Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT02906930
Sponsor trial ID:	NN9924-4233
Official title of study:	Efficacy and safety of oral semaglutide versus placebo in subjects with type 2 diabetes mellitus treated with diet and exercise only
Document date:	20 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: NN9924-4233

PIONEER 1 – Monotherapy

Efficacy and safety of oral semaglutide versus placebo in subjects with type 2 diabetes mellitus treated with diet and exercise only

Author:

[redacted]

Table of contents

Page

List of abbreviations

AACE	American Association of Clinical Endocrinologists
ADA	American Diabetes Association
AE	adverse event
ANCOVA	analysis of covariance
ATC	anatomical therapeutic chemical
BG	blood glucose
BMI	body mass index
CI	confidence interval
CRF	case report form
CRP	c-reactive protein
CTR	clinical trial report
EAC	event adjudication committee
ECG	electrocardiogram
EMA	European Medicines Agency
FAS	full analysis set
FPG	fasting plasma glucose
GLP-1	glucagon-like peptide-1
HbA_{1c}	glycosylated haemoglobin
HDL	high-density lipoprotein
HRQoL	health-related quality of life
IWQOL-Lite-CT	Impact of Weight on Quality of Life Clinical Trials
IWRS	interactive web response system
LDL	low-density lipoprotein
LLoQ	lower limit of quantification
LOCF	last observation carried forward
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MMRM	mixed model for repeated measurements
PG	plasma glucose
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Status
PK	pharmacokinetics
PRO	patient reported outcomes
SAP	statistical analysis plan
SAS	safety analysis set

s.c.	subcutaneous
SF-36v2	Short Form 36 version 2 Health Survey®
SMPG	self-measured plasma glucose
TE	treatment effect
TEAE	treatment-emergent adverse events

1 Introduction

1.1 Trial information

This is a 26-week, randomised, double-blinded, placebo-controlled, four-armed, parallel-group, multi-centre, multi-national trial comparing the efficacy and safety of three dose levels of once-daily oral semaglutide and placebo in subjects with type 2 diabetes mellitus treated with diet and exercise only.

Primary objective
To compare the effects of three dose levels of once-daily oral semaglutide (3, 7 and 14 mg) versus once-daily placebo on glycaemic control in subjects with type 2 diabetes mellitus treated with diet and exercise only.

Secondary objectives
To compare the effects of three dose levels of once-daily oral semaglutide (3, 7 and 14 mg) versus once-daily placebo on body weight in subjects with type 2 diabetes mellitus treated with diet and exercise only.

To compare the safety and tolerability of three dose levels of once-daily oral semaglutide (3, 7 and 14 mg) versus once-daily placebo in subjects with type 2 diabetes mellitus treated with diet and exercise only.

Trial design
Subjects with type 2 diabetes treated with diet and exercise only will after approximately 2 weeks screening period be randomised in a 1:1:1:1 manner to receive either a dose of 3, 7, or 14 mg of oral semaglutide or placebo once daily. After the treatment period of 26 weeks, all subjects enter a follow-up period of 5 weeks ended by a follow-up visit. The total trial duration for the individual subject will be approximately 33 weeks. For further details, see the trial protocol.

1.2 Scope of the statistical analysis plan

This statistical analysis plan (SAP) is based on the protocol for trial NN9924-4233 "Efficacy and safety of oral semaglutide versus placebo in subjects with type 2 diabetes mellitus treated with diet and exercise only", version 3.0 (18 November 2016), and includes more detailed procedures for executing the statistical analyses of the primary and secondary endpoints. Statistical analyses and a number of clarifications additional to those specified in the trial protocol are pre-specified with this SAP. All changes to the statistical analyses planned in the trial protocol are documented in section 3.

Novo Nordisk will be responsible for the statistical analyses and reporting.

2 Statistical considerations

General considerations

The blinding of the randomised treatments will be maintained until the database has been released for statistical analysis.

Data from all sites will be analysed and reported together.

1. The randomisation is stratified based on descent (Japanese subjects/non-Japanese subjects). The information regarding descent will be included based on country details from the IWRS. Descent (Japanese subjects/non-Japanese subjects) will be included in the statistical analyses as part of region. The regions are: Europe (Bulgaria, Czech Republic, Russian Federation, Serbia, and Turkey), North America (United States), South America (Mexico), Africa (Algeria), and Asia (Japan).

The latest available measurement, at or prior to the randomisation visit, will be used as the baseline measurement. If no measurement(s) have been obtained, at or prior to randomisation, the baseline value will be left missing.

Laboratory values below the lower limit of quantification (LLoQ) will be set to ½LLoQ. Number of values below LLoQ by treatment and visit will be summarised if deemed relevant.

Results from a statistical analysis will as a minimum be presented by the estimated treatment contrasts for the below three comparisons with associated two-sided 95% confidence intervals and p-values corresponding to two-sided tests of no difference:

- Oral semaglutide 14 mg vs. placebo
- Oral semaglutide 7 mg vs. placebo
- Oral semaglutide 3 mg vs. placebo

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

Primary and secondary estimands

Two estimands addressing different aspects of the primary trial objective will be defined as follow:

- Primary estimand – 'Treatment policy'
  - treatment difference (oral semaglutide versus placebo) at week 26 for all randomised subjects regardless of adherence to randomised treatment and initiation of rescue medication

The treatment policy estimand assesses the expected glycaemic benefit in a future population that results from subjects initiating treatment with oral semaglutide including potential rescue

medication(s). Generalisation of this estimand depends among other things on the extent to which the use of rescue medication in this trial reflects clinical practice and the treatment adherence reflects the behaviour of the target population. Accordingly, data collected regardless of discontinuation of trial product or initiation of rescue medication(s) will be used to draw inference.

- Secondary estimand – 'Hypothetical'
  - treatment difference (oral semaglutide versus placebo) at week 26 for all randomised subjects if all subjects adhered to treatment and did not initiate rescue medication

The hypothetical estimand assesses the glycaemic benefit a future subject is expected to achieve if initiating and continuing treatment with oral semaglutide. It is considered a clinically relevant estimand as it provides information to treating clinicians about the expected glycaemic efficacy of oral semaglutide for purposes of treating individual subjects. Generalisation of this estimand depends among other things on the extent to which the adherence to trial product administration in this trial reflects the behaviour of the target population. 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.

Analogously, two estimands will be pre-defined for the remaining secondary endpoints addressing the secondary objective.

Missing data considerations at week 26

When estimating the primary estimand, the proportion of missing data, i.e., data that do not exist even though subjects are intended to stay in the trial regardless of treatment status and initiation of rescue medication, is expected to be maximum 10% based on the oral semaglutide phase 2 trial (NN9924-3790). Thus, missing data will mainly be due to withdrawal from trial or lost to follow-up.

When estimating the secondary estimand, the proportion of missing data is expected to be higher (20%) since data collected after discontinuation of trial product or initiation of rescue medication(s) will be set to missing. The 20% of missing data is based on the oral semaglutide phase 2 trial (NN9924-3790), that indicates that a low starting dose with gradual dose escalation diminishes gastrointestinal AEs compared with more aggressive dosing regimens. Across treatment arms the main reasons for missing data are expected to be early treatment discontinuation due to gastrointestinal AEs and potential initiation of rescue medication. Initiation of rescue medication is expected to be more frequent in the placebo arm and in the oral semaglutide 3 mg arm than for the two highest dose levels of oral semaglutide. A higher proportion of subjects are expected to discontinue treatment due to AEs in the oral semaglutide 14 mg arm, compared to the other treatment arms. So overall the frequency of missing data is expected to be similar across treatment arms.

Descriptive summaries and graphical representation of extent, reason(s) for and pattern of missing data will be presented by treatment arm.

2.1 Sample size calculation

Both the primary endpoint, change from baseline to week 26 in HbA_{1c} and the confirmatory secondary endpoint, change from baseline to week 26 in body weight are planned to be tested for superiority of oral semaglutide vs. placebo at each dose level (3 mg, 7 mg, and 14 mg).

The sample size calculation is made to ensure a power of at least 90% to jointly confirm HbA_{1c} superiority of oral semaglutide vs. placebo at each dose level out of the six pre-specified confirmatory hypotheses shown in Figure 2–1. The closed testing procedure described in Bretz et al.[1] is used to control the overall type I error at a nominal two-sided 5% level. The statistical testing strategy is built on the following two principles:

- Within a dose level, glycaemic effect must be established in terms of HbA_{1c} superiority before testing for added benefits in terms of body weight superiority.
- Glycaemic effect in terms of HbA_{1c} superiority must be established on all higher dose levels before continuing testing hypotheses on lower dose levels.

The sample size is calculated using the calcPower function in the R package, gMCP[2] using 10000 simulations. All of the six pre-specified confirmatory tests are assumed to be independent. Since some of the tests are positively correlated, the assumption of independence is viewed as conservative.

The sample size assumptions for treatment effects (TE), adjusted treatment effects and the common standard deviations used across dose levels are given in Table 2–1. These are primarily based on the oral semaglutide phase 2 results (NN9924-3790) and supported by results from the s.c. semaglutide phase 2 trial (NN9535-1821).

Since the equalising effect of rescue medication will be included in the primary analysis as well as a conservative approach for handling of missing data will be performed, an adjustment in treatment effect will be implemented for the 10% of subjects who are expected to discontinue trial product or initiate rescue medication and for the 10% of subjects who are expected to have actual missing data. The treatment effects used in the sample size calculation will be adjusted according to a 75% smaller effect in these subjects. The adjusted treatment effects for testing superiority are defined as $0.8 \times TE + 0.2 \times TE \times 0.25$.

Table 2–1 Assumptions used in the sample size calculation

Oral semaglutide vs. placebo	HbA_{1c} (%-point)			Body weight (kg)		
Treatment dose	14 mg	7 mg	3 mg	14 mg	7 mg	3 mg
Treatment effect (TE)	-1.0	-0.75	-0.45	-3.0	-2.0	-1.0
Adjusted TE, superiority	-0.85	-0.6375	-0.3825	-2.55	-1.70	-0.85
Standard deviation	1.1	1.1	1.1	4.0	4.0	4.0

With the above assumptions, allocating 176 subjects to each of the oral semaglutide arms and the placebo arm provides 90% power to confirm HbA_{1c} superiority of oral semaglutide vs. placebo at all dose levels. Calculated powers for individual hypotheses are presented in Table 2–2. In total $4 \times 176 = 704$ subjects are planned to be randomised.

Table 2–2 Calculated powers for individual hypotheses

Statistical test	HbA_{1c} superiority			Body weight superiority		
Treatment dose	14 mg	7 mg	3 mg	14 mg	7 mg	3 mg
Power (%)	> 99%	> 99%	90%	> 99%	96%	46%

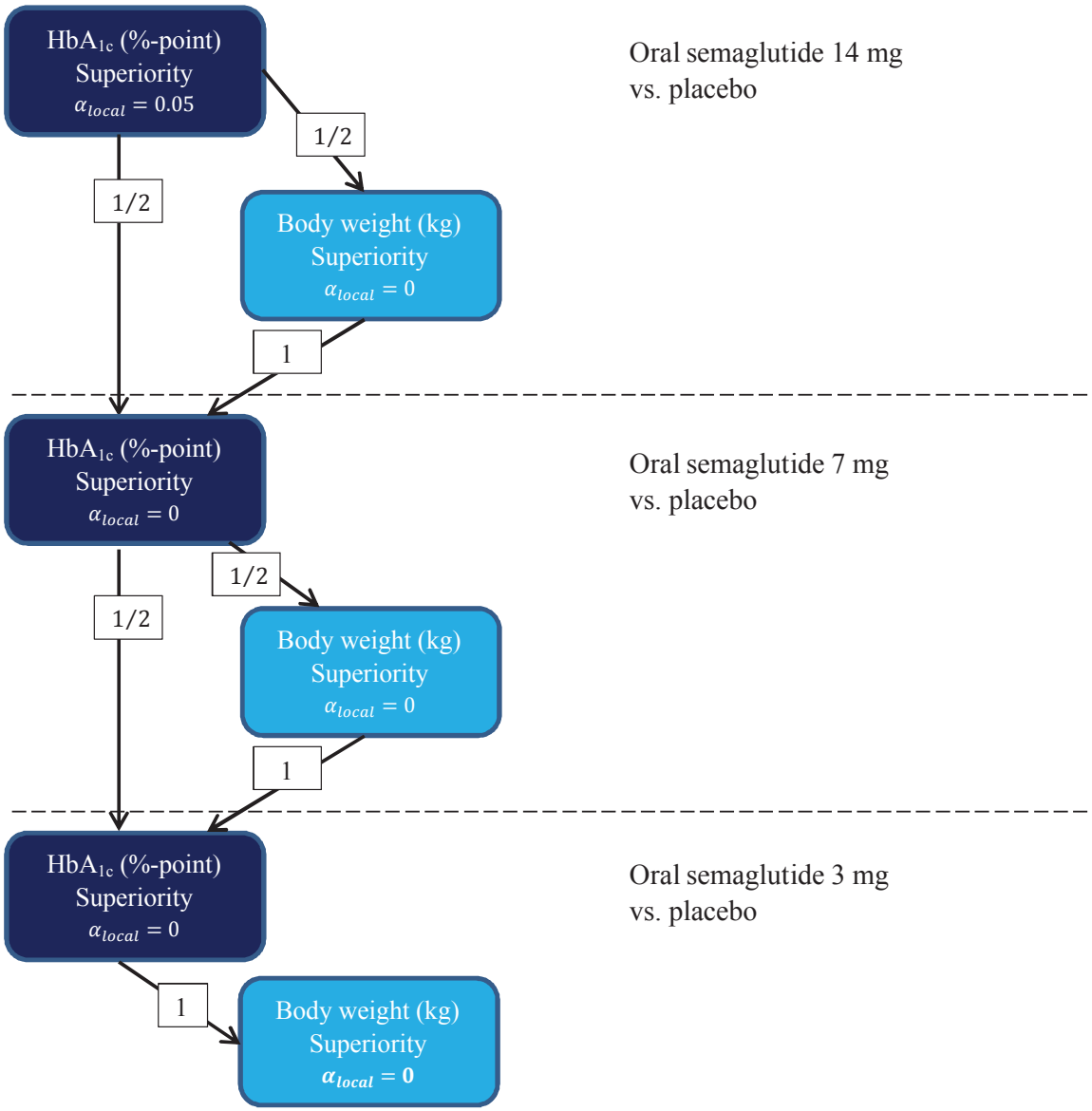


Figure 2–1 Graphical illustration of the closed testing procedure

The overall significance level of α = 0.05 (two-sided) is initially allocated to the HbA_{1c} superiority test on the highest dose level. The local significance level (α-local) will be reallocated if a hypothesis is confirmed according to the weight given by the directed edges between nodes (hypotheses). The sample size is based on the hypotheses in the dark boxes.

2.2 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 the 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 where they were on treatment. This will be referred to as contributing to the evaluation "as treated".

Data selections and observation periods

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 (V9) for subjects on trial product
- The latest occurring visit of the end-of-treatment visit (V8) or the follow-up premature discontinuation visit (V9A), for subjects who have discontinued trial product prematurely

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

Definition of the observation periods:

In-trial: This observation period represents the time period 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
- 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

On-treatment: This observation period represents the time period where subjects are considered treated with the trial product. The observation period is a subset 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. For adjudicated events, ECGs, eye examination category, anti-

semaglutide antibodies, and AEs including hypoglycaemic episodes, the observation period ends at the first date of any of the following:

- The follow-up visit (V9)
- The follow-up prematurely discontinuation visit (V9A)
- The last date on trial product + 38 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 oral semaglutide. The visit window for the follow-up visit is + 3 days.

For efficacy and other safety assessments (laboratory assessments, physical examination and vital signs) the observation period ends at the last date on trial product + 3 days. This will be used in order to ensure specificity to reversible effects of treatment.

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

- The last dose of trial product + 3 days
- Initiation of rescue medication

The in-trial observation period will be the primary observation period for estimating the primary estimand. The on-treatment without rescue medication observation period will be the primary observation period when estimating the secondary estimand. Safety will be evaluated based on the in-trial and the on-treatment observation periods.

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. For adjudicated events, the onset date will be the event adjudication committee (EAC) adjudicated onset date.

Before data are locked for statistical analysis and the randomisation code is broken, 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 should 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 must 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 (CTR).

Confirmatory hypotheses

For the primary HbA_{1c} endpoint and the confirmatory secondary body weight endpoint the following confirmatory one-sided hypotheses are planned to be tested at each dose level of oral semaglutide versus placebo. Let the mean treatment difference be defined as μ = (oral semaglutide minus placebo):

- HbA_{1c} superiority
  - H_0: $\mu \geq 0.0$ %-point against H_A: $\mu < 0.0$ %-point

- Body weight superiority
  - H_0: $\mu \geq 0.0$ kg against H_A: $\mu < 0.0$ kg

Operationally the hypotheses will be evaluated by two-sided tests at the 5% significance level.

Multiplicity and criteria for confirming hypotheses

The type I error for testing the six confirmatory hypotheses related to the HbA_{1c} and body weight endpoints will be preserved in the strong sense at 5% (two-sided) using the weighted Bonferroni-based closed testing procedure described in Bretz et al.[1] and outlined in Figure 2–1.

The first hypothesis to be tested is superiority of HbA_{1c} at the highest dose level. It will be tested at the overall significance level (5%) while allocating 0% local significance level to the remaining of the hypotheses. For this hypothesis, and in general, if a hypothesis is confirmed, then the significance level 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 updated local significance level (α-local). This process will be repeated until no further hypotheses can be confirmed.

Superiority will be considered confirmed if the mean treatment difference is supporting the corresponding alternative hypothesis and the two-sided p-value from the primary analysis of the primary estimand is strictly below its local two-sided significance level as defined by the closed testing procedure in Figure 2–1. This is equivalent to using a one-sided p-value (nominal $\alpha = 0.025$) and a one-sided 2.5% overall significance level in the closed testing procedure.

2.3 Primary endpoint

The primary endpoint is change from baseline to week 26 in HbA_{1c}.

2.3.1 Primary analysis for the primary estimand

The primary estimand will be estimated based on the FAS using week 26 measurements from the in-trial observation period. The primary statistical analysis will be a pattern mixture model using multiple imputation to handle missing data assuming that the missing data mechanism is missing at

random (MAR) within the groups used for imputation. Imputation of missing data at week 26 will be done within 8 groups of subjects defined by randomised treatment arm, and whether subjects at week 26: (i) have discontinued treatment or initiated rescue medication or; (ii) are still on treatment and have not initiated rescue medication. 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 26 are similar in terms of randomised treatment arm and treatment adherence/rescue medication status.

Missing values for each group will be imputed as follows:

- An analysis of covariance (ANCOVA) with region as a categorical fixed effect and baseline HbA_{1c} measurement as a covariate will be fitted to observed values of the change from baseline in HbA_{1c} at week 26.
- The estimated parameters for location and dispersion will be used to impute 1000 values for each subject with missing week 26 data based on region and baseline HbA_{1c}. Thus, 1000 complete data sets will be generated including observed and imputed values.

The number of subjects in the groups who at week 26 have discontinued trial product or initiated rescue medication are expected to be relatively low. Therefore the region variable included in the imputation model will be reduced in levels avoiding estimation problems due to sparse data. The regions to be used in these imputations are defined as North America and Other.

Analysis used for confirming superiority versus placebo at week 26:

For each of the 1000 (now complete) imputed data sets the change from baseline to week 26 in HbA_{1c} will be analysed using an ANCOVA with treatment and region as categorical fixed effects and baseline HbA_{1c} as covariate. The results obtained from analysing the datasets will be combined using Rubin's rule[3] to draw inference.

From this analysis the estimated treatment differences between each of the oral semaglutide dose levels and placebo together with associated two-sided 95% confidence intervals and unadjusted two-sided p-values for testing no difference from zero will be presented.

2.3.2 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 without rescue medication observation period. The primary analysis for the secondary estimand will be a Mixed Model for Repeated Measurements (MMRM). A restricted maximum likelihood will be used. The model will include all post-baseline HbA_{1c} measurements collected at scheduled visits up to and including week 26 as dependent variables. The independent effects included in the model will be treatment and region as categorical fixed effects and baseline HbA_{1c} as a covariate, 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 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 same as the observed data.

For subjects who do not have post-baseline assessments for planned visits available in the on-treatment without rescue medication period, the baseline value will be carried forward to the first planned visit to ensure that all randomised subjects will contribute to the statistical analysis.

2.3.3 Sensitivity analyses

To investigate the sensitivity of the primary analysis results, complementary and separate analyses will be performed for the primary and secondary estimand. In line with EMA recommendations[4] and with a report from the US National Research Council[5], these analyses will evaluate the sensitivity of the results due to the impact of missing data.

The evaluation of the robustness of the primary analysis results will be based on a pattern mixture model approach using multiple imputation. An overview of the sensitivity analyses for each of the estimands are specified below followed by a more detailed description of the three different pattern mixture models used.

Sensitivity analyses for the primary estimand

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

- A comparator multiple imputation analysis based on FAS using the in-trial observation period
- A comparator multiple imputation analysis differentiating between reasons for discontinuing treatment prematurely based on FAS using the in-trial observation period
- A tipping-point multiple imputation analysis based on FAS using the in-trial observation period

Sensitivity analyses for the secondary estimand

The estimation of the secondary estimand will be repeated using the following sensitivity analyses:

- A tipping-point multiple imputation analysis based on FAS using the on-treatment without rescue medication observation period.

2.3.3.1 Pattern mixture models

Common for the three pattern mixture model sensitivity analyses is that they all aim to stress-test the primary HbA_{1c} results by changing the assumptions for part or all missing data in the oral

semaglutide treatment arms, while maintaining the missing at random data assumption for the placebo arm:

- *Comparator multiple imputation analysis:* In this sensitivity analysis missing data at week 26 for all subjects will be imputed to resemble the distribution of the week 26 values observed in the placebo treatment arm. In effect this imputation approach removes the treatment difference between oral semaglutide and placebo for all subjects randomised to oral semaglutide, given that oral semaglutide is better than placebo.
- *Comparator multiple imputation analysis differentiating between reasons for discontinuing treatment prematurely:* In this sensitivity analysis only missing data at week 26 for subjects who discontinue oral semaglutide treatment due to treatment related AEs will be imputed to resemble the distribution of the week 26 values observed in the placebo treatment arm. Treatment related AEs are defined as AEs classified as possible or probable related to trial product as reported by the investigator. In effect this imputation approach removes the treatment difference between oral semaglutide and placebo for this selected group of subjects randomised to oral semaglutide.
- *Tipping-point multiple imputation analysis:* In this sensitivity analysis, missing data will first be imputed according to the primary analysis for the treatment policy estimand. For the hypothetical estimand imputation will be done as described below for the binary endpoints (see section 2.4.2.1). Secondly, for all oral semaglutide treatment arms a penalty will be added to the imputed values at week 26. The approach is to gradually increase this penalty until all confirmed HbA_{1c} conclusions from the primary analysis are changed. For each hypothesis tested the specific value of the penalty that reverses the conclusion will be used to evaluate the robustness of the primary analysis results.

2.3.3.2 Assessment of sensitivity analyses

The results from the sensitivity analyses will be collectively used to interpret the robustness of the trial results for HbA_{1c}. Due to the sensitivity analyses inherent conservative nature, it will not be a requirement that all confirmatory hypotheses are consistently confirmed across the sensitivity analyses. Thus, no absolute success criteria will be pre-defined for each sensitivity analysis. The sensitivity results in totality will be used to substantiate the credibility of the trial results.

2.4 Secondary endpoints

2.4.1 Confirmatory secondary endpoints

Change from baseline to week 26 in body weight (kg) will be a confirmatory secondary endpoint.

The primary and secondary estimands will be estimated using the same approaches as described for the primary HbA_{1c} endpoint. Baseline body weight will be used as a covariate instead of baseline HbA_{1c} in both the multiple imputation and MMRM analysis models.

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 the primary estimand is strictly below its local two-sided significance level resulting from the closed testing procedure in Figure 2–1. Sensitivity analyses similar to the ones pre-specified for the primary HbA_{1c} endpoint will be made to evaluate the robustness of the body weight results.

2.4.2 Supportive secondary endpoints

2.4.2.1 Efficacy endpoints

The below supportive secondary efficacy endpoints will be evaluated for:

- The primary estimand based on FAS using the in-trial observation period
- The secondary estimand based on FAS using the on-treatment without rescue medication observation period

No sensitivity analyses are planned for these.

Continuous efficacy endpoints

Change from baseline to week 26 in:

- FPG
- Fasting C-peptide
- Fasting insulin and proinsulin
- Fasting glucagon
- Insulin resistance (homeostatic model assessment index of insulin resistance (HOMA-IR)) and beta-cell function (homeostatic model assessment index of beta-cell function (HOMA-B))
- CRP
- Body weight (%)
- BMI
- Waist circumference
- Fasting lipid profile (total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides)

BMI will be calculated based on body weight and height based on the formulae:
BMI kg/m^2 = body weight (kg)/(height (m) × height (m)) or (kg/m^2 = [lb/in^2 × 703])

Change from baseline to week 26 in the below derived endpoints from the SMPG 7-point profile:

- Mean 7-point profile. Defined as the area under the profile, calculated using the trapezoidal method, divided by the measurement time
- Mean postprandial increment (over all meals)

The above continuous endpoints will be analysed separately using similar model approaches as for the primary endpoint with the associated baseline value as a covariate. The following endpoints will be log-transformed prior to analysis with the associated log-transformed baseline value as a covariate: Fasting C-peptide, fasting insulin and proinsulin, fasting glucagon, insulin resistance, CRP, and fasting lipid profile.

For evaluation of the primary estimand the analyses will be performed at week 26. This will result in imputation of missing data within 8 groups as described for the week 26 evaluation in section 2.3.1.

For evaluation of the secondary estimand the MMRM based primary analysis will include all scheduled post-baseline measurement up to and including week 26. From this model the estimated treatment differences (ratios) will be presented at week 26 with 95% confidence intervals and two-sided p-values for test of no difference. The baseline will not be carried forward to first planned visit if the first planned visit falls later than 8 weeks after randomisation.

Binary efficacy endpoints

If a subject after week 26 achieves (yes/no):

- $HbA_{1c} < 7.0$ % (53 mmol/mol) (ADA target)
- $HbA_{1c} \leq 6.5$ % (48 mmol/mol) (AACE target)
- Body weight loss ≥ 5 %
- Body weight loss ≥ 10 %
- $HbA_{1c} < 7.0$ % (53 mmol/mol) without hypoglycaemia (treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes) and no weight gain
- HbA_{1c} reduction ≥ 1 %-point (10.9 mmol/mol) and weight loss ≥ 3 %

Missing data for the above six binary endpoints will be accounted for using multiple imputation techniques. For the treatment policy estimand the binary endpoints will be calculated as dichotomisations of the 1000 multiple imputations underlying the primary MI analysis. For the hypothetical estimand the model will be implemented using a sequential imputation approach assuming MAR. The imputation will be done as described below:

- Intermittent missing values in the on-treatment without rescue observation period 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 group separately and 1000 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 planned end of treatment visit. For each treatment group an analysis of covariance model

will be used to impute missing values at each planned visit. The model will include region as categorical effects and baseline and post-baseline values prior to the visit in question as covariates.

The binary endpoints will be derived as dichotomisations of the 1000 multiple imputations from the sequential imputation.

For both estimands, each of the 1000 datasets will be analysed using a logistic regression model with treatment and region as fixed effects and baseline value as covariate (i.e. baseline HbA_{1c} for binary HbA_{1c} endpoints, baseline body weight for body weight endpoints and both baseline HbA_{1c} and baseline body weight for the composite binary endpoints that comprise both parameters). The results will be combined using Rubin's rule to draw inference.

Only observed data within the corresponding observation period will be included for the 'without hypoglycaemia' component of the composite endpoint. Because the number of hypoglycaemic episodes are expected to be very low in this trial, the observed data is considered sufficient when addressing both estimands.

For the composite endpoints involving both HbA_{1c} and body weight the imputed datasets will be combined by imputation number.

Time to event endpoints

- Time to additional anti-diabetic medication (to support the treatment policy estimand)
- Time to rescue medication (to support the hypothetical estimand)

Definition of additional anti-diabetic medication: New anti-diabetic medication initiated at or after randomisation and before (planned) end-of-treatment.

Definition of rescue medication: New anti-diabetic medication initiated at or after randomisation and before last date on trial product. This is a subset of the additional anti-diabetic medication.

The following rule will be applied based on the concomitant medication data reported by the investigator, to determine whether or not the recorded anti-diabetic medication can be considered an new anti-diabetic medication:

1. **New anti-diabetic medication:** Anti-diabetic medication (4^{th} level ATC code) that is initiated at or after randomisation and with a dosing duration of more than 21 days

More than 21 days are chosen as a minimum duration for the medication to be considered as 'anti-diabetic medication'. This threshold is set to ensure that short-term durations (i.e., $\leq$ 21 days) of anti-diabetic medication (e.g., in connection with concurrent illnesses) are not included because such intensifications are not likely to affect the effect endpoints.

The analysis supporting the treatment policy estimand is addressed for the FAS using the in-trial observation period and additional anti-diabetic medication will be considered an event regardless of whether or not subjects prematurely discontinued treatment. Time from randomisation to additional anti-diabetic medication will be analysed using a Cox proportional hazards model with treatment and region as categorical fixed effects and the baseline HbA_{1c} value as a covariate. From this analysis the estimated hazard ratios between each of the oral semaglutide dose levels and placebo together with associated two-sided 95% CIs and unadjusted two-sided p-values will be presented. The endpoint aims to address the need of additional anti-diabetic medication regardless of whether this is due to lack of effect or related to tolerability of the trial product. Switching to another anti-diabetic treatment is therefore also considered an event and withdrawn subjects or subject lost to follow-up will be considered as having an event (started on additional anti-diabetic medication) on the day of withdrawal. Subjects will be censored on the day before the planned end-of-treatment visit.

The analysis supporting the hypothetical estimand is addressed for the FAS using the on-treatment without rescue medication observation period. Time from first dose of trial product to initiation of rescue medication will be analysed using the same model as described above. The endpoint aims to address a lack of effect of treatment with trial product. Only initiation of rescue medication as add-on to randomised treatment is considered an event; switching to another anti-diabetic treatment is not considered an event (initiation of rescue medication) and as a consequence subjects will be censored on the day before date of last trial product. Potential events occurring between randomisation and first date on trial product will be included in the analysis as events on first date of trial product, in order to account for all events of rescue medication initiation.

2.4.2.2 Safety endpoints and safety assessments

The safety endpoints and safety assessments will be evaluated based on SAS using the on-treatment observation period and based on SAS using the in-trial observation period unless otherwise stated. The following endpoints and assessments are used to support the safety objectives.

Adverse events

- Number of treatment-emergent adverse events (TEAEs) during exposure to trial product, assessed up to approximately 31 weeks

All AEs will be coded using version 20.1 of the Medical Dictionary for Regulatory Activities (MedDRA) coding.

A treatment-emergent AE is defined as an AE with onset in the on-treatment observation period (see definition of observation periods in section 2.2).

TEAEs will be summarised 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 patient years of observation time (R) for the on-treatment observation period. Supportive summaries of AEs will be made for the in-trial observation period. The development over time in gastrointestinal AEs will be presented graphically.

Other safety endpoints

Change from baseline to week 26 in:

- Amylase
- Lipase
- Pulse
- Systolic blood pressure
- Diastolic blood pressure

The above safety endpoints will be evaluated using the primary analysis for the primary estimand based on SAS using the in-trial observation period and using the primary analysis for the secondary estimand based on SAS using the on-treatment observation period. Endpoints will be analysed separately as described above for continuous efficacy endpoints. Results will be presented at week 26. Amylase and lipase endpoints will be log-transformed prior to analysis with the associated log-transformed baseline value as a covariate.

Change from baseline to week 26 in:

- ECG evaluation
- Physical examination
- Eye examination category

Any occurrence of anti-semaglutide antibodies (yes/no) up to approximately 31 weeks:

- Anti-semaglutide binding antibodies
- Anti-semaglutide neutralising antibodies
- Anti-semaglutide binding antibodies cross reacting with native GLP-1
- Anti-semaglutide neutralising antibodies cross reacting with native GLP-1

Anti-semaglutide binding antibodies up to approximately 31 weeks:

- Anti-semaglutide binding antibody levels

Other safety assessments

Change from baseline to week 26 in:

- Haematology
- Biochemistry (except for amylase and lipase)

- Calcitonin

Change from pre-dose to post-dose (25 and 40 minutes) at week 4 and 26 in:

- Lactate

The above safety endpoints and assessments will be summarised descriptively by treatment arm and visit. Categorical safety and assessments endpoints will be summarised as counts and relative frequencies. Calcitonin will also be presented by gender.

Hypoglycaemia endpoints

- Number of treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed up to approximately 31 weeks
- Treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed up to approximately 31 weeks (yes/no)

Classification of hypoglycaemia

Hypoglycaemic episodes will be summarised for the SAS and the on-treatment observation period only.

Treatment-emergent: hypoglycaemic episodes will be defined as treatment-emergent if the onset of the episode occurs within the on-treatment observation period (see definition of observation periods in section 2.2).

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 and the ADA classification of hypoglycaemia (see Figure 2–2).

Novo Nordisk classification of hypoglycaemia

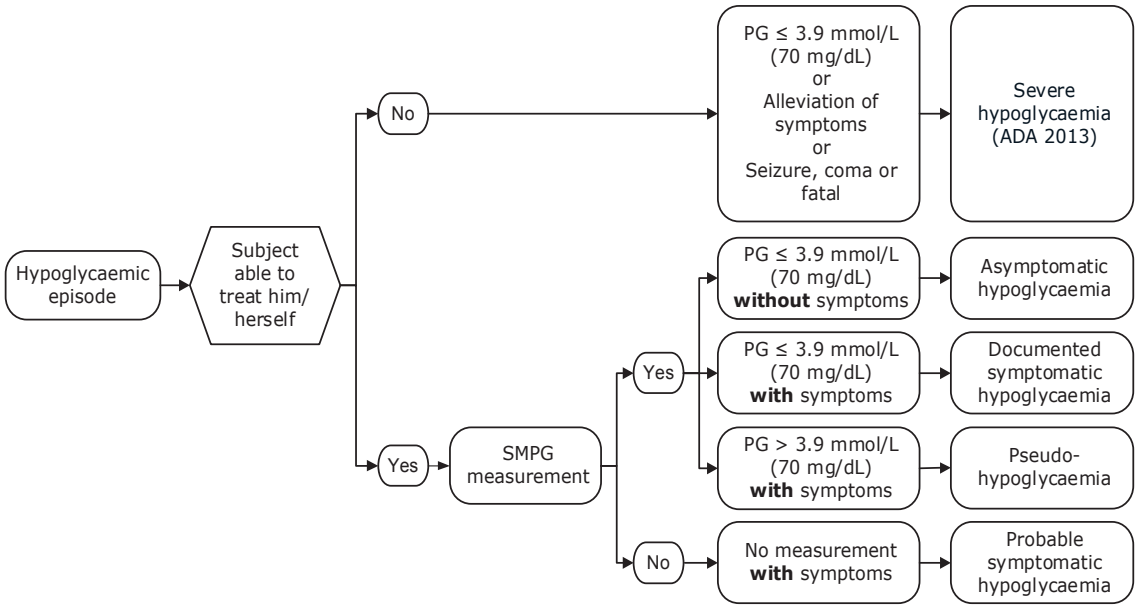
In normal physiology, symptoms of hypoglycaemia occur below a PG level of 3.1 mmol/L (56 mg/dL)[6]. Therefore, Novo Nordisk has included hypoglycaemia with PG levels below this cut-off point in the definition of BG-confirmed hypoglycaemia.

Novo Nordisk uses the following classification in addition to the ADA classification:

- Severe or BG-confirmed symptomatic hypoglycaemia: An episode that is severe according to the ADA classification[7] or BG-confirmed by a PG value < 3.1 mmol/L (56 mg/dL) **with** symptoms consistent with hypoglycaemia.

ADA classification[7] of hypoglycaemia

- Severe hypoglycaemia: An episode requiring assistance of another person to actively administer carbohydrate, glucagon, or take other corrective actions. PG concentrations may not be available during an event, but neurological recovery following the return of PG to normal is considered sufficient evidence that the event was induced by a low PG concentration.
- Asymptomatic hypoglycaemia: An episode not accompanied by typical symptoms of hypoglycaemia, but with a measured PG concentration ≤ 3.9 mmol/L (70 mg/dL).
- Documented symptomatic hypoglycaemia: An episode during which typical symptoms o f hypoglycaemia are accompanied by a measured PG 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 PG 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 PG determination but that was presumably caused by a PG 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 2–2 ADA classification of hypoglycaemia

Data on treatment-emergent hypoglycaemic episodes will be 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 patient years of observation time.

Severe or BG-confirmed symptomatic hypoglycaemic endpoints

Because the number of hypoglycaemic episodes are expected to be very low in this trial, statistical analyses of severe or BG-confirmed symptomatic hypoglycaemia will not be performed.

2.4.2.3 Pharmacokinetic endpoints

- Semaglutide plasma concentrations for population PK analysis
- SNAC plasma concentrations

The semaglutide plasma concentrations and SNAC plasma concentrations collected in this trial will be evaluated using relevant summary statistics. In addition, the semaglutide plasma concentration will be part of a meta-analysis across the oral semaglutide phase 3a trials, see more details in section 2.6.

2.5 Interim analysis

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

2.6 Pharmacokinetic and/or pharmacodynamic modelling

Data from this trial will be evaluated using population pharmacokinetic analysis and exposure-response for semaglutide. The purpose of the population pharmacokinetic analysis will be:

- To describe the covariate factors (such as weight, age, gender, race and ethnicity) that influence semaglutide exposure
- To estimate a steady-state exposure level for each subject with pharmacokinetic data, in order to facilitate subsequent exposure-response analyses

The purpose of the exposure-response analyses will be to support the recommended dose, by investigating response and potentially side effects across the exposure range.

The population pharmacokinetic (PK) and exposure-response analyses will be conducted as a meta-analysis, including all relevant oral semaglutide phase 3a trials with PK assessments. A separate modelling analysis plan will be prepared before first database lock in the oral semaglutide phase 3a programme, outlining details of the analyses. The modelling will be performed by Quantitative Clinical Pharmacology at Novo Nordisk A/S and will be reported separately from the clinical trial report.

2.7 Health economics and/or patient reported outcomes

Change from baseline to week 26 in:

- SF-36v2™ (acute version) health survey: Scores from the 8 domains and summary of the physical component score and the mental component score
- IWQOL-Lite Clinical Trial Version: Total score of the 22 items
- PGI-S Items: Scores of the two individual items
- PGI-C Items: Scores of the two individual items

A more detailed description of the handling of the four patient reported outcomes (PRO) questionnaires used in the trial is provided in the following sections.

No multiplicity adjustments will be done for the PRO questionnaires.

2.7.1 SF-36v2® (acute version) health survey

The SF-36v2® Health Survey (SF-36v2) (acute version) instrument is a commonly used generic instrument measuring health-related quality of life (HRQoL)/general health status across disease areas including diabetes. The SF-36v2 is a PRO questionnaire for adults and contains 36 items (see Table 2–3).

A total of 35 items measure eight domains of functional health and well-being as well as two summary domains with a 1-week recall period: physical functioning (10 items), role limitation due to physical health problems (4 items), bodily pain (2 items), general health perceptions (5 items), vitality (4 items), social functioning (2 items), role limitations due to emotional problems (3 items) and general mental health (5 items), mental component score (MCS), physical component score (PCS). There is an additional single item giving information on health change over the past week.

Domain scores

Norm-based scores (NBS) will be derived using the QualityMetric Health Outcomes™ Scoring Software[8] including the 2009 US general population norm used for all the oral semaglutide phase 3a programme trials. Version 4.5 of the QualityMetric Health Outcomes™ Scoring Software available at time of licensing will be used. Table 2–3 provides an overview of the domains. NBS standardises domain and component scores into T-scores using the means and standard deviations from the US general population. Higher scores on all domains and component summary measures (PCS and MCS) indicate better HRQoL/general health status. Item 2 (i.e. Question 2 in the eCRF) is not included in any score.

Table 2–3 Overview of domains for SF-36v2® (acute version)

Domain	Items numbers of items included in domain	Comment
Physical Functioning (PF)	Items 3a-j	
Role Limitations Due to Physical Health (Role-Physical; RP)	Items 4a-d	
Bodily Pain (BP)	Items 7, 8	Both item scores reversed
General Health Perceptions (General Health; GH)	Items 1, 11a-d	Item scores 1, 11b and 11d reversed
Vitality (VT)	Items 9a, 9e, 9g, 9i	Item scores 9a and 9e reversed
Social Functioning (SF)	Items 6, 10	Item score 6 reversed
Role Limitations Due To Emotional Problems (Role-Emotional; RE)	Items 5a-c	
Mental Health (MH)	Items 9b, 9c, 9d, 9f, 9h	Item scores 9d and 9h reversed
Physical component summary (PCS)	NA	The PCS score is a weighted average of the 8 domain scores.
Mental component summary (MCS)	NA	The MCS score is also a weighted average of the 8 domain scores. Weights differ from PCS to MCS.

Missing data at instrument level will be handled using the Maximum Data Recovery method: The method applies a value to a domain item rendered missing if at least one of the items in that domain has valid data. A domain score is considered missing if all item values in the domain are missing. PCS and MCS are calculated when at least seven of the eight domains have valid data, either actual or estimated. However, to calculate PCS, the PF domain must be one of the seven domains having valid data. Also, to calculate MCS, the MH domain must be one of the seven domains having valid data.

The domains will be evaluated using the primary analysis for the treatment policy estimand for the CTR. The domains will be evaluated using the primary analysis for the hypothetical estimand in a report separate from the CTR.

Responder threshold values

The responder threshold values, see section 2.7.4, in terms of T-score points for change from baseline are defined in Table 2–4[8].

Table 2–4 Responder thresholds for SF-36v2 (acute version)

Domain	Responder threshold
Physical Functioning (PF)	4.3
Role Limitations Due to Physical Health (Role-Physical; RP)	4.0
Bodily Pain (BP)	5.5
General Health Perceptions (General Health; GH)	7.0
Vitality (VT)	6.7
Social Functioning (SF)	6.2
Role Limitations Due To Emotional Problems (Role-Emotional; RE)	4.6
Mental Health (MH)	6.7
Physical component summary (PCS)	3.8
Mental component summary (MCS)	4.6

2.7.2 Impact of Weight on Quality of Life Clinical Trials Version (IWQOL-Lite-CT)

The IWQOL-Lite-CT is designed to assess the impact of changes in weight on patients' quality of life within the context of clinical trials. The items of the IWQOL-Lite-CT pertain to physical functioning and psychosocial domains and all items employ a 5-point graded response scale (never, rarely, sometimes, usually, always; or not at all true, a little true, moderately true, mostly true, completely true).

Composite scores

All IWQOL-Lite-CT composite scores range from 0 to 100, with higher scores reflecting better levels of functioning. Table 2–5 provides an overview of domains.

Table 2–5 Overview of domains for IWQOL-Lite-CT

Domain	Items numbers included in domain
Psychosocial	6-8, 10-16, 19, 21, 22
Physical	1-5, 17, 18
Physical Function	1-3, 17, 18
Pain/Discomfort	4, 5
IWQOL-Lite-CT Total	1-8, 10-19, 21, 22

Missing data at instrument level will be handled in the following way. Raw scores for each subscale are computed if a minimum of 50% of the items for that subscale are non-missing, and for the IWQOL-Lite-CT total score if a minimum of 75% of all items are non-missing.

The scoring is done in three steps:

1. If the minimum number of items are answered for a composite, compute the average of the valid item responses for that composite where 1 = ”never” or “not at all true” and 5 = ”always” or “completely true.” This average must be a number between 1 and 5, inclusive.

2. Multiply that average by the total number of items in that composite. The total number of items in each IWQOL-Lite-CT composite is as follows: Psychosocial = 13, Physical = 7, Physical Function = 5, Pain/Discomfort = 2, IWQOL-Lite-CT Total = 20. Round to the nearest integer.

3. To convert the IWQOL-Lite-CT raw score (as calculated above) to the 0 (worst) to 100 (best) metric:

    a. Subtract the raw score from the maximum raw score for each composite. The maximum raw scores are as follows: Psychosocial = 13x5 = 65, Physical = 7x5 = 35, Physical Function = 5x5 = 25, Pain/Discomfort = 2x5 = 10, IWQOL-Lite-CT Total = 20x5 = 100.

    b. Divide the difference in (a) by the raw score range for each composite. The ranges of the raw scores are as follows: Psychosocial = 65-13 = 52, Physical = 35-7 = 28, Physical Function = 25-5 = 20, Pain/Discomfort = 10-2 = 8, IWQOL-Lite-CT Total = 100-20 = 80.

    c. Multiply the total in (b) by 100.

Responder threshold values

The data from this trial is used as part of the confirmation of the psychometric properties of IWQOL-Lite-CT in patients with type 2 diabetes mellitus. To estimate potential responder threshold values, see section 2.7.4, for the composite scores in IWQOL-Lite-CT anchor-based and distribution-based methods are applied to data in this trial.

As a consequence, it is not possible to specify responder threshold values before database lock for this trial.

IWQOL-Lite-CT data will be presented using summary statistics at the item level. Scoring and analyses will be reported after the CTR. These analyses will be done for all domains in Table 2–5. The domains will be evaluated using the primary analysis for the both the treatment policy estimand and the hypothetical estimand in a report separate from the CTR.

2.7.3 Patient Global Impression of Status (PGI-S) and Patient Global Impression of Change (PGI-C)

PGI-S assesses patients’ impression of physical functioning and mental health status during the clinical trial. The PGI-S contains two items evaluated on a 5-point graded response scale.

PGI-C assesses patients' impression of change in physical functioning and mental health during the clinical trial compared to when they entered the trial. The PGI-C contains two items evaluated on a 7-point graded response scale.

Descriptive statistics (frequencies of answers per response scale point) will be given for PGI-S and PGI-C item level data. No scores will be derived and as a consequence scores will not be analysed.

2.7.4 Responder analyses for PRO to be reported in a separate report

Responder analyses will be reported in a report separate from the CTR for SF-36v2® and IWQOL-Lite-CT. These additional analyses were not in scope during the development process of the protocols for the NN9924 phase 3a programme. Responder analyses will be conducted for both estimands, for the same time points that are defined for the analyses of PRO endpoints and separately for each domain.

For descriptive statistics the following subject responder categorisation is applied for all relevant time points and domains:

- Responder (improvement): Individual change from baseline in score ≥ positive responder threshold
- Non-responder (no change): Individual change from baseline in score > negative responder threshold value and < positive responder threshold value
- Non-responder (worsening): Individual change from baseline in score ≤ negative responder threshold value

The following binary subject responder definition is applied for all relevant time points and scores:

- Responder: Individual change from baseline in score ≥ positive responder threshold
- Non-responder: Individual change from baseline in score < positive responder threshold

The binary responder endpoints will be analysed as the other supportive secondary binary effect endpoints (section 2.4.2.1). Estimated proportions and differences in proportions will be reported in addition to odds and odds ratios.

3 Changes to the statistical analyses planned in the protocol

The main analyses were described in the protocol for trial NN9924-4233. However, clarifications, more detailed descriptions of endpoints and analyses are provided in this SAP. The changes from protocol NN9924-4233 are summarised below:

- It has been specified which countries belong to which regions.
- The primary and secondary estimand has changed name from de-facto and de-jure to treatment policy and hypothetical, respectively.
- The MMRM sensitivity analysis of the primary estimand has been omitted in section 2.3.3. It is considered sufficient to keep the two current sensitivity analyses to stress test the primary results.
- For the MMRM analyses, it is specified that for subjects who do not have post-baseline assessments for planned visits available in the on-treatment without rescue medication period, the baseline value will be carried forward to the first planned visit, if the first planned visit do not fall later than 8 weeks after randomisation, to ensure that all randomised subjects will contribute to the statistical analyses.
- The three MI sensitivity analyses of the secondary estimand have been omitted in section 2.3.3. It is considered sufficient to keep the tipping point sensitivity analysis for the secondary hypothetical estimand as it can be considered as a progressive stress-testing to assess how severe departures from missing at random must be in order to reverse the conclusions from the primary MMRM analysis used to address the hypothetical estimand.
- The LOCF sensitivity analysis specified in the trial protocol (section 17.3.3.2) has been omitted, as it is not realistic that subjects with missing data would have had stable results from the point of dropout to trial completion.
- The statistical analysis of the two binary effect endpoints: HbA_{1c} reduction $\geq$ 1 %-point (10.9 mmol/mol); and body weight loss $\geq$ 3 %, have been omitted, because they are being analysed as a part of the two composite binary effect endpoints.
- For the binary efficacy endpoints, it has been specified how missing data in the analyses for the hypothetical estimand will be imputed using a sequential imputation approach assuming MAR.
- It has been specified which assessments will be analysed on the logarithmic scale.

- The definitions of initiation of rescue medication and additional anti-diabetic medication used for the time-to-event endpoints as well as the accompanying statistical analyses have been further clarified.

- It has been specified that all safety laboratory results (except amylase and lipase) are safety assessments and not safety endpoints as written in the trial protocol.

- The pre-specified analyses of severe or BG-confirmed symptomatic hypoglycaemic episodes will not be performed due to sparse data.

- Only one of the four PRO questionnaires, SF-36v2® Health Survey (SF-36v2) (acute version), will be analysed statistically, and only for the primary analysis of the primary estimand.

  The responder analyses, and the primary analysis for the secondary estimand of SF-36v2® will be presented in a report separate from and after finalisation of the CTR. The three other PRO questionnaires, IWQOL-Lite-CT and PGI-S/PGI-C, will be presented using summary statistics in the CTR.

  Statistical analysis of the IWQOL-Lite-CT questionnaires will be presented in a report separate from and after finalisation of the CTR.

4 References

1. Bretz F, Posch M, Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. Biometrical Journal. 2011;53(6):894-913.

2. Rohmeyer K, Klinglmueller F. gMCP: Graph Based Multiple Test Procedures. R package version 0.8-8. 3 Oct 2014.

3. Little RJA, Rubin DB. Statistical analysis with missing data: New York: John Wiley & Sons. 1987.

4. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP). Guideline on Missing Data in Confirmatory Clinical Trials (EMA/CPMP/EWP/1776/99 Rev. 1). 2 Jul 2010.

5. National Academy of Sciences (NAS). The Prevention and Treatment of Missing Data in Clinical Trials. Washington D.C.: The National Academies Press. 2010.

6. 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.

7. 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.

8. Maruish ME, (Ed.). User's manual for the SF-36v2 Health Survey (3rd ed.). Lincoln, RI: QualityMetric Incorporated. 2011.

16.1.9.1 Pre-defined MedDRA search – list of preferred terms

Table of contents

Page

Trial: NN9924-4233

16-1-09-statistical methods, Version 1.0 dated 20 June 2018

Overview of deleted pages

Pages	Section	Title
2-211	16.1.9.1	Pre-defined MedDRA search – list of preferred terms