NON-INTERVENTIONAL STUDY PROTOCOL

TITLE	Database Linkage Study to Evaluate the Risk of Medullary Thyroid Carcinoma
PROTOCOL/STUDY NO.	CCI Eli Lilly: I8F-MC-B014
VERSION	V 2.0 PASS Protocol Previous version 1.0 Amendment #: N/A
CCI	
CCI	

This protocol contains confidential information that should only be disclosed to those persons responsible for execution and organization of the study and on condition that all such persons agree not to further disseminate it.

Redacted Protocol

Document ID: VV-PVG-111959

Post-Authorization Safety Study (PASS) Information

Title	Database Linkage Study to Evaluate the Risk of Medullary Thyroid Carcinoma
Study Identifier	CCI Eli Lilly: I8F-MC-B014
Version Identifier	Number 2.0
Date of Last Version	20 September 2024
EU PAS Register No:	EUPAS1000000513
Active Substance	*Glucagon-like peptide-1, Glucagon-like peptide-1 receptor agonists*
Medicinal Product(s)	Eli Lilly products: dulaglutide (Trulicity®) and tirzepatide (Mounjaro®, Zepbound®) CCI Final list to be included in the statistical analysis plan (SAP) and/or final study report as appropriate
Product Reference	See product list above
Procedure Number	EMEA/H/C/005620
Marketing Authorization Holder(s)	Eli Lilly and Company CCI
Joint PASS	Yes
Research Question and Objectives	The primary objectives are to: • estimate the incidence of medullary thyroid carcinoma (MTC) among adults (18 years of age and older) in the US (hereafter referred to as adult patients) who are exposed to long-acting glucagon-like peptide-1 receptor agonists (LA GLP-1 RA) therapies, as compared to adult patients initiating an active comparator medication using incidence rate ratios (IRRs) and 95% confidence intervals (CIs), and • characterize adult patients exposed to LA GLP-1 RA therapies, and active comparator cohorts using demographic characteristics and other clinical characteristics, including selected prescription medications dispensed during the baseline period, and duration of LA GLP-1 RA therapy use. The secondary objective is to: • evaluate trends in the annual incidence of MTC in adult patients in the US for identification of any possible increase related to the introduction of LA GLP-1 RA therapies, into the US market.
Country of study	United States

Abbreviations: LA GLP-1 RA = long-acting glucagon-like peptide-1 receptor agonists; No. = number; PAS = Post-Authorization Study; PASS = Post-Authorization Safety Study.

TABLE OF CONTENTS

Section **Page**

List of Tables

Table **Page**

List of Figures

Figure **Page**

PPD

Study Synopsis

Full Study Title: Database Linkage Study to Evaluate the Risk of Medullary Thyroid Carcinoma			
Phase:	Not applicable	**Type:**	Observational
Number of Patients: TBD		**Duration of Patient Participation**: TBD	
Number of Sites: TBD		**Duration of study**: ~5 years	

Background/Rationale: Glucagon-like peptide-1 receptor agonists (GLP-1 RAs) were first indicated as an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus (T2D). Some long-acting GLP-1 RAs were subsequently approved as an adjunct to a reduced-calorie diet and increased physical activity for chronic weight management in adult patients that meet body mass index (BMI) and comorbidity criteria.

GLP-1 RAs stimulate glucose-dependent insulin release, slow gastric emptying, inhibit inappropriate postprandial glucagon release, and reduce food intake. These effects are mediated by a G-protein coupled receptor, GLP-1R, which is widely distributed throughout a variety of tissues. Nonclinical studies in rodents of clinically relevant doses of GLP-1 RAs showed dose-related and treatment-duration dependent increases in the incidence of thyroid C-cell tumors (adenomas and carcinomas). The clinical relevance of rodent thyroid findings observed with GLP-1 RAs is unknown.

Medullary Thyroid Carcinoma (MTC) is the human equivalent of C-cell carcinoma in rodents. MTC is a rare form of human cancer. In the US, thyroid cancer represents approximately 2% of all cancer types and MTC accounts for a small percentage of thyroid cancer overall, with estimates of the proportion ranging from approximately 3% to 4%. In addition, approximately 1,000 people are diagnosed with MTC each year in the US, translating to approximately 0.2 to 0.3 cases per 100,000 adults.

MTC manifests in a sporadic (75%) or hereditary form (25%) including multiple endocrine neoplasia syndromes (MEN) MEN2A and 2B or familial MTC (FMTC). Activating rearranged during transfection (RET) proto-oncogene mutations generally are present in the hereditary MTC syndromes; and these same germline mutations may be found in approximately 6% of patients, with apparently sporadic varieties. The overall 5-year survival rate of MTC is between 65% and 89% and the 10-year survival rate is approximately 75% to 85%; however, given the sheer rarity of the disease, the accuracy of the existing survival rates is unknown.

This is a database linkage study utilizing real-world data (RWD). The aim in this study is to estimate the incidence of MTC among adult patients who initiated treatment with GLP-1 RA therapies and the glucose-dependent insulinotropic polypeptide/glucagon-like peptide-1 receptor agonist (GIP/GLP-1 RA) tirzepatide, incretin-based therapies, hereafter referred to as long-acting (LA) GLP-1 RA therapies, as compared to 3 active comparator cohorts using incidence rate ratios (IRRs) and 95% confidence interval (CI). The use of RWD access to many patients exposed to LA GLP-1 RA therapies, longitudinal data capture during follow-up, and validated cancer outcomes from State Cancer Registry (SCR) data. Clinical information from both open and closed claims sources will be used to conduct separate sensitivity analyses.

Objectives:

The primary objectives are to:

- estimate the incidence of MTC among adults (18 years of age and older) in the US (hereafter referred to as adult patients) who are exposed to LA GLP-1 RA therapies, as compared to adult patients initiating an active comparator medication using IRRs and 95% CIs, and
- characterize adult patients exposed to LA GLP-1 RA therapies, and active comparator cohorts using demographic and other clinical characteristics, including selected prescription medications dispensed during the baseline period, and duration of LA GLP-1 RA therapy use.

The secondary objective is to:

- evaluate trends in the annual incidence of MTC in adult patients in the US for identification of any possible increase related to the introduction of LA GLP-1 RA therapies, into the US market.

Study design: This is a database linkage study with an active comparator new user study design. This study will use dispensed prescription claims and SCR data to assess the risk of MTC related to use of LA GLP-1 RAs in adults ≥18 years of age between 01 January 2010, and 31 December 2023 (study patient selection period). Data on treatment exposures will be obtained from dispensed prescription claims using a US real-world claims database, namely, the IQVIA Longitudinal Prescriptions Database (LRx). MTC diagnosis information will be obtained from the US SCR data. Linkage of patients identified in LRx to SCR data for MTC outcomes is scheduled to occur in 2025. Individuals are expected to have data until 31 December 2023, and only a subset of patients may have data available past December 2023 due to SCR data lags, which can extend up to 2 years. A 24-month baseline period prior to the index date will be applied and therefore, data will be extracted from 01 January 2008. Comparative analyses will be conducted using propensity score (PS) weighting methodology. The PS for the primary analysis will include demographic variables, payer type, index year, a count of therapeutic classes dispensed and a proxy for diabetes severity (determined by time since the first antidiabetic therapy) during baseline. Additional covariates will be included in the sensitivity analysis such as history of radiation use, history of cancer (not including MTC), inpatient visits, outpatient visits, diabetes severity (determined by clinical information), Charlson comorbidity index (CCI) (modified to exclude diabetes, prior to the index date) and thyroid ultrasound.

Study Population:

All eligible patients must fulfill the following requirements:

- ≥18 years or older during the year of index (date of qualifying medication)
- In addition to index medication, ≥1 dispensed medication (any class) within 24 months prior to the index date[1]
- ≥1 dispensed medication (any class) after the index date and within 12 months of index date (12-month post-index period)
- No evidence of cohort qualifying medication during 12-month before the index date
- Reside in the US, including the District of Columbia (DC) during the study period
- No missing values for year of birth

Patient eligibility for inclusion by study cohorts, in addition to the above criteria are captured below.

LA GLP-1 RA therapies exposed cohort:

- ≥1 dispensed prescription for a LA GLP-1 RA therapy during the study patient selection period

T2D Active Comparator 1 Cohort:

- ≥1 dispensed prescription for sodium-glucose transport protein 2 (SGLT2) or dipeptidyl peptidase IV (DPP-4) inhibitors during the study patient selection period

T2D Active Comparator 2 Cohort:

- ≥1 dispensed prescription for any antidiabetic medication, other than LA GLP-1 RA therapies, during the study patient selection period

Overweight/Obesity Active Comparator Cohort:

- ≥1 dispensed prescription for any anti-obesity medication, other than LA GLP-1 RA therapies, during the study patient selection period

Patients with a diagnosis of MTC in the linked SCR data before their index date will not be included in the study.

Data collection/Data Sources: This will be a database linkage study utilizing the IQVIA Longitudinal Prescription Database (LRx) and linked to the US SCR data. The IQVIA Open Claims (Dx) and PharMetrics® Plus (P+) databases will be used for sensitivity analyses; no patients will be excluded due to missing DX and P+ data linkage.

Data Management and Quality Assurance: The IQVIA personnel are responsible for the integrity of the data reported to the clients. Datasets and analytic programs will be stored according to IQVIA procedures with access restricted to study personnel. Data provided by the SCRs will be destroyed following data destruction procedures specified by the SCRs and agreed to by IQVIA.

[1] For the exposed cohort: the date of the first LA GLP-1 RA therapy prescription dispensed medications during the study patient selection period; for the un-exposed cohorts: the date of the first comparator prescription medications dispensed.

Safety: This is a non-interventional study based on secondary data use, and therefore, no individual case safety reports (ICSRs) are required. The study protocol-defined adverse events (AEs) include: MTC. All protocol-defined adverse events collected will be summarized in the final study report. No other AEs will be collected.

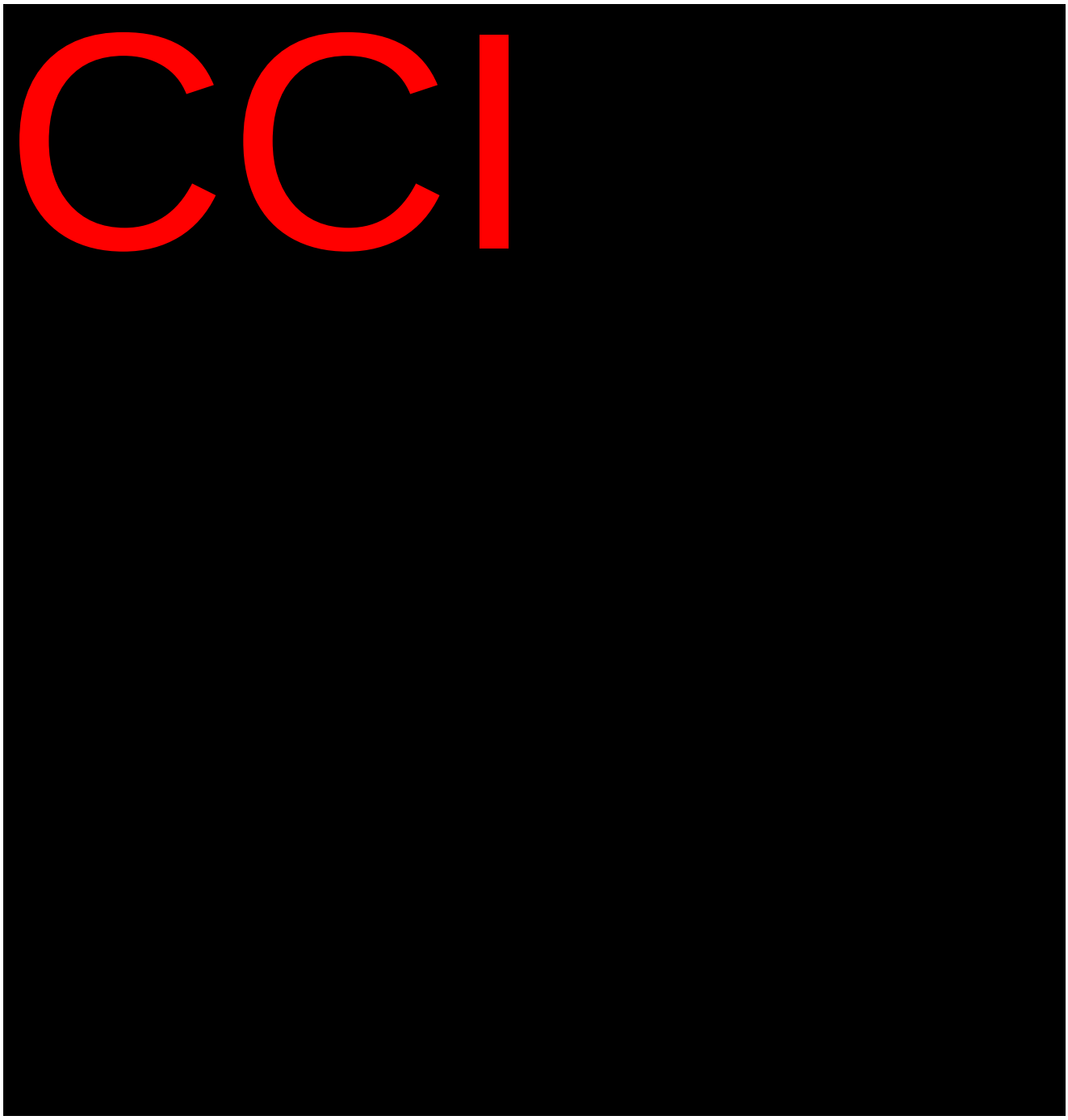
Statistical Considerations: Study cohorts will be adjusted for SCR participation using two different methods which will create two sets of four study cohorts. All study estimates from primary, exploratory and sensitivity analyses will be estimated for both SCR adjusted sets of study cohorts. For the primary analysis, the incidence rate (IR), IRR, and 95% CI for MTC occurrence among LA GLP-1 RA therapy users and active comparators will be estimated using Negative Binomial regression, Poisson regression, or both, (depending on data fit and model assumptions). Descriptive statistics will be calculated, unless otherwise specified, for the LA GLP-1 RA therapy-exposed cohorts and the three comparator cohorts (that is, T2D active comparator 1, T2D active comparator 2 and overweight/obesity active comparator cohort). The official federal cancer statistics program, the United States Cancer Statistics, will be used to provide data for estimating annual incidence of MTC in the US during the study period. To understand the annual MTC incidence trends over time, the IRs will be stratified by year and graphically presented in the final study report.

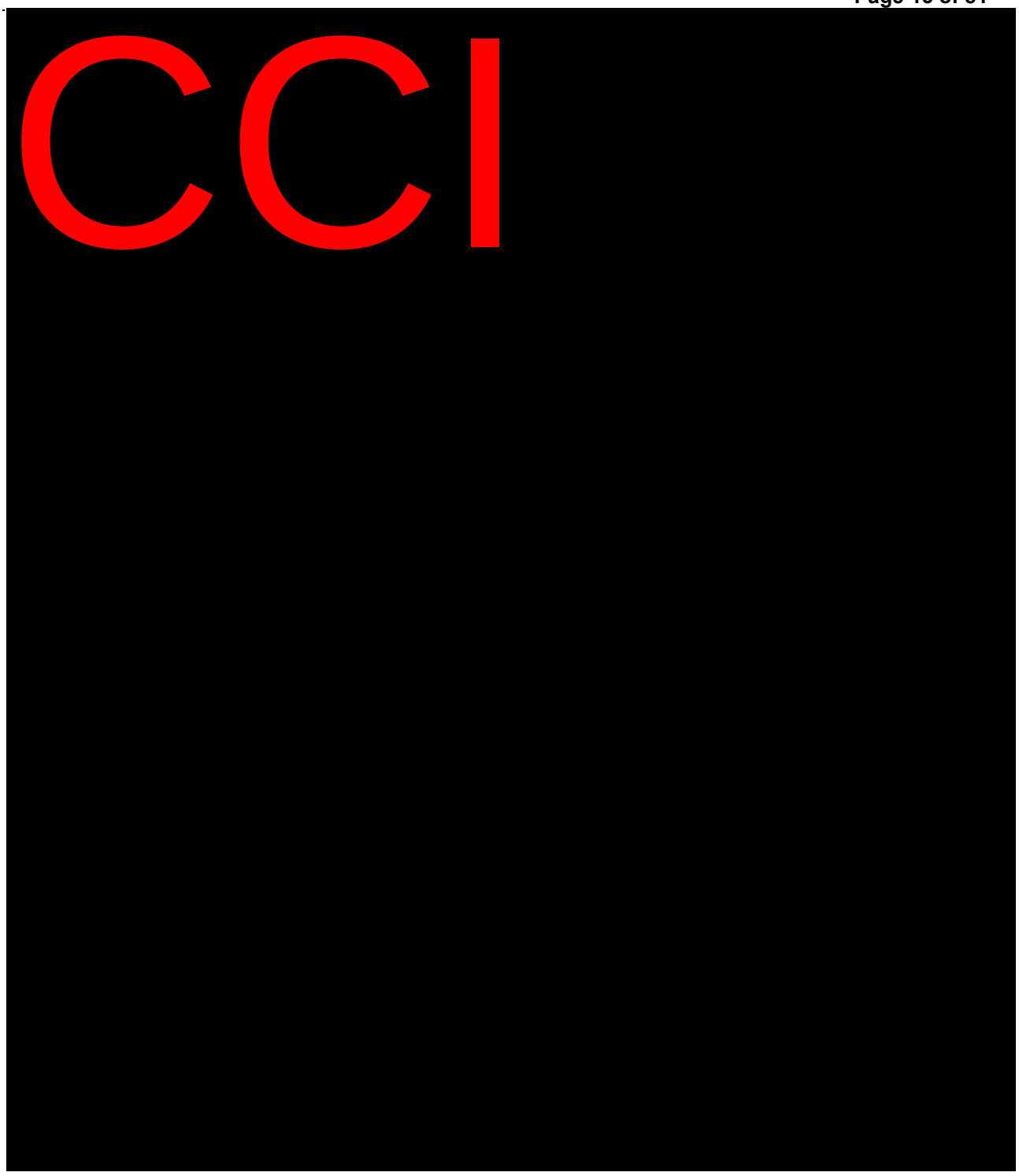
Sample size: Based on data from a feasibility assessment, it is estimated that within LRx, approximately 13.4 million total unique patients had a dispensed prescription for a LA GLP-1 RA therapy product between January 2010 and December 2023. An attrition table will be reported for each cohort comparison and time period showing the number of patients remaining after each inclusion criterion. The sample sizes are expected to reduce with application of additional selection criteria. Assuming approximately 13.4 million GLP-1 RA users between 2010 and 2023, up to 4 times the sample size of the exposed cohort (up to 53.6 million comparator patients in each comparator group), α=0.05, two-sided test, and a background MTC IR of 0.2 events per 100,000 person-years, the study would achieve 80% power to detect a hazard ratio (HR) of 1.2. If the number of LA GLP-1 RA users is higher (that is, >13.4 million), the smaller of an effect becomes detectable (HR <1.2). If the number of LA GLP-1RA users are smaller than 13.4 million or the size of comparator cohorts is smaller than 53.6 million, then the larger hazard ratio HR is detectable (>1.2).

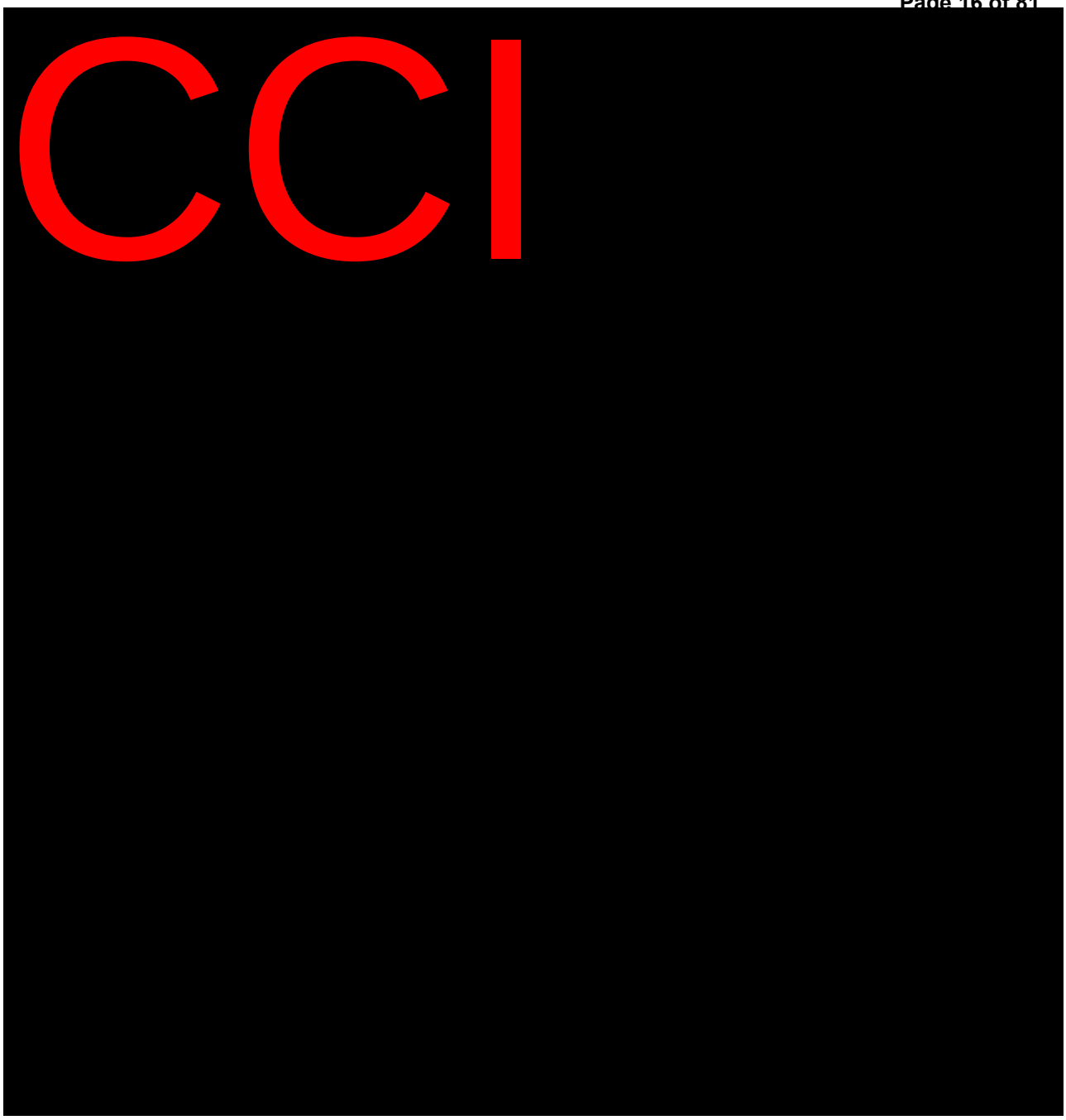
Final Analyses: A final study report will be prepared after completion of the database linkage and analysis and is planned for Q1 2027. The final study report will encompass all planned analyses, including a description of the complete study population and patient outcomes, as described in the SAP. All reporting will be consistent with the STROBE (STrengthening the Reporting of OBservational studies in Epidemiology).

Ethical and Regulatory Considerations: This non-interventional study will be conducted in accordance with the protocol and all applicable laws and regulations including, but not limited to good pharmacoepidemiology practices (GPP), good pharmacovigilance practice (GVP) module VIII for post-authorization safety studies rev. 3, and the ethical principles that have their origins in the Declaration of Helsinki and applicable privacy laws. Data protection and privacy regulations will be strictly observed in capturing, forwarding, processing, and storing patient data. Every effort will be made to protect patient confidentiality according to the Directive 95/46/EC on the protection of individuals, and in compliance with Safe Harbor privacy principles. An International Review Board (IRB)/ Independent Ethics Committee (IEC) must review and approve the protocol before any patients are enrolled.

Documentation of Protocol Amendments







1. MILESTONES

Milestone	Planned date
Protocol Submission to the CCI)	Q3 2023
Start date of data collection[2]	Q4 2025
End of data collection	Q4 2025
Study progress report [n] if applicable	Not applicable
Registration in the EU PAS register	Within 3 months after protocol finalization
Final report of study results submission	Q1 2027

[2] The start of data collection for secondary use of database studies is the date from which any data extraction starts. The end of data collection is the date from which the analytical dataset is completely available. Data collection and extraction dates are contingent on SCR IRB approvals and execution of data use agreements.

2. BACKGROUND

Glucagon-like peptide-1 receptor agonists (GLP-1 RAs) were first indicated as an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus (T2D). Some long-acting GLP-1 RAs were subsequently approved as an adjunct to a reduced-calorie diet and increased physical activity for chronic weight management in adult patients that meet body mass index BMI and comorbidity criteria.

GLP-1 RAs stimulate glucose-dependent insulin release, slow gastric emptying, inhibit inappropriate postprandial glucagon release, and reduce food intake. These effects are mediated by a G-protein coupled receptor, GLP-1R, which is widely distributed throughout a variety of tissues. Nonclinical studies in rodents of clinically relevant doses of GLP-1 RAs showed dose-related and treatment-duration dependent increases in the incidence of thyroid C-cell tumors (adenomas and carcinomas).[1,2] The clinical relevance of rodent thyroid findings observed with GLP-1 RAs is unknown.

Medullary Thyroid Carcinoma (MTC) is the human equivalent of C-cell carcinoma in rodents. MTC is a rare form of human cancer. In the US, thyroid cancer represents approximately 2% of all cancer types and MTC accounts for a small percentage of thyroid cancer overall, with estimates of the proportion ranging from approximately 3% to 4%.[3] In addition, approximately 1,000 people are diagnosed with MTC each year in the US, which translates to approximately 0.2 to 0.3 cases per 100,000 adults.[4]

MTC manifests in a sporadic (75%) or hereditary form (25%) including multiple endocrine neoplasia syndromes (MEN) (MEN2A and 2B) or familial MTC.[5,6] Activating rearranged during transfection (RET) proto-oncogene mutations generally are present in the-hereditary MTC syndromes; and these same germline mutations may be found in approximately 6% of patients, with apparently sporadic varieties.[7,8] The overall 5-year survival rate of MTC is between 65% and 89% and the 10-year survival rate is approximately 75% to 85%; however, given the sheer rarity of the disease, the accuracy of the existing survival rates is unknown.[3]

3. RATIONALE

This is a database linkage study utilizing real-world data (RWD). The aim in this study is to estimate the incidence of MTC among adult patients who initiated treatment with GLP-1 RAs and the glucose-dependent insulinotropic polypeptide/glucagon-like peptide-1 receptor agonist (GIP/GLP-1 RA) tirzepatide, hereafter referred to as long-acting (LA) GLP-1 RA therapies, as compared to 3 active comparator cohorts using incidence rate ratios (IRRs) and 95% confidence intervals (CIs). The use of RWD provides access to many patients exposed to LA GLP-1 RA therapies, longitudinal data capture during follow-up, and validated cancer outcomes from state cancer registry (SCR) data. Clinical information from both open and closed claims sources will be used to conduct separate sensitivity analyses.

4. OBJECTIVES

The primary objectives are to:

- estimate the incidence of MTC among adults (18 years of age and older) in the US (hereafter referred to as adult patients) who are exposed to LA GLP-1 RA therapies, as compared to adult patients initiating an active comparator medication using IRRs and 95% CIs.
- characterize adult patients exposed to LA GLP-1 RA therapies, and active comparator cohorts using demographic characteristics and other clinical characteristics, including selected prescription medications dispensed during the baseline period, and duration of LA GLP-1 RA therapy use.

The secondary objective is to:

- evaluate trends in the annual incidence of MTC in adult patients in the US for identification of any possible increase related to the introduction of LA GLP-1 RA therapies, into the US market.[3]

[3] Trends of annual incidence of MTC will not be submitted to Health Authorities on an annual basis but will support the interpretation of the totality of data and included with submission of the final report.

5. STUDY DESIGN

5.1. Study Overview

This is a database linkage study with an active comparator new user study design. The study will use dispensed prescription claims and SCR data to assess the risk of MTC related to use of LA GLP-1 RA therapy in adults ≥18 years of age between 01 January 2010 and 31 December 2023 (study patient selection period). Data on treatment exposures will be obtained from dispensed prescription claims using a US real-world claims database, namely, the IQVIA Longitudinal Prescriptions Database (LRx). MTC diagnosis information will be obtained from the US SCR data. Linkage of patients identified in LRx to SCR data for MTC outcomes is scheduled to occur in 2025. Individuals are expected to have data until 31 December 2023, and only a subset of patients may have data available past December 2023 due to SCR data lags, which can extend up to 2 years. A 24-month baseline period prior to the index date will be applied and therefore, data will be extracted from 01 January 2008 (Table 1).

The study cohorts will be created using LRx and then linked to SCR data to determine if patients have been diagnosed with MTC (primary analysis, see Section 6.2.2). The exposed cohort will be comprised of patients with a pharmacy-dispensed prescription for a LA GLP-1 RA therapy during the study patient selection period. The comparator cohorts will be defined as follows: patients with ≥1 dispensed prescription of SGLT2 or DPP-4 inhibitors will be eligible for the T2D active comparator 1 cohort; patients with ≥1 dispensed prescription for any antidiabetic medication (ADM) (including SGLT2 or DPP-4 inhibitors) other than LA GLP-1 RA therapies will be eligible for the T2D active comparator 2 cohort; patients with ≥1 dispensed prescription for anti-obesity medication (AOM) other than LA GLP-1 RA therapies will be eligible for the overweight/obesity active comparator cohort. Eligible patients will index on the dispensing date of the cohort-qualifying prescription. The incidence rate (IR), IRR, and their respective 95% CIs for MTC occurrence in LA GLP-1 RA therapy users and active comparators will be estimated using Negative Binomial regression, Poisson regression, or both (depending on data fit and model assumptions). Comparative analyses will be conducted using PS weighting methodology. The PS will incorporate demographic variables, payer type, index year, a count of therapeutic classes dispensed and a proxy for diabetes severity (determined by time since the first antidiabetic therapy) during baseline.

The study will use the observational parallel to the intention-to-treat analytic approach. Patients exposed to GLP-1 RA therapy will be first identified and sampled without replacement. As a result, the patients remaining in the comparator pool will not have a documented GLP-1 RA exposure during the study patient selection period. As the next step, the active comparator cohorts will be identified with replacement from the comparator pool (see Section 5.2.2 for details). Switching between different GLP-1 RA therapies in the exposed cohort will be described in a sensitivity analysis, as described in Section 6.2.2.6.1.

Table 1. Study Time Periods

Study time periods	Dates
Study period	01 January 2008 through 31 December 2025
Study patient selection period	01 January 2010 through 31 December 2023
Study baseline period (24-month baseline)[a]	A period of up to 24 months prior to patient index date (01 January 2008 through 30 December 2023)
Study lookback period for patient selection (12-month lookback)[b]	A period of up to 12 months prior to the patient index date (01 January 2009 through 30 December 2023)
Study follow-up period for outcome identification[2]	01 January 2010 to 31 December 2025

[a] Study baseline period and lookback period will begin prior to the index date (not including index date); study baseline period can be up to 24 months long for covariate identification and study lookback period can be up to 12 months long for patient selection requirement.

[b] Study follow-up period will start from the index date and end with the end of the study period, or MTC diagnosis, whichever comes first.

5.1.1. Database Selection

The databases used in this linkage study were carefully selected to provide the most robust study structure using RWD to answer the research question: is there an association between LA GLP-1 RA use and MTC? MTC is a very rare disease with an incidence falling between 0.2 to 0.3 cases per 100,000 adults per year. Open claims databases have access to large numbers of patients, which is important for evaluation of a rare disease such as MTC and increases the level of precision in study effect estimates.

The study databases span a variety of RWD sources, all with critical attributes that add value to the study design. Open claims data sources provide substantial coverage of the US population, a wide variety of payer types (commercial insurance, cash, Medicaid, Medicare part D) and short data lags (meaning data is available within days). The open claims databases selected for this study have broad US coverage; however, they do not capture when a patient enrolls with a particular payer. While it is acknowledged that missing data in open claims can't be routinely quantified, the large breadth of coverage of open claims data offers confidence in its strength in capturing patient activities both with their insurers and via other avenues (self-pay care, for instance). To address some of these limitations, the study design includes a closed claims data source, which captures fewer payer types but offer increased depth to patient medical history within a single insurer, including enrollment information. This linkage study has planned sensitivity analyses in both an open and closed clinical claims data source to support the primary analyses. For outcome identification, SCRs will be leveraged. SCRs are population-based cancer registries offering the most comprehensive view of observational cancer data with fully validated outcomes in the US that is representative of the state populations (and the entire US, if pooled).

Specifically, this database linkage study will maximize the study population by utilizing an open claims data source (LRx) for the patient selection and cohort creation linked to SCRs for MTC

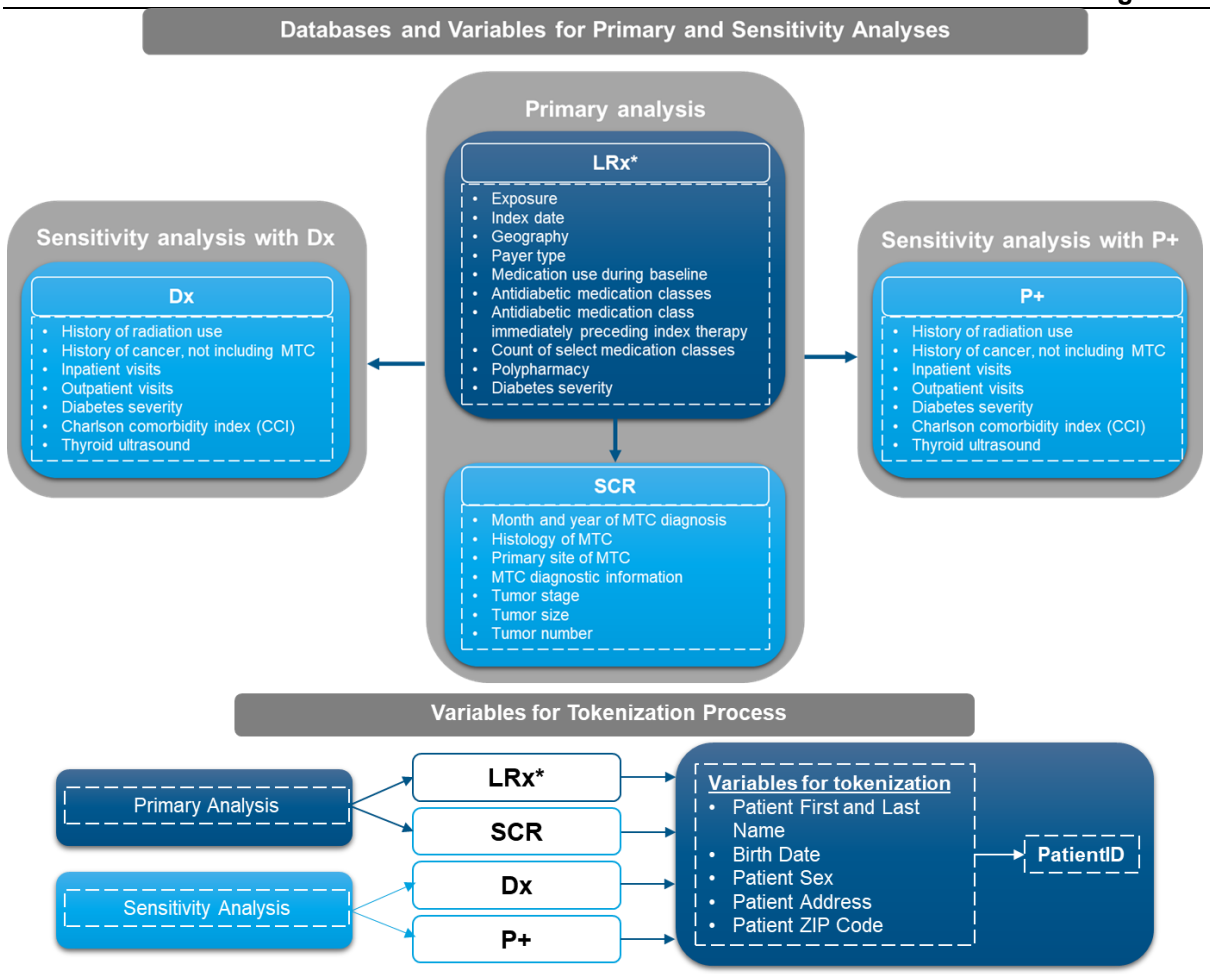
identification and additional cancer attributes as the **primary linkage**. Both an open (Dx) and closed (P+) claims data source for clinical information will be used in **sensitivity analyses** (Figure 1).

Primary linkage:

IQVIA LRx is an open claims database and contains fully adjudicated pharmacy claims by the payers. It is derived from broad-based healthcare sources (practice management systems, clearinghouse, pharmacies, and software vendors).[9] LRx will be used to source the study cohorts and linked to SCR data as the primary linkage for this study to evaluate the IRR of MTC for LA GLP-1 RA – exposed patients compared to the 3 active comparator cohorts. Detailed information on LRx and SCR data can be found below in Sections 5.1.2.1 and 5.1.2.2.

Linkages for Sensitivity analyses:

The LRx-SCR linked data (primary analysis) will be additionally linked to clinical information from an open claims and closed claims source in sensitivity analyses. These linkages will occur independently such that 2 independent files will be created [LRx-SCR-Dx (1) and LRx-SCR-P+ (2)]. The IQVIA open medical claims (Dx) data provides clinical information derived from broad-based healthcare sources (practice management systems, clearinghouses, pharmacies, and software vendors). IQVIA's Pharmetrics Plus (P+) database, provides clinical information from closed claims, and covers the commercially insured US population with 210+ million enrollees. It contains fully adjudicated health plan claims and a complete view of the patient medical history across care settings. Detailed information on Dx and P+ can be found below in Sections 5.1.3.1 and 5.1.3.2.



* LRx will be the anchor database for linkages with SCR, Dx and P+ databases.
Abbreviations: LRx – Longitudinal Prescription database; P+ - PharMetrics® Plus database; SCR – State Cancer registries.

Figure 1. Study database linkages and variables for primary and sensitivity analyses.

5.1.2. Databases Used in Primary Analysis

5.1.2.1. IQVIA Longitudinal Prescription (LRx) Database

Description of LRx

The IQVIA LRx database will be used as the anchor, or primary database from which to source the study population in the primary analysis. The LRx database is an open claims database that contains electronic dispensed and adjudicated prescription records in the US at the anonymized patient-level collected from retail, long-term care (LTC), specialty and mail order pharmacies. Prescriptions in LRx mainly come directly from pharmacies instead of from switch data processors/clearinghouses, a major strength of this data source as the information feeding into

LRx is timely, accurate and not processed by an intermediary.[9,10] LRx data enables researchers to track patient prescription information longitudinally across a breadth of payer types for new medication use, continued medication use and medication switching. All data are Health Insurance Portability and Accountability Act (HIPAA) compliant to protect patient privacy.

LRx Coverage of the US

Through agreements with a variety of data contributors, IQVIA receives more than 4 billion prescription claims per year from retail or mail order prescription claims. This represents dispensed prescriptions for approximately 85% of all US pharmacies, including dispensed prescriptions for approximately 94% of the retail pharmacy channel, 74% of specialty and mail order, and 74% of LTC in the US. The database contains information for over 252 million unique de-identified patients and 1 million physicians. It provides the breadth necessary to measure prescribing behavior at the territory and provider level. The coverage of prescriptions filled in the US population (94% of all retail prescriptions and 85% of all prescriptions in the US) provides a dataset representative of the US.

LRx Claims in Detail

LRx prescription claims are adjudicated fully by the payer (as evidenced by dispensed vs. prescribed prescriptions) and fill rates for data elements are in line with industry standards for pharmacy (National Council for Prescription Drug Programs [NCPDP]) and medical (837-P/I) claims for claims adjudication. These elements include, among many others, service dates, identifiers for providers (National Provider Identifier [NPI], NCPDP), diagnosis (International Classification of Disease [ICD]–9/10), product (national drug code [NDC]), procedure (Current Procedural Terminology [CPT]/Healthcare Common Procedure Coding System [HCPCS], ICD), and payer, that are enhanced with attributes (for example, Provider Specialty) from IQVIA's core reference masters, and tokenized patient demographics that enable longitudinal studies. The LRx database goes through the IQVIA standard quality control processes to ensure that the data transactions are considered final and that they could be used for research purposes (see Section 5.1.4).

Payer types contributing to the LRx database include third party payer, Medicare, Medicaid, and cash payments, including patients without insurance who pay out-of-pocket for their care. Available from April 2001, approximately 95% of claims are available for analyses within 12 days of being dispensed. LRx database covers 94% of all retail pharmacies, that increases the confidence of capturing a representative sample of GLP-1 RA users in the study with low risk of misclassification. In addition, Medicare representation in the LRx database is approximately 31.7% Medicare Part D, 0.3% Medicare, and 3.5% Medicaid as of 2023, while the remaining prescription dispense claims are captured by other payer types (for example, self-pay/cash). Based on the payer contributions assessment over the past 10 years (2014 to 2023), third party payers represent the highest proportion of claims per year (58% to 62%), followed by Medicare (27% to 31%), cash payments (4% to 7%) and Medicaid (4% to 6%).

Based on the most recent LRx data available during analysis (August 2023) with any prescription claims, all patients (100%) contribute data from the last 12 months (1 year), 79% from the last 24 months (2 years), 65% from the last 36 months (3 years), 54% from the last 48 months (4 years), and 48% from the last 60 months (5 years) ensuring a high availability of capture for the study.

LRx Data Used in MTC Linkage Study

Attributes and metrics within the LRx data include payer, payer types, product information, age, sex, 3-digit zip code as well as the prescriptions relevant information including prescriber, date of service, refill number, quantity dispensed, and day supply. All age groups are well-represented in the LRx database, including patients aged 40 to 60 years of age, known to have the most utilization of GLP-1 RA therapies as well as highest incidence of MTC.[3,4,11,12] In order to link LRx with the SCR data for capturing MTC diagnosis, the following LRx data fields will need to be tokenized: first name, last name, date of birth, sex, street address, and ZIP code. The remaining study variables not included in the token will not be tokenized. For a full list of variables, please refer to Section 5.4.

Several published retrospective database studies in various therapeutic areas have utilized LRx database.[13-15]

5.1.2.2. State Cancer Registries (SCRs)

Description of SCRs

Population-based cancer registries record all cancer cases observed in a population, typically at the state level in the US. Cancer registries are designed with the goal to make their cancer data available for epidemiology studies such as the MTC linkage study. SCRs are important tools for capturing cancer outcomes due to the reliability of the information; cases are validated and incorporated into each state's datafile following rigorous validation activities at the end of each calendar year that can take up to 2 years to complete. Healthcare providers are required to report all cancers to their state-wide cancer registry under state-specific laws (for example, for New York, this law is Public Health Law Section 2401). As a result, SCR data are robust and reliable datasets with rich information on state-wide cancer data that can be used for epidemiology studies of all kinds, including association studies and surveillance.

SCR Coverage of the US

Every state in the United States has its own cancer register, as do the US territories. Given the strict reporting rules in each state, it is anticipated that the cancer registers provide representative data for their state. If pooled, the SCR data would be representative of the US. In addition, the North American Association of Central Cancer Registries (NAACCR) annually reviews data from cancer registries to assure the quality, accuracy, and completeness of cancer incidence data based on pre-determined objective and independent state cancer registry certification criteria.[16] For the present study, the SCRs from all 50 US states and the District of Columbia, will be approached for participation. Minor outlaying islands and territories will not be included in this

study due to extremely small populations and the rarity of MTC. With respect to the anticipated capture of MTC, it is expected that all MTC cases within participating SCRs will be captured. MTC cases that fall in non-participating SCRs will not be captured. This study has incorporated adjustments to account for less than full participation of SCRs. Please see Section 6.2.2.2. During years 2010 to 2020, a total of 9,192[4] MTC cases were reported in the US[20]. The goal for this study is to recruit sufficient number of SCRs to cover a minimum of 65% of the US population aged 18 years and older during the MTC linkage study period of 2010 to 2025. Please see Section 6.2.2.2 for more information on the derivation of this threshold.

SCR Data Used in MTC Linkage Study

SCR data include demographic variables for tokenization and linking, MTC diagnosis codes (that is, histology, as coded by International Classification of Diseases for Oncology, 3rd edition codes [ICD-O-3] or most current addition), primary site, tumor size (when available), tumor number, diagnostic confirmation, and month (when available), and year of MTC diagnosis. Due to SCR reporting, data collection, and adjudication processes, there is a 9-to-18-month lag in data availability after the close of the calendar year among SCRs.

Each SCR which agrees to participate will create a data file containing all the MTC cases diagnosed in their state during the study period. The IQVIA tokenization software will be used to de-identify patients and to facilitate linkages. The prepared data file will include demographic variables used for tokenization (that is, first name, last name, date of birth, sex, street address, and ZIP code), and study variables such as MTC diagnosis codes, primary site, diagnostic confirmation, and month (when available), and year of MTC diagnosis. Further details regarding tokenization and data linkage processes are presented in Sections 7.1 and 7.2.

5.1.3. Databases Used in Sensitivity Analyses

5.1.3.1. IQVIA Open Medical Claims (Dx)

Description of Dx

The IQVIA Dx database will be used to extract clinical information for the subset of patients linkable to Dx in the sensitivity analysis. The Dx database is an open claims database that includes anonymized patient level diagnoses, procedures, and in-office treatments (for visits to US office-based professionals), ambulatory and general healthcare site visits. The data are sourced from 837p transaction (raw data format of the electronic version of claims) or Centers for Medicare & Medicaid Services (CMS)–1500 forms (paper-based version of claims), which are the standard reimbursement formats used by healthcare professionals to transmit healthcare

[4] The MTC count of 9,192 includes all US states and does not account for attrition due to linkages or selection based on inclusion/exclusion criteria.

claims for payment. Each transaction or submission of the forms is considered a pre-adjudicated claim[5]. Dx data enables researchers to gain insights into what takes place during a patient's visit with their physician, including the conditions a patient is diagnosed with, the procedures performed, and the drugs administered; all of which are important in understanding patient populations and the reason for initiating or changing treatments.

Dx Coverage of the US

The Dx database is comprised of approximately 1.6 billion professional fee claims (that is, unadjudicated or pre-adjudicated claims) at the anonymized patient-level, representing over 1.2 million providers (approximately 65% of all physicians) and 168 million patients per year. There is representation from approximately 236 physician specialties (for example, American Medical Association [AMA] classifications such as Family Medicine physician, Pediatrician, Radiologist, Urologist, etc.) as well as representation of non-physician practitioners (for example, Nurse Practitioners and Physicians Assistants). Dx captures 94% of AMA providers showcasing the capture of the gold standard association for American medical providers and contains both patient diagnosis and procedure details providing insight into treatment patterns.

Dx Data Used in MTC Linkage Study

Attributes and metrics within the Dx data include diagnoses, medical procedures, health care encounters, and patient demographic information (age, sex, 3-digit zip code, and geography) and payer type. In the Dx database, diagnosis and medical procedures can be identified using ICD-9/10 diagnosis codes and CPT/HCPCS codes, respectively. In order to link Dx with the LRx-SCR linked data for capturing clinical information of the subset of linkable patients, the following Dx data fields will need to be tokenized: first name, last name, date of birth, sex, street address, and ZIP code. The Dx database will be used in this sensitivity analysis to obtain clinical covariates such as history of radiation use or cancer, healthcare utilization (for example, outpatient and inpatient visits), diabetes severity, Charlson comorbidity index (CCI), and thyroid ultrasound procedures (see Table 3 for full list of covariates).

Based on a feasibility assessment involving the LRx and Dx linkage, approximately 8.7 million unique patients treated with any LA GLP-1 RA therapies identified in LRx during the patient selection period (January 2010 and December 2023) were linked to Dx database. This feasibility assessment shows a linkage rate of 65.2% between LRx and Dx databases for the planned sensitivity analysis in this study; however, it is expected that the linkage rate will reduce with the additional study requirements including the application of IQVIA's tokenization/linkage process

[5] Pre-adjudicated claims are submitted for reimbursement on the medical claim. It is unclear how the claim was paid by the insurer (paid in full, amount paid, or denied); however, there is confirmation the claim contains the services that were performed, and for which diagnoses the patient was seen. As with any data asset, there is potential for manual error or coding error, although this is to be expected and very minimal.

and remaining inclusion/exclusion criteria applied for the full data analysis. Several published retrospective database studies in various therapeutic areas have utilized the Dx database linked to the LRx database for research purposes.[13-15] Therefore, the LRx-SCR-Dx linkage sensitivity analysis will provide clinical insights not available in the LRx-SCR primary analysis.

5.1.3.2. IQVIA PharMetrics® Plus (P+) Database

Description of P+

The IQVIA P+ database will be used to extract clinical information for the subset of patients linkable to P+ in the sensitivity analysis. While both Dx and P+ provide clinical information on patients, the P+ database is a closed claims database as opposed to an open claims data source and can provide enrollment information and additional depth into patient medical histories within payers. P+ is comprised of adjudicated claims, and data are available from 2006 onwards, with a typical 3- to 4-month data lag due to claims adjudication by payers. The data included in P+ are anonymized patient-level information on inpatient and outpatient diagnoses and procedures, as well as retail and mail order prescription records. P+ also has detailed information on the pharmacy and medical benefit (copayment, deductible), inpatient stay (admission type and source, discharge status) and provider details (specialty, provider ID). Amounts charged by providers and amounts allowed and paid by health plans are available for all services rendered, as well as dates of services for all claims. Other data elements include demographic variables (age, sex, and geographic region), product type (for example, health maintenance organization [HMO], preferred provider organization [PPO]), payer type (for example, commercial, self-pay), and start and stop dates of health plan enrollment. Like the Dx database, the P+ database also utilizes ICD-9/10 codes to identify medical diagnoses and procedures. Data contributions are subjected to a series of quality checks by IQVIA to ensure a standardized format and to minimize error rates. All data are HIPAA compliant to protect patient privacy. P+ enables researchers to gain insights into medication use, every healthcare facility interaction billed through insurance, and true cost data including coinsurance, copayment, and deductible amounts. These data are important to understand the patient's medical history as they visit different doctors, hospitals, and pharmacies.

P+ Coverage of the US

P+ data has adjudicated claims for more than 210 million unique patients across the US. The data have a diverse representation of geography, employers, payers, providers, and therapy areas and are representative of the commercially insured population under 65 years of age. Although only select payer types are included in the database, the data itself is comprehensive for patients enrolled and captures in-depth information for all covered healthcare interactions. The data are also longitudinal, with 3 or more years of continuous enrollment. Patients in each 3-digit zip code and every Metropolitan Statistical Area of the US are included, with coverage of data from 90% of US hospitals, and 80% of all US doctors. This allows for more granular patient segmentation and comparisons by geography. The data do not include quality of life data, race, ethnicity, or immigration status.

P+ Data Used in the MTC Linkage Study

Attributes and metrics within the P+ data include demographic variables (age, sex, and geographic region), payer product type (for example, HMO, PPO), payer type (for example, commercial, self-pay), and start and stop dates of health plan enrollment. To link P+ with the LRx-SCR data, the following P+ fields will need to be tokenized: patient first name, last name, date of birth, sex, street address, and ZIP code. The P+ database will be used in this sensitivity analysis to obtain clinical covariates including history of radiation use or cancer, healthcare utilization (for example, outpatient and inpatient visits), diabetes severity, CCI and thyroid ultrasound procedure (see Table 3 for full list of covariates). While the linkage rate between LRx-SCR-P+ is expected to be much lower than the LRx-SCR-Dx linkage sensitivity analysis, the linkage to P+ will provide valuable insights using a closed claims database and will validate the primary data analysis.

Each of these databases and their corresponding contributing variables are identified in Table 2.

Table 2. Study Databases and Their Corresponding Contributing Variables

Data source	Variable
***Required variables across databases for Tokenization**	Patient First and Last Name
	Birth Date[6]
	Sex
	Patient Address Line 1
	Patient ZIP Code
Longitudinal Prescription (LRx) database	Required tokenization variables
	Exposure (duration of use, number of dispensings)
	Index/Treatment date
	Geography
	Payer type
	Medication use
	ADM classes during baseline period
	ADM class immediately preceding index therapy
	Count of medication classes during baseline period
	Polypharmacy
	Diabetes severity (assessed using proxy)
State Cancer Registries (SCRs)	Required tokenization variables
	Month and year of MTC diagnosis
	Histology of MTC
	Primary site of MTC
	MTC Diagnostic confirmation
	Tumor stage
	Tumor size
	Tumor sequence number
Medical Claims (Dx) database	Required tokenization variables
	History of radiation use
	History of cancer, not including MTC
	Inpatient visits
	Outpatient visits
	Diabetes severity
	CCI[17], modified to exclude diabetes, prior to the index date
	Thyroid ultrasound
PharMetrics® Plus (P+) Database	Required tokenization variables
	History of radiation use
	History of cancer, not including MTC
	Inpatient visits
	Outpatient visits
	Diabetes severity
	CCI[17], modified to exclude diabetes, prior to the index date
	Thyroid ultrasound

5.1.4. Open Claims Database (LRx and Dx) Quality Assurance (QA)/Quality Control (QC) Process

The QA/QC process for the open claims at IQVIA involves optimizing overall data quality, reducing data duplication, and ensuring adherence to IQVIA policies and procedures. IQVIA collects pharmacy claims, and medical claims from various entities in the healthcare ecosystem. IQVIA uses standard data layouts to ensure key data elements are captured to enable a wide range of analytics. During data onboarding process and as part of the ongoing production process, quality checks are performed on the data to ensure that data is populated correctly, and fields are valid including industry standard fields like Provider/Pharmacy NPI, NDC, Diagnosis Codes, and Procedure Codes. IQVIA study teams will apply methodology to deduplicate claims as part of their analysis using best practice approach and knowledge of key fields in the data. After linkages, the demographic information will be pulled from the index claim, which will be from LRx data. In addition, QC is performed to confirm that data trends are aligned with provider and market expectations. The claims data is bridged to and supplemented with industry standard references (for example, ICD-10 diagnosis, CPT codes, etc.). The data are stored in a central database and are provided to clients using various analytic tools, data formats and data aggregations/segmentations. The claims data are processed through the IQVIA tokenization/de-identification engine prior to IQVIA receipt. This process enables IQVIA to leverage the tokens to assign a longitudinal patient ID on the claims. The longitudinal patient ID enables privacy compliant linking to across IQVIA's LRx and Dx databases (additional details on IQVIA's Tokenization and Encryption process is detailed in Section 7.1).

5.2. Study Population

All eligible patients must fulfill the following requirements in LRx database:

- ≥18 years or older during the year of index (date of qualifying medication)
- In addition to index medication, ≥1 dispensed medication (any class) within 24 months prior to the index date[7]
- ≥1 dispensed medication (any class) after the index date and within 12 months of index date (12-month post-index period)[8]
- No evidence of cohort qualifying medication during 12-month before the index date (12-month lookback period)[9]

[6] Birth date will be used to calculate age in years at the index date and will be reported as both continuous and categorical measure (age groups).
[7] Date of qualifying medication dispensing.
[8] Although unlikely, there is the potential for a patient to have a medication dispensation 12 months post-index and immortal time between the index date and that medication dispensing.
[9] For each study cohort, a lookback period of 12 months will ensure that patients indexed on ADMs or AOMs do not have the cohort qualifying medication during 12 months prior to index date.

- Reside in the US and the District of Columbia (DC) during the study period, and
- No missing values for year of birth.

Patient eligibility for inclusion by study cohorts, in addition to the above criteria, are captured below.

LA GLP-1 RA therapies exposed cohort:

- ≥1 dispensed prescription for a LA GLP-1 RA therapy during the study patient selection period.

T2D Active Comparator 1 Cohort:

- ≥1 dispensed prescription for any sodium-glucose transport protein 2 (SGLT2) or dipeptidyl peptidase IV (DPP-4) inhibitors during the study patient selection period.

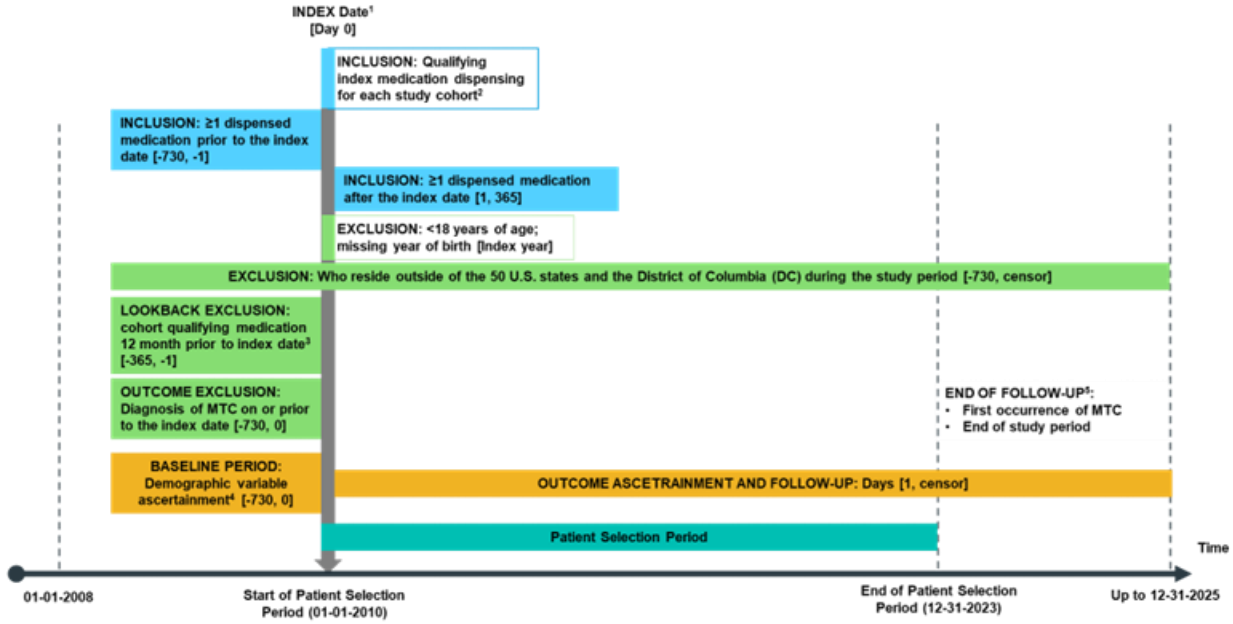
T2D Active Comparator 2 Cohort:

- ≥1 dispensed prescription for any ADM[10], other than LA GLP-1 RA therapies, during the study patient selection period.

Overweight/obesity Active Comparator Cohort:

- ≥1 dispensed prescription for any AOM[11], other than LA GLP-1 RA therapies, during the study patient selection period.

[10] See Section 5.3.2 for list of anti-diabetic medications.
[11] See section 5.3.2 for list of anti-obesity medications



[1]Earliest possible index date in the study would be 01-01-2010
[2]Qualifying index medication dispensing by study cohorts:
- LA GLP-1 RA therapies cohort - LA GLP-1 RA therapy
- SGLT2 or DPP-4 inhibitors cohort - SGLT2 or DPP-4 inhibitors
- Other ADM cohort - any antidiabetic medication, other than LA GLP-1 RA therapies
- Other AOM cohort – any anti-obesity medication, other than LA GLP-1 RA therapies

[3]No evidence of cohort qualifying medication during 12-month before the index date
[4]≥1 dispensed medication (any class) within 24 months prior to the index date.
[5]In a sensitivity analysis using the subset of data linked to Dx (medical claims) and PharMetrics Plus, end of follow-up will either be a censoring event defined as evidence of end-of-life care (proxies) and/or death, the end of the study period, or MTC diagnosis, whichever comes first

Figure 2. Study period schematic (2008 to 2025).

5.2.1. LA GLP-1 RA Exposed Cohorts

Patients will be initially eligible for inclusion in the LA GLP-1 RA exposed cohort if they have at least 1 dispensed prescription for a LA GLP-1 RA therapy during the study patient selection period. The earliest evidence of qualifying medication for this cohort will be identified as index medication and the date of the first dispensed qualifying medication will be defined as the index date. The specific exposures of interest for the LA GLP-1 RA therapies cohort include once-weekly exenatide (Bydureon®, Bydureon Bcise®), liraglutide (Victoza®, Saxenda®), albiglutide (Tanzeum®), dulaglutide (Trulicity®), injectable semaglutide (Ozempic®, Wegovy®), oral semaglutide (Rybelsus®), and tirzepatide (Mounjaro®, Zepbound®). LA GLP-1 RA therapies in the above list will be included in the study for all companies that produce such products. Any newly approved drugs during the study period will be captured in the SAP and final study report as necessary.

5.2.2. Active Comparator Cohorts

Patients treated using LA GLP-1 RA therapies (exposed) will be compared to three active comparator cohorts: T2D active comparator 1, T2D active comparator 2 and overweight/obesity active comparator cohorts. The three comparator cohorts will be defined as follows:

T2D Active Comparator 1 Cohort: Patients will be included in the T2D active comparator 1 cohort if they fill at least one dispensed prescription for SGLT2 or DPP-4 inhibitors during the study patient selection period with no evidence of SGLT2 or DPP-4 inhibitors and without LA GLP-1 RA use during the prior 12-month lookback period. The earliest evidence of qualifying medication for this cohort will be identified as index medication and its date of first dispensing will be defined as the index date. Medications of interest may include, but are not limited to bexagliflozin (Brenzavvy™), canagliflozin (Invokana®), dapagliflozin (Farxiga®), empagliflozin (Jardiance®), ertugliflozin (Steglatro®), sitagliptin (Januvia®), saxagliptin (Onglyza®), linagliptin (Tradjenta®), alogliptin (Nesina®). SGLT2 and DPP-4 inhibitors in the above list will be included in the study for all companies that produce such products. Any newly approved drugs during the study period will be captured in the SAP and final study report as necessary.

T2D Active Comparator 2 Cohort: Patients will be included in the T2D active comparator 2 cohort if they fill at least 1 dispensed prescription for any ADMs (including SGLT2 or DPP-4 inhibitors), other than LA GLP-1 RA therapies during the study patient selection period with no evidence of the cohort qualifying medication during the prior 12-month lookback period. The earliest evidence of the qualifying medication for this cohort will be identified as index medication and its date of first dispensing will be defined as the index date. The medications of interest may include, but not limited to SGLT2 inhibitors, DPP-4 inhibitors, meglitinides, thiazolidinediones, insulin, sulfonylureas, or combination of oral ADMs.

Overweight/Obesity Active Comparator Cohort: Patients will be included in the overweight/obesity active comparator cohort if they fill at least 1 dispensed prescription for anti-obesity medications, other than LA GLP-1 RA therapies during the study patient selection period with no evidence of the cohort qualifying medication during the prior 12-month lookback period. The earliest evidence of the qualifying medication for this cohort will be identified as index medication and its date of first dispensing will be defined as the index date. The medications of interest may include, but not limited to orlistat (Xenical®), phentermine-topiramate (Qysmia®), phentermine (Lomaira®, Apidex-P®), naltrexone-bupropion (Contrave®).

5.2.2.1. Additional Requirement for the Study Cohorts

Patients in either the exposed or comparator cohorts will be identified using NDCs in the LRx database for dispensed prescriptions. This design improves the accuracy of identifying drug exposure by eliminating recall bias. Since new medications may be approved during the study period, the final list of medications for exposures of interest will be included in the SAP and final study report.

An observational parallel to the intent-to- treat approach will be used for the primary analysis where patients will be sampled without replacement for the exposed and comparator cohorts. The exposed patients will be identified first from the LRx database, and the comparator cohorts will be identified as the next step without replacement. As a result, the patients included in the comparator cohorts will not have a documented GLP-1 RA exposure during the study patient selection period. Once a patient indexes on LA GLP-1 RA therapies, they are considered

exposed for the duration of the study follow-up period. Similarly, once a patient is selected as a comparator (selection will occur after the exposed patients are removed from the sampling pool), they are considered unexposed to LA GLP 1 RAs for the duration of the study period. The process for selecting comparators is restarted for the three comparator groups starting with the same pool of eligible patients available once the exposed cohort has been removed. After cohort development and linkage to SCRs, to assess temporality, any patient with an MTC diagnosis on or prior to their index date (for LA GLP-1 RA therapy or comparator drug [for example, SGLT2 or DPP-4 or other ADM or AOM]) will be removed. This step will occur post SCR linkage (using SCR MTC diagnosis date information).[12]

Although it is expected that the SCR will contain MTC diagnosis date information, including month and year, the extent of missing information in the data pertinent to this study is currently undetermined. However, as a data source, cancer registries are highly accurate and complete in part because healthcare providers are required to report all cancers to their state-wide cancer registry and are reviewed by NAACCR (see Section 5.1.2.2 for more details). The SAP will include detailed plan to address any missing month information. For instance, if month of diagnosis is missing from date information, the mid-year date may be used, see Section 6.2.4 and the SAP for details.

5.3. Time Periods

The study patient selection period will begin from 01 January 2010 and end 31 December 2023, with a 24-month baseline period beginning as early as 01 January 2008 (Figure 2). Identification of MTC cases will continue into 2025 and potentially up to 31 December 2025 (depending on SCR data lags). The index date will be the date of the earliest identified dispensed prescription for either the LA GLP-1 RA therapy dispensing, SGLT2 or DPP-4 inhibitor dispensing, other ADM dispensings, or other AOM dispensings.

This study will include linkages between LRx and SCR data for the primary analysis and a linkage to the Dx and P+ databases for 2 separate sensitivity analyses that will add clinical covariate information. A linkage is expected to occur in Q4 2025 and will include data from 01 January 2008 up to 31 December 2025 (depending on SCR data lags). The linkage dates are contingent on SCR involvement and participation.

The observation period, or the time at risk, will be defined as the time post-entry into the study cohort. Among patients with the MTC outcome, the observation period will begin the day after the index date and will end at the date of MTC diagnosis rather than the linkage date. In the

[12] Exclusion criteria will be further detailed in the study SAP. For example, details regarding handling of patients indexing on a LA GLP-1 RA therapy in the same year as their MTC diagnosis who are also missing month of MTC diagnosis will be provided.

sensitivity analyses with linkages to Dx and P+, a published claims-based mortality algorithm will be used to estimate patient mortality as an additional censoring criterion in this study.[18] Further details on the application of the mortality algorithm can be found in Section 6.2.2.6.4 and in the SAP.

5.4. Variables

Table 3 outlines the variables included in the study, the assessment window, description, and source. The detailed operational definitions of study variables will be described in the study SAP.

Table 3. Example Study Variables

Variable	Assessment Window	Description	Source
Baseline			
Patient ID	Index month/year	Unique patient identifier from tokenization process	LRx
Index date	Index month/year	Patients with index year will be reported starting from 2010 to 2023	LRx
Age	Index month/year	Age at index date Age 18 to 44, 45 to 64, 65+ Age groups may further be informed by the data	LRx
Sex	Index month/year	Female Male Other Missing	LRx
Geography	Index month/year	3-digit zip code State (including District of Columbia) US regions: Northeast, Midwest, South, West	LRx
Payer type for index prescription dispense[13]	Index month/year	Third party Medicare Medicaid Self-Pay/Cash Other	LRx
Medication use during the 24-month baseline period	24-month baseline period	NDC codes mapped to generic name Prior antidiabetic therapy, AOMs; anti-hypertensive medication use; statin use; antidepressant use (excluding bupropion); opioids use; treatment with therapies associated with carcinoma (for example, oral chemotherapy) This variable will be used as a proxy for health status	LRx

[13] For the comparator patients, defined as most frequent payer type during their Index Month

Variable	Assessment Window	Description	Source
ADM classes during baseline period	24-month baseline period	NDC codes will be used to identify below ADM classes: sulfonylureas, meglitinides, metformin (a biguanide), thiazolidinediones, alpha glucosidase inhibitors, DPP-4 inhibitors, bile acid sequestrants, dopamine agonists, SGLT2 inhibitors and GLP-1 RA therapies	LRx
ADM class immediately preceding index therapy	24-month baseline period	NDC codes will be used to identify below ADM classes: sulfonylureas, meglitinides, metformin (a biguanide), thiazolidinediones, alpha glucosidase inhibitors, DPP-4 inhibitors, bile acid sequestrants, dopamine agonists, SGLT2 inhibitors and GLP-1 RA therapies	LRx
Count of medication classes	24-month baseline period	Count of medication classes, described under medication use variable, during the 24 months prior to the index date	LRx
Polypharmacy index	24-month baseline period	Polypharmacy will be defined as patients with five or more medication use during baseline period[19]	LRx
History of radiation use	24-month baseline period	CPT and HCPCS codes prior to the index date Identified using CPT/HCPCS codes: 77371-3, 77401-9, 77411-4, 77416, 77418, 77422, 77423, 77432, 77470, 77750, 77761-3, 7776-8, 77781-4, 77789	Dx, P+
History of cancer	24-month baseline period	Identified using ICD-9-CM codes 140.xx-209.xx or ICD-10-CM codes C00-C97 prior to the index date	Dx, P+
Number of inpatient and outpatient visits prior to the index date	24-month baseline period	Inpatient and outpatient medical claims representing visits on unique days prior to the index date	Dx, P+
Diabetes severity	24-month baseline period	Time since the first antidiabetic therapy Time since first observed diagnosis of diabetes Prior diagnosis for diabetic complications (for example, retinopathy, nephropathy, neuropathy, etc.) These variables will be used as proxies for disease severity	LRx, Dx, P+
Thyroid ultrasound	Follow-up period	CPT codes will be used This variable will be used as a proxy for thyroid cancer/disorder screening	Dx, P+
CCI, modified to exclude diabetes, prior to the index date	24-month baseline period	Inpatient and outpatient medical claims. This variable will be used as a proxy for 1-year mortality risk	Dx

Variable	Assessment Window	Description	Source
Exposure			
Index Medication	Index month/year	NDC codes mapped to generic name Dispensing dates Days' supply for each dispensing Quantity dispensed Dosage form and strength Duration of use (≤1, 1 to 3, >3 years)[20] Number of prescription dispensings	LRx
Cohort	Index month/year	LA GLP-1 RA therapy-exposed cohort: • Adult patients with dispensed prescription of LA GLP-1 RA therapy during the study patient selection period (first observed will define the index date) T2D Active Comparator 1 Cohort: • Adult patients with a first dispensed prescription for SGLT2 or DPP-4 inhibitors T2D Active Comparator 2 Cohort: • Adult patients with a first dispensed prescription for ADMs other than LA GLP-1 RA therapy Overweight/Obesity Active Comparator Cohort: • Adult patients with a first dispensed prescription for AOMs other than LA GLP-1 RA therapy	LRx
Outcome			
MTC diagnosis (primary outcome)	Study follow-up from index date to linkage in December 2025[14]	ICD-O-3 diagnosis codes: 8345 (medullary carcinoma with amyloid stroma), 8346 (mixed medullary-follicular carcinoma), 8347 (mixed medullary-papillary carcinoma), 8510 (medullary carcinoma, not otherwise specified), 8512 (medullary carcinoma with lymphoid stroma), 8513 (atypical medullary carcinoma) Date of diagnosis (month when available and year) MTC diagnostic confirmation Tumor stage, size, primary site, and tumor sequence number (when available) MedDRA code for MTC: C0238462[15]	SCR

[14] Due to lag, data will be available through 2023.
[15] The current study is planned to use US data sources, which do not use MedDRA codes, but it is added here for comprehensiveness.

6. STATISTICAL METHODS

6.1. Sample Size

Based on data from a feasibility assessment, it is projected that there will be approximately 13.4 million total unique patients treated with LA GLP-1 RA therapy products between 2010 and 2023. An attrition table will be reported for each cohort comparison and time period showing the number of patients remaining after each inclusion criterion. The sample sizes are expected to reduce with application of additional selection criteria. Using the LRx and SCR data, assuming approximately 13.4 million GLP-1 RA users between 2010 and 2023, up to 4 times the sample size of the exposed cohort (up to 53.6 million comparator patients in each comparator group), α=0.05, two-sided test, and a background MTC IR of 0.2 events per 100,000 person-years, the study would achieve 80% power to detect a HR of 1.2. If the number of LA GLP-1-RA users is higher (that is, >13.4 million), the smaller of an effect becomes detectable (HR <1.2). If the number of LA GLP-1-RA users are smaller than 13.4 million, then the larger HR is detectable (>1.2). It is expected that the T2D active comparator cohorts 1 and 2 will reach the highest number of comparators. However, given the smaller pool of anti-obesity medications, a comparator group of up to 4 may not be achieved. The minimum detectable effect for up to 3 and up to 2 active comparators in the anti-obesity comparator group are HR = 1.26 and 1.28 (40.2 and 26.8 million patients), respectively, assuming α=0.05, two-sided test, and a background MTC IR of 0.2 events per 100,000 person-years. Inclusion of comparators up to 4 times the sample size of the exposed cohort could help increase statistical power of the study and reduce bias in estimates of treatment effects. Please note that the HR is being used to report effect sizes as a proxy for IR due to the rare outcome of MTC. Study power will be addressed thoroughly in the SAP, particularly with regard to potential attrition following linkage. The following counts did not include all relevant exclusion criteria that would be applied for this study (Table 4).

Table 4. LA GLP-1 RA Therapy Feasibility Counts in the LRx Database: January 2010 to December 2023

Years	Any GLP-1 RA	Adlyxin	Bydureon	Byetta	Mounjaro	Ozempic	Rybelsus	Saxenda	Soliqua 100/33	Tanzeum	Trulicity	Victoza	Wegovy	Xultophy 100/3.6
2010	193,258	0	0	111,362	0	0	0	0	0	0	0	90,915	0	0
2011	338,524	0	0	144,049	0	0	0	0	0	0	0	208,203	0	0
2012	478,826	0	54,211	141,819	0	0	0	0	0	0	0	313,477	0	0
2013	589,335	0	102,213	99,885	0	0	0	0	0	0	0	415,805	0	0
2014	661,094	0	145,881	78,157	0	0	0	0	0	9,079	2,582	459,352	0	0
2015	881,190	0	184,654	66,786	0	0	0	20,085	0	68,167	85,338	525,075	0	0
2016	1,133,637	0	183,375	49,051	0	0	0	47,133	0	97,425	257,386	584,104	0	0
2017	1,435,505	129	190,798	41,382	0	0	0	67,422	22,794	55,733	499,965	653,923	0	7,417
2018	1,854,974	324	234,464	31,288	0	115,274	0	87,566	45,158	18,219	737,005	717,045	0	22,537
2019	2,344,192	553	212,018	23,116	0	476,391	10,213	104,374	57,826	676	966,133	647,926	0	28,658
2020	2,866,752	372	183,857	16,937	0	794,977	148,491	109,952	63,997	3	1,148,886	555,149	0	25,723
2021	3,935,114	488	138,074	12,498	0	1,336,382	320,178	134,108	67,439	0	1,541,102	512,885	115,525	21,454
2022	6,151,159	357	88,200	8,709	813,268	2,457,106	503,505	212,336	72,599	0	1,983,523	438,237	199,331	17,612
2023	9,115,963	28	65,135	7,141	1,832,494	4,071,421	673,186	219,901	79,296	0	2,109,988	424,795	924,714	17,154
2010-2023	**13,423,211**	**1,504**	**867,639**	**422,134**	**2,046,225**	**5,451,639**	**1,092,966**	**633,526**	**187,042**	**150,090**	**3,997,470**	**2,447,258**	**1,064,902**	**58,202**

The GLP-1 users identified in LRx database during each year (2010-2023) are presented in rows and the number of patients at the bottom of the table represent total unique adult patients with at least one LA GLP-1 RA use, 24-months baseline period, up to 12-month post-index period and no missing region information between January 2010 and December 2023.

To understand the sample size further regarding LRx and Dx linkage, an early feasibility assessment involving a LRx and Dx linkage was performed. Approximately 8.7 million unique patients treated with any LA GLP-1 RA therapy identified in LRx during the patient selection period were linked to Dx database (Table 5). Overall, approximately 8.7 million patients out of the 13.4 million estimated total exposed patients were linkable to Dx with a 24-month baseline period and up to 12 months follow-up period. This feasibility assessment shows a linkage rate of 65.2% between LRx and Dx databases for the planned sensitivity analysis in this study. The estimated linkage of 65.2% may reduce further with any remaining selection criteria and the additional linkage to SCRs with application of IQVIA's tokenization/linkage process at the time of analysis to ensure robust linkages.

Table 5. LA GLP-1-RA Therapy Feasibility Counts in the LRx and Dx Databases: January 2010 to December 2023

Years	Any GLP-1 RA	Adlyxin	Bydureon	Byetta	Mounjaro	Ozempic	Rybelsus	Saxenda	Soliqua 100/33	Tanzeum	Trulicity	Victoza	Wegovy	Xultophy 100/3.6
2010	130,647	0	0	74,874	0	0	0	0	0	0	0	62,227	0	0
2011	229,214	0	0	96,376	0	0	0	0	0	0	0	142,486	0	0
2012	329,239	0	38,247	95,581	0	0	0	0	0	0	0	217,146	0	0
2013	416,133	0	73,190	68,672	0	0	0	0	0	0	0	294,809	0	0
2014	472,252	0	105,387	54,216	0	0	0	0	0	6,557	1,965	328,872	0	0
2015	634,259	0	134,880	46,688	0	0	0	14,325	0	46,931	62,423	379,460	0	0
2016	814,181	0	133,508	34,150	0	0	0	33,252	0	66,809	188,196	421,022	0	0
2017	1,031,437	96	137,712	28,356	0	0	0	47,114	16,032	39,118	362,706	471,286	0	5,269
2018	1,329,540	236	168,203	21,272	0	81,592	0	60,522	32,085	13,141	531,552	517,443	0	16,074
2019	1,665,698	403	151,175	15,283	0	336,806	6,893	71,132	41,197	483	690,908	464,552	0	20,468
2020	2,043,376	261	131,703	11,043	0	566,908	101,981	75,643	46,080	1	824,390	399,131	0	18,440
2021	2,785,841	366	99,044	8,344	0	947,647	223,745	91,888	48,685	0	1,101,129	361,774	80,329	15,360
2022	4,272,486	264	63,833	5,839	536,865	1,717,516	349,981	143,655	52,716	0	1,410,353	303,620	135,085	12,659
2023	5,923,064	20	46,310	4,641	1,174,104	2,675,967	439,637	138,360	57,046	0	1,445,970	271,921	568,133	12,275
2010-2023	**8,748,105**	**1,095**	**614,855**	**284,542**	**1,309,184**	**3,603,141**	**727,450**	**416,951**	**132,398**	**103,533**	**2,748,095**	**1,674,950**	**662,246**	**40,780**

The GLP-1 users identified in LRx and Dx databases during each year (2010 to 2023) are presented in rows and the number of patients at the bottom of the table represent total unique adult patients with at least one LA GLP-1 RA use, 24-months baseline period, up to 12-month post-index period and no missing region information in LRx and linkable to Dx data with baseline and post-index requirement between January 2010 and December 2023.

To understand the sample size further for LRx and P+ linkage, a feasibility assessment involving a LRx and P+ linkage was performed with the most recent available data for P+ database (up to November 2023). Approximately 1 million unique patients treated with any LA GLP-1 RA therapy identified in LRx during the available patient selection period (January 2010 to November 2023) were linked to P+ database (Table 6). Overall, approximately 1 million patients out of the 13.2 million estimated total exposed patients were linkable to P+ with continuous medical enrollment for the index month and 24-month baseline period. This feasibility assessment with nearly 1 million patients shows a linkage rate of 7.3% between LRx and P+ databases for the planned sensitivity analysis in this study. The estimated linkage of 7.3% may reduce further with any remaining selection criteria and the additional linkage to SCRs. Given the lower coverage of P+ in the US with a focus on commercial insurance, the linkage rate with LRx is low.

Table 6. LA GLP-1 RA Therapy Feasibility Counts in the LRx and P+ Databases: January 2010 to November 2023

Years	Any GLP-1 RA	Adlyxin	Bydureon	Byetta	Mounjaro	Ozempic	Rybelsus	Saxenda	Soliqua 100/33	Tanzeum	Trulicity	Victoza	Wegovy	Xultophy 100/3.6
2010	6,455	0	0	3,712	0	0	0	0	0	0	0	3,092	0	0
2011	10,715	0	0	4,194	0	0	0	0	0	0	0	6,938	0	0
2012	13,915	0	1,690	3,811	0	0	0	0	0	0	0	9,475	0	0
2013	16,412	0	2,992	2,375	0	0	0	0	0	0	0	11,801	0	0
2014	16,846	0	3,343	1,671	0	0	0	0	0	154	73	12,292	0	0
2015	19,692	0	3,851	1,306	0	0	0	229	0	1,147	1,688	12,879	0	0
2016	24,049	0	3,923	1,061	0	0	0	514	0	1,675	4,911	13,830	0	0
2017	54,313	6	6,368	1,138	0	0	0	2,972	1,214	1,282	19,297	25,006	0	305
2018	83,050	8	9,342	938	0	7,210	0	4,729	2,181	368	33,075	30,579	0	948
2019	115,640	19	8,450	662	0	29,016	1,196	5,991	2,541	11	47,087	27,966	0	1,228
2020	182,172	13	7,953	508	0	58,378	18,743	10,062	3,187	0	67,095	25,219	0	1,302
2021	288,003	19	5,985	386	0	110,847	32,332	14,745	3,548	0	97,502	23,095	16,534	1,211
2022	521,442	19	3,935	296	109,843	217,894	46,862	25,439	3,842	0	127,308	19,154	26,632	995
2023	733,505	0	2,573	171	194,216	326,548	53,949	24,146	4,208	0	125,752	17,219	98,916	938
2010-2023	**960,586**	**54**	**28,945**	**10,790**	**220,378**	**433,958**	**94,846**	**58,199**	**9,326**	**2,775**	**223,393**	**87,917**	**118,054**	**2,789**

The GLP-1 users identified in LRx and P+ databases during each year (2010-2023) are presented in rows and the number of patients at the bottom of the table represent total unique adult patients with at least one LA GLP-1 RA use, 24-months baseline period, up to 12-month post-index period and no missing region information in LRx and linkable to P+ data with baseline and post-index requirement between January 2010 and November 2023.

In a preliminary feasibility assessment, using data from January 2010 to September 2023, 12.9 million adult patients (≥18 years) initiated LA GLP-1 RA therapies in the LRx database and 34% (4.4 million) of these patients are linkable to patients available in the closed claims P+ database. After incorporating additional patient selection criteria, the proportion of patients included in the LRx-SCR-P+ linkage sensitivity analysis could be reduced as much as approximately 25%. Although this sensitivity analysis including P+ (LRx-SCR-P+) will be conducted on a subset of patients included in the primary analysis (LRx-SCR) linkage, the resulting closed claims-linked (P+) patient sample will include elements not available in open claims, such as enrollment information and a comprehensive view of patient activity as it pertains to commercial claims. Characteristics of the patient sample in the closed claims sensitivity analysis (LRx-SCR P+) can be compared to a synonymous patient sample in the open claims sensitivity analysis (LRx-SCR-Dx), providing support for determination of overall completeness of patient information in open claims. See Section 6.2.2.6.5 of this protocol for additional detail.

6.2. Data Analyses

6.2.1. General Considerations

All data analyses 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 summarized in tables and figures in Microsoft® Excel format. Counts of less than five will be reported as "N<5" within study results.

The analysis plan will be fully described in a written and approved SAP.

6.2.2. Planned Analyses (Primary Objective Analysis)

6.2.2.1. Descriptive Analyses

Descriptive statistics will be calculated for all variables listed in Table 3, unless otherwise specified, for the LA GLP-1 RA therapy-exposed cohorts and the 3 comparator cohorts (that is, T2D active comparator 1, T2D active comparator 2 and overweight/obesity active comparator cohort). In addition, as an exploratory analysis, descriptive statistics will be calculated for each LA GLP-1 RA therapy separately. Categorical variables will be reported using frequency distributions. Ordinal and continuous variables will be reported using means, standard deviations, medians, minimums, maximums, 25th percentiles, and 75th percentiles, unless otherwise specified. Baseline characteristics will be reported in total and stratified by reporting age categories (if the sample size allows). In addition, an attrition table will be provided showing how patients qualified for each analysis and the number of patients remaining after each inclusion criteria are applied. Descriptive analyses will be stratified by specific medication where applicable.

6.2.2.2. Adjustments for SCR Participation

All US SCRs will be invited to participate; however, it is anticipated that not all state registries will participate. As a result, the participating SCRs will cover a percentage of the US population aged ≥18 years during the observation period. Based on learnings from a previous study,[25] it is assumed that participating cancer registries will cover approximately 65% of the US population aged ≥18 years. Two analytical methods will be used to account for the proportion of the US population covered by participating SCRs: 1) calculation of IRs and IRRs using a cohort restricted to patients in states with participating registries; 2) using a coverage fraction that represents the percentage of MTC cases captured in this study (based on participating registries) divided by the total number of MTC cases expected.

Two sets of study cohorts will result from adjustment for SCR participation described above in this section. Cohorts will have the same selection criteria under each SCR approach as described in Section 5.2. Both sets of study cohorts will undergo PS weighting as described in Section 6.2.2.3.

6.2.2.3. Propensity Score Methods

This is a linkage database cohort design to compare the incidence of MTC among a LA GLP-1 RA therapy-exposed cohort to active comparator cohorts (that is, T2D active comparator 1, T2D active comparator 2 and overweight/obesity active comparator cohort). Propensity Score (PS) weighting in conjunction with model adjustment for residual confounding will account for measured risk factors and potential confounders and will ultimately balance the groups with respect to baseline covariates.[16] Additionally, the use of the PS weights will prioritize retaining all selected patients, rather than only those in a matched set (that is, when using direct matching methods). Inverse Probability of Treatment Weighting (IPTW) methods have been shown to be less restrictive than matching. During analysis, we will explore stabilized weights and/or trimming methods for extreme scores to reduce variance, and this topic will be further addressed in the study SAP.

Details of the statistical methods for PS estimation and weighting will be described in the SAP. To control for confounding factors, PS will be constructed from baseline covariates for comparison of the LA GLP-1-RA therapy cohort with each of three active comparator cohorts (propensity scores will be calculated for each cohort and respective weights applied). The PS is the estimated probability of receiving a study drug dispensing (that is, LA GLP-1 RA therapy), conditional on a set of observed covariates. PS methods will be applied to both sets of study cohort comparisons (a total of 6 cohort comparisons for 2 sets of cohorts post SCR adjustments) resulting from 2 methods of adjustment for SCR participation (see

[16] Additional model adjustments will be performed if the PS does not effectively balance a given covariate. This will be determined in the analysis phase and details will be available in study SAP.

Section 6.2.2.2). Propensity scores will be estimated with covariate information collected during the baseline period for each group and each comparison using logistic regression models. Covariates selected will be variables that are related to the outcome, based on literature and clinical judgment as well as sufficiently prevalent in the population as to not bias the associations under study.[21] Baseline variables to be considered in the model for computing PSs will include age group (5-year age categories), sex, geography (three digit ZIP code), payer type (commercial plans, Medicare, Medicaid, other third parties, and self-pay/cash payments), index year, and count of select unique dispensed prescriptions grouped by therapeutic class (Table 3). Patients will be grouped in 5-year age categories up to 80 years of age, where all patients aged 80 and older will be combined. In each of the sensitivity analyses using Dx and P+, additional variables (for example, comorbid conditions, radiation exposures, etc.), will be considered, if feasible, and described in detail in the SAP and/or final study report as applicable. If it is not feasible to include all baseline variables in PS weighting, either due to missing information in variables or PS model performance, additional confounder adjustment may be performed post-PS weighting in the regression models. Further details on confounder adjustment using PS weighting and model adjustment will be detailed in the SAP.

For each pair-wise comparison of LA GLP-1 RA therapies exposed cohort and active comparator cohorts, a separate PS model will be fitted and the contributing subjects will be assigned induvial PS accordingly.[22,23] Each patient in the exposed and active comparator groups will be weighted individually by the inverse probability of receiving their actual treatment which will provide the average of the individual treatment effects of the study population (that is, average treatment effect [ATE]). LA GLP-1 RA therapy-exposed patients with a lower probability of exposure and T2D active comparator 1, T2D active comparator 2 and overweight/obesity active comparator cohort patients with a higher probability of exposure will be assigned larger weights so their influence on the comparison is increased. In addition to inverse probability weighting, other (that is, greedy nearest neighbor matching 1: k) algorithms may be considered if adequate balance is not achieved.

The PS weighted groups will be evaluated for balance and any patient restrictions will be described. For each patient characteristic, the prevalence (categorical variables) or mean (continuous variables) will be calculated in each cohort. Absolute standardized differences, the difference in means or proportions divided by the pooled standard deviation, will be computed for each covariate to check its distribution balance within exposure. For covariates with an absolute standardized difference greater than 0.10, residual differences will be further explored in the analysis phase.

Table 7. Study Variables to be Considered in PS Weighting

Data source	Variable
Longitudinal Prescription (LRx) database	Index year
	Age (years)[17]
	Sex
	Geography
	Payer type
	Medication use during baseline period
	ADM classes during baseline period
	ADM class immediately preceding index therapy
	Polypharmacy
	Count of medication classes during baseline period
	Diabetes severity (determined by time since first antidiabetic therapy)
Medical Claims (Dx) and P+ database (two separate sensitivity analyses)	History of radiation use
	History of cancer, not including MTC
	Inpatient visits
	Outpatient visits
	Diabetes severity (determined by clinical information)
	CCI[17], modified to exclude diabetes, prior to the index date

6.2.2.4. Primary Comparative Analyses

6.2.2.4.1. Incidence Rates of MTC Before PS Weighting

Incidence of MTC among the LA GLP-1 RA therapy cohort will be estimated as the number of LA GLP-1 RA therapy-exposed with a diagnosis of MTC during the study period divided by total person-years of follow-up before PS weighting.

6.2.2.4.2. Comparative IR/IRR After PS Weighting

IR and IRR for MTC will be estimated among LA GLP-1 RA treated patients versus the active comparator cohorts in the linked LRx-SCR patients, and in 2 sensitivity analyses data linkage cohorts linking LRx-SCR-Dx and LRx-SCR-P+ (see Figure 1 for details). The outcome, MTC, will be identified using predefined ICD-O-3 codes in SCR data (see Table 3). These codes are currently used for the MTC Registry.

For the primary analysis, the IRR and 95% CI for MTC occurrence in LA GLP-1 RA therapy users and active comparators will be estimated using Negative Binomial regression, Poisson regression, or other approaches will be considered to account for zeros (depending on data fit and model assumptions). The following describes the IRR for the study, a ratio of the IR of

[17] Age in years will be calculated at the index date using birth date information and will be reported as continuous as well as categorical measures (age groups).

MTC for patients treated with LA GLP-1 RA therapy vs. the IR of MTC for one of the comparator cohorts:

- PS weighted incidence of MTC among the LA GLP-1 RA therapy cohort will be estimated as the sum of the weights for the number of LA GLP-1 RA therapy-exposed patients with a diagnosis of MTC during the study period (weighted events) divided by sum of weights times the total person-years of follow-up (weighted person-years).
- PS weighted incidence of MTC among the comparator cohorts will be estimated as the sum of the weights for the number of comparators with a diagnosis of MTC during the study period (weighted events) divided by sum of the weights for the total person-years of follow-up among comparators (weighted person-years).

An IRR will be estimated as the incidence of MTC among the LA GLP-1 RA therapy cohort divided by incidence of MTC among the comparators. If sample size permits, hazard ratios and their corresponding 95% CIs will be estimated using Cox proportional hazards models to compare the LA GLP-1 RA therapy-exposed cohorts to the comparator cohorts. If Cox proportional hazards models are used, appropriate model checking diagnostics will be initiated.

6.2.2.5. Exploratory Analyses

For all exploratory analyses, estimates will be reported for both sets of study cohorts resulting from the 2 methods of restriction and adjustment for SCR participation (see Section 6.2.2.6.4). No additional PS re-estimations will be conducted for exploratory analyses.

6.2.2.5.1. Product-Specific Analysis

A subgroup analysis will include an evaluation of exposures and outcomes of MTC for each LA GLP-1 RA therapy drug product separately if the sample size allows. The IRs from each LA GLP-1 RA therapy stratification group will not be statistically compared. Counts of less than 5 will be reported as “N<5” within study results and all counts will be labeled as “not reported” if calculation of individual product subgroups is not possible.

6.2.2.5.2. Dose-Response Relationship Analysis

An analysis will be conducted to assess the dose-response relationship between the exposure (the index LA GLP-1 RA therapy) and the outcome (MTC incidence). The number of LA GLP-1 RA prescriptions will be used as a proxy to evaluate the cumulative exposure to LA GLP-1 RA therapy and an increased number of prescriptions would equate to an increased cumulative exposure. Additional analysis details will be included in the study SAP.

6.2.2.6. Sensitivity Analyses

Data-driven sensitivity analyses are planned. The need for data-driven sensitivity analyses will be driven by the number of identified MTC cases (for example, if there are few to no cases, there is little need for sensitivity analyses). The proposed secondary sensitivity analyses

will provide another dimension to the study and will add to the robustness of the overall interpretation of study findings.

All estimates in the sensitivity analyses will be reported for both sets of study cohorts resulting from the 2 methods of restriction and adjustment for SCR participation (see Section 6.2.2.2).

6.2.2.6.1. Exposure Misclassification

- **Requiring Two of the Same Index Medication of LA GLP-1 RA Therapy Dispensed Prescriptions**

A single dispensed prescription does not necessarily mean the patient took the medicine; however, the likelihood that the patient took the medicine increases if a second dispensed prescription for the same medication exists. This sensitivity analysis requires two of the same LA GLP-1 RA prescriptions to define LA GLP-1 RA exposure. In this subset of LA GLP-1 RA therapy-exposed patients, ≥2 prescription fills within 90 days of their initial LA GLP-1 RA fill will be required.[24] Similar definition would be required for the active comparator cohorts. Calculation of the IRR will then be performed in this group. Further details will be described in the study SAP.

- **Requiring Two of Any LA GLP-1 RA Therapy Dispensed Prescriptions**

A single dispensed prescription does not necessarily mean the patient took the medicine; however, the likelihood that the patient took the medicine increases if a second dispensed prescription for a LA GLP-1 RA exists. The analysis will be performed on LA GLP-1 RA therapy-exposed patients with ≥2 prescriptions within 90 days of the initial LA GLP-1 fill.[24] Similar definition would be required for the active comparator cohorts. Further details will be described in the study SAP.

- **Requiring Any LA GLP-1 RA Therapy Dispensed Prescriptions for 12-Month Post Index Without Discontinuation**

To capture patients with >2 prescriptions of any LA GLP-1 RA therapy, this sensitivity analysis will include patients with continued LA GLP-1 RA therapy use without discontinuation from index date to 12-month follow-up. The index date will be defined as the second dispensed prescription for a LA GLP-1 RA therapy. The LA GLP-1 RA treatment will be considered as discontinued when there is a gap of ≥90 days in the medication supply (that is, gap after the end of the days' supply of the last fill) during the 12-months following the first GLP-1-RA therapy.[24] Similar definition would be required for the active comparator cohorts. In addition, treatment switching between different LA GLP-1 RA therapies will be allowed in this analysis.

6.2.2.6.2. Outcome Misclassification – MTC Without Any Mixed Histological Types

A sensitivity analysis will be conducted to estimate MTC incidence with exclusion of histology codes for any mixed histological types, that include mixed medullary-follicular carcinoma (8346) and mixed medullary- papillary carcinoma (8347).

6.2.2.6.3. No Latency Assumption - Implementing a 6- and 12-Month Lag Period

The primary estimate of the IRR and 95% CI assumes that there is no lag time between treatment initiation and MTC to occur following the index date and does not account for the latency of MTC. Therefore, a sensitivity analysis of the IRR and 95% CI will be performed, allowing for 6- and 12-month latency periods following the index date if there are enough available patients. For this sensitivity analysis, follow-up time will be recalculated starting at 6 and 12 months after the index date rather than starting the day after the index date. This will decrease the amount of follow-up time in all cohorts, and it is assumed that the decrease will be non-differential across cohorts. Implementing the lag period will account for the biologically plausible risk period between LA GLP-1 RA use and MTC diagnosis.

LA GLP-1 RA therapy-exposed patients and their comparators who do not have at least 6 and 12 months of follow-up from their original index date and those who had a diagnosis of MTC prior to their revised index date, will be excluded from the sensitivity analysis. Other details will be described in the study SAP.

LA GLP-1 RA therapy-exposed patients, and the patients in the active comparator cohorts who have a diagnosis of MTC prior to their revised index date (accommodating the 6- and 12-month follow-up periods) will be excluded during analysis. This step will occur after cohort formation from LRx and linkage to the SCRs when the full dataset has been compiled.

6.2.2.6.4. Mortality

- **Mortality Adjustments**

Health outcomes vary by age, and subsequently the effect of the populations' age distributions will be considered as mortality generally increases with age. Since mortality files will not be used, mortality adjustments using Centers for Disease Control (CDC) and Prevention published rates will be applied to estimate appropriate time to censor each patient's person-years. A sensitivity analysis that includes assumptions about differential mortality between the cohorts will be conducted. This will be done by assuming up to 10% higher mortality for the LA GLP-1 RA therapy cohort and by calculating the percent differential mortality that would be necessary for the IRR to be statistically significantly elevated. The assumptions of 2%, 4%, 6%, 8%, and 10% will be calculated, respectively. This differential mortality analysis will be repeated with a lower mortality assumption and further detailed in the study SAP.

Further, as a sensitivity analysis to account for a mortality adjustment among older patients, follow-up will not continue to the end of the study for all patients. Follow-up duration for the study cohorts will be derived from the CDC *United States Life Tables, 2020* (https://www.cdc.gov/nchs/data/nvsr/nvsr71/nvsr71-01.pdf).

For all ages up to 69, the expected number of years of life remaining is >15 years of age, and the entire study observation period is 15 years. Therefore, all patients <70 years of age at their index date will be followed until the first of the other censoring criteria. For patients in the ≥70-year age group, the years of life remaining for 85-year-olds will be used as a proxy for duration of follow-up. Follow-up time for people in this age group will accumulate to the earliest of the other censoring criteria, or 5.6 years for men and 6.4 years for women.

Note: If a patient has an MTC diagnosis date (per state cancer registry) that occurs *after* the mortality-adjusted end of follow-up, then follow-up time will be accumulated until the date of the MTC diagnosis and *not* truncated earlier.

- **Claims-Based Mortality Algorithm**

As the claims databases do not include mortality information, a published claims-based algorithm will be employed to identify patients with evidence of mortality during the follow-up period in the Dx and P+ linkage sensitivity analyses.[18] Based on the presence of and the time from the last fatal event code observed to the last claim in the data, a mortality flag will be assigned to patients. The date of the last claim will be considered as the potential mortality date among patients with a mortality flag assigned. The list of fatal event codes will include event codes such as brain death, cardiac arrest, respiratory arrest, myocardial infarction, palliative care, do not resuscitate, and mechanical ventilation. An analysis of censoring patients with a mortality flag will be conducted. Further details will be contained in the study SAP, including the full algorithm.

6.2.2.6.5. Adjustment for Potential Confounders with Additional Linkage with Dx and P+ in Two Separate Sensitivity Analyses

For sensitivity analyses, an attempt will be made to link all study cohorts to the Dx and P+ database separately, for ascertainment of medical claims representing baseline patient characteristics. Patients from linked LRx-SCR database for the primary analysis will be linked to Dx and P+ databases separately with a few additional requirements.

LRx-SCR patients linkable to Dx would be required to meet the following requirements:

- ≥1 medical record in Dx within 24 months prior to the index date
- ≥1 medical record after the index date and within 12 months of index date

- No evidence of thyroidectomy (identified using diagnosis or procedure codes)[18] during 24 months prior to the index date

LRx-SCR patients linkable to P+ would be required to meet the following requirements:

- With continuous medical enrollment during the 24 months prior to index date
- With continuous medical enrollment during the index month
- No evidence of thyroidectomy (identified using diagnosis or procedure codes) during 24 months prior to the index date

The preliminary linkage between LRx and Dx databases for this patient population was estimated to be 65.2% (see Section 6.1 for feasibility results), which may reduce after application of linkage methodology and remaining selection criteria. Following cohort selection from LRx and SCR linkage (the final analytical dataset), linkages to Dx and P+ will commence. Confounder adjustment using variables from Dx and P+ for all linked patients with available Dx and P+ data will be explored through PS weighting and model adjustment for the primary objective. Within the sub-cohort with LRx-SCR-Dx and LRx-SR-P+ linkages, baseline covariates such as history of radiation exposure, history of cancer, and CCI will be explored to characterize the cohorts (second primary objective). The LRx-SCR-P+ linked subset will also be used to assess the patient characteristics between LRx and P+ databases in order to support the determination of overall completeness of patient information in LRx database and correct exposure classification. These characteristics will include, but not be limited to, an assessment of completeness of variables in LRx and P+ and longitudinal information in each database during similar timeframes. Additional details on these comparisons will be outlined in the SAP.

Further methods regarding this data linkage and subsequent baseline covariate investigation will be described in detail in the SAP.

6.2.2.7. Quantitative Bias Assessment for Unmeasured Confounding

A quantitative bias assessment (QBA) will be conducted with a detailed plan developed in an independent document. The QBA will assess the impact of unmeasured confounding in the study.

While the large size of the present analysis provides justification for the assumption that random error in study results will be minimized, it is recognized that systemic error by way of certain unobserved confounders is not corrected simply by study population size. MTC has very few risk factors, the majority of which are herein captured (age, sex, radiation history,

[18] Thyroidectomy variable will be an exclusion and not an analytic variable and thus has not been added to the variables list.

diabetes severity (as applicable to thyroid cancer, impact on MTC specifically is unknown). However, 25%[19] of MTC cases are linked to familial history and a mutation of the RET gene, a factor that is not observable in administrative pharmacy (LRx) and medical (Dx/P+) claims and is not a routine component of state cancer registry capture. For this reason, a QBA is planned to estimate the minimum strength of association of unmeasured confounding with exposure and outcome to fully explain away the observed exposure-outcome association, and to estimate the impact of unmeasured confounding bias from familial MTC (FMTC) on the patient level IRR estimates observed in the sensitivity (LRx-SCR-Dx linked data) analyses.

Given the size and complexity of the QBA planned for this MTC linkage study, a separate document describing the QBA plan will be developed independent of the study protocol and SAP. The QBA plan will capture all relevant details required to fully determine deviations from study effect estimates from estimates produced from a hypothetical RCT. Operationally, the QBA will include an E-value and a patient-level adjustment of protocol IRRs excluding patients with a positive history of FMTC using a proxied variable for status of history of FMTC.

6.2.3. Assessing MTC Incidence in the US During the Study Period (Secondary Objective Analysis)

The official federal cancer statistics program, the US Cancer Statistics, will be used to provide data for annual incidence of MTC cases for the study period. This data source includes cancer registry data from the CDC National Program of Cancer Registries (NPCR), the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. The US Cancer Statistics data source includes information on newly diagnosed cancer cases and cancer deaths for the entire US population. Through NCPR, CDC supports central cancer registries in 46 states, the DC, Puerto Rico, the US Pacific Island Jurisdictions, and the US Virgin Islands. SEER collects and publishes cancer incidence and survival data from population-based cancer registries in 22 US geographic areas.

To understand the annual MTC incidence trends over time, the IRs will be stratified by year and graphically presented. In addition, to supplement these results, a literature review of publicly available data will be performed to further understand LA GLP-1 RA therapy market trends. Trends of annual incidence of MTC will not be submitted to Health Authorities on an annual basis but will support the interpretation of the totality of data and included with submission of the final report.

Specific methods for this analysis will be developed in the SAP.

[19] 75% of MTC cases are sporadic in nature, of which the etiology is unknown

6.2.4. Handling of Missing Data

Missing data will be explored to determine if imputation is necessary. For example, for any missing month values, imputation to the mid-year date (for example 01 July) may be explored and will be further described in the SAP.

The LA GLP-1 RA therapy-exposed cohort with missing or invalid days' supply and quantity dispensed values on one or more LA GLP-1 RA therapy dispensed prescription claims between index and the earlier of the end of study period or date on which MTC is detected, whichever comes first, will not be included in assessment of cumulative LA GLP-1 RA therapy exposure. However, missing data on days' supply and quantity dispensed is not expected to be common in LRx database.

LRx and Dx databases are open claims databases, where continuous patient enrollment in health plans cannot be confirmed. To that end, the full treatment journey of a patient during the study period may not be captured if care is received outside of offices/hospitals/pharmacies captured in the study databases. To address this limitation, our study will ensure patient stability in the LRx database using evidence of consistent data availability during the baseline and post-index period (see Section 5.2 for details).

6.2.5. Strengths and Limitations of Research Methods

6.2.5.1. Strengths and Limitations

This study aims to provide novel and comprehensive findings on the risk of MTC in relation to the treatment (LA GLP-1 RA therapies) that has not yet been characterized using RWD. The LRx database represents US retail, specialty, and mail order prescriptions, as well as prescriptions filled at LTC facilities across therapeutic areas (see Section 5.1.2.1). While the open nature of the database has its own unique set of challenges, we will be able to include adult patients of all ages (including age group with highest MTC incidence [40 to 60 years of age]), geographic locations, and health plan types (or lack thereof). As a result, the findings of the study will be generalizable to the broader population of patients treated with LA GLP-1 RA therapies in the US. Additionally, the linkage of LRx and SCR databases is the gold standard approach to ascertain the study outcome (MTC diagnosis) in relation to the treatment.

The LRx database is an open database; therefore, if a patient fills a prescription at a pharmacy that did not report to IQVIA (or reported inconsistently), those data will not be captured, resulting in incomplete data and possible misclassification of exposure duration. Additionally, it is possible that patients who are dispensed a prescription did not actually take the medication, resulting in misclassification of exposure. However, to maximize the population studied, the main study analysis will require only one dispensing of a LA GLP-1 RA therapy to be considered as ever exposed. This may overestimate exposures and a sensitivity analysis is planned to evaluate this. LRx is also strengthened by its size (covering 94% of retail pharmacies and 82% of claims across payer types), the adjudicated nature of the claims,

longitudinality of patient time in LRx and the sourcing of LRx data direct from pharmacies instead of from third party switch data centers, which would mitigate any misclassification of exposure (also non-differential by cohort).[9,10]

The LRx and Dx open claims data houses a wealth of patient health information, with specific analytic variables described in the variables table (Section 5.4, Table 3). While patients may change providers over the course of their healthcare utilization in both LRx and Dx, it is not expected that these changes would be differential by study cohort and would thus not affect study effect measures.

On another note, pharmacy claims data are not captured for research purposes, but for billing, and SCRs collect data in part for public health purposes including monitoring of cancer trends. Both data sources were not specifically designed for the conduct of this research study and therefore, linkage rates between the data sources (LRx and SCR data) are not expected to be 100%. However, a deterministic data linkage method has been used in a previous study to link LRx to SCR data with a linkage rate of 89%.[26] Finally, the outcome ascertainment is based on linkage with participating SCRs. Not all cancer registries will participate which will not allow for ascertainment of 100% of MTC cases in the US. Two methods will be applied to attempt to adjust for this. The first will include only including data from participating registries, and the second will be to apply a coverage fraction. Although all cases may not be included, many cases in the US will be captured, including representation of Medicare and Medicaid recipients.

6.2.5.2. Confounding and Bias

A confounding factor is an independent risk factor for MTC that is associated with LA GLP-1 RA therapy exposure. There are, however, few established risk factors for MTC, including family history of MTC (which is not available in the LRx, Dx or SCRs). Patient characteristics will be balanced between groups by virtue of the PS weighting; however, residual confounding by unmeasured confounders, such as family history of MTC for example, may be present. A personal history of cancer irrespective of type is available in SCRs, but family history of cancer is not available in SCRs. To account for familial history of MTC and any other unmeasured confounding, a QBA analysis is included, and detailed methods for this approach is described in Section 6.2.2.7 and will be further delineated in the SAP and an independent QBA study plan. Further, PS weighting estimates average treatment effects at the population level for the study groups that could help reduce the impact of detection bias as we would be controlling for patient characteristics that would otherwise differ between the study cohorts. The study team will exercise flexibility in the PS weighting scheme to ensure optimal balance between the exposed and active comparator populations indexed and consider most PS adjustment in the model-based analyses where necessary.

In addition to the strengths of the chosen databases, the active comparator design reduces bias from indication, making 2 groups comparable as all patients are in need of initiating new treatment. In this study, the use of ADMs and AOMs are assumed to follow the approved

indications of diabetes and obesity. Any off-label use of ADMs and AOMs will not be considered in this study.

Reporting bias will also be mitigated using the single, reliable, and robust source of MTC incidence in the US. The SCRs report data following the NAACCR guidelines and most participating SCRs have a gold data standard certification level (see Section 5.1.2.2 for more details).

An additional limitation to be addressed in the study is channeling bias. Channeling bias can occur when drugs that have similar indications are prescribed to patients that have different prognoses. In the case of LA GLP-1 RAs, it is possible that patients with genetic risk for MTC may be "channeled" away from GLPs and therefore an increased number of patients with familial history of MTC could result in the comparator groups. Because this will not be observable by design of the study (that is, genetic information is not available for patients), this bias will be mitigated in 2 ways. First, this study will evaluate MTC by age group; peak onset of the disease appears to be between 40 and 60 years of age. Stratifying by age group in the present analysis will help to tease out differences in age groups that could also be due to presentation of familial MTC. Second, as noted in Section 6.2.2.5, a QBA is planned for this study that will account for familial history and produce an adjusted study estimate accounting for family history of MTC. Additional details for the QBA and how familial history will be accounted for in that analysis will be captured in a separate QBA study plan. Importantly, in a recent study conducted in the United Kingdom's Clinical Practice Research Datalink (CPRD), channeling bias was investigated between LA GLP-1 RA and DPP-4, and no channeling bias was found when comparing to initiators of insulin and sulfonylureas.[27]

7. STUDY MANAGEMENT

This study will be performed by IQVIA with guidance, input, review, and approval of the Client, including development of materials, ethical review, training, and management of SCRs, electronic data capture and data management, linkage, and analysis.

7.1. Database Linkage Process

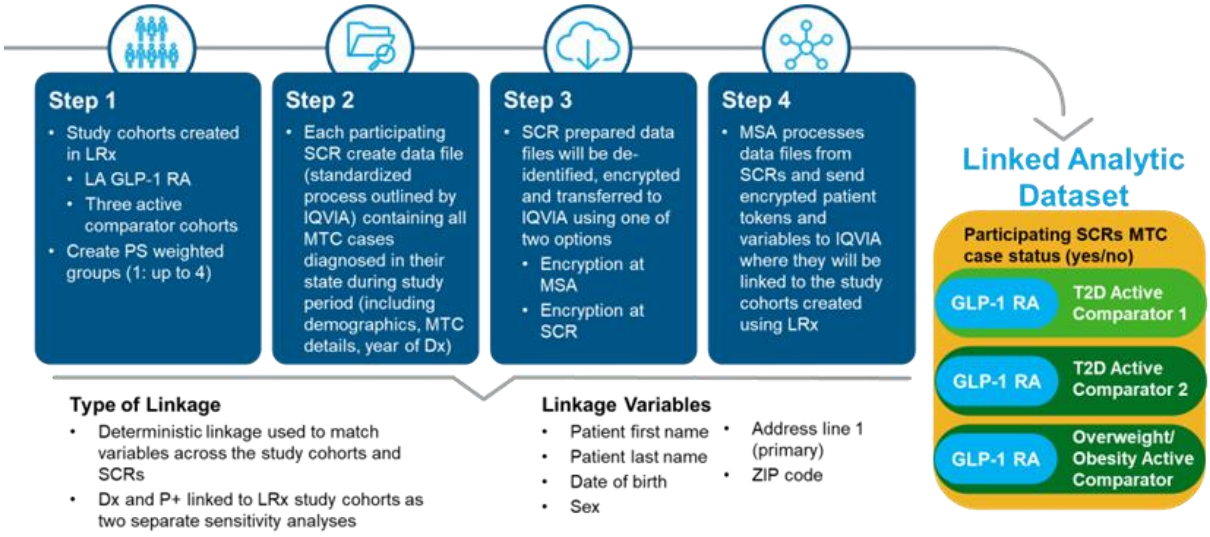
Patients treated with LA GLP-1 RA therapy and active comparators will be selected from IQVIA LRx data that will be linked to SCRs and clinical data (Dx and P+) via a tokenization process. This study will use the standard IQVIA de-identification and linkage process.[28] Tokenization transforms identifiable patient information into an irreversible random string of characters that has no meaningful value but can be used to perform de-identified longitudinal analysis of patients. Applying a uniform tokenization method across disparate datasets enables patient-level linking and longitudinal analysis across datasets. Linking variables will be used to perform a deterministic data linkage (if feasible) which will link the SCR data to an anonymous IQVIA patient ID number. If deterministic data linkage is not feasible, due to lack of linking variable availability, probabilistic data linkage will be undertaken. This process is certified as HIPAA compliant and is depicted below (Figure 3). As noted previously, LRx will be the anchor dataset for this analysis, meaning that the study cohorts will be sourced from LRx and the study population assembled. Tokenized variables will be created for first and last name, sex, address including zip code, and date of birth. Tokenized (meaning hashed) variables, which are simply alphanumeric strings, are created in place of the patient data. The same tokenization procedure produces the same hash in the dataset that will be linked. This means that the datasets can be joined together on those patient attributes. The variables remain hashed throughout analysis, protecting patient privacy.

Patient matching controls for changes in patient demographics caused by normal life events and minimizes patient confusion that improves match quality by minimizing false positives and false negatives. For example, based on identifiable information, 2 patient records appear to be the same person that had a last name change after marriage. Using IQVIA's process of tokenization, the generated patient IDs are compared after matching resulting in matched patient. This matching process also controls for "snowbirds", name changes via marriage, and variations on name (that is, "William" versus "Bill"), amongst others. During the tokenization process, it is a small possibility that patients can be over or under-linked to IQVIA patient IDs. But over-linking of patient IDs occurs approximately 0.94% of the time and under linking occurs far less to be able to specifically measure. Further details on IQVIA tokenization process are publicly available in a fact sheet.[28]

Tokenization Process Rules for LRx, SCR, Dx and P+ Database Linkage

- The IQVIA software uses various combinations of patient demographics to create irreversible hashed tokens that IQVIA can use to assign an anonymous Patient ID.
    - first name, last name, date of birth, sex, address, zip code

- Patient zip code is limited to 3 digits and low-density zip codes are removed
- Patients >85 years of age automatically have birth year hidden
- The process has been reviewed by a third party privacy expert to meet HIPAA's Expert Determination Method of de-identification, and
- If additional data elements are to be provided to IQVIA for analytics, IQVIA Information Governance team would review and a re-identification risk determination (RRD) is required to be performed on the data IQVIA's Privacy Analytics or other third party privacy expert can be leverage for(RRD). If the (RRD)requires data edits, those edits would be done at the data partner's site or a trusted third party prior to coming to IQVIA.



Abbreviations: Dx = Medical claims database; GLP-1 RA= glucagon-like peptide-1 receptor agonists; LA = long-acting; LRx = IQVIA Longitudinal Prescriptions Database; MTC = medullary thyroid carcinoma; MSA = Management Science Associates; PS = propensity score; P+ = PharMetrics® Plus database; SCR = State Cancer Registry; T2D = type 2 diabetes mellitus.

Figure 3. Data linkage process.

7.2. Database Processing and Transfer

All SCRs in the US will be invited to participate in the study, which will involve preparation of a dataset of patients diagnosed with MTC during the study period for de-identification and subsequent use in the specified analyses. It is anticipated that not all cancer registries will be able or willing to participate due to lack of resources and/or regulations that prohibit them from sending identifiable data to third parties.

SCRs will be asked to follow a standard process for preparing the linkage file. The file will be sent to the IQVIA's trusted third party data processor, Management Science Associates (MSA, https://www.msa.com/), for de-identification of variables necessary for linkage. MSA

is an IQVIA vendor that handles encrypted and secured data files and can combine these files into a single de-identified datafile by simply compiling the datasets from SCRs. The file will also contain MTC tumor-specific information to allow the study team to better understand and describe cases of MTC that link to the study cohorts.

Participating SCRs will transfer data using one of the two options described and depicted below. The following steps describe the data transfer process:

Step 1: Prepare a file with a minimum the following variables:

Variables used in combination to create tokens for linkage:

- patient first and last name
- birth date
- patient sex
- patient address 1 (patient's primary correspondence address 1)
- patient ZIP code (patient's primary correspondence zip code)

Variables utilized for study analyses (not tokenized):

- MTC diagnosis codes (see Table 3)
- primary site
- month and year of MTC diagnoses
- MTC diagnostic confirmation
- other clinical/pathological information (for example, stage, size), and
- tumor number

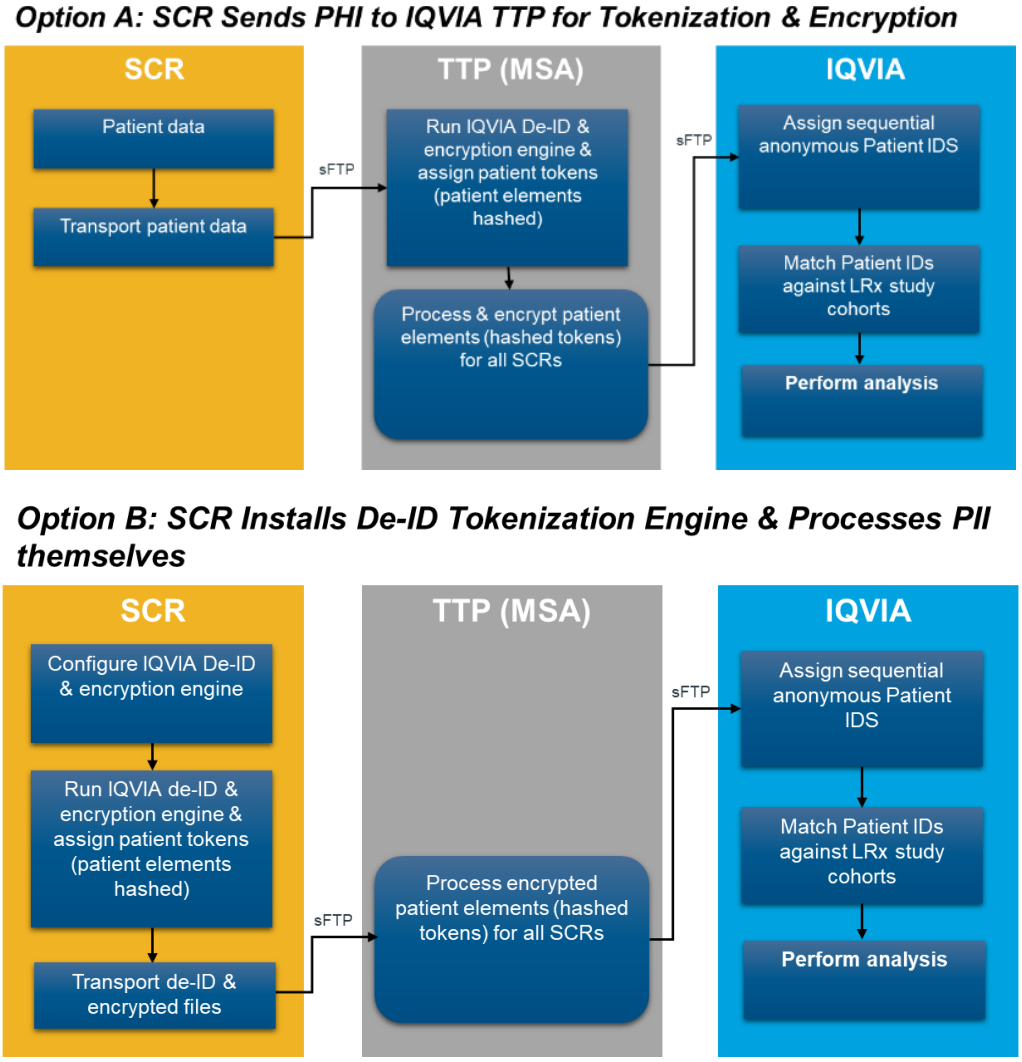
SCRs will be asked to follow a standard process for preparing the linkage file, although files sent to the third party can be in different layouts if all data elements are represented, and the format/layout agreed with the third party a priori.

Step 2: De-identify the file using IQVIA encryption engine and one of the following options:

Option A: Prepare data files containing prespecified information for MTC patients and send resulting files to MSA for variable de-identification necessary for linkage.

Option B: Run the encryption engine locally at the state cancer registry and transfer resulting encrypted patient token to MSA.

Step 3: The trusted third party will send the encrypted tokens to the research team at IQVIA where they will be linked to encrypted study cohorts (GLP-1 RA exposed cohort and the 3 active comparator cohorts) created using the LRx database. Figure 4 describes the steps for linkage for each option.



Abbreviations: MSA = Management Science Associates; SCR = State cancer registry; TTP = trusted third party. Option A. SCR Data Transmission/Linkage: Encryption at the Trusted Third Party; Option B. SCR Data Transmission/Linkage: Encryption at the SCR.

Figure 4. Linkage options.

7.3. Data Management

The IQVIA is responsible for the integrity of the data reported to the Client. Datasets and analytic programs will be stored according to IQVIA procedures with access restricted to study personnel. Data provided by the SCRs will be destroyed following data destruction procedures specified by the cancer registries and agreed to by IQVIA.

The IQVIA confidentiality agreements are signed by all employees and include data protection and strict prohibitions on re-identification attempts. All aspects of the study will be conducted within the framework of the IQVIA Quality Management System. A QC plan for

the study will be developed and executed. The IQVIA will document and retain a quality review of all final deliverables.

7.4. QA and Monitoring

All aspects of the study from protocol development to the reporting of the results will be conducted within the framework of the IQVIA Quality Management System. A QC checklist for the study has been developed and executed on the study protocol. Furthermore:

- The study QC checklist will establish ownership for the execution of the individual QC steps
- The Principal in Charge of the study project will ensure that individuals responsible for the execution of specific QC steps will have the knowledge, capability, and experience necessary to perform the assigned tasks, and
- The result of the execution of the individual steps of the QC plan will be documented, and will include the required corrective actions, if any. The execution of any required corrective action will also be documented.

The QC checklist will be subjected to a final review and approval for sufficiency and completeness from the Principal in Charge.

7.5. Plans for Disseminating and Communicating Study Results

A final report will be submitted to regulatory agencies. The study, including the final report, may also be registered in the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) registry. The study findings may be submitted to a scientific congress and/or to a peer-reviewed journal.

8. SAFETY REPORTING

8.1. Secondary Data Use Study

This is a non-interventional study based on secondary data use, and therefore, no ICSR is required. The study protocol-defined AEs include: MTC. The MedDRA code for MTC is C0238462[20]. This MedDRA code will also be reviewed and updated, as needed during final study report authoring. All captured MTC cases will be summarized in the final study report. No other AEs will be collected.

[20] The current study is planned to use US data sources, which do not use MedDRA codes, but it is added here for comprehensiveness.

9. ETHICAL AND REGULATORY CONSIDERATIONS

9.1. Guiding Principles

To ensure the quality and integrity of research, this study will be conducted under the guidelines GPPs issued by the International Society for Pharmacoepidemiology (ISPE), the Declaration of Helsinki and its amendments, and any applicable national guidelines.

The study will be conducted in compliance with the US FDA Title 21 CFR Part 50 – Protection of Human Patients and/or Part 56 – Institutional Review Boards; the International Conference on Harmonization (ICH) for Good Clinical Practice (GCP) E6 (R2) guidelines (15 December 2016) as they apply to non-interventional studies; the Declaration of Helsinki and its amendments; and HIPAA.

The final study report will be written in accordance with the GVP guidelines module VIII, (EMA/813938/2011), and the RECORD-PE checklist (ref: https://www.record-statement.org/checklist-pe.php).

9.2. Patient Confidentiality

To maintain patient confidentiality, each patient will be assigned a unique patient identifier upon data extraction. This patient identifier will be used in place of patient name for the purpose of data analysis and reporting. Medical record number or other local reference identifiers are not collected as part of the database. All parties will ensure protection of patient personal data and will not include patient names on any study forms, reports, publications, or in any other disclosures. Data protection and privacy regulations will be observed in capturing, forwarding, processing, and storing patient data. Every effort will be made to protect patient confidentiality according to the Directive 95/46/EC on the protection of individuals, and in compliance with Safe Harbor privacy principles.

The study database will be housed at the IQVIA in a physically and logically secure computer system maintained by the IQVIA in accordance with a written security policy. The system meets approved established standards for the security of health information and is validated. The system also meets the standards of the ICH GCP E6 guideline (revision 2) regarding electronic study data handling and is available for audit upon request. Patient confidentiality will be strictly maintained.

9.3. IEC/Institutional Review Board

Consistent with local regulations, the study protocol will be submitted to the responsible central and state IRB/IECs for review to utilize the SCR data (note, LRx and Dx data sources do not require IRB/IEC approval for use). Data extraction will not start before the Client has obtained written confirmation of a favorable opinion/approval from the relevant central or local IRBs/IECs. The IRBs/IECs will be asked to provide documentation of the date of the

meeting at which the favorable opinion/approval was given that clearly identifies the study and the protocol version reviewed.

Before implementation of any substantial changes to the protocol, protocol amendments will also be submitted to the relevant IRBs/IECs in a manner consistent with local regulations. Pertinent safety information will be submitted to the relevant IECs during the study in accordance with local regulations and requirements. It is the responsibility of the investigator to have prospective approval of the study protocol, protocol amendments, and other relevant documents, if applicable, from their local IRB/IEC and provide documentation of approval to the Client. All correspondence with the IRBs/IECs should be retained in the investigator file.

Should the study be terminated early for any unanticipated reason, the investigator will be responsible for informing the IRBs/IECs of the early termination.

9.4. Changes to the Protocol

Changes to the protocol will be documented in written protocol amendments. Major (that is, substantial, significant) amendments will usually require submission to the relevant IRB/IECs for approval or favorable opinion and the FDA, if applicable. In such cases, the amendment will be implemented only after approval or favorable opinion has been obtained.

Minor (non-substantial) protocol amendments, including administrative changes, will be filed submitted to the relevant IRB/IEC or regulatory authorities where required by pertinent regulations.

10. REFERENCES

1. 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.
2. Madsen LW, Knauf JA, Gotfredsen C, et al. GLP-1 receptor agonists and the thyroid: C-cell effects in mice are mediated via the GLP-1 receptor and not associated with RET activation. *Endocrinology.* 2012;153(3):1538-1547.
3. Stamatakos M, Paraskeva P, Stefanaki C, et al. Medullary thyroid carcinoma: The third most common thyroid cancer reviewed. *Oncol Lett.* 2011;2(1):49-53.
4. National Cancer Institute. Medullary Thyroid Cancer (MTC). https://www.cancer.gov/pediatric-adult-rare-tumor/rare-tumors/rare-endocrine-tumor/medullary-thyroid-cancer#:~:text=Thyroid%20cancer%20is%20fairly%20common,each%20year%20in%20the%20U.S. Published 2019. Accessed.
5. Milićević S, Bergant D, Žagar T, Perić B. Crude annual incidence rate of medullary thyroid cancer and RET mutation frequency. *Croat Med J.* 2021;62(2):110-119.
6. Fialkowski EA, Moley JF. Current approaches to medullary thyroid carcinoma, sporadic and familial. *J Surg Oncol.* 2006;94(8):737-747.
7. Elisei R, Romei C, Cosci B, et al. RET genetic screening in patients with medullary thyroid cancer and their relatives: experience with 807 individuals at one center. *J Clin Endocrinol Metab.* 2007;92(12):4725-4729.
8. Jiménez C, Hu MI, Gagel RF. Management of medullary thyroid carcinoma. *Endocrinol Metab Clin North Am.* 2008;37(2):481-496, x-xi.
9. Baser O, Samayoa G, Yapar N, Baser E, Mete F. Use of Open Claims vs Closed Claims in Health Outcomes Research. *J Health Econ Outcomes Res.* 2023;10(2):44-52.
10. Franklin JM, Lin KJ, Gatto NM, Rassen JA, Glynn RJ, Schneeweiss S. Real-World Evidence for Assessing Pharmaceutical Treatments in the Context of COVID-19. *Clin Pharmacol Ther.* 2021;109(4):816-828.
11. Kosiborod M, Aroda VR, Broe Honore J, et al. Trends in initiation of GLP-1 RA in patients with type 2 diabetes during 2014–2019: a US database study. *European Heart Journal.* 2021;42(Supplement_1).
12. Palanca A, Ampudia-Blasco FJ, Calderón JM, et al. Real-World Evaluation of GLP-1 Receptor Agonist Therapy Persistence, Adherence and Therapeutic Inertia Among Obese Adults with Type 2 Diabetes. *Diabetes Ther.* 2023;14(4):723-736.
13. Blonde L, Burudpakdee C, Divino V, et al. The impact of non-medical switch on type 2 diabetes patients treated with canagliflozin in the commercially insured US population. *Current Medical Research and Opinion.* 2018;34(8):1501-1511.
14. Portillo EC, Pollack M, Lee I, et al. Characteristics of initiators of budesonide/glycopyrrolate/formoterol for treatment of chronic obstructive pulmonary

disease (COPD) in the United States: the AURA study. *Ther Adv Respir Dis.* 2023;17:17534666231164534.

15. Otto T, Myland M, Jung H, Lebrec J, Richter H, Norrbacka K. Utilization patterns of glucagon-like peptide-1 receptor agonists in patients with type 2 diabetes mellitus in Germany: a retrospective cohort study. *Current Medical Research and Opinion.* 2019;35(5):893-901.
16. NAACCR Cancer Registry Standards Certification Criteria. North American Association of Central Cancer Registries https://www.naaccr.org/certification-criteria/. Published 2022. Accessed 08/22/2023.
17. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol.* 2011;173(6):676-682.
18. Wade RL, Nunna S, Zhou Z, Sun K. A Rigorous Examination of the Coding Algorithms Used for Identifying Mortality in Administrative Claims Data. *Academy of Managed Care Pharmacy (AMCP) Nexus* 2020.
19. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? A systematic review of definitions. *BMC Geriatr.* 2017;17(1):230.
20. Bezin J, Gouverneur A, Pénichon M, et al. GLP-1 Receptor Agonists and the Risk of Thyroid Cancer. *Diabetes Care.* 2022;46(2):384-390.
21. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *Am J Epidemiol.* 2006;163(12):1149-1156.
22. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* 2009;28(25):3083-3107.
23. Rassen JA, Glynn RJ, Rothman KJ, Setoguchi S, Schneeweiss S. Applying propensity scores estimated in a full cohort to adjust for confounding in subgroup analyses. *Pharmacoepidemiol Drug Saf.* 2012;21(7):697-709.
24. Weiss T, Carr RD, Pal S, et al. Real-World Adherence and Discontinuation of Glucagon-Like Peptide-1 Receptor Agonists Therapy in Type 2 Diabetes Mellitus Patients in the United States. *Patient Prefer Adherence.* 2020;14:2337-2345.
25. Kellier-Steele N, Casso D, Anderson A, Oliveria SA, Motsko S. Assessing the incidence of osteosarcoma among teriparatide-treated patients using linkage of commercial pharmacy and state cancer registry data, contributing to the removal of boxed warning and other labeling changes. *Bone.* 2022;160:116394.
26. Eli Lilly and Company. *B3D-MC-GHBX (2.3b): Observational study assessing incidence of osteosarcoma among Forteo(teriparatide) users by linking state cancer registry data to large national pharmacy database data.* 2018.
27. Ankarfeldt MZ, Thorsted BL, Groenwold RH, Adalsteinsson E, Ali MS, Klungel OH. Assessment of channeling bias among initiators of glucose-lowering drugs: A UK cohort study. *Clin Epidemiol.* 2017;9:19-30.

28. *Patient Tokenization: Enabling patient-level linkages in a privacy protected manner.* IQVIA2021.

11. APPENDICES

Table 8. List of Abbreviations

Abbreviation	Description
ADM	Anti-Diabetic Medication
AE	Adverse event
AMA	American Medical Association
AOM	Anti-Obesity Medication
ATE	Average treatment effect
AUC	Area under the curve
BMI	Body mass index
CCI	Charlson Comorbidity Index
CDC	Centers for Disease Control & Prevention
CI	Confidence Interval
CMS	Centers for Medicare & Medicaid Services
CPRD	United Kingdom's Clinical Practice Research Datalink
CRO	Contract research organization
DC	District of Columbia
DPP-4	Dipeptidyl peptidase IV
Dx	IQVIA Open Medical Claims Database
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
FDA	US Food and Drug Administration
FMTC	Familial medullary thyroid carcinoma
GCP	Good clinical practice
GIP/GLP-1 RA	Glucose-dependent insulinotropic polypeptide/glucagon-like peptide-1 receptor agonist
GLP-1 RA	glucagon-like peptide-1 receptor agonist
GPP	good pharmacology practices
GVP	good pharmacovigilance practices
HCPCS	Healthcare Common Procedure Coding System
HMO	Health maintenance organization
HIPAA	Health Insurance Portability and Accountability Act of 1996
HR	Hazard Ratio
ICD-10	International Classification of Diseases, Tenth Revision
ICD-9	International Classification of Diseases, Ninth Revision
ICD-O-3	International Classification of Diseases for Oncology, 3rd edition
ICH	International Conference on Harmonization
ICSR	Individual case safety report
IEC	Independent ethics committee

Abbreviation	Description
IPTW	Inverse Probability of Treatment Weighting
IR	Incidence Rate
IRB	Institutional Review Board
IRR	Incidence Rate Ratio
ISPE	International Society for Pharmacoepidemiology
LA GLP-1 RA	long-acting GLP-1 RA
LRx	IQVIA Longitudinal Prescriptions Database
LTC	Long-term care
MedDRA	Medical Dictionary for Regulatory Activities
MEN2A and MEN2B	Multiple endocrine neoplasia syndromes
MSA	Management Science Associates
MTC	Medullary Thyroid Carcinoma
NCPDP	National Council for Prescription Drug Programs
NDC	National Drug Code
NPCR	National Program of Cancer Registries
NPI	National Provider Identifier
P+	IQVIA PharMetrics® Plus Database
PAS	Post-Authorization Study
PASS	Post-Authorization Safety Study
PLC	Publicly Limited Company
PPO	Preferred provider organization
PRAC	Pharmacovigilance and Risk Assessment Committee
PS	Propensity Score
QA	Quality Assurance
QBA	Quantitative Bias Assessment
QC	Quality Control
RCT	Randomized controlled trial
RET	Rearranged during transfection
RRD	Re-identification risk determination
RWD	Real-World Data
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SCR	State Cancer Registry data
SEER	National Cancer Institute Surveillance, Epidemiology, and End Results program
SGLT-2	Sodium-glucose cotransporter-2
SmPC	Summary of product characteristics
STROBE	STrengthening the Reporting of OBservational studies in Epidemiology
T2D	Type 2 diabetes mellitus

Abbreviation	Description
TBD	To be determined
TZD	Thiazolidinediones
US	United States
zip	Zone improvement plan

Table 9. ENCePP Checklist

Study title: Database Linkage Study to Evaluate the Risk of Medullary Thyroid Carcinoma

EU PAS Register® number: Study not yet registered Study reference number (if applicable): Study not yet registered

Section 1: Milestones	Yes	No	N/A	Section Number
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection21	☒	☐	☐	1
1.1.2 End of data collection22	☒	☐	☐	1
1.1.3 Progress report(s)	☐	☐	☒	1
1.1.4 Interim report(s)	☐	☐	☒	1
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	1
1.1.6 Final report of study results	☒	☐	☐	1

Comments:

Section 2: Research question	Yes	No	N/A	Section Number
2.1 Does the formulation of the research question and objectives clearly explain:				

21 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.

22 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)	☒	☐	☐	4
2.1.2 The objective(s) of the study?	☒	☐	☐	4
2.1.3 The target population? (i.e., population or subgroup to whom the study results are intended to be generalized)	☒	☐	☐	5.3
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)	☒	☐	☐	5
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	5
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	6.2
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))	☒	☐	☐	6.2
3.5 Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g., AEs that will not be collected in case of primary data collection)	☒	☐	☐	8

Comments:

Section 4: Source and study populations	Yes	No	N/A	Section Number
4.1 Is the source population described?	☒	☐	☐	5.3
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period	☒	☐	☐	5.1
4.2.2 Age and sex	☒	☐	☐	5.5
4.2.3 Country of origin	☒	☐	☐	5.1
4.2.4 Disease/indication	☒	☐	☐	5.3
4.2.5 Duration of follow-up	☒	☐	☐	5.4
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g., event or inclusion/exclusion criteria)	☒	☐	☐	5.3

Comments:

Section 5: Exposure definition and measurement	Yes	No	N/A	Section Number
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g., operational details for defining and categorizing exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	5.3, 5.5
5.2 Does the protocol address the validity of the exposure measurement? (e.g., precision, accuracy, use of validation sub-study)	☒	☐	☐	5.2
5.3 Is exposure categorized according to time windows?	☒	☐	☐	5.4

Section 5: Exposure definition and measurement	Yes	No	N/A	Section Number
5.4 Is intensity of exposure addressed? (e.g., dose, duration)	☒	☐	☐	5.5
5.5 Is exposure categorized based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	5.3.2

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?	☒	☐	☐	5.5
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	5.5
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 utilization, 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)	☒	☐	☐	5.3.3
7.2 Does the protocol address selection bias? (e.g., healthy user/adherer bias)	☒	☐	☐	6.2.7
7.3 Does the protocol address information bias? (e.g., misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	6.2.7

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, subgroup analyses, anticipated direction of effect)	☒	☐	☐	6.2.7

Comments:

Section 9: Data sources	Yes	No	N/A	Section Number
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g., pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	5.2.1
9.1.2 Outcomes? (e.g., clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	5.2.2

Section 9: Data sources	Yes	No	N/A	Section Number
9.1.3 Covariates and other characteristics?	☒	☐	☐	5.2.1, 5.2.3
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g., date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	5.5
9.2.2 Outcomes? (e.g., date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	5.5
9.2.3 Covariates and other characteristics? (e.g., age, sex, clinical and drug use history, comorbidity, co-medications, lifestyle)	☒	☐	☐	5.5
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g., WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	5.2.1, 5.5
9.3.2 Outcomes? (e.g., ICD, Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	5.2.2, 5.5
9.3.3 Covariates and other characteristics?	☒	☐	☐	5.2.1, 5.2.3, 5.5
9.4 Is a linkage method between data sources described? (e.g., based on a unique identifier or other)	☒	☐	☐	7.1

Comments:

Separate subsections for each country.

Section 10: Analysis plan	Yes	No	N/A	Section Number
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	6.2

Section 10: Analysis plan	Yes	No	N/A	Section Number
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	6.1
10.3 Are descriptive analyses included?	☒	☐	☐	6.2.3
10.4 Are stratified analyses included?	☒	☐	☐	6.2.7
10.5 Does the plan describe methods for analytic control of confounding?	☐	☒	☐	
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	6.2.8
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	6.2.7

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)	☒	☐	☐	7.3
11.2 Are methods of quality assurance described?	☒	☐	☐	7.4
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations		Yes	No	N/A	Section Number
12.1 Does the protocol discuss the impact on the study results of:					
12.1.1 Selection bias?		☒	☐	☐	6.2.9
12.1.2 Information bias?		☒	☐	☐	6.2.9
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).					6.2.9
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)		☒	☐	☐	6.1

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?	☒	☐	☐	9
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	
13.3 Have data protection requirements been described?	☒	☐	☐	7.3, 9.2

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?	☐	☐	☒	

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)?	☒	☐	☐	7.5
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	7.5

Comments:

Name of the primary investigator: _____

Date:

Signature: _____