Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT04074161
Sponsor trial ID:	NN9536-4576
Official title of study:	Effect and safety of subcutaneous semaglutide 2.4 mg once weekly compared to liraglutide 3.0 mg once daily on weight management in subjects with overweight or obesity
Document date*:	08 July 2021

16.1.9 Documentation of statistical methods

List of contents

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

.

[redacted] [redacted]

.

Statistical Analysis Plan

Trial ID: NN9536-4576

Effect and safety of subcutaneous semaglutide 2.4 mg once weekly compared to liraglutide 3.0 mg once daily on weight management in subjects with overweight or obesity

Author

[redacted] and [redacted]

Table of contents Page

List of abbreviations

AD	available but discontinued
AE	adverse event
ANCOVA	analysis of covariance
AT	available on randomised treatment
BMI	body mass index
CI	confidence interval
COVID-19	Coronavirus disease 2019
dBP	diastolic blood pressure
FAS	full analysis set
FFA	free fatty acids
FPG	fasting plasma glucose
HbA_{1c}	glycated haemoglobin
HDL	high density lipoprotein
hsCRP	high sensitive C-reactive protein
J2R-MI	jump to reference multiple imputation
LAO-OT	last available observation on-treatment
LDL	low density lipoprotein
MD	missing and discontinued
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model for repeated measurements
MT	missing on randomised treatment
OR	odds ratio
PYE	patient years of exposure
PYO	patient years of observation
RD-MI	multiple imputation using retrieved subjects
SAE	serious adverse event
SAP	statistical analysis plan
SAS	safety analysis set
sBP	systolic blood pressure
s.c.	subcutaneus
SD	standard deviation
TEAE	treatment emergent adverse event
VLDL	very low density lipoprotein
WC	waist circumference

1. Introduction

1.1 Trial information

This is a 68-week, randomised, open label, pairwise placebo-controlled, multi-centre, US only clinical trial comparing semaglutide s.c. 2.4 mg once weekly with liraglutide s.c. 3.0 mg once daily in subjects with overweight or obesity. Semaglutide once weekly vs liraglutide once daily treatment will be open label, but each of the two active treatment arms will be double blinded against placebo administered at the same dosing frequency.

Primary objective
To show the superiority of semaglutide s.c. 2.4 mg once weekly versus liraglutide 3.0 mg once daily both as adjuncts to a reduced-calorie diet and increased physical activity in subjects with obesity or with overweight and at least one weight related comorbidity on body weight.

Secondary objectives
To compare the safety and tolerability of semaglutide s.c. 2.4 mg once weekly versus liraglutide 3.0 mg once daily both as adjuncts to reduced-calorie diet and increased physical activity in subjects with obesity or with overweight and at least one weight related comorbidity.

To compare the effect of semaglutide s.c. 2.4 mg once weekly versus liraglutide 3.0 mg once daily both as adjuncts to a reduced-calorie diet and increased physical activity in subjects with obesity or with overweight and at least one weight related comorbidity on cardiovascular risk factors and glucose metabolism

To show the superiority of semaglutide s.c. 2.4 mg once weekly versus placebo both as adjuncts to a reduced-calorie diet and increased physical activity in subjects with obesity or with overweight and at least one weight related comorbidity on body weight.

To show the superiority of liraglutide s.c. 3.0 mg once daily versus placebo both as adjuncts to a reduced-calorie diet and increased physical activity in subjects with obesity or with overweight and at least one weight related comorbidity on body weight.

Primary estimand
The estimand will quantify the average treatment effect of semaglutide relative to liraglutide after 68 weeks, as adjuncts to a reduced-calorie diet and increased physical activity, in all randomised subjects regardless of adherence to treatment and regardless of starting rescue interventions (weight management drugs or bariatric surgery) (“treatment policy” estimand). The estimand will cover all effect-related objectives.

The following expansion of the primary estimand will cover objectives related to weight. The estimand will quantify the average treatment effect of semaglutide s.c. 2.4 mg relative to placebo after 68 weeks, as an adjunct to reduced-calorie diet and increased physical activity, in all

randomised subjects regardless of adherence to treatment and regardless of starting rescue interventions.

The following expansion of the primary estimand will cover objectives related to weight. The estimand will quantify the average treatment effect of liraglutide s.c. 3.0 mg relative to placebo after 68 weeks, as an adjunct to reduced-calorie diet and increased physical activity, in all randomised subjects regardless of adherence to treatment and regardless of starting rescue interventions.

Secondary estimand
The estimand will quantify the average treatment effect of semaglutide relative to liraglutide after 68 weeks, as an adjunct to a reduced-calorie diet and increased physical activity, in all randomised subjects had they remained on their randomised treatment for the entire planned duration of the trial and not started any rescue intervention (weight management drugs or bariatric surgery) ("hypothetical" estimand). The estimand will cover all effect-related objectives.

The handling of intercurrent events with respect to data collection and analysis is specified in Table 1-1 for the primary endpoint. Apart from the listed intercurrent events, missing data will occur due to death, or if subjects withdraw consent, become lost to follow-up, or continue to be followed without being ascertained for the endpoint.

Table 1-1 Handling of intercurrent events for the primary endpoint

Intercurrent event	Data collection	Data analysis
Premature treatment discontinuation	Subjects will be followed and data collected after intercurrent events	**Primary estimand:** data collected after intercurrent events used in analysis in line with a treatment-policy strategy **Secondary estimand:** data collected after intercurrent events treated as missing in line with a hypothetical strategy
Initiation of other weight management drugs or bariatric surgery		

1.1.1 Endpoints

1.1.1.1 Primary endpoint

- Change from baseline (week 0) to week 68 in body weight (%)

1.1.1.2 Secondary endpoints

Confirmatory secondary endpoints

- Subject who from baseline (week 0) to week 68 achieve (yes/no):
  - Body weight reduction ≥10%
  - Body weight reduction ≥15%

- Body weight reduction ≥20%

Supportive secondary endpoints

The supportive secondary endpoints are used to compare the effect of semaglutide s.c. 2.4 mg once-weekly versus liraglutide 3.0 mg once-daily unless otherwise stated:

- Change from baseline (week 0) to week 68 in waist circumference (cm)
- Change from baseline (week 0) to week 68 in body weight (kg)
- Change from baseline (week 0) to week 68 in body weight (%) (semaglutide s.c. 2.4 mg once-weekly versus placebo and liraglutide s.c. 3.0 mg once-daily versus placebo)
- Change from baseline (week 0) to week 68 in body weight (kg) (semaglutide s.c. 2.4 mg once-weekly versus placebo and liraglutide s.c. 3.0 mg once-daily versus placebo)
- Change from baseline (week 0) to week 68 in:
  - systolic blood pressure (mmHg)
  - diastolic blood pressure (mmHg)
  - lipids (mg/dL, mmol/L)
    - Total cholesterol
    - High density lipoprotein (HDL) cholesterol
    - Low density lipoprotein (LDL) cholesterol
    - Very low density lipoprotein (VLDL) cholesterol
    - Free fatty acids (FFA)
    - Triglycerides
  - hsCRP (mg/L)
  - HbA1c (%, mmol/mol)
  - fasting plasma glucose (mg/dL, mmol/L)
  - fasting serum insulin (mIU/L, pmol/L)
  - glycaemic category (normo-glycaemia, pre-diabetes, T2D)
- Subjects who from baseline (week 0) to week 68 have permanently discontinued randomised trial product (yes/no)
- Number of treatment emergent adverse events (TEAEs) from baseline (week 0) to week 75
- Number of serious adverse events (SAEs) from baseline (week 0) to week 75

1.1.1.3 Exploratory endpoints

- Subject who from baseline (week 0) to week 68 achieve (yes/no):
  - Body weight reduction >5%

1.2 Scope of the statistical analysis plan

This statistical analysis plan (SAP) is based on the protocol for trial NN9536-4576 "Effect and safety of subcutaneous semaglutide 2.4 mg once weekly compared to liraglutide 3.0 mg once daily on weight management in subjects with overweight or obesity", version 1.

2. Statistical considerations

General considerations

A statistical analysis plan (SAP) will be written, including a more technical and detailed elaboration of the statistical analyses. The SAP will be finalized before breaking the blind to treatment assignment.

The placebo arms will be pooled in the statistical analyses.

The last available and eligible observation at or before randomisation is used as the baseline (randomisation) value. If no assessments are available, the mean value at randomisation across all subjects is used as the baseline value.

Results from statistical analyses will generally be accompanied by two-sided 95% confidence intervals and corresponding p-values.

Taxonomy of week 68 assessments

For each subject a given assessment at week 68 may be available or missing and Table 2-1 describes the taxonomy for this. Note, this is done per assessment and per subject; subjects may be a different type for different assessments (a subject may have "available on randomised treatment (AT)" for body weight but "missing on randomised treatment (MT)" for waist circumference).

Table 2-1 Taxonomy for subjects based on week 68 assessments

Assessment at week 68	Subjects on randomised treatment at week 68	Type description	Type Abbreviation
Available	Yes	**Available on randomised treatment**: Subjects who complete the trial on randomised treatment with an assessment at week 68: Includes those that stop and restart trial product.	AT
	No	**Available but discontinued** Subjects who discontinued randomised treatment prematurely but returned to have an assessment at week 68. These are also called retrieved subjects	AD
Missing	Yes	**Missing on randomised treatment**: Subjects who complete the trial on randomised treatment without an assessment at week 68: Includes those that stop and restart trial product.	MT
	No	**Missing and discontinued**: Subjects who discontinued randomised treatment prematurely and did not return to have an assessment at week 68. These are also called non-retrieved subjects	MD

2.1 Sample size determination

The tests of superiority of semaglutide 2.4 mg to liraglutide 3.0 mg for the primary and confirmatory secondary endpoints are performed using the fixed-sequence statistical strategy. This

strategy tests the endpoints using a predefined hierarchical order, all at the significance level of 5%, moving to test the next endpoint only after a statistically significant superiority result (p-value < 5%) on the previous endpoint. The test hierarchy is given in Table 2-2 with underlying assumptions, marginal power and effective power. The effective power is calculated under the assumption of independence of endpoints by multiplying the respective marginal powers successively.

The trial is designed with an effective power of 92% to detect differences on all endpoints in the statistical test hierarchy at one-sided alpha of 0.025 (equivalent to a two-sided alpha of 0.05). Specifically, the trial tests the null hypothesis of equal % change in body weight (or equal proportions achieving body weight loss of 10%, 15%, and 20%) against the one-sided alternative of greater % body weight loss in the semaglutide 2.4 mg arm (or greater proportions achieving body weight loss of 10%, 15%, and 20%) compared to the liraglutide 3.0 mg arm. The calculations for the primary endpoint are based on a t test on the mean difference assuming equal variances, whereas those for the confirmatory secondary endpoints are based on the Pearson chi-square test for two independent proportions. Assumptions for these calculations are based on findings from NN8022 (SCALE) and trial NN9536-4153. These assumptions are further explored across a range of sample sizes in Table 2-3. Under these assumptions and a 3:1:3:1 randomisation ratio (see section 5 in the protocol), the desired power of more than 90% is obtained with 126 subjects randomized to each active drug arm and 42 subjects randomized to each placebo arm.

Furthermore, a sample size of 84 subjects in the pooled placebo group (42 each in the semaglutide placebo and liraglutide placebo arms) gives a power of >99% for the comparison between semaglutide 2.4 mg once-weekly and pooled placebo on the primary endpoint, as well as a power of at least 80% for the comparison between liraglutide 3.0 mg once-daily and pooled placebo on the primary endpoint.

Table 2-2 Assumptions, marginal power and effective power for each endpoint in the hierarchical testing procedure given an anticipated number of 336 randomised subjects

Order	Endpoint	Expected mean (±SD) or proportion		Expected difference or proportion ratio	Marginal power (%)	Effective power (%)
		Semaglutide 2.4 mg	Liraglutide 3.0 mg			
1	% weight change #	12.5 (±10)	7.0 (±10)	5.5%-points	99	99
2	10% responders	61%	37%	1.6	97	96
3	15% responders	39%	18%	2.2	96	93
4	20% responders	27%	6%	4.5	99	92

SD = standard deviation; # shown as a positive number

All tests in the hierarchy are based on the primary estimand. Since all tests are tests of superiority of semaglutide 2.4 mg to liraglutide 3.0 mg, power is only shown for this comparison.

Table 2-3 Assumptions and effective power for the primary and confirmatory secondary endpoints across a range of number of randomised subjects

Expected mean (±SD)		Expected proportions of 10%/15%/20% body weight loss responders		N per active arm	Total N randomised	Effective power (%)
Semaglutide 2.4 mg	**Liraglutide 3.0 mg**	**Semaglutide 2.4 mg**	**Liraglutide 3.0 mg**			
12.5 (±10)	7.0 (±10)	61%/39%/27%	37%/18%/6%	99	264	81
				126	**336**	**92**
				150	400	97

SD = standard deviation
Mean weight change shown as a positive number.

2.2 Definition of analysis sets

Two analysis sets are defined:

The *full analysis set* (FAS) includes all randomised subjects according to the intention-to-treat principle. Subjects in the FAS will contribute to evaluation "as randomised".

The *safety analysis set* (SAS) includes all randomised subjects exposed to at least one dose of randomised treatment. Subjects in the SAS will contribute to evaluation "as treated".

Any observation excluded from the analysis database will be documented before database lock with the reason for exclusion provided.

Two observation periods are defined for each subject:

In-trial: The *in-trial period* is defined as the uninterrupted time interval from date of randomisation to date of last contact with trial site.

On-treatment (with trial product): A time-point is considered as 'on-treatment' if any dose of trial product has been administered within the prior 2 weeks (14 days). The *on-treatment period* is defined as all times which are considered on-treatment.

In general, the *on-treatment period* will therefore be from the date of first trial product administration to date of last trial product administration (+14 days) excluding potential off-treatment time intervals triggered by at least two consecutive missed doses.

For the evaluation of AEs the lag time for each on-treatment time interval is 7 weeks (49 days).

The in-trial and on-treatment periods define the patient years of observation (PYO) and patient years of exposure (PYE), respectively, as the total time duration in the periods.

Effect endpoints will be analysed using the FAS; safety endpoints will be analysed using the SAS.

2.3 Statistical analyses

Handling of missing baseline data

The last available and eligible observation at or before randomisation is used as the baseline value. If no assessments are available, the mean of baseline values across all subjects is used as the baseline value.

2.3.1 Primary endpoint

Definition of primary endpoint: % weight change

Change from baseline (week 0) to week 68 in body weight (%) is defined as

$$\% \text{ weight change} = \frac{(\text{body weight at week 68} - \text{body weight at baseline})}{\text{body weight at baseline}} \times 100.$$

Analyses addressing the primary estimand

The following statistical analyses and imputation methods are designed to address the primary estimand.

The analysis model for % weight change is a linear regression (ANCOVA) of % weight change with randomised treatment as a factor and baseline body weight (kg) as covariate. Randomised treatment will be coded in 3 levels: semaglutide 2.4 mg, liraglutide 3.0 mg, and pooled placebo. The estimated treatment difference between semaglutide 2.4 mg and liraglutide 3.0 mg will be reported together with the associated two-sided 95% confidence interval (CI) and corresponding p-value. Estimated treatment differences between semaglutide 2.4 mg and pooled placebo, as well as between liraglutide 3.0 mg and pooled placebo, will also be reported together with associated two-sided 95% confidence intervals (CI).

The superiority tests of semaglutide 2.4 mg vs. liraglutide 3.0 mg will be carried out as follows.

Let $\mu_{\text{semaglutide}}$ and $\mu_{\text{liraglutide}}$ denote the true mean of % weight change for semaglutide 2.4 mg and liraglutide 3.0 mg group, respectively. The null and alternative hypotheses tested are

$$H: \mu_{semaglutide} \geq \mu_{liraglutide} \; vs$$
$$H_A: \mu_{semaglutide} < \mu_{liraglutide}.$$

The hypothesis will be rejected and superiority claimed, if the upper limit of the estimated two-sided 95% CI is below 0.

Handling of missing week 68 values for the primary estimand

All available data at week 68 are used and missing values at week 68 will be imputed and the endpoints will be derived from the imputed values. The imputation approach for the primary analysis is a multiple imputation similar to the one described by McEvoy[1]. For subjects in the semaglutide 2.4 mg, liraglutide 3.0 mg and pooled placebo groups, missing body weight measurements at week 68 for non-retrieved subjects (MD) are imputed using assessments from retrieved subjects (AD) in each treatment group. This will be done according to the timing of last available observation during the on-treatment period (LAO-OT) of body weight prior to week 68. Missing body weight measurements at week 68 for subjects on randomised treatment (MT) are imputed by sampling from available measurements at week 68 from subjects on randomised treatment (AT) in the relevant randomised treatment arm. The multiple imputation approach is done in three steps:

1. **Imputation:** Defines an imputation model using retrieved subjects (AD) from FAS and done within groups defined by randomised treatment and the timing of the LAO-OT of body weight. The model will be a linear regression of body weight (kg) at week 68 with sex (male/female), baseline BMI (kg/m^2) (in categories -<35, 35-<40, ≥40) as factors and baseline body weight (kg) and LAO-OT of body weight (kg) as covariates. No interactions will be included. The grouping of timing will be done by quarters (intervals of 17 weeks). If timing by quarters is too restrictive, halves (intervals of 34 weeks) or excluding timing will be used. The timing by quarters or halves is defined as too restrictive if the imputation model cannot be fit due to inadequate numbers of retrieved subjects in 1 or more groups. If the imputation model still cannot be fit after excluding timing then the model will be further reduced until the model can be fit. Reduction will be done in a fixed order by first removing sex, then collapsing the two highest baseline BMI-groups into one (≥35) and finally removing baseline BMI-group. If no LAO-OT exists post-baseline then the LAO-OT will be the baseline body weight and the timing will be the first interval. If any subjects are MT, an imputation model for missing body weight measurements at week 68 for MT subjects will also be defined using AT subjects in a similar way. The estimated posterior distribution for the parameters (regression coefficients and variances) in the imputation models are then used to impute missing week 68 body weight values for each randomised treatment arm. This will be done 1,000 times and results in 1,000 complete data sets.

2. **Analysis**: Analysis of each of the 1,000 complete data sets, using the analysis models (ANCOVA and logistic regression) results in 1,000 times 2 estimations.

3. **Pooling**: Integrates the 1,000 times 2 estimation results into two final results using Rubin's formula.

Based on NN9536-4153 phase 2 results 1,000 copies should be sufficient to establish stable results. If 1,000 copies are insufficient, 10,000 copies will be used. The multiple imputations will be

generated using Novo Nordisk trial number 95364576 as seed number. In addition to the seed number, it is specified that the dataset is sorted by subject ID.

Sensitivity analyses

Tipping-point multiple imputation analysis: First, missing data are imputed according to the primary multiple imputation approach. Then, a penalty is added to the imputed values at week 68. The approach is to explore a range of penalties for both the semaglutide 2.4 mg and liraglutide 3.0 mg arms, and the impact these would have on the study conclusions. The 2-dimensional space of penalties covering the range from -30% to 30% will be explored for both arms. Penalties for the pooled placebo group will not be considered, as this sensitivity analysis evaluates the robustness of the superiority conclusions to departures from the observed change in body weight in both active treatment groups.

Jump to reference multiple imputation approach (J2R-MI): Missing values of body weight at week 68 (MT and MD) for both the semaglutide 2.4 mg and liraglutide 3.0 mg group are imputed by sampling among all available assessments at week 68 in the pooled placebo group (AT and AD). This approach makes the assumption that subjects instantly after discontinuation lose any effect of randomised treatment beyond what can be expected from placebo treatment as adjunct to reduced-calorie diet and increased physical activity[1].. The multiple imputation approach is done as above with the first step replaced by:

- **Imputation**: Defines an imputation model using pooled placebo subjects from FAS with a week 68 measurement (AT and AD). The model will be a linear regression of body weight (kg) at week 68 with sex (male/female), BMI (kg/m^2) (in categories -<35, 35-<40, ≥40) as factors and baseline body weight (kg) as covariate. No interactions will be included. If the imputation model cannot be fit due to inadequate numbers of retrieved subjects in one or more groups, then the imputation model will be reduced until the model can be fit. Reduction will be done in a fixed order by first removing sex, then collapsing the two highest baseline BMI groups into one (≥35) and finally removing baseline BMI group. The estimated posterior distribution for the parameters (regression coefficients and variances) in the imputation models are then used to impute missing week 68 body weight values for each randomised treatment arm. This will be done 1,000 times and results in 1,000 complete data sets.

In-house body weight assessments: The analysis of primary endpoint addressing primary estimand will be repeated with exclusion of body weight measurements at week 68 assessed with mobile scale. Excluded data points at week 68 will be treated as missing and subsequently imputed together with other missing values using the same multiple imputations approach as described for primary analysis.

Analysis addressing the secondary estimand

The secondary estimand for % weight change addresses the efficacy of semaglutide 2.4 mg and will be assessed using a Mixed Model for Repeated Measurements (MMRM) approach. Week 68 assessments for retrieved subjects are not used in this analysis. The MMRM will use assessments only from subjects who are taking the randomised treatment until end of treatment or until first discontinuation of randomised treatment. For subjects who initiate rescue interventions before completion or first discontinuing of randomised treatment, the date of starting weight management drugs or undergoing bariatric surgery will be used as latest date for using assessments in this MMRM. The MMRM will be fitted using % weight change and the same factor and covariate as for the primary analysis all nested within visit. An unstructured covariance matrix for measurements within the same subject will be employed, assuming that measurements for different subjects are independent.

An overview of all analysis and imputation methods to address the primary and secondary estimands for the primary endpoints is given in Table 2-4.

Table 2-4 Analysis and imputation methods to address the primary and secondary estimands for the primary and confirmatory secondary endpoints in the statistical testing hierarchy

Objective	Endpoint	Test order	Endpoint type	Estimand	Analysis set	Statistical model	Imputation approach	Sensitivity analyses
Primary endpoints								
Primary	% weight change	1	Continuous	Primary	FAS	ANCOVA	RD-MI	TP-MI In-house J2R-MI
				Secondary	FAS	MMRM	-	-
Confirmatory secondary endpoints								
Primary	10% responders	2	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-
Primary	15% responders	3	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-
Primary	20% responders	4	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-

FAS = full analysis set; ANCOVA = analysis of covariance; RD-MI = multiple imputation using retrieved subjects; TP-MI = tipping point multiple imputation; MMRM = mixed model for repeated measurements; LR = logistic regression; Test order refers to the order of the endpoint in the statistical test hierarchy outlined in Table 2-2.

2.3.2 Secondary endpoints

2.3.2.1 Confirmatory secondary endpoints

Confirmatory secondary endpoints are listed in Section 1.1.1.2and are all included in the fixed-sequence statistical strategy, see above. All tests are tests of superiority of semaglutide 2.4 mg to liraglutide 3.0 mg.

Analyses addressing the primary estimand

All confirmatory secondary endpoints will be analysed using the primary imputation approach used for the primary endpoint and to address the primary estimand. The statistical model for body weight responder endpoints is a logistic regression using randomised treatment as a factor and baseline body weight (kg) as covariate. Randomised treatment will be coded in 3 levels: semaglutide 2.4 mg, liraglutide 3.0 mg, and pooled placebo. The estimated odds ratio (OR) between semaglutide 2.4 mg and liraglutide 3.0 mg will be reported together with the associated two-sided 95% confidence interval and corresponding p-value. Estimated odds ratios between semaglutide 2.4 mg and pooled placebo, as well as between liraglutide 3.0 mg and pooled placebo, will also be reported together with associated two-sided 95% confidence intervals.

Analyses addressing the secondary estimand

The confirmatory secondary endpoints will be analysed to address the secondary estimand using the same MMRM described for the primary endpoint. From the MMRM individually predicted values for % weight change at week 68 will be used to classify each subject as a responder or not. This classification will then be analysed using a logistic regression model with treatment as factor and baseline body weight (kg) as covariate.

Sensitivity analyses for confirmatory secondary endpoints

For all confirmatory secondary endpoints a sensitivity analysis using non-retrieved subjects as non-responders will be carried out.

An overview of all analysis and imputation methods to address the primary and secondary estimands for confirmatory secondary endpoints is given in Table 2-4.

2.3.2.2 Supportive secondary endpoints

Supportive secondary endpoints are listed in Section 1.1.1.2. All tests are tests of superiority of semaglutide 2.4 mg to liraglutide 3.0 mg unless otherwise stated.

Analyses addressing the primary estimand

The effect-related supportive secondary endpoints will be analysed using the same imputation approach as used for the primary endpoints and to address the primary estimand. The imputation model is the same as for the primary endpoints with body weight replaced by assessments of the endpoint to be analysed. For fasting serum insulin, collected only at baseline and week 68, the LAO-OT covariate is not relevant and will be excluded from the imputation model. The statistical model for continuous endpoints will be ANCOVA with factor and covariate as for the primary endpoint % weight change with baseline body weight replaced by the baseline assessment of the endpoint to be analysed.

For lipids, fasting serum insulin and hsCRP a multiplicative model will be used, i.e. the ratio between post randomisation measurements and baseline will be calculated instead of differences, and both the dependent variable and covariate will be log-transformed.

For glycaemic status and treatment discontinuation no analysis will be performed. Observed data will be summarised by descriptive statistics.

Analyses addressing the secondary estimand

The supportive secondary endpoints which relate to the primary objective will be analysed to address the secondary estimand using the same MMRM described for the primary endpoints.

Sensitivity analyses for supportive secondary endpoints

For supportive secondary endpoints no sensitivity analysis will be carried out.

Analysis of safety endpoints

An overview of all analysis and imputation methods to address the primary and secondary estimands for supportive secondary endpoints is given in Table 2-5.

Table 2-5 Analysis and imputation methods to address the primary and secondary estimands for supportive secondary endpoints

Adverse events will be defined as “treatment-emergent” (TEAE), if the onset of the event occurs in the on-treatment period (see definition in Section 2.2). TEAEs and SAEs will be summarised by descriptive statistics, such as frequencies and rates. No formal statistical inference will be carried out based on the number of TEAEs and SAEs. All AEs will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA).

Objective	Endpoint	Endpoint type	Estimand	Analysis set	Statistical model	Imputation approach	Sensitivity analyses
Supportive secondary endpoints (effect related)							
Primary	WC change (cm)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Primary	Weight change (kg)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	Weight change (kg)*	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	Weight change (kg)**	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	% weight change*	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	% weight change**	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	sBP change (mmHg)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	dBP change (mmHg)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	Total cholesterol change (mg/dL, mmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	HDL change (mg/dL, mmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	LDL change (mg/dL, mmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	VLDL change (mg/dL, mmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	FFA change (mg/dL, mmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	Triglycerides change (mg/dL, mmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	hsCRP change (mg/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	HbA1c change (%, mmol/mol)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	FPG change (mg/dL, mmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	Fasting serum insulin change (mIU/L, pmol/L)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	Glycaemic category change	Categorical	Primary	FAS	Descriptive statistics	-	-
Secondary	Subjects who have permanently discontinued randomised trial product	Binary	Primary	FAS	Descriptive statistics	-	-
Supportive secondary endpoints (safety related)							
Secondary	Number of TEAEs	Continuous	-	SAS	-	-	-
Secondary	Number of SAEs	Continuous	-	SAS	-	-	-

WC = waist circumference; FAS = full analysis set; ANCOVA = analysis of covariance; RD-MI = multiple imputation using retrieved subjects; MMRM = mixed model for repeated measurements; sBP = systolic blood pressure; dBP = diastolic blood pressure; HDL = high density lipoprotein; LDL = low density lipoprotein; VLDL = very low density lipoprotein; FFA = free fatty acids; hsCRP = high sensitivity C-Reactive Protein; HbA1c = Hemoglobin A1c; FPG = fasting plasma glucose; TEAEs = treatment emergent adverse events; SAEs = serious adverse events; *comparison of semaglutide 2.4 mg once-weekly vs pooled placebo; **comparison of liraglutide 3.0 mg once-daily vs pooled placebo;

2.3.3 Exploratory endpoints

Exploratory endpoints are listed in Section 1.1.1.3. Observed data for exploratory endpoints will be summarised by descriptive statistics.

2.4 Interim analysis

This section is not applicable in this trial.

3. Changes to the statistical analyses planned in the protocol

The analyses for the primary and confirmatory secondary endpoints were described in the protocol for the trial NN9536-4576. Analysis of supportive secondary endpoints are described in this SAP. Additionally, clarifications and more detailed descriptions of endpoints and analyses are provided in this SAP. The changes from the protocol of NN9536-4576 are summarised below:

- It has been corrected that the secondary estimand will cover all effect-related objectives
- Supportive secondary endpoints of change in body weight (kg), diastolic blood pressure (mmHg) and fasting serum insulin (mIU/L, pmol/L) have been added.
- Following units have been added to the respective endpoints: “mmol/L” for lipids and FPG and “mmol/mol” for HbA1c.
- It was clarified that subjects in the FAS/SAS will be evaluated “as randomised”/“as treated”.
- Analyses updated to include added endpoints: fasting serum insulin (mIU/L, pmol/L), diastolic blood pressure (mmHg) and body weight (kg) and units: “mmol/L” for lipids and FPG and “mmol/mol” for HbA_{1c}.
- In the text describing that “In general, the on-treatment period will therefore be from the date of first trial product administration to date of last trial product administration” the following has been added “(+14 days)” to emphasize that the lag-time after last trial product administration is included in the on-treatment period.
- The text explaining how to handle missing baseline values has been added.
- Description of the taxonomy of week 68 assessments has been added.
- The hypothesis for the primary endpoint has been defined.
- Details of the multiple imputation approach are described.
- It is clarified that the LAO-OT must be prior to the landmark visit (week 68).
- The BMI-grouping “27-<35” has been changed to “-<35”, since subjects may lose weight between the screening and the randomisation visit, and therefore have a BMI below 27 kg/m2 at the time of randomisation.
- Further steps of imputation model reduction have been described.
- It is clarified that if no post-baseline LAO-OT exist, then LAO-OT will be the baseline value and the timing of LAO-OT will be the first interval.
- J2R-MI sensitivity analysis on primary endpoint addressing primary estimand was added.
- In-house body weight sensitivity analysis on primary endpoint addressing primary estimand was added.
- Analysis of supportive secondary endpoints are described.
- Tables specifying analysis for primary and confirmatory endpoints has been added.

- Proportion of subjects losing at least 5% of baseline body weight at week 68 has been defined as an exploratory endpoint.
- It has been corrected that the logistic regression model for confirmatory endpoints addressing secondary estimand will include both randomised treatment as factor and baseline body weight (kg) as covariate.
- It has been clarified that for the fasting serum insulin endpoint the LAO-OT covariate will not be included in the imputation model due to data being collected at baseline and week 68 only.

4. Change log

SAP change log

Version	Reason for change
1.0	New
2.0	J2R-MI sensitivity analysis on primary endpoint addressing primary estimand was added to investigate robustness of primary analysis. In-house body weight sensitivity analysis on primary endpoint addressing primary estimand was added to investigate the impact of mobile scale assessments – an option made available to subjects at week 68 in an effort to limit COVID-19 impact on missing body weight data. Correction was made to specification of the confirmatory secondary endpoints analysis addressing secondary estimand clarifying that the logistic regression model should include both randomised treatment as factor and baseline body weight (kg) as a covariate. This adjustment is in line with respective analysis performed in previously reported STEP1-4 trials. A number of adjustments has been made to align with previously reported STEP1-4 trials. These cover: • correction that the secondary estimand will cover all effect-related objectives • addition of change in body weight (kg), diastolic blood pressure (mmHg) and fasting serum insulin (mIU/L, pmol/L) as supportive secondary endpoints • inclusion of additional units: “mmol/L” for lipids and FPG and “mmol/mol” for HbA1c • update to the analysis to include added endpoints and units • clarification to the on-treatment period definition that it includes +14 day lag from last dosing

	• explanation how to handle missing baseline values • clarification that LAO-OT must be prior to the landmark visit • change in BMI-grouping • description of further steps in the reduction procedure for imputation model • clarification on when baseline value can be used as LAO-OT • adjustment of imputation model for fasting serum insulin endpoint. The changes are described in more details in section 3.

5. References

1. McEvoy BW. Missing data in clinical trials for weight management. J Biopharm Stat. 2016;26(1):30-6.

Clinical Trial Report

Trial ID: NN9536-4576 (STEP 8)

MedDRA searches within safety focus areas

Author
Novo Nordisk

Table of contents

Page

List of abbreviations and definitions of terms

AE	adverse event
HLT	high level term
MedDRA	Medical Dictionary for Regulatory Activities
NEC	not elsewhere classified
NNMQ	Novo Nordisk MedDRA query
PT	preferred term
SMQ	standardised MedDRA query
SOC	system organ class

1 MedDRA searches for safety focus areas in project NN9536

The MedDRA search strings in this document (ordered alphabetically) were used for the NN9536 submission documents. The MedDRA version used was 23.1.

2 Abuse and Misuse

Custom query (NNMQ Abuse and Misuse):

- SMQ Drug abuse and dependence, narrow terms only
- HLT Intentional product misuses
- Additional PTs:
    - Poisoning deliberate
    - Intentional dose omission
    - Performance enhancing product use
    - Completed suicide
    - Intentional self-injury
    - Suicide attempt
    - Assisted suicide
    - Suspected suicide attempt
    - Suspected suicide.

3 Acute renal failure

SMQ Acute renal failure, narrow terms only

4 Allergic reactions

Custom Query (NNMQ Allergic reactions) – only narrow terms from the following:

- SMQ Anaphylactic reaction
- SMQ Angioedema
- SMQ Severe cutaneous adverse reactions
- SMQ Anaphylactic/anaphylactoid shock conditions
- SMQ Hypersensitivity
-

5 Cardiovascular disorders

Custom query (NNMQ Cardiovascular disorders). Broad and narrow terms from the following:

- SMQ Central nervous system vascular disorders
- SMQ Vasculitis
- SMQ Ischaemic heart disease
- SMQ Cardiac arrhythmias
- SMQ Cardiac failure
- SMQ Cardiomyopathy

- SMQ Embolic and thrombotic events
- SMQ Shock
- SMQ Torsade de pointes/QT prolongation

6 Drug-related hepatic disorders

SMQ Drug related hepatic disorders - comprehensive search

7 Gallbladder-related disorders

Custom query (NNMQ Gallbladder-related disorders). Narrow terms from the following:

- SMQ Functional, inflammatory and gallstone related biliary disorders
- SMQ Infectious biliary disorders

8 Gastrointestinal disorders

Custom query (NNMQ Gastrointestinal disorders SOC):

- SOC Gastrointestinal disorders, primary terms only

9 Hypoglycaemia

SMQ Hypoglycaemia, narrow terms only

10 Injection site reactions

Custom query (NNMQ Injection site reactions), both primary and secondary terms from the following:

- HLT Administration site reactions NEC
- HLT Application and instillation site reactions
- HLT Infusion site reactions
- HLT Injection site reactions

11 Malignant tumours

SMQ Malignant tumours

12 Medication errors

SMQ Medication errors.

13 Neoplasms

Custom query (NNMQ Neoplasms)

- SOC Neoplasms benign, malignant and unspecified (incl cysts and polyps), primary and secondary terms
- SMQ Biliary neoplasms
- SMQ Breast neoplasms, malignant and unspecified
- SMQ Liver neoplasms, benign (incl cysts and polyps)
- SMQ Liver neoplasms, malignant and unspecified
- SMQ Malignancies
- SMQ Malignant lymphomas
- SMQ Oropharyngeal neoplasms
- SMQ Ovarian neoplasms, malignant and unspecified
- SMQ Premalignant disorders
- SMQ Prostate neoplasms, malignant and unspecified
- SMQ Skin neoplasms, malignant and unspecified
- SMQ Uterine and fallopian tube neoplasms, malignant and unspecified

14 Overdose

Custom query (NNMQ Overdose):

- HLT Overdoses NEC
- Additional PTs:
    - Accidental overdose
    - Completed suicide
    - Suicide attempt
    - Suspected suicide attempt
    - Suspected suicide

15 Pancreatitis

Custom query (NNMQ Pancreatitis:

- SMQ Acute pancreatitis), narrow terms only
- HLT Acute and chronic pancreatitis, primary and secondary terms

16 Psychiatric disorders

Custom query:

- SOC Psychiatric disorders, primary terms only

17 Rare events

Custom query (NNMQ Rare events) excluding events that are included in other safety focus areas:

- SMQ Agranulocytosis, narrow terms only

- SMQ Guillain-Barre syndrome, narrow terms only
- SMQ Haematopoietic cytopenias affecting more than one type of blood cell, broad and narrow terms
- SMQ Haematopoietic leukopenia, broad and narrow terms
- SMQ Haematopoietic thrombocytopenia, narrow terms only
- SMQ Interstitial lung disease, narrow terms only
- SMQ Neuroleptic malignant syndrome, narrow terms only
- SMQ Pseudomembranous colitis, narrow terms only
- SMQ Retroperitoneal fibrosis, narrow terms only
- SOC Congenital, familial and genetic disorders, (all terms are primary PTs)
- HLT Angioedemas, primary and secondary routed PTs
- HLT Glomerulonephritis and nephrotic syndrome, primary and secondary routed PTs
- HLT Nephritis NEC, primary and secondary routed PTs
- Additional PTs:
    - Disseminated intravascular coagulation
    - Hepatic lymphocytic infiltration
    - Multiple organ dysfunction syndrome