Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT04970654
Sponsor trial ID:	NN8640-4468
Official title of study:	A trial comparing the efficacy and safety of once weekly dosing of somapacitan with daily Norditropin® in Chinese children with growth hormone deficiency.
Document date:	15 December 2023

Statistical Analysis Plan

Protocol title: A trial comparing the efficacy and safety of once weekly dosing of somapacitan with daily Norditropin® in Chinese children with growth hormone deficiency

Substance: Somapacitan

Table of contents

Page

Table of figures

Page

Table of tables

Page

Version History

This Statistical Analysis Plan (SAP) for study NN8640-4468 is based on the protocol version 1.0 dated 01 July 2020.

SAP Version	Date	Change	Rationale
1.0	28 April 2023	Not Applicable	Original version
2.0	15 December 2023	Section 4.3.2: Secondary endpoints, height SDS	As Chinese standards data did not had monthly LMS values, they were calculated using the available reference data. SAP has been updated by describing the methodology used to calculate monthly LMS values. Height SDS was then calculated using these derived LMS values. No changes were made to primary analysis.

List of abbreviations

ADHD	attention deficit hyperactivity disorder
AE	adverse event
AESI	adverse event of special interest
ANCOVA	analysis of covariance
CI	confidence interval
CTR	clinical trial report
ECG	electrocardiogram
FAS	full analysis set
FDAAA	the Food and Drug Administration Amendments Act
GH	growth hormone
GHD	growth hormone deficiency
HbA1c	glycated haemoglobin
hGH	human growth hormone
HV	height velocity
ICH	International Council on Harmonization
IGF-I	insulin-like growth factor I
IGF BP-3	insulin-like growth factor binding protein 3
MAR	missing at random
MedDRA	medical dictionary for regulatory activities
MMRM	mixed model for repeated measurements
PD	pharmacodynamics

PK	pharmacokinetics
PP	per protocol analysis set
PRO	patient reported outcome
SAE	serious adverse event
SAP	Statistical Analysis Plan
SAS	safety analysis set
SDS	standard deviation score
TFL	tables, figures and listings

1 Introduction

Primary analysis of the primary endpoint addressing the primary estimand was defined in protocol before first patient first visit (FPFV). There are no changes to the analysis described in the protocol.

This SAP is based on the protocol: *A trial comparing the efficacy and safety of once weekly dosing of somapacitan with daily Norditropin® in Chinese children with growth hormone deficiency*. REAL 6, version 1.0 (dated 01 July 2020). This SAP covers specification of statistical considerations and analyses for efficacy and safety data. SAP 1.0 was created prior to unblinding but post FPFV. Sensitivity analysis of primary endpoint has been included to investigate possible impact of delayed visits due to COVID.

1.1 Objectives, Endpoints, and Estimands

1.1.1 Primary objective:

To compare efficacy of somapacitan vs Norditropin® on longitudinal growth in Chinese children with GHD.

1.1.2 Secondary objective:

To compare safety of somapacitan vs Norditropin® in Chinese children with GHD

1.1.3 Primary endpoint

Table 1 Primary endpoint

Endpoint title	Time frame	Unit
Height Velocity	Height velocity (annualised) at week 52	cm / year

1.1.4 Secondary endpoints

1.1.4.1 Supportive secondary endpoints

Table 2 Supportive secondary: Efficacy endpoints

Endpoint title	Time frame	Unit
Change in bone age	From visit 1 to week 52	Years
Change in Height Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10

Change in Height Velocity Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10

Table 3 Supportive secondary: Safety endpoints

Endpoint title	Time frame	Unit
Change in fasting plasma glucose	From baseline (week 0) to week 52	mmol/l
Change in HbA1c	From baseline (week 0) to week 52	%

Table 4 Supportive secondary: Pharmacodynamics

Endpoint title	Time frame	Unit
Change in IGF-I Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10
Change in IGFBP-3 Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10

1.1.5 Primary estimand

Hypothetical strategy - ancillary therapy not available: The treatment difference between somapacitan and Norditropin® in mean annualised HV at week 52 if ancillary therapy had not been available prior to week 52 (i.e. assuming no initiation of ancillary therapy) in children with GHD. The use of ancillary therapy may lead to attenuation of the treatment effect of interest or even exaggerate the treatment effect and the estimand thus aims to reflect the treatment difference attributable to the initially randomised treatments.

1.2 Study Design

A randomised open-labelled two arm phase 3 trial designed to confirm non-inferiority of efficacy and investigate safety of once weekly subcutaneous treatment of somapacitan compared to daily subcutaneous growth hormone (Norditropin®) treatment in Chinese prepubertal children with growth hormone deficiency.

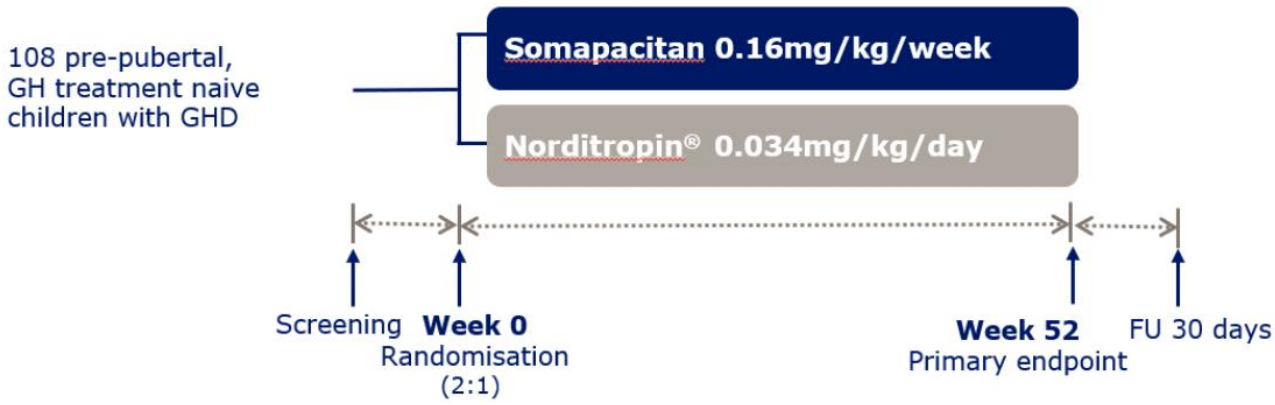
The total trial duration for a subject will be up to 70 weeks approximately.

The trial duration includes a 2 to 14 weeks of screening period, a 52 weeks of treatment period and a minimum 30 days of follow up period.

Eligible subjects will be randomised in a 2:1 manner to receive either somapacitan or Norditropin®

The non-inferiority margin of -2.0 cm/year is used in the trial. The randomisation will be stratified by age (<6 versus ≥6 years), gender (boys versus girls) and GH peak (< 7 versus ≥ 7 ng/ml) to minimize bias on the primary endpoint.

Figure 1 Study design



2 Statistical Hypothesis

Hypothesis testing for the primary endpoint will be done by testing H0: D≤ -2 cm/year vs HA: D>-2 cm/year, where D is the mean treatment difference (somapacitan – Norditropin®). Non-inferiority of somapacitan will be considered confirmed if the lower boundary of the two-sided 95% confidence interval is above -2 cm/year. As only one confirmatory hypothesis is to be tested, no further control for multiplicity is needed.

2.1 Multiplicity Adjustment

Not applicable

3 Analysis Sets

The following populations are defined:

Population	Description
Full analysis set (FAS)	All subjects randomised. Exclusion of data from analyses should be used restrictively, and normally no data should be excluded from the FAS. Subjects will be analysed according to the randomised treatment
Safety analysis set (SAS)	All subjects randomly assigned to trial treatment and who take at least 1 dose of trial product. Subjects are analysed according to the treatment they actually received.
Per protocol analysis set (PP)	Subjects from FAS who have not violated any inclusion/exclusion criteria and have used the randomised treatment for at least 47 weeks (for subjects receiving somapacitan) or 329 days (for subjects receiving Norditropin) corresponding to 90% of the planned exposure. Subjects are analysed according to the treatment they actually received.

The subjects or observations to be excluded, and the reasons for their exclusion must be documented before unblinding. The subjects and observations excluded from analysis sets, and the reason for this, will be described in the CTR. Two observation periods are defined:

• on-treatment: from first administration and up until last trial contact, visit 7 or 14 days after last administration, whichever comes first

• in-trial: from first administration and up until last trial contact or visit 8, whichever comes first Analysis based on the ‘in-trial’ observation period is to be viewed as supplemental analysis to the analysis based on the ‘on-treatment’ analysis.

4 Statistical Analyses

4.1 General Considerations

The one-sided test used for the primary endpoint is based on an alpha level of 2.5%. All other statistical tests conducted will be two-sided on the 5% significance level.
Two factors are defined as follows:

• Age group: Age at randomisation (<6 versus ≥6 years)
• GH peak group: Growth hormone peak level (< 7 versus ≥ 7 ng/ml)

All efficacy endpoints will be analysed using FAS and all safety endpoints will be analysed using SAS.
The primary endpoint will additionally be analysed using PP as a support to the results achieved using FAS under the hypothetical strategy.

4.2 Primary Endpoint Analysis

The primary analysis of the primary endpoint Height velocity (cm/year) at week 52, addressing the primary estimand is based on the FAS but data assessed after discontinuation of randomised treatment will be disregarded in the analysis.

4.2.1 Definition of Endpoint

The primary endpoint: Height velocity (cm/year) (HV) at week 52

- HV will be derived from height measurements taken at baseline and the week 52 visit (landmark visit) in the following way:

HV = (height at 52 weeks visit - height at baseline)/(time from baseline to 52 weeks visit in years).

Annualized HV at the intermediate visits: Week 13, 26 and 39, will be derived analogously to the week 52 HV: HV = (height at j weeks visit - height at baseline)/(time from baseline to j weeks visit in years), j=13, 26, 39.

4.2.2 Main Analytical Approach

In order to estimate the primary estimand, a mixed model for repeated measurements (MMRM) with an unstructured covariance matrix is conducted on HV data (annualized HV at planned visits at week 13, week 26, week 39 and week 52) up to discontinuation of randomised treatment for each treatment arm using all randomised subjects and assuming missing at random (MAR) for both treatment arms. The MMRM will include sex, age group, GH peak group and sex by age group interaction term as factors and baseline height as a covariate, all nested within week as a factor. From this analysis an estimate of the treatment difference at week 52 with corresponding 95% CI will be presented.

Non-inferiority of somapacitan will be considered confirmed if the lower boundary of the two-sided 95% confidence interval is above -2 cm/year.

The primary endpoint will be analysed based on the 'on-treatment' observation period.

4.2.3 Sensitivity Analysis

4.2.3.1 Tipping point analysis

A tipping point analysis will be conducted as a sensitivity analysis for the analysis of the primary endpoint for the primary estimand. For this analysis, the missing data are expected to be mainly due to subjects that are withdrawn from the trial or discontinue randomised treatment. The sensitivity analyses described below will be used to investigate whether the results from the primary analysis are robust against departures from the assumption of MAR.

Let δ be defined as the difference between the mean of the observed data and the mean of the unobserved data μobs-μunobs, adjusted for other observed data. Under an MAR analysis, δ is assumed to be 0. Positive values of δ indicate that subjects with missing endpoint values have smaller HV than subjects with observed endpoint values. If subjects primarily withdraw or discontinue randomised treatment/start ancillary therapy due to a perceived lack of efficacy, then this could be the most likely direction of departure from MAR. Let f1 and f0 be the fractions of subjects with unobserved endpoint data in the somapacitan and Norditropin® arms, respectively. The sensitivity analysis is done by subtracting a quantity Δ from the treatment effect estimate under the MAR assumption, where Δ= f1δ if data depart from MAR in the somapacitan arm only, Δ= -f0δ if data depart from MAR in the Norditropin® arm only, and Δ= (f1-f0)δ if data depart from MAR in the

same way in both arms. The calculations for the 3 scenarios will use a range of δ values increasing from 0 until the resulting 95% CI no longer is completely above -2cm/year for the most conservative evaluation (data depart from MAR in the somapacitan arm only) and the approximation that the standard error of the treatment difference is unaffected by the sensitivity analysis[1]. All subjects from the FAS can be viewed as included in this analysis as subjects with missing endpoint data will be contributing to one of the fraction values f1 and f0.

4.2.3.2 Subjects whose height measurement at visit 7 (week 52) was performed later than planned

Additional sensitivity analysis of primary endpoint using single imputation will be performed to investigate the impact of delayed visit on height velocity after 52 weeks of treatment. The single imputation will be done by making a linear interpolation between that last observed height prior to week 52, meaning the visit 6 (week 39) height measurement, and the observed delayed visit 7 height measurement. From this linear interpolation we find the height at week 52 and calculate a week 52 HV and data will be analysed using the same analysis model as was used for analysing the primary endpoint for the primary estimand. From this analysis an estimate of the treatment difference at week 52 with corresponding 95% CI will be presented.

4.2.3.3 Adjusting for baseline IGF-I SDS

As an additional sensitivity analysis the primary endpoint for the primary estimand will also be analysed adding baseline IGF-I SDS as a covariate in the primary analysis model. From this analysis an estimate of the treatment difference at week 52 with corresponding 95% CI will be presented

4.2.4 Supplementary Analysis

An analysis of the primary endpoint using the same analysis model as was used for the primary analysis under the primary estimand but based on PP instead of FAS will be conducted as a supplementary analysis.

4.3 Secondary Endpoints Analysis

4.3.1 Confirmatory Secondary Endpoints

Not applicable

4.3.2 Supportive Secondary Endpoints

Table Efficacy

Endpoint title	Time frame	Unit
Change in bone age	From visit 1 to week 52	Years
Change in Height Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10
Change in Height Velocity Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10

The efficacy endpoints will be analysed based on the 'on-treatment' observation period as well as the 'in-trial' observation period within the trial period (52 weeks of treatment).

Height SDS will be derived using Chinese general population standards[2] and HV SDS will be derived using Prader standards[3] as reference data.

The formula to calculate height SDS is as below:

Height SDS = ((Height / M)**L-1) / (L*S)

Height: height at the time of assessment,

L: The sex and age-specific power in the Box-Cox transformation,

M: The sex and age-specific median,

S: The sex and age-specific generalized coefficient of variation

The Chinese standards data contained three heights in cm corresponding to the median and plus/minus 2 SD for ages 0 to 18 years every six month, separately for girls and boys. Using these data as input, LMS values were calculated using the LMSfit function in R4.2.0 at these ages.

With the calculated LMS values as input the smooth.spline function with degrees of freedom 10 for L, 15 for M and 25 for S was used to obtain smoothed monthly estimates of LMS, separately for girls and boys. Finally, height SDS was calculated using above formula based on the smoothed monthly estimates.

Change in height SDS and HV SDS will be analysed using the same analysis model as was used for analysing the primary endpoint for the primary estimand except for using baseline height SDS and baseline HV SDS, respectively, as a covariate in the model instead of baseline height. The estimate for the treatment difference at week 52 will be reported with corresponding 95% CI and p-value.

Bone age will be analysed using an ANCOVA model on bone age/chronological age assessed at week 52 and the model will include treatment, sex, age group, GH peak group and sex by age group

interaction term as factors and baseline bone age/chronological age as a covariate. The treatment difference estimate will be reported with corresponding 95% CI and p-value.

Table Safety

Endpoint title	Time frame	Unit
Change in fasting plasma glucose	From baseline (week 0) to week 52	mmol/l
Change in HbA1c	From baseline (week 0) to week 52	%

The safety endpoints will be analysed using descriptive statistics based on the 'on –treatment' observation period and the 'in-trial' observation period.

Table Pharmacodynamics

Endpoint title	Time frame	Unit
Change in IGF-I Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10
Change in IGFBP-3 Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10

The PD endpoints will be analysed based on the 'on –treatment' observation period and the 'in-trial' observation period.

Change in IGF-I SDS and IGFBP-3 SDS will be analysed using an MMRM with an unstructured covariance matrix on all relevant post-baseline change from baseline values as dependant variables. The model will include treatment, sex, age group, region, GH peak group and sex by age group interaction term as factors and baseline value as a covariate, all nested within week as a factor. From the MMRM, the treatment difference at week 52 will be estimated and the corresponding 95% CI and p-value will be reported for each endpoint.

4.4 Exploratory Endpoints Analysis

Not applicable

4.5 Other Safety Analysis

The subsections refer to analysis of other safety parameters. For calculation of derived safetyparameters please see appendix 1, section 6.1.

4.5.1 Extent of Exposure

Not applicable

4.5.2 Adverse Events

Adverse events will be analysed using descriptive statistics based on the 'on –treatment' observation period (primary evaluation) and 'in-trial' observation period (secondary evaluation) within the main trial period (52 weeks of treatment). The adverse events will be summarised by treatment, MedDRA (Medical Dictionary for Regulatory Activities) system organ class and MedDRA preferred term. The descriptive statistics will include the number and percentage of subjects who experienced adverse events, the number of events and rate. Adverse events will be listed by treatment and subject with information on severity, relationship to trial product and demographics based on the 'on –treatment' observation period. Adverse events with onset 14 days or more after last trial drug administration will be reported in a separate listing. Adverse events with onset before first dosing will be reported in a separate listing.

4.5.3 Additional Safety Assessments

Safety laboratory assessments

Safety laboratory assessments (biochemistry, haematology, glucose metabolism, hormones, antibodies and lipids) will be summarised by treatment and time of assessment. All abnormal values will be listed.

Vital signs and Electrocardiograms

Vital signs (pulse rate, systolic and diastolic blood pressure) will be summarised by treatment and time of assessment. The investigator's evaluation of ECG classifications will be tabulated by treatment and time of assessment.

Antibodies

For participants randomised to somapacitan with confirmed anti-somapacitan antibody positive samples, the anti-somapacitan positive cross-reactivity to hGH antibodies and in vitro neutralising

effect will be tabulated by time of assessment.

4.6 Other Analyses

The subsections refer to analysis of other parameters. For calculation of other parameters please see appendix 1, section 6.1.

4.6.1 PRO analysis

The following Patient reported outcome questionnaires would be collected in the trial:

• GHD-CIM (Growth Hormone Deficiency – Child Impact Measure)

• GHD-CTB (Growth Hormone Deficiency – Child Treatment Burden)

• GHD-PTB (Growth Hormone Deficiency – Parent Treatment Burden)

The PRO data will be analysed based on the 'on –treatment' observation period only.

Title	Time frame	Unit
Change in GHD-CIM (domain and total scores)	From Baseline (Week 0) to visit 5 (Week 26), Baseline (Week 0) to Week 52	-100 to 100
GHD – CTB (domain and total scores)	At visit 5 (Week 26) At visit 7(Week 52)	0 to 100
GHD – PTB (domain and total scores)	At visit 5 (Week 26) At visit 7(Week 52)	0 to 100

GHD – CTB scores and GHD – PTB scores will be analysed using an MMRM with an unstructured covariance matrix on all relevant post-baseline values as dependant variables. The model will include treatment, sex, age group, GH peak group and sex by age group interaction term as factors, all nested within week as a factor. From the MMRM, the treatment differences will be estimated and the corresponding 95% CI and p-values will be reported for week 26 and week 52, respectively.

Changes from baseline to week 26 and week 52 in GHD-CIM scores will be analysed using an MMRM with an unstructured covariance matrix. The model will include treatment, sex, age group, GH peak group and sex by age group interaction term as factors and baseline value as a covariate, all nested within week as a factor. From the MMRM, the treatment differences will be estimated and the corresponding 95% CI and p-values will be reported for week 26 and week 52, respectively.

In addition, the following questions of the physical domain of the GHD-CTB PRO will be analysed using Wilcoxon rank sum tests:

- In the past week, how much did your child's injections hurt
- In the past week, how much did your child's injections sting
- In the past week, how much bruising did your child have from their injections
- In the past week, how much soreness did your child have at their injection site

4.6.2 Pharmacokinetic and/or pharmacodynamic modelling

Somapacitan and IGF-I serum concentration data may be used for population PK, population PK/PD and exposure-response modelling, potentially as a joint analysis of data from multiple trials. Other exploratory PK/PD and exposure-response analyses for this trial may be performed if deemed relevant. If conducted, a more technical and detailed elaboration of the population PK, population PK/PD and exposure-response analyses will be given in a prospective modelling analysis plan.

4.6.3 Subgroup Analysis

Not applicable

4.7 Interim Analysis

Not applicable

4.8 Changes to Protocol-planned Analysis

There are no changes to the protocol-planned analysis.

5 Sample size determination

Approximately 108 subjects will be randomly assigned to trial product.

The sample size calculation is based on the primary estimand. It is expected based on phase 2 trial data (NN8640-4172) that the proportion of subjects with no landmark visit data or who discontinued randomised treatment before landmark visit is 10% with similar withdrawal reasons in the two treatment arms. NN8640-4172 is a global trial, including Japanese subjects, and therefore trial information from 4172 is expected to also be of use when planning for a somapacitan phase 3 trial in China. It is expected that subjects discontinuing their randomised treatment will start on ancillary treatment, if no medical reasons prohibit this. Thus, data assessed after discontinuation of the randomised treatment will not be used for the primary analysis of the primary endpoint based on the primary estimand. Assuming the same proportions of subjects with no landmark visit and subjects discontinuing randomised treatment but have landmark visit data in the two arms leads to the following sample size calculation.

The sample size is determined using a non-inferiority margin of -2.0 cm/year and a one sided two-group t-test with a significance level of 2.5% for a 2:1 randomisation ratio between somapacitan and Norditropin®.

Based on data from NN8640-4172, a standard deviation for HV at week 52 was chosen (SD=2.6 cm/year) for the sample size calculation giving a sample size of 108 subjects. Different SD scenarios with power calculation for the primary analysis are presented in the table below under the assumption of a true difference in annualized HV of 0 cm/year between the two treatment arms.

Table 9 Calculated power with 108 subjects randomised 2:1

SD	2.5 cm/year	2.6 cm/year	2.7 cm/year	2.8 cm/year
	95%	**94%**	**92%**	**90%**

The SD candidates are based on reported SD values from clinical trials: Valtropin phase 3 trial[3], 52 weeks (SD=2.8, 3.0), OPKO phase 2 trial, 52 weeks (SD=2.1, 2.3, 2.6, 3.5), and NN8640-4172,

phase 2 trial, 52 weeks (SD=2.3, 2.6).

For sensitivity analysis, a one-way tipping point analysis is planned. Based on a penalty of 2 cm/year for subjects in the somapacitan arm who have missing landmark visit data or discontinue randomised treatment before landmark visit, this gives a power of ~92% (adjusted treatment effect 0.9*0 – 0.1*2 = -0.2).

If the per protocol analysis set (PP) is assumed to consist of ~85% of the trial subjects (same proportion across treatment groups), giving 60 subjects in the somapacitan arm and 30 subjects in the Norditropin® arm, then repeating the primary analysis based on the primary estimand on the PP should result in a power of 92% for confirming non-inferiority of somapacitan compared to Norditropin®, under the assumption of no true treatment difference between the two treatments.

6 Supporting Documentation

6.1 Appendix 1: Definition and calculation of endpoints, assessments and derivations

Type	Title	Time frame	Unit	Details
Primary endpoint	Height velocity	Height velocity (annualised) at week 52	cm / year	HV = (height at 52 weeks visit - height at baseline)/(time from baseline to 52 weeks visit in years). Annualized HV at the intermediate visits: Week 13, 26 and 39, will be derived analogously to the week 52 HV: HV = (height at j weeks visit - height at baseline)/(time from baseline to j weeks visit in years), j=13, 26, 39.
Supportive secondary efficacy endpoint	Change in bone age	From visit 1 to week 52	Years	
Supportive secondary efficacy endpoint	Change in Height Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10	Height SDS will be derived using Chinese general population standards and HV SDS will be derived using Prader standards[2] as reference data
Supportive secondary efficacy endpoint	Change in Height Velocity Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10	
Supportive secondary safety endpoint	Change in fasting plasma glucose	From baseline (week 0) to week 52	mmol/l	

Type	Title	Time frame	Unit	Details
Supportive secondary safety endpoint	Change in HbA1c	From baseline (week 0) to week 52	%	
Supportive secondary endpoint	Change in IGF-I Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10	
Supportive secondary endpoint	Change in IGFBP-3 Standard Deviation Score	From baseline (week 0) to week 52	-10 to +10	
Patient reported outcomes	Change in GHD-CIM (domain and total scores)	From Baseline (Week 0) to visit 5 (Week 26), Baseline (Week 0) to Week 52	-100 to 100	
Patient reported outcomes	GHD – CTB (domain and total scores)	At visit 5 (Week 26) At visit 7(Week 52)	0 to 100	
Patient reported outcomes	GHD – PTB (domain and total scores)	At visit 5 (Week 26) At visit 7(Week 52)	0 to 100	

7 References

1. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials. 2012;9(4):396-407.
2. LI Hui, JI Cheng-ye, ZONG Xin-nan, ZHANG Ya-qin. Height and weight standardized growth charts for Chinese children and adolescents aged 0 to18 years. Chinese Journal of Pediatrics,2009,47(7):487-492
3. Prader. Physical growth of Swiss children from birth to 20 years of age: first Zurich longitudinal study of growth and development. Helv Paediatr Acta Suppl. 52:1-1251989 1989.
4. Peterkova V, Arslanoglu I, Bolshova-Zubkovskaya E, Romer T, Zdravkovic D, Kratzsch J, et al. A randomized, double-blind study to assess the efficacy and safety of valtropin, a biosimilar growth hormone, in children with growth hormone deficiency. Horm Res. 2007;68(6):288-93