Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT05564117
Sponsor trial ID:	NN9932-4954
Official title of study:	Efficacy and safety of oral semaglutide 25 mg once daily in adults with overweight or obesity (OASIS 4)
Document date*:	30 May 2024

*Document date refers to the date on which the document was most recently updated.

Statistical Analysis Plan

Efficacy and safety of oral semaglutide 25 mg once daily in adults with overweight or obesity (OASIS 4)

Semaglutide

[redacted]

Data Science - Biostatistics and Programming, Global Business Services (GBS)

Table of contents

Page

Table of tables

Page

Version History

This Statistical Analysis Plan (SAP) for study NN9932-4954 is based on the protocol version 1.0 dated 28 FEB 2022 and protocol version 2.0 dated 16 AUG 2022 (applicable to Germany only)

SAP Version	Date	Change	Rationale
1.0		Not Applicable	Original version

List of abbreviations

AD	*available but treatment discontinued*
AE	*adverse event*
ANCOVA	*analysis of covariance*
AT	*available on randomised treatment*
BMI	*body mass index*
bpm	*beats per minute*
CI	*confidence interval*
cm	*centimetre*
COA	*clinical outcome assessment*
CVD	*cardiovascular disease*
FAS	*full analysis set*
FPG	*fasting plasma glucose*
HbA1c	*glycated haemoglobin*
HDL	*high density lipoprotein*
ICH	*International Council on Harmonization*
kg	*kilogram*
LAO-OT	*last available observation during the on-treatment period*
LDL	*low density lipoprotein*
LR	*logistic regression*
MAR	*missing at random*
MD	*missing and treatment discontinued*
MedDRA	*medical dictionary for regulatory activities*
mg	*milligrams*
mg/dL	*milligrams per decilitre*
MI	*multiple imputation*
mmHg	*millimetre of mercury*
mmol/mol	*millimoles per mol*
MMRM	*mixed model for repeated measurements*
MT	*missing on randomised treatment*
OR	*odds ratio*
RD-MI	*multiple imputation using retrieved participants*
SAP	*statistical analysis plan*
SAS	*safety analysis set*
TFL	*tables, figures and listings*
VLDL	*very low density lipoprotein*

1 Introduction

Changes to the protocol-planned analyses are described in Section 4.8.

1.1 Objectives, Endpoints, and Estimands

The primary and secondary objectives of the study are listed below followed by an introduction of the estimands used to address the efficacy-related objectives.

The objectives and endpoints are summarised in **Table 1-1**.

Table 1-1 Objectives and Endpoints

Objectives	**Endpoints**		
Primary	Title	Time frame	Unit
• To confirm superior efficacy on body weight reduction from baseline (week 0) to end-of-treatment (week 64) of oral semaglutide 25 mg once daily versus placebo as an adjunct to reduced-calorie diet and increased physical activity in adults with overweight or obesity.	Co-primary:		
	Relative change in body weight	From baseline (week 0) to end of treatment (week 64)	%
	Achievement of body weight reduction ≥ 5% (Yes/No)	At end of treatment (week 64)	Count of participants
	Confirmatory Secondary:		
	Achievement of body weight reduction ≥ 10% (Yes/No)	At end of treatment (week 64)	Count of participants
	Achievement of body weight reduction ≥ 15% (Yes/No)	At end of treatment (week 64)	Count of participants
	Achievement of body weight reduction ≥ 20% (Yes/No)	At end of treatment (week 64)	Count of participants
Secondary	Title	Time frame	Unit
• To confirm superior efficacy on physical function from baseline (week 0) to end-of-treatment (week 64) of oral semaglutide 25 mg once daily versus placebo as an adjunct to reduced-calorie diet and increased physical activity in adults with overweight or obesity.	Confirmatory secondary		
	Change in Physical function domain (5-items) score (IWQOL-Lite-CT)	From baseline (week 0) to end of treatment (week 64)	Score points
	Supportive secondary		
	Change in IWQOL-Lite-CT PFD score ≥14.6 (Yes/No)	From baseline (week 0) to end-of-treatment (week 64)	Count of participant
• To estimate the efficacy on cardio-metabolic parameters from baseline (week 0) to end-of-treatment (week 64) of oral semaglutide 25 mg once daily versus placebo as an adjunct to reduced-calorie diet and increased physical activity	Supportive secondary		
	Change in body weight	From baseline (week 0) to end of treatment (week 64)	kg
	Change in body mass index (BMI)	From baseline (week 0) to end of treatment (week 64)	kg/m^2

Objectives	Endpoints		
in adults with overweight or obesity.	Change in waist circumference	From baseline (week 0) to end of treatment (week 64)	cm
	Change in systolic blood pressure	From baseline (week 0) to end of treatment (week 64)	mmHg
	Change in diastolic blood pressure	From randomisation (week 0) to end of treatment (week 64)	mmHg
	Change in HbA_{1c}	From baseline (week 0) to end of treatment (week 64)	% - point
	Change in lipids: • Total cholesterol • HDL cholesterol • LDL cholesterol • VLDL cholesterol • Triglycerides • Free fatty acids	From baseline (week 0) to end of treatment (week 64)	Ratio to baseline
	Change in high sensitivity C-Reactive Protein	From baseline (week 0) to end of treatment (week 64)	Ratio to baseline
	Change in fasting plasma glucose (FPG)	From baseline (week 0) to end of treatment (week 64)	mg/dL
	Change in fasting serum insulin	From baseline (week 0) to end of treatment (week 64)	Ratio to baseline
	BMI≥30 at baseline and BMI<30 at week 64 (yes/no)	From baseline (week 0) to end-of-treatment (week 64)	Count of participant
	Change in glycaemic status	From baseline (week 0) to end-of-treatment (week 64)	Count of participant
	Exploratory		
	Change in waist-height ratio (WtHR)	From baseline (week 0) to end-of-treatment (week 64)	No units
	BMI at week 64 < 27(Yes/No)	At end of treatment (week 64)	Count of participant
	Waist-height ratio (WtHR) at week 64 < 0.53 (Yes/No)	At end of treatment (week 64)	Count of participant
• To compare the safety and tolerability from baseline (week 0) to end-of-study (week 71) of oral semaglutide 25 mg once daily versus placebo as an adjunct to	Number of treatment emergent adverse events	From baseline (week 0) to end of study (week 71)	Count of events
	Number of serious treatment emergent adverse events	From baseline (week 0) to end of study (week 71)	Count of events

Objectives	Endpoints		
reduced-calorie diet and increased physical activity in adults with overweight or obesity.	Change in pulse	From baseline (week 0) to end-of-treatment (week 64)	Beats/min
Exploratory	Title	Time frame	Unit
• To estimate the efficacy on clinical outcome assessments baseline (week 0) to end-of-treatment (week 64) of oral semaglutide 25 mg once daily versus placebo as an adjunct to reduced-calorie diet and increased physical activity in adults with overweight or obesity.	Change in IWQOL-Lite-CT • Physical domain score • Psychosocial domain score • Total score	From baseline (week 0) to end of treatment (week 64)	Score points
	Change in Impact of Weight on Daily Activities Questionnaire (IWDAQ) total score	From randomisation (week 0) to end-of-treatment (week 64)	Score points
	Change in Control of Eating Questionnaire (COEQ): • Craving Control domain score • Positive Mood domain score • Craving for Sweet domain score • Craving for Savoury domain score • Hunger item score • Fullness item score	From baseline (week 0) to end of treatment (week 64)	Score points
	Change in Three-factor eating Questionnaire: • cognitive restraint • uncontrolled eating • emotional eating	From baseline (week 0) to end of treatment (week 64)	Score points

The estimands for the primary objective are introduced below and the attributes of the estimands are summarised in Table 1-2. Additional details are available in Section 4.

The secondary estimands with confirmatory and supportive secondary endpoints are similar to the co-primary estimands except for the endpoint attribute. The secondary estimands with continuous endpoints for secondary objectives are similar to the co-primary estimand for relative weight change, with the exception of endpoints with units of ratio to baseline, for which the population-level summary is the ratio between treatment conditions. The secondary estimands with binary endpoints for secondary objectives are similar to the co-primary estimand for body weight reduction ≥ 5%.

Table 1-2 Estimands

Objective	Attributes

Estimand category		Treatment condition	Variables / endpoints	Population of interest	Intercurrent events and strategy	Population-level summary measure
Primary objective: To confirm superior efficacy on body weight reduction from baseline (week 0) to end-of-treatment (week 64) of oral semaglutide 25 mg once daily versus placebo as adjuncts to reduced-calorie diet and increased physical activity in adults with overweight or obesity	Co-primary	The effect of oral semaglutide and regardless of initiating other anti-obesity therapies (weight management drugs or bariatric surgery) versus the effect of placebo and regardless of initiating other anti-obesity therapies (weight management drugs or bariatric surgery) each as adjunct to reduced-calorie diet and increased physical activity	From baseline to week 64: • Relative change in body weight • Achievement of body weight reduction ≥ 5% (yes/no)	Participants with overweight or obesity	**Treatment policy** strategy for: • Premature randomised treatment discontinuation • Initiation of rescue intervention (other anti-obesity therapies or bariatric surgery)	• Difference in means in relative weight change • Treatment odds-ratio for proportion of participants achieving body weight reduction ≥ 5%
	Additional*	The effect of oral semaglutide without other anti-obesity therapies versus the effect of placebo without other anti-obesity therapies , each as adjunct to reduce-			**Hypothetical strategy** for: • Premature randomised treatment discontinuation • Initiation of rescue intervention (other anti-obesity therapies or	

		calorie diet and increased physical activity			bariatric surgery)	

* Not related to the confirmatory hypotheses

1.2 Study Design

The study design is provided in protocol section 4.1.

2 Statistical Hypotheses

The confirmatory hypotheses to be tested are superiority of oral semaglutide 25 mg once daily vs placebo for the co-primary endpoints, as well as superiority of oral semaglutide 25 mg once daily vs placebo for the confirmatory secondary endpoints. Superiority needs to be confirmed for both primary endpoints to confirm the primary objective.
For the primary estimands with the co-primary endpoints, 1) change in body weight (%) from baseline to end of treatment (week 64) and 2) ≥ 5% body weight reduction (yes/no) from baseline (week 0) to end of treatment (week 64), the following confirmatory 1-sided hypotheses are planned to be tested for semaglutide 25 mg versus placebo. Operationally, the hypotheses will be evaluated by 2-sided tests.

The superiority tests of oral semaglutide 25 mg once daily vs. placebo will be carried out as follows:
Let $\mu_{semaglutide}$ and $\mu_{placebo}$ denote the true mean of % weight change for oral semaglutide 25 mg once daily and placebo group, respectively. The hypothesis and alternative hypothesis tested are

$$H_0: \mu_{semaglutide} \geq \mu_{placebo} \; vs$$
$$H_A: \mu_{semaglutide} < \mu_{placebo}$$

The null hypothesis will be rejected, and superiority claimed, if the upper limit of the estimated two-sided 95% CI for the true treatment difference ($\mu_{semaglutide}$ minus $\mu_{placebo}$) is below 0.
Let $OR_{semaglutide/placebo}$ denote the true odds ratio between oral semaglutide 25 mg once daily and placebo. The hypothesis and alternative hypothesis tested are:

$$H_0: OR_{semaglutide/placebo} \leq 1 \; vs$$
$$H_A: OR_{semaglutide/placebo} > 1.$$

The null hypothesis will be rejected, and superiority claimed, if the lower limit of the estimated two-sided 95% CI is above 1.

2.1 Multiplicity Adjustment

The type I error will be controlled in the strong sense using a hierarchical (fixed sequence) testing procedure. This is based on priority ordering of the null hypotheses and testing them in this order using the 2-sided 95% confidence interval approach until an insignificant result appears. Consequently, all previous hypotheses must be rejected in order to proceed to the next hypothesis test thereby preserving the type I error rate. The order of hypothesis tests is shown in Table 2-1. All hypothesis tests are for the superiority of semaglutide 25 mg versus placebo.

Table 2-1 Hierarchical order for hypothesis testing

Test order	Endpoint	Target	Comparator
1	Relative change in body weight	Semaglutide	Placebo
2	Achievement of body weight reduction ≥ 5% (Yes/No)	Semaglutide	Placebo

3	Achievement of body weight reduction ≥ 10% (Yes/No)	Semaglutide	Placebo
4	Achievement of body weight reduction ≥ 15% (Yes/No)	Semaglutide	Placebo
5	Achievement of body weight reduction ≥ 20% (Yes/No)	Semaglutide	Placebo
6	Change in IWQOL-Lite for CT Physical Function Score	Semaglutide	Placebo

3 Analysis Sets

The following populations are defined:

Table 3-1 Analyses Sets

Participant Analysis Set	Description
Full analysis set (FAS)	All participants randomised. Participants will be analysed according to the randomised treatment.
Safety analysis set (SAS)	All participants randomly assigned to study treatment and who take at least 1 dose of IMP. Participants are analysed according to the treatment they actually received.

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

Two observation periods are defined for each participant:

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

On-treatment (with IMP): In general, the on-treatment period will be from the date of first IMP administration to date of last IMP administration plus three days, except when randomised treatment is temporarily discontinued. If randomised treatment is temporarily discontinued, the on-treatment period ends 3 days after the treatment discontinuation and resumes on the day randomised treatment is resumed. Hence, the on-treatment period can consist of several disjoint periods.

In general, the on-treatment period will therefore be from the date of first IMP administration to date of last IMP administration excluding potential off-treatment time intervals of more than 3 consecutive days.

For the evaluation of adverse events, the lag time for each on-treatment time interval is 7 weeks.

The in-study 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.

The estimands using the treatment policy strategy for intercurrent events use all data from the in-trial observation period and estimands using the hypothetical strategy use data from the on treatment observation period until first treatment discontinuation or initiation of other anti-obesity therapy.

4 Statistical Analyses

4.1 General Considerations

Results from statistical analyses will generally be accompanied by two-sided 95% confidence intervals (CIs) and corresponding p-values. Superiority will be claimed if p-values are less than 5% and the estimated treatment contrasts favours oral semaglutide 25 mg once daily.

Taxonomy of week 64 assessments

For each participant, a given week 64 assessment may be available or missing as specified in Table 4-1. The assessment availability is defined by participant and by assessment; thus, for body weight at week 64, a participant may be characterised as 'available on randomised treatment (AT)', whereas for waist circumference, the participant may be characterised as 'missing on randomised treatment (MT)'.

Table 4-1 Taxonomy for participants based on week 64 assessments

Availability	Participants on randomised treatment at week 64	Description	Abbreviation
Available	Yes	**Available on randomised treatment**: Participants who complete the study on randomised treatment with an assessment at week 64: Includes those that stop and restart IMP.	AT
	No	**Available but discontinued:** Participants who discontinued randomised treatment prematurely but returned to have an assessment at week 64. These are also called retrieved participants	AD
Missing	Yes	**Missing on randomised treatment**: Participants who complete the study on randomised treatment without an assessment at week 64: Includes those that stop and restart IMP.	MT
	No	**Missing and discontinued**: Participants who discontinued randomised treatment prematurely and did not return to have an assessment at week 64. These are also called non-retrieved participants	MD

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 value at randomisation across all participants is used as the baseline value.

4.2 Primary endpoints/Estimands Analysis

4.2.1 Definition of Endpoints

Relative change from baseline (week 0) to week 64 in body weight (%)

Relative change from baseline (week 0) to week 64 in body weight (%) is defined as:

$$\% \text{ weight change} = \frac{(\text{body weight at week 64} - \text{body weight at baseline})}{\text{body weight at baseline}} \times 100.$$

Achievement of body weight reduction ≥ 5%
A body weight reduction of at least X% from baseline (week 0) to week 64 is defined as

$$\text{X\% responder} = \begin{cases} 1 \text{ if } \% \text{ weight change} \leq -\text{X\%} \\ 0 \text{ if } \% \text{ weight change} > -\text{X\%} \end{cases}$$

4.2.2 Main Analytical Approach

Analyses addressing co-primary estimands

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 a covariate. The estimated treatment differences between oral semaglutide 25 mg once daily and semaglutide placebo will be reported together with the associated two-sided 95% confidence intervals (CI) and corresponding pvalues.

The analysis model for the 5% responder endpoint is a logistic regression using randomised treatment as a factor and baseline body weight (kg) as covariate. The estimated odds ratios (OR) between oral semaglutide 25 mg once daily and semaglutide placebo will be reported together with the associated two-sided 95% CIs and corresponding p-values.

Where response rates close to 0% or 100% in any treatment group lead to non-convergence, Firth's maximum-likelihood estimation will be used when performing the logistic regression.

In addition to the estimated OR, the estimated treatment differences (ETDs) will be provided by calculating the responder probabilities and treatment differences between responder probabilities based on the logistic regression model, with confidence intervals for treatment differences obtained using the delta method.

Handling of missing week 64 values for the treatment policy strategy

All available data at week 64 (AT and AD) are used and missing values (MT and MD) at week 64 will be imputed and the endpoints will be derived from the imputed values. Several approaches for imputation will be applied. First, a description of the primary imputation approach to address the estimands applying the treatment policy strategy for the primary endpoints is given followed by a description of the sensitivity analyses used to assess the robustness of the conclusions. The sensitivity analyses investigate how assumptions on body weight development after discontinuation of randomised trial treatment impact the estimated treatment contrasts between oral semaglutide 25 mg once daily and semaglutide placebo.

Primary imputation approach for the treatment policy strategy

Multiple imputation approach using retrieved drop-outs (RD-MI): The primary imputation approach for estimands applying the treatment policy strategy is a multiple imputation similar to the one described by McEvoy[1]. Missing body weight measurement at week 64 for non-retrieved participants (MD) are imputed using assessments from retrieved participants (AD) in each randomised treatment arm. This will be done according to the timing of last available observation on treatment (LAO-OT) of body weight prior to week 64. Missing body weight measurements at week 64 for participants on randomised treatment (MT) are imputed by sampling from available measurements at week 64 from participants 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 participants (AD) from FAS and applied 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 64 with gender (male/female) as factor and baseline body weight (kg), timing of LAO-OT of body weight and LAO-OT of body weight (kg) as covariates. No interactions will be included. If the imputation model cannot be fit the model will be reduced until the model can be fit. Reduction will be done in a fixed order by first removing gender and then removing baseline body weight (kg). If the imputation model with only LAO-OT of body weight (kg) cannot be fit, the imputation will be done regardless of the randomised treatment arm. If no LAO-OT exists post-baseline then LAO-OT will be the baseline body weight. If any participants are MT, an imputation model for missing body weight measurements at week 64 for MT participants will also be defined using AT participants 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 64 body weight values for each randomised treatment arm. This will be done 1,000 times and results in 1,000 complete data sets.

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

2. **Pooling**: The results obtained from analysing the datasets will be combined using Rubin's formula[2].

The multiple imputations will be generated using Novo Nordisk study number 99324954 as seed number.

Analysis addressing the hypothetical strategy

The estimand applying the hypothetical strategy for % weight change addresses the efficacy of semaglutide 25 mg once daily and will be assessed using a 'MMRM for efficacy'. Week 64 assessments for retrieved drop-outs (AD) are not used in this analysis. The MMRM for efficacy will use assessments only from participants who are taking the randomised treatment until end-of-treatment or until first discontinuation of randomised treatment. The date of the last dose before first discontinuation of randomised treatment plus 3 days will be used as the latest date for using assessments in this MMRM. The assessment closest in time and before the date of the last dose

before first discontinuation of randomised treatment plus 3 days will be used as last assessment on randomised treatment.

For participants who initiate any other anti-obesity therapies before completion or first discontinuation 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. Similarly, the assessment closest in time and before the date of starting weight management drugs or undergoing bariatric surgery will be used as last assessment on randomised treatment. The MMRM for efficacy will be fitted using % weight change with randomised treatment as a factor and baseline body weight (kg) as a covariate all nested within visit. An unstructured covariance matrix for measurements within the same participant will be employed, assuming that measurements for different participants are independent.

The estimand applying the hypothetical strategy for 5% responders will be assessed using the same MMRM for efficacy. For participants with missing body weight at week 64, individual values for body weight will be predicted from the MMRM and used to classify each subject as 5% responder or not. This classification will then be analysed using a logistic regression model with randomised treatment as a factor and baseline body weight (kg) as covariate.

4.2.3 Sensitivity Analysis

Jump to reference multiple imputation approach (J2R-MI): Missing values of body weight at week 64 (MT and MD) for all treatment groups are imputed by sampling among all available assessments at week 64 in the placebo group (AT and AD). This approach is based on the assumption that participants 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[3]. The multiple imputation approach is done as above with the first imputation step replaced by the following:

1. **Imputation**: Defines an imputation model using placebo participants from FAS with a week 64 measurement (AT and AD). The model will be a linear regression of body weight (kg) at week 64 with gender (male/female) as factor 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 by removing gender. The estimated posterior distribution for the parameters (regression coefficients and variances) in the imputation models are then used to impute missing week 64 body weight values for each randomised treatment arm. This will be done 1,000 times and results in 1,000 complete data sets.

The jump to reference approach is the basis for the sample size calculations.

Tipping-point multiple imputation analysis (TP-MI): This sensitivity analysis evaluates the robustness of the superiority conclusions to violations of the MAR assumption. First, missing body weight data are imputed according to the primary multiple imputation approach for the treatment policy strategy. Then, a penalty is added to the imputed values at week 64. The approach is to explore a range of penalties for both treatment groups 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 treatment groups.

Mixed model for repeated measurements (MMRM): This 'MMRM for effectiveness' will use all assessments regardless of adherence to randomised treatment, including assessments at week 64 for retrieved drop-outs (AD). The MMRM for effectiveness will be fitted using the same factor and covariate as for the primary analyses all nested within visit. An unstructured covariance matrix for measurements within the same participant will be employed, assuming that measurements for different participants are independent. For the 5% responder analysis, the same MMRM will be applied except that body weight (kg) will be used as response variable in the model. Individual missing values for body weight at week 64 will be predicted from the MMRM and used to classify each subject as 5% responder or not. This classification will then be analysed using the same logistic regression model as in the primary analysis.

Non-retrieved participants as non-responders: For binary responder endpoints an analysis using non-retrieved participants as non-responders in the logistic regressions will be done.

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

Table 4-2 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	Strategy	Analysis set	Statistical model	Imputation approach	Sensitivity analyses
Primary endpoints								
Primary	% weight change	1	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	J2R-MI TP-MI MMRM
				Hypothetical	FAS	MMRM	-	-
Primary	5% responders	2	Binary	Treatment policy	FAS	LR	RD-MI	J2R-MI TP-MI MMRM Non-responder
				Hypothetical	FAS	LR	MMRM	-
Confirmatory secondary endpoints								
Primary	10% responders	3	Binary	Treatment policy	FAS	LR	RD-MI	J2R-MI MMRM Non-responder
				Hypothetical	FAS	LR	MMRM	-
Primary	15% responders	4	Binary	Treatment policy	FAS	LR	RD-MI	J2R-MI MMRM Non-responder
Primary	20% responders	5	Binary	Treatment policy	FAS	LR	RD-MI	J2R-MI MMRM Non-responder
				Hypothetical	FAS	LR	MMRM	-
Secondary	IWQOL-Lite-CT PFD score change	6	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	J2R-MI MMRM
				Hypothetical	FAS	MMRM	-	-

FAS = full analysis set; ANCOVA = analysis of covariance; RD-MI = multiple imputation using retrieved participants; J2R-MI = jump to reference multiple imputation; TP-MI = tipping point multiple imputation; MMRM = mixed model for repeated measurements; LR = logistic regression; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite for Clinical Trials; PFD = physical function domain;

Test order refers to the order of the endpoint in the statistical test hierarchy outlined in Table 2-1

4.3 Secondary endpoints Analyses

4.3.1 Confirmatory Secondary endpoints

4.3.1.1 Main Analytical Approach

The continuous confirmatory secondary endpoints will be analysed using the same analysis and imputation models as used to address the primary estimand for the primary continuous endpoint.

The binary confirmatory secondary endpoints will be analysed using the same analysis and imputation models as used to address the primary estimand for the primary binary endpoint and also the same sensitivity analyses will be done.

4.3.1.2 Sensitivity Analyses

The confirmatory secondary endpoints will be analysed using the same sensitivity analyses as used to address the primary estimand. The tipping-point sensitivity analysis will not be performed for confirmatory secondary endpoints.

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

4.3.2 Supportive Secondary Endpoints

All supportive secondary endpoints are tested for superiority of oral semaglutide 25 mg versus semaglutide placebo. Supportive secondary endpoints are listed in Section 1.1.

Analyses addressing the treatment policy strategy

The effect-related supportive secondary endpoints will be analysed using the same imputation approach as used for the primary endpoints using a treatment policy. The imputation model is the same as for the primary endpoints with body weight replaced by assessments of the endpoint to be analysed. 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 endpoints where the endpoint is ‘ratio to baseline’, e.g. lipids, 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.

The statistical model for responder endpoints relating to COAs will be logistic regression with randomised treatment as a factor and the baseline assessment of the endpoint to be analysed as covariate.

Analyses addressing the hypothetical strategy

The supportive secondary endpoints which relate to the primary objective will be analysed using a hypothetical strategy with the same MMRM for efficacy 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

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 3). 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 27.0).

An overview of all analysis and imputation methods to address the secondary estimands for supportive secondary endpoints is given in Table 4-3.

Table 4-3 Analysis and imputation for supportive secondary endpoints

Objective	Endpoint	Endpoint type	Strategy	Analysis set	Statistical model	Imputation approach	Sensitivity analyses
Supportive secondary endpoints (effect related)							
Primary	Weight change (kg)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Primary	Waist circumference (cm)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Primary	BMI change (kg/m^2)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	HbA_{1c} change (%, mmol/mol)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	FPG change (mg/dL, mmol/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	Fasting serum insulin change (mIU/L, pmol/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	sBP change (mmHg)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	dBP change (mmHg)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary		Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-

Objective	Endpoint	Endpoint type	Strategy	Analysis set	Statistical model	Imputation approach	Sensitivity analyses
	Total cholesterol change (mg/dL, mmol/L)		Hypothetical	FAS	MMRM	-	-
Secondary	HDL change (mg/dL, mmol/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	LDL change (mg/dL, mmol/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	VLDL change (mg/dL, mmol/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	FFA change (mg/dL, mmol/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	Triglycerides change (mg/dL, mmol/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	hsCRP change (mg/L)	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	IWQOL-Lite-CT PFD score responder[#]	Binary	Treatment policy	FAS	LR	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	BMI responder[##]	Binary	Treatment policy	FAS	LR	RD-MI	-
			Hypothetical	FAS	MMRM	-	-
Secondary	Glycaemic status change[####]	Categorical	Treatment policy	FAS	-	-	-
Supportive secondary endpoints (safety related)							
Secondary	Number of TEAEs	Continuous	-	SAS	-	-	-
Secondary	Number of SAEs	Continuous	-	SAS	-	-	-
Secondary	Pulse change (bpm)	Continuous	Hypothetical	SAS	MMRM	-	-

FAS = full analysis set; ANCOVA = analysis of covariance; RD-MI = multiple imputation using retrieved participants; MMRM = mixed model for repeated measurements; BMI = body mass index; HbA_{1c} = Haemoglobin A_{1c}; FPG = fasting plasma glucose; dBP = diastolic blood pressure; HDL = high density lipoprotein; LDL = low density lipoprotein; VLDL = very low density lipoprotein; FFA = free fatty acids; LR = logistic regression; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite for Clinical Trials; PFD = physical function domain; TEAEs = treatment emergent adverse events; SAEs = serious adverse events; # responder value = 14.6; ## Analysis performed in participants with BMI≥30 at baseline and response defined as BMI<30 at week 64 (yes/no);

Shift in glycaemic status will be presented as descriptive statistics.

4.4 Exploratory Endpoints Analyses

Exploratory estimands similar to the co-primary and secondary estimands are used to address the exploratory endpoints summarised in Table 1-1. Analysis and imputation approach for the exploratory and corresponding additional estimands are summarised in Table 4-4.

Table 4-4 Analysis and imputation of exploratory endpoints

Objective	Endpoint	Endpoint type	Strategy	Analysis set	Statistical model	Imputation approach	Sensitivity analyses	Other
Exploratory endpoints (effect related)								
Exploratory	IWQOL-Lite-CT PD score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	IWQOL-Lite-CT PSD score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	IWQOL-Lite-CT total score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	COEQ Craving control domain score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	COEQ Positive Mood domain score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	COEQ Craving for Sweet domain score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	COEQ Craving for Savoury domain Score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	COEQ Hunger item score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	COEQ Fullness item score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	TFEQ Cognitive restraint score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	TFEQ Uncontrolled eating score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	TFEQ Emotional eating score change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	WtHR change	Continuous	Treatment policy	FAS	ANCOVA	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	WtHR responder[#]	Binary	Treatment policy	FAS	LR	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	BMI responder[##]	Binary	Treatment policy	FAS	LR	RD-MI	-	
			Hypothetical	FAS	MMRM	-	-	
Exploratory	IWDAQ total score change	Continuous	-	FAS	-	-	-	Descriptive statistics

FAS = full analysis set; ANCOVA = analysis of covariance; RD-MI = multiple imputation using retrieved participants; MMRM = mixed model for repeated measurements; BMI = body mass index; HbA_{1c} = Haemoglobin A_{1c}; FPG = fasting plasma glucose; dBP = diastolic blood pressure; HDL = high density lipoprotein; LDL = low density lipoprotein; VLDL = very low density lipoprotein; FFA = free fatty acids; LR = logistic regression; IWQoL-Lite = Impact of Weight on Quality of Life-Lite for Clinical Trials; PD = physical domain; PSD = psychosocial domain; IWDAQ = Impact of Weight on Daily Activities Questionnaire; WtHR = Waist-Height Ratio; # Analysis performed in participants with WtHR<0.53; ## Analysis performed in participants with BMI<27

The IWDAQ is a measure under development and validation. OASIS 4 data will be used for psychometric evaluation of the IWDAQ including development of a scoring algorithm. Calculation of the IWDAQ total score is based on a preliminary scoring of the measure. The IWDAQ total score will only be presented as descriptive statistics.

4.5 Other Safety Analyses

Observed data for safety assessments will be summarised by descriptive statistics.

4.6 Other Analyses

All collected data that were not defined as endpoints will be summarised by descriptive statistics.

4.7 Interim Analysis

Not applicable for this study.

4.8 Changes to Protocol-planned Analysis

- Updated estimand description including:
    - Updated naming of primary estimand for co-primary endpoints to co-primary estimands.
    - Addition of estimand attributes.
    - Imputation approach has been updated based on the strategy for handling intercurrent events instead of estimand.
- In general “subject” is replaced by “participant” when referring to the study population. For the analysis of binary endpoints the following additions have been added:
  Where response rates close to 0% or 100% in any treatment group lead to non-convergence, Firth’s maximum-likelihood estimation will be used when performing the logistic regression.
- In addition to the estimated OR, the estimated treatment differences (ETDs) will be provided by calculating the responder probabilities and treatment differences between responder probabilities based on the logistic regression model, with confidence intervals for treatment differences obtained using the delta method.
- It has been clarified how the imputation model in the primary imputation approach for the treatment policy strategy will be reduced if the model cannot fit. It has also been clarified how the J2R-MI is reduced if the imputation model cannot fit.
- It has been clarified that the J2R-MI sensitivity will be aligned with the RD-MI and not include BMI group as a factor in the imputation model.
- It has been clarified how the 5% responder analysis will be assessed using the MMRM for efficacy and MMRM for effectiveness.
- The following supportive secondary endpoints have been included in Table 4-3
    - Body weight, change from baseline, kg
    - IWQOL-Lite-CT PFD score responders

- BMI responders (BMI≥30 at baseline and BMI<30 at week 64 (yes/no))
- Glycaemic status change
- Pulse change

Additionally, it has been clarified that change from baseline (mg/dL) will be reported for FPG instead of ratio to baseline.

- The following exploratory endpoints have been included in Table 4-4
  - Change in Three-factor eating Questionnaire (cognitive restraint, uncontrolled eating, emotional eating)
  - Change in waist-height ratio (WtHR)
  - BMI responders (BMI < 27 at week 64 (yes/no))
  - WtHR responders (WtHR < 0.53 at week 64 (yes/no))
- The tipping-point sensitivity analysis will only be performed for the primary endpoints.
- Changed the on-treatment period description from "*the date of first trial product administration to date of last trial product administration excluding potential off-treatment time intervals of at* ***least*** *3 consecutive days*" to "*the date of first trial product administration to date of last trial product administration excluding potential off-treatment time intervals of* ***more than*** *3 consecutive days.*"
- A description of the sensitivity analysis of the 5% responder endpoint (primary estimand) using MMRM has been included.

5 Sample size determination

See protocol section 9.5.

6 References

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

2. Mislevy RJ, Little RJA, Rubin DB. Statistical Analysis with Missing Data. Journal of Educational Statistics. 1991;16(2).

3. Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. J Biopharm Stat. 2013;23(6):1352-71.