Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT03811574
Sponsor trial ID:	NN9536-4382
Official title of study:	Effect and safety of semaglutide once-weekly in East Asian subjects with overweight or obesity
Document date*	19 March 2021

*Document date refers to the date on which the document was most recently updated.
Note: The date in the header of Page 2 is the date of compilation of the documents and not of an update to content.

16.1.9 Documentation of statistical methods

List of contents

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

.

[redacted] [redacted]

.

NN9536-4382 (STEP 6)

Statistical Analysis Plan

[redacted]
Clinical Statistics

[redacted]

Table of contents

Page

Table of figures

Page

Table of tables

Page

Version history

This Statistical Analysis Plan (SAP) for trial NN9536-4382 is based on the protocol version 2 dated 08 August 2018.

Table 1 SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	14 January 2021	Not Applicable	Original version
2		Clarification of visceral fat area used for analysis	This SAP version 2 is made after unblinding, but prior to a post-database lock update of visceral fat area (VFA) data. The purpose of version 2 is to clarify prior to this update which data points are to be used to define the VFA endpoints.

1 Introduction

This SAP includes 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 3.2.

Specifications of tables, figures and listings (TFL) and other specifications not included in this SAP will be described in the mock TFL.

1.1 Objectives and endpoints

1.1.1 Primary, secondary and exploratory objectives

1.1.1.1 Primary objective

To compare the effect of semaglutide s.c. 2.4 mg once-weekly versus semaglutide placebo as an adjunct to a reduced calorie diet and increased physical activity in subjects with overweight or obesity on body weight.

1.1.1.2 Secondary objectives

To compare the effect of semaglutide s.c. 2.4 mg once-weekly versus semaglutide placebo as an adjunct to a reduced calorie diet and increased physical activity in subjects with overweight or obesity on:

- Cardiovascular risk factors
- Glucose metabolism
- Clinical Outcome Assessments (COA)
- Other factors related to body weight

To compare the safety and tolerability of semaglutide s.c. 2.4 mg once-weekly versus semaglutide placebo as an adjunct to a reduced calorie diet and increased physical activity in subjects with overweight or obesity.

To compare the effect of semaglutide s.c. 1.7 mg once-weekly versus semaglutide placebo as an adjunct to a reduced calorie diet and increased physical activity in subjects with overweight or obesity on body weight.

To compare the effect of semaglutide s.c. 1.7 mg once-weekly versus semaglutide placebo as an adjunct to a reduced calorie diet and increased physical activity in subjects with overweight or obesity on:

- Cardiovascular risk factors
- Glucose metabolism
- Clinical Outcome Assessments (COA)
- Other factors related to body weight

To compare the safety and tolerability of semaglutide s.c. 1.7 mg once-weekly versus semaglutide placebo as an adjunct to a reduced calorie diet and increased physical activity in subjects with overweight or obesity.

1.1.1.3 Exploratory objectives

To compare the effect of semaglutide s.c. 2.4 mg and 1.7 mg once-weekly versus semaglutide placebo as an adjunct to a reduced calorie diet and increased physical activity in subjects with overweight or obesity on:

- Glycaemic status (apply to subjects with no T2D at baseline)
- Use of medication for hypertension and dyslipidaemia
- Use of oral antidiabetic drug (OAD) (apply to subjects with T2D at baseline)
- Treatment discontinuation

1.1.2 Estimands

1.1.2.1 Primary estimand

The estimand will quantify the average treatment effect of semaglutide relative to semaglutide placebo after 68 weeks, as an adjunct to a reduced calorie diet and increased physical activity, in all randomised subjects regardless of adherence to treatment and regardless of initiating other anti-obesity therapies (weight management drugs or bariatric surgery) ("treatment policy" estimand). The estimand will cover all effect-related objectives.

1.1.2.2 Secondary estimand

The estimand will quantify the average treatment effect of semaglutide relative to semaglutide placebo 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 initiated other anti-obesity therapies (weight management drugs or bariatric surgery) ("hypothetical" estimand). The estimand will cover all effect-related objectives.

1.1.3 Primary, secondary and exploratory endpoints

All endpoints are being compared between semaglutide 2.4 mg vs placebo and semaglutide 1.7 mg vs placebo.

1.1.3.1 Primary endpoint

The primary endpoints addressing the primary and secondary objectives:

- Change from baseline at week 0 to week 68 in body weight (%)
- Subjects who after 68 weeks achieve (yes/no):
  - Body weight reduction ≥ 5% from baseline at week 0

1.1.3.2 Secondary endpoints

The confirmatory and supportive secondary endpoints addressing the primary and secondary objectives are listed below.

Confirmatory secondary endpoints

- Subjects who after 68 weeks achieve (yes/no):
  - Body weight reduction ≥ 10% from baseline at week 0
  - Body weight reduction ≥ 15% from baseline at week 0
- Change from baseline at week 0 to week 68 in:
  - Waist circumference (cm) measured midway between the lower rib margin and the iliac crest

Supportive secondary endpoints

Effect endpoints

- Change from baseline at week 0 to week 68 in:
  - Body weight (kg)
  - BMI (kg/m^2)
  - Waist circumference (cm) measured according to the JASSO guideline[1]
  - Visceral Fat Area (VFA) (%, cm^2) measured by CT scan in a subset of the Japanese trial population
  - HbA1c (%, mmol /mol)
  - Fasting plasma glucose (FPG) (mg/dL)
  - Fasting serum insulin (µIU/mL)
  - Systolic blood pressure (mmHg)
  - Diastolic blood pressure (mmHg)
  - Lipids (mg/dL)
    - Total cholesterol
    - High density lipoprotein (HDL) cholesterol
    - Low density lipoprotein (LDL) cholesterol
    - Very low density lipoprotein (VLDL) cholesterol
    - Free fatty acids
    - Triglycerides
  - High sensitivity C-Reactive Protein (hsCRP) (mg/L)
  - Plasminogen Activator Inhibitor-1 (PAI-1) (AU/mL)
  - Short Form-36 (SF-36) (range of score 1-100)
    - role-physical score
    - bodily pain score
    - general health score
    - vitality score
    - physical functioning score
    - social functioning score
    - role-emotional score
    - mental health score
    - physical component summary
    - mental component summary
  - Impact of Weight on Quality of Life-Lite for Clinical Trials (IWQoL-Lite for CT) (range of score 1-20)

- physical function domain (5-items) score
- pain/discomfort domain score
- psychosocial domain score
- total score

- Subjects who after 68 weeks achieve (yes/no):
  - Responder definition value for SF-36 physical functioning score
  - Responder definition value for IWQoL-Lite for CT physical function domain (5-items) score
  - Body weight reduction ≥ 20% from baseline at week 0

The following supportive secondary endpoints are used for subjects with T2D at baseline:

- Subjects who after 68 weeks achieve (yes/no):
  - HbA1c < 7.0% (53 mmol/mol)
  - HbA1c ≤ 6.5% (48 mmol/mol)

Safety endpoints

- Number of treatment emergent adverse events (TEAEs) from baseline at week 0 to week 75
- Number of serious adverse events (SAEs) from baseline at week 0 to week 75
- Number of treatment emergent severe or blood glucose confirmed symptomatic hypoglycaemia episodes (yes/no) from baseline to week 75 (only apply to subjects with T2D at baseline)
- Change from baseline at week 0 to week 68 in:
  - Pulse (bpm)
  - Amylase (U/L)
  - Lipase (U/L)
  - Calcitonin (ng/L)
  - QTcF interval (corrected QT interval using Fridericia's correction) (msec)

1.1.3.3 Exploratory endpoints

The exploratory endpoints addressing the explorative objectives:

- Change from baseline at week 0 to week 68 in:
  - Glycaemic category (normo-glycaemia, pre-diabetes, T2D) (only apply to subjects with no T2D at baseline)
  - Antihypertensive medication (decrease, no change, increase)
  - Lipid-lowering medication (decrease, no change, increase)
  - Concomitant OAD (decrease, no change, increase) (only apply to subjects with T2D at baseline)
- Subjects who from randomisation at week 0 to week 68 have permanently discontinued randomised trial product (yes/no)
- Time to permanent discontinuation of randomised trial product (weeks)

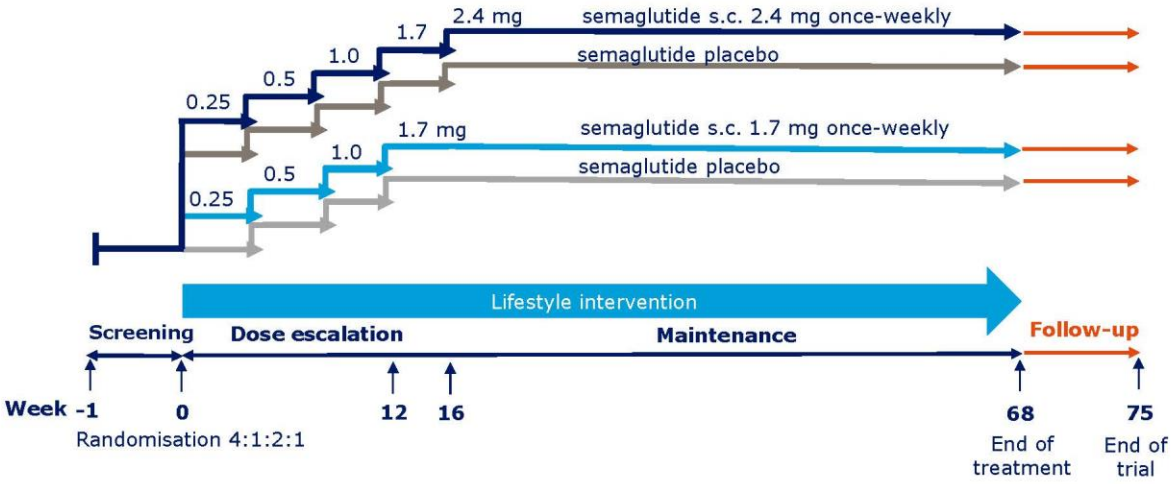
1.2 Trial design

This is a 68-week, randomised, double-blind, placebo-controlled, four-armed, parallel group, multicentre, multinational clinical trial comparing semaglutide s.c. 2.4 mg once-weekly with semaglutide placebo once-weekly and semaglutide s.c. 1.7 mg once-weekly with semaglutide placebo once-weekly in subjects with overweight or obesity.

The trial includes a screening visit to assess the subject's eligibility, randomisation visit (week 0), followed by visits/phone contacts every 2nd week during dose escalation. From week 20, visits/phone contacts will take place every 4th week for the remaining maintenance period until end of treatment (week 68). A follow-up visit ('End of trial') for safety assessments is scheduled 7 weeks after end of treatment to account for the exposure to the long half-life of semaglutide.

The trial design is outlined in Figure 1.

Figure 1 A schematic diagram of the trial design with the escalation, duration of the trial periods including follow-up period and intervention



Eligible subjects fulfilling all randomisation criteria at visit 2 will be randomised in a 4:1:2:1 manner to receive either semaglutide s.c. 2.4 mg once-weekly, semaglutide placebo once-weekly, semaglutide s.c. 1.7 mg once-weekly or semaglutide placebo once-weekly as an adjunct to a reduced calorie diet and increased physical activity.

A subset of maximum 180 randomised Japanese subjects will have VFA assessed by CT scan at selected sites at randomisation and end-of-treatment to demonstrate the size of VFA after 68 weeks of treatment in accordance with JASSO guideline.[1] A maximum of 25 % of all subjects undergoing CT scan are expected to have T2D corresponding to 45 subjects. Randomisation will be stratified according to T2D diagnosis at screening and planned CT scan.

2 Statistical considerations

Taxonomy of week 68 assessments

For each subject a given assessment at week 68 may be available or missing and Table 2 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 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 sample size and thereby the power for this trial is primarily defined to support safety. However, no formal statistical inference is planned based on number of adverse events. Given the trial sample size, the power of statistical tests for effect endpoints is described below.

The tests of superiority of semaglutide 2.4 mg to semaglutide placebo or semaglutide 1.7 mg to semaglutide placebo 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 3 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. As the two primary endpoints are included in the statistical testing hierarchy, significant superiority of semaglutide 2.4 mg vs. semaglutide placebo must be demonstrated for each of the primary endpoints.

In the analysis approach addressing the primary estimand, week 68 assessments from retrieved subjects (AD) are used. These data are also used to impute missing measurements at week 68 for non-retrieved subjects (MD). The imputation is done separately within each treatment arm (see description below). However, for the power calculations missing values (MT and MD), regardless of treatment arm, are assumed to be similar to semaglutide placebo subjects. These assumptions are likely conservative with respect to the power, and correspond to the jump to reference sensitivity analysis planned below.

Assumptions

The common assumptions for the power calculations are:

- The significance level is 5%
- The randomisation ratio is 4:1:2:1
- For continuous endpoints the t-test on the mean difference assuming equal variances is used
- For binary endpoints the Pearson chi-square test for two independent proportions is used
- Based on data from NN9536-4153
  - 20% of subjects discontinue permanently and
  - 60% of these are retrieved (AD) at week 68
- 100 subjects have T2D
- All subjects in the semaglutide placebo arm are assumed to have same effect as subjects who complete the trial on semaglutide placebo (AT)
- Retrieved subjects (AD) in the semaglutide 2.4 mg (or semaglutide 1.7 mg) arm are assumed to have an effect corresponding to half the treatment difference (compared to semaglutide placebo) of subjects who complete the trial on semaglutide 2.4 mg (or semaglutide 1.7 mg) (AT)
- Non-retrieved subjects (MD) in the semaglutide 2.4 mg (or semaglutide 1.7 mg) arm are assumed to have an effect corresponding to semaglutide placebo

Further assumptions made to calculate the power for each of the primary and confirmatory secondary endpoints are based on findings from other projects conducted by Novo Nordisk (NN8022 (SCALE), NN9535 (SUSTAIN), NN9924 (PIONEER) and trial NN9536-4153 and are presented in Table 3.

Given these assumptions, the sample size of 400 subjects (200 in the semaglutide s.c. 2.4 mg once-weekly, 100 in the semaglutide s.c. 1.7 mg once-weekly and 100 (50+50) in the semaglutide placebo arm), gives an effective power (marginal powers multiplied) of 84%. As sample size is primarily driven by safety, additional scenarios for assumptions are not included due to the overall high power.

Table 3 Assumptions, marginal power and effective power for each endpoint in the hierarchical testing procedure given an anticipated number of 400 randomised subjects

Order	Endpoint	Assumed mean (±SD) or proportion for treatment completers (AT) [non-T2D]|[T2D]		Expected mean (±SD) / proportion		Expected difference or proportion ratio	Marginal power (%)	Effective power (%)
		Semaglutide 2.4 / 1.7 mg	Semaglutide placebo	Semaglutide 2.4 / 1.7 mg	Semaglutide placebo			
Semaglutide 2.4 mg vs semaglutide placebo								
1	% weight change#	[14.0] [11.6] (±10)	[3.0] [1.7] (±10)	11.9 (±11)	2.7 (±11)	9.2%-points	> 99	> 99
2	5% responders	[82%] [75%]	[42%] [37%]	75%	41%	1.8	> 99	> 99
3	10% responders	[66%] [56%]	[24%] [20%]	58%	23%	2.5	> 99	> 99
4	15% responders	[46%] [37%]	[12%] [9%]	39%	11%	3.5	> 99	> 99
5	WC change (cm)#	[11.0] [9.1] (±10)	[4.0] [2.8] (±10)	9.6 (±11)	3.7 (±11)	5.9 cm	99	99
Semaglutide 1.7 mg vs semaglutide placebo								
6	% weight change#	[12.8] [10.4] (±10)	[3.0] [1.7] (±10)	10.9 (±11)	2.7 (±11)	8.2%-points	> 99	99
7	5% responders	[78%] [71%]	[42%] [37%]	72%	41%	1.8	> 99	98
8	10% responders	[61%] [52%]	[24%] [20%]	54%	23%	2.3	> 99	98
9	15% responders	[41%] [32%]	[12%] [9%]	35%	11%	3.2	98	96
10	WC change (cm)#	[9.8] [7.9] (±10)	[4.0] [2.8] (±10)	8.5 (±11)	3.7 (±11)	4.8 cm	87	84

SD = standard deviation; WC = waist circumference; # shown as a positive number.

All tests in the hierarchy are based on the primary estimand.

2.2 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 the 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 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.
  - The on-treatment period as described above (i.e. employing a lag time of 2 weeks [14 days]) applies to all effect assessments, safety laboratory assessments, physical examination, pulse and ECG (including QTc related assessments).
  - For the evaluation of adverse events, hypoglycaemic episodes, adjudicated events, eye examination and antibodies 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.

2.3 Statistical analyses

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

The two randomised semaglutide placebo arms (corresponding to target doses of 2.4 mg and 1.7 mg) will be pooled in all statistical analyses. All endpoints will be compared between semaglutide 2.4 mg vs semaglutide placebo and between semaglutide 1.7 mg vs semaglutide placebo. Results from statistical analyses will generally be accompanied by two-sided 95% confidence intervals (CI) and corresponding p-values. Superiority will be claimed if p-values are less than 5% and the estimated treatment contrasts favours semaglutide 2.4 mg (or semaglutide 1.7 mg).

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 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:

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

Definition of primary endpoint: 5% responders

A body weight reduction of at least 5% from baseline (week 0) to week 68 is defined as:

$$5\%\ \text{responder} = \begin{cases} 1 \text{ if } \%\ \text{weight change} \leq -5\% \\ 0 \text{ if } \%\ \text{weight change} > -5\% \end{cases}$$

Analyses addressing the primary estimand

The following statistical analyses and imputation methods are designed to address the primary estimand, i.e. to assess the effectiveness of semaglutide 2.4 mg and semaglutide 1.7 mg.

The analysis model for % weight change is a linear regression (ANCOVA) of % weight change with randomised treatment and stratification group as factors and baseline body weight (kg) as covariate. The stratification group is defined by T2D status stratification category. The estimated treatment difference between semaglutide 2.4 mg (or semaglutide 1.7 mg) and semaglutide placebo will be reported together with the associated two-sided 95% CI and corresponding p-value.

The analysis model for the 5% responder endpoint is a logistic regression using randomised treatment and stratification group as factors and baseline body weight (kg) as covariate. The stratification group is defined by T2D status stratification category. If the model cannot be fit due to low frequency of responders in 1 or more treatment groups, Firth's maximum likelihood estimation will be used to prevent non-convergence. The estimated odds ratio (OR) between semaglutide 2.4 mg (or semaglutide 1.7 mg) and semaglutide placebo will be reported together with the associated two-sided 95% CI and corresponding p-value.

The superiority tests of semaglutide 2.4 mg (or semaglutide 1.7 mg) vs. semaglutide placebo will be carried out as follows for the two analysis models.

Let $\mu_{\text{semaglutide}}$ and $\mu_{\text{semaglutide placebo}}$ denote the true mean of % weight change for semaglutide 2.4 mg (or semaglutide 1.7 mg) and semaglutide placebo group, respectively. The null and alternative hypotheses tested are:

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

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

Let $OR_{\text{semaglutide/semaglutide placebo}}$ denote the true odds ratio between semaglutide 2.4 mg (or semaglutide 1.7 mg) and semaglutide placebo. The null and alternative hypotheses tested are:

[redacted]

$$H: OR_{semaglutide/semaglutide\ placebo} \leq 1\ vs$$
$$H_A: OR_{semaglutide/semaglutide\ placebo} > 1$$

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

Handling of missing week 68 values for the primary estimand

All available data at week 68 (AT and AD) are used and missing values (MT and MD) at week 68 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 primary estimand for the primary endpoints is given followed by a description of the sensitivity analyses used to assess the robustness of the primary analysis results. The sensitivity analyses investigate how assumptions on body weight development after discontinuation of randomised treatment impact the estimated treatment contrasts between semaglutide 2.4 mg (or semaglutide 1.7 mg) and semaglutide placebo. An illustration of all imputation approaches for the primary estimand is given in Figure 2.

Primary imputation approach for the primary estimand

Multiple imputation approach using retrieved subjects (RD-MI): The primary imputation approach for the primary estimand is a multiple imputation similar to the one described by McEvoy.[2] Missing body weight measurement at week 68 for non-retrieved subjects (MD) are imputed using assessments from retrieved subjects (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 68. Missing body weight measurements at week 68 for subjects on randomised treatment (MT) are imputed in a similar way 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. The model will be a linear regression of body weight (kg) at week 68 with gender (male/female), baseline BMI (kg/m^2) (in categories <35, ≥ 35), timing of the LAO-OT of body weight and stratification group (defined by stratification categories for T2D status) 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 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 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 one 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 gender, then removing stratification factor (T2D status), then removing baseline BMI group, and finally removing baseline body weight (kg). If the imputation model with only LAO-OT of body weight (kg) cannot be fit, the imputation model will be defined by pooling two semaglutide doses (i.e. two groups: semaglutide active and placebo). If the model still cannot be fit, the imputation will be done regardless of the randomised treatment arm. 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 95364382 as seed number. In addition to the seed number, it is specified that the dataset is sorted by subject ID.

Sensitivity analyses

Jump to reference multiple imputation approach (J2R-MI): Missing values of body weight at week 68 (MT and MD) for the semaglutide 2.4 mg (or semaglutide 1.7 mg) and semaglutide placebo group are imputed by sampling among all available assessments at week 68 in the semaglutide 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 semaglutide placebo treatment as adjunct to reduced calorie diet and increased physical activity.[3] The multiple imputation approach is done as above with the first step replaced by:

1. **Imputation**: Defines an imputation model using semaglutide 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 gender (male/female), BMI (kg/m^2) (in categories < 35, ≥ 35) and stratification group (defined by stratification categories for T2D status) 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 gender, then removing stratification factor (T2D status) 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.

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

A single imputation approach as done by Sacks[4] (S1-SI and S2-SI): Missing weight measurements at week 68 for non-retrieved subjects (MD) are imputed using a weight regain rate of 0.3 kg/month after LAO but truncated at no change from baseline whenever the extrapolation would lead to a positive weight gain relative to baseline. If a subject's weight at drug discontinuation represented a gain in weight relative to baseline, no additional gain will be imputed, and the unfavourable gain is carried forward to week 68. The weight regain imputation will be done for all randomised arms (S1-

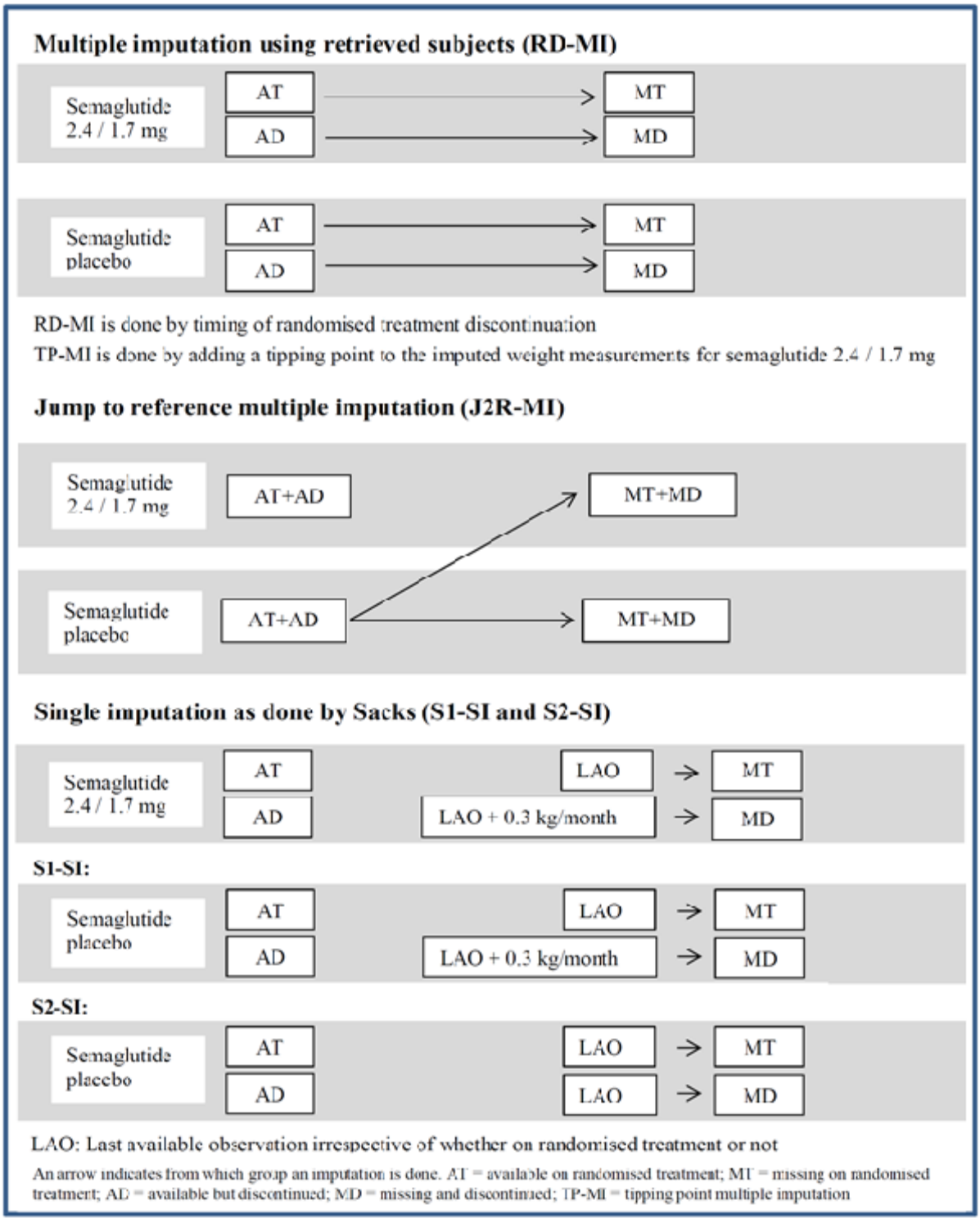
SI). Additionally, a version where only the semaglutide 2.4 mg and 1.7 mg arms use the regain rate while the semaglutide placebo arm uses last available observation (corresponding to a weight regain rate of 0 kg/month) will be performed (S2-SI). For both versions, missing weight measurements at week 68 for subjects on MT are imputed by using LAO. The semaglutide 1.7 mg arm will also be kept for both analyses, but treatment contrast will be presented only for semaglutide 2.4 mg vs placebo.

Tipping-point multiple imputation analysis (TP-MI): 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 semaglutide 2.4 mg and placebo 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. This sensitivity analysis evaluates the robustness of the superiority conclusions to departures from the observed change in body weight in both treatment groups. The semaglutide 1.7 mg arm will also be kept for the analysis, but the tipping-point will be assessed only for semaglutide 2.4 mg vs placebo.

Mixed model for repeated measurements (MMRM): This 'MMRM for effectiveness' will use all assessments regardless of adherence to randomised treatment, including assessments at week 68 for retrieved drop-outs (AD). The MMRM for effectiveness will be fitted using the same factors and covariate as for the primary analyses 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. 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 68 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 of the primary estimand. The semaglutide 1.7 mg arm will also be kept for the analysis, but treatment contrast will be presented only for semaglutide 2.4 mg vs placebo.

Subjects with missing week 68 assessment as non-responders: For the 5% responder analysis an analysis using subjects with missing week 68 assessment as non-responders in the logistic regressions will be done.

Figure 2 Illustration of imputation approaches for the primary estimand



Analysis addressing the secondary estimand

The secondary estimand for % weight change addresses the efficacy of semaglutide 2.4 mg and semaglutide 1.7 mg and will be assessed using a 'MMRM for efficacy'. Week 68 assessments for retrieved drop-outs (AD) are not used in this analysis. The MMRM for efficacy will use assessments only from subjects who are taking the randomised treatment until end of treatment or until first discontinuing of randomised treatment. The derived date of the second consecutive

missed dose will be used as the latest date for using assessments in this MMRM. The assessment closest in time and before the derived date of the second consecutive missed dose will be used as last assessment on randomised treatment. For subjects who initiate other anti-obesity therapies (i.e., weight management drugs or bariatric surgery) 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. 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 and the same factors and covariate as for the primary analyses 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.

The secondary estimand for 5% responders will be assessed using the same MMRM for efficacy except that body weight (kg) will be used as response variable in the model. For subjects with missing body weight at week 68, 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 the same logistic regression model as in the primary analysis of the primary estimand.

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

2.4 Secondary endpoints

2.4.1 Confirmatory secondary endpoints

Confirmatory secondary endpoints are listed in section 1.1.3.2 and are all included in the fixed-sequence statistical strategy, see Table 3. All tests are tests of superiority of semaglutide 2.4 mg (or semaglutide 1.7 mg) to semaglutide placebo.

Analyses addressing the primary estimand

All confirmatory 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. The statistical model for continuous endpoints will be ANCOVA with factors and covariate as for the primary endpoint % weight change with baseline body weight replaced by the baseline assessment of the endpoint to be analysed. The statistical model for body weight responder endpoints will be logistic regression with factors and covariate as for the primary endpoint 5% responders.

Analyses addressing the secondary estimand

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

Sensitivity analyses for confirmatory secondary endpoints

For all continuous confirmatory secondary endpoints a sensitivity analysis using jump to reference as imputation approach will be carried out. For all binary confirmatory secondary endpoints a sensitivity analysis using subjects with missing week 68 assessment 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 4.

Table 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								
Semaglutide 2.4 mg vs semaglutide placebo								
Primary	% weight change	1	Continuous	Primary	FAS	ANCOVA	RD-MI	J2R-MI S1-SI S2-SI TP-MI MMRM
				Secondary	FAS	MMRM	-	-
Primary	5% responders	2	Binary	Primary	FAS	LR	RD-MI	J2R-MI S1-SI S2-SI TP-MI MMRM Non-responder
				Secondary	FAS	LR	MMRM	-
Confirmatory secondary endpoints								
Semaglutide 2.4 mg vs semaglutide placebo								
Primary	10% responders	3	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-
Primary	15% responders	4	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-
Primary	WC change (cm)	5	Continuous	Primary	FAS	ANCOVA	RD-MI	J2R-MI
				Secondary	FAS	MMRM	-	-
Semaglutide 1.7 mg vs semaglutide placebo								
Primary	% weight change	6	Continuous	Primary	FAS	ANCOVA	RD-MI	J2R-MI
				Secondary	FAS	MMRM	-	-
Primary	5% responders	7	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-
Primary	10% responders	8	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-
Primary	15% responders	9	Binary	Primary	FAS	LR	RD-MI	Non-responders
				Secondary	FAS	LR	MMRM	-
Primary	WC change (cm)	10	Continuous	Primary	FAS	ANCOVA	RD-MI	J2R-MI
				Secondary	FAS	MMRM	-	-

FAS = full analysis set; ANCOVA = analysis of covariance; RD-MI = multiple imputation using retrieved subjects; J2R-MI = jump to reference multiple imputation; S1-SI and S2-SI = single imputation as done by Sacks; TP-MI = tipping point multiple imputation; MMRM = mixed model for repeated measurements; LR = logistic regression; WC = waist circumference.

Test order refers to the order of the endpoint in the statistical test hierarchy outlined in Table 3.

2.4.2 Supportive secondary endpoints

Supportive secondary endpoints are listed in Section 1.1.3.2. All tests are tests of superiority of semaglutide 2.4 mg (or semaglutide 1.7 mg) to semaglutide placebo.

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, unless otherwise specified. 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 covariate in the imputation model will be the LAO-OT value only. The statistical model for continuous endpoints will be ANCOVA with factors and covariate as for the primary endpoint % weight change with baseline body weight replaced by the baseline assessment of the endpoint to be analysed.

The statistical model for HbA1c responder endpoints for subjects with T2D at baseline will be logistic regression with randomised treatment as factor and the baseline HbA1c as covariate. Missing HbA1c data at week 68 will be imputed using J2R due to the low number of expected retrieved subjects in this subset. For responder endpoints relating to COAs the statistical model will be logistic regression with randomised treatment and stratification groups (defined by stratification categories for T2D status) as factors and the baseline assessment of the endpoint to be analysed as covariate.

For lipids and biomarkers 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.

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 for efficacy described for the primary endpoints. The effect-related supportive secondary endpoints related to the secondary objective will also be analysed to address the secondary estimand using the same MMRM for efficacy described for the primary endpoints. For change in HbA1c and FPG, whose planned measurement visits differ depending on T2D status at baseline, only the visits where the measurements are planned for all subjects will be included in the MMRM.

HbA1c responders will be analysed on the subjects with T2D at baseline only using a logistic regression model with randomised treatment as factor and the baseline HbA1c as covariate. For subjects with missing HbA1c at week 68, individual values will be predicted from the MMRM and used to classify each subjects as responder or not. The MMRM will include all scheduled visit data from the T2D subjects as response, randomised treatment as a fixed factor, and baseline HbA1c as covariate.

Sensitivity analyses for supportive secondary endpoints

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

Additional considerations for statistical analyses

The supportive secondary endpoint change from baseline to week 68 in VFA (cm^2 and %) as measured by CT scan will be defined as absolute change from baseline (cm^2) and %change from baseline (%). VFA for three individual CT slices including lumbar vertebral level 4 (L4) are to be

measured (i.e. one with L4 in the center, one above L4 and one below), and the assessments at L4 will be used to derive the endpoints. These endpoints will be analysed using an ANCOVA with factors and covariate as for the primary endpoints with baseline body weight replaced by baseline value of VFA. Missing VFA (cm^2) data at week 68 will be imputed using J2R due to the low number of expected retrieved subjects in this subset. Assuming that 144 (72 on semaglutide s.c. 2.4 mg, 36 on semaglutide s.c. 1.7 mg and 36 on semaglutide placebo) of the planned 180 Japanese subjects for CT scan will have a CT scan at week 68, then an expected treatment difference in the range of 6% to 10%-points (SD=10) between semaglutide 2.4 mg (or semaglutide 1.7 mg) and semaglutide placebo will provide a power between 83% to > 99% or 70% to 99% respectively. In addition, the correlation of VFA (cm^2) at baseline and waist circumference (cm) at baseline will be evaluated across treatment arms. The coefficient of correlation (R) will be provided.

Analysis of safety endpoints

The safety endpoint pulse will be analysed using an MMRM for efficacy as described in Section 2.3.1. For amylase, lipase and calcitonin descriptive statistics will be provided. The descriptive statistics for calcitonin will be stratified by gender.

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.

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

For QTcF interval, in addition to the descriptive statistics, the categorical summaries will be produced by visit using the following categories:

- Absolute QTcF interval:
  - > 450 msec
  - > 480 msec
  - > 500 msec
- Change from baseline in QTc interval:
  - > 30 msec
  - > 60 msec

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

Objective	Endpoint	Endpoint type	Estimand	Analysis set	Statistical model	Imputation approach	Sensitivity analyses
Supportive secondary endpoints (effect related)							
Primary	Weight change (kg)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Primary	BMI change (kg/m^2)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Primary	WC change (cm)*1	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Primary	VFA (kg)	Continuous	Primary	FAS (CT)*2	ANCOVA	J2R-MI	-

[redacted]

			Secondary	FAS (CT)*2	MMRM	-	-
Primary	VFA (%)	Continuous	Primary	FAS (CT)*2	ANCOVA	J2R-MI	-
			Secondary	FAS (CT)*2	MMRM	-	-
Secondary	HbA_{1c} 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 (µIU/mL, pmol/L)	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	Free fatty acids 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	PAI-1 change (AU/mL)	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 PF score responders #	Binary	Primary	FAS	LR	RD-MI	-
			Secondary	FAS	LR	MMRM	-
Secondary	SF-36 PF score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 RP score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 BP score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 GH score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 VT score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 SF score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 RE score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 MH score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 PCS score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	SF-36 MCS score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	IWQoL-Lite for CT PFD score responders##	Binary	Primary	FAS	LR	RD-MI	-
			Secondary	FAS	LR	MMRM	-
Secondary	IWQoL-Lite for CT PFD score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	IWQoL-Lite for CT PDD score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	IWQoL-Lite for CT PSD score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	IWQoL-Lite for CT total score change	Continuous	Primary	FAS	ANCOVA	RD-MI	-
			Secondary	FAS	MMRM	-	-
Secondary	HbA1c < 7.0% responders	Binary	Primary	FAS (T2D)*3	LR	J2R-MI	-
			Secondary	FAS (T2D)*3	LR	MMRM	-
Secondary	HbA1c ≤ 6.5% responders	Binary	Primary	FAS (T2D)*3	LR	J2R-MI	-
			Secondary	FAS (T2D)*3	LR	MMRM	-

Supportive secondary endpoints (safety related)							
Secondary	Number of TEAEs	Continuous	-	SAS	-	-	-
Secondary	Number of SAEs	Continuous	-	SAS	-	-	-
Secondary	Number of TE hypoglycaemia episodes	Continuous	-	SAS (T2D)*3	-	-	-
Secondary	Pulse change (bpm)	Continuous	-	SAS	MMRM	-	-
Secondary	Amylase change (U/L)	Continuous	-	SAS	Descriptive statistics	-	-
Secondary	Lipase change (U/L)	Continuous	-	SAS	Descriptive statistics	-	-
Secondary	Calcitonin change (ng/L)	Continuous	-	SAS	Descriptive statistics	-	-

FAS = full analysis set; ANCOVA = analysis of covariance; RD-MI = multiple imputation using retrieved subjects; MMRM = mixed model for repeated measurements; LR = logistic regression; WC = waist circumference; VFA = visceral fat area; sBP = systolic blood pressure; dBP = diastolic blood pressure; PF= Physical Functioning; RP = Role-Physical; BP = Bodily Pain; GH = General Health; VT = Vitality; SF = Social Functioning; RE = Role-Emotional; MH = Mental Health; PCS = Physical component summary; MCS = Mental component summary; PFD = physical function domain; PDD = pain/discomfort domain; PSD = psychosocial domain; *1 according to JASSO guideline; *2 based on subset of subjects undergoing CT-scan; *3 based on subjects with T2D at screening; # responder value = 3.7; ## responder value = 14.6

2.4.3 Exploratory endpoints

Exploratory endpoints are listed in Section 1.1.3.3. Observed data for exploratory endpoints will be summarised by descriptive statistics.

2.4.4 Explorative statistical analysis for pharmacogenetics and biomarkers

The statistical analysis of biomarker endpoints is described under Section 2.4.2.

2.4.5 Other analyses

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

2.4.6 Pharmacokinetic and/or pharmacodynamic modelling

Pop-PK and exposure-response analyses will be used as supportive evidence for the evaluation of efficacy and safety and further to support the recommended dose of semaglutide in subjects with obesity. First, plasma semaglutide concentrations will be analysed using a population pharmacokinetic model, quantifying covariate (such as baseline body weight, age, gender, race, ethnicity and injection site) effects on semaglutide exposure. Second, model based estimates of steady-state average concentrations will be derived for each subject, in order to facilitate subsequent exposure-response analyses. Relevant efficacy and safety endpoints will be related to steady-state average concentrations and subjected to model based analysis.

A modelling analysis plan will be prepared before first database lock.

3 Supporting documentation

3.1 Appendix 1: List of abbreviations

AD	available but discontinued
AE	adverse event
ANCOVA	analysis of covariance
AT	available on randomised treatment
BMI	body mass index
BP	Bodily Pain
CI	confidence interval
COA	Clinical Outcome Assessments
dBP	diastolic blood pressure
CT	computed tomography
ECG	electrocardiogram
FAS	full analysis set
FDA	U.S. Food and Drug Administration
GH	General Health
HbA1c	glycated haemoglobin
HDL	high density lipoprotein
hsCRP	high sensitive C-reactive protein
ICH	International Council for Harmonisation
IWQoL-Lite for CT	Impact of Weight on Quality of Life-Lite for Clinical Trials
JASSO	Japan Society for the Study of Obesity
LAO-OT	last available observation on-treatment
LDL	low-density lipoprotein
LR	logistic regression
MCS	Mental component summary

MD	missing and discontinued
MH	Mental Health
MMRM	mixed model for repeated measurements
MT	missing on randomised treatment
OAD	oral antidiabetic drug
OR	odds ratio
PAI-1	plasminogen activator inhibitor-1
PCS	Physical component summary
PDD	pain/discomfort domain
PF	Physical Functioning
PFD	physical function domain
PK	pharmacokinetic
Pop	population
PSD	psychosocial domain
PYE	patient years of exposure
PYO	patient years of observation
QTcF interval	corrected QT interval using Fridericia's correction
RE	Role-Emotional
RP	Role-Physical
SAE	serious adverse event
SAP	statistical analysis plan
SAS	safety analysis set
sBP	systolic blood pressure
s.c.	subcutaneus
SD	standard deviation
SF	Social Functioning

SF-36	Short Form-36 v2.0 acute
T2D	type 2 diabetes
TE	treatment emergent
TEAE	treatment emergent adverse event
TFL	tables, figures and listings
VFA	Visceral Fat Area
VT	Vitality
WC	waist circumference

3.2 Appendix 2: Changes to protocol-planned analyses

The main analyses were described in the protocol for the trial NN9536-4382. However, clarifications and more detailed descriptions of endpoints and analyses are provided in this SAP. The changes from the protocol of NN9536-4382 are summarised in Table 6.

Table 6 Changes made in the statistical analysis plan to the planned statistical analysis in the protocol

Change to planned statistical analysis	Rationale for change and possible implications
In general, the terminology 'rescue interventions' has been changed to 'other anti-obesity therapies'.	To align with other trials in STEP phase 3 program. This change will not impact the actual definition.
The terminologies 'effectiveness' and 'efficacy' in the estimands have been deleted.	To adopt the terminologies in the addendum to ICH E9 (R1). This change will not impact the actual definition.
It has been added that the secondary estimand will cover all effect-related objectives. Accordingly, the analyses addressing the secondary estimand have been specified for all effect-related supportive secondary endpoints.	Results from analysis of the secondary estimand are relevant for all objectives. These results will provide useful context to the results from analysis of the primary estimand for all endpoints, though they will not impact the conclusions based on these endpoints.
The text has been updated so that the primary endpoints address the secondary objective.	To clarify that the primary endpoints also address the secondary weight-related objective for semaglutide 1.7 mg vs placebo.
Unit for PAI-1 has been corrected to AU/mL.	Correction.
The supportive secondary endpoint "Body weight reduction ≥ 20% from baseline at week 0" has been added.	To be able to assess the proportion of subjects with the greatest weight loss.
The supportive secondary endpoint "change from baseline at week 0 to week 68 in QTcF interval" has been added and the associated analysis method has been specified.	To pre-specify analysis of QTc interval prolongation considering regulatory importance.
The text has been updated so that the exploratory endpoint "glycaemic category" will be applied to only subjects with no T2D at baseline	Corrected to reflect actual data collection.

Change to planned statistical analysis	Rationale for change and possible implications
It has been clarified that subjects in the FAS/SAS will be evaluated "as randomised"/"as treated".	To clarify how patients are analysed according to treatment arm.
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.
It has been specified that, for the following assessments, a time-point is considered 'on-treatment' if any dose has been administered within the prior 2 weeks (14 days): • all effect assessments and for safety laboratory assessments, physical examination, pulse and ECG (including QTc related assessments)	To clarify how the on-treatment period is defined for each assessment.
It has been specified that, for the following assessments, a time-point is considered 'on-treatment' if any dose has been administered within the prior 7 weeks (49 days): • adverse events, hypoglycaemic episodes, adjudicated events, eye examination and antibodies	To clarify how the on-treatment period is defined for each assessment.
It has been specified that if the model cannot be fit in a logistic regression model due to low frequency of responders in 1 or more treatment groups, Firth's maximum likelihood estimation will be used to prevent non-convergence.	To pre-specify the alternative approach if the model cannot fit.
It has been clarified that RD-MI imputation is performed according to the timing of last available observation during the on-treatment period (LAO-OT) for all endpoints. Furthermore, it has been clarified that the LAO-OT must be prior to the landmark visit (week 68).	To clarify that the grouping of subjects according to timing is as in the publication by McEvoy.[2]
The baseline BMI-grouping "27-<35" in RD-MI and J2R-MI has been changed to "-<35".	To accommodate the fact that subjects may loose weight between the screening and the randomisation visit
The two highest baseline BMI-groups (35 - <40 and ≥40) in RD-MI and J2R-MI have been collapsed into one (≥35).	Because it has been found out that there were only a few subjects with baseline BMI ≥40 during blind data review.
In grouping of retrieved subjects by timing of LAO-OT in the RD-MI procedure, it has been clarified that 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.	To clarify the criteria on which a decision to reduce the imputation model will be based.
In the RD-MI imputation model, it has been specified that model reduction will be done in the following order: • First removing gender, then removing stratification factor (T2D status), then removing baseline BMI group, and finally removing baseline body weight (kg). • If the model cannot fit after above reduction, first pooling two semaglutide doses (i.e. two groups: semaglutide active and placebo). If the model still cannot be fit, the imputation will be done regardless of the randomised treatment arm.	To pre-specify the model reduction approach.
It has been clarified that if no post-baseline LAO-OT exist, then LAO-OT will be the baseline value and the timing of LAO-OT were to be the first interval.	To clarify the definition of the last available observation during the on-treatment period

Change to planned statistical analysis	Rationale for change and possible implications
In all multiple imputation procedures, in addition to the seed number, it has been specified that the dataset is sorted by subject ID.	To ensure reproducibility of the imputed datasets.
In the J2R-MI imputation model, it has been specified that model reduction will be done in the following order: First removing gender, then removing stratification factor (T2D status), and finally removing baseline BMI group.	To pre-specify the model reduction approach.
It has been clarified that the sensitivity analyses S1-SI, S2-SI, TP-MI and MMRM for effectiveness will address sensitivity only for primary comparison of semaglutide 2.4 mg vs placebo.	To specify clearly the treatment contrast of interest in the sensitivity analyses.
The TP-MI procedure has been updated to be a 2-way tipping point analysis in which penalties are applied to both treatment groups (semaglutide 2.4 mg and placebo).	Based on feedback received from the FDA (response letter 17 May 2018): To confirm the robustness of superiority conclusions using a tipping point analysis, we believe that a 2-way tipping point analysis represents the real-world situation for missing data from both treatment arms (semaglutide and placebo). We would like to see departures from the treatment difference by varying both treatment arms rather than only adding a penalty to the active treatment arm (semaglutide). Additionally, please include interpretations for the varying scenarios and how likely they would be seen in a real-world setting.
A description has been included of the sensitivity analysis of the 5% responder endpoint (primary estimand) using MMRM.	To clarify precisely how the MMRM will be parameterized, and how missing values will be imputed.
It has been clarified that the non-responder analysis includes subjects with missing body weight assessment at week 68 as non-responders.	To clarify that all subjects with missing body weight assessment at week 68 will be considered as non-responders in this sensitivity analysis.
It has been clarified that the 5% responder analysis using MMRM for the secondary estimand is predicting individual values for % weight change only when % weight change is missing at week 68. Furthermore, it is clarified that the logistic regression includes the same factors and covariate as for the analysis of the primary estimand.	To clarify precisely which values are imputed, and how the analysis model is parameterized.
For fasting serum insulin collected only at baseline and week 68, it has been clarified that the covariate in the RD-MI will be the LAO-OT value only.	To clarify which values are used in the imputation.
For analyses of HbA1c responder endpoints, the imputation approach has been changed to J2R.	Because the low number of retrieved subjects is expected in this subpopulation (subjects with T2D at baseline).
It has been clarified that for change in HbA1c and FPG, whose planned measurement visits differ depending on T2D status at baseline, only the visits where the measurements are planned for all subjects will be included in the MMRM	To avoid including the visits which are not planned for subset of subjects.
The derivation of change in VFA (cm^2 and %) has been added.	To clarify how the endpoints will be derived.
For missing VFA data at week 68, the J2R approach has been specified instead of analysis on only subjects who are on-treatment at week 68 and having available VFA measurement at week 68.	To include all subjects who underwent CT-scan. The J2R will be used due to the low number of expected retrieved subjects in this subset.

Change to planned statistical analysis	Rationale for change and possible implications
For correlation of VFA at baseline and waist circumference at baseline, it has been updated so that only VFA in cm^2 will be used instead of both in cm^2 and %.	Because VFA in % is not defined at baseline.
It has been added that the responder definition value is 3.7 for SF-36 physical functioning score and 14.6 for IWQoL-Lite for CT physical function domain (5 items) score	To specify the threshold for the responders for the SF-36 physical functioning score and IWQoL-Lite for CT physical function domain, which was not available at the time of protocol preparation.
It has been clarified that for VFA endpoints, the assessments at L4 are used for derivation.	This clarification has been made after unblinding, but prior to a post-database lock data correction of VFA assessments. It is to clarify before the data are updated which slice is to be used to define the VFA endpoints, since the updated data are expected to include measurements from 3 single slices.

4 Reference

[1] Japan Society for the Study of Obesity. Guidelines for the management of obesity disease. 2016.
[2] McEvoy BW. Missing data in clinical trials for weight management. J Biopharm Stat. 2016;26(1):30-6.
[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.
[4] Sacks FM, Bray GA, Carey VJ, Smith SR, Ryan DH, Anton SD, et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. N Engl J Med. 2009;360(9):859-73.

Clinical Trial Report

Trial ID: NN9536-4382 (STEP 6)

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

18 Retinal disorders

Custom query (NNMQ Retinal disorders and visual impairment):

- SMQ Retinal disorders, narrow terms only
- HLT Visual impairment and blindness (excl colour blindness), primary terms only