Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT01541215
Sponsor trial ID:	NN2211-3659
Official title of study:	Efficacy and safety of liraglutide in combination with metformin versus metformin monotherapy on glycaemic control in children and adolescents with type 2 diabetes A 26-week double-blind, randomised, parallel group, placebo controlled multi-centre trial followed by a 26-week open-label extension
Document date:	25 June 2018

Statistical Analysis Plan

Trial ID: NN2211-3659

Efficacy and Safety of liraglutide in combination with metformin monotherapy on glycaemic control in children and adolescents with type 2 diabetes

A 26-week double-blind, randomised, parallel group, placebo controlled multi-centre trial followed by a 26-week open-label extension

Trial Phase: 3a

Redacted statistical analysis plan
Includes redaction of personal identifiable information only.

Author

[redacted], [redacted]

2129 Biostatistics GLP-1 Diabetes

Table of Contents

Page

List of Abbreviations

ADA	American Diabetes Association
AE	adverse event
ALAT	alanine aminotransferase
ANCOVA	analysis of covariance
ASAT	aspartate aminotransferase
BMI	body mass index
CEA	carcinoembryonic antigen
CTR	clinical trial report
DHEAS	dehydroepiandrosterone sulfate
ECG	electrocardiogram
FAS	full analysis set
FPG	fasting plasma glucose
FSH	follicle stimulating hormone
HbA_{1C}	glycosylated haemoglobin A_{1c}
HDL	high density lipoprotein
HOMA	homeostasis model assessment
ID	identification
IGF-1	insulin-growth factor-1
IGFBP-3	insulin-growth factor-binding protein-3
LDL	low density lipoprotein
LH	luteinizing hormone
LLOQ	lower limit of quantification

LOCF	last observation carried forward
LSMean	least square mean
MAP	modelling analysis plan
MedDRA	Medical Dictionary for Regulatory Activities
N	number of subjects
NTX	N-telopeptide
PG	plasma glucose
PK	pharmacokinetics
SAE	serious adverse event
SAP	statistical analysis plan
SDS	standard deviation score
SMPG	self-measured plasma glucose
SPS	Statistical Programming Specification
TSH	thyroid stimulating hormone
VLDV	very low density lipoprotein

1 Introduction

1.1 Trial information

Trial design

This is a multi-centre, 26-week randomised double-blind, parallel-group, placebo-controlled clinical trial followed by a 26-week open-label extension in subject' aged 10–17 years with type 2 diabetes. After being titrated to 2000 mg of metformin or maximum tolerated dose (MTD) (metformin dose must be ≥1000 mg and ≤2000 mg) subjects will be randomised 1:1 to receive liraglutide (1.8 mg or MTD) or liraglutide placebo. Subjects treated with basal insulin will continue treatment with basal insulin.

Subjects already treated with 2000 mg or more of metformin and with a stable dose for at least 56 days prior to Visit 1 may advance directly to randomisation (Visit 7) when eligibility according to the inclusion and exclusion criteria has been confirmed. Subjects who are treated with basal insulin should in addition to the stable dose of metformin have a stable dose of basal insulin for at least 56 days in order to advance directly to Visit 7.

After 26 weeks of blinded treatment, the treatment allocation will be unblinded. Subjects treated with liraglutide will continue their trial medication until end of treatment. Subjects treated with liraglutide placebo will discontinue their liraglutide placebo treatment. Rescue treatment will be allowed for subjects in both treatment groups experiencing confirmed hyperglycaemia. Subjects on rescue medication will stay in the trial.

Subjects treated with liraglutide for more than 3 months will complete 1 and 2 year follow-up visits.

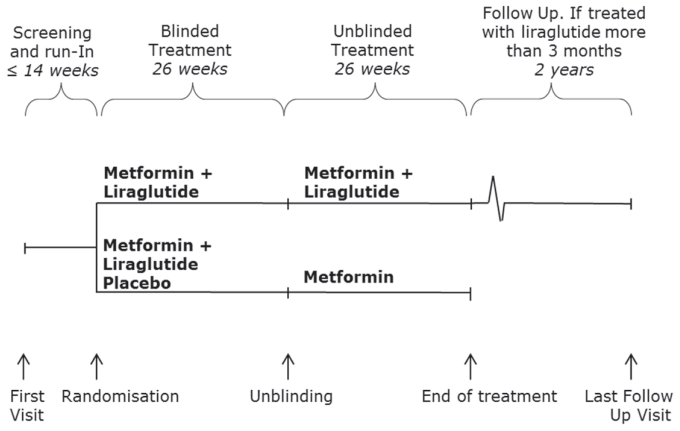

Figure 1 Overview of Trial Design

Primary objective
To confirm the superiority of liraglutide at the maximum tolerated dose (0.6 mg, 1.2 mg or 1.8 mg) versus placebo when added to metformin with or without basal insulin in controlling glycaemia in children and adolescent (ages 10–17 years) with type 2 diabetes.

Secondary objectives
To assess and compare the effect of liraglutide versus placebo in combination with metformin with or without basal insulin treatment on:

- Parameters of glycaemic control
- Parameters of beta-cell function
- Parameters of body composition
- Vital signs
- Growth velocity (if subject is still growing)
- Safety and tolerability
- Growth and pubertal development at 1 and 2 year follow up after trial drug cessation at week 52 (only applicable for subjects on active liraglutide treatment for more than 3 months)

Number of subjects to be investigated

It is planned to randomise 94 subjects in this trial.

Please refer to the protocol for further details.

1.2 Scope of the Statistical Analysis Plan

The planned analyses for trial NN2211-3659 "Efficacy and safety of liraglutide in combination with metformin monotherapy on glycaemic control in children and adolescents with type 2 diabetes. A 26-week double-blind, randomised, parallel group, placebo controlled multi-centre trial followed by a 26-week open-label extension" are described in the trial protocol, version 7.0.

The Statistical Analysis Plan (SAP) version 1 was finalised before the protocol version 7.0. The update to protocol version 7.0 included a major revision of the statistical analysis for the primary endpoint as requested by the FDA.

The description of the analyses in this SAP, version 2.0, is according to that in the protocol version 7.0 and includes additional details and clarifications on the planned statistical analyses. The SAP version 1.0 can be disregarded.

This SAP is finalised prior to database lock of the main period of the trial and release of the randomisation code. The 1- and 2-year follow-up after the 52 weeks of treatment is not part of the database lock of the main period.

2 Endpoints

2.1 Primary endpoint

Change in HbA_{1c} from baseline to week 26*

2.2 Secondary endpoints

2.2.1 Confirmatory secondary endpoints to be assessed at 26 weeks of treatment

- Change from baseline in FPG*
- HbA_{1c} <7.0% (yes/no)*
- Change from baseline in BMI standard deviation score (SDS)*

2.2.2 Supportive secondary efficacy endpoints to be assessed at 26 and 52 weeks of treatment unless otherwise stated

- HbA_{1c} <7.0% (yes/no) at 52 weeks*
- HbA_{1c} ≤6.5% (yes/no)*
- HbA_{1c} <7.0% without severe or minor hypoglycaemic episodes (yes/no)*
- HbA_{1c}<7.5% (yes/no)*
- 7-point self-measured plasma glucose (SMPG)
- Basal insulin dose

Change from baseline at 26 and 52 weeks of treatment unless otherwise stated:

- HbA_{1c} at 52 weeks*
- FPG at 52 weeks*
- Mean 7-point SMPG*
- Post-prandial increments*
- Body weight*
- Waist circumference

- Body mass index (BMI)
- BMI SDS at 52 weeks*
- BMI percentile (age and gender adjusted)
- Systolic and diastolic blood pressure*

Ratio to baseline at 26 and 52 weeks of treatment unless otherwise stated:

- Fasting insulin, fasting pro-insulin, pro-insulin to insulin ratio, fasting glucagon, fasting C-peptide, and homeostasis model assessment (HOMA-B and HOMA-IR)*
- Fasting lipid profile (cholesterol, low density lipoprotein (LDL), very low density lipoprotein (VLDL), high density lipoprotein (HDL), triglycerides and free fatty acids*

2.2.3 Secondary safety endpoints

Change from baseline at 26 and 52 weeks of treatment unless otherwise stated

- Clinical evaluations (physical examination including fundoscopy [fundoscopy at 26 and 52 weeks])
- Electrocardiography (ECG) with rhythm strip (at 26 and 52 weeks)
- Pulse*
- Laboratory tests:
  - Haematology (haemoglobin, haematocrit, thrombocytes, erythrocytes, leucocytes and differential count (eosinophils, neutrophils, basophils, lymphocytes and monocytes)
  - Biochemistry (creatinine, creatine kinase, urea, albumin, bilirubins (total), alanine aminotransferase (ALAT), aspartate aminotransferase (ASAT), alkaline phosphatase, sodium, potassium, calcium, calcium (albumin corrected), amylase and lipase)
  - Hormones (calcitonin, prolactin, follicle stimulating hormone (FSH), estradiol, luteinizing hormone (LH), testosterone, dehydroepiandrosterone sulfate (DHEAS), carcinoembryonic antigen (CEA) and thyroid stimulating hormone (TSH), insulin-like growth factor 1 (IGF-1), insulin–like growth factor binding protein 3 (IGFBP-3)
  - First morning urinalysis (micro albumin, creatinine, albumin:creatinine ratio calculated, protein, ketone, glucose, pH)

  - Biochemical parameters of bone metabolism: Alkaline Phosphatase, N-telopeptide (NTX), C-telopeptide (CTX), serum type 1 procollagen (P1NP)
  - Formation of anti liraglutide antibodies (at 26 and 53 weeks)
- Height SDS*
- Bone age assessment (x-ray of left hand and wrist) at 52 weeks*
- Pubertal assessment/ progression (Tanner staging)*

In addition the following will be assessed at 26 and 52 weeks:

- Assessment of compliance (questioning of subjects and subjects legally acceptable representative)
- Growth (height) velocity in cm/year and height velocity SDS (if subject is still growing). A growth velocity < 1.0 cm/year is defined as no longer growing.*
- Hypoglycaemic episodes*
- AEs and serious adverse events (SAEs)*

Safety follow-up at 1 and 2 years after trial drug cessation at week 52 (only applicable for subjects on active liraglutide treatment for more than 3 months)

- AEs and SAEs*
- Growth (height) velocity in cm/year (if subject is still growing)*
- Height velocity SDS (if subject is still growing)

Change in:

- Height SDS
- Pubertal assessment/progression (Tanner staging)*

- Bone Age assessment (x-ray of left hand and wrist)*

* Key secondary endpoint prospectively selected for disclosure (e.g. clinicaltrials.gov and EudraCT)

3 Statistical considerations

3.1 Evaluability of subjects for analysis

Full analysis set (FAS): includes all randomised subjects receiving at least one dose of liraglutide/liraglutide placebo. The statistical evaluation of the FAS will follow the intention-to-treat principle and subjects will contribute to the evaluation "as randomised".
In exceptional cases subjects or observations from the FAS may be eliminated. The subjects or observations to be excluded, and the reasons for their exclusion must be documented and signed by those responsible before database lock. The subjects and observations excluded from FAS, and the reason for this, will be described in the CTR.

Safety Analysis Set: includes all subjects receiving at least one dose of liraglutide/liraglutide placebo. Subjects in the safety set will contribute to the evaluation "as treated".

Randomised subjects who are lost to follow up and where no exposure information of the investigational product or comparators is available after randomisation will be handled as unexposed.

3.2 Handling of missing data

If an assessment has been made both at screening and randomisation, and if not otherwise specified, the value from the randomisation visit will be used as the baseline value. If the value measured at the randomisation visit is missing and the assessment also has been made at screening, then the screening value will be used as the baseline value.

For efficacy variables missing values will be imputed as described in protocol section 17.3 and 17.4 (section 3.5 and 3.6 in this SAP).

Data collected after treatment discontinuation or initiation of rescue treatment (see protocol section 5.3.5) will be considered missing in some of the efficacy analyses as described in protocol section 17.3 and 17.4 (section 3.5 and 3.6 in this SAP). Data collected after treatment discontinuation will be data from withdrawals collected at either Visit 17 or Visit 25 (see protocol section 8.1.7) in case the measurement is within the visit window (± 5 days) of a planned visit.

For safety variables missing values will not be imputed, except for 'pulse' which will be imputed as described for the secondary efficacy endpoints (section 3.6).
For subjects who received rescue medication, all data collected after initiation of rescue medication will be included.

Laboratory values below the lower limit of quantification (LLOQ) will be set to ½LLOQ.

3.3 General considerations

Novo Nordisk will be responsible for the statistical analysis and reporting. Analysis and reporting will be based on pooled data from all sites.

The statistical analyses will be performed with a significance level of 5% (two-sided tests).

For all endpoints analysed statistically, estimated mean treatment differences will be presented together with two-sided 95% confidence intervals and p-values.

The fixed effects in the models described in the following sections are defined as:

- 'age group' is '10-14' and '>14 years' at end of trial, where age at end of trial is calculated as age at baseline + 1 year.
- 'concomitant diabetes treatment at baseline' is determined by the prescribed insulin dose at baseline. The fixed effect will be 'Y' for all subjects with a prescribed insulin dose above zero at baseline, and 'N' otherwise.
- 'region' is defined as 'Asia', 'Europe', 'North America', 'South America' and 'Rest of the world'. A list of the countries in each region can be found in the NN2211 project statistical programming specification (4).

3.4 Sample size calculation

Please refer to the protocol.

3.5 Primary endpoint

The primary endpoint is change from baseline to week 26 in HbA_{1c}. The FAS will be used for this analysis. All available data will be used, including data collected after treatment discontinuation and rescue initiation. A pattern mixture model using multiple imputation will be used.

The null hypothesis is no difference between the changes from baseline in HbA_{1c} (%) after 26 weeks of randomised treatment with liraglutide + metformin with or without basal insulin and liraglutide placebo + metformin with or without basal insulin (H0: d= liraglutide – placebo =0). The alternative hypothesis is that there is a difference between the two treatment arms (HA: $d \neq 0$). Superiority of liraglutide over liraglutide placebo will be concluded if the 95% confidence interval for the treatment difference for change from baseline in HbA_{1c} (%) after 26 weeks of randomised treatment lies entirely below 0%; implying that the two sided p-value is less than 5%.

The multiple imputation procedure will be as follows:
For subjects in the liraglutide arm who are missing their week 26 HbA_{1c} measurement, measurements will be imputed using the subjects' baseline HbA1c under a regression model based on the completers from the liraglutide placebo arm. Likewise, the same imputation model will be used for subjects in the liraglutide placebo arm with missing week 26 data. Multiple imputation of

missing week 26 HbA_{1c} data will be performed by utilizing the relationship between HbA_{1c} measured at baseline and weeks 10 and 14 with that measured at week 26 in placebo patients. The following four regression models will be built from the liraglutide placebo completers group for this purpose:

- Model 1: Only baseline covariates (baseline HbA_{1c}, stratification group (gender*age group), concomitant diabetes treatment at baseline)
- Model 2: Baseline covariates and week 10 HbA_{1c} as covariates
- Model 3: Baseline covariates and week 14 HbA_{1c} as covariates
- Model 4: Baseline covariates, week 10 HbA_{1c}, and week 14 HbA_{1c} as covariates

Missing week 26 data will be imputed by selecting a random observation from a normal distribution centered at the value predicted by the regression model and with variance analogous to predicting a new observation in regression analysis. For subjects on the placebo arm, the model used will be dependent on the subjects' available HbA_{1c} data throughout the trial. For example, if a subject had HbA_{1c} measurements only at baseline and week 14, then model 3 would be used. For subjects on the liraglutide arm, the measurements will be imputed using only the subjects' baseline HbA_{1c} (model 1).

The imputation procedure will be iterated 10,000 times, thus generating 10,000 complete data sets including observed and imputed values.

For each of the imputed data sets the change in HbA_{1c} from baseline to week 26 will be analysed using an ANCOVA with treatment and stratification groups (gender*age group) as categorical fixed effects and baseline HbA1c as covariate. The results obtained from analysing the datasets will be combined using Rubin's rule (1) to draw inference.

The estimated treatment difference between liraglutide and liraglutide placebo together with two-sided 95% CI and p-value for the test of no difference in effect will be presented.

Sensitivity analyses will be performed for the primary endpoint. They will include:

1. an analysis of covariance (ANCOVA) model with LOCF imputation for missing data. Data collected after treatment discontinuation or initiation of rescue medication will be handled as missing data. Effects in the model are treatment, stratification groups (gender*age group), and baseline HbA_{1c} as a covariate. The LOCF will be performed using data collected at time points in the treatment period (on treatment) and before initiation of rescue medication, or data at baseline if there is no data on treatment without rescue medication post baseline.
2. an ANCOVA model including data after treatment discontinuation or initiation of rescue treatment and LOCF imputation for missing data. Effects in the model are treatment, stratification groups (gender*age group), and baseline HbA_{1c} as a covariate. The LOCF will be performed using baseline data, if this is the only available data, and otherwise data collected at time points during the trial.

3. multiple imputation of missing values in both treatment groups (data collected after treatment discontinuation or initiation of rescue treatment considered as missing) based on parameters estimated from the placebo group. The same model as for the primary analysis will be used.
4. a repeated measurement analysis (MMRM) model using all data points for all visits, but excluding data collected after treatment discontinuation or initiation of rescue medication. Fixed factors in the model are treatment, stratification groups (gender*age group), and baseline HbA_{1c} included as a covariate, all nested within visit. If possible an unstructured covariance matrix will be used.
5. an MMRM model as described above, but including age of the subjects at baseline as covariate, as well as concomitant diabetes treatment at baseline and region. All will be nested within visit.
6. multiple imputation of missing values in both treatment groups based on parameters estimated from the placebo group (the same model as for the primary analysis) based on all randomised subjects

The HbA_{1c} over time (at randomisation, week 10, 14, 26 and 52) will be presented in figures as:
- mean (SE) from the primary analysis multiple imputation, combined using Rubin's rule, at week 26 and 52
- Least square mean (SE) estimated from the MMRM model (no. 4 above) at week 10, 14, 26 and 52. The mean values of the covariate effects will be used for computation of the least square means.

The change in HbA_{1c} over time will in addition be summarised by age group and concomitant diabetes treatment at baseline (see section 3.3 for the definitions of age group and concomitant diabetes treatment at baseline). The summary of change in HbA_{1c} over time will also be shown in plots as mean (SE) both grouped by treatment alone, age group by treatment and subjects having a run-in period by treatment

3.6 Secondary endpoints

3.6.1 Confirmatory secondary endpoints

All confirmatory secondary endpoints will be presented and analysed using FAS, unless otherwise specified.

The primary analysis for the confirmatory secondary endpoints is the pattern mixture model using multiple imputation and using all available data, including data collected after treatment discontinuation and initiation of rescue medication. The model is described in detail in the primary analysis for the primary endpoint, except that the baseline value of the analysed endpoint will be included instead of the baseline HbA_{1c}. value For the dichotomous endpoint (HbA_{1c}<7.0%), a

logistic regression model will be used with missing values imputed from the pattern mixture model of change from baseline in HbA_{1c}

In addition, the confirmatory secondary endpoints will be analysed by an MMRM model using all data points for all visits, but excluding data collected after treatment discontinuation or initiation of rescue medication. Fixed factors in the model are treatment and stratification groups (gender*age group),with baseline value of the analysed endpoint included as covariate., Both fixed factors and the covariate will be nested within visit (sensitivity analysis no. 4 for the primary endpoint). For measurements within a subject an unstructured covariance will be used. For the dichotomous endpoint (HbA_{1c}<7.0%), a logistic regression model will be used with missing values imputed from the corresponding MMRM model of change from baseline in HbA_{1c}.

The continues confirmatory endpoints will be presented as described for the primary analysis.

Change from baseline in FPG after 26 weeks of treatment
The change from baseline in FPG after 26 weeks of treatment will be analysed with the same method as the primary efficacy endpoint (as described above). Baseline FPG will be included in the analyses instead of baseline HbA_{1c}.

HbA_{1c} < 7.0% after 26 weeks of treatment (yes/no)
This dichotomous endpoint will be analysed by a logistic regression model. Effects in the model are treatment, stratification groups (gender*age group), and baseline HbA_{1c} as a covariate. The results will include the 95% confidence interval for the odds ratio (liraglutide over liraglutide placebo) and the p-value for test of no difference between the groups as part of the presentation.
Missing data for HbA_{1c} at week 26 will be imputed from the multiple imputation procedure of the primary endpoint (as described above).

Change from baseline in BMI SDS after 26 weeks of treatment
The change from baseline in BMI SDS after 26 weeks of treatment will be analysed with the same method as the primary efficacy endpoint (as described above). Baseline BMI SDS will be included in the analyses instead of baseline HbA_{1c}.

BMI SDS will be calculated using the following formula:

$$Z = \frac{\left(\frac{measurement}{M}\right)^{L} - 1}{S * L}$$

where L,M and S are median (M), skewness (L) and variation coefficient (S) of children's BMI provided for each sex and age . For each subject a standard deviation score Z (SDS) will be calculated based on age and sex referring to the values L, M and S. The method is described in the WHO Multicentre Growth Reference (3) Values for M, L and S by age and sex is also found in the WHO reference. For Z (SDS) scores below -3 and above 3 the score will be adjusted as described in the WHO instruction.

Hierarchical testing
The primary endpoint and confirmatory secondary endpoints will be analysed in a hierarchical manner in the following order:

1. Primary efficacy endpoint
2. Change from baseline in FPG after 26 weeks of treatment
3. HbA_{1c} < 7.0% after 26 weeks of treatment (yes/no)
4. Change from baseline in BMI SDS after 26 weeks of treatment

In order to be able to conclude significance for an endpoint in the hierarchical list above, the test for that endpoint and the tests for the endpoints higher up in the hierarchy must all be concluded significant with a difference in favour of the liraglutide group.

The hierarchical testing is chosen to maintain the family wise type I error of 5%.

3.6.2 Supportive secondary endpoints

All supportive endpoints will be presented and analysed using FAS.

The supportive secondary endpoints will be analysed using the models as described for the confirmatory secondary analyses (see section 3.6.1).

In the analyses at week 52 the pattern mixture model is extended with the following models:

- Model 5: Baseline covariates and week 26 data of the endpoint as covariates
- Model 6: Baseline covariates and week 10, and week 26 data of the endpoint as covariates
- Model 7: Baseline covariates, week 14 and week 26 data of the endpoint as covariates
- Model 8: Baseline covariates, week 10, week 14 and week 26 data of the endpoint as covariates

Data at baseline, week 10, 14, 26 and 52 will be used in the pattern mixture model.

In the MMRM analysis for the supportive secondary endpoints data points at all visits , but excluding data collected after treatment discontinuation or initiation of rescue medication will be used.

The following dichotomous endpoints are defined:
At 26 and 52 weeks of treatment unless otherwise stated:
HbA_{1c} <7.0% (yes/no) at 52 weeks
HbA_{1c} ≤6.5% (yes/no)
HbA_{1c} <7.0% without severe or minor hypoglycaemic episodes (yes/no)
HbA_{1c} <7.5% (yes/no)

The dichotomous endpoints will be analysed by a logistic regression model. Effects in the model are treatment, stratification groups (gender*age group), and baseline HbA_{1c} as a covariate. The results

will include the 95% confidence interval for the odds ratio (liraglutide over liraglutide placebo) and the p-value for test of no difference between the groups as part of the presentation.

Missing data at week 26 and week 52 will be imputed

- using the multiple imputation procedure as described above
- and in addition using imputations from the MMRM sensitivity analysis no. 4 of the primary endpoint.

The following continuous endpoints are defined:
Change from baseline at 26 and 52 weeks of treatment unless otherwise stated:

- HbA_{1c} at 52 weeks
- FPG at 52 weeks
- 7-point SMPG
  - mean 7-point SMPG, defined as the area under the profile (calculated using the trapezoidal method) divided by the measurement time
  - post-prandial increments after breakfast, lunch and dinner (from before meal to 90 min after breakfast, lunch and dinner)
  - mean post-prandial increment for all meals (breakfast, lunch, and dinner)
- Fasting insulin, fasting pro-insulin, pro-insulin to insulin ratio, fasting glucagon, fasting C-peptide, and homeostasis model assessment (HOMA-B and HOMA-IR). The values will be logarithmically transformed before analysis, and the actual values will be modelled instead of change from baseline.
- Fasting lipid profile (total cholesterol, LDL cholesterol, VLDL cholesterol, HDL cholesterol, triglycerides and free fatty acids. The values will be logarithmically transformed before analysis, and the actual values will be modelled instead of change from baseline.
- Body weight
- Waist circumference
- BMI
- BMI SDS at 52 weeks
- BMI percentile (age and gender adjusted)
- Systolic and diastolic blood pressure

Throughout the trial:

- Basal insulin dose

For calculation of BMI SDS please refer to section 3.6.1 above.
BMI percentile (age and gender adjusted) is the result of the cumulative normal distribution function based on the respective Z-value (SDS value).

The continuous endpoints described above will be analysed with a similar method as for the primary endpoint (detailed above), except for change in basal insulin dose that will only be presented by descriptive statistics. The BMI percentile will also only be summarised (descriptive statistics) as it is uniquely determined from the BMI SDS. The post prandial increments after breakfast, lunch, and dinner will also only be presented by descriptive statistics, whereas the mean post-prandial increment will be analysed as the other endpoints.

If data has been log transformed, the ratio to baseline and treatment ratio will be considered instead of the change from baseline and treatment difference.

3.6.3 Safety endpoints

All safety endpoints will be presented using the safety analysis set. Unless otherwise specified safety endpoints will be presented only by descriptive statistics.

Change from baseline at 26 and 52 weeks of treatment unless otherwise stated

- Clinical evaluations (physical examination including fundoscopy [fundoscopy at 26 and 52 weeks])
- ECG with rhythm strip (at 26 and 52 weeks)
- Laboratory tests:
    - Haematology (haemoglobin, haematocrit, thrombocytes, erythrocytes, leucocytes and differential count (eosinophils, neutrophils, basophils, lymphocytes and monocytes) Biochemistry (creatinine, creatine kinase, urea, albumin, bilirubin (total), ALAT, ASAT, sodium, potassium, alkaline phosphatase, calcium, calcium (albumin corrected), amylase and lipase)
    - Hormones (calcitonin, prolactin, FSH, estradiol, LH, testosterone, DHEAS, CEA and TSH, IGF-1, IGFBP-3
    - First morning urinalysis (Micro albumin, creatinine, albumin:creatinine ratio calculated, protein, ketone, glucose, pH)
    - Biochemical parameters of bone metabolism: Alkaline Phosphatase, NTX, CPX, P1NP
    - Formation of anti-liraglutide antibodies (at 26 and 53 weeks)
- Pulse
- Pubertal assessment/ progression (Tanner staging), DHEAS, LH, FSH, estradiol in females and testosterone in males
- Height SDS versus baseline at 26 and 52 weeks
- Bone age assessment (x-ray of left hand and wrist) at 52 weeks

In addition the following will be assessed at 26, and 52 weeks:

- Assessment of compliance (questioning of subjects and subjects legally acceptable representative)

- Growth (height) velocity in cm/year and height (growth) velocity SDS (if subject is still growing). A growth (height) velocity < 1.0 cm/year is defined as no longer growing.

Collected throughout the trial:

- Hypoglycaemic episodes (see details below)
- AEs and serious adverse events (SAEs) (see details below)

Pulse
Pulse will be analysed with a similar method as for the primary analysis for the primary endpoint (as described for the secondary confirmatory and supportive efficacy endpoints).

Height SDS
Height SDS will be calculated using the same formula as described for BMI SDS (see section 3.6.1) The values for M, S and L are downloaded from http://www.who.int/growthref/hfa_boys_perc_WHO2007_exp.txt and http://www.who.int/growthref/hfa_girls_perc_WHO2007_exp.txt on 2nd February, 2018)

Height velocity
Growth (height) velocity is the change in height per year and is measured in cm/year (height velocity is the same as growth velocity).
The height velocity is calculated as the difference between current height and baseline divided by time duration in days between those measurement time points and multiplied by 365 days.

Height velocity SDS
Height velocity SDS will be calculated using the same formula as described for BMI SDS (see section 3.6.1) The reference values for M, S and L are downloaded from https://academic.oup.com/jcem/article/99/6/2104/2537746. Height velocity SDS at week 26 will only be used to find extreme outliers as it is not recommended to calculate this value based on height change over less than one year.

Hypoglycaemic episodes
Hypoglycaemic episodes are recorded by subjects in their trial diaries throughout the trial. The information collected includes plasma glucose (PG) before treating the episode and whether the subject was able to treat him/herself. This information is used by Novo Nordisk to classify an episode according to the ADA definition (severe, documented symptomatic, asymptomatic, probable symptomatic and relative) and the additional minor category as defined by Novo Nordisk (also named as BG confirmed). A hypoglycaemic episode is defined as treatment emergent if the onset of the episode is on or after the first day of exposure to randomised treatment and no later than one day after the last day on randomised treatment. All hypoglycaemic episodes will be summarised by treatment and severity.

Nocturnal hypoglycaemic episodes, defined as episodes with time of onset between 00:01 and 05:59, will also be summarised by severity. In addition the hypoglycaemic episodes will be

summarised by events occurring at periods in the trial when the subjects are on basal insulin and/or bolus insulin, and at periods without insulin.

The hypoglycaemic episodes will be summarised for the blinded period of the trial (up to week 26) and for the total exposure period.

Adverse events
All AEs will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA). A treatment emergent adverse event (TEAE) is defined as an event that has onset date (or increase in severity) on or after the first day of exposure to randomised treatment and no later than seven days after the last day of randomised treatment. In the open period of the trial the randomised treatment for placebo is metformin alone.

The TEAEs will be summarised by treatment for all TEAEs by system organ class, preferred term, seriousness, severity, outcome and relation to trial product. TEAEs are summarised descriptively in terms of the number of subjects with at least one event (N), the percentage of subjects with at least one event (%), the number of events (E) and the event rate per 1000 years (R).

The most frequent TEAEs will be defined as preferred terms that are experienced by at least 5% of the subjects in any of the treatment arms. The most frequent TEAEs will be summarised and presented in a plot with number of subjects with an event and rate of events (number per 1000 years of exposure).

The TEAEs will be summarised for the blinded period of the trial (up to week 26) and for the total exposure period.

Non-treatment emergent AEs will be presented in listings.

Medical events of special interest (MESI) will be summarised separately.

MedDRA search terms have been defined to find AEs of special interest in the following categories: Cardiac events, Pancreatitis, Medication error, Neoplasm, Thyroid disease, Altered renal function, Injection site reactions, Overdose, Gallstone, Cardiac arrhythmia, Hyperglycaemia, Allergic reaction, Elevated amylase, Elevated calcitonin, Suspected transmission of infectious agents, Rare adverse events, Depression adverse events, Device issue events and Immune adverse events. These MedDRA search terms are defined prior to DBL (5). The AEs of special interest will be listed by category.

3.6.4 Safety follow-up at 1 and 2 years after trial drug cessation at week 52 (only applicable for subjects on active liraglutide treatment for more than 3 months)

- AEs and SAEs
- Growth (height) velocity in cm/year (if subject is still growing)

- Height velocity SDS (height velocity is the same as growth velocity) (if subject is still growing)

Change in:

- Height SDS
- Pubertal assessment/progression by Tanner staging
- Bone age assessment (x-ray of left hand and wrist)

For calculation of height SDS please refer to section 3.6.2.

3.7 Exploratory analyses

To evaluate the influence of intake of insulin on the treatment effect on change in HbA_{1c} at week 26 and 52 a repeated measurement analysis (MMRM) model including treatment interaction with insulin at baseline will be performed. All data points for all visits, but excluding data collected after treatment discontinuation or initiation of rescue medication, will be used. Fixed effects in the model are treatment, stratification groups (gender*age group), diabetes medication at baseline and treatment by diabetes medication at baseline interaction. The baseline HbA_{1c} will be included as a covariate. All factors and covariates will be nested within visit. If possible, an unstructured covariance matrix will be used.

3.8 Interim analysis

NA

3.9 Sequential safety analysis/safety monitoring

An external data monitoring committee (DMC) was established to independently review and evaluate accumulated unblinded safety data from the trial in order to protect the safety of the subjects and to evaluate the evolving risk-benefit if required. See protocol for details.

Novo Nordisk constituted an internal safety committee to perform ongoing safety surveillance. The safety committee conducted ongoing monitoring of blinded safety data. The safety committee could recommend unblinding of any data for further analysis, and in this case an independent ad hoc group would be established in order to maintain the blinding of the trial personnel.

3.10 Explorative statistical analysis for pharmacogenetics and biomarkers

NA

3.11 PK data

The liraglutide concentrations will be listed.

4 Changes to the statistical analyses planned in the protocol

This statistical analysis plan follows the protocol with the following clarifications:

4.1 Sensitivity analysis for the primary endpoint on all randomised subjects

The primary analysis is based on FAS which includes all randomised subjects receiving at least one dose of liraglutide/liraglutide placebo. Due to an FDA request to change in the primary analysis, the primary estimand is the "treatment policy" estimand, and therefore it is appropriate to add a sensitivity analysis of the primary endpoint including all randomised subjects.

4.2 Handling of missing data

For safety variables it was stated in the protocol that missing values (including intermittent missing values) would be imputed using the last observation carried forward (LOCF) on post-baseline measurements. The analyses on safety parameters will be performed using a pattern mixture model with multiple imputation and missing values will therefore not be imputed using the LOCF method.

4.3 Definition of fixed effects

The fixed effect 'region' was not specified in the protocol, but in a sensitivity analysis the regions were described as: North America, Asia including Russia, Europe including Israel, South America including Mexico. To align with other NN2211 trials, the regions are defined as has been done overall for the NN2211 project. The only difference from the description in the protocol is that Russia is not assigned to region Asia, but is assigned to region Europe.

The fixed effect 'concomitant diabetes treatment at baseline' was described as 'diet and exercise alone vs. diet and exercise plus metformin and/or basal insulin'. As all subjects were to take metformin at baseline the definition is updated to with/without basal insulin at baseline (see section 3.3)

4.4 Supportive secondary endpoints

The mean post prandial increment is added as an endpoint, as this endpoint has been presented in other trials in project NN2211. The mean post-prandial increment will be analysed as the other supportive secondary endpoints. The post prandial increments after breakfast, lunch and dinner will only be presented by descriptive statistics, whereas in the protocol they were planned to be analysed.
The BMI percentile will also only be summarised, as it is closely related to the BMI SDS.

4.5 Statistical analyses of the confirmatory and supportive secondary endpoints

In the protocol it is stated the confirmatory and supportive secondary endpoints 'will be analysed with the same method as the primary efficacy endpoint'. It has been clarified that the endpoints analysed with the same method as the primary efficacy endpoint will be analysed using the primary analysis for the primary endpoint (the pattern mixture model) and, in addition, using the MMRM sensitivity analysis described as no. 4 for the primary analysis (see section 3.5). These analyses have been described in detail in sections 3.6.1 and 3.6.2.

4.6 Exploratory analyses

To evaluate the influence of intake of insulin on the treatment effect on change in HbA1c at week 26 and 52 an exploratory analysis including treatment and diabetes medication at baseline interaction has been added.

4.7 Miscellaneous

Some wordings have been updated without changing the content, these are not mentioned here. In addition, details have been added with regards to how data will be summarised and presented which are also not mentioned here.

5 References

1. Rubin, D. B. (1987), Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons
2. Ratitch, B., O'Kelly (2011), M. Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures, PharmaSUG2011 - Paper SP04
3. WHO Multicentre Growth Reference Study Group. 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 (312 pages) (http://www.who.int/childgrowth/standards/technical_report/en/)
4. nn2211-phase3b-4-project-statistical-programming-specification, Version 1.0 (Novo Docs ID 004154390)
5. SAS dataset with MedDRA search terms for AEs of special interest (3659_meddra, Novo Docs ID 004247169)