, swallowing, seizures, and ocular movement). A higher mDRS score indicates a clinical worsening, while a lower mDRS score indicates a clinical improvement. A 0-point change in the mDRS score represents a stabilization of disease progression, which is also considered a significant clinical improvement.

The mDRS composite score and individual domains, including change values, will be summarized using descriptive statistics at each visit for Extension phase.

The primary endpoint in this Extension Phase is success measured on the mDRS score from the Extension Phase baseline (Visit 7) to the end of treatment in the Extension Phase (Visit 9). Success is defined as no change or a decrease in mDRS score from Visit 7 to Visit 9.

The primary endpoint will be summarized using a frequency table. The primary statistical analysis of comparing the proportion success with the null proportion of 0.10 will be done using a one-sided Fisher's Exact Test, expecting an increase in the success rate. The statistical result (one-sided p-value and exact 90% confidence interval[2]) will be added to the frequency table. This analysis is considered to be exploratory.

In addition, a spaghetti plot presenting the total mDRS scores of each patient through time will be created.

6.6.2 Key secondary endpoints Extension phase

As PK is handled by an external company, there are no analyses for secondary endpoints to add here.

6.6.3 Exploratory Endpoints

All exploratory endpoints for the Extension phase will be presented using the mITTe population, and all individual data will be listed as well. For all exploratory endpoints, the change from Visit

[2] The exact CI and the p-value using the Fisher's Exact test can contradict each other, especially in small sample sizes. Therefore, the exact CI will include a mid-p adjustment.

7 to Visit 9 will be evaluated. No statistical analysis other than descriptive statistics will be done. No plots will be created.

Measurement of Ataxia and Functioning

- SARA total score
- SCAFI total score
  - 9HPT-D
  - 9HPT-ND
  - 8MWT
  - PATA test

Descriptive statistics will be used to summarize the endpoints above at each visit.

The subscores of the SARA score (gait, stance, sitting, speech disturbance, finger chase, nose-finger test, fast alternating hand movements, heel-shin slide) will not be summarized separately, and only listed.

The average SCAFI subscores (8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg) as determined according to Section 10.1, following the scoring manual, will also be summarized with descriptive statistics.

The changes in SARA total score, SCAFI total score, and the SCAFI subscores (8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg) between Visit 7 and Visit 9, and Visit 10 and Visit 12, will be descriptively summarized in tabular format as well.

Measurement of Health-related Quality of Life

The results of EQ-5D-5L as well as the EQ-5D-Y will be combined into a 5-digit number presenting the health status of the subject. Frequency tables will be presented per visit for the 5 domains, as well as for the 5-digit number of the EQ-5D-5L and the EQ-5D-Y separately.

Note that if a patient turns 18 before Visit 7, this patient can have results on EQ-5D-Y in the Parent Study, and results on EQ-5D-5L in the Extension phase Visit 7 to 10. For this reason, separate QoL presentations will be created for Extension phase I, using Visit 7 measurements as baselines. The EQ-VAS score will be summarized by visit using descriptive statistics in a combined manner for patients aged < and ≥18 years (using Visit 7 age values).

Measurement of Global Impression

The three CGI-S scores (scored by treating physician, caregiver, and patient) will be summarized using descriptive statistics and frequency tables for each visit.

Descriptive statistics and frequency tables will also be used to summarize the three CGI-C scores (treating physician, caregiver, and patient) between Visit 7 and Visit 9.

Note that CGI-C is mentioned as CGI-I in the eCRF.

It is important to realize that the CGI-S and CGI-C scores per visit for each patient as provided by the caregiver may be based on varying caregivers and hence additional variation is introduced. No correction for this variation can be done and a footnote will be added to the table to describe this situation.

6.6.4 Additional exploratory endpoints

For all exploratory endpoints as described in 4.2.4, descriptive presentations will be made and the change from Visit 7 to Visit 10, Visit 9 to Visit 10, Visit 10 to Visit 12, and Visit 1 to Visit 12 will be evaluated descriptively. No plots will be created, and no statistical analysis done.

These presentations include the endpoints as mentioned above in sections 6.6.1, 6.6.2, and 6.6.3 that will be presented as exploratory using only descriptive statistics for the Extension phase on data collected after Visit 10. For EQ-5D-5L as well as the EQ-5D-Y, a different questionnaire could be used between Visits 10 and 12 if a patient turns 18 before Visit 10, and for this reason separate QoL presentations will be created for Extension phase II, using Visit 10 measurements as baselines. The EQ-VAS score will be summarized by visit using descriptive statistics in a combined manner for patients aged < and ≥18 years (using Visit 10 age values).

For presentations on efficacy data collected after Visit 10, no imputations or LOCF is considered necessary.

6.6.5 Other additional statistical analyses

No other additional statistical analyses are foreseen.

6.7 Safety and tolerability evaluation

The safety and tolerability evaluation for the extension study is similar to what is described for the Parent study (see Section 5.7), with the remark that the evaluation is limited to the Extension phase only and Visit 7 is considered the baseline. For the AE presentation per treatment cycle, the cycles for Extension phase are as follows:

- Treatment cycle Extension phase I: the period between start of first treatment at Visit 7 and 24 hours after last dose at Visit 9.
- Washout cycle Extension phase I: starts 24 hours after last dose at Visit 9 until and includes Visit 10.
- Treatment cycle Extension phase II: the period between start of first treatment at Visit 10 and 24 hours after last dose at Visit 12.
- No washout cycle is defined for Extension phase II.

Note that for those patients that do not transition seamlessly to the Extension phase, additional prior medication/adverse events may be collected at the start of the Extension phase, and these will be presented separately from the prior medication/adverse events of the Parent phase.

Physical examination, which is collected twice during the Extension phase, will be presented as follows: general physical examination data will be tabulated and listed per visit. The summary table will include number/percentage of patients with normal or abnormal observations and NCS/CS frequencies/percentages for abnormal observations per visit.

6.8 Scheduled visits, dosing, and treatment compliance

The presentation for the Extension phase is similar to what is described for the Parent study phase (see Section 5.8), with the remark that the evaluation is limited to the Extension phase only and Visit 7 is considered the baseline.

7 CHANGES TO PROTOCOL AND OTHER RELEVANT REMARKS

Instead of Bland-Altman analyses of outcome scores at Visit 1 vs Visit 2, Visit 3 vs Visit 4, and Visit 5 vs Visit 6 as described in the protocol for the Parent study, an ICC method for agreement

between the visits will be used instead to assess the stability of the outcome, as described in section 5.6.4. This is because no per-rater information is available in the dataset to be received, only an averaged or agreed value.

Regarding CI-S, the protocol is stating that for each assessment, the videos will be labeled as Video A, B, C, D, E, F. This is however not present in the DTA and test data from [redacted] and is therefore removed from the SAP.

For Quality-of-Life measures, since the scoring is done for children with the EQ-5D-Y and for adults with the EQ-5D-5L, tabulations and analyses will be done separately for the corresponding age groups. The presentations on the VAS will be combined for children and adults.

For the Parent phase, sparse PK sampling will be collected for biochemical analysis to characterize the pharmacokinetics of N-Acetyl-L-Leucine in patients with GM2 Gangliosidosis. For the Extension phase, full PK profiles will be collected at Visit 7 and Visit 9. Creation of descriptive displays and reporting of these data is outside the scope of this statistical analysis plan. PK concentration data and PK parameters (following a population PK analysis) will be included in SDTM domains, after receipt. For the Parent phase a listing of PK concentration data will be created for inclusion in the Parent phase CSR, since the full PK report will only be available after the end of the Extension phase. Furthermore, urine tests for N-Acetyl-D-Leucine are handled by the same lab but will be included in the clinical database and will therefore be part of the CDISC datasets.

Regarding the sensitivity analysis (imputation of value -1) for missing data in the primary and secondary analyses: this is only applied to the primary analysis of the Parent phase.
Regarding the LOCF/assumption of value 0 for missing data: this is only applied to the primary and one key secondary endpoint as specified in the relevant sections in this SAP.

The investigation of the intra-rater correlation cannot be performed, since the [redacted] data do not differentiate the data per rater but rather provide an average or agreed value. Therefore, this analysis has been removed from the SAP.

Considering the primary endpoint of the Extension phase is a different one then the primary endpoint of the Parent Study, the trial is called successful only when the primary endpoint of the Parent Study is positive. All analyses of the Extension phase on data collected until Visit 10 are considered secondary to ones of the Parent Study. All presentations in the Extension phase on data collected after Visit 10 are exploratory and descriptive.

The analysis on ASIS as described in the protocol has been removed from the SAP, as there are too limited datapoints available to perform the analysis.

For some patients, repeat visits were scheduled if measurements could not be fully handled during the agreed visit. Repeat visits will be listed for BDRM, and during BDRM or Clean File meeting it will be decided how the repeat data will be handled for analysis. Decisions will be documented prior to Database lock.

The addition of a short section of Covid-19 details (with reference to an addendum) has been added to the SAP and is not mentioned in the protocol.

8 COVID-19 DETAILS

This study is conducted during the COVID-19 outbreak. The following data capture decisions have been taken due to national, local, or site-specific restrictions imposed due to COVID-19, with the primary need to prioritize the safety of clinical study teams and patients, and to maintain the integrity of the study/data collection.

- All protocol deviations relating to COVID-19 are collected.
- If Visit 2 is performed 8 weeks (56 days) or more after Visit 1, some Visit 1 assessments must be repeated.
- Visits 3 and 5 can be converted to remote visits (i.e. phone, video conference).
- Visits 4 and 6 can be postponed until the in-person visit is feasible and should be performed in person. Where applicable, remote data collection should be collected on the date of the original visit. The postponed Visits 4 and 6 will collect all data per protocol.
- For the Extension Phase, no remote visits are allowed.

Currently, the impact of COVID-19 on the proposed statistical analyses is unknown. The statistical analysis already allows some flexibility for handling missing data, however, specific items/changes related to COVID-19 are being detailed in a separate document (Addendum 1), which will be maintained as a living document throughout the study continuation and adjusted based on the duration and extent of the impact of the COVID-19 situation.

9 DATA RECEIPT

All clinical eCRF data received from the Data Management provider per transfer agreement will be mapped to SDTM files. The information regarding minor/major violations will be received as excel file (a so-called “protocol deviations log”) and will be transferred to SDTM as well. The blinded results on the video-recordings (blinded rater’s CI-CS and CI-S) will be received from [redacted] as provider following the transfer agreement and will be mapped to SDTM. The received data will include relevant categorizations on missing data.

The SDTM files will be recoded to ADaM format, where necessary. Listings will be programmed on the SDTM and ADaM datasets, statistical analyses, tables, and figures will be programmed on ADaM datasets.

For all received data, a split will be made between data for the Parent study and the data for the Extension phase, to allow submission of the CDISC data and documentation for the Parent study.

10 TECHNICAL DETAILS

10.1 Programming conventions

Programming conventions apply to both the Parent study as well as the Extension phase, unless otherwise specified.

Demographics

BMI will be calculated in SAS as follows: weight/height^2, with weight in kg and height in meter (unit kg/m^2). Durations will be programmed as stated in the respective analysis sections. Where applicable, weight and height measurements will be converted to units cm and kg, using the following conversion factors: from inch to cm: multiply by 2.54, from lb to kg: multiply by 0.45359237. Variables with values representing a missing observation (such as -99, 99, 888 or 999) will be recoded to missing in SAS.

GM2 Diagnosis

Time since diagnosis will be calculated as the difference in days between first dose date and the date of diagnosis +1. Date of diagnosis will be taken as the earliest date from the following two records: the (earliest) date of Medical History with MHDECOD= "Tay-Sachs disease" or "Sandhoff disease" and the date of genetic confirmation of disease as recorded in the Primary Genetic Diagnosis of GM2 Gangliosidosis data.
If considered more feasible, the duration may be presented in weeks (calculated time divided by 7), months (calculated time divided by 365.25/12) or years (calculated time divided by 365.25) instead of days. For historical data however, it is likely that partial dates are available. If date of diagnosis is (partially) missing, the following will be applied:

- If only day is missing, the first day of the month will be used.
- If day and month are both missing, 1 January of the same year will be used.

If full date is missing, no imputation will be done.

Age calculation

Age at diagnosis will be based on the date of diagnosis (see above for determination) and the birth date as collected in the database. For a partial missing date of diagnosis, the rules as stated above will be applied. For partially missing birthdate: depending upon the level of precision of the collected birth date: if only month and year is available, then imputation of day 15 is used for calculation of age. If only birth year is available, 15 June of that year will be used as birthdate.

Age is not collected during the Extension phase, and therefore whenever age is necessary for any calculations or presentations at Extension phase visits, it should be calculated using the respective visit date combined with the date of birth as collected at Visit 1. The same imputation rule as stated above (for calculation of age at diagnosis) will be applied.

Blinded data

The blinded results on the video-recordings (blinded rater's CI-CS and CI-S) will be received from [redacted] according to the agreed transfer specifications. Recoding needs to be done using the visit schedule provided with the blinded results, since the visit review order is randomly assigned. The LOCF for video comparison is already available in the [redacted] data.

SCAFI

For 8MWT, assistance of another person or using the wall as support is not allowed. However, timed 8MWT trials for participants who used "unallowable support" (another person or the wall) are entered in the database, considering these measurements are also necessary due to their relation to the primary and a number of secondary endpoints of the study based on videos of the test. Therefore, for the SCAFI analysis, the data for these patients that used either the wall or person support during the 8MWT should be made missing.

Furthermore, patients for 9HPT, the following clinical rules apply:

- if proband cannot complete trial in 5 minutes (i.e. 300 seconds) with dominant hand, move on to the trials with non-dominant hand
- if proband cannot complete one trial in 5 minutes (i.e. 300 seconds) with non-dominant hand, discontinue 9HPT

As a consequence, if in the database values >300 are entered for Trial 1 and/or 2, the approach for handling these data in regards to SCAFI will be:

- In the case Trial 1 has a value <300 and Trial 2 has a value of >300 seconds, only the trial 1 with value <300 will be used for further analysis.
- In the case Trial 1 has a value of >300 seconds and Trial 2 has a value of <300 seconds, the data will count as missing and unable to perform due to physical limitations (as per the discontinuation rules defined above, Trial 2 should not have been conducted)
- In other instances, where only one trial is performed with a value >300 seconds, or two trials are performed with both values >300 seconds, the data will be considered missing due to physical limitations and replaced with a missing value for further analysis.

After the adjustments made for 8MWT and 9HPT as above, SCAFI total score will be calculated as follows:
The mean of the trial1 and trial2 scores are determined for each subtest/hand, and referred to as: 8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg. In case a subtest is only performed once, this value is taken as the mean. Then, the 8MW and 9HPT performance times are converted to the same dimension as the PATA rate, as follows: 8MW_recipr = 1/8MW_avg, 9HPT_recipr = (1/9HPT-D_avg + 1/9HPT-ND_avg)/2. Note that 8MW and 9HPT performance times cannot be 0.
Then each is converted into a Z-score using the following algorithms:

- 8MW_zscore = (8MW_recipr - BL(mean 8MW_recipr)) / BL(SD 8MW_recipr)
- 9HPT_zscore = (9HPT_recipr - BL(mean 9HPT_recipr)) / BL(SD 9HPT_recipr)
- PATA_zscore = (PATA_avg - BL(mean PATA_avg)) / BL(SD PATA_avg)

where BL means baseline value and SD means standard deviation. The mean 8MW_recipr and SD 8MW_recipr are determined using the SAF analysis population.

In case of missing values, the following applies:

- The cases with either 8MW, 9HPT (in one or both hands) or PATA tests not performed at all (“unable to perform due to physical limitations” or “not performed for other reason”) are excluded from the calculation of baseline means.
- Only when an 8MW score, 9HPT score or PATA score is missing due to “unable to perform due to physical limitations”, then a substitution will be applied as follows:
    - When 8MW tests were not performed, then 8MW_recipr will be substituted by 1/1800 s.
    - When 9HPT tests were not performed: if only data for one hand is missing, then 9HPT_avg=3000 s for that specific hand; when data for both hands are missing, then 9HPT_recipr=1/3000 s.
    - When PATA tests were not performed, then PATA_avg is substituted with 0.

For agreement on “unable to perform due to physical limitations” or “not performed for other reason”, the sponsor will be contacted. The missings created for 8MWT and 9HPT as stated above are considered “unable to perform due to physical limitations”.

According to the “SCAFI rating manual”, the SCAFI total score is to be determined as the arithmetic mean of the 3 Z-scores. However, following explanation of IntraBio and Dr. Tanja Schmitz-Huebsch, the SCAFI score needs to be calculated as the mean of the non-missing Z-scores. The SCAFI score can only be determined if at least one test is actually performed, i.e. in case all tests were imputed because of missing values, no SCAFI score will be determined. This contradicts the “SCAFI rating manual”.

SARA

The overall SARA score is reported as provided in the eCRF. For presentation of the subscore results, the individual results as provided in the eCRF will be used, where for the subscores finger chase, nose-finger test, fast alternating hand movements, heel-shin slide this is the average value for the two sides (left and right) as determined per eCRF.

mDRS

The mDRS composite score is calculated as a sum of the separate subscores divided by 24. If one of the subscores is missing, the mDRS composite score is missing as well. As compared to the official questionnaire, the eCRF allows values of 0 ('normal') for the subscores.

Laboratory

Laboratory safety data measurements that are denoted as smaller/larger than a certain value (e.g. < x.x), will be imputed with this value (e.g. x.x) for calculations of descriptive statistics. Listings will show the value as collected, and a footnote will be added to the summary tables for explanation.

Comments

The following data are either not collected or only collected as free text fields in the eCRF/database or lack relevant information in the database, and it is therefore necessary that these details are added in the programming code manually. The decisions to drive this manual coding may be documented and finalized during the BDRM but can already be maintained/discussed during the course of the trial for efficiency.

- The reason for video unavailability is recorded as free text field, and it needs to be deducted if any of these reasons consider a "technical issue". The sponsor will provide input.
- PATA score missing due to "unable to perform due to physical limitations" is not available as such in the database but needs to be deducted from a free comment field. The sponsor will provide input.
- Age is not collected during the Extension phase and full birthdate is not provided in the Parent study; therefore, the sponsor will provide details which patients turned 13 between Visit 7 and 8, and between Visit 10 and 11.
- SCAFI comments as stated above on SCAFI calculations. The sponsor will provide input to which comments relate to "unable to perform due to physical limitations" and which relate to "not performed for other reason".
- Information to match the supplied bottles and/or sachets of study drug with the returned bottles and/or sachets, to be able to differentiate bottles/sachets used for the Parent phase and the Extension phase to allow appropriate compliance calculations.

Other

Durations, determination of baseline values and handling of missing data will be programmed as stated in the respective analysis sections.

Any other programming conventions that are not foreseen in preparation of this SAP, will be handled when encountered and documented separately.

10.2 Coding

Coding of adverse events, concomitant medication and medical history will be performed by the Data Management provider. Adverse events and medical history are coded with the MedDRA

coding system (version as provided by the DM vendor). Concomitant medication is coded according to the WHO drug code and the ATC class code, using WHO ATC/DDD Index (version as provided by the DM vendor). Coding will be supplied as part of the data transfer.

10.3 Analysis software

The statistical analysis and reporting will be done using SAS® for Windows™ version 9.4. SAS tabular output (tables and listings) will be saved in RTF format. SAS graphs will be saved in PNG format. Both will be imported into PDF and/or Word® and supplied to the Medical Writer for use in the clinical study report. When the sponsor wants to receive the output before the study report, then a PDF document is supplied.

Furthermore, a number of excels will be provided to the client when providing the output. This may concern the CI-CS Visit 2 to Visit 6 for each patient and the CI-S primary anchor data for each patient at each visit from the Parent Study, and the excels for creation of Forest plots of the Parent Study and Extension Phase.

10.4 Presentation of tables, listings, graphs

All output will be generated as SAS tables, graphs and listings.

All tables and listings will be created such that they fit landscape pages. The tables for the end-of-text and listings for the appendix will be created using SAS, following the specifications according to the TLF template document that will be created separately.

For graphs, output will be as created as PNG plot. Graphs are preferably created using black, grey, and white color only, to facilitate black-and-white printing. Different line patterns and symbols will be used to differentiate between classification or treatment levels. If certain plots can only be visually improved by using colours, then this will be done using colors as distinct as possible. Graphs will be created such (i.e. taking into account line thickness and font size) that they can be presented as two (2) per page in the clinical study report.

11 TABLES, LISTINGS, GRAPHS – PARENT STUDY

11.1 General

A detailed list of tables, graphs and listings is presented, if applicable, per report section in Sections 11.2, 11.3 and 11.4.
Template tables and listings as well as *example* plots (as received from client or extracted from a relevant paper, if available) will be used as a reference for creation of all output, and a separate document will be created for this. Table/graph/listing numbering will be followed, however, if the data give cause for combining or splitting tables or listings, the numbering may be adapted as necessary. In case there are no data available for a certain table/listing, an empty table will be created and ‘No data available’ will be stated.

11.2 In-text tables and graphs

Not applicable. Will be designed or extracted by the Medical Writer during creation of the Clinical Study Report, based on the tables and graphs created for Section 14 of the CSR.

11.3 End-of-text tables and graphs

Following ICH E3 guidelines, all tables and graphs mentioned here will be presented in Section 14 of the CSR, and tables will be prepared in the order and with section number as stated.

Table/graph number	Contents of table/graph
14.1 Demographic Data Summary figures and tables	
14.1.1-14.1.4	Demographics (ITT, mITT, PPS and SAF population)
14.1.5-14.1.7	Baseline characteristics (ITT, mITT and PPS population)
14.1.7	Disposition
14.1.8	Medical history
14.1.9	GM2 Gangliosidosis specific history
14.1.10	Compliance
14.2 Efficacy Data Summary figures and Tables	
14.2.1	Result of t-test or Wilcoxon test on the primary endpoint (mITT and PPS), including sensitivity analysis (only mITT).
14.2.2	Frequency table of the CI-CS score Visit 4 vs Visit 2 minus the CI-CS score Visit 6 vs Visit 4, combined with the frequency table of the CI-CS score Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.3	Descriptive statistics of the CI-CS score Visit 4 vs Visit 2 minus the CI-CS score Visit 6 vs Visit 4, combined with the descriptive statistics of the CI-CS score Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 (mITT and PPS population).
14.2.4	Bar chart of CI-CS score Visit 4 vs Visit 2 minus CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.5	Frequency table of the CI-CS score Visit 4 vs Visit 2 minus the CI-CS score Visit 6 vs Visit 4 according to subgroups, combined with the frequency table of the CI-CS score Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 according to subgroups (mITT population)
14.2.6	Descriptive statistics of the CI-CS score Visit 4 vs Visit 2 minus the CI-CS score Visit 6 vs Visit 4 according to subgroups, combined with the descriptive statistics of the CI-CS score Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 according to subgroups (mITT population)
14.2.7	Bar chart of CI-CS score Visit 4 vs Visit 2 and separately, CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.8	Pearson rank correlation coefficients (mITT population)
14.2.9	ICC analysis (mITT population)

14.2.10	Result of t-test or Wilcoxon test on the key secondary endpoint 'Change in severity score based on averaged CI-S' (mITT and PPS population),
14.2.11	Frequency table of the change in severity score based on averaged CI-S scores between baseline and treatment minus the change between treatment and washout (mITT and PPS population)
14.2.12	Descriptive statistics of the change in severity score based on averaged CI-S scores between baseline and treatment minus the change between treatment and washout (mITT and PPS population)
14.2.13	Bar chart of change in severity score based on averaged CI-S scores between baseline and treatment minus the change between treatment and washout (mITT and PPS population)
14.2.14	Frequency table of the change in severity score based on averaged CI-S scores between baseline and treatment minus the change between treatment and washout according to subgroups (mITT population)
14.2.15	Descriptive statistics of change in severity score based on averaged CI-S scores between baseline and treatment minus the change between treatment and washout according to subgroups (mITT population)
14.2.16	Result of Wilcoxon test on the key secondary endpoint 'CI-CS score reclassified' (mITT and PPS)
14.2.17	Frequency table of the CI-CS score reclassified on a 3-point scale Visit 4 vs Visit 2 and separately, the reclassified CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.18	Descriptive statistics of the CI-CS score reclassified on a 3-point scale Visit 4 vs Visit 2 and separately, the reclassified CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.19	Bar chart of CI-CS score reclassified on a 3-point scale Visit 4 vs Visit 2 and separately, CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.20	Frequency table of the CI-CS score reclassified on a 3-point scale Visit 4 vs Visit 2 and separately, the reclassified CI-CS score Visit 6 vs Visit 4 according to subgroups (mITT population)
14.2.21	Descriptive statistics of the CI-CS score reclassified on a 3-point scale Visit 4 vs Visit 2 and separately, the reclassified CI-CS score Visit 6 vs Visit 4 according to subgroups (mITT population)
14.2.22	Frequency table of the CI-CS score for the non-primary anchor test Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.23	Descriptive statistics of the CI-CS score for the non-primary anchor test Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 (mITT and PPS population)
14.2.24	Frequency table of the CI-CS score for the non-primary anchor test Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 according to subgroups (mITT population)

14.2.25	Descriptive statistics of the CI-CS score for the non-primary anchor test Visit 4 vs Visit 2 and separately, the CI-CS score Visit 6 vs Visit 4 according to subgroups (mITT population)
14.2.26	Result of t-test or Wilcoxon test on the change in SARA total score between Visit 2 and Visit 4, as well as Visit 4 and Visit 6 (mITT and PPS population)
14.2.27	Descriptive statistics of the SARA total score and 8 subscores at each visit (mITT and PPS population)
14.2.28	Mean or median plot versus time of the SARA total score (mITT and PPS population)
14.2.29	Stacked bar chart of the 8 SARA subscores at each visit (mITT and PPS population)
14.2.30	Spaghetti plot of SARA total score at each visit (mITT and PPS population)
14.2.31	Descriptive statistics of the SARA total score and 8 subscores at each visit according to subgroups (mITT population)
14.2.32	Result of t-test or Wilcoxon test on the change in SCAFI total score, and subscores 8MW_avg, 9HPT-D_avg and PATA_avg between Visit 2 and Visit 4, as well as Visit 4 and Visit 6 (mITT and PPS population)
14.2.33	Descriptive statistics of the SCAFI total score and subscores 8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg at each visit (mITT and PPS population)
14.2.34	Mean or median plot of SCAFI total score versus time (mITT and PPS population)
14.2.35	Spaghetti plot of SCAFI total score versus time (mITT and PPS population)
14.2.36	Descriptive statistics of the SCAFI total score and subscores 8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg at each visit according to subgroups (mITT population)
14.2.37	Frequency table of the EQ-5D-5L and EQ-5D-Y health status and 5 subdomains at each visit (mITT and PPS population)
14.2.38	Frequency table of the EQ-5D-5L and EQ-5D-Y health status and 5 subdomains at each visit according to subgroups (mITT population)
14.2.39	Descriptive statistics of the EQ-VAS for -5L and -Y combined at each visit (mITT and PPS population)
14.2.40	Mean or median plot of EQ-VAS versus time for -5L and -Y combined (mITT and PPS population)
14.2.41	Spaghetti plot of EQ-VAS versus time for -5L and -Y combined (mITT and PPS population)
14.2.42	Descriptive statistics of the EQ-VAS for -5L and -Y combined at each visit according to subgroups (mITT population)

14.2.43	Descriptive statistics on the change in mDRS composite score between Visit 2 and Visit 4, as well as Visit 4 and Visit 6 (mITT and PPS population)
14.2.44	Descriptive statistics of the mDRS composite score and 6 components at each visit (mITT and PPS population)
14.2.45	Mean or median plot versus time of the mDRS composite score (mITT and PPS population)
14.2.46	Spaghetti plot of the mDRS composite score versus time (mITT and PPS population)
14.2.47	Descriptive statistics of the mDRS composite score and 6 components at each visit according to subgroups (mITT population)
14.2.48	Frequency table of the 3 CGI-S scores (treating physician, caregiver and patient) at each visit (mITT and PPS population)
14.2.49	Descriptive statistics of the 3 CGI-S scores (treating physician, caregiver and patient) at each visit (mITT and PPS population)
14.2.50	Bar charts of the 3 CGI-S scores (treating physician, caregiver and patient) at each visit (mITT and PPS population)
14.2.51	Mean or median plot versus time of the 3 CGI-S scores (treating physician, caregiver and patient) (mITT and PPS population)
14.2.52	Spaghetti plot of the 3 CGI-S scores (treating physician, caregiver and patient) versus time (mITT and PPS population)
14.2.53	Frequency table of the 3 CGI-S scores (treating physician, caregiver and patient) at each visit according to subgroups (mITT population)
14.2.54	Descriptive statistics of the 3 CGI-S scores (treating physician, caregiver and patient) at each visit according to subgroups (mITT population)
14.2.55	Frequency table of the 3 CGI-C scores (treating physician, caregiver and patient) between Visit 2 vs Visit 4, and between Visit 4 and Visit 6 (mITT and PPS population)
14.2.56	Descriptive statistics of the 3 CGI-C scores (treating physician, caregiver and patient) between Visit 2 vs Visit 4, between Visit 4 and Visit 6, and Visit 4 vs Visit 2 minus the CI-CS score Visit 6 vs Visit 4 (mITT and PPS population).
14.2.57	Result of t-test or Wilcoxon test on the change of CGI-C score for treating physician Visit 4 vs Visit 2 minus Visit 6 vs Visit 4 (mITT and PPS population)
14.2.58	Result of t-test or Wilcoxon test on the change of CGI-C score for treating physician between Visit 2 and Visit 4, as well as Visit 4 and Visit 6 (mITT and PPS population)
14.2.59	Bar charts of the 3 CGI-C scores (treating physician, caregiver and patient) between Visit 2 vs Visit 4, and between Visit 4 and Visit 6 (mITT and PPS population)

14.2.60	Frequency table of the 3 CGI-C scores (treating physician, caregiver and patient) between Visit 2 vs Visit 4, and between Visit 4 and Visit 6 according to subgroups (mITT population)
14.2.61	Descriptive statistics of the 3 CGI-C scores (treating physician, caregiver and patient) between Visit 2 vs Visit 4, and between Visit 4 and Visit 6 according to subgroups (mITT population)
14.2.62	Forest plots
14.3 Safety Data Summary figures and tables – 14.3.1 Displays of Adverse Events	
14.3.1.1	Overview adverse events
14.3.1.2	Treatment emergent adverse events
14.3.1.3	Treatment emergent adverse events by severity grade
14.3.1.4	Treatment emergent adverse events by relationship
14.3.1.5	Treatment emergent adverse events leading to premature discontinuation
14.3.1.6	Treatment emergent adverse events by treatment cycle
14.3.1.7	Treatment emergent adverse events by severity grade by treatment cycle
14.3.1.8	Treatment emergent adverse events by relationship by treatment cycle
14.3.1.9	Treatment emergent adverse events leading to premature discontinuation by treatment cycle
14.3 Safety Data Summary figures and tables – 14.3.2 Listings of Deaths, Other Serious and Significant Adverse Events	
14.3.2.1	SAEs
14.3.2.2	Deaths
14.3 Safety Data Summary figures and tables – 14.3.4 Abnormal Laboratory Value Listing (each patient)	
14.3.4.1	Out of range clinical laboratory
14.3.4.2	Clinical laboratory – haematology
14.3.4.3	Clinical laboratory – clinical chemistry
14.3.4.4	Clinical laboratory – urinalysis
14.3.5	Vital signs
14.3.6	ECG
14.3.7	Prior and concomitant medication and therapies

11.4 Listings

Following ICH E3 guidelines, all listings mentioned here will be presented in Section 16.2 of the CSR, and listings will be prepared in the order and with section number as stated.

Individual listings will be prepared of the data collected in the database, following SDTM data format. No combining of data other than mentioned in this paragraph will be performed. The key variables in the listings (except for a few displaying screening data) will be patient number, and age/weight/dosing group. If applicable, visit number will be listed additionally. Furthermore, a listing containing study visit dates will be presented.

Listing number	**Contents of listing**
16.2.1 Discontinued patients	
16.2.1.1	Inclusion/exclusion criteria – deviations
16.2.1.2	Patient disposition
16.2.2 Protocol deviations	
16.2.2	Protocol deviations
16.2.3 Demographic and baseline data	
16.2.3.1	Demographics
16.2.3.2	Baseline characteristics
16.2.3.3	Medical history
16.2.3.4	GM2 Gangliosidosis specific history
16.2.4 Compliance and/or treatment data	
16.2.4.1	Drug administration
16.2.4.2	Drug accountability
16.2.4.3	Study visits
16.2.5 Individual efficacy response data	
16.2.5.1	Individual CI-S and CI-CS scores
16.2.5.2	Individual scores on ataxia and functioning - SARA test
16.2.5.3	Individual scores on ataxia and functioning - SCAFI test
16.2.5.4	Individual health related quality of life – EQ-5D-5L, EQ-5D-Y and the respective EQ-VAS

16.2.5.5	Individual overall neurological status – mDRS test
16.2.5.6	Individual global impression scores – CGI-S (treating physician, caregiver, and patient)
16.2.5.7	Individual change in global impression scores – CGI-C (treating physician, caregiver, and patient)
16.2.6 Adverse event listings	
16.2.6.1	Adverse events
16.2.6.2	AEs leading to discontinuation from the study
16.2.6.3	Vital signs
16.2.6.4	ECG
16.2.6.5	Prior and concomitant medication and therapies
16.2.7 Listing of individual laboratory measurements by patient	
16.2.7.1	Laboratory safety data – haematology
16.2.7.2	Laboratory safety data – clinical chemistry
16.2.7.3	Laboratory safety/pregnancy data – urinalysis
16.2.7.4	Listing of urine test for N-Acetyl-D-Leucine
16.2.7.5	Listing of PK concentration data
16.2.7.6	General comments

12 TABLES, LISTINGS, GRAPHS – EXTENSION PHASE

12.1 General

A detailed list of tables, graphs and listings is presented, if applicable, per report section in Sections 12.2, 12.3 and 12.4.
Template tables and listings as well as *example* plots (as received from client, a previous study or extracted from a relevant paper, if available) will be used as a reference for creation of all output, and a separate document will be created for this. Table/graph/listing numbering will be followed, however, if the data give cause for combining or splitting tables or listings, the numbering may be adapted as necessary. In case there are no data available for a certain table/listing, an empty table will be created and 'No data available' will be stated.

12.2 In-text tables and graphs

Not applicable. Will be designed or extracted by the Medical Writer during creation of the Clinical Study Report, based on the tables and graphs created for Section 14 of the CSR.

12.3 End-of-text tables and graphs

Following ICH E3 guidelines, all tables and graphs mentioned here will be presented in Section 14 of the CSR, and tables will be prepared in the order and with section number as stated.

Table/graph number	Contents of table/graph
14.1 Demographic Data Summary figures and tables	
14.1.11-14.1.13	Demographics (mITTe and SAFe)
14.1.14-14.1.15	Baseline characteristics (mITTe)
14.1.16	Disposition
14.1.17	Medical history
14.1.18	Compliance
14.2 Efficacy Data Summary figures and Tables	
14.2.63	Frequency table, p-value, and CI on the primary endpoint CI-CS success rate (mITTe)
14.2.64	Descriptive statistics on CI-CS scores on the primary anchor, including change values between visits (mITTe).
14.2.65	Frequency table of the primary endpoint CI-CS success rate according to subgroups (mITTe).
14.2.66	Descriptive statistics of the primary endpoint CI-CS, including change values between visits, according to subgroups (mITTe).
14.2.67	Frequency table of the success rate of CI-CS score for the non-primary anchor test (mITTe)

14.2.68	Descriptive statistics of the CI-CS score for the non-primary anchor test, including change values between visits (mITTe)
14.2.69	Frequency table of the success rate of CI-CS score for the non-primary anchor test according to subgroups (mITTe)
14.2.70	Descriptive statistics of the CI-CS score for the non-primary anchor test, including change values between visits, according to subgroups (mITTe)
14.2.71	Descriptive statistics of the SARA total score and 8 subscores at each visit, including change values between visits (mITTe)
14.2.72	Descriptive statistics of the SARA total score and 8 subscores at each visit, including change values between visits, according to subgroups (mITTe)
14.2.73	Descriptive statistics of the SCAFI total score and subscores 8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg at each visit, including change values between visits (mITTe)
14.2.74	Descriptive statistics of the SCAFI total score and subscores 8MW_avg, 9HPT-D_avg, 9HPT-ND_avg and PATA_avg at each visit, including change values between visits, according to subgroups (mITTe)
14.2.75	Frequency table of the EQ-5D-5L and EQ-5D-Y health status and 5 subdomains at each visit (mITTe)
14.2.76	Frequency table of the EQ-5D-5L and EQ-5D-Y health status and 5 subdomains at each visit according to subgroups (mITTe)
14.2.77	Descriptive statistics of the EQ-VAS for -5L and -Y combined at each visit (mITTe)
14.2.78	Descriptive statistics of the mDRS total score and 8 subscores at each visit, including change values between visits (mITTe)
14.2.79	Descriptive statistics of the mDRS total score and 8 subscores at each visit, including change values between visits, according to subgroups (mITTe)
14.2.80	Frequency table of the 3 CGI-S scores (treating physician, caregiver, and patient) at each visit (mITTe)
14.2.81	Descriptive statistics of the 3 CGI-S scores (treating physician, caregiver, and patient) at each visit, including change values between visits (mITTe)
14.2.82	Frequency table of the 3 CGI-S scores (treating physician, caregiver, and patient) at each visit according to subgroups (mITTe)
14.2.83	Descriptive statistics of the 3 CGI-S scores (treating physician, caregiver, and patient) at each visit, including change values between visits, according to subgroups (mITTe)
14.3 Safety Data Summary figures and tables – 14.3.1 Displays of Adverse Events	
14.3.1.10	Overview adverse events
14.3.1.11	Treatment emergent adverse events

14.3.1.12	Treatment emergent adverse events by severity grade
14.3.1.13	Treatment emergent adverse events by relationship
14.3.1.14	Treatment emergent adverse events leading to premature discontinuation
14.3.1.15	Treatment emergent adverse events by treatment cycle
14.3.1.16	Treatment emergent adverse events by severity grade by treatment cycle
14.3.1.17	Treatment emergent adverse events by relationship by treatment cycle
14.3.1.18	Treatment emergent adverse events leading to premature discontinuation by treatment cycle
14.3.1.19	Treatment emergent events incidence rates
14.3 Safety Data Summary figures and tables – 14.3.2 Listings of Deaths, Other Serious and Significant Adverse Events	
14.3.2.3	SAEs
14.3.2.4	Deaths
14.3 Safety Data Summary figures and tables – 14.3.4 Abnormal Laboratory Value Listing (each patient)	
14.3.4.8	Out of range clinical laboratory
14.3.4.9	Clinical laboratory – haematology
14.3.4.10	Clinical laboratory – clinical chemistry
14.3.4.11	Clinical laboratory – urinalysis
14.3.12	Vital signs
14.3.13	ECG
14.3.14	Prior and concomitant medication and therapies
14.3.15	Physical examination

12.4 Listings

Following ICH E3 guidelines, all listings mentioned here will be presented in Section 16.2 of the CSR, and listings will be prepared in the order and with section number as stated.

Individual listings will be prepared of the data collected in the database, following SDTM data format. No combining of data other than mentioned in this paragraph will be performed. The key variables in the listings (except for a few displaying screening data) will be patient number, and age/weight/dosing group. If applicable, visit number will be listed additionally. Furthermore, a listing containing study visit dates will be presented.

16.2.3 Demographic and baseline data	
16.2.3.5	Demographics
16.2.3.6	Baseline characteristics
16.2.3.7	Medical history
16.2.4 Compliance and/or treatment data	
16.2.4.4	Drug administration
16.2.4.5	Drug accountability
16.2.4.6	Study visits
16.2.5 Individual efficacy response data	
16.2.5.9	Individual CI-CS scores
16.2.5.10	Individual scores on ataxia and functioning - SARA test
16.2.5.11	Individual scores on ataxia and functioning - SCAFI test
16.2.5.12	Individual health related quality of life – EQ-5D-5L, EQ-5D-Y and the respective EQ-VAS
16.2.5.13	Individual scores on mDRS
16.2.5.14	Individual global impression scores – CGI-S (treating physician, caregiver, and patient)
16.2.5.15	Individual change in global impression scores – CGI-C (treating physician, caregiver, and patient)
16.2.6 Adverse event listings	
16.2.6.6	Adverse events
16.2.6.7	AEs leading to discontinuation from the study
16.2.6.8	Vital signs
16.2.6.9	ECG
16.2.6.10	Prior and concomitant medication and therapies
16.2.6.11	Physical examination
16.2.7 Listing of individual laboratory measurements by patient	
16.2.7.5	Laboratory safety data – haematology

16.2.7.6	Laboratory safety data – clinical chemistry
16.2.7.7	Laboratory safety/pregnancy data – urinalysis
16.2.7.8	Listing of urine test for N-Acetyl-D-Leucine
16.2.7.9	Listing of PK concentration data
16.2.7.10	General comments