Protocol

Study ID: NN9535-4447

Epidemiological assessment of the risk for pancreatic cancer associated with the use of semaglutide in patients with type 2 diabetes
- A cohort study based on Nordic registry data

Redacted protocol
Includes redaction of personal identifiable information only.

Non-interventional post-authorisation safety study (PASS)

Protocol originator:
Name: [redacted]
Department: Epidemiology, 2527

PASS information

Title	Epidemiological assessment of the risk for pancreatic cancer associated with the use of semaglutide in patients with type 2 diabetes - A cohort study based on Nordic registry data.
Protocol version identifier	2.0
Date of last version of protocol	14 April 2021
EU PAS Register number	EUPAS37258
Active substance	Glucagon-like peptide-1 (GLP-1) analogues; semaglutide (A10BJ06)
Medicinal product	Ozempic®; Rybelsus®
Product reference	EU/1/17/1251/001, EU/1/20/1430/001
Procedure number	EMEA/H/C/004174
Marketing authorisation holder(s)	Novo Nordisk A/S Novo Allé DK-2880 Bagsværd Denmark
Joint Post-Authorisation Safety Study (PASS)	No
Research question and objectives	The aim of this study is to evaluate whether exposure to semaglutide influences the risk of pancreatic cancer in patients with type 2 diabetes. This is achieved by estimating the risk of pancreatic cancer associated with semaglutide use as compared to use of other non-incretin antidiabetic drugs used at a similar stage as Ozempic® or Rybelsus® in the treatment of type 2 diabetes.
Countries of study	Denmark, Sweden, and Norway
Authors	[redacted]

	[redacted] [redacted] epidemiologist, Novo Nordisk A/S

Marketing authorisation holder

Marketing authorisation holder (MAH)	Novo Nordisk A/S Novo Allé DK-2880 Bagsværd Denmark
MAH contact person	[redacted]

1 Table of contents

Page

2 List of abbreviations

AERS	Adverse Event Reporting System
ATC	Anatomical Therapeutic Chemical (Classification System)
BMI	Body Mass Index
CV	Curriculum Vitae
CVOT	Cardiovascular Outcome Trial
DDD	Defined Daily Dose
DPP-4	Dipeptidyl Peptidase 4
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU PAS	The EU electronic register of Post-Authorisation Studies maintained by the European Medicines Agency
FDA	(US) Food and Drug Administration
GLP-1	Glucagon-Like Peptide-1
GPP	Good Pharmacoepidemiology Practice
GVP	Good Pharmacovigilance Practice
ICD-10	International Classification of Diseases, 10th version
ICD-O-3	International Classification of Diseases for Oncology, 3rd version
ICPC	International Classification of Primary Care
MAH	Marketing Authorisation Holder
MEN1	Multiple Endocrine Neoplasia Type 1
NOMESCO	Nordic Medico-Statistical Committee
NSAID	Non-Steroidal Anti-Inflammatory Drugs
OTC	Over-The-Counter
PASS	Post-Authorisation Safety Study
PCOS	Polycystic Ovary Syndrome
RCPNS	Research collaborating partners from Norway and Sweden
SAP	Statistical Analysis Plan

S.C.	Subcutaneous
SGLT-2	Sodium-Glucose co-Transporter 2
SSRI	Selective Serotonin Reuptake Inhibitor
SU	Sulphonylurea
TZD	Thiazolidinedione
T2DM	Type 2 Diabetes Mellitus
WHO	World Health Organization

3 Responsible parties

Novo Nordisk A/S is the sponsor of the study.

The study is conducted according to Good Pharmacoepidemiology Practice (GPP) by research groups belonging to the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP), in accordance with the ENCePP Code of Conduct. The ENCePP Study Seal will be applied for.

The study is conducted based on a research collaboration between the sponsor, Novo Nordisk A/S, and the coordinating study entity, [redacted]. The research collaborators are the [redacted], and the [redacted]

4 Abstract

4.1 Title

Epidemiological assessment of the risk for pancreatic cancer associated with the use of semaglutide in patients with type 2 diabetes - A cohort study based on Nordic registry data (PASS).

Protocol version 2.0

14 April 2021

Main authors: [redacted]

Novo Nordisk A/S.

4.2 Rationale and background

As part of the joint risk management plan for Ozempic® (semaglutide once-weekly subcutaneous injection) and Rybelsus® (semaglutide once-daily oral formulation), the aim of this study is to estimate the risk of pancreatic cancer in users of semaglutide.

Semaglutide is a Glucagon-like peptide-1 analogue indicated for the treatment of adults with insufficiently controlled type 2 diabetes mellitus as an adjunct to diet and exercise. Ozempic® and Rybelsus® are approved for use as monotherapy when metformin is considered inappropriate due to intolerance or contraindications, or in addition to other medicinal products for the treatment of diabetes.
In 2010, a potential risk of pancreatic cancer was hypothesized for the incretin mimetic class of antidiabetic drugs (including the Glucagon-like peptide-1 analogues). Up until now, a causal relationship between treatment with semaglutide and pancreatic cancer is not supported by the totality of available clinical data. However, in spite of large-scale phase 3a programmes (4,798 subjects, and 6,322 person years of observation in total for semaglutide s.c; 5,707 subjects and 6,820 person years of observation in total for semaglutide oral), the mean follow-up of patients of up to two years in the semaglutide development programme was relatively short for the assessment of pancreatic cancer.

Thus, an epidemiological study with longer follow-up time and a substantially larger patient population is warranted to estimate the risk of pancreatic cancer in users of semaglutide.

4.3 Research question and objectives

The aim of this study is to evaluate whether exposure to semaglutide influences the risk of pancreatic cancer in patients with type 2 diabetes. This is achieved by estimating the risk of pancreatic cancer associated with semaglutide use as compared to use of other non-incretin

antidiabetic drugs used at a similar stage as Ozempic® or Rybelsus® in the treatment of type 2 diabetes.

4.4 Study design

This study is a post-authorisation safety study (PASS). A multi-national, non-interventional study based on health care data from national registers from Denmark, Sweden, and Norway is conducted based on data covering the period 2018-2023. A cohort study design is used comparing new users of semaglutide with new users of other antidiabetic drugs used at a similar stage as Ozempic® or Rybelsus® in the treatment of type 2 diabetes (active comparators). Active comparators will include the following non-incretin antidiabetic agents: sulphonylureas, sodium-glucose co-transporter 2 inhibitors, and insulin subdivided into i) basal insulin only and ii) basal + bolus insulin or premix insulin.

Propensity scores are used to match new users of semaglutide with new users of active comparators and ensure that only patients who are at the same stage of their diabetes treatment are compared, thus limiting the potential for confounding.

Due to the potential long induction time and the latency period for pancreatic cancer as well as the possibility of reverse causation, a one-year lag-time period is used. Patients are followed from one year after treatment initiation until they are censored due to a first occurrence of pancreatic cancer, conditions serving as exclusion criteria (except acute and chronic pancreatitis), death, migration, or end of study period (December 31, 2023), whichever comes first.

4.5 Population

The study population consists of new users of semaglutide (Ozempic® or Rybelsus®) and new users of active comparators if they initiate treatment with semaglutide or active comparators from the time the first subject is prescribed semaglutide in the respective country until December 31, 2022; fill at least two prescriptions of semaglutide or active comparators with the second prescription filled less than one year after the initial prescription; are ≥18 years old at the date of the initial prescription; and have at least 10 years of continuous residency in the respective country before the first prescription of either semaglutide or active comparators.

Patients will be excluded if they have rare but strong risk factors for developing pancreatic cancer before initiating treatment (e.g., cystic fibrosis, Peutz-Jeghers syndrome, etc.). Furthermore, patients who initially started insulin as first-line antidiabetic treatment and who did not initiate non-insulin treatment within 3 months will be excluded (to limit inclusion of patients with type 1 diabetes).

4.6 Variables

The primary outcome is new pancreatic cancer cases as defined by relevant diagnostic codes in national cancer registries, supplemented by data from national patient registries for the year 2023 not covered by the cancer registries in Denmark and Sweden due to lag time in data availability.

Exposure is defined as the filling of at least two prescriptions within one year of either semaglutide or active comparators in the national prescription registries. 'Ever use' (i.e. having filled ≥2 prescriptions of semaglutide or active comparators) will be considered as the main exposure. Known risk factors for pancreatic cancer and potential confounders such as indicators of diabetes severity, diabetes treatment duration and comorbidities will be included in the propensity score model.

4.7 Data sources

This multi-national study is based on data from health care registries in Denmark, Sweden, and Norway. National prescription-, cancer-, patient- and population registers are used to identify exposure to antidiabetic agents, pancreatic cancer cases, and covariates to be used in propensity score matching. Cause of death registers will also be used in supplementary analyses.

4.8 Study size

The statistical precision of the study is strongly dependent on the population exposed to semaglutide as users of active comparator drugs are more abundant. Based on projected sales figures, there will be a cumulative exposure to semaglutide by the end of the study of at least 280,000 person years, after excluding a lag-time of one year for each user. Given an expected crude incidence rate of 72 per 100,000 person years for pancreatic cancer for a study population with type 2 diabetes, there would be approximately 204 outcomes among semaglutide users over a five year study duration (under the assumption of a null association), which would allow a 95% confidence interval for a null estimate of 0.87-1.15 in 2023.

Based on the sample size calculation, it is estimated that after the study duration of five years, the study will have 80% power to detect a relative hazard of 1.25 (or a 25% increase in risk) between cohorts exposed to semaglutide and active comparators, respectively.

4.9 Data analysis

National analyses are conducted separately in each country according to a common data model and a meta-analysis is performed by the [redacted] to obtain the final estimates.

The propensity for initiating treatment with semaglutide as opposed to active comparators is estimated using logistic regression to determine the association between covariates and semaglutide initiation. The propensity score is used to match new users of semaglutide to new users of active comparators. Ozempic® and Rybelsus® initiation will be handled in two separate propensity score

models. Further, patients who are treated/not treated with semaglutide as opposed to active comparators against prediction are removed by asymmetric trimming of the tails of the propensity score.

In the final study sample, the hazard ratios with 95% confidence intervals for pancreatic cancer comparing users of semaglutide to users of active comparators is estimated by using a Cox proportional hazards model.

4.10 Milestones

Start of data collection (time of first data extraction in Denmark)	Q4 2020
End of data collection (time of full data extraction in Denmark, Norway and Sweden)	Q4 2024
Final report of study results	Q1 2026

5 Amendments and updates

Amendment or update no	Date	Section of study protocol	Amendment or update	Reason
1	08 July 2020	Entire protocol	The list of authors and responsible parties have been updated.	Inclusion of research collaborating partners from Norway and Sweden (RCPNS).
		Entire protocol	Inclusion of oral semaglutide in the study protocol. As such, users of oral semaglutide and s.c. semaglutide will comprise one cohort of semaglutide users. Patients who shift from s.c. semaglutide to oral semaglutide or vice versa will be considered a continued user in the semaglutide cohort. Initiation of s.c. semaglutide and oral semaglutide will be handled in two separate propensity score models which will be further detailed in the statistical analysis plan. This update is further supplemented by the addition of an Ozempic®-specific supplementary analysis and an update of the sample size calculation to include all semaglutide users.	Recommendation by PRAC.
		Section 6	The milestones have been updated and the submission of the final report has been postponed to Q1 2026.	To align with update of statistical methods and the possibility to include Swedish cancer data up to and including 2022.

Amendment or update no	Date	Section of study protocol	Amendment or update	Reason
		Section 8	Update of research objective to evaluate whether exposure to semaglutide "… influences the risk of pancreatic cancer…" instead of "…increases the risk of pancreatic cancer…".	Comments from RCPNS to align with a statistical approach with a two-sided confidence interval.
		Section 9.1	Split of insulin into 1) basal insulin and 2) basal + bolus insulin or premix insulin in main analysis.	Internal discussions with RCPNS.
		Section 9.1, 9.3.3	Specification of the inclusion of "diabetes treatment duration" as a potential confounder to be included in the propensity score.	Comments from RCPNS.
		Section 9.2.3	Update of exclusion criteria to reflect available ICD-10 codes.	Internal reevaluation of available ICD-10 codes in the national patient registers.
		Section 9.2.3	Update of exclusion criterion where patients who initiate insulin as first-line antidiabetic treatment are excluded. Patients who use insulin in short term will no longer be excluded.	It is not uncommon that insulin may be used short term as part of diagnostics. Such patients might include patients with type 2 diabetes and should not be excluded from the study.
		Section 9.4	The data sources have been updated and additional data sources, e.g. data from cause of death registers, primary care data from Norway, and socioeconomic data from Norway and Sweden have been added.	Comments from RCPNS and the opportunity to include more data.

Amendment or update no	Date	Section of study protocol	Amendment or update	Reason
		Section 9.5.3	The detectable relative hazard for pancreatic cancer has been updated and elaborated	Due to inclusion of oral semaglutide as recommended by PRAC and updated semaglutide sales forecasts
		Section 9.7.2	Construction of propensity score models for specific calendar intervals instead of fixed six month intervals.	Comments from RCPNS.
		Section 9.7.2	Addition of a rule of matching specifying which matching ratio to use if >5% of semaglutide users are being excluded after trimming.	Comments from RCPNS.
		Section 9.6, 9.7.1, 9.7.2, 9.8	Separate analysis of data in each country and performance of a meta-analysis to obtain the final study results.	Due to legislative changes in data sharing between Denmark and Norway/Sweden it is no longer possible to pool individual-level data from all sites prior to statistical analysis.
		Section 9.7.3.1	A supplementary analysis 'Semaglutide versus other incretin-mimetic antidiabetic drugs' has been added.	Suggestion by PRAC.
		Section 9.7.3.1	Update of supplementary analysis 'Death from pancreatic cancer as primary outcome' to include a sensitivity analysis where deaths from pancreatic cancer will be identified by using the national cause of death registers.	Comments from RCPNS to ensure more valid data on the cause of death.

Amendment or update no	Date	Section of study protocol	Amendment or update	Reason
		Section 9.7.3.1	A supplementary analysis 'All-cause mortality as primary outcome' has been added.	Suggestion by PRAC.
		Section 9.7.3.2	A sensitivity analysis 'No trimming of the propensity score' has been added.	Comments from RCPNS to address the limitations of trimming the propensity score.
		Section 9.7.3.2	A sensitivity analysis 'Competing risk' has been added.	Comments from RCPNS and PRAC to address the potential influence of competing risk.
		Section 9.7.3.2	A sensitivity analysis 'Restricting to verified cancer cases' has been added.	Comments from RCPNS to increase the validity of study results.
		Annex 2	Update of ENCePP checklist	Update of study protocol
2	01 December 2020	Section 9.2.4	A clarification of the necessity of the exclusion criteria as well as an evaluation and quantification of the expected impact of the criteria on the number of patients available for analysis has been added. The power calculation in the protocol is an estimate based on several assumptions. Since the impact of exclusion criteria is expected to be low (~10%), the power calculation in section 9.5 has not been updated.	Suggestion by PRAC

6 Milestones

Milestone	Planned date
First subject dispensed semaglutide in Denmark, Sweden, and Norway	2018
Application for data access and approvals in Denmark	Q2 2020
Registration in the EU PAS Register and ENCePP seal registration	Q3 2020
Finalization of common Statistical Analysis Plan (SAP)	Q4 2020
Start of data collection (Time of first data extraction in Denmark)	Q4 2020
Application for ethical approval* in Norway and Sweden	Q4 2020
Last subject prescribed semaglutide to be included in the study	Q4 2022
Applying for data access in Norway	Q4 2023
Applying for data access in Sweden	Q1 2024
End of data collection (Time of full data extraction in Denmark, Norway and Sweden)	Q4 2024
Adaptation of data into common data model (CDM)	Q4 2024
Running of national analyses	Q1-Q2 2025
Meta-analysis completed	Q3 2025
Study progress report 1	Q4 2020
Study progress report 2	Q3 2021
Study progress report 3	Q3 2022
Study progress report 4	Q3 2023
Final report of study results	Q1 2026

*Not required in Denmark

7 Rationale and background

As part of the joint risk management plan for Ozempic® and Rybelsus®, the aim of this study is to estimate the risk of pancreatic cancer associated with the use of semaglutide. The study is a post-authorisation safety study (PASS).

Semaglutide is a Glucagon-like peptide-1 (GLP-1) analogue, indicated for the treatment of adults with insufficiently controlled type 2 diabetes mellitus (T2DM) as an adjunct to diet and exercise. Ozempic® (semaglutide once-weekly subcutaneous injection) and Rybelsus® (semaglutide once-daily oral formulation) may be used as monotherapy when metformin is considered inappropriate due to intolerance or contraindications, or in addition to other medicinal products for the treatment of diabetes. Current guidelines for the treatment of T2DM recommend metformin monotherapy as first line therapy. GLP-1 analogues may be prescribed as second or third line therapy or as part of combination injection therapy depending on patients' tolerability, concomitant medical history and availability of the medications (1,2).

In 2010, a potential risk of pancreatic cancer was hypothesized for the incretin mimetic class of antidiabetic drugs (GLP-1 analogues and Dipeptidyl Peptidase 4 (DPP-4) inhibitors) by Butler et al (3). Butler et al suggested that based on the mode of action of incretin mimetic drugs and the pancreatic metaplastic changes seen in animal models following administration of incretin mimetic drugs, prolonged exposure to incretin mimetic drugs may lead to an increased risk of pancreatic cancer.

Following this, Elashoff et al demonstrated increased rates of pancreatic cancer in patients treated with exenatide and sitagliptin compared to other antidiabetic drugs (including sulphonylureas (SUs) and thiazolidinediones (TZDs)) identified from spontaneously reported adverse events recorded in the US Food and Drug Administration (FDA) Adverse Event Reporting System (AERS) database (4). Butler supported these findings and postulated a causal association between incretin mimetic drugs and pancreatic cancer (5). After an extensive review of all available nonclinical and clinical trial data, FDA and the European Medicines Agency (EMA) published a joint commentary stating that assertions concerning a causal association between incretin-based drugs and pancreatitis or pancreatic cancer were inconsistent with the then available data (6). In addition, a meta-analysis of the Cardiovascular Outcome Trials (CVOTs) for incretin mimetic drugs (with exposure times up to 4.5 years), found no increased risk of pancreatic cancer with a relative risk of 0.95 (confidence interval, 0.49-1.81) for all GLP-1 analogues (7).

Large, observational studies suggest no statistically significant increase in risk of pancreatic cancer associated with the use of incretin-based drugs (8,9). Recently, another epidemiological study observed a significant increase in pancreatic cancer risk with incretin-based therapy; however the authors conclude this finding to be likely due to reverse causation bias (10). This and other sources of bias and confounding in observational studies of drug-cancer associations call for further and carefully designed studies.

The semaglutide development programme has not supported a causal relationship between treatment with semaglutide and pancreatic cancer. Animal studies showed no indications of semaglutide-induced acute pancreatitis, pancreatic adenocarcinomas, glucagonomas, or other pancreatic proliferative lesions in any of the repeat dose toxicity studies or the 2-year carcinogenicity studies (NN9535: Non-clinical overview 2.4, Section 2.4.4.1, seq 0000; NN9924: Non-clinical overview 2.4, Section 2.4.4.2 and 2.4.5.4, seq 0000). Further, in the phase 3 programme of semaglutide (s.c./oral), there was no increase in the number of pancreatic cancer cases when compared to comparators. In the 3a key efficacy trials, where the maximum follow-up time was 1 year for semaglutide s.c. and 78 weeks for sematglutide, the rates of pancreatic cancers were comparable for semaglutide and the comparator groups. In addition, in the CVOTs with longer mean follow up periods of up to 2 years, the rates were comparable between semaglutide and the placebo group. In these trials, the population was older, and had more advanced diabetes compared to the efficacy trials, thus representing a population that would be at higher risk of developing pancreatic cancer (NN9535: Summary 2.7.4, Section 2.11.5, seq 0000; NN9924: 2.7.4 Summary of clinical safety). However, in spite of large-scale phase 3a programmes (4,798 subjects, and 6,322 person years of observation in total for semaglutide s.c.; 5,707 subjects and 6,820 person years of observation in total for semaglutide oral), the mean follow-up of up to two years in the semaglutide development programme was relatively short for the assessment of pancreatic cancer.

Thus, an epidemiological study with longer follow-up time and a substantially larger patient population, as the study described in the present protocol, is warranted.

8 Research question and objectives

The aim of this study is to evaluate whether exposure to semaglutide influences the risk of pancreatic cancer in patients with T2DM.

8.1 Primary objective

To estimate the risk of pancreatic cancer associated with semaglutide use as compared to use of other non-incretin antidiabetic drugs used at a similar stage as Ozempic® or Rybelsus® in the treatment of type 2 diabetes.

9 Research methods

9.1 Study design

This is a multi-national, non-interventional PASS based on data from national registers in three Nordic countries: Denmark, Sweden, and Norway during the period 2018-2023.

In this registry-based cohort study, the risk of pancreatic cancer is compared in a cohort of new users of semaglutide (Ozempic® or Rybelsus®) relative to a cohort of new users of other non-incretin antidiabetic drugs used at a similar stage as Ozempic® or Rybelsus® in the treatment of type 2 diabetes (referred to collectively as "active comparators" (11)). The following non-incretin antidiabetic agents are used as active comparators in this study: i) sulphonylureas (SUs), ii) sodium-glucose co-transporter 2 (SGLT-2) inhibitors, and insulin subdivided into iii) basal insulin (long acting and intermediate acting insulin) only and iv) basal + bolus insulin or premix insulin (i.e. intermediate- or long-acting combined with fast acting). In supplementary analyses, semaglutide use will be compared to each of these three antidiabetic drug classes individually (see section 9.7.3.1). Users of Ozempic® and Rybelsus® will be combined into one group (semaglutide cohort) as the active ingredient is identical and the only difference between the two are the route of administration. Patients who shift from Ozempic® to Rybelsus® or vice versa will be considered a continued user in the semaglutide cohort.
In a supplementary and Ozempic®-specific analysis, new users of Rybelsus® will be excluded from the main analysis (see section 9.7.3.1).

The use of active comparators balances covariates related to the indication for diabetes treatment between the two cohorts (i.e. the semaglutide cohort and the active comparator cohort) and thereby minimizes confounding by indication. Furthermore, the use of active comparators ensures the clinical relevance of the research question, that is, whether choosing Ozempic® or Rybelsus® over active comparators leads to a difference in the risk of pancreatic cancer.

The study follows a new-user design and is restricted to new users (12). This design prevents prevalent user bias and avoids adjustment for covariates in the causal pathway as all covariates used for controlling confounding are measured prior to treatment initiation. In addition, by only including new users of semaglutide and active comparators it is ensured that follow-up is started when patients in both cohorts are in need of treatment intensification corresponding to initiation of Ozempic® or Rybelsus® or active comparators. New users are identified based on the last 10 years of prescription data and includes patients with no prior use of semaglutide or of a specific drug in the active comparator drug classes at the level of the active ingredient name (see section 9.7.3.2).

Propensity score matching and trimming (13,14) are used to control for potential confounders that may not be adequately controlled by the use of active comparators alone. Propensity scores are used to match new users of semaglutide with new users of active comparators on a number of covariates

that may act as potential confounders including indicators of diabetes severity and diabetes treatment duration. Ozempic® and Rybelsus® initiation will be handled in two separate propensity score models. The use of propensity score matching allows for a transparent data analysis by allowing a direct assessment of whether the confounder adjustment works, i.e. whether measured covariates are balanced between cohorts. Trimming of the tails of the propensity score is used to remove patients who are treated/not treated with Ozempic® or Rybelsus® versus active comparators contrary to prediction.

A one-year lag-time period is used to avoid reverse causation. Hence, patients are followed from one year after they initiate treatment with semaglutide or the active comparators.

9.1.1 Primary endpoint

The endpoint used to address the primary objective is the occurrence of first time (i.e. incident) malignant neoplasm of pancreas as defined by relevant diagnostic codes (see section 9.3.1).

9.2 Setting

This non-interventional study is based on health care registries from three Nordic countries: Denmark, Sweden, and Norway.

9.2.1 Study Population

Time period for the study: the study period runs from when the first subject was dispensed semaglutide in Denmark, Sweden, and Norway (2018) until December 31, 2023.

The study population comprises users of semaglutide and active comparators from Denmark, Norway and Sweden. A cohort of new users of semaglutide and a cohort of new users of active comparators are defined in the registries. To be included in either of the cohorts, new users have to initiate treatment between the marketing date of semaglutide (Ozempic®) in the country of residence and December 31, 2022 and to comply with the inclusion and exclusion criteria given below (see section 9.2.2 and 9.2.3).

The term “new users” is used throughout the protocol to refer to patients who are first-time ever users of semaglutide or of an active comparator drug (at the level of the active ingredient name) based on the past 10 years of prescription data.

Patients are followed from one year after the date they initiate treatment with semaglutide or active comparators (see section 9.3.2) and until they are censored due to a first occurrence of pancreatic cancer, non-pancreatic cancer (except non-melanoma skin cancer), other conditions serving as exclusion criteria (except acute and chronic pancreatitis) (see section 9.2.3), death, emigration, switch from active comparators to semaglutide, or end of study period (December 31, 2023), whichever comes first. Outcomes occurring within the first year after cohort entry date (defined as

the dispensing date of the initial prescription) will be disregarded in the analysis (and consequently the patients will be excluded).

Patients who experience acute or chronic pancreatitis during follow-up (i.e., after cohort entry date) are not censored because a diagnosis of pancreatitis may be a precursor for pancreatic cancer. Censoring of such individuals might lead to incorrect attenuation of effect estimates.

To maintain a clear comparison, new users of active comparators are censored in the event of initiating semaglutide. Such patients will, however, subsequently contribute follow-up time in the semaglutide cohort.

As it is expected that a substantial proportion of new users of semaglutide have been treated with other incretin mimetic drugs before initiating semaglutide, patients in both cohorts are allowed to have been exposed to incretin mimetic drugs (GLP-1 analogues and DPP-4 inhibitors) prior to the cohort entry date. These criteria are subject to extensive sensitivity analyses (see section 9.7.3.2).

9.2.2 Inclusion criteria

To be eligible for the study, patients are required to meet the following inclusion criteria:

- Use of semaglutide or initiation of active comparators from when the first subject is dispensed semaglutide in the respective country (2018) until December 31, 2022. Initiation is defined as no previous prescription of a drug with the same active ingredient in the past 10 years as recorded in the national prescription registries. Previous use of the same class of drug is however accepted. Further, a patient who shift from Ozempic® to Rybelsus® or vice versa will be considered a continued user in the semaglutide group.
- At least two prescriptions of either semaglutide or active comparators (at the level of the active ingredient) with the second prescription filled less than one year after the initial prescription. Two prescriptions filled at first day of treatment will, however, only be counted as one prescription.
- ≥18 years old at the cohort entry date. The cohort entry date is defined as the dispensing date of the initial prescription (i.e. treatment initiation).
- Ten years or more of continuous residency in the country of residence before initiation of semaglutide or active comparators.

9.2.3 Exclusion criteria

Patients with rare but strong risk factors for developing pancreatic cancer are excluded (15,16). Therefore, patients with any history of the following conditions (identified according to 10th version of International Classification of Diseases [ICD-10]) before cohort entry date (based on the last 10 years of data or more) are excluded:

- A history of any cancer except non-melanoma skin cancer

- Acute pancreatitis
- Alcohol-induced chronic pancreatitis
- Other chronic pancreatitis
- Cystic fibrosis
- Other phacomatoses, not elsewhere classified (including Peutz-Jeghers syndrome and Von Hippel-Lindau syndrome)
- Neurofibromatosis (non-malignant, Von Recklinghausen disease)

In addition, the following patients are excluded:

- Patients with a history of pancreatic cancer as these patients are no longer at risk for an incident pancreatic cancer
- Patients who initially started insulin as first-line antidiabetic treatment and who did not add on or switch to non-insulin treatment within 3 months. Those initiating insulin as first-line treatment during the study period is disregarded in terms of study inclusion, although they may be included upon initiating additional therapy, provided it is started within three months.
- Patients with polycystic ovary syndrome (PCOS)
- Patients with gestational diabetes
- Patients who meet one of the conditions used as exclusion criteria listed above (except acute and chronic pancreatitis) after cohort entry date but before start of follow-up (i.e. during the first year after treatment initiation)

9.2.4 Rationale for selection criteria

Patients are required to fill at least two prescriptions of semaglutide or active comparators within one year to ensure actual use of the antidiabetic agent and to limit the possibility of non-adherence where patients fill a prescription but do not administer the drug. In addition, patients who only fill one prescription of semaglutide will only have had a limited exposure to semaglutide.

Patients are required to have 10 years or more of continuous residency in the respective countries before initiation of semaglutide or active comparators to ensure at least 10 years of medical history in the health care registries. This ensures correct identification of new users of semaglutide and active comparators (based on the last 10 years of data), as well as exclusion criteria, T2DM severity, and other risk factors and covariates used for confounder adjustment.

Patients with strong risk factors for developing pancreatic cancer at cohort entry date and before follow-up are excluded as these patients are at substantially increased risk for developing the primary outcome compared to other patients (17).

The extent to which these criteria will lead to exclusions in the present study is estimated based on baseline patient characteristics obtained from a recent Danish nationwide cohort study, during 2008-2018, including new users of GLP-1 analogues (N: 27,808 GLP-1 users) (18). From the data

material underlying this study, it was for the purpose of this protocol identified that 7.6 % of GLP-1 initiators had a history of cancer (excl. skin cancer), 1.5 % and 0.6 % had a history of acute- and chronic pancreatitis, respectively, 0.8 % had a history of PCOS, and finally 0 % (N: 5) of the population had a history of gestational diabetes (within 365 days of GLP-1 initiation). The patterns observed in this Danish study are expected to be similar in Norway and Sweden. Remaining exclusion criteria are generally rare and their contribution to the overall proportion of semaglutide users that will be excluded is considered negligible.

For transparency in the data analysis, a description of patients excluded due to fulfilling any of these exclusion criteria will be presented. Furthermore, an attrition figure will be produced showing patients who are excluded from the study.

To ensure that study findings are applicable to all users of semaglutide, a sensitivity analysis is planned a priory to include the largest group of patients expected to be excluded due to exclusion criteria, which is patients with "a history of any cancer except non melanoma skin cancer" (sensitivity analysis 9, section 9.7.3.2).

Patients are not required to have an "administrative" diagnosis of T2DM in the health care registries (based on hospital contacts) to be included in the study. There are two reasons for this. First, T2DM patients are incompletely registered in the hospital-based patient registries due to the substantial number of T2DM patients diagnosed and treated in the primary health care sector only. Restriction to those with a registered T2DM diagnosis would thereby strongly diminish the study population. Second, the two cohorts of new users of semaglutide and active comparators will almost exclusively comprise patients with T2DM as they are defined by drugs that are all (except insulin) only licensed for use in T2DM. Patients who initially started insulin as first-line antidiabetic treatment and who did not initiate non-insulin treatment within 3 months are excluded to limit inclusion of patients with type 1 diabetes. Patients receiving antidiabetic treatment due to PCOS or gestational diabetes are excluded as these patients are not part of the target population, i.e. they do not have T2DM. Based on the inclusion and exclusion criteria, the study population is considered to be fully representative of the target population of T2DM patients who qualify for treatment with semaglutide or active comparators in Denmark, Norway and Sweden.

9.3 Variables

9.3.1 Outcome

The outcome is an incident pancreatic cancer diagnosis in the national cancer and/or patient registries as defined by the following diagnostic codes according to the International Classification of Diseases 10th version (ICD-10):

- C25: Malignant neoplasm of pancreas
  - C25.0: Head of pancreas

- C25.1: Body of pancreas
- C25.2: Tail of Pancreas
- C25.3: Pancreatic duct
- C25.4: Endocrine pancreas
- C25.7: Other parts of pancreas
- C25.8: Overlapping lesion of pancreas
- C25.9: Pancreas, unspecified

An incident cancer diagnosis is defined based on the last 10 years of data or more.

MedDRA terms for the outcome are included in Annex 1.

9.3.2 Exposure

The national prescription registries in each of the three countries are used to identify exposure defined by filled prescriptions for semaglutide or active comparators.

'Ever use' (i.e. having filled ≥2 prescriptions of semaglutide or active comparators within one year) is considered as the main exposure.

As part of the definition of what may be considered relevant "exposed time" for the development of pancreatic cancer, a one-year lag-time period is applied to avoid the possibility of reverse causation (19) and because of the assumed long induction time and latency period for pancreatic cancer (20-22). It is unlikely that a pancreatic cancer occurring in the first year after cohort entry date (i.e. treatment initiation) can be causally attributed to exposure to semaglutide or active comparators. It has previously been shown that a lag-time period of at least 6 months is enough to reduce bias from reverse causation in general (23) and specifically for pancreatic cancer (24).

Considerations regarding lag-time and exposure definitions are revisited in sensitivity and supplementary analyses (see section 9.7.3).

9.3.3 Covariates

It is expected that the active comparator cohort to a large extent minimizes confounding related to diabetes (confounding by indication). However, to adjust for confounding not handled by use of the active comparator cohort, known risk factors for pancreatic cancer (15,16), other confounders, and proxies for diabetes severity and diabetes treatment duration (e.g., type of diabetes treatment, number of antidiabetic agents used prior to cohort entry date and days since the dispensing date of first antidiabetic prescription) are included in the propensity score model.

The list below provides a crude overview of the known risk factors for pancreatic cancer and potential confounders that are included in the propensity score model. The final list of covariates to be included in the propensity score model and corresponding definitions (diagnosis codes etc.) will

be compiled in collaboration with research partners from Denmark, Sweden and Norway and included in the common SAP (see section 9.6).

- Demographic and sociodemographic variables:
  - Sex
  - Age
  - Education
  - Income
- Indicators of diabetes severity and diabetes treatment duration:
  - Diabetes complications
  - Number of hospital (in- and outpatient) contacts for diabetes
  - Duration of antidiabetic treatment (based on the date of filling the first antidiabetic agent in the national prescription registries)
  - Number of different antidiabetic agents used 180 days before cohort entry date
  - Type of antidiabetic agents used 180 days before cohort entry date
- Medication use (including medications known to or suspected to be associated with pancreatic cancer), e.g.:
  - Non-steroidal anti-inflammatory drugs (NSAIDs)
  - Low-dose aspirin
  - Statins
  - Selective serotonin reuptake inhibitors (SSRIs)
  - Medications used for alcoholism
- Comorbidities (including indicators for frailty and proxies for lifestyle risk factors):
  - Comorbidity index
  - Cardiovascular disease, including congestive heart failure
  - Gall stone disease (risk factor for pancreatic cancer)
  - Helicobacter pylori infection (risk factor for pancreatic cancer)
  - Obesity (diagnosis codes)
  - Alcohol-related diseases (proxy for alcohol consumption)
  - Smoking-related diseases (proxy for smoking)

To minimise confounding, it is essential that new users of semaglutide and new users of active comparators are matched according to their T2DM treatment stage, which is considered a strong indicator of T2DM severity. The inclusion of the markers of T2DM severity listed above in the propensity score model ensures the comparison of patients in the semaglutide cohort and the active comparator cohort, who are at a similar stage in the treatment of their T2DM. Importantly, the number and type of antidiabetic medications used 180 days prior to cohort entry and the duration of antidiabetic treatment before cohort entry is included in the propensity score model.

The covariates listed above are captured in the relevant national patient and prescription registries in the three countries. Data from Statistics Denmark, Statistics Norway and Statistics Sweden are used

to capture demographic and sociodemographic variables. Comorbidities are measured up to 10 years prior to cohort entry date, while current medication use is measured up to 180 days prior to cohort entry date. Duration of antidiabetic treatment is measured based on all data available prior to cohort entry date. Number of hospital diabetes contacts and income are measured in the one year prior to cohort entry date.

9.4 Data sources

National prescription-, cancer-, patient-, and cause of death registers in Denmark, Sweden, and Norway are used in this study (Table 9–1). In addition, data from Statistics Denmark, Statistics Norway and Statistics Sweden are used to capture demographic and socioeconomic variables. These countries are chosen due to the high data quality, the long history of health care registrations, the possibility of linkage between registers using the unique person identifier in each of the three countries, and the nationwide coverage of the registries.

The unique person identifier was introduced more than 50 years ago in the Nordic countries (25). It is assigned to all residents and makes it possible to conduct accurate population-based register-linkage studies with linkage of e.g., prescription and diagnosis data.

Data collection

In Denmark, data in the health care registries are gathered by Statistics Denmark, which is a government institution collecting electronic records for a broad spectrum of statistical and scientific purposes. The Danish Civil Person Register covers every Danish resident and contains data on vital status (date of birth and death) and migrations to and from Denmark since 1968 (26).

In Sweden, individual data regarding health care is obtained from the National Board of Health and Welfare (27) which is responsible for the Swedish Patient Register, the Swedish Cancer Register, and the Swedish Prescribed Drug Register. The Population Registers (LISA and total population register) have been registered with Statistics Sweden since 1968 and covers data on the total population, e.g. individual-level data on death and migrations (total population register) and income and education (LISA) (28).

In Norway, data are gathered by different institutions depending on the given register. The Norwegian Prescription Database (NorPD) and the Cause of Death register is gathered by the Norwegian Institute of Public Health. The Patient Registry (NPR) and Norwegian Registry for Primary Health Care (KUHR/KPR) is gathered by the Directorate of Health. Data on cancer is gathered by the Norwegian Cancer Registry. Socioeconomic data is gathered by Statistics Norway. The National Registry in Norway (Folkeregisteret) covers all Norwegian residents and contains data on vital status (date of birth and death) and migrations to and from Norway (29).

Informed consent is not needed for secondary analyses of already collected data in any of the three countries.

Prescription registers

The national prescription registries in Denmark, Sweden and Norway are very similar with regard to content and data structure (25,30). They are all nationwide and cover all prescription drugs dispensed to the entire population at pharmacies. Importantly, the completeness and accuracy of recording is high.

The registries contain a wide range of information on the individual prescriptions filled which fall into four main categories: patient-specific data, prescriber data, drug data, and pharmacy data. Of relevance to the present study, the national prescription registries in the three countries contain data on the unique person identifier, date of dispensing, the Anatomical Therapeutic Chemical (ATC) code of the dispensed drug, product name, formulation, and the volume of the dispensed drug in Defined Daily Doses (DDD).

Non-prescription over-the-counter (OTC) drugs or drugs administered at hospitals or in nursing homes are not included in the prescription registries, and the indication for drug use and prescribed dose is not easily available for research purposes in any of the three countries.

The nationwide coverage of the national prescription registries makes it possible to conduct studies based on these data sources with no selection bias. In addition, the long history of the prescription registries in all three countries makes it possible to apply a new-user design based on at least 10 years of data.

Cancer registers

The Nordic cancer registries are among the oldest population-based registries in the world (31). Registration of cancer is mandatory in the Nordic countries and high validity has been documented (31). Due to a close collaboration there is a high degree of similarity between the Nordic cancer registries and cancer data from the Nordic countries are overall comparable (31). The cancer registries in each of the three countries receive information on cancer cases from multiple data sources (e.g. public hospitals, private clinicians, etc.) with some variation between the three countries (e.g. dentists, death certificates, etc.) (31). Some of the variables included in the cancer registries in Denmark, Sweden and Norway are the unique person identifier, date of death, date (or month and year) of diagnosis, topography, morphology, and behavior/malignancy of the tumor (31). In all the Nordic countries, topography, morphology, and behavior/malignancy of the tumor are currently coded using the International Classification of Diseases for Oncology, 3rd version (ICD-O-3) (31). Older cancer cases might be coded according to older versions of the ICD (31). The topography section in the cancer registries use the same codes as the malignant neoplasm section of the ICD-10 classification system (C00-C80) (32).There is a high completeness of key data on incident cancer cases in the Nordic cancer registries and around 94-98% of cases in the cancer registries in Denmark, Sweden and Norway are microscopically verified (31).

The long history and nationwide coverage of the national cancer registries makes it possible to identify true incident cancer cases.

Patient registers

The national patient registries in Denmark, Sweden, and Norway contain administrative data on all hospital admissions and contacts to outpatient clinics or specialist visits. The registries contain information on date of contact/admission, date of closure of contact/discharge, diagnoses (including cancer diagnoses) and procedure codes. Diagnoses are coded using the ICD-10 classification system and procedures are coded using the Nordic Medico-Statistical Committee (NOMESCO) classification of surgical procedures.

Except for the Norwegian Patient Registry, the patient registries in Denmark and Sweden do not contain information on visits to private practicing specialists working outside hospitals.

Besides data from the Norwegian Patient Registry, data in Norway will also include data from the Norwegian Registry for Primary Health Care which includes information on all visits to general practitioners including International Classification of Primary Care (ICPC) diagnostic codes. This will provide information on patients diagnosed by general practitioners and will give further information on different covariates, e.g. cardiovascular disease. The registry can also give information on patients diagnosed with diabetes in primary care. The long history and nationwide coverage of the national patient registers makes it possible to accurately assess comorbidities that could act as potential confounders, and to accurately exclude patients who are already at an increased risk of pancreatic cancer prior to cohort entry date and start of follow-up.

Data on socioeconomic factors and cause of death

Data on socioeconomic factors (education and income) are available from socioeconomic registers at Statistics Denmark (33), Statistics Norway and Statistics Sweden. Data on cause of death are available from cause of death registers in the three countries.

Table 9–1 Overview of health care and population registers in Denmark, Sweden, and Norway of relevance for this study

Country	Coverage (size of population)	Registers	Start of register
Denmark	Nationwide (5.7 million)	The Registry of Medicinal Products Statistics (34)	1995
		The Danish National Patient Registry (35,36)	1977 (hospital admissions) 1995 (outpatient contacts)
		The Danish Civil Person Register (26)	1968
		The Danish Cancer Registry (31,37)	1942 (complete from 1943)
		Cause of death registry	1970
		Income and education registries from Statistics Denmark	-
Sweden	Nationwide (10.2 million)	The Swedish Prescribed Drug Register (25)	July 2005
		The Swedish National Patient Register (38)	1987 (hospital admissions) 2001 (outpatient contacts)
		The Population Registers (28)	1985 (education) 1968 (income)
		The Swedish Cancer Register (31,39)	1958
		Cause of death register (40)	1952
Norway	Nationwide (5.3 million)	The Norwegian Prescription Database (25)	2004
		The Norwegian Patient Register (41,42)	2008 (hospital admissions and outpatient contacts)
		The National Registry (Folkeregisteret) (29)	1964
		The Cancer Registry of Norway (31,43)	1952 (complete from 1953)
		Norwegian Registry for Primary Health Care (KUHR/KPR) (42)	2006

Country	Coverage (size of population)	Registers	Start of register
		Cause of death registry (44)	1951
		Income, education and status (immigration, emigration, death) data from Statistics Norway	-

In this study, pancreatic cancers are identified from the national cancer- and patient registries listed above. The national cancer registries in Denmark, Sweden, and Norway are used to identify pancreatic cancer cases and the national patient registers are used to identify pancreatic cancer cases from the date (expected December 31, 2022) when data from the cancer registries are no longer available due to latency in data availability.

The cause of death registers will be used in a supplementary analysis where death from pancreatic cancer is considered the main outcome (see section 9.7.3.1).

9.4.1 Data validity

There is a long tradition for registry-based research in the Nordic countries and all registers have been widely used for research purposes within pharmacoepidemiology (25). The data quality and validity of the national prescription-, patient-, cancer-, and cause of death registers mentioned above are generally expected to be high. However, to assess the validity of the use of the national patient registries for outcome ascertainment during the last year of the study period (i.e. 2023) (as described above), data on pancreatic cancer is abstracted from the national cancer registries for the years available, and a cross tabulation is performed comparing the rates obtained in the cancer registries to those obtained using the patient registries. This allows for a quantification of the potential under- or over ascertainment of outcomes when using the national patient registries to identify cases of pancreatic cancer during the last year of the study period.

Furthermore, a sensitivity analysis will be conducted where the analysis is based only on data available from the cancer registers (see section 9.7.3.2).

9.5 Study size

The assessment of the statistical precision is based on an estimate of the number of persons who are expected to develop pancreatic cancer while being exposed to semaglutide overall (including users of Ozempic® or Rybelsus®), according to the exposure definition and assuming a null association. A conventional power calculation as done in studies with primary data collection is not performed here, as the data sources used in this study are nationwide registries with little or no possibility of recruiting further patients exposed to semaglutide over the duration of the study. Hence, the

objective of the power assessment is rather to assess what level of statistical precision on the primary estimate is achievable within the available data.

The study population will be the cohort of all new users of semaglutide or active comparators in Denmark, Norway and Sweden initiating treatment from the marketing date of semaglutide in the respective country (2018) until December 31, 2022. It will be possible to enroll new users of the active comparator drugs in far larger numbers than semaglutide initiators. As a consequence, the dominant contribution of statistical uncertainty in the study stems from the number of pancreatic cancer cases occurring among semaglutide users, while the contribution of uncertainty from active comparator users is negligible. Hence, the estimation of the expected number of pancreatic cancer cases among users of semaglutide forms the cornerstone of the sample size estimation presented here, and is carried out by making realistic assumptions concerning expected background incidence of pancreatic cancer in the target population and taking into consideration observed as well as projected market uptake of semaglutide (Ozempic® and Rybelsus®) as described in further detail in the following sections.

9.5.1 Background incidence of pancreatic cancer in the target population

The target population is T2DM patients in scope for treatment with semaglutide or active comparators in the Nordic countries. The expected background incidence in this target population is estimated to be approximately 72 events of pancreatic cancer per 100,000 person years. This estimate is based on the background incidence of pancreatic cancer in the relevant age group in the general population multiplied by the relative risk for pancreatic cancer in patients with T2DM as compared to the general population, as outlined below:

- Based on Danish MEDSTAT data (45) from 2018, the majority of semaglutide users are in the age interval 50-74 years, and this is therefore considered the relevant age group. Background age- and sex specific incidence rates of pancreatic cancers in the general population (in Denmark, Sweden, Norway, Finland and Iceland) are provided in the NORDCAN data source (46). The age-specific population incidence rate is 10 per 100,000 person-years for the 50-year-olds and 80 per 100,000 person-years for the 74-year-olds, and hence a conservative average estimate of the incidence rate for the relevant age group for semaglutide would then be approximately 40 per 100,000 person-years. As the incidence rate for pancreatic cancer appears to be largely independent of sex, sex-differences in pancreatic cancer incidence or semaglutide utilization will not be taken into account in the power calculations.

- Finally, the incidence rate of pancreatic cancer among patients with T2DM is estimated to be 1.82 times higher than that of the background population (47). Therefore, the expected background incidence of pancreatic cancer in the target population is expected to be 40/100,000 person-years * 1.82 = 72.8 per 100,000 person years.

9.5.2 Observed and projected market uptake of semaglutide

The expected market uptake of semaglutide in Denmark, Sweden and Norway is inevitably subject to some uncertainty. The observed and projected sales of semaglutide in the three countries for the years 2018-2022 are listed in Table 9–2 below:

Table 9–2 Observed and projected sales of semaglutide* in Denmark, Sweden and Norway (2018-2022)

Year	Person years of semaglutide treatment in each year
2018	2,000
2019	24,000
2020	52,000
2021	85,000
2022	117,000

*Covering both Ozempic® and Rybelsus®

The figures presented in the table above are the observed and projected person-years of treatment (assuming patients treated for a full year) and are based on sales figures and sales forecasts (volume estimates). The numbers do therefore not correspond to total number of patients initiating semaglutide but reflect the hypothetical scenario where all patients initiating semaglutide remain treated. The sample size calculations presented in Table 9–3 below are based on the above assumption (i.e. that a patient initiating semaglutide in e.g. 2019 will stay treated for the remainder of the year as well as for consecutive years).

9.5.3 Sample size calculation

The sample size calculation is based on the assumptions about background incidence of pancreatic cancer and the observed and expected market uptake of semaglutide outlined above. It is furthermore assumed that the number of pancreatic cancer cases occurring among semaglutide users will conform to a Poisson distribution, as there is an underlying Poisson process; each individual in the population has a small but finite probability of developing pancreatic cancer.

The sample size calculation is presented in Table 9–3 and explained in corresponding steps below:

1. The observed and projected market uptake of semaglutide (Table 9–2) is used to calculate the cumulative follow-up time (in person years) among semaglutide users. The lag-time of one year is applied such that e.g. for 2019, only 2,000 out of 24,000 users have relevant observation time for pancreatic cancer. The observed and projected market uptake of Ozempic® and Rybelsus® are added to present one cohort of semaglutide users under the assumption that 1/3 of Rybelsus® initiators are previous Ozempic® users.

2. The estimated incidence of pancreatic cancer in the target population is applied, 72.8 per 100,000 person years, thereby yielding the expected number of pancreatic cancers among semaglutide users.
3. The confidence interval of a null-estimate is calculated by applying an exact Poisson confidence interval around the expected count of pancreatic cancer cases among semaglutide users.
4. Finally, the detectable relative hazard for pancreatic cancer for semaglutide vs. active comparators is calculated using the Method of Schoenfeld (1982) for a proportional hazards regression model and a two-sided test, under the assumptions of 80% power, a statistical significance level of 5% and that semaglutide users are matched 1:4 to active comparators.

Table 9–3 Sample size calculation and detectable relative hazard

Step	(1)	(1)	(1)	(2)	(3)	(4)
Year	Annual projected amount of semaglutide exposure in Denmark, Norway, and Sweden, Person-years	Excl. lag-time: Annual projected amount of semaglutide exposure after excluding a 1-year lag-time, Person-years	Running total of cumulative semaglutide exposure excluding lag-time, Person-years	Expected number of pancreatic cancers among semaglutide exposed, cumulative	Confidence interval for a null-estimate (hazard ratio semaglutide vs. active comparators)	Detectable relative hazard (semaglutide vs. active comparators)
2018	2,000	0	0	0		
2019	24,000	2,000	2,000	1	1.00 (0.03 - 5.57)	14.12
2020	52,000	24,000	26,000	19	1.00 (0.60 - 1.56)	2.06
2021	85,000	52,000	78,000	57	1.00 (0.76 - 1.30)	1.52
2022	117,000	85,000	163,000	119	1.00 (0.83 – 1.20)	1.33
2023		117,000	280,000	204	1.00 (0.87 – 1.15)	1.25

Based on the expected number of outcomes occurring among the semaglutide users, an acceptable precision is reached after five years study duration. The expected confidence interval for a null estimate is 0.87-1.15 in 2023. The proposed five-year study duration will thus provide a sufficient precision of the primary estimate.

Based on the sample size calculation, it is estimated that after five years, the study will have 80% power to detect a relative hazard of 1.25 (or a 25% increase in risk) between cohorts exposed to semaglutide and active comparators, respectively.

As aforementioned, the power of this study is largely driven by the number of pancreatic cancer cases occurring among semaglutide users and hence the number of patients exposed to semaglutide in Denmark, Norway and Sweden. The sample size calculations listed in Table 9–3 are based on the assumption that patients who are prescribed semaglutide remain treated for the remainder of the follow-up, which results in the lowest possible estimate of the total number of patients exposed to semaglutide. Based on experience from other GLP-1 analogues (48), a substantial proportion of semaglutide initiators are however expected to be short-term users. Therefore, the total number of patients initiating semaglutide will most likely be higher than assumed in the sample size calculation. For this reason, the achievable precision calculated above (Table 9–3) is considered a conservative estimate.

9.6 Data management

The data extraction will follow a common SAP for all three countries which will fully detail the conduct of the study, including the analytical approach, operational definitions, as well as shell tables and figures to be presented. All extraction and cleaning of data occurs locally in each country. Each country performs their own national analysis by using a common data model and finally a meta-analysis is performed by the coordinating study entity. Stata Version 15 or higher ([redacted]) and SAS® Version 9.4 ([redacted]) is used for statistical analysis.

In the initial phase of the project, a data extraction is performed in Denmark. Based on this data extraction, a common data model and programming of analysis is developed by the coordinating study entity in close collaboration with research collaborators in Norway and Sweden. In the final phase of the project, a full data extraction is performed in all countries and national analyses are conducted. Finally, based on the three national analyses, a meta-analysis is performed by the coordinating study entity to obtain the final study results (see section 6 (Milestones)).

9.7 Data analysis

9.7.1 Definition of analysis sets

The three national datasets used for analysis contains all patients in Denmark, Sweden and Norway, respectively, who are new users (based on the last 10 years of prescription data) of semaglutide or active comparator drugs from the time the first subject is prescribed semaglutide in the respective country (2018) until December 31, 2022.
The datasets contain prescription and diagnosis data on the study population from the beginning of each register and until December 31, 2023. Due to latency in cancer data availability (in Sweden and Denmark), cancer data are available from the cancer registers until the end of 2022 in Denmark and Sweden, and until the end of 2023 in Norway. For the year 2023, cancer data from Denmark and Sweden are extracted from the national patient registers.

New users will be followed from one year after the date they fill the first prescription in the study period until they are censored due to a diagnosis of pancreatic cancer, non-pancreatic cancer (except non-melanoma skin cancer), other conditions serving as exclusion criteria (except acute and chronic pancreatitis), death, emigration, or end of study period (December 31, 2023), whichever comes first.

For each subject included in the study population, the dataset will contain information from the last 10 years or more on all hospital discharge diagnoses and all co-medication filled at pharmacies by the patients.

In addition, the dataset will contain information on vital status, and socioeconomic and demographic factors (sex and age), migration and cause of death.

9.7.2 Statistical methods

The datasets from Denmark, Sweden, and Norway are considered homogeneous due to a uniform organization of the hospital sector and similar characteristics of health care databases. Data is analyzed separately in each country and a meta-analysis is performed.

Tables of the established cohort of new users of semaglutide and the matched cohort of new users of active comparators are presented using descriptive statistics.

New users of semaglutide are matched to new users of active comparators by the propensity score which is modelled on the covariates listed in section 9.3.3. Of note, Ozempic® or Rybelsus® initiation will be handled in two separate propensity score models. In the construction of the propensity score models, users of insulin are subdivided into two types: those using i) basal insulin only and ii) those using basal + bolus insulin or premix insulin. This ensures balance in the use of these types of insulin and not only the overall use of insulin, which is considered necessary in order to balance the baseline disease severity according to disease stage reflected in the specific insulin type.

The propensity score is calculated separately for patients in Denmark, Sweden and Norway and separate propensity score models are constructed for specific calendar intervals as variation over time in how the specified covariates affect the propensity to receive semaglutide versus active comparators is expected.

The specific propensity for initiating treatment with semaglutide (as opposed to the active comparators) is estimated using logistic regression, with semaglutide initiation as the outcome and the potential confounders and covariates outlined in section 9.3.3 as independent variables.

Each patient initiating semaglutide is matched 1:4 to patients initiating active comparators. This ratio is used unless this results in >5% of semaglutide patients being excluded after trimming. In this case, a matching ratio of 1:3 is used. If this still results in >5% of semaglutide patients being excluded after trimming a ratio of 1:2 is used. This rule is applied in each specific country, that is, the matching ratio will be allowed to differ between countries. Pairwise nearest neighbor matching with a caliper of 0.01 (without replacement) is used. This means that the distance between the propensity score in those initiated with semaglutide and those initiated with active comparators can be up to 0.01.

The performance of the propensity score for handling confounding is judged via estimation of standardized mean differences of the covariates after matching. A difference of less than 0.1 has been taken to indicate a negligible difference in the mean and can thus be considered indicative of balanced covariates (49).

Besides matching on the propensity score, which leads to trimming of non-overlapping regions of the propensity score distributions, patients who are treated/not treated with semaglutide versus

active comparators contrary to prediction are removed by "asymmetric" trimming of the tails of the propensity score. By using a cut-off of 2.5%, patients below the 2.5th percentile for the semaglutide-treated and above the 97.5th percentile for the active comparator-treated are excluded from the analysis. For transparency in the data analysis, a description of patients excluded by trimming, and patients entering the trimmed and untrimmed cohorts as well as matched and unmatched cohorts is presented. Furthermore, an attrition figure will be produced showing patients who are excluded from the study.

The incidence rates of pancreatic cancer are estimated in the semaglutide and active comparator cohorts and used to estimate both the relative and absolute risk of pancreatic cancer associated with use of semaglutide compared to use of active comparators.

Following trimming and matching, the hazard ratio with two-sided 95% confidence intervals for pancreatic cancer are estimated by using a Cox proportional hazards model. As covariates are balanced via the matching procedure, no further adjustment for confounding is included in this model.

Data analysis and statistical analysis will be described in further detail in the common SAP (see section 9.6). In-depth knowledge on the included registries is pivotal to the development of the SAP, and therefore it will be completed in collaboration between research collaborators in Denmark, Norway and Sweden.

9.7.3 Supplementary and sensitivity analyses

9.7.3.1 Supplementary analyses

A number of supplementary analyses are performed.

1. Semaglutide versus individual active comparators:

To facilitate the interpretation of results, the main analysis is split into four sub-analyses, where new users of the different active comparators are compared separately to new users of semaglutide. In these analyses, new users of semaglutide are compared to new users of i) SUs, ii) SGLT-2 inhibitors, and insulin subdivided into iii) basal insulin only and iv) basal + bolus insulin or premix insulin. The rationale for the subdivision of insulin is to enable separate analyses covering both the scenario where Ozempic® or Rybelsus® is used at the same treatment stage as basal insulin and the scenario, where Ozempic® or Rybelsus® is used as add-on to basal insulin (i.e. comparable to basal + bolus insulin or premix insulin treatment). For these four sub-analyses, new propensity score models are estimated separately for each of the active comparators.

2. Semaglutide versus other incretin-mimetic antidiabetic drugs:

To contextualize results from the main analysis, users of semaglutide are compared to new users of other incretin-mimetic antidiabetic drugs (GLP-1 analogues and DPP-4 inhibitors) in two separate analyses (2.1 and 2.2 below). If the sample size allows it (i.e. sufficient number of users of other GLP-1 analogues) analysis 2.1 and 2.2 below are performed. If, however, there are too few new users of other GLP-1 analogues, pragmatically defined as the total number of eligible users (after exclusions) being lower than the number of eligible users of semaglutide, only analysis 2.2 is performed. If the number of users allows it, users are matched 1:4 as in the main analysis. Otherwise, a lower matching ratio (e.g. 1:2) is used, using the same decision rule for matching as outlined in the main analysis above.

2.1. New users of semaglutide are compared to new users of any other GLP-1 analogue (if the number of users of other GLP-1 analogues allows it).

2.2. New users of semaglutide are compared to new users of any other incretin-mimetic antidiabetic drug (GLP-1 analogues and DPP-4 inhibitors).

3. Death from pancreatic cancer as primary outcome:

For the purpose of this supplementary analysis, the primary outcome is changed to death from pancreatic cancer. Due to a concern of incretin mimetic drugs causing pancreatic cancer, it is possible that patients initiating semaglutide are followed more closely with regards to the development of pancreatic cancer. Similarly, known gastrointestinal side effects to semaglutide might be mistaken as symptoms that lead to diagnostic workup revealing a pancreatic cancer. For these reasons, a pancreatic cancer might be detected earlier in patients using semaglutide, thus moving a diagnosis of pancreatic cancer forward in time. This may lead to a spurious association between use of semaglutide and pancreatic cancer. By using death from pancreatic cancer as the primary outcome, detection bias stemming from an increased diagnostic activity may be reduced. This is because for the majority of patients diagnosed with pancreatic cancer it can be assumed that death from pancreatic cancer is usually not delayed by earlier detection and medical interventions, or at least, in most cases, only to a very limited extent.

Most patients with pancreatic cancer are detected in the advantaged stage (50). The only potential curative treatment of pancreatic cancer is surgical resection, however, only around 10-20% of patients are surgically resectable (50,51). Less than 20% of patients who undergo surgery survive five years, and the median survival after surgery is 12 – 19 months (51). Adjuvant chemotherapy after surgery improves survival by two to three months (51). Accordingly, data from the Danish Cancer Society estimates a survival rate of 7-9% five years after a pancreatic cancer diagnosis (52). This five-year survival rate is further supported by Rawla et al (50). Therefore, from a practical viewpoint, pancreatic cancer deaths are defined as all deaths in patients with a pancreatic cancer diagnosis. As a sensitivity analysis, deaths from pancreatic cancer will be identified based on the cause of death registers.

4. All-cause mortality as primary outcome:

In order to put the results for pancreatic cancer into a broader benefit/risk perspective, the primary outcome is changed from pancreatic cancer to all-cause mortality.

5. Dose-response analysis (cumulative dose and cumulative duration):

A dose-response analysis is performed where cumulative dose of semaglutide (instead of 'ever use' as in the main analysis) and 'ever use' of active comparators will be considered as the reference. Associations between drugs and cancers are, in case of a causal association, expected to be dependent on cumulative drug exposure such that higher exposure leads to a larger risk increase (17). To this end, follow-up in the semaglutide cohort is stratified by cumulative dose, defined as the amount of semaglutide filled while still applying a one year lag-time period, that is, users begin contributing person-time to a category of cumulative dose one year after crossing the threshold into that category.
Cumulative dose of Ozempic® is defined as very low (0-12.9 mg), low (13-25.9 mg), medium (26-52 mg), and high (>52 mg). The chosen cut-offs for very low, low, medium, and high cumulative dose reflect doses corresponding to <6 months of use of 0.5 mg once-weekly (the lowest recommended maintenance dose), 6 months-1 year of use of 0.5 mg once-weekly, 1-2 years of use of 0.5 mg once-weekly, and >2 years of use of 0.5 mg once-weekly, respectively. Of note, those using the higher dose (1.0 mg once-weekly) thus enter the higher strata of cumulative dose faster than those using lower doses. Cumulative dose of Rybelsus® will be transformed into cumulative dose of Ozempic® by dividing the cumulative dose of Rybelsus® by a factor 196. This is based on the US package insert of Rybelsus® indicating that 0.5mg Ozempic® corresponds to 14mg Rybelsus®. Thus, 1mg Ozempic® once-weekly corresponds to 196 mg ((14 mg*7)/0.5) Rybelsus® a week.

In addition to the dose-response analysis based on cumulative dose, an analysis considering cumulative duration of use is also performed. In this analysis, each filled prescription of Ozempic® (regardless of dose) is assigned a fixed duration of four weeks per package of Ozempic®. For Rybelsus®, the duration is based on the package size, i.e. the total number of tablets. The following package sizes are available: 30 and 90 tables, corresponding to 1 month and 3 months of treatment. Duration is calculated based on the number of tablets and adding 20% to take non-compliance into account. The cumulative duration of semaglutide use is defined as very short (<6 months), short (6 months-1 year), intermediate (1-2 years), and long (>2 years), while still applying a one-year lag-time period.

6. Ozempic-specific analysis:

An analysis is performed where new users of Rybelsus® are excluded from the study and only users of Ozempic® are included. In this analysis, users of Ozempic® switching to Rybelsus® will be censored upon switching.

7. Stratification by age and sex:

Finally, supplementary analyses investigating the primary outcome stratified by age and sex will be conducted.

9.7.3.2 Sensitivity analyses

A number of sensitivity analyses are performed to check the influence from the analytical/design choices on the study findings.

1. Censoring approach:

The censoring approach on the exposure in the main analysis is changed in two ways (referred to as I and II below).

I. A sensitivity analysis is performed where new users of active comparator drugs are censored not only upon initiating semaglutide (as in the main analysis), but also upon the initiation of other incretin mimetic drugs. In addition, contrary to the main analysis, users of semaglutide and users of active comparators are not allowed to have been exposed to incretin mimetic drugs (GLP-1 analogues or DPP-4 inhibitors) before cohort entry date. This sensitivity analysis is a stricter analysis that does not allow noise from a possible class-effect of incretin mimetic drugs to influence on the risk of pancreatic cancer, although at the cost of a reduced study sample size.

II. One potential problem in the main and sensitivity analysis I is that a substantial amount of follow-up may be lost due to censoring and (for the main analysis) that the handling of censoring is not symmetrical for semaglutide and its active comparators, since users of the active comparators are censored upon initiating semaglutide, while the opposite is not the case. Therefore, another sensitivity analysis is performed where new users of active comparators are not censored upon initiating semaglutide or other incretin mimetic drugs. Instead, users are classified as ever-users of both drug categories from the switch and onwards. For example, subjects who switch from SU to semaglutide are categorized as ever-exposed to both SU and semaglutide in an entirely time-dependent manner, i.e. first as ever-exposed to SU and then, from the date of the switching, also as ever-exposed to semaglutide. The cumulative amount dispensed for each drug category is tallied in a time-dependent manner, as the subjects cross new thresholds for cumulative exposure for each drug. In the analysis, the effect of semaglutide is estimated by multivariable Cox regression, i.e., with adjustment for use of the active comparators and for potential confounders.

2. No trimming of the propensity score:

Instead of first trimming and then matching on the propensity score, as done in the main analysis, an analysis is performed where trimming is omitted.

3. Restricting to patients ≤ 75 years of age:

The study population is restricted to patients ≤ 75 years of age. Elderly patients with T2DM might be treated differently compared to younger patients, and the diagnostic workup in relation to suspicion of a cancer diagnosis might be different. The incidence rate of pancreatic cancer provides evidence that the incidence of pancreatic cancer starts levelling off around the age of 75 years in the Nordic countries possibly reflecting that some frail patients are not diagnosed with their cancer (46).

4. Restricting to users with a record of metformin use:

The study population is restricted to study subjects with a record of metformin use in the national prescription registries based on the last 10 years before cohort entry (i.e. metformin ever users). This analysis aims to remove potential off-label users of semaglutide as well as semaglutide users with strong contraindications for some oral antidiabetics (e.g. kidney failure).

5. Restricting to current users of metformin:

The study is restricted to users of metformin at baseline, defined as users with one or more metformin prescriptions within 180 days before initiating semaglutide or active comparators. Users of metformin at baseline may be using other antidiabetic drugs in combination with metformin. This aims at restricting the study population to those that either add-on semaglutide or active comparators to metformin or change from metformin to semaglutide or active comparators.

6. Changing the lag time period:

The lag-time period is extended from one year to two years due to uncertainty in the assumptions regarding empirical induction time for pancreatic cancer.

7. Competing risk:

To address the potential influence of competing risks, death is included as a competing risk.

8. Continuous residency:

The inclusion criterion of “ten years or more of continuous residency in the country of residence before initiation of semaglutide or active comparators“ is changed to a criterion of at least two years or more of continuous residency. This sensitivity analysis aims to ensure that study findings are applicable to all users of semaglutide, including those migrating in, out and between the Nordic countries.

9. Allowing patients to have a previous cancer diagnosis:

The exclusion criterion of "a history of any cancer except non melanoma skin cancer" is removed from the analysis. This sensitivity analysis aims to ensure that study findings are applicable to all users of semaglutide, including those with a previous cancer diagnosis (except for pancreatic cancer).

10. Restricting to verified cancer cases:

The analysis is based only on data available from the national cancer registries. This sensitivity analysis will include only the verified pancreatic cancer cases from the national cancer registers.

9.8 Quality control

Data extraction will be performed locally by local data holders and thereafter delivered to the research collaborators within each country. Each country performs their own analysis according to a common data model and meta-analysis is performed by the coordinating study entity.

The coordinating study entity will collaborate extensively with research collaborators in order to ensure the highest level of quality of the data. Furthermore, by retrieving as much information as possible about how data are generated and extracted, the close collaboration with research collaborators will ensure a correct interpretation of data, taking into account any differences between data sources.

The research collaborators and the coordinating study entity will ensure that the standard operating procedures applicable for this type of study in their research unit are followed. The standard operating procedures include internal quality audits, rules for confidential data storage and storage of local study documents. Local study documents are data application and abstraction forms. All documents are written in English and will undergo local quality control as well as review by research collaborators. Finally, they will be reviewed and approved by the coordinating study entity.

SAPs and records of statistical programming performed to generate the results will be stored by the coordinating study entity.

The research collaborators and the coordinating study entity will file the necessary documents that would allow a replication of all data extractions and analyses performed in the course of the study. The documentation should clearly show what was done at which time point, to meet the requirements for full transparency. The documentation will be kept on file for at least 5 years after sign-off of the final study report.

9.8.1 Critical documents

Before the coordinating study entity performs the first data extract, the following documents must be available to Novo Nordisk:

- Regulatory approval of protocol and/or notification as required
- Documentation of the coordinating study entity's qualifications according to GPP (53) (for instance a short curriculum vitae (CV) or authorisation for each of the individuals involved)
- Signed and dated agreement on the final protocol
- Signed and dated agreement on any amendment(s), if applicable
- Non-interventional study agreement

9.8.2 Retention of study documentation

Novo Nordisk will comply with GPP and relevant national legislation related to archiving of study documentation.

The coordinating study entity must agree to archive the documentation pertaining to the study in an archive for at least 5 years after final report/first publication of the study, whichever comes later. The coordinating study entity should not destroy any documents without prior permission from Novo Nordisk.

Novo Nordisk will retain the documentation pertaining to the study according to company procedure and in accordance with national regulations if they require a longer retention period.

9.9 Limitations of the research methods

Limitations related to data sources

Drug exposure in this study is based on pharmacy data of dispensed drugs recorded in the national prescription registries. While data recording is expected to be complete and valid it is not possible to assess whether patients actually take/administer the dispensed drug (secondary adherence or non-adherence). However, as outlined in section 9.2.4, an inclusion criterion of two prescriptions within one year is expected to overcome this limitation to a large extent.

Furthermore, the indication for drug use and prescribed dose is not easily available for research purposes in the national prescription registries. Thus, this study relies on assumptions about indications for antidiabetic drug use. However, as outlined in section 9.2.4, there is good reason to assume that the study population to a large extent only includes patients with T2DM. Furthermore, the potential off-label use of semaglutide is addressed in sensitivity analyses (9.7.3.2). Should there be a small degree of erroneous inclusion of type 1 diabetes patients, this is unlikely to confer a bias as there is for in practice no difference in risk for pancreatic cancer between patients with T2DM and type 1 diabetes (15).

Limitations related to confounding

As this is a non-interventional study, potential confounding factors cannot be ruled out. Data collection will reflect routine clinical practice rather than mandatory assessments at pre-specified time points, which may have an impact on the amount of data and its interpretation.

As outlined above, control for confounding in the study is done through the use of an active comparator cohort and by use of the propensity score matching. However, residual confounding may still be present if covariates (e.g. unmeasured confounders) that are unevenly distributed between users of semaglutide and users of active comparators are not included in the propensity score model, or if they are included but only poorly measured in the health care registries.

Lifestyle risk factors such as alcohol use, smoking and high Body Mass Index (BMI) are considered as risk factors for pancreatic cancer (15,16). While alcohol use and BMI are considered weak risk factors for pancreatic cancer (15,54,55), smoking is considered a major risk factor (with an attributable risk of 25%) (16).

All of these lifestyle risk factors are, however, poorly measured in the health care registries. Hospital diagnosis of obesity, alcohol-related diseases, and smoking-related diseases are included in the propensity score model as crude proxies (see section 9.3.3). Some of these lifestyle risk factors will, besides diagnoses, be defined by the filling of certain prescription drugs. However, as they are only crude proxies, differences in these lifestyle risk factors between users of semaglutide and users of active comparators may not be fully accounted and controlled for. Importantly, however, there is no reason to suspect that users of semaglutide have a markedly different smoking history or different alcohol consumption compared to users of active comparators. Thus, the potential for confounding from smoking and alcohol use is mainly handled by use of the active comparator cohort.

High BMI might, on the other hand, be unequally distributed between users of semaglutide and active comparators (specifically users of insulin and SUs) as GLP-1 analogues are specifically known to induce weight loss in patients with T2DM. However, as high BMI is only a weak risk factor for pancreatic cancer, confounding by BMI is expected to be limited.

Limitations related to analytical methods

Although the propensity score offers many advantages there are a number of limitations. When matching on the propensity score, patients with non-overlapping propensity scores are excluded as there is no matching counterpart. This can lead to loss of information and a decrease in the precision of the estimated association. By trimming the tails of the propensity score and excluding more than the non-overlapping regions of the propensity score, patients in the outer range of the propensity score are excluded. This is appropriate if patients who are excluded are those with an absolute contraindication or absolute indication. However, trimming and matching may, in any case, limit the representativeness of the study population. This limitation should, however, be held up against the potential advantages of trimming and matching. In this study, patients excluded by trimming are carefully described, and trimmed and untrimmed as well as matched and unmatched populations are presented.

10 Protection of human subjects

The present study is a non-interventional study based on already existing registry data. Pseudonymized data of relevance to the objectives of the study is extracted for analysis.

The study is conducted in accordance with GPP (53) and any local regulations. Special attention is paid to data privacy protection and data protection as reflected in Directive (EU) 2016/680 of 27 April 2016 and its implementation in the national legislation by May 2018. Novo Nordisk will get access to the results of the study, but not to the data used in the study.

10.1 Informed consent form for study patients

In the Nordic countries, studies based solely on registry data do not require informed consent from individuals in the study population. Investigators involved in the study are governed by regional rules that guarantee the integrity of data and the privacy of individuals.

10.2 Institutional Review Boards/Independent Ethics Committee, health authorities and other relevant national institutions/bodies

Study specific documentation (study protocol, amendments, and the non-interventional study report) will be submitted to regulatory authorities as required by national requirements. Approval of the study protocol will be sought from EMA.

It is the responsibility of the research collaborators in each country to perform the required submissions and get the necessary approvals for the study from the relevant authorities, the national Data Protection Agencies and the relevant data owners. Ethical approval is not required in Denmark for purely registry-based studies. Ethical approval is required in Sweden and Norway.

11 Managing and reporting of adverse events/adverse reactions

This study is based on data already available in existing databases (secondary data) and single case collection and reporting from such studies is not required according to the current European Pharmacovigilance Regulations (Module VI (rev.2)).

12 Plans for disseminating and communicating study results

Study progress reports will be submitted to the EMA annually during the conduct of the study. Further, a final study report will be submitted to the EMA upon study completion (6).

12.1 Registration of study information

This study is subject to registration no later than at the time of first data extraction according to Novo Nordisk requirements for non-interventional study disclosure.

Non-interventional PASS must be registered in the EU Electronic Register of Post-Authorisation Studies (EU PAS Register) maintained by the European Medicines Agency and accessible through the European Medicines Agency's web portal.

Note: Study registration is regarded as the publication of an internationally agreed set of information (which can be found at the World Health Organization (WHO) homepage) about the design, conduct and administration of non-interventional studies. These details are published on a publicly accessible website managed by a registry conforming to WHO standards (also found at the WHO homepage); for example, www.clinicaltrials.gov.

12.2 Communication and publication

Novo Nordisk commits to communicating or otherwise making available for public disclosure results of studies regardless of outcome. Public disclosure includes publication of a paper in a scientific journal, abstract submission with a poster or oral presentation at a scientific meeting, or by other means.

At the end of the study, one or more public disclosures for publication may be prepared by the coordinating study entity and research collaborators in collaboration with Novo Nordisk. Novo Nordisk reserves the right to postpone publication and/or communication for 60 days to protect intellectual property. The results of this study will be subject to public disclosure on external web sites according to international regulations, as reflected in the Novo Nordisk Commitment to share information about clinical studies.

In accordance with Novo Nordisk commitment to transparency in clinical activities, this study will be registered by Novo Nordisk at www.clinicaltrials.gov and www.novonordisk-trials.com in accordance with the Novo Nordisk Commitment to share information about clinical studies.

In all cases, the study results must be reported in an objective, accurate, balanced and complete manner, with a discussion of the strengths and limitations of the study. The communication of study findings will follow ENCePP code of conduct and EnCePP study seal requirements. Authorship of publications should be in accordance with guidelines from The International Committee of Medical Journal Editors' Uniform Requirements (also referred to as the Vancouver Criteria) (56).

A contractual agreement is in place between Novo Nordisk and the coordinating study entity which details the publication of study results. Novo Nordisk maintains the right to be informed of any plans for publication by the coordinating study entity and to review any scientific paper, presentation, communication or other information concerning the investigation described in this protocol. Any such communication must be submitted in writing to the Novo Nordisk study manager 30 days prior to submission for comments. Comments will be given within the agreed timeframe (30 days) from receipt of the planned communication. Requests that interpretation of the results or their presentation be changed should be based on sound scientific reasons and follows the rules of EnCePP code of conduct.

13 References

1. American Diabetes Association. 8. Pharmacologic Approaches to Glycemic Treatment: *Standards of Medical Care in Diabetes—2018*. Diabetes Care. 2018 Jan;41(Supplement 1):S73–85.

2. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, et al. Management of hyperglycaemia in type 2 diabetes, 2015: a patient-centred approach. Update to a Position Statement of the American Diabetes Association and the European Association for the Study of Diabetes. Diabetologia. 2015 Mar;58(3):429–42.

3. Butler PC, Matveyenko AV, Dry S, Bhushan A, Elashoff R. Glucagon-like peptide-1 therapy and the exocrine pancreas: innocent bystander or friendly fire? Diabetologia. 2010 Jan;53(1):1–6.

4. Elashoff M, Matveyenko AV, Gier B, Elashoff R, Butler PC. Pancreatitis, Pancreatic, and Thyroid Cancer With Glucagon-Like Peptide-1–Based Therapies. Gastroenterology. 2011 Jul;141(1):150–6.

5. Butler PC, Elashoff M, Elashoff R, Gale EAM. A Critical Analysis of the Clinical Use of Incretin-Based Therapies: Are the GLP-1 therapies safe? Diabetes Care. 2013 Jul 1;36(7):2118–25.

6. Egan AG, Blind E, Dunder K, de Graeff PA, Hummer BT, Bourcier T, et al. Pancreatic Safety of Incretin-Based Drugs — FDA and EMA Assessment. N Engl J Med. 2014 Feb 27;370(9):794–7.

7. Nauck MA, Meier JJ, Cavender MA, Abd El Aziz M, Drucker DJ. Cardiovascular Actions and Clinical Outcomes With Glucagon-Like Peptide-1 Receptor Agonists and Dipeptidyl Peptidase-4 Inhibitors. Circulation. 2017 Aug 29;136(9):849–70.

8. Azoulay L, Filion KB, Platt RW, Dahl M, Dormuth CR, Clemens KK, et al. Incretin based drugs and the risk of pancreatic cancer: international multicentre cohort study. BMJ. 2016 Feb 17;i581.

9. Knapen LM, van Dalem J, Keulemans YC, van Erp NP, Bazelier MT, De Bruin ML, et al. Use of incretin agents and risk of pancreatic cancer: a population-based cohort study. Diabetes Obes Metab. 2016 Mar;18(3):258–65.

10. Boniol M, Franchi M, Bota M, Leclercq A, Guillaume J, van Damme N, et al. Incretin-Based Therapies and the Short-term Risk of Pancreatic Cancer: Results From Two Retrospective Cohort Studies. Diabetes Care. 2018 Feb;41(2):286–92.

11. Lund JL, Richardson DB, Stürmer T. The Active Comparator, New User Study Design in Pharmacoepidemiology: Historical Foundations and Contemporary Application. Curr Epidemiol Rep. 2015 Dec;2(4):221–8.

12. Ray WA. Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. Am J Epidemiol. 2003 Nov 1;158(9):915–20.

13. Stürmer T, Wyss R, Glynn RJ, Brookhart MA. Propensity scores for confounder adjustment when assessing the effects of medical interventions using nonexperimental study designs. J Intern Med. 2014 Jun;275(6):570–80.

14. Glynn RJ, Schneeweiss S, Sturmer T. Indications for Propensity Scores and Review of their Use in Pharmacoepidemiology. Basic Htmlent Glyphamp Asciiamp Clin Pharmacol Htmlent Glyphamp Asciiamp Toxicol. 2006 Mar;98(3):253–9.

15. Maisonneuve P, Lowenfels AB. Epidemiology of Pancreatic Cancer: An Update. Dig Dis. 2010;28(4–5):645–56.

16. Raimondi S, Maisonneuve P, Lowenfels AB. Epidemiology of pancreatic cancer: an overview. Nat Rev Gastroenterol Hepatol. 2009 Oct 6;6:699–708.

17. Pottegård A, Friis S, Stürmer T, Hallas J, Bahmanyar S. Considerations for Pharmacoepidemiological Studies of Drug-Cancer Associations. Basic Clin Pharmacol Toxicol [Internet]. 2018 Jan 15 [cited 2018 Jan 30]; Available from: http://doi.wiley.com/10.1111/bcpt.12946

18. Pottegård A, Lund LC, Henriksen DP, Folkestad L, Hellfritzsch M, Hallas J, et al. Use of dipeptidyl peptidase 4 inhibitors and risk of splanchnic vein thrombosis: A Danish nationwide new user active comparator cohort study. Diabetes Obes Metab. 2020 Nov 12;dom.14253.

19. Csizmadi I, Collet J-P, Boivin J-F. Bias and Confounding in Pharmacoepidemiology. In: Strom BL, editor. Pharmacoepidemiology [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2006 [cited 2018 Feb 27]. p. 791–809. Available from: http://doi.wiley.com/10.1002/9780470059876.ch47

20. Nadler DL, Zurbenko IG. Estimating Cancer Latency Times Using a Weibull Model. Adv Epidemiol. 2014;2014:1–8.

21. Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature. 2010 Oct 28;467(7319):1114–7.

22. Yu J, Blackford AL, dal Molin M, Wolfgang CL, Goggins M. Time to progression of pancreatic ductal adenocarcinoma from low-to-high tumour stages. Gut. 2015 Nov;64(11):1783–9.

23. Pottegård A, Hallas J. New use of prescription drugs prior to a cancer diagnosis: Use of Prescription Drugs Prior to Cancer Diagnosis. Pharmacoepidemiol Drug Saf. 2017 Feb;26(2):223–7.

24. Hicks B, Friis S, Pottegård A. Use of proton pump inhibitors and risk of pancreatic cancer. Pharmacoepidemiol Drug Saf. (in press).

25. Furu K, Wettermark B, Andersen M, Martikainen JE, Almarsdottir AB, Sørensen HT. The Nordic countries as a cohort for pharmacoepidemiological research. Basic Clin Pharmacol Toxicol. 2010 Feb;106(2):86–94.

26. Pedersen CB. The Danish Civil Registration System. Scand J Public Health. 2011 Jul;39(7 Suppl):22–5.

27. Socialstyrelsen. Socialstyrelsen [Internet]. Available from: http://www.socialstyrelsen.se

28. Statistika centralbyrån (SCB). Mikrodata för Registret över totalbefolkningen (RTB) [Internet]. Available from: https://www.scb.se/vara-tjanster/bestalla-mikrodata/vilka-mikrodata-finns/individregister/registret-over-totalbefolkningen-rtb/

29. Skatteetaten. Folkeregister [Internet]. Available from: https://www.skatteetaten.no/person/folkeregister/

30. Wettermark B, Zoëga H, Furu K, Korhonen M, Hallas J, Nørgaard M, et al. The Nordic prescription databases as a resource for pharmacoepidemiological research--a literature review. Pharmacoepidemiol Drug Saf. 2013 Jul;22(7):691–9.

31. Pukkala E, Engholm G, Højsgaard Schmidt LK, Storm H, Khan S, Lambe M, et al. Nordic Cancer Registries – an overview of their procedures and data comparability. Acta Oncol. 2017 Dec 11;1–16.

32. World Health Organization (WHO). International Classification of Diseases for Oncology. Third edition. [Internet]. Available from: http://apps.who.int/iris/bitstream/handle/10665/96612/9789241548496_eng.pdf;jsessionid=1E947DF48ACA0C1E94072E74F799551B?sequence=1

33. Statistics Denmark. Statistics Denmark [Internet]. Available from: dst.dk/en

34. Pottegård A, Schmidt SAJ, Wallach-Kildemoes H, Sørensen HT, Hallas J, Schmidt M. Data Resource Profile: The Danish National Prescription Registry. Int J Epidemiol. 2016 Oct 27;

35. Thygesen LC, Daasnes C, Thaulow I, Bronnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: Structure, access, legislation, and archiving. Scand J Public Health. 2011 Jul 1;39(7 Suppl):12–6.

36. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol. 2015;7:449–90.

37. Gjerstorff ML. The Danish Cancer Registry. Scand J Public Health. 2011 Jul 1;39(7 Suppl):42–5.

38. Socialstyrelsen. Patientregistret [Internet]. Available from: http://www.socialstyrelsen.se/register/halsodataregister/patientregistret

39. Socialstyrelsen. Swedish Cancer Registry [Internet]. Available from: http://www.socialstyrelsen.se/register/halsodataregister/cancerregistret/inenglish

40. Brooke HL, Talbäck M, Hörnblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Eur J Epidemiol. 2017 Sep;32(9):765–73.

41. Helsedirektoratet. Norwegian Patient Registry [Internet]. Available from: https://helsedirektoratet.no/english/norwegian-patient-registry

42. Bakken IJ, Ariansen AMS, Knudsen GP, Johansen KI, Vollset SE. The Norwegian Patient Registry and the Norwegian Registry for Primary Health Care: Research potential of two nationwide health-care registries. Scand J Public Health. 2020 Feb;48(1):49–55.

43. Larsen IK, Småstuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity and timeliness. Eur J Cancer. 2009 May;45(7):1218–31.

44. Pedersen AG, Ellingsen CL. Datakvaliteten i Dødsårsaksregisteret. Tidsskr Den Nor Legeforening. 2015;135(8):768–70.

45. The Danish Health Data Authority. MEDSTAT.dk [Internet]. Available from: http://medstat.dk/

46. Association of the Nordic Cancer Registries. NORDCAN [Internet]. Available from: http://www-dep.iarc.fr/NORDCAN/english/frame.asp

47. Huxley R, Ansary-Moghaddam A, Berrington de González A, Barzi F, Woodward M. Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. Br J Cancer. 2005 Jun 6;92(11):2076–83.

48. Pottegård A, Bjerregaard BK, Larsen MD, Larsen KS, Hallas J, Knop FK, et al. Use of exenatide and liraglutide in Denmark: a drug utilization study. Eur J Clin Pharmacol. 2014 Feb;70(2):205–14.

49. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivar Behav Res. 2011 May 31;46(3):399–424.

50. Rawla P, Sunkara T, Gaduputi V. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. World J Oncol. 2019 Feb;10(1):10–27.

51. De La Cruz MSD, Young AP, Ruffin MT. Diagnosis and management of pancreatic cancer. Am Fam Physician. 2014 Apr 15;89(8):626–32.

52. The Danish Cancer Society. [Statistics on pancreatic cancer] (in Danish) [Internet]. Available from: https://www.cancer.dk/bugspytkirtelkraeft-pancreascancer/statistik-bugspytkirtelkraeft/

53. Public Policy Committee, International Society of Pharmacoepidemiology. Guidelines for good pharmacoepidemiology practice (GPP): Guidelines for good pharmacoepidemiology practice. Pharmacoepidemiol Drug Saf. 2016 Jan;25(1):2–10.

54. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. The Lancet. 2014 Aug;384(9945):755–65.

55. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. The Lancet. 2008 Feb;371(9612):569–78.

56. International Committee of Medical Journal Editors (ICMJE). Defining the Role of Auhtors and Contributors [Internet]. Available from: http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html

ANNEX 1. MedDRA codes

MedDRA codes

Acinar cell carcinoma of pancreas
Adenocarcinoma pancreas
Carcinoid tumor of the pancreas
Ductal adenocarcinoma of pancreas
Gastrinoma malignant
Glucagonoma
Intraductal papillary-mucinous carcinoma of pancreas
Malignant neoplasm of islets of Langerhans
Metastases to pancreas
Metastatic glucagonoma
Mucinous cystadenocarcinoma of pancreas
Pancreatic carcinoma
Pancreatic carcinoma metastatic
Pancreatic carcinoma recurrent
Pancreatic carcinoma stage 0
Pancreatic carcinoma stage I
Pancreatic carcinoma stage II
Pancreatic carcinoma stage III
Pancreatic carcinoma stage IV
Pancreatic neuroendocrine tumor
Pancreatic neuroendocrine tumor metastatic
Pancreatic sarcoma
Pancreatoblastoma
Serous cystadenocarcinoma of pancreas
Solid pseudopapillary tumor of the pancreas

ANNEX 2. ENCePP Checklist for Study Protocols

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Doc.Ref. EMA/540136/2009

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

ENCePP Checklist for Study Protocols (Revision 3)

Adopted by the ENCePP Steering Group on 01/07/2016

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is “Yes”, the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer ‘N/A’ (Not Applicable) can be checked and the “Comments” field included for each section should be used to explain why. The “Comments” field can also be used to elaborate on a “No” answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS as recommended in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title: Epidemiological assessment of the risk for pancreatic cancer associated with the use of semaglutide in patients with type 2 diabetes - A cohort study based on Nordic registry data

Study reference number: NN9535-4447

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	6
1.1.2 End of data collection[2]	☒	☐	☐	6
1.1.3 Study progress report(s)	☒	☐	☐	6
1.1.4 Interim progress report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS register	☒	☐	☐	6
1.1.6 Final report of study results.	☒	☐	☐	6

Comments:

Re: 1.1.4) No interim reports will be developed, as data will be too limited during the conduct of the study for interim analyses. Study progress reports will however be developed.

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	8
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	7
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	8

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.
[2] Date from which the analytical dataset is completely available.

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1.4 Which hypothesis(-es) is (are) to be tested?	☒	☐	☐	8
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Re: 2.1.5) The aim of this study is to evaluate whether, and if so, to what extent, exposure to semaglutide influences the risk of pancreatic cancer in patients with T2DM. This is, however, not specifically stated as a hypothesis.

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, new or alternative design)	☒	☐	☐	9.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.1, 9.4
3.3 Does the protocol specify measures of occurrence? (e.g. incidence rate, absolute risk)	☒	☐	☐	9.7.2
3.4 Does the protocol specify measure(s) of association? (e.g. relative risk, odds ratio, excess risk, incidence rate ratio, hazard ratio, number needed to harm (NNH) per year)	☒	☐	☐	9.7.2
3.5 Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☐	☒	11

Comments:

Re: 3.5) This study is based on data already available in existing databases (secondary data) and single case collection and reporting from such studies is not required according to the current European Pharmacovigilance Regulations (Module VI (rev.2)).

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1 Is the source population described?	☒	☐	☐	9.2.1
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period?	☒	☐	☐	9.2.1
4.2.2 Age and sex?	☒	☐	☐	9.2.1, 9.2.2
4.2.3 Country of origin?	☒	☐	☐	9.2.1
4.2.4 Disease/indication?	☒	☐	☐	8
4.2.5 Duration of follow-up?	☒	☐	☐	9.2.1
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	9.2.2, 9.2.3

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.3.2
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	9.2.4
5.3 Is exposure classified according to time windows? (e.g. current user, former user, non-use)	☒	☐	☐	9.3.2
5.4 Is exposure classified based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☒	☐	☐	9.2.4, 9.3.2, 9.7.3.1

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.1.1
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.1.1, 9.3.1
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, prospective or retrospective ascertainment, use of validation sub-study)	☒	☐	☐	9.4.1
6.4 Does the protocol describe specific endpoints relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease, disease management)	☐	☒	☐	

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol describe how confounding will be addressed in the study?	☒	☐	☐	9.1, 9.3.3
7.1.1. Does the protocol address confounding by indication if applicable?	☒	☐	☐	9.1, 9.3.3
7.2 Does the protocol address:	☐	☐	☐	

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.2.1. Selection biases (e.g. healthy user bias)	☒	☐	☐	9.1, 9.4, 9.2.4
7.2.2. Information biases (e.g. misclassification of exposure and endpoints, time-related bias)	☒	☐	☐	9.4
7.3 Does the protocol address the validity of the study covariates?	☐	☒	☐	

Comments:

Re: 7.3) The protocol does not directly address the validity of study covariates. These are, however, expected to be valid as they come from registries with high validity.

Section 8: Effect modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☒	☐	☐	

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.4, 9.3.2

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.4, 9.3.1
9.1.3 Covariates?	☒	☐	☐	9.4, 9.3.3
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9.4
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9.4
9.2.3 Covariates? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9.4
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9.4
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD)-10, Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9.4, 9.3.1, annex 1
9.3.3 Covariates?	☒	☐	☐	9.4
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9.4

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Is the choice of statistical techniques described?	☒	☐	☐	9.7.2
10.2 Are descriptive analyses included?	☒	☐	☐	9.7.2
10.3 Are stratified analyses included?	☒	☐	☐	9.7.3.1
10.4 Does the plan describe methods for adjusting for confounding?	☒	☐	☐	9.1, 9.3.3, 9.7.2
10.5 Does the plan describe methods for handling missing data?	☐	☐	☒	
10.6 Is sample size and/or statistical power estimated?	☒	☐	☐	9.5

Comments:

Re: 10.5) There are no missing data in the data sources to be used for this study.

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.8
11.2 Are methods of quality assurance described?	☒	☐	☐	9.8
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	9.9
12.1.2 Information bias?	☐	☒	☐	
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	☒	☐	☐	9.1, 9.9
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure, duration of follow-up in a cohort study, patient recruitment)	☒	☐	☐	9.5

Comments:

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	10
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☒	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	10

Comments:

Re: 13.2) Ethical approval is not required in Denmark for purely registry-based studies. Ethical approval is required in Sweden and Norway and will be applied for at a later stage.

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	5

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	12
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	12

Comments:

Name of the main author of the protocol: [redacted]

Date: [redacted]

Signature: [redacted]

List of figures and tables

Page