Statistical Analysis Plan

NCT Number: NCT02412787

Title: An Open Label Extension of Study HGT-HIT-094 Evaluating Long Term Safety and Clinical Outcomes of Intrathecal Idursulfase Administered in Conjunction with Elaprase® in Patients with Hunter Syndrome and Cognitive Impairment

Study Number: SHP609-302

Document Version and Date: Version 2.0, 21 February 2023

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

STATISTICAL ANALYSIS PLAN FOR SHP609-302 Final Abbreviated CSR

Protocol Title:	An Open-Label Extension of Study HGT-HIT-094 Evaluating Long-Term Safety and Clinical Outcomes of Intrathecal Idursulfase Administered in Conjunction with Elaprase® in Patients with Hunter Syndrome and Cognitive Impairment
Protocol Number:	SHP609-302 (Amendment 4)
Protocol Date:	09 Oct 2018
Investigation Product:	idursulfase for intrathecal use (idursulfase-IT [HGT-2310]),
Device Name:	SOPH-A-PORT® Mini S, Implantable Access Port, Spinal, Mini Unattached, with Guidewire (SOPH-A-PORT® Mini S)
SAP Version:	2.0, an addendum to SAP for interim analysis 1.1
Date:	21 Feb 2023
Prepared by:	[redacted]

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AE	Adverse Event
CRF	case report form
CRO	contract research organization
CSF	cerebrospinal fluid
CSR	clinical study report
ECG	electrocardiogram
GAG	Glycosaminoglycan(s)
HEOR	health economics and outcomes research
IA	interim analysis
IDDD	intrathecal drug delivery device
IT	intrathecal
IV	intravenous
PD	pharmacodynamic
PK	pharmacokinetic
SAP	statistical analysis plan
SAS	Statistical Analysis System©
SQS	Statistical & Quantitative Sciences

1. INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to provide specifications for the final analysis which will appear in the final abbreviated clinical study report (CSR).

Unless otherwise specified, the analyses in this SAP will apply the same methodologies as specified in the SAP for interim analysis (IA) (v1.1, dated 14 Mar 2019) and addendum 1 for interim analysis 2 (dated 26 Mar 2020). All available data in the extended treatment phase will be included in the final analyses. Only safety and pharmacodynamic (PD) analyses will be presented in the final abbreviated CSR.

2. OBJECTIVES AND ENDPOINTS

2.1 Objectives

The primary objective of this study is to evaluate long-term safety in patients with Hunter syndrome and cognitive impairment who are receiving intrathecal idursulfase-IT (SHP609) and intravenous (IV) Elaprase® enzyme replacement therapy (ERT).

The secondary and exploratory objectives are as noted in the SAP for IA in section 4.1. The final analyses will not address efficacy, exploratory, pharmacokinetic (PK) or health economics and outcomes research (HEOR) objectives.

The device objective is to evaluate the safety and performance of the SOPH-A-PORT Mini S.

2.2 Endpoints

The endpoints are defined in section 4.1 of the SAP for IA. The final analyses will focus on the safety and PD related endpoints as stated in section 4.1 of the SAP for IA. The endpoints for the final abbreviated CSR are reiterated below.

The primary endpoint of this study is the safety of intrathecal idursulfase-IT administration. Safety will be measured by adverse events (AEs, by type, severity, and relationship to treatment [idursulfase-IT, the intrathecal drug delivery device (IDDD), device surgical procedure, or IT administration process] and IV Elaprase infusion), changes in clinical laboratory testing (serum chemistry, hematology, urinalysis), vital signs, twelve-lead electrocardiogram (ECG) recordings, cerebral spinal fluid (CSF) parameters (chemistries, cell counts), anti-idursulfase antibodies in CSF and serum, including determination of anti-idursulfase antibodies having enzyme neutralizing activity.

The secondary PD endpoints include CSF concentration of idursulfase, the change from baseline in concentration of glycosaminoglycan (GAG) in CSF and the change from baseline in the concentration of GAG in urine.

The SOPH-A-PORT Mini S device will be evaluated using assessments of device implantation, device function, device longevity, and AEs associated with the implant surgery or device. These data will be collected on the patient's eCRF from the time of initial implantation.

3. ANALYSIS SETS

All PD and safety data analysis will be performed using the Safety Population, which is defined as follows: All subjects in study SHP609-302 who underwent IDDD implantation or received at least 1 dose of study drug (full or partial). The definitions of each analysis population are stated in section 6.2 of the SAP for IA; however, the focus of the final analyses will be safety and PD endpoints.

4. STATISTICAL ANALYSIS

All statistical analyses will be performed by Statistics & Quantitative Sciences (SQS) Department at Takeda or its designated contract research organization (CRO) unless otherwise specified, using SAS® software version 9.4 or higher (SAS Institute, Cary, N.C., USA).

Summary tables will be tabulated as noted in section 6.1 of the SAP for IA. The complete results from initial and extended treatment phases will be described in the final abbreviated CSR.

4.1 Patient Disposition

Patient disposition will be presented as described in section 6.3 of the SAP for IA.

4.2 Protocol Deviations

Protocol deviations will be presented in a listing as noted in section 6.4 of the SAP for IA.

4.3 Demographic and Baseline Characteristics

Demographic and baseline characteristics (as detailed in section 6.5 of the SAP for IA) will be presented in summary using the Safety Population by the Early IT group, the Delayed IT group for the antecedent pivotal study, antecedent substudy and overall.

4.4 Treatment Compliance and Extent of Exposure

Treatment compliance and extent of exposure will use the same methodology as in section 6.6 of the SAP for IA. These will be summarized by the Early IT group, the Delayed IT group for the antecedent pivotal study, antecedent substudy and overall using the Safety Population.

4.5 Secondary Pharmacodynamic Analyses

The CSF and urine GAG levels will be presented in graphs (mean observed values and mean percent change by visit) and descriptive statistics will be in summary tables (observed values, change from baseline and percent change from baseline) as noted in section 6.7.6 of the SAP for IA.

Analysis of both CSF total GAG and CSF heparan sulfate if available will be performed. CFS albumin concentrations, CSF/serum albumin quotients observed levels, the change from baseline and the percent change from baseline will be summarized as noted in section 6.7.6 of the SAP for IA.

Summaries will be presented by the Early IT group, the Delayed IT group and overall by visit using the Safety Population. All available data in the extended treatment phase to the end of study will be presented.

4.6 Analysis of Safety

Similar methodology as in the SAP for IA section 6.8 will be used for the Safety Population. All summaries will include all available scheduled visits to the end of the study.

The primary assessments of this study are the safety of idursulfase-IT administration. Safety will be measured by AEs, changes in clinical laboratory testing, vital signs, 12-lead ECG, CSF chemistries, anti-idursulfase antibodies and antibodies having enzyme neutralizing activity in CSF and serum. All analyses of safety data will be descriptive and presented by the Early IT group, the Delayed IT group and overall. Any former HGT-HIT-094 Substudy subjects will be included in the Early IT group.

AEs will be presented as detailed in section 6.8.1 of the SAP for IA. The clinical and CSF laboratory results will be presented as detailed in section 6.8.2 of the SAP for IA. The 12-lead ECG results will be presented as detailed in section 6.8.3 of the SAP for IA. The injection (IT) vital signs and regular vital signs will be listed as detailed in section 6.8.4 of the SAP for IA.

Similar methodology as in SAP for IA section 6.8.6 will be applied to the Safety Population for:

- Anti-idursulfase antibody formation (specifics in section 6.8.6.1)
- The Idursulfase enzyme levels in CSF (specifics in section 6.8.6.2)
- Height (cm), weight (kg), and head circumference (cm) (specifics in section 6.8.6.3)
- Brain MRI (specifics in section 6.8.6.4)
- Hearing Assessments (specifics in section 6.5.6.5).

The details are in sections of the SAP for IA mentioned above and presentations will include all available scheduled study visits to the end of study.

All safety analyses will be descriptive, no statistical testing will be performed.

4.7 SOPH-A-PORT Mini S Device Performance

See section 6.8.6.6 of the SAP for IA for details regarding device-related terminology as well as well as device performance analyses to be presented.

4.8 Concomitant Medications

See section 6.8.7 of the SAP for IA regarding presentations relating to concomitant medications.

5. STATISTICAL/ANALYTIC ISSUES

See section 8 of the SAP for IA. This section has 9 subsections which give details on adjustment for covariates, handling of dropouts/missing data, IA and data monitoring, multicenter studies, multiple comparisons/multiplicity, examination of interactions, sensitivity analyses, windowing visits, and data listings.

6. REFERENCES

1. Statistical Analysis Plan for Interim Analyses (for Protocol SHP609-302, amendment 4). Version 1.1. 14 March 2019.

7. APPENDICES

See section 10 of the SAP for IA for the appendix I: list of statistical outputs, and appendix II: protocol violations definitions. Other appendices are not applicable for the final CSR.

STATISTICAL ANALYSIS PLAN
FOR INTERIM ANALYSIS

Protocol Title:	An Open Label Extension of Study HGT-HIT-094 Evaluating Long Term Safety and Clinical Outcomes of Intrathecal Idursulfase Administered in Conjunction with Elaprase® in Patients with Hunter Syndrome and Cognitive Impairment
Protocol Number:	SHP609-302 (Amendment 4)
Protocol Date:	09 October 2018
Investigational Product and Device:	idursulfase for intrathecal use (idursulfase-IT, SHP609, HGT-2310), SOPH-A-PORT® Mini S, Implantable Access Port, Spinal, Mini Unattached, with Guidewire (SOPH-A-PORT® Mini S)
SAP Author:	[redacted]
SAP Version:	1.1
Release Date:	14 March 2019
Shire Medical Monitor:	[redacted], DO [redacted] 300 Shire Way Lexington, MA 02421, USA Telephone: [redacted] Fax: [redacted] USA)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Definition
ABR	auditory brainstem response
AE	adverse event
ATC	Anatomic therapeutic class
$AUC_{0-\infty}$	Area under the curve extrapolated to infinity, calculated using the observed value of the last non-zero concentration
AUC_{0-t}	Area under the curve from the time of dosing to the last measureable concentration
BBB	Blood brain barrier
BLQ	Below the lower limit of quantification
BOT-2	Bruininks-Oseretsky Test of Motor Proficiency, Second Edition
BRI	Behavioral Regulation Index
BRIEF	Behavior Rating Inventory of Executive Function
BRIEF-P	Behavior Rating Inventory of Executive Function-Preschool
BSID-III	Bayley Scales of Infant Development, Third Edition
CBC	complete blood count
CI	confidence interval
CL	Total body clearance for intravenous administration
CL/F	Total body clearance for extravascular administration divided by fraction of dose absorbed
C_{max}	Maximum serum concentration at t_{max}
CNS	central nervous system
CS	clinically significant
CSF	cerebrospinal fluid
CSR	Clinical study report
DAS-II	Differential Ability Scales, Second Edition
DQ	Developmental Quotient
DS	dermatan sulfate
eCRF	Electronic case report form
eCDF	Empirical Cumulative Distribution Function
EC	Emotional Control
ECG	Electrocardiogram
EMI	Emergent Metacognition Index
EOS	end of study
ERT	enzyme replacement therapy
EY	Early Years

Abbreviation	Definition
FI	Flexibility Index
FMQ	Fine Motor Quotient
GAG	glycosaminoglycan
GCA	General Conceptual Ability
GCA EY	DAS-II GCA standard scores from Early Years battery
GEC	Global Executive Composite
GMQ	Gross Motor Quotient
HS	heparan sulfate
IA	interim analysis
ICP	intracranial pressure
IDDD	intrathecal drug delivery device
ILP	Inter-peak latency
ISCI	Inhibitor Self-Control Index
IT	Intrathecal
IV	Intravenous
KM	Kaplan-Meier
LLOQ	Lower limit of quantification
LP	Lumbar puncture
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram(s)
MI	Megacognition Index
Min	Minimum
Max	Maximum
MMRM	Mixed Model for Repeated Measures
MRI	magnetic resonance imaging
MRT	Mean residence time
MSE	Mean squared error
NA	Not applicable
NCS	not clinically significant
PDMS-2	Peabody Developmental Motor Scales-2
PK	pharmacokinetic(s)
PO	Plan/Organize
PORT-A-CATH	PORT-A-CATH® II Low Profile™ Intrathecal Implantable Access System
PT	prothrombin time
PTA	Pure tone average

Abbreviation	Definition
PTT	partial thromboplastin time
Q1	25th percentile
Q3	75th percentile
QTc	corrected QT interval
SA	School Age
SAE	serious adverse event
SAP	Statistical Analysis Plan
SIB-R	Scale of Independent Behavior-revised
SNC	Special Nonverbal Composite
SOC	system organ class
SOPH-A-PORT Mini S	SOPH-A-PORT® Mini S, Implantable Access Port, Spinal, Mini Unattached, with Guidewire
SD	Standard deviation
Std. Err.	Standard error
SS	Standard score
$t_{1/2}$	Terminal half-life
T4	Thyroxine
t_{max}	Time to maximum serum concentration during a dosing interval
TEAE	Treatment-emergent adverse event
TMQ	Total Motor Quotient
V_{ss}	Apparent volume of distribution at steady-state
V_{ss}(%BW)	V_{ss} normalized for body weight
V_z	Volume of distribution associated with the terminal slope
V_z/F	Volume of distribution associated with the terminal slope following extravascular administration divided by the fraction of dose absorbed
WHO	World Health Organization
WM	Working Memory
λ_z	First order rate constant associated with the terminal (log-linear) portion of the curve

2. INTRODUCTION

2.1 Background

Hunter syndrome (Mucopolysaccharidosis II; MPS II) is an extremely rare disease with an estimated incidence of 1 in 162,000 live births worldwide. [1, 2] It is expected that approximately 67%-77% of these patients will present with the severe phenotype of the disease that includes progressive cognitive impairment (communicating hydrocephalus, increased intracranial pressure, seizures and hearing problems) and serious somatic disease.

2.2 Study Rationale

Elaprase is approved globally including the EU, USA, Canada, and Japan as an intravenously (IV) administered enzyme replacement therapy (ERT) for patients with Hunter syndrome. Large proteins such as Elaprase are not expected to cross the blood brain barrier (BBB) in sufficient amounts to be therapeutically beneficial. Therefore, it is not possible to treat the progressive brain disease in severe Hunter syndrome with Elaprase, and direct administration of the active enzyme to the central nervous system is required. Although idursulfase-IT (SHP609) contains the same active substance as Elaprase, their respective compositions differ in that idursulfase-IT (SHP609) contains only three of the five excipients contained in Elaprase. In contrast to Elaprase, idursulfase-IT (SHP609) is specially formulated for and compatible with direct introduction into the IT space, since it is isotonic and contains excipients suitable for IT administration. Note that Elaprase and idursulfase-IT (SHP609) are specifically formulated for the IV and IT compartments respectively; they must not be interchanged for safety reasons. A formulation appropriate for intrathecal administration of idursulfase (idursulfase-IT, SHP609, HGT-2310), using an intrathecal drug delivery device (IDDD) the SOPH-A-PORT® Mini S (Implantable Access Port, Spinal, Mini Unattached, with Guidewire), has been developed to address this unmet medical need in an extremely rare population. The advantage of using an IDDD in a chronic disease such as Hunter syndrome is the potential to obviate the need for multiple lumbar punctures for drug delivery.

Nonclinical experience with IT administration of idursulfase-IT has demonstrated wide distribution of idursulfase-IT to the CNS tissues. Idursulfase-IT has been shown to be well tolerated in several species and to be active in a murine disease model of idursulfase deficiency.

The safety and tolerability of ascending doses (1, 10, or 30 mg) of intrathecally administered idursulfase-IT (SHP609) were investigated in the first-in-human study HGT-HIT-045, a randomized, open-label, no-treatment controlled Phase I/II study in which idursulfase-IT (SHP609) was administered once monthly to pediatric MPS II patients via a surgically implanted IDDD (PORT-A-CATH® II Low Profile™ Intrathecal Implantable Access System [PORT-A-CATH]) for 6 months in conjunction with once weekly IV infusion of Elaprase. Based on the data available from HGT-HIT-045 and the extension study HGT-HIT-046, idursulfase-IT (SHP609) has been found to be well tolerated at all doses administered without safety concerns related to the investigational product.

The current study is an extension study to the pivotal Phase II/III study, HGT-HIT-094. Study HGT-HIT-094 is a controlled, randomized, two-arm, open-label, assessor-blinded, multicenter

study to determine the effect on clinical parameters of neurodevelopmental status of monthly IT administration of idursulfase-IT (SHP609) for 12 months in pediatric patients with Hunter syndrome and early cognitive impairment who have previously received and tolerated a minimum of 4 months of weekly therapy with Elaprase. This study serves to extend the treatment, as well as safety and efficacy monitoring, for patients who completed study HGT-HIT-094, which is designed to include the pediatric patients randomized to receive idursulfase-IT (SHP609) monthly, or no IT treatment, plus potentially a number of pediatric patients younger than 3 years old in the open-label, single-arm substudy also included in HGT-HIT-094.

3. PURPOSE OF STATISTICAL ANALYSIS PLAN

The interim analysis (IA) of study data will be conducted when all SHP609-302 subjects complete 12 months of idursulfase-IT (SHP609) treatment from the time of enrollment in SHP609-302 at visit Month 25/Week 52 or discontinue in study SHP609-302. The visit Month 25/Week 52 is the Visit Month 25 in the study SHP609-302 Extended Treatment Phase for subjects who were randomized to the idursulfase-IT (SHP609) treatment cohort in study HGT-HIT-094 and continued in study SHP609-302, and the Visit Week 52 in the study SHP609-302 Initial Treatment Phase for subjects who were randomized to the control cohort (No IT treatment) in study HGT-HIT-094 and began IT treatment in study SHP609-302. The purpose of this interim analysis statistical analysis plan (SAP) is to document technical and detailed specifications for the interim analysis of data collected for protocol SHP609-302 Amendment 4. Results of the interim analysis described in this SAP will be included in the interim clinical study report (CSR). Additionally, the planned analyses identified in this SAP may be included in regulatory submissions or future manuscripts. Any post-hoc, or unplanned analyses performed to provide results for inclusion in the interim CSR but not identified in this prospective SAP will be clearly identified in the interim CSR.

4. SUMMARY OF CLINICAL TRIAL FEATURES

4.1 General Description

Study Objectives	The primary objective of this study is: • To evaluate long-term safety in patients with Hunter syndrome and cognitive impairment who are receiving intrathecal idursulfase-IT (SHP609) and intravenous (IV) Elaprase® enzyme replacement therapy (ERT) The secondary objectives of this study are to evaluate long-term clinical efficacy outcomes in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609), in conjunction with Elaprase therapy with respect to the following: • Cognitive function as measured by General Conceptual Ability (GCA), the cluster areas and subtests of the Differential Abilities Scale, Second Edition (DAS-II), or domains of the Bayley Scales of Infant Development, Third Edition (BSID-III) • Adaptive behavior as measured by the Adaptive Behavior Composite (ABC) score and standard domain scores of the Vineland Adaptive Behavior Scales, Second Edition (VABS-II) • Brain structure volume as measured by magnetic resonance imaging (MRI) The exploratory objectives of this study are: • To evaluate the long-term effects of intrathecal idursulfase-IT (SHP609), given in conjunction with Elaprase, on clinical parameters (e.g., physiological assessments, neurocognitive assessments, neurologic function, and brain structure volumes) • To explore potential relationships between biomarkers and CNS symptomatology • To determine whether monthly idursulfase-IT (SHP609) administrations results in accumulation of idursulfase within the CSF compartment by measuring idursulfase levels in CSF immediately prior to idursulfase-IT (SHP609) administration • To determine the safety and performance of the SOPH-A-PORT Mini S The Pharmacokinetic and Pharmacodynamic Objectives are: • To evaluate the concentration of idursulfase and determine PK parameters in serum after IT administration in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy • To evaluate the concentration of idursulfase in CSF in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy • To evaluate the concentration of GAG in CSF in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy • To evaluate the concentration of GAG in urine in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy The Health Economics and Outcomes Research Objectives are: • To evaluate health status as measured by the EuroQol-5D-5L (EQ-5D-5L) instrument, in patients with Hunter syndrome and cognitive impairment who are

	treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy • To evaluate healthcare resource utilization, as measured by the Healthcare Utilization Questionnaire (HCUQ)in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy • Health Related Quality of Life (HRQoL) for caregivers in the Caregiver Impact Questionnaire (CIQ), in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy • To evaluate functional status as measured by the HS-FOCUS (Hunter Syndrome Functional Outcomes for Clinical Understanding Scale) instrument, in patients with Hunter syndrome and cognitive impairment who are treated with idursulfase-IT (SHP609) in conjunction with Elaprase therapy The SOPH-A-PORT Mini S Device Objective is: • To evaluate the safety and performance of the SOPH-A-PORT Mini S
Study Endpoints	The primary safety endpoints of this study are: • Adverse events (AEs; by type, severity, and relationship to treatment [idursulfase-IT, the intrathecal drug delivery device (IDDD), device surgical procedure, or IT administration process] and IV Elaprase infusion) • Changes in clinical laboratory testing (serum chemistry, hematology, urinalysis) • Vital signs • Twelve-lead ECG recordings • CSF laboratory parameters (chemistries, cell counts) • Anti-idursulfase antibodies in CSF and serum, including determination of anti-idursulfase antibodies having enzyme neutralizing activity The secondary efficacy endpoints of this study are: • Change from baseline (simple subtraction) in DAS-II standard scores at scheduled visits (including both Early Years and School Age batteries): GCA and Special Nonverbal Composite (SNC); and standard cluster scores: Verbal, Nonverbal, Spatial; and/or the age equivalents and Development Quotient (DQ) from the BSID-III domains: Cognitive and Language • Change from baseline (simple subtraction) in standard composite score of the VABS-II at scheduled visits: ABC; and domain standard scores: Communication, Daily Living Skills, Socialization, and Motor Skills • Change from baseline (simple subtraction) in age equivalents, developmental quotients, and T scores for the core subtests of the DAS-II at scheduled visits: Verbal Comprehension, Picture Similarities, Naming Vocabulary, Pattern Construction, Matrices, and Copying for the DAS-II Early Years battery and Recall of Designs, Word Definitions, Pattern Construction, Matrices, Verbal Similarities, and Sequential and Quantitative Reasoning for the DAS-II School Age battery • Change from baseline (simple subtraction) in age equivalents, developmental quotients, and v-scale scores of the VABS-II subdomains at scheduled visits: Communication (Receptive, Expressive, Written), Daily Living Skills (Personal, Domestic, Community), Socialization (Interpersonal Relationships, Play and Leisure Time, Coping Skills), Motor Skills (Gross, Fine) • Change from baseline (simple subtraction) in the v-scale scores and observed maladaptive levels of the VABS-II Maladaptive Behavior Index and its subscales (Internalizing, Externalizing) at scheduled visits • Change from baseline (simple subtraction) in Brain Total Intracranial Volume, Brain Total Tissue Volume, Brain Total White Matter, Brain Total Gray Matter, and Total CSF Volume at scheduled visits

	The exploratory efficacy endpoints of this study are:
	• Change in scores from the time of enrollment in SHP609-302 to 12 months during study SHP609-302, estimated by linear regression from Visit Month 13/Baseline to Visit Month 25/Week 52 in study SHP609-302, for:
	  • DAS-II GCA standard scores from Early Years battery
	  • DAS-II T scores from Early Years battery
	  • DAS-II standard cluster scores and SNC composite scores from Early Years battery
	  • DAS-II GCA standard scores (including both Early Years and School Age batteries)
	  • VABS-II ABC standard scores
	• Ordered categorical outcomes for each subject at 12-month visit (i.e., Visit Month 25/Week 52) in study SHP609-302, for:
	  • DAS-II GCA standard scores from Early Years battery
	  • DAS-II GCA standard scores (including both Early Years and School Age batteries)
	• Binary unreversed floor effect outcome for each subject at 12-month visit (i.e., Visit Month 25/Week 52) in study SHP609-302, for:
	  • DAS-II GCA standard scores from Early Years battery
	  • DAS-II T scores from Early Years battery
	  • DAS-II standard cluster scores and SNC composite scores from Early Years battery
	  • DAS-II GCA standard scores (including both Early Years and School Age batteries)
	  • VABS-II standard ABC scores
	
	The pharmacokinetic and pharmacodynamic endpoints of this study are:
	• Serum concentration of idursulfase and serum PK parameters after IT administration
	• CSF concentration of idursulfase
	• Change from baseline in the concentration of GAG in CSF
	• Change from baseline in the concentration of GAG in urine
	
	The health economics and outcomes research endpoints for this study are:
	• Health status dimensions as obtained by the EQ-5D questionnaire
	• Health care resource utilization will be assessed using the HCUQ
	• Health-related quality of life (HRQoL) of the caregiver using CIQ
	• Functional status as obtained by the HS-FOCUS form
	
	SOPH-A-PORT Mini S assessments will include:
	• The SOPH-A-PORT Mini S device will be evaluated using assessments of device implantation, device function, device longevity, and AEs associated with the implant surgery or device. These data will be collected on the patient's eCRF from the time of initial implantation.

Study Design	This is an open-label, non-randomized study for patients who completed Study HGT-HIT-094; all patients in this study will be treated with intrathecal idursulfase-IT (SHP609) in conjunction with Elaprase therapy. Patients who complete Visit Week 52 assessments of Study HGT-HIT-094 and who meet

	the eligibility criteria and for whom informed consent is provided will be enrolled in this extension study. All patients will receive idursulfase-IT (SHP609) at the same dose and frequency (10 mg once every 28 days) as selected for the HGT-HIT-094 study. Patients who are younger than 3 years of age will continue to receive an adjusted dose as specified in the protocol (Amendment 3) Section 6.3. The assessments in this study are similar to those performed in Study HGT-HIT-094 and will be performed at regular intervals; the frequency of assessments will vary depending on the treatment phase of the study. Initial Treatment Phase – Only Patients Not Treated with Idursulfase-IT (SHP609) in Study HGT-HIT-094: Patients who participated in the control arm (no treatment with idursulfase-IT) in Study HGT-HIT-094 will begin this study in the Initial Treatment Phase. These patients will have the intrathecal drug delivery device (IDDD) implanted following enrollment in order to begin treatment with idursulfase-IT (SHP609). These patients will, during the Initial Treatment Phase of this study (i.e., the first 12 months), undergo treatment and assessments similar to those performed for patients who were treated in Study HGT-HIT-094. After completion of the Initial Treatment Phase and if there are no safety concerns, patients may continue receiving monthly idursulfase-IT (SHP609) in the Extended Treatment Phase of this study. Extended Treatment Phase – All Patients: Patients who received treatment with idursulfase-IT (SHP609) in the HGT-HIT-094 study will begin this study in the Extended Treatment Phase, which reflects a reduced frequency of assessments compared with the schedule of assessments in Study HGT-HIT-094. Standardized neurodevelopmental assessments and health economic and outcomes research assessments will be performed every 24 weeks and CSF assessments (other than standard chemistry) and clinical laboratory tests of blood and urine will be performed every 12 weeks. Brain MRI and 12-lead ECG will be performed every 48 weeks. A pharmacokinetic assessment will be performed once during this phase of the study. Patients will receive their monthly doses of intrathecal idursulfase-IT (SHP609) at the main study site or at a local site. All patients will receive their weekly IV infusions of Elaprase throughout the study. It is anticipated that the IDDD will be used to collect CSF samples and to deliver IT injections of idursulfase-IT (SHP609) and preservative-free saline flushes. No other medication will be administered through the device. If the IDDD appears to be non-functional, or if its use is precluded on a scheduled day of dosing, site personnel will refer to the IDDD Manual, which provides details on the investigation and management of any IDDD-related issues. This includes possible partial revision or complete replacement of the IDDD as indicated. If the IT space is not accessible via the IDDD, investigational product may be administered by LP. Should the IDDD become clogged, undergo mechanical complications or otherwise not be accessible, the CSF sample may also be obtained by LP. General anesthesia may be required for injections of study drug and some evaluations, and can be used at the discretion of the Investigator as indicated in the study manuals. It is anticipated that study procedures such as lumbar puncture, MRI, and audiometry will have to be performed with sedation/anesthesiology support. Patients will have the IDDD removed when they discontinue from or complete the study unless the patient is continuing to receive product or treatment through another mechanism (e.g., extension study, expanded access program).

Number of Patients	Pediatric patients (plus potentially any pediatric patients younger than 3 years of age in the antecedent substudy) who completed Study HGT-HIT-094, and meet the eligibility criteria may enroll in this study. Up to 49 patients from the HGT-HIT-094 pivotal study and 9 patients from the HGT-HIT-094 substudy are expected.
Study Product	Idursulfase for intrathecal use (idursulfase-IT, SHP609, HGT-2310) 10 mg
Intrathecal drug delivery device	SOPH-A-PORT® Mini S, Implantable Access Port, Spinal, Mini Unattached, with Guidewire (SOPH-A-PORT® Mini S)
Treatment and Study Duration	Patients will participate in this extension study for a maximum duration of 10 years of treatment across studies HGT-HIT-094 and SHP609-302, unless they discontinue the study or Shire discontinues the study. The study will conclude after the last patient has completed his last visit.
Randomization and Blinding	This is an open-label, non-randomized study. All patients will receive idursulfase-IT (SHP609) at the same dose and frequency (10 mg once every 28 days) as selected for the HGT-HIT-094 study.

4.2 Determination of Sample Size/Randomization

This is an open-label extension trial of Study HGT-HIT-094. Only eligible patients who participated in Study HGT-HIT-094 may enroll in SHP609-302, and therefore no sample size calculation was performed. The number of patients is described above in Section 4.1.

5. EFFICACY AND SAFETY VARIABLES

5.1 Schedule of Evaluations

This extension study has 2 phases: an Initial Treatment Phase (only patients not treated with idursulfase-it in study HGT-HIT-094) and an Extended Treatment Phase (all patients). The detailed schedule of evaluations is located in the SHP609-302 protocol (Amendment 4).

5.2 Primary Safety Endpoints

The primary safety endpoints are described above in Section 4.1.

5.3 Secondary Efficacy Endpoints

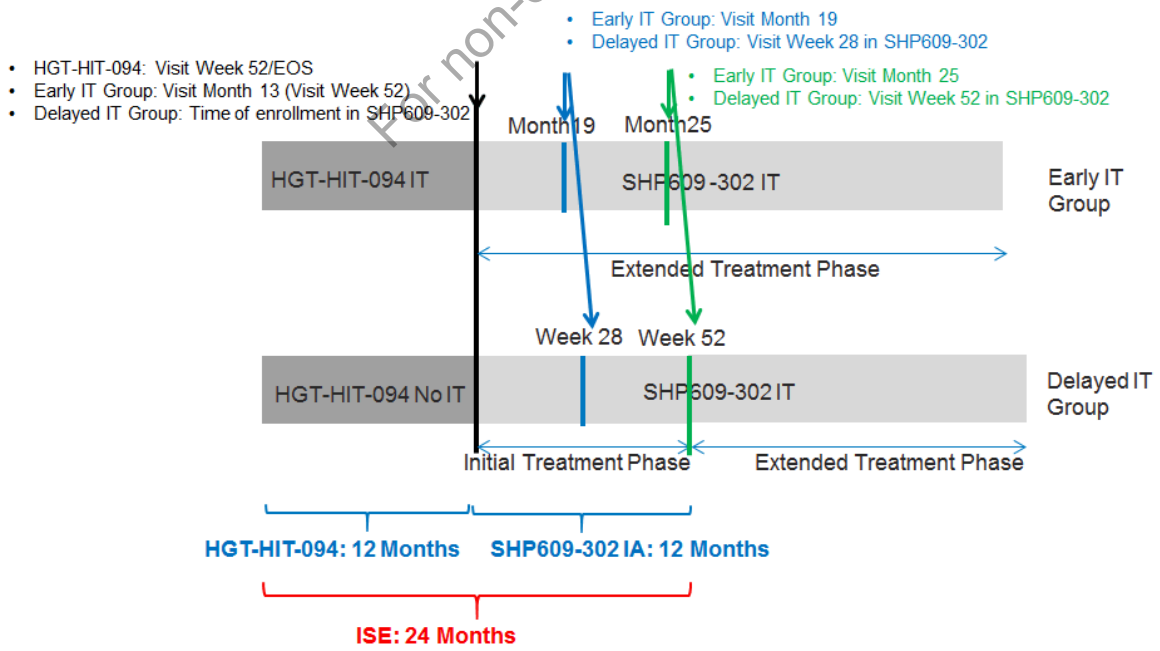
The secondary efficacy endpoints are described above in Section 4.1.

5.4 Exploratory Efficacy Endpoints

5.4.1 Study Visits Used to Define Exploratory Efficacy Endpoints

For exploratory efficacy endpoints described above in Section 4.1, the treatment effect is evaluated at the following scheduled study visits in study SHP609-302 Interim Analysis. The Early IT and Delayed IT groups are defined in Section 6.2, and their designations are based on the treatment regimen (idursulfase-IT or No IT treatment) in the antecedent study HGT-HIT-094.

Figure 1 Scheduled Visits and Duration of Treatment in Studies HGT-HIT-094 and SHP609-302



Scheduled study visits in the Early IT group of study SHP609-302 IA:

- Visit Month 13 (Visit Week 52) in the Extended Treatment Phase of study SHP609-302 for subjects in the Early IT group is the scheduled visit when those subjects have completed a 12-month period of IT treatment in the antecedent study HGT-HIT-094, and comprises of one month for surgical IDDD implantation at the beginning of the study followed by 12 months of IT treatment during HGT-HIT-094.
- Visit Month 19 in the Extended Treatment Phase of study SHP609-302 for subjects in the Early IT group is the scheduled visit when those subjects have completed a 18-month period of IT treatment, and comprises of one month for surgical IDDD implantation at the beginning of HGT-HIT-094, 12 months of IT treatment during HGT-HIT-094 and an additional 6 months of IT treatment in the Extended Treatment Phase of SHP609-302.
- Visit Month 25 in the Extended Treatment Phase of study SHP609-302 for subjects in the Early IT group is the scheduled visit when those subjects have completed a 24-month period (i.e., two 12-month periods) of IT treatment, and comprises of one month for surgical IDDD implantation at the beginning of HGT-HIT-094, 12 months of IT treatment during HGT-HIT-094 and an additional 12 months of IT treatment in the Extended Treatment Phase of SHP609-302.

Scheduled study visits in the Delayed IT group of study SHP609-302:

- Time of Enrollment (Baseline) in SHP609-302 in delayed IT group is the closest available assessment prior to the initial IDDD implant date, which takes place in study SHP609-302, for subjects in the Delayed IT group. This could potentially include EOS visit data from Study HGT-HIT-094.
- Visit Week 16 in SHP609-302 in the Initial Treatment Phase of study SHP609-302 for subjects in the Delayed IT group is the scheduled visit when those subjects have completed a 15-month period (i.e., 12 months of No IT treatment and 3 months of IT treatment), and comprises of 12 months of No IT treatment during HGT-HIT-094, one month for surgical IDDD implantation at the beginning of SHP609-302 and 3 months of IT treatment in the Initial Treatment Phase of SHP609-302.
- Visit Week 28 in SHP609-302 in the Initial Treatment Phase of study SHP609-302 for subjects in the Delayed IT group is the scheduled visit when those subjects have completed a 18-month period (i.e., 12 months of No IT treatment and 6 months of IT treatment), and comprises of 12 months of No IT treatment during HGT-HIT-094, one month for surgical IDDD implantation at the beginning of SHP609-302 and 6 months of IT treatment in the Initial Treatment Phase of SHP609-302.
- Visit Week 40 in SHP609-302 in the Initial Treatment Phase of study SHP609-302 for subjects in the Delayed IT group is the scheduled visit when those subjects have completed a 21-month period (i.e., 12 months of No IT treatment and 9 months of IT treatment), and comprises of 12 months of No IT treatment during HGT-HIT-094, one month for surgical IDDD implantation at the beginning of SHP609-302 and 9 months of IT treatment in the Initial Treatment Phase of SHP609-302.

- Visit Week 52 in SHP609-302 in the Initial Treatment Phase of study SHP609-302 for subjects in the Delayed IT group is the scheduled visit when those subjects have completed a 24-month period (i.e., 12 months of No IT treatment and 12 months of IT treatment), and comprises of 12 months of No IT treatment during HGT-HIT-094, one month for surgical IDDD implantation at the beginning of SHP609-302 and 12 months of IT treatment in the Initial Treatment Phase of SHP609-302.

Matched visits and comparisons between Early IT group and Delayed IT group in study SHP609-302 IA:

- Visit Month 13 vs. Time of Enrollment in SHP609-302 (12-month comparison in HGT-HIT-094), black vertical lines in Figure 1, Visit Month 13 in the Extended Treatment Phase of study SHP609-302 (i.e., 52 weeks from study HGT-HIT-094 baseline) for subjects in the Early IT group versus time of enrollment in SHP609-302 for subjects in the Delayed IT group. This is considered as the beginning of study SHP609-302. For simplicity, the visit names will be referred to as 'Visit Month 13/ Time of Enrollment in SHP609-302' hereafter.
- Visit Month 19 vs. Visit Week 28 in SHP609-302 (18-month comparison), blue vertical lines in Figure 1, Visit Month 19 in the Extended Treatment Phase of study SHP609-302 (i.e., 76 weeks from study HGT-HIT-094 baseline) for subjects in the Early IT group versus Visit Week 28 in the Initial Treatment Phase of study SHP609-302 (i.e., 80 weeks from study HGT-HIT-094 baseline) for subjects in the Delayed IT group. For simplicity, the visit names will be referred to as 'Visit Month 19/ Week 28 in SHP609-302' hereafter.
- Visit Month 25 vs. Visit Week 52 in SHP609-302 (24-month comparison), green vertical lines in Figure 1, Visit Month 25 in the Extended Treatment Phase of study SHP609-302 (i.e., 100 weeks from study HGT-HIT-094 baseline) for subjects in the Early IT group versus Visit Week 52 in the Initial Treatment Phase of study SHP609-302 (i.e., 104 weeks from study HGT-HIT-094 baseline) for subjects in the Delayed IT group. For simplicity, the visit names will be referred to as 'Visit Month 25/ Week 52 in SHP609-302' hereafter.
- Visit Month 19 vs. Visit Week 28 from Visit Month 13/ Time of Enrollment in SHP609-302 (6-month comparison in SHP609-302), from the black vertical line to the blue vertical lines in Figure 1, from Visit Month 13 to Visit Month 19 in the Extended Treatment Phase of study SHP609-302 for subjects in the Early IT group versus from time of enrollment in study SHP609-302 to Visit Week 28 in SHP609-302 in the Initial Treatment Phase of study SHP609-302 for subjects in the Delayed IT group. For simplicity, the visit names will be referred to as 'From Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 19/ Week 28 in SHP609-302' hereafter.
- Visit Month 25 vs. Visit Week 52 from Visit Month 13/ Time of Enrollment in SHP609-302 (12-months comparison in SHP609-302), from the black vertical lines to the green vertical lines in Figure 1, from Visit Month 13 to Visit Month 25 in the Extended Treatment Phase of study SHP609-302 for subjects in the Early IT group versus from time of enrollment in study SHP609-302 to Visit Week 52 in SHP609-302 in the Initial Treatment Phase of study SHP609-302 for subjects in the Delayed IT group. For

simplicity, the visit names will be referred to as 'From Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302' hereafter.

It is noted that Visit Month 25 vs. Visit Week 52 in SHP609-302 (also Visit Month 19 vs. Visit Week 28 in SHP609-302) are 4 weeks apart for the two treatment groups. This difference is due to the need for subjects in the Delayed IT group to undergo surgical IDDD implantation of the device at the beginning of study SHP609-302. Therefore, the first dose of idursulfase-IT was administered 4 weeks later to subjects in the Delayed IT group compared with subjects in the Early IT group who were randomized to IT treatment and underwent IDDD implantation in HGT-HIT-094 prior to enrolling in the Extended Treatment Phase of SHP609-302. For evaluation of the long-term efficacy of IT treatment in SHP609-302, this 4-week difference is considered to have minimal impact on neurodevelopmental assessment scores.

5.4.2 Thresholds of Clinically Meaningful Change

For each of the neurodevelopmental assessment scores, a putatively clinically meaningful change threshold is established using standard error of measurement (SEM), based on test-retest reliability between the baseline and the first post-baseline visit (Month 3 visit) of all pooled SHP609 clinical trial data. The pooled clinical trial data include the natural history study HGT-HIT-090, the Phase I/II study HGT-HIT-045 and the Phase II/III pivotal study HGT-HIT-094. Data from extension studies HGT-HIT-046 and SHP609-302 are not included for establishing the SEM since those subjects are the same subjects from the studies HGT-HIT-045 and HGT-HIT-094, respectively.

Using 2 SEMs to denote a putatively clinically meaningful change, the threshold for each neurodevelopmental assessment score is summarized in Table 1 below. For example, for the GCA standard scores, a putatively clinically meaningful change is considered to be 10 points, where SEM is calculated as 4.5 (rounded up to 5 points).

These putatively clinically meaningful change thresholds are also applied to define the ordered categorical outcomes (3 categories) endpoints in Section 5.4.4 and the binary unreversed floor effect outcome endpoints in Section 5.4.5.

Patient level profile plots of each neurodevelopmental assessment scores will be provided to assess GCA scores and other neurodevelopmental assessment scores during the 24-month period.

Table 1 Putatively Clinically Meaningful Change and Floor Effect Threshold for DAS-II and VABS-II Assessment Scores

Neurodevelopment Assessment Scores	SEM (Approximately)	Putatively Clinically Meaningful Change	Lowest Reportable Score	Floor Effect Threshold
DAS-II GCA standard scores from Early Years battery	5	10	30	40
DAS-II T scores from Early Years battery	2.5	5	10	15
DAS-II standard clusters scores and composite scores from Early Years battery	5	10	30	40
DAS-II GCA standard scores (including both Early Years and School Age batteries)	5	10	30	40
VABS-II ABC standard scores	4	8	20	28

Note: SEM= Standard Error of Measurement

5.4.3 Change over a 12-Month Period Endpoints

Change in scores from the time of enrollment in SHP609-302 to 12 months during study SHP609-302, estimated by linear regression from Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302, for:

- DAS-II GCA standard scores from Early Years battery
- DAS-II T scores from Early Years battery: Verbal Comprehension, Picture Similarities, Naming Vocabulary, Pattern Construction, Matrices, and Copying
- DAS-II standard cluster scores and composite scores from Early Years battery: Verbal, Nonverbal, Spatial and Special Nonverbal Composite (SNC)
- DAS-II GCA standard scores (including both Early Years and School Age batteries)
- Vineland Adaptive Behavior Scales, Second Edition (VABS-II) Adaptive Behavior Composite (ABC) standard scores

5.4.4 Ordered Categorical Outcomes (3 Categories) Endpoints

The ordered categorical outcomes (3 categories: above average cognitive development, average cognitive development, and below average cognitive development) for each subject at 12-month visit in SHP609-302 (i.e., Visit Month 25/Week 52 in SHP609-302) from the time of enrollment in SHP609-302 are defined for the following assessments

- DAS-II GCA standard scores from Early Years battery
- DAS-II GCA standard scores (including both Early Years and School Age batteries)

The categorical outcomes will be determined by comparing the baseline values of neurodevelopment assessment score from the time of enrollment in SHP609-302 +/- the putatively clinical meaningful change thresholds to the 12-month post-baseline values of

corresponding neurodevelopment assessment score. For DAS-II GCA standard scores, the categorical outcome endpoints are defined below:

- Above average cognitive development (Category 1) is defined as subjects who had the observed GCA at Visit Month 25/Week 52 in SHP609-302 more than 10 points (2 SEM) higher than observed GCA at Visit Month 13/ Time of Enrollment in SHP609-302, i.e.,
  - GCA (Month 25/Week 52) > GCA (Month 13/ Time of Enrollment in SHP609-302) + 2 × SEM (10 points for GCA scores)
- Average cognitive development (Category 2) is defined as subjects who had the observed GCA at Visit Month 25/Week 52 in SHP609-302 within the range of ± 10 points (2 SEM) (inclusive) of the observed GCA at Visit Month 13/ Time of Enrollment in SHP609-302, i.e.,
  - GCA (Month 13/ Time of Enrollment in SHP609-302) -2 × SEM (10 points for GCA scores) ≤ GCA (Month 25/Week 52 in SHP609-302) ≤ GCA (Month 13/ Time of Enrollment in SHP609-302) +2 × SEM (10 points for GCA scores)
- Below average cognitive development (Category 3) is defined as subjects who had the observed GCA at Visit Month 25/Week 52 in SHP609-302 more than 10 points (2 SEM) below the observed GCA at Visit Month 13/ Time of Enrollment in SHP609-302, i.e.,
  - GCA (Month 25/Week 52 in SHP609-302) < GCA (Month 13/ Time of Enrollment in SHP609-302) – 2 × SEM (10 points for GCA scores)

5.4.5 Binary Unreversed Floor Effect Outcomes Endpoints

For each of the neurodevelopmental assessment scores, there may be a "floor effect". "Floor effect" is considered to be the lower limitation of the assessment tool, below which the assessment may not be reliable or meaningful. The term "floor effect" is used to define an endpoint, where we classify the status of unreversed neurodevelopment decline. Subjects who reach this score are not expected to achieve improvement of neurodevelopment ability.

The binary unreversed floor effect outcome is defined as whether or not a subject has reached an unreversed floor effect status by 12 months (assessed in SHP609-302) from the time of enrollment in SHP609-302. An unreversed floor effect status for a subject means that the subject has achieved a decline in the neurodevelopmental assessment score to below the floor effect threshold (Table 1 in Section 5.4.2) and has not reversed to equal or above the threshold as of the 12-month visit (i.e., Visit Month 25/Week 52 in SHP609-302). Of note, if a subject reached below the floor effect threshold at the 12-month visit only, the subject will still be considered as having reached the unreversed floor effect by 12 months. For missing data, the last available score carried forward approach will be implemented.

The binary unreversed floor effect outcomes will be defined for the following assessments:

- DAS-II GCA standard scores from Early Years battery
- DAS-II T scores from Early Years battery
- DAS-II standard cluster scores and composite scores from Early Years battery
- DAS-II GCA standard scores (including both Early Years and School Age batteries)
- VABS-II ABC standard scores

5.5 Health Economics and Outcomes Research Endpoints

The health economics and outcomes research endpoints are described above in Section 4.1.

5.6 Pharmacokinetic and Pharmacodynamic Endpoints

The pharmacokinetic and pharmacodynamic endpoints are described above in Section 4.1.

5.7 SOPH-A-PORT Mini S Assessments

The SOPH-A-PORT Mini S device will be evaluated using assessments of device implantation, device function, device longevity, and AEs associated with the implant surgery or device. These data will be collected on the patient's case report form (CRF) from the time of initial implantation.

6. STATISTICAL ANALYSIS

6.1 General Methodology

All statistical analyses will be performed by the Biostatistics and Statistical Programming Department at Shire or its designated CRO unless otherwise specified, using SAS® software version 9.3 or higher (SAS Institute, Cary, N.C., USA).

Data from studies SHP609-302 and HGT-HIT-094 will be integrated for efficacy and safety analyses. Baseline of study SHP609-302 for the subjects previously IT treated in HGT-HIT-094 will be the same as the baseline defined in HGT-HIT-094. Baseline of study SHP609-302 for the previously IT untreated subjects in HGT-HIT-094 will be the closest available assessment prior to the initial IDDD implant date, which takes place in Study SHP609-302, unless otherwise specified. This could potentially include EOS visit data from Study HGT-HIT-094. The analyses presented here will include the data measured at and after baseline of study SHP609-302 as defined above.

Safety data descriptive summaries will be presented by the Early IT group, the Delayed IT group and overall. The Early IT group and Delayed IT group designations are based on the treatment regimen (idursulfase-IT or No IT treatment) in the antecedent study (HGT-HIT-094). For analysis of safety, medical history, concomitant medications and concomitant therapies, any former HGT-HIT-094 substudy subjects will be included in the Early IT treated group.

Efficacy data descriptive summaries will be presented separately for subjects who enrolled from the pivotal study or substudy of HGT-HIT-094. For subjects enrolled from the pivotal study, descriptive statistics will be presented by the Early IT group, the Delayed IT group and overall. For secondary efficacy endpoints, the mean difference in the change at each time point between the two treatment groups (Early IT group and Delayed IT group) and the corresponding 90% confidence interval of the mean difference will be presented where appropriate. The mean values (± SD) for all efficacy endpoints will be graphed over time where appropriate.

Summary statistics for continuous variables will include the n, mean, standard deviation, median, minimum and maximum. Categorical variables will be summarized in a contingency table by the frequency and percentage of subjects in each category.

For subjects enrolled from HGT-HIT-094 substudy, their efficacy data will be presented in listings.

Listing will be provided as per Shire's standard recommended listings. Additional listings may be produced as deemed appropriate.

An interim clinical study report is planned to describe the results of the interim analysis.

6.2 Data Integration Strategy and Analysis Populations

6.2.1 Data Integration Strategy

Data from studies SHP609-302 and HGT-HIT-094 will be integrated for data analyses.

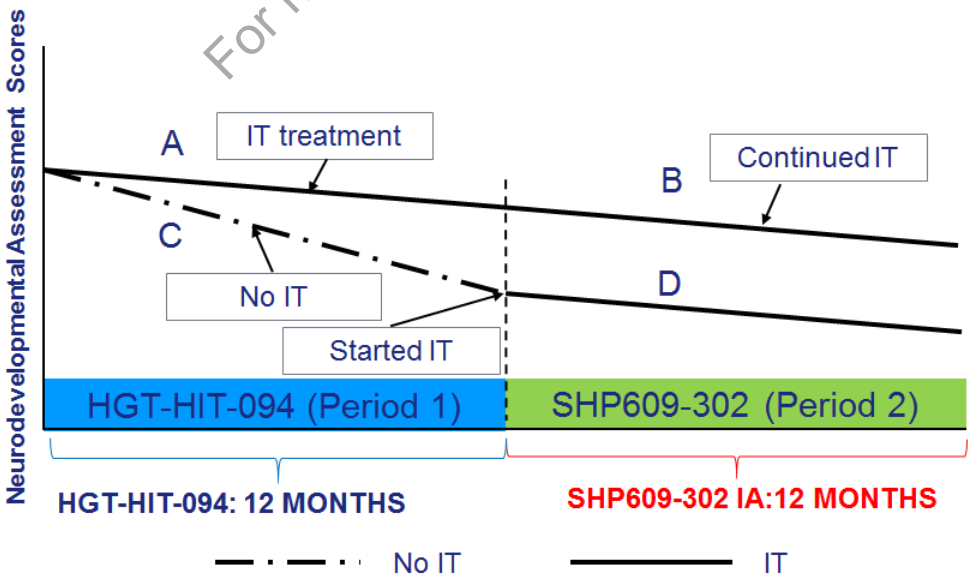
For safety and efficacy descriptive summary analyses using integrated data, the two treatment groups are Early IT group and Delayed IT group based on the HGT-HIT-094 treatment regimen (idursulfase-IT or No IT Treatment), defined as follows:

- Early IT group: subjects who were randomized to the idursulfase-IT treatment cohort in study HGT-HIT-094 and continued in study SHP609-302. HGT-HIT-094 substudy will be included in the Early IT group.
- Delayed IT group: subjects who were randomized to the control cohort (No IT treatment) in study HGT-HIT-094 and began IT treatment in study SHP609-302

Baseline of study SHP609-302 for the subjects in Early IT group will be the same as the baseline defined in HGT-HIT-094. Baseline of study SHP609-302 for subjects in Delayed IT group will be the closest available assessment prior to the initial IDDD implant date, which takes place in study SHP609-302. This could potentially include EOS visit data from Study HGT-HIT-094.

For exploratory efficacy analyses, only data from the HGT-HIT-094 pivotal study subjects will be included because the substudy subjects were evaluated by BSID-III in the 12-month HGT-HIT-094 substudy and their GCA scores at the HGT-HIT-094 baseline and the time of enrollment in SHP609-302 may not available. Only 12-month data collected in study SHP609-302 from the time of enrollment in SHP609-302 at Visit Month 25/Week 52 in SHP609-302 will be used, except the correlation analysis between different neurodevelopmental assessments (section 6.7.4.2); the two treatment groups are the Early IT group (excluding substudy subjects) and Delayed IT group. The 12-month data from Visit Month 13/Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302 will be used for Early IT group and Delayed IT group, respectively.

Figure 2 Schematic Relationship in Studies HGT-HIT-094 and SHP609-302



The following nomenclature will be used to identify “A”, “B”, “C”, and “D” in Figure 2.

- Observation Groups A and C identify the period of observation during study HGT-HIT-094, for subjects randomized to IT treatment and No IT treatment, respectively.
- Observation Groups B and D identify the period of observation during study SHP609-302 up to completion of one year, for subjects who were randomized in HGT-HIT-094 to receive IT treatment and No IT treatment, respectively, noting that all subjects in SHP609-302 received IT treatment.

The subjects in "B" are the same subjects as in "A" with periods of observation from studies SHP609-302 and HGT-HIT-094, respectively, as we can view "B" as a continuation of "A". Subjects in "A" and "B" constitute the Early IT group. The subjects in "D" are the same subjects as in "C" with periods of observation from studies SHP609-302 and HGT-HIT-094, respectively, as we can view "D" as a continuation of "C". Subjects in "C" and "D" constitute the Delayed IT group.

"D" subjects are therefore similar to "A" subjects since these subjects all transitioned from receiving weekly IV infusion of Elaprase, to also receive monthly IT idursulfase. Since study SHP609-302 did not have baseline GCA inclusion criteria at the time of enrollment in SHP609-302, such as employed for Study HGT-HIT-094, "B" and "D" subjects could differ with regard to their baseline GCA scores at the beginning of the observation period. Therefore, for consideration of "B" and "D", only subjects meeting the key inclusion criteria to HGT-HIT-094 will be included in efficacy analyses, i.e., with GCA scores between 55 and 85 inclusive at the time of enrollment in SHP609-302 (i.e., Visit Month 13/ Time of Enrollment in SHP609-302).

6.2.2 Analysis Populations

All descriptive summary analyses of safety and efficacy data will be based on the Safety Population, which is defined as all subjects in study SHP609-302 who underwent IDDD implantation or received at least 1 dose of study drug (full or partial). Device-related analyses will be conducted in the set of patients in the Safety Population who had the device implant procedure performed. Data from "A", "B" and "D" will be used.

Descriptive analyses of health economics and outcomes research endpoints and pharmacodynamic endpoints will be performed using the Safety Population and presented separately for subjects enrolled from the pivotal study or substudy of HGT-HIT-094. Data from "A", "B" and "D" will be used.

All pharmacokinetic data analyses will be performed using the PK population. Data from "A", "B" and "D" will be used.

The exploratory efficacy analyses, i.e., the rate of change (weighted) analysis, the ordered categorical outcomes analysis and the floor effect analysis, will be conducted on "HGT-HIT-094 Comparable Set", "HGT-HIT-094 Comparable Subset 1" and "HGT-HIT-094 Comparable Subset 2". Data from "B" and "D" will be used. The correlation analysis will be conducted on the Safety Population.

The primary efficacy analysis population is HGT-HIT-094 Comparable Subset 1, which is the most informative patient population (e.g. youngest and least affected patients at baseline) with the most potential to benefit from the idursulfase-IT treatment.

Based on the analysis results from study HGT-HIT-094, we do not expect idursulfase IT to provide a treatment benefit for the older subjects (≥6 years) and for those subjects with cognitive abilities below a GCA of 55 at baseline. Therefore the efficacy analyses in the ISE will focus on the younger subjects (< 6 years) with cognitive abilities above a GCA of 55 at baseline, referred to as HGT-HIT-094 Comparable Subset 1. The more general population (subjects with GCA scores between 55 and 85 at the time of enrollment in SHP609-302) is defined in the SHP609-302 IA, the HGT-HIT-094 Comparable Set as an exploratory overall analysis population. The primary efficacy analysis population is the HGT-HIT-094 Comparable Subset 1, defined as a subgroup of HGT-HIT-094 Comparable Set, limited to subjects of age <6 years old at the time of enrollment in SHP609-302. The HGT-HIT-094 Comparable Subset 2 is a subgroup of even younger subjects, i.e., age <55 months old at the time of enrollment in SHP609-302.

The definition of each analysis populations used in study SHP609-302 IA are listed in Table 2.

Table 2 Definition of analysis populations in SHP609-302 IA SAP

Analysis Population	Definition	Intended Usage
Safety Population	All subjects in study SHP609-302 who underwent IDDD implantation or received at least 1 dose of study drug (full or partial)	Primary safety population
HGT-HIT-094 Comparable Set	"B" and "D" subjects in the Safety Population with GCA scores between 55 and 85 inclusive at the time of enrollment in SHP609-302 (i.e., Visit Month 13/Baseline)	Exploratory overall analysis population
HGT-HIT-094 Comparable Subset 1	"B" and "D" subjects in the Safety Population with GCA scores between 55 and 85 inclusive and age <6 years old at the time of enrollment in SHP609-302 (i.e., Visit Month 13/Baseline)	**Primary efficacy analysis population**
HGT-HIT-094 Comparable Subset 2	"B" and "D" subjects in the Safety Population with GCA scores between 55 and 85 inclusive and age <55 months old at the time of enrollment in SHP609-302 (i.e., Visit Month 13/Baseline)	Subgroup of primary efficacy analysis population
Pharmacokinetic Population	All subjects in SHP609-302 who received study drug and participated in the scheduled pharmacokinetic studies, and for whom at least 1 post-dose PK blood sample was collected	PK population

6.3 Patient Disposition

Patient disposition (signed informed consent, patient population [Safety Population and PK Population], treatment status [completed, discontinued/withdrew]) will be presented in summary tables using number and percentage of subjects by the Early IT group, the Delayed IT group for antecedent pivotal study, antecedent substudy and overall. Reasons for discontinuation/withdrawal will be presented.

6.4 Protocol Deviations

An incident involving noncompliance with the protocol, but one which typically does not have significant effects on the patient's rights, safety, or welfare, or the integrity of the resultant data will be considered a protocol deviation.

Protocol violations will be defined as any major protocol deviation that affects study evaluations. Patients will be examined on a case-by-case basis prior to final database lock to determine if conditions set forth in the study protocol have been violated. The complete list of protocol deviations will not be summarized; however, a list of major protocol deviation will be presented.

6.5 Demographic and Other Baseline Characteristics

Demographic and baseline characteristics (e.g., age [years], race, ethnicity, weight [kg], height [cm], and baseline disease status) will be reported in summary tables using the Safety Population by the Early IT group, the Delayed IT group for the antecedent pivotal study, antecedent substudy and overall.

For patients previously receiving active treatment in study HGT-HIT-094, age (in years) = (date of event from study HIT-094 – birth date + 1)/365.25. Date of event is defined as randomization date for the antecedent pivotal study patients, and date of eligibility (enrollment date) for the antecedent substudy patients. For patients previously receiving no IT treatment in study HIT-094, age (in years) = (date of informed consent for the Initial Treatment Phase of study SHP609-302 – birth date + 1)/365.25. When age is presented in the table, it will be rounded down to the first decimal point. Since height and weight may be measured for multiple times during a visit, the average, if any, of these measurements for the same visit will be used as the value for that visit.

6.6 Treatment Compliance and Extent of Exposure

The total number of doses of study drug, the number of doses received via IDDD, the number of doses received via LP, the actual average dose, the actual average duration of IT administrations, the duration of idursulfase-IT treatment, and the percent doses received will be summarized by the Early IT group, the Delayed IT group for the antecedent pivotal study, antecedent substudy and overall using the Safety Population.

The duration of idursulfase-IT treatment, summarized in months, is defined as the time from the first IT administration to the last administration.

The duration for each idursulfase-IT administrations (in minutes) is calculated by subtracting the administration start time from the administration end time. The actual average duration of idursulfase-IT administration (in minutes) will be calculated as an average value of the idursulfase-IT administrations across all available administrations for the same subject.

Treatment compliance will be summarized by the Early IT group, the Delayed IT group for the antecedent pivotal study, antecedent substudy and overall using the Safety population. Treatment compliance is defined as the percent of scheduled doses received for each subject, which is

[(Number of Complete IT injections Received) ÷ (Expected Number of IT injections at data cutoff date)] * 100.

6.7 Analysis of Efficacy

All efficacy descriptive summaries described below refer to the Safety Population who enrolled from the pivotal study of HGT-HIT-094. All exploratory efficacy analyses described below refer to the HGT-HIT-094 Comparable Set, HGT-HIT-094 Comparable Subsets 1 and 2, except the correlation analysis using the Safety Population (section 6.7.4.2). All efficacy data for subjects who enrolled from the antecedent substudy will be listed, including measurements from the BSID-III, VABS-II, DAS-II if available, health economics and outcomes research data, and pharmacodynamics data.

6.7.1 Primary Efficacy Analysis

Not applicable.

6.7.2 Secondary Efficacy Analysis

For the parameters listed in each subsection below except as noted, the following analyses will be done:

- The observed values and change from baseline will be summarized at the scheduled study visits, unless specified otherwise. The mean difference in the change at each time point between the two treatment groups (the Early IT group, the Delayed IT group) and the corresponding 90% confidence interval of the mean difference will be presented.
    - During the HGT-HIT-094/Initial Treatment Phase, descriptive statistics of observed values and changes from baseline will be summarized at Screening/Baseline, on Pre-Treatment Day 1 of IT Dosing Week 16, 28 and 40 visits.
    - During the Extended Treatment Phase, descriptive statistics of observed values and changes from baseline will be summarized on Pre-Treatment of IT Day 1 of Month 13 and at 6-month intervals on the Pre Treatment Day of Monthly IT Dosing Weeks (i.e., at Months 19, 25, 31, 37, 43, 49, 55 and 61).
- The mean value (± SD) for each endpoint over time will be graphed.
- A spaghetti plot of the age equivalent scores for individual subjects will be plotted against chronological age.
- A spaghetti plot of each endpoint over time for individual subjects will be plotted.

Chronological age, when presented, is the age of the subject at the time of the test. It will be calculated by (test date – birth date + 1)/365.25 and 2 decimals will be kept in the dataset.

6.7.2.1 Differential Ability Scale, second edition (DAS-II) and Bayley Scales of Infant Development (BSID-III) Analyses

DAS-II

Descriptive statistics at observed values and changes from baseline will be summarized at scheduled visits for the following:

- Standard scores in subjects who had taken DAS-II test (including both Early Years and School Age batteries) for each cluster and composite:
  - Verbal
  - Nonverbal
  - Spatial
  - GCA
  - SNC

- T scores, age equivalents and developmental quotients in subjects who had taken a DAS-II test for Early Years for each core subtest, where DQ=Age equivalent*100/Chronological age:

  - Verbal Comprehension (Vcom)
  - Picture Similarities (Psim)
  - Naming Vocabulary (Nvoc)
  - Pattern Construction (Pcon)
  - Matrices (Mat)
  - Coping (Copy)

- T scores, age equivalents and developmental quotients in subjects who had a DAS-II test for School Age for each core subtest:
  - Recall of Designs (RDes)
  - Word Definitions (WDef)
  - Pattern Construction
  - Matrices (Mat)
  - Verbal Similarities (VSim)
  - Sequential & Quantitative Reasoning (SQR)

A summary table of neurodevelopment assessment type for each subject will be presented.

Mean values (± SD) by the Early IT group and the Delayed IT group will be plotted over time. This includes the following endpoints:

- The observed values and change from baseline to each scheduled visit in the standard scores for GCA, SNC, each cluster, T scores, age equivalents and DQ for the core subtests of the DAS-II Early Years and the DAS-II School Years

The observed values of age equivalents will be plotted for individual subjects against chronological age. This includes the following endpoints:

- Age equivalents for the core subtests of the DAS-II Early Years and DAS-II/School Years

The observed values and change from baseline values of each endpoint for individual subjects will be plotted over time. This includes the following endpoints:

- Standard scores for each cluster, T scores, age equivalents and DQ for the core subtests of the DAS-II/Early Years and the DAS-II/School Years

Empirical cumulative distribution function (eCDF) plots for the change from Visit Month 13/Time of Enrollment in SHP609-302 at 12 months (Visit Month 25/Week 52 in SHP609-302) in DAS-II GCA standard scores from Early Years battery (GCA EY), between the Early IT and Delayed IT groups will be provided for each analysis population HGT-HIT-094 Comparable Set, Comparable Subset 1 and 2.

Listings will be presented for the above parameters as well as the parameters below:

- Standard scores / percentiles in each subject who had taken a DAS-II test for GCA, SNC and each cluster
- T scores, percentiles of the T scores and age equivalents (directly from eCRF) versus chronological age as well as Developmental Quotient (DQ) in each subject who had taken a DAS-II test for Early Years for each core subtest
- T scores, percentiles of the T scores and age equivalents versus chronological age as well as DQ in each subject who had taken a DAS-II test for School Age for each core subtest

A listing of neurodevelopment assessment type for each subject will be presented.

BSID-III

A listing for the antecedent substudy population will be presented by subject and visit for the following parameters:

- Raw score, age equivalents versus chronological age, as well as DQ, composite score and percentile for each subject evaluated by the BSID-III test and each domain
    - Motor (Fine, Gross)
    - Language (Receptive, Expressive)
    - Cognitive

6.7.2.2 Vineland Adaptive Behavior Scales, Second Edition (VABS-II)

Descriptive statistics at observed values and changes from baseline will be summarized at scheduled visits for the following:

- Standard scores of the VABS-II composite and domains:
  - ABC
  - Communication
  - Daily Living Skills
  - Socialization
  - Motor Skills
- Age equivalents, developmental quotients, and v-scale scores of the VABS-II subdomains:
  - Communication (Receptive, Expressive, Written)
  - Daily Living Skills (Personal, Domestic, Community)
  - Socialization (Interpersonal Relationships, Play and Leisure Time, Coping Skills)
  - Motor Skills (Gross, Fine)
- v-scale scores and observed maladaptive levels of the VABS-II Maladaptive Behavior Index and its subscales (Internalizing, Externalizing)

Mean values (± SD) by the Early IT group and the Delayed IT group will be plotted over time. This includes the following endpoints:

- The observed values and change from baseline to each scheduled visit in standard scores for ABC, each domain, v-scale scores, age equivalents and DQ for the subdomains of the VABS-II
- The observed values and change from baseline to each scheduled visit in v-scale scores and the observed maladaptive levels of the VABS-II Maladaptive Behavior Index and its sub-scales

The observed values of age equivalents for the subdomains of the VABS-II will be plotted for individual subjects against chronological age.

The observed values and change from baseline values of each endpoint for individual subjects will be plotted over time. This includes the following endpoints:

- Standard scores in subjects evaluated by the VABS-II test for ABC and each domain
- v-scale scores, age equivalents, chronological age and DQ for subjects evaluated by the VABS-II test for each subdomain

Listings will be presented for the following parameters:

- Standard scores and percentile in subjects evaluated by the VABS-II test for ABC and each domain
- v-scale scores, age equivalents, chronological age and DQ for subjects evaluated by the VABS-II test for each subdomain

- v-scale scores and Maladaptive levels of Maladaptive Behavior Index and its sub-scales (Internalizing, Externalizing) in VABS-II

6.7.3 Subgroup Analyses

Subgroup descriptive summary analyses of the secondary efficacy endpoints, change from baseline in GCA (including both Early Years and School Age batteries) and ABC scores at scheduled visits will be performed for baseline GCA classification groups (either ≤70 or >70), baseline age groups (either <6 years or ≥6 years and either <55 months or ≥ 55 months). Descriptive summaries within these subgroups and plots of mean values (± SD) over time will be presented. The baseline of study SHP609-302 is defined in Section 6.2.1.

Subgroup analyses of the exploratory rate of change (weighted) analyses are planned for the following subgroups:

- HGT-HIT-094 Comparable Subset 2 (age <55 months old at the time of enrollment in SHP609-302), which is a subgroup of primary efficacy analysis population;
- GCA classification groups (either ≤70 or >70) at the time of enrollment in SHP609-302 (i.e., Visit Month 13/ Baseline) using HGT-HIT-094 Comparable Set (exploratory overall analysis population).

HGT-HIT-094 Comparable Subset 1 (age <6 years old at the time of enrollment in SHP609-302) is considered as the primary efficacy analysis population. It is a subgroup of HGT-HIT-094 Comparable Set.

6.7.4 Exploratory Efficacy Analyses

6.7.4.1 Exploratory Ordered Categorical Outcome Analysis

Descriptive summary of ordered categorical outcomes will be provided for the 12-month visit in SHP609-302 (i.e., Visit Month 25/Week 52 in SHP609-302) from the time of enrollment in SHP609-302 by the observation group (either "B" or "D") for the HGT-HIT-094 Comparable Set, HGT-HIT-094 Comparable Subsets 1 and 2. The definitions of ordered categorical outcomes endpoints for each subject are described in Section 5.4.4. Sensitivity analysis using a 5-, 15-, and 20- point score change to supplement the proposed 10-point change threshold for the Ordered Categorical Outcomes for DAS-II GCA standard scores from Early Years battery (GCA EY) will be performed as specified below.

For each sensitivity-change threshold (SENS=5, 15 or 20) the categorical outcomes will be determined for each subject by comparing the baseline values of GCA EY from study SHP609-302 +/- SENS to the post-baseline values of GCA EY. The ordered categorical outcomes for each SENS are defined as:

- Above average cognitive development (Sens-Category 1) is defined for subjects who had the observed GCA EY at post baseline visit more than SENS points higher than observed study SHP609-302 baseline GCA EY, i.e.,
    - GCA EY (Post-Baseline) > GCA EY (Baseline) +SENS

- Average cognitive development (Sens-Category 2) is defined for subjects who had the observed GCA EY at a post-baseline visit within the range of ± SENS (inclusive) of the observed study SHP609-302 baseline GCA EY, i.e.,
    - GCA EY (Baseline) -SENS ≤ GCA EY (Post-Baseline) ≤ GCA EY (Baseline) + SENS
- Below average cognitive development (Sens-Category 3) is defined for subjects who had the observed GCA EY at a post-baseline visit more than SENS points below the observed study SHP609-302 baseline GCA EY, i.e.,
    - GCA EY (Post-Baseline) < GCA EY (Baseline) – SENS

Descriptive summary of ordered categorical outcomes defined for each SENS value (5-, 15- and 20-points) will be provided at post-baseline of 12 months (i.e., Month 25/Week 52 in SHP609-302) by Early IT and Delayed IT groups for the HGT-HIT-094 Comparable Set, Comparable Subset 1 and 2.

The 5-point change threshold is approximately 1 SEM and the 15-point change is approximately a 90% CI for Sdiff estimates, providing psychometric and clinical meaning for these changes scores. Shire is not aware of any statistical, clinical or psychometric evidence to support using a 20-point change from baseline threshold for the DAS-II GCA scores. On the contrary, a 20-point change over this time period is considered by neuropsychologists as a quite extreme change in cognitive development. While this SAP is incorporating the suggested sensitivity analyses with all of the thresholds proposed by the Agency (5, 15 and 20 points), the limitations of their clinical interpretations should be noted.

6.7.4.2 Exploratory Analysis to Assess Floor Effect

The number and percentage of subjects who reached an unreversed floor effect by 12-month visit in SHP609-302 (i.e., Month 25/Week 52 in SHP609-302) from the time of enrollment in SHP609-302 will be summarized by the observation group (either “B” or “D”) and overall using the 12-month SHP609-302 data “B” and “D” for the HGT-HIT-094 Comparable Set, HGT-HIT-094 Comparable Subsets 1 and 2. The definitions of binary unreversed floor effect outcomes endpoints are described in Section 5.4.5.

6.7.4.3 Exploratory Rate of Change (Weighted) Analysis

The main exploratory efficacy analyses to examine the 12-month IT treatment effect in study SHP609-302 will be the rate of change (weighted) analyses on the exploratory efficacy endpoints, change in scores from the time of enrollment in SHP609-302 to 12 months during study SHP609-302, estimated by linear regression from Visit Month 13/Baseline to Visit Month 25/Week 52 in study SHP609-302, for the select neurodevelopment assessment scores.

The 12-month rate of change (weighted) analysis on DAS-II GCA standard scores from Early Years battery can be considered as a two-stage procedure [3] and is described below. For other neurodevelopmental assessment scores, a similar two-stage procedure will be followed.

- The first stage involves estimation of the rate parameter (slope) for each subject separately, which is the linear regression slope of GCA scores from Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302 (12-month SHP609-302 observation period "B" or "D") for each individual subject. For "B" or "D", the subjects who meet any criteria below will not be included in the linear regression analyses:
    - Have less than 3 available GCA scores in that period. We require 3 or more points for the conventional fit of linear regression. A linear regression with only 2 available scores will result in zero degree of freedom for errors, so we will not have a weight estimate.
        - For sensitivity analysis, subjects with two available GCA scores, at the beginning and at the end of the 12-month period, (i.e., Visit Month 13/ Time of Enrollment in SHP609-302 and Visit Month 25/ Week 52 in SHP609-302) will be included and assigned the smallest weight observed among subjects with at least 3 observations in the second stage.
    - Have missing GCA scores either at the beginning or at the end of the 12-month period (i.e., Visit Month 13/ Time of Enrollment in SHP609-302 or Visit Month 25/Week 52 in SHP609-302). This criterion is used to avoid estimation of 12-month change using 6-month or 9-month data, essentially extrapolating a 12-month result..
    - Have reached GCA 'floor effect' at the beginning of the 12-month period (i.e., Visit Month 13/Baseline for "B" or "D"). The term "floor effect" is for the purpose of evaluating the baseline eligibility of subjects entering SHP609-302 for inclusion in the analyses for the 12 months endpoints. For example, GCA scores below 30 points are not possible and GCA baseline scores below 40 points will not lead to a reliable estimate of change from baseline at 12 months in study SHP609-302, potentially falsely indicating stability of disease. This subject exclusion is for subjects at or near the bottom of the neurodevelopment assessment scale at the time of enrollment in SHP609-302. Those subjects have little potential to measure further decline using the DAS II, so they would inappropriately appear to have stable cognitive ability based on test scores due to the inability of the instrument to differentiate scores near the bottom of the scale. Including those subjects could potentially lead to analysis bias in favor of treatment effect. This is a limitation of all cognitive ability tests, which cannot provide meaningful measurement below a floor effect threshold. The exclusion of these subjects only impacts the 12 months endpoints.

The GCA score change over a 12-month period from Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302 is estimated by the slope estimate multiplied by 52 weeks directly. The slope unit is change in GCA scores per nominal week.

- The second stage uses the individual rate parameter estimates for treatment comparison based on weighted generalized linear models. The weights are the inverse of estimated variances of GCA score change over a 12-month period from Visit Month 13/Baseline to Visit Month 25/ Week 52, estimated by the inverse of the variance of the slope from each

linear regression divided by squared 52 weeks. When the variance is 0, we replace it with the minimum of the non-zero variances of the slopes across all subjects and periods.

$$\text{Weight} = \frac{1}{\text{Variance (GCA change from Visit Month 13/Baseline to Visit Month 25/Week 52)}} = \frac{1}{52^{\cdot}} \times \frac{1}{\text{Variance (slope)}}$$

To examine the sensitivity of the estimated treatment effect to the weighting scheme, two shrinkage estimators of variance of GCA score change over a 12-month period from Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302 in the form of the James-Stein estimator [4] are used to construct shrinkage weights to check that the estimated treatment effects will not be dominated by a few values with large weights. The sample SAS code is presented in Appendix 10.3.3.

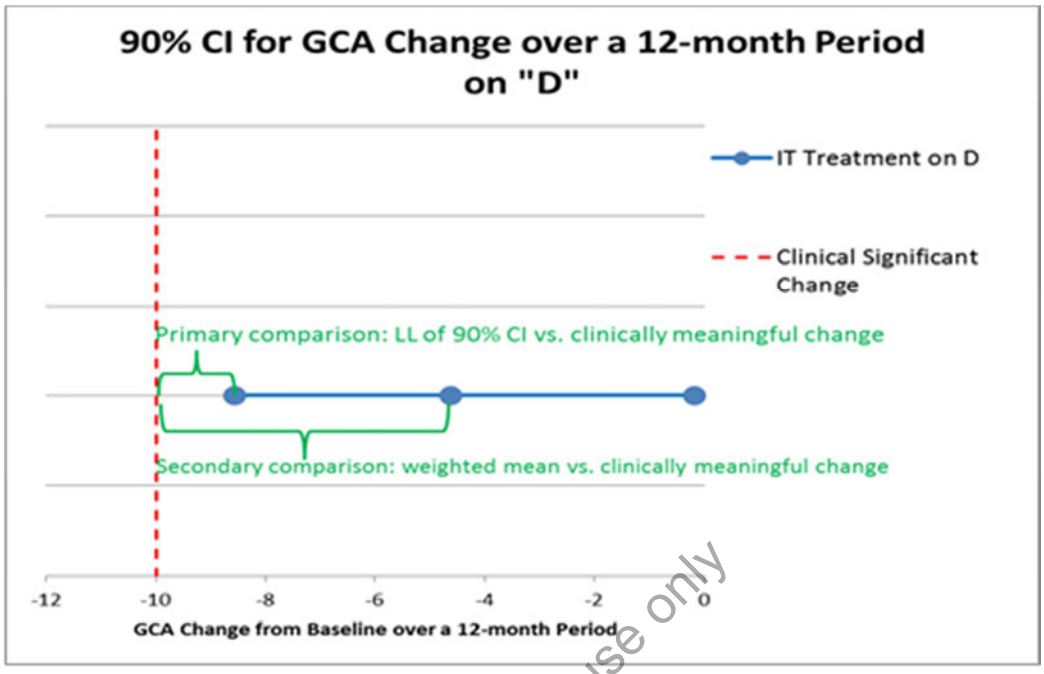
Two analyses will be conducted in the second stage – Analysis 1 and Analysis 2. Analysis 1 is the main exploratory analysis defined in the SHP609-302 IA and Analysis 2 is the supportive exploratory analysis to evaluate 12-month IT treatment effect in study SHP609-302.

Analysis 1 (IT treatment effect in "D"): Examine the initial IT treatment effect in SHP609-302 using weighted ANOVA model. The model will include the fixed categorical effect for GCA classification groups (either ≤70 or >70) at the time of enrollment in SHP609-302 (i.e., Visit Month 13/ Time of Enrollment in SHP609-302).

- Primary comparison: if the lower bound (LL) of 90% confidence interval (CI) of estimated change over a 12-month period from Visit Month 13/Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302 in "D" based on weighted ANOVA model is greater than a clinically meaningful change (Table 1), then the IT treatment would be considered showing effect in "D".
- Secondary comparison: in view of the small sample size and wide confidence interval, if the estimated mean change over a 12-month period in "D" based on weighted ANOVA model is greater than a clinically meaningful change, then the IT treatment would be considered showing effect in "D".

The hypothetical illustration of the primary comparison using lower bound of 90% confidence interval and the secondary comparison using a point estimate of 90% confined interval of Analysis 1 on "D", the main exploratory analysis approach in the SHP609-302 IA SAP, is demonstrated in Figure 3.

Figure 3 Hypothetical Illustration of Comparisons of Analysis 1 on "D"



Note: The primary comparison based on a 1-sample test is to compare the LL of the 90% CI for the estimated change over a 12-month period and the clinically meaningful change (e.g., -10 points for GCA scores). The secondary comparison is to compare the estimated mean change over a 12-month period and the same threshold.

Analysis 2 (IT treatment effect in "B"∪"D"): Compare subjects' initial 12 months of IT treatment with subjects' second 12 months of IT treatment using the rate of change over a 12 month period and examine the overall IT treatment effect in SHP609-302 using weighted ANOVA model. The model will include a group indicator of "B" vs "D" as well as a fixed categorical effect for GCA classification groups (either ≤70 or >70) at the time of enrollment in SHP609-302 (i.e., Visit Month 13/ Time of Enrollment in SHP609-302) in the weighted ANOVA model. In this scenario, the treatment by the group interaction cannot be included because the treatment (Early IT or Delayed IT) and group indicator ("B" or "D") are identical.

- Primary comparison: if the LL of 90% CI of estimated change over a 12-month period in "B" ∪"D" from Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302 based on weighted ANOVA model is greater than the pre-defined clinically meaningful change (Table 1), then the IT treatment would be considered showing effect in "B"∪ "D".
- Secondary comparison: in view of the small sample size and wide confidence interval, if the estimated mean change over a 12-month period in "B" ∪"D" based on weighted ANOVA model is greater than the pre-defined clinically meaningful change, then the IT treatment would be considered showing effect in "B"∪"D".

The premises of the rate of change (weighted) analysis will be examined prior to the analysis. The premises of the rate of change (weighted) analysis include:

- No extreme systematic departure of linearity on residual distribution of overall linear fits of GCA scores change over a 12-month period from Visit Month 13/ Time of Enrollment in SHP609-302 to Visit Month 25/Week 52 in SHP609-302 for each individual subject meeting the first stage criteria
    - Approximately normally distributed residuals centered near zero at each visit time
    - Variance homogeneity of residuals at each visit time
- Overall agreement between the estimated GCA score change over a 12-month period as estimated by a linear regression and the observed GCA score change over a 12-month period in SHP609-302 as calculated by a simple subtraction
- No correlation between weights and estimated GCA score change over a 12-month period in SHP609-302

As model diagnostics, the distribution property of linearity, variance homogeneity at each visit and approximate normality at each visit will be examined using residual plots, boxplots of residuals and normal probability plots of residuals. The overall agreement between estimated GCA score change over a 12-month period by a linear regression and observed GCA score change over a 12-month period in SHP609-302 by a simple subtraction will be examined using scatter plots. The correlation between weights and estimated GCA score change over a 12-month period in SHP609-302 (also between weight and slopes) will be examined using scatter plots. Linear regression lines for individual subjects will be plotted. For rate of change (weighted) analysis on other select neurodevelopmental assessment scores, similar model premises have to be examined prior to the analysis.

The rate of change (weighted) analysis will be performed on HGT-HIT-094 Comparable Set, HGT-HIT-094 Comparable Subsets 1 and 2. The estimated treatment effect and 90% confidence intervals on the treatment effect from the weighted ANOVA model will be reported without reporting p-values and presented in Forest plots.

The sample SAS code is presented in. Appendix 10.3.1

6.7.4.4 Correlation between Different Neurodevelopmental Assessment Scores

The correlation analysis between different neurodevelopment assessment scores is a supportive exploratory analysis performed on the Safety Population who enrolled from the pivotal study of HGT-HIT-094. Data from periods A, B and D in Figure 1 will be used.

- The following correlation coefficients for observed values and change from baseline will be estimated from a linear mixed effects model that takes into account the repeated measures on each subject using the method described by Hamlett et. al.[5] SAS Proc Mixed with restricted maximum likelihood estimation (REML) and an unstructured between and within-subject covariance structure for the random effects will be used for this model. Each pair of parameters will also be graphed as scatter plots. DAS-II (including both Early Years and School Age batteries) GCA standard score and

  - DAS-II Verbal standard score
  - DAS-II Nonverbal standard score
  - DAS-II Spatial standard score
  - DAS-II SNC standard score
  - VABS-II ABC standard score
  - VABS-II Communication standard score
  - VABS-II Daily Living Skills standard score
  - VABS-II Socialization standard score
  - VABS-II Motor Skills standard score
  - CSF GAG levels (natural log scale)

- VABS-II ABC standard score and
  - VABS-II Communication standard score
  - VABS-II Daily Living Skills standard score
  - VABS-II Socialization standard score
  - VABS-II Motor Skills standard score
  - CSF GAG levels (natural log scale)

Each pair of parameters will also be graphed as scatter plots. The sample SAS code is presented in Appendix 10.3.2.

Other exploratory analyses may include assessment of the correlation between PD and efficacy endpoints; and between composite scores and their components (i.e., correlations between the GCA and ABC scores and their respective cluster/domain scores).

Table 3 Summary of Exploratory Efficacy Analysis in the SHP609-302 IA SAP

Analysis	Comparison /Evaluation	Endpoints	Rationale and Additional Considerations	SAP

Supportive Exploratory:
 	
12-month Ordered Categorical Outcome (3 categories) Analysis
	"B" and "D"	**Ordered categorical outcomes** (3 categories: above average cognitive development, average cognitive development, and below average cognitive development) at 12 months (assessed in SHP609-302) from the time of enrollment in SHP609-302, for:
		• **Main**
		  ○ DAS-II GCA scores from Early Years battery
		• **Supportive**
		  ○ DAS-II GCA scores (including both Early Years and School Age batteries)
			Rationale:
			• Descriptive summary of subjects change during the second 12 months of IT treatment ("B")
			• Descriptive summary of subjects change during the first 12 months of IT treatment ("D")
				SHP609-302 IA
				
				Section 6.7.4.1

Supportive Exploratory:
 	
12-month Floor Effect Analysis
	"B" and "D"	**Binary outcomes** for each subject: whether or not the subject reached an unreversed floor effect at 12 months (assessed in SHP609-302) from the time of enrollment in SHP609-302, for:
		• **Main**
		  ○ DAS-II GCA scores from Early Years battery
		• **Supportive**
		  ○ DAS-II T scores from Early Years battery
		  ○ DAS-II standard cluster scores and composite scores from Early Years battery
		  ○ DAS-II GCA scores (including both Early Years and School Age batteries)
		  ○ VABS-II ABC scores
			Rationale:
			• Descriptive summary of subjects reaching an unreversed floor effect during the second 12 months of IT treatment ("B")
			• Descriptive summary of subjects reaching an unreversed floor effect during the first 12 months of IT treatment ("D")
			Additional consideration:
			• "Floor effect" is considered to be the lower limitation of the assessment tool, below which the assessment may not be reliable or meaningful.
			• The summarized number and percentage of subjects who reached the unreversed floor effect will be used to evaluate the floor effects on the efficacy analyses results.
				SHP609-302 IA
				
				Section 6.7.4.2

Table 3 Summary of Exploratory Efficacy Analysis in the SHP609-302 IA SAP

Analysis	Comparison /Evaluation	Endpoints	Rationale and Additional Considerations	SAP

Main Exploratory: 12-month Rate of Change (Weighted) Analysis	"D" (Analysis 1 – Primary Comparison)	Change in scores from the time of enrollment in SHP609-302 to 12 months during SHP609-302 ("D"), **estimated by linear regression**, for:
		• **Main**
		  ○ DAS-II GCA scores from Early Years battery
		• **Supportive**
		  ○ DAS-II T scores from Early Years battery
		  ○ DAS-II standard cluster scores and composite scores from Early Years battery
		  ○ DAS-II GCA scores (including both Early Years and School Age batteries)
		  ○ VABS-II ABC scores
			Rationale:
			• Examine the treatment effect in "D" alone.
			Additional consideration:
			• The primary comparison will be based on a 1-sample test, i.e., between the lower bound of the 90% CI for the weighted mean of change from baseline and a clinically meaningful change.
			• Due to the very small projected sample size in "D", the CI may be too wide. Therefore, a secondary comparison will be made using only a point estimate, i.e., the weighted mean of change from baseline, and compare it to a clinically meaningful change.
			• Since study SHP609-302 did not have baseline GCA inclusion criteria range at the time of enrollment in SHP609-302, such as employed for study HGT-HIT-094, "D" subjects could differ with regard to their baseline GCA scores at the beginning of the observation period. **Therefore, for consideration of "D", only "D" subjects with GCA scores between 55 and 85 inclusive at the time of enrollment in SHP609-302 will be included.**
				SHP609-302 IA
				Section 6.7.4.3

Table 3 Summary of Exploratory Efficacy Analysis in the SHP609-302 IA SAP

Analysis	Comparison /Evaluation	Endpoints	Rationale and Additional Considerations	SAP

Supportive Exploratory: 12-month Rate of Change (Weighted) Analysis	"B" ∪ "D" (Analysis 2 – Supportive Comparison)	Change in scores from the time of enrollment in SHP609-302 to 12 months during SHP609-302 ("B" and "D"), **estimated by linear regression**, for:
		• **Main**
		  ○ DAS-II GCA scores from Early Years battery
		• **Supportive**
		  ○ DAS-II T scores from Early Years battery
		  ○ DAS-II standard cluster scores and composite scores from Early Years battery
		  ○ DAS-II GCA scores (including both Early Years and School Age batteries)
		  ○ VABS-II ABC scores
			Rationale:
			• Compare subjects' initial 12 months of IT treatment ("D") with subjects' second 12 months of IT treatment ("B") using the rate of change over a 12 month period. Note: this is a parallel group comparison. Within subject comparisons (intra-subject) will be performed in the ISE.
			• Estimate the treatment effect using all IT treatment data in SHP609-302 (increased sample size) if "B" and "D" appear similar.
			Additional consideration:
			Method of analysis will be the same as for the main exploratory analysis above.Only "B" and "D" subjects with GCA scores between 55 and 85 inclusive at the time of enrollment in SHP609-302 will be included.
				SHP609-302 IA
				Section 6.7.4.3

Note: The exploratory analyses listed in Table 1 will be performed on HGT-HIT-094 Comparable Set, HGT-HIT-094 Comparable Subsets 1 and 2.

6.7.5 Health Economics and Outcomes Research Endpoint Analyses

The health status of subjects will be assessed using the EQ-5D questionnaire. The EQ-5D measures 5 dimensions of health status: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. For each dimension, there are 5 levels of response:

- No problems
- Slight problems
- Moderate problems
- Severe problems
- Unable to do / Extreme problems

The number and percent of subjects with each response will be presented by dimension at each visit. The visual analogue scale (VAS) records the subject's health rated by parent/caregiver on a 0 (worst health) to 100 (best health) scale. The VAS score, as well as the change from baseline will be summarized at each time point.

Additional pharmacoeconomic analyses may be performed by the Sponsor staff in the Health Economics and Outcomes Research group or designee and reported separately in a pharmacoeconomic report to be appended to the Clinical Study Report. Accordingly, any planned pharmacoeconomic analyses related to this data may be described in a separate document.

6.7.6 Pharmacodynamic Outcome

- CSF and Urine GAG

CSF and Urine GAG observed levels, the change from baseline and percent change from baseline will be summarized by the Early IT group, the Delayed IT group and overall and by visit. Mean observed values (+/- SE) and mean percent change compared to baseline will be plotted over time. The percent change compared to baseline is defined as the value at the corresponding visit multiplied by 100 and divided by the baseline value. One decimal will be kept for the percent compared to baseline. Analysis of both CSF total GAG and CSF heparan sulfate if available will be performed.

- Albumin Concentration and Albumin Quotients

CSF albumin concentrations, CSF/serum albumin quotients (QAlb = CSF albumin/serum albumin) observed levels, the change and the percent change from baseline will be summarized by the Early IT group, the Delayed IT group and overall and by visit and to assess the blood-CSF barrier function.

- During the HGT-HIT-094/Initial Treatment Phase, descriptive statistics at observed values, changes from baseline, and the percent change from baseline will be summarized at Screening/Baseline, Week 4, 16, 28, and 40.
- During the Extended Treatment Phase, descriptive statistics at observed values, changes from baseline, and the percent change from baseline will be summarized on month 13

and at 3-month intervals (i.e., at Months 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58 and 61).

6.8 Analysis of Safety

All analyses of safety data will be descriptive and presented by the Early IT group, the Delayed IT group and overall. The analysis of safety will be based on the Safety population. Any former HGT-HIT-094 Substudy subjects will be included in the Early IT group.

6.8.1 Adverse Events

AEs will be recorded throughout the study and at early termination. AEs and medical conditions will be coded using the MedDRA version 16.1 or higher.

For IA analysis, treatment-emergent AEs (TEAE) are defined as all AEs occurring on or after the date of the first IDDD implant surgery or first dose (whichever is earlier) and on or before the EOS visit (+30 days) or 2 weeks after the removal of the last IDDD if early termination (whichever is later). In general, an AE will be deemed TEAE if it cannot be definitively categorized by the available components (day, month, year) of the AE onset date with respect to the date of intervention (the date of the first IDDD implant surgery or first dose [whichever is earlier]). Summaries for the following TEAE categories will be presented:

1) Patients who experienced no AEs,
2) Patients who experienced at least one AE
3) Patients who discontinued due to an AE(s)
4) Patients who died
5) Patients who experienced at least one serious adverse event (SAE),
6) Patients who experienced at least one severe/life-threatening AE,
7) Patients who experienced at least one IV Elaprase infusion-related adverse event
8) Patients who experienced at least one idursulfase-IT-related adverse event
9) Patients who experienced at least one idursulfase-IT or IV Elaprase infusion-related adverse event
10) Patients who experienced at least one IT Treatment Regimen-related adverse event (i.e., related to one or more of the following: study drug, IDDD, device surgical procedure, and IT-administration process)
11) Patients who experienced at least one IDDD surgical procedure-related adverse event
12) Patients who experienced at least one IDDD-related adverse event
13) Patients who experienced at least one IT Administration Process-related adverse event

The AE categories 1-13 will be presented by each treatment group and overall, while the AE categories 8-13 will be presented for patients in the Safety Population who received idursulfase-IT. IT administration process-related AEs will be summarized by IT administration method (LP or IDDD) as well. The AE categories 11-12 will be presented for patients in the Safety Population with the device implanted.

The number and percentage of patients, and the number of corresponding AEs will be summarized in overall summary tables. In addition, the number and percentage of patients having any treatment-emergent AE within the AE categories 2, 5, 7-13 and the number of corresponding AEs will be displayed by MedDRA system organ class (SOC), preferred term (PT) and treatment group. The most common treatment-emergent AEs which happened in >10% patients in either treatment group will also be summarized by SOC and PT. The most common non-serious treatment-emergent AEs which happened in >5% patients in either treatment group will also be summarized by SOC and PT.

The number and percentage of patients having any treatment-emergent AE, and the number of corresponding AEs will be summarized by severity.

The number and percentage of patients having any treatment-emergent AE, and the number of corresponding AEs will be summarized by relationship to IV Elaprase, idursulfase-IT, and at least one IT treatment regimen, and displayed by MedDRA SOC, PT and treatment dose.

Furthermore, listings of all the AEs of categories 3-6 will be provided if applicable.

6.8.2 Clinical and CSF Laboratory Evaluations

Chemistry and hematology serum laboratory and CSF (standard chemistries, glucose, protein, and cell counts) values will be summarized in terms of the absolute value and change from baseline at each scheduled study visit by the Early IT group, the Delayed IT group and overall.

- During the HGT-HIT-094/Initial Treatment Phase, descriptive statistics at observed values and changes from baseline will be summarized at Screening/Baseline, Pre-Surgery Day 1 of Week 2, at Week 4, 16, 28 and 40 visit.
- During the Extended Treatment Phase, descriptive statistics at observed values and changes from baseline will be summarized on Pre-Treatment of IT Day 1 of Month 13 and at 3-month intervals on the Pre-Treatment Day of Monthly IT Dosing Weeks (i.e., at Months 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55 and 61).

All the laboratory (chemistry, hematology, urinalysis and CSF) values will be categorized as a patient having had (1) a Clinically Significant (CS) value at any time during the study post baseline, and (2) no CS values at any time during the study post baseline. The number and percentage of patients in each category will be presented by the Early IT group, the Delayed IT group and overall. Patients with a Normal or non-CS value for a particular laboratory parameter at baseline who experience a change to CS for that laboratory parameter post-baseline will be identified and listed separately along with the corresponding laboratory values over time. Baseline for the patients previously treated in HGT-HIT-094 is defined as the screening measurement in HGT-HIT-094; if it is not available, the Pre-Surgery Day 1 of Week 1 measurement will be used as baseline. Baseline for the untreated patients in HGT-HIT-094 is the closest available assessment prior to the initial IDDD implant date, which takes place in Study SHP609-302. This could potentially include EOS visit data from Study HGT-HIT-094.

If a particular laboratory measurement has been either repeated or retested, then the repeated or retested measurement for that laboratory parameter, with respect to date/time, will be used in the

statistical analysis unless this value is invalid/missing. The handling of repeated or retested laboratory measurements should only consider the specific laboratory measurement that was repeated or retested.

6.8.3 12-Lead ECG Evaluations

The 12-lead ECG parameters (heart rate [bpm], PR interval [msec], QRS interval [msec], QT interval [msec] and the corrected QT interval (QTc) [msec]) will be summarized in terms of absolute value and change from baseline by the Early IT group, the Delayed IT group and visit. The analysis will be based on the Safety population. Any former HGT-HIT-094 Substudy patients will be included in the Early IT group. The QTc interval will be calculated using Bazett's formula as QT divided by the square root of RR interval. The number and percentage of patients with ECG abnormalities post baseline will be presented by the Early IT group and the Delayed IT group.

- During the HGT-HIT-094/Initial Treatment Phase, descriptive statistics at observed values and changes from baseline will be summarized at Screening/Baseline, Pre-Surgery Day 1 of Week 2, and at 3-month intervals of after injection of study drug on Week 4, 16, 28 and 40.
- During the Extended Treatment Phase, descriptive statistics of observed values and changes from baseline will be summarized on Pre-Treatment of IT Day 1 of Month 13, and at 12-month intervals of after injection of study drug on Monthly IT Dosing Weeks (i.e., at Months 25, 37, 49 and 61).

6.8.4 Vital Signs

The injection (IT) vital signs and regular vital signs (temperature [C], pulse [bpm], blood pressure [systolic and diastolic, mmHg], oxygen saturation, and respiration rate [per min] will be listed.

- During the HGT-HIT-094/Initial Treatment Phase, vital signs will be collected at Screening/Baseline, during Surgery Week 2 at Day 1, Day 2, Days 3-6 (i.e., during IV infusion of Elaprase), Pre-Treatment Day 1 of each IT Dosing Week, Day 2 and Days 3-7 (i.e., during IV infusion of Elaprase) of each IT Dosing Week after injection of the study drug.

At a minimum, vital signs will be collected at the following time points on IT administration of idursulfase-IT: within 15 minutes prior to IT administration, 30 minutes (±10 minutes) post end of IT administration, 60 minutes (±10 minutes) post end of IT administration, 120 minutes (±10 minutes) post end of IT administration, and 4 hours (±10 minutes) post end of IT administration.

- During the Extended Treatment Phase, vital signs will be collected on Pre-Treatment Day 1 and on the IT Injection Day of Month 13 and on Pre-Treatment Day 1 and on the IT Injection Day on Monthly IT Dosing Weeks.

At a minimum, vital signs will be collected at the following time points on IT administration of idursulfase-IT: within 15 minutes prior to IT administration, 30 minutes (±10 minutes) post end

of IT administration, 60 minutes (±10 minutes) post end of IT administration, 120 minutes (±10 minutes) post end of IT administration, and 4 hours (±10 minutes) post end of IT administration.

Vital signs will be plotted for individual patients over the time by scheduled study visit. The IT vital signs will be plotted for individual patients over the time by scheduled study visit from the start of the IT injection for each IT injection.

6.8.5 Physical and Neurological Findings

Clinically significant physical and neurological examination findings will be recorded and summarized either as part of the medical history or adverse event data. No additional summary or listing will be provided.

6.8.6 Other Observations Related to Safety

6.8.6.1 Immunogenicity

Anti-idursulfase antibody formation will be monitored throughout the study for both serum and CSF. The number and percentage of patients testing anti-idursulfase antibody positive and negative at each scheduled time point will be summarized by the Early IT group, the Delayed IT group and overall. The analysis will be based on the Safety population. Any former HGT-HIT-094 Substudy patients will be included in the Early IT group.

Antibody titer values will be plotted for individual patients over time by scheduled study visit in patients with positive antibodies at or prior to each scheduled visit. The neutralizing antibodies (NAb) titer will be plotted similarly for patients who developed positive neutralizing antibodies at or prior to each scheduled visit.

6.8.6.2 CSF Enzyme Level

The observed idursulfase enzyme levels in CSF will be plotted for individual patients before each IT dose over time by scheduled study visit. If a measurement is less than the lower level of quantification (LLOQ), then it will be replaced by zero.

6.8.6.3 Height, Weight, and Head Circumference

Height (cm), weight (kg), and head circumference (cm) observed values, and the change from baseline will be summarized by the Early IT group and the Delayed IT group at scheduled study visits. The analysis will be based on the Safety population. Any former HGT-HIT-094 Substudy patients will be included in the Early IT group. If multiple measurements were taken for the same visit, their mean will be the measurement for that visit.

6.8.6.4 Brain MRI

All patients will have MRI of the brain at Screening and at 12-month intervals, i.e., at months 13, 25, 37, 49 and 61. The analysis will be based on the Safety population. Any former HGT-HIT-094 Substudy patients will be included in the idursulfase-IT treated group. Descriptive statistics at observed values and changes from baseline will be summarized by the Early IT group, the Delayed IT group and overall for the following parameters:

- Brain Total Intracranial Volume (cm3)
- Brain Total Tissue Volume (cm3)
- Brain Total White Matter (cm3)
- Brain Total Gray Matter (cm3)
- Total CSF Volume (cm3)

6.8.6.5 Hearing Assessments

The number and percentage of patients in the hearing assessment categories (right ear hearing aid, left ear hearing aid, and sufficient hearing for study assessments) will be summarized by the Early IT group, the Delayed IT group and visit. The analysis will be based on the Safety population. Any former HGT-HIT-094 Substudy patients will be included in the Early IT group.

6.8.6.6 SOPH-A-PORT Mini S Device Performance

Device-related Terminology

- Initial device implant: The first IDDD implant that a patient ever receives
- Partial device revision: Surgical revision/replacement of one or more component(s) of the device; other component(s) of the original device remain implanted and are not affected (e.g., port revision).
- Full device revision: The device is removed (explanted) in its entirety and a completely new device is implanted, including complete device removal with immediate replacement and delayed device implant after previous removal.
- Complete device removal with immediate replacement: Immediately implant of a new device after a device had been completely removed.
- Delayed device implant after previous removal: Implant of a new device after a previous device had been completely removed without immediate replacement on a separate and earlier occasion.
- Complete device removal without replacement: All parts of the device (both port and catheter) are removed and there is no new device implant.
- Device adjustment: surgical procedure of the device that is not a port replacement or full device revision.
- Device malfunction: The device does not perform as intended, based on the description in the device's Instructions for Use, but does not require either a partial or full device revision. If at the time of a scheduled dosing it is not possible to administer a full medication dosage as per the standard administration steps detailed in the device's Instructions for Use due to a device related issue, the IDDD will be declared a device

malfunction. A device malfunction could be resolved without the need for a full or partial surgical revision. Programmatically it is when the malfunction date is present and outcome of malfunction is resolved.

- Device failure: When the device irreversibly fails to perform as intended and cannot be corrected without a device surgical intervention of either a partial or full device revision or removal. The IDDD will be declared a device failure, starting from the date of the initial malfunction that persists leading to the above surgical intervention. IDDDs that are considered to be malfunctioning at the end of the study will be categorized as Failures. For programming purposes, a device failure is when the malfunction date is present and outcome of malfunction is either ongoing or device failure (Surgical Procedure). The date of the device failure is the date of the initial malfunction.
- IDDD longevity: Total number of IDDD failures and explant divided by the total time at risk across all subjects in weeks (sum of time to IDDD failure or the last injection if not failed at the end of the data cut or explant date, from initial implantation, delayed implantation, or revision for all IDDDs).

Device Performance Analyses

SOPH-A-PORT® safety and performance will be summarized for IDDD-implanted patients in Safety Population. The number and proportion of patients of the following categories and the corresponding event count and event percentage will be summarized.

- Patients who had the initial implant only (i.e., no additional surgeries)
- Patients who had any post-initial implantation device surgeries
- Patients who had any type of difficulty associated with the implant procedure (e.g., difficulty accessing spinal canal, etc.)

IDDD-related surgical procedure details for initial implants, complete removal with immediate replacement, and delayed implantation across all IDDD implantations for all patients will be summarized by failed or not failed IDDDs and overall, including incision region (paramedian vs. other), identification of the catheter passer used (Phoenix Neuro vs. other), number of suture wings implanted, suture wing configuration, interspace for catheter insertion into lumbar spine (L1-2, L2-3, L3-4, L4-5, L5-S1), spinal vertebral level of catheter tip (cervical, thoracic, lumbar and sacral), use of a purse string suture, use of the guidewire, orientation of the Tuohy needle and clinical site.

The number and proportion of patients who had at least one abnormal IDDD radiological assessment finding and the number of abnormal findings from the IDDD radiological assessments will be also summarized by types of the abnormality.

The number and proportion of patients and IDDDs with total malfunctions (including failure and malfunctions) as well as the corresponding event numbers will be presented. The types of total malfunctions and the reasons for IDDD failures reported by the site and the Adjudication Committee will be summarized at the patient, IDDD and event level.

The annual event rate of IDDD failures and malfunctions will be calculated for each patient and the descriptive statistics will be summarized. The overall IDDD failure rate and its 95% CI will be presented. The overall IDDD failure rate is calculated as the total number of IDDD failures for all patients divided by the total IDDD time at risk, which is defined as the total time to IDDD failure or the last injection if IDDD is not failed at the end of the study, from initial implantation, or full device revision for all IDDDs. The overall IDDD malfunction rate is defined in the same manner.

The IDDD longevity (time to failure in weeks) and time to first total malfunction, for all implanted IDDDs, will be analyzed using the KM method. A new port identifies a new IDDD, starting from the date of implantation (either an initial implantation, a partial or full revision, or a delayed device implant after previous removal). The time to IDDD failure (weeks) or total malfunction will be obtained by subtracting the date of the IDDD kit implantation from the date of IDDD failure (i.e., the initial malfunction date that persists leading to surgical intervention) or first total malfunction plus 1, and divided by 7, one decimal will be kept. IDDDs which did not fail or malfunction will be censored at the last study drug injection date for each IDDD. The number of IDDDs at risk, the cumulative number of IDDDs failed, censored, and the cumulative probability of failure with its standard error will be summarized in a table at each event time.

The proportion of patients for whom a successful first injection of investigational drug product occurred will be reported among those for whom a first injection was attempted (i.e., those who had an apparently successful implantation and did not suffer a device removal or revision prior to first scheduled injection). An injection is defined as study drug administered using IDDD. A successful injection is defined as the completion of study drug injection via IDDD without a malfunction. The proportion of patients who had no unsuccessful injection attempts during the study will also be summarized. Injections that are not successful for patient-related reasons (e.g., patient uncooperative, competing medical issue) will not be included in the determination of the injection success rate. The detailed list of the "Other" malfunction types will be reviewed by the Shire physician to determine if the injection was not given for patient reasons. If the injection is not given for patient reasons then it would not be counted as an unsuccessful injection attempt.

Whether any device component was implanted will be summarized for all IDDD kits which were opened. For the implanted IDDDs, their status (i.e., removed or not) at the end of the study will be reported. If an IDDD was removed, whether the catheter was removed along with the port will also be summarized.

An IDDD timeline of events (in weeks from implantation) including the timing of device surgeries (device adjustment, partial or full revision, and complete device removal without replacement), and malfunction and failure by IDDDs will be plotted by patient sorted by the duration since the initial implantation, longest on top.

The number of surgeries per patient will be presented by a bar chart. The types of IDDD malfunctions and failure per patient, as well as the reasons for the failures for the IDDDs reported by the site and the Adjudication Committee will be presented similarly by a bar chart.

The time from initial implant surgery to first IDDD failure and the time to first malfunction will be analyzed using the Kaplan Meier (KM) method for each patient who is in the Safety Population with IDDD implanted. The time to first IDDD failure or malfunction will be obtained by subtracting the date of the first IDDD implantation for the patient from the date of first failure or malfunction plus 1, divided by 7, and one decimal will be kept. Patients without an IDDD failure or malfunction will be censored at their last study drug injection date. The 25th percentile (Q1), median and 75th percentile (Q3) of the time to first failure or malfunction distribution and the 95% confidence interval (CI) will be presented together with the KM plots. A by-patient listing of the device failure and malfunction data will be displayed.

The IDDD longevity (time to failure in weeks) and time to first total malfunction, for all implanted IDDDs, will be plotted using the KM method. The 25th percentile (Q1), median and 75th percentile (Q3) of the time to failure or total malfunction distribution and the 95% CI will also be presented together with the KM plots.

The time (weeks) from IDDD implantation to failure/continuing function will be plotted as a horizontal bar chart for each device sorted by duration of each device implantation (time from the implantation of the corresponding IDDD), longest on top.

6.8.7 Concomitant Medications

For IA analysis, concomitant medications are defined as all medications taken on or after the date of the first IDDD implant surgery or first dose (whichever is earlier) and on or before the EOS visit (+30 days) or 2 weeks after the removal of the last IDDD if early termination (whichever is later).

Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary version March 2013 or higher. Concomitant medication use will be summarized by Therapeutic Class and Preferred Term by the Early IT group, the Delayed IT group and overall for the Safety Population. Any former HGT-HIT-094 Substudy patients will be included in the Early IT group.

6.8.8 Concomitant Therapies

For IA analysis, concomitant therapies are defined as all therapies taken on or after the date of the first IDDD implant surgery or first dose (whichever is earlier) and on or before the EOS visit (+30 days) or 2 weeks after the removal of the last IDDD if early termination (whichever is later).

Concomitant therapies will be coded using the MedDRA Version 16.1 or higher and summarized by Therapeutic Class and Preferred Term by the Early IT group, the Delayed IT group and overall for the Safety Population. Any former HGT-HIT-094 Substudy patients will be included in the Early IT group.

6.9 Analysis of Pharmacokinetic Data

All pharmacokinetic analyses will be performed using the PK Population. Serum and CSF concentrations as well as pharmacokinetic parameters will be listed and summarized by the Early IT group, the Delayed IT group and visit.

6.9.1 Concentration Data

Blood samples will be collected for determination of idursulfase serum concentration-time profiles and serum pharmacokinetic parameters after IT administration. Idursulfase concentrations in serum and CSF will be determined using a validated Enzyme-Linked Immunosorbent Assay (ELISA) method which was used for the previous idursulfase IT clinical studies. The idursulfase concentrations in serum will also be determined by a validated liquid chromatography tandem mass spectrometry (LC-MS/MS) method. The LC-MS/MS method is expected to be less susceptible to interference by anti-idursulfase antibodies.

6.9.2 Handling BLQ Values

The following procedures will be used for serum concentrations below the lower limit of quantification (LLOQ) (reported as BLQ):

- Serum samples that are BLQ are reported as zero on the data listings.
- Samples that are BLQ are treated as zero in the calculation of summary statistics (e.g., mean, SD, etc.) for the serum concentrations at individual time points.
- Mean concentrations are reported as zero if all values are BLQ, and no descriptive statistics are reported. If the calculated mean (± SD) concentration is less than the LLOQ, the value will be reported as calculated. The mean values derived using these conventions will be used to create the mean serum concentration versus time plots.
- For calculation of area under the serum concentration curve (AUC), BLQ values are set equal to zero in the dataset loaded into WinNonlin for pharmacokinetic analysis. WinNonlin uses the zero values that occur before the first time point with a concentration greater than LLOQ, but WinNonlin excludes the zero values from AUC calculation for all later time points.

6.9.3 Pharmacokinetic Parameters

The pharmacokinetic analysis will be conducted by the pharmacokinetics CRO for the Clinical Pharmacology and Pharmacokinetics Department of Shire Pharmaceuticals using Phoenix WinNonlin version 6.2 or higher (Pharsight Corporation, Mountain View, California, USA).

Pharmacokinetic parameters will be determined from serum concentration-time data using noncompartmental methods and all calculations will be based on actual sampling times. Serum concentration vs. time will be plotted for each patient. Mean serum concentration vs. time curves will also be presented by dose (5 mg, 7.5 mg, and 10 mg) and visit (Week 4, Week 48, and Week 100) with Week 4 and 48 assessments only made for subjects who received no IT treatment in HGT-HIT-094 and initiated IT treatment in this study.

The pharmacokinetic parameters will include, but not be limited to, the following:

- $AUC_{0-\infty}$ - Area under the curve extrapolated to infinity, calculated using the observed value of the last non-zero concentration
- AUC_{0-t} - Area under the curve from the time of dosing to the last measurable concentration
- C_{max} - Maximum concentration occurring at t_{max}
- t_{max} - Time of maximum observed concentration sampled during a dosing interval
- CL/F – Total body clearance for extravascular administration divided by the fraction of dose absorbed.
- V_z/F – Volume of distribution associated with the terminal slope following extravascular administration divided by the fraction of dose absorbed.
- λ_z – First order rate constant associated with the terminal (log-linear) portion of the curve
- t½ - Terminal half-life
- $MRT_{0-\infty}$ - Mean residence time

CL/F and Vz/F will also be normalized for body weight.

Summary statistics (number of observations, mean, standard deviation, coefficient of variation, median, maximum, minimum, and geometric mean) will be determined for all pharmacokinetic parameters and presented by bioanalytical method (ELISA, LC-MS/MS), dose and visit. Serum and CSF concentrations of idursulfase at each nominal sampling time will also be summarized by bioanalytical method, dose and visit using descriptive statistics.

Protocol Amendment 3 provides the investigator with the option of administering the IV Elaprase infusion and idursulfase-IT on the same day during the Extended Treatment Phase. If same-day dosing is elected, idursulfase-IT will be administered first and the Elaprase infusion will be administered second. Pharmacokinetic assessments are to occur at the first study visit at which same-day dosing is elected. At this visit, administration of the Elaprase infusion should start within approximately 90 minutes of completion of idursulfase-IT administration. Note that on the occasion of the first same-day administration, the duration of IV Elaprase infusion is to be standardized to 3 hours. The pharmacokinetic assessments for the same-day dosing option will be analyzed and summarized separately from the days with IT-only administration.

The association of the presence of anti-idursulfase antibodies and idursulfase concentration-time profiles and pharmacokinetic parameters will be evaluated, if applicable.

6.10 Drug-Drug and Drug-Disease Interactions

This study was not designed to evaluate drug-drug or drug-disease interactions and therefore the interactions will not be assessed in the study.

7. CHANGES IN THE CONDUCT OF THE STUDY OR PLANNED ANALYSES

7.1 Changes in the Conduct of the Study

Please see the amendment summaries in the SHP609-302 protocol Amendment 4.

7.2 Changes from the Analyses Planned in the Protocol

Not applicable

8. STATISTICAL/ANALYTIC ISSUES

8.1 Adjustment for Covariates

The rate of change (weighted) analysis will adjust for GCA classification factor (either ≤70 or >70) at the time of enrollment in SHP609-302 (i.e., Visit Month 13/Time of Enrollment in SHP609-302) in the weighted generalized linear model.

8.2 Handling of Dropouts or Missing Data

No data imputation will be performed.

8.3 Interim Analyses and Data Monitoring

One interim analysis is planned when all SHP609-302 subjects complete visit Month 25/Week 52 in SHP609-302 or discontinue.

Patient safety data in this study will be monitored by an independent DMC until the last patient completes his last scheduled study visit/assessment. The DMC will be an external group overseeing the safety of the study treatment, including both the investigational product and the IDDD, and will operate according to a charter determining the scope of its activities and frequency of meetings.

8.4 Multicenter Studies

This is a multicenter study. No subset analyses of center effect will be conducted. No adjustment due to multi-centers will be utilized in the analyses.

8.5 Multiple Comparisons/Multiplicity

No multiple comparison procedure or multiplicity adjustment will be performed for the secondary and exploratory efficacy endpoints.

8.6 Examination of Interactions

Interaction effects will not be examined for the outcome measurements due to the small sample size.

8.7 Sensitivity Analyses

For the exploratory ordered categorical outcome analysis sensitivity analysis will be conducted as described in Section 6.7.4.1.

For the 12-month rate of change (weighted) analysis, the following sensitivity analyses will be conducted (see Section 6.7.4.1 for further details):

- Subjects with only two available GCA scores, one at the beginning and one at the end of the 12-month period
- Two shrinkage weights in the form of the James-Stein estimator

8.8 Windowing Visits

Although there is a visit window of +/- 7 days around the expected visit date, nominal visits will be used for the per-visit analyses. The data obtained outside the scheduled visits for certain procedures could be assigned to nominal visits, for instance, CSF GAGs or Neurodevelopment assessments. For subjects who withdraw from the study prematurely, if the early termination visit falls into the window of a scheduled visit as defined in the protocol, the early termination visit is also summarized for that scheduled visit, unless the scheduled visit already took place.

8.9 Data Listings

All data will be presented as SAS datasets in CDISC format. Unless specifically stated in above sections, no other by-subject data listings will be provided.

9. REFERENCES

1. Meikle PJ, Hopwood JJ, Clague AE, Carey WF. Prevalence of lysosomal storage disorders. JAMA 1999;281:249-54.
2. Young ID, Harper PS. Incidence of Hunter's syndrome. Hum Genet 1982;60:391-2.
3. Singer, J. D., & Willett, J. B. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. Oxford: Oxford University Press.
4. Efron, B. & Morris, C.(1977). Stein's paradox in statistics. Scientific American 236 (5), 119-127.
5. Hamlett A, Ryan L, Serrano-Trespalacios P, Wolfinger R. Mixed models for assessing correlation in the presence of replication. Journal of the Air & Waste Management Association (1995) 2003;53:442-50.

10. APPENDICES

10.1 Appendix I – List of Statistical Outputs

There is a separate document called "Table Shells" including Tables, Figures and Listings.

10.2 Appendix II – Protocol Violations Definitions

A list of possible protocol violations is given below. A more comprehensive list will be constructed once the review of patient data has been completed, immediately prior to locking the database.

Violations

- Violation of admission (inclusion/exclusion) criteria (i.e., eligibility violation) for which no exemption was obtained
- Occurrence of a treatment dispensing error

10.3 Appendix III – Sample SAS Code

10.3.1 Sample SAS Code for Rate of Change (Weighted) Analysis

```
/*Linear regression fit for individual subjects*/;
ods output  ANOVA=ANOVA ParameterEstimates=Parms;
proc reg data = subject_data ;
    model GCA = AVISITN ;
run ;

data slope;  /*calculate slope and estimated change from baseline*/;
        set parms;
        if variable=' AVISITN';
        CFBL= estimate*52;  /* GCA score change from Visit Month 13/ Time of Enrollment in
        SHP609-302 to Visit Month 25/ Week 52 in SHP609-302*/;
        keep usubjid variable estimate CFBL;
run;

/*If subjects have 5 non-missing visits, then n=5, calculate the sum of square (x_i-x_bar) for
each subject, x_i is 0, 16, 28, 40 and 52. If n=4, x_i is 0, 52 and any two values from 16, 28, and
40 depends on the missing visit. If n=3, x_i is 0, 52 and any one value from 16, 28, and 40
depends on the missing visits.*/;
proc univariate data=subqs noprint;
  var avisitn;
  output out=parameters mean=mean;
run;
proc sql noprint;
  select mean
  into : avg_avisitn
  from parameters;
```

```
quit;

data subqs;
  set subqs;
  sqavt = (avisitn - &avg_avisitn)**2;
run;
proc univariate data=subqs noprint;
  var sqavt;
  output out=parameters sum=sum;
run;
proc sql noprint;
  select sum
  into :sum_sqavt
  from parameters;
quit;

data weight; /*calculate weight=inverse of variance of change over 12-month period*/;
  set anova;
  weight = 1/((52**2) * ms/sum_sqavt);  /*mse is unbiased estimator of sigma2*/;
run;

/*Analysis 1 - weighted generalized linear model to test treatment effect in "D" */;
proc glm data=data;
      class gcascatn_m13_bl; /* GCA classification factor at Visit Month 13/Baseline (either ≤70
    vs. >70)*/;
      model CFBL= gcascatn_m13_bl /solution clparm alpha = 0.1;
      weight weight;
      estimate 'Pooled Mean' intercept 1;
run;

/*Analysis 2 - weighted generalized linear model to test treatment effect in "B"∪"D"  */;
proc glm data=data;
      class observation_group gcascatn_m13_bl;
      model CFBL= observation_group gcascatn_m13_bl /solution clparm alpha = 0.1 ;
      weight weight;
      estimate 'Pooled Mean' intercept 1;
      estimate 'Diff between observation group' observation_group 1 -1;
      lsmeans observation_group / stderr pdiff CL alpha = 0.1;
 run;
```

10.3.2 Sample SAS Code for Correlation Analysis Using Repeated Measures

```
ods output vcorr=vcorr;
proc mixed;
    class usubjid vtype replicate;
    model response = vtype / ddfm=kr;
    random vtype / type=un subject=usubjid v vcorr cl;
    repeated vtype / type=un subject=replicate(usubjid) r rcorr;
    run;

/*usubjid corresponds to subject number; vtype refers to the two variables assessed, which are
coded as 1 and 2; Response corresponds to the values of the two variables; and Replicate
corresponds to the number of repeated measurements for each subject, which need not be the
same.*/;
```

10.3.3 Sample SAS Code for James-Stein/Empirical Bayes Estimator

```
proc univariate data=slope_data noprint;
  var rms;  /* rms = sqrt(ms) positive square root of mse*/;
  output out=parameters mean=mean;
run;

proc sql noprint;
  select mean
  into : avg_rms
  from parameters;
quit;

data weight_js;
  set slope_data;
  c=0.8;    /* using a fixed shrinkage factor c=0.8*/
  z = &avg_rms + (rms - &avg_rms) * c;   /* JS estimator of square root of mse */;
  weight_js = 1/((z**2) / sum_sqavt * (52**2));  /* JS estimator of weight */;
run;
```


10.4 Appendix IV – James-Stein/Empirical Bayes Estimator of Weight

The James-Stein form estimator is used to construct the shrinkage estimators of variance of GCA score change over a 12-month period, and then the shrinkage estimators of weights used in the weighted generalized linear model. The James-Stein estimator (a.k.a. empirical Bayes estimator) used information from all linear regressions to provide potentially better estimates of variance of GCA score change over a 12-month period for each linear regression.

There are three steps to construct James-Stein estimator of weight.

1. Generate James-Stein estimator of square root of mean squared error (MSE), i.e., RMSE
2. Plug the James-Stein estimator of RMSE into the variance of GCA score change over a 12-month period, which is a function of RMSE

3. Plug the James-Stein estimator of the variance of GCA score change over a 12-month period into the weight, which is the inverse of the variance estimator

Here are the details of each step.

1. Generate James-Stein estimator of square root of MSE (RMSE)

Let's first look at the relationship between MSE (or RMSE) and the variances of GCA score change over a 12-month from a simple linear regression.
For the simple linear regression model: $. _i = b_0 + b_1 x_i + e_i$, where $e_i \sim iid\ Normal\ (0, {\sigma_e}^2), i = 1, \ldots, n$. For the linear regression of each subject, x_i is the nominal visit week and y_i is the GCA score at each visit week.

Estimator of slope in the linear regression is

$$\widehat{b_1} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2$$

Variance of slope estimator in the linear regression is

$$\mathrm{VAR}\{\widehat{b_1}\} = {\sigma_e}^2 \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2$$

The GCA score change over a 12-month period is estimated by the slope estimates multiplied by 52 weeks directly. Therefore, estimated GCA score change over a 12-month period is

$$\mathrm{est\_CFBL} = 52 \times \widehat{b_1}.$$

Variances of GCA score change over a 12-month period is

$$\mathrm{VAR}\{\mathrm{est\_CFBL}\} = \mathrm{VAR}\{52 \times \widehat{b_1}\} = 52^2 \times \mathrm{VAR}\{\widehat{b_1}\}$$

Means Squared Error (MSE) from the linear regression is

$$\mathrm{MSE} = \frac{SSE}{n-2} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \Big/ (n-2) = \sum_{i=1}^{n} {e_i}^2 \Big/ (n-2)$$

MSE is an unbiased estimator for ${\sigma_e}^2$, i.e., E{MSE}= ${\sigma_e}^2$. An estimator of the standard deviation σ_e is simply positive square root of MSE, i.e., RMSE $= \sqrt{MSE}$.

We can estimate the variance of slope estimator $\mathrm{VAR}\{\widehat{b_1}\}$ by replacing the parameter ${\sigma_e}^2$ with MSE:

$$\widehat{\mathrm{VAR}}\{\widehat{b_1}\} = MSE \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2 = RMSE^2 \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2$$

Then the variance of GCA score change over a 12-month period can be estimated as

$$\widehat{\mathrm{VAR}}\{\mathrm{est\_CFBL}\} = 52^2 \times \widehat{\mathrm{VAR}}\{\widehat{b_1}\} = 52^2 \times RMSE^2 \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2$$

The variance of GCA score change over a 12-month period is a function of RMSE.
For regression model, $SSE/\sigma_e{}^2$ is distributed as χ^2 with n-2 degrees of freedom, $\chi^2(n-2)$.

We have $\frac{\widehat{\mathrm{VAR}}\{\widehat{b_1}\}}{\mathrm{VAR}\{\widehat{b_1}\}} = \frac{\frac{MSE}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\frac{\sigma_e{}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}} = \frac{MSE}{\sigma_e{}^2} = \frac{\frac{SSE}{n-2}}{\sigma_e{}^2} = \frac{SSE}{\sigma_e{}^2(n-2)} \sim \frac{\chi^2(n-2)}{n-2}$.

Therefore $MSE \sim \frac{\sigma_e{}^2}{n-2}\chi^2(n-2)$, and RMSE is approximately normally distributed.

2. Plug the James-Stein estimator of RMSE into the variance of GCA score change over a 12-month period, which is a function of RMSE

Based on step 1, in order to construct the James-Stein estimator of variance of GCA score change over a 12-month period, $\widehat{\mathrm{VAR}}_{JS}(\mathrm{est\_CFBL})$. We first need to construct the James-Stein estimator of RMSE, $RMSE_{JS}$.

The general formula of the James-Stein estimator is $z = \bar{y} + c(y - \bar{y})$, where $c = 1 - \frac{(k-3)\sigma^2}{\sum(y-\bar{y})^2}$ is the shrinking factor, y is the average of a single set of data, $\bar{y}$ is the grand average of averages, k is the number of averages (number of y's), and σ^2 is the square of the standard deviation and $\sum(y-\bar{y})^2$ is the sum of the squared deviation of the individual average y from the grand average $\bar{y}$.

We consider 10% and 20% shrinkage are sufficient for the weights to check that the estimated treatment effects from weighted generalized linear model will not be dominated by a few values with large weights. The following parameters will be plugged into general formula of James-Stein estimator to construct James-Stein estimator of RMSE, $RMSE_{JS}$:

- y is the observed individual RMSE from linear regressions,
- $\bar{y}$ is the grand average of observed RMSEs from linear regressions (i.e., observed mean of all RMSEs),
- c is the shrinking factor, c=0.8 (20% shrinkage) or 0.9 (10% shrinkage)

3. Plug the James-Stein estimator of the variance of GCA score change over a 12-month period into the weight, which is the inverse of the variance

Once we get the James-Stein estimator of RMSE, we can plug $RMSE_{JS}$ into the formula of estimating the variance of slope estimator $\mathrm{VAR}\{\widehat{b_1}\}$

$$\widehat{\mathrm{VAR}}_{JS}\{\widehat{b_1}\} = RMSE_{JS}{}^2 / \sum_{i=1}^{n}(x_i - \bar{x})^2$$

When subjects have 5 non-missing visits, n=5 and

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=(0-27.2)^2+(16-27.2)^2+(28-27.2)^2+(40-27.2)^2+(52-27.2)^2$$
$$=1644.8$$

We can also calculate the $\sum_{i=1}^{n}(x_i-\bar{x})^2$ for subjects with 3 or 4 non-missing visits.

Then the James-Stein estimator of the variance of GCA score change over a 12-month period is

$$\widehat{\mathrm{VAR}}_{JS}\{\mathrm{est\_CFBL}\}=52^2\times\widehat{\mathrm{VAR}}_{JS}\{\hat{b}_1\}=52^2\times {RMSE_{JS}}^2/1644.8$$

The weight is the inverse of estimated variances of GCA score change over a 12-month period, estimated by the inverse of the variance of the slope from each linear egression divided by squared 52 weeks.

$$\mathrm{Weight}=\frac{1}{\widehat{\mathrm{VAR}}\{\mathrm{est\_CFBL}\}}=\frac{1}{52^2}\times\frac{1}{\widehat{\mathrm{VAR}}\{\hat{b}_1\}}$$

The James-Stein estimator of weights is the inverse of James-Stein estimator of the variance of GCA score change over a 12-month period

$$\mathrm{Weight}_{JS}=\frac{1}{\widehat{\mathrm{VAR}}_{JS}(\mathrm{est\_CFBL})}=\frac{1}{52^2}\times\frac{1}{\widehat{\mathrm{VAR}}_{JS}\{\hat{b}_1\}}$$

ADDENDUM
STATISTICAL ANALYSIS PLAN
FOR INTERIM ANALYSIS 2
(24-Month Data for Study SHP609-302)

Investigational Product and Device:	An Open Label Extension of Study HGT-HIT-094 Evaluating Long Term Safety and Clinical Outcomes of Intrathecal Idursulfase Administered in Conjunction with Elaprase® in Patients with Hunter Syndrome and Cognitive Impairment
Protocol Number:	SHP609-302 (Amendment 4)
Protocol Date:	09 October 2018
SAP Addendum Author:	[redacted]
Release Date:	26 March 2020
Shire Medical Monitor:	[redacted], DO [redacted] 300 Shire Way Lexington, MA 02421, USA Telephone: [redacted] Fax: [redacted] (USA)

Table of Contents

1. ABBREVIATIONS

ERT	enzyme replacement therapy
IA2	Interim analysis 2
IV	Intravenous
SAP	statistical analysis plan

2. PURPOSE OF THE STATISTICAL ANALYSIS PLAN ADDENDUM

The interim analysis 2 (IA2) of study data will be conducted when all SHP609-302 subjects complete 24 months of idursulfase-IT (SHP609) treatment from the time of enrollment in SHP609-302 (at visit Month 37/Month 25) or discontinue in study SHP609-302.

The purpose of this statistical analysis plan (SAP) addendum for IA2 is to document the technical and detailed specifications for the analysis on the 24-month data of study SHP609-302 with cut-off date of 30th July 2019 in addition to the SAP for interim analysis v1.1 dated on 14 March 2019 (referred to as SHP609-302 IA SAP v1.1 hereafter) which provided a technical and detailed elaboration of the planned interim analysis based on the 12-month data with cut-off date of 6th September 2018 for protocol SHP609-302 Amendment 4.0.

Unless otherwise specified, the analyses in this SAP addendum will apply the same methodologies as specified in the SHP609-302 IA SAP v1.1 for 12-month data. Detailed methods for additional analyses, if not included in the SHP609-302 IA SAP v1.1, will be specified in this SAP addendum.

3. OBJECTIVE AND ENDPOINTS

3.1 Objective

To evaluate long-term safety in patients with Hunter syndrome and cognitive impairment who are receiving intrathecal idursulfase-IT (SHP609) and intravenous (IV) Elaprase® enzyme replacement therapy (ERT) as specified in SHP609-302 IA SAP v1.1 Section 4.1.

3.2 Efficacy Endpoints

Primary, secondary and exploratory efficacy endpoints are defined similarly to those in Sections 4.1 of the SHP609-302 IA SAP v1.1, where the duration will be extended by an additional 12 months, to a total of at least 24 months for each subject during study SHP609-302, with the exception that “Change over a 12-Month Period Endpoints”, “Ordered Categorical Outcomes (3 Categories)”, “Binary Unreversed Floor Effect Outcomes” (SHP609-302 IA SAP v1.1 Section 5.4.3, 5.4.4 and 5.4.5, respectively) will not be included in this SAP addendum because those analyses only pertain to 12-month data during study SHP609-302.

4. STATISTICAL ANALYSIS

4.1 Analysis Populations

Safety Population and Pharmacokinetic Population will be used in SHP609-302 IA2. The definition of the populations will be the same as in SHP609-302 IA SAP v1.1.

Definition of analysis populations in SHP609-302 IA2 SAP Addendum

Analysis Population	Definition	Intended Usage
Safety Population	All subjects in study SHP609-302 who underwent IDDD implantation or received at least 1 dose of study drug (full or partial)	Primary safety population
Pharmacokinetic Population	All subjects in SHP609-302 who received study drug and participated in the scheduled pharmacokinetic studies, and for whom at least 1 post-dose PK blood sample was collected	PK population

4.2 Primary Efficacy Analysis

Not applicable.

4.3 Secondary Efficacy Analysis

The same methodology as in SHP609-302 IA SAP v1.1 Section 6.7.2 will be used with additional data in Extended Treatment Phase included.

4.4 Subgroup Analysis

The same methodology as in SHP609-302 IA SAP v1.1 Section 6.7.3 will be used with additional data in Extended Treatment Phase included.

4.5 Exploratory Efficacy Analyses

The same methodology as in SHP609-302 IA SAP v1.1 Section 6.7.4 will be used with additional data in Extended Treatment Phase included.

The following analyses will not be included in this SAP addendum:

- Rate of Change (Weighted) Analysis
- Exploratory Ordered Categorical Outcome Analysis
- Exploratory Analysis to Assess Floor Effect

4.6 Health Economics and Outcomes Research Endpoint Analyses

The same methodology as in SHP609-302 IA SAP v1.1 Section 6.7.5 will be used with additional data in Extended Treatment Phase included.

4.7 Pharmacodynamic Outcome

The same methodology as in SHP609-302 IA SAP v1.1 Section 6.7.6 will be used with additional data in Extended Treatment Phase included.

4.8 Analysis of Safety

The same methodology as in SHP609-302 IA SAP Section 6.8 will be used with additional data in Extended Treatment Phase included.

4.9 Analysis of Pharmacokinetic Data

The same methodology as in SHP609-302 IA SAP Section 6.9 will be used with additional data in Extended Treatment Phase included.