Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT02906917
Sponsor trial ID:	NN5401-4266
Official title of study:	A 38 week trial comparing effect and safety of insulin degludec/insulin aspart vs. insulin glargine plus insulin aspart in subjects with type 2 diabetes treated with basal insulin with or without oral antidiabetic treatment in need of treatment intensification
Document date:	22 June 2018

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: NN5401-4266

A 38 week trial comparing effect and safety of insulin degludec/insulin aspart vs. insulin glargine plus insulin aspart in subjects with type 2 diabetes treated with basal insulin with or without oral antidiabetic treatment in need of treatment intensification

Trial phase: 3b

Author:
[redacted] ([redacted])
Biostatistics Degludec and Devices

Table of contents

Page

List of abbreviations

ADA	American Diabetes Association
AE	adverse event
ALT	alanine amino transferase
ANCOVA	analysis of covariance
AP	alkaline phosphatase
AST	aspartate aminotransferase
BG	blood glucose
BHQ	baseline hypo questionaire
BID	two times a day
CAS	Completer analysis set
CPMP	Committee for Proprietary Medicinal Products
EOT	end of treatment
FAS	full analysis set
FPG	fasting plasma glucose
HbA_{1c}	glycosylated haemoglobin
IAsp	insulin aspart
ICH	International Conference on Harmonisation
IDeg	insulin degludec
IDegAsp	insulin degludec/insulin aspart
IGlar	insulin glargine
ITT	intention-to-treat
MMRM	mixed model for Repeated Measurement
NA	not applicable
NI	non-inferiority
OAD	oral antidiabetic drug
OD	once daily
PG	plasma glucose

PP	per protocol
REML	restricted maximum likelihood
SAP	statistical analysis plan
SD	standard deviation
SF-36v2®	Short-Form Health Survey 36 version 2
SMPG	self-measured plasma glucose
TEAE	treatment-emergent adverse event
TID	three times daily
TE	treatment effect
TRIM-D	Treatment Related Impact Measure-Diabetes
T2DM	type 2 diabetes mellitus

1 Introduction

1.1 Trial information

Objectives and endpoints

The primary objective

- To confirm the effect of insulin degludec/insulin aspart once daily versus insulin glargine once daily in combination with insulin aspart once daily in controlling glycaemia after 26 weeks in subjects with type 2 diabetes mellitus treated with basal insulin with or without oral antidiabetic treatment in need of treatment intensification.

The secondary objectives

- To compare other effect parameters and safety of insulin degludec/insulin aspart once daily versus insulin glargine once daily in combination with insulin aspart once daily after 26 weeks in subjects with type 2 diabetes mellitus treated with basal insulin with or without oral antidiabetic treatment in need of treatment intensification.
- To compare the effect and safety of insulin degludec/insulin aspart once daily or twice daily versus insulin glargine once daily in combination with insulin aspart (1-3 times daily) after 38 weeks in subjects with type 2 diabetes mellitus treated with basal insulin with or without oral antidiabetic treatment in need of treatment intensification.

Primary endpoint

- Change from baseline in glycosylated haemoglobin (HbA_{1c}) (%) after 26 weeks

Supportive secondary efficacy endpoints:

- Change from baseline in HbA_{1c}(%) after 38 weeks
- Change from baseline in fasting plasma glucose (FPG) after 26 weeks
- Change from baseline in fasting plasma glucose (FPG) after 38 weeks
- Responder (Yes/No) for HbA_{1c} after 26 weeks: $HbA_{1c} < 7\%$
- Responder (Yes/No) for HbA_{1c} after 38 weeks: $HbA_{1c} < 7\%$
- Responder (Yes/No) for HbA_{1c} after 26 weeks: $HbA_{1c} < 7\%$ without severe or BG confirmed symptomatic hypoglycaemia
- Responder (Yes/No) for HbA_{1c} after 38 weeks: $HbA_{1c} < 7\%$ without severe or BG confirmed symptomatic hypoglycaemia
- Self-measured plasma glucose measurements (SMPG) after 26 weeks: Mean pre-breakfast measurements used for titration

- Self-measured plasma glucose measurements (SMPG) after 38 weeks: Mean pre-breakfast measurements used for titration
- Self-measured plasma glucose measurements (SMPG) after 26 weeks: 9-point profile - Post-prandial increments
- Self-measured plasma glucose measurements (SMPG) after 38 weeks: 9-point profile - Post-prandial increments

Supportive secondary safety endpoints:

- Incidence of treatment-emergent adverse events (TEAEs) from baseline to week 26
- Incidence of treatment-emergent adverse events (TEAEs) during 38 weeks
- Incidence of treatment-emergent adverse events (TEAEs) during intensification period (week 26-38)
- Number of nocturnal, treatment-emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 26 weeks
- Number of nocturnal, treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 38 weeks
- Number of nocturnal, treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during maintenance (week 16-26)
- Number of treatment-emergent severe or blood glucose (BG) confirmed symptomatic hypoglycaemic episodes during 26 weeks
- Number of treatment-emergent severe or blood glucose (BG) confirmed symptomatic hypoglycaemic episodes during 38 weeks
- Number of treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during maintenance (week 16-26)
- Total insulin dose at week 26
- Total insulin dose at week 38
- Change in weight from baseline to week 26
- Change in weight from baseline to week 38

Explorative endpoints

- Change from baseline to week 26 for:
    - Health-related quality of life as evaluated by SF-36v2® Health Survey
    - Treatment related impact as evaluated by the Treatment Related Impact Measure-Diabetes (TRIM-D)

- Change from baseline to week 38 for:
    - Health-related quality of life as evaluated by the SF-36v2® Health Survey

- Treatment related impact as evaluated by the TRIM-D

Trial design:

This is a 38-week, multinational, open-label, two-arm, randomised (1:1), controlled, parallel-group, treat-to-target trial in subjects with type 2 diabetes mellitus treated with basal insulin with or without oral antidiabetic drug(s) in need of treatment intensification. Subjects will be randomised in a 1:1 manner to receive either insulin degludec/insulin aspart (IDegAsp) OD or insulin glargine (IGlar) OD in combination with insulin aspart (IAsp) OD for the first 26 weeks followed by a 12-weeks intensification period. The randomisation will be stratified based on the pre-trial basal insulin treatment regimen: OD or BID/TID (two/three times daily) dosing.

Subjects will remain in the same treatment arm throughout the trial, with the possibility of intensification of their treatment. At week 26 (V28), subjects not achieving sufficient glycaemic control (V27 $HbA_{1c} \geq 7\%$) will be further intensified from either IDegAsp OD to IDegAsp BID or IGlar OD + IAsp OD to IGlar OD + IAsp BID, whereas subjects in good glycaemic control (V27 $HbA_{1c} < 7\%$) will continue their insulin treatment regimen.

At week 32 (V34), glycaemic control will be evaluated again, and subjects in good glycaemic control (V33 $HbA_{1c} < 7\%$) will continue their insulin treatment regimen. Subjects on IDegAsp OD not reaching the glycaemic target (V33 $HbA_{1c} \geq 7\%$) will be intensified to IDegAsp BID, whereas subjects in the IGlar arm not reaching glycaemic control will be intensified with one additional IAsp dose to IGlar OD + IAsp BID or TID daily for the remaining six weeks of the intensification period.

Further details

For further details on trial design and endpoints please see the NN5401-4266 protocol version 4.0.

1.2 Scope of the statistical analysis plan

This SAP is based on the NN5401-4266 protocol version 4.0.

2 Statistical considerations

Novo Nordisk will analyse and report data from all sites together.

Unless otherwise specified, all continuous measurements will be summarised descriptively at each visit by treatment using in-trial data. Endpoints that are analysed untransformed and endpoints that are not formally analysed are summarised by the arithmetic mean, standard deviation (SD), median, and minimum and maximum value. Endpoints that are analysed log-transformed are supplemented with the geometric mean and coefficient of variation. For measurements over time, mean values will be plotted to explore the trajectory over time. The geometric mean values will be plotted for

endpoints that are analysed log-transformed. All descriptive summaries and plots will be based on in-trial data unless otherwise specified.

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 has also been made at screening, then the screening value will be used as the baseline value. For summaries and statistical analysis of all endpoints the first non-missing values is used. The value should be taken according to the specification in the protocol, e.g. non-fasting FPG-values are not used. Re-tests of non-missing values are only included in subject specific listings.

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

Presentation of results from a statistical analysis will include the estimated mean treatment effects for absolute values and change from baseline. The estimated means are either obtained by Rubin's rule weighing together estimated means from multiple imputation or directly from estimation in a parameterized statistical model. In addition estimated mean treatment differences (or ratios) will be presented together with two-sided 95% confidence intervals and corresponding two-sided p-values.

2.1 Primary and secondary estimands

The primary estimand will be the difference in change from baseline at week 26 or week 38 depending on the endpoint considered between IDegAsp OD and IGlar OD + IAsp OD, for all randomised subjects regardless of whether the subjects remained on initially assigned treatment until week 26 or week 38. This estimand is a de facto estimand addressing effectiveness.

The secondary estimand will be the difference in change from baseline at week 26 or week 38 depending on the endpoint considered between IDegAsp OD and IGlar OD + IAsp OD if all randomised subjects had adhered to treatment until week 26 or week 38. This estimand is a de jure estimand addressing efficacy.

For hypoglycaemic episode endpoints, only the de jure estimand will be used. The reason for this is that data for hypoglycaemic episodes will not be systematically collected after treatment discontinuation.

2.2 Sample size calculation

The sample size is based on confirming non-inferiority of the primary estimand for the primary endpoint assessed at week 26 using the full analysis set (FAS). Sample sizes are also given for confirming non-inferiority for the primary estimand of the primary endpoint on the per-protocol set.

The primary objective is to confirm the effect of IDegAsp OD versus IGlar OD + IAsp OD in controlling glycaemia in basal insulin-treated subjects with T2DM.

This will be done by comparing the difference between IDegAsp OD and IGlar OD + IAsp OD in change from baseline in HbA_{1c} after 26 weeks to a non-inferiority limit of 0.4%. The non-inferiority limit of 0.4% is aligned with the phase 3a programme and is considered appropriate to ensure a relevant effect of IDegAsp compared to placebo since a superior treatment effect of −0.85% [−1.04; −0.66] of IGlar versus placebo was reported in a placebo controlled clinical trial in T2DM subjects[1]. Supplementing IGlar with IAsp OD would further increase the treatment effect compared to placebo.

Formally, let D be the mean treatment difference (IDegAsp OD minus IGlar OD + IAsp OD) in change from baseline in HbA_{1c}. The null-hypothesis of IDegAsp OD inferior by 0.4% or more will be tested against the alternative hypothesis of non-inferiority as given by

H_0: D ≥0.40% against H_A: D <0.40%

Non-inferiority will be considered confirmed if the upper bound of the two-sided 95% confidence interval for D is strictly below 0.4%. This is equivalent to using a one-sided test of size 2.5%, which means that the Type 1 error rate is controlled at 2.5% level.

Accordingly, the sample size is calculated using a t-statistic under the assumptions of a one-sided test of size 2.5%, a mean treatment difference of 0.0%, a standard deviation of SD=1.3%, and a non-inferiority limit of 0.40%. It is assumed that 15% of the randomised subjects will discontinue randomised treatment before week 26 or have missing data at week 26, and that 10% of the randomised subjects will be excluded from the per protocol (PP) analysis set. The assumptions are based on the results from a phase 3a trial (NN5401-3593).

When estimating the primary estimand for the evaluation of the non-inferiority hypothesis, missing data will be imputed based on in-trial data. A penalty will be added to the change in HbA_{1c} at week 26 for all subjects in the IDegAsp OD arm who discontinued trial product prior to week 26 or who have missing data at week 26. The penalty corresponds to the non-inferiority limit of 0.4%.

In total an expected 15% will have missing on-treatment data at week 26, 10% will not be part of the PP analysis set. This means that 5/90*100%= 5.56% of the PP analysis set is expected to be non-completers, hence the expected treatment differences (adjusted treatment effect (TE) resulting from the applied penalty) for the non-inferiority hypothesis of the primary estimand is:

1) FAS: 0.15*0.4=0.06

2) PP: 0.0556*0.4=0.0222

The sample size assumptions for TE and the common SD are given in Table 2–1.

From these assumptions, and based on a 1:1 randomisation, the sample size is set to 264 subjects per treatment arm, in total 528 subjects. This will ensure a nominal power of at least 85% for confirming the primary objective based on the FAS as well as on the PP analysis set. The sample size calculation was done using SAS 9.4.

Table 2–1 Assumptions for sample size calculation

Primary Estimand Analysis set	Non-inferiority margin	SD	TE	Adjusted TE	Randomisation Scheme	Required Power
FAS	0.4%	1.3%	0.0%	0.06%	1:1	85%
PP	0.4%	1.3%	0.0%	0.0222%	1:1	85%

Abbreviations: SD = standard deviation; TE = Treatment effect

Table 2–2 Sensitivity of sample size to variations in SD and power

#subjects in total		SD=1.1%	SD=1.2%	SD=1.3%	SD=1.4%
Power 80%	**FAS**	332	394	462	536
	PP	300	356	416	482
Power 85%	**FAS**	378	450	528	612
	PP	342	406	476	552
Power 90%	**FAS**	442	526	618	716
	PP	400	474	556	644
Power 95%	**FAS**	546	650	762	884
	PP	494	586	686	796

2.3 Definition of analysis sets

The following analysis sets are defined in accordance with the ICH-E9[2].

- Full Analysis Set (FAS): includes all randomised subjects. In exceptional cases, subjects may be excluded. In such cases the elimination will be justified and documented. The statistical evaluation of the FAS will follow the intention-to-treat (ITT) principle and subjects will contribute to the evaluation “as randomised”.

- Per-Protocol (PP) analysis set: includes all subjects in the Full Analysis Set who fulfils the following criteria:
    - Have not violated any inclusion criteria
    - Have not fulfilled any exclusion criteria
    - Have a non-missing HbA_{1c} at screening or randomisation
    - Have at least one non-missing HbA_{1c} after 12 weeks of exposure
    - Have at least 12 weeks of exposure

Subjects in the PP will contribute to the evaluation "as treated".

- Safety Analysis Set (SAS): includes all subjects receiving at least one dose of the investigational product or comparator. Subjects in the safety set will contribute to the evaluation "as treated".
- Trial completer analysis set: Includes all randomised subjects who has attended either V40 on treatment, or V40A, 38 weeks after randomisation according to protocol .
- Treatment completer analysis set: Includes all randomised subjects who have completed the trial without discontinuation from randomised treatment.

Before data are released for statistical analysis, a review of all data will take place to ensure a sufficient data quality and to ensure the planned statistical analysis are applicable. Any data decisions not foreseen in the protocol will be documented before database lock.
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.

The decision to exclude any subject or observation from the statistical analysis is the joint responsibility of the study group. The subjects or observations to be excluded, and the reasons for their exclusion will be documented before database lock. The subjects and observations excluded from the analysis sets, and the reasons for this, will be described in the clinical trial report.

Definition of the observation periods:

In-trial: This observation period will include information collected at or after date of randomisation and up to the last scheduled subject-site contact. For subjects who withdraw their informed consent, or is lost to follow-up, the end of the in-trial period is defined as the date of the last scheduled subject-investigator contact (site or phone visit). In case a subject dies during the trial, the date of death will be the end-date of the in-trial observation period regardless of the above defined end-dates. This observation period will be used for estimating the effectiveness estimand.

On-treatment: This observation period is a subset of the in-trial observation period and represents the time period in which a subject is considered exposed to trial product. The period starts on day of the first dose of trial product and ends 7 days after last dose of trial product is taken

2.4 Primary endpoint

Primary statistical analysis for primary estimand

The primary estimand will be estimated based on the FAS using all post baseline HbA_{1c} measurements obtained on planned visits up until week 26, including week 26 measurements from subjects withdrawing from randomised treatment (in-trial observation period), mimicking an ITT scenario. To estimate this estimand for evaluation of non-inferiority of $HbA_{1c,}$ multiple imputation of missing values with separate imputation from each treatment arm will be applied. The use of data from prematurely treatment discontinued subjects in analyses will equalise treatment effects and the inclusion of a penalty equal to the non-inferiority margin will remedy this. Imputation will be done as follows:

- In the first step, 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 treatment arm separately and 1000 copies of the dataset will be generated.

- In the second step, for each of the 1000 copies of the dataset, an analysis of covariance (ANCOVA) for each treatment arm with region, sex, previous insulin regimen and previous OAD treatment as categorical fixed effects and baseline HbA_{1c} measurement and age as a covariate will be fitted to the change in HbA_{1c} from baseline to week 8 (visit 10). The estimated parameters and their variance from this model are used to impute missing values at week 8 for subjects in the respective treatment arms based on region, sex, previous insulin regimen, previous OAD treatment, age and HbA_{1c} at baseline.

- In the third step, for each of the 1000 copies of the dataset, an analysis of covariance model with region, sex, previous insulin regimen and previous OAD treatment as categorical fixed effects, and age, baseline HbA_{1c} and HbA_{1c} at week 8 (Visit 10) as covariates is fitted to the change in HbA_{1c} from baseline to week 12 (Visit 14) for each treatment arm. The estimated parameters and their variances from this model are used to impute missing values at week 12 for subjects in the respective treatment arms, based on region, sex, previous insulin regimen, previous OAD treatment, age and HbA_{1c} at baseline and week 8.

- Step three is repeated over the available planned visits, adding one visit at a time until week 26. In each step the model is expanded to include a covariate from the previous visit.

- Finally, a penalty equal to the non-inferiority margin 0.4% is added to the week 26 (visit 28) HbA_{1c} value for any prematurely treatment discontinued subject in the IDegAsp OD arm, irrespective if the value was retrieved (V28A) data or an imputed value from the sequential imputation, as well as for other subject in IDegAsp arm with missing HbA_{1c} measurement after 26 weeks.

For each of the 1000 (now complete) imputed data sets the change from baseline to week 26 will be analysed using an ANCOVA with treatment, region, sex, previous insulin treatment regimen and previous OAD treatment as categorical fixed effects and baseline HbA_{1c} and age as covariate. The estimates and standard deviations for the 1000 data sets are pooled to one estimate and associated standard deviation using Rubin's rule. From these pooled estimates the confidence interval for the treatment differences and the associated p-value are calculated.

Non-inferiority of IDegAsp OD vs. IGlar OD + IAsp OD will be considered confirmed if the 95% confidence interval for the mean treatment difference lies entirely below 0.40%. A conclusion on non-inferiority will be based on the FAS and the primary estimand. The addition of the 0.4% penalty in the last imputation step corresponds to the non-inferiority null method, as described by Koch[3].

Sensitivity analyses for the primary estimand

The following sensitivity analyses will be performed:

- The primary analysis will be repeated based on the PP analysis set in line with the Committee for Proprietary Medicinal Products (CPMP) Points to Consider (CPMP/EWP/482/99).

- The primary analysis will be repeated on the FAS but without adding the penalty for subjects in the IDegAsp OD arm. The result of this analysis is regarded to give the most fair and unbiased estimate of the difference between the two treatments, in contrast to the primary analysis that due to the penalty is biased in favour of the IGlar treatment.

- A tipping point analysis based on the FAS and the primary analysis for the primary estimand will be performed. In this analysis subjects who withdraw from the IDegAsp OD arm or have missing values at visit 26 are assumed to have an effect that is inferior to the de facto effect of IDegAsp OD. The extent of the inferiority (also termed a 'penalty') will be gradually increased to evaluate at which value IDegAsp OD is no longer non-inferior to IGlar OD + IAsp OD. This penalty value, also known as the tipping point, corresponds to the hypothetical degree of efficacy deterioration in withdrawn subjects that would change the conclusion of non-inferiority.

Primary analysis for the secondary estimand

The secondary estimand will be estimated based on the FAS using post-baseline measurements up to and including week 26 from the on-treatment observation period. To estimate this estimand, the change from baseline in HbA_{1c} after 26 weeks of treatment will be analysed using the mixed model for Repeated Measurement (MMRM) method with a restricted maximum likelihood (REML). The model will include treatment, sex, previous insulin treatment regimen, previous OAD treatment and region as factors and baseline HbA_{1c} and age as covariates, all nested within visit. An unstructured covariance matrix for HbA_{1c} measurements within the same subject will be employed, assuming measurements from different subjects are independent.

The MMRM is a well-established method that accounts for the uncertainty pertaining to missing data. This analysis assumes that the missing data mechanism is missing at random (MAR). Under this assumption the statistical behaviour of the missing data (given the observed responses and model fixed effects and covariates) is assumed to be the same as the observed data.

In a non-inferiority trial the efficacy estimand may be more appropriate than the effectiveness estimand, since it is not confounded by the effects of changing treatment from the randomised treatment to other treatments.

Sensitivity analyses for the secondary estimand

The following sensitivity analyses will be performed:

- The above MMRM analysis will be repeated on the PP using the on-treatment period, as well on the subset of subjects who remains on treatment up until week 26 visit. These analyses will be done to align with the reporting of the phase 3a trials.

- Using the FAS, the change in HbA1c from baseline after 26 weeks of treatment will be analysed using an analysis of variance (ANOVA) method with treatment, sex, previous insulin treatment regimen, previous OAD treatment and region as fixed factors, and age and baseline HbA_{1c} as covariates, and where missing values will be imputed using the Last Observation Carried Forward (LOCF) method. This is done to align with the reporting of the phase 3a trials.

2.5 Supportive secondary endpoints

2.5.1 Efficacy endpoints

For the following continuous supportive efficacy endpoints, analyses similar to those for the primary estimand of the primary endpoint will be done, with the adjustment that the non-inferiority-penalty (NI-penalty) in step five of multiple imputation is not applied (this is only considered

relevant for non-inferiority testing). Continuous safety endpoints that are analysed statistically will also follow this procedure. A further change to the above specified model is that the model will include the baseline measurement of the endpoint as a covariate instead of the baseline HbA_{1c} measurement.

The following supportive secondary endpoints will be analysed through the statistical analysis described above.

HbA_{1c}

Change from baseline in HbA_{1c} (%) after 38 weeks.

In addition to the analysis specified for the primary estimand, the analysis specified for the secondary estimand will also be performed for this endpoint to align with analysis of HbA_{1c} after 26 weeks.

Fasting plasma glucose

- Change from baseline in FPG after 26 weeks
- Change from baseline in FPG after 38 weeks

Self-measured plasma glucose measurements (SMPG)

- **SMPG after 26 weeks**
  - Change from baseline in mean pre-breakfast measurements after 26 weeks
  - Change from baseline in post-prandial increments after 26 weeks
- **SMPG after 38 weeks**
  - Change from baseline in mean pre-breakfast measurements after 38 weeks
    Change from baseline in post-prandial increments after 38 weeks

Furthermore, the mean PG before breakfast used for dose-titration and the 9-point profile will be calculated at each visit using the available data.

Responder for HbA_{1c}

- **Responder (Yes/No) for HbA_{1c} after 26 weeks:**
  - $HbA_{1c} < 7\%$
  - $HbA_{1c} < 7\%$ without severe or BG confirmed symptomatic hypoglycaemia
- **Responder (Yes/No) for HbA_{1c} after 38 weeks:**
  - $HbA_{1c} < 7\%$
  - $HbA_{1c} < 7\%$ without severe or BG confirmed symptomatic hypoglycaemia

Subjects with less than 16 weeks of treatment will be set as non-responders. Hypoglycaemic episodes will be based on the available data.

Analysis of each of the responder endpoints will be based on a logistic regression model which will include treatment, sex, previous insulin treatment regimen, previous OAD treatment and region as fixed factors, and age and baseline HbA_{1c} as covariates. The analysis for the responder endpoints will be derived using the HbA_{1c} values from the 1000 datasets originating from multiple imputation analysis of HbA_{1c}, without applying the NI-penalty.

Each of these datasets will be analysed using the above logistic regression model and the results will then be pooled using Rubin's rule.

Explorative endpoints

The following endpoint will be analysed through the same statistical analysis for continuous endpoint:

- SMPG after 38 weeks
    - Change from baseline in post-prandial increments after 38 weeks

- Change from baseline to week 26 for:
    - Health-related quality of life as evaluated by SF-36v2® Health Survey
    - Treatment related impact as evaluated by the Treatment Related Impact Measure-Diabetes (TRIM-D)

- Change from baseline to week 38 for:
    - Health-related quality of life as evaluated by the SF-36v2® Health Survey
    - Treatment related impact as evaluated by the TRIM-D

Assessments

The following efficacy assessments will be summarised descriptively but with no statistical analyses performed:

- Responder (Yes/No) for HbA_{1c} after 26 weeks:
    - $HbA_{1c} < 7\%$ without nocturnal (00:01-05:59) severe or BG confirmed symptomatic hypoglycaemia

- Responder (Yes/No) for HbA_{1c} after 38 weeks:
    - $HbA_{1c} < 7\%$ without nocturnal (00:01-05:59) severe or BG confirmed symptomatic hypoglycaemia

2.5.2 Safety endpoints

Adverse events

The following endpoint on adverse events will be reported:

- Incidence of treatment-emergent adverse events (TEAEs) from baseline to week 26
- Incidence of treatment-emergent adverse events (TEAEs) during 38 weeks
- Incidence of treatment-emergent adverse events (TEAEs) during intensification period (week 26-38)

The AEs will be coded using the most recent version of the Medical Dictionary for Regulatory Activities coding.

A treatment-emergent adverse event (TEAE) is defined as an event that has onset date on or after the first day of exposure to randomised treatment and no later than seven days after the last day of randomised treatment. If the event has onset date before the first day of exposure on randomised treatment and increases in severity during the treatment period and until seven days after the last drug date, then this event should also be considered as a TEAE.

TEAEs are summarised descriptively, whereas non-treatment-emergent AE's are presented in listings. TEAE data will be displayed 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 of exposure (R).

Summaries of TEAEs and serious TEAEs will be presented as an overview including all AEs, serious AEs, number of deaths, AEs by severity, AEs by relation to treatment as assessed by the investigator and AEs of special interest including AEs leading to withdrawal and AEs leading to drug withdrawn. .

Furthermore summary tables based on system organ class and preferred terms are made for:

- All TEAEs
- Serious TEAEs
- Possibly and probably related TEAEs
- Severe, moderate and mild TEAEs
- TEAEs reported by safety area of interest
- TEAEs with preferred term that are experienced by at least 5% (1%) of the subjects in any treatment arm or by at least 5% (1%) of all subjects
- TEAEs leading to withdrawal of subject or drug withdrawn

A listing of non-treatment-emergent events with onset date before the first day of exposure to randomised treatment will be presented. A listing will also be presented for non-treatment-emergent adverse events collected after the treatment-emergent period according to the definition of the TEAE.

Hypoglycaemic episodes

The definition and classification of hypoglycaemic episodes is given after the description of the statistical analysis.

Below assessments will be summarized descriptively.

- Number of nocturnal, treatment-emergent severe or BG confirmed hypoglycaemic episodes during 26 weeks
- Number of treatment-emergent severe or BG confirmed hypoglycaemic episodes during 26 weeks
- Number of nocturnal, treatment-emergent severe or BG confirmed hypoglycaemic episodes during maintenance (week 16-26)
- Number of treatment-emergent severe or BG confirmed hypoglycaemic episodes during maintenance (week 16-26)
- Number of nocturnal, treatment-emergent severe or BG confirmed hypoglycaemic episodes during 38 weeks
- Number of treatment-emergent severe or BG confirmed hypoglycaemic episodes during 38 weeks

Data on treatment-emergent hypoglycaemic episodes are presented 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). Separate summaries are made by severity considering all confirmed hypoglycaemic episodes, nocturnal confirmed hypoglycaemic episodes and the ADA classification of hypoglycaemia.

In addition, to the extent where data allows, the following endpoints are analysed statistically.

During 26 weeks:

- Number of nocturnal, treatment-emergent severe or blood glucose (BG) confirmed symptomatic hypoglycaemic episodes during 26 weeks
- Number of treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 26 weeks
- Number of nocturnal, treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during maintenance (week 16-26)

- Number of treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during maintenance (week 16-26)

During 38 weeks:

- Number of nocturnal, treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 38 weeks
- Number of treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 38 weeks

For endpoints addressing the hypoglycaemic rate in the 26 weeks or the 38 weeks period, a negative binomial model will be used, including treatment, previous insulin treatment regimen, previous OAD treatment, sex, and region as factors and age as covariate. The model will use a log-link function and the logarithm of the time period in which a hypoglycaemic episode was considered treatment emergent as offset. The treatment rate ratios will be estimated and 95% confidence intervals will be calculated.

The rationale for investigating the hypoglycaemic rate during the maintenance is to reflect long term use, as prior to this period (the titration period) the glycaemic control might be unstable. For endpoints addressing hypoglycaemic rate in the maintenance period, subjects that discontinue randomised treatment during the maintenance period will contribute with the available data from the maintenance period. For subjects that discontinue randomised treatment during the 16 week titration period, the number of events in the maintenance period will be imputed based on data from the maintenance period for subjects that discontinued randomised treatment during the maintenance period. For these subjects, the number of events will be imputed corresponding to a 10 weeks maintenance period as follows:

- In the first step, 1000 samples from the posterior distribution of model parameters are extracted. The model is fitted to the on-treatment maintenance data for subjects that discontinued randomised treatment during the maintenance period for each treatment arm separately. This is done using a Bayes negative binomial log-link model with the same factors, covariates and offset as in the analyses of hypoglycaemic episodes described above. The 1000 samples are used as sampled parameter estimates in the next step.
- For each sample of model parameters, the total number of hypoglycaemic events in the maintenance period for subjects that discontinued randomised treatment in the titration period is imputed as a random number of events from a negative binomial distribution using the sampled parameters.
- For each of the 1000 imputed data sets, that now have maintenance period data for all randomised subjects, the mean treatment ratio is estimated using the same negative binomial model as in the analyses of hypoglycaemic episodes described above.

- The estimates and standard deviations for the 1000 imputed data sets are pooled to one estimate and associated standard deviation using Rubin's formula. From these the 95% confidence interval for the treatment ratio and the associated p-value are calculated.

If the above model cannot fit due to sparse data, the model in step 1 is fitted to the on-treatment data for subjects that discontinued randomised treatment during either titration or maintenance, i.e. do not complete the scheduled 26 weeks of treatment. If this model cannot fit either, factors will be left out one by-one in the model using data from all subjects discontinuing randomised treatment before week 26 in the following order, until a model fits:

- sex
- age
- region
- previous insulin treatment regimen
- previous OAD treatment

The analyses and data presentations of hypoglycaemic events will be based on the hypoglycaemic episode form.

Classification of Hypoglycaemia:

Hypoglycaemic episodes are recorded by the subjects in their trial diaries throughout the trial. The information collected includes PG before treating the episode and whether the subject was able treat him/herself. This information is used by Novo Nordisk to classify episodes according to the confirmed hypoglycaemia definition and the ADA definition (severe, documented symptomatic, asymptomatic, probably symptomatic and pseudo).

- Treatment-emergent: hypoglycaemic episodes will be defined as treatment-emergent if the onset of the episode occurs on or after the first day of trial product administration, and no later than 7 calendar days after the last day on trial product

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

- Hypoglycaemic episodes are classified according to the Novo Nordisk classification of hypoglycaemia (see Figure 2–1) and the ADA classification of hypoglycaemia (see Figure 2–2).

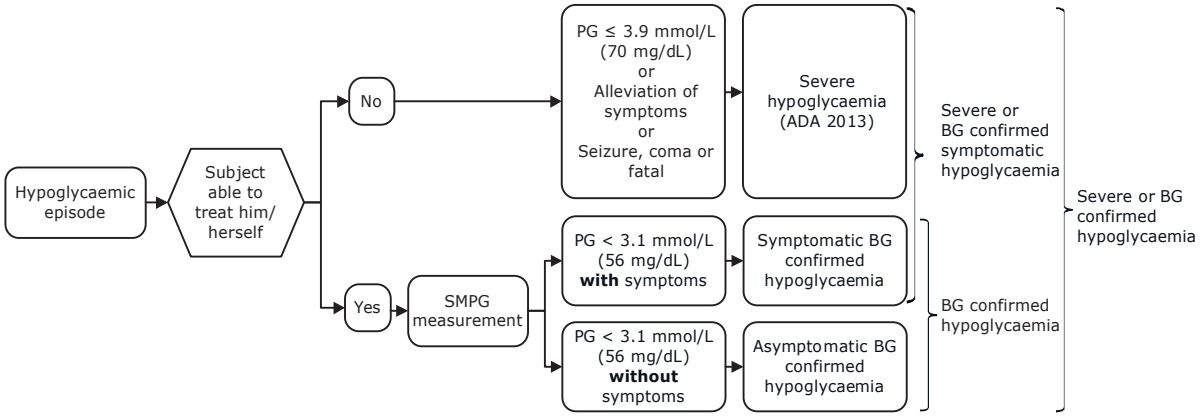
Novo Nordisk classification of hypoglycaemia

In normal physiology, symptoms of hypoglycaemia occur below a plasma glucose level of 3.1 mmol/L (56 mg/dL)[4]. 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 2–1) in addition to the ADA classification:

- Severe hypoglycaemia according to the ADA classification[5]
- Symptomatic BG confirmed hypoglycaemia: An episode that is BG confirmed by plasma glucose value <3.1 mmol/L (56 mg/dL) **with** symptoms consistent with hypoglycaemia.
- Asymptomatic BG confirmed hypoglycaemia: An episode that is BG confirmed by plasma glucose value <3.1 mmol/L (56 mg/dL) **without** symptoms consistent with hypoglycaemia.
- Severe or BG confirmed symptomatic hypoglycaemia: An episode that is severe according to the ADA classification[5] or BG confirmed by a plasma glucose value <3.1 mmol/L (56 mg/dL) **with** symptoms consistent with hypoglycaemia.
- BG confirmed hypoglycaemia: An episode that is BG confirmed by a plasma glucose value <3.1 mmol/L (56 mg/dL) **with** or **without** symptoms consistent with hypoglycaemia.

Severe or BG confirmed hypoglycaemia: An episode that is severe according to the ADA classification[5] or BG confirmed by a plasma glucose value <3.1 mmol/L (56 mg/dL) **with** or **without** symptoms consistent with hypoglycaemia



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

Figure 2–1 Novo Nordisk classification of hypoglycaemia

ADA classification[5] of hypoglycaemia

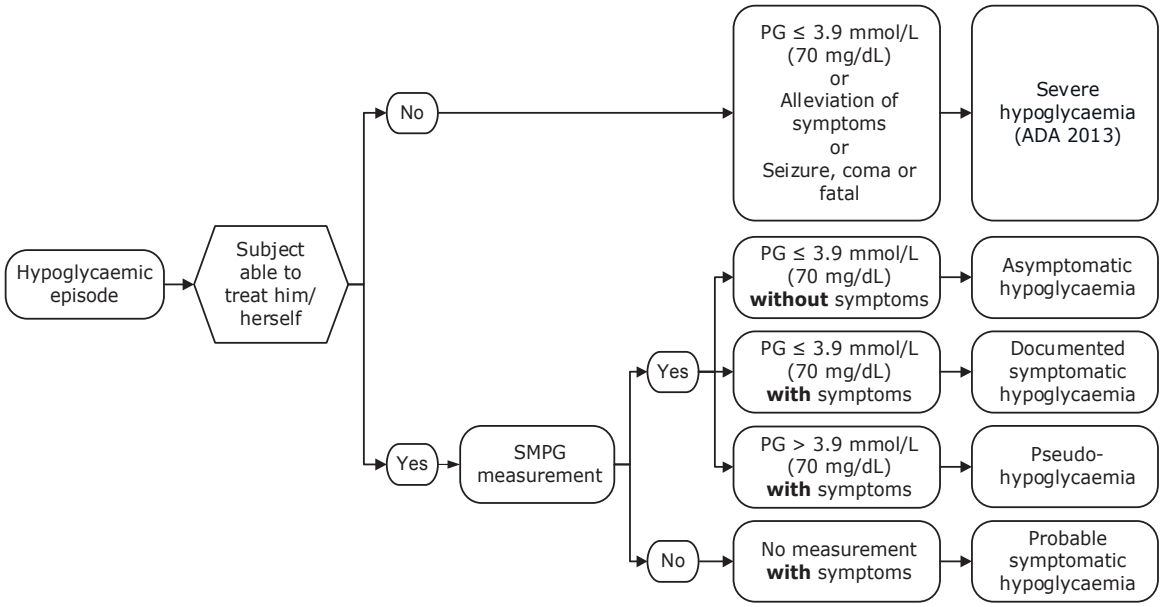
- Severe hypoglycaemia: An episode requiring assistance of another person to actively administer carbohydrate, glucagon, or take other corrective actions. Plasma glucose concentrations may not be available during an event, but neurological recovery following the return of plasma glucose to normal is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

Figure 2–2 ADA classification of hypoglycaemia

Insulin dose

- Total insulin dose at week 26
- Total insulin dose at week 38

Observed values of total insulin dose at week 26 and week 38 will be summarised by the arithmetic mean, SD, median, and minimum and maximum value. Separate summaries will also be made for the basal insulin dose and the bolus insulin dose.

Weight

- Change in weight from baseline to week 26
- Change in weight from baseline to week 38

are summarised descriptively and analysed statistically using the primary statistical analysis for the primary estimand.

The following safety assessments will be reported:

Clinical evaluations at week 26 and week 38

- Physical examination
- Vital signs (blood pressure and pulse)
- Dilated fundoscopy or fundus photography
- 12-lead electrocardiogram

will be summarised descriptively including:

- Change from baseline to week 26
- Change from baseline to week 38

Any findings in the physical examination evaluation at screening will be presented as listings. Any clinically significant deterioration of a pre-existing condition after the screening visit, as well as any new clinically findings will be recorded as adverse events.

Laboratory assessments

- Haematology (haemoglobin, haematocrit, erythrocytes, thrombocytes, leucocytes)
- Biochemistry (creatinine, ALT, AST, AP, sodium, potassium, albumin, total bilirubin)
- Insulin antibodies

All laboratory parameters will be summarised descriptively.

The following tables will be presented based on observed data:

- Shift tables from baseline to week 26
- Shift tables from baseline to week 38
- Change from baseline to week 26
- Change from baseline to week 38
- Proportion of subjects with measurements outside reference range by treatment and week.

For each laboratory parameter, individual values outside the reference range (abnormal values) will be listed.

2.6 Health economics and/or patient reported outcomes

The following patient reported outcome exploratory endpoints are included in the trial:

Change from baseline to week 26 for:

- Health-related quality of life as evaluated by SF-36v2® Health Survey
- Treatment related impact as evaluated by the Treatment Related Impact Measure-Diabetes (TRIM-D)

Change from baseline to week 38 for:

- Health-related quality of life as evaluated by the SF-36v2® Health Survey
- Treatment related impact as evaluated by the TRIM-D

The domain and overall scores for both SF-36 and TRIM-D will be summarised descriptively based on observed data and explorative statistical analysis may be performed.

The data collected in the BHQ will not be included in the clinical trial report, nor will the data collected in the hypoglycaemic resource use interview questionnaire. Data from the latter will be used to assess costs associated with hypoglycaemia and provide information to be used in health economic analyses.

3 Changes to the statistical analyses planned in the protocol

3.1 Clarification of primary statistical analysis, primary estimand:

In the protocol version 4.0 the penalisation of data from subjects on treatment but with missing data at week 26 is only described in the sample size section, not in section 17.4 when describing the analysis in details. This is updated to clarify that also imputed values for subjects in IDegAsp arm on treatment for 26 weeks but with missing HbA_{1c} assessment at week 26 will be penalised.

Rationale:

Alignment of protocol sections.

3.2 Clarification of sensitivity analyses to secondary estimand for primary analysis,

The sensitivity analysis for the secondary estimand specified in the protocol version 4.0 includes subjects that complete the first 26 weeks of the trial on treatment and have on-treatment HbA_{1c} assessment at V28.

Rationale:

This is in alignment with the declared aim for the sensitivity analysis, namely to align with the phase 3a program.

3.3 Changes to supportive secondary endpoints and explorative endpoints

The following endpoints are considered secondary supportive and explorative endpoints, respectively:

Supportive secondary endpoints

Efficacy endpoints:

- Change from baseline in HbA_{1c}(%) after 38 weeks (defined as key supportive secondary efficacy endpoint in protocol)
- Change from baseline in fasting plasma glucose (FPG) after 26 weeks
- Change from baseline in fasting plasma glucose (FPG) after 38 weeks
- Responder (Yes/No) for HbA_{1c} after 26 weeks: $HbA_{1c} < 7\%$ (defined as key supportive secondary efficacy endpoint in protocol)
- Responder (Yes/No) for HbA_{1c} after 38 weeks: $HbA_{1c} < 7\%$ (defined as key supportive secondary efficacy endpoint in protocol)
- Responder (Yes/No) for HbA_{1c} after 26 weeks: $HbA_{1c} < 7\%$ without severe or BG confirmed symptomatic hypoglycaemia
- Responder (Yes/No) for HbA_{1c} after 38 weeks: $HbA_{1c} < 7\%$ without severe or BG confirmed symptomatic hypoglycaemia
- Self-measured plasma glucose measurements (SMPG) after 26 weeks: Mean pre-breakfast measurements used for titration
- Self-measured plasma glucose measurements (SMPG) after 38 weeks: Mean pre-breakfast measurements used for titration
- Self-measured plasma glucose measurements (SMPG) after 26 weeks: 9-point profile - Post-prandial increments
- Self-measured plasma glucose measurements (SMPG) after 38 weeks: 9-point profile - Post-prandial increments

Safety endpoints:

- Incidence of treatment-emergent adverse events (TEAEs) from baseline to week 26
- Incidence of treatment-emergent adverse events (TEAEs) during 38 weeks (defined as key supportive secondary safety endpoint in protocol)
- Incidence of treatment-emergent adverse events (TEAEs) during intensification period (week 26-38)

- Number of nocturnal, treatment-emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 26 weeks
- Number of nocturnal, treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during 38 weeks
- Number of nocturnal, treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during maintenance (week 16-26)
- Number of treatment-emergent severe or blood glucose (BG) confirmed symptomatic hypoglycaemic episodes during 26 weeks (defined as key supportive secondary safety endpoint in protocol)
- Number of treatment-emergent severe or blood glucose (BG) confirmed symptomatic hypoglycaemic episodes during 38 weeks (defined as key supportive secondary safety endpoint in protocol)
- Number of treatment emergent severe or BG confirmed symptomatic hypoglycaemic episodes during maintenance (week 16-26)
- Total insulin dose at week 26
- Total insulin dose at week 38
- Change in weight from baseline to week 26
- Change in weight from baseline to week 38

Explorative endpoints

-
- Change from baseline to week 26 for:
    - Health-related quality of life as evaluated by SF-36v2® Health Survey
    - Treatment related impact as evaluated by the Treatment Related Impact Measure-Diabetes (TRIM-D)
- Change from baseline to week 38 for:
    - Health-related quality of life as evaluated by the SF-36v2® Health Survey
    - Treatment related impact as evaluated by the TRIM-D

Remaining endpoints in protocol version 4.0 are considered assessments.

Rationale:
Clarification on distinction between supportive secondary and exploratory endpoints.

3.4 Statistical analyses of the following assessments will not be performed

Efficacy assessments:

- HbA_{1c} < 7% without nocturnal (00:01-05:59) severe or BG confirmed symptomatic hypoglycaemia after 26 weeks
- HbA_{1c} < 7% without nocturnal (00:01-05:59) severe or BG confirmed symptomatic hypoglycaemia after 38 weeks

Safety assessments:

- Number of nocturnal, treatment-emergent severe or BG confirmed hypoglycaemic episodes during 26 weeks
- Number of treatment-emergent severe or BG confirmed hypoglycaemic episodes during 26 weeks
- Number of nocturnal, treatment-emergent severe or BG confirmed hypoglycaemic episodes during maintenance (week 16-26)
- Number of treatment-emergent severe or BG confirmed hypoglycaemic episodes during maintenance (week 16-26)
- Number of nocturnal, treatment-emergent severe or BG confirmed hypoglycaemic episodes during 38 weeks
- Number of treatment-emergent severe or BG confirmed hypoglycaemic episodes during 38 weeks

All above assessments will be summarized descriptively.

Rationale:

These assessments are supportive and no statistical analyses are needed.

3.5 Statistical analyses based on the secondary estimand will not be performed for the following endpoints

Efficacy endpoints:

- Change from baseline in FPG after 26 weeks
- Change from baseline in FPG after 38 weeks
- Change from baseline in mean pre-breakfast measurements after 26 weeks
- Change from baseline in post-prandial increments after 26 weeks
- Change from baseline in mean pre-breakfast measurements after 38 weeks
- Change from baseline in post-prandial increments after 38 weeks
- Responder (Yes/No) for HbA_{1c}< 7% after 26 weeks
- Responder (Yes/No) for HbA_{1c} < 7% without severe or BG confirmed symptomatic hypoglycaemia

- Responder (Yes/No) for HbA_{1c}< 7% after 38 weeks
- Responder (Yes/No) for HbA_{1c} < 7% without severe or BG confirmed symptomatic hypoglycaemia after 38 weeks

Safety endpoints:

- Change in weight from baseline to week 26
- Change in weight from baseline to week 38

Rationale:

For the above endpoints only the primary estimand are relevant.

3.6 Changes to reporting of safety end points related to adverse events

Below mentioned end point is added:

- Incidence of treatment-emergent adverse events (TEAEs) during intensification period (week 26-38)

Rationale:

Safety when intensifying from OD to BID/TID is of interest for the trial.

The following sentence of protocol 4.0 will be deleted: “The EAC confirmed adverse events will be listed and summarised descriptively.”

Rationale:

Included in protocol by mistake, as no adjudication is done.

3.7 Changes to reporting of insulin dose

The following sentence of protocol 4.0 will be deleted: “and supplemented with the geometric mean and coefficient of variation”

Rationale:

Geometric mean and coefficient of variation for doses is not project standard anymore.

3.8 Change to reporting of laboratory assessments

The following sentence of protocol 4.0 will be deleted: “Laboratory values will be presented graphically as box plots by treatment and week.”

Rationale:

Empirical distribution plots are evaluated as a better graphical presentation of laboratory data.

3.9 In-trial period and on-treatment period

The definition of in-trial period will be replaced by:"This observation period will include information collected at or after date of randomisation and up to the last scheduled subject-site contact. For subjects who withdraw their informed consent, or is lost to follow-up, the end of the in-trial period is defined as the date of the last scheduled subject-investigator contact (site or phone visit). In case a subject dies during the trial, the date of death will be the end-date of the in-trial observation period regardless of the above defined end-dates. This observation period will be used for estimating the effectiveness estimand."

The definition of On-treatment period will be replaced by: "This observation period is a subset of the in-trial observation period and represents the time period in which a subject is considered exposed to trial product. The period starts on day of the first dose of trial product and ends 7 days after last dose of trial product is taken."

Rationale:

Alignment to current project standards.

3.10 Update of completer analyses set

The definition of completer analyses set (CAS) will be replaced by:

- "Trial completer analysis set: Includes all randomised subjects who has attended either V40 on treatment, or V40A, 38 weeks after randomisation according to protocol ."
- Treatment completer analysis set: Includes all randomised subjects who have completed the trial without discontinuation from randomised treatment. "

Rationale:

The protocol does not give a clear definition for trial completers.

4 References

1. Russell-Jones D, Vaag A, Schmitz O, Sethi BK, Lalic N, Antic S, et al. Liraglutide vs insulin glargine and placebo in combination with metformin and sulfonylurea therapy in type 2 diabetes mellitus (LEAD-5 met+SU): a randomised controlled trial. Diabetologia. 2009;52(10):2046-55.

2. International Conference of Harmonisation. ICH Harmonised Tripartite Guideline E9. Statistical Principles for Clinical Trials, Step 4 Version. 5 Feb 1998.

3. Koch GG. Comments on 'Current issues in non-inferiority trials' by Thomas R. Fleming, Statistics in Medicine, DOI: 10.1002/sim.2855. Stat Med. 2008;27(3):333-42.

4. 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.

5. 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.