Protocol H9X-MC-B013(c)
Dulaglutide and Potential Risks of Pancreatic Cancer and Thyroid Cancer: A Non-Interventional Post-Authorisation Safety Study (PASS)

Confidential Information

Document ID: VV-PVG-103946

Protocol Electronically Signed and Approved by Lilly: 14-Sep-2020 GMT
Amendment (a) Electronically Signed and Approved by Lilly: 22-Apr-2021 GMT
Amendment (b) Electronically Signed and Approved by Lilly: 14-Jul-2022
Amendment (c) Electronically Signed and Approved by Lilly: approval date provided below

Post-Authorisation Safety Study (PASS) Information

Title	Dulaglutide and Potential Risks of Pancreatic Cancer and Thyroid Cancer: A Non-Interventional PASS
Study Identifier	H9X-MC-B013
Version identifier	Protocol amendment (c)
Date of last version	14 July 2022
EU PAS Register No:	EUPAS32646
Active substance	Dulaglutide (ATC code: A10BJ05)
Medicinal product(s):	Trulicity 0.75-mg solution for injection Trulicity 1.5-mg solution for injection Trulicity 3.0-mg solution for injection Trulicity 4.5-mg solution for injection
Product reference:	EU/1/14/956
Procedure number:	EMEA/H/C/002825
Marketing authorisation holder(s)	Eli Lilly and Company
Joint PASS	No
Research question and objectives	This study aims to evaluate the incidence of pancreatic cancer and thyroid cancer in association with dulaglutide treatment compared to other second-line anti-diabetes medications (ADMs) among patients with type 2 diabetes mellitus (T2DM). The primary objective is to • estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer (including subtypes: Papillary, Follicular, and Medullary [C-cell tumour]) for patients with T2DM who initiated dulaglutide compared to those who initiated other non-incretin second-line ADMs. The secondary objectives are to • estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer for patients with T2DM who initiated dulaglutide compared to those who initiated other glucagon-like peptide-1 receptor agonist (GLP-1 RA), and • estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer for patients with T2DM who initiated GLP-1 RAs compared to those who initiated other non-incretin second-line ADMs.
Country(ies) of study	Finland, Sweden, United States
Author	PPD

Marketing Authorisation Holder

Marketing authorisation holder (MAH)	Eli Lilly Nederland B.V. Papendorpseweg 83, 3528 BJ Utrecht The Netherlands
MAH contact person	PPD

1. Table of Contents

Section **Page**

List of Tables

Table **Page**

List of Figures

Figure **Page**

List of Annexes

Annex **Page**

2. List of Abbreviations

Term	Definition
ADM	antidiabetes medication
AE	adverse event
AHLTA	Armed Forces Health Longitudinal Technology Application
AR	adverse reaction
ATC	Anatomical Therapeutic Chemical Classification System
AvoHILMO	Register of Primary Health Care Visits
BMI	body mass index
CI	confidence interval
CPT-4	Current Procedural Terminology, 4th edition
DEERS	Defense Enrollment Eligibility Reporting System
DoD	Department of Defense
DPP-4-I	Dipeptidyl peptidase-4 inhibitors
DRG	Diagnosis-Related Group
EMA	European Medicines Agency
EMR	electronic medical record
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EPS	exposure propensity score
ERB	ethical review board
EU PAS	European Union Post-Authorisation Studies
FDA	Food and Drug Administration
GDPR	General Data Protection Regulation

GLP-1 RA	Glucagon-like peptide-1 receptor agonist
GPP	Good Pharmacoepidemiology Practices
GVP	Good Pharmacovigilance Practices
HbA1c	hemoglobin A1c (glycated hemoglobin)
HILMO	Finnish Care Register for Health Care
HR	hazard ratio
ICD	International Classification of Disease
ICD-10-CM	International Classification of Disease, 10th revision, Clinical Modification
ICD-9-CM	International Classification of Disease, 9th revision, Clinical Modification
ICD-O-3	International Classification of Diseases for Oncology, 3rd Edition
ICSR	Individual case safety report
INN	International Non-proprietary Names
IR	incidence rate
ITT	intention-to-treat
MDR	Military Health System Data Repository
MHS	Military Health System
NDC	National Drug Code
PanIN	Pancreatic intraepithelial neoplasia
PASS	post-authorisation safety study
PIN	population identification number
PPV	positive predictive value
PRAC	Pharmacovigilance Risk Assessment Committee
PRC	Population Register Centre

QMS	Quality Management System
RWD	real-world data
SGLT-2I	sodium-glucose cotransporter-2 inhibitor
SID	study identification number
T2DM	type 2 diabetes mellitus
TZD	thiazolidinedione
VNR	Nordic article number

3. Responsible Parties

4. Abstract

Title

Dulaglutide and Potential Risks of Pancreatic Cancer and Thyroid Cancer: A Non-interventional PASS (H9X-MC-B013)

Rationale and background

Dulaglutide is a long-acting glucagon-like peptide-1 receptor agonist (GLP-1 RA) approved for use in 2014 (first GLP-1 RA [exenatide] was approved in 2005) as a monotherapy or in combination with other antidiabetic therapies by the European Medicines Agency (EMA) and the United States (US) Food and Drug Administration (FDA) for the treatment of type 2 diabetes mellitus (T2DM).

At present there is no evidence to support an increase in pancreatic cancer with long-term GLP-1-RA therapy; however, some epidemiological data have suggested an elevated risk of pancreatic cancer after exposure to GLP-1 RA therapy. More research is needed to better understand the long-term safety profile of GLP-1 RA therapies for the treatment of T2DM. In 2014, the FDA and the EMA announced ongoing efforts to evaluate the potential relationship between incretin-based therapies and pancreatic cancer.

Concerns also exist for a potentially increased risk of thyroid C-cell tumours, a rare condition (accounting for 1% to 2% of all thyroid cancers) and challenging to study since it is difficult to achieve large enough sample sizes to obtain meaningful effect estimates. Studies with large sample sizes, long follow-up, and adequate confounder control are required to better delineate these risks.

The marketing authorisation holder has committed to the EMA to conduct a retrospective non-interventional post-authorisation safety study (PASS) to further evaluate the long-term safety profile of dulaglutide in routine clinical care settings.

Research question and objectives

This study aims to evaluate the incidence of pancreatic cancer and thyroid cancer in association with dulaglutide treatment compared to other second-line anti-diabetes medications (ADMs) among patients with T2DM.

The primary study objective is to

- estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer (including subtypes: Papillary, Follicular, and Medullary [C-cell tumour]) for patients with T2DM who initiated dulaglutide compared to those who initiated other non-incretin second-line ADMs.

The secondary study objectives are to

1. estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer for patients with T2DM who initiated dulaglutide compared to those who initiated other GLP-1 RA, and
2. estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer for patients with T2DM who initiated GLP-1 RAs compared to those who initiated other non-incretin second-line ADMs.

Study design

This study is a retrospective, non-interventional PASS utilising real-world data (RWD) in 3 countries (2 in Europe and 1 in North America) to address the study objectives. Patients with T2DM who initiated a second-line ADM during the observation period will comprise the study cohort. The exposure groups of interest will be patients initiating dulaglutide, patients initiating GLP-1 RA therapies, and patients initiating other non-incretin second-line ADMs. All patients will be followed for first occurrence of the outcome of interest: pancreatic cancer and thyroid cancer.

Population

The study will be conducted using patient clinical data extracted from the respective databases in the 3 countries (Finland, Sweden, and the US) and will cover up to approximately 15 years of post-launch data accrual for dulaglutide (since dulaglutide launch in 2014) and up to 24 years of post-launch data accrual for the GLP-1 RA class (since exenatide launch in 2005).

Data extraction/collection will be staggered by country and will be performed between 31 March 2024 and 31 March 2030.

Variables

The exposures of interest include second-line ADMs, including meglitinides, thiazolidinediones (TZDs), sodium-glucose cotransporter-2 inhibitors (SGLT-2-Is), insulin, GLP-1 RAs, alpha glucosidase inhibitors, amylin mimetics, or combinations of ADMs. Exposure status will be determined from medication dispensings.

The outcomes of interest include the first incident diagnosis of pancreatic cancer or thyroid cancer during follow-up and will be identified via validated claims-based algorithms (United States) or from linked cancer registry data (Sweden and Finland).

In addition to study exposures and outcomes of interest, other covariates will be collected during the study baseline period, including patient demographics (e.g., age, sex, race/ethnicity, socioeconomic status, and geographic region), clinical characteristics, comorbidities, prior treatment with therapies associated with prior cancer, diabetes severity indicators, and diagnostic and therapeutic procedures and tests received that are relevant to the outcome of interest.

Data sources

Data sources include:

- United States: The Military Health System Data Repository (MDR),

- Sweden: National Registries, and
- Finland: National Registries.

Study size

Based on projected patient counts for the study period (76,500; 41,800; and 15,700 new users of dulaglutide with mean follow-up time of 4.8, 5.5, and 4.9 years in the US, Sweden, and Finland, respectively) and under the assumptions of a 1:2 matching ratio of dulaglutide versus comparators; a false-positive rate $\alpha = 0.05$; a 1-sided test; a background pancreatic cancer incidence rate of 22.5, 69.66, and 77.94 events per 100,000 person-years in patients with T2DM for the US, Sweden, and Finland, respectively; and a background thyroid cancer incidence rate of 15.96, 12.84, and 19.68 events per 100,000 person-years in patients with T2DM for the US, Sweden, and Finland, respectively (background Medullary Thyroid Carcinoma [MTC] incidence rate of 0.25 events per 100,000 person-years in patients with T2DM); the study will have 80% power to rule out a hazard ratio (HR) of 1.13 or higher for pancreatic cancer, and a HR of 1.16 or higher for thyroid cancer in the meta-analysis approach based on 3 data sources. For the MTC meta-analysis approach, the study will have 80% power to rule out a HR of 5.7 or higher. The MTC meta-analysis will be restricted to the Finnish and Swedish data due to those data sources having a method to identify MTC. The US data source will not be included in the meta-analysis for MTC due to lack of information in claims data to differentiate tumour subtypes and limited access to EMRs for a complete case verification.

Data analysis

The primary analysis will be an intention-to-treat (ITT) approach, in which pancreatic cancer and thyroid cancer will be assessed any time after the end of an exposure latency period of 3 years.

Sensitivity analysis will be conducted to address the limitation of ITT analysis using a stricter exposure definition and time-varying approaches.

Among patients in each of the exposure groups, the incidence rate of thyroid cancer (overall and by sub-type, if available) and pancreatic cancer will be calculated.

Cox proportional hazards regression with matching on the exposure propensity score (EPS) to control for the potential confounders will be applied to compare exposure groups with respect to the outcomes of interest. The EPS will represent the probability of exposure on the index date from baseline covariates. HRs with corresponding 95% confidence intervals (CIs) will be calculated for each comparison of interest. In addition to calculating and presenting the results for each country separately, the HRs for pancreatic cancer and thyroid cancer (analysed separately) from the 3 countries and the HRs for MTC from Finland and Sweden will be combined in a meta-analysis approach using an inverse variance-weighted, fixed-effect model. Additionally, a combined HR result for pancreatic cancer and a combined HR for thyroid cancer will be presented for the EU databases (Finland and Sweden) to contextualise the findings in the EU compared to the US.

Milestones

Upon EMA approval of the Study Protocol, IQVIA will carry out data management and the development of a statistical analysis plan that will outline analytic activities in the respective data sources, including propensity score matching and descriptive and comparative analyses to support the development of 2 Interim Reports and a Final Report. Interim Report 1, Interim Report 2, and the Final Report will be submitted to the EMA by 31 December 2024, 31 December 2027, and 31 December 2030, respectively. The interim reports will comprise of details regarding data collection metrics (including accrual of sample size, availability, and measurement of study variables of interest), and presenting the progress of the study outcomes in terms of crude unadjusted incidence rates.

5. Amendments and Updates

The key changes to the content are updated in this table. Minor editorial changes made throughout the document are not captured in this table.

Amendment or Update Identifier	Date	Section of Study Protocol	Amendment or Update	Reason
Amendment (c)	October 2022	3. Responsible Parties	The main author of this protocol amendment is Dr. Camelia Graham, Principal, Epidemiology and Drug Safety, Real-World Solutions, IQVIA.	Administrative reasons
		4. Abstract – Data Analysis	Correction to the follow-up period of the 'as treated (time-varying) approach' in the sensitivity analysis.	A 12-month of follow-up is not adequate for non-acute outcomes such as cancer.
		9.2.1.2. Exclusion Criteria	Removal of familial adenomatous polyposis (FAP) from the exclusion criteria.	No unique claim codes to differentiate FAP from sporadic polyps/adenomas available.
		9.3.1.4. Sensitivity Definition of Exposure	A stricter exposure definition (≥4 prescriptions within a latency period) is added.	To further complement the limitation of ITT analysis.
		9.3.1.4. Sensitivity Definition of Exposure	Additional censoring rules were removed.	The language regarding censoring for the as-treated approach does not align with the language that is described under the time-varying exposure methodology.
		9.7.1. Exposure Lags	The shortest lag period of 6 months will be replaced with no lag period (zero month) in additional sensitivity analyses using different latency periods.	For a more exhaustive description of cancer occurring in the study population.
		9.7.2. Outcome Assessment Period	The same outcome assessment period of the ITT (i.e., any time after the end of the exposure	A 12 month of follow-up is not adequate for non-acute outcomes such as cancer.

Amendment or Update Identifier	Date	Section of Study Protocol	Amendment or Update	Reason
			lag period) will be applied to ascertain the outcome for the exposure definitions.	
		Figure 3	Title changes	The outcome assessment period is applicable to both the ITT analysis and the 'as-treated' approach.
		Figure 4	Removal of Figure 4	Figure 3 applies to the 'as-treated' approach.
		9.7.4 Primary Objective: Association of Dulaglutide and Outcomes of Interest	Language was added to clarify the detailed analysis methods to be described in the statistical analysis plan.	To improve clarity of text.

6. Milestones

Milestone	Planned submission date (data lock point)
Start of data extraction	31 March 2024
End of data extraction	31 March 2030
Interim report 1[a]	31 December 2024 (31 March 2024)
Interim report 2[a]	31 December 2027 (31 March 2027)
Registration in the EU PAS register	24 November 2021
Final report of study results	31 December 2030 (31 March 2030)

Abbreviation: EU PAS = European Union Post-Authorisation Studies.

[a] The interim reports will comprise details regarding data collection metrics (including accrual of sample size, availability, and measurement of study variables of interest), and presenting the progress of the study outcomes in terms of crude unadjusted incidence rates.

7. Rationale and Background

Dulaglutide is a long-acting glucagon-like peptide-1 receptor agonist (GLP-1 RA) approved in 2014 by the United States (US) Food and Drug Administration (FDA) and the European Medicines Agency (EMA) for the treatment of type 2 diabetes mellitus (T2DM) as an adjunct to diet and exercise (first GLP-1 RA was approved in 2005 by the FDA and in 2006 by the EMA). The medication is approved for use as a monotherapy when metformin is considered inappropriate due to intolerance or contraindication, or as an add-on therapy to other glucose-lowering therapies, including insulin. It is administered subcutaneously at approved doses of 0.75 mg or 1.5 mg once per week. Two additional doses of 3.0 mg and 4.5 mg once per week are approved by the FDA and the EMA.[1]

The association between T2DM and pancreatic cancer is complex since T2DM has been shown to be a risk factor, a manifestation and a prognostic factor for pancreatic cancer.[2-5] Individuals who have lived with diabetes for 5 or more years are between 1.5 and 2 times more likely to develop pancreatic cancer.[6] At present there is no direct evidence to support an increase in pancreatic cancer with long-term GLP-1 RA therapy; however, some epidemiological data have suggested an elevated risk of pancreatic cancer after exposure to GLP-1 RA therapy.[7] Rodent studies have indicated that pancreatic intraepithelial neoplasia (PanIN) lesions (precursor lesions of pancreatic ductal adenocarcinoma) and pancreatic duct glands express GLP-1 receptors and undergo proliferation in response to incretin mimetics.[8,9] More research is needed to better understand the long-term safety profile of GLP-1 RA therapies for the treatment of T2DM. In 2014, the FDA and the EMA concluded that ongoing efforts to evaluate the potential relationship between incretin-based therapies and adverse pancreatic outcomes are needed.[10]

In rodent studies, exposures to dulaglutide, exenatide, and liraglutide were associated with C-cell hyperplasia and tumours, and calcitonin levels.[11-13] However, these results have not been confirmed in humans or primates.[7,11,14,15] Thyroid C-cell tumour is a rare condition (accounting for 1% to 2% of all thyroid cancers[16]) for which it is difficult to achieve large enough sample sizes to obtain meaningful effect estimates. Studies with large sample sizes, long follow-up, and adequate confounder control are required to better delineate these risks.

The marketing authorisation holder (MAH) has committed to the EMA to conduct a retrospective non-interventional post-authorisation safety study (PASS) to further evaluate the long-term safety profile of dulaglutide in a routine clinical care setting. The current protocol details a study designed to evaluate the potential risks of pancreatic cancer and thyroid cancer among adults with T2DM who initiated dulaglutide or other GLP-1 RA compared to initiators of other second-line non-incretin anti-diabetes medications (ADMs).

8. Research Question and Objectives

This study aims to evaluate the incidence of pancreatic cancer and thyroid cancer in association with dulaglutide treatment compared to other second-line ADMs among patients with type 2 diabetes mellitus (T2DM).

The primary study objective is to

- estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer (including subtypes: Papillary, Follicular, and Medullary [C-cell tumour]) for patients with T2DM who initiated dulaglutide compared to those who initiated other non-incretin second-line ADMs.

The secondary study objectives are to

- estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer for patients with T2DM who initiated dulaglutide compared to those who initiated other GLP-1 RA, and
- estimate the incidence rates and evaluate the potential association of pancreatic cancer and thyroid cancer for patients with T2DM who initiated GLP-1 RAs compared to those who initiated other non-incretin second-line ADMs.

9. Research Methods

9.1. Study Design

This study is a retrospective, non-interventional PASS utilising real-world data (RWD) in 3 countries (Finland, Sweden, and the US) to address the study objectives. Patients with T2DM who initiated a second-line ADM during the observation period will comprise the study cohort. The exposure groups of interest will be patients initiating dulaglutide, patients initiating GLP-1 RA, and patients initiating other non-incretin second-line ADMs. All patients will be followed for first occurrence of each of the following study outcomes: pancreatic cancer and thyroid cancer.

9.2. Setting

This study will be conducted using electronic longitudinal databases, including data collected in outpatient and inpatient settings of the target countries.

9.2.1. Study Population

The study will be conducted using patient clinical data extracted from the respective databases in the 3 countries (Finland, Sweden, and the US) and will cover up to approximately 15 years of post-launch data accrual for dulaglutide and up to 24 years of post-launch data accrual for the GLP-1 RA class.

Data extraction/collection will be staggered by country and will be performed 3 times between 31 March 2024 and 31 March 2030.

Patient selection criteria are described in the following sections and Figure 1.

9.2.1.1. Inclusion Criteria

The following are the inclusion criteria for the study cohort:

- ≥1 dispensing for a second-line ADM (see Section 9.3.1 Exposures) with the date of the first dispensing as the patient's **index date**
- ≥12 months of available medical history prior to the **index date**
- confirmed diagnosis for T2DM
    - ≥1 diagnosis code for T2DM in the US Military Health System Data Repository
    - confirmation of T2DM in the Diabetes Register for Sweden, or
    - reimbursement code for T2DM in Finnish National Registry, and
- adult patients ≥18 years of age at **index date**.

9.2.1.2. Exclusion Criteria

The following are the exclusion criteria for the study cohort:

1. Confirmed diagnosis for type 1 diabetes on the **index date** or any time prior to the **index date**; or
    1.1. ≥1 diagnosis code for type 1 diabetes in the US Military Health System Data Repository; or

1.2. Confirmation of type 1 diabetes in the Diabetes Register for Sweden; or
1.3. Reimbursement code for type 1 diabetes in Finnish National Registry.

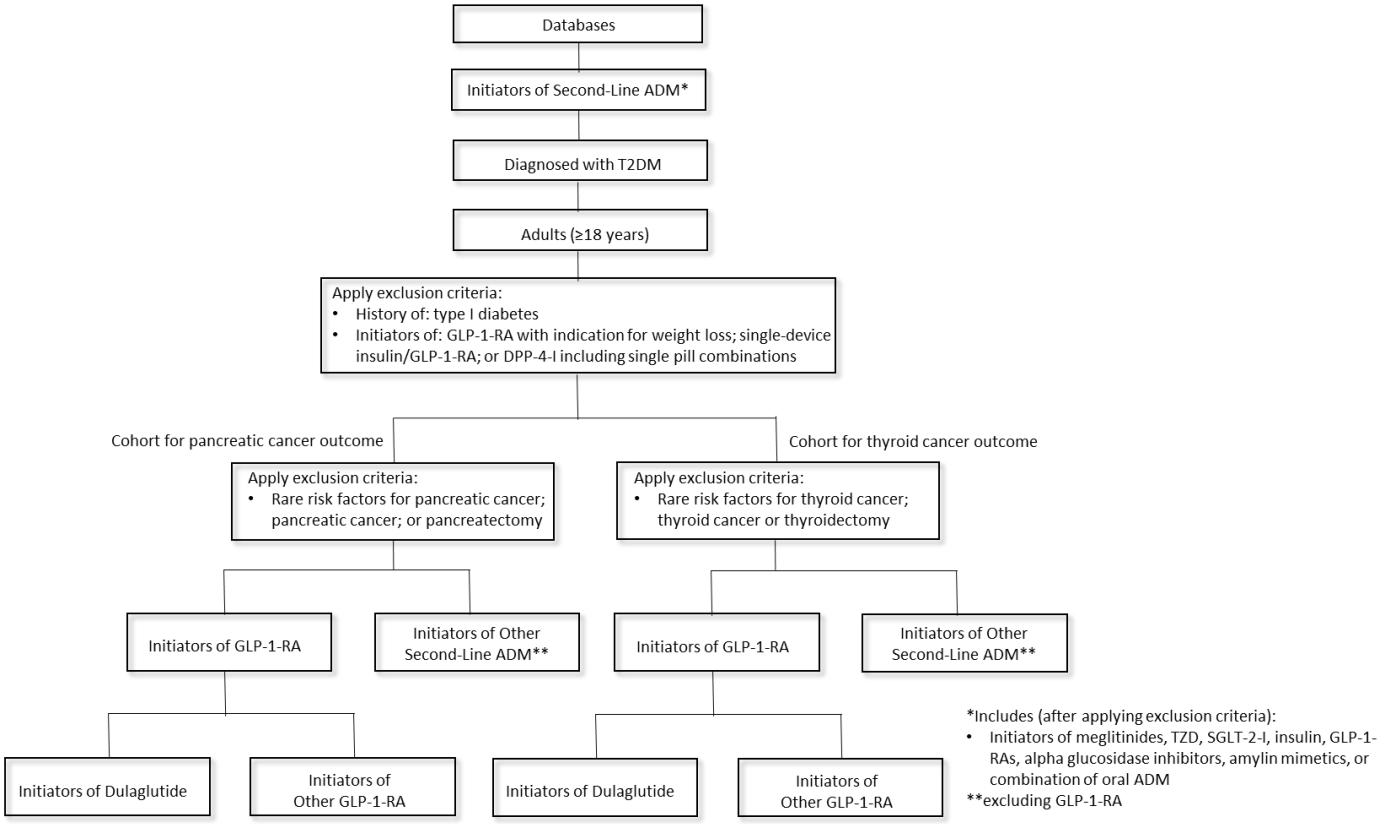
2. ≥1 dispensing for GLP-1 RA with weight-loss indication (e.g., liraglutide [Saxenda®], semaglutide [Wegovy®]) on the **index date** or any time prior to the **index date**; or
3. ≥1 dispensing for single-device insulin/GLP-1 RA injection (e.g., insulin degludec/liraglutide [Xultophy®], insulin glargine/lixisenatide [Soliqua®]) on the **index date** or any time prior to the **index date**; or
4. ≥1 dispensing for dipeptidyl peptidase-4 inhibitors (DPP-4-I), an incretin-based therapy, including single-pill combination formulations, on the **index date** or any time prior to the **index date**. Exclusion of DPP-4-I is considered to minimise misclassification bias and potential effect modification; or
5. ≥1 diagnosis code for human immunodeficiency virus; or
6. Treatment with highly active antiretroviral therapy.

For the pancreatic cancer outcome analyses, the following additional exclusion criteria will be applied:

7. Rare risk factors for pancreatic cancer recorded on the index date or any time prior to the index date; or
   7.1. ≥1 diagnosis code for any type of chronic pancreatitis; or
   7.2. ≥1 diagnosis code indicative of a congenital defect of the pancreas; or
   7.3. ≥1 diagnosis code for cystic fibrosis; or
   7.4. ≥1 diagnosis code for multiple endocrine neoplasia type 1; or
   7.5. ≥1 diagnosis code for Peutz-Jeghers syndrome.
8. ≥1 diagnosis code for pancreatic cancer on the index date or any time prior to the index date; or
9. ≥1 diagnosis or procedure code for pancreatectomy.

For the thyroid cancer outcome analyses, the following additional exclusion criteria will be applied:

10. Rare risk factors for thyroid cancer recorded on the index date or any time prior to the index date; or
    10.1. ≥1 diagnosis code for lupus erythematosus; or
    10.2. Selected hereditary conditions that predispose patients to thyroid cancer risk; or
       10.2.1. Cowden syndrome; or
       10.2.2. Carney complex; or
       10.2.3. Multiple endocrine neoplasia type 1 and type 2.
11. ≥1 diagnosis code for thyroid cancer on the index date or any time prior to the index date; or
12. ≥1 diagnosis or procedure code for thyroidectomy.

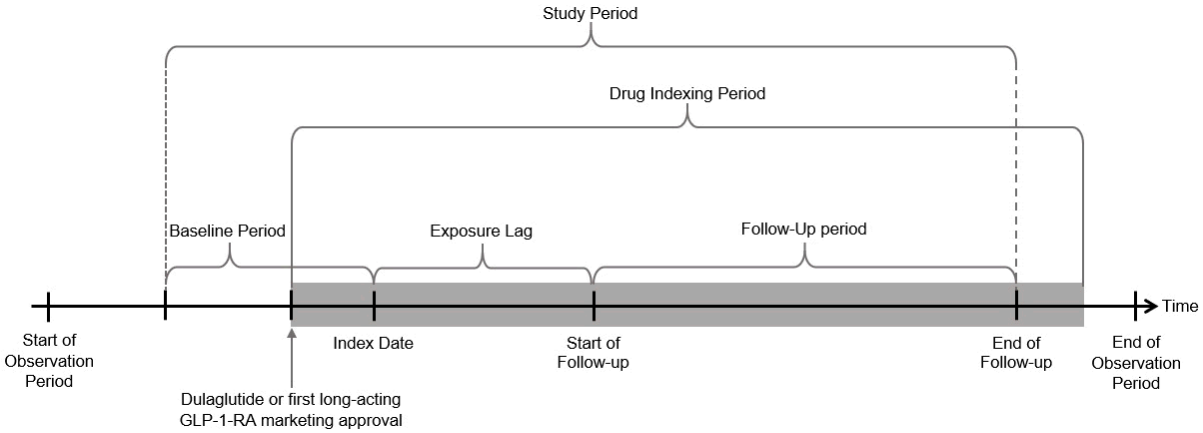


Abbreviations: ADM = anti-diabetes medication; DPP-4-I = dipeptidyl peptidase-4 inhibitors; GLP-1-RA = glucagon-like peptide-1 receptor agonist; SGLT-2-I = sodium-glucose cotransporter-2 inhibitor; T2DM = type 2 diabetes mellitus; TZD = thiazolidinediones.

Figure 1. Study cohort selection diagram.

9.2.2. Primary Objectives Study Period

Each study country will have its own observation period (defined with respect to market availability), consisting of a unique drug indexing period and individual baseline and follow-periods for each patient (Figure 2). The observation period will be defined as the 12 months prior to the date of dulaglutide (or first GLP-1 RA) approval in the study country to the last available date in the database at the time of extraction (dulaglutide was approved in the US and EU in 2014, and the first GLP-1 RA [exenatide] was approved in the US and EU in 2005 and 2006, respectively). The drug indexing period will be used to ascertain ADM initiation and will be defined as the date of dulaglutide (and first GLP-1 RA) approval in the study country until 6 months prior to the last available date in the data at the time of extraction. The 6-month period prior to the end of data is required to prevent selection bias due to informative censoring.[17] Each patient will be assigned an **index date** and this date will form the basis for their unique **study period**, consisting of the 12-month period prior to the index drug date for the given patient (baseline period) until the end of follow-up after exposure lag period (Section 9.2.4 [Follow-up Period and Censoring Criteria] and Section 9.7.1 [Exposure Lags]).



Abbreviation: GLP-1 RA = glucagon-like peptide-1 receptor agonist.

Figure 2. Illustration of study period.

9.2.3. Secondary Objectives Study Period

Thyroid cancer takes between 5 to 10 years from initiation to clinical presentation;[18] however, scientific literature is limited with regard to understanding pancreatic cancer latency (pancreatic cancer takes up to 20 years from initiation of tumorigenesis to patient death[19]). The current body of literature acknowledges that studies with larger sample size and longer duration of follow-up are required to further characterise the potential risk of pancreatic cancer in relation to exposure to the GLP-1 RA class,[20] the proposed study contributes additional longitudinal data and sample size to characterise the potential risks of interest. For the secondary objectives, the study period will be extended to include additional data years for patients who used GLP-1 RAs approved prior to dulaglutide approval (e.g., exenatide, approved in 2005 in the US, and in 2006 in the EU); however, operationalisation of drug indexing and observation periods will remain similar to the primary objectives (Figure 1).

9.2.4. Follow-up Period and Censoring Criteria

Patients will be followed up from the end of the latency period specified in Section 9.7.1 [Exposure Lags] until the end of the study observation period or the occurrence of the following events, whichever occurs first:

- death as a censoring criterion; or
- end of patient data (e.g., due to emigration out of Sweden or Finland or end of enrolment in the US database); or
- first occurrence of the outcome of interest (Section 9.3.2 [Outcomes]).
    - pancreatic cancer for the pancreatic cancer outcome analyses or
    - thyroid cancer for the thyroid cancer outcome analyses.

Further detail is described in Section 9.3.1.3 (Main Definition of Exposure) and Section 9.3.1.4 (Sensitivity Definition of Exposure).

9.3. Variables

The following sections describe the study exposures, outcomes, and covariates (including patient demographics [e.g., age, sex, race/ethnicity, socioeconomic status, and geographic region], clinical characteristics, comorbidities, and concomitant medications).

9.3.1. Exposures

Initiators of second-line ADMs of interest will include initiators of meglitinides, thiazolidinediones (TZD), sodium-glucose cotransporter-2 inhibitors (SGLT-2-I), insulin, GLP-1 Ras, alpha glucosidase inhibitors, amylin mimetics, or combination of ADM; or patients who switch to or add-on second-line ADMs after failure with metformin or sulphonylurea monotherapy (Figure 1). The GLP-1 RA class will include dulaglutide (Trulicity®), exenatide (Byetta® and Bydureon®), lixisenatide (Adlyxin®), liraglutide (Victoza®), albiglutide (Tanzeum®), and semaglutide (Ozempic®, and oral Rybelsus®). Exposure status will be determined from the first medication dispensing of a qualifying second-line ADM among patients who are second-line ADM treatment naïve. To minimise misclassification bias and potential effect modification, initiators of DPP-4-I as incretin-based ADMs are excluded from exposure groups of interest in the primary analyses and will be added to the other second-line ADMs as a comparison group in the sensitivity analyses (Section 9.3.1.2).

9.3.1.1. Main Exposure Groups

To address the primary and secondary objectives, 3 mutually exclusive exposure groups will be defined (dulaglutide initiators, all GLP-1 RA initiators, and other non-incretin second-line ADM initiators) and categorised into 3 independent comparisons.

- Comparison 1: dulaglutide initiators compared to non-incretin second-line ADM initiators
- Comparison 2: dulaglutide initiators compared to other GLP-1 RA initiators, and
- Comparison 3: all GLP-1 RA initiators compared to non-incretin second-line ADM initiators.

9.3.1.2. Sensitivity Exposure Groups

To understand the magnitude of potential effect modification by DPP-4-I, initiators of DPP-4-I will be included in the initiators of "other second-line ADMs" exposure group as a secondary analysis for both the primary and secondary objectives. Furthermore, to minimise confounding by indication, subgroup analyses will be performed by comparing dulaglutide initiators (primary objective) and all GLP-1 RA initiators (secondary objective) to a subgroup of the "other second-line ADM initiators" comparison group, such as insulin initiators and SGLT-2-I initiators.

For all exposure groups, exposure status will be determined using both a time-fixed (intention-to-treat [ITT]) and time-varying (as-treated) approaches, as detailed in Section 9.7 (Data Analysis). Both approaches will be applied to primary and secondary exposure groups defined above.

9.3.1.3. Main Definition of Exposure

Patients' *initial* exposure status will be assigned to one of the exposure groups listed above based on the patient's index dispensing. As described in Section 9.2.4 (Follow-up Period and Censoring Criteria), this initial exposure status will be carried forward and patients will be followed up from the end of the latency period (Section 9.7.1 [Exposure Lags]) until the occurrence of the study outcomes of interest, death as censoring criterion, end of patient data, or end of follow-up (Section 9.7.2 [Outcome Assessment Period]), whichever occurred first, regardless of whether they discontinue, switch, restart, or initiate another ADM (analogous to an ITT approach).

9.3.1.4. Sensitivity Definition of Exposure

Sensitivity analyses will be conducted to address the limitations of ITT analysis.

First, using a stricter exposure defined as follows:

- Study exposure groups will be redefined using a stricter exposure definition. Only patients who receive at least 4 dispensing of a medication within the same exposure group (i.e., dulaglutide initiators, all GLP-1 RA initiators, and other non-incretin second-line ADM initiators) during the latency period (Section 9.7.1 [Exposure Lags]) will be included in the sensitivity analysis. As with the main definition of exposure, study drug initiators will be followed up from the end of the latency period until the occurrence of any of the above censoring criteria (Section 9.3.1.3 [Main Definition of Exposure]).

Second, in the time-varying (as-treated) approach, the following definitions of exposure will be used[21-23]:

- **Time-varying exposure.** Patients will contribute person-time of exposure to another exposure group if they change exposure status during follow-up.
- **Cumulative duration of exposure.** Exposure will be defined as the sum of the duration of dulaglutide, all GLP-1 RAs, or non-incretin second-line ADM use between the index date and end of follow-up as defined in Section 9.2.4 (Follow-Up Period and Censoring Criteria). Duration of use will be categorised and the effect of longer durations of exposure on incidence of the outcomes of interest will be compared.

In the time-varying (as-treated) analysis, termination of exposure will be defined with a grace period of 30 days between refills and a risk window of 30 days after the last refill. Duration of exposure will be calculated from the provided days supplied data fields in respective study databases; and fields with 0 are assumed to be 1 day. Patients will contribute person-time of exposure to the exposure group of interest once they are qualified as initiators in the respective group. Additional details on the exposure definitions will be provided in the study's statistical analysis plan.

9.3.2. Outcomes

Table 1 summarises the primary outcomes and their definitions.

Table 1. Definition of Study Outcomes by Country

Outcome	Definition of Outcomes by Country		
	United States	Sweden	Finland
Pancreatic cancer	Claims-based algorithm based on the study by Wu *et al.* [24] requiring at least 1 inpatient ICD-9/ICD-10 diagnosis code for pancreatic cancer (PPV = 79%). Validation will be performed for a subset of patients for whom electronic medical records are accessible, and PPV will be calculated.	Linked cancer registry	Linked cancer registry
Thyroid cancer	Claims-based algorithm based on the study by Funch *et al.*[25] requiring thyroid surgery during follow-up and ≥2 ICD-9/ICD-10 diagnoses codes for thyroid cancer within 90 days of surgery (PPV = 91%), Validation will be performed for a subset of patients for whom electronic medical records are accessible, and PPV will be calculated.	Linked cancer registry	Linked cancer registry

Abbreviations: ICD-9 = International Classification of Diseases, Ninth Revision; ICD-10 = International Classification of Diseases, Tenth Revision; PPV = Positive Predictive Value.

NOTE: Thyroid cancer is not currently divided in the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3) dictionary by subtype. Papillary, follicular, and medullary subtypes will be identifiable in Finnish and Swedish cancer registries based on morphological codes.[26-30] To the extent that it is available and recorded, thyroid cancer morphologic subtypes may be collected from the EMR (AHLTA) chart review of cancer cases identified by the algorithm in the US data source.

9.3.3. Covariates

Table 2 summarises baseline covariate data to be collected, on or at any time prior to the index date, and the source from which they will be collected. (Note: The length of time patients have for assessment prior to their index date will vary by patient but will be, at a minimum, 12 months per study inclusion criterion.)

Table 2. Definition of Study Covariates by Country

Variable Type	Variable*	Levels	Country		
			United States	Sweden	Finland
Cohort entry year	Cohort entry year	YYYY date	✔	✔	✔
Patient demographics	Age	Numeric	✔	✔	✔
	Sex	Male Female	✔	✔	✔
Clinical characteristic	Smoking status	Current Former	✔ For a subset	✔ Since 2015	✔ Since 2011

Variable Type	Variable*	Levels	Country		
			United States	**Sweden**	**Finland**
		Non-smoker			
	BMI	<18.5 ≥18.5 - <25.0 ≥25.0 - <30.0 ≥30.0	✔ For a subset	✔ Since 2015	✔ Since 2011
	Alcohol use	Yes/No	✔ For a subset	NA	NA
	Alcohol use disorder	Yes/No	✔	✔	✔
Comorbidity	Prior cancer (excluding nonmelanoma skin cancer) Haematological solid organ NOTE: An additional sensitivity analysis will be conducted that will exclude all patients with a history of *any* type of cancer (excluding non-melanoma skin cancer) instead of only excluding patients with a history of the study outcome of interest (pancreatic cancer or thyroid cancer).	Yes/No	✔	✔	✔
	Cerebrovascular disease	Yes/No	✔	✔	✔
	Coronary artery disease, congestive heart failure, ventricular tachycardia/fibrillation	Yes/No	✔	✔	✔
	Allergic rhinitis/hay fever	Yes/No	✔	✔	✔
	Asthma	Yes/No	✔	✔	✔
	Chronic obstructive pulmonary disease/bronchitis	Yes/No	✔	✔	✔
	Gastrointestinal disease	Yes/No	✔	✔	✔
	Cirrhosis	Yes/No	✔	✔	✔
	Bile duct and gallbladder disease	Yes/No	✔	✔	✔
	Haemochromatosis	Yes/No	✔	✔	✔
	Hypercalcaemia	Yes/No	✔	✔	✔
	Hypertriglyceridaemia or hyperlipidaemia	Yes/No	✔	✔	✔
	Hyperparathyroidism	Yes/No	✔	✔	✔
	Hypertension	Yes/No	✔	✔	✔
	Infectious disease	Yes/No	✔	✔	✔
	Hepatitis B virus infection	Yes/No	✔	✔	✔
	Hepatitis C virus infection	Yes/No	✔	✔	✔

Variable Type	Variable*	Levels	Country		
			United States	Sweden	Finland
	Rheumatoid arthritis	Yes/No	✔	✔	✔
	Serum triglycerides	≤1000 mg/dL >1000 mg/dL	NA	✔	✔
Prior treatment associated with prior cancer	Radiation therapy	Yes/No	✔	✔	✔
	Chemotherapy (e.g., Asparaginase, Cytarabine, Mercaptopurine, Mesalamine)	Yes/No	✔	✔	✔
	Immunotherapy (e.g., Nivolumab, Pembrolizumab)	Yes/No	✔	✔	✔
Diabetes severity indicators	Disease duration	Years since first diagnosis code for T2DM (will be reported using mean, standard deviation and cut points: <5, 5 – 10, >10)	✔	✔	✔
	Diabetic retinopathy	Yes/No	✔	✔	✔
	Diabetic nephropathy	Yes/No	✔	✔	✔
	Diabetic neuropathy	Yes/No	✔	✔	✔
	Peripheral arteriopathy	Yes/No	✔	✔	✔
	Prior renal disease	Yes/No	✔	✔	✔
	HbA1c level	Will be reported using mean and standard deviation	NA	✔	✔
Diagnostic and therapeutic procedures and tests	Procedures and tests relevant to outcomes (thyroid, head and neck, pancreas, liver, gallbladder, bile duct, and abdomen)	Yes/No	✔	✔	✔
Concomitant medications	Medications that are known or suspected risk factors for pancreatic cancer (e.g., aspirin, NSAIDs)	Yes/No	✔	✔	✔
* Some variables will be defined only for the pancreatic cancer outcome analyses or only for the thyroid cancer outcome analyses.					

Abbreviations: BMI = body mass index; NA = not available; HbA1c = haemoglobin A1c (glycated haemoglobin); NSAID = nonsteroidal anti-inflammatory drug; T2DM = type 2 diabetes mellitus.

9.4. Data Sources

9.4.1. United States: The Military Health System Data Repository (MDR)

The Military Health System (MHS) is a comprehensive medical network within the US Department of Defense (DoD) that provides health care to all US military personnel, their dependents, and retirees. MHS operates the largest cradle-to-grave health care system in the US, with over 10 million patients actively receiving care on an annual basis. Patients enrolled in the MHS receive benefits through the TRICARE nationwide managed care program, which combines health care from DoD facilities with those from the private sector. Individuals must be registered in the Defense Enrollment Eligibility Reporting System (DEERS) to receive TRICARE benefits. Patients are not required to use military medical facilities, and many use their MHS coverage to obtain care in civilian facilities. Thirteen percent of the MHS population are active duty military, meaning that most enrolees are non-military. Males represent 51% of the population and the age distribution of patients in the MHS is similar to the United States,[31,32] Sweden, and Finland populations. The current study will use data from the MHS Data Repository (MDR). Electronic medical record (EMR) chart reviews from the Armed Forces Health Longitudinal Technology Application (AHLTA) database will be conducted for a subset of cancer cases identified in the MDR using validated algorithms. Patients have unique identifiers allowing for their records to be linked across the sources within the MHS. The MHS database contains a death code and the date of death. The death data in the MHS database is maintained in DEERS and comes from the following sources: a contracted, recurring Social Security Death Index feed from the Social Security Administration, reported deaths in both military and civilian facilities, combat related deaths, and survivor self-report directly to TRICARE. All death data are processed and direct-linked to the beneficiary within a master death file.

MHS Data Repository (MDR): Comprehensive medical care is recorded in the MDR, which is comprised of the claims data from TRICARE and supplementary data from the DoD's longitudinal EMR system AHLTA database. The continually updated MDR has been capturing and integrating all health care events for the entire DoD network since 2000. AHLTA is linked to the Composite Health Care System, which allows access to laboratory, pathology, and radiology orders and results for a subset of patients receiving care in military treatment facilities (15% to 20% of encounters). Currently, the MDR contains data on more than 10 million active TRICARE beneficiaries receiving care at more than 65 hospitals and over 500 military clinics, as well as at private hospitals and clinics throughout the country. Examples of data recorded in such encounters include demographic data (e.g., age, sex, race), provider information (e.g., provider ID, specialty, and facility), diagnostic codes (International Classification of Disease, 10th revision, Clinical Modification [ICD-10-CM] and International Classification of Disease, 9th revision, Clinical Modification [ICD-9-CM] codes for outpatient encounters; Diagnosis-Related Group [DRG] classifications for inpatient encounters), Current Procedural Terminology, 4th edition (CPT-4) codes for procedures, and, among visits within military treatment facilities (15%

to 20% of the encounters), additional vitals and lifestyle information (e.g., body mass index [BMI], smoking behaviours, alcohol use, and chemistry laboratory results). Characteristics of TRICARE and AHLTA databases, including key data elements, and temporal and population coverage, are summarised in Annex 3.

All prescribing and dispensing details, including mail order and inpatient medications, are electronically coded in the MDR. Details include the prescribed drug name and national drug code (NDC), the treating facility, department(s) rendering care, the treating provider(s), start and end dates of the prescription order, amount dispensed, date dispensed, route of medication, units, number of refills, and remaining refills. The DoD manages a drug formulary that has been well established for decades.

In general, DoD providers follow the recommendations of the DoD formulary. However, the approval process to prescribe a non-formulary drug is historically expedient, allowing providers to prescribe what they consider to be the optimal treatment regimen. In general, a DoD provider can prescribe most, if not all, available medications in a therapeutic class to a DoD beneficiary. In the civilian network, in which patients utilise TRICARE insurance in the pharmacy setting (i.e., retail and mail order), there are even fewer restrictions on prescription drugs as prescribed by the provider. Notably, it is not uncommon for drugs recently approved by the FDA to be prescribed in the DoD pharmacy system much faster than in other integrated delivery networks and commercial insurance plans.

The MDR is updated continuously and has a data lag of approximately 60-90 days.

Abstraction of Patient Charts: A data collection instrument will be developed for the MDR, and case adjudication will take place for 155 patient charts (Approximately 20% of estimated new cases of pancreatic and thyroid cancer across 5 years). The administrative health database algorithms summarized in Table 1[24,25] will be used to identify incident cases of pancreatic cancer and incident cases of thyroid cancer in the MDR database. These administrative claims-based algorithms were developed using health insurance claims-based databases similar to the MDR database and were shown to perform well to identify true incident pancreatic and thyroid cancer cases. The algorithm for identifying thyroid cancer was validated through chart review and was shown to have a high PPV of 91%. The algorithm for identifying pancreatic cancer was validated using a cancer registry and demonstrated a PPV of 79%. The cases identified using the previously validated algorithms will be confirmed or refuted through medical record review using pre-determined case confirmation rules, and PPV will be calculated as a measure of accuracy of the algorithms in this study. Stage and histology information will be collected as available in the obtained charts.

9.4.2. Sweden: National Registries

The study will use the following 6 large Swedish registers linked via patients' personal identification numbers: the Swedish National Patient Register, the Swedish Prescribed Drug Register, National Cancer Registry, the National Diabetes Register, the National Cause of Death

Register, and the Total Population Register. Patients have unique identifiers allowing for their records to be linked across the national registries.

Swedish National Patient Register: The Swedish National Patient Register includes information on inpatient and outpatient non-primary (i.e., specialist) encounters across Sweden. Coverage for inpatient admissions became nationwide in 1987 and is near 100% today (the current population of Sweden is ~10 million).[33] Coverage for outpatient non-primary care began nationwide in 2001 and is around 87% today.[34] Key variables include diagnoses, surgery, external causes of injury (E-codes), age, sex, hometown, hospital, specialty, and information related to hospital admissions and discharges (e.g., dates, main and contributory diagnoses, mode of discharge). A number of diagnoses have been reported to have high positive predictive values in the inpatient register.[35] The Swedish National Patient Register is updated monthly.

Swedish Prescribed Drug Register: The Prescribed Drug Register, maintained by the National Board of Health and Welfare since 2005, contains all dispensed medication for patients in Sweden. Coverage of dispensing is close to 100%, however, inpatient medication administrations are not captured and there is only partial capture of drugs administered in hospital-based outpatient visits. Information included in the Prescribed Drug Register include basic demographic characteristics such as age, sex, and residency, as well as medication-specific information such as the prescribed/dispensed drug (e.g., Anatomical Therapeutic Chemical Classification System [ATC] code, International Non-proprietary Names [INN]), dispensing date, pack size, dispensed amount, formulation, dosage, prescribing health care practitioner, and costs. The Prescribed Drug Register is updated every 2 weeks. The data for the previous calendar year are usually available in January.

National Cancer Registry: The National Cancer Registry, founded in 1958, is run by the Swedish National Board of Health and Welfare, and covers ~100% of the population of Sweden. In Sweden, it is compulsory for every health care provider to report newly detected cancer cases to the registry. A report is required to be sent for every cancer case diagnosed via clinical-, morphological-, other laboratory examinations as well as cases diagnosed at autopsy. Key variables include demographic data (e.g., age, sex, place of residence); medical data (e.g., site of tumour [ICD-O-3], histological type [ICD-O-3], stage [TNM classification; since 2004; except for brain, cranial tumours, lymphoma, and leukaemia], date of diagnosis, reporting hospital and department, and reporting pathology/cytology department); and follow-up data (e.g., date of death, cause of death, date of migration). Some variables in the Swedish Cancer Registry are updated monthly (e.g., diagnosis, address, date of death) and other variables are updated annually (e.g., cause of death, emigration).

National Diabetes Register: The National Diabetes Register was launched in 1996 and has been an integral part of Swedish diabetes care for the past 18 years. The registry contains information about medicines, diagnoses, laboratory results, primary care, closed care, specialised outpatient care, patient-reported outcome measures, and other self-reported health outcomes. BMI and smoking status are available in the register since 2015 however alcohol consumption is not recorded. The database has engaged the participation of both hospitals and primary care clinics with 98% and 92% coverage, respectively. The register offers a unique opportunity to monitor

the quality of care in terms of risk factors and potential complications of diabetes, as well as the evolution of treatment. The National Diabetes Register is updated monthly with information from the prescription and patient registers and is updated annually for cause of death.

National Cause of Death Register: The National Cause of Death Register has captured death information for residents of Sweden since 1953. Key variables include hometown, sex, date of death, cause of death (ICD-9/10 codes have been used since 1988), and intent (in cases of injury or poisoning). The National Cause of Death Register is updated every year in December for the cause of death variable, however dates of death are available in the register earlier.

Total Population Register: This register is managed by Statistics Sweden and contains information on all persons registered in Sweden. The register was established in 1968 and contains data on life events including birth, death, name change, marital status, family relationships and migrations, residence, citizenship, and country of birth. The information in the register is updated daily. The data in the register can be accessed for research through a standardised application procedure to Statistics Sweden.

9.4.3. Finland: National Registries

Finland has a well-developed, population-wide register system with longitudinal follow-up data. Patients are identified in the registers with a unique patient ID; thus, enabling patient records to be linked across registers. The study will include the following 7 primary registers: the Finnish Care Register for Health Care (HILMO), the Register of Primary Health Care Visits (AvoHILMO), the Finnish Prescription Registers, the Finnish Cancer Registry, the Finnish Causes of Death Register, the Population Register and regional laboratory data. Patients have unique identifiers allowing for their records to be linked across the sources within the national registries.

In Finland there are 2 patient registers: HILMO and AvoHILMO, which are managed by the National Institute for Health and Welfare. HILMO contains data since 1994. The database contains information on secondary care (in- and outpatient care) such as duration of hospitalisation, diagnoses (ICD-10 codes), and medical procedures. AvoHILMO contains information about public primary care visits. For example, the time and place of treatment, as well as diagnoses (ICD-10 codes) and procedures, are recorded. BMI and smoking behaviours have been captured in AvoHILMO since 2011 (alcohol use is not captured, and alcohol abuse is only captured if diagnosed). Coverage for both patient registers is nationwide and near 100% (the current population of Finland is ~5.5 million). Medical treatment is only recorded via procedure codes. Quality of the data is considered high, but there is variation, for example, in the reporting rate and accuracy of secondary diagnoses. Data are updated annually and data for the previous year are available in September.

Finnish Prescription Registers: There are 2 prescription registers available in Finland: the traditional Prescription register and the new e-Prescription register. The traditional Prescription register is managed by the Social Insurance Institute and contains data for all dispensed drugs that were both prescribed and reimbursed since 1994. The e-Prescription register is held by the Social Insurance Institute but managed by the National Institute for Health and Welfare. The e-

Prescription register captures data on all dispensed drugs including drugs that were not reimbursed. Use of e-prescriptions has been compulsory in public health care since April 2013, and in private health care since January 2015. Both databases capture dispensing from retail pharmacies across Finland. Coverage of pharmacy dispensing are close to 100%, however, inpatient medication administrations are not captured and there is only partial capture of medications administered in hospital-based outpatient visits. Key variables include basic demographic characteristics such as age, sex, and residency, as well as medication-specific information such as the prescribed/dispensed drug (e.g., ATC code, Nordic article number [VNR code]), prescription dispensing date, pack size, dispensed amount, formulation, dosage, prescribing health care practitioner, and costs. Complete data from the traditional Prescription register is available for the previous year in March of the next year. The e-prescription register is updated monthly, and data is available without lag time.

Finnish Cancer Registry: The Finnish Cancer Registry was founded in 1952, is maintained by the National Institute for Health and Welfare and consists of a national registry of all cancer cases since 1953 in Finland. Since 1961, notification of cancer cases to the registry has been compulsory (i.e., coverage is ~100%). The registry provides information on cancer burden and its determinants, patients' survival, and cancer predictions. Key variables include demographic information (e.g., age, sex), medical data (e.g., date of diagnosis, topography [ICD-O-3], morphology [ICD-O-3], laterality, basis [method] of diagnosis, stage, treatment type), and follow-up information (e.g., cause of death, date of death or migration). It also provides information on aetiology, the recommended changes in cancer screening, and the late effects of cancer care and treatment. Some variables in the Finnish Cancer Registry are updated monthly (e.g., diagnosis, geographic area, date of death) and other variables are updated annually (e.g., cause of death, emigration) with a lag time of approximately 16 months.

Finnish Causes of Death Register: The Finnish Causes of Death Register has captured death information for residents of Finland since 1969. Key variables include hometown, sex, date of death, cause of death (ICD-10 codes have been used since 1996), and intent (in cases of injury or poisoning). The Finnish Causes of Death Register is updated every year in December for the cause of death variable, however dates of death are available in the register earlier.

Population Register Centre: The Population Register Centre (PRC) collects demographic data on the population, including information on migration and civil status. The records are available from 1969 (from 1971 in electronic format). The PRC covers all registered individuals in Finland and their unique patient IDs with 100% coverage.

Laboratory Data: Primary registries will be linked to regional laboratory databases containing data on laboratory measurements taken in all domains of the public health care in Finland.

Requirements for accessing the laboratory data vary between the numerous regional databases. In the current study, the regional laboratory databases of 7 regions in Finland, representing 68% of the Finnish population, will be contacted separately to obtain access to their data. Thus, regional laboratory data will be available for a subset of the study cohort.

9.5. Study Size

Based on projected patient counts for the study period (Sections 9.5.1 to 9.5.3) and under the assumptions of uniform patient accrual over 12 years of effective follow-up (from 2015 to 2029 minus the 3-year latency period), exponential distributions for the probability of study outcomes; a 1:2 matching ratio of dulaglutide versus comparators; a false-positive rate $\alpha = 0.05$; a 1-sided test; a background pancreatic cancer incidence rate of 22.5, 69.66, and 77.94 events per 100,000 person-years for the US, Sweden, and Finland respectively; and a background thyroid cancer incidence rate of 15.96, 12.84 and 19.68 events per 100,000 person-years in the US, Sweden, and Finland, respectively (background Medullary Thyroid Carcinoma [MTC] incidence rate of 0.25 events per 100,000 person-years), the study will have 80% power to rule out a hazard ratio (HR) of 1.13 or higher for pancreatic cancer, and a HR of 1.16 or higher for thyroid cancer for the meta-analysis based on 3 data sources. A meta-analysis for MTC, based on the 3 data sources, will not be conducted due to the US data source not having a method to comprehensively identify MTC in the MDR. Additionally, the study will have 80% power to rule out a HR of 1.18 or higher for pancreatic cancer, and a HR of 1.40 or higher for thyroid cancer, and an HR of 5.7 or higher for MTC in the meta-analysis approach for the EU databases (Sweden and Finland). In each country, the study will have 80% power to rule out a HR of 1.15, 1.21, and 1.33 or higher for pancreatic cancer; and a HR of 1.17, 1.52, and 1.70 or higher for thyroid cancer for the US, Sweden, and Finland, respectively; and a HR of 6.9, and 13.8 or higher for MTC for Sweden and Finland, respectively.

Summary of the methodology used for projecting sample size and estimating mean follow-up is presented in Annex 1.

9.5.1. MDR US

Preliminary patient counts from 2015 through 2020 show that the exposed cohort will include approximately 76,500 new users of dulaglutide and approximately 151,300 new users of any GLP-1 RA. Additionally, there are approximately 806,500 new users of other non-incretin second-line ADMs. The projected number of incident users of dulaglutide expected by the end of patient accrual in 2026 (i.e., the end of the 15-year study in 2029 minus the 3-year latency period) is estimated to be approximately 269,000 with a minimum follow-up of 3 years and a mean follow-up time of 4.8 years.

Pancreatic cancer: The observed rate of new cases of pancreatic cancer in the general US population is 12.5 per 100,000 person-years.[36] Individuals with T2DM have an 80% increased risk of pancreatic cancer versus individuals without diabetes.[35] Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background pancreatic cancer incidence rate of 22.5 per 100,000 person-years (12.5 multiplied by 1.8),[36,37] the study will have 80% power to rule out an HR of 1.15 or higher for pancreatic cancer.

Thyroid cancer: The observed rate of new cases of thyroid cancer in the general US population is 13.3 per 100,000 person-years.[38] Individuals with T2DM have a 20% increased risk of thyroid cancer versus individuals without diabetes.[39] Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background thyroid cancer incidence rate of 15.96 per 100,000 person-years (13.3

multiplied by 1.2),[38,39] the study will have 80% power to rule out an HR of 1.17 or higher for thyroid cancer.

MTC: The study power and detectable effect size were not estimated for MTC in the US database due to not having a method to comprehensively identify MTC in the MDR.

Note: Unlike in Sweden and Finland, age-specific cancer incidence rates were not available in the US.

9.5.2. Swedish Registers, Sweden

- In Sweden, the projected number of patients expected by 2030 are to include approximately 41,800 new users of dulaglutide with a mean follow-up period of 5.5 years, about 352,900 new users of any GLP-1 RA, and approximately 1.2 million new users of other non-incretin second-line ADMs.

Pancreatic cancer: The observed rate of new cases of pancreatic cancer in the Swedish population ≥40 years old is 38.7 per 100,000 person-years (the majority of patients with T2DM in Sweden is ≥40 years of age). Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background pancreatic cancer incidence rate of 69.66 per 100,000 person-years (38.7 multiplied by 1.8),[37,40] the study will have 80% power to rule out an HR of 1.21 or higher for pancreatic cancer.

Thyroid cancer: The observed rate of new cases of thyroid cancer in the Swedish population ≥40 years of age is 10.7 per 100,000 person-years. Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background thyroid cancer incidence rate of 12.84 per 100,000 person-years (10.7 multiplied by 1.2),[39,40] the study will have 80% power to rule out an HR of 1.52 or higher for thyroid cancer.

MTC: Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background MTC incidence rate of 0.25 per 100,000 person-years,[39,41] the study will have 80% power to rule out an HR of 6.9 or higher for MTC. The incidence rate of MTC in the US was used in the sample size calculation in the absence of published incidence rates of MTC in Sweden.

9.5.3. Finnish Registers, Finland

- In Finland, the projected number of patients expected by 2030 are estimated to include approximately 15,700 new users of dulaglutide with a mean follow-up period of 4.9 years, about 123,900 new users of any GLP-1 RA, and approximately 1 million new users of other non-incretin second-line ADMs.

Pancreatic cancer: The observed rate of new cases of pancreatic cancer in the Finnish population ≥40 years of age is 43.3 per 100,000 person-years (the majority of patients with T2DM in Finland is ≥40 years of age). Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background pancreatic cancer incidence rate of 77.94 per 100,000 person-years (43.3 multiplied by 1.8),[37,40] the study will have 80% power to rule out an HR of 1.33 or higher for pancreatic cancer.

Thyroid cancer: The observed rate of new cases of thyroid cancer in the Finnish population ≥40 years of age is 16.4 per 100,000 person-years.[39] Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background thyroid cancer incidence rate of 19.68 per 100,000 person-years (16.4 multiplied by 1.2),[39,40] the study will have 80% power to rule out an HR of 1.70 or higher for thyroid cancer.

MTC: Assuming a 1:2 matching ratio, $\alpha = 0.05$, 1-sided test, and background MTC incidence rate of 0.25 per 100,000 person-years,[39,41] the study will have 80% power to rule out an HR of 13.8 or higher for MTC. The incidence rate of MTC in the US was used in the sample size calculation in the absence of published incidence rates of MTC in Finland.

9.6. Data Management

IQVIA Archived records are stored in either onsite or secure offsite facilities (e.g., Iron Mountain, Recall) and are reviewed prior to secure destruction at the end of the specified retention period. Records not archived are destroyed in a secure, confidential manner according to local site security procedures.

IQVIA record retention policy states that records are retained only for as long as required by legal requirements, contractual commitments or IQVIA Global Record Retention Schedules. IQVIA follows a standard data destruction process once the data are released by the client to be destroyed. The destruction meets or exceeds DoD data destruction processes. The process includes but is not limited to degaussing, shredding and other physical destruction depending on the media.

IQVIA has policies and procedures in place to protect the confidentiality of individually identifiable information in accordance with applicable laws and regulations, regardless of the nature, source, or form of the information. IQVIA completed Certification to US-EU Safe Harbor Effective January 5, 2005, with annual recertification. IQVIA's global Council on Data Protection is chartered to monitor global data protection laws and regulations and compliance with these laws and regulations. The Council has implemented a global “Privacy Awareness Basic Training” course for all IQVIA employees on the company’s privacy policy and procedures. IQVIA vendor contracts have provisions to safeguard any confidential and proprietary information disclosed and vendors must sign a “Vendor Privacy Certification Standard” which sets out IQVIA privacy and security requirements. IQVIA vendors are subject to assessment of their privacy and security practices as part of the selection and qualification process and are subject to performance audits.

In addition, IQVIA™ has corporate policies and procedures in place to verify data privacy and completed Certification to EU-US Privacy Shield Framework for HR, with annual recertification. IQVIA™ global Council on Data Protection is chartered to monitor global data protection laws and regulations and compliance with these laws and regulations.

All study permit approvals and access to the study data in Sweden and Finland will be applied for by StatFinn-EPID Research, an IQVIA company. After the identification of the study population from the National Healthcare registers, data from each relevant register will be

extracted according to the population identification number (PIN). This will be performed by the register maintainers, according to their own standards. As the data from the national registers consist of data registered independent of the current study, the data collection process cannot affect the research question.

Once all relevant data have been extracted by the register holders and prior to data delivery to StatFinn-EPID Research, each PIN will be replaced with a unique dummy study identification number (SID). StatFinn-EPID Research will then receive pseudonymised raw data without PINs. The SIDs will be used for data linkage on the individual level. The study sponsor or any other parties outside StatFinn-EPID Research cannot receive access to individual-level data. Only aggregated results will be presented to the sponsor or otherwise published. Therefore, only study programmers and statisticians at StatFinn-EPID Research have rights to files and directories that contain individual-level data where individuals cannot be directly identified.

Study data cannot be used for other purposes than described in the study protocol. All requests to use the study data for other purposes must be subjected to appropriate data permit processes.

StatFinn-EPID Research will maintain information on the study individuals securely on site according to up-to-date standard operating procedures. StatFinn-EPID Research will also maintain appropriate data storage, including periodic backup of files and archiving procedures. These procedures will comply with the General Data Protection Regulation (GDPR) and country legislations that include checking electronic files, maintaining security and data confidentiality, following analyses plans, and performing quality checks for all programs. After completion of the study, StatFinn-EPID Research shall erase or destroy all the personal data and existing copies unless EU law or Member State law requires storage of the personal data.

9.7. Data Analysis

Unless otherwise specified, results will be provided as descriptive statistics. All analyses will be conducted within each study country and results may be pooled across countries in a meta-analytical approach. Categorical variables will be reported using frequency distributions.

Continuous variables will be reported using means, standard deviations, medians, minimums, maximums, 25th percentiles, and 75th percentiles, unless otherwise specified. For inferential statistics, both crude and propensity score-adjusted analyses will be reported for study outcomes.

An attrition table will be provided showing how patients qualified for each analysis.

Missing values will be reported as missing, and no imputation will be undertaken. All data analysis will be conducted using the most recent version of SAS® software (SAS Institute Inc., Cary, NC) in use by IQVIA on Windows®. Results will be summarised in tables and figures in Microsoft® (MS) Excel format.

9.7.1. Exposure Lags

As an exploratory analysis, the frequency of cancer endpoints during the first 3 years after the index date will be compared between exposure groups and stratified by three 12-month periods,

which will provide a basis for the assessment of potential detection bias and protopathic bias that may exist between exposure groups.[42]

Given long progression of the cancer outcomes of interest, and to minimise reverse causation, all analyses for the pancreatic cancer and thyroid cancer endpoints will be conducted with lagged exposure times (i.e., latency periods). A primary latency period of 3-years for the cancer endpoints will be used, with additional sensitivity analyses using 0-months, 1-year, and 5-years lag periods. These sensitivity analyses will provide more information on the magnitude of protopathic bias and detection bias.[42] In each lagged analysis, patients without complete follow-up during the latency period (e.g., 5 years of available medical history for the 5-year lagged analysis) or with an outcome during the latency period will be excluded.

9.7.2. Outcome Assessment Period

The primary analysis will be an ITT approach as described in Section 9.3.1.3 (Main Definition of Exposure), in which pancreatic cancer and thyroid cancer will be assessed any time after the end of the exposure lag period described in Section 9.7.1 (Exposure Lags).

The same outcome assessment period of the ITT (that is, any time after the end of the exposure lag period) will be applied to ascertain the outcome for the exposure definitions that is described in Section 9.3.1.4 (Sensitivity Definition of Exposure).

Outcome assessment periods are depicted in Figure 3.

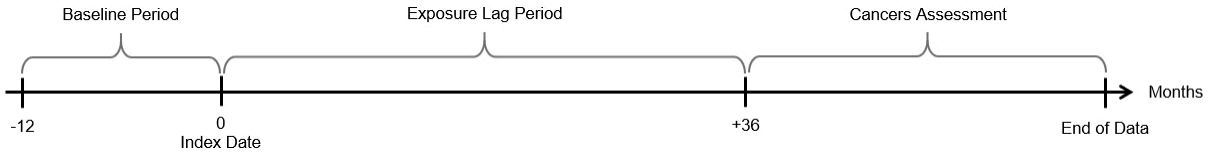


Figure 3. Illustration of outcome assessment periods.

9.7.3. Primary Objective: Incidence Rates

In each of the outcome specific cohorts (pancreatic cancer and thyroid cancer), the incidence rate (IR) and 95% CIs of the outcome of interest (overall and by sub-type, if available) will be calculated among patients in each of the exposure groups (Section 9.3.1 [Exposures]),within the 3 mutually exclusive exposure groups, along with the number of events, total number of individuals, and accrued person-time. Incidence rates will be calculated as the number of incident events of interest within the follow-up period divided by the total person-time at-risk.

9.7.4. Primary Objective: Association of Dulaglutide and Outcomes of Interest

Cox proportional hazards regression with matching on the exposure propensity score (EPS) to control for the potential confounders will be applied to compare exposure groups with respect to the outcomes of interest at the database level. The EPS will represent the probability of exposure on the index date from the baseline covariates listed in Table 2. The EPS matching between exposure groups will be performed using 1:2 nearest neighbour matching with a maximum matching caliper of 1%. In addition, multiple matching ratios and caliper width will be assessed. In addition to graphical depictions of EPS distributions, the standardised differences in proportions and means of baseline covariates will be estimated to examine comparability of exposure groups. HRs with corresponding 95% CIs will be calculated for each comparison of interest. In addition to calculating and presenting the results for each country separately, the HRs from the 3 countries (US, Finland, and Sweden) will be combined in a meta-analysis approach using an inverse variance-weighted, fixed-effect model. Additionally, a combined result in meta-analysis will be presented for the EU databases (Finland and Sweden) to contextualise the findings in the EU compared to the US.

In countries without complete information for the listed confounders (e.g., missing BMI, smoking, and alcohol use information for patients treated outside of military treatment facilities in the US), primary analyses will exclude these confounders from the propensity score model and sensitivity analyses will include the variables among a subset of patients with available information.

In addition to EPS matching, stratification analysis by the EPS quintiles will be performed to obtain quintile-specific and weighted average HRs. The weighted average HRs from the 3 countries will be combined, analogous to the meta-analysis approach in the primary analysis. Similarly, for the EU countries, the weighted average HRs will be pooled in a meta-analysis approach. The stratified analysis will provide more information on the magnitude of confounding by indication.[43]

As an additional sensitivity analysis, competing risks analysis will be explored to investigate the impact of all-cause mortality as an event potentially competing with the occurrence of study outcomes of interest.

Additional details for the analysis of the primary objective and the analysis plan for the secondary objectives will be included in the statistical analysis plan.

9.8. Quality Control

This study will be conducted according to the International Society for Pharmacoepidemiology's 'Guidelines for Good Pharmacoepidemiology Practices (GPP)' (source available from: https://www.pharmacoepi.org/resources/policies/guidelines-08027/ and 'Guideline on good pharmacovigilance practices (GVP), Module VIII – Post-authorisation safety studies (Rev 3)', 09 Oct 2017 (EMA/813938/2011 Rev 3).

At the study level, all aspects of the study from protocol development to the reporting of the results are conducted within the framework of the IQVIA Quality Management System (QMS) and in accordance to the following manual, operating procedures, and work instructions:

- RWI_MAN_RWW0007: Real-World Quality Manual
- RWI_OP_PM0004 Real World Project Quality Control
- RWI_OP_PM0020 Real-World Records Management
- RWI_OP_PM0003 Post-Authorisation Safety Studies (PASS)
- RWI_WI_EPI0005: Protocol Development; and
- RWI_WI_EPI0004: Quality Control of Biostatistics and Epidemiology Deliverables

A Quality Control plan for the study will be developed and executed, which will include quality control on study methodology, statistical analysis plan, programming, data management and analysis, study results, conclusions, and study report. Furthermore:

1) The study Quality Control plan will establish ownership for the execution of the individual Quality Control steps.
2) The Principal in Charge of the study will ensure that individuals responsible for the execution of specific Quality Control steps will have the knowledge, capability, and experience necessary to perform the assigned tasks.
3) The result of the execution of the individual steps of the Quality Control plan will be documented, and will include the required corrective actions, if any. The execution of any required corrective action will be documented.
4) The Quality Control plan will be subjected to a final review and approval for sufficiency and completeness from the Principal on Charge of the study.

9.9. Limitations of the Research Methods

As with any secondary data study, anticipated challenges include possible misclassification. For example, a patient may not necessarily use a prescription that they have filled. To address this limitation, we have proposed sensitivity analyses restricting to patients with sustained use of the exposure medications. Claims-based algorithms will be used to identify the outcomes of interest in the US data source. Although the proposed algorithms have shown an adequate PPV in ascertaining pancreatic and thyroid cancer in previous validation studies of similar data source and population characteristics, there is no information in claims data to differentiate tumour subtypes. While it is possible to identify pancreatic and thyroid cancer in the cancer registries (Sweden and Finland), thyroid cancer is not currently divided in the ICD-O-3 dictionary by subtype. Further, it is currently unknown whether this level of detail is available in EMR data, given that the subtypes are not clearly identified in ICD-O-3 guidelines.

Due to the rare nature of the outcomes and the long induction period for pancreatic and thyroid cancer, the number of patients experiencing these outcomes may be relatively low. To address this limitation, and to explore the GLP-1 RA class effect, the study will include an assessment of all GLP-1 RAs, including those with earlier approval dates (and thus longer follow-up periods).

Additionally, we chose an incident user design which has the limitation of excluding prevalent users of dulaglutide thereby reducing patient sample size. The advantage of an incident user design is that it reduces biases that can impact non-randomised studies, especially when using RWD from healthcare databases.[44] We expect the reduction in sample size, due to an incident user design, to be minimal since the study indexing period includes the time period when dulaglutide (and other first GLP-1 RA) first entered the market in each of the study countries.

We also chose an ITT approach for the primary analysis because it has the advantage of eliminating certain types of biases by preserving the prognostic balance obtained through the exposure propensity score and maintaining sample size.[45,46] Even though ITT analysis is the preferred method of analysis, it is sometimes considered to be “too conservative” and has the possibility of exposure misclassification.[45,46] To address these issues, we will conduct a sensitivity analysis using an as-treated approach that will censor individuals upon discontinuation of index treatment. However, limitations of an as-treated analysis include the possibility of introducing certain biases including differential censoring, informative censoring, or time-dependent confounding biases.[45,46]

Finally, while most databases contain substantial information for inclusion/exclusion criteria and confounder control, medical conditions or a family history of medical conditions are only ascertainable where established diagnoses (ICD-10 codes) and procedures for those conditions exist. Additionally, some key covariates may not be available in all databases (e.g., alcohol abuse is not routinely recorded in Nordic databases, Finland does not have a diabetes register with laboratory data recorded, the MDR only captures BMI and laboratory values for the subset of encounters occurring in military treatment facilities). To address this limitation, the study will incorporate sensitivity analyses among sub-groups with available confounder information available (e.g., a primary analysis in the MDR of all patients meeting inclusion and exclusion

criteria and a sensitivity analysis among only those with laboratory data available). This approach will help to both maintain sample size and ensure that the effects of key confounders are considered. However, environmental, and occupational risk factors are not available in healthcare databases; therefore, residual confounding by these factors will likely be present.

10. Protection of Human Subjects

This study will be conducted in accordance with applicable laws and regulations of the countries where the study is being conducted, as appropriate.

This observational, non-interventional study does not affect the treatment of the patients. The study is conducted by following the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Code of Conduct as well as the Guidelines for Good Pharmacoepidemiology Practice. Eli Lilly and Company, IQVIA, the other participating entities and individuals acting on their behalf commit to adhere to the rules of the ENCePP Code of Conduct in their entirety.

Approval from relevant ethical review boards (ERBs) and other local authorities will be sought in each study country prior to data extraction and analysis. A progress report will be submitted to ERBs and regulatory authorities as required by local laws and regulations.

11. Management and Reporting of Adverse Events/Adverse Reactions

This is a non-interventional study based on secondary data use, and therefore no individual case safety report (ICSR) reporting is required. The study outcomes of interest are the protocol-defined adverse events (AEs) and a summary of AEs will be included in the final study report as planned.

During the course of observational research using existing secondary databases, the proposed study will not involve chart review or validation to obtain additional information on the AEs other than the study outcomes of interest. Thus, Lilly is not expecting to report any AEs or adverse reactions (ARs). Researchers will report all ARs with attribution explicitly stated in the individual patient records to the appropriate party (for example, regulators or marketing authorisation holder) as they would in normal practice as required by applicable laws, regulations, and practices.

12. Plans for Disseminating and Communicating Study Results

The study, including the final report, will be registered in the ENCePP Registry. Two study interim reports will be prepared during data collection.

The final study report will be submitted to the Pharmacovigilance Risk Assessment Committee (PRAC) of the EMA according to the milestones.

13. References

1. U.S. Food and Drug Administration (FDA). Dulaglutide Supplement Approval [press release]. 3 September 2020.
2. De Souza A, Irfan K, Masud F, Saif MW. Diabetes Type 2 and Pancreatic Cancer: A History Unfolding. *JOP*. 2016;17(2):144-148.
3. Bartosch-Harlid A, Andersson R. Diabetes mellitus in pancreatic cancer and the need for diagnosis of asymptomatic disease. *Pancreatology*. 2010;10(4):423-428.
4. Karnevi E, Said K, Andersson R, Rosendahl AH. Metformin-mediated growth inhibition involves suppression of the IGF-I receptor signalling pathway in human pancreatic cancer cells. *BMC Cancer*. 2013;13:235.
5. Pannala R, Basu A, Petersen GM, Chari ST. New-onset diabetes: a potential clue to the early diagnosis of pancreatic cancer. *Lancet Oncol*. 2009;10(1):88-95.
6. Li D. Diabetes and pancreatic cancer. *Mol Carcinog*. 2012;51(1):64-74.
7. Elashoff M, Matveyenko AV, Gier B, Elashoff R, Butler PC. Pancreatitis, pancreatic, and thyroid cancer with glucagon-like peptide-1-based therapies. *Gastroenterology*. 2011;141(1):150-156.
8. Gier B, Matveyenko AV, Kirakossian D, Dawson D, Dry SM, Butler PC. Chronic GLP-1 receptor activation by exendin-4 induces expansion of pancreatic duct glands in rats and accelerates formation of dysplastic lesions and chronic pancreatitis in the Kras(G12D) mouse model. *Diabetes*. 2012;61(5):1250-1262.
9. Nachnani JS, Bulchandani DG, Nookala A, et al. Biochemical and histological effects of exendin-4 (exenatide) on the rat pancreas. *Diabetologia*. 2010;53(1):153-159.
10. Egan AG, Blind E, Dunder K, et al. Pancreatic safety of incretin-based drugs--FDA and EMA assessment. *N Engl J Med*. 2014;370(9):794-797.
11. Bjerre Knudsen L, Madsen LW, Andersen S, et al. Glucagon-like Peptide-1 receptor agonists activate rodent thyroid C-cells causing calcitonin release and C-cell proliferation. *Endocrinology*. 2010;151(4):1473-1486.
12. Bulchandani D, Nachnani JS, Herndon B, et al. Effect of exendin (exenatide)--GLP 1 receptor agonist on the thyroid and parathyroid gland in a rat model. *Eur J Pharmacol*. 2012;691(1-3):292-296.
13. Byrd RA, Sorden SD, Ryan T, et al. Chronic toxicity and carcinogenicity studies of the long-acting GLP-1 receptor agonist dulaglutide in rodents. *Endocrinology*. 2015;156(7):2417-2428.
14. Alves C, Batel-Marques F, Macedo AF. A meta-analysis of serious adverse events reported with exenatide and liraglutide: acute pancreatitis and cancer. *Diabetes Res Clin Pract*. 2012;98(2):271-284.
15. Vahle JL, Byrd RA, Blackbourne JL, et al. Effects of Dulaglutide on Thyroid C Cells and Serum Calcitonin in Male Monkeys. *Endocrinology*. 2015;156(7):2409-2416.

16. Wells SA, Jr., Asa SL, Dralle H, et al. Revised American Thyroid Association guidelines for the management of medullary thyroid carcinoma. *Thyroid*. 2015;25(6):567-610.

17. Ali AK. Methodological challenges in observational research: a pharmacoepidemiological perspective. British Journal of Pharmaceutical Research. 2013;3(3):15.

18. Iglesias ML, Schmidt A, Ghuzlan AA, et al. Radiation exposure and thyroid cancer: a review. *Arch Endocrinol Metab*. 2017;61(2):180-187.

19. Yachida S, Jones S, Bozic I, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*. 2010;467(7319):1114-1117.

20. Pinto LC, Falcetta MR, Rados DV, Leitao CB, Gross JL. Glucagon-like peptide-1 receptor agonists and pancreatic cancer: a meta-analysis with trial sequential analysis. *Sci Rep*. 2019;9(1):2375.

21. Dore DD, Seeger JD, Arnold Chan K. Use of a claims-based active drug safety surveillance system to assess the risk of acute pancreatitis with exenatide or sitagliptin compared to metformin or glyburide. *Curr Med Res Opin*. 2009;25(4):1019-1027.

22. Romley JA, Goldman DP, Solomon M, McFadden D, Peters AL. Exenatide therapy and the risk of pancreatitis and pancreatic cancer in a privately insured population. *Diabetes Technol Ther*. 2012;14(10):904-911.

23. Abrahami D, Douros A, Yin H, et al. Incretin based drugs and risk of cholangiocarcinoma among patients with type 2 diabetes: population based cohort study. *BMJ*. 2018;363:k4880.

24. Wu JW, Azoulay L, Huang A, et al. Identification of incident pancreatic cancer in Ontario administrative health data: A validation study. *Pharmacoepidemiol Drug Saf*. 2020;29 Suppl 1:78-85.

25. Funch D, Ross D, Gardstein BM, et al. Performance of claims-based algorithms for identifying incident thyroid cancer in commercial health plan enrollees receiving antidiabetic drug therapies. *BMC Health Serv Res*. 2017;17(1):330.

26. Falconer H, Yin L, Bellocco R, Altman D. Thyroid cancer after hysterectomy on benign indications: Findings from an observational cohort study in Sweden. *Int J Cancer*. 2017;140(8):1796-1801.

27. Franssila KO. Prognosis in thyroid carcinoma. *Cancer*. 1975;36(3):1138-1146.

28. Galanti MR, Ekbom A, Grimelius L, Yuen J. Parental cancer and risk of papillary and follicular thyroid carcinoma. *Br J Cancer*. 1997;75(3):451-456.

29. Hemminki K, Dong C. Familial relationships in thyroid cancer by histo-pathological type. Int J Cancer. 2000;85(2):201-205.

30. Luoto R, Grenman S, Salonen S, Pukkala E. Increased risk of thyroid cancer among women with hysterectomies. *Am J Obstet Gynecol*. 2003;188(1):45-48.

31. Dorrance KA, Ramchandani S, Neil N, Fisher H. Leveraging the military health system as a laboratory for health care reform. *Mil Med*. 2013;178(2):142-145.

32. Gimbel RW, Pangaro L, Barbour G. America's "undiscovered" laboratory for health services research. *Med Care*. 2010;48(8):751-756.

33. Eriksson JK, Neovius M, Ernestam S, Lindblad S, Simard JF, Askling J. Incidence of rheumatoid arthritis in Sweden: a nationwide population-based assessment of incidence, its determinants, and treatment penetration. *Arthritis Care Res* (Hoboken). 2013;65(6):870-878.

34. Waldenlind K, Eriksson JK, Grewin B, Askling J. Validation of the rheumatoid arthritis diagnosis in the Swedish National Patient Register: a cohort study from Stockholm County. *BMC Musculoskelet Disord*. 2014;15:432.

35. Jakobsson GL, Sternegard E, Olen O, et al. Validating inflammatory bowel disease (IBD) in the Swedish National Patient Register and the Swedish Quality Register for IBD (SWIBREG). *Scand J Gastroenterol*. 2017;52(2):216-221.

36. National Cancer Institute (NIH): Surveillance, Epidemiology, and End Results Program. Cancer Stat Facts: Pancreatic Cancer.

37. Brodovicz KG, Kou TD, Alexander CM, et al. Impact of diabetes duration and chronic pancreatitis on the association between type 2 diabetes and pancreatic cancer risk. Diabetes Obes Metab. 2012;14(12):1123-1128.

38. National Cancer Institute (NIH): Surveillance E, and End Results Program. Cancer Stat Facts: Thyroid Cancer. https://seer.cancer.gov/statfacts/html/thyro.html. Accessed March 14, 2019.

39. Li H, Qian J. Association of diabetes mellitus with thyroid cancer risk: A meta-analysis of cohort studies. *Medicine* (Baltimore). 2017;96(47):e8230.

40. Commission E. ECIS - European Cancer Information System. 2008. https://ecis.jrc.ec.europa.eu/. Accessed January 8, 2020.

41. Randle RW, Balentine CJ, Leverson GE, et al. Trends in the presentation, treatment, and survival of patients with medullary thyroid cancer over the past 30 years. *Surgery*. 2017;161(1):137-146.

42. Ali AK. Biases related to prescribing decisions in retrospective database research in diabetes. *Values and Outcomes Spotlight*. 2015;1(4).

43. Ali AKH, A.G. *Post-Authorization Safety Studies of Medicinal Products: The PASS Book*. Academic Press; 2018.

44. Johnson ES, Bartman BA, Briesacher BA, et al. The incident user design in comparative effectiveness research. *Pharmacoepidemiol Drug Saf*. 2013;22(1):1-6.

45. Park Y, Bateman BT, Kim DH, et al. Use of haloperidol versus atypical antipsychotics and risk of in-hospital death in patients with acute myocardial infarction: cohort study. *BMJ*. 2018;360:k1218.

46. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Intention-to-treat versus per-protocol analysis. *Perspect Clin Res*. 2016;7(3):144-146.

Annex 1. List of Standalone Documents

Methodology Used for Projecting Sample Size and Estimating Mean Follow-up

Study: Dulaglutide and Potential Risks of Pancreatic Cancer and Thyroid Cancer: A Non-Interventional PASS (H9X-MC-B013)

Methodology Used for Projecting Sample Size and Estimating Mean Follow-up

Approval Date: 23-Apr-2021 GMT

Methodology Used for Projecting Sample Size and Estimating Mean Follow-up

TABLE OF CONTENTS

LIST OF TABLES

Methodology Used for Projecting Sample Size and Estimating Mean Follow-up

1 METHODS SUMMARY

1.1 Overview

Access to data from the United States (US) Military Health System (MHS) Data Repository (MDR) and the National Registries in Sweden and Finland is not allowed until a final approved protocol is in place and ethics applications are submitted and approved. Limited summary data can be shared on current patient counts by antidiabetic medication drug categories of interest.

The following summary tables were provided to Lilly:

Table 1. Patient Counts from the US MDR by Treatment Category

Year	Dulaglutide	Other GLP-1 Receptor Agonists	Other Treatments for Type 2 Diabetes Mellitus (excluding GLP-1 Receptor Agonists)
2015	3329	43,723	617,269
2016	4657	48,849	620,372
2017	7010	52,243	601,889
2018	23,965	47,957	612,429
2019	47,315	40,619	591,325
2020	62,385	40,979	569,297
Total unique patients	**76,454**	**74,846**	**806,536**

Abbreviations: GLP-1 = glucagon-like peptide 1; US MDR = United States Military Health System Data Repository.

Table 2. Patient Counts from the Swedish National Registry by Treatment Category

Year	Dulaglutide	Other GLP-1 Receptor Agonists	Other Treatments for Type 2 Diabetes Mellitus (excluding GLP-1 Receptor Agonists)
2015	696	18,468	415,828
2016	2650	21,457	429,634
2017	4911	25,544	444,245
2018	6944	32,741	457,330
2019	7946	49,005	463,460
2020	8292	65,122	467,668

Abbreviation: GLP-1 = glucagon-like peptide 1.

Methodology Used for Projecting Sample Size and Estimating Mean Follow-up

Table 3. Patient Counts from the Finnish National Registry by Treatment Category

Year	Dulaglutide	Other GLP-1 Receptor Agonists	Other Treatments for Type 2 Diabetes Mellitus (excluding GLP-1 Receptor Agonists)
2015	0	13,241	347,417
2016	66	15,622	348,326
2017	930	14,186	357,976
2018	1,813	15,704	366,204
2019	3,470	22,759	367,791
2020	Not yet available	Not yet available	Not yet available

Abbreviation: GLP-1 = glucagon-like peptide 1.

To compute total exposure to dulaglutide over the accrual period, each new user of dulaglutide in a given year (incident user) should be associated with the remaining years of follow-up in the study, thus, providing the respective patient exposure (in patient-years). Those patient-specific exposures are summed up over all incident users to yield the total exposure available in the study.

1.2 Methodology for Projecting Patient Accrual and to Derive the Expected Mean Follow-up Time

1.2.1 US MDR

The number of incident users per year was not available in any of the 3 databases, only the estimated numbers of patients exposed to dulaglutide in each year (prevalent users) were available. However, for the US database, the total number of unique patients during 2015 to 2020 was provided (Table 1). The information allowed an estimation of the potential overall ratio of incident users to prevalent users, calculating as 76,454 / 148,661 = 0.514. Based on the assumption of a constant rate over all years, the year-specific incident users (column 3 in Table 4) were estimated by multiplying the corresponding year-specific prevalent users (column 2) by 0.514. To project the number of dulaglutide initiators beyond 2020, we assumed a flat forecast or stable market for uptake for the calendar years 2021 to 2026. Given the study follow-up until 2030 and accounting for the 3-year latency period to be implemented in the analysis, the maximum follow-up period (column 4) was obtained and was subsequently multiplied by the number of incident users to derive total patient exposure as the sum of all patient-years (column 5). The mean follow-up time will therefore be equal to 1,286,162 / 268,807 = 4.8 years.

Methodology Used for Projecting Sample Size and Estimating Mean Follow-up

Table 4. Projecting Patient Accrual

Year	# Prevalent users (from US MDR)	# Estimated incident users	Maximum follow-up (years)	Retained # incident users for total exposure	Patient-years
2015	3329	1711	12	1711	20,533
2016	4657	2394	11	2394	26,331
2017	7010	3603	10	3603	36,031
2018	23,965	12,318	9	12,318	110,862
2019	47,315	24,320	8	24,320	194,559
2020	62,385	32,066	7	32,066	224,461
2021	62,385	32,066	6	32,066	192,395
2022	62,385	32,066	5	32,066	160,329
2023	62,385	32,066	4	32,066	128,264
2024	62,385	32,066	3	32,066	96,198
2025	62,385	32,066	2	32,066	64,132
2026	62,385	32,066	1	32,066	32,066
2027	62,385	32,066	-	-	-
2028	62,385	32,066	-	-	-
2029	62,385	32,066	-	-	-
Total				268,807	1,286,162

Abbreviation: # = number of; US MDR = United States Military Health System Data Repository.

Although the assumption of a constant incident/prevalent ratio over the years is simplistic, further simulations based on persistence rates as described in study by Divino et al. using real-world data from 5 European countries and Canada (persistence on dulaglutide at 1 year was 36.8% to 67.2%) lead to a similar mean follow-up times of approximately 5 years. The assumption about a flat uptake scenario beyond 2020 is uncertain but not unrealistic given the current competitive landscape.

1.2.2 National Registries in Sweden and Finland

Since the total number of unique patients for each of the treatment categories of interest was not reported for Sweden and Finland, the same ratio between incident and prevalent users (0.514) that was calculated for the US MDR was applied to estimated patient counts for Sweden (Table 2) and Finland (Table 3). Based on this, the projected number of incident users of dulaglutide was calculated to be 41,755 for Sweden and 15,721 for Finland, and the expected mean follow-up time was 5.5 years and 4.9 years for Sweden and Finland, respectively.

2 REFERENCE

1. Divino, V., et al., *GLP-1 RA Treatment and Dosing Patterns Among Type 2 Diabetes Patients in Six Countries: A Retrospective Analysis of Pharmacy Claims Data.* Diabetes Ther, 2019. **10**(3): p. 1067-1088.

Leo Document ID = a2df49f4-4300-4bd7-a711-a57fe19ea8bd

Approver: PPD

Approval Date & Time: 23-Apr-2021 14:24:05 GMT

Signature meaning: Approved

Approver: PPD

Approval Date & Time: 23-Apr-2021 15:17:36 GMT

Signature meaning: Approved

Approver: PPD

Approval Date & Time: 23-Apr-2021 15:37:36 GMT

Signature meaning: Approved

Approver: PPD

Approval Date & Time: 23-Apr-2021 16:02:07 GMT

Signature meaning: Approved

Annex 2. ENCePP Checklist for Study Protocols

Study title: Dulaglutide and Potential Risks of Pancreatic Cancer and Thyroid Cancer: A Non-Interventional PASS

EU PAS Register® number: EUPAS32646

Study reference number (if applicable): H9X-MC-B013

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection2	✓	☐	☐	6
1.1.2 End of data collection3	✓	☐	☐	6
1.1.3 Progress report(s)	✓	☐	☐	6
1.1.4 Interim report(s)	✓	☐	☐	6
1.1.5 Registration in the EU PAS Register®	✓	☐	☐	6
1.1.6 Final report of study results.	✓	☐	☐	6

Comments:

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] 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.

[3] Date from which the analytical dataset is completely available.

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
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) 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)	✓	☐	☐	9.2.1
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	✓	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	✓	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g., cohort, case-control, cross-sectional, other 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., rate, risk, prevalence)	✓	☐	☐	9.7.3
3.4 Does the protocol specify measure(s) of association? (e.g., risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	✓	☐	☐	9.7.4
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:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1 Is the source population described?	✓	☐	☐	9.4
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period	✓	☐	☐	9.2.3
4.2.2 Age and sex	✓	☐	☐	9.2.1.1 & 9.2.1.2
4.2.3 Country of origin	✓	☐	☐	9.2.1
4.2.4 Disease/indication	✓	☐	☐	9.2.1.1
4.2.5 Duration of follow-up	✓	☐	☐	9.2.4
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	✓	☐	☐	9.2.1.1 & 9.2.1.2

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)	✓	☐	☐	9.3.1

5.2 Does the protocol address the validity of the exposure measurement? (e.g., precision, accuracy, use of validation sub-study)	✓	☐	☐	9.3.1.2 & 9.3.1.4
5.3 Is exposure categorised according to time windows?	✓	☐	☐	9.3.1.3 & 9.3.1.4
5.4 Is intensity of exposure addressed? (e.g., dose, duration)	✓	☐	☐	9.3.1.4

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	✓	☐	☐	9.3.1
5.6 Is (are) (an) appropriate comparator(s) identified?	✓	☐	☐	9.3.1.1

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?	✓	☐	☐	9.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	✓	☐	☐	9.3.2
6.3 Does the protocol address the validity of outcome measurement? (e.g., precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☐	☐	✓	
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	✓	

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g., confounding by indication)				

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
	✓	☐	☐	9.3.1.2 & 9.7.4
7.2 Does the protocol address selection bias? (e.g., healthy user/adherer bias)	✓	☐	☐	9.2.1.2; 9.2.2; 9.3.1 & 9.7.1
7.3 Does the protocol address information bias? (e.g., misclassification of exposure and outcomes, time-related bias)	✓	☐	☐	9.2.1.2 & 9.3.1

Comments:

Section 8: Effect measure 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)	✓	☐	☐	9.2.1.2; 9.3.1 & 9.3.1.2

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:	✓			9.4
9.1.1 Exposure? (e.g., pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3

9.1.2 Outcomes? (e.g., clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3
9.1.3 Covariates and other characteristics?	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
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)	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3
9.2.2 Outcomes? (e.g., date of occurrence, multiple events, severity measures related to event)	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3
9.2.3 Covariates and other characteristics? (e.g., age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g., WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3
9.3.2 Outcomes? (e.g., International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3
9.3.3 Covariates and other characteristics?	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3
9.4 Is a linkage method between data sources described? (e.g., based on a unique identifier or other)	✓	☐	☐	9.4.1; 9.4.2 & 9.4.3

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	✓	☐	☐	9.7

10.2 Is study size and/or statistical precision estimated?	✓	☐	☐	9.5

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.3 Are descriptive analyses included?	✓	☐	☐	9.7
10.4 Are stratified analyses included?	✓	☐	☐	9.7.1 & 9.7.4
10.5 Does the plan describe methods for analytical control of confounding?	✓	☐	☐	9.3.1.2 & 9.7.4
10.6 Does the plan describe methods for analytical control of outcome misclassification?	☐	☐	✓	
10.7 Does the plan describe methods for handling missing data?	✓	☐	☐	9.7
10.8 Are relevant sensitivity analyses described?	✓	☐	☐	9.3.1.2; 9.3.1.4; 9.7.1 & 9.7.2

Comments:

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)	✓	☐	☐	9.6 & 9.7
11.2 Are methods of quality assurance described?	✓	☐	☐	9.8
11.3 Is there a system in place for independent review of study results?	✓	☐	☐	9.8

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1.1 Selection bias? 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).	✓ ✓ ✓	☐ ☐ ☐	☐ ☐ ☐	9.9
12.2 Does the protocol discuss study feasibility? (e.g., study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	✓	☐	☐	9.5

Comments:

Section 13: Ethical/data protection 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?	✓	☐	☐	9.6

Comments:

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: PPD

Date: dd/Month/year

Signature: _____

Annex 3. Characteristics of US Data Source

Study title: Dulaglutide and Potential Risks of Pancreatic Cancer and Thyroid Cancer: A Non-Interventional PASS

EU PAS Register® number: EUPAS32646

Study reference number (if applicable): H9X-MC-B013

Table 3. Overview of Characteristics of US Data Source – Military Health System Data Repository (MDR)

Description	TRICARE	Armed Forces Health Longitudinal Technology Application (AHLTA)
Year available	2000 or earlier	2000
Database type	The TRICARE nationwide insurance program provides benefits to the comprehensive medical network within the US Department of Defense (DoD), which covers all military personnel, their dependents and retirees	AHLTA is the DoD's EMR system that covers only military treatment facilities; it links to encounter vitals, laboratory, pathology, and radiology orders and results (available for 15% to 20% of encounter data)
Purpose	Administrative/Claims	EMR
Database size	Over 10 million patients actively receiving care annually from DoD facilities (65 hospitals and 500+ clinics) and/or civilian facilities; roughly 2/3 of care treated in civilian sector via TRICARE	Over 10 million patients actively receiving care annually from DoD facilities (65 hospitals and 500+ clinics) and/or civilian facilities; roughly 1/3 of care treated in military treatment facilities; Approximately 800 estimated new cases of thyroid and pancreatic cancer across 5 years
Representativeness of patients	13% of beneficiaries are active duty US military; 51% are male; the age distribution is similar to the US population	Patients using US military care skews younger, as civilian care is largely driven by beneficiaries who retire and reside is areas in which military facilities are not necessarily available
Type of health care contact or source of data	Inpatient hospital, outpatient hospital, emergency room, physician's office	Inpatient hospital, outpatient hospital, emergency room, physician's office
Data on medications	All dispensing details, including hospital / facility outpatient pharmacies, retail, and mail order	All prescribing details, including hospital / facility outpatient pharmacies, retail, and mail order

Medication information available	Dispensed drug name and NDC, the treating facility, department(s) rendering care, the treating provider(s), amount dispensed, date dispensed, route of medication, units, number of refills, and remaining refills	Prescribed drug name and NDC, the treating facility, department(s) rendering care, the treating provider(s), start and end dates of the prescription order, amount dispensed, date dispensed, route of medication, units, number of refills, and remaining refills
Dose	Based on pharmacy dispensing data	Based on physician prescribing data
Duration	Based on pharmacy dispensing data	Based on physician prescribing data
Drug dictionary codes	NDC, HCPCS	NDC, HCPCS
Outpatient diagnosis	100%	100%, available for military facility encounters only
Hospital diagnosis	100%	100%, available for military facility encounters only
Diagnoses/clinical indication/disease codes	ICD-9-CM, ICD-10-CM, DRG	ICD-9-CM, ICD-10-CM, DRG
Procedure codes	CPT-4, HCPCS	CPT-4, HCPCS
Diagnostic examinations	CPT-4, HCPCS	CPT-4, HCPCS
Laboratory tests	Orders	Orders and results
Lifestyle risk factors (e.g., BMI, smoking and alcohol history)	Not available	Available, if documented in the charts
Source of data for validation	Medical charts in AHLTA	N/A
Linkage to other data sources	AHLTA	TRICARE
Approximate time lag (the frequency of update)	60 to 90 days (continuously)	60 to 90 days (quarterly)

CPT = Current Procedural Terminology; DRG = Diagnosis-Related Groups; EMR = electronic medical record; HCPCS = Healthcare Financing Administration Common Procedure Coding System; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, 10th Revision; ICD-10-CM = International Classification of Diseases, 10th Revision, Clinical Modification; ICD-9-CM = International Classification of Diseases, 9th Revision, Clinical Modification; N/A = not applicable; NDC = National Drug Codes.