Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT02453711
Sponsor trial ID:	NN9536-4153
Official title of study:	Investigation of safety and efficacy of once-daily semaglutide in obese subjects without diabetes mellitus
Document date:	24 October 2017

16.1.9 Documentation of statistical methods

List of contents

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

.

Statistical Analysis Plan

Trial ID: NN9536-4153

Investigation of safety and efficacy of once-daily semaglutide in obese subjects without diabetes mellitus

A 52-week, randomised, double-blind, placebo-controlled, sixteen-armed, parallel group, multi-centre, multinational trial with liraglutide 3.0 mg as active comparator

Author:

[redacted]

Table of contents Page

List of abbreviations

ADA	American Diabetes Association
AE	adverse event
ALT	Alanine aminotransaminase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
BG	blood glucose
BMI	body mass index
BMR	basal metabolic rate
CPK	creatine phosphokinase
C-SSRS	Columbia Suicidality Severity Rating Scale
CRF	case report form
CTR	clinical trial report
DPP-4	dipeptidyl peptidase 4
DUN	dispensing unit number
EAC	event adjudication committee
ECG	electrocardiogram
eCRF	electronic case report form
FAS	full analysis set
FPFV	first patient first visit
FPG	fasting plasma glucose
GCP	Good Clinical Practice
GLP-1	glucagon-like peptide-1
GLP-1 RA	glucagon-like peptide-1 receptor agonist
HbA_{1c}	glycosylated haemoglobin
HRQoL	health-related quality of life
hsCRP	High-sensitivity C reactive protein
ICMJE	International Committee of Medical Journal Editors

IEC	independent ethics committee
IRB	institutional review board
IWRS	interactive web response system
IWQoL-Lite	Impact of Weight on Quality of Life Lite
LDL	low density lipoprotein
LPLV	last patient last visit
MAR	missing at random
MESI	medical event of special interest
MHP	mental health professional
MMRM	mixed model for repeated measurements
NRS	numeric rating scale
PD	pharmacodynamic
PHQ-9	Patient Health Questionnaire-9
PK	pharmacokinetic
PRO	patient reported outcome
RET	re-arranged during transfection
SAE	serious adverse event
s.c.	subcutaneous(ly)
SF-36	Short Form-36
SMBG	self-measured blood glucose
SUSAR	suspected unexpected serious adverse reaction
T2DM	type 2 diabetes mellitus
TEE	total energy expenditure
TMM	Trial Materials Manual
TSH	thyroid-stimulating hormone
UTN	Universal Trial Number

1 Introduction

1.1 Trial information

The trial is a 52-week, randomised, double-blind, placebo-controlled, sixteen-armed, parallel group, multi-centre, multinational trial comparing once-daily subcutaneous administration of semaglutide in five different doses (ranging from 0.05 mg/day to 0.4 mg/day) with placebo in obese subjects without diabetes mellitus. Once-daily administration of two doses of semaglutide (0.3 mg/day and 0.4 mg/day) will be tested in a fast escalation regimen to investigate the effect of a different regimen on efficacy, safety and tolerability. Additionally, liraglutide 3.0 mg/day is included as an active comparator. The total trial duration for the individual subjects will be approximately 60 weeks. The trial includes a 1-week screening period, followed by a 52-week treatment period and a follow-up visit after 59 weeks.

1.2 Scope of the statistical analysis plan

This SAP is based on the protocol "Investigation of safety and efficacy of once-daily semaglutide in obese subjects without diabetes mellitus", version 3.0, and amendment 3.

The SAP contains the statistical section 17 from the protocol with some clarifications as a few non-substantial changes where needed. Technical and detailed elaborations of the statistical analyses are specified in the protocol and the SPS.

2 Statistical considerations

If necessary, a statistical analysis plan (SAP) may be written in addition to the protocol, including a more technical and detailed elaboration of the statistical analyses. The SAP will be finalised before database lock and unblinding of the trial.

Definition of Estimands

Effectiveness estimand

The primary estimand is an effectiveness estimand (de facto) quantifying the average treatment effect of once-daily semaglutide relative to placebo and liraglutide 3.0 mg after 52 weeks, as add-on to nutritional and physical activity counselling, in all randomised subjects regardless of adherence to treatment.

Efficacy estimand

In addition, an efficacy estimand (de jure) is quantifying the average treatment effect of once-daily semaglutide relative to placebo and liraglutide 3.0 mg after 52 weeks, as add-on to nutritional and physical activity counselling, if all randomised subjects had adhered to the assigned treatment regimen for the entire planned duration of the trial.

Results from the statistical analysis will generally be presented by treatment differences with two-sided 95% confidence intervals.

The full analysis set (FAS) will be used in the analysis of efficacy endpoints. For safety endpoints the safety analysis set will be used.

The 8 different placebo arms will be pooled into one placebo treatment arm in the main analyses. This pooling assumes that there is no substantial effect of different placebo volumes or different dose escalation on the efficacy and safety endpoints. The validity of this assumption will be checked for the primary endpoint by plotting mean data for the 8 placebo arms separately, and by evaluating summaries of adverse events for each placebo arm.

For the statistical analysis of the primary endpoint, comparisons between each semaglutide dose/escalation arm, the liraglutide arm and their corresponding placebo arms will be performed in addition to comparisons with the pooled placebo arm. Statistical inference and data presentations will be separated into two parts. Part A concerns identifying the optimal dose and includes inference for the liraglutide arm, the semaglutide arms with dose escalation every fourth week, the corresponding placebo arms and the pool of the placebo arms. Part B concerns identifying the optimal dose escalation regime and includes inference for the semaglutide arms with dose escalation every second week, the corresponding semaglutide arms (with regards to dose) with dose escalation every fourth week, the corresponding placebo arms and the pool of the placebo arms.

Descriptive statistics for all efficacy and safety endpoints are always presented for each of the randomised treatment arms and the pool of the placebo arms. Unless otherwise specified, the following sections describe the planned analyses using one pooled placebo arm.

The baseline value will be defined as the last measured and available value from visit 1 and 2.

Laboratory values below the lower limit of quantification (LLOQ) will be set to ½LLOQ.

2.1 Sample size calculation

The sample size calculation is based on the primary endpoint; change from baseline in body weight (%) at 52 weeks.

Part A is a dose-finding trial examining five doses of semaglutide and placebo, and an active comparator. The sample size calculation is based on the relative change after 52 weeks treatment in the primary endpoint, body weight. In the sample size calculations, it is presumed that the placebo groups will be pooled in the analysis, assuming no correlation between body weight change after 52 weeks and placebo-injected volume. By trial design we expect to have 135 subjects on placebo and 100 in each active group. In the unfortunate situation that we should not be able to pool all placebo arms but only be able to pool the (semaglutide) placebo arms with dose escalation every fourth

week and (liraglutide) placebo, we will still expect to have at least 100 subjects in the relevant pool of placebo subjects. Hence the argumentation below only assumes 100 subjects in the placebo arm in a conservative manner and for simplicity. However, at time of analysis we plan to include all subjects on placebo as one group in the analysis as described above, if possible.

The following assume 100 subjects randomised in a balanced manner to receive each active treatment and being conservative at least 100 evaluable subjects in the placebo arm. In trial NN8022-1839, in obese subjects without T2DM a standard deviation of just below 7% was seen for observed weight loss (in %) in the liraglutide 3.0 mg arm. A conservative estimate of the dropout rate is 40%. A standard deviation of 7% and a sample size of 100 in each treatment arm will allow the 95% confidence interval for the estimated difference between two semaglutide doses, with 90% probability, to be contained within ±2.5% of the estimate, which is considered to be a sufficient precision for determining which doses to use for the continued development of semaglutide in the weight management indication.

For the primary endpoint change in body weight after 52 weeks of treatment, a difference (semaglutide minus placebo) of 9.5% is expected for completers (12% in the optimal dose group for semaglutide vs. 2.5% in placebo). For the withdrawn subjects, who are anticipated to constitute up to 40% of the total trial population, the treatment difference (semaglutide minus placebo) is assumed to be 0% giving an overall expected treatment difference of 5.7% (8.2% in semaglutide vs. 2.5% in placebo). The standard deviation will also be increased using the MI approach. The standard deviation in the final data is assumed to be up to 8.4% (8.4% in semaglutide vs. 7.0% in placebo. A standard deviation of up to 8.4% together with an expected difference of 5.7% results in a power of more than 99%, which is not corrected for multiple comparisons between different semaglutide arms and placebo.

For part B, in total two times 100 subjects will be randomised to the dose escalation every second week. Combined with placebo and the two doses corresponding to the ‘every second week’ arms (i.e. 0.3 mg/day and 0.4 mg/day) we will have more than five hundred subjects for the inference of the dose escalation finding part of the trial. In trial NN8022-1839, nausea, vomiting, and constipation were the most common gastrointestinal AEs with incidences between 15% and 40% of all subjects on liraglutide 3.0 mg. These AEs are not expected to be less frequent with semaglutide. This part of the trial is exploratory in nature and is intended to evaluate the overall safety profile with respect to the different types of events and when they occur compared to the escalation steps and with a view towards the efficacy response as well. With the given number of subjects, we have a reasonable sample size to detect marked clinical relevant difference between the arms.

2.2 Definition of analysis sets

The following analysis sets are defined in accordance with the ICH-E9[1] guidance:

- The full analysis set (FAS) will include all randomised subjects. Only in exceptional cases may subjects be excluded from the FAS. In such cases the reason for exclusion will be justified and documented. Subjects in the FAS will contribute to the evaluation 'as randomised'.
- The safety analysis set will include all subjects receiving at least one dose of randomised treatment. Subjects in the safety analysis set will contribute to the evaluation 'as treated'.

Any subjects or observations excluded from analysis, and the reason for exclusion will be described in the CTR.

2.3 Primary endpoint

The primary endpoint, relative change from baseline in body weight (%) at 52 weeks, will be investigated using the following main analysis to compare between the randomised treatment arms using a multiple imputation (MI) analysis. The main analysis of the primary endpoint will also be referred to as the primary analysis as opposed to the sensitivity analysis of the primary endpoint. Week 52 data from subjects discontinued from trial product that return for visit 22x will be included. In this pattern mixture model approach withdrawn subjects without visit 22x from all treatment arms are assumed to respond as if treated with placebo for the entire trial. Multiple copies (100 copies) of the full dataset will be generated by imputing missing values (body weight (kg) at 52 weeks) based on estimated parameters for the placebo group. This will be done as follows:

- In the first step, 100 copies of the dataset will be generated
- In the second step, an analysis of covariance (ANCOVA) model with region and sex as factors and baseline body weight as covariates is fitted to body weight (kg) at 52 weeks using only placebo subjects with non-missing body weight measurements at baseline and week 52
- In the third step, for each of the 100 copies of the dataset the estimated parameters, and their variances, from this model are used to impute missing values at 52 weeks for subjects in all treatment arms, based on their region, sex and body weight at baseline from the model in step two
- For each of the 100 complete data sets, the change from baseline in body weight (%) at 52 weeks is analysed using an ANCOVA model with treatment (nine classes including one for pooled placebo), region, and sex as factors, and baseline body weight as a covariate
- The estimates and standard deviations for the 100 data sets are pooled into one estimate and associated standard deviation using Rubin's formula:

$$m_{MI} = \frac{1}{100}\sum_{i=1}^{100} m_i, \quad SD_{MI} = \sqrt{\frac{1}{100}\sum_{i=1}^{100} SD_i^2 + \left(1+\frac{1}{100}\right)\left(\frac{1}{100-1}\right)\sum_{i=1}^{100}(m_i - m_{MI})^2},$$

where m_i and SD_i are the estimated means and standard deviations for the 100 copies of the dataset, and m_{MI}, SD_{MI} are the pooled estimates.

- From m_{MI} and SD_{MI}, the 95% confidence interval for the treatment differences and the associated p-value are calculated

If 100 copies are not sufficient to establish stable results, a higher number will be used. The multiple imputations will be generated using Novo Nordisk trial number 95364153 as seed number.

Pairwise treatment differences between semaglutide doses and placebo, liraglutide and placebo, different semaglutide doses, and between semaglutide doses and liraglutide at week 52 will be estimated from the model and 95% confidence intervals will be calculated.

In part A, the comparisons of semaglutide doses vs. placebo will have the family wise type I error protected in the strong sense. This will be achieved by using Dunnett's method in which simultaneous confidence intervals will be calculated. A significance level of 5% will be applied. Further, the focus of this part of the trial is to examine the dose response relationship. In part B, no multiplicity adjustment will be performed.

The dose of semaglutide providing a weight loss corresponding to liraglutide 3.0 mg will be estimated by fitting a linear approximation to the log dose vs. estimated means for the semaglutide doses and compare this to the estimated mean for liraglutide 3.0 mg. This analysis will be based on the estimated means and the covariance matrix for the means obtained from the MI analysis.

The delta method will be used to calculate 95% confidence limits for the estimated dose of semaglutide corresponding to liraglutide 3.0 mg. If a linear approximation does not describe the log(dose)-response relationship well, a different approximation (e.g. a sigmoidal curve) may be investigated.

The MI method does not assume missing at random. It assumes that withdrawn subjects and subjects with missing endpoint at week 52 in the placebo arm have a response similar to the completers in the placebo arm given similar baseline characteristics. In the active treatment arms, the assumption is that withdrawn subjects and subjects with missing endpoint at week 52 behave as if they have been in the placebo arm during entire trial regardless of the time of discontinuation. In this way the assumptions are differential and conservative for estimating the treatment effect.

Based on previous trials in weight management the withdrawal rate from randomised treatment is expected to be up to 40%. Semaglutide treatment has in previous (T2DM) trials been effective with regard to weight loss, and this should reduce the number of withdrawals due to ineffective therapy. Based on previous experience, a higher rate of withdrawal of consent is expected in the placebo group compared to active treatment. This difference may be due to lack of efficacy with placebo treatment. A higher withdrawal rate due to gastrointestinal adverse events is expected in the high

dose semaglutide treatment arms and the liraglutide 3.0 mg arm compared to placebo. Apart from this, missing data due to adverse events (AEs) is expected to be similar across groups. This emphasises the validity of the primary analysis as a conservative analysis of the treatment effect of semaglutide.

The main analysis of the primary endpoint using the pool of the placebo arms will be repeated using the 8 different placebo arms. This supportive analysis includes comparisons between each semaglutide dose/escalation arm, the liraglutide arm and their corresponding placebo arms. As in the primary analysis, week 52 data from subjects discontinued from trial product that return for visit 22x will be included, and withdrawn subjects without visit 22x from all treatment arms are assumed to respond as if treated with placebo. In contrast to the primary analysis, a single imputation (SI) approach instead of MI will be used to handle missing body weight measurements at week 52, where subjects in each active treatment arm are imputed with the average of the corresponding placebo arm.

Pairwise treatment differences between each semaglutide dose/escalation arm, the liraglutide arm and their corresponding placebo arms at week 52 will be estimated from the model and the 95% confidence interval and associated p-value will be calculated.

For the supportive analysis, a dose-response analysis similar as for the primary analysis will be performed.

The sensitivity of the results with regard to different assumptions for missing data from withdrawn subjects will be investigated by plotting weight loss data for withdrawn subjects and compare this with plots of weight loss data for subjects completing 52 weeks of treatment. Further, for subjects that discontinue treatment but provide data at the week 52 visit (22x) plots will compare weight loss at last visit on treatment with weight loss at visit 22x. In addition, several sensitivity analyses will be performed where different assumptions are made with regard to withdrawn subjects.

The following sensitivity analyses will be performed to address the effectiveness estimand:

An ANCOVA model based on a multiple imputation approach as described by McEvoy[2], where missing body weight measurements at week 52 for discontinuing subjects are imputed by sampling from values obtained from retrieved subjects in each randomisation arm and according to the timing (monthly) of last available observation on randomised treatment (LAO-OT). Missing body weight measurements at week 52 for subjects on treatment are imputed by sampling from subjects completing treatment in the relevant randomisation arm. Thus, the imputation model for each randomised treatment arm and timing of LAO-OT is a linear regression of body weight (kg) at week 52 on the factors and covariates mentioned above (except treatment) with no interactions and including LAO-OT of body weight as covariate. If timing by month is too restricted, quarters, half-years, or excluding timing will be used.

The first sensitivity analysis assumes that withdrawn subjects, who return for visit 22x, are representative for subjects, who withdrew from the same randomised treatment but are not retrieved at week 52. Similarly, it is assumed that subjects, who complete the randomised treatment, are representative for subjects, who were randomised to the same treatment arm and have a missing week 52 body weight measurement due to other reasons than withdrawal from randomised treatment.

A weighted ANCOVA model where returning treatment discontinuing subjects are up-weighted relative to their proportion of all withdrawn subjects to account for the subjects not returning for assessments at week 52[2]. Similar subjects with measurements at week 52 on treatment are up-weighted relative to their proportion of all subjects who completed treatment with trial product. The up-weighing is done by randomisation arm and the timing of LAO-OT. Subjects who are missing the body weight measurement at week 52 are assigned a weight of 0.

The second sensitivity analysis is based on the same assumptions as the first sensitivity analysis. In contrast, this weighted analysis can be carried out in situations where it is not possible to realise multiple imputation according to McEvoy[2] due to e.g. too few available body weight measurements from retrieved subjects or the small size of the eight separate placebo arms.

An ANCOVA model is based on a single imputation approach as done by Sacks[3]. Missing body weight measurements at week 52 for withdrawn subjects are imputed using a weight regain rate of 0.3 kg/month after last available observation (LAO). Change from baseline is truncated whenever the extrapolation would lead to a positive weight gain relative to baseline. When a subject's body weight at discontinuation represented a gain in weight relative to baseline, no additional gain will be imputed, but the unfavourable gain is carried forward to week 52. Missing body weight measurements at week 52 for subjects on treatment will be imputed using LAO. The weight regain imputation will be done for all randomised arms. Additionally, a version where only the active arms

use the regain rate while the placebo arms use LAO (corresponding to a weight regain rate of 0 kg/month) will be performed.

The third sensitivity analysis assumes that subjects, who withdraw from randomised active or placebo treatment, lose any treatment effect linearly after discontinuation. In the additional version subjects discontinuing placebo treatment are assumed to experience no change in treatment effect since LAO.

An ANCOVA model is based on a tipping point approach. In a similar manner as above for a range of weight regain rates (starting from 0.1 kg/month and in intervals of 0.1 kg/month) for subjects in the active treatment arms, who discontinued treatment with trial product but were not retrieved at week 52, will be used to define a tipping point in which superiority of semaglutide disappears. In this analysis, subjects on placebo will be imputed by LAO.

The fourth sensitivity analysis assumes that subjects, who withdraw from randomised active treatment, lose any treatment effect linearly after discontinuation and that subjects, who withdraw from randomised placebo treatment, experience no change in treatment effect since LAO.

Dose-response analysis will be repeated based on estimates derived from the first three sensitivity analyses.

The following model will be performed to address the efficacy estimand:

- A mixed model for repeated measurements (MMRM) comparing the change from baseline in body weight (%) at 52 weeks between treatments. All post randomisation measurements at planned visits up to week 52 and obtained before withdrawal from treatment will be included in the model as dependent variables. Treatment, region, and sex will be included as fixed factors, and the baseline body weight will be included as a covariate. All factors and the covariate will be nested under the factor visit. An unstructured covariance matrix will be used to describe the variability for the repeated measurements for a subject. Subjects without post randomisation measurements of weight will be excluded from the analysis.

The MMRM model assumes that withdrawn subjects, had they completed the trial, would not have behaved differently than completing subjects from the same treatment arm with the same baseline characteristic and change in body weight at time of withdrawal.

2.4 Secondary endpoints

For statistical analysis of secondary endpoints, when mentioned below, the analyses methods reproduce the main analysis of the primary endpoint by the same MI approach. The endpoint at baseline will replace body weight at baseline as covariate. The statistical methodology depends on

the type of endpoint, but the model factors and covariates are similar to those specified for the primary analysis.

2.4.1 Efficacy endpoints

Descriptive statistics for efficacy endpoint will be tabulated using on-treatment period defined as the period from first trial product administration to last trial product administration.

Endpoints addressing weight loss

- Proportion (%) of subjects with weight loss of ≥ 5% of baseline body weight at 52 weeks
- Proportion (%) of subjects with weight loss of ≥ 10% of baseline body weight at 52 weeks

These two dichotomous endpoints will be compared between the treatment arms using a MI approach as in the primary analysis based on a logistic regression. The datasets from the primary analysis will be reused for this analysis where imputed values for change in body weight will be used to generate the dichotomous endpoints. Pairwise treatment differences between treatments will be estimated from the model and 95% confidence intervals will be calculated.

Change from baseline to 52 weeks in:

- Body weight (kg)
- Waist circumference (cm)
- Waist to hip circumference ratio (waist (cm)/hip (cm))
- BMI (kg/m^2)

These endpoints will be compared between treatments using the MI approach used for the main analysis of the primary endpoint (with the corresponding baseline value as covariate).

Endpoints addressing glucose metabolism

Change from baseline to 52 weeks in:

- HbA_{1c}
- FPG
- Glycaemic category (normoglycaemia, pre-diabetes, T2DM)

The endpoints HbA1c and FPG will be compared between treatments using the MI approach used for the main analysis of the primary endpoint (with the corresponding baseline value as covariate).

Endpoints addressing cardiovascular risk factors

Change from baseline to 52 weeks in:

- Systolic and diastolic blood pressure
- Lipids (total cholesterol [TC], low density lipoprotein cholesterol [LDL cholesterol], high density lipoprotein cholesterol [HDL cholesterol], very low density lipoprotein cholesterol [VLDL cholesterol], triglycerides [TG], free fatty acids [FFA])
- Cardiovascular biomarker (high sensitivity C reactive protein [hsCRP])

These endpoints will be compared between treatments using the MI approach used for the main analysis of the primary endpoint (with the corresponding baseline value as covariate). For lipids and hsCRP a multiplicative model will be used, i.e. the ratio between post randomisation measurements and baseline will be calculated instead of differences, and both the dependent variable and covariate will be log-transformed. Estimates and CI will be presented as percentage change from baseline.

Endpoints addressing patient reported weight-related quality of life and general health status

Change from baseline to 52 weeks in:

- Impact of Weight on Quality of Life-Lite (IWQoL-Lite) for Clinical Trials: individual items
- Short form-36 (SF-36): Physical and mental component summary scores and scores on the individual sub-domains: Physical functioning, role functioning, bodily pain, general health, vitality, social functioning, role emotional and mental health

These endpoints will be compared between treatments using the MI approach used for the main analysis of the primary endpoint (with the baseline total score as covariate).

Endpoints addressing changes in antihypertensive and lipid-lowering medical treatment

Change from baseline to 52 weeks in:

Proportion of subjects with change in concomitant medications:

- Antihypertensive medications
- Lipid-lowering medications

The endpoint is based on the evaluations made by the investigators and recorded according to the description in the protocol.The proportion of subjects with any change (decrease / no change / increase) in dose and/or drug within the two above classes will be calculated and described.

Endpoint addressing nutritional compliance

Compliance with nutritional counselling will be summarised by week.

Analysis identifying early responders

- Predictability of weight loss of more than 5% at week 52 by early weight loss response (3%, 4%, and 5%) after 12, 16, and 20 weeks

The ability of early weight loss of 3%, 4%, and 5% at week 12, 16, 20 to predict long term weight loss (5% at week 52) will be described by sensitivity, specificity, positive predictive value, and negative predictive value. Receiver operating curve (ROC) will be presented for weight loss at week 12, 16 and 20. The analysis will be repeated seperately for each treatment arm.

2.4.2 Safety endpoints

All adverse events, hypoglycaemic episodes as well as nausea, vomiting, diarrhoea and constipation events will be classified and analysed as 'in-trial' and 'on-treatment'.

In-trial is defined as the observation period from randomisation to last contact with trial site. On-treatment is defined as the observation period from first trial product administration to last trial product administration with a 7 weeks ascertainment window.

The endpoint "Number of adverse events during the trial" will be extensively described using descriptive statistics and listings.

All adverse events will be coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA).

Adverse events will be summarised by system organ class, preferred term, seriousness, severity and relation to trial product.

Endpoint addressing hypoglycaemic episodes:

The endpoint "Number of hypoglycaemic episodes" will be tabulated according to the ADA definition below.

ADA Classification of hypoglycaemic episodes

The hypoglycaemic episodes will be categorised based on the ADA classification[4] of hypoglycaemia:

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

- Asymptomatic hypoglycaemia: An episode not accompanied by typical symptoms of hypoglycaemia, but with a measured plasma glucose concentration ≤ 3.9 mmol/L (70 mg/dL).
- Documented symptomatic hypoglycaemia: An episode during which typical symptoms of hypoglycaemia are accompanied by a measured plasma glucose 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 plasma glucose 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 plasma glucose determination but that was presumably caused by a plasma glucose concentration ≤ 3.9 mmol/L (70 mg/dL).

Given that the trial population does not have T2DM at inclusion, the risk of developing hypoglycaemia is considered low and hence no blood glucose monitoring will be instituted. However, in case of severe hypoglycaemia where third party assistance is needed or in case of a low blood glucose value detected by scheduled blood sampling accompanied by relevant symptoms, the hypoglycaemic episode will qualify for the endpoint analysis.

Treatment-emergent: hypoglycaemic episodes will be defined as treatment-emergent if the onset of the episode occurs within the on-treatment observation period.

Safety endpoints continued

The endpoint "Number of new and ongoing nausea, vomiting, diarrhoea and constipation events by week" will be summarised by week.

Nausea:
- Individual scores of nausea questionnaire
- Severity by numeric rating scale (NRS) score

Nausea questionnaire and NRS score will be summarised by week.

Changes from baseline to 52 weeks in:
- EGG
- Pulse

- Haematology (haemoglobin, haematocrit, thrombocytes, erythrocytes, leucocytes, differential count)
- Biochemistry (creatinine, CPK, urea, albumin, bilirubin [total], ALT, AST, alkaline phosphatase, sodium, potassium, calcium [total], amylase, lipase, calcitonin, TSH)
- Mental health assessed by Columbia Suicidality Severity Rating Scale (C-SSRS) and Patient Health Questionnaire-9 (PHQ-9)

will be summarised and described for each treatment arm. Changes in pulse will be compared between treatments using an MMRM model as described under sensitivity analysis (with the baseline pulse value as covariate) based on the safety analysis set. For amylase and lipase two statistical analyses will be applied, respectively. The relative change (100*value/baseline) will be analysed with an MMRM model as described under sensitivity analysis. The relative change and baseline values will be log-transformed prior to the analysis.

The endpoint "Anti-semaglutide antibodies during and after treatment" will be described by summarising the number and percentage of subjects with antibodies in the different treatment arms. Similarly, subjects with semaglutide antibodies with neutralising effect and with cross-reactivity against endogenous GLP-1 will be described by summaries. The primary endpoint will be summarised by anti-semaglutide antibody status (positive or negative) at follow-up.

2.5 Pharmacokinetic and pharmacodynamic modelling

Exploratory population PK and PK/PD modelling will be used to evaluate the semaglutide dose-exposure, the effects of pre-specified covariates on the exposure and the semaglutide exposure-response on selected efficacy and safety parameters. For the covariate analysis, covariates such as sex, body weight and age will be explored.

The population PK modelling will include data from all randomised subjects that were exposed to semaglutide, excluding data records with concentration values missing or below LLOQ, and data records with incomplete or ambiguous dosing information. Actual time points for dose administration and PK sampling will be used. PK/PD modelling will include data from subjects included in the population PK modelling, with relevant PD assessments available.

Results of the modelling analysis will be presented using criteria which will be pre-specified in a modelling analysis plan that is to be finalised before database lock (DBL). The modelling will be performed by Quantitative Clinical Pharmacology at Novo Nordisk A/S and will be reported separately from the CTR.

3 Changes to the statistical analyses planned in the protocol

3.1 Deviation from analyses as described in the protocol:

- No summary/analysis of IWQoL scores since questionnaire is not validated yet (will be validated after DBL)

The following text:

"*Impact of Weight on Quality of Life-Lite (IWQoL-Lite) for Clinical Trials: Total score and scores on the individual sub-domains*"

has been replaced by:

"*Impact of Weight on Quality of Life-Lite (IWQoL-Lite) for Clinical Trials: Individual items*"

- No logistic regression for amylase/lipase outliers since there are either no or too few observations in category >3x UNR

The following text has been deleted:

"*Subjects having a measurement above >3x UNR anytime during treatment (yes/no) will be analysed using a logistic regression. For the evaluation of the response, all measurements obtained during treatment will be included and these measurements are defined as any scheduled or unscheduled measurements obtained from, but not including, baseline and until, and including, end of treatment. Separate analyses will be made for amylase and lipase. The results will be presented as odds ratios together with the associated 95% confidence intervals.*"

- No ordinal logistic regression for change in antihypertensive or lipid-lowering medications due to too few subjects changing this type of medication during the trial

The following text has been deleted:

"*These endpoints will be compared between treatments using the proportional odds model (or ordinal logistic regression). Only on-treatment data will be considered. Subjects not using drugs within the specified categories will count as no-change. The model factors and covariates will be identical to the main analysis model.*"

- Glycaemic categorisation will not be reported, since the categorisation is not aligned with the usual diagnosis criteria which require repeated testing to confirm the diagnosis and allows for

the diagnosis to be made based on random glucose assessments and/or 2-hour glucose assessments during an oral glucose tolerance test.

The following text:

"*Change from baseline to 52 weeks in:*
- *HbA_{1c}*
- *FPG*
- *Glycaemic category (normoglycaemia, pre-diabetes, T2DM)*"

has been replaced by

"*Change from baseline to 52 weeks in:*
- *HbA_{1c}*
- *FPG*"

Furthermore the following text has been deleted:

"***Glycaemic categories***

Normoglycaemia:
- *FPG < 5.6 mmol/L (100 mg/dL)* ***and*** *HbA_{1c} < 5.7%*

Pre-diabetes:
- *FPG 5.6-6.9 mmol/L (100-125 mg/dL) (both inclusive)* ***and*** *$HbA_{1c} \leq 6.4\%$*

or
- *FPG ≤ 6.9 mmol/L (125 mg/dL)* ***and*** *HbA_{1c} 5.7-6.4% (both inclusive)*

T2DM:
- *FPG ≥ 7.0 mmol/L (126 mg/dL)* ***and/or*** *$HbA_{1c} \geq 6.5\%$"*

and the following text has been deleted:

"*Shift in glycaemic category will be compared between treatments using the proportional odds model (or ordinal logistic regression) and imputing missing measurements in the same fashion as used for the main analysis of the primary endpoint.*"

- Delta method instead of Fieller's method to calculate confidence intervals in dose-response modelling due to problem with the implementation of Fieller's method

The following text:

"*Fieller's method will be used to calculate 95% confidence limits for the estimated dose of semaglutide corresponding to liraglutide 3.0 mg.*"

has been replaced by:

"The delta *method will be used to calculate 95% confidence limits for the estimated dose of semaglutide corresponding to liraglutide 3.0 mg.*"

- Clarification that the on-treatment observation period for safety endpoints includes an ascertainment window, whereas the on-treatment observation for efficacy endpoints does not.

The following text has been added:

"*Descriptive statistics for efficacy endpoint will be tabulated using on-treatment period defined as the period from first trial product administration to last trial product administration.*"

3.2 Additional analyses not described in the protocol:

- Dose-response modelling for GI AEs and body weight responder to support risk-benefit discussion

Analyses of the dose-response for the proportion of subjects meeting three different response criteria:

- Discontinued due to a gastro-intestinal adverse event (on-treatment).
- Having a gastro-intestinal adverse event (on-treatment)
- Meeting the weight loss threshold of 5% at week 52 (in-trial)

Data from the fast escalation regimens will not be used in these analyses. Pooled data from the placebo arms (corresponding to a dose of zero mg) will be included if deemed appropriate.

The model will be developed ad hoc after un-blinding of the data and aim for an appropriate fit of the dose-response relation versus dose or log(dose) with a logistic dose-response relation as the default option. If placebo is included in the analyses and log(dose) provides the best fit the dose of placebo will arbitrarily be set to a small value (e.g. 0.001 mg).

For the on-treatment analyses, the dose-response analysis will be modelled via a logistic regression of the response status (Y/N) per patient with dose level as covariate. The response status of each patient will be based on what is observed while the patient is on treatment, i.e. on whether a patient discontinues treatment due to a gastro-intestinal AE or experiences such an AE while on treatment. Hence, no imputation will be conducted and intercept and regression coefficient will be estimated

from a single analysis of the observed data. The resulting coefficients of interest comprise the intercept and the regression coefficient on dose from the logistic regression model.

The in-trial dose-response analysis of weight loss exceeding 5% will be based on estimated treatment log-odds obtained via logistic regression on datasets generated for the primary analysis using multiple imputations for change in body weight. The coefficients for the logistic dose-response model will be obtained from a linear regression using dose as an independent variable and estimated log-odds as the dependent variable.

If the logistic regression for 5% body weight responders cannot be fitted (e.g. due to too many subjects fulfilling this criteria) the 5% threshold will be replaced by a threshold of 10%.

4 References

1. Phillips A, Haudiquet V. ICH E9 guideline 'Statistical principles for clinical trials': a case study. Stat Med. 2003;22(1):1-11; discussion 3-7.
2. McEvoy BW. Missing data in clinical trials for weight management. J Biopharm Stat. 2016;26(1):30-6.
3. 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.
4. 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.

Statistical documentation – sample size calculation

Assumptions and results

The sample size calculation is based on the primary endpoint; change from baseline in body weight (%) at 52 weeks.

Part A is a dose-finding trial examining five doses of semaglutide and placebo, and an active comparator. The sample size calculation is based on the relative change after 52 weeks treatment in the primary endpoint, body weight. In the sample size calculations, it is presumed that the placebo groups will be pooled in the analysis, assuming no correlation between body weight change after 52 weeks and placebo-injected volume. By trial design we expect to have 135 subjects on placebo and 100 in each active group. In the unfortunate situation that we should not be able to pool all placebo arms but only be able to pool the (semaglutide) placebo arms with dose escalation every fourth week and (liraglutide) placebo, we will still expect to have at least 100 subjects in the relevant pool of placebo subjects. Hence the argumentation below only assumes 100 subjects in the placebo arm in a conservative manner and for simplicity. However, at time of analysis we plan to include all subjects on placebo as one group in the analysis as described above, if possible.

The following assume 100 subjects randomised in a balanced manner to receive each active treatment and being conservative at least 100 evaluable subjects in the placebo arm. In trial NN8022-1839, in obese subjects without T2DM a standard deviation of just below 7% was seen for observed weight loss (in %) in the liraglutide 3.0 mg arm. A conservative estimate of the dropout rate is 40%. A standard deviation of 7% and a sample size of 100 in each treatment arm will allow the 95% confidence interval for the estimated difference between two semaglutide doses, with 90% probability, to be contained within ±2.5% of the estimate, which is considered to be a sufficient precision for determining which doses to use for the continued development of semaglutide in the weight management indication.

For the primary endpoint change in body weight after 52 weeks of treatment, a difference (semaglutide minus placebo) of 9.5% is expected for completers (12% in the optimal dose group for semaglutide vs. 2.5% in placebo). For the withdrawn subjects, who are anticipated to constitute up to 40% of the total trial population, the treatment difference (semaglutide minus placebo) is assumed to be 0% giving an overall expected treatment difference of 5.7% (8.2% in semaglutide vs. 2.5% in placebo). The standard deviation will also be increased using the MI approach. The standard deviation in the final data is assumed to be up to 8.4% (8.4% in semaglutide vs. 7% in placebo). A standard deviation of up to 8.4% together with an expected difference of 5.7% results in a power of more than 99%, which is not corrected for multiple comparisons between different semaglutide arms and placebo.

For part B, in total two times 100 subjects will be randomised to the dose escalation every second week. Combined with placebo and the two doses corresponding to the 'every second week' arms (i.e. 0.3 mg/day and 0.4 mg/day) we will have more than five hundred subjects for the inference of the dose escalation finding part of the trial. In trial NN8022-1839, nausea, vomiting, and constipation were the most common gastrointestinal AEs with incidences between 15% and 40% of all subjects on liraglutide 3.0 mg. These AEs are not expected to be less frequent with semaglutide. This part of the trial is exploratory in nature and is intended to evaluate the overall safety profile with respect to the different types of events and when they occur compared to the escalation steps and with a view towards the efficacy response as well. With the given number of subjects, we have a reasonable sample size to detect marked clinical relevant difference between the arms.

SAS program code

Path to program:
P:\nn9536\nn9536-4153\current\stats\document\0100_protocol\0110_sample_size

```
/*************************************************************************************************
 Project    :  Semaglutide obesity
 Study      :  NN9536-4153
 Name       :  powerph2_NN9536-4153.sas
 Description:  Sample size calculation for the NN9536-4153 trial phase II
               trial for semaglutide obesity
-------------------------------------------------------------------------------------------------
               Init
 Programmer  : HHg Apr2014
 Text        :
 UPdate log  : LKOO 2014-12-08

*************************************************************************************************/

/*
dm log 'clear' wpgm;
dm output 'clear' wpgm;

options symbolgen;
ods output file = "";
*/

/*********************************************************************
* Power for phase 2 trial dose selection of semaglutide for obesity
*********************************************************************/

/*************************************************************************************************
** Standard assumptions:
** A standard deviation of 7% (in completers/placebo)
** a sample size of 100 in each treatment arm
** 40% WD,
** an expected effect of 10% (best Sema) using LOCF [MMLA want us to keep 10, even TPP says
12]
** (8% on lira using LOCF),
** 2.5 in placebo  (using 3.0 for placebo completers, does not make a noticable difference)
** Missing data: MI using copy placebo/reference
**
*************************************************************************************************/

%let sd_pbo=7;
%let meanP=2.5;
%let WD=0.4;
%let meanS_locf=10;
%let meanL_locf=8;

/*
first calculations of emplied completer effects and recalculate derived means with MI and SD
with MI
(usually we assume that LOCF give ½ETD in WD subjects)

This gives the following mean effects in completers
sema_All = (1-WD)*Sema_comp + WD * (placebo + ½*(sema_comp-placebo) = 10
The implied effect in Sema_completers is:
Sema_comp = (10-½*WD*placebo)/(1-½*WD)
*/

%let meanSC=%sysevalf( (&meanS_locf-0.5*&WD*&meanP)/(1-0.5*&WD)  );
%let meanLC=%sysevalf( (&meanL_locf-0.5*&WD*&meanP)/(1-0.5*&WD)  );

%put Sema completer: &meanSC;
%put Lira completer: &meanLC;

/*************************************************************************************************
** Thus we are assuming (approximately) a
** 12% WL in Sema completers on optimal dose
** 9.5% WL in Lira completers (similar to 1839 data)
*************************************************************************************************/
```

```
%let meanSC=12;
%let meanLC=9.5;

/*
With MI the ALL WL with 40 WD rate will then be:
*/
%let meanS=%sysevalf( (1-&WD)*&meanSC + &WD*&meanP   );
%let meanL=%sysevalf( (1-&WD)*&meanLC + &WD*&meanP   );
%put Placebo: &meanP. ;
%put Sema: &meanS;
%put Lira: &meanL;

/*****************************************************************************************
****************************************
** Thus we expect LSmean weight loss (40% WD)
** 8.2% WL in Sema all on optimal dose
** 6.7% WL in Lira all
*****************************************************************************************
*************************************/

* regarding SD to use: *;

%let ETDc=%sysevalf(&meanSC-&meanP);

data b;
  sd_s= sqrt(&sd_pbo*&sd_pbo + &WD*(1-&WD)*&ETDc*&ETDc);
call symput ('sd_s', sd_s);
  sd_l= sqrt(&sd_pbo*&sd_pbo + &WD*(1-&WD)*(&meanLC-&meanP)*(&meanLC-&meanP));
call symput ('sd_l', sd_l);
run;

title 'SD in Sema and Lira for full population WD=40pct';
proc print data=b;
run;

/*****************************************************************************************
** Thus we expect sd in Sema arms (40% WD) is
** When the missing data procedure is MI (jump to reference/copy placebo) the SD in the active
group become sqrt ( sd^2 +p(1-p)*ETD^2 )
** sd_s=8.4
** sd_s=7.8
*****************************************************************************************/

************************ CALCULATIONS ***************************************************;

/*****************************************************************************************
* with the above assumptions
* will allow the 95% confidence interval
*  for the estimated difference between two arms, with 90% probability, to be contained within
±2.5%;
*****************************************************************************************/

proc power;
title 'Estimating the probability for the width of the confidence interval for the estimated
difference';
  twosamplemeans CI = diff
  alpha = 0.05
  stddev =&sd s
  groupweights = 1 | 1
  halfwidth = 2.5
  probwidth = .
  ntotal = 200;
run;

/*************************************************************************
  Marginal power for primary endpoint: relative change in body weight
  4W/QD arms WD rate 40%
  We have high power: >99.9%
*************************************************************************/

%let diff=%sysevalf(&meanS - &meanP ) ;


proc power;
title "Power for semaglutide vs. placebo - relative change in body weight";
            twosamplemeans
            alpha=0.05
```

```
            dist=normal
            test=diff_satt
            gmeans =&meanP | &meanS
            nulldiff=0
            sides=1
            GROUPSTDDEVS = (&sd_pbo  &sd_s)
            power=.
            ntotal = 200;
run;

/*************************************************************************
  Marginal power for primary endpoint: relative change in body weight
  2W/QD arms WD rate 50%
  We still have high power: 99.6%
*************************************************************************/

%let WD2=0.5;
data c;
  sd_s2= sqrt(&sd_pbo*&sd_pbo + &WD2*(1-&WD2)*&ETDc*&ETDc);
  call symput ('sd s2', sd s2);
  meanS2=%sysevalf( (1-&WD2)*&meanSC + &WD2*&meanP   );
  call symput ('meanS2', meanS2);
run;

proc print data = c;
run;

proc power;
title "Power for semaglutide vs. placebo - relative change in body weight (WD=50%)";
            twosamplemeans
            alpha=0.05
            dist=normal
            test=diff_satt
            gmeans =&meanP | &meanS2
            nulldiff=0
            sides=1
            GROUPSTDDEVS = &sd_pbo | &sd_s2
            power=.
            ntotal = 200;
run;


/*************************************************************************
  Marginal powers for secondary endpoints: proportion of subjects loosing
  5% respectively 10% of initial body weight at end of treatment
*************************************************************************/

/* relative weight loss are expected to follow a normal distribution, for semaglutide
   we expect a mean of 12% in completers and a standard deviation of 7% (in completers).

   This leads to an expected proportion in completers of subjects loosing 10% about 60%
   The expected proportion loosing 5% or more is about 80%; both calculated below
   For Placebo the numbers are set to 35% and 10% (1839)

   With MI and WD=40%, gives proportion of 5% responder: 64%, and proportion of 10% responder:
40%

   For Lira the corresponding numberes are: proportion of 5% responder: 58%, and proportion of
10% responder: 32%
 */

data prob;
  prob_placebo05 = 0.35 /*1 - cdf('NORMAL', 5, 2.5, 7) */ ;
  prob sema05 c  = 1 - cdf('NORMAL', 5, &meanSC, &sd pbo) ;
  prob_sema05    = (1-&WD)*(prob_sema05_c) + &WD*(prob_placebo05);
  prob_placebo10 = 0.1 /* 1 - cdf('NORMAL', 10, 2.5, 7) */ ;
  prob_sema10_c  = 1 - cdf('NORMAL', 10, &meanSC, &sd_pbo);
  prob_sema10    = (1-&WD)*(prob_sema10_c) + &WD*(prob_placebo10);
  prob_lira05_c  = 1 - cdf('NORMAL', 5, &meanLC, &sd_pbo);
  prob lira05    = (1-&WD)*(prob lira05 c) + &WD*(prob placebo05);
  prob_lira10_c  = 1 - cdf('NORMAL', 10, &meanLC, &sd_pbo);
  prob_lira10    = (1-&WD)*(prob_lira10_c) + &WD*(prob_placebo10);
  call symput('placebo_5pct', prob_placebo05);
  call symput ('sema_5pct', prob_sema05);
  call symput('placebo_10pct', prob_placebo10);
  call symput ('sema_10pct', prob_sema10);
run;
```

```
title 'Estimating the proportion of subjects loosing 5%/10% or more based on the assumption of
normal distribution with mean 12 (sema, completer) respectively 2.5 (placebo) and SD of 7% and
WD=40pct';
proc print data=prob;
run;

%macro chi(sema= , placebo= , pct= , title= );
* chi square test;
title &title.;
proc power;
  twosamplefreq test=pchi
  grouppproportions = (&sema. &placebo.)
  nullproportiondiff = 0
  power = .
  npergroup = 100;
  ods output output=Power;
run;

data power&pct.;
  set power;
  sema_prob=Proportion1;
  placebo_prob=Proportion2;
  drop error info Proportion1 Proportion2 index;
run;

proc print data = power&pct.;
run;
%mend chi;

%chi(sema=&sema_5pct., placebo=&placebo_5pct., pct=5, title="Power for semaglutide vs. placebo
- 5pct Responders");
%chi(sema=&sema_10pct., placebo=&placebo_10pct., pct=10, title="Power for semaglutide vs.
placebo - 10pct Respondres");

* Power for 5% responder: 99%
* Power for 10% responder: >99.9%
*;

/*************************************************************************
  power calculation for comparison between liraglutide 3.0 and semaglutide
*************************************************************************/

proc power;
title "Power for semaglutide vs. Lira 3.0 - relative change in body weight";
            twosamplemeans
            alpha=0.05
            dist=normal
            test=diff_SATT
            gmeans =&meanL | &meanS
            nulldiff=0
            sides=1
            GROUPSTDDEVS = &sd_l | &sd_s
            power=.
            ntotal = 200;
run;

proc power;
title "Sample Size for semaglutide vs. Lira 3.0 - relative change in body weight";
            twosamplemeans
            alpha=0.05
            dist=normal
            test=diff_SATT
            gmeans =&meanL | &meanS
            nulldiff=0
            sides=1
            GROUPSTDDEVS = &sd_l | &sd_s
            power=.90
            ntotal =.;
run;

* Power for Sema vs Lira 3.0:  36.7%
* samplesize for Sema vs Lira 3.0: 500 ptt pr arm
*;
```

```
/*************************************************************************
  power calculation for comparison GI side effect between two semaglutide arms
  (50% chosen as a worst case number)
*************************************************************************/


title 'Power for GI side effects between arms. Conservative one arm have 50pct event rate';
proc power;
  twosamplefreq test=pchi
  grouppproportions = (.5 .305)
  nullproportiondiff = 0
  power = .
  npergroup = 100;
  ods output output=Power;
run;

* so 80% power to show a difference from 50% to 30% *;

proc power;
  twosamplefreq test=pchi
  grouppproportions = (.5 .28)
  nullproportiondiff = 0
  power = .
  npergroup = 100;
  ods output output=Power;
run;

* so 90% power to show a difference from 50% to 28% *;

/*************************************************************************************************
**
  power calculation for comparison dropu out due to GI side effect between two semaglutide
arms
  (15% chosen as a relistic conservative number)
*************************************************************************************************
**/

title 'Power for GI side effects between arms. Conservative one arm have 15pct event rate';
proc power;
  twosamplefreq test=pchi
  grouppproportions = (.15 .035)
  nullproportiondiff = 0
  power = .
  npergroup = 100;
  ods output output=Power;
run;

* so 80% power to show a difference from 15% to 3.5% *;

* Conclussion: WRT nausea we will need to look at the overall picture and do biomodeling and
common sense
*;

/*************************************************************************
** update of TPP numbers assuming 12% with LOCF
*************************************************************************/

* caluculation of MI sema WL =X if LOCF WL=12%
*;

%let meanS_locf2=12;
%let meanSC2=%sysevalf( (&meanS_locf2-0.5*&WD*&meanP)/(1-0.5*&WD)  );
%put Sema completer: &meanSC2;
%let meanSC=14.4;
%let meanS2=%sysevalf( (1-&WD)*&meanSC2 + &WD*&meanP   );
%put Sema: &meanS2;

* thus 12% with LOCF correspond to 9.6 with MI and 40% WD
* and to 14.4% in completers
*;

data prob2;
  prob_placebo05 = 0.35 /*1 - cdf('NORMAL', 5, 2.5, 7) */ ;
  prob_placebo10 = 0.1 /* 1 - cdf('NORMAL', 10, 2.5, 7) */ ;
  prob_sema05_c  = 1 - cdf('NORMAL', 5, &meanSC2, &sd_pbo) ;
  prob_sema05    = (1-&WD)*(prob_sema05_c) + &WD*(prob_placebo05);
  prob_sema10_c  = 1 - cdf('NORMAL', 10, &meanSC2, &sd_pbo);
```

```
  prob_sema10     = (1-&WD)*(prob_sema10_c) + &WD*(prob_placebo10);
run;

title 'Estimating the proportion of subjects loosing 5%/10% or more based on the assumption of
normal distribution with mean 14.4 (sema, completer) respectively 2.5 (placebo) and SD of 7%
and WD=40pct';
proc print data=prob2;
run;


* thus 12% with LOCF correspond to the following with MI and 40% WD
* 5% responders: 68%
* 10% responders: 48%
*;
```

SAS program output

SD in Sema and Lira for full population WD=40pct

Obs	sd_s	sd_l
1	8.40595	7.79487

Estimating the probability for the width of the confidence interval for the estimated difference

The POWER Procedure
Confidence Interval for Mean Difference

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Alpha	0.05
CI Half-Width	2.5
Standard Deviation	8.40595
Group 1 Weight	1
Group 2 Weight	1
Total Sample Size	200
Number of Sides	2
Prob Type	Conditional

Computed Prob(Width)
Prob(Width)
0.908

Power for semaglutide vs. placebo - relative change in body weight

The POWER Procedure
Two-Sample t Test for Mean Difference with Unequal Variances

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Number of Sides	1
Null Difference	0
Nominal Alpha	0.05
Group 1 Mean	2.5
Group 2 Mean	8.2
Group 1 Standard Deviation	7
Group 2 Standard Deviation	8.40595
Total Sample Size	200
Group 1 Weight	1
Group 2 Weight	1

Computed Power	
Actual Alpha	Power
0.05	>.999

Power for semaglutide vs. placebo - relative change in body weight

Obs	sd_s2	meanS2
1	8.45946	7.25

Power for semaglutide vs. placebo - relative change in body weight (WD=50%)

The POWER Procedure
Two-Sample t Test for Mean Difference with Unequal Variances

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Number of Sides	1
Null Difference	0
Nominal Alpha	0.05
Group 1 Mean	2.5
Group 2 Mean	7.25
Group 1 Standard Deviation	7
Group 2 Standard Deviation	8.459462
Total Sample Size	200
Group 1 Weight	1
Group 2 Weight	1

Computed Power	
Actual Alpha	Power
0.05	0.996

Estimating the proportion of subjects loosing 5%/10% or more based on the assumption of normal distribution with mean 12 (sema, completer) respectively 2.5 (placebo) and SD of 7% and WD=40pct

Obs	prob_placebo05	prob_sema05_c	prob_sema05	prob_placebo10	prob_sema10_c	prob_sema10	prob_lira05_c	prob_lira05	prob_lira10_c	prob_lira10
1	0.35	0.84134	0.64481	0.1	0.61245	0.40747	0.73984	0.58390	0.47153	0.32292

Power for semaglutide vs. placebo - 5pct Responders

The POWER Procedure
Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements	
Distribution	Asymptotic normal
Method	Normal approximation
Null Proportion Difference	0
Group 1 Proportion	0.644807
Group 2 Proportion	0.35
Sample Size per Group	100
Number of Sides	2
Alpha	0.05

Computed Power
Power
0.990

Power for semaglutide vs. placebo - 5pct Responders

Obs	Analysis	NullProportionDiff	NPerGroup	Sides	Alpha	Power	sema_prob	placebo_prob
1	TwoSampleFreq	0	1002		0.05	0.990	0.64481	0.35

Power for semaglutide vs. placebo - 10pct Respondres

The POWER Procedure
Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements	
Distribution	Asymptotic normal
Method	Normal approximation
Null Proportion Difference	0
Group 1 Proportion	0.407471
Group 2 Proportion	0.1
Sample Size per Group	100
Number of Sides	2
Alpha	0.05

Computed Power
Power
>.999

Power for semaglutide vs. placebo - 10pct Respondres

Obs	Analysis	NullProportionDiff	NPerGroup	Sides	Alpha	Power	sema_prob	placebo_prob
1	TwoSampleFreq	0	1002	2	0.05	0.999	0.40747	0.1

Power for semaglutide vs. Lira 3.0 - relative change in body weight

The POWER Procedure
Two-Sample t Test for Mean Difference with Unequal Variances

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Number of Sides	1
Null Difference	0
Nominal Alpha	0.05
Group 1 Mean	6.7
Group 2 Mean	8.2
Group 1 Standard Deviation	7.79487
Group 2 Standard Deviation	8.40595
Total Sample Size	200
Group 1 Weight	1
Group 2 Weight	1

Computed Power	
Actual Alpha	Power
0.05	0.367

Sample Size for semaglutide vs. Lira 3.0 - relative change in body weight

The POWER Procedure
Two-Sample t Test for Mean Difference with Unequal Variances

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Number of Sides	1
Null Difference	0
Nominal Alpha	0.05
Group 1 Mean	6.7
Group 2 Mean	8.2
Group 1 Standard Deviation	7.79487
Group 2 Standard Deviation	8.40595
Nominal Power	0.9
Group 1 Weight	1
Group 2 Weight	1

Computed N Total		
Actual Alpha	Actual Power	N Total
0.05	0.900	1002

Power for GI side effects between arms. Conservative one arm have 50pct event rate

The POWER Procedure
Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements	
Distribution	Asymptotic normal
Method	Normal approximation
Null Proportion Difference	0
Group 1 Proportion	0.5
Group 2 Proportion	0.305
Sample Size per Group	100
Number of Sides	2
Alpha	0.05

Computed Power
Power
0.808

Power for GI side effects between arms. Conservative one arm have 50pct event rate

The POWER Procedure
Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements	
Distribution	Asymptotic normal
Method	Normal approximation
Null Proportion Difference	0
Group 1 Proportion	0.5
Group 2 Proportion	0.28
Sample Size per Group	100
Number of Sides	2
Alpha	0.05

Computed Power
Power
0.897

Power for GI side effects between arms. Conservative one arm have 15pct event rate

The POWER Procedure
Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements	
Distribution	Asymptotic normal
Method	Normal approximation
Null Proportion Difference	0
Group 1 Proportion	0.15
Group 2 Proportion	0.035
Sample Size per Group	100
Number of Sides	2
Alpha	0.05

Computed Power
Power
0.806

Estimating the proportion of subjects loosing 5%/10% or more based on the assumption of normal distribution with mean 14.4 (sema, completer) respectively 2.5 (placebo) and SD of 7% and WD=40pct

Obs	prob_placebo05	prob_placebo10	prob_sema05_c	prob_sema05	prob_sema10_c	prob_sema10
1	0.35	0.1	0.90976	0.68586	0.73401	0.48041

MedDRA searches in NN9536-4153

Author:
Nina Engberg Jungdal,
Safety Surveillance GLP-1 & Obesity

Terms used for predefined MedDRA search for adverse events of acute renal failure, allergic reactions, cardiac arrhythmia, gallbladder disorders, gastrointestinal adverse events, hepatic adverse events, injection site reactions, medication errors, neoplasms, rare events, transmission of an infectious agent via trial product and thyroid disorders

Adverse event search	Search criteria
Acute renal failure	SMQ 'Acute renal failure' (narrow scope)
Allergic reactions	NNMQ 'Allergic reactions' including: SMQ Anaphylactic reaction (narrow scope) SMQ Angioedema (narrow scope) SMQ Severe cutaneous adverse reactions (narrow scope) SMQ Anaphylactic/anaphylactoid shock conditions (narrow scope) SMQ Hypersensitivity (narrow scope)
Cardiac arrhythmia	SMQ 'Cardiac arrhythmias' including: SMQ Arrhythmia related investigations, signs and symptoms SMQ Cardiac arrhythmia terms (incl bradyarrhythmias and tachyarrhythmias) SMQ Congenital and neonatal arrhythmias
Gallbladder disorders	NNMQ 'Gallbladder related disorders' including: SMQ Biliary tract disorders SMQ Biliary system related investigations, signs and symptoms SMQ Gallbladder related disorders SMQ Gallstone related disorders SMQ Infectious biliary disorders
Gastrointestinal adverse events	SOC 'Gastrointestinal disorders'

SMQ: standard MedDRA query, NNMQ: Novo Nordisk MedDRA query, SOC: system organ class

Terms used for predefined MedDRA search for adverse events of acute renal failure, allergic reactions, cardiac arrhythmia, gallbladder disorders, gastrointestinal adverse events, hepatic adverse events, injection site reactions, medication errors, neoplasms, rare events, transmission of an infectious agent via trial product and thyroid disorders

Adverse event search	Search criteria
Hepatic adverse events	SMQ 'Drug related hepatic disorders – comprehensive search' including: SMQ Cholestasis and jaundice of hepatic origin SMQ Drug related hepatic disorders – severe events only SMQ Liver related investigations, signs and symptoms SMQ Liver-related coagulation and bleeding disturbances
Injection site reactions	NNMQ 'Injection site reactions' including: HLT Administration site reactions HLT Application and instillation site reactions HLT Infusion site reactions HLT Injection site reactions
Medication error	NNMQ 'Medication error' including: HLGT Medication errors HLGT Device issues HLGT Product Quality Issue HLT Complications associated with device NEC HLT Overdoses NEC HLT Underdoses NEC PT Contraindicated drug administered PT Drug administered to patient of inappropriate age
Neoplasms	NNMQ 'Neoplasm' including: SOC Neoplasms benign, malignant and unspecified (including cysts and polyps) (primary and secondary terms) SMQ Biliary neoplasms SMQ Breast neoplasms, malignant and unspecified

SMQ: standard MedDRA query, NNMQ: Novo Nordisk MedDRA query, HLT: high level term, HLGT: high level group term, NEC: not elsewhere classified, PT: preferred term, SOC: system organ class

Terms used for predefined MedDRA search for adverse events of acute renal failure, allergic reactions, cardiac arrhythmia, gallbladder disorders, gastrointestinal adverse events, hepatic adverse events, injection site reactions, medication errors, neoplasms, rare events, transmission of an infectious agent via trial product and thyroid disorders

Adverse event search	Search criteria
	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
Rare events	NNMQ 'Rare' excluding all search strings included in any of the other medical areas NNMQ 'Rare' includes: SMQ Acute renal failure (narrow scope) SMQ Agranulocytosis (narrow scope) SMQ Anaphylactic reaction (narrow scope) SMQ Cardiac arrhythmias SMQ Cholestasis and jaundice of hepatic origin SMQ Guillain-Barre syndrome (narrow scope) SMQ Haematopoietic cytopenias affecting more than one type of blood cell SMQ Haematopoietic leukopenia SMQ Haematopoietic thrombocytopenia (narrow scope) SMQ Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (narrow scope) SMQ Hepatitis, non-infectious SMQ Severe cutaneous adverse reactions (narrow scope) SMQ Interstitial lung disease (narrow scope)

SMQ: standard MedDRA query, NNMQ: Novo Nordisk MedDRA query

Terms used for predefined MedDRA search for adverse events of acute renal failure, allergic reactions, cardiac arrhythmia, gallbladder disorders, gastrointestinal adverse events, hepatic adverse events, injection site reactions, medication errors, neoplasms, rare events, transmission of infectious agent via trial product and thyroid disorders

Adverse event search	Search criteria
	SMQ Neuroleptic malignant syndrome (narrow scope) SMQ Pseudomembranous colitis (narrow scope) SMQ Retroperitoneal fibrosis (narrow scope) SMQ Acute Pancreatitis (narrow scope) SOC Congenital, familial and genetic disorders (primary and secondary routed PTs) HLT Acute and chronic pancreatitis (primary and secondary routed 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) PT Disseminated intravascular coagulation PT Multi-organ failure
Suspected transmission of an infectious agent	NNMQ 'Suspected transmission of an infectious agent' including: HLT Infectious disorders carrier (only primary terms) HLT Infectious transmissions (only primary terms)
Thyroid disorders	NNMQ 'Thyroid disorders' including: SMQ Hyperthyroidism SMQ Hypothyroidism HLGT Thyroid gland disorders (incl both primary and secondary linked PTs) PT Thyroglossal cyst PT Blood calcitonin abnormal PT Blood calcitonin increased PT Ectopic calcitonin production PT Hypercalcitoninaemia PT Blood calcitonin PT Calcitonin secretion disorder

SMQ: standard MedDRA query, HLT: high level term, NEC: not elsewhere classified, PT: preferred term, NNMQ: Novo Nordisk MedDRA query, HLGT: high level group term,