Cover Page for Statistical Analysis Plan

Sponsor name:	Novo Nordisk A/S
NCT number	NCT03596450
Sponsor trial ID:	NN9535-4416
Official title of study:	Long Term Comparative Effectiveness of Once Weekly Semaglutide Versus Standard of Care in a Real World Adult US Population With Type 2 Diabetes - a Randomized Pragmatic Trial
Document date*	26 October 2021

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

16.1.9 Documentation of statistical methods

16.1.9.1 Statistical Analysis Plan

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

.

Statistical Analysis Plan for Long term comparative effectiveness of once weekly semaglutide versus standard of care in a real world adult US population with type 2 diabetes - a randomized pragmatic clinical trial

Protocol No. NN9535-4416

Novo Nordisk

July 5, 2018

Version 1.0

[redacted]

[redacted]

Signature Page for Statistical Analysis Plan
Protocol No: NN9535-4416
05-JUL-2018

[redacted] Approval:

Name: [redacted]
Title: [redacted] Scientist, Biostatistics

Signature: [redacted]

Date: [redacted]

Name: [redacted]
Title: [redacted] Scientist

Signature: [redacted]

Date: [redacted]

Novo Nordisk Approval:

Name: [redacted]
Title: [redacted]

Signature: [redacted]

Date: [redacted]

Name: [redacted]
Title: [redacted]

Signature: [redacted]

Date: [redacted]

[redacted]

TABLE OF CONTENTS

1.0 INTRODUCTION

This document describes the Statistical Analysis Plan (SAP) for Protocol Number NN9535-4416, Long term comparative effectiveness of once weekly semaglutide versus standard of care in a real world adult US population with type 2 diabetes – a randomized pragmatic clinical trial. This is a 2-year, multi-center, randomized, open label, parallel group, active comparator pragmatic clinical trial (PCT) comparing semaglutide versus standard of care (SOC) when added to metformin monotherapy as treatment intensification among adult type 2 diabetes mellitus (T2DM) patients in the course of routine clinical practice.

This SAP is based on the most recent version of the protocol: Version 1.0, March 13, 2018.

Protocol History:

Number	Version	Date
1.0	Original Protocol	March 13, 2018

2.0 STUDY OBJECTIVES

2.1 Primary Objective

The primary objective is to demonstrate superior long term effects of treatment with semaglutide compared to SOC each added to metformin on glycemic control when used as intensification in routine clinical practice in adult patients with T2DM.

2.2 Secondary Objectives

The secondary objectives are to compare the long term effect of semaglutide versus SOC each added to metformin and used as intensification in routine clinical practice in adult patients with T2DM with regards to:

- Weight loss
- Patient reported outcomes (PROs) and clinician reported outcomes (ClinROs)
- Hypoglycemia
- Healthcare Resource Utilization (HCRU)
- Adherence and persistence to treatment

2.3 Estimands

The estimand informs choices about data foundation and statistical analysis including possible imputation of missing data, hereby ensuring that randomization is preserved as a sound basis for statistical inference; i.e. estimation of effect size, associated uncertainty and statistical testing. Two estimands, each targeting a unique research question, have been defined to adequately describe the effect of semaglutide as compared to SOC (excluding semaglutide).

2.3.1 Primary Estimand

The primary estimand for all objectives will be the "intention-to-treat" (ITT) estimand evaluating the effectiveness of randomized treatment intervention irrespective of adherence to this randomized intervention or changes to other antidiabetic medication.

2.3.2 Secondary Estimand

The secondary estimand for all objectives with the exception of the adherence and persistence to treatment objective, is the "if all patients had adhered" estimand. This estimand evaluates the effect of randomized treatment intervention for all randomized patients if all patients had adhered to randomized treatment, regardless of changes to other antidiabetic medication.

3.0 STATISTICAL METHODOLOGY

3.1 General Considerations

Statistical analyses will be performed using SAS® version 9.4 or higher computer software. Analyses will utilize prospectively collected data from study sites as well as secondary data collection utilizing administrative claims data from the HealthCore Integrated Research Database (HIRD®). Data from both sources will be integrated into one dataset for analysis.

Patient flow, patient characteristics, treatments, and outcomes will be tabulated and summarized with descriptive statistics based on observed values only. All descriptive data summaries will include means, medians, standard deviation (SD), and ranges for continuous variables and absolute/relative frequencies for categorical data. In addition, for each variable the number (count) of values that are missing will also be reported for dedicated study visit data. Missing data are data that are planned and can be collected but are absent. Statistics will be summarized for the study population overall and by treatment arm. Two sets of descriptive summaries will be produced: one based on the primary estimand (ITT) and one based on the secondary estimand ("if all patients had adhered").

Raw data (i.e., minimum and maximum values presented for range in continuous variables) will be reported out to the precision with which it was collected. Means will be reported to 1 decimal place more than the raw data. SD will be reported to 1 decimal place more than the mean. Percentages will be reported to 1 decimal place. Trailing zeros will be presented to maintain a consistent level of precision, e.g. 2.0 rather than 2.

Inferential tests will be performed at the 5% level of significance (two-sided). All p-values will be rounded to 3 decimal places. If a rounded p-value is 0.000 (i.e., the actual p-value is less than 0.0005), then the p-value will be presented as 'p< 0.001.'

3.2 Study Populations

3.2.1 Analysis Population

The following analysis set will be defined:

Full analysis set (FAS): Includes all randomized patients analyzed according to the treatment group to which they were assigned at randomization.

3.2.2 Safety Population

The safety population includes all randomized patients initiated on study drug.

3.3 Evaluation Schedule and Definitions

Dedicated study visits are at randomization, at year 1, and at year 2. The study will also capture data collected at the sites during routine care visits, i.e., office visits and other patient contacts that occur as part of routine clinical practice. Routine care visits will occur per study physician's routine clinical practice, therefore the number of visits and data available may differ from site to site and patient to patient.

For analysis, baseline includes assessments conducted and patient data collected at or prior to randomization. Baseline is defined as ≤ 90 days prior to randomization visit (week 0) for HbA1c. For secondary endpoint assessments, baseline is defined as ≤ 4 weeks prior to the randomization visit (week 0). All baseline values will be derived from the last assessment taken prior to or at the randomization visit.

The dedicated year 1 visit is at 52 weeks (±6 weeks) after randomization. Year 1 endpoint data will be derived from the dedicated year 1 study visit if the dedicated study visit is ±10 weeks of 52 weeks post-randomization. If year 1 endpoint data is missing or outside of 52 ±10 weeks, then the routine care data closest to 52 weeks post-randomization ±10 weeks will be used. If no routine care data are available 52 weeks post-randomization ±10 weeks, then year 1 endpoint data will be considered missing and imputed as described in section 4.5, if applicable.

The dedicated year 2/EOS visit is 104 weeks (±6 weeks) after randomization. Year 2 endpoint data will be derived from the dedicated year 1 study visit if the dedicated study visit is ±10 weeks of 104 weeks post-randomization. If year 2 endpoint data is missing or outside of 104 ±10 weeks, then the routine care data closest to 104 weeks post-randomization ±10 weeks will be used. If no routine care data are available 104 weeks post-randomization ±10 weeks, then year 2 endpoint data will be considered missing and imputed as described in section 4.5, if applicable.

3.4 Confirmatory Endpoints and Hypotheses

The primary endpoint is HbA1c <7.0% (53 mmol/mol) at year 1 (yes/no).

Confirmatory secondary endpoints include:

- Change in HbA1c (%-point) from baseline to year 1
- HbA1c <7.0% (53 mmol/mol) at year 2 (yes/no)
- Change in HbA1c (%-point) from baseline to year 2

This study is designed to have 90% power to jointly confirm superiority of the primary endpoint and the above three confirmatory secondary endpoints based on primary estimand analyses. The primary and confirmatory secondary endpoints will all be tested for superiority under multiplicity control via a hierarchical testing scheme according to the order given in the above bullet list. Confirmatory testing will only be performed for the primary estimand with the secondary estimand being supportive. The testing procedure will be stopped the first time an analysis fails to confirm superiority of the endpoint in question using a two-sided significance level of 5%.

Superiority on HbA1c <7.0% at year 1 and at year 2 will be evaluated with respect to the odds ratio (OR) (odds semaglutide / odds SOC):

H_0: OR≤1 against H_a: OR>1

Superiority on change in HbA1c (year 1 – baseline; year 2 – baseline) will be evaluated with respect to the mean treatment difference (TD) (semaglutide – SOC):

H_0: TD≥0 against H_a: TD<0

Statistical analysis of all other endpoints will be unadjusted for multiplicity.

3.5 Sample Size Estimation

3.5.1 Power and Sample Size for Primary Objective

Assumptions for the sample size were based on input from the HIRD claims/laboratory results database. In line with the primary ITT estimand, the assumptions for the proportion of patients with HbA1c < 7.0% at year 1 and for the change from baseline in HbA1c are based on the claims/laboratory results data within the Anthem population for all patients initiating treatment intensification regardless of whether patients adhered to this treatment. Specifically, for semaglutide, assumptions are based on data for liraglutide and dulaglutide. For SOC, assumptions are based on all intensifications in the claims database. The data from the claims database are shown in Table 3-1. Because the baseline HbA1c was higher in the all intensifications population compared to the liraglutide and dulaglutide groups, the assumptions used for the sample size calculation were adjusted accordingly.

Table 3-1: HbAlc Change Data from Anthem Population via HealthCore Integrated Research Database on Patients Intensifying Diabetes Treatment after Metformin

Follow-up	Treatment intensification after metformin	Number of patients with HbA1c Follow-up data	Proportion Patients Hba1c <7% at Follow-up	Number of patients with Baseline and Follow-up HbA1c data	HbA1c (%) Baseline mean	HbA1c (%) Change from Baseline mean	SD
1y	All intensifications	6,522	47%	4,068	8.9	-1.54	2.3
	dulaglutide	411	62%	115	8.6	-1.47	2.2
	liraglutide	1,762	62%	394	8.4	-1.34	2.0
2y	All intensifications	4,047	49%	2,401	8.9	-1.42	2.3
	dulaglutide	32	71%	≤10	8.8	-1.35	1.2
	liraglutide	893	55%	217	8.6	-1.24	2.1

Source: Data from the HealthCore Integrated Research Database. Data on file.

Table 3-2 presents several sample size scenarios together with assumptions and calculated power. The sample size calculation aimed for 90% power for simultaneously confirming superiority of semaglutide versus SOC on all of the primary endpoint and the secondary confirmatory endpoints at both year 1 and year 2 (4 hypotheses in total). For the sample size calculation the likelihood ratio chi-square test was assumed for analyses of the binary presence of HbA1c < 7.0% endpoints and a t-test for the continuous change in HbA1c endpoints. An overall alpha level of 5% with two-sided tests was assumed.

Table 3-2 Assumptions used in the sample size and calculated total power.

Total Number Randomized	Proportion Patients with HbA1c <7% at 1 and 2 Years		HbA1c ETD at 1 and 2 years	HbA1c change SD	Power total
	SOC	Semaglutide			
1849	50%	60%	0.5	2.3	80%
2250	50%	58%	0.5	2.3	72%
2250	**50%**	**60%**	**0.5**	**2.3**	**90%**
2250	50%	60%	0.5	2.0	93%
2250	50%	60%	0.4	2.3	77%

ETD = Estimated Treatment Difference, SD=Standard Deviation
Bold marked numbers indicate the chosen assumptions and numbers marked in underline indicate variations

The proportion of missing data for the confirmatory endpoints was estimated to be 25% after one year and 44% after two years. In the sample size calculation, it was assumed that only non-missing data at year 1 and year 2 would be used for the respective analyses. This is considered conservative, since the use of imputed data in the actual primary analysis (see 4.5) will increase the power. When accounting for missing data, randomizing 2250 patients will contribute 1687 patients for the year 1 analyses and 1260 patients for the year 2 analyses, achieving a total power of 90% for confirming all 4 confirmatory hypotheses. The corresponding marginal powers for presence of HbA1c < 7.0% at year 1, change in HbA1c to 1 year, presence of HbA1c < 7.0% at year 2 and change in HbA1c to 2 years are 99%, 99%, 95% and 97% respectively.

3.6 Missing Data

Except where noted in section 4.5 for HbA1c, weight, systolic blood pressure (SBP), and diastolic blood pressure (DBP) endpoints, missing data will not be imputed and will be excluded from calculations.

3.7 Data Definitions and Calculations

3.7.1 General

All absolute change from baseline variables will be calculated as:

Change = Endpoint – Baseline

All relative change from baseline variables will be calculated as:

Relative Change (%) = [(Endpoint– Baseline)/Baseline] x 100

If day and/or month is missing from dates, the first day of the month and/or the first month of the year will be assumed when calculating duration between dates. If imputation of an incomplete date associated with a post-randomization data element results in a date prior to randomization, the first day of the month immediately following randomization will be assumed. E.g., if the randomization date is 08/15/18 (mm/dd/yy) and the unknown date is 08/UNK/18 or UNK/10/18 or UNK/UNK/18, the imputed date would be 09/01/18.

The data sources for each endpoint and variables used in analysis are summarized in Appendix A.

3.7.2 Demographics

Age will be presented in years based on the patient's birth date and randomization visit date.

3.7.3 Diabetes History

Duration of diabetes will be presented in years based on the patient's diagnosis date and randomization visit date.

3.7.4 Treatment Variables

Study drug is defined as the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization. A patient can start and stop study drug throughout the study. At any time during the duration of the study, if a patient is taking study drug as defined, they will be considered "on study drug" regardless of prior study drug discontinuations. At any time during the duration of the study, if a patient is not taking study drug as defined, they will be considered "off study drug." On and off study drug classification will be based on the eCRF data.

The first study drug discontinuation date will be defined as the date of the first time a patient discontinues study drug such that they are "off study drug."

For analysis of year 1 and year 2 endpoints, patients will categorized as "on study drug" or "off study drug" according to their study drug use at the endpoint visit, regardless of prior use/non-use. If study drug status is missing for the year 1 or year 2 endpoint, the last registered study drug status prior to the endpoint visit will be used.

3.7.4.1 Semaglutide Study Drug

For the semaglutide treatment group, study drug is defined as semaglutide. Therefore at any time during the study, patients randomized to the semaglutide group will be categorized as:

On Study Drug: Taking Semaglutide (regardless of any other antidiabetic medication use or previous discontinuations)

Off Study Drug: Not currently taking Semaglutide

The following definitions will also apply to individual patient semaglutide dosing during the 2 year study period, or until time of study drug discontinuation:

Most Common: The dose associated with the longest period of use from randomization to year 2/EOS visit.
Highest: The highest dose reported at any time from randomization to year 2/EOS visit.
Final: The latest dose reported from randomization to year 2/EOS visit.

3.7.4.2 Standard of Care (SOC) Study Drug

For the SOC treatment group, study drug is defined as the *drug class* of the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization (excluding semaglutide). If a fixed dose combination (FDC) product is prescribed, one of the drugs will be specified by the study physician to be the study drug and on/off drug categorization will be based on the study physician specified drug only. Therefore, at any time during the study, patients randomized to the SOC treatment group will be categorized as:

On Study Drug: Taking a drug in the same drug class as their first study drug, even if it is not the same individual drug (regardless of any other antidiabetic medication use or previous discontinuations except semaglutide). This also applies to treatment intensification with a FDC product if one of the drugs is in the same drug class as the first study drug.

Off Study Drug: Not currently taking a drug in the same drug class as their first study drug.

3.7.4.3 Treatment Intensification or Change

Treatment intensification (add-on) or change (switch) in study drug after randomization will be identified and categorized.

Treatment intensification is defined as initiation of an antidiabetic medication (besides metformin) in addition to study drug, or initiation of more than one antidiabetic medication after discontinuation of study drug.

The first treatment intensification date will be defined as the date of the first occurrence of an antidiabetic treatment intensification.

Treatment change is defined as discontinuation of study drug and initiation of another antidiabetic medication (besides metformin).

The first treatment change date will be defined as the date of the first occurrence of an antidiabetic treatment change.

3.7.5 Measures of Glycemic Control

Physicians will set a patient's individualized HbA1c target prior to randomization. Achieving target at year 1 and year 2 is defined as:

Endpoint HbA1c $\leq$ Individualized HbA1c target

Attainment of the 2018 Healthcare Effectiveness Data and Information Set (HEDIS) HbA1c will be defined at year 1 and year 2 as:

HbA1c <8.0% if age $\geq$ 65 years at endpoint or with defined comorbidities (Appendix B), else <7.0%

HbA1c data for the HEDIS endpoints will be obtained from the eCRF. HEDIS comorbidities as defined in Appendix B will be obtained from the claims data and apply to the year prior to endpoint. HEDIS criteria at year 1 will be based on comorbidities identified between the randomization and year 1 visit. HEDIS criteria at year 2 will be based on comorbidities identified between the year 1 and year 2 visits.

3.7.6 Weight

No weight gain is defined as: Endpoint Weight $\leq$ Baseline Weight

3.7.7 Adherence and Persistence to Treatment

Medication adherence refers to a patient's conformance to the provider's recommendation with respect to timing, dosage, and frequency of medication taken during the prescribed length of time.

For this study, the medication possession ratio (MPR) will be used to assess adherence. MPR will be calculated for the study drug from claims data for baseline to year 1 and for baseline to year 2 as follows:

$$\text{MPR (\%)} = \frac{\text{Sum of days supply for all prescription fills}}{\text{Total number of days in time period}} \times 100$$

MPR will be capped at 100%.

Medication persistence refers to whether a patient stays on therapy or the time from initiation to discontinuation of therapy. Medication persistence will be assessed from baseline to year 2.

For this study, medication persistence will be defined as duration of time in days from initiation of study drug to first study drug discontinuation from eCRF data as defined in section 3.7.4.

A second persistence measure will be defined as duration of time in days from the date of initiation of study drug to the date of the first intensification or change in study drug as defined in section 3.7.4.3.

3.7.8 Patient Reported Outcomes (PROs)

3.7.8.1 Diabetes Treatment Satisfaction Questionnaire (DTSQ)

The Diabetes Treatment Satisfaction Questionnaire (DTSQ) is a validated instrument that is widely used in clinical trials and for routine clinical monitoring. It is useful in assessing outcomes of diabetes care. [1] The questionnaire consists of eight items -- six items assess treatment satisfaction, where each item is scored on a scale of 0–6, with six representing the greatest satisfaction, and two items assess the perceived frequency of hyperglycemia and hypoglycemia, also scored on a scale of 0-6, where "0" corresponds to none of the time and "6" corresponds to most of the time.

The original DTSQ is now referred to as the status version (DTSQs) in order to distinguish it from the DTSQ change version (DTSQc), which was developed to overcome potential ceiling effects (i.e. where respondents score maximum or near-maximum satisfaction at baseline and can show little or no improvement at follow-up). [2] This study will use the DTSQs at randomization and the DTSQc at the dedicated year 1 and year 2 study visits.

Scoring the DTSQs

The DTSQs has 8 items and produces the following measures:

1. Treatment Satisfaction: Items 1, 4, 5, 6, 7 & 8 are summed to produce a Treatment Satisfaction score (range: 0 to 36). The higher the score, the greater the satisfaction with treatment.
2. Individual satisfaction with treatment items (items 1, 4, 5, 6, 7 & 8) can be considered separately: All rated: 6 (very satisfied, convenient, flexible, etc.) to 0 (very dissatisfied, inconvenient, inflexible, etc.). The higher the score, the greater the satisfaction with each aspect of treatment.
3. 'Perceived frequency of hyperglycemia' (item 2) & 'Perceived frequency of hypoglycemia' (item 3): Both rated: 6 ('most of the time') to 0 ('none of the time'). Here, lower scores indicate blood glucose levels closer to the ideal. Higher scores indicate problems.

Scoring the DTSQc

The change version has the same 8 items as the status version, with a small alteration to the wording of Item 7. The DTSQc instructions and response options differ from those of the DTSQs and produce measures of relative change in satisfaction rather than measures of absolute satisfaction.

1. Treatment Satisfaction (Change): Items 1, 4, 5, 6, 7 and 8 are summed to produce a Treatment Satisfaction (change) score (range: +18 to –18). The higher the score, the greater the improvement in satisfaction with treatment; the lower the score, the greater the deterioration in satisfaction with treatment. A score of 0 represents no change.
2. Individual satisfaction with treatment change items (items 1, 4, 5, 6, 7 and 8) can be considered separately: All rated: +3 ('much more satisfied', 'much more convenient', 'much more flexible', etc.) to –3 ('much less satisfied', 'much less convenient', 'much less flexible', etc.). The higher the score, the greater the improvement in satisfaction with each aspect of treatment and the lower the score, the greater the deterioration in satisfaction with each aspect of treatment.
3. Two remaining items ('Perceived change in frequency of hyperglycemia' (item 2) and 'Perceived change in frequency of hypoglycemia' (item 3)) are treated individually: Both rated: +3 ('much more of the time now') to –3 ('much less of the time now'). Here, negative scores indicate fewer problems with blood glucose levels. Positive scores indicate more problems than before.

3.7.8.2 Short Form 12-Item Version 2 (SF-12 v2)

The Short Form 12-item Version 2 (SF-12v2) is a generic health-related quality of live (HRQoL) measure that consists of 12 items from the Short Form 36-Item (SF-36) health survey and

measures physical and mental dimensions of health. [3] Eight subscale scores (physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health) and two summary scores (physical component summary (PCS) and mental component summary (MCS)) can be obtained using the proprietary scoring software provided by Optum. Scores are calculated as norm-based standardized scores ranging from 0 to 100 with means of 50 and SDs of 10, where a score of "0" indicates the lowest level of health measured by the scales, "100" indicates the highest level of health, and "50" indicates the mean or norm-score for the general US population. [4]

For this study, the PCS-12 and MCS-12 summary scores will be used. The PCS-12 and MCS-12 will be calculated at baseline, year 1, and year 2. Change from baseline will be calculated for the PCS-12 and MCS-12 at year 1 and year 2.

3.7.8.3 Work Productivity and Activity Impairment, General Health (WPAI-GH)

The Work Productivity and Activity Impairment (WPAI) questionnaire is a 6-item instrument that measures absenteeism (work time missed), presenteeism (impairment at work / reduced on-the-job effectiveness), work productivity loss (overall work impairment / absenteeism plus presenteeism), and activity impairment because of health problems in the past seven days. [5] There are two versions of the WPAI. In the general health version (WPAI:GH) individuals are asked questions about work and activity impairment due to health problems. In the specific health problem version (WPAI:SHP), individuals are asked questions concerning impairment due to a target health problem (e.g., arthritis). This study will use the general health version of the WPAI (WPAI:GH), per general guidelines from the website that indicate the general health version is more appropriate for a disease like diabetes.

WPAI:GH Scoring Information

General Information:

In general, WPAI outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity, i.e., worse outcomes. The recall period is the past seven days, excluding the current day.

Questions:

Q1 = currently employed (Yes=works full-/part-time; self-employed; works in family business; on vacation from paid employment/No=does not work for pay; only does volunteer work; usually works, but has been laid-off or unemployed during the past seven days; retired; seasonal workers not currently working)

Q2 = hours missed due to health problems
Q3 = hours missed for other reasons
Q4 = hours actually worked
Q5 = degree health affected productivity while working
Q6 = degree health affected regular activities

Scores:

Multiply scores by 100 to express in percentages

1. Percent work time missed due to health (Absenteeism): Q2/(Q2 + Q4)
2. Percent impairment while working due to health (Presenteeism): Q5 /10
3. Percent overall work impairment due to health (Work Productivity Loss): Q2/(Q2 + Q4) + [(1 – (Q2/(Q2 + Q4))) x (Q5/10)]
4. Percent activity impairment due to health (Activity Impairment): Q6/10

Absenteeism, presenteeism, work productivity loss, and activity impairment will be calculated at baseline, year 1, and year 2. Change from baseline will be calculated for each score at year 1 and year 2.

3.7.8.4 Patient Global Impressions of Disease Severity (PGI-S) and Change (PGI-C)

The patient global impression of disease severity (PGI-S) is a single item that the patient will complete at the baseline randomization visit that assesses the patient's perception of the severity of their diabetes at the time of the randomization visit. The patient global assessment of change (PGI-C) is a single item that the patient will complete at the dedicated year 1 and year 2 study visits that assesses their perception of their current disease status relative to the way it was before they started the study drug.

3.7.8.5 Clinician Global Impressions of Disease Severity (CGI-S) and Change (CGI-C)

The clinician global impression of disease severity (CGI-S) is a single item that the clinician will complete after the patient's baseline randomization visit that assesses the clinician's perception of the patient's severity of their diabetes at the randomization visit. The clinician global assessment of change (CGI-C) is a single item that the clinician will complete at the patient's dedicated year 1 and year 2 study visits that assesses the clinician's perception of the patient's current disease status relative to the way it was before they started the study drug.

3.7.9 Healthcare Resource Utilization (HCRU)

All-cause and T2DM-related HCRU will be defined from baseline to year 2. These data will come from the HIRD. HCRU will include inpatient admissions, ER encounters, physician office visits (OV), other outpatient encounters, and pharmacy prescription fills.

All medical claims (inpatient admissions, ER visits, physician OVs, or other outpatient encounters) will be included in the "all-cause" analysis. All pharmacy claims will be included in the "all-cause" HCRU. Inpatient admissions, ER visits, physician OVs, or other outpatient encounters will be considered to be "T2DM-related" if the associated medical claim(s) has an International Classification of Diseases, Tenth Revision (ICD-10) T2DM diagnosis code of E11.% in the primary position or secondary position. Pharmacy claims for antidiabetic medications will be included in the "T2DM-related" HCRU.

- All-cause inpatient admission is defined by a claim or claims with inpatient place of service. Facility and provider claims are rolled into episodes with unique admission and discharge dates. ER visit claims on one day that are directly followed by an inpatient admission are considered as one inpatient admission and all HCRU from the encounter is combined.
- T2MD-related inpatient admission is defined by a claim or claims with inpatient place of service and primary or secondary diagnosis of T2DM. Facility and provider claims are rolled into episodes with unique admission and discharge dates. ER visit claims on one day that are directly followed by an inpatient admission with a primary or secondary diagnosis of T2DM (regardless of the diagnoses associated with the ER visit claim) are considered as one T2MD-related inpatient admission and all HCRU from the encounter is combined.
- All-cause ER visit is defined by a medical claim with ER place of service.
- T2DM-related ER visit is defined by a medical claim with ER place of service and primary or secondary diagnosis of T2DM.
- All-cause outpatient physician OV is defined by a medical claim with outpatient place of service and an evaluation and management (E&M) code that indicates physician OV.
- T2DM-related outpatient physician OV is defined by a medical claim with outpatient place of service and an E&M code that indicates physician OV with a primary or secondary diagnosis of T2DM.
- All-cause other outpatient service is defined by a medical claim with outpatient place of service and code other than E&M, overall and by the following categories:
    - Tests – Lab

  - Imaging
  - Procedures
  - Occupational Therapy (OT)/Speech
  - Medication and Related Services
  - Durable Medical Equipment
  - Physician Other Services
  - Tests – Other
  - Other

- T2DM-related other outpatient service is defined by a medical claim with outpatient place of service, with a primary or secondary diagnosis of T2DM, and code other than E&M, overall and by the following categories:
  - Tests – Lab
  - Imaging
  - Procedures
  - OT/Speech
  - Medication and Related Services
  - Durable Medical Equipment
  - Physician Other Services
  - Tests – Other
  - Other
- All-cause pharmacy prescription fill is defined by a pharmacy claim.
- T2DM-related pharmacy prescription fill is defined by a pharmacy claim for an antidiabetic medication.

The length of stay (LOS) for individual all-cause and T2DM-related inpatient admissions is defined as:

LOS (days) = Discharge Date – Admission Date

The cumulative LOS will be the sum of all all-cause inpatient admissions and all T2DM-related inpatient admissions a patient experiences from baseline to year 2.

3.8 Examination of Subgroups

Data will be presented overall and by treatment group. Additionally, subgroup analysis under the primary estimand (ITT) of the primary endpoint will be conducted on the following subgroups:

- Age (<55 years versus ≥55 years)
- Gender

- Baseline HbA1c (<8.0% versus ≥8.0%)
- Baseline body mass index (BMI) (<30 kg/m^2, ≥30 kg/m^2)
- Duration of T2DM diagnosis at baseline (<5 years versus ≥5 years)

4.0 STATISTICAL ANALYSES

4.1 Overview

The primary analysis for this study is a year 1 analysis. Once data collection for year 1 has completed, a database lock and year 1 analysis will be performed. The year 1 analysis will not be integrated with HCRU and will be limited to year 1 endpoints derived from eCRF data, including PRO data. The year 1 analysis results will be presented in a clinical study report and an internal results meeting. To maintain study integrity for the remaining study period, data from year 1 will be used for limited and confidential communications while complying with public disclosure requirements. All other analyses will be conducted following a second database lock once data collection for the entire study is complete.

No changes to the study design will be made based on the results of the year 1 analysis.

Descriptive summaries will be produced for both the primary (ITT) and secondary ("if all patients adhered") estimands, with the exception of antidiabetic treatment patterns which will be based on the primary estimand only. All statistical analyses will be conducted under both the primary estimand (ITT) and secondary estimand ("if all patients had adhered"), with the exception of the adherence and persistence measures, which will be based on the primary estimand only.

4.2 Patient Disposition and Accountability

Patient disposition will be descriptively summarized for all randomized patients, including the number and percentage patients in the FAS and safety populations, patients who complete the dedicated year 2 study visit, and the primary reason for not completing the dedicated year 2 study visit for patients who terminated participation in the study early. Summary metrics for number of screen failures by study site will also be presented.

4.3 Demographics and Baseline Clinical Characteristics

Demographic and baseline clinical characteristics will be descriptively summarized for the FAS overall and by treatment group. For baseline comorbid conditions, the number and percentage of patients with at least one of the comorbid conditions collected in the study will also be summarized along with the relative frequencies of type of comorbid condition. The same

approach will be followed for concomitant cardiovascular medications. In addition to eCRF data, type of insurance at randomization (obtained from claims) and geographic region (from study site location) will be summarized.

4.4 Protocol Deviations

Major protocol deviations will be summarized by type (informed consent, randomization error, inclusion/exclusion criteria error, unreported serious adverse events (SAEs), procedure not per protocol, study visit not per protocol, other).

4.5 Missing Value Imputation Methodology

Missing endpoint data will be imputed for HbA1c, weight, SBP, and DBP. The data of patients who, in violation of the protocol, initiate semaglutide in the SOC treatment group, will be censored analytically following initiation of semaglutide. The patient will remain in the analysis, but data following their initiation of semaglutide will be censored and treated as missing.

Prior to analysis of binary endpoints derived by dichotomizing HbA1c and body weight, missing data will be imputed by multiple imputation on the continuous scale and dichotomized as required for the endpoint. 500 complete data sets will be generated to adequately account for the uncertainty due to missing data.

4.5.1 Primary Estimand Imputation Model

Missing value imputation for the primary estimand will be based on data from all patients with observations at the endpoint visit (year 1 for the primary endpoint). Missing endpoint data will be imputed separately by treatment group and based on similar patients with available endpoint data according to study drug treatment status, i.e., according to whether these patients are on or off study drug at the endpoint visit. Data will be imputed based on the assumption that, within treatment groups, patients with missing endpoint data will behave like patients with the same study drug treatment status at endpoint as the missing patients' last registered treatment status prior to the missing endpoint. For example, for patients with missing data who have discontinued treatment with study drug, data will be imputed based on the assumption that these patients will behave like patients with available data who are no longer receiving study drug.

Technically, the imputation model will be an analysis of covariance (ANCOVA) for the endpoint data. The ANCOVA will include the baseline value of the imputed variable (HbA1c, weight, SBP, or DBP), diabetes duration, age, and sex as independent variables. After this model has been used to predict missing values, each of the now 500 complete data sets will be analyzed as described for the primary analysis in section 4.6.1. Finally, the multiple analysis results will be combined using Rubin's rule. [6] For the OR, the results will be combined on the logarithm

scale.

This process will be repeated for year 2 endpoints.

4.5.2 Secondary Estimand Imputation Model

Missing value imputation for the secondary estimand will be based only on data from the subset of patients who are receiving study drug (on study drug) at the endpoint visit (year 1 for the primary endpoint) in order to estimate the treatment effect if all patients had continued treatment. The endpoint data of patients who are not receiving study drug at the endpoint visit (off study drug) will be censored analytically and imputed together with missing data. Collectively, missing and censored data will be imputed separately by treatment group based on all patients who are on study drug at the endpoint visit. Data will be imputed based on the assumption that patients with missing data will behave like patients with available data who are on study drug.

The technical aspects of missing data imputation based on multiple imputation, the statistical analysis of the multiple complete data sets, and combination of the multiple results will be the same as the ones described for the primary estimand.

This process will be repeated for year 2 endpoints.

4.6 Effectiveness Analysis

4.6.1 Statistical Analysis for the Primary Estimand

The primary effectiveness analysis is the primary estimand of the primary endpoint, HbA1c <7.0% at year 1, and confirmatory secondary endpoints of change in HbA1c (%-point) from baseline to year 1, HbA1c <7.0% at year 2, and change in HbA1c (%-point) at year 2. This analysis will be based on the FAS with missing data imputation as described in the primary estimand imputation model.

Year 1 Analysis

The primary endpoint, HbA1c < 7.0% at year 1, will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide versus SOC) will be presented.

The confirmatory secondary continuous endpoint of change in HbA1c (%-point) at year 1 will be analyzed using ANCOVA that will include the treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide versus SOC) will be presented.

The estimated treatment effect from each of these analyses will be complemented with associated 95% CI and two-sided p-value for testing the null-hypothesis of no difference.

Superiority for the primary endpoint, HbA1c <7.0% at year 1, will be considered established if the OR 95% confidence interval (CI) is greater than 1, or similarly if the two-sided p-value is significant on a 5% level and the treatment OR is in favor of semaglutide.

If the hierarchical testing scheme allows, superiority for change in HbA1c from baseline to endpoint visit will be considered established if the 95% CI for the estimated treatment difference is smaller than 0, or similarly if the two-sided p-value is significant on a 5% level and the treatment difference is in favor of semaglutide.

Year 2 Analysis

The Year 1 analysis described above will be repeated for the confirmatory secondary endpoints of HbA1c <7.0% at year 2 and change in HbA1c (%-point) at year 2. If the hierarchical testing scheme allows, superiority will be tested first for HbA1c <7.0% at year 2 followed by change in HbA1c (%-point) at year 2 as described for the year 1 counterparts above.

4.6.2 Statistical Analysis for the Secondary Estimand

The primary effectiveness analysis will be repeated with the secondary estimand. This analysis will be based on the FAS with missing data imputation as described in the secondary estimand imputation model. With the exception of the data used and the imputation of missing data, which follows the secondary estimand, the statistical analysis is the same as the primary effectiveness analysis.

The binary endpoints of HbA1c <7.0% at endpoint visit will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide versus SOC) and 95% CI will be presented.

The continuous endpoints of change in HbA1c (%-point) from baseline to endpoint visit will be analyzed using ANCOVA with treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide versus SOC) will be presented along with associated 95% CI.

4.6.3 Additional Analyses of Glycemic Control

4.6.3.1 Descriptive Summaries of HbA1c

Observed HbA1c values will be descriptively summarized at baseline, year 1, and year 2. Additionally, the following categories of observed HbA1c will be descriptively summarized at baseline, year 1, and year 2:

- <7.0, 7.0-8.0, 8.1-9.0, 9.1-10.0, >10.0%

- <8.0 versus ≥8.0%
- <9.0 versus ≥9.0%

4.6.3.2 Supportive Analyses of Glycemic Control

The following supportive endpoints of glycemic control are binary endpoints. These analyses will compare the proportion of patients achieving the secondary endpoints related to glycemic control between semaglutide and SOC treatment arms via a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide versus SOC) will be presented.

- Individualized HbA1c target attained at year 1 (yes/no)
- HbA1c <7.0% (53mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 1 (yes/no)
- HbA1c target attainment per Healthcare Effectiveness Data and Information Set (HEDIS) criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 1 (yes/no)
- Individualized HbA1c target attained at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 2 (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 1 (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 2 (yes/no)
- HbA1c target attainment per HEDIS criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <7.0% (53 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)

4.6.4 Supplementary Analyses

The HbA1c analyses described for the primary and secondary effectiveness analyses described above will be complemented with two complete case analyses based on the subset of patients

without missing endpoint data, i.e., patients that do not have their endpoint imputed in the two analyses used for the primary and secondary estimand.

4.6.4.1 Complete Case Analysis

This analysis will be based only on patients with available measurements at the dedicated study visit including measurements irrespective of whether patients discontinued study drug or not.

The binary endpoints of HbA1c < 7.0% at year 1 and HbA1c < 7.0% at year 2 will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide versus SOC) will be presented along with associated 95% CI.

The continuous endpoints of change in HbA1c (%-point) at year 1 and change in HbA1c (%-point) at year 2 will be analyzed using ANCOVA with treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide versus SOC) will be presented along with associated 95% CI.

4.6.4.2 Complete Case On-Study Drug Analysis

The analysis will be based only on patients with available measurements at the dedicated study visit including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed. The analysis will use the same analysis models as described for complete case analysis above.

4.7 PRO Analysis

PROs and ClinROs will be measured with the instruments described in section 3.7.8. Analysis of these measures will address the secondary objective of this study to compare semaglutide versus SOC in the study's patient population as is relates to PROs and ClinROs, i.e., treatment satisfaction, generic health outcomes, work productivity, and patient and physician global assessment measures, over one and two year observation periods. PRO analysis will descriptively summarize these measures at baseline, year 1 and year 2, as well as compare semaglutide versus SOC for change from baseline to year 1 and year 2.

No imputation of missing PRO measures is planned. Therefore, the primary estimand for PRO analysis will be based on the FAS with endpoint data. The secondary estimand for PRO analysis will be based only on patients with endpoint data including only measurements for patients still on study drug. In the secondary estimand, patients who discontinue study drug will be censored and missing data will not be imputed.

Data regarding all PROs will be reported for the entire sample by treatment group. Percentages of completed PROs will be reported as well. Descriptive statistics will be reported for aggregate scores for all appropriate measures by total score and subscale score(s) if applicable. Semaglutide and SOC treatment groups will be compared by independent t-tests for PRO measures in which the data are normally distributed or the Mann-Whitney U-test for PRO measures in which the data are not normally distributed.

- Diabetes Treatment Satisfaction Questionnaire, change version (DTSQc), Total treatment satisfaction score measured at year 1
- DTSQc, Total treatment satisfaction score measured at year 2
- Change from baseline in Short Form 12-Item Version 2 Survey (SF-12 v2), Physical component summary (PCS-12) score at year 1
- Change from baseline in SF-12 v2, PCS-12 score at year 2
- Change from baseline in SF-12 v2, Mental component summary (MCS-12) score at year 1
- Change from baseline in SF-12 v2, MCS-12 score at year 2
- Change from baseline in Work Productivity and Activity Impairment, General Health questionnaire (WPAI-GH) Absenteeism (work time missed) score at year 1
- Change from baseline in WPAI-GH Absenteeism (work time missed) score at year 2
- Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 1
- Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 2
- Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 1
- Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 2
- Change from baseline in WPAI-GH Activity Impairment score at year 1
- Change from baseline in WPAI-GH Activity Impairment score at year 2

In addition to the above endpoints, PGI-S and CGI-S will be descriptively summarized at baseline overall and by treatment group. PGI-C and CGI-C will be descriptively summarized at year 1 and year 2 overall and by treatment group.

During later analysis on the primary endpoints, PRO measures can be used in two ways:

- Correlational analysis of the PRO measures with the dependent variable. These correlations will include the appropriate measure of Pearson's 'r', Spearman's rho,

Kendall's tau statistics or other appropriate measures that might fit the nature of the variable.

- Should the correlational analysis from step 1 find that patient characteristics are linked to primary analysis endpoints, multivariate analysis or Logistic Regression will be conducted to determine if there are any patient related variables from the PROs that predict the effectiveness of therapy in combination with other clinical variables. Data reported will include model specifics (H-L Goodness of fit; R2, beta coefficients for each variable as well as ORs with 95% CIs.)

4.8 Concomitant Medications

The number and percentages patients reporting pre-specified concomitant cardiovascular medications during year 1 and during year 2 will be summarized for the population overall and by treatment group.

4.9 Other Analyses

The following analyses further support the primary objective to compare semaglutide versus SOC in glycemic control. They will address the secondary objectives of this study to compare semaglutide versus SOC in the study's patient population over one and two year observation periods as is relates to body weight loss, hypoglycemia, HCRU, adherence and persistence to treatment, antidiabetic medication treatment patterns, and safety.

Except where otherwise noted, the following analyses will be conducted for the primary estimand and the secondary estimand. For endpoints that do not include HbA1c, weight, SBP, or DBP, no missing data will be imputed. Therefore, the primary estimand for these endpoints will be based on the FAS with endpoint data. The secondary estimand for these endpoints will be based only on patients with endpoint data including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed.

4.9.1 Weight Loss

Change in patient weight from baseline to endpoint will be calculated in absolute change (pounds) and relative change (percentage). Mean change in weight will be compared between semaglutide and SOC treatment groups following the ANCOVA model described for continuous endpoints in the primary (4.6.1) and secondary estimand (4.6.2). This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

- Change in body weight (%) from baseline to year 1
- Change in body weight (lb) from baseline to year 1
- Change in body weight (%) from baseline to year 2

- Change in body weight (lb) from baseline to year 2

4.9.2 Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP)

Absolute change in patient SBP and DBP from baseline to year 1 and baseline to year 2 will be calculated. Mean change in SBP and DBP will be compared between the semaglutide and SOC treatment groups following the ANCOVA model described for continuous endpoints in the primary (4.6.1) and secondary estimand (4.6.2). This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

4.9.3 Hypoglycemia

The total number of hypoglycemic episodes leading to an inpatient admission or ER encounter will be summed per patient from baseline to year 2 and compared by semaglutide and SOC treatment groups utilizing a negative binomial model with fixed effect of treatment and baseline HbA1c as a covariate. Exploratory analysis will assess the impact of SOC study drug (defined in section 3.7.4.2) on the rate of hypoglycemia.

Additionally, the frequency and percentage of patients reporting hypoglycemic episodes leading to an inpatient admission or ER encounter from baseline to year 1 and from baseline to year 2 will be reported and compared between treatment groups via chi-square test.

Hypoglycemic episodes that are considered SAEs will also be included in the safety analysis (section 5.4).

4.9.4 Healthcare Resource Utilization (HCRU)

HCRU analyses will compare all-cause and diabetes-related HCRU (inpatient admissions, ER encounters, physician OV, other outpatient encounters, and pharmacy utilization) as defined in section 3.7.9 from the HIRD by semaglutide and SOC treatment arm from baseline to year 2. The number of each type of encounter (all-cause and T2DM-related: inpatient admission, ER encounter, physician OV, other outpatient encounter (overall and by specific categories listed in section 3.7.9), pharmacy prescription fills) a patient experiences from baseline to year 2 will be summed.

HCRU will be summarized by setting (i.e., inpatient, ER, OV, outpatient, pharmacy) and cause (i.e., all-cause versus diabetes-related), for the study population overall and by treatment group. The occurrence (yes/no) and number of each type of encounter or prescription fill will be reported. LOS will be reported per admission and cumulative (baseline to year 2). LOS for inpatient admissions (days) per inpatient admission will be an encounter level analysis of the average inpatient admission LOS based on the total number of inpatient admissions; therefore,

patients may contribute more than once. Statistics for cumulative LOS for inpatient admissions (days) will only be produced for patients with at least one inpatient admission.

Treatment group comparisons will utilize the Student t-test for continuous variables and chi-squared test (or Fisher's exact test if outcome count less than 5) for categorical variables.

4.9.5 Composite Endpoints

The following composite endpoints will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c and baseline weight as covariates. This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 1 (yes/no)
- Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 2 (yes/no)
- Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 2 (yes/no)

4.9.6 Antidiabetic Treatment Patterns

4.9.6.1 Study Drug Regimen

Study drug regimen will be descriptively summarized by treatment arm.

The SOC group's randomized treatment regimen will be summarized by medication and class for study drug, i.e., the first antidiabetic medication prescribed for treatment intensification following randomization.

For the semaglutide group, relative frequency by dose category (0.25, 0.5, 1.0 mg) will be calculated at randomization, year 1, and year 2. Additionally, the relative frequency by semaglutide dose category will be calculated for most common, highest, and final as defined in section 3.7.4.1.

Treatment intensification and treatment change (definition 3.7.4.3) will also be descriptively summarized as follows:

- Treatment intensification from baseline to year 1 (yes/no)
- Treatment intensification from baseline to year 2 (yes/no)
- Treatment change from baseline to year 1 (yes/no)
- Treatment change from baseline to year 2 (yes/no)

4.9.6.2 Adherence and Persistence to Treatment

Adherence and persistence to the study drug will be calculated and compared between semaglutide and SOC treatment groups. This analysis will be conducted for the primary estimand only.

Study drug medication adherence will be summarized by the MPR (%) using the claims data. The MPR (%) for the first year of the study and the MPR (%) for the two year study period will be compared between treatment arms using the test of independent proportions.

Medication persistence, as calculated using the eCRF data, will be summarized by:

- Time to first study drug discontinuation during 2 years (day)
- Time to first treatment intensification (add-on) or change (switch) after randomization during 2 years (day)

Medication persistence measures will be compared between semaglutide and SOC treatment groups via the Cox proportional hazards model (with no covariates) to estimate the hazard ratio (HR) (semaglutide versus SOC) and 95% CI. Patients with no event will be censored at year 2 or last contact if withdrawn or lost to follow-up.

4.9.6.3 Other Antidiabetic Treatments

Background Medication

Metformin dose will be descriptively summarized overall and by treatment group at randomization, year 1, and year 2.

Other Antidiabetic Treatments

Antidiabetic treatment patterns will be assessed and compared between semaglutide and SOC treatment groups. This analysis will summarize the number and classification type of antidiabetic medications taken during the study period.

Descriptive statistics will be presented overall and by treatment group for the average number of antidiabetic medications patients report at year 1 and year 2. Relative frequencies of number of

antidiabetic medications will also be summarized for year 1 and year 2. Relative frequencies by antidiabetic class will also be presented at year 1 and year 2.

4.10 Study Site Recruitment and Participation

To contextualize the participating study sites within the pool of Anthem providers, tables summarizing the identification, recruitment, and participation of providers will be presented.

Site identification for this study begins with identifying the pool of potentially eligible patients in the HIRD using very broad claims-based inclusion/exclusion criteria: currently active members with at least one medical claim with an ICD-9 or ICD-10 diagnosis code for T2DM during the identification period, pharmacy carve-in benefits and pharmacy carve-out benefits at any time during the identification period, commercial plan coverage only, ≥18 years old, at least one pharmacy claim for metformin in the past 6 months, and no pharmacy claims for other antidiabetic medications within the past 30 days. These patients are then mapped back to the provider sites with which they had diabetes-related office visits during the identification period. From this process, a list of providers with number of potentially eligible patients is generated. The overall pool of Anthem providers serving the target study population will be summarized with a table presenting patient densities by number of providers.

A flow chart summarizing each step of the process will be produced to describe site recruitment and participation:

1) Total number of providers with at least one potentially eligible patient identified.
2) Total number of providers approached (I.e., providers who met predetermined threshold (x number of patients) for outreach.)
3) Disposition of providers approached
    a. Interested, invited to participate
    b. Interested, rejected from participation
    c. Not Interested
    d. No Response

A table summarizing characteristics of providers (geographic region, provider specialty, and claims-based patient density) by disposition status will be produced.

4.11 Evaluation of the Study Population

To evaluate the generalizability of the study results, an analysis of the study population will be performed. The primary objective of this analysis is to evaluate the external validity of the study and to understand how the study population fits into the larger T2DM populations within study sites and Anthem as a whole. Two claims-based T2DM patient populations will be described and

compared to the study population: (1) all research-eligible commercially-insured Anthem T2DM patients treated with metformin who undergo an antidiabetic treatment intensification and (2) within the study sites, commercially-insured Anthem T2DM patients treated with metformin who are not enrolled in the study, but who undergo an antidiabetic treatment intensification.

Identification of Claims-Based Comparator Populations

The claims-based comparator populations and analysis will be based on data from the HIRD. Claims based inclusion/exclusion criteria will be imposed on the two comparator populations as best as can be identified via the claims data and available lab result data to reflect the inclusion/criteria of the study population. All patients meeting the inclusion/exclusion criteria for each of the claims-based T2DM populations will be included in the analysis.

- Claims patient identification period = July 2018 until date last patient enrolled. (To align with the dates of the PCT study period).
- Claims index date = Patient's first prescription date associated with antidiabetic treatment intensification during the patient identification period.
- Claims study period = July 2017 until 12 months following date of last patient enrolled. (To allow for 12 months of claims data before and after the identified antidiabetic treatment intensification.)

Inclusion Criteria for Population #1

- ≥ 1 medical claim with an ICD-10 diagnosis code for T2DM during patient identification period
- ≥ 1 pharmacy claim for antidiabetic treatment in addition to metformin, as described in the PCT, during the patient identification period
    - First date of antidiabetic treatment pharmacy claim is index date
- ≥ 12 months medical and pharmacy eligibility prior to index date
- ≥ 1 pharmacy claim for metformin in 6 months prior to index date
- Age ≥ 18 years on index date
- HbA1c >7.0% within 90 days prior to and including index date

Exclusion Criteria for Population #1

- ≥ 1 pharmacy claim for any antidiabetic medication other than metformin for 30 days prior to index date.
- Randomized into the current PCT.

Inclusion Criteria for Population #2

The inclusion criteria for population #2 will include all of the inclusion criteria as population #1, as well as:

- Receiving antidiabetic treatment at a study site

Exclusion Criteria for Population #2

The exclusion criteria for population #2 is the same as the exclusion criteria for population #1.

All analyses evaluating the study population will be descriptive only and will help to contextualize study results within the T2DM population broadly. All analyses will be based on the overall population only; analyses by treatment group will not be conducted.

The analysis will descriptively compare the overall baseline demographics between the PCT study population and the two comparator populations. Baseline HbA1c (defined as closest HbAlc up to and including 90 days prior to randomization (for PCT population) or index date (for comparators)) will also be compared.

Observed HCRU (all-cause and T2DM-related; inpatient, ER, outpatient physician OV, and other outpatient services, pharmacy utilizations) during the 12 month period prior to randomization (for study population) or prior to index date (for comparators) will also be compared between the PCT study population and the two comparator populations. The PCT study population for this comparison will be limited to study patients with 12 months of health plan eligibility prior to randomization.

Additionally, treatment patterns of non-enrolled T2DM patients within the practices from which the study patients are recruited will be evaluated to identify any relevant patterns of care suggesting channeling of certain types of patients away from the study. This analysis be based on the subset of patients with 12 months post-index data medical and pharmacy eligibility and will present the relative frequency of antidiabetic medications by class during the 12 months following randomization or index date for the PCT study population and comparator population #2.

5.0 SAFETY

5.1 AEs

For the purpose of this study, AEs that do not meet the definition of an SAE (protocol section

7.2) will only be required to be collected in the eCRF if they lead to study drug discontinuation.

5.2 SAEs

All AEs meeting the definition of an SAE (protocol section 7.2) will be collected in the eCRF.

5.3 Pregnancy

Any abnormal pregnancy outcome (e.g., spontaneous miscarriage, fetal death, congenital anomaly/birth defect, etc.) is considered an SAE. For the purposes of this study, any pregnancies in participating female patients will be reported, along with pregnancy outcome and any AEs or SAEs observed in the fetus or newborn until 1 month of age.

5.4 Analysis

No formal safety analyses are planned for this study. SAEs and AEs leading to study drug discontinuation will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and descriptively summarized by System Organ Class and Preferred Term (PT). Pregnancies occurring during the study will be also be descriptively summarized.

6.0 Changes from Protocol

The protocol specified conducting PRO Analysis (Protocol Section 6.6) and Other Analyses (Protocol Section 6.8) as described for the primary estimand, secondary estimand, and supplementary analysis (complete case and complete case on-study drug). However, missing data imputation will only be completed for HbA1c, weight, SBP, and DBP. Therefore, for endpoints that do not include HbA1c, weight, SBP, or DBP data, the primary estimand will be based on the FAS with endpoint data. The secondary estimand for these endpoints will be based only on patients with endpoint data including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed. Supplementary analyses will not be completed for these endpoints.

7.0 SAP Change Log

Summary of SAP Changes since First Patient First Visit

Number	Section/Page	Change	Rationale

8.0 APPENDIX A: Study Endpoints and Other Variables by Data Source

Endpoint/Variable	Data Source	
	eCRF	HIRD (Claims)
Study Drug Variables		
On/Off Study Drug	X	
Treatment Intensification	X	
Treatment Change	X	
Antidiabetic treatment patterns	X	
Primary Endpoint		
HbA1c <7.0% at year 1 (yes/no)	X	
Confirmatory Endpoints		
Change in HbA1c (%-point) from baseline to year 1	X	
HbA1c <7.0% at year 2 (yes/no)	X	
Change in HbA1c (%-point) from baseline to year 1	X	
Supportive Secondary Endpoint Assessment		
Individualized HbA1c target attained at year 1 (yes/no)	X	
HbA1c <7.0% (53mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 1 (yes/no)	X	
HbA1c target attainment per Healthcare Effectiveness Data and Information Set (HEDIS) criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 1 (yes/no)	X	X
Change in body weight (lb) from baseline to year 1	X	
Change in body weight (%) from baseline to year 1	X	
Change in systolic blood pressure (SBP; mm Hg) from baseline to year 1	X	
Change in diastolic blood pressure (DBP; mm Hg) from baseline to year 1	X	

[redacted]

Time to first study drug discontinuation during 2 years (day)	X	
Time to first treatment intensification (add-on) or change (switch) after randomization during 2 years (day)	X	
Study drug medication adherence for the first year of the study, as measured by medication possession ratio (MPR) (%)		X
Number of hypoglycemic episodes leading to an inpatient admission or emergency room (ER) encounter from baseline to year 2	X	
Diabetes Treatment Satisfaction Questionnaire, change version (DTSQc), Total treatment satisfaction score measured at year 1	X	
DTSQc, Total treatment satisfaction score measured at year 2	X	
Change from baseline in Short Form 12-Item Version 2 Survey (SF-12 v2), Physical component summary (PCS-12) score at year 1	X	
Change from baseline in SF-12 v2, PCS-12 score at year 2	X	
Change from baseline in SF-12 v2, Mental component summary (MCS-12) score at year 1	X	
Change from baseline in SF-12 v2, MCS-12 score at year 2	X	
Change from baseline in Work Productivity and Activity Impairment, General Health questionnaire (WPAI-GH) Absenteeism (work time missed) score at year 1	X	
Change from baseline in WPAI-GH Absenteeism (work time missed) score at year 2	X	
Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 1	X	
Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 2	X	
Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 1	X	
Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 2	X	
Change from baseline in WPAI-GH Activity Impairment score at year 1	X	
Change from baseline in WPAI-GH Activity Impairment score at year 2	X	

All cause HCRU from baseline to year 2		
Number of inpatient admissions		X
LOS for inpatient admissions (days) per inpatient admission		X
Cumulative length of stay for inpatient admissions (days)		X
Number of ER encounters		X
Number of physician OV		X
Number of other outpatient encounters (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Number of medications		X
Occurrence of inpatient admission (yes/no)		X
Occurrence of ER encounter (yes/no)		X
Occurrence of physician OV (yes/no)		X
Occurrence of other outpatient encounter (yes/no) (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Diabetes related HCRU from baseline to year 2		
Number of diabetes related inpatient admissions		X
LOS for diabetes related inpatient admissions (days) per diabetes related inpatient admission		X
Cumulative length of stay for diabetes related inpatient admissions (days)		X
Number of diabetes related ER encounters		X
Number of diabetes related physician OV		X
Number of diabetes related other outpatient encounters (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X

[redacted]

Number of diabetes related medications		X
Occurrence of diabetes related inpatient admission (yes/no)		X
Occurrence of diabetes related ER encounter (yes/no)		X
Occurrence of diabetes related physician OV (yes/no)		X
Occurrence of diabetes related outpatient encounter (yes/no) (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Additional Derived Outcome Variables for Supportive Analyses		
Supportive Measures of Glycemic Control		
Individualized HbA1c target attained at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 2 (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 1 (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 2 (yes/no)	X	
HbA1c target attainment per HEDIS criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 2 (yes/no)	X	X
HbA1c <7.0% (53 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)	X	

[redacted]

Body Weight Loss		
Change in body weight (%) from baseline to year 2	X	
Change in body weight (lb) from baseline to year 2	X	
Blood Pressure		
Change in SBP (mm Hg) from baseline to year 2	X	
Change in DPB (mm Hg) from baseline to year 2	X	
Hypoglycemia		
Reported hypoglycemia leading to inpatient admission or ER encounter during year 1 (yes/no)	X	
Reported hypoglycemia leading to inpatient admission or ER encounter during year 2 (yes/no)	X	
Composite Variables		
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 1 (yes/no)	X	
Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 2 (yes/no)	X	
Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 2 (yes/no)	X	

Adherence to Treatment		
Study drug medication adherence for the two years of the study, as measured by the medication possession ratio (MPR) (%)		X

9.0 APPENDIX B: 2018 HEDIS Criteria

Denominator

The eligible population.

Note: *The eligible population for the HbA1c Control <7% for a Selected Population indicator is reported after required exclusions are applied.*

Required exclusions for HbA1c Control <7% for a Selected Population indicator

Exclude members who meet any of the following criteria:

- 65 years of age and older as of December 31 of the measurement year.
- *CABG.* Members who had CABG (CABG Value Set) in any setting during the measurement year or the year prior to the measurement year.
- *PCI.* Members who had PCI (PCI Value Set), in any setting, during the measurement year or the year prior to the measurement year.
- *IVD.* Members who met at least one of the following criteria during both the measurement year and the year prior to the measurement year. Criteria need not be the same across both years.
  - At least one outpatient visit (Outpatient Value Set) with an IVD diagnosis (IVD Value Set).
  - At least one acute inpatient encounter (Acute Inpatient Value Set) with an IVD diagnosis (IVD Value Set).
- *Thoracic aortic aneurysm.* Members who met at least one of the following criteria during both the measurement year and the year prior to the measurement year. Criteria need not be the same across both years.
  - At least one outpatient visit (Outpatient Value Set), with a diagnosis of thoracic aortic aneurysm (Thoracic Aortic Aneurysm Value Set).
  - At least one acute inpatient encounter (Acute Inpatient Value Set), with a diagnosis of thoracic aortic aneurysm (Thoracic Aortic Aneurysm Value Set).
- Any of the following, in any setting, any time during the member's history through December 31 of the measurement year.
  - *Chronic heart failure.* A diagnosis of chronic heart failure (Chronic Heart Failure Value Set).
  - *Prior MI.* A diagnosis of MI (MI Value Set).
  - *ESRD.* ESRD (ESRD Value Set; ESRD Obsolete Value Set).
  - *Chronic kidney disease (stage 4).* Stage 4 chronic kidney disease (CKD Stage 4 Value Set).
  - *Dementia.* A diagnosis of dementia (Dementia Value Set; Frontotemporal Dementia Value Set).
  - *Blindness.* A diagnosis of blindness (Blindness Value Set).
  - *Amputation (lower extremity).* Lower extremity amputation (Lower Extremity Amputation Value Set).

10.0 REFERENCES

[1] C. Bradley, "Diabetes Treatment Satistcation Questionnaire: DTSQ," in *Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Practice*, C. Bradley, Ed., Abington, Routledge, 1994, pp. 111-132.

[2] C. Bradley, "The Diabetes Treatment Satisfaction Questionnaire (DTSQ): change version for use alongside status version provides appropriate solution where ceiling effects occur," *Diabetes Care,* vol. 22, no. 3, pp. 530-532, 1999.

[3] J. Ware, M. Kosinski and S. Keller, "A 12-item short-form healch survey: construction of scales and preliminary tests of reliability and validity," *Med Care,* vol. 34, no. 3, pp. 220-233, 1996.

[4] J. Ware, M. Kosinski, D. Turner-Bowker and B. Gandek, How to score version 2 of the SF-12 health survey (with a supplement documenting version 1), Lincoln, RI: QualityMetric, Incorporated, 2002.

[5] M. Reilly, A. Zbrozek and E. Dukes, "The validity and reproducibility of a work productivity and activity impairment instrument," *Pharmacoeconomics,* vol. 4, no. 5, pp. 353-365, 1993.

[6] R. Little and D. Rubin, Statistical analysis with missing data, New York: John Wiley and Sons, 1987.

Statistical Analysis Plan for Long term comparative effectiveness of once weekly semaglutide versus standard of care in a real world adult US population with type 2 diabetes - a randomized pragmatic clinical trial

Protocol No. NN9535-4416

Novo Nordisk

June 30, 2020

Version 2.0

[redacted]

~~CONFIDENTIAL~~

[redacted]

Signature Page for Statistical Analysis Plan
Protocol No: NN9535-4416
30-JUN-2020

[redacted] **Approval:**

Name: [redacted]
Title: [redacted], Research Biostatistics

DocuSigned by:
Signature: [redacted]
Date: 30-Jul-2020

Name: [redacted]
Title: [redacted] Scientist

DocuSigned by:
Signature: [redacted]
Date: 04-Aug-2020

Novo Nordisk Approval:

Name: [redacted]
Title: [redacted]

DocuSigned by:
Signature: [redacted]
Date: 25-Aug-2020

Name: [redacted]
Title: [redacted]

DocuSigned by:
Signature: [redacted]
Date: 25-Aug-2020

[redacted]

TABLE OF CONTENTS

[redacted]

[redacted]

1.0 INTRODUCTION

This document describes the Statistical Analysis Plan (SAP) for Protocol Number NN9535-4416, Long term comparative effectiveness of once weekly semaglutide versus standard of care in a real world adult US population with type 2 diabetes – a randomized pragmatic clinical trial. This is a 2-year, multi-center, randomized, open label, parallel group, active comparator pragmatic clinical trial (PCT) comparing semaglutide s.c. versus standard of care (SOC) when added to up to 2 oral antidiabetic medications as treatment intensification among adult type 2 diabetes mellitus (T2DM) patients in the course of routine clinical practice.

This SAP is based on the most recent version of the protocol: Version 4.0, December 5, 2019.

Protocol History:

Number	Version	Date
1.0	Original Protocol	March 13, 2018
2.0	Version 2.0	March 27, 2019
3.0	Version 3.0	August 21, 2019
4.0	Version 4.0	December 5, 2019

2.0 STUDY OBJECTIVES

2.1 Primary Objective

The primary objective is to demonstrate superior long term effects of treatment with semaglutide s.c. compared to SOC each added to up to 2 oral antidiabetic medications on glycemic control when used as intensification in routine clinical practice in adult patients with T2DM.

2.2 Secondary Objectives

The secondary objectives are to compare the long term effect of semaglutide s.c. versus SOC each added to up to 2 oral antidiabetic medications and used as intensification in routine clinical practice in adult patients with T2DM with regards to:

- Weight loss
- Patient reported outcomes (PROs) and clinician reported outcomes (ClinROs)
- Hypoglycemia

- Healthcare Resource Utilization (HCRU)
- Adherence and persistence to treatment

2.3 Estimands

The estimand informs choices about data foundation and statistical analysis including possible imputation of missing data, hereby ensuring that randomization is preserved as a sound basis for statistical inference; i.e. estimation of effect size, associated uncertainty and statistical testing. Two estimands, each targeting a unique research question, have been defined to adequately describe the effect of semaglutide s.c. as compared to SOC (excluding any semaglutide).

2.3.1 Primary Estimand

The primary estimand for all objectives will be the "intention-to-treat" (ITT) estimand evaluating the effectiveness of randomized treatment intervention irrespective of adherence to this randomized intervention or changes to other antidiabetic medication.

2.3.2 Secondary Estimand

The secondary estimand for all objectives with the exception of the adherence and persistence to treatment objective, is the "if all patients had adhered" estimand. This estimand evaluates the effect of randomized treatment intervention for all randomized patients if all patients had adhered to randomized treatment, regardless of changes to other antidiabetic medication.

3.0 STATISTICAL METHODOLOGY

3.1 General Considerations

Statistical analyses will be performed using SAS® version 9.4 or higher computer software. Analyses will utilize prospectively collected data from study sites as well as secondary data collection utilizing administrative claims data from the HealthCore Integrated Research Database (HIRD®). Data from both sources will be integrated into one dataset for analysis.

Patient flow, patient characteristics, treatments, and outcomes will be tabulated and summarized with descriptive statistics based on observed values only. All descriptive data summaries will include means, medians, standard deviation (SD), and ranges for continuous variables and absolute/relative frequencies for categorical data. In addition, for each variable the number (count) of values that are missing will also be reported for dedicated study visit data. Missing data are data that are planned and can be collected but are absent. Statistics will be summarized for the study population overall and by treatment arm. Two sets of descriptive summaries will be

produced: one based on the primary estimand (ITT) and one based on the secondary estimand ("if all patients had adhered").

Raw data (i.e., minimum and maximum values presented for range in continuous variables) will be reported out to the precision with which it was collected. Means will be reported to 1 decimal place more than the raw data. SD will be reported to 1 decimal place more than the mean. Percentages will be reported to 1 decimal place. Trailing zeros will be presented to maintain a consistent level of precision, e.g. 2.0 rather than 2.

Inferential tests will be performed at the 5% level of significance (two-sided). All p-values will be rounded to 3 decimal places. If a rounded p-value is 0.000 (i.e., the actual p-value is less than 0.0005), then the p-value will be presented as '$p< 0.001$.'

3.2 Study Populations

3.2.1 Analysis Population

The following analysis set will be defined:

Full analysis set (FAS): Includes all randomized patients analyzed according to the treatment group to which they were assigned at randomization.

3.2.2 Safety Population

The safety population includes all randomized patients initiated on study drug.

3.3 Evaluation Schedule and Definitions

Dedicated study visits are at randomization, at year 1, and at year 2. The study will also capture data collected at the sites during routine diabetic care visits, i.e., office visits and other patient contacts that occur as part of routine clinical practice. Routine diabetic care visits will occur per study physician's routine clinical practice, therefore the number of visits and data available may differ from site to site and patient to patient.

For analysis, baseline includes assessments conducted and patient data collected at or prior to randomization. Baseline is defined as ≤ 90 days prior to randomization visit (week 0) for HbA1c. For secondary endpoint assessments, baseline is defined as ≤ 4 weeks prior to the randomization visit (week 0). All baseline values will be derived from the last assessment taken prior to or at the randomization visit.

The dedicated year 1 visit is at 52 weeks (±6 weeks) after randomization. Year 1 endpoint data will be derived from the dedicated year 1 study visit if the dedicated study visit is ±10 weeks of 52 weeks post-randomization. If year 1 endpoint data is missing or outside of 52 ±10 weeks, then the routine diabetic care data closest to 52 weeks post-randomization ±10 weeks will be used. If

no routine diabetic care data are available 52 weeks post-randomization ±10 weeks, then year 1 endpoint data will be considered missing and imputed as described in section 4.5, if applicable.

The dedicated year 2/EOS visit is 104 weeks (±6 weeks) after randomization. Year 2 endpoint data will be derived from the dedicated year 2 study visit if the dedicated study visit is ±10 weeks of 104 weeks post-randomization. If year 2 endpoint data is missing or outside of 104 ±10 weeks, then the routine diabetic care data closest to 104 weeks post-randomization ±10 weeks will be used. If no routine diabetic care data are available 104 weeks post-randomization ±10 weeks, then year 2 endpoint data will be considered missing and imputed as described in section 4.5, if applicable.

3.4 Confirmatory Endpoints and Hypotheses

The primary endpoint is HbA1c <7.0% (53 mmol/mol) at year 1 (yes/no).

Confirmatory secondary endpoints include:

- Change in HbA1c (%-point) from baseline to year 1
- HbA1c <7.0% (53 mmol/mol) at year 2 (yes/no)
- Change in HbA1c (%-point) from baseline to year 2

This study is designed to have 90% power to confirm superiority of the primary endpoint and 85% power to also confirm superiority of the first confirmatory secondary endpoint based on primary estimand analyses. The primary and confirmatory secondary endpoints will all be tested for superiority under multiplicity control via a hierarchical testing scheme according to the order given in the above bullet list. Confirmatory testing will only be performed for the primary estimand with the secondary estimand being supportive. The testing procedure will be stopped the first time an analysis fails to confirm superiority of the endpoint in question using a two-sided significance level of 5%.

Superiority on HbA1c <7.0% at year 1 and at year 2 will be evaluated with respect to the odds ratio (OR) (odds semaglutide / odds SOC):

H_0: OR≤1 against H_a: OR>1

Superiority on change in HbA1c (year 1 – baseline; year 2 – baseline) will be evaluated with respect to the mean treatment difference (TD) (semaglutide – SOC):

H_0: TD≥0 against H_a: TD<0

Statistical analysis of all other endpoints will be unadjusted for multiplicity.

3.5 Sample Size Estimation

3.5.1 Power and Sample Size for Primary Objective

Assumptions for the sample size were based on input from the HIRD claims/laboratory results database. In line with the primary ITT estimand, the assumptions for the proportion of patients with HbA1c < 7.0% at year 1 and year 2, and for the change from baseline in HbA1c are based on the claims/laboratory results data within the [redacted] population for all patients initiating treatment intensification regardless of whether patients adhered to this treatment. Specifically, for semaglutide s.c., assumptions are based on data for liraglutide and dulaglutide. For SOC, assumptions are based on all intensifications in the claims database. The data from the claims database are shown in Table 3-1. Because the baseline HbA1c was higher in the all intensifications population compared to the liraglutide and dulaglutide groups, the assumptions used for the sample size calculation were adjusted accordingly.

Table 3-1: HbAlc Change Data from [redacted] Population via HealthCore Integrated Research Database on Patients Intensifying Diabetes Treatment after Metformin

Follow-up	Treatment intensification after metformin	Number of patients with HbA1c Follow-up data	Proportion Patients Hba1c <7% at Follow-up	Number of patients with Baseline and Follow-up HbA1c data	HbA1c (%) Baseline mean	HbA1c (%) Change from Baseline	
						mean	SD
1y	All intensifications	6,522	47%	4,068	8.9	-1.54	2.3
	dulaglutide	411	62%	115	8.6	-1.47	2.2
	liraglutide	1,762	62%	394	8.4	-1.34	2.0
2y	All intensifications	4,047	49%	2,401	8.9	-1.42	2.3
	dulaglutide	32	71%	≤10	8.8	-1.35	1.2
	liraglutide	893	55%	217	8.6	-1.24	2.1

Source: Data from the HealthCore Integrated Research Database. Data on file.

Table 3-2 presents several sample size scenarios together with assumptions and calculated power. The sample size calculation aimed for 90% power for confirming superiority of semaglutide s.c. versus SOC on the primary endpoint and 85% power to also confirm superiority of the first confirmatory secondary endpoint based on an analysis of the primary estimand for each of the endpoints. For the sample size calculation the likelihood ratio chi-square test was assumed for analyses of the binary presence of HbA1c < 7.0% endpoints and a t-test for the continuous change in HbA1c endpoints. An overall alpha level of 5% with two-sided tests was assumed.

Table 3-2 Assumptions used in the sample size and calculated total power.

Total Number Randomized	Proportion Patients with HbA1c <7% at 1 and 2 Years		HbA1c ETD at 1 and 2 years	HbA1c change SD	Power total
	SOC	Semaglutide			
1849	50%	60%	0.5	2.3	80%
2250	50%	58%	0.5	2.3	72%
2250	**50%**	**60%**	**0.5**	**2.3**	**90%**
2250	50%	60%	0.5	2.0	93%
2250	50%	60%	0.4	2.3	77%

ETD = Estimated Treatment Difference, SD=Standard Deviation
Bold marked numbers indicate the chosen assumptions and numbers marked in underline indicate variations

The proportion of missing data for the confirmatory endpoints was estimated to be 25% after one year and 44% after two years. In the sample size calculation, it was assumed that only non-missing data at year 1 and year 2 would be used for the respective analyses. This is considered conservative, since the use of imputed data in the actual primary analysis (see 4.5) will increase the power. When accounting for missing data, randomizing 1387 patients will contribute 1040 patients for the year 1 analyses and 780 patients for the year 2 analyses, achieving a total power of 85% for confirming the two year 1 confirmatory hypotheses. The joint power for confirming all 4 confirmatory hypotheses is 58%. The corresponding marginal powers for presence of HbA1c < 7.0% at year 1, change in HbA1c to 1 year, presence of HbA1c < 7.0% at year 2, and change in HbA1c to 2 years are 90%, 94%, 80% and 86% respectively.

3.6 Missing Data

Except where noted in section 4.5 for HbA1c, weight, systolic blood pressure (SBP), and diastolic blood pressure (DBP) endpoints, missing data will not be imputed and will be excluded from calculations.

3.7 Data Definitions and Calculations

3.7.1 General

All absolute change from baseline variables will be calculated as:

Change = Endpoint – Baseline

All relative change from baseline variables will be calculated as:

Relative Change (%) = [(Endpoint– Baseline)/Baseline] x 100

If day and/or month is missing from dates, the first day of the month and/or the first month of the year will be assumed when calculating duration between dates. If imputation of an incomplete date associated with a post-randomization data element results in a date prior to randomization, the first day of the month immediately following randomization will be assumed. E.g., if the randomization date is 08/15/18 (mm/dd/yy) and the unknown date is 08/UNK/18 or UNK/10/18 or UNK/UNK/18, the imputed date would be 09/01/18.

The data sources for each endpoint and variables used in analysis are summarized in Appendix A.

3.7.2 Demographics

Age will be presented in years based on the patient's birth date and randomization visit date.

3.7.3 Diabetes History

Duration of diabetes will be presented in years based on the patient's diagnosis date and randomization visit date.

3.7.4 Treatment Variables

Study drug is defined as the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization. A patient can start and stop study drug throughout the study. At any time during the duration of the study, if a patient is taking study drug as defined, they will be considered "on study drug" regardless of prior study drug discontinuations. At any time during the duration of the study, if a patient is not taking study drug as defined, they will be considered "off study drug." On and off study drug classification will be based on the eCRF data.

The first study drug discontinuation date will be defined as the date of the first time a patient discontinues study drug such that they are "off study drug."

For analysis of year 1 and year 2 endpoints, patients will categorized as "on study drug" or "off study drug" according to their study drug use at the endpoint visit, regardless of prior use/non-use. If study drug status is missing for the year 1 or year 2 endpoint, the last registered study drug status prior to the endpoint visit will be used.

3.7.4.1 Semaglutide Study Drug

For the semaglutide treatment group, study drug is defined as semaglutide s.c.. Therefore at any time during the study, patients randomized to the semaglutide group will be categorized as:

On Study Drug: Taking Semaglutide s.c. (regardless of any other antidiabetic medication use or previous discontinuations)

Off Study Drug: Not currently taking Semaglutide s.c.

The following definitions will also apply to individual patient semaglutide s.c. dosing during the 2 year study period, or until time of study drug discontinuation:

Most Common: The dose associated with the longest period of use from randomization to year 2/EOS visit.
Highest: The highest dose reported at any time from randomization to year 2/EOS visit.
Final: The latest dose reported from randomization to year 2/EOS visit.

3.7.4.2 Standard of Care (SOC) Study Drug

For the SOC treatment group, study drug is defined as the *drug class* of the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization (excluding any formulation of semaglutide). If a fixed dose combination (FDC) product is prescribed, one of the drugs will be specified by the study physician to be the study drug and on/off drug categorization will be based on the study physician specified drug only. Therefore, at any time during the study, patients randomized to the SOC treatment group will be categorized as:

On Study Drug: Taking a drug in the same drug class as their first study drug, even if it is not the same individual drug (regardless of any other antidiabetic medication use or previous discontinuations except semaglutide). This also applies to treatment intensification with a FDC product if one of the drugs is in the same drug class as the first study drug.

Off Study Drug: Not currently taking a drug in the same drug class as their first study drug.

3.7.4.3 Treatment Intensification or Change

Treatment intensification (add-on) or change (switch) in study drug after randomization will be identified and categorized.

Treatment intensification is defined as initiation of an antidiabetic medication (besides the up to 2 oral antidiabetic medication(s) taken prior to randomization) in addition to study drug, or initiation of more than one antidiabetic medication after discontinuation of study drug.

The first treatment intensification date will be defined as the date of the first occurrence of an antidiabetic treatment intensification.

Treatment change is defined as discontinuation of study drug and initiation of another antidiabetic medication (besides the up to 2 oral antidiabetic medication(s) taken prior to randomization).

The first treatment change date will be defined as the date of the first occurrence of an antidiabetic treatment change.

3.7.5 Measures of Glycemic Control

Physicians will set a patient's individualized HbA1c target prior to randomization. Achieving target at year 1 and year 2 is defined as:

Endpoint HbA1c ≤ Individualized HbA1c target

Attainment of the 2018 Healthcare Effectiveness Data and Information Set (HEDIS) HbA1c will be defined at year 1 and year 2 as:

HbA1c <8.0% if age ≥ 65 years at endpoint or with defined comorbidities (Appendix B), else <7.0%

HbA1c data for the HEDIS endpoints will be obtained from the eCRF. HEDIS comorbidities as defined in Appendix B will be obtained from the claims data and apply to the year prior to endpoint. HEDIS criteria at year 1 will be based on comorbidities identified between the randomization and year 1 visit. HEDIS criteria at year 2 will be based on comorbidities identified between the year 1 and year 2 visits.

3.7.6 Weight

No weight gain is defined as: Endpoint Weight ≤ Baseline Weight

3.7.7 Adherence and Persistence to Treatment

Medication adherence refers to a patient's conformance to the provider's recommendation with respect to timing, dosage, and frequency of medication taken during the prescribed length of time.

For this study, the medication possession ratio (MPR) will be used to assess adherence. MPR will be calculated for the study drug from claims data for baseline to year 1 and for baseline to year 2 as follows:

$$\text{MPR (\%)} = \frac{\text{Sum of days supply for all prescription fills}}{\text{Total number of days in time period}} \times 100$$

MPR will be capped at 100%.

Medication persistence refers to whether a patient stays on therapy or the time from initiation to discontinuation of therapy. Medication persistence will be assessed from baseline to year 2.

For this study, medication persistence will be defined as duration of time in days from initiation of study drug to first study drug discontinuation from eCRF data as defined in section 3.7.4.

A second persistence measure will be defined as duration of time in days from the date of initiation of study drug to the date of the first intensification or change in study drug as defined in section 3.7.4.3.

3.7.8 Patient Reported Outcomes (PROs)

3.7.8.1 Diabetes Treatment Satisfaction Questionnaire (DTSQ)

The Diabetes Treatment Satisfaction Questionnaire (DTSQ) is a validated instrument that is widely used in clinical trials and for routine clinical monitoring. It is useful in assessing outcomes of diabetes care. [1] The questionnaire consists of eight items -- six items assess treatment satisfaction, where each item is scored on a scale of 0–6, with six representing the greatest satisfaction, and two items assess the perceived frequency of hyperglycemia and hypoglycemia, also scored on a scale of 0-6, where "0" corresponds to none of the time and "6" corresponds to most of the time.

The original DTSQ is now referred to as the status version (DTSQs) in order to distinguish it from the DTSQ change version (DTSQc), which was developed to overcome potential ceiling effects (i.e. where respondents score maximum or near-maximum satisfaction at baseline and can show little or no improvement at follow-up). [2] This study will use the DTSQs at randomization and the DTSQc at the dedicated year 1 and year 2 study visits.

Scoring the DTSQs

The DTSQs has 8 items and produces the following measures:

1. Treatment Satisfaction: Items 1, 4, 5, 6, 7 & 8 are summed to produce a Treatment Satisfaction score (range: 0 to 36). The higher the score, the greater the satisfaction with treatment.
2. Individual satisfaction with treatment items (items 1, 4, 5, 6, 7 & 8) can be considered separately: All rated: 6 (very satisfied, convenient, flexible, etc.) to 0 (very dissatisfied, inconvenient, inflexible, etc.). The higher the score, the greater the satisfaction with each aspect of treatment.
3. 'Perceived frequency of hyperglycemia' (item 2) & 'Perceived frequency of hypoglycemia' (item 3): Both rated: 6 (‘most of the time’) to 0 (‘none of the time’). Here, lower scores indicate blood glucose levels closer to the ideal. Higher scores indicate problems.

Scoring the DTSQc

The change version has the same 8 items as the status version, with a small alteration to the wording of Item 7. The DTSQc instructions and response options differ from those of the DTSQs and produce measures of relative change in satisfaction rather than measures of absolute satisfaction.

1. Treatment Satisfaction (Change): Items 1, 4, 5, 6, 7 and 8 are summed to produce a Treatment Satisfaction (change) score (range: +18 to –18). The higher the score, the greater the improvement in satisfaction with treatment; the lower the score, the greater the deterioration in satisfaction with treatment. A score of 0 represents no change.
2. Individual satisfaction with treatment change items (items 1, 4, 5, 6, 7 and 8) can be considered separately: All rated: +3 ('much more satisfied', 'much more convenient', ‘much more flexible', etc.) to –3 ('much less satisfied', 'much less convenient', 'much less flexible', etc.). The higher the score, the greater the improvement in satisfaction with each aspect of treatment and the lower the score, the greater the deterioration in satisfaction with each aspect of treatment.
3. Two remaining items ('Perceived change in frequency of hyperglycemia' (item 2) and 'Perceived change in frequency of hypoglycemia' (item 3)) are treated individually: Both rated: +3 (‘much more of the time now’) to –3 (‘much less of the time now’). Here, negative scores indicate fewer problems with blood glucose levels. Positive scores indicate more problems than before.

3.7.8.2 Short Form 12-Item Version 2 (SF-12 v2)

The Short Form 12-item Version 2 (SF-12v2) is a generic health-related quality of live (HRQoL) measure that consists of 12 items from the Short Form 36-Item (SF-36) health survey and measures physical and mental dimensions of health. [3] Eight subscale scores (physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health) and two summary scores (physical component summary (PCS) and mental component summary (MCS)) can be obtained using the proprietary scoring software provided by Optum. Scores are calculated as norm-based standardized scores ranging from 0 to 100 with means of 50 and SDs of 10, where a score of "0" indicates the lowest level of health measured by the scales, "100" indicates the highest level of health, and "50" indicates the mean or norm-score for the general US population. [4]

For this study, the PCS-12 and MCS-12 summary scores will be used. The PCS-12 and MCS-12 will be calculated at baseline, year 1, and year 2. Change from baseline will be calculated for the PCS-12 and MCS-12 at year 1 and year 2.

3.7.8.3 Work Productivity and Activity Impairment, General Health (WPAI-GH)

The Work Productivity and Activity Impairment (WPAI) questionnaire is a 6-item instrument that measures absenteeism (work time missed), presenteeism (impairment at work / reduced on-the-job effectiveness), work productivity loss (overall work impairment / absenteeism plus presenteeism), and activity impairment because of health problems in the past seven days. [5] There are two versions of the WPAI. In the general health version (WPAI:GH) individuals are asked questions about work and activity impairment due to health problems. In the specific health problem version (WPAI:SHP), individuals are asked questions concerning impairment due to a target health problem (e.g., arthritis). This study will use the general health version of the WPAI (WPAI:GH), per general guidelines from the website that indicate the general health version is more appropriate for a disease like diabetes.

WPAI:GH Scoring Information

General Information:

In general, WPAI outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity, i.e., worse outcomes. The recall period is the past seven days, excluding the current day.

Questions:

Q1 = currently employed (Yes=works full-/part-time; self-employed; works in family business; on vacation from paid employment/No=does not work for pay; only does volunteer work; usually works, but has been laid-off or unemployed during the past seven days; retired; seasonal workers not currently working)
Q2 = hours missed due to health problems
Q3 = hours missed for other reasons
Q4 = hours actually worked
Q5 = degree health affected productivity while working
Q6 = degree health affected regular activities

Scores:

Multiply scores by 100 to express in percentages

1. Percent work time missed due to health (Absenteeism): Q2/(Q2 + Q4)
2. Percent impairment while working due to health (Presenteeism): Q5 /10
3. Percent overall work impairment due to health (Work Productivity Loss): Q2/(Q2 + Q4) + [(1 – (Q2/(Q2 + Q4))) x (Q5/10)]
4. Percent activity impairment due to health (Activity Impairment): Q6/10

Absenteeism, presenteeism, work productivity loss, and activity impairment will be calculated at baseline, year 1, and year 2. Change from baseline will be calculated for each score at year 1 and year 2.

3.7.8.4 Patient Global Impressions of Disease Severity (PGI-S) and Change (PGI-C)

The patient global impression of disease severity (PGI-S) is a single item that the patient will complete at the baseline randomization visit that assesses the patient's perception of the severity of their diabetes at the time of the randomization visit. The patient global assessment of change (PGI-C) is a single item that the patient will complete at the dedicated year 1 and year 2 study visits that assesses their perception of their current disease status relative to the way it was before they started the study drug.

3.7.8.5 Clinician Global Impressions of Disease Severity (CGI-S) and Change (CGI-C)

The clinician global impression of disease severity (CGI-S) is a single item that the clinician will complete after the patient's baseline randomization visit that assesses the clinician's perception of the patient's severity of their diabetes at the randomization visit. The clinician global assessment of change (CGI-C) is a single item that the clinician will complete at the patient's

dedicated year 1 and year 2 study visits that assesses the clinician's perception of the patient's current disease status relative to the way it was before they started the study drug.

3.7.9 Healthcare Resource Utilization (HCRU)

All-cause and T2DM-related HCRU will be defined from baseline to year 2. These data will come from the HIRD for [redacted] patients or requested from a set of other health insurers for non-[redacted] patients. HCRU will include inpatient admissions, ER encounters, physician office visits (OV), other outpatient encounters, and pharmacy prescription fills.

All medical claims (inpatient admissions, ER visits, physician OVs, or other outpatient encounters) will be included in the "all-cause" analysis. All pharmacy claims will be included in the "all-cause" HCRU. Inpatient admissions, ER visits, physician OVs, or other outpatient encounters will be considered to be "T2DM-related" if the associated medical claim(s) has an International Classification of Diseases, Tenth Revision (ICD-10) T2DM diagnosis code of E11.% in the primary position or secondary position. Pharmacy claims for antidiabetic medications will be included in the "T2DM-related" HCRU.

- All-cause inpatient admission is defined by a claim or claims with inpatient place of service. Facility and provider claims are rolled into episodes with unique admission and discharge dates. ER visit claims on one day that are directly followed by an inpatient admission are considered as one inpatient admission and all HCRU from the encounter is combined.
- T2MD-related inpatient admission is defined by a claim or claims with inpatient place of service and primary or secondary diagnosis of T2DM. Facility and provider claims are rolled into episodes with unique admission and discharge dates. ER visit claims on one day that are directly followed by an inpatient admission with a primary or secondary diagnosis of T2DM (regardless of the diagnoses associated with the ER visit claim) are considered as one T2MD-related inpatient admission and all HCRU from the encounter is combined.
- All-cause ER visit is defined by a medical claim with ER place of service.
- T2DM-related ER visit is defined by a medical claim with ER place of service and primary or secondary diagnosis of T2DM.
- All-cause outpatient physician OV is defined by a medical claim with outpatient place of service and an evaluation and management (E&M) code that indicates physician OV.

- T2DM-related outpatient physician OV is defined by a medical claim with outpatient place of service and an E&M code that indicates physician OV with a primary or secondary diagnosis of T2DM.
- All-cause other outpatient service is defined by a medical claim with outpatient place of service and code other than E&M, overall and by the following categories:
    - Tests – Lab
    - Imaging
    - Procedures
    - Occupational Therapy (OT)/Speech
    - Medication and Related Services
    - Durable Medical Equipment
    - Physician Other Services
    - Tests – Other
    - Other
- T2DM-related other outpatient service is defined by a medical claim with outpatient place of service, with a primary or secondary diagnosis of T2DM, and code other than E&M, overall and by the following categories:
    - Tests – Lab
    - Imaging
    - Procedures
    - OT/Speech
    - Medication and Related Services
    - Durable Medical Equipment
    - Physician Other Services
    - Tests – Other
    - Other
- All-cause pharmacy prescription fill is defined by a pharmacy claim.
- T2DM-related pharmacy prescription fill is defined by a pharmacy claim for an antidiabetic medication.

The length of stay (LOS) for individual all-cause and T2DM-related inpatient admissions is defined as:

LOS (days) = Discharge Date – Admission Date

The cumulative LOS will be the sum of all all-cause inpatient admissions and all T2DM-related inpatient admissions a patient experiences from baseline to year 2.

3.8 Examination of Subgroups

Data will be presented overall and by treatment group. Additionally, subgroup analysis under the primary estimand (ITT) of the primary endpoint will be conducted on the following subgroups:

- Age (<55 years versus ≥55 years)
- Gender
- Baseline HbA1c (<8.0% versus ≥8.0%)
- Baseline body mass index (BMI) (<30 kg/m^2, ≥30 kg/m^2)
- Duration of T2DM diagnosis at baseline (<5 years versus ≥5 years)

4.0 STATISTICAL ANALYSES

4.1 Overview

The primary analysis for this study is a year 1 analysis. Once data collection for year 1 has completed, a database lock and year 1 analysis will be performed. The year 1 analysis will not be integrated with HCRU and will be limited to year 1 endpoints derived from eCRF data, including PRO data. The year 1 analysis results will be presented in a clinical study report and an internal results meeting. To maintain study integrity for the remaining study period, data from year 1 will be used for limited and confidential communications while complying with public disclosure requirements. All other analyses will be conducted following a second database lock once data collection for the entire study is complete.

No changes to the study design will be made based on the results of the year 1 analysis.

Descriptive summaries will be produced for both the primary (ITT) and secondary ("if all patients adhered") estimands, with the exception of antidiabetic treatment patterns which will be based on the primary estimand only. All statistical analyses will be conducted under both the primary estimand (ITT) and secondary estimand ("if all patients had adhered"), with the exception of the adherence and persistence measures, which will be based on the primary estimand only.

4.2 Patient Disposition and Accountability

Patient disposition will be descriptively summarized for all randomized patients, including the number and percentage patients in the FAS and safety populations, patients who complete the dedicated year 2 study visit, and the primary reason for not completing the dedicated year 2 study visit, and the primary reason for not completing the dedicated year 2 study visit for patients who terminated participation in the study early. Summary metrics for number of screen failures by study site will also be presented.

4.3 Demographics and Baseline Clinical Characteristics

Demographic and baseline clinical characteristics will be descriptively summarized for the FAS overall and by treatment group. For baseline comorbid conditions, the number and percentage of patients with at least one of the comorbid conditions collected in the study will also be summarized along with the relative frequencies of type of comorbid condition. The same approach will be followed for concomitant cardiovascular medications. In addition to eCRF data, type of insurance at randomization (obtained from claims) and geographic region (from study site location) will be summarized.

4.4 Protocol Deviations

Major protocol deviations will be summarized by type (informed consent, randomization error, inclusion/exclusion criteria error, unreported serious adverse events (SAEs), procedure not per protocol, study visit not per protocol, other).

4.5 Missing Value Imputation Methodology

Missing endpoint data will be imputed for HbA1c, weight, SBP, and DBP. The data of patients who, in violation of the protocol, initiate any formulation of semaglutide in the SOC treatment group, will be censored analytically following initiation of semaglutide. The patient will remain in the analysis, but data following their initiation of semaglutide will be censored and treated as missing.

Prior to analysis of binary endpoints derived by dichotomizing HbA1c and body weight, missing data will be imputed by multiple imputation on the continuous scale and dichotomized as required for the endpoint. 500 complete data sets will be generated to adequately account for the uncertainty due to missing data.

4.5.1 Primary Estimand Imputation Model

Missing value imputation for the primary estimand will be based on data from all patients with observations at the endpoint visit (year 1 for the primary endpoint). Missing endpoint data will be imputed separately by treatment group and based on similar patients with available endpoint data according to study drug treatment status, i.e., according to whether these patients are on or off study drug at the endpoint visit. Data will be imputed based on the assumption that, within treatment groups, patients with missing endpoint data will behave like patients with the same study drug treatment status at endpoint as the missing patients' last registered treatment status prior to the missing endpoint. For example, for patients with missing data who have discontinued treatment with study drug, data will be imputed based on the assumption that these patients will behave like patients with available data who are no longer receiving study drug.

Technically, the imputation model will be an analysis of covariance (ANCOVA) for the endpoint data. The ANCOVA will include the baseline value of the imputed variable (HbA1c, weight, SBP, or DBP), diabetes duration, age, and sex as independent variables. After this model has been used to predict missing values, each of the now 500 complete data sets will be analyzed as described for the primary analysis in section 4.6.1. Finally, the multiple analysis results will be combined using Rubin's rule. [6] For the OR, the results will be combined on the logarithm scale.

This process will be repeated for year 2 endpoints.

4.5.2 Secondary Estimand Imputation Model

Missing value imputation for the secondary estimand will be based only on data from the subset of patients who are receiving study drug (on study drug) at the endpoint visit (year 1 for the primary endpoint) in order to estimate the treatment effect if all patients had continued treatment. The endpoint data of patients who are not receiving study drug at the endpoint visit (off study drug) will be censored analytically and imputed together with missing data. Collectively, missing and censored data will be imputed separately by treatment group based on all patients who are on study drug at the endpoint visit. Data will be imputed based on the assumption that patients with missing data will behave like patients with available data who are on study drug.

The technical aspects of missing data imputation based on multiple imputation, the statistical analysis of the multiple complete data sets, and combination of the multiple results will be the same as the ones described for the primary estimand.

This process will be repeated for year 2 endpoints.

4.6 Effectiveness Analysis

4.6.1 Statistical Analysis for the Primary Estimand

The primary effectiveness analysis is the primary estimand of the primary endpoint, HbA1c <7.0% at year 1, and confirmatory secondary endpoints of change in HbA1c (%-point) from baseline to year 1, HbA1c <7.0% at year 2, and change in HbA1c (%-point) at year 2. This analysis will be based on the FAS with missing data imputation as described in the primary estimand imputation model.

Year 1 Analysis

The primary endpoint, HbA1c < 7.0% at year 1, will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) will be presented.

The confirmatory secondary continuous endpoint of change in HbA1c (%-point) at year 1 will be analyzed using ANCOVA that will include the treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide s.c. versus SOC) will be presented.

The estimated treatment effect from each of these analyses will be complemented with associated 95% CI and two-sided p-value for testing the null-hypothesis of no difference.

Superiority for the primary endpoint, HbA1c <7.0% at year 1, will be considered established if the OR 95% confidence interval (CI) is greater than 1, or similarly if the two-sided p-value is significant on a 5% level and the treatment OR is in favor of semaglutide.

If the hierarchical testing scheme allows, superiority for change in HbA1c from baseline to endpoint visit will be considered established if the 95% CI for the estimated treatment difference is smaller than 0, or similarly if the two-sided p-value is significant on a 5% level and the treatment difference is in favor of semaglutide.

Year 2 Analysis

The Year 1 analysis described above will be repeated for the confirmatory secondary endpoints of HbA1c <7.0% at year 2 and change in HbA1c (%-point) at year 2. If the hierarchical testing scheme allows, superiority will be tested first for HbA1c <7.0% at year 2 followed by change in HbA1c (%-point) at year 2 as described for the year 1 counterparts above.

4.6.2 Statistical Analysis for the Secondary Estimand

The primary effectiveness analysis will be repeated with the secondary estimand. This analysis will be based on the FAS with missing data imputation as described in the secondary estimand imputation model. With the exception of the data used and the imputation of missing data, which follows the secondary estimand, the statistical analysis is the same as the primary effectiveness analysis.

The binary endpoints of HbA1c <7.0% at endpoint visit will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) and 95% CI will be presented.

The continuous endpoints of change in HbA1c (%-point) from baseline to endpoint visit will be analyzed using ANCOVA with treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide s.c. versus SOC) will be presented along with associated 95% CI.

4.6.3 Additional Analyses of Glycemic Control

4.6.3.1 Descriptive Summaries of HbA1c

Observed HbA1c values will be descriptively summarized at baseline, year 1, and year 2. Additionally, the following categories of observed HbA1c will be descriptively summarized at baseline, year 1, and year 2:

- <7.0, 7.0-8.0, 8.1-9.0, 9.1-10.0, >10.0%
- <8.0 versus ≥8.0%
- <9.0 versus ≥9.0%

4.6.3.2 Supportive Analyses of Glycemic Control

The following supportive endpoints of glycemic control are binary endpoints. These analyses will compare the proportion of patients achieving the secondary endpoints related to glycemic control between semaglutide and SOC treatment arms via a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) will be presented.

- Individualized HbA1c target attained at year 1 (yes/no)
- HbA1c <7.0% (53mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 1 (yes/no)
- HbA1c target attainment per Healthcare Effectiveness Data and Information Set (HEDIS) criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 1 (yes/no)
- Individualized HbA1c target attained at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 2 (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 1 (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 2 (yes/no)

- HbA1c target attainment per HEDIS criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <7.0% (53 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)

4.6.4 Supplementary Analyses

The HbA1c analyses described for the primary and secondary effectiveness analyses described above will be complemented with two complete case analyses based on the subset of patients without missing endpoint data, i.e., patients that do not have their endpoint imputed in the two analyses used for the primary and secondary estimand.

4.6.4.1 Complete Case Analysis

This analysis will be based only on patients with available measurements at the dedicated study visit including measurements irrespective of whether patients discontinued study drug or not.

The binary endpoints of HbA1c < 7.0% at year 1 and HbA1c < 7.0% at year 2 will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) will be presented along with associated 95% CI.

The continuous endpoints of change in HbA1c (%-point) at year 1 and change in HbA1c (%-point) at year 2 will be analyzed using ANCOVA with treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide s.c. versus SOC) will be presented along with associated 95% CI.

4.6.4.2 Complete Case On-Study Drug Analysis

The analysis will be based only on patients with available measurements at the dedicated study visit including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed. The analysis will use the same analysis models as described for complete case analysis above.

4.7 PRO Analysis

PROs and ClinROs will be measured with the instruments described in section 3.7.8. Analysis of these measures will address the secondary objective of this study to compare semaglutide s.c. versus SOC in the study's patient population as is relates to PROs and ClinROs, i.e., treatment satisfaction, generic health outcomes, work productivity, and patient and physician global

assessment measures, over one and two year observation periods. PRO analysis will descriptively summarize these measures at baseline, year 1 and year 2, as well as compare semaglutide s.c. versus SOC for change from baseline to year 1 and year 2.

No imputation of missing PRO measures is planned. Therefore, the primary estimand for PRO analysis will be based on the FAS with endpoint data. The secondary estimand for PRO analysis will be based only on patients with endpoint data including only measurements for patients still on study drug. In the secondary estimand, patients who discontinue study drug will be censored and missing data will not be imputed.

Data regarding all PROs will be reported for the entire sample by treatment group. Percentages of completed PROs will be reported as well. Descriptive statistics will be reported for aggregate scores for all appropriate measures by total score and subscale score(s) if applicable. Semaglutide and SOC treatment groups will be compared by independent t-tests for PRO measures in which the data are normally distributed or the Mann-Whitney U-test for PRO measures in which the data are not normally distributed.

- Diabetes Treatment Satisfaction Questionnaire, change version (DTSQc), Total treatment satisfaction score measured at year 1
- DTSQc, Total treatment satisfaction score measured at year 2
- Change from baseline in Short Form 12-Item Version 2 Survey (SF-12 v2), Physical component summary (PCS-12) score at year 1
- Change from baseline in SF-12 v2, PCS-12 score at year 2
- Change from baseline in SF-12 v2, Mental component summary (MCS-12) score at year 1
- Change from baseline in SF-12 v2, MCS-12 score at year 2
- Change from baseline in Work Productivity and Activity Impairment, General Health questionnaire (WPAI-GH) Absenteeism (work time missed) score at year 1
- Change from baseline in WPAI-GH Absenteeism (work time missed) score at year 2
- Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 1
- Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 2
- Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 1
- Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 2

- Change from baseline in WPAI-GH Activity Impairment score at year 1
- Change from baseline in WPAI-GH Activity Impairment score at year 2

In addition to the above endpoints, PGI-S and CGI-S will be descriptively summarized at baseline overall and by treatment group. PGI-C and CGI-C will be descriptively summarized at year 1 and year 2 overall and by treatment group.

During later analysis on the primary endpoints, PRO measures can be used in two ways:

- Correlational analysis of the PRO measures with the dependent variable. These correlations will include the appropriate measure of Pearson's 'r', Spearman's rho, Kendall's tau statistics or other appropriate measures that might fit the nature of the variable.
- Should the correlational analysis from step 1 find that patient characteristics are linked to primary analysis endpoints, multivariate analysis or Logistic Regression will be conducted to determine if there are any patient related variables from the PROs that predict the effectiveness of therapy in combination with other clinical variables. Data reported will include model specifics (H-L Goodness of fit; R2, beta coefficients for each variable as well as ORs with 95% CIs.)

4.8 Concomitant Medications

The number and percentages patients reporting pre-specified concomitant cardiovascular medications during year 1 and during year 2 will be summarized for the population overall and by treatment group.

4.9 Other Analyses

The following analyses further support the primary objective to compare semaglutide s.c. versus SOC in glycemic control. They will address the secondary objectives of this study to compare semaglutide s.c. versus SOC in the study's patient population over one and two year observation periods as is relates to body weight loss, hypoglycemia, HCRU, adherence and persistence to treatment, antidiabetic medication treatment patterns, and safety.

Except where otherwise noted, the following analyses will be conducted for the primary estimand and the secondary estimand. For endpoints that do not include HbA1c, weight, SBP, or DBP, no missing data will be imputed. Therefore, the primary estimand for these endpoints will be based on the FAS with endpoint data. The secondary estimand for these endpoints will be based only on patients with endpoint data including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed.

4.9.1 Weight Loss

Change in patient weight from baseline to endpoint will be calculated in absolute change (pounds) and relative change (percentage). Mean change in weight will be compared between semaglutide and SOC treatment groups following the ANCOVA model described for continuous endpoints in the primary (4.6.1) and secondary estimand (4.6.2). This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

- Change in body weight (%) from baseline to year 1
- Change in body weight (lb) from baseline to year 1
- Change in body weight (%) from baseline to year 2
- Change in body weight (lb) from baseline to year 2

4.9.2 Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP)

Absolute change in patient SBP and DBP from baseline to year 1 and baseline to year 2 will be calculated. Mean change in SBP and DBP will be compared between the semaglutide and SOC treatment groups following the ANCOVA model described for continuous endpoints in the primary (4.6.1) and secondary estimand (4.6.2). This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

4.9.3 Hypoglycemia

The total number of hypoglycemic episodes leading to an inpatient admission or ER encounter will be summed per patient from baseline to year 2 and compared by semaglutide and SOC treatment groups utilizing a negative binomial model with fixed effect of treatment and baseline HbA1c as a covariate. Exploratory analysis will assess the impact of SOC study drug (defined in section 3.7.4.2) on the rate of hypoglycemia.

Additionally, the frequency and percentage of patients reporting hypoglycemic episodes leading to an inpatient admission or ER encounter from baseline to year 1 and from baseline to year 2 will be reported and compared between treatment groups via chi-square test.

Hypoglycemic episodes that are considered SAEs will also be included in the safety analysis (section 5.4).

4.9.4 Healthcare Resource Utilization (HCRU)

HCRU analyses will compare all-cause and diabetes-related HCRU (inpatient admissions, ER encounters, physician OV, other outpatient encounters, and pharmacy utilization) as defined in section 3.7.9 from claims data by semaglutide and SOC treatment arm from baseline to year 2. The number of each type of encounter (all-cause and T2DM-related: inpatient admission, ER

encounter, physician OV, other outpatient encounter (overall and by specific categories listed in section 3.7.9), pharmacy prescription fills) a patient experiences from baseline to year 2 will be summed.

HCRU will be summarized by setting (i.e., inpatient, ER, OV, outpatient, pharmacy) and cause (i.e., all-cause versus diabetes-related), for the study population overall and by treatment group. The occurrence (yes/no) and number of each type of encounter or prescription fill will be reported. LOS will be reported per admission and cumulative (baseline to year 2). LOS for inpatient admissions (days) per inpatient admission will be an encounter level analysis of the average inpatient admission LOS based on the total number of inpatient admissions; therefore, patients may contribute more than once. Statistics for cumulative LOS for inpatient admissions (days) will only be produced for patients with at least one inpatient admission.

Treatment group comparisons will utilize the Student t-test for continuous variables and chi-squared test (or Fisher's exact test if outcome count less than 5) for categorical variables.

4.9.5 Composite Endpoints

The following composite endpoints will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c and baseline weight as covariates. This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 1 (yes/no)
- Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 2 (yes/no)
- Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 2 (yes/no)

4.9.6 Antidiabetic Treatment Patterns

4.9.6.1 Study Drug Regimen

Study drug regimen will be descriptively summarized by treatment arm.

The SOC group's randomized treatment regimen will be summarized by medication and class for study drug, i.e., the first antidiabetic medication prescribed for treatment intensification following randomization.

For the semaglutide group, relative frequency by dose category (0.25, 0.5, 1.0 mg) will be calculated at randomization, year 1, and year 2. Additionally, the relative frequency by semaglutide dose category will be calculated for most common, highest, and final as defined in section 3.7.4.1.

Treatment intensification and treatment change (definition 3.7.4.3) will also be descriptively summarized as follows:

- Treatment intensification from baseline to year 1 (yes/no)
- Treatment intensification from baseline to year 2 (yes/no)
- Treatment change from baseline to year 1 (yes/no)
- Treatment change from baseline to year 2 (yes/no)

4.9.6.2 Adherence and Persistence to Treatment

Adherence and persistence to the study drug will be calculated and compared between semaglutide and SOC treatment groups. This analysis will be conducted for the primary estimand only.

Study drug medication adherence will be summarized by the MPR (%) using the claims data. The MPR (%) for the first year of the study and the MPR (%) for the two year study period will be compared between treatment arms using the test of independent proportions.

Medication persistence, as calculated using the eCRF data, will be summarized by:

- Time to first study drug discontinuation during 2 years (day)
- Time to first treatment intensification (add-on) or change (switch) after randomization during 2 years (day)

Medication persistence measures will be compared between semaglutide and SOC treatment groups via the Cox proportional hazards model (with no covariates) to estimate the hazard ratio (HR) (semaglutide s.c. versus SOC) and 95% CI. Patients with no event will be censored at year 2 or last contact if withdrawn or lost to follow-up.

4.9.6.3 Other Antidiabetic Treatments

Antidiabetic treatment patterns will be assessed and compared between semaglutide and SOC treatment groups. This analysis will summarize the number and classification type of antidiabetic medications taken during the study period.

Descriptive statistics will be presented overall and by treatment group for the average number of antidiabetic medications patients report at randomization, year 1 and year 2. Relative frequencies of number of antidiabetic medications will also be summarized for randomization, year 1 and year 2. Relative frequencies by antidiabetic class will also be presented at randomization, year 1 and year 2.

4.10 Study Site Recruitment and Participation

To contextualize the participating study sites within the pool of [redacted] providers, tables summarizing the identification, recruitment, and participation of providers will be presented.

Site identification for this study begins with identifying the pool of potentially eligible patients in the HIRD using very broad claims-based inclusion/exclusion criteria: currently active members with at least one medical claim with an ICD-9 or ICD-10 diagnosis code for T2DM during the identification period, pharmacy carve-in benefits and pharmacy carve-out benefits at any time during the identification period, commercial plan coverage only, ≥18 years old, at least one pharmacy claim for metformin in the past 6 months, and no pharmacy claims for other antidiabetic medications within the past 30 days. These patients are then mapped back to the provider sites with which they had diabetes-related office visits during the identification period. From this process, a list of providers with number of potentially eligible patients is generated. The overall pool of [redacted] providers serving the target study population will be summarized with a table presenting patient densities by number of providers.

A flow chart summarizing each step of the process will be produced to describe site recruitment and participation:

1) Total number of providers with at least one potentially eligible patient identified.
2) Total number of providers approached (I.e., providers who met predetermined threshold (x number of patients) for outreach.)
3) Disposition of providers approached
    a. Interested, invited to participate
    b. Interested, rejected from participation
    c. Not Interested
    d. No Response

A table summarizing characteristics of providers (geographic region, provider specialty, and claims-based patient density) by disposition status will be produced.

4.11 Evaluation of the Study Population

To evaluate the generalizability of the study results, an analysis of the study population will be performed. The primary objective of this analysis is to evaluate the external validity of the study and to understand how the study population fits into the larger T2DM populations within study sites and commercially-insured population as a whole. Two claims-based T2DM patient populations will be described and compared to the study population: (1) all research-eligible commercially-insured and Medicare Advantage [redacted] T2DM patients treated with up to 2 oral antidiabetic medication, excluding oral semaglutide, who undergo an antidiabetic treatment intensification and (2) within the study sites, commercially-insured and Medicare Advantage [redacted] T2DM patients treated with up to 2 oral antidiabetic medications, excluding oral semaglutide, who are not enrolled in the study, but who undergo an antidiabetic treatment intensification.

Identification of Claims-Based Comparator Populations

The claims-based comparator populations and analysis will be based on data from the HIRD. Claims based inclusion/exclusion criteria will be imposed on the two comparator populations as best as can be identified via the claims data and available lab result data to reflect the inclusion/criteria of the study population. All patients meeting the inclusion/exclusion criteria for each of the claims-based T2DM populations will be included in the analysis.

- Claims patient identification period = July 2018 until date last patient enrolled. (To align with the dates of the PCT study period).
- Claims index date = Patient's first prescription date associated with antidiabetic treatment intensification during the patient identification period.
- Claims study period = July 2017 until 12 months following date of last patient enrolled. (To allow for 12 months of claims data before and after the identified antidiabetic treatment intensification.)

Inclusion Criteria for Population #1

- ≥ 1 medical claim with an ICD-10 diagnosis code for T2DM during patient identification period
- A pharmacy claim(s) for 1 or 2 oral antidiabetic treatments, excluding oral semaglutide, as described in the PCT, during the patient identification period
    - First date of antidiabetic treatment pharmacy claim is index date

- ≥ 12 months medical and pharmacy eligibility prior to index date
- Age ≥ 18 years on index date
- HbA1c >7.0% within 90 days prior to and including index date

Exclusion Criteria for Population #1

- Pharmacy claims for ≥2 oral antidiabetic medications, oral semaglutide, or any injectable antidiabetic medicationfor 30 days prior to index date.
- Randomized into the current PCT.

Inclusion Criteria for Population #2

The inclusion criteria for population #2 will include all of the inclusion criteria as population #1, as well as:

- Receiving antidiabetic treatment at a study site

Exclusion Criteria for Population #2

The exclusion criteria for population #2 is the same as the exclusion criteria for population #1.

All analyses evaluating the study population will be descriptive only and will help to contextualize study results within the T2DM population broadly. All analyses will be based on the overall population only; analyses by treatment group will not be conducted.

The analysis will descriptively compare the overall baseline demographics between the PCT study population and the two comparator populations. Baseline HbA1c (defined as closest HbAlc up to and including 90 days prior to randomization (for PCT population) or index date (for comparators)) will also be compared.

Observed HCRU (all-cause and T2DM-related; inpatient, ER, outpatient physician OV, and other outpatient services, pharmacy utilizations) during the 12 month period prior to randomization (for study population) or prior to index date (for comparators) will also be compared between the PCT study population and the two comparator populations. The PCT study population for this comparison will be limited to study patients with 12 months of health plan eligibility prior to randomization.

Additionally, treatment patterns of non-enrolled T2DM patients within the practices from which the study patients are recruited will be evaluated to identify any relevant patterns of care suggesting channeling of certain types of patients away from the study. This analysis be based on

the subset of patients with 12 months post-index data medical and pharmacy eligibility and will present the relative frequency of antidiabetic medications by class during the 12 months following randomization or index date for the PCT study population and comparator population #2.

5.0 SAFETY

5.1 AEs

For the purpose of this study, AEs that do not meet the definition of an SAE (protocol section 7.2) will only be required to be collected in the eCRF if they lead to study drug discontinuation.

5.2 SAEs

All AEs meeting the definition of an SAE (protocol section 7.2) will be collected in the eCRF.

5.3 Pregnancy

Any abnormal pregnancy outcome (e.g., spontaneous miscarriage, fetal death, congenital anomaly/birth defect, etc.) is considered an SAE. For the purposes of this study, any pregnancies in participating female patients will be reported, along with pregnancy outcome and any AEs or SAEs observed in the fetus or newborn until 1 month of age.

5.4 Analysis

No formal safety analyses are planned for this study. SAEs and AEs leading to study drug discontinuation will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and descriptively summarized by System Organ Class and Preferred Term (PT). Pregnancies occurring during the study will be also be descriptively summarized.

6.0 Changes from Protocol

The protocol specified conducting PRO Analysis (Protocol Section 6.6) and Other Analyses (Protocol Section 6.8) as described for the primary estimand, secondary estimand, and supplementary analysis (complete case and complete case on-study drug). However, missing data imputation will only be completed for HbA1c, weight, SBP, and DBP. Therefore, for endpoints that do not include HbA1c, weight, SBP, or DBP data, the primary estimand will be based on the FAS with endpoint data. The secondary estimand for these endpoints will be based only on patients with endpoint data including only measurements for patients still on study drug.

Patients who discontinue study drug will be censored and missing data will not be imputed. Supplementary analyses will not be completed for these endpoints.

7.0 SAP Change Log

Summary of SAP Changes since First Patient First Visit

Number	Section/Page	Change	Rationale
1	Global Change, Study Population	Allow enrollment of patients treated with either 1 or 2 oral antidiabetic medications	Expand study population from patients treated with metformin monotherapy to patients treated with either 1 or 2 oral antidiabetic medications to facilitate enrollment.
2	Global Change, Study Population	Allow enrollment of patients with any commercial or Medicare insurance with pharmacy benefits.	Expand study population to facilitate enrollment.
3	Global Change	~~July~~**May ~~5~~20, ~~2018~~2020** Version ~~1~~2.0	Update protocol date and version.
4	Study Approvals, p. 2	**Name:** [redacted] **Title:** [redacted] Research Biostatistics ~~**Name:**~~ [redacted] ~~**Title:**~~ [redacted] ~~Scientist~~	Update to [redacted] study statistician.
5	Study Approvals, p. 2	**Name:** [redacted] **Title:** [redacted] ~~**Name:**~~ [redacted] ~~**Title:**~~ [redacted]	Update to Novo Nordisk study statistician.
6	Global Change, Study Treatments	Semaglutide s.c.	Clarify formulation of semaglutide allowed for study treatment is subcutaneous only.
7	Global Change	routine diabetic care visits	Clarify data collection between study visits for sites.
8	Global Change, Study Treatments	…semaglutide s.c. as compared to SOC (excluding any semaglutide).	No formulation of semaglutide is allowed in SOC group. Clarify that main study comparison excludes any formulation of semaglutide from the SOC group.
9	Section 1.0, Introduction	This is a 2-year, multi-center, randomized, open label, parallel group, active comparator	Update study design to reflect expanded study population.

		pragmatic clinical trial (PCT) comparing semaglutide s.c. versus standard of care (SOC) when added to ~~metformin monotherapy~~up to 2 oral antidiabetic medications as treatment intensification among adult type 2 diabetes mellitus (T2DM) patients in the course of routine clinical practice.	
10	Section 2.1, Primary Objective	The primary objective is to demonstrate superior long term effects of treatment with semaglutide s.c. compared to SOC each added to ~~metformin~~ up to 2 oral antidiabetic medications on glycemic control when used as intensification in routine clinical practice in adult patients with T2DM.	Update primary objective to reflect expanded study population.
11	Section 2.2, Secondary Objectives	The secondary objectives are to compare the long term effect of semaglutide s.c. versus SOC each added to ~~metformin~~ up to 2 oral antidiabetic medications and used as intensification in routine clinical practice in adult patients with T2DM with regards to:	Update secondary objectives to reflect expanded study population.
12	Section 3.3, Evaluation Schedule and Definitions	Year 2 endpoint data will be derived from the dedicated year ~~1~~ 2 study visit if the dedicated study visit is ±10 weeks of 104 weeks post-randomization.	Correction
13	Section 3.4, Confirmatory Endpoint and Hypothesis	This study is designed to have 90% power to ~~jointly~~ confirm superiority of the primary endpoint and 85% power to also confirm superiority of the first ~~the above three~~ confirmatory secondary endpoints based on primary estimand analyses.	Update to power calculations.
14	Section 3.5.1 Power and Sample Size for Primary Objective	In line with the primary ITT estimand, the assumptions for the proportion of patients with HbA1c < 7.0% at year 1 and year 2, and for the change from baseline in HbA1c are based on the claims/laboratory results data within the [redacted] population for all patients initiating treatment intensification regardless of whether patients adhered to this treatment.	Update to reflect protocol.
15	Section 3.5.1 Power and Sample Size for Primary Objective	The sample size calculation aimed for 90% power for ~~simultaneously~~ confirming superiority of semaglutide s.c. versus SOC on ~~all of~~ the primary endpoint and 85% power to also confirm superiority of the first	Update to power calculations.

		confirmatory secondary endpoint based on an analysis of the primary estimand for each of the endpoints~~the secondary confirmatory endpoints at both year 1 and year 2 (4 hypotheses in total)~~.	
16	Section 3.5.1 Power and Sample Size for Primary Objective	When accounting for missing data, randomizing ~~2250~~ 1387 patients will contribute ~~1687~~ 1040 patients for the year 1 analyses and ~~1260~~ 780 patients for the year 2 analyses, achieving a total power of ~~90%~~ 85% for confirming the two year 1 confirmatory hypotheses. The joint power for confirming all 4 confirmatory hypotheses is 58%. The corresponding marginal powers for presence of HbA1c < 7.0% at year 1, change in HbA1c to 1 year, presence of HbA1c < 7.0% at year 2, and change in HbA1c to 2 years are ~~99%~~90%, ~~99%~~94%, ~~95%~~80%, and ~~97%~~86% respectively.	Update to planned enrollment and power calculations.
17	Section 3.7.4.2 Standard of Care (SOC) Study Drug	For the SOC treatment group, study drug is defined as the drug class of the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization (excluding any formulation of semaglutide).	Semaglutide in any formulation is not allowed during the study period for the SOC group.
18	Section 3.7.4.3 Treatment Intensification or Change	Treatment intensification is defined as initiation of an antidiabetic medication (besides ~~metformin~~the up to 2 oral antidiabetic medication(s) taken prior to randomization) in addition to study drug, or initiation of more than one antidiabetic medication after discontinuation of study drug.	Update definition to reflect expanded study population.
19	Section 3.7.4.3 Treatment Intensification or Change	Treatment change is defined as discontinuation of study drug and initiation of another antidiabetic medication (besides ~~metformin~~the up to 2 oral antidiabetic medication(s) taken prior to randomization).	Update definition to reflect expanded study population.
20	Section 3.7.9 Healthcare Resource Utilization (HCRU)	These data will come from the HIRD for [redacted] patients or requested from a set of other health insurers for non-[redacted] patients.	Updated definition to reflect expanded study population and collection of claims data.
21	Section 4.5	Missing endpoint data will be imputed for HbA1c, weight, SBP, and DBP. The data of	Semaglutide in any formulation is not allowed

	Missing Value Imputation Methodology	patients who, in violation of the protocol, initiate any formulation of semaglutide in the SOC treatment group, will be censored analytically following initiation of semaglutide. The patient will remain in the analysis, but data following their initiation of semaglutide will be censored and treated as missing.	during the study period for the SOC group.
22	Section 4.9.4 Healthcare Resource Utilization (HCRU)	HCRU analyses will compare all-cause and diabetes-related HCRU (inpatient admissions, ER encounters, physician OV, other outpatient encounters, and pharmacy utilization) as defined in section 3.7.9 from ~~the HIRD~~claims data by semaglutide and SOC treatment arm from baseline to year 2.	Update to reflect expanded study population and collection of claims data.
23	Section 4.9.6.3 Other Antidiabetic Treatments	~~Background Medication Metformin dose will be descriptively summarized overall and by treatment group at randomization, year 1, and year 2. Other Antidiabetic Treatments~~ … Descriptive statistics will be presented overall and by treatment group for the average number of antidiabetic medications patients report at randomization, year 1 and year 2. Relative frequencies of number of antidiabetic medications will also be summarized for randomization, year 1 and year 2. Relative frequencies by antidiabetic class will also be presented at randomization, year 1 and year 2.	Update descriptive summary of other antidiabetic treatments to reflect expanded study population.
24	Section 4.11 Evaluation of the Study Population	The primary objective of this analysis is to evaluate the external validity of the study and to understand how the study population fits into the larger T2DM populations within study sites and [redacted] commercially-insured population as a whole. Two claims-based T2DM patient populations will be described and compared to the study population: (1) all research-eligible commercially-insured and Medicare Advantage [redacted] T2DM patients treated with ~~metformin~~up to 2 oral antidiabetic medication, excluding oral semaglutide, who undergo an antidiabetic treatment intensification and (2) within the study sites, commercially-insured and	Update to reflect expanded study population.

		Medicare Advantage [redacted] T2DM patients treated with ~~metformin~~ up to 2 oral antidiabetic medications, excluding oral semaglutide, who are not enrolled in the study, but who undergo an antidiabetic treatment intensification.	
25	Section 4.11 Evaluation of the Study Population	Inclusion Criteria for Population #1 • ≥ 1 medical claim with an ICD-10 diagnosis code for T2DM during patient identification period • ~~≥ 1~~ A pharmacy claim(s) for 1 or 2 oral antidiabetic treatments, excluding oral semaglutide, ~~in addition to metformin,~~ as described in the PCT, during the patient identification period o First date of antidiabetic treatment pharmacy claim is index date • ≥ 12 months medical and pharmacy eligibility prior to index date ~~• ≥ 1 pharmacy claim for metformin in 6 months prior to index date~~ • Age ≥ 18 years on index date • HbA1c >7.0% within 90 days prior to and including index date	Update to reflect expanded study population.
26	Section 4.11 Evaluation of the Study Population	Exclusion Criteria for Population #1 • ~~≥ 1 pharmacy~~ Pharmacy claims for ≥2 ~~any~~ oral antidiabetic medications, oral semaglutide, or any injectable antidiabetic medication ~~other than metformin~~ for 30 days prior to index date. • Randomized into the current PCT.	Update to reflect expanded study population.

8.0 APPENDIX A: Study Endpoints and Other Variables by Data Source

Endpoint/Variable	Data Source	
	eCRF	Administrative Claims
Study Drug Variables		
On/Off Study Drug	X	
Treatment Intensification	X	
Treatment Change	X	
Antidiabetic treatment patterns	X	
Primary Endpoint		
HbA1c <7.0% at year 1 (yes/no)	X	
Confirmatory Endpoints		
Change in HbA1c (%-point) from baseline to year 1	X	
HbA1c <7.0% at year 2 (yes/no)	X	
Change in HbA1c (%-point) from baseline to year 1	X	
Supportive Secondary Endpoint Assessment		
Individualized HbA1c target attained at year 1 (yes/no)	X	
HbA1c <7.0% (53mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 1 (yes/no)	X	
HbA1c target attainment per Healthcare Effectiveness Data and Information Set (HEDIS) criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 1 (yes/no)	X	X
Change in body weight (lb) from baseline to year 1	X	
Change in body weight (%) from baseline to year 1	X	
Change in systolic blood pressure (SBP; mm Hg) from baseline to year 1	X	

Change in diastolic blood pressure (DBP; mm Hg) from baseline to year 1	X	
Time to first study drug discontinuation during 2 years (day)	X	
Time to first treatment intensification (add-on) or change (switch) after randomization during 2 years (day)	X	
Study drug medication adherence for the first year of the study, as measured by medication possession ratio (MPR) (%)		X
Number of hypoglycemic episodes leading to an inpatient admission or emergency room (ER) encounter from baseline to year 2	X	
Diabetes Treatment Satisfaction Questionnaire, change version (DTSQc), Total treatment satisfaction score measured at year 1	X	
DTSQc, Total treatment satisfaction score measured at year 2	X	
Change from baseline in Short Form 12-Item Version 2 Survey (SF-12 v2), Physical component summary (PCS-12) score at year 1	X	
Change from baseline in SF-12 v2, PCS-12 score at year 2	X	
Change from baseline in SF-12 v2, Mental component summary (MCS-12) score at year 1	X	
Change from baseline in SF-12 v2, MCS-12 score at year 2	X	
Change from baseline in Work Productivity and Activity Impairment, General Health questionnaire (WPAI-GH) Absenteeism (work time missed) score at year 1	X	
Change from baseline in WPAI-GH Absenteeism (work time missed) score at year 2	X	
Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 1	X	
Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 2	X	
Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 1	X	
Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 2	X	
Change from baseline in WPAI-GH Activity Impairment score at year 1	X	
Change from baseline in WPAI-GH Activity Impairment score at year 2	X	

All cause HCRU from baseline to year 2		
Number of inpatient admissions		X
LOS for inpatient admissions (days) per inpatient admission		X
Cumulative length of stay for inpatient admissions (days)		X
Number of ER encounters		X
Number of physician OV		X
Number of other outpatient encounters (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Number of medications		X
Occurrence of inpatient admission (yes/no)		X
Occurrence of ER encounter (yes/no)		X
Occurrence of physician OV (yes/no)		X
Occurrence of other outpatient encounter (yes/no) (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Diabetes related HCRU from baseline to year 2		
Number of diabetes related inpatient admissions		X
LOS for diabetes related inpatient admissions (days) per diabetes related inpatient admission		X
Cumulative length of stay for diabetes related inpatient admissions (days)		X
Number of diabetes related ER encounters		X
Number of diabetes related physician OV		X
Number of diabetes related other outpatient encounters (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X

Number of diabetes related medications		X
Occurrence of diabetes related inpatient admission (yes/no)		X
Occurrence of diabetes related ER encounter (yes/no)		X
Occurrence of diabetes related physician OV (yes/no)		X
Occurrence of diabetes related outpatient encounter (yes/no) (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Additional Derived Outcome Variables for Supportive Analyses		
Supportive Measures of Glycemic Control		
Individualized HbA1c target attained at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 2 (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 1 (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 2 (yes/no)	X	
HbA1c target attainment per HEDIS criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 2 (yes/no)	X	X
HbA1c <7.0% (53 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)	X	

Body Weight Loss		
Change in body weight (%) from baseline to year 2	X	
Change in body weight (lb) from baseline to year 2	X	
Blood Pressure		
Change in SBP (mm Hg) from baseline to year 2	X	
Change in DPB (mm Hg) from baseline to year 2	X	
Hypoglycemia		
Reported hypoglycemia leading to inpatient admission or ER encounter during year 1 (yes/no)	X	
Reported hypoglycemia leading to inpatient admission or ER encounter during year 2 (yes/no)	X	
Composite Variables		
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 1 (yes/no)	X	
Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 2 (yes/no)	X	
Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 2 (yes/no)	X	

Adherence to Treatment		
Study drug medication adherence for the two years of the study, as measured by the medication possession ratio (MPR) (%)		X

9.0 APPENDIX B: 2018 HEDIS Criteria

Denominator

The eligible population.

Note: *The eligible population for the HbA1c Control <7% for a Selected Population indicator is reported after required exclusions are applied.*

Required exclusions for HbA1c Control <7% for a Selected Population indicator

Exclude members who meet any of the following criteria:

- 65 years of age and older as of December 31 of the measurement year.
- *CABG.* Members who had CABG (CABG Value Set) in any setting during the measurement year or the year prior to the measurement year.
- *PCI*. Members who had PCI (PCI Value Set), in any setting, during the measurement year or the year prior to the measurement year.
- *IVD.* Members who met at least one of the following criteria during both the measurement year and the year prior to the measurement year. Criteria need not be the same across both years.
  - At least one outpatient visit (Outpatient Value Set) with an IVD diagnosis (IVD Value Set).
  - At least one acute inpatient encounter (Acute Inpatient Value Set) with an IVD diagnosis (IVD Value Set).
- *Thoracic aortic aneurysm.* Members who met at least one of the following criteria during both the measurement year and the year prior to the measurement year. Criteria need not be the same across both years.
  - At least one outpatient visit (Outpatient Value Set), with a diagnosis of thoracic aortic aneurysm (Thoracic Aortic Aneurysm Value Set).
  - At least one acute inpatient encounter (Acute Inpatient Value Set), with a diagnosis of thoracic aortic aneurysm (Thoracic Aortic Aneurysm Value Set).
- Any of the following, in any setting, any time during the member's history through December 31 of the measurement year.
  - *Chronic heart failure*. A diagnosis of chronic heart failure (Chronic Heart Failure Value Set).
  - *Prior MI*. A diagnosis of MI (MI Value Set).
  - *ESRD*. ESRD (ESRD Value Set; ESRD Obsolete Value Set).
  - *Chronic kidney disease (stage 4)*. Stage 4 chronic kidney disease (CKD Stage 4 Value Set).
  - *Dementia*. A diagnosis of dementia (Dementia Value Set; Frontotemporal Dementia Value Set).
  - *Blindness*. A diagnosis of blindness (Blindness Value Set).
  - *Amputation (lower extremity)*. Lower extremity amputation (Lower Extremity Amputation Value Set).

10.0 REFERENCES

[1] C. Bradley, "Diabetes Treatment Satistcation Questionnaire: DTSQ," in *Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Practice*, C. Bradley, Ed., Abington, Routledge, 1994, pp. 111-132.

[2] C. Bradley, "The Diabetes Treatment Satisfaction Questionnaire (DTSQ): change version for use alongside status version provides appropriate solution where ceiling effects occur," *Diabetes Care,* vol. 22, no. 3, pp. 530-532, 1999.

[3] J. Ware, M. Kosinski and S. Keller, "A 12-item short-form healch survey: construction of scales and preliminary tests of reliability and validity," *Med Care,* vol. 34, no. 3, pp. 220-233, 1996.

[4] J. Ware, M. Kosinski, D. Turner-Bowker and B. Gandek, How to score version 2 of the SF-12 health survey (with a supplement documenting version 1), Lincoln, RI: QualityMetric, Incorporated, 2002.

[5] M. Reilly, A. Zbrozek and E. Dukes, "The validity and reproducibility of a work productivity and activity impairment instrument," *Pharmacoeconomics,* vol. 4, no. 5, pp. 353-365, 1993.

[6] R. Little and D. Rubin, Statistical analysis with missing data, New York: John Wiley and Sons, 1987.

Statistical Analysis Plan for Long term comparative effectiveness of once weekly semaglutide versus standard of care in a real world adult US population with type 2 diabetes - a randomized pragmatic clinical trial

Protocol No. NN9535-4416

Novo Nordisk

October 26, 2021

Version 3.0

[redacted]

[redacted]

~~CONFIDENTIAL~~

[redacted]

Signature Page for Statistical Analysis Plan
Protocol No: NN9535-4416
26-OCT-2021

[redacted] **Approval:**

Name: [redacted]
Title: [redacted], Research Biostatistics

DocuSigned by:
Signature: [redacted]

Date: 27 October 2021

Name: [redacted]
Title: [redacted] Scientist

DocuSigned by:
Signature: [redacted]

Date: 01 November 2021

Novo Nordisk Approval:

Name: [redacted]
Title: [redacted]

DocuSigned by:
Signature: [redacted]

Date: 05 November 2021

Name: [redacted]
Title: [redacted] Innovative Trials

DocuSigned by:
Signature: [redacted]

Date: 09 November 2021

TABLE OF CONTENTS

[redacted]

1.0 INTRODUCTION

This document describes the Statistical Analysis Plan (SAP) for Protocol Number NN9535-4416, Long term comparative effectiveness of once weekly semaglutide versus standard of care in a real world adult US population with type 2 diabetes – a randomized pragmatic clinical trial. This is a 2-year, multi-center, randomized, open label, parallel group, active comparator pragmatic clinical trial (PCT) comparing semaglutide s.c. versus standard of care (SOC) when added to up to 2 oral antidiabetic medications as treatment intensification among adult type 2 diabetes mellitus (T2DM) patients in the course of routine clinical practice.

This SAP is based on the most recent version of the protocol: Version 4.0, December 5, 2019.

Protocol History:

Number	Version	Date
1.0	Original Protocol	March 13, 2018
2.0	Version 2.0	March 27, 2019
3.0	Version 3.0	August 21, 2019
4.0	Version 4.0	December 5, 2019

2.0 STUDY OBJECTIVES

2.1 Primary Objective

The primary objective is to demonstrate superior long term effects of treatment with semaglutide s.c. compared to SOC each added to up to 2 oral antidiabetic medications on glycemic control when used as intensification in routine clinical practice in adult patients with T2DM.

2.2 Secondary Objectives

The secondary objectives are to compare the long term effect of semaglutide s.c. versus SOC each added to up to 2 oral antidiabetic medications and used as intensification in routine clinical practice in adult patients with T2DM with regards to:

- Weight loss
- Patient reported outcomes (PROs) and clinician reported outcomes (ClinROs)
- Hypoglycemia

- Healthcare Resource Utilization (HCRU)
- Adherence and persistence to treatment

2.3 Estimands

The estimand informs choices about data foundation and statistical analysis including possible imputation of missing data, hereby ensuring that randomization is preserved as a sound basis for statistical inference; i.e. estimation of effect size, associated uncertainty and statistical testing. Two estimands, each targeting a unique research question, have been defined to adequately describe the effect of semaglutide s.c. as compared to SOC (excluding any semaglutide).

2.3.1 Primary Estimand

The primary estimand for all objectives will be the "intention-to-treat" (ITT) estimand evaluating the effectiveness of randomized treatment intervention irrespective of adherence to this randomized intervention or changes to other antidiabetic medication.

2.3.2 Secondary Estimand

The secondary estimand for all objectives with the exception of the adherence and persistence to treatment objective, is the "if all patients had adhered" estimand. This estimand evaluates the effect of randomized treatment intervention for all randomized patients if all patients had adhered to randomized treatment, regardless of changes to other antidiabetic medication.

3.0 STATISTICAL METHODOLOGY

3.1 General Considerations

Statistical analyses will be performed using SAS® version 9.4 or higher computer software. Analyses will utilize prospectively collected data from study sites as well as secondary data collection utilizing administrative claims data from the HealthCore Integrated Research Database (HIRD®). Data from both sources will be integrated into one dataset for analysis.

Patient flow, patient characteristics, treatments, and outcomes will be tabulated and summarized with descriptive statistics based on observed values only. All descriptive data summaries will include means, medians, standard deviation (SD), and ranges for continuous variables and absolute/relative frequencies for categorical data. In addition, for each variable the number (count) of values that are missing will also be reported for dedicated study visit data. Missing data are data that are planned and can be collected but are absent. Statistics will be summarized for the study population overall and by treatment arm. Two sets of descriptive summaries will be

produced: one based on the primary estimand (ITT) and one based on the secondary estimand ("if all patients had adhered").

Raw data (i.e., minimum and maximum values presented for range in continuous variables) will be reported out to the precision with which it was collected. Means will be reported to 1 decimal place more than the raw data. SD will be reported to 1 decimal place more than the mean. Percentages will be reported to 1 decimal place. Trailing zeros will be presented to maintain a consistent level of precision, e.g. 2.0 rather than 2.

Inferential tests will be performed at the 5% level of significance (two-sided). All p-values will be rounded to 3 decimal places. If a rounded p-value is 0.000 (i.e., the actual p-value is less than 0.0005), then the p-value will be presented as 'p< 0.001.'

3.2 Study Populations

3.2.1 Analysis Population

The following analysis set will be defined:

Full analysis set (FAS): Includes all randomized patients analyzed according to the treatment group to which they were assigned at randomization.

3.2.2 Safety Population

The safety population includes all randomized patients initiated on study drug analyzed according to the treatment group to which they received.

3.3 Evaluation Schedule and Definitions

Dedicated study visits are at randomization, at year 1, and at year 2. The study will also capture data collected at the sites during routine diabetic care visits, i.e., office visits and other patient contacts that occur as part of routine clinical practice. Routine diabetic care visits will occur per study physician's routine clinical practice, therefore the number of visits and data available may differ from site to site and patient to patient.

For analysis, baseline includes assessments conducted and patient data collected at or prior to randomization. Baseline is defined as ≤ 90 days prior to randomization visit (week 0) for HbA1c. For secondary endpoint assessments, baseline is defined as ≤ 4 weeks prior to the randomization visit (week 0). All baseline values will be derived from the last assessment taken prior to or at the randomization visit.

The dedicated year 1 visit is at 52 weeks (±6 weeks) after randomization. Year 1 endpoint data will be derived from the dedicated year 1 study visit if the dedicated study visit is ±10 weeks of 52 weeks post-randomization. If year 1 endpoint data is missing or outside of 52 ±10 weeks, then

the routine diabetic care data closest to 52 weeks post-randomization ±10 weeks will be used. If no routine diabetic care data are available 52 weeks post-randomization ±10 weeks, then year 1 endpoint data will be considered missing and imputed as described in section 4.5, if applicable.

The dedicated year 2/EOS visit is 104 weeks (±6 weeks) after randomization. Year 2 endpoint data will be derived from the dedicated year 2 study visit if the dedicated study visit is ±10 weeks of 104 weeks post-randomization. If year 2 endpoint data is missing or outside of 104 ±10 weeks, then the routine diabetic care data closest to 104 weeks post-randomization ±10 weeks will be used. If no routine diabetic care data are available 104 weeks post-randomization ±10 weeks, then year 2 endpoint data will be considered missing and imputed as described in section 4.5, if applicable.

3.4 Confirmatory Endpoints and Hypotheses

The primary endpoint is HbA1c <7.0% (53 mmol/mol) at year 1 (yes/no).

Confirmatory secondary endpoints include:

- Change in HbA1c (%-point) from baseline to year 1
- HbA1c <7.0% (53 mmol/mol) at year 2 (yes/no)
- Change in HbA1c (%-point) from baseline to year 2

This study is designed to have 90% power to confirm superiority of semaglutide s.c. versus SOC on the primary endpoint and 85% power to also confirm superiority of the first confirmatory secondary endpoint based on primary estimand analyses. The primary and confirmatory secondary endpoints will all be tested for superiority of semaglutide under multiplicity control via a hierarchical testing scheme according to the order given in the above bullet list. Confirmatory testing will only be performed for the primary estimand with the secondary estimand being supportive. The testing procedure will be stopped the first time an analysis fails to confirm superiority of semaglutide for the endpoint in question using a two-sided significance level of 5%.

Superiority on HbA1c <7.0% at year 1 and at year 2 will be evaluated with respect to the odds ratio (OR) (odds semaglutide / odds SOC):

H_0: OR≤1 against H_a: OR>1

Superiority on change in HbA1c (year 1 – baseline; year 2 – baseline) will be evaluated with respect to the mean treatment difference (TD) (semaglutide – SOC):

H_0: TD≥0 against H_a: TD<0

Statistical analysis of all other endpoints will be unadjusted for multiplicity.

3.5 Sample Size Estimation

3.5.1 Power and Sample Size for Primary Objective

Assumptions for the sample size were based on input from the HIRD claims/laboratory results database. In line with the primary ITT estimand, the assumptions for the proportion of patients with HbA1c < 7.0% at year 1 and year 2, and for the change from baseline in HbA1c are based on the claims/laboratory results data within the [redacted] population for all patients initiating treatment intensification regardless of whether patients adhered to this treatment. Specifically, for semaglutide s.c., assumptions are based on data for liraglutide and dulaglutide. For SOC, assumptions are based on all intensifications in the claims database. The data from the claims database are shown in Table 3-1. Because the baseline HbA1c was higher in the all intensifications population compared to the liraglutide and dulaglutide groups, the assumptions used for the sample size calculation were adjusted accordingly.

Table 3-1: HbAlc Change Data from [redacted] Population via HealthCore Integrated Research Database on Patients Intensifying Diabetes Treatment after Metformin

Follow-up	Treatment intensification after metformin	Number of patients with HbA1c Follow-up data	Proportion Patients Hba1c <7% at Follow-up	Number of patients with Baseline and Follow-up HbA1c data	HbA1c (%) Baseline mean	HbA1c (%) Change from Baseline	
						mean	SD
1y	All intensifications	6,522	47%	4,068	8.9	-1.54	2.3
	dulaglutide	411	62%	115	8.6	-1.47	2.2
	liraglutide	1,762	62%	394	8.4	-1.34	2.0
2y	All intensifications	4,047	49%	2,401	8.9	-1.42	2.3
	dulaglutide	32	71%	≤10	8.8	-1.35	1.2
	liraglutide	893	55%	217	8.6	-1.24	2.1

Source: Data from the HealthCore Integrated Research Database. Data on file.

Table 3-2 presents several sample size scenarios together with assumptions and calculated power. The sample size calculation aimed for 90% power for confirming superiority of semaglutide s.c. versus SOC on the primary endpoint and 85% power to also confirm superiority of the first confirmatory secondary endpoint based on an analysis of the primary estimand for each of the endpoints. For the sample size calculation the likelihood ratio chi-square test was assumed for analyses of the binary presence of HbA1c < 7.0% endpoints and a t-test for the continuous change in HbA1c endpoints. An overall alpha level of 5% with two-sided tests was assumed.

Table 3-2 Assumptions used in the sample size and calculated total power.

Total Number Randomized	Proportion Patients with HbA1c <7% at 1 and 2 Years		HbA1c ETD at 1 and 2 years	HbA1c change SD	Power total
	SOC	Semaglutide			
1849	50%	60%	0.5	2.3	80%
2250	50%	58%	0.5	2.3	72%
2250	**50%**	**60%**	**0.5**	**2.3**	**90%**
2250	50%	60%	0.5	2.0	93%
2250	50%	60%	0.4	2.3	77%

ETD = Estimated Treatment Difference, SD=Standard Deviation
Bold marked numbers indicate the chosen assumptions and numbers marked in underline indicate variations

The proportion of missing data for the confirmatory endpoints was estimated to be 25% after one year and 44% after two years. In the sample size calculation, it was assumed that only non-missing data at year 1 and year 2 would be used for the respective analyses. This is considered conservative, since the use of imputed data in the actual primary analysis (see 4.5) will increase the power. When accounting for missing data, randomizing 1387 patients will contribute 1040 patients for the year 1 analyses and 780 patients for the year 2 analyses, achieving a total power of 85% for confirming the two year 1 confirmatory hypotheses. The joint power for confirming all 4 confirmatory hypotheses is 58%. The corresponding marginal powers for presence of

HbA1c < 7.0% at year 1, change in HbA1c to 1 year, presence of HbA1c < 7.0% at year 2, and change in HbA1c to 2 years are 90%, 94%, 80% and 86% respectively.

3.6 Missing Data

Except where noted in section 4.5 for HbA1c, weight, systolic blood pressure (SBP), and diastolic blood pressure (DBP) endpoints, missing data will not be imputed and will be excluded from calculations.

3.7 Data Definitions and Calculations

3.7.1 General

All absolute change from baseline variables will be calculated as:

Change = Endpoint Visit – Baseline

All relative change from baseline variables will be calculated as:

Relative Change (%) = [(Endpoint Visit – Baseline)/Baseline] x 100

If day and/or month is missing from dates, the first day of the month and/or the first month of the year will be assumed when calculating duration between dates. If imputation of an incomplete date associated with a post-randomization data element results in a date prior to randomization, the first day of the month immediately following randomization will be assumed. E.g., if the randomization date is 08/15/18 (mm/dd/yy) and the unknown date is 08/UNK/18 or UNK/10/18 or UNK/UNK/18, the imputed date would be 09/01/18.

The data sources for each endpoint and variables used in analysis are summarized in Appendix A.

3.7.2 Demographics

Age will be presented in years based on the patient's birth date and randomization visit date.

3.7.3 Diabetes History

Duration of diabetes will be presented in years based on the patient's diagnosis date and randomization visit date.

3.7.4 Treatment Variables

Study drug is defined as the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization. A patient can start and stop study drug throughout the study. At any time during the duration of the study, if a patient is taking study drug as defined, they will be considered "on study drug" regardless of prior study drug

discontinuations. At any time during the duration of the study, if a patient is not taking study drug as defined, they will be considered "off study drug." On and off study drug classification will be based on the eCRF data.

The first study drug discontinuation date will be defined as the date of the first time a patient discontinues study drug such that they are "off study drug."

For analysis of year 1 and year 2 endpoints, patients will be categorized as "on study drug" or "off study drug" according to their study drug use at the endpoint visit, regardless of prior use/non-use. If study drug status is missing for the year 1 or year 2 endpoint, the last registered study drug status prior to the endpoint visit will be used.

3.7.4.1 Semaglutide Study Drug

For the semaglutide treatment group, study drug is defined as semaglutide s.c.. Therefore at any time during the study, patients randomized to the semaglutide group will be categorized as:

On Study Drug: Taking Semaglutide s.c. (regardless of any other antidiabetic medication use or previous discontinuations)

Off Study Drug: Not currently taking Semaglutide s.c.

The following definitions will also apply to individual patient semaglutide s.c. dosing during the 2 year study period, or until time of study drug discontinuation:

Most Common: The dose associated with the longest period of use from randomization to year 2/EOS visit.
Highest: The highest dose reported at any time from randomization to year 2/EOS visit.
Final: The latest dose reported from randomization to year 2/EOS visit.

3.7.4.2 Standard of Care (SOC) Study Drug

For the SOC treatment group, study drug is defined as the *drug class* of the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization (excluding any formulation of semaglutide). If a fixed dose combination (FDC) product is prescribed, one of the drugs will be specified by the study physician to be the study drug and on/off drug categorization will be based on the study physician specified drug only. Therefore, at any time during the study, patients randomized to the SOC treatment group will be categorized as:

On Study Drug: Taking a drug in the same drug class as their first study drug, even if it is not the same individual drug (regardless of any other antidiabetic medication use or previous discontinuations except semaglutide). This also applies to treatment

intensification with a FDC product if one of the drugs is in the same drug class as the first study drug.

Off Study Drug: Not currently taking a drug in the same drug class as their first study drug.

3.7.4.3 Treatment Intensification, Change or Permanent Early Discontinuation

Treatment intensification (add-on), change (switch) in study drug, or permanent early discontinuation after randomization will be identified and categorized.

Treatment intensification is defined as initiation of an antidiabetic medication (besides the up to 2 oral antidiabetic medication(s) taken prior to randomization) in addition to study drug, or initiation of more than one antidiabetic medication after discontinuation of study drug.

The first treatment intensification date will be defined as the date of the first occurrence of an antidiabetic treatment intensification.

Treatment change is defined as discontinuation of study drug and initiation of another antidiabetic medication (besides the up to 2 oral antidiabetic medication(s) taken prior to randomization).

The first treatment change date will be defined as the date of the first occurrence of an antidiabetic treatment change.

Permanent early treatment discontinuation is defined as discontinuation of study drug as described in section 3.7.4 and no initiation of another antidiabetic medication (besides the up to 2 oral antidiabetic medication(s) taken prior to randomization) for the remainder of the subject's time in study.

3.7.5 Measures of Glycemic Control

Physicians will set a patient's individualized HbA1c target prior to randomization. Achieving target at year 1 and year 2 is defined as:

Endpoint Visit HbA1c ≤ Individualized HbA1c target

Attainment of the 2018 Healthcare Effectiveness Data and Information Set (HEDIS) HbA1c will be defined at year 1 and year 2 as:

HbA1c <8.0% if age ≥ 65 years at endpoint visit or with defined comorbidities (Appendix B), else <7.0%

[redacted]

HbA1c data for the HEDIS endpoints will be obtained from the eCRF. HEDIS comorbidities as defined in Appendix B will be obtained from the claims data and apply to the year prior to endpoint. HEDIS criteria at year 1 will be based on comorbidities identified between the randomization and year 1 visit. HEDIS criteria at year 2 will be based on comorbidities identified between the year 1 and year 2 visits.

3.7.6 Weight

No weight gain is defined as: Endpoint Visit Weight ≤ Baseline Weight

3.7.7 Adherence and Persistence to Treatment

Medication adherence refers to a patient's conformance to the provider's recommendation with respect to timing, dosage, and frequency of medication taken during the prescribed length of time.

For this study, the medication possession ratio (MPR) will be used to assess adherence. MPR will be calculated for the study drug from claims data for baseline to year 1 and for baseline to year 2 as follows:

$$\text{MPR (\%)} = \frac{\text{Sum of days supply for all prescription fills}}{\text{Total number of days in time period}} \times 100$$

MPR will be capped at 100%.

Medication persistence refers to whether a patient stays on therapy or the time from initiation to discontinuation of therapy. Medication persistence will be assessed from baseline to year 2.

For this study, medication persistence will be defined as duration of time in days from initiation of study drug to first study drug discontinuation from eCRF data as defined in section 3.7.4.

A second persistence measure will be defined as duration of time in days from the date of initiation of study drug to the date of the first intensification or change in study drug as defined in section 3.7.4.3 or study drug discontinuation as defined in section 3.7.4, whichever occurs first.

3.7.8 Patient Reported Outcomes (PROs)

3.7.8.1 Diabetes Treatment Satisfaction Questionnaire (DTSQ)

The Diabetes Treatment Satisfaction Questionnaire (DTSQ) is a validated instrument that is widely used in clinical trials and for routine clinical monitoring. It is useful in assessing outcomes of diabetes care. [1] The questionnaire consists of eight items -- six items assess treatment satisfaction, where each item is scored on a scale of 0–6, with six representing the

greatest satisfaction, and two items assess the perceived frequency of hyperglycemia and hypoglycemia, also scored on a scale of 0-6, where "0" corresponds to none of the time and "6" corresponds to most of the time.

The original DTSQ is now referred to as the status version (DTSQs) in order to distinguish it from the DTSQ change version (DTSQc), which was developed to overcome potential ceiling effects (i.e. where respondents score maximum or near-maximum satisfaction at baseline and can show little or no improvement at follow-up). [2] This study will use the DTSQs at randomization and the DTSQc at the dedicated year 1 and year 2 study visits.

Scoring the DTSQs

The DTSQs has 8 items and produces the following measures:

1. Treatment Satisfaction: Items 1, 4, 5, 6, 7 & 8 are summed to produce a Treatment Satisfaction score (range: 0 to 36). The higher the score, the greater the satisfaction with treatment.
2. Individual satisfaction with treatment items (items 1, 4, 5, 6, 7 & 8) can be considered separately: All rated: 6 (very satisfied, convenient, flexible, etc.) to 0 (very dissatisfied, inconvenient, inflexible, etc.). The higher the score, the greater the satisfaction with each aspect of treatment.
3. 'Perceived frequency of hyperglycemia' (item 2) & 'Perceived frequency of hypoglycemia' (item 3): Both rated: 6 ('most of the time') to 0 ('none of the time'). Here, lower scores indicate blood glucose levels closer to the ideal. Higher scores indicate problems.

Scoring the DTSQc

The change version has the same 8 items as the status version, with a small alteration to the wording of Item 7. The DTSQc instructions and response options differ from those of the DTSQs and produce measures of relative change in satisfaction rather than measures of absolute satisfaction.

1. Treatment Satisfaction (Change): Items 1, 4, 5, 6, 7 and 8 are summed to produce a Treatment Satisfaction (change) score (range: +18 to –18). The higher the score, the greater the improvement in satisfaction with treatment; the lower the score, the greater the deterioration in satisfaction with treatment. A score of 0 represents no change.
2. Individual satisfaction with treatment change items (items 1, 4, 5, 6, 7 and 8) can be considered separately: All rated: +3 ('much more satisfied', 'much more convenient', 'much more flexible', etc.) to –3 ('much less satisfied', 'much less convenient', 'much less flexible',

etc.). The higher the score, the greater the improvement in satisfaction with each aspect of treatment and the lower the score, the greater the deterioration in satisfaction with each aspect of treatment.

3. Two remaining items ('Perceived change in frequency of hyperglycemia' (item 2) and 'Perceived change in frequency of hypoglycemia' (item 3)) are treated individually: Both rated: +3 (‘much more of the time now’) to –3 (‘much less of the time now’). Here, negative scores indicate fewer problems with blood glucose levels. Positive scores indicate more problems than before.

3.7.8.2 Short Form 12-Item Version 2 (SF-12 v2)

The Short Form 12-item Version 2 (SF-12v2) is a generic health-related quality of live (HRQoL) measure that consists of 12 items from the Short Form 36-Item (SF-36) health survey and measures physical and mental dimensions of health. [3] Eight subscale scores (physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health) and two summary scores (physical component summary (PCS) and mental component summary (MCS)) can be obtained using the proprietary scoring software provided by Optum. Scores are calculated as norm-based standardized scores ranging from 0 to 100 with means of 50 and SDs of 10, where a score of “0” indicates the lowest level of health measured by the scales, “100” indicates the highest level of health, and “50” indicates the mean or norm-score for the general US population. [4]

For this study, the PCS-12 and MCS-12 summary scores will be used. The PCS-12 and MCS-12 will be calculated at baseline, year 1, and year 2. Change from baseline will be calculated for the PCS-12 and MCS-12 at year 1 and year 2.

3.7.8.3 Work Productivity and Activity Impairment, General Health (WPAI-GH)

The Work Productivity and Activity Impairment (WPAI) questionnaire is a 6-item instrument that measures absenteeism (work time missed), presenteeism (impairment at work / reduced on-the-job effectiveness), work productivity loss (overall work impairment / absenteeism plus presenteeism), and activity impairment because of health problems in the past seven days. [5] There are two versions of the WPAI. In the general health version (WPAI:GH) individuals are asked questions about work and activity impairment due to health problems. In the specific health problem version (WPAI:SHP), individuals are asked questions concerning impairment due to a target health problem (e.g., arthritis). This study will use the general health version of the WPAI (WPAI:GH), per general guidelines from the website that indicate the general health version is more appropriate for a disease like diabetes.

WPAI:GH Scoring Information

General Information:

In general, WPAI outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity, i.e., worse outcomes. The recall period is the past seven days, excluding the current day.

Questions:

Q1 = currently employed (Yes=works full-/part-time; self-employed; works in family business; on vacation from paid employment/No=does not work for pay; only does volunteer work; usually works, but has been laid-off or unemployed during the past seven days; retired; seasonal workers not currently working)
Q2 = hours missed due to health problems
Q3 = hours missed for other reasons
Q4 = hours actually worked
Q5 = degree health affected productivity while working
Q6 = degree health affected regular activities

Scores:

Multiply scores by 100 to express in percentages

1. Percent work time missed due to health (Absenteeism): Q2/(Q2 + Q4)
2. Percent impairment while working due to health (Presenteeism): Q5 /10
3. Percent overall work impairment due to health (Work Productivity Loss): Q2/(Q2 + Q4) + [(1 – (Q2/(Q2 + Q4))) x (Q5/10)]
4. Percent activity impairment due to health (Activity Impairment): Q6/10

Absenteeism, presenteeism, work productivity loss, and activity impairment will be calculated at baseline, year 1, and year 2. Change from baseline will be calculated for each score at year 1 and year 2.

3.7.8.4 Patient Global Impressions of Disease Severity (PGI-S) and Change (PGI-C)

The patient global impression of disease severity (PGI-S) is a single item that the patient will complete at the baseline randomization visit that assesses the patient's perception of the severity of their diabetes at the time of the randomization visit. The patient global assessment of change (PGI-C) is a single item that the patient will complete at the dedicated year 1 and year 2 study

visits that assesses their perception of their current disease status relative to the way it was before they started the study drug.

3.7.8.5 Clinician Global Impressions of Disease Severity (CGI-S) and Change (CGI-C)

The clinician global impression of disease severity (CGI-S) is a single item that the clinician will complete after the patient's baseline randomization visit that assesses the clinician's perception of the patient's severity of their diabetes at the randomization visit. The clinician global assessment of change (CGI-C) is a single item that the clinician will complete at the patient's dedicated year 1 and year 2 study visits that assesses the clinician's perception of the patient's current disease status relative to the way it was before they started the study drug.

3.7.9 Healthcare Resource Utilization (HCRU)

All-cause and T2DM-related HCRU will be defined from baseline to year 2. These data will come from the HIRD for [redacted] patients or requested from a set of other health insurers for non-[redacted] patients. HCRU will include inpatient admissions, ER encounters, physician office visits (OV), other outpatient encounters, and pharmacy prescription fills.

All medical claims (inpatient admissions, ER visits, physician OVs, or other outpatient encounters) will be included in the "all-cause" analysis. All pharmacy claims will be included in the "all-cause" HCRU. Inpatient admissions, ER visits, physician OVs, or other outpatient encounters will be considered to be "T2DM-related" if the associated medical claim(s) has an International Classification of Diseases, Tenth Revision (ICD-10) T2DM diagnosis code of E11.% in the primary position or secondary position. Pharmacy claims for antidiabetic medications will be included in the "T2DM-related" HCRU.

- All-cause inpatient admission is defined by a claim or claims with inpatient place of service. Facility and provider claims are rolled into episodes with unique admission and discharge dates. ER visit claims on one day that are directly followed by an inpatient admission are considered as one inpatient admission and all HCRU from the encounter is combined.
- T2MD-related inpatient admission is defined by a claim or claims with inpatient place of service and primary or secondary diagnosis of T2DM. Facility and provider claims are rolled into episodes with unique admission and discharge dates. ER visit claims on one day that are directly followed by an inpatient admission with a primary or secondary diagnosis of T2DM (regardless of the diagnoses associated with the ER visit claim) are considered as one T2MD-related inpatient admission and all HCRU from the encounter is combined.

- All-cause ER visit is defined by a medical claim with ER place of service.
- T2DM-related ER visit is defined by a medical claim with ER place of service and primary or secondary diagnosis of T2DM.
- All-cause outpatient physician OV is defined by a medical claim with outpatient place of service and an evaluation and management (E&M) code that indicates physician OV.
- T2DM-related outpatient physician OV is defined by a medical claim with outpatient place of service and an E&M code that indicates physician OV with a primary or secondary diagnosis of T2DM.
- All-cause other outpatient service is defined by a medical claim with outpatient place of service and code other than E&M, overall and by the following categories:
    - Tests – Lab
    - Imaging
    - Procedures
    - Occupational Therapy (OT)/Speech
    - Medication and Related Services
    - Durable Medical Equipment
    - Physician Other Services
    - Tests – Other
    - Other
- T2DM-related other outpatient service is defined by a medical claim with outpatient place of service, with a primary or secondary diagnosis of T2DM, and code other than E&M, overall and by the following categories:
    - Tests – Lab
    - Imaging
    - Procedures
    - OT/Speech
    - Medication and Related Services
    - Durable Medical Equipment
    - Physician Other Services
    - Tests – Other
    - Other
- All-cause pharmacy prescription fill is defined by a pharmacy claim.
- T2DM-related pharmacy prescription fill is defined by a pharmacy claim for an antidiabetic medication.

The length of stay (LOS) for individual all-cause and T2DM-related inpatient admissions is defined as:

LOS (days) = Discharge Date – Admission Date

The cumulative LOS will be the sum of all all-cause inpatient admissions and all T2DM-related inpatient admissions a patient experiences from baseline to year 2.

3.8 Examination of Subgroups

Data will be presented overall and by treatment group. Subgroup analyses under the primary estimand (ITT) of the primary endpoint and of the first confirmatory secondary endpoint (change in HbA1c from baseline to year 1) will be conducted on the following subgroups:

- Age (<55 years versus ≥55 years)
- Gender
- Baseline HbA1c (<8.0% versus ≥8.0%)
- Baseline body mass index (BMI) (<30 kg/m^2, ≥30 kg/m^2)
- Duration of T2DM diagnosis at baseline (<5 years versus ≥5 years)

Analysis of each subgroup will be based on a logistic regression model with logit link function for the primary endpoint and an ANCOVA model for the first confirmatory secondary endpoint. Models will also include treatment and subgroup as categorical effects, baseline HbA1c as covariate, and an interaction term composed of treatment and subgroup.

Additionally, analysis of the above subgroups will be repeated for the treatment adherence and persistence endpoints using the FAS and an ANCOVA model for the adherence measures and a Cox proportional hazards model for the persistence measures. Models will also include treatment and subgroup as categorical effects, and an interaction term composed of treatment and subgroup.

For all interaction models, the estimated treatment effect for semaglutide s.c. versus SOC and 95% confidence interval within each subgroup level and the p-value for the interaction term will be presented from each model.

4.0 STATISTICAL ANALYSES

4.1 Overview

The primary analysis for this study is a year 1 analysis. Once data collection for year 1 has completed, a database lock and year 1 analysis will be performed. The year 1 analysis will not be integrated with HCRU and will be limited to year 1 endpoints derived from eCRF data, including PRO data. The year 1 analysis results will be presented in a clinical study report and an internal

results meeting. To maintain study integrity for the remaining study period, data from year 1 will be used for limited and confidential communications while complying with public disclosure requirements. All other analyses will be conducted following a second database lock once data collection for the entire study is complete.

No changes to the study design will be made based on the results of the year 1 analysis.

4.2 Patient Disposition and Accountability

Patient disposition will be descriptively summarized for all randomized patients, including the number and percentage patients in the FAS and safety populations, patients who complete the dedicated year 2 study visit, and the primary reason for not completing the dedicated year 2 study visit for patients who terminated participation in the study early. Summary metrics for number of screen failures by study site will also be presented.

4.3 Demographics and Baseline Clinical Characteristics

Demographic and baseline clinical characteristics will be descriptively summarized for the FAS overall and by treatment group. For baseline comorbid conditions, the number and percentage of patients with at least one of the comorbid conditions collected in the study will also be summarized along with the relative frequencies of type of comorbid condition. The same approach will be followed for concomitant cardiovascular medications. In addition to eCRF data, type of insurance at randomization (obtained from claims) and geographic region (from study site location) will be summarized.

4.4 Protocol Deviations

Major protocol deviations will be summarized by type (informed consent, randomization error, inclusion/exclusion criteria error, unreported serious adverse events (SAEs), procedure not per protocol, study visit not per protocol, other).

4.5 Missing Value Imputation Methodology

Missing endpoint data will be imputed for HbA1c, weight, SBP, and DBP. The data of patients who, in violation of the protocol, initiate any formulation of semaglutide in the SOC treatment group, will be censored analytically following initiation of semaglutide. The patient will remain in the analysis, but data following their initiation of semaglutide will be censored and treated as missing.

Prior to analysis of binary endpoints derived by dichotomizing HbA1c and body weight, missing data will be imputed by multiple imputation on the continuous scale and dichotomized as required for the endpoint. 500 complete data sets will be generated to adequately account for the

uncertainty due to missing data.

4.5.1 Primary Estimand Imputation Model

Missing value imputation for the primary estimand will be based on data from all patients with observations at the endpoint visit (year 1 for the primary endpoint). Missing endpoint data will be imputed separately by treatment group and based on similar patients with available endpoint data according to study drug treatment status, i.e., according to whether these patients are on or off study drug at the endpoint visit, resulting in four separate imputation models. Data will be imputed based on the assumption that, within treatment groups, patients with missing endpoint data will behave like patients with the same study drug treatment status at endpoint as the missing patients' last registered treatment status prior to the missing endpoint. For example, for patients with missing data who have discontinued treatment with study drug, data will be imputed based on the assumption that these patients will behave like patients with available data who are no longer receiving study drug.

Technically, the imputation model will be an analysis of covariance (ANCOVA) for the endpoint data. The ANCOVA will include the baseline value of the imputed variable (HbA1c, weight, SBP, or DBP), diabetes duration, age, and sex as independent variables. After this model has been used to predict missing values, each of the now 500 complete data sets will be analyzed as described for the primary analysis in section 4.6.1. Finally, the multiple analysis results will be combined using Rubin's rule. [6] For the OR, the results will be combined on the logarithm scale.

This process will be repeated for year 2 endpoints.

4.5.2 Secondary Estimand Imputation Model

Missing value imputation for the secondary estimand will be based only on data from the subset of patients who are receiving study drug (on study drug) at the endpoint visit (year 1 for the primary endpoint) in order to estimate the treatment effect if all patients had continued treatment. The endpoint data of patients who are not receiving study drug at the endpoint visit (off study drug) will be censored analytically and imputed together with missing data. Collectively, missing and censored data will be imputed separately by treatment group based on all patients who are on study drug at the endpoint visit. Data will be imputed based on the assumption that patients with missing data will behave like patients with available data who are on study drug.

The technical aspects of missing data imputation based on multiple imputation, the statistical analysis of the multiple complete data sets, and combination of the multiple results will be the same as the ones described for the primary estimand.

This process will be repeated for year 2 endpoints.

4.6 Effectiveness Analysis

4.6.1 Statistical Analysis for the Primary Estimand

The primary effectiveness analysis is the primary estimand of the primary endpoint, HbA1c <7.0% at year 1, and confirmatory secondary endpoints of change in HbA1c (%-point) from baseline to year 1, HbA1c <7.0% at year 2, and change in HbA1c (%-point) at year 2. This analysis will be based on the FAS with missing data imputation as described in the primary estimand imputation model.

Year 1 Analysis

The primary endpoint, HbA1c < 7.0% at year 1, will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) will be presented.

The confirmatory secondary continuous endpoint of change in HbA1c (%-point) at year 1 will be analyzed using ANCOVA that will include the treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide s.c. versus SOC) will be presented.

The estimated treatment effect from each of these analyses will be complemented with associated 95% CI and two-sided p-value for testing the null-hypothesis of no difference.

Superiority for the primary endpoint, HbA1c <7.0% at year 1, will be considered established if the OR 95% confidence interval (CI) is greater than 1, or similarly if the two-sided p-value is significant on a 5% level and the treatment OR is in favor of semaglutide.

If the hierarchical testing scheme allows, superiority for change in HbA1c from baseline to endpoint visit will be considered established if the 95% CI for the estimated treatment difference is smaller than 0, or similarly if the two-sided p-value is significant on a 5% level and the treatment difference is in favor of semaglutide.

Year 2 Analysis

The Year 1 analysis described above will be repeated for the confirmatory secondary endpoints of HbA1c <7.0% at year 2 and change in HbA1c (%-point) at year 2. If the hierarchical testing scheme allows, superiority will be tested first for HbA1c <7.0% at year 2 followed by change in HbA1c (%-point) at year 2 as described for the year 1 counterparts above.

4.6.2 Statistical Analysis for the Secondary Estimand

The primary effectiveness analysis will be repeated with the secondary estimand. This analysis will be based on the FAS with missing data imputation as described in the secondary estimand imputation model. With the exception of the data used and the imputation of missing data, which follows the secondary estimand, the statistical analysis is the same as the primary effectiveness analysis.

The binary endpoints of HbA1c <7.0% at endpoint visit will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) and 95% CI will be presented.

The continuous endpoints of change in HbA1c (%-point) from baseline to endpoint visit will be analyzed using ANCOVA with treatment and baseline HbA1c as independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide s.c. versus SOC) will be presented along with associated 95% CI.

4.6.3 Additional Analyses of Glycemic Control

4.6.3.1 Descriptive Summaries of HbA1c

Observed HbA1c values will be descriptively summarized at baseline, year 1, and year 2. Additionally, the following categories of observed HbA1c will be descriptively summarized at baseline, year 1, and year 2:

- <7.0, 7.0-<7.5, 7.5-<8.0, 8.0-<8.5, 8.5-<9.0, and ≥9.0%
- <8.0 versus ≥8.0%
- <9.0 versus ≥9.0%

4.6.3.2 Supportive Analyses of Glycemic Control

The following supportive endpoints of glycemic control are binary endpoints. These analyses will compare the proportion of patients achieving the secondary endpoints related to glycemic control between semaglutide and SOC treatment arms via a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) will be presented.

- Individualized HbA1c target attained at year 1 (yes/no)
- HbA1c <7.0% (53mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 1 (yes/no)

- HbA1c target attainment per Healthcare Effectiveness Data and Information Set (HEDIS) criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 1 (yes/no)
- Individualized HbA1c target attained at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 2 (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 1 (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 2 (yes/no)
- HbA1c target attainment per HEDIS criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <7.0% (53 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)
- HbA1c <8.0% (64 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)

4.6.4 Supplementary Analyses

The HbA1c analyses described for the primary and secondary effectiveness analyses described above will be complemented with two complete case analyses based on the subset of patients without missing endpoint data, i.e., patients that do not have their endpoint imputed in the two analyses used for the primary and secondary estimand.

4.6.4.1 Complete Case Analysis

This analysis will be based only on patients with available measurements at the dedicated study visit including measurements irrespective of whether patients discontinued study drug or not.

The binary endpoints of HbA1c < 7.0% at year 1 and HbA1c < 7.0% at year 2 will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c as covariate. From the model the estimated OR (semaglutide s.c. versus SOC) will be presented along with associated 95% CI.

The continuous endpoints of change in HbA1c (%-point) at year 1 and change in HbA1c (%-point) at year 2 will be analyzed using ANCOVA with treatment and baseline HbA1c as

independent variables. From the model, the estimated mean difference in change from baseline to endpoint visit (semaglutide s.c. versus SOC) will be presented along with associated 95% CI.

4.6.4.2 Complete Case On-Study Drug Analysis

The analysis will be based only on patients with available measurements at the dedicated study visit including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed. The analysis will use the same analysis models as described for complete case analysis above.

4.7 PRO Analysis

PROs and ClinROs will be measured with the instruments described in section 3.7.8. Analysis of these measures will address the secondary objective of this study to compare semaglutide s.c. versus SOC in the study's patient population as is relates to PROs and ClinROs, i.e., treatment satisfaction, generic health outcomes, work productivity, and patient and physician global assessment measures, over one and two year observation periods. PRO analysis will descriptively summarize these measures at baseline, year 1 and year 2, as well as compare semaglutide s.c. versus SOC for change from baseline to year 1 and year 2.

No imputation of missing PRO measures is planned. The continuous change from baseline endpoints will be analyzed using the same analysis approach as described under the complete case and complete case on-study drug analyses in Section 4.6.4.

Data regarding all PROs will be reported for the entire sample by treatment group. Percentages of completed PROs will be reported as well. Descriptive statistics will be reported for aggregate scores for all appropriate measures by total score and subscale score(s) if applicable. Semaglutide and SOC treatment groups will be compared by independent t-tests for PRO measures in which the data are normally distributed or the Mann-Whitney U-test for PRO measures in which the data are not normally distributed.

- Diabetes Treatment Satisfaction Questionnaire, change version (DTSQc), Total treatment satisfaction score measured at year 1
- DTSQc, Total treatment satisfaction score measured at year 2
- Change from baseline in Short Form 12-Item Version 2 Survey (SF-12 v2), Physical component summary (PCS-12) score at year 1
- Change from baseline in SF-12 v2, PCS-12 score at year 2
- Change from baseline in SF-12 v2, Mental component summary (MCS-12) score at year 1
- Change from baseline in SF-12 v2, MCS-12 score at year 2

- Change from baseline in Work Productivity and Activity Impairment, General Health questionnaire (WPAI-GH) Absenteeism (work time missed) score at year 1
- Change from baseline in WPAI-GH Absenteeism (work time missed) score at year 2
- Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 1
- Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 2
- Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 1
- Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 2
- Change from baseline in WPAI-GH Activity Impairment score at year 1
- Change from baseline in WPAI-GH Activity Impairment score at year 2

In addition to the above endpoints, PGI-S and CGI-S will be descriptively summarized at baseline overall and by treatment group. PGI-C and CGI-C will be descriptively summarized at year 1 and year 2 overall and by treatment group.

During later analysis on the primary endpoints, PRO measures can be used in two ways:

- Correlational analysis of the PRO measures with the dependent variable. These correlations will include the appropriate measure of Pearson's 'r', Spearman's rho, Kendall's tau statistics or other appropriate measures that might fit the nature of the variable.
- Should the correlational analysis from step 1 find that patient characteristics are linked to primary analysis endpoints, multivariate analysis or Logistic Regression will be conducted to determine if there are any patient related variables from the PROs that predict the effectiveness of therapy in combination with other clinical variables. Data reported will include model specifics (H-L Goodness of fit; R2, beta coefficients for each variable as well as ORs with 95% CIs.)

4.8 Concomitant Medications

The number and percentages patients reporting pre-specified concomitant cardiovascular medications during year 1 and during year 2 will be summarized for the FAS overall and by treatment group.

4.9 Other Analyses

The following analyses further support the primary objective to compare semaglutide s.c. versus SOC in glycemic control. They will address the secondary objectives of this study to compare semaglutide s.c. versus SOC in the study's patient population over one and two year observation periods as is relates to body weight loss, hypoglycemia, HCRU, adherence and persistence to treatment, antidiabetic medication treatment patterns, and safety.

The binary and continuous endpoints that include HbA1c, weight, SBP, or DBP data will be analyzed using the same analysis approach as described under the primary estimand, secondary estimand, and supplementary analysis sections. All other endpoints will be analyzed using the same analysis approach as described under the complete case and complete case on-study drug analyses in Section 4.6.4, except where otherwise noted.

4.9.1 Weight Loss

Change in patient weight from baseline to endpoint will be calculated in absolute change (pounds) and relative change (percentage). Mean change in weight will be compared between semaglutide and SOC treatment groups following the ANCOVA model described for continuous endpoints in the primary (4.6.1) and secondary estimand (4.6.2). This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

- Change in body weight (%) from baseline to year 1
- Change in body weight (lb) from baseline to year 1
- Change in body weight (%) from baseline to year 2
- Change in body weight (lb) from baseline to year 2

4.9.2 Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP)

Absolute change in patient SBP and DBP from baseline to year 1 and baseline to year 2 will be calculated. Mean change in SBP and DBP will be compared between the semaglutide and SOC treatment groups following the ANCOVA model described for continuous endpoints in the primary (4.6.1) and secondary estimand (4.6.2). This analysis will be repeated for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

4.9.3 Hypoglycemia

The total number of hypoglycemic episodes leading to an inpatient admission or ER encounter during year 1 (from baseline to year 1), during year 2 (from year 1 to year 2), and the cumulative sum per patient from baseline to year 2 will be compared by semaglutide and SOC treatment groups utilizing a negative binomial model with treatment group as a factor and baseline HbA1c as a covariate in the model and observation time included as an offset variable. Subjects

reporting one or more episodes during year 2 who missed the year 1 study visit will use the expected date of the year 1 study visit (365 days from date of randomization) in calculating observation time during year 2. Exploratory analysis will assess the impact of SOC study drug (defined in section 3.7.4.2) on the rate of hypoglycemia.

Additionally, the frequency and percentage of patients reporting hypoglycemic episodes leading to an inpatient admission or ER encounter during year 1, during year 2, and cumulatively from baseline to year 2 will be reported and compared between treatment groups via chi-square test.

The above described analyses will also be repeated separately for those hypoglycemic episodes leading to an inpatient admission only and for those hypoglycemic episodes leading to an ER encounter only.

All analyses will be performed as described for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2). Hypoglycemic episodes that are considered SAEs will also be included in the safety analysis (section 5.4).

4.9.4 Healthcare Resource Utilization (HCRU)

HCRU analyses will compare all-cause and diabetes-related HCRU (inpatient admissions, ER encounters, physician OV, other outpatient encounters, and pharmacy utilization) as defined in section 3.7.9 from claims data by semaglutide and SOC treatment arm from baseline to year 2. The number of each type of encounter (all-cause and T2DM-related: inpatient admission, ER encounter, physician OV, other outpatient encounter (overall and by specific categories listed in section 3.7.9), pharmacy prescription fills) a patient experiences from baseline to year 2 will be summed.

HCRU will be summarized by setting (i.e., inpatient, ER, OV, outpatient, pharmacy) and cause (i.e., all-cause versus diabetes-related), for the study population overall and by treatment group. The occurrence (yes/no) and number of each type of encounter or prescription fill will be reported. LOS will be reported per admission and cumulative (baseline to year 2). LOS for inpatient admissions (days) per inpatient admission will be an encounter level analysis of the average inpatient admission LOS based on the total number of inpatient admissions; therefore, patients may contribute more than once. Statistics for cumulative LOS for inpatient admissions (days) will only be produced for patients with at least one inpatient admission.

Treatment group comparisons will utilize the Student t-test for continuous variables and chi-squared test (or Fisher's exact test if outcome count less than 5) for categorical variables.

4.9.5 Composite Endpoints

The following composite endpoints will be analyzed with a logistic regression model with a logit link function, treatment as a categorical effect, and baseline HbA1c and baseline weight as covariates. This analysis will be performed as described for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).

- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 1 (yes/no)
- Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 2 (yes/no)
- Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 2 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 1 (yes/no)
- HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 2 (yes/no)

4.9.6 Antidiabetic Treatment Patterns

4.9.6.1 Study Drug Regimen

Study drug regimen will be descriptively summarized by treatment arm.

The SOC group's randomized treatment regimen will be summarized by medication and class for study drug, i.e., the first antidiabetic medication prescribed for treatment intensification following randomization.

For the semaglutide group, relative frequency by dose category (0.25, 0.5, 1.0 mg) will be calculated at randomization, year 1, and year 2. Additionally, the relative frequency by semaglutide dose category will be calculated for most common, highest, and final as defined in section 3.7.4.1.

Treatment intensification,treatment change and permanent early treatment discontinuation (definition 3.7.4.3) will also be descriptively summarized as follows:

- Treatment intensification from baseline to year 1 (yes/no)

- Treatment intensification from baseline to year 2 (yes/no)
- Treatment change from baseline to year 1 (yes/no)
- Treatment change from baseline to year 2 (yes/no)
- Treatment discontinuation from baseline to year 1 (yes/no)
- Treatment discontinuation from baseline to year 2 (yes/no)

4.9.6.2 Adherence and Persistence to Treatment

Adherence and persistence to the study drug will be calculated and summarized for the FAS and compared between semaglutide and SOC treatment groups.

Study drug medication adherence will be summarized by the MPR (%) using the claims data. The MPR (%) for the first year of the study and the MPR (%) for the two year study period will be compared between treatment arms using a Student t-test.

Medication persistence, as calculated using the eCRF data, will be summarized by:

- Time to first study drug discontinuation during 2 years (day)
- Time to first treatment intensification (add-on), change (switch) or discontination after randomization during 2 years (day), whichever occurs first

Medication persistence measures will be compared between semaglutide and SOC treatment groups via the Cox proportional hazards model (with no covariates) to estimate the hazard ratio (HR) (semaglutide s.c. versus SOC), 95% CI, and p-value. Patients with no event will be censored at year 2 or last contact if withdrawn or lost to follow-up.

4.9.6.3 Other Antidiabetic Treatments

Antidiabetic treatment patterns will be assessed and compared between semaglutide and SOC treatment groups. This analysis will summarize the number and classification type of antidiabetic medications taken during the study period.

Descriptive statistics will be presented overall and by treatment group for the average number of antidiabetic medications patients report at randomization, year 1 and year 2. Relative frequencies of number of antidiabetic medications will also be summarized for randomization, year 1 and year 2. Relative frequencies by antidiabetic class will also be presented at randomization, year 1 and year 2.

4.10 Study Site Recruitment and Participation

To contextualize the participating study sites within the pool of [redacted] providers, tables summarizing the identification, recruitment, and participation of providers will be presented.

Site identification for this study begins with identifying the pool of potentially eligible patients in the HIRD using very broad claims-based inclusion/exclusion criteria: currently active members with at least one medical claim with an ICD-9 or ICD-10 diagnosis code for T2DM during the identification period, pharmacy carve-in benefits and pharmacy carve-out benefits at any time during the identification period, commercial plan coverage only, ≥18 years old, at least one pharmacy claim for metformin in the past 6 months, and no pharmacy claims for other antidiabetic medications within the past 30 days. These patients are then mapped back to the provider sites with which they had diabetes-related office visits during the identification period. From this process, a list of providers with number of potentially eligible patients is generated. The overall pool of [redacted] providers serving the target study population will be summarized with a table presenting patient densities by number of providers.

A flow chart summarizing each step of the process will be produced to describe site recruitment and participation:

1) Total number of providers with at least one potentially eligible patient identified.
2) Total number of providers approached (I.e., providers who met predetermined threshold (x number of patients) for outreach.)
3) Disposition of providers approached
    a. Interested, invited to participate
    b. Interested, rejected from participation
    c. Not Interested
    d. No Response

A table summarizing characteristics of providers (geographic region, provider specialty, and claims-based patient density) by disposition status will be produced.

4.11 Evaluation of the Study Population

To evaluate the generalizability of the study results, an analysis of the study population will be performed. The primary objective of this analysis is to evaluate the external validity of the study and to understand how the study population fits into the larger T2DM populations within study sites and commercially-insured population as a whole. Two claims-based T2DM patient populations will be described and compared to the study population: (1) all research-eligible commercially-insured and Medicare Advantage [redacted] T2DM patients treated with up to 2 oral antidiabetic medication, excluding oral semaglutide, who undergo an antidiabetic treatment intensification and (2) within the study sites, commercially-insured and Medicare Advantage [redacted] T2DM patients treated with up to 2 oral antidiabetic medications, excluding oral

semaglutide, who are not enrolled in the study, but who undergo an antidiabetic treatment intensification.

Identification of Claims-Based Comparator Populations

The claims-based comparator populations and analysis will be based on data from the HIRD. Claims based inclusion/exclusion criteria will be imposed on the two comparator populations as best as can be identified via the claims data and available lab result data to reflect the inclusion/criteria of the study population. All patients meeting the inclusion/exclusion criteria for each of the claims-based T2DM populations will be included in the analysis.

- Claims patient identification period = July 2018 until date last patient enrolled. (To align with the dates of the PCT study period).
- Claims index date = Patient's first prescription date associated with antidiabetic treatment intensification during the patient identification period.
- Claims study period = July 2017 until 12 months following date of last patient enrolled. (To allow for 12 months of claims data before and after the identified antidiabetic treatment intensification.)

Inclusion Criteria for Population #1

- ≥ 1 medical claim with an ICD-10 diagnosis code for T2DM during patient identification period
- A pharmacy claim(s) for 1 or 2 oral antidiabetic treatments, excluding oral semaglutide, as described in the PCT, during the patient identification period
    - First date of antidiabetic treatment pharmacy claim is index date
- ≥ 12 months medical and pharmacy eligibility prior to index date
- Age ≥ 18 years on index date
- HbA1c >7.0% within 90 days prior to and including index date

Exclusion Criteria for Population #1

- Pharmacy claims for ≥2 oral antidiabetic medications, oral semaglutide, or any injectable antidiabetic medicationfor 30 days prior to index date.
- Randomized into the current PCT.

Inclusion Criteria for Population #2

The inclusion criteria for population #2 will include all of the inclusion criteria as population #1, as well as:

- Receiving antidiabetic treatment at a study site

Exclusion Criteria for Population #2

The exclusion criteria for population #2 is the same as the exclusion criteria for population #1.

All analyses evaluating the study population will be descriptive only and will help to contextualize study results within the T2DM population broadly. All analyses will be based on the overall population only; analyses by treatment group will not be conducted.

The analysis will descriptively compare the overall baseline demographics between the PCT study population and the two comparator populations. Baseline HbA1c (defined as closest HbAlc up to and including 90 days prior to randomization (for PCT population) or index date (for comparators)) will also be compared.

Observed HCRU (all-cause and T2DM-related; inpatient, ER, outpatient physician OV, and other outpatient services, pharmacy utilizations) during the 12 month period prior to randomization (for study population) or prior to index date (for comparators) will also be compared between the PCT study population and the two comparator populations. The PCT study population for this comparison will be limited to study patients with 12 months of health plan eligibility prior to randomization.

Additionally, treatment patterns of non-enrolled T2DM patients within the practices from which the study patients are recruited will be evaluated to identify any relevant patterns of care suggesting channeling of certain types of patients away from the study. This analysis be based on the subset of patients with 12 months post-index data medical and pharmacy eligibility and will present the relative frequency of antidiabetic medications by class during the 12 months following randomization or index date for the PCT study population and comparator population #2.

5.0 SAFETY

5.1 AEs

For the purpose of this study, AEs that do not meet the definition of an SAE (protocol section 7.2) will only be required to be collected in the eCRF if they lead to study drug discontinuation.

5.2 SAEs

All AEs meeting the definition of an SAE (protocol section 7.2) will be collected in the eCRF.

5.3 Pregnancy

Any abnormal pregnancy outcome (e.g., spontaneous miscarriage, fetal death, congenital anomaly/birth defect, etc.) is considered an SAE. For the purposes of this study, any pregnancies in participating female patients will be reported, along with pregnancy outcome and any AEs or SAEs observed in the fetus or newborn until 1 month of age.

5.4 Analysis

No formal safety analyses are planned for this study. SAEs and AEs leading to study drug discontinuation will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and descriptively summarized by System Organ Class and Preferred Term (PT). Pregnancies occurring during the study will be also be descriptively summarized.

6.0 Changes from Protocol

Version 1.0 of the protocol specified conducting PRO Analysis (Protocol Section 6.6) and Other Analyses (Protocol Section 6.8) as described for the primary estimand, secondary estimand, and supplementary analysis (complete case and complete case on-study drug). However, missing data imputation will only be completed for HbA1c, weight, SBP, and DBP. Therefore, for endpoints that do not include HbA1c, weight, SBP, or DBP data, endpoints will be analyzed using the same analysis approach as described under the complete case and complete case on-study drug analyses in Section 4.6.4, except where otherwise noted. Relevant text was corrected in NN95359-4416 Protocol Version 2.0, dated 27MAR2019.

7.0 SAP Change Log

Summary of SAP Changes since First Patient First Visit
NN9535-4416 SAP Version 2.0, dated 30JUN2020

Number	Section/Page	Change	Rationale
1	Global Change, Study Population	Allow enrollment of patients treated with either 1 or 2 oral antidiabetic medications	Expand study population from patients treated with metformin monotherapy to patients treated with either 1 or 2 oral antidiabetic medications to facilitate enrollment.
2	Global Change, Study Population	Allow enrollment of patients with any commercial or Medicare insurance with pharmacy benefits.	Expand study population to facilitate enrollment.
3	Global Change	~~**July**~~**June ~~5~~30, ~~2018~~2020** Version ~~1~~2.0	Update SAP date and version.
4	Study Approvals, p. 2	**Name:** [redacted] **Title:** [redacted], Research Biostatistics ~~**Name:**~~ [redacted] ~~**Title:**~~ [redacted] ~~Scientist~~	Update to [redacted] study statistician.
5	Study Approvals, p. 2	**Name:** [redacted] **Title:** [redacted] ~~**Name:**~~ [redacted] ~~**Title:**~~ [redacted]	Update to Novo Nordisk study statistician.
6	Global Change, Study Treatments	Semaglutide s.c.	Clarify formulation of semaglutide allowed for study treatment is subcutaneous only.
7	Global Change	routine diabetic care visits	Clarify data collection between study visits for sites.
8	Global Change, Study Treatments	…semaglutide s.c. as compared to SOC (excluding any semaglutide).	No formulation of semaglutide is allowed in SOC group. Clarify that main study comparison excludes any formulation of semaglutide from the SOC group.

9	Section 1.0, Introduction	This is a 2-year, multi-center, randomized, open label, parallel group, active comparator pragmatic clinical trial (PCT) comparing semaglutide s.c. versus standard of care (SOC) when added to ~~metformin monotherapy~~up to 2 oral antidiabetic medications as treatment intensification among adult type 2 diabetes mellitus (T2DM) patients in the course of routine clinical practice.	Update study design to reflect expanded study population.
10	Section 2.1, Primary Objective	The primary objective is to demonstrate superior long term effects of treatment with semaglutide s.c. compared to SOC each added to ~~metformin~~ up to 2 oral antidiabetic medications on glycemic control when used as intensification in routine clinical practice in adult patients with T2DM.	Update primary objective to reflect expanded study population.
11	Section 2.2, Secondary Objectives	The secondary objectives are to compare the long term effect of semaglutide s.c. versus SOC each added to ~~metformin~~ up to 2 oral antidiabetic medications and used as intensification in routine clinical practice in adult patients with T2DM with regards to:	Update secondary objectives to reflect expanded study population.
12	Section 3.3, Evaluation Schedule and Definitions	Year 2 endpoint data will be derived from the dedicated year ~~1~~ 2 study visit if the dedicated study visit is ±10 weeks of 104 weeks post-randomization.	Correction
13	Section 3.4, Confirmatory Endpoint and Hypothesis	This study is designed to have 90% power to ~~jointly~~ confirm superiority of the primary endpoint and 85% power to also confirm superiority of the first ~~the above three~~ confirmatory secondary endpoints based on primary estimand analyses.	Update to power calculations.
14	Section 3.5.1 Power and Sample Size for Primary Objective	In line with the primary ITT estimand, the assumptions for the proportion of patients with HbA1c < 7.0% at year 1 and year 2, and for the change from baseline in HbA1c are based on the claims/laboratory results data within the [redacted] population for all patients initiating treatment intensification regardless of whether patients adhered to this treatment.	Update to reflect protocol.
15	Section 3.5.1	The sample size calculation aimed for 90% power for ~~simultaneously~~ confirming superiority of semaglutide s.c. versus SOC on	Update to power calculations.

[redacted]

	Power and Sample Size for Primary Objective	~~all of~~ the primary endpoint and 85% power to also confirm superiority of the first confirmatory secondary endpoint based on an analysis of the primary estimand for each of the endpoints~~the secondary confirmatory endpoints at both year 1 and year 2 (4 hypotheses in total)~~.	
16	Section 3.5.1 Power and Sample Size for Primary Objective	When accounting for missing data, randomizing ~~2250~~ 1387 patients will contribute ~~1687~~ 1040 patients for the year 1 analyses and ~~1260~~ 780 patients for the year 2 analyses, achieving a total power of ~~90%~~ 85% for confirming the two year 1 confirmatory hypotheses. The joint power for confirming all 4 confirmatory hypotheses is 58%. The corresponding marginal powers for presence of HbA1c < 7.0% at year 1, change in HbA1c to 1 year, presence of HbA1c < 7.0% at year 2, and change in HbA1c to 2 years are ~~99%~~90%, ~~99%~~94%, ~~95%~~80%, and ~~97%~~86% respectively.	Update to planned enrollment and power calculations.
17	Section 3.7.4.2 Standard of Care (SOC) Study Drug	For the SOC treatment group, study drug is defined as the drug class of the first antidiabetic oral or injectable medication prescribed for treatment intensification following randomization (excluding any formulation of semaglutide).	Semaglutide in any formulation is not allowed during the study period for the SOC group.
18	Section 3.7.4.3 Treatment Intensification or Change	Treatment intensification is defined as initiation of an antidiabetic medication (besides ~~metformin~~the up to 2 oral antidiabetic medication(s) taken prior to randomization) in addition to study drug, or initiation of more than one antidiabetic medication after discontinuation of study drug.	Update definition to reflect expanded study population.
19	Section 3.7.4.3 Treatment Intensification or Change	Treatment change is defined as discontinuation of study drug and initiation of another antidiabetic medication (besides ~~metformin~~the up to 2 oral antidiabetic medication(s) taken prior to randomization).	Update definition to reflect expanded study population.
20	Section 3.7.9 Healthcare Resource Utilization (HCRU)	These data will come from the HIRD for [redacted] patients or requested from a set of other health insurers for non-[redacted] patients.	Updated definition to reflect expanded study population and collection of claims data.

21	Section 4.5 Missing Value Imputation Methodology	Missing endpoint data will be imputed for HbA1c, weight, SBP, and DBP. The data of patients who, in violation of the protocol, initiate any formulation of semaglutide in the SOC treatment group, will be censored analytically following initiation of semaglutide. The patient will remain in the analysis, but data following their initiation of semaglutide will be censored and treated as missing.	Semaglutide in any formulation is not allowed during the study period for the SOC group.
22	Section 4.9.4 Healthcare Resource Utilization (HCRU)	HCRU analyses will compare all-cause and diabetes-related HCRU (inpatient admissions, ER encounters, physician OV, other outpatient encounters, and pharmacy utilization) as defined in section 3.7.9 from ~~the HIRD~~claims data by semaglutide and SOC treatment arm from baseline to year 2.	Update to reflect expanded study population and collection of claims data.
23	Section 4.9.6.3 Other Antidiabetic Treatments	~~Background Medication~~ ~~Metformin dose will be descriptively summarized overall and by treatment group at randomization, year 1, and year 2.~~ ~~Other Antidiabetic Treatments~~ … Descriptive statistics will be presented overall and by treatment group for the average number of antidiabetic medications patients report at randomization, year 1 and year 2. Relative frequencies of number of antidiabetic medications will also be summarized for randomization, year 1 and year 2. Relative frequencies by antidiabetic class will also be presented at randomization, year 1 and year 2.	Update descriptive summary of other antidiabetic treatments to reflect expanded study population.
24	Section 4.11 Evaluation of the Study Population	The primary objective of this analysis is to evaluate the external validity of the study and to understand how the study population fits into the larger T2DM populations within study sites and [redacted] commercially-insured population as a whole. Two claims-based T2DM patient populations will be described and compared to the study population: (1) all research-eligible commercially-insured and Medicare Advantage [redacted] T2DM patients treated with ~~metformin~~up to 2 oral antidiabetic medication, excluding oral semaglutide, who undergo an antidiabetic	Update to reflect expanded study population.

		treatment intensification and (2) within the study sites, commercially-insured and Medicare Advantage [redacted] T2DM patients treated with ~~metformin~~ up to 2 oral antidiabetic medications, excluding oral semaglutide, who are not enrolled in the study, but who undergo an antidiabetic treatment intensification.	
25	Section 4.11 Evaluation of the Study Population	Inclusion Criteria for Population #1 • ≥ 1 medical claim with an ICD-10 diagnosis code for T2DM during patient identification period • ~~≥ 1~~ A pharmacy claim(s) for 1 or 2 oral antidiabetic treatments, excluding oral semaglutide, ~~in addition to metformin,~~ as described in the PCT, during the patient identification period o First date of antidiabetic treatment pharmacy claim is index date • ≥ 12 months medical and pharmacy eligibility prior to index date ~~• ≥ 1 pharmacy claim for metformin in 6 months prior to index date~~ • Age ≥ 18 years on index date • HbA1c >7.0% within 90 days prior to and including index date	Update to reflect expanded study population.
26	Section 4.11 Evaluation of the Study Population	Exclusion Criteria for Population #1 • ~~≥ 1 pharmacy~~ Pharmacy claims for ≥2 ~~any~~ oral antidiabetic medications, oral semaglutide, or any injectable antidiabetic medication ~~other than metformin~~ for 30 days prior to index date. • Randomized into the current PCT.	Update to reflect expanded study population.

[redacted]

Summary of SAP Changes since First Patient First Visit
NN9535-4416 SAP Version 3.0, dated 26-OCT-2021

Number	Section/Page	Change	Rationale
1	Global Change	~~June~~ October ~~26~~30, 202~~1~~0 **Version ~~2~~3.0**	Update SAP date and version.
2	Study Approvals, p. 2	**Name:** [redacted], **Title:** [redacted] **Name:** [redacted] **Title:** [redacted] Innovative Trials	Update to Novo Nordisk [redacted] and study statistician.
3	Section 3.2.2, Safety Population	The safety population includes all randomized patients initiated on study drug analyzed according to the treatment group to which they received.	Text added to clarify definition of study population.
4	Section 3.4, Confirmatory Endpoints and Hypotheses	This study is designed to have 90% power to confirm superiority of semaglutide s.c. versus SOC on ~~of~~ the primary endpoint and 85% power to also confirm superiority of the first confirmatory secondary endpoint based on primary estimand analyses. The primary and confirmatory secondary endpoints will all be tested for superiority of semaglutide under multiplicity control via a hierarchical testing scheme according to the order given in the above bullet list. Confirmatory testing will only be performed for the primary estimand with the secondary estimand being supportive. The testing procedure will be stopped the first time an analysis fails to confirm superiority of semaglutide ~~of~~ for the endpoint in question using a two-sided significance level of 5%.	Text revised to clarify that references to 'superiority' specifically mean superiority of semaglutide s.c. in comparison to SOC.
5	Section 3.7.4.3, Treatment Intensification, Change or Permanent Early Discontinuation ~~or Change~~	Permanent early treatment discontinuation is defined as discontinuation of study drug as described in section 3.7.4 and no initiation of another antidiabetic medication (besides the up to 2 oral antidiabetic medication(s) taken prior to randomization) for the remainder of the subject's time in study.	Section revised to include permanent early treatment discontinuation.
6	Section 3.7.7, Adherence and Persistence to Treatment	A second persistence measure will be defined as duration of time in days from the date of initiation of study drug to the date of the first intensification or change in study drug as defined in section 3.7.4.3 or study drug discontinuation as defined in section 3.7.4, whichever occurs first.	Text added to clarify the definition of the second persistence measure.

7	Section 3.8, Examination of Subgroups	Data will be presented overall and by treatment group. ~~Additionally, s~~Subgroup ~~analysis~~ analyses under the primary estimand (ITT) of the primary endpoint and of the first confirmatory secondary endpoint (change in HbA1c from baseline to year 1) will be conducted on the following subgroups: • Age (<55 years versus ≥55 years) • Gender • Baseline HbA1c (<8.0% versus ≥8.0%) • Baseline body mass index (BMI) (<30 kg/m2, ≥30 kg/m2) • Duration of T2DM diagnosis at baseline (<5 years versus ≥5 years) Analysis of each subgroup will be based on a logistic regression model with logit link function for the primary endpoint and an ANCOVA model for the first confirmatory secondary endpoint. Models will also include treatment and subgroup as categorical effects, baseline HbA1c as covariate, and an interaction term composed of treatment and subgroup. Additionally, analysis of the above subgroups will be repeated for the treatment adherence and persistence endpoints using the FAS and an ANCOVA model for the adherence measures and a Cox proportional hazards model for the persistence measures. Models will also include treatment and subgroup as categorical effects, and an interaction term composed of treatment and subgroup. For all interaction models, the estimated treatment effect for semaglutide s.c. versus SOC and 95% confidence interval within each subgroup level and the p-value for the interaction term will be presented from each model.	The analysis plan for the examination of subgroups was revised to include the first confirmatory endpoint (change in HbA1c from baseline to year 1). In addition, further text added to describe the subgroup analysis for the treatment adherence and persistence measures noted in protocol Section 6.8.8 which were accidentally omitted in the previous version.
8	Section 4.1, Overview	~~Descriptive summaries will be produced for both the primary (ITT) and secondary ("if all patients adhered") estimands, with the exception of antidiabetic treatment patterns which will be based on the primary estimand only. All statistical analyses will be conducted under both the primary estimand (ITT) and secondary estimand ("if all patients had adhered"), with the exception of the~~	Text removed. Specific details regarding which descriptive summaries and/or analyses will utilize the primary and secondary estimands are detailed within each subsection of

[redacted]

		~~adherence and persistence measures, which will be based on the primary estimand only.~~	Section 4.0 (Statistical Analyses).
9	Section 4.5.1, Primary Estimand Imputation Model	Missing endpoint data will be imputed separately by treatment group and based on similar patients with available endpoint data according to study drug treatment status, i.e., according to whether these patients are on or off study drug at the endpoint visit, resulting in four separate imputation models.	Text added to note that the current approach will result in four separate imputation models.
10	Section 4.6.3.1, Descriptive Summaries of HbA1c	Observed HbA1c values will be descriptively summarized at baseline, year 1, and year 2. Additionally, the following categories of observed HbA1c will be descriptively summarized at baseline, year 1, and year 2: • ~~<7.0, 7.0-8.0, 8.1-9.0, 9.1-10.0, >10.0%~~ • <7.0, 7.0-<7.5, 7.5-<8.0, 8.0-<8.5, 8.5-<9.0, and ≥9.0% • <8.0 versus ≥8.0% • <9.0 versus ≥9.0%	Cut-points use to categorized observed HbA1c were revised from 1.0% to 0.5% increments to facilitate the comparison with other clinical research studies.
11	Section 4.7, PRO Analysis	No imputation of missing PRO measures is planned. The continuous change from baseline endpoints will be analyzed using the same analysis approach as described under the complete case and complete case on-study drug analyses in Section 4.6.4. ~~Therefore, the primary estimand for PRO analysis will be based on the FAS with endpoint data. The secondary estimand for PRO analysis will be based only on patients with endpoint data including only measurements for patients still on study drug. In the secondary estimand, patients who discontinue study drug will be censored and missing data will not be imputed.~~	Text revised to clarify the planned analysis of PRO data.
12	Section 4.8, Concomitant Medications	The number and percentages patients reporting pre-specified concomitant cardiovascular medications during year 1 and during year 2 will be summarized for the ~~population~~ FAS overall and by treatment group.	Clarification
13	Section 4.9, Other Analyses	The binary and continuous endpoints that include HbA1c, weight, SBP, or DBP data will be analyzed using the same analysis approach as described under the primary estimand, secondary estimand, and supplementary analysis sections. All other endpoints will be analyzed using the same analysis approach as described under the complete case and complete case on-	Revisions were made to the text to better align with protocol and to clarify the planned analyses for the endpoint measures.

		study drug analyses in Section 4.6.4, except where otherwise noted. ~~Except where otherwise noted, the following analyses will be conducted for the primary estimand and the secondary estimand. For endpoints that do not include HbA1c, weight, SBP, or DBP, no missing data will be imputed. Therefore, the primary estimand for these endpoints will be based on the FAS with endpoint data. The secondary estimand for these endpoints will be based only on patients with endpoint data including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed.~~	
14	Section 4.9.3, Hypoglycemia	The total number of hypoglycemic episodes leading to an inpatient admission or ER encounter during year 1 (from baseline to year 1), during year 2 (from year 1 to year 2), and the cumulative sum ~~will be summed~~ per patient from baseline to year 2 ~~and~~ will be compared by semaglutide and SOC treatment groups utilizing a negative binomial model with ~~fixed effect of~~ treatment group as a factor and baseline HbA1c as a covariate in the model and observation time included as an offset variable. Subjects reporting one or more episodes during year 2 who missed the year 1 study visit will use the expected date of the year 1 study visit (365 days from date of randomization) in calculating observation time during year 2. Exploratory analysis will assess the impact of SOC study drug (defined in section 3.7.4.2) on the rate of hypoglycemia. Additionally, the frequency and percentage of patients reporting hypoglycemic episodes leading to an inpatient admission or ER encounter during year 1~~ (from baseline to year 1)~~, during year 2, and cumulatively from baseline to year 2 will be reported and compared between treatment groups via chi-square test. The above described analyses will also be repeated separately for those hypoglycemic episodes leading to an inpatient admission only and for those hypoglycemic episodes leading to an ER encounter only.	Analysis plan updated to further defined in more detail the hypoglycemic endpoints and time periods to be assessed, and the statistical models to be used for analysis.

		All analyses will be performed as described for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2). Hypoglycemic episodes that are considered SAEs will also be included in the safety analysis (section 5.4).	
15	Section 4.9.5, Composite Endpoints	This analysis will be ~~repeated~~ performed as described for the complete case (4.6.4.1) and complete case on-study drug (4.6.4.2).	Correction.
16	Section 4.9.6.1, Study Drug Regimen	Treatment intensification, ~~and~~ treatment change and permanent early treatment discontinuation (definition 3.7.4.3) will also be descriptively summarized as follows: • Treatment intensification from baseline to year 1 (yes/no) • Treatment intensification from baseline to year 2 (yes/no) • Treatment change from baseline to year 1 (yes/no) • Treatment change from baseline to year 2 (yes/no) • Treatment discontinuation from baseline to year 1 (yes/no) • Treatment discontinuation from baseline to year 2 (yes/no)	Section updated to include the analysis of permanent early treatment discontinuation.
17	Section 4.9.6.2, Adherence and Persistence to Treatment	Adherence and persistence to the study drug will be calculated and summarized for the FAS and compared between semaglutide and SOC treatment groups. ~~This analysis will be conducted for the primary estimand only.~~	Clarification as to the analysis population to be used for the analysis of adherence and persistence to treatment.
18	Section 4.9.6.2, Adherence and Persistence to Treatment	Study drug medication adherence will be summarized by the MPR (%) using the claims data. The MPR (%) for the first year of the study and the MPR (%) for the two year study period will be compared between treatment	Correction made to the statistical test that will be used for the analysis of MPR.

		arms using ~~the~~ a ~~test of independent proportions~~Student t-test.	
19	Section 4.9.6.2, Adherence and Persistence to Treatment	Medication persistence, as calculated using the eCRF data, will be summarized by: • Time to first study drug discontinuation during 2 years (day) • Time to first treatment intensification (add-on) ~~or~~ , change (switch) or discontination after randomization during 2 years (day), whichever occurs first Medication persistence measures will be compared between semaglutide and SOC treatment groups via the Cox proportional hazards model (with no covariates) to estimate the hazard ratio (HR) (semaglutide s.c. versus SOC), ~~and~~ 95% CI, and p-value. Patients with no event will be censored at year 2 or last contact if withdrawn or lost to follow-up.	Text added to further define the analysis and presentation of medication persistence.
20	Section 6.0, Changes from Protocol	~~The~~ Version 1.0 of the protocol specified conducting PRO Analysis (Protocol Section 6.6) and Other Analyses (Protocol Section 6.8) as described for the primary estimand, secondary estimand, and supplementary analysis (complete case and complete case on-study drug). However, missing data imputation will only be completed for HbA1c, weight, SBP, and DBP. Therefore, for endpoints that do not include HbA1c, weight, SBP, or DBP data, endpoints will be analyzed using the same analysis approach as described under the complete case and complete case on-study drug analyses in Section 4.6.4, except where otherwise noted. ~~the primary estimand will be based on the FAS with endpoint data. The secondary estimand for these endpoints will be based only on patients with endpoint data including only measurements for patients still on study drug. Patients who discontinue study drug will be censored and missing data will not be imputed. Supplementary analyses will not be completed for these endpoints.~~ Relevant text was corrected in NN95359-4416 Protocol Version 2.0, dated 27MAR2019.	Section updated to note that relevant text was corrected in version 2.0 of the NN95359-4416 Protocol, dated 27Mar2019.

[redacted]

8.0 APPENDIX A: Study Endpoints and Other Variables by Data Source

Endpoint/Variable	Data Source	
	eCRF	Administrative Claims
Study Drug Variables		
On/Off Study Drug	X	
Treatment Intensification	X	
Treatment Change	X	
Antidiabetic treatment patterns	X	
Primary Endpoint		
HbA1c <7.0% at year 1 (yes/no)	X	
Confirmatory Endpoints		
Change in HbA1c (%-point) from baseline to year 1	X	
HbA1c <7.0% at year 2 (yes/no)	X	
Change in HbA1c (%-point) from baseline to year 1	X	
Supportive Secondary Endpoint Assessment		
Individualized HbA1c target attained at year 1 (yes/no)	X	
HbA1c <7.0% (53mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 1 (yes/no)	X	
HbA1c target attainment per Healthcare Effectiveness Data and Information Set (HEDIS) criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 1 (yes/no)	X	X
Change in body weight (lb) from baseline to year 1	X	
Change in body weight (%) from baseline to year 1	X	
Change in systolic blood pressure (SBP; mm Hg) from baseline to year 1	X	

Change in diastolic blood pressure (DBP; mm Hg) from baseline to year 1	X	
Time to first study drug discontinuation during 2 years (day)	X	
Time to first treatment intensification (add-on) or change (switch) after randomization during 2 years (day)	X	
Study drug medication adherence for the first year of the study, as measured by medication possession ratio (MPR) (%)		X
Number of hypoglycemic episodes leading to an inpatient admission or emergency room (ER) encounter from baseline to year 2	X	
Diabetes Treatment Satisfaction Questionnaire, change version (DTSQc), Total treatment satisfaction score measured at year 1	X	
DTSQc, Total treatment satisfaction score measured at year 2	X	
Change from baseline in Short Form 12-Item Version 2 Survey (SF-12 v2), Physical component summary (PCS-12) score at year 1	X	
Change from baseline in SF-12 v2, PCS-12 score at year 2	X	
Change from baseline in SF-12 v2, Mental component summary (MCS-12) score at year 1	X	
Change from baseline in SF-12 v2, MCS-12 score at year 2	X	
Change from baseline in Work Productivity and Activity Impairment, General Health questionnaire (WPAI-GH) Absenteeism (work time missed) score at year 1	X	
Change from baseline in WPAI-GH Absenteeism (work time missed) score at year 2	X	
Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 1	X	
Change from baseline in WPAI-GH Presenteeism (impairment at work / reduced on-the-job effectiveness) score at year 2	X	
Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 1	X	
Change from baseline in WPAI-GH Work productivity loss (overall work impairment / absenteeism plus presenteeism) score at year 2	X	
Change from baseline in WPAI-GH Activity Impairment score at year 1	X	
Change from baseline in WPAI-GH Activity Impairment score at year 2	X	

All cause HCRU from baseline to year 2		
Number of inpatient admissions		X
LOS for inpatient admissions (days) per inpatient admission		X
Cumulative length of stay for inpatient admissions (days)		X
Number of ER encounters		X
Number of physician OV		X
Number of other outpatient encounters (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Number of medications		X
Occurrence of inpatient admission (yes/no)		X
Occurrence of ER encounter (yes/no)		X
Occurrence of physician OV (yes/no)		X
Occurrence of other outpatient encounter (yes/no) (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Diabetes related HCRU from baseline to year 2		
Number of diabetes related inpatient admissions		X
LOS for diabetes related inpatient admissions (days) per diabetes related inpatient admission		X
Cumulative length of stay for diabetes related inpatient admissions (days)		X
Number of diabetes related ER encounters		X
Number of diabetes related physician OV		X
Number of diabetes related other outpatient encounters (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X

[redacted]

Number of diabetes related medications		X
Occurrence of diabetes related inpatient admission (yes/no)		X
Occurrence of diabetes related ER encounter (yes/no)		X
Occurrence of diabetes related physician OV (yes/no)		X
Occurrence of diabetes related outpatient encounter (yes/no) (overall, and by category: Tests – Lab, Imaging, Procedures, OT/Speech, Medication and Related Services, Durable Medication Equipment, Physician Other Services, Tests – Other, Other)		X
Additional Derived Outcome Variables for Supportive Analyses		
Supportive Measures of Glycemic Control		
Individualized HbA1c target attained at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) or at least 1% point improvement in HbA1c compared to baseline at year 2 (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 1 (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) and no further antidiabetic medication intensification after randomization at year 2 (yes/no)	X	
HbA1c target attainment per HEDIS criteria (<8.0% if age ≥65 years or with defined comorbidities, otherwise <7.0%) at year 2 (yes/no)	X	X
HbA1c <7.0% (53 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 1 in patients with HbA1c >9.0% at baseline (yes/no)	X	
HbA1c <8.0% (64 mmol/mol) at year 2 in patients with HbA1c >9.0% at baseline (yes/no)	X	

Body Weight Loss		
Change in body weight (%) from baseline to year 2	X	
Change in body weight (lb) from baseline to year 2	X	
Blood Pressure		
Change in SBP (mm Hg) from baseline to year 2	X	
Change in DPB (mm Hg) from baseline to year 2	X	
Hypoglycemia		
Reported hypoglycemia leading to inpatient admission or ER encounter during year 1 (yes/no)	X	
Reported hypoglycemia leading to inpatient admission or ER encounter during year 2 (yes/no)	X	
Composite Variables		
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 1 (yes/no)	X	
Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and body weight loss of ≥5% vs baseline at year 2 (yes/no)	X	
Absolute HbA1c reduction of ≥0.5% without experiencing hypoglycemia leading to inpatient admission or ER encounter and a body weight loss of ≥5% vs baseline at year 2 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 1 (yes/no)	X	
HbA1c <7.0% (53 mmol/mol) without experiencing hypoglycemia leading to inpatient admission or ER encounter and no body weight gain vs baseline at year 2 (yes/no)	X	

Adherence to Treatment		
Study drug medication adherence for the two years of the study, as measured by the medication possession ratio (MPR) (%)		X

[redacted]

9.0 APPENDIX B: 2018 HEDIS Criteria

Denominator

The eligible population.

Note: *The eligible population for the HbA1c Control <7% for a Selected Population indicator is reported after required exclusions are applied.*

Required exclusions for HbA1c Control <7% for a Selected Population indicator

Exclude members who meet any of the following criteria:

- 65 years of age and older as of December 31 of the measurement year.
- *CABG.* Members who had CABG (CABG Value Set) in any setting during the measurement year or the year prior to the measurement year.
- *PCI*. Members who had PCI (PCI Value Set), in any setting, during the measurement year or the year prior to the measurement year.
- *IVD.* Members who met at least one of the following criteria during both the measurement year and the year prior to the measurement year. Criteria need not be the same across both years.
  - At least one outpatient visit (Outpatient Value Set) with an IVD diagnosis (IVD Value Set).
  - At least one acute inpatient encounter (Acute Inpatient Value Set) with an IVD diagnosis (IVD Value Set).
- *Thoracic aortic aneurysm.* Members who met at least one of the following criteria during both the measurement year and the year prior to the measurement year. Criteria need not be the same across both years.
  - At least one outpatient visit (Outpatient Value Set), with a diagnosis of thoracic aortic aneurysm (Thoracic Aortic Aneurysm Value Set).
  - At least one acute inpatient encounter (Acute Inpatient Value Set), with a diagnosis of thoracic aortic aneurysm (Thoracic Aortic Aneurysm Value Set).
- Any of the following, in any setting, any time during the member's history through December 31 of the measurement year.
  - *Chronic heart failure*. A diagnosis of chronic heart failure (Chronic Heart Failure Value Set).
  - *Prior MI*. A diagnosis of MI (MI Value Set).
  - *ESRD*. ESRD (ESRD Value Set; ESRD Obsolete Value Set).
  - *Chronic kidney disease (stage 4)*. Stage 4 chronic kidney disease (CKD Stage 4 Value Set).
  - *Dementia*. A diagnosis of dementia (Dementia Value Set; Frontotemporal Dementia Value Set).
  - *Blindness*. A diagnosis of blindness (Blindness Value Set).
  - *Amputation (lower extremity)*. Lower extremity amputation (Lower Extremity Amputation Value Set).

[redacted]

10.0 REFERENCES

[1] C. Bradley, "Diabetes Treatment Satistcation Questionnaire: DTSQ," in *Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Practice*, C. Bradley, Ed., Abington, Routledge, 1994, pp. 111-132.

[2] C. Bradley, "The Diabetes Treatment Satisfaction Questionnaire (DTSQ): change version for use alongside status version provides appropriate solution where ceiling effects occur," *Diabetes Care,* vol. 22, no. 3, pp. 530-532, 1999.

[3] J. Ware, M. Kosinski and S. Keller, "A 12-item short-form healch survey: construction of scales and preliminary tests of reliability and validity," *Med Care,* vol. 34, no. 3, pp. 220-233, 1996.

[4] J. Ware, M. Kosinski, D. Turner-Bowker and B. Gandek, How to score version 2 of the SF-12 health survey (with a supplement documenting version 1), Lincoln, RI: QualityMetric, Incorporated, 2002.

[5] M. Reilly, A. Zbrozek and E. Dukes, "The validity and reproducibility of a work productivity and activity impairment instrument," *Pharmacoeconomics,* vol. 4, no. 5, pp. 353-365, 1993.

[6] R. Little and D. Rubin, Statistical analysis with missing data, New York: John Wiley and Sons, 1987.

[redacted]