Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT02849080
Sponsor trial ID:	NN9924-4257
Official title of study:	PIONEER 7 – Flexible dose adjustment. Efficacy and Safety of Oral Semaglutide Using a Flexible Dose Adjustment Based on Clinical Evaluation versus Sitagliptin in Subjects with Type 2 Diabetes Mellitus
Document date:	25 October 2018

16.1.9 Documentation of statistical methods

List of contents

Statistical analysis plan ..... Link

Pre-defined MedDRA search – list of preferred terms ..... Link

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

.

Statistical Analysis Plan

Trial ID: NN9924-4257

Efficacy and Safety of Oral Semaglutide Using a Flexible Dose Adjustment Based on Clinical Evaluation versus Sitagliptin in Subjects with Type 2 Diabetes Mellitus

A 52-week Randomised, Open-label, Active-controlled Trial with a 52-week Extension Phase

Author:

[redacted]

Table of contents

Page

Table of contents **2**

List of abbreviations **3**

1 Introduction **5**
1.1 Trial information 5
1.1.1 Main phase 5
1.1.2 Extension phase 5
1.1.2.1 Extension phase (sustainability) 5
1.1.2.2 Extension phase (switch) 5
1.2 Trial duration 6
1.3 Objectives 6
1.3.1 Primary objective 6
1.3.1.1 Main phase 6
1.3.2 Secondary objectives 6
1.3.2.1 Main phase 6
1.3.2.2 Extension phase (sustainability) 7
1.3.2.3 Extension phase (switch) 7
1.4 Scope of the statistical analysis plan 7

2 Statistical considerations **8**
2.1 Sample size calculation 11
2.2 Definition of analysis sets 13
2.3 Primary endpoint 19
2.3.1 Primary analysis for the primary estimand 19
2.3.2 Primary analysis for the secondary estimand 20
2.3.3 Sensitivity analyses 20
2.3.3.1 Pattern mixture models 21
2.4 Secondary endpoints 22
2.4.1 Confirmatory secondary endpoints 22
2.4.2 Supportive secondary endpoints 23
2.4.2.1 Efficacy endpoints 23
2.4.2.2 Safety endpoints and safety assessments 28
2.5 Interim analysis 33
2.6 Patient reported outcomes 33
2.6.1 SF-36v2® (acute version) health survey 34
2.6.2 Diabetes Treatment Satisfaction Questionnaire – status version (DTSQs) 36
2.6.3 Responder analyses 37

3 Changes to the statistical analyses planned in the protocol **38**

4 References **39**

List of abbreviations

AACE	American Association of Clinical Endocrinologists
ADA	American Diabetes Association
AE	adverse event
ANCOVA	analysis of covariance
ATC	anatomical therapeutic chemical
BG	blood glucose
BMI	body mass index
BP	bodily pain
CHMP	European Medicines Agency
CI	confidence interval
CRF	case report form
DTSQs	Diabetes Treatment Satisfaction Questionnaire status version
EAC	event adjudication committee
ECG	electrocardiogram
eCRF	electronic case report form
FAS	full analysis set
FPG	fasting plasma glucose
GH	general health
HbA1c	glycosylated haemoglobin
HRQoL	health-related quality of life
IWRS	interactive web response system
LLoQ	lower limit of quantification
LOCF	last observation carried forward
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MCS	mental component score
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model repeated measures
MNAR	missing not at random
MH	mental health
NA	not applicable
NAS	US National Research Council
NBS	norm-based score
OAD	oral anti-diabetic drug
OR	odds ratio
PCS	physical component score
PF	physical functioning

PG	plasma glucose
PRO	patient reported outcome
RE	role-emotional
RP	role-physical
SAP	statistical analysis plan
SAS	safety analysis set
s.c.	subcutaneous
SD	standard deviation
SF	social functioning
SF-36v2	Short Form-36 version 2
SGLT-2	sodium glucose co-transporter 2
SU	sulfonylureas
T2DM	Type 2 diabetes mellitus
TE	treatment effect
TEAE	treatment-emergent adverse event
TZD	thiazolidinediones
VT	vitality

1 Introduction

1.1 Trial information

The trial consists of two 52-week treatment periods. The first 52-week treatment period is referred to as the main phase and the second 52-week treatment period as the extension phase. The extension phase consists of two parts; these are referred to as the extension phase (sustainability) and the extension phase (switch).

1.1.1 Main phase

The main phase is a 52 week randomised, open-label, active-controlled, parallel-group, multi-centre, multi-national treatment phase with two arms comparing the efficacy and safety of oral semaglutide using a flexible dose adjustment versus sitagliptin 100 mg once-daily, in subjects with T2DM treated with 1-2 OADs (metformin, sulfonylureas [SU], thiazolidinediones [TZD], sodium glucose co-transporter 2 [SGLT-2] inhibitors).

Subjects will be randomised 1:1 to receive one of the following treatments as add-on to their anti-diabetic background medication:

- flexible dosing (3, 7 or 14 mg) of oral semaglutide once-daily
- 100 mg sitagliptin once-daily

1.1.2 Extension phase

The extension phase is a 52-week open-label, 3-arm, parallel-group treatment period following the main phase. The extension phase (switch) is active-controlled.

1.1.2.1 Extension phase (sustainability)

Subjects randomised to oral semaglutide in the main phase and still on trial product will continue treatment with oral semaglutide using a flexible dose adjustment in the 52-week extension phase.

1.1.2.2 Extension phase (switch)

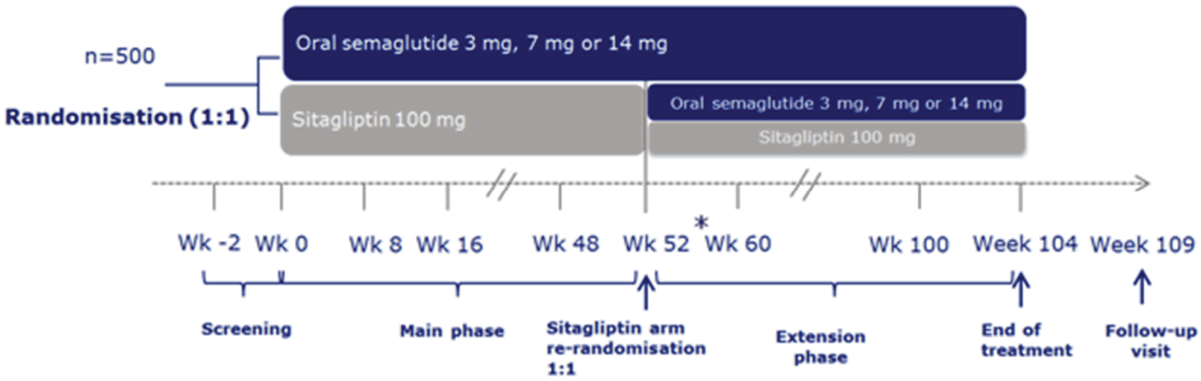
Subjects randomised to sitagliptin in the main phase and still on trial product will be re randomised 1:1 at week 52 to receive one of the following treatments in the 52-week extension phase as add-on to their anti-diabetic background medication:

- flexible dosing (3, 7 or 14 mg) of oral semaglutide once daily
- 100 mg sitagliptin once daily

1.2 Trial duration

The total trial duration of the main phase for the individual subject will be either approximately 54 weeks or approximately 59 weeks depending on whether or not the subject continues into the extension phase. The trial includes a 2 week screening period, a 52 week treatment period and, for subjects not continuing in the extension phase, a 5 week follow-up period. For subjects continuing in the extension phase, additional 52 weeks will be included and the total trial duration will be approximately 111 weeks (2 weeks screening, 104 weeks treatment and 5 weeks follow-up). The trial is designed to have visits every 8 weeks and a phone contact at week 4 and week 56.

A schematic illustration of the trial design is shown in Figure 1–1



* Subjects who have completed treatment in the main phase and are not continuing in the extension phase must attend the Follow-up visit (Week 57) 5 week after the last date on trial product.

Figure 1–1 Trial design

1.3 Objectives

1.3.1 Primary objective

1.3.1.1 Main phase

To compare the effect of once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation versus sitagliptin once daily, both in combination with 1-2 OADs on glycaemic control in subjects with T2DM.

1.3.2 Secondary objectives

1.3.2.1 Main phase

To compare the effect of once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation versus sitagliptin once daily, both in combination with 1-2 OADs on body weight in subjects with T2DM.

To compare the safety and tolerability of once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation versus sitagliptin once daily, both in combination with 1-2 OADs in subjects with T2DM.

1.3.2.2 Extension phase (sustainability)

To evaluate the sustainability on glycaemic control and body weight reduction of once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation in subjects with T2DM.

To evaluate the long term safety of once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation in subjects with T2DM.

1.3.2.3 Extension phase (switch)

To compare the effect on glycaemic control of switching to once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation versus staying on once-daily sitagliptin in subjects with T2DM.

To compare the effect on body weight of switching to once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation versus staying on once-daily sitagliptin in subjects with T2DM.

To compare the safety and tolerability of switching to once-daily dosing of oral semaglutide using a flexible dose adjustment based on clinical evaluation versus staying on once-daily sitagliptin in subjects with T2DM.

1.4 Scope of the statistical analysis plan

This SAP is based on the protocol "Efficacy and Safety of Oral Semaglutide Using a Flexible Dose Adjustment Based on Clinical Evaluation versus Sitagliptin in Subjects with Type 2 Diabetes Mellitus - A 52-week Randomised, Open-label, Active-controlled Trial with a 52-week Extension Phase", version 3.0.

2 Statistical considerations

General considerations

The results from the trial will be reported in two clinical trial reports; the first clinical trial report will report the results from the main phase and the second clinical trial report will report the results of the entire trial period for subjects who continued in the extension phase. The analyses from the main phase will not be repeated in the second clinical trial report.

A database lock for the main phase will be performed when all subjects have completed the main phase to be able to report the results in the first clinical trial report. After completion of the extension phase, there will be a second database lock followed by the reporting of the results of the entire trial in the second clinical trial report for subjects who continued in the extension phase.

Data from all sites will be analysed and reported together. The data from the different dose levels for each oral semaglutide arm will be pooled as one oral semaglutide arm and not analysed by dose level. However, the two oral semaglutide arms in the extension phase will not be pooled.

In statistical analyses where stratification is included, the stratification factors will be included based on the actual information collected through the eCRF via the concomitant medication form and the central laboratory results for HbA_{1c} (extension phase only). The stratification factor in the main phase will be anti-diabetic background medication at screening (with and without SU). In case of missing eCRF information for the anti-diabetic background medication the information collected from the IWRS system will be used. The stratification factors in the extension phase will be HbA_{1c} below 7% (yes/no) and if the subject is currently on prescribed rescue medication and continuing on the same medication in the extension phase (yes/no). Rescue medication is defined as intensification of existing background anti-diabetic medication compared to V1 and/or initiation of new anti-diabetic medication.

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

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

The results will be presented as follows.

Main phase

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

- Oral semaglutide (3, 7 or 14 mg)
- Sitagliptin (100 mg)

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

Extension phase (sustainability)

Relevant descriptive statistics will be presented, including graphical presentations, to evaluate the data from the below treatment arm for subjects randomised to oral semaglutide and continuing on oral semaglutide in the extension phase.

- Oral semaglutide (3, 7 or 14 mg) during 104 weeks

Extension phase (switch)

The results from the extension phase (switch) will be presented as for the main phase; however, the comparison will be between the below two treatment arms in which subjects were randomised to sitagliptin at trial entry and re-randomised at week 52 to either continue on sitagliptin or switch to oral semaglutide.

- Oral semaglutide (3, 7 or 14 mg), switching from sitagliptin at week 52
- Sitagliptin (100 mg), continuing on sitagliptin after week 52

Primary and secondary estimands

Two estimands addressing different aspects of the trial objective will be defined; a primary de-facto (effectiveness) estimand and a secondary de-jure (efficacy) estimand:

- Primary estimand
  - de-facto treatment effect for all randomised subjects regardless of adherence to randomised treatment and initiation of rescue medication

The primary de-facto estimand assesses the expected glycaemic benefit in a future population that results from subjects initiating treatment with oral semaglutide including potential rescue medication(s) as compared to initiating treatment with sitagliptin including potential rescue medication(s). Generalisation of this estimand depends among other things on the extent to which

the use of rescue medication in this trial reflects clinical practice and the adherence to trial product administration in this trial reflects the behaviour of the target population. Accordingly, data collected regardless of discontinuation of trial product or initiation of rescue medication(s) will be used to draw inference.

- Secondary estimand
  - de-jure treatment effect for all randomised subjects if all subjects adhered to treatment and did not initiate rescue medication

The secondary de-jure estimand assesses the glycaemic benefit a future subject is expected to achieve if initiating and continuing treatment with oral semaglutide as compared to sitagliptin. It is considered a clinically relevant estimand as it provides information to treating clinicians about the expected glycaemic efficacy of oral semaglutide compared to sitagliptin for purposes of treating individual subjects. Generalisation of this estimand depends among other things on the extent to which the adherence to trial product administration in this trial reflects the behaviour of the target population. Accordingly, only data collected prior to discontinuation of trial product or initiation of rescue medication will be used to draw inference. For the two treatment arms in the extension phase (switch), the current prescribed rescue medication (at V10) in the main phase will be treated as background medication in the extension phase and only data collected after initiation of rescue medication in the extension phase will be excluded. This will avoid confounding from rescue medication.

Main phase
For both estimands, the treatment effect refers to an odds ratio (OR) for the primary endpoint of achieving $HbA_{1c} < 7.0\%$ at week 52 (yes/no). The treatment effect refers to a treatment difference for the confirmatory secondary endpoint; change from baseline to week 52 in body weight.

Extension phase (switch)
For both estimands, the treatment effect refers to a treatment difference for both confirmatory secondary endpoints; change from week 52 to week 104 in HbA_{1c} and in body weight.

Missing data considerations at week 52 and at week 104

When estimating the primary estimand, the proportion of missing data, i.e., data that do not exist even though subjects are intended to stay in the trial regardless of treatment status and initiation of rescue medication(s), is expected to be maximum 15%. Missing data will mainly be due to withdrawal from trial or lost to follow-up.

When estimating the secondary estimand, the proportion of missing data is expected to be higher (20-30%) since data collected after discontinuation of trial product or initiation of rescue medication(s) will be set to missing. The 20-30% of missing data is based on the sitagliptin phase 3

trials[3], and the oral semaglutide phase 2 trial (NN9924-3790), that indicates that a low starting dose with gradual dose-escalation diminishes gastrointestinal AEs compared with more aggressive dosing regimens. The possibility to reduce the dose due to issues with tolerability is also expected to reduce the number of subjects withdrawing from the oral semaglutide arm. Across treatment arms the main reasons for missing data are expected to be early treatment discontinuation due to gastrointestinal AEs and initiation of rescue medication. Initiation of rescue medication is expected to be more frequent in the sitagliptin than in the oral semaglutide arm. A higher proportion of subjects are expected to discontinue treatment due to AEs in the oral semaglutide arm, compared to the sitagliptin treatment arm. So overall the frequency of missing data is expected to be similar across treatment arms.

The proportion of missing data from week 52 to week 104 is expected to be similar as for the first 52 weeks in the two treatment arms of re-randomised subjects. The rate of missing data is expected to decline over time in the oral semaglutide arm in which subjects were randomised to oral semaglutide at baseline due to decreased tolerability issues over time.

It is expected that 190 subjects will enter the extension phase (switch). This is based on the assumption that maximum 25% of the subjects do not continue in the extension phase (switch) due to withdrawal from trial or they are lost to follow up, off treatment or unwilling to continue.

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

2.1 Sample size calculation

The following four hypotheses are planned to be tested:

Main phase:
Superiority of $HbA_{1c} < 7\%$ (yes/no) at week 52
Superiority of change from baseline to week 52 in body weight

Extension phase (switch):
Superiority of change from week 52 to week 104 in HbA_{1c}
Superiority of change from week 52 to week 104 in body weight

The sample size calculation is made to ensure at least 90% power to confirm superiority of $HbA_{1c} < 7\%$ (yes/no) at week 52 of oral semaglutide versus sitagliptin. The four pre-specified confirmatory hypotheses are shown in Figure 2–1. The hierarchal testing strategy is used to control the overall type-1 error at a nominal two-sided 5% level for the four confirmatory tests.

The statistical testing strategy is built on the principle that glycaemic effect will have to be established in terms of HbA_{1c} superiority before testing for added benefits in terms of body weight

superiority. The hypotheses related to the main phase will be tested and reported in the first clinical trial report. If these confirmatory hypotheses are confirmed, the two confirmatory hypotheses related to the extension phase (switch) will be tested in the hierarchical order shown in Figure 2–1 and reported in the final clinical trial report.

The assumptions used in the sample size calculation are based on the oral semaglutide phase 2 results (trial NN9924-3790)[1], sitagliptin phase 3a trial results[3], and supported by results from the s.c. semaglutide phase 3 trial, SUSTAIN 2[4].

The sample size calculation is based on a 5% (two-sided) significance level and Fishers exact test. The sample size depends on the proportion of responders and the absolute difference in proportions between semaglutide and sitagliptin. Assuming an absolute difference in proportions of 15 percentage-points (taking the retrieved and imputed data into account) and that the proportion of sitagliptin responders are distributed around 20 to 50%, the power for confirming superiority on the primary endpoint, if a subject after week 52 achieves (yes/no) $HbA_{1c} < 7\%$ (53 mmol/mol) will be at least 90% with 250 subjects per arm. In total 500 subjects are planned to be randomised.

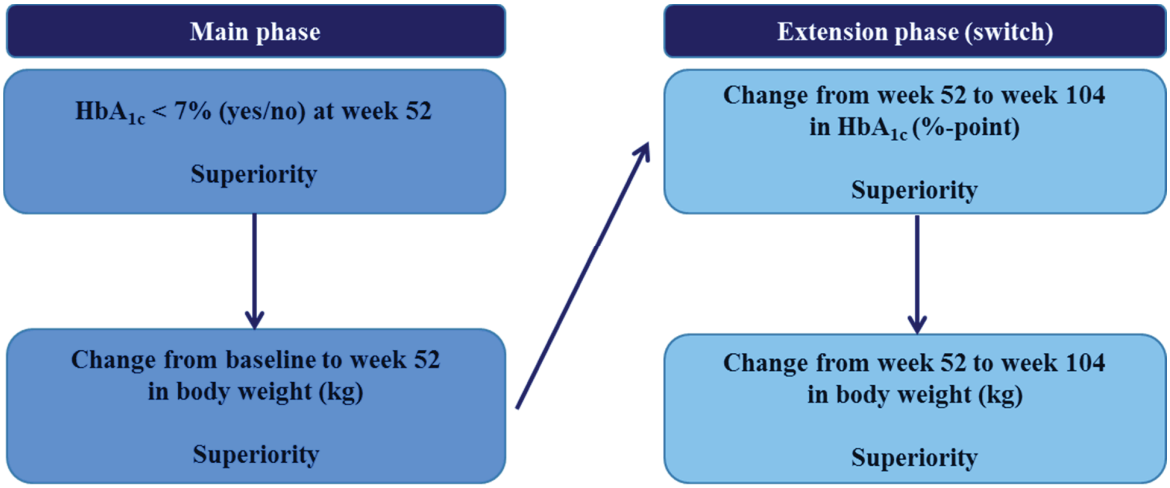


Figure 2–1 Graphical illustration of the closed testing procedure

The overall significance level of $\alpha = 0.05$ (two-sided) is initially allocated to the HbA_{1c} superiority test at week 52. The local significance level will be reallocated to the second hypothesis if the first hypothesis is confirmed and so on. The sample size is based on the first HbA_{1c} hypothesis (main phase).

The sample size assumptions for treatment effects (TE), adjusted TE and the standard deviations (SD) used to calculate the power for the additional three confirmatory hypotheses are presented in Table 2–1.

Because the equalising effect of rescue medication will be included in the analyses and because a conservative approach for handling of missing data will be applied, an adjustment of the TE will be implemented for the 15% of subjects who are expected to either discontinue trial product prematurely or initiate rescue medication as well as for the 15% of subjects who are expected to have missing data. The TEs used in the sample size calculation will be adjusted according to a 75% smaller effect in these subjects. The adjusted TE is calculated as follows:

$$0.7 \times TE + 0.3 \times TE \times 0.25$$

The calculated marginal and conditional powers for each of the four tests are also presented in Table 2–1. All the confirmatory hypotheses are assumed to be independent. Because positive correlation amongst the test is expected, the assumption of independence is viewed as conservative.

Table 2–1 Assumptions used in sample size and power calculation

Endpoint	Treatment effect (TE)	Adjusted TE	SD	Number of subjects	Marginal power	Conditional power
Main phase:						
HbA_{1c} < 7% (yes/no)	15%-point difference			500	90%	90%
Change in body weight (kg)	2.5	1.94	4.0	500	>99%	90%
Extension phase (switch):						
Change in HbA_{1c} (%-point)	0.4	0.31	1.1	190	49%	44%
Change in body weight (kg)	2.5	1.94	4.0	190	91%	40%

2.2 Definition of analysis sets

The following analysis sets will be defined:

Main phase

Full analysis set (FAS): Includes all randomised subjects. Subjects in the FAS will contribute to the evaluation "as randomised".

Safety analysis set (SAS): Includes all subjects exposed to at least one dose of trial product. Subjects in the SAS will contribute to the evaluation "as treated".

Extension phase (sustainability)

Extension (sustainability) FAS: Includes all subjects randomised to oral semaglutide in the main phase and continuing on oral semaglutide in the extension phase. Subjects in the extension (sustainability) FAS will contribute to the evaluation "as randomised".

Extension (sustainability) SAS: Includes all subjects randomised to oral semaglutide in the main phase and exposed to at least one dose of oral semaglutide in the extension phase. Subjects in the extension (sustainability) SAS will contribute to the evaluation "as treated".

Extension phase (switch)
Extension (switch) FAS: Includes all subjects randomised to sitagliptin in the main phase and re-randomised to either continue on sitagliptin or switch to oral semaglutide in the extension phase. Subjects in the extension (switch) FAS will contribute to the evaluation "as randomised".

Extension (switch) SAS: Includes all subjects randomised to sitagliptin in the main phase and re-randomised and exposed with at least one dose of either sitagliptin or oral semaglutide in the extension phase. Subjects in the extension (switch) SAS will contribute to the evaluation "as treated".

Data selections and observation periods

Unless subjects withdraw their informed consent, data collection will continue for the full duration of the trial. The full duration of the trial is defined as up to and including:

- the follow-up visit V11 for subjects who do not discontinue trial product in the main phase, but decide not to continue in the extension phase
- the latest occurring visit of the end-of-treatment visit V10 or the follow-up premature treatment discontinuation visit V11A for subjects who discontinue trial product prematurely in the main phase
- the follow-up visit V20 for subjects who continue in the extension phase and do not discontinue trial product in the extension phase
- the latest occurring visit of the end-of-treatment visit V19 or the follow-up premature treatment discontinuation visit V20A for subjects who discontinue trial product prematurely in the extension phase

Subjects and data to be used in an analysis will be selected in a two-step manner:

- Firstly, subjects will be selected based on the specified analysis set
- Secondly, data points on the selected subjects from the first step will be selected based on the specified observation period

Definition of the observation periods:

In-trial: This observation period represents the time period where subjects are considered to be in the trial, regardless of discontinuation of trial product or initiation of rescue medication.

Main phase
The in-trial observation period starts at randomisation (as registered in the IWRS) and ends at the date of:

- the last direct subject-site contact, which is scheduled to take place 5 weeks after planned last dose of trial product at the follow-up visit V11 for subjects who decide not to continue in the extension phase
- V10 for subjects who continue in the extension phase
- withdrawal for subjects who withdraw their informed consent
- the last subject-investigator contact as defined by the investigator for subjects who are lost to follow-up
- death for subjects who dies before any of the above

Extension phase (sustainability)
The in-trial observation period starts at randomisation in main phase (as registered in the IWRS) and ends at the date of:

- the last direct subject-site contact, which is scheduled to take place 5 weeks after planned last dose of trial product at the follow-up visit V20 for subjects who continue in the extension phase
- withdrawal for subjects who withdraw their informed consent
- the last subject-investigator contact as defined by the investigator for subjects who are lost to follow-up
- death for subjects who dies before any of the above

Extension phase (switch)
The in-trial observation period starts at re-randomisation (as registered in the IWRS) planned at V10 and ends at the date of:

- the last direct subject-site contact, which is scheduled to take place 5 weeks after planned last dose of trial product at the follow-up visit V20 for subjects who continue in the extension phase
- withdrawal for subjects who withdraw their informed consent
- the last subject-investigator contact as defined by the investigator for subjects who are lost to follow-up
- death for subjects who dies before any of the above

On-treatment: This observation period represents the time period where subjects are considered treated with the trial product. The observation period is a subset of the in-trial observation period.

Main phase
The on-treatment observation period starts at the date of first dose of trial product. Two slightly different end dates will be needed to cover all assessments appropriately.

For adjudicated events, ECGs, eye examination category and AEs including hypoglycaemic episodes, the observation period ends at the first date of any of the following:

- the follow-up visit V11 for subjects on trial product who decide not to continue in the extension phase
- the follow-up prematurely discontinuation visit V11A
- the last date on trial product +38 days
- the end-date for the in-trial observation period

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

Extension phase (sustainability)
The on-treatment observation period starts at the date of first dose of trial product.

For adjudicated events, ECGs, eye examination, and AEs including hypoglycaemic episodes, the observation period ends at the first date of any of the following:

- the follow-up visit V20 for subjects on trial product who continue in the extension phase
- the follow-up prematurely discontinuation visit V20A
- the last date on trial product +38 days
- the end-date for the in-trial observation period

For efficacy and other safety assessments (laboratory assessments, physical examination and vital signs) the observation period ends at the last date on trial product +3 days.

Extension phase (switch)
The on-treatment observation period starts at the date of first dose of trial product in the extension phase planned at V10.

For adjudicated events, ECGs, eye examination, and AEs including hypoglycaemic episodes, the observation period ends at the first date of any of the following:

- the follow-up visit V20 for subjects on trial product who continue in the extension phase
- the follow-up prematurely discontinuation visit V20A
- the last date on trial product +38 days
- the end-date for the in-trial observation period

For efficacy and other safety assessments (laboratory assessments, physical examination and vital signs) the observation period ends at the last date on trial product +3 days.

The follow-up visit (V11/V11A/V20/V20A) is scheduled to take place 5 weeks after the last date on trial product corresponding to approximately five half-lives of oral semaglutide. The visit window for the follow-up visit is +3 days.

On-treatment without rescue medication: This observation period is a subset of the on-treatment observation period, where subjects are considered treated with trial product, but have not initiated any rescue medications.

Main phase
The on-treatment without rescue medication observation period starts at the date of first dose of trial product and ends at the first date of any of the following:

- the last dose of trial product +3 days
- initiation of rescue medication

Extension phase (sustainability)
The on-treatment without rescue medication observation period starts at the date of first dose of trial product and ends at the first date of any of the following:

- the last dose of trial product +3 days
- initiation of rescue medication

Extension phase (switch)
For the two treatment arms in the extension phase (switch), the current prescribed rescue medication (at V10) in the main phase will be treated as background medication in the extension phase and only data collected after initiation of rescue medication in the extension phase will be excluded.

Specifically the on-treatment without rescue medication observation period starts at the date of first dose of trial product in the extension phase, planned at V10, and ends at the first date of any of the following:

- the last dose of trial product +3 days
- initiation of rescue medication in the extension phase

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

Data points collected outside an observation period will be treated as missing in the analysis. Baseline data will always be included in an observation period. For adjudicated events, the onset date will be the EAC adjudicated onset date.

Before data are locked for statistical analysis and the randomisation code is broken, a review of all available data will take place. Any decision to exclude either a subject or single observations from the statistical analysis is the joint responsibility of the members of the Novo Nordisk study group.

Exclusion of data from analyses should be used restrictively, and normally no data should be excluded from the FAS. The subjects or observations to be excluded, and the reasons for their exclusion must be documented and signed by those responsible before database lock. The subjects and observations excluded from analysis sets, and the reason for this, will be described in the clinical trial report.

Confirmatory hypotheses

For the primary HbA_{1c} endpoint the following confirmatory one-sided hypothesis are planned to be tested where OR refers to the odds ratio for oral semaglutide versus sitagliptin.

- H_0: OR ≤ 1 against H_a: OR >1

For the confirmatory secondary endpoint in the main phase (change from baseline to week 52 in body weight), the following hypothesis is planned to be tested, where μ refers to the treatment difference between oral semaglutide minus sitagliptin.

- H_0: $\mu \geq 0.0$ kg against H_a: $\mu < 0.0$ kg

For the two confirmatory endpoints in the extension phase (switch; change from week 52 to week 104 in HbA_{1c} and change from week 52 to week 104 in body weight), the following confirmatory one-sided hypotheses are planned to be tested for oral semaglutide versus sitagliptin.

- H_0: $\mu \geq 0.0$ %-point against H_a: $\mu < 0.0$ %-point

- H_0: $\mu \geq 0.0$ kg against H_a: $\mu < 0.0$ kg

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

Multiplicity and criteria for confirming hypotheses

The type I error for testing the four confirmatory hypotheses related to the HbA_{1c} and body weight endpoints will be preserved in the strong sense at 5% (two-sided) using an hierarchical testing strategy as outlined in Figure 2–1.

Superiority of the primary hypotheses will be considered confirmed if the odds ratio is supporting the corresponding alternative hypothesis and the two-sided p-value from the primary analysis of the primary estimand is strictly below the 5% two-sided significance level. This is equivalent to using a one-sided p-value (nominal $\alpha = 0.025$) and a one-sided 2.5% overall significance level. If the primary hypothesis is confirmed, the testing strategy will continue by testing the following confirmatory secondary hypothesis at a 5% two-sided significance level. Superiority of the following secondary confirmatory hypotheses will then be considered confirmed if the treatment difference is supporting the corresponding alternative hypothesis and the two-sided p-value from the primary analysis of the primary estimand is strictly below the 5% two-sided significance level and so on. If for any of the four confirmatory hypotheses, superiority is not confirmed the testing will terminate.

2.3 Primary endpoint

The primary endpoint is if a subject after week 52 achieves (yes/no) $HbA_{1c} < 7\%$ (53 mmol/mol) ADA target.

2.3.1 Primary analysis for the primary estimand

The primary estimand for the primary endpoint will be estimated based on the FAS using week 52 measurements from the in-trial observation period. The primary statistical analysis will be a logistic regression model with treatment, region, and stratification factor as fixed effects and baseline HbA_{1c} measurement as a covariate. Firstly, a pattern mixture model using multiple imputation is used to impute missing values for continuous HbA_{1c} assessments assuming that the missing data mechanism is missing at random (MAR) within the groups used for imputation. Imputation of missing data at week 52 will be done within 4 groups of subjects defined by randomised treatment arm, and whether subjects at week 52; (i) have discontinued treatment or initiated rescue medication or (ii) are still on treatment and have not initiated rescue medication. It is hereby assumed that the likely values of what the missing data would have been if available are best described by information from subjects who at week 52 are similar in terms of randomised treatment arm and treatment adherence/rescue status.

Missing values for each group will be imputed as follows:

- An analysis of covariance (ANCOVA) with region and stratification factor as categorical fixed effects and baseline HbA_{1c} measurement as a covariate will be fitted to observed values of the change from baseline to week 52 in HbA_{1c}.
- The estimated parameters for location and dispersion will be used to impute 1000 values for each subject with missing week 52 data based on region and stratification factor as categorical fixed effects and baseline HbA_{1c} measurement as a covariate. Thus, 1000 complete data sets will be generated including observed and imputed values.
- The binary endpoint will be created for each of the 1000 complete data sets.
- A logistic regression model with treatment, region, and stratification factor as fixed effects and baseline HbA_{1c} measurement as a covariate is fitted for each of the imputed datasets.
- The resulting 1000 estimates and variances will be combined using Rubin's rule[5], and the estimated odds ratio between oral semaglutide and sitagliptin together with two-sided 95 % CI and two-sided p-value for testing no difference from one will be presented.

2.3.2 Primary analysis for the secondary estimand

The secondary estimand for the primary endpoint will be estimated based on the FAS using post-baseline measurements up to and including week 52 from the on-treatment without rescue medication observation period. The primary statistical analysis will be a logistic regression model with treatment, region, and stratification factor as fixed effects and baseline HbA_{1c} measurement as covariate. First; intermittent missing values for continuous HbA_{1c} assessments will be imputed using Markov Chain Monte Carlo (MCMC) method to obtain a monotone missing data pattern. This imputation is done for each treatment group separately and 1000 copies of the dataset will be generated. Second; based on these 1000 datasets the remaining missing values will be imputed sequentially. For each treatment group an analysis of covariance model will be used to impute missing values at each planned visit. The model will include region and stratification factor as categorical effects and baseline and post-baseline HbA_{1c} values prior to the visit in question as covariates. Third; the binary endpoint will be created, analysed and the resulting 1000 estimates and variances will be combined using Rubin's rule[46].

2.3.3 Sensitivity analyses

To investigate the sensitivity of the primary analysis results, complementary and separate analyses will be performed for the primary and secondary estimand. In line with European Medicines Agency recommendations (CHMP 2010)[6] and with a report from the US National Research Council (NAS 2010)[7], these analyses will primarily evaluate the sensitivity of the results due to the impact of missing data.

The evaluation of the robustness of the primary analysis will in all cases be a logistic regression model with treatment, region, and stratification factor as fixed effects and baseline HbA_{1c} measurement as a covariate. However missing data will be based on different pattern mixture model

approaches using multiple imputation. The binary endpoint will be created using the imputed continuous data. An overview of the sensitivity analyses for each of the estimands are specified below followed by a more detailed description of the different pattern mixture models used.

Sensitivity analyses for the primary estimand

The estimation of the primary estimand will be repeated using the following sensitivity analyses for imputing missing continuous HbA_{1c} assessments before testing the binary primary endpoint using logistic regression:

- A comparator multiple imputation analysis based on FAS using the in-trial observation period
- A tipping-point multiple imputation analysis based on FAS using the in-trial observation period.

Sensitivity analyses for the secondary estimand

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

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

2.3.3.1 Pattern mixture models

Common for the pattern mixture model sensitivity analyses is that they all aim to stress-test the primary HbA_{1c} results by changing the assumptions for missing data in the oral semaglutide treatment arms, while maintaining the missing at random data assumption for the sitagliptin arm:

- *Comparator multiple imputation analysis:* In this sensitivity analysis missing data at week 52 for all subjects will be imputed to resemble the distribution of the week 52 values observed in the sitagliptin treatment arm. In effect, this imputation approach removes the treatment difference between oral semaglutide and sitagliptin for all subjects randomised to oral semaglutide, given that oral semaglutide is better than sitagliptin.
- *Tipping-point multiple imputation analysis:* In this sensitivity analysis, missing data will first be imputed according to the primary analysis. Secondly, for the oral semaglutide treatment arm a penalty will be added to the imputed values at week 52. The approach is to gradually increase this penalty until a confirmed HbA_{1c} conclusion from the primary analysis is changed. The specific value of the penalty that changes the conclusion will be used to evaluate the robustness of the primary analysis result.

Assessment of sensitivity analyses

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

2.4 Secondary endpoints

2.4.1 Confirmatory secondary endpoints

Main phase

Change from baseline to week 52 in body weight (kg)

The primary estimand will be estimated using the same approaches as described for the primary HbA_{1c} endpoint without dichotomizing the endpoint and with a linear normal regression model instead of the logistic regression model. Baseline body weight will be used as a covariate instead of baseline HbA_{1c} for the endpoints related to body weight in both the multiple imputation and analysis models.

The analyses for the secondary estimand will be a mixed model for repeated measure (MMRM). The MMRM based analysis will use a restricted maximum likelihood (REML) and include all post-baseline measurements collected at scheduled visits up to and including week 52 as dependent variables. The independent effects included in the model will be treatment, region, and stratification factor as categorical fixed effects and baseline value as a covariate, all nested within visit. An unstructured covariance matrix for measurements within the same subject will be employed, assuming measurements from different subjects are independent.

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

Extension phase (switch)

Change from week 52 to week 104 in HbA_{1c}

Change from week 52 to week 104 in body weight (kg)

The primary estimand will be based on extension phase (switch) FAS using the in-trial extension phase (switch) observation period. Whereas, the secondary estimand will be based on extension

phase (switch) FAS using the on-treatment without rescue medication, extension phase (switch) observation period.

The estimands will be estimated using the same approaches as described for the primary HbA_{1c} endpoint without dichotomizing the endpoint and with a linear normal regression model instead of the logistic regression model. Baseline body weight will be used as a covariate instead of baseline $HbA_{1c,}$ in the analyses of the body weight endpoint, in both the multiple imputation and analysis models.

The stratification factors, HbA_{1c} below 7% (yes/no) and if the subject is currently on prescribed rescue medication and continuing rescue medication in the extension phase (yes/no), and the interaction of these will be included as fixed effects in the analyses.

Superiority will be considered confirmed if the mean treatment difference is supporting the corresponding hypothesis and the two-sided p-value from the analysis of the primary estimand is strictly below its local two-sided significance level resulting from the closed testing procedure in Figure 2–1. Sensitivity analyses similar to the ones pre-specified for testing superiority for the primary HbA_{1c} endpoint will be made to evaluate the robustness of the confirmatory body weight results for the main phase. The tipping point analysis will be used as a sensitivity analysis for the confirmatory results for the extension phase (switch).

2.4.2 Supportive secondary endpoints

2.4.2.1 Efficacy endpoints

The below supportive secondary efficacy endpoints will be evaluated for:

Main phase

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

Extension phase (sustainability)

- extension phase (sustainability) FAS using the in-trial extension phase (sustainability) observation period
- extension phase (sustainability) FAS using the on-treatment without rescue medication, extension phase (sustainability) observation period

Extension phase (switch)

- the primary estimand based on extension phase (switch) FAS using the in-trial extension phase (switch) observation period
- the secondary estimand based on extension phase (switch) FAS using the on-treatment without rescue medication, extension phase (switch) observation period

No sensitivity analyses are planned for these.

Continuous efficacy endpoints

Main phase

Change from baseline to week 52 in:

- HbA_{1c}
- FPG
- Body weight (%)
- BMI
- Waist circumference
- Fasting lipid profile (total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides)

Extension phase (sustainability)

Change from baseline to week 104 in:

- HbA_{1c}
- FPG
- Body weight (kg)
- Body weight (%)
- BMI
- Waist circumference

Extension phase (switch)

Change from week 52 to week 104 in:

- FPG
- Body weight (%)
- BMI
- Waist circumference

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

Main and extension phase (switch)

The above continuous endpoints will be analysed separately using similar model approaches as for the confirmatory secondary endpoints for the main phase and the extension phase (switch), respectively, with the associated baseline response as a covariate. Fasting lipid profile endpoints (main phase only) will be log-transformed prior to analysis with the associated log-transformed baseline value as a covariate.

Extension phase (sustainability)

The endpoints will be summarised descriptively by visit.
Furthermore, the change from baseline to week 52 in HbA_{1c} and in body weight will be presented using both the main phase FAS and the extension phase (sustainability) FAS to explore differences in HbA_{1c} and in body weight for subjects not continuing in the extension phase (sustainability).

Binary efficacy endpoints

Main phase

If a subject after 52 weeks achieves (yes/no):

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

Extension phase (sustainability)

If a subject after 104 weeks achieves (yes/no):

- $HbA_{1c} < 7\%$ (53 mmol/mol) ADA target
- $HbA_{1c} \leq 6.5\%$ (48 mmol/mol) AACE target
- $HbA_{1c} < 7\%$ (53 mmol/mol) or HbA_{1c} reduction ≥ 1%-point (10.9 mmol/mol)
- Weight loss ≥ 5%
- $HbA_{1c} < 7\%$ (53 mmol/mol) without treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes and no weight gain

Extension phase (switch)

If a subject after 104 weeks achieves (yes/no):

- HbA_{1c} < 7% (53 mmol/mol) ADA target
- HbA_{1c} ≤ 6.5% (48 mmol/mol) AACE target
- Weight loss ≥ 5% compared to week 52
- HbA_{1c} < 7% (53 mmol/mol) without treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes after week 52 and no weight gain compared to week 52
- HbA_{1c} < 7% (53 mmol/mol) and no need for rescue medication after week 52
- No need for rescue medication after week 52

Main and extension phase (switch)

The above binary endpoints will be analysed separately using similar model approaches as for the primary endpoint with the associated baseline response, HbA_{1c} or body weight, as a covariate. When addressing the primary estimand for the composite endpoint the 'without hypoglycaemic episodes' component will include non-treatment-emergent events of severe or BG-confirmed symptomatic hypoglycaemia as data collected regardless of discontinuation of trial product or initiation of rescue medication(s) is used.

The analysis of the endpoint for rescue medication in extension phase (switch) will not include a baseline covariate as rescue medication in the main phase (yes/no) is used as a stratification factor in the model.

Extension phase (sustainability)

The endpoints will be summarised descriptively by visit.

Time to event endpoint

The analyses of time to event endpoints will be based on the below two definitions:

- *Additional anti-diabetic medication:* New anti-diabetic medication and/or intensification of anti-diabetic medication initiated at or after randomisation and before (planned) end-of-treatment
- *Rescue medication:* New anti-diabetic medication and/or intensification of anti-diabetic medication initiated at or after randomisation and before last date on trial product. This is a subset of the *additional anti-diabetic medication*.

The following rules will be applied based on the concomitant medication data reported by the investigator, to determine whether or not the recorded anti-diabetic medication is either 1) new anti-diabetic medication or 2) intensification of anti-diabetic medication

1. *New anti-diabetic medication:* Anti-diabetic medication (4th-level ATC code) that is initiated after randomisation and is new compared to the anti-diabetic background medication at randomisation and with a dosing duration of more than 21 days
2. *Intensification of anti-diabetic medication:* A more than 20% increase in the dose of anti-diabetic medication after randomisation as compared to the anti-diabetic medication dose at randomisation (5th-level ATC code not changed) and with a dosing duration of more than 21 days.

Main phase

- Time to event

For the primary estimand the above time to event endpoint will be analysed based on FAS using the in-trial observation period. Additional anti-diabetic medication will be considered an event regardless of treatment adherence. Time from randomisation to additional anti-diabetic medication will be analysed using a Cox proportional hazards model with treatment, region and stratification factor as categorical fixed effects and baseline HbA_{1c} as a covariate. From this analysis the estimated hazard ratios between oral semaglutide and sitagliptin together with associated two-sided 95% CIs and unadjusted two-sided p-values will be presented. The analysis aims to address the need of additional anti-diabetic medication regardless of if this is due to lack of effect or tolerability. Switch to other anti-diabetic treatment is therefore also considered an event and withdrawn subjects or subjects lost to follow-up will be considered as having an event on the day of withdrawal. Subjects will be censored on the day before planned end-of-treatment visit.

For the secondary estimand the above time to event endpoint will be analysed based on FAS using the on-treatment without rescue medication observation period. Time from first dose of trial product to initiation of rescue medication will be analysed using the same model as described above. The analysis aims to address lack of effect and only initiation of rescue medication as add-on to

randomised treatment is considered an event. Switch to other anti-diabetic treatment is not considered an event and as a consequence subjects will be censored on the day before last date on trial product. Events occurring between randomisation and first date on trial product will be included in the analysis as events at day 0, in order to count all events of rescue medication.

Extension phase (switch)

- Time to event after week 52

For the primary estimand the above time to event endpoint will be analysed based on FAS using the in-trial extension phase (switch) observation period. Additional anti-diabetic medication will be considered an event regardless of treatment adherence. Time from re-randomisation to additional anti-diabetic medication will be analysed using a Cox proportional hazards model with treatment, region, stratification factors and the interaction between the two stratification factors as categorical fixed effects and baseline HbA_{1c} as a covariate.

For the secondary estimand the above time to event endpoint will be analysed based on FAS using the on-treatment without rescue medication extension phase (switch) observation period. Time from first dose of trial product in the extension phase to initiation of rescue medication will be analysed using the same model as described above. For the extension phase (switch), rescue medication initiated in the main phase will be treated as background medication in the extension phase and only addition of rescue medication after week 52 (V10) will be counted in the analyses.

2.4.2.2 Safety endpoints and safety assessments

The safety endpoints and safety assessments for the main phase in the first clinical trial report will be evaluated based on SAS using the on-treatment observation period and based on SAS using the in-trial observation period unless otherwise stated.

The safety endpoints and safety assessments for the extension phase in the second clinical trial report will be evaluated based on extension phase (switch) and (sustainability) SAS using the on-treatment extension phase (switch) and (sustainability) observation period and based on extension phase (switch) and (sustainability) SAS using the in-trial extension phase (switch) and (sustainability) observation period unless otherwise stated.

The following endpoints and assessments are used to support the safety objectives.

Adverse events

Main phase

- Number of TEAEs during exposure to trial product, assessed up to approximately 52 weeks

Extension phase (sustainability)

- Number of TEAEs during exposure to trial product, assessed up to approximately 109 weeks

Extension phase (switch)

- Number of TEAEs during exposure to trial product, assessed from week 52 up to approximately 109 weeks

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

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

TEAEs will be summarised in terms of the number of subjects with at least one event (N), the percentage of subjects with at least one event (%), the number of events (E) and the event rate per 100 patient years of observation time (R) for the on-treatment observation period. Supportive summaries of AEs will be made for the in-trial observation period. The development over time in gastrointestinal AEs will be presented graphically.

Other safety endpoints

Main phase

Change from baseline to week 52 in:

- Amylase (part of biochemistry)
- Lipase (part of biochemistry)
- Pulse
- Systolic blood pressure
- Diastolic blood pressure

Extension phase (sustainability)

Change from baseline to week 104 in:

- Amylase (part of biochemistry)
- Lipase (part of biochemistry)
- Pulse
- Systolic blood pressure
- Diastolic blood pressure

Extension phase (switch)

Change from week 52 to week 104 in:

- Amylase (part of biochemistry)
- Lipase (part of biochemistry)
- Pulse
- Systolic blood pressure
- Diastolic blood pressure

Main and extension phase (switch)

The above safety endpoints will be analysed as described above for continuous efficacy endpoints. Amylase and lipase endpoints will be log-transformed prior to analysis with the associated log-transformed baseline value as a covariate.

Extension phase (sustainability)

The endpoints will be summarised descriptively by visit.

The following safety assessments will be calculated similarly as for the above other safety endpoints:

Main phase

The safety assessments for the main phase will furthermore be calculated for the following:

- Haematology
- Biochemistry (except for amylase and lipase)
- Calcitonin
- ECG category
- Physical examination
- Eye examination category

Extension phase

The safety endpoints for the extension phase (switch) and extension phase (sustainability) will furthermore be calculated for the following:

- Haematology
- Biochemistry (except for amylase and lipase)
- Calcitonin
- Eye examination category

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

Hypoglycaemia

Main phase

- Number of treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed up to approximately 52 weeks
- Treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed up to approximately 52 weeks (yes/no)

Extension phase (sustainability)

- Number of treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed up to approximately 109 weeks
- Treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed up to approximately 109 weeks (yes/no)

Extension phase (switch)

- Number of treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed from week 52 up to approximately 109 weeks
- Treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during exposure to trial product, assessed from week 52 up to approximately 109 weeks (yes/no)

Classification of hypoglycaemia

Hypoglycaemic episodes will be summarised for the SAS (main phase), extension phase (sustainability) SAS and extension phase (switch) SAS, using only the on-treatment (main phase), on-treatment extension phase (sustainability) and on-treatment extension phase (switch) observation period, respectively (referred to as the on-treatment observation periods in the following).

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

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

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

Novo Nordisk classification of hypoglycaemia

In normal physiology, symptoms of hypoglycaemia occur below a PG level of 3.1 mmol/L (56 mg/dL)[8]. Therefore, Novo Nordisk has included hypoglycaemia with PG levels below this cut-off point in the definition of BG-confirmed hypoglycaemia.

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

Severe or BG-confirmed symptomatic hypoglycaemia: An episode that is severe according to the ADA classification[2] or BG-confirmed by a PG value < 3.1 mmol/L (56 mg/dL) **with** symptoms consistent with hypoglycaemia.

ADA classification[2] of hypoglycaemia

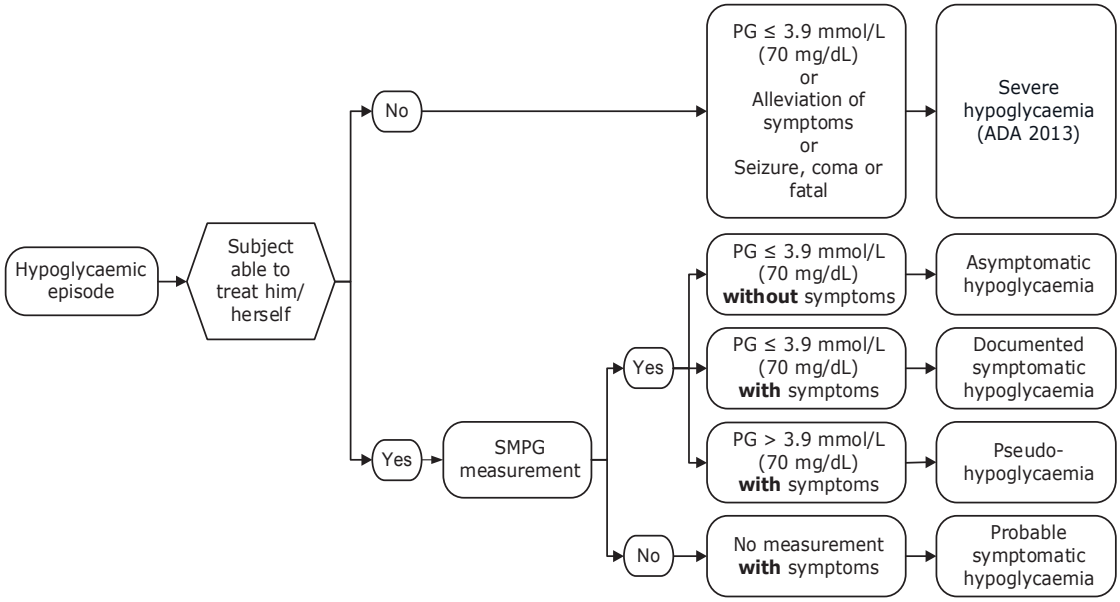
Severe hypoglycaemia: An episode requiring assistance of another person to actively administer carbohydrate, glucagon, or take other corrective actions. PG concentrations may not be available during an event, but neurological recovery following the return of PG to normal is considered sufficient evidence that the event was induced by a low PG concentration.

Asymptomatic hypoglycaemia: An episode not accompanied by typical symptoms of hypoglycaemia, but with a measured PG concentration ≤ 3.9 mmol/L (70 mg/dL).

Documented symptomatic hypoglycaemia: An episode during which typical symptoms of hypoglycaemia are accompanied by a measured PG concentration ≤ 3.9 mmol/L (70 mg/dL).

Pseudo-hypoglycaemia: An episode during which the person with diabetes reports any of the typical symptoms of hypoglycaemia with a measured PG concentration > 3.9 mmol/L (70 mg/dL) but approaching that level.

Probable symptomatic hypoglycaemia: An episode during which symptoms of hypoglycaemia are not accompanied by a PG determination but that was presumably caused by a PG concentration ≤ 3.9 mmol/L (70 mg/dL).



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

PG: plasma glucose SMPG: Self-measured plasma glucose

Figure 2–2 **ADA classification of hypoglycaemia**

Data on treatment-emergent hypoglycaemic episodes will be presented in terms of the number of subjects with at least one episode, the percentage of subjects with at least one episode (%), the total number of episodes and the episode rate per 100 patient years of observation time.

Analysis of severe or BG-confirmed symptomatic hypoglycaemic episodes

Main and extension phase (switch)

The endpoints related to number of treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episodes during the on-treatment period will be analysed using a negative binomial regression model with a log-link function and the logarithm of the duration of the subject's on-treatment observation period as offset. The model will include treatment, stratification factor and region as fixed factors and baseline HbA_{1c} as covariate.

The binary endpoints showing whether a subject has at least one treatment-emergent severe or BG-confirmed symptomatic hypoglycaemic episode will be analysed using a logistic regression model with treatment, stratification factor and region as fixed factors and baseline HbA_{1c} as covariate.

Extension phase (sustainability)

The endpoints will be summarised descriptively.

2.5 Interim analysis

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

A database lock for the main phase will be performed when all subjects have completed the main phase to be able to report the results in the first clinical trial report. After completion of the extension phase, there will be a second database lock followed by the reporting of the results of the entire trial in the second clinical trial report.

2.6 Patient reported outcomes

Main phase

Change from baseline to week 52 in:

- SF-36v2® Health Survey (acute version) (SF-36v2 (acute version)): scores from the 8 domains and the physical component score and mental component score summary scores
- Diabetes Treatment Satisfaction Questionnaire – status (DTSQs): individual items and total treatment satisfaction score (6 of the 8 items summed)

Extension phase (sustainability)

Change from baseline to week 104 in:

- SF-36v2 (acute version): scores from the 8 domains, the physical component summary score and mental component summary scores
- DTSQs: individual items and total treatment satisfaction score (6 of the 8 items summed)

Extension phase (switch)

Change from week 52 to week 104 in:

- SF-36v2 (acute version): scores from the 8 domains, the physical component summary score and mental component summary scores
- DTSQs: individual items and total treatment satisfaction score (6 of the 8 items summed)

The PRO endpoints will be based on FAS, extension phase (sustainability) FAS and the extension phase (switch) FAS, using the on-treatment without rescue medication, on-treatment without rescue medication extension phase (sustainability) and on-treatment without rescue medication extension phase (switch) observation periods, respectively.

Main and extension phase (switch)

All of the above individual items and scores will be analysed separately as the other supportive continuous efficacy endpoints with the associated baseline response as a covariate. For SF-36v2 the primary analysis for the hypothetical estimand will be reported in a report separate from the CTR.

Extension phase (sustainability)

The endpoints will be summarised descriptively by visit.

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

The SF-36v2® Health Survey (SF-36v2) (acute version) instrument is a commonly used generic instrument measuring health-related quality of life (HRQoL)/general health status across disease areas including diabetes[9]. The SF-36v2 (acute version) for adults with a 1-week recall period contains 36 items. A total of 35 items measure eight domains of functional health and well-being as well as two component summary scores: physical functioning (10 items), role limitation due to physical health problems (4 items), bodily pain (2 items), general health perceptions (5 items), vitality (4 items), social functioning (2 items), role limitations due to emotional problems (3 items) and general mental health (5 items), mental component summary (MCS) score, physical component summary (PCS) score. There is an additional single item giving information on health change over the past week.

Norm-based scores (NBS) will be derived using the QualityMetric Health Outcomes™ Scoring Software[9] including the 2009 US general population norm. The most recent version of the

QualityMetric Health Outcomes™ Scoring Software available at time of licensing was used for the specific trial (version 4.5 for PIONEER 7). Table 2–2 provides an overview of the domains. NBS standardises domain and component scores into T-scores using the means and standard deviations from the US general population. Higher scores on all domains and component summary measures (PCS and MCS) indicate better HRQoL/general health status. Item 2 (i.e. Question 2 in CRF) is not included in any score.

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

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

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

The responder threshold values, in terms of T-score points for change from baseline are defined in Table 2–3.

Table 2–3 Responder thresholds for SF-36v2 (acute version)

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

Responder analyses will be based on the responder threshold values and are described in section 2.6.3.

2.6.2 Diabetes Treatment Satisfaction Questionnaire – status version (DTSQs)

The DTSQs questionnaire will be used to assess subject's treatment satisfaction. This questionnaire contains 8 items that measures the treatment satisfaction for subjects' diabetes treatment in terms of convenience, flexibility and general feelings regarding treatment.

The DTSQs items are scored on a 7-point graded response scale ranging from 6 to 0. Higher scores indicate higher levels of treatment satisfaction for DTSQs items 1, 4 -8. For items 2 and 3 a higher score indicates a higher patient perceived experience of hyperglycaemia and hypoglycaemia, respectively. Thus, lower scores indicate a perception of blood glucose levels being "none of the time" unacceptably high (item 2) or low (item 3). If data are missing for an item, the item score is treated as missing. No reversal of item scores will be done.

The domain score of total treatment satisfaction (total treatment sat isfaction score) is computed by adding the six items scores 1, 4-8. The score has a minimum of zero and a maximum of 36. A higher treatment satisfaction score indicates a higher level of treatment satisfaction. No reversals of items are necessary prior to computing the treatment satisfaction score.

Missing data at instrument level will be handled in the following way. For computing the total treatment satisfaction score consisting of six items, missing data from one item is allowed.

Scoring algorithm:

- Step 1: Sum the existing item scores (i.e. either 5 or 6 item scores)
- Step 2: Divide this sum by the number of existing item scores
- Step 3: Multiply by 6 (the number of items in the total treatment satisfaction scale)

Half of a standard deviation (SD) of the baseline DTSQs item and domain scores per trial were used as distribution-based approach defining the responder thresholds. The thresholds are derived from baseline DTSQs data across trial arms per trial. Responder analyses will be based on the responder threshold values and are described in section 2.6.3.

2.6.3 Responder analyses

Responder analyses will be conducted for both estimands, for the same time points that are defined for the analyses of PRO endpoints (see protocols) and separately for each score.

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

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

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

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

The binary responder endpoints will be analysed as the other supportive secondary binary efficacy endpoints. Estimated proportions and differences in proportions will be reported in addition to odds and odds ratios.

The responder analyses will not be included in the CTR, but in a separate PRO report.

3 Changes to the statistical analyses planned in the protocol

In section 2 it has been specified in more detail what the second CTR will include.

In section 2.3.2 the analyses for the secondary estimand was specified incorrectly. The correct analysis has been described.

In section 2.3.3 the amount of sensitivity analyses has reduced for the trial. The tipping point analysis has been kept for both the primary and the secondary estimand as an approach for performing a sensitivity analysis under the missing at not random (MNAR) assumption. In other words, the tipping point approach is like a progressive stress-testing to assess how severe departures from missing at random (MAR) must be in order to overturn conclusions from the primary analysis. This sensitivity analysis is considered sufficient to stress-test the secondary estimand. In addition to the tipping point analysis one sensitivity analyses for primary estimand is included, to get a better understanding of the primary results.

In regards to other sensitivity analysis, the LOCF analysis will not be performed anyway. This analysis is not judged to add additional value to the evaluation of the effect as it seems unrealistic that data from a subject with missing data, would have been stable from the point of dropout to trial completion.

In section 2.4.2.1 the statistical analyses of the two binary effect endpoints (HbA_{1c} reduction $\geq$ 1%-point (10.9 mmol/mol) and body weight loss $\geq$ 3%) have been omitted, because they are being analysed as a part of the two composite binary effect endpoints. The analyses of the composite endpoints including the 'without hypoglycaemic episodes' component and the analyses of the time to event endpoints have been described in more detail.

In section 2.4.2.2 the listings of adverse events and hypoglycaemic episodes reported from week 52 (V10) to week 57 (V11) for subjects not continuing into extension has been removed. All adverse events and hypoglycaemic episodes for these subjects are a subset of other listings, hence all events are included.

In section 2.6 details on the analyses of PRO endpoints have been included.

4 References

1. Novo Nordisk A/S. Investigator's Brochure for oral semaglutide (NN9924), edition 8 or any updates hereof. 2017.

2. Seaquist ER, Anderson J, Childs B, Cryer P, Dagogo-Jack S, Fish L, et al. Hypoglycemia and diabetes: a report of a workgroup of the American Diabetes Association and the Endocrine Society. Diabetes Care. 2013;36(5):1384-95.

3. Merck Sharp & Dohme Limited. JANUVIA - European Summary of Product Characteristics Annex 1. 26 May 2015.

4. Ahrén B, Comas LM, Kumar H, Sargin M, Derving Karsbøl J, Jacobsen SH, et al. Efficacy and Safety of Once-weekly Semaglutide vs Sitagliptin as add-on to Metformin and/or Thiazolidinediones After 56 Weeks in Subjects With Type 2 Diabetes (SUSTAIN 2). European Association for the Study of Diabetes, 52nd meeting, ePoster #767 2016.

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

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

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

8. Schwartz NS, Clutter WE, Shah SD, Cryer PE. Glycemic thresholds for activation of glucose counterregulatory systems are higher than the threshold for symptoms. J Clin Invest. 1987;79(3):777-81.

9. Maruish, M. E. (2011). *User's manual for the SF-36v2 Health Survey (3rd ed.)*. Lincoln, RI: QualityMetric Incorporated.

16.1.9.1 Pre-defined MedDRA search – list of preferred terms

Table of contents

Page

16.1.9.1 Pre-defined MedDRA search – list of preferred terms 2

Trial ID: NN9924-4257
16.1.9 Documentation of statistical methods, Version 1.0, dated 19-June-2018
Overview of deleted pages

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