# Cover Page for Statistical Analysis Plan

| | |
|---|---|
| Sponsor name: | Novo Nordisk A/S |
| NCT number | NCT03191396 |
| Sponsor trial ID: | NN9535-4339 |
| Official title of study: | Efficacy and safety of semaglutide 1.0 mg once-weekly versus liraglutide 1.2 mg once-daily as add-on to 1–3 oral anti-diabetic drugs (OADs) in subjects with type 2 diabetes. (SUSTAIN 10) |
| Document date: | 27 May 2019 |

### 16.1.9 Documentation of statistical methods

# List of contents



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

.

**Statistical Analysis Plan**

# Trial ID: NN9535-4339

# SUSTAIN 10
# Efficacy and safety of semaglutide 1.0 mg once-weekly versus liraglutide 1.2 mg once-daily as add-on to 1-3 oral anti-diabetic drugs (OADs) in subjects with type 2 diabetes

**Trial phase:** 3b

**Author:**
[redacted], Statistician
Biostatistics Semaglutide s.c.

# Table of contents Page

# List of abbreviations

| | |
|---|---|
| AE | adverse event |
| ANCOVA | analysis of covariance |
| BMI | body mass index |
| FAS | full analysis set |
| MedDRA | Medical Dictionary for Regulatory Activities |
| PP | per protocol |
| SAS | safety analysis set |
| SAP | statistical analysis plan |

# 1 Introduction

## 1.1 Trial information

This is a 30-week, confirmatory, randomised, multicentre, multinational, active-controlled, parallel groups, open label, 2-armed trial.

### The primary objective

To compare the effect of semaglutide s.c. 1.0 mg once-weekly versus liraglutide s.c. 1.2 mg once-daily on glycaemic control after 30 weeks of treatment in subjects with type 2 diabetes.

### The secondary objectives

To compare the effect of semaglutide s.c. 1.0 mg once-weekly versus liraglutide s.c. 1.2 mg once-daily after 30 weeks of treatment on body weight, efficacy parameters, safety and tolerability in subjects with type 2 diabetes.

### Trial design

Subjects with T2D inadequately controlled on 1-3 oral anti-diabetic drug(s) (OAD) will be randomised in a 1:1 manner to receive either semaglutide s.c. 1.0 mg once-weekly or liraglutide s.c. 1.2 mg once-daily. The randomisation will be stratified based on subjects background medication of SU and sodium-glucose cotransporter-2 (SGLT-2) inhibitors:

- SU (+/- metformin)
- SGLT-2 inhibitors (+/- metformin)
- SU + SGLT-2 inhibitors (+/- metformin)
- No SU and no SGLT-2 inhibitors (metformin monotherapy)

See protocol for trial NN9535-4339 for further details.

## 1.2 Scope of the statistical analysis plan

This SAP is based on the protocol "SUSTAIN 10: Efficacy and safety of semaglutide 1.0 mg once-weekly versus liraglutide 1.2 mg once-daily as add-on to 1-3 oral anti-diabetic drugs (OADs) in subjects with type 2 diabetes", version 2.0.

# 2 Statistical considerations

## 2.1 General considerations

No interim analyses or other analyses of un-blinded data will be performed before the database is locked.

If necessary, a statistical analysis plan may be written in addition to the protocol, including a more technical and detailed elaboration of the statistical analyses. The statistical analysis plan will be finalised before database lock.

Results from a statistical analysis will be presented by the estimated treatment contrasts at week 30 with associated two-sided 95% confidence intervals and p-values corresponding to two-sided tests of no difference if not otherwise specified.

The comparison presented from a statistical analysis will be semaglutide 1.0 mg versus liraglutide 1.2 mg.

If no statistical analysis is specified, data will be presented using relevant summary statistics.

Data from all trial sites will be analysed and reported together.

### 2.1.1 Primary estimand

To further detail the trial objective an estimand is defined which is a *de-jure* (efficacy) estimand:

- The treatment difference between semaglutide and liraglutide at week 30 for all randomised subjects if all subjects completed treatment and did not initiate rescue medication

This primary *de-jure* estimand is considered clinically relevant as it assesses the glycaemic benefit a person with T2D is expected to achieve if initiating and continuing treatment with semaglutide compared to liraglutide. Accordingly, only data collected prior to discontinuation of trial product or initiation of rescue medication will be used to draw inference. This will avoid confounding from rescue medication.

### 2.1.2 Missing data considerations at week 30

The overall rate of missing data at week 30 is expected to be not more than 10% based on the rate of subjects with measurements of the primary and the confirmatory secondary endpoints in the s.c. semaglutide phase 3a clinical development programme. The frequency and reasons of missing data is expected to be similar in the semaglutide and the liraglutide groups.

When estimating the primary estimand, the combined rate of missing data and, subjects discontinuing treatment prematurely or initiating rescue medication on top of trial product, is

expected to be maximum 20%. This is based on the results from the s.c. semaglutide phase 3a clinical development programme. Based on these data, premature treatment discontinuation due to gastrointestinal adverse events is expected to be low. Other possible reasons for premature discontinuing treatment are assumed to be unrelated to treatment and to therefore occur at similar rates between the two treatment groups. Thus, overall, the frequency of missing data or data not used at week 30 in the primary analysis is expected to be similar between the treatment groups.

To document the extent and reason(s) for missing data, descriptive summaries and graphical representation of extent, reason(s) for and pattern of missing data for the primary and secondary endpoints will be presented by treatment group.

### 2.1.3 Confirmatory hypotheses

The mean treatment difference is defined as $\mu$ = (semaglutide minus liraglutide). Three confirmatory hypotheses are to be tested:

1. $HbA_{1c}$ non-inferiority of semaglutide 1.0 mg vs. liraglutide 1.2 mg with a non-inferiority margin of 0.3
    - $H_0$: $\mu \geq 0.3$%-point against Ha: $\mu < 0.3$%-point
2. Body weight superiority of semaglutide 1.0 mg vs. liraglutide 1.2 mg
    - $H_0$: $\mu \geq 0.0$ kg against Ha: $\mu < 0.0$ kg
3. $HbA_{1c}$ superiority of semaglutide 1.0 mg vs.liraglutide 1.2 mg
    - $H_0$: $\mu \geq 0.0$%-point against Ha: $\mu < 0.0$%-point

The non-inferiority margin of 0.3 is chosen based on the diabetes guideline[1, 2]. The effect of liraglutide was investigated in various trials including the LEAD programme that included a series of six randomised controlled phase 3 trials. In these trials, a substantial and sustained reduction in HbA1c was obtained with liraglutide treatment across the continuum of care in patients with type 2 diabetes[3]. Based on the Lead 1 & Lead 4 studies, Liraglutide 1.2 mg showed HbA1c treatment difference to placebo of -1.3%[4] and -0.9%[5] respectively. Hence, based on these trials, and considering the LEAD programme results collectively that assures the effect of liraglutide, the chosen non-inferiority margin of 0.3 provides assurance that semaglutide has an effect greater than 0 with a clinically relevant size. With regards to the constancy assumption, controlled clinical trials have consistently established that liraglutide is an effective anti-diabetic drug. Therefore, lack of trial sensitivity with liraglutide as comparator is not anticipated to be an issue in this trial.

### 2.1.4 Multiplicity and criteria for confirming hypotheses

The Type-I error rate for testing the three confirmatory hypotheses related to the HbA1c and body weight endpoints will be preserved in the strong sense at an overall alpha (α) level of 2.5% (one-sided), corresponding to a two-sided 5% level hierarchical test, using the closed testing procedure described in Bretz et al. The overall α-level of 2.5% is initially allocated to the HbA1c non-inferiority test. For this hypothesis, and in general, if a hypothesis is confirmed the local α-level (α-local) will be reallocated according to the weight and the direction of the edges going from the confirmed hypothesis to the next hypotheses as specified in Figure 2–1.

Each of the following hypotheses will be tested at their local α level. This process will be repeated until no further hypotheses can be confirmed.

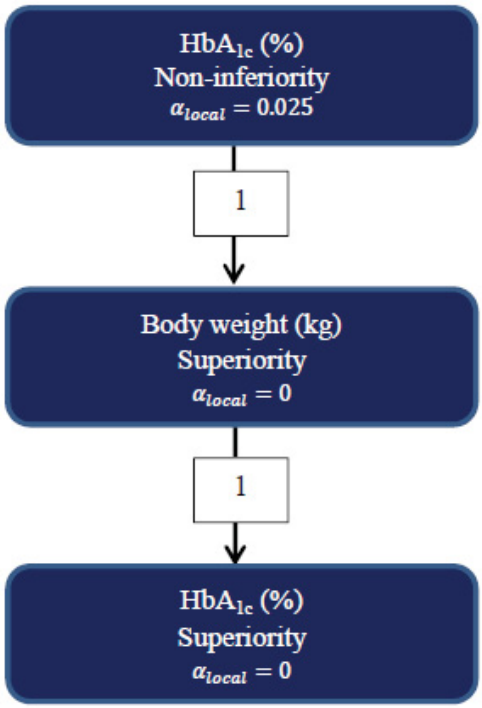



**Figure 2–1 Graphical illustration of the closed testing procedure**

## 2.2 Sample size calculation

The sample size calculation will assume a power of 90% for confirming all three confirmatory hypotheses (see Section 2.1.3) across efficacy and in-trial assumptions. Please see Table 2–2 for calculated powers for efficacy as estimated by the primary analysis for the primary estimand (see Section 2.4.1) and for the in-trial effect as estimated by the treatment policy sensitivity analysis (see Section 2.4.4)

The sample size assumptions for efficacy based on 'on-treatment without rescue medication' data and a treatment effect based on in-trial data (see Section 2.4.2) together with the standard are given in Table 2–1.These are based on the efficacy results and an observed reduction of up to 15% in the treatment effect based on in-trial data compared to efficacy based on 'on-treatment without rescue medication' data in the s.c. semaglutide phase 3a clinical development programme. A similar reduction in the in-trial treatment effect compared to efficacy is assumed with liraglutide as comparator.

**Table 2–1 Assumptions used in the sample size calculation**

| **Semaglutide vs. liraglutide** | **$HbA_{1c}$** | **Body weight** |
|---|---|---|
| Efficacy (treatment difference) | -0.35% | -3.3 kg |
| In-trial treatment effect (treatment difference) | -0.30 % | -2.38 kg |
| Standard deviation | 1.1% | 4.0 kg |

Tests statistics for treatment differences are assumed to follow normal distributions. The sample size is calculated using the calcPower function in the R package, gMCP[6], using 10,000 simulations. All of the three pre-specified confirmatory tests are assumed to be independent. Since some of these tests are positively correlated, the assumption of independence is viewed as conservative.

With the above assumptions, allocating 288 subjects to each of the semaglutide and liraglutide groups (576 subjects in total) provides at least 90% power to reject all three confirmatory hypotheses and thus confirm $HbA_{1c}$ superiority and body weight superiority of semaglutide vs. liraglutide across efficacy and in-trial assumptions. Please see Table 2–2.

**Table 2–2 Calculated powers for meeting individual hypotheses**

| **Statistical test** | **$HbA_{1c}$ non-inferiority** | **$HbA_{1c}$ superiority** | **Body weight superiority** | **All** |
|---|---|---|---|---|
| Efficacy power (%) | >99% | 97% | >99% | 97% |
| In-trial effect power (%) | >99% | 90% | >99% | 90% |

## 2.3 Data definitions

### 2.3.1 Data selection

Subjects and data to be used in an analysis will be selected in a two-step manner:

- Firstly, subjects will be selected based on the specified analysis set

- Secondly, data points on the selected subjects from first step will be selected based on the specified observation period

#### 2.3.1.1 Definition of analysis sets

The following analysis sets will be defined:

**Full analysis set (FAS)**: includes all randomised subjects. Subjects in the FAS will contribute to evaluation "as randomised".

**Safety analysis set (SAS)**: includes all subjects exposed to at least one dose of trial product. Subjects in the SAS will contribute to the evaluation based on the trial product received for the majority of the period they were on treatment. This will be referred to as contributing to the evaluation "as treated".

**Per protocol (PP) analysis set:** includes all subjects in the FAS who fulfil the following criteria:

- Have not violated any inclusion criteria
- Have not fulfilled any exclusion criteria
- Have a non-missing $HbA_{1c}$ measurement at screening and/or randomisation
- Have a treatment duration of at least 16 weeks
- Have at least one non-missing $HbA_{1c}$ measurement at or after week 16

Subjects in the PP analysis set will contribute to the analysis "as treated" as defined for the SAS.

#### 2.3.1.2 Definition of observation periods

Definition of the observation periods:

**In-trial:** This observation period represents the time period after randomisation where subjects are considered to be in the trial, regardless of discontinuation of trial product or initiation of rescue medication. The in-trial observation period starts at randomisation (as registered in IWRS) and ends at the date of:

- The last direct subject-site contact, which is scheduled to take place 5 weeks after planned last dose of trial product at a follow-up visit (phone visit)
- Withdrawal for subjects who withdraw their informed consent
- The last subject-investigator contact as defined by the investigator for subjects who are lost to follow-up
- Death for subjects who dies before any of the above

For subjects not randomised but exposed to trial product the in-trial period starts at the date of first dose of trial product.

**On-treatment:** This observation period represents the time period where subjects are considered exposed to trial product. The observation period is a sub-set of the in-trial observation period. It starts at the date of first dose of trial product. Two slightly different end dates will be needed to cover all assessments appropriately according to the flow chart. For AEs including hypoglycaemic episodes, the observation period ends at the first date of any of the following:

- The follow-up visit (P10)
- The premature discontinuation follow-up visit (P10A)
- The last date on trial product + 42 days
- The end-date for the in-trial observation period

The follow-up visit is scheduled to take place 5 weeks after the last date on trial product corresponding to approximately five half-lives of s.c. semaglutide. The visit window for the follow-up visit is + 7 days, which is the reason for the 42 days specified in the bullet above. Hence, for AEs including hypoglycaemic episodes, this period reflects the period in which subjects are exposed.

For efficacy and other safety assessments (laboratory assessments, physical examination and vital signs) the observation period ends at the last date on trial product with a visit window of + 7 days. This ascertainment window corresponds to the dosing interval and will be used to avoid attenuation of a potential treatment effect on endpoints for which the effect is reversible shortly after treatment discontinuation. Hence, for those assessments this period reflects the period in which subjects are treated.

**On-treatment without rescue medication:** This observation period is a sub-set of the on-treatment observation period, where subjects are considered treated with trial product, but have not initiated rescue medication. Specifically, the period starts at the date of first dose of trial product and ends at the first date of any of the following:

- The last dose of trial product with a visit window of +7 days
- Initiation of rescue medication

The 'on-treatment without rescue medication' observation period will be the primary observation period for efficacy evaluations. The in-trial observation period will be considered supportive for efficacy evaluation. Safety will be evaluated based on the in-trial and the on-treatment observation periods unless otherwise specified.

Data points collected outside an observation period will be treated as missing in the analysis. Baseline data will always be included in an observation period. Before data are locked for statistical analysis, a review of all data will take place. Any decision to exclude either a subject or single

observations from the statistical analysis is the joint responsibility of the members of the Novo Nordisk study group.

Exclusion of data from analyses will be used restrictively and normally no data should be excluded from the FAS. The subjects or observations to be excluded, and the reasons for their exclusion will be documented and signed by those responsible before database lock. The subjects and observations excluded from analysis sets, and the reason for this, will be described in the clinical trial report.

### 2.3.2 Data transformations

A number of the continuous parameters will be log-transformed prior to statistical analysis. The output tables and figures will show the results of the analysis back-transformed to the original scale, implying that log-treatment-differences are reported as treatment ratios. Confidence intervals for the treatment ratios will be calculated as exponentiated upper and lower limits for log-treatment difference confidence intervals. The standard errors of the back-transformed mean and ratio to baseline estimates are also provided; these SEs are calculated using the delta-method (first order Taylor approximation), whereby the SE on the original scale is calculated as the product of the SE on log-scale and the exponentiated estimate of the mean (geometric mean).

Moreover, a number of binary endpoints will be analysed using a logistic regression. From this analysis, the estimated odds will be presented, calculated as the exponentiated estimate of the log-odds. The standard errors for the estimated odds are calculated as the standard error of the estimated log-odds times the estimated odds by using the delta method.

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

### 2.3.3 Definition of baseline

For each assessment, the baseline assessment is defined as the latest available measurement at or prior to the randomisation visit (V2). This specifically implies that if a visit 2 assessment is missing (whether it was planned or not planned) then the screening assessment (from visit 1), if available, will be used as the baseline assessment.

### 2.3.4 Trial completion

Unless subjects withdraw their informed consent, data collection will continue for the full duration of the trial. The full duration of the trial is defined as up to and including the follow-up visit (P10). Subjects completing the follow-up visit (P10) will be considered trial completers.

### 2.3.5 Treatment completion

Treatment period completion is defined as when the subject has received the required treatment, and attended the ‘End of Treatment’ (V9).

## 2.4 Primary endpoint

The primary endpoint is change from baseline to week 30 in $HbA_{1c}$.

### 2.4.1 Primary analysis

The primary estimand will be estimated based on the FAS using post-baseline measurements up to and including week 30 from the 'on-treatment without rescue medication' observation period. Imputation of missing data will be handled using multiple imputation assuming that missing data is missing at random (MAR). Missing data will be imputed using observed data within the same group defined by the randomised treatment (semaglutide/liraglutide). It is hereby assumed that the likely values of what the missing data would have been if available are best described by information from subjects who receive the same treatment.

Technically, missing values will be imputed as follows:

- Intermittent missing values are imputed using a Markov Chain Monte Carlo (MCMC) method, in order to obtain a monotone missing data pattern. This imputation is done for each of the treatment groups separately and 500 copies of the dataset will be generated
- A sequential regression approach for imputing monotone missing values at planned visits will be implemented starting with the first visit after baseline and sequentially continuing to the last planned visit at week 30. A model used to impute missing values at each planned visit will be fitted for each of the treatment groups using observed and imputed data. The model will include the SU and SGLT-2 inhibitors anti-diabetic background medication stratification factor ( SU +/- metformin, SGLT-2 inhibitors +/- metformin, SU + SGLT-2 inhibitors +/- metformin, no SU and no SGLT-2 inhibitors (metformin monotherapy)) as categorical effects and the baseline and post-baseline $HbA_{1c}$ values observed prior to the visit in question as covariates.
- An analysis of covariance (ANCOVA) with treatment and the SU and SGLT-2 inhibitors anti-diabetic background medication stratification factor (SU +/- metformin, SGLT-2 inhibitors +/- metformin, SU + SGLT-2 inhibitors +/- metformin, no SU and no SGLT-2 inhibitors (metformin monotherapy)) as categorical effects and baseline $HbA_{1c}$ as a covariate will be used to analyse $HbA_{1c}$ values at week 30 for each of the 500 complete data sets generated as part of the imputation of missing values. Rubin's rule will be used to combine the analysis results in order to draw inference.

From this analysis, the estimated treatment difference between semaglutide and liraglutide at week 30 will be presented together with the associated two-sided 95% confidence interval and unadjusted two sided p-values.

The one-sided non-inferiority and superiority hypotheses will be confirmed using the overall significance level of 2.5% in line with the closed testing procedure in Figure 2–1. Operationally, non-inferiority and subsequent superiority will be considered confirmed if the mean treatment

difference is supporting the corresponding alternative hypothesis and the two-sided p-value (non-inferiority H0: u=0.3 vs. Ha: u≠0.3, superiority H0: u=0 vs Ha: u≠0 respectively), from the primary analysis of the primary estimand, is strictly below two times its local significance level resulting from the closed testing procedure.

### 2.4.2 Sensitivity analyses

In order to investigate the robustness of the conclusions from the primary analysis and to stress test the MAR assumption for missing data tipping point sensitivity analyses will be performed for the primary estimand both for the sensitivity of the non-inferiority and the superiority $HbA_{1c}$ hypotheses. No tipping point analysis will be performed on the non-inferiority test if the superiority test is confirmed.

### 2.4.3 Sensitivity analyses for the primary estimand

The estimation of the primary estimand will be repeated using the following sensitivity analysis:

- Tipping-point analysis (pattern mixture model based) based on the FAS using the 'on-treatment without rescue medication' observation period. In this analysis, subjects from the semaglutide group with missing observations will be given a penalty, i.e., it is assumed that subjects with missing observations who are randomised to semaglutide will receive a treatment that is less beneficial than subjects with observed values who are randomised to semaglutide. The idea is to gradually increase the penalty to evaluate at which level the superiority conclusion of the analyses in terms of statistical significance is changed. The tipping point is the penalty level, at which the magnitude of efficacy reduction in subjects with missing data creates a shift in the treatment effect of semaglutide from being statistically significantly better than liraglutide to being non-statistically significantly better for the superiority test and similarly for the non-inferiority test. Technically, this analysis will be implemented by replicating the primary analysis including the assumption of MAR but subsequently adding increasing penalty values at week 30 to imputed observations in the semaglutide group before applying ANCOVA on the 500 complete data sets.

### 2.4.4 Other sensitivity analyses

The following additional sensitivity analyses are specified:

- Retrieved dropout analysis based on the FAS using post-baseline measurements up to and including week 30 from the in-trial observation period. Missing data will be imputed using the same approach as described for the primary analysis of the primary estimand. However the imputation will be done within the same group defined not only by the randomised treatment (semaglutide/ liraglutide) but also by the status of treatment completion (still on randomised treatment at week 30 yes/no) (4 groups in total). It is hereby assumed that the likely values of what the missing data would have been if available are best described by information from

subjects who at week 30 are similar in terms of randomised treatment and treatment completion status. In addition in the imputation step stratification factor is not included in the model in order to avoid potential issues with sparse data. The ANCOVA performed in the analysis step will include the same categorical effects and covariates as the primary analysis. This analysis could be considered addressing an effectiveness estimand. This analysis will be carried out for the superiority test only.

- Per protocol (PP) analysis based on the PP data set using the 'on-treatment without rescue medication' observation period. This analysis will be carried out for non-inferiority testing only. The statistical analysis will be the same as the primary analysis for the primary estimand.

## 2.5 Secondary endpoints

### 2.5.1 Confirmatory secondary endpoints

The confirmatory secondary endpoint is change from baseline to week 30 in body weight (kg).

The primary estimand will be estimated for body weight using the same approach as described in Section 2.4.1. Body weight will be tested for superiority. Baseline and post-baseline body weight will be used as covariates instead of HbA1c.

The one-side hypothesis of superiority will be considered confirmed if the mean treatment difference is supporting the corresponding hypothesis and the two-sided p-value from the analysis of body weight is strictly below two times its local two-sided significance level resulting from the closed testing procedure in Figure 2–1.

The tipping point sensitivity analysis pre-specified to evaluate the robustness of the conclusions from the primary analysis of $HbA_{1c}$ will also be performed to evaluate the robustness of the conclusion from the body weight superiority test. In addition, the retrieved dropout sensitivity analysis will also be performed for body weight.

### 2.5.2 Supportive secondary endpoints

No sensitivity analyses are planned for the supportive secondary endpoints.

#### 2.5.2.1 Efficacy endpoints

**Continuous endpoints**

The continuous endpoints are change from baseline to week 30 in:

- FPG
- SMPG, 7 point profile
  - Mean 7-point profile
  - Mean post prandial increment (over all meals)

- Fasting blood lipids (total cholesterol, LDL-cholesterol, HDL-cholesterol, triglycerides)
- BMI
- Waist circumference
- Systolic and diastolic blood pressure
- Body weight (%)

The above continuous endpoints will be analysed separately using a similar model approach as for the primary endpoint with the associated baseline value as covariates instead of $HbA_{1c}$ for their respective analyses.

Fasting lipid profile endpoints will be log-transformed prior to analysis with the associated log-transformed baseline value as a covariate.

**Mean 7-point profile (SMPG) definition**

Subjects will be asked to perform SMPG measurements before and 90 minutes after breakfast, lunch, dinner, and at bedtime.

Mean of the 7-point profile is defined as the area under the profile, calculated using the trapezoidal method, and divided by the measurement time.

**Binary endpoints**

The binary endpoints are subjects who after 30 weeks treatment achieve (yes/no):

- $HbA_{1c} < 7.0\%$ (53 mmol/mol), ADA target
- $HbA_{1c} \leq 6.5\%$ (48 mmol/mol), AACE target
- Weight loss $\geq 3\%$
- Weight loss $\geq 5\%$
- Weight loss $\geq 10\%$
- $HbA_{1c}$ <7.0% (53 mmol/mol) without severe or blood glucose confirmed symptomatic hypoglycaemia episodes and no weight gain
- $HbA_{1c}$ reduction $\geq 1\%$
- $HbA_{1c}$ reduction $\geq 1\%$ and weight loss $\geq 3\%$
- $HbA_{1c}$ reduction $\geq 1\%$ and weight loss $\geq 5\%$
- $HbA_{1c}$ reduction $\geq 1\%$ and weight loss $\geq 10\%$

The above 10 endpoints will be analysed for the primary estimand. The analyses for the primary estimand for all 10 endpoints will be based on the 'on-treatment without rescue medication' observation period. They will be analysed separately using the same type of logistic regression model with treatment and the SU and SGLT-2 inhibitors anti-diabetic background medication

stratification factor (SU +/- metformin, SGLT-2 inhibitors +/- metformin, SU + SGLT-2 inhibitors +/- metformin, no SU and no SGLT-2 inhibitors (metformin monotherapy)) as categorical effects and associated baseline and post-baseline response(s) (i.e. $HbA_{1c}$ responses for $HbA_{1c}$ endpoints, body weight responses for weight endpoints and both $HbA_{1c}$ and body weight responses for the binary endpoints that combine both parameters) as covariates.

To account for missing data, the analysis will be made using a sequential multiple imputation approach as described below:

- Multiple imputed data sets (500) will be created in which missing values for the underlying continuous assessments are imputed by treatment group assuming MAR similar to the approach described for the primary analysis in Section 2.4.1
- The binary endpoint will be created for each of the 500 complete data sets
- Each of the created complete data sets will be analysed with the logistic regression model. Estimated odds ratios will be log transformed and inference will be drawn using Rubin's rule[7].

The results after applying Rubin's rule will be back-transformed and described by the odds ratio between treatments and the associated 95% confidence interval and p-value for no treatment difference.

#### 2.5.2.2 Safety endpoints

The safety endpoints will be evaluated based on SAS using the on-treatment observation period and the in-trial observation period unless otherwise stated.

**Adverse Events**

The following endpoint related to AEs is used to support the safety objective;

- Number of treatment emergent adverse events

A treatment-emergent AE is an event that has onset date (or increase in severity) during the on-treatment observation period. These will therefore be referred to as 'on-treatment AEs' hereafter. On-treatment AEs 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 100 years (R). These summaries are replicated by outputs including all 'in-trial' AEs (i.e., AEs with onset date [or increase in severity] during the 'in-trial' observation period). AEs with onset after the end of the 'in-trial' observation period will be reported in a listing. The development over time in gastrointestinal AEs will be presented graphically.

The most frequent AEs will be defined as preferred terms that are experienced by at least 5% of the subjects in any of the treatment arms.

All AEs will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA) coding.

**Hypoglycaemic episodes**

The following two endpoints related to hypoglycaemic episodes are used to support the safety objective:

- Number of treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes
- Treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes (yes/no)

Data on treatment-emergent hypoglycaemic episodes are presented in terms of the number of subjects with at least one episode, the percentage of subjects with at least one episode (%), the total number of episodes and the episode rate per 100 years of exposure. Summaries of treatment-emergent hypoglycaemic episodes will be presented as an overview including all episodes and episodes by severity.

**Classification of Hypoglycaemia:**

Treatment emergent: hypoglycaemic episodes will be defined as treatment emergent if the onset is in the on-treatment observation period (see Section 2.3.1.2)

Nocturnal hypoglycaemic episodes: are episodes occurring between 00:01 and 05.59 both inclusive.

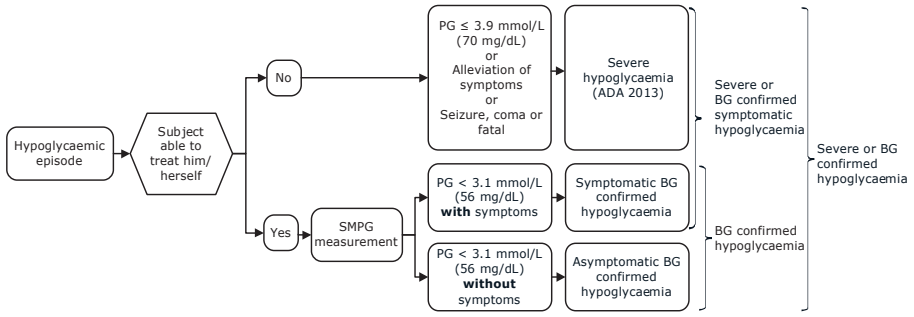
Hypoglycaemic episodes are classified according to the Novo Nordisk classification of hypoglycaemia (see Figure 2–2) and the ADA classification of hypoglycaemia (see Figure 2–3).

**Novo Nordisk classification of hypoglycaemia**

In normal physiology, symptoms of hypoglycaemia occur below a plasma glucose level of 3.1 mmol/L (56 mg/dL) [8]. Therefore, Novo Nordisk has included hypoglycaemia with plasma glucose levels below this cut-off point in the definition of BG confirmed hypoglycaemia.

Novo Nordisk uses the following classification (see Figure 2–2) in addition to the ADA classification:

- Severe or BG confirmed symptomatic hypoglycaemia: An episode that is severe according to the ADA classification[9] or BG confirmed by a plasma glucose value <3.1 mmol/L (56 mg/dL) **with** symptoms consistent with hypoglycaemia.




Note: Glucose measurements are performed with capillary blood calibrated to plasma equivalent glucose values

BG: blood glucose PG: plasma glucose SMPG: Self-measured plasma glucose

**Figure 2–2 Novo Nordisk classification of hypoglycaemia**

## ADA classification[9] of hypoglycaemia

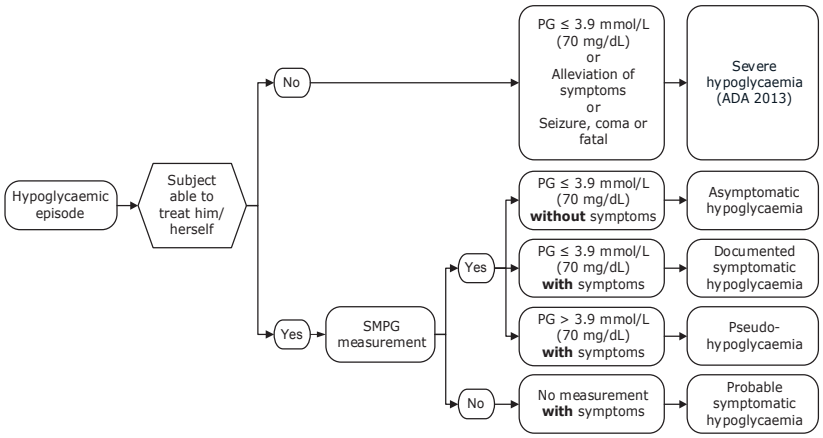
- Severe hypoglycaemia: An episode requiring assistance of another person to actively administer carbohydrate, glucagon, or take other corrective actions. Plasma glucose concentrations may not be available during an event, but neurological recovery following the return of plasma glucose to normal is considered sufficient evidence that the event was induced by a low plasma glucose concentration.

- Asymptomatic hypoglycaemia: An episode not accompanied by typical symptoms of hypoglycaemia, but with a measured plasma glucose concentration ≤ 3.9 mmol/L (70 mg/dL).

- Documented symptomatic hypoglycaemia: An episode during which typical symptoms of hypoglycaemia are accompanied by a measured plasma glucose concentration ≤ 3.9 mmol/L (70 mg/dL).

- Pseudo-hypoglycaemia: An episode during which the person with diabetes reports any of the typical symptoms of hypoglycaemia with a measured plasma glucose concentration > 3.9 mmol/L (70 mg/dL) but approaching that level.

- Probable symptomatic hypoglycaemia: An episode during which symptoms of hypoglycaemia are not accompanied by a plasma glucose determination but that was presumably caused by a plasma glucose concentration ≤ 3.9 mmol/L (70 mg/dL).




Note: Glucose measurements are performed with capillary blood calibrated to plasma equivalent glucose values

PG: plasma glucose SMPG: Self-measured plasma glucose

**Figure 2–3 ADA classification of hypoglycaemia**

### Number of treatment emergent severe or blood glucose (BG) confirmed symptomatic hypoglycaemic episodes

Number of treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 35 weeks will be analysed using a negative binomial regression model with a log -link function and the logarithm of the time period, from the randomisation and up to the time point in which an occurrence of a hypoglycaemic episode is considered treatment emergent as offset assuming MAR. The model will include factors for treatment and the SU and SGLT-2 inhibitors anti-diabetic background medication stratification factor (SU +/- metformin, SGLT-2 inhibitors +/- metformin, SU + SGLT-2 inhibitors +/- metformin, no SU and no SGLT-2 inhibitors (metformin monotherapy)) as categorical factors and baseline $HbA_{1c}$ as covariate. In case of convergence issues, due to sparse data, the stratification factor will be omitted from the model. The SAS will be used for the analysis.

The results will be described by the rate ratio between treatments and the associated 95% confidence interval and p-value for no treatment difference.

### Treatment emergent severe or blood glucose confirmed symptomatic hypoglycaemia episodes (yes/no)

The binary endpoint indicating whether a subject has no treatment-emergent severe or BG confirmed symptomatic hypoglycaemic episodes or at least one will be analysed using a logistic regression model. The model will include factors for treatment and the SU and SGLT-2 inhibitors anti-diabetic background medication stratification factor (SU +/- metformin, SGLT-2 inhibitors +/-

metformin, SU + SGLT-2 inhibitors +/- metformin, no SU and no SGLT-2 inhibitors (metformin monotherapy)) as categorical factors and baseline $HbA_{1c}$ as covariate. In case of convergence issues, due to sparse data, the stratification factor will be omitted from the model. The SAS will be used for the analysis.

The results will be described by the odds ratio between treatments and the associated 95% confidence interval and p-value for no treatment difference.

**Laboratory assessments**

The laboratory assessments supporting the safety objective are change from baseline to week 30 in:

- Haematology
- Biochemistry
- Calcitonin

The above continuous laboratory assessments will be summarised and evaluated by descriptive statistics.

In addition amylase and lipase will be analysed separately using an analysis similar to the primary analysis of the primary endpoint. However this analysis will be based on SAS using the on-treatment observation period.

Both analyses will use the associated log-transformed baseline value as covariates instead of $HbA_{1c}$. Lipase and amylase values will be log-transformed prior to the analysis.

**Pulse rate**

Change from baseline to week 30 in pulse rate will be analysed separately with the same model approach as for amylase and lipase, but with the pulse rate (not log-transformed) value at baseline as covariate.

**Categorical safety assessments**

The categorical assessments supporting the safety objective are change from baseline to week 30 in:

- ECG category
- Physical examination category
- Eye examination category

The above assessments will be summarised descriptively

## 2.6 Health economics and/or PROs

Change from baseline to week 30 in:

- Scores for selected patient reported outcomes:
  - SF-36v2™ (standard version): Total summary scores (physical component and mental component) and scores from the 8 domains
  - DTSQs: Treatment satisfaction score (sum of 6 of 8 items) and the 8 items separately

The PRO questionnaires, SF-36v2™, DTSQs will be used to evaluate the objective regarding Quality of Life. Each of the PRO endpoints will be analysed separately as the other continuous efficacy endpoints using a similar model approach as for the primary endpoint with the associated baseline value as covariates.

## 2.7 Interim analysis

No interim analyses will be performed before the database is locked.

# 3 Changes to the statistical analyses planned in the protocol

The changes to the statistical analyses planned in the protocol in described in Table 3–1 below.

**Table 3–1 Overview of changes**

| Change to planned statistical analyses | Rationale for change |
|---|---|
| Addition of PP analysis set | The PP analysis set was not pre-specified in the protocol so details have been provided in this SAP. |
| Sensitivity analysis on body weight:<br><br>Retrieved drop-out analysis and tipping point analysis | The sensitivity analyses to be conducted to test the robustness of the results of the confirmatory superiority hypothesis on body weight were not pre-specified in the protocol so details have been provided in this SAP. |
| Wording on analysis on pulse rate | Clarification of that pulse rate should not be log-transformed prior to analysis |
| Stratification factor in hypoglycaemic episode analyses | For both the negative binomial model and the logistic regression model for analysing hypoglycaemic episodes, it has been specified that the stratification factor will be omitted as a covariate in case of convergence issues due to few hypoglycaemic episodes. |
| Wording on retrieved drop-out analysis on body weight | Per new preferred terminology, this is no longer called an 'in-trial' analysis, but rather a retrieved drop-out analysis. |
| Wording on analyses of amylase and lipase | Clarification of that baseline values of amylase and lipase should be log-transformed before being used as covariates in the analyses. |
| Clarification of retrieved drop-out analysis | Clarification of that the covariates to be used in the ANCOVA step of the retrieved drop-out analysis, is the same ones as for the primary analysis.<br><br>Clarification that the retrieved drop-out analysis will only be conducted to test the robustness of the superiority hypotheses. |
| Clarification of overall alpha level in testing approach | It has been clarified that the overall alpha level of 2.5% is one-sided, corresponding to a 5% two-sided alpha level. |

# 4 References

1. U.S.Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Guidance for Industry, Diabetes Mellitus - Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes. December 2008.
2. U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for Industry. Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention, Draft Guidance. 2008.
3. Sethi B, Viswanathan, V., Kumar, A., Chatterjee, S., Chandalia, HB. Liraglutide in Clinical Practice: Insights from LEAD Programme. JAPI; 2010. p. 18-22.
4. Marre M, Shaw J, Brandle M, Bebakar WM, Kamaruddin NA, Strand J, et al. Liraglutide, a once-daily human GLP-1 analogue, added to a sulphonylurea over 26 weeks produces greater improvements in glycaemic and weight control compared with adding rosiglitazone or placebo in subjects with Type 2 diabetes (LEAD-1 SU). Diabet Med. 2009;26(3):268-78.
5. Zinman B, Gerich J, Buse JB, Lewin A, Schwartz S, Raskin P, et al. Efficacy and safety of the human glucagon-like peptide-1 analog liraglutide in combination with metformin and thiazolidinedione in patients with type 2 diabetes (LEAD -4 Met+TZD). Diabetes Care. 2009;32(7):1224-30.
6. Rohmeyer K, Klinglmueller F. gMCP: Graph Based Multiple Test Procedures. R package version 0.8-8. 3 Oct 2014.
7. Little RJA, Rubin DB. Statistical analysis with missing data: New York: John Wiley & Sons. 1987.
8. Schwartz NS, Clutter WE, Shah SD, Cryer PE. Glycemic thresholds for activation of glucose counterregulatory systems are higher than the threshold for symptoms. J Clin Invest. 1987;79(3):777-81.
9. Seaquist ER, Anderson J, Childs B, Cryer P, Dagogo-Jack S, Fish L, et al. Hypoglycemia and diabetes: a report of a workgroup of the American Diabetes Association and the Endocrine Society. Diabetes Care. 2013;36(5):1384-95.