santhera	**Statistical Analysis Plan**

PHASE II OPEN-LABEL, MULTIPLE DOSE STUDY TO ASSESS THE SAFETY, TOLERABILITY, PHARMACOKINETICS, PHARMACODYNAMICS, AND EXPLORATORY EFFICACY OF VAMOROLONE IN BOYS AGES 2 TO <4 YEARS AND 7 TO <18 YEARS WITH DUCHENNE MUSCULAR DYSTROPHY (DMD)

Study Code: VBP15-006

Study Phase: II

VERSION: 2.0

2024-06-25

STATISTICAL ANALYSIS PLAN

1 Abbreviations

AE	Adverse Event
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the concentration-time Curve
BMI	Body Mass Index
BLQ	Below the Limit of Quantification
BP	Blood Pressure
CDC	Centers for Disease Control and Prevention
CS	Clinically Significant
CSR	Clinical Study Report
CRF	Case Report Form
DMD	Duchenne Muscular Dystrophy
DNA	Deoxyribonucleic acid
EAP	Expanded Access Protocol
ECG	Electrocardiogram
GMSS	Gross Motor Scaled Score
IOP	Intraocular pressure (IOP)
LLOQ	Lower Limit of Quantification
MedDRA	Medical Dictionary for Regulatory Activities
PARS	Personal Adjustment and Role Skills
PD	Pharmacodynamics
PK	Pharmacokinetics
PODCI	Pediatric Outcomes Data Collection Instrument
POSNA	Pediatric Orthopaedic Society of North America
PT	Preferred Term
PUL	Performance of Upper Limb
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SEM	Standard Error of the Mean
SI	International System of Units
SOC	System Organ Class
SoC	Standard of Care
TEAE	Treatment-Emergent Adverse Event
WHO-DD	World Health Organization Drug Dictionary

2 Introduction

This first version of the Statistical Analysis Plan (SAP) for study VBP15-006 is based on Amendment 3 of the Clinical Study Protocol dated 11 April 2023.

The SAP describes the statistical approach to reach the objectives of the study.

3 Study objectives

Primary objective:

- To evaluate the safety and tolerability of vamorolone administered orally at daily doses of 2.0 mg/kg and 6.0 mg/kg over a 3-month treatment period in boys ages 2 to <4 and 7 to <18 years with Duchenne Muscular Dystrophy (DMD).

Secondary objectives:

- To evaluate the pharmacokinetics (PK) of vamorolone administered orally in boys ages 2 to <4 and 7 to <18 years with DMD.
- To confirm the vamorolone exposure in boys ages 2 to < 4 and 7 to <18 years with DMD at 2.0 and 6.0 mg/kg and to adjust the doses if appropriate to achieve similar vamorolone areas under the concentration-time curve (AUCs) across the entire pediatric age range.

Exploratory objectives:

- To compare the efficacy, as measured by the effect on muscle function, of vamorolone administered orally at daily doses of 2.0 mg/kg versus 6.0 mg/kg over a 3-month treatment period in boys ages 2 to <4 and 7 to <18 years with DMD;
- To evaluate the effect of vamorolone administered orally at daily doses of 2.0 mg/kg and 6.0 mg/kg over a 3-month treatment period on behavior and neuropsychology in boys ages 2 to <4 and 7 to <18 years with DMD;
- To evaluate the effect of vamorolone administered orally at daily doses of 2.0 mg/kg and 6.0 mg/kg over a 3-month treatment period on physical functioning in boys ages 2 to <4 and 7 to <18 years with DMD;
- To evaluate the ease of administration of vamorolone in boys ages 2 to <4 years with DMD and study medication acceptability of vamorolone in boys ages 7 to <18 years with DMD at daily oral doses of 2.0 mg/kg and 6.0 mg/kg;
- To investigate the effects of vamorolone administered orally at daily doses of 2.0 mg/kg and 6.0 mg/kg over a 3-month treatment period on pharmacodynamic (PD) biomarkers of safety and efficacy in boys ages 2 to <4 and 7 to <18 years with DMD.

4 Endpoints

This study does not test any formal statistical hypothesis and therefore the endpoints will be assessed without a pre-specified hierarchy.

4.1 Safety endpoints (endpoints related to the primary objective)

The following safety endpoints will be evaluated for all age groups:

- **Treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs)**: For more details on adverse event (AE) endpoints, see Section 15.3.

- **Change from baseline to Week 12 in Weight (kg, percentile and z-score)**

  The z-score is calculated using the following formula:

  if $L \neq 0$ then $z = ((\text{raw value} / M)^{L} - 1) / (L \times S)$

  else $z = \text{Ln}(\text{raw value} / M) / S$,

  where L, M, and S parameters are the median (M), the generalized coefficient of variation (S), and the power in the Box-Cox transformation (L) as provided in the reference tables describing the evolution of the BMI in the general population. The growth charts by Centers for Disease Control and Prevention (CDC) will be used for derivation of z-scores (and percentiles).

  Reference data will be extracted from:
  https://www.cdc.gov/growthcharts/percentile_data_files.htm.

- **Change from baseline to Week 12 in Height and Body Mass Index (BMI) (absolute, percentile and z-score)**:

  BMI will be derived from weight and height assessments as weight (kg) / height^2 (m^2). Height will be derived from either the ulna length or from standing height as follows:

  > Within an individual, the height will either be taken from CRF as the standing height or estimated based on ulna length. A mix of the two methods will not be used in the analyses within the same individual. For subjects in the 7 to <18 years age group the estimate based on ulna length will be used by default. However, if the baseline value is available only as standing height, standing height will be used instead. The height will be estimated from ulna length using the formula by Gauld (2004): Height (male) = ((4.605 x ulna length [cm]) + (1.308 x age [years]) + 28.003. In case the ulna length decreases compared to the previous visit, the ulna length from the previous visit will be imputed to the visit at which the ulna length decreased. If the age is >18 years, a value of 18 years will be used in the formula. The age will be calculated as date of assessment minus date of birth (+1 day) divided by 365.25.

  The BMI will be calculated only if both weight and height (either ulna length or standing height) have been measured at the same visit. Otherwise, the BMI value will be set as missing for the visit in question.

  As a sensitivity analysis implausible height (either standing or derived from ulnar length) values and their corresponding BMI values are removed from the analysis. Definition of implausible height is growth >0.8cm/week or decrease of >0.04cm/week from previous visit (starting from post-Baseline) to the next visit.

- **Change from baseline to each of the scheduled on-treatment and post-treatment assessment time points in vital signs:**

  The following endpoints will be evaluated: sitting blood pressure, heart rate, respiratory rate, and body temperature. For more details on vital signs endpoints, see Section 14.5.

- **Cushingoid features:**
  To determine whether subject has Cushingoid features at Baseline, medical history records are searched for Preferred Term = "Cushingoid". For Week 6 and 12 whether subject has Cushingoid features is based on Physical Examination. Number of subjects having Cushingoid at Baseline, Week 6 and 12 are presented together with other Physical Examination parameters.

- **Change from baseline to each of the scheduled on-treatment and post-treatment assessment time points in clinical laboratory values:**
  - Hematology and clinical chemistry
  - Lipid profile (triglycerides, total cholesterol, low density lipoprotein [LDL], high density lipoprotein [HDL])
  - Vitamin D level

  For more details on Laboratory endpoints, see Section 14.4.

- **Change from baseline to each of the scheduled on-treatment and post-treatment assessment time points in 12-lead ECG**

  The ECG endpoints include heart rate (HR), QRS duration, PR interval, RR interval, QT interval and QTc interval. The QTc interval will be derived based on QT and RR interval, using the correction of Fridericia ($QT / RR^{1/3}$) using RR derived from HR (60/HR). For more details on ECG endpoints, see Section 14.5.

- **Eye examination for detection of clinically significant abnormalities**

  The proportion of eyes with cataracts and the proportion of eyes with glaucoma will be evaluated. Furthermore, the intraocular pressure (IOP) will be summarized on eye-level.

  For more details on eye endpoints, see Section 14.5.

4.2 Pharmacokinetic endpoints (endpoints related to a secondary objective)

The pre-dose and post-dose plasma concentration measurements of Vamorolone at Day 1 and Week 2 will be described with descriptive statistics, figures and listings.

If a value is considered abnormal due to being inconsistent with the expected PK profile, the pharmacokineticist may exclude the value from the PK descriptive statistical analysis. However, the exclusion of any data will be documented in the CSR.

For the concentration summary and mean concentration figure the following rules will be applied:

• Mean concentration at any individual time point will only be calculated on valid values (i.e., quantifiable and not missing).
• In cases where a mean concentration is not calculated due to the above criterion not being met, the mean concentration will be set to missing.
• BLQ values will be set to zero.

Individual plasma concentrations (ng/ml) of vamorolone will be listed and summarized at each nominal time for the PK population. Individual plasma concentration (ng/ml) will be plotted on a linear and semi-log scale against actual sampling time points relative to dosing time.

For the summary plot, arithmetic mean (+/- SD) of concentrations (ng/ml) will be presented on a linear and geometric mean (*SD) of concentrations will be plotted in semi-log scale against nominal time.

LLOQ will be plotted as a reference line.

Separate popPk report (including AUCs and other relevant PK parameters) related to the other secondary endpoint is produced outside the context of this SAP.

4.3 Endpoints related to exploratory objectives

Efficacy endpoints

The efficacy endpoints include:

- Gross Motor Scaled Score (GMSS) based on Bayley-III Gross Motor scale (2 to <4 years age group only): The Bayley-III Gross Motor scale is comprised of 72 items that assess developmental functioning and include movement of the limbs and torso, static positioning (e.g., sitting, standing), dynamic movement including locomotion and coordination, balance, and motor planning. GMSS is derived as defined in Table A.1 of the Bayley-III Administration Manual (Bayley, 2006). The change from baseline to Week 12 will be calculated for the GMSS.

- Total upper extremity functional score and shoulder, middle and distal level domain scores based on the Performance of Upper Limb (PUL) scale (7 to <18 years age group only): The PUL version 2.0 includes an entry item to define the broad starting functional level, and 22 items subdivided into shoulder level (6 items), mid-level (9 items) and distal level (7 items). The entry items range from score 0 (no useful hand function) to score 6 (full shoulder abduction - no weakness). For weaker subjects a low score on the entry item means high level items do not need to be performed as they would not be achieved. The maximum score for the shoulder level is 12, for the mid-level 17 and for the distal level 13. The total score is achieved by adding the three-dimension level scores (max total score 42). The change from baseline to Week 12 will be calculated for all PUL scores.

Behavior, neuropsychology and physical functioning endpoints

Following endpoints are considered:

- Total and subscale scores of PARS III: The PARS III scale includes 28 items which all use a 4-point interval rating scale: ‘never or rarely’, ‘sometimes’, ‘often’ or ‘always or almost always’.

Of the 28 items, 20 items are reverse scored. Item responses are assigned 1–4 points on a likert-type scale, and the summation of all item responses results in the total score. Summation of specific items also yields six factor-derived psychosocial subscales:

- peer relations
- dependency
- hostility
- productivity
- anxiety/depression
- withdrawal.

Higher scores indicate better adjustment. Changes from baseline to Week 12 will be calculated for the total score and the subscale scores.

In addition to the raw total and subscale scores, the proportion of subjects with abnormal values will be calculated. A clinically abnormal value is defined as a z-score of -1 or less (i.e., z-score of at least 1 SD below the mean normative value). The z-score will be calculated as (Score–Mean)/SD, where Mean and SD represent the values presented in Table 1 and Score represent the individual subscale score from the present study.

The data will be normalized based on the data published by Hendriksen et al 2009. The cross-sectional data from the 287 boys with DMD, aged 5-18 years are summarized in Table 1.

Table 1 PARS-III by subscale in a cross-section study of 287 boys with DMD

Subscale	Mean	SD
Peer relations	8.93	3.09
Dependency	12.01	2.35
Hostility	18.88	4.11
Productivity	10.30	2.87
Anxiety/depression	20.02	2.98
Withdrawal	14.34	1.95

Source: Hendriksen et al 2009

- Subscale scores based on Pediatric Outcomes Data Collection Instrument (PODCI): PODCI consists of scores on seven core scales, four encompassing physical function and three assessing psychological well-being. Physical function core scales include: 1) Upper Extremity and Physical Function, 2) Transfer and Basic Mobility, 3) Sports and Physical Functioning, 4) Pain/Comfort. Psychological well-being core scales include: 1) Happiness, 2) Satisfaction and 3) Expectations. The scores will be calculated as defined in the POSNA (PODCI) Child Proxy Scoring Outcomes Instrument. The changes from baseline to Week 12 will be calculated for all PODCI scores.

Acceptability endpoints

The endpoints assessing study medication acceptability of vamorolone include:

- Ease of Study Medication Administration Assessment (2 to <4 years age group only): each question of the questionnaire will be summarized separately

- Study Medication Acceptability Assessment (7 to <18 years age group only): each question of the questionnaire will be summarized separately

Pharmacodynamic biomarkers of safety

The PD biomarkers of safety include:

- Adrenal suppression (Morning Cortisol)
- Bone turnover (Osteocalcin, serum P1NP, serum CTX)
- Insulin resistance (HbA1c, fasting glucose and insulin)
- Exploratory biomarkers for aspects of safety and efficacy (LH, FSH, TSH, FT4)

5 Sample size considerations

This is an open-label, parallel group, multiple dose study assessing safety and tolerability of vamorolone as the primary objective. Because of this, there is no formal sample size calculation. The sample size is considered sufficient to detect drug safety concerns in a pediatric population of ages 2 to <4 and 7 to <18 years with DMD based on clinical judgment and prior study of vamorolone at these doses. In addition, this sample size is sufficiently large to determine PK parameters.

A total of approximately 20 subjects will be enrolled within the 2-<4 years age group, as follows:
- Vamorolone 2.0 mg/kg/day (n=10); enrolled first
- Vamorolone 6.0 mg/kg/day (n=10); enrolled after previous dose group.

A total of approximately 34 subjects will be enrolled to treatment within the 7-<18 years age group, with 2.0 mg/kg/day groups enrolled first, as follows:
- Vamorolone 2.0 mg/kg/day, steroid untreated at entry (n=6)
- Vamorolone 2.0 mg/kg/day, steroid treated at entry (n=6)
- Vamorolone 6.0 mg/kg/day, steroid untreated at entry (n=6)
- Vamorolone 6.0 mg/kg/day, steroid treated at entry (n=6)
- Vamorolone 6.0 mg/kg/day, 12 to <18 years and steroid treated at entry (n=10; additional group).

6 Statistical hypotheses

This is an open-label study assessing the safety and tolerability of vamorolone as the primary objective. No formal statistical hypotheses will be tested.

7 Analysis sets

In general, the analyses of this study will be conducted separately for the following four subject groups:

- 2-<4 years age group
- 7-<18 years age group; steroid treated at entry
- 7-<18 years age group; steroid untreated at entry.
- 12-<18 years age group; steroid treated at entry.

Within these four subject groups the results are shown by dose level and overall (2.0 mg/kg, 6.0 mg/kg and Total) if not stated otherwise.

For selected tabulations (disposition, demographics, extent of exposure, compliance and adverse events), the data from the groups are also pooled together and shown as total Safety Set.

The following datasets will be evaluated:

- Screened Analysis Set: The Screened analysis set will include all subjects who have consented for the study, including subjects who failed the screening. Unless specified otherwise, the Screened analysis set will be used for subject listings and for the summary of subject disposition.

- Safety Set: All subjects who receive at least one dose of study medication will be included in the Safety Set. The Safety Set is the primary analysis population for safety and PD assessments. Results will be presented as treated.

- Pharmacokinetic (PK) Set: All subjects who receive at least one dose of vamorolone study medication and have sufficient data for PK analysis will be included in the PK Set.

8 General statistical considerations

All data from all subjects entered into the database will be included in subject data listings. The listings will be generally sorted by age, dose, prior steroid usage, centre and subject number (and by visit, if applicable).

Continuous variables will be summarized using the number of observations (n), mean, standard deviation (SD), Q1, median, Q3, minimum, and maximum. Standard error of the mean (SEM) will also be provided for summaries of efficacy data, where relevant.

Descriptive statistics for categorical data will include absolute (counts) and relative frequencies (percentages). All planned categories will be displayed if applicable. In case of event-based analyses (like Adverse Events) the total number of subjects in the treatment group (N) will be used as the denominator for percent calculations, regardless of subjects with missing values, unless stated otherwise. For categorical data presented by visit or timepoint the denominator is based on the observed ‘n’ per visit or timepoint, unless stated otherwise.

This study does not include any formal statistical testing and therefore no p-values testing statistical hypotheses will be presented. Data can be summarized with confidence intervals and for this purpose, two-sided 95% coverage will be used.

In the event that a subject is terminated early from this study, the early termination visit data will be analyzed as the next following scheduled visit where the assessments were to be measured.

Unless otherwise noted, baseline is defined as the last measurement taken prior to first exposure to study treatment.

Values from unscheduled visits are not included in the analyses if not otherwise specified. If measurement from unscheduled visit is the last before study treatment start it is used as baseline (except local laboratory measurements). Data gathered at unscheduled visits will be included in by-subject data listings.

All statistical analyses and summaries will be produced using SAS version 9.4 or higher. Deviations from the statistical plan will be reported in the clinical study report, including the rationale for the deviation.

9 Subject disposition

The subject disposition will be summarized as follows and presented for each age and dose group, as applicable, and overall.

- Screening (% calculated from the Screening Analysis Set)
  - The number of subjects screened
  - The number (%) of subjects who failed screening, including the distribution of reasons for failing the screening (including only the last screening per subject)

- Treatment phase (% calculated from the Safety Set)
  - Number of subjects who got treatment group assigned and at least one dose of study medication (Safety Set)
  - Number (%) of subjects who completed the study
  - Number (%) of subjects who discontinued the study prematurely, including the distribution of reasons for premature discontinuations

- Treatment Continuation (% calculated from the number of subjects who completed the study)
  - Number (%) of subjects by the continuation treatment (e.g., EAP or standard glucocorticoid treatment)

- Datasets used for other analysis sets (% calculated from the Safety Set)
  - Number (%) of subjects included in the PK analysis set

10 Demographic and other baseline characteristics

Demographic and other baseline characteristics will be summarized descriptively for the Safety set and for PK set if it is different from Safety set. For categorical variables, a missing data category will be presented in case of missing baseline data.
The following variables will be summarized:

- Demographics: age (calculated as date of informed consent minus date of birth*, divided by
- 365.25), ethnicity, race (subjects that indicate more than 1 race will be counted under “Multiple”), height, height percentile and z-score, weight, weight percentile and z-score, body mass index (BMI), BMI percentile and z-score.

- In case a subject has multiple informed consent dates, the date closest to and prior to baseline (i.e. the latest date) will be used in the calculations.

- DMD history: age at first signs of DMD symptoms.

- Mutation type and proportion of subjects with potential leaky/non-null mutations (i.e., exon 44 skippable, exon 8 skippable and splice site mutations), ambulatory status

All demographic and baseline information, including DNA and muscle biopsy diagnosis information will be presented in by-subject listings.

* = Due to country/regional regulations, subject's dates of birth have not been entered to the study database, however age is recorded in two parts, age in years and months. Date of Birth is derived by converting age reported at baseline in years and months to days +15, after which Date of Birth = Date of Informed Consent - Age in days.

11 Prior and concomitant medication/ concomitant non-drug treatments

A categorical summary of prior and all concomitant medications and non-pharmacological treatments taken prior to and during the course of the study will be presented in tabular form summarizing the number and percentage of subjects by Anatomical Therapeutic Chemical (ATC) drug classification level and preferred name using the World Health Organization (WHO) Drug classification for the Safety set. At each level of tabulation subjects will be counted once if they had one or more instance of medication usage.

A concomitant medication is defined as any medication with stop date after the day of first exposure to study drug and start date before final study visit.

Prior medications are defined as any medications with end date prior to or at the day of first exposure to study drug.

For partial dates following imputation rules are followed:

Start date:

- Missing day, month, and year: No imputation done
- Missing day and month:
  - If year of Start Date is equal to year of Study Treatment Start then day of Study Treatment Start and month of Study Treatment Start is imputed for missing day and month respectively.
  - If year of Start Date is earlier or after than the year of Study Treatment Start then first of January is imputed.

- Missing day only:
  - If month and year of Start Date are same as month and year of Study Treatment Start then day of Study Treatment Start is imputed for missing day.
  - Else first of month is imputed for missing day.

Note: If End Date is not missing and imputed Start Date would be greater than End Date then Imputed Start Date is set to be equal to End Date.

End Date:

- Missing day, month, and year: No imputation done.
- Missing day and month:
  - 31st of December is imputed.
- Missing day only:
  - last day of the month (28/29/30/31 depending on the month) is imputed for missing day

All prior medications, concomitant medications, and non-pharmacological treatments will be presented in by-subject listings.

12 Prior use of Glucocorticoids for DMD

For steroid-treated subjects in 7-<18 years age group (and separately 12-<18 years), the prior use of standard of care glucocorticoids for DMD is summarized by the prior medication (Deflazacort, Prednisone and in Total).
The following variables will be summarized with descriptive statistics:

- Number of subjects (%) using each prior medication
- Duration (months) of the use of each prior medication
- Dose at study start (mg/kg), derived by dividing the actual dose by weight at Baseline
- Dose at study start (% of recommended dose), derived as $mg/kg_{(observed)}$ / $mg/kg_{(recommended)}$ * 100, where recommended dose is 0.9mg/kg for Deflazacort and 0.75mg/kg for Prednisone.

13 Medical history

Subject medical history is collected at the screening visit. Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The number and percentage of subjects with any medical history events will be summarized by system organ class (SOC) and preferred term (PT) for the Safety analysis set by group. At each level of tabulation subjects will be counted once if they had 1 or more of each such event.

Subject medical history data will be presented in a by-subject listing.

14 Analysis of safety and tolerability

The Safety Set will be used for all safety analyses.

14.1 Extent of exposure

The following information will be summarized:

- The number of subjects exposed to study treatment,

- Duration of exposure (weeks): (date of last dose of study medication – date of first dose of study medication + 1 day) / 7,
    - The durations will be summarized with descriptive statistics.

- Total exposure to study treatment, expressed as patient-years (sum of duration of exposure to study treatment over all subjects in days divided by 365.25).

14.2 Compliance

Suspension administration and compliance will be calculated based on the weight of suspension in grams. For each subject, the total suspension administered will be estimated from the total weight of the suspension bottles dispensed and returned (amount dispensed g – amount returned g), as collected on the CRF. Further, the amount of suspension prescribed is captured on the CRF in mL. This data will be used to calculate the total suspension prescribed for each subject and converted to grams (1 mL suspension = 1 g suspension). Compliance percentage will then be calculated as the total suspension administered g / total suspension prescribed g x 100%.

If subject completed the study without returning all the bottles dispensed the compliance is not calculated. However, if subject discontinues for example between Weeks 6 and 12, subject's compliance will be calculated up to Week 6 if all the bottles are returned at Week 6. If drug is spilled or dose is missed, compliance is not calculated for the corresponding time interval. But if study treatment has been interrupted, number of days are corrected to account for the interruption.

Another compliance parameter (Dose % of Expected) compares the actual daily dose received with the expected dose (the treatment group 2 or 6 mg/kg/day). Actual daily dose is calculated followingly: Daily Dose = Total Suspension Administered / Weight(kg) / Days Administered. If the subject's weight changes (and affects dosing) the Daily Dose is first calculated per interval where the weight is considered constant. Then the Daily Dose for the whole duration of the study is calculated as weighted average (weighted by days) of multiple interval Daily Doses. If the subject's weight rises to 50kg, weight in the equation is set to 50kg. Finally, Dose % of Expected = Daily Dose / Expected Dose * 100%.

Both compliance parameters will be summarized with descriptive statistics. Compliance categories < 80%, 80-120% and > 120% will also be summarized with subject count and % of observed compliances. Also, listings are provided.

14.3 Adverse events

All adverse events (AEs) will be coded with MedDRA. Treatment-emergent AEs (TEAEs) are defined as any new or worsening events after initiation of the study treatment.

If not stated otherwise Adverse Event tables include frequency (number of events), Rate (numbers of events per subject-year) and Incidence (number and % of subjects).

The following summaries will be provided:

- An overall summary table of TEAEs include following categories: Any TEAEs, drug-related TEAEs, severe TEAEs, serious TEAEs (other than death), TEAEs leading to permanent study

treatment discontinuation, TEAEs leading to temporary dose interruption, and TEAEs leading to death.

- TEAEs broken down by SOC and PT.
- TEAEs broken down by PT, sorted in descending order of PTs in the overall group.
- TEAEs broken down by Causality, SOC and PT. For Incidence (number and % of subjects) subject is counted per SOC and PT category only for the most causal case.
- TEAEs broken down by Severity, SOC and PT. For Incidence (number and % of subjects) subject is counted per SOC and PT category only for the most severe case.
- TEAEs leading to death broken down by SOC and PT.
- Listing of serious TEAEs
- Listing of TEAEs leading to withdrawal from study

14.4 Laboratory safety variables

All clinical laboratory parameters will be converted to consistent units according to the International System of Units (SI) before summarization.

Age dependent reference ranges applicable to the pediatric population included in the study were defined by medical professional based on literature (Provided in Appendix 2). The age at the final laboratory measurement per subject and parameter is used to determine the reference range for all visits (including for biomarkers).

The following will be summarized:

- Numeric laboratory parameters: Actual values and change from baseline for each parameter will be summarized with descriptive statistics. These summary tables will include only central laboratory data.
- Values for liver function parameters (ALT, AST, Total Bilirubin, and GGT) are summarized by giving number and percentage (out of population N) of subjects having measurements over ULN for baseline and post-baseline. Values are categorized in following categories:
  >ULN
  >1.5 x ULN (for Bilirubin only)
  >2 x ULN
  >3 x ULN
  >5 x ULN
  For post-baseline each subject is included only in the most severe category that their measurement reaches post-baseline. For total bilirubin post-baseline values >ULN will only be presented in the table if the baseline value was within the normal range, ie. treatment emergence increase. All bilirubin values for subjects with high bilirubin values at baseline, including direct bilirubin, if available, will be presented in listings.

- Number and percentage (out of population N) of subjects meeting Hy's law criteria at least once during post-baseline are described in binomial categories 'Yes' or 'No'.
    - Hy's law criteria: ALT or AST > 3 times above upper limit of normal (ULN) and simultaneously total bilirubin >2 times above ULN.

- All other laboratory parameters which have an upper or lower reference range: Number and percentage of subjects with low, normal or high (i.e., below, within or above reference range) values for each parameter will be summarized
    - For post-baseline only the most extreme case per subject is kept. If subject has both Low and High value post-baseline this will be depicted in 'Low and High' category.
    - Post-baseline values will be presented as a shift table, i.e. the distribution of the response categories at post-baseline will be classified by the baseline category. The shift tables will include also local laboratory data.

- Categorical laboratory parameters are presented in the listings.

Laboratory results that are below the lower limit of quantification will be divided by 2 for analysis (e.g., a laboratory response of "<0.1" will be analyzed as 0.05). Lab results that are above the upper level of quantification will be analyzed using the upper level of quantification (e.g., a response of ">1500" will be analyzed as 1500).

14.5 Other safety endpoints

Vital signs

Vital signs (blood pressure (including z-score and percentile), heart rate, respiratory rate parameters) will be summarized at each assessment time point using descriptive statistics and presented for observed response as well as change from baseline.

ECG

The ECG parameters (QRS duration, HR, and the PR, RR, QT, QTc intervals) will be summarized by visit with descriptive statistics.

Overall ECG interpretation will be summarized with listing.

In addition, the number and proportion of subjects with elevated QTc values will be summarized with categorical tables at baseline and at Week 12. The following categories will be used in the tables: QTc ≤450 ms, >450 – 480 ms, >480 – 500 ms and >500 ms. Also, change from baseline for QTc is summarized (number and proportion of subjects) in following categories: <30ms, 30-60ms and >60ms.

Height, weight and BMI

Body weight, height and BMI will be summarized at each assessment time point using descriptive statistics and presented for observed response as well as change from baseline. All parameters will be presented as absolute values and as percentiles and z-scores.

As a sensitivity analysis implausible values on height (and BMIs affected by that) are excluded from the analysis.

Eye examination

The frequency and percentage of eyes with a cataract or a glaucoma will be provided by visit.

IOP per individual eye is summarized with descriptive statistics by visit. Also, proportion of eyes with IOP above 21 mmHg are summarized by visit. If eye has multiple IOP values on the same visit, the lowest is used.

Other collected eye examination information (e.g. Cup to disc ratio, severity of cataracts) are described in listings.

Physical examination

In the eCRF it is possible to answer "NO NEW FINDINGS" for Physical examinations as a whole. In these cases, the results from previous visits are duplicated to the current visit.

Physical examination results are described with descriptive statistics (count and %) by body system and visit. Proportion of subjects having Cushingoid features (from Medical history for baseline and from Physical Examination CRF page for on-treatment visits) and Deterioration of Muscle Condition are presented together with physical examination results.

Biomarkers

Serum PD biomarkers will be summarized with descriptive statistics and with tables presenting categorical changes and with figures as described below.

Morning cortisol: shift tables for morning cortisol will be presented using the categories low, normal, high, as per the normal range for pediatric populations.

Glucose: shift tables for glucose will be presented using the categories low, normal, high, as per the normal range for pediatric populations.

Glycosylated haemoglobin: Proportion of subjects presenting with changes from baseline at Visit using the following categories: ≥10% decrease from baseline, no change, ≥10% increase from baseline (cut offs defined as per Diabetes Care 2015, Staging Pre-symptomatic Type 1 Diabetes).

Bone turnover markers: shift tables for each of the bone turnover markers using the following categories low, normal, high as per the normal ranges for pediatric populations, i.e. adjusted by age.

Insulin resistance (HbA1c, fasting glucose and insulin): Observed values are presented with individual line plots.

LH and FSH: Observed values are presented with individual line plots. Also, possible reach of Tanner Stage 2 or Tanner Stage 3 (or above) are presented per parameter and combined (with

categories “Stage 1”, “Stage 2” and “Stage 3 or above”). Minimum criteria for Tanner Stage 2 for LH is 0.26 IU/L and for FSH 0.72 IU/L, and for Tanner Stage 3 0.64 IU/L and 1.24 IU/L respectively (Howard 2021). Values below Stage 2 criteria are categorized to “Stage 1”. For presentation of parameters combined the maximum Stage from the two parameters is selected.

TSH and FT4: Observed values are presented with individual line plots.

Protocol Deviations

The data set of Protocol deviations is produced based on the “Cumulative protocol deviations” listing received from TRiNDS and including Important / Non important designation.
Only PDs classified as Important will be extracted along with their “Category for important designation” becoming PD Category:

- Inappropriate consent process
- Participant entered study without satisfying entry criteria
- Participant developed withdrawal criteria during the study but were not withdrawn
- Participant received wrong treatment or incorrect dose
- Participant received excluded concomitant medication
- Participant data not reported
- Accuracy and reliability of the study data
- Other

In addition, in each PD category, sub-categories may be added by the Clinical scientist to increase granularity of PDs categorization.

Important protocol deviations will be summarized with categorical descriptive statistics by PD category and sub-category, and also with a listing.

Analysis of efficacy

All efficacy endpoints (including behavior, neuropsychology and physical functioning) will be summarized descriptively for each group in the Safety Set.

The endpoints assessing acceptability will be tabulated with descriptive statistics.

Deviations from the analyses planned in the study protocol

Differences between clinical study protocol and statistical analysis plan

Protocol	SAP	Reason for change
mITT population included	No mITT	mITT deemed irrelevant
MMRM and ANCOVA analyses defined	No MMRM or ANCOVA analyses	For study this small and short the statistical models defined in the protocol are not suitable

Execution of statistical analysis

The statistical analyses of this study will be performed by Santhera Pharmaceuticals.

Hardware and software

Statistical analysis, tables and subject data listings will be performed with SAS® System for Windows (SAS Institute Inc., Cary, NC, USA) version 9.4 or higher.

Revision History

Version	Version Date	Changes
0.1	2023-05-23	Initial draft
1.0	2024-05-02	Version 1.0
2.0	2024-06-25	New tables/listings added

References

Gauld LM, Kappers J, Carlin JB, Robertson CF. Height prediction from ulna length. Dev Med Child Neurol. 2004 Jul;46(7):475-80.

Hendriksen JG, Poysky JT, Schrans DG, Schouten EG, Aldenkamp AP, Vles JS. Psychosocial adjustment in males with Duchenne muscular dystrophy: psychometric properties and clinical utility of a parent-report questionnaire. J Pediatr Psychol. 2009 Jan-Feb;34(1):69-78.

Signatures

Function / Role	Name	Date/Signature
Author Statistician	PI	PI
Owner Program Statistician	PI	PI
PPD	PI	PI

Appendix 1 – Statistical Output Table of Content

Appendix 2 – Laboratory Reference Ranges

VBP15-006 labrefs_withAges.xlsx

SAP VBP-006 v2-0

Final Audit Report 2024-07-11

Created:	2024-06-25 (Eastern European Time)
By:	PI (PI)
Status:	Signed
Transaction ID:	PI

"SAP VBP-006 v2-0" History

Document created by PI (PI)
2024-06-25 - 1:45:12 PM GMT+3

Document emailed to PI for signature
2024-06-25 - 1:46:55 PM GMT+3

Email viewed by PI
2024-06-25 - 1:49:01 PM GMT+3

Signer PI entered name at signing as PI
2024-06-25 - 1:49:59 PM GMT+3

Document e-signed by PI (PI)
Signature Date: 2024-06-25 - 1:50:01 PM GMT+3 - Time Source: server

Document emailed to PI for signature
2024-06-25 - 1:50:03 PM GMT+3

Email viewed by PI
2024-06-25 - 2:42:46 PM GMT+3

Signer PI entered name at signing as PI
2024-06-25 - 2:43:52 PM GMT+3

Document e-signed by PI (PI)
Signature Date: 2024-06-25 - 2:43:54 PM GMT+3 - Time Source: server

Document emailed to PI for signature
2024-06-25 - 2:43:55 PM GMT+3

Email viewed by PI
2024-07-11 - 12:52:07 PM GMT+3

Adobe Acrobat Sign

New document URL requested by [redacted]
2024-07-11 - 12:52:12 PM GMT+3

Signer [redacted] entered name at signing as [redacted]
2024-07-11 - 12:53:43 PM GMT+3

Document e-signed by [redacted]
Signature Date: 2024-07-11 - 12:53:45 PM GMT+3 - Time Source: server

Agreement completed.
2024-07-11 - 12:53:45 PM GMT+3