STATISTICAL ANALYSIS PLAN

NCT Number: NCT02386839

Study Title: Long-term Outcome of Children Enrolled in Study ROPP-2008-01 Previously Treated with rhIGF-1/rhIGFBP-3 for the Prevention of Retinopathy of Prematurity (ROP) or Who Received Standard Neonatal Care

Study Number: SHP607-201

SAP Version and Date:

Version 1.0: 01 September 2020

STATISTICAL ANALYSIS PLAN

Protocol No.:	SHP-607-201
Protocol Title:	Long-term Outcome of Children Enrolled in Study ROPP-2008-01 Previously Treated with rhIGF-1/rhIGFBP-3 for the Prevention of Retinopathy of Prematurity (ROP) or Who Received Standard Neonatal Care
Drug:	SHP607, Mecasermin rinfabate (rhIGF-1/rhIGFBP-3)
Sponsor:	Premacure AB, A Member of the Shire Group of Companies 300 Shire Way Lexington, MA 02421 USA [redacted]
Version No. and Date	Version 1.0, 1 September 2020

Version No:	Document History Description of Update	Author(s)	Effective Date
1.0	Final	[redacted]	1 September 2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS

ADHD	attention-deficit hyperactivity disorder
ADHD-RS	Attention-Deficit/Hyperactivity Disorder Rating Scale
AE	adverse event
AESI	adverse event of special interest
ASD	Autism Spectrum Disorder
ATC	Anatomic Therapeutic Chemical
BPD	bronchopulmonary dysplasia
BSID-III	Bayley Scales of Infant and Toddler Development, Third Edition
CA	corrected age
CBCL	Child Behavior Checklist
CI	confidence interval
CRF	case report form
eCRF	electronic case report form
EOS	end of study
ER	emergency room
[redacted]	[redacted]
GA	gestational age
HCRU	health care resource use
HRQoL	health-related quality of life
HSCS-PS	Health Status Classification System-Preschool
HUI	Health Utilities Index
HUI2/3	Health Utilities Index Mark 2 and Mark 3
IGF	insulin-like growth factor
IGFBP-3	insulin-like growth factor binding protein-3
IP	investigational product
IVH	intraventricular hemorrhage
KM	Kaplan-Meier
LV	left ventricle
MedDRA	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
OD	right eye
OS	left eye
OU	both eyes
PCS	potentially clinically significant
PedsQL	Pediatric Quality of Life Inventory
PMA	post menstrual age
PT	preferred term
PVL	periventricular leukomalacia

ROP	retinopathy of prematurity
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Statistical Analysis System
SCQ	Social Communication Questionnaire
SD	standard deviation
SOC	system organ class
VABS-II	Vineland Adaptive Behavior Scales, Second Edition
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary
WPPSI	Wechsler Preschool and Primary Scale of Intelligence

1. INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed elaboration of the statistical analyses of efficacy and safety data as described in the study protocol amendment 4, dated 09 April 2018. Specifications for tables, figures, and listings are contained in a separate document.

2. OBJECTIVES AND ENDPOINTS

2.1 Objectives

2.1.1 Primary Objectives

The primary objectives of this study are:

- To evaluate the long-term efficacy outcomes following short-term exposure to rhIGF-1/rhIGFBP-3 versus standard neonatal care in Study ROPP-2008-01 (Section D) as assessed by ROP-associated visual outcomes
- To evaluate the long-term safety outcomes following short-term exposure to rhIGF-1/rhIGFBP-3 versus standard neonatal care in Study ROPP-2008-01 (Section D)

2.1.2 Secondary Objectives

The secondary objectives of this study are to evaluate the effect following short-term exposure to rhIGF-1/rhIGFBP-3 versus standard neonatal care in Study ROPP-2008-01 (Section D) on:

- Growth parameters
- Cognitive development
- Physical development
- Child behavior
- Pulmonary morbidity
- Survival
- Health-related quality of life (HRQoL)
- Health utility
- Health care resource use (HCRU)

2.1.3 Exploratory Objective

[redacted]

2.2 Endpoints

2.2.1 Primary Efficacy Endpoints

The primary efficacy endpoints of this study are:

- Visual acuity as assessed by an age-appropriate method

- Ocular alignment and ocular motor examination in primary gaze and in as many of 9 positions of gaze as possible as assessed by corneal light reflex and by the cover test
- Assessment of nystagmus by observation
- Refraction as assessed by retinoscopy with cycloplegia
- Stereoacuity as assessed with the Lang Stereotest

2.2.2 Secondary Efficacy Endpoints

The secondary efficacy endpoints of this study are:

- Growth parameters including body weight, body length (or height), and head circumference
- Cognitive development as assessed by the following standardized, age-appropriate tools:
  - Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III)
  - Wechsler Preschool and Primary Scale of Intelligence (WPPSI)
- Physical development as assessed by the following standardized, age-appropriate tools:
  - Neurological examination (as part of physical examination) for assessment of cerebral palsy
  - Hearing assessment
- Child behavior as assessed by the following:
  - Vineland Adaptive Behavior Scales, Second Edition (VABS-II)
  - Child Behavior Checklist (CBCL; 1 ½ to 5)
  - Attention-Deficit/Hyperactivity Disorder Rating Scale (ADHD-RS) for the assessment of symptoms of attention-deficit/hyperactivity disorder (ADHD)
  - Social Communication Questionnaire (SCQ) for screening of Autism Spectrum Disorder (ASD)
- Pulmonary morbidity data (eg, hospitalizations, emergency room [ER] visits, pulmonary medications)
- Survival as assessed by death during the study due to any cause

2.2.3 Exploratory [redacted] Endpoint

[redacted]

2.2.4 Health Economic Outcome Research Endpoints

The health economic outcome research endpoints of this study are:

- Health Related Quality of Life (HRQoL) as assessed by the Pediatric Quality of Life Inventory (PedsQL™) Scales appropriate for the child's age of development with the

Total Scale Score and 5 domains within Physical Health (Physical Functioning and Physical Symptoms) and Psychosocial Health Scores (Emotional, Social, and Cognitive Functioning, respectively)

- Health status (eg, health utility) as measured by the Health Status Classification System-Preschool (HSCS-PS) and the Health Utilities Index Mark 2 (HUI2) and Mark 3 (HUI3)
- Resource use associated with inpatient visits, outpatient visits, medical utilization, and pharmacy utilization as measured by Health Care Resource Use (HCRU)

2.2.5 Safety Endpoints

The safety endpoints of this study are:

- Adverse events (AEs), as follows:
    - those considered related to rhIGF-1/rhIGFBP-3 (as administered in Study ROPP-2008-01, Section D)
    - those considered related to procedures performed in this study (Study SHP-607-201)
    - specified targeted medical events regardless of causality
    - fatal serious adverse events (SAEs) regardless of causality
- Cardiac size as assessed by echocardiogram
- Kidney and spleen size and any other gross abnormalities as assessed by abdominal ultrasound
- Physical examination including tonsil examination
- Blood pressure, heart rate, and respiratory rate

3. STUDY DESIGN

3.1 General Description

This is a Phase II, multicenter, long-term developmental outcome study enrolling subjects who were randomized in Section D of Study ROPP-2008-01 to receive either rhIGF-1/rhIGFBP-3 (treated) or standard neonatal care (control). Enrolled subjects in this study will be followed through age 5 years CA. This study will enroll both subjects who were in the treated (received rhIGF-1/rhIGFBP-3) and control (received standard neonatal care) groups of Study ROPP-2008-01 (Section D) to enable assessment of rhIGF-1/rhIGFBP-3 long-term efficacy and safety outcomes versus standard neonatal care. No investigational product will be administered in this study. In the descriptions of summary statistics, treatment groups refer to the 2 treatment assignments in the antecedent study, ROPP-2008-01, Section D (rhIGF-1/rhIGFBP-3 or standard neonatal care).

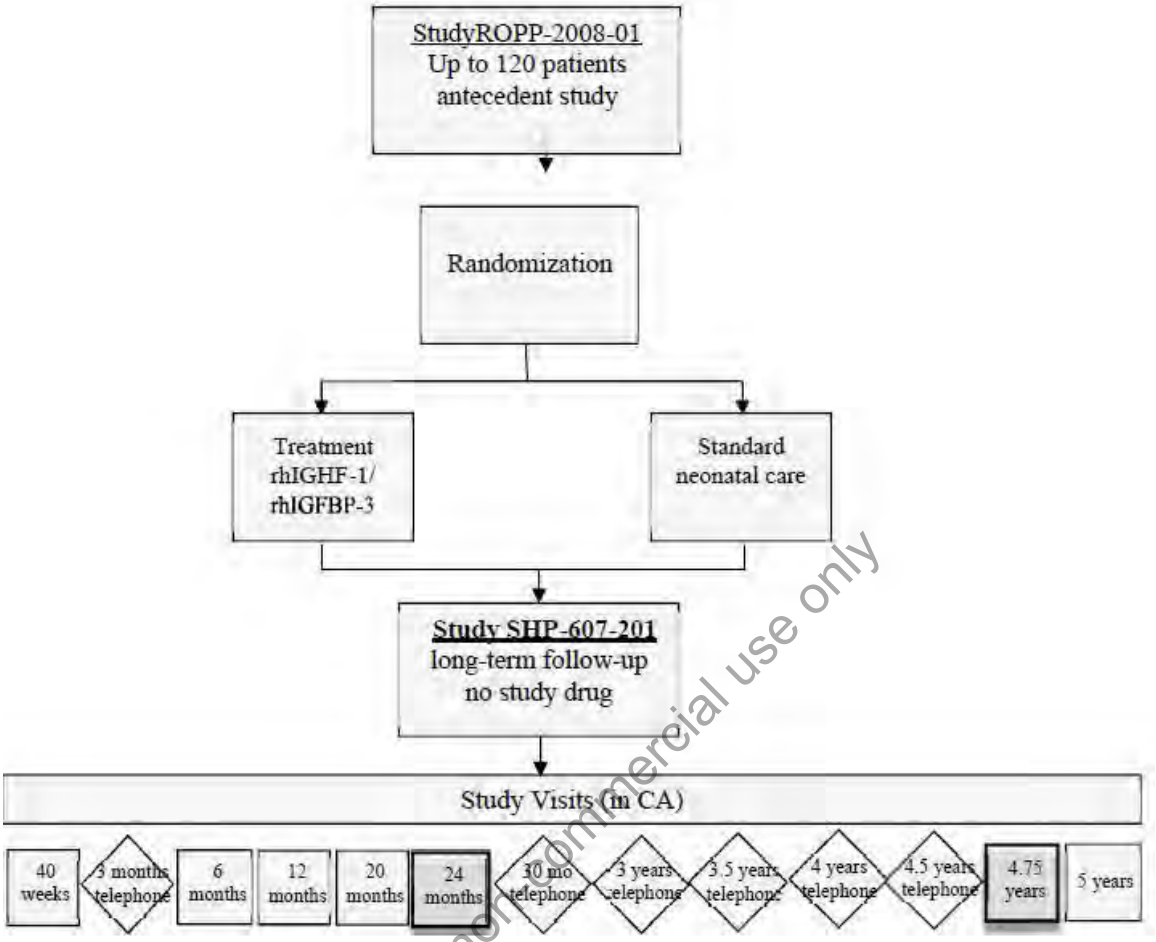
In this study, the Initial Visit will occur at 40 weeks corrected age (CA; ie, term equivalent) or upon discontinuation from Study ROPP-2008-01, or may occur any time until 2 years CA +3 months. Subjects are no longer eligible to participate in this study after they turn 2 years CA +3 months. Subjects in Study ROPP-2008-01 are premature infants enrolling at gestational age (GA) of 23 weeks+0 days to 27 weeks +6 days.

Time points for assessments have been chosen based on standard premature infant follow-up periods and represent important developmental ages for premature infant follow-up. Both telephone and clinical site visits are included to help maintain contact with subjects throughout the 5-year duration of the study.

Subjects will be evaluated at appropriate follow-up site locations with expertise in the assessment of the developmental outcomes of premature infants. Pediatric ophthalmology expertise will also be required.

The overall study design is outlined in Figure 3-1.

Figure 3-1 Overview of Study Design



CA=corrected age
Note: Visits conducted by telephone are indicated with a diamond shape. Visits conducted at the study site are indicated with rectangles. Visit windows are provided in Appendix 1.

3.2 Randomization

Not applicable.

3.3 Masking

Not applicable.

3.4 Sample Size and Power Considerations

No formal sample size calculation was performed for this study because this is a follow-up study to Section D of Study ROPP-2008-01. Any subjects enrolled in Study ROPP-2008-01 are eligible to enroll in this study. There are up to 120 subjects who will be eligible to enroll in this long-term developmental outcome study.

4. STATISTICAL ANALYSIS SETS

4.1 Enrolled Set

The Enrolled Set will consist of all subjects for whom written informed consent has been provided for this long-term outcome study.

4.2 Safety Set

The Safety Set will consist of all subjects in the Enrolled Set who have safety follow-up data in this long-term outcome study.

5. STUDY SUBJECTS

5.1 Disposition of Subjects

This section describes subject disposition for both the analysis sets and the study status.

The number of subjects included in each defined analysis set will be summarized by treatment group and overall. Percentages will be provided using the Enrolled Set as the denominator.

The number and percentage of subjects who completed through the study visits (3-month, 6-month, 12-month, 24-month and 5-year CA visits), who completed the study, and who prematurely discontinued from the study will be presented for each treatment group and overall for the Enrolled Set. Primary reasons for premature discontinuation as recorded on the Study Completion page of the eCRF will be summarized (number and percentage) by treatment group and overall. All subjects who prematurely discontinued will be listed for the Enrolled Set.

A by-subject listing of study completion information, including the reason for premature study discontinuation, where applicable, will be presented.

5.2 Demographic and Other Baseline Characteristics

Descriptive summaries of demographic and baseline characteristics will be presented by treatment group and overall for the Enrolled Set.

Continuous variables will be summarized using descriptive statistics, including number of observations, mean, standard deviation, median, minimum and maximum values. Categorical variables will be summarized using number of subjects and percentages.

The following demographic and baseline characteristics will be summarized: chronological age (weeks) at enrollment, CA at enrollment (weeks), GA at birth (weeks), GA group (<26 weeks, ≥26 weeks), sex, ethnicity, race, weight at birth (kg), body length at birth (cm), head circumference at birth (cm), way of delivery (Vaginal, Caesarean Section), bronchopulmonary dysplasia (BPD) status (None/Mild, Moderate/Severe), intraventricular hemorrhage (IVH) status (Grade 0-II, Grade III-IV), and periventricular leukomalacia (PVL, severe or cystic). The history of BPD, IVH, and PVL status will be obtained from the antecedent study (ROPP-2008-01). CA (weeks) is calculated as [chronological age (weeks) – 40 + GA (weeks)].

A listing will be presented to show all the demographic and baseline characteristics for each subject in the Enrolled Set.

5.3 Medical History

Medical history, including maternal and perinatal medical history, will be obtained from the antecedent study (ROPP-2008-01) and summarized by medication for each treatment group and overall for the Enrolled Set. In addition, medical history will be provided in data listings for the Enrolled Set.

5.4 Concomitant Medication

Concomitant medications will be coded using WHO-Drug Dictionary (WHO-DD) dated June 2013. All medications administered to the subjects collected from the time of informed consent through the 5-year CA visit (or until the subject withdraws or is discontinued) are considered as concomitant medications. The therapeutic class is coded using the highest anatomic therapeutic chemical (ATC) level that is available (eg, ATC level 3 will be used if ATC level 4 is not applicable).

Concomitant medications will be summarized by the number and proportion of subjects in each treatment group and overall by therapeutic class and preferred term for the Enrolled Set. Multiple medication usage by a subject in the same category will be counted only once.

All concomitant medications will be listed for the Enrolled Set.

5.5 Exposure to Investigational Product

Exposure to investigational product is not applicable as no investigational product is administered in this study.

5.6 Protocol Violations and Deviations

Protocol violations and deviations will be recorded by the site separately from the clinical database. All protocol violations and deviations will be summarized by category and site for each treatment group and overall, for the Enrolled Set. A by-subject listing will be provided for protocol violations and deviations for the Enrolled Set.

6. EFFICACY ANALYSES

All efficacy analyses will be based on the Enrolled Set and summarized by treatment group unless stated otherwise. Baseline is defined as the first assessment in this study (SHP-607-201) unless stated otherwise.

All efficacy analyses will be conducted according to the antecedent treatment assigned in Study ROPP-2008-01, Section D (rhIGF-1/rhIGFBP-3 or standard neonatal care).

Continuous variables will be summarized using descriptive statistics, including number of subjects, mean, standard deviation, median, minimum, and maximum values. Categorical variables will be summarized using number of subjects and percentages.

All statistical tests will be 2-sided hypothesis tests performed at the 5% level of significance for main effects, unless stated otherwise. All confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise.

6.1 Analyses of Primary Efficacy Endpoints

6.1.1 Visual Acuity

Visual acuity will be measured at the 6-month, 12-month, 20-month, 24-month, 4.75-year and 5-year CA visits. It will be assessed by an age-appropriate method summarized in Table A1 and categorized as the following:

- normal (measurable acuity ≥20/40 or ≥15 cycles/degree)
- below normal (20/200 ≤ measurable acuity <20/40 or 3 cycles/degree ≤measurable acuity <15 cycles/degree)
- poor (measurable acuity ≤ 20/200 or ≤3 cycles/degree)
- blind/low vision (only the ability to detect the 2.2 cm wide stripes on the low-vision Teller acuity card and at any location in the visual field).

Acuity results will be further classified as favorable outcomes (normal, below normal) and unfavorable outcomes (poor, blind/low vision). The number and proportion of subjects within each category listed above will be summarized for the left eye (OS), right eye (OD), and both eyes (OU) by treatment group and visit. In addition, a shift table of favorable outcomes from baseline (first assessment during this study) to each of the subsequent assessments, including the last assessment, will be presented by treatment group for OS, OD and OU.

The measurement results will be presented in a by-subject listing.

6.1.2 Ocular Alignment and Oculomotor Exam (Motility)

Ocular alignment and motility will be assessed at the 12-month, 24-month and 5-year CA visits. Findings from the ocular motility assessment will be either presence or absence of strabismus (esotropia, exotropia, hypertropia or hypotropia). Tabular summaries by treatment group and

visit will include the frequency and the percentage in each category. In addition, a shift table from baseline (first assessment during this study) to each of the subsequent assessments, including the last assessment will be provided by treatment group for each category.

All the assessment results will be presented in a by-subject listing.

6.1.3 Nystagmus

The assessment of nystagmus will be recorded at the 12-month, 24-month and 5-year CA visits during the ocular alignment assessments. The presence or absence of nystagmus will be summarized by treatment group and visit and presented in a by-subject listing.

6.1.4 Refraction with Cycloplegia

Refraction with cycloplegia will be performed at the 6-month, 12-month, 20-month and 4.75-year CA visits as part of the corrective lens determination procedure. Descriptive statistics will be presented for each measurement (sphere, cylinder, axis, prism) and each eye individually (OS and OD) by treatment group and visit.

All measurement results will be presented in a by-subject listing.

6.1.5 Stereoacuity

Stereoacuity will be performed at the 5-year CA visit. Presence or absence of stereopsis will be summarized by treatment group and presented in a by-subject listing.

6.1.6 Sensitivity Analyses of Primary Efficacy Endpoints

Not applicable.

6.2 Analyses of Secondary Efficacy Endpoints

6.2.1 Growth Parameters

For growth parameters, baseline is defined as the first assessment in the antecedent study (ROPP-2008-01). Body weight will be collected at the 6-month, 12-month, 24-month and 5-year CA visits. Body length will be collected at the 6-month, 12-month and 24-month CA visits, and height will be collected at the 5-year CA visit. Head circumference will be collected at the 6-month, 12-month and 24-month CA visits. For height and length, 2 measures should be taken and recorded. The average height and length will be used in the analysis.

A standard Z-score for each growth parameter (body weight, body length/height, and head circumference) will be calculated at each time point by adjusting age- and sex- matched means and standard deviations (norm) utilizing WHO child growth standards. Refer to Appendix 2 for reference files and the detailed calculation on the Z-score.

Descriptive statistics of the observed value and Z-score for each growth parameter and the corresponding change from baseline at each post-baseline visit and the last visit will be summarized by treatment group. A 95% confidence interval (CI) for the corresponding mean

change within each treatment group, the difference between two treatment groups and the corresponding 95% CI will also be presented.

The Shapiro-Wilk test will be used to test the normality of the change from baseline for each of the above parameters. If the normality assumption can be satisfied, the mean change will be utilized to measure the treatment difference at each post-baseline visit, and the associated 95% CI will be estimated from the two-sample t-test. Otherwise, the median difference will be utilized, and the associated 95% CI will be estimated from the Hodges-Lehmann estimator.

All the growth parameters with Z-scores will be provided in a data listing. The growth parameters and the corresponding Z-scores against scheduled visits will be plotted by treatment group.

6.2.2 Cognitive Development

Cognitive development will be assessed by the BSID-III and WPPSI. All the questionnaire assessments and results will be provided in the data listings.

6.2.2.1 Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III)

The BSID-III will be performed at the 12-month and 24-month CA visits. There are 3 domains of development (cognitive, language and motor) and 5 subtests. The cognitive subtest stands alone while the 2 language subtests (expressive and receptive communication) combine to make a total language score and the 2 motor subtests (fine and gross motor) form a combined motor scale. The cognitive and language composite scores <85 and the motor composite scores <70 are classified as moderate to severe developmental delay.

The raw score and scale score for each subtest will be summarized by treatment group using descriptive statistics of the observed value at each visit and change from the 12-month CA visit. The composite score for each domain will be summarized in the same manner. In addition, the number and percentage of subjects with the developmental status (no/mild delay, moderate/severe delay) classified by the composite score will be summarized by treatment group and visit.

6.2.2.2 Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV)

The WPPSI-IV will be performed at the 5-year CA visit. It is composed of 15 subtests: Block Design, Information, Matrix Reasoning, Bug Search, Picture Memory, Similarities, Picture Concepts, Cancellation, Zoo Locations, Object Assembly, Vocabulary, Animal Coding, Comprehension, Receptive Vocabulary, and Picture Naming.

For children aged 4 years through 7 years 7 months (4:0–7:7), the test framework of the WPPSI-IV (Table 6-1) is organized into five Primary Index scales, with each scale composed of two core subtests: Verbal Comprehension (Information and Similarities), Visual Spatial (Block Design and Object Assembly), Fluid Reasoning (Matrix Reasoning and Picture Concepts), Working Memory (Picture Memory and Zoo Locations), and Processing Speed (Bug Search and

Cancellation). The Full Scale includes all of the subtests in each scale at the Primary Index scale level, as well as any additional supplemental subtests that may be used to derive the Full Scale IQ. At the Full Scale level, there are six core subtests (Information, Similarities, Block Design, Matrix Reasoning, Picture Memory, and Bug Search) and seven supplemental subtests (Vocabulary, Comprehension, Object Assembly, Picture Concepts, Zoo Locations, Cancellation, and Animal Coding). The sum of scaled scores will be used to derive the composite score for each Primary Index scale and the Full Scale.

Table 6-1 Test Framework for Ages 4:0–7:7

	Verbal Comprehension	***Visual Spatial***	***Fluid Reasoning***	***Working Memory***	***Processing Speed***
Full Scale	Information Similarities Vocabulary* Comprehension*	Block Design Object Assembly*	Matrix Reasoning Picture Concepts*	Picture Memory Zoo Locations*	Bug Search Cancellation* Animal Coding*
Primary Index Scale	***Verbal Comprehension***	***Visual Spatial***	***Fluid Reasoning***	***Working Memory***	***Processing Speed***
	Information Similarities	Block Design Object Assembly	Matrix Reasoning Picture Concepts	Picture Memory Zoo Locations	Bug Search Cancellation

*Supplemental subtest

The total raw score and scaled score for each subtest will be summarized by treatment group using descriptive statistics. The sum of scaled scores and the composite score for each Primary Index scale and the Full Scale will also be summarized by treatment group using descriptive statistics. The mean of the scaled score for each subtest will be plotted by treatment group.

6.2.3 Physical Development

The cerebral palsy will be assessed at the 24-month CA visit. Descriptive statistics will be summarized for cerebral palsy assessment details as collected in the Cerebral Palsy CRF page by treatment group.

The hearing assessments will be recorded at the 6-month (historical hearing test data may be recorded at any time prior to the 6-month visit) and 5-year CA visits. The number and proportion of subjects with hearing screening status (normal, abnormal) will be summarized by treatment group.

All the assessment results will be presented in the data listings.

6.2.4 Child Behavior

Child behavior will be assessed by the ADHD-RS, SCQ, VABS-II, and CBCL. All the questionnaire assessments and results will be provided in the data listings.

6.2.4.1 Attention Deficit/Hyperactivity Disorder Rating Scale, Version IV (ADHD-RS-IV)

The ADHD-RS-IV will be performed at the 5-year CA visit. It consists of 18 items designed to reflect current symptomatology of ADHD based on DSM-IV criteria. Each item is scored on a 4-point scale ranging from 0 (reflecting no symptoms) to 3 (reflecting severe symptoms) with total scores ranging from 0 to 54. The 18 items are grouped into 2 subscales: hyperactivity/impulsivity (even numbered items, ranging from 0 to 27) and inattentiveness (odd numbered items, ranging from 0 to 27). The total score and subscale scores (hyperactivity/impulsivity and inattentiveness) will be summarized by treatment group using descriptive statistics.

6.2.4.2 Social Communication Questionnaire (SCQ)

The SCQ (Lifetime version) will be performed at the 5-year CA visit. It has 40 dichotomous (yes, no) items, and each scored item receives a value of 1 point for "abnormal behavior" and 0 point for "absence of abnormal behavior/normal behavior." It yields a total score ranging from 0 to 39, with higher scores representing more social communication impairment. The first item, "Is she/he now able to talk using short phrases or sentences?", is not scored, but rather determines whether six items relating to abnormal language are assigned. Only "verbal" children (ie, children with a "yes" response to the first question) are assigned the six items relating to abnormal language and can score a total of 0 to 39 points; "non-verbal" children (ie, children with a "no" response to the first question) are not assigned the six items in relation to abnormal language and can score a total of 0 to 33 points. These items focus on three areas of functioning: Reciprocal Social Interaction; Communication; and Restricted, Repetitive, and Stereotyped Patterns of Behavior. Refer to Appendix 3 for SCQ items listed under each domain.

The total score and the domain (Reciprocal Social Interaction, Communication, and Repetition/Stereotyped Behavior) scores will be summarized by treatment group using descriptive statistics.

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

The VABS-II (Sparrow et al., 2005) will be performed at the 6-month, 12-month, 24-month, and 5-year CA visits. These scales are organized into three domains and three subdomains within each domain: Communication (Receptive, Expressive, and Written), Daily Living Skills (Personal, Domestic, and Community), and Socialization (Interpersonal Relationships, Play and Leisure Time, and Coping Skills). In addition, the VABS-II offers a Motor Skills Domain and an optional Maladaptive Behavior Index. The Motor Skills domain with Gross and Fine subdomains will be assessed, if applicable. The four domains combine to form the Adaptive Behavior Composite. Standard scores and adaptive levels will be used to interpret domains and Adaptive Behavior Composite, while raw scores, v-scaled scores, adaptive levels, and age equivalents will be used to interpret subdomains.

The following descriptive summaries will be provided by treatment group and visit for each domain and/or subdomain, if applicable:

- Descriptive statistics of raw scores, v-scaled scores, and age equivalents for subdomains
- Descriptive statistics of standard scores for domains and Adaptive Behavior Composite
- The number and percentage of subjects by percentile-based adaptive skill level (low, moderately low, adequate, moderately high, and high) for domains, subdomains and Adaptive Behavior Composite

6.2.4.4 Child Behavior Checklist (CBCL) 1½-5

The CBCL (Achenbach et al., 2000) is a parent-reported outcome measure used to assess behavioral, emotional, and social functioning of toddlers and preschool children aged 18-60 months. The CBCL will be performed at the 24-month and 5-year CA visits. It is composed of 7 syndromes: Emotionally Reactive, Anxious/Depressed, Somatic Complaints,

Withdrawn, Sleep Problems, Attention Problems, and Aggressive Behavior. These preschool syndromes are submitted to a second-order factor analysis, two broad-band problems, labeled "Internalizing" and "Externalizing". The former includes the syndromes of Emotionally Reactive, Anxious/Depressed, Somatic Complaints, and Withdrawn, while the latter includes syndromes of Aggressive Behavior and Attention Problems. The syndrome of Sleep Problems stands alone.

The CBCL raw scores are converted to the standardized T-scores, which have a normative group mean of 50 and standard deviation of 10. Descriptive statistics of the observed value at each visit and change from the 24-month CA visit will be summarized for the seven syndrome scales, Internalizing, Externalizing, and Total Problems by treatment group using the raw score and normalized T score, respectively.

For each syndrome, Internalizing, Externalizing, and Total Problems, scores can be interpreted as falling in the non-clinical, borderline, or clinical range. The non-clinical T-score range for all syndromes is 65 or less, the borderline clinical range is from 66 to 70, while the clinical range is 71 or greater; For Internalizing, Externalizing, and Total Problems, the cut-off point for the non-clinical range is a T score less than 60, borderline is from 60 to 63, and the clinical range is 64 or greater.

The number and proportion of subjects within each category listed above will be summarized for each syndrome, Internalizing, Externalizing, and Total Problems by treatment group and visit.

6.2.5 Pulmonary Morbidity Assessment

Pulmonary morbidity is assessed with questions related to family history and smoking status as well as diagnosis of select pulmonary symptoms, conditions and related hospitalizations. Pulmonary morbidity assessment will be performed as outlined in the Study Schedule of Events (see Appendix 1) with both clinical site visits and phone interviews. Different questionnaires will be provided for different type of visits. Items below will be summarized by treatment group and visit using descriptive statistics:

- Had asthma, wheezing, BPD exacerbation, or flare-up (yes/no)
- Had bronchiolitis, bronchitis, or pneumonia diagnosed (yes/no)
- Had to use oxygen at home (yes/no)
- Had to visit ER or urgent care (yes/no)
- Had to stay in a hospital overnight (yes/no)

Questionnaires of pulmonary morbidity assessment for clinical visits and phone interviews will be presented in the data listings separately.

6.2.6 Survival

Time to all-cause mortality in days will be calculated as (date of death – date of birth +1). Subjects who have completed or who have terminated the study early without experiencing the event will be censored at the date of study completion or early termination.

The cumulative probability of all-cause mortality will be summarized and plotted over time by treatment group using the Kaplan-Meier (KM) method. The mortality status for all subjects will be provided in a data listing.

6.3 Analyses of Exploratory [redacted] Endpoint

[redacted]

6.4 Subgroup Analyses

The baseline data that are considered clinically relevant will be evaluated by subgroup analyses. Analysis of below efficacy outcomes will be repeated:

- BSID-III composite scores by excluding site [redacted] from the overall population, the baseline IVH/PVL status, and the baseline BPD status, separately
- [redacted]

7. SAFETY ANALYSIS

The safety analysis will be performed using the Safety Set. Safety data include AEs, abdominal ultrasound, echocardiogram, physical examination including tonsil examination, blood pressure, heart rate, and respiratory rate. Safety summaries will be based on all assessments post-baseline.

All safety analyses will be conducted according to the treatment the subject actually received in Study ROPP-2008-01 (Section D).

7.1 Adverse Events

Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA) Version 16.0. AEs and SAEs will be collected and recorded throughout the study, from the time of informed consent is signed through the last visit. Ongoing targeted medical events regardless of causality, and ongoing study drug-related AEs (including SAEs) collected as part of Study ROPP-2008-01 (Section D) will be recorded as AEs in this study.

For the purposes of this study only the following adverse events will be collected:

- Those considered related to investigational product (rhIGF-1/rhIGFBP-3 as administered in Study ROPP -2008-01, Section D)
- Those considered related to procedures performed in this study (SHP-607-201)
- AEs related to ROP
- AEs related to congenital malformations not identified at birth which may impact neurocognitive development
- Additional illnesses present at the time when informed consent is given are to be regarded as AEs
- Specified targeted medical events (protocol Section 7.9.1.4) regardless of causality

If it is determined that any of the specified targeted medical events have been experienced by a subject, they will be recorded as AEs or SAEs, as appropriate, regardless of relationship to investigational product (rhIGF-1/rhIGFBP-3 as administered in Study ROPP-2008-01, Section D). The targeted medical events are identified based on a list of pre-specified preferred terms and are further classified into the following Adverse Events of Special Interest (AESI) class:

- Cardiac size increase (including cardiomegaly, cardiac hypertrophy, cardiac septal hypertrophy, ventricular hypertrophy, atrial hypertrophy, left ventricular hypertrophy, right ventricular hypertrophy, left atrial hypertrophy, right atrial hypertrophy, cardiomyopathy, cardiomyopathy neonatal, hypertrophic cardiomyopathy)
- Tonsillar hypertrophy (including tonsillar hypertrophy, adenoidal hypertrophy, thymus enlargement, hypoacusis, snoring, ear tube insertion, tonsillectomy, adenoidectomy, sleep apnoea syndrome)

- Intracranial hypertension (including intracranial pressure increased, benign intracranial hypertension, fontanelle bulging, delayed fontanelle closure)
- Hypoglycemia (including hypoglycaemia, hypoglycaemia neonatal, blood glucose decreased)

An overall summary of the number of subjects as well as the number of events in each treatment group and overall will be presented, including the number and percentage of subjects with:

- Any AE
- Any SAE
- Any severe AE
- Any AE related to IP as administered in Study ROPP-2008-01 (Section D)
- Any SAE related to IP as administered in Study ROPP-2008-01 (Section D)
- Any severe AE related to IP as administered in Study ROPP-2008-01 (Section D)
- Any AE related to procedures performed in this study (SHP-607-201)
- Any SAE related to procedures performed in this study (SHP-607-201)
- Any severe AE related to procedures performed in this study (SHP-607-201)
- Any AE leading to death
- Any AESI

The number and percentage of subjects reporting AEs, as well as the number of events (except for tables by the maximum severity), in each treatment group and overall will be tabulated, respectively, in following ways:

- By system organ class (SOC) and preferred term (PT) for all AEs
- By the maximum severity, SOC and PT for all AEs
- By SOC and PT for all AEs considered related to IP as administered in Study ROPP-2008-01 (Section D)
- By SOC and PT for all AEs considered related to study procedures
- By SOC and PT for all AEs leading to death
- By SOC and PT for all SAEs
- By AESI class and PT for AESI
- By AESI class and PT for serious AESI

In tables by severity or relationship, missing severity or missing relationship imputation will be applied, and handling of those missing cases are detailed in Section 10.4.1 and Section 10.4.2, respectively. If a subject had multiple occurrences of the same AEs, the subject will be counted only once per SOC and once per PT at the maximum severity.

All AEs will be presented in a data listing. Separate data listings will be presented for AEs related to the antecedent IP, AEs related to study procedures, AEs leading to death, SAEs and AESI.

7.2 Cardiac Size

Echocardiographic examination will be performed for the evaluation of cardiac size at the 6-month CA visit. The following measurements will be summarized by treatment group using descriptive statistics.

- Interventricular septal thickness
- Posterior wall thickness
- Intracavity volume (both in end diastole and end systole)

The measurement results will be summarized by treatment group using descriptive statistics. By-subject listings with the echocardiogram measurement results will be presented.

7.3 Kidney and Spleen Size

An abdominal ultrasound will be performed to assess the size of the kidneys and spleen at the 6-month CA visit.

The measurements of the length of each kidney (left, right) and spleen will be summarized by treatment group using descriptive statistics. The measurement results will be categorized into a binary response (normal, abnormal) and summarized within each category by the number and proportion of subjects in each treatment group. A 2-sided 95% CI will be estimated for subjects with a normal status using the Wald test.

A by-subject listing with the measurement results by abdominal ultrasound and status of normality will be presented.

7.4 Physical Examination

Any abnormal findings from physical examination (including tonsil examination) will be summarized by treatment group and visit.

All the physical exam results will be listed by subject in a data listing.

7.5 Blood Pressure, Heart Rate, and Respiratory Rate

Blood pressure, heart rate and respiratory rate will be summarized by treatment group and visit using descriptive statistics and listed by subject in a data listing.

8. OTHER ANALYSES

The Enrolled Set will be used for health economics and outcomes research analyses. All the analyses will be conducted and summarized by the treatment assigned in Study ROPP-2008-01 (Section D). All the questionnaire assessments and results will be presented in the data listings.

8.1 Pediatric Quality of Life Inventory (PedsQL)

The PedsQL will be performed at the 3-month, 6-month, 12-month, 24-month, 3-year, 4-year and 5-year CA visits. The infant scales that apply to ages less than 24 months include Physical Functioning, Physical Symptoms, Emotional Functioning, Social Functioning, and Cognitive Functioning. The generic core scales that apply to toddlers (ages 24 months to 4 years) and young children (ages 5 to 7 years) include Physical Functioning, Emotional Functioning, Social Functioning, and School Functioning. The scale scores, the summary scores (Physical Health Summary Score and Psychosocial Health Summary Score), and the total score will be summarized by treatment group using descriptive statistics of the observed value at each visit and change from the first assessment for both infants (ages less than 24 months) and toddlers/young children (ages 24 months to 7 years), separately.

8.2 Health Utility

8.2.1 Health Status Classification System-Preschool (HSCS-PS)

The HSCS-PS will be performed at the 24-month, 3-year, 4-year and 5-year CA visits. The instrument is composed of 12 domains (Vision, Hearing, Speech, Mobility, Dexterity, Self-care, Emotion, Learn/Remember, Think/Problem Solve, Pain, General Health, and Behavior). Each domain has up to 6 levels (0-5) to indicate the health status (with higher numbers indicating better status). Descriptive statistics will be presented for each domain by treatment group and visit.

These domains will be grouped into four categories: Neurosensory (Vision and Hearing), Motor (Mobility, Dexterity, and Self-care), Learning/Remembering (Speech, Learn/Remember, and Think/Problem Solve) and Quality of Life (Emotion, Pain, General Health, and Behavior). For each category, the data will be recorded into the following levels: no problem (scoring 0 on any attribute); a mild problem (scoring 1 on a scale of 0 to 3, or 1 to 2 on a scale of 0 to 5 for any attribute); or a moderate or severe problem (scoring >1 on a scale of 0 to 3, or >2 on a scale of 0 to 5 for any attribute). The number and percentage of subjects in each level for each category will be summarized by treatment group and visit.

8.2.2 Health Utilities Index Mark 2 and Mark 3 (HUI2/3)

The HUI2/3 will be performed at the 5-year CA visit. The HUI2 classification system scores 6 attributes: Sensation, Mobility, Cognition, Self-Care, Emotion, and Pain. The HUI3 classification system comprises 8 attributes: Vision, Hearing, Speech, Cognition, Ambulation, Dexterity, Emotion, and Pain. Each attribute has up to 6 levels to indicate the severity (with higher numbers indicating worse levels). Single-attribute scores of morbidity are defined on a scale such the worst level has a score of 0.00 and the best level has a score of 1.00.

The number and percentage of subjects in each level for each attribute will be summarized by treatment group. The utility score for each attribute and the overall utility score will be summarized by treatment group using descriptive statistics.

8.3 Health Care Resource Use (HCRU)

The HCRU will be performed at the 3-month, 6-month, 12-month, 24-month, 3-year, 4-year and 5-year CA visits. The following items will be summarized descriptively by treatment group and visit:

- The number of visits to health care providers (including Pediatrician, General Practitioner or Family Doctor, Pediatric Ophthalmologist or Ophthalmologist, Pediatric Pulmonologist or Pulmonologist, Physical Therapist, Occupational Therapist, Developmental Therapist, Speech Therapist, and Other) in private office/clinic and hospital outpatient
- The number of visits to the Emergency Department / Urgent Center / Casualty / Accident or Emergency
- The duration of hospitalization (days)
- Use of prescription eyeglasses (yes/no)
- Receive educational support (yes/no)

The duration/number of visits will be calculated as the sum of all durations/number of visits at each visit window. Subjects who were not hospitalized/made no visits at the corresponding visit window will not be counted.

9. INTERIM ANALYSIS

No formal interim analysis is planned. A snapshot of the data was reviewed after all enrolled subjects in this study completed 2-year follow-up (24-month visit) assessments or prematurely withdrew from the study before completing 2 years of follow up, and after data were entered into the database, queried, and discrepancies resolved.

A selective descriptive analysis of the data with focus on the 24-month CA visit for this interim review was performed and used to understand the progress of the study, completion of the data, safety monitoring, regulatory communication, and general planning purposes.

10. DATA HANDLING CONVENTIONS

10.1 General Data Reporting Conventions

Continuous variables will be summarized using the following descriptive statistics: n, mean, median, standard deviation, minimum and maximum. Categorical and count variables will be summarized by the number of subjects (n) and the percent of subjects in each category.

Unless specified otherwise, median and mean will use 1 decimal place beyond the precision of data used for the measurement; standard deviation and standard error of the mean will use 2 decimal places beyond the precision; minimum and maximum values will use the same number of decimal places as the precision. P-values will be rounded to 3 decimal places and presented as <0.001 for values less than 0.001.

10.2 Definition of Visit Windows

Assessments will be assigned to visits based on the completed CRF page.

10.3 Repeated or Unscheduled Assessments

If End-of-Study (EOS) assessments are repeated or unscheduled, the last post-baseline assessment will be used as the end of study assessment for generating descriptive statistics if needed. However, all assessments will be presented in the data listings.

10.4 Handling of Missing, Unused, and Spurious Data

10.4.1 Missing Severity Assessment for Adverse Events

If severity is missing for an AE, a severity of "Severe" will be assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be used in data listings.

10.4.2 Missing Relationship for Adverse Events

If the relationship to investigational product (as administered in Study ROPP-2008-01, Section D) is missing for an AE, a causality of "Related" will be assigned. If the relationship to study procedure performed in this study (SHP-607-201) is missing for an AE starting on or after the date of informed consent, a causality of "Related" will be assigned. The imputed values for relationship will be used for incidence summaries, while the actual values will be presented in data listings.

11. ANALYSIS SOFTWARE

Statistical analyses will be performed using Version 9.4 (or newer) of SAS® on a suitably qualified environment.

12. CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL

The following changes were incorporated in Protocol Amendment 4, dated 09 April 2018:

- Changes in physical examination (including tonsil examination) from the secondary efficacy endpoint to the safety endpoint
- Changes in blood pressure, heart rate, and respiratory rate from the secondary efficacy endpoint to the safety endpoint
- [redacted]
- [redacted]

13. REFERENCES

Achenbach T, Rescorla L. Manual for the ASEBA preschool forms and profiles. Burlington, VT: University of Vermont; 2000.Johnson S, Moore T, Marlow N. Using the Bayley-III to assess neurodevelopmental delay: which cut-off should be used? Pediatr Res. 2014; 75(5): 670-674.

Marvin AR, Marvin DJ, Lipkin PH, Law JK. Analysis of Social Communication Questionnaire (SCQ) Screening for Children Less Than Age 4. Curr Dev Disord Rep. 2017; 4(4): 137-144.

Rutter M, Bailey A, Lord C. Manual for the Social Communication Questionnaire. 2003.

Schiariti V, Klassen AF, Houbé JS, Synnes A, Lisonkova S, Lee SK. Perinatal characteristics and parents' perspective of health status of NICU graduates born at term. J Perinatol. 2008; 28(5): 368-376.

Sparrow SS, Cicchetti, DV, Balla, DA. The Vineland Adaptive Behavior Scales, 2nd edition (VABS-II). NCS Pearson Inc. 2005.

Wechsler D. Wechsler preschool and primary scale of intelligence 4th ed. Bloomington, MN: Pearson; 2012.

World Health Organization. WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development. Geneva: World Health Organization; 2006: 301-304.

Appendix 1 Study Schedule of Events

Procedures	Initial Study Visit[e]	Months (CA)[k]						Years (CA)[k]					
	40 weeks (CA)/term equivalent	3[f] ± 2 wks	6 ± 1 mth	12 ± 3 mths	20 -1 mth[g]	24 ± 3 mth	30[f] ± 3 mth	3[f] ± 3 mths	3.5[f] ± 3 mth	4[f] ± 3 mths	4.5[f] ± 3 mth	4.75 -1 mth[g]	5 + 6 mths
Informed Consent	•												
Eligibility Criteria	•												
Demographics	•												
Visual acuity[a]			•	•	•	•						•	•
Corrective lens determination[h]				•	•[g]							•[g]	
Ocular alignment and motility				•		•							•
Refraction with cycloplegia[h]			•										
Stereoacuity													•
Length			•	•		•							
Height													•
Weight			•	•		•							•
Head Circumference			•	•		•							
BSID-III				•		•							
WPPSI													•
CBCL						•							•
VABS-II			•	•		•							•
ADHD-RS													•
SCQ													•
Physical Examination including tonsil examination			•	•		•							•
Blood Pressure, Heart Rate, and Respiratory Rate													•
Cerebral Palsy Assessment						•							
Hearing Assessment History[b]			•										•
Pulmonary Morbidity Assessment 1			•	•									
Pulmonary Morbidity Assessment 2 (clinical site visit)						•							•
Pulmonary Morbidity Assessment (phone interview)							•	•	•	•	•		

Procedures	Initial Study Visit[e]	Months (CA)[k]						Years (CA)[k]					
	40 weeks (CA)/term equivalent	3[f] ± 2 wks	6 ± 1 mth	12 ± 3 mths	20 -1 mth[g]	24 ± 3 mth	30[f] ± 3 mth	3[f] ± 3 mths	3.5[f] ± 3 mth	4[f] ± 3 mths	4.5[f] ± 3 mth	4.75 -1 mth[g]	5 ± 6 mths
Survival assessment		•	•	•		•	•	•	•	•	•		•
Cerebral MRI													[redacted]
[redacted]													•
HRQoL[c]		•	•[i]	•[i]		•[i]		•		•			•[i]
HCRU		•	•[i]	•[i]		•[i]		•		•			•[i]
HSCS-PS						•[i]		•		•			•[i]
HUI2/3													•
Abdominal Ultrasound			•										
Echocardiogram			•										
Assessment of Participation in Other Clinical Studies		•	•	•		•	•	•	•	•	•		•
Medications		•	•	•		•	•	•	•	•	•		•
Adverse events[d]	•[j]	•	•	•		•	•	•	•	•	•		•

ADHD-RS=Attention-Deficit/Hyperactivity Disorder Rating Scale; BSID-III=Bayley Scales of Infant and Toddler Development-Third Edition, CBCL=Child Behavior Checklist; CA=corrected age; HCRU=health care resource use; HRQoL=health-related quality of life; HSCS-PS=Health Status Classification System; HUI2/HUI3=Health Utilities Index Mark 2 and Mark 3; mth(s)=months; PedsQL=Pediatric Quality of Life Inventory; SCQ=Social Communication Questionnaire; VABS-II=Vineland Adaptive Behavior Scales, Second Edition; wks=weeks; WPPSI=Wechsler Preschool and Primary Scale of Intelligence

[a] The tools used to assess visual acuity will change as the subject ages during their participation in the study. The tools that will be used in this study and are summarized by applicable study visit in Table A1.

[b] Historical hearing test data may be recorded at any time during the study prior to the 6-month visit.

[c] HRQoL will be assessed via the validated PedsQL™ scales appropriate for the child's age of development as specified in the Study Operations Manual.

[d] Adverse event collection will include an assessment of the specified targeted medical events.

[e] The Initial Visit may be performed prior to 40 weeks CA for any subject who discontinued from Study ROPP-2008-01 and, for all subjects, any time after 40 weeks CA. If the Initial Visit does not occur at or before 40 weeks CA, the subject may still be enrolled until 2 years CA +3 months. Subjects are no longer eligible to participate in this study after they turn 2 years CA +3 months.

[f] Visits at 3 months, 30 months, 3 years, 3.5 years, 4 years, and 4.5 years CA will be conducted by telephone.

Procedures	Initial Study Visit[e]	Months (CA)[k]						Years (CA)[k]					
	40 weeks (CA)/term equivalent	3[f] ± 2 wks	6 ± 1 mth	12 ± 3 mths	20 -1 mth[g]	24 ± 3 mth	30[f] ± 3 mth	3[f] ± 3 mths	3.5[f] ± 3 mth	4[f] ± 3 mths	4.5[f] ± 3 mth	4.75 -1 mth[g]	5 + 6 mths

[g] The 20-month visit and the 4.75-year (4 years, 9 months) visit must occur at least 1 month prior to the 24-month and 5-year visits, respectively. Any prescribed corrective lenses must be worn for at least 1 month prior to the 24-month and 5-year assessments.

[h] Refraction with cycloplegia will be performed as part of the corrective lens determination procedure.

[i] The HRQoL, HCRU, and HSCS-PS assessments for the 6-month, 12-month, 24-month and 5-year visits may be performed through clinical site staff if there are time constraints during the on-site visit. At the 3-month, 3-year, and 4-year visits, these assessments will be performed through clinical site staff and may be performed at any time within the visit window.

[j] The following, collected as part of the ROPP-2008-01 study, will be used as part of this study (SHP-607-201): any ongoing targeted medical events regardless of causality and any ongoing study drug-related AEs, including SAEs.

[k] Multiple visits and/or phone contacts are allowed to complete all assessments, if needed.

Table A1 Summary of Visual Acuity Assessments

Visual Acuity Assessment Tool	Description	Unit of Measure	Applicable Age/Study Visit (CA)
Teller acuity cards	Resolution acuity – discrimination of a grating optotype from a background with the same mean luminance	cycles per degree	6 months 12 months
LEA Symbols Test	Recognition acuity - tests recognition of 4 optotypes	Snellen fraction (eg. 20/20)	24 months[a]
LEA Symbols (ETDRS-style) chart	Distance visual acuity test	Snellen fraction (eg. 20/20)	5 years[a]

CA=corrected age; ETDRS=Early Treatment of Diabetic Retinopathy Study

[a] At ages 24 months and 5 years CA an attempt should be made to assess visual acuity by the LEA Symbols Test and LEA Symbols Chart, respectively. If during this attempt, the examiner determines that it will not be possible to perform an accurate assessment with the relevant LEA tool, visual acuity should be assessed with the Teller acuity cards.

Appendix 2 WHO Child Growth Standards

For z-score derivation on growth parameters, WHO Child growth standards will be utilized as reference files and links below will be used as source data.

		Source link
Weight (Birth to 5 years)	Girls	https://www.who.int/childgrowth/standards/wfa_girls_0_5_zscores.txt
	Boys	https://www.who.int/childgrowth/standards/wfa_boys_0_5_zscores.txt
Length (Birth to 2 years)	Girls	https://www.who.int/childgrowth/standards/lhfa_girls_0_2_zscores.txt
	Boys	https://www.who.int/childgrowth/standards/lhfa_boys_0_2_zscores.txt
Height (2 to 5 years)	Girls	https://www.who.int/childgrowth/standards/lhfa_girls_2_5_zscores.txt
	Boys	https://www.who.int/childgrowth/standards/lhfa_boys_2_5_zscores.txt
Head Circumference (Birth to 5 years)	Girls	https://www.who.int/childgrowth/standards/second_set/tab_hcfa_girls_z_0_5.txt
	Boys	https://www.who.int/childgrowth/standards/second_set/tab_hcfa_boys_z_0_5.txt

The following procedure is recommended to calculate a z-score for an individual child with measurement y at age t:

1. Calculate

$$\ldots = \frac{[{}^{y}/_{M(t)}]^{L(t)} - 1}{S(t)L(t)}$$

where the tabulated fitted values of Box-Cox power, median and coefficient of variation corresponding to age t are denoted by L(t), M(t) and S(t), respectively.

2. Compute the final z-score (Z^*_{ind}) if the child for that indicator as:

$$Z^*_{ind} \begin{cases} Z_{ind} & if\ |Z_{ind}| \leq 3 \\ 3 + \left(\frac{y - SD3pos}{SD23pos}\right) & if\ \ Z_{ind} > 3 \\ -3 + \left(\frac{y - SD3neg}{SD23neg}\right) & if\ \ Z_{ind} < -3 \end{cases}$$

where SD3*pos* is the cut-off 3 SD calculated at t by the LMS method:

$$.. \quad 3... \quad = M(t)[1 + L(t) * S(t) * (3)]^{1/L(t)}$$

SD3*neg* is the cut-off -3 SD calculated at t by the LMS method:

$$SD3neg = M(t)[1 + L(t) * S(t) * (-3)]^{1/L(t)}$$

SD23*pos* is the difference between the cut-offs 3 SD and 2 SD calculated at t by the LMS method:

$$SD23pos = M(t)[1 + L(t) * S(t) * (3)]^{1/L(t)} - M(t)[1 + L(t) * S(t) * (2)]^{1/L(t)}$$

and SD23*neg* is the difference between the cut-offs -2 SD and -3 SD calculated at t by the LMS method:

$$SD23neg = M(t)[1 + L(t) * S(t) * (-2)]^{1/L(t)} - M(t)[1 + L(t) * S(t) * (-3)]^{1/L(t)}$$

Example 1: 44-month-old boy with weight=23.5. L = -0.0993, M = 15.6828, S = 0.12531.

$$Z_{ind} = \frac{[23.5/15.6828]^{-0.0993} - 1}{0.12531\ x\ (-0.0993)} = 3.16 \quad (Z_{ind} > 3)$$

SD3*pos* = 23.0, SD2*pos* = 20.2, SD23*pos* = 23.0 – 20.2 = 2.8.

$$Z^*_{ind} = 3 + \left(\frac{23.5 - 23.0}{2.8}\right) = 3.18$$

Example 2: 24-month-old boy with weight=8.2. L = -0.0137, M = 12.1515, S = 0.11426.

$$\ldots = \frac{[{}^{8.2}/_{12.1515}]^{-0.0137} - 1}{0.11426\ x\ (-0.0137)} = -3.45 \quad (Z_{ind} < -3)$$

SD3*neg* = 8.6, SD2*neg* = 9.7, SD23*neg* = 9.7 – 8.6 = 1.1.

$$Z^{*}_{ind} = -3 + \left(\frac{8.2 - 8.6}{1.1}\right) = -3.36$$

Example 3: 56-month-old boy with weight=19.8. L = -0.1387, M = 17.6768, S = 0.13261.

$$Z_{ind} = \frac{[{}^{19.8}/_{17.6768}]^{-0.1387} - 1}{0.13261\ x\ (-0.1387)} = 0.85 \quad (|Z_{ind}| \leq 3)$$

Detailed derivation formulas can be found in *WHO Child Growth Standards, Section 7*:
https://www.who.int/childgrowth/standards/Technical_report.pdf

Appendix 3 SCQ Items Listed Under the Domains

Reciprocal Social Interaction domain
9. Inappropriate facial expression
10. Use of other's body to communicate
19. Friends
26. Eye gaze
27. Social smiling
28. Showing and directing attention
29. Offering to share
30. Seeking to share enjoyment
31. Offering comfort
32. Quality of social overtures
33. Range of facial expression
36. Interest in children
37. Response to other children's approaches
39. Imaginative play with peers
40. Group play

Communication domain
2. Conversation
3. Stereotyped utterances
4. Inappropriate questions
5. Pronoun reversal
6. Neologisms
20. Social chat
21. Imitation
22. Pointing to express interest
23. Gestures
24. Nodding to mean yes
25. Head shaking to mean no
34. Imitative social play
35. Imaginative play

Restricted, Repetitive, and Stereotyped Patterns of Behavior domain
7. Verbal rituals
8. Compulsions and rituals
11. Unusual preoccupations
12. Repetitive use of objects
13. Circumscribed interests
14. Unusual sensory interests
15. Hand and finger mannerisms
16. Complex body mannerisms