Redacted protocol

Incidence of Pancreatic Malignancy and Thyroid Neoplasm in Type 2 Diabetes Mellitus Patients who Initiate Exenatide Compared to Other Antihyperglycemic Drugs—Phase 2 (Extended Accrual and Follow-Up)

Redacted protocol based on the original final protocol of 18 December 2015

Table of Contents

1 BACKGROUND

1.1 Background

Exenatide is a glucagon-like peptide-1 (GLP-1) receptor agonist, approved by the United States (US) Food and Drug Administration (FDA) for the treatment of type 2 diabetes (T2D). Exenatide BID (Byetta) was approved on 28 April 2005 and is self-administered via subcutaneous injection twice daily. Exenatide once weekly (Bydureon) is administered through subcutaneous injection once weekly and was approved on 27 January 2012. Exenatide facilitates glucose control through enhancement of glucose-dependent insulin secretion by pancreatic beta cells, reduction of gluconeogenesis via suppression of excess glucagon secretion, and slowing of gastric emptying [1].

There are several studies of acute pancreatitis or pancreatic cancer in association with use of GLP-1 receptor agonists. Acute pancreatitis was reported among users of exenatide, principally through post-marketing spontaneous reports [2,3]. Subsequent epidemiologic investigations had conflicting findings. Some found no association between use of exenatide and acute pancreatitis [4,5], while another found an excess risk of acute pancreatitis among users of exenatide or sitagliptin [6]. Results of the latter study were not reported for exenatide alone. Little epidemiologic information exists on the potential association between incretin-modulating therapies and pancreatic cancer. A recent study by Funch et al. reported no excess risk of pancreatic cancer among users of liraglutide relative to users of other antihyperglycemic therapies [7]. A study using Medicare claims data observed no increased pancreatic cancer risk for patients with sitagliptin relative to thiazolidinedione (TZD) and sulfonylurea (SU) and for patients with exenatide relative to long acting insulin [8].

In rodent carcinogenicity studies, a statistically significant increase in thyroid C-cell tumor incidence (adenomas and/or carcinomas) was observed with exenatide once weekly, which is similar to changes described for liraglutide leading to concern about effects of GLP-1 pharmacotherapy on medullary thyroid cancer (MTC) [12]. In contrast to the effects on thyroid C-cells in the rat model, chronic monkey studies with GLP-1 agonists have not identified changes in thyroid C-cells. These data are consistent with those reported for liraglutide, in which no effects on thyroid C-cells were noted in monkey studies of up to 87 weeks [12-14]. These studies included a quantitative assessment of thyroid C-cells in a 52-week study. The available data on GLP-1 receptor expression suggests strong expression in the thyroid of rats and mice and little to no expression in humans or monkeys [13-16]. Furthermore, calcitonin (MTC is a C-cell cancer that produces excessive calcitonin [9]) concentrations were unaffected in humans with diabetes following up to 2 years of clinical exposure to exenatide once weekly or liraglutide. Data from clinical studies and post-marketing experience showed no increased risk of thyroid malignancy in general, and no cases of medullary thyroid cancer have been reported with either exenatide formulation.

Optum has completed a cohort study with patients accrued from 01 June 2005 to 31 July 2010 to estimate the association between use of exenatide BID and the occurrence of pancreatic cancer or thyroid neoplasms [17]. There were few outcomes, and the resulting lack of statistical power prevented satisfactory interpretation of the results. The current study is an extension with more recent data and a pooled analysis of the ORD and the Impact National Benchmark Database to increase statistical power.

2 OBJECTIVES

2.1 Primary Objectives

The primary objectives are to estimate the absolute and relative incidence of pancreatic cancer and thyroid cancer that occurs at least one year after initiation of exenatide BID or once weekly (hereafter exenatide) or initiation of other antidiabetic drugs (OADs)—overall and by duration of follow-up and duration of exposure.

2.2 Secondary Objectives

The secondary objectives are to estimate incidence rates (IRs) of benign thyroid neoplasm, MTC, and non-MTC neoplasms that occur at least one year after initiation of exenatide or OADs.

3 METHODS OF DESIGN

3.1 Overview of Study Design

This project is a retrospective cohort study that compares IRs of pancreatic cancer and thyroid neoplasm between initiators of exenatide and initiators of OADs using 2 administrative databases from commercial health plans in the US. In this extension, Optum will recreate the study cohorts to include patients accrued from 01 June 2005 through 31 December 2015. Initiators will be matched 1:1 or 1:2 (exenatide:OAD) on propensity scores within 6-month calendar blocks.The matched cohorts, when aggregated, will form the analytic population. The analyses of outcomes will account for the source databases and matching ratios through statistical conditioning. In a new analysis, data from the 2 databases will be combined to increase statistical precision.

Pancreatic cancer and thyroid neoplasm will be identified via patterns of claims using the “restricted” algorithms applied in the previous study. A second validation of the restricted algorithms will be conducted within a sample of medical records of persons in the ORD. Clinical characteristics that are captured poorly in the claims data will be abstracted from the medical records. Estimation of effects will involve time-fixed and time-dependent, cumulative classifications of exposure. A nested case-control analysis will be performed to account for potential confounders that are captured poorly in the claims data, if sample size allows.The results will be presented as follows:

- Combined person-time from the ORD and the Impact Database:
  - Cohort analysis of pancreatic cancer
  - Cohort analysis of thyroid cancer
  - Cohort analysis of the thyroid neoplasm subgroups
  - Sensitivity analyses, excluding nested-case control analysis
- ORD cohorts only
  - Cohort analysis of pancreatic cancer
  - Cohort analysis of thyroid cancer
  - Nested case-control analysis (malignancies only)
- Impact Database cohorts only
  - Cohort analysis of pancreatic cancer
  - Cohort analysis of thyroid cancer

The specific tasks to accomplish for the study objectives are:

1. Identify a cohort of T2D patients who initiated exenatide and characterize them in terms of demographics, diagnoses, medical procedures, drug use, and health care services utilization.
2. Identify a contemporaneous cohort of T2D patients who initiated an OAD and characterize them in terms of demographics, diagnoses, medical procedures, drug use, and health care services utilization.
3. Match the exenatide cohort and the OAD cohort within 6-month calendar blocks on patterns of demographics, diagnoses, drug use, procedures, and health care utilization in the time preceding cohort entry.
4. Combine the matched initiators across calendar blocks in each database.
5. Combine the matched initiators from the 2 databases into one matched exenatide cohort and one matched OAD cohort, while keeping the original matching intact.
6. Follow each cohort, identify outcomes using the restricted algorithms.
7. Validate a sample of algorithm-identified outcomes during the extension period (01 August 2010 to 31 December 2015).
8. Estimate IRs of pancreatic and thyroid cancers, overall and by duration of follow-up, within each database and in the combined dataset.
9. Estimate IRs of pancreatic and thyroid cancers by cumulative duration and dose of exenatide, within each database and in the combined dataset.
10. Estimate IRs of benign thyroid neoplasm, MTC, and non-MTC neoplasms in the combined dataset.
11. Perform sensitivity analyses in the combined dataset.
12. Perform a nested case-control study within a subset of the ORD population if sample size allows.

3.2 Data Sources

The patients included in this study will be drawn from proprietary research databases containing eligibility and pharmacy and medical claims data for enrollees of commercial health plans in the US.

3.2.1 Optum Research Database

Optum has access to a proprietary research database containing medical claims, pharmacy claims, and laboratory results (for a subset) with linked enrollment information covering the period from 1993 to the present. For 2013, data relating to approximately 12.7 million individuals with both medical and pharmacy coverage are available. An additional 11.9 million enrollees with medical benefits only are available. The population is geographically diverse and representative of the US population on age from birth through age 64 years but under-represents people 65 years of age or older.

Claims for pharmacy services are typically submitted electronically by the pharmacy at the time prescriptions are filled. These claims include all outpatient prescription pharmacy services provided and covered by the health plan. Pharmacy claims data include drug name, dosage form, drug strength, fill date, days of supply, financial information, and de-identified patient and prescriber codes, allowing for longitudinal tracking of medication refill patterns and changes in medications.

Medical claims or encounter data are collected from all available health care sites (inpatient hospital, outpatient hospital, emergency room, physician's office, surgery center, etc.) for virtually all types of provided services, including specialty, preventive, and office-based treatments. Medical claims and coding conform to insurance industry standards. Claims for ambulatory services submitted by individual providers, e.g., physicians, use the CMS-1500 format. Claims for facility services submitted by institutions, e.g., hospitals, use the UB-82, UB-92, UB-04, or CMS-1450 formats. Medical claims include: multiple diagnosis codes recorded with the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis codes; procedures recorded with ICD-9-CM procedure codes, Current Procedural Terminology (CPT), or Healthcare Common Procedure Coding System (HCPCS) codes; site of service codes; provider specialty codes; revenue codes (for facilities); paid amounts; and other information. Typically, facility claims do not include medications dispensed in hospital.

Most pharmacy claims are added to the research database within 6 weeks of dispensing. After approximately 6 months following the delivery of services, the medical data are complete.

3.2.2 Impact National Benchmark Database™

The Impact National Benchmark Database is a comprehensive, de-identified US health care claims database that, similar to the ORD, is representative of the non-elderly, commercially-insured population in the US. The database contains inpatient, outpatient, and pharmacy claims, lab results, and enrollment information. More than 75% of patients in the database have both medical and pharmacy benefits and, on average, 25.1 months of enrollment and claims information; the annual attrition rate is roughly 15-25%. The data are collected from more than 46 health plans, covering 9 census regions. The Impact Database is a fully de-identified, HIPAA-compliant dataset. Membership in the Impact Database may overlap somewhat with membership in the ORD. Internal Optum processes are used to identify overlapping periods of membership prior to integration with the ORD so that subjects appearing in the ORD do not also appear in the Impact Database, After combining the 2 databases, the resulting dataset is de-duplicated, and therefore, subjects are not counted more than once in analyses.

3.3 Institutional Review Board / Privacy Board Approvals

This study will use de-identified insurance claims data and seek medical records for a subset of patients with identifiable information to confirm the diagnosis of potential cases and abstract clinical variables. Optum will seek a waiver of Patient Authorization for access to protected health information from a Privacy Board and approval from an Institutional Review Board. Confidentiality of patient records will be maintained at all times. All study reports will contain aggregate data only and will not identify individual patients or physicians.

3.4 Study Population

The study population will consist of patients with T2D who had at least 9 months of continuous enrollment in their health plan between 01 September 2004 and 31 December 2015. The initial date of cohort eligibility is 01 June 2005, the launch date of exenatide BID in the US.

Eligible patients will have:

- Complete medical and pharmacy benefits and at least 9 months of continuous enrollment in the health plan prior to the cohort entry date
- A diagnosis of T2D (ICD-9-CM 250.x0, 250.x2) during the 9-month baseline period, inclusive of the cohort entry date
- A dispensing of at least one antidiabetic drug other than the initiating drug during the 9-month baseline period, inclusive of the cohort entry date

Subjects will be excluded if they have:

- A dispensing of the same class of drugs as the initiating drug during the 9-month baseline period
- Claims associated with pancreatic or thyroid neoplasm (including benign and malignant neoplasms) during the 9-month baseline period
- A dispensing of dipeptidyl peptidase-4 (DPP-4) inhibitors/GLP-1 receptor agonists (including exenatide) during the 9-month baseline period, inclusive of the cohort entry date

Appendix I includes a list of OADs to be included in this study.

3.5 Study Cohorts

3.5.1 Initiators of Exenatide

Between 01 June 2005 and 31 December 2015, the pharmacy claims will be searched for exenatide dispensings. The date of cohort eligibility will be the date of the first dispensing of exenatide without an exenatide dispensing in the previous 9 months (exclusive of cohort entry date), but with at least one OAD dispensing in the previous 9-month period inclusive of the cohort entry date.

3.5.2 Comparison Cohort: OAD Initiators

A contemporaneous comparison cohort of new users of OADs will be identified in the same manner as the exenatide cohort. The date of cohort entry will be the date of the first dispensing of an OAD with no dispensing of the same drug, or another drug from the same class, in the previous 9 months (exclusive of cohort entry date). OAD initiators must have at least one dispensing of a different OAD in the previous 9 months (inclusive of cohort entry date).

For patients initiating multiple antidiabetic medications during the study period, users of exenatide will be chosen first, such that a person who initiates exenatide and metformin during the study period will be assigned to the exenatide cohort even if s/he initiates exenatide later than metformin. This hierarchical cohort selection will allow for attribution of exenatide-exposed person-time to the exenatide cohort in the primary analysis, because that analysis is time-fixed. This approach did not introduce appreciable immortal time bias in our previous work on exenatide [21].

3.5.3 **Matching of the Comparison Cohort to the Exenatide Cohort**

Within each database, each exenatide initiator will be matched within 6-month calendar blocks to up to 2 OAD initiators on the estimated propensity scores. Matching within blocks of calendar time will account for potential changes in exenatide prescribing behavior over time. Six-month blocks will be used based on the experience with the previous exenatide study that demonstrated that a shorter block would better capture the variability of exenatide prescribing behavior over time. Patients will be assigned to calendar blocks based on their index date (date of cohort entry). It is expected there will be up to 20 calendar blocks in each database.

The propensity score is the probability of being a member of the exenatide cohort, given membership in the study population and the covariate pattern. Matching on the propensity score results in exposed and unexposed cohorts with balanced covariate distributions. In this context, balanced means that the exposed and unexposed cohorts will have, in expectation, the same distribution of all covariates modeled to estimate the propensity score [22].

Within each calendar block, propensity scores will be estimated from baseline covariates described in Section 3.10 using logistic regression modeling. To fit the propensity score model efficiently within each calendar block, a number of steps will be automated, including initial estimation, covariate balance checking, and modification/re-estimation of the propensity scores as needed. A common set of variables will apply, including a set of variables defined *a priori* and the 200 most prevalent diagnoses, procedures, and drug dispensings identified in the baseline period. The following steps will be applied, with automated repeated estimation, as needed:

1. Identify a set of variables to be forced into the models. Include variables with the 10 highest univariate c-statistics within the specific time block, along with all variables specified *a priori* as clinically important.
2. From the remaining covariates, identify predictors of exenatide initiation via a stepwise selection process. Set the stepwise criteria initially as a p-value of 0.2 for model entry and 0.3 for model retention.
3. Estimate the propensity score via an unconditional logistic regression model.
4. Exclude (“trim”) patients who have the lowest 2% of propensity scores in the exenatide cohort or the top 2% of propensity scores in the OAD cohort. Trimming on extreme values of propensity scores may reduce residual confounding from unmeasured attributes of patients or their context of care [23].
5. Match each exenatide initiator to one or 2 OAD initiators on the estimated propensity score using a “greedy” matching algorithm. Greedy matching is a linear, sequential matching algorithm that will identify patients in the OAD cohort who have propensity scores similar to those in the exenatide cohort. This algorithm identifies matches from the pool of possible matches without replacement. When identifying 2 matches per exposed subject, the greedy algorithm finds the closest match for each exposed subject before returning and identifying the second match. Once an exenatide initiator has been

matched with 2 OAD initiators, the triplet is removed from further consideration [24]. Specifically, the algorithm matches exposed and unexposed persons iteratively identifying the closest matches first, where closest is defined as exenatide-OAD pairs who match on the estimated propensity score at the 8th decimal point. With subsequent iterations, the algorithm identifies matches with less precision, decreasing by one decimal point at each iteration (i.e., 8, 7, 6, etc. decimal points), ending once matches at 0.1 of the propensity score are identified. The matching procedure is “greedy” in the sense that it preserves sample size by accepting matches on calipers as wide as 0.1 of the propensity score but only after identifying all possible matches with greater precision. Thus, greedy matching balances residual bias that could be introduced through inexact matching with preservation of statistical power. The greedy matching algorithm has been used extensively, including within Optum, and its details have been published [25].

6. Calculate the standardized difference of the mean of each covariate comparing exenatide initiators with OAD initiators to evaluate the balancing performance of the propensity score matching. Covariate imbalance will be defined as an absolute standardized difference > 0.1 (difference between the 2 mean values of the covariate divided by the standard deviation) [26]. If necessary, modify the propensity score to reduce any imbalance of covariates found within each calendar block. Given the complexity of the covariate patterns and the intent to balance the covariates within each calendar block, the following steps will be taken within each block of time if necessary to resolve imbalance in covariates:
    a. Evaluate whether the imbalanced covariate is a potential confounder by calculating its empirical association with the outcomes among the unexposed. If not associated with the outcomes, this variable will be removed from the propensity score model.
    b. Re-estimate the propensity score with the updated covariate list and repeat trimming, matching, and balance checking.
    c. If there are new or remaining imbalanced covariates, repeat Steps 6a and 6b up to 10 times.
    d. If imbalanced covariates remain after Step 6c, add interaction terms of the imbalanced covariates and finer blocks of calendar time (quarterly or bimonthly blocks) to the propensity score model.
    e. If imbalanced variables remain after Step 6d, adjust for these variables in the final regression models for the association between exenatide use and outcome events.

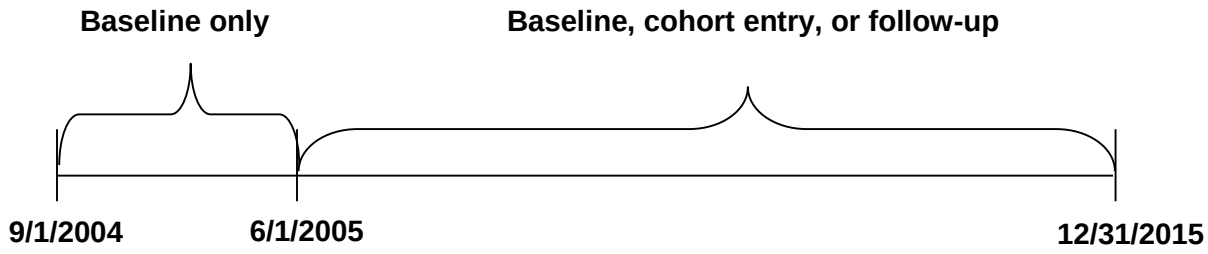
The propensity score matching will be performed within each database separately. Matched initiators will be compiled across calendar blocks to form the final analytic cohorts.

3.6 Follow-up

Follow-up time for each cohort member will extend from one year after study drug initiation until the first occurrence of a study outcome, disenrollment from the health plan (a gap of >32 days in membership), or the end of the study period (31 December 2015) (Figure 1).

Follow-up time in the OAD cohort will be censored at the time of a dispensing of DPP-4 inhibitors/GLP-1 receptor agonists (including exenatide).

Figure 1. Time frame for all cohorts



3.7 Exposure Definitions

3.7.1 Time-fixed Exposure Classification

For the primary analysis, exposure to exenatide or OADs will be defined based on the drug initiation that qualified each subject for cohort entry. That is, the exposure status at baseline will be carried forward for the duration of follow-up regardless of actual patterns of exposure.

3.7.2 Cumulative Exposure Classification

In secondary analyses, the measure of exenatide exposure after the initial dispensing will be a metric of cumulative duration and, separately, cumulative dose of exenatide received during follow-up. These measures will be handled as time-varying covariates during follow-up. Estimates of the daily dose will come from the number of units or volume dispensed, days supplied, and micrograms of exenatide per volume. Cumulative duration will be measured by summing the days supply of each dispensing over time. Exenatide duration will be calculated as the sum of all days supply for each user. The cumulative exenatide dose will be the sum of daily doses for all days supplied. For comparison, the cumulative duration and dose of OAD exposure will not be measured. Instead, an assumption will be made that use of OADs does not affect the study outcomes. For analyses of cumulative exposure, IRs will be calculated within categories of person time defined by cumulative dose or duration of exposure and compared to IRs in the OAD cohort using all exenatide unexposed person-time.

Person-time at-risk for each subject will begin one year after drug initiation. However, the first year of person-time will be quantified in cumulative exposure measures. The person-time of exenatide cumulative exposure will be distributed in a time-dependent fashion such that each person-day of follow-up will be attributed to the exposure category with the highest cumulative exposure (cumulative duration or dose) as of that day.

By design, the exenatide-unexposed person-time will be assumed to be exposed to one or more OADs. The use and switching status of OADs over time will be tracked for both exenatide and OAD initiator cohorts, so that concomitant OAD use can be adjusted for as a time-dependent covariate.

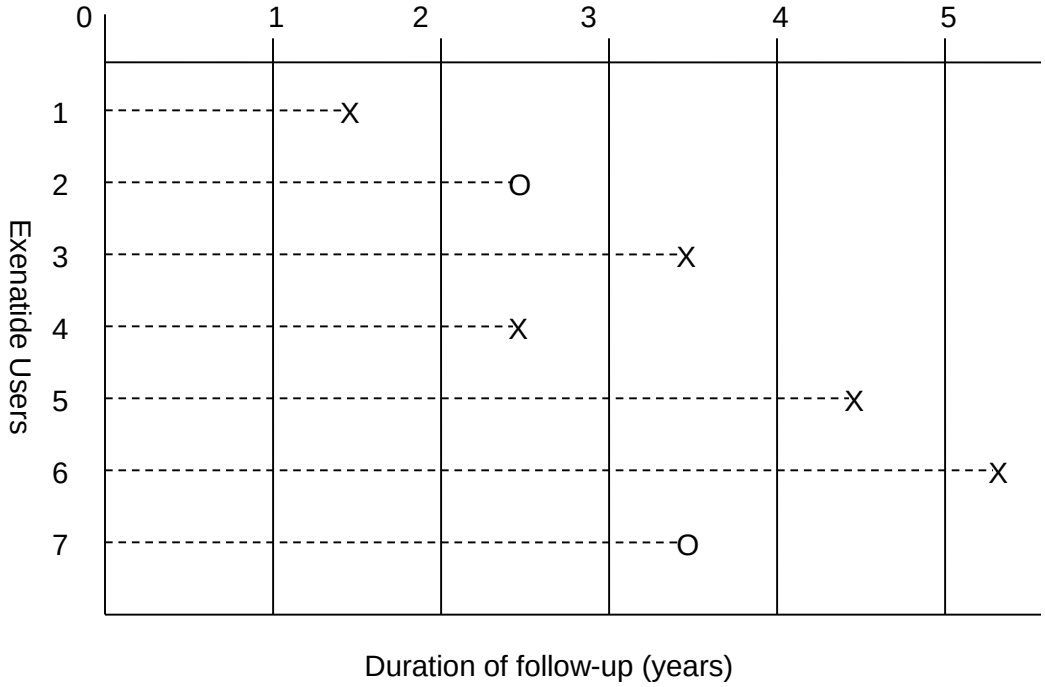
For example, consider a patient who initiates exenatide, and the first dispensing has a 30 day supply. The patient switches to another OAD and does not receive a second exenatide dispensing until 6 months after initial dispensing. The second dispensing, at 6 months, is a 30 day supply and is the last exenatide dispensing the patient receives during follow-up. Thus, the

person-time contribution for this patient will be 6 months at 30 days cumulative exposure and the remainder of follow-up time (approximately one year) will be 60 days cumulative exposure. In other words, this patient will contribute person-time to 2 cumulative exposure categories.

Events and person-time in the first year of follow-up will not be counted, since incidence will be evaluated beginning one year after drug initiation. Therefore, the 5 categories of post-initiation cumulative exposure duration are: 1 to <2 years, 2 to <3 years, 3 to <4 years, 4 to <5 years, and 5+ years. A patient can contribute person-time to multiple exposure duration categories [27]; allowing for the assessment of risk that may change with increasing time since exenatide initiation. The cumulative exposure duration categories will be modeled as a time-dependent covariate in the analysis.

Examples of cumulative exenatide exposure duration are given in Figure 2, where a patient is followed until a study outcome (X), or censoring (O). In this example, Person 1 is followed for 1.5 years total and contributes 0.5 person-years to the category of 1 to <2 years of exenatide exposure. Similarly, Person 2, whose overall follow-up is 2.5 years, contributes one person-year to the category of 1 to <2 years and 0.5 person-years to the category of 2 to <3 years of exenatide exposure. The total person-time across these 7 exenatide users is 6.5 person-years in the category of 1 to <2 years, 5 person-years in the category of 2 to <3 years, 3 person-years in the category of 3 to <4 years, 1.5 person-years in the category of 4 to <5 years, and 0.5 person-years in the category of 5+ years.

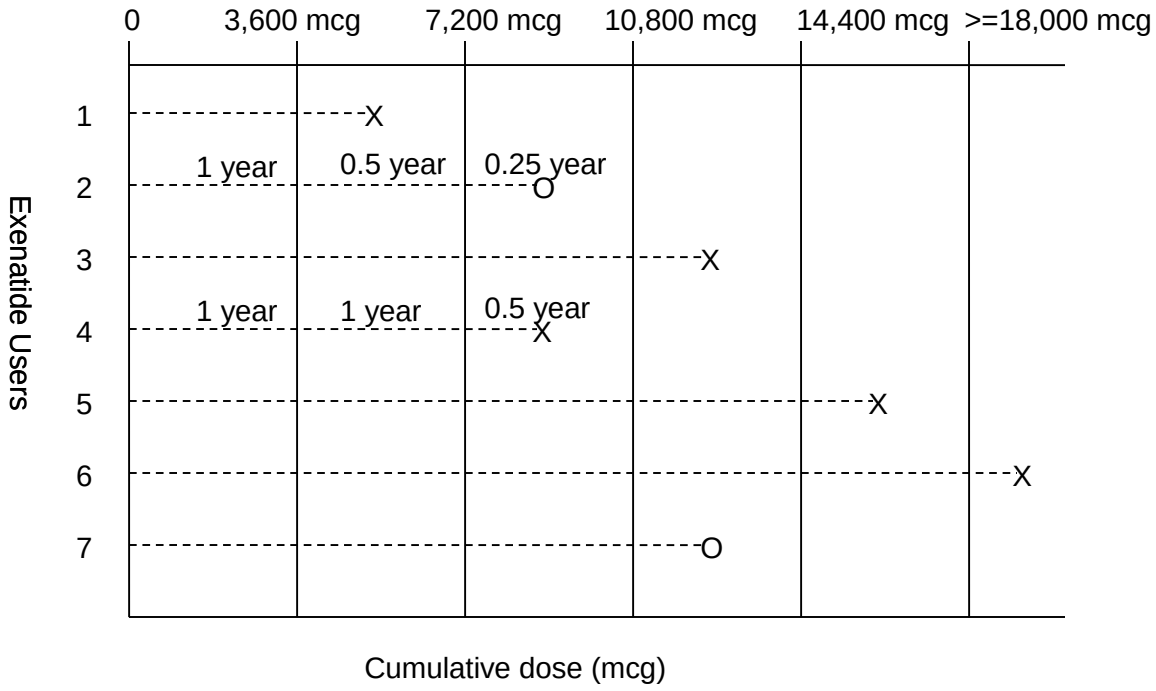
Figure 2. Schema Illustrating the Allocation of Person-years of Exenatide Exposure by Duration of Follow-up. X, Outcome of interest; O, Disenrollment



The cumulative of exenatide dose, operationally, will be a function of the exenatide formulation. Exenatide BID has 2 formulations: separate pens for 5 micrograms (mcg) and 10 mcg per dose, typically administered twice daily. Exenatide once weekly is available in a single type of pen designed to deliver a weekly dose of 2 mg. Each pen typically lasts 30 days (60 doses) so that a person generally receives 300 mcg per month of the 5 mcg per dose pen or 600 mcg per month of the 10 mcg pen. As above, person-time within the first year of follow-up will be classified as not-at-risk, yet exenatide dosage will be summed during this period. The dose of exenatide once weekly, 2 mg per week or 8 mg per month, will be assumed equivalent to 10 mcg of exenatide BID.

Cumulative exenatide dose will be calculated at each dispensing, summed across all dispensings, and categorized. The 5 categories of post-initiation cumulative exenatide dose are: 3600 to < 7200 mcg, 7200 to <10,800 mcg, 10,800 to 14,400 mcg, 14,400 to <18,000 mcg, and 18,000+ mcg. Examples of how cumulative exenatide dose is calculated are given in Figure 3, where a patient is followed until a study outcome (X), or censoring (O). In this example, Person 2 receives 10 mcg exenatide per day (i.e., 5 mcg BID), for the first year, and 20 mcg per daily thereafter (i.e., 10 mcg BID), and is followed for a total of 1.75 years. This person will contribute one year to the category of <3,600 mcg (on 5 mcg BID), 0.5 years to the category of cumulative dose 3,600 to <7,200 mcg (on 10 mcg BID), and 0.25 years to the category of cumulative dose 7,200 to <10,800 mcg (on 10 mcg BID). Person 4 receives 5 mcg BID and is followed for a total of 2.5 years. This person will contribute one year to the category of cumulative dose 3,600 to <7,200 mcg and 0.5 years to the category of cumulative dose 7,200 to <10,800 mcg. This metric of exposure will allow for the evaluation of risk that may change with increasing dose of exenatide use.

Figure 3. Schema Illustrating the Allocation of Person-years of Exenatide Exposure by Cumulative Dose. X, Outcome of interest; O, Disenrollment



3.8 Outcome Identification

Outcomes of pancreatic cancer and thyroid neoplasms will be ascertained from the claims data using the restricted algorithms described in the Optum report of 25 July 2013. Appendix II contains the ICD-9-CM codes, and Appendix III includes the detailed algorithms.

3.9 Covariates

Covariates derived from claims data will include baseline characteristics including demographics, diagnoses, medical procedures, drug use, and health care utilization.

All members of the study cohorts will be classified according to covariates, including the following:

- Demographics
  - Age, sex
  - Geographic area
  - Cohort entry year
- Diabetes severity indicators
  - Use of oral antidiabetic medication
  - Dispensings of one, 2, or 3 study medications within 45 days of cohort entry
  - Peripheral neuropathy
  - Nephropathy
  - Retinopathy

- Cardiovascular disease indicators
  - Hypertension
  - Hyperlipidemia
  - Hypertriglyceridemia
  - Ischemic heart disease
  - Myocardial infarction
  - Congestive heart failure
  - Stroke
- Other
  - Health care utilization (e.g. the number of days hospitalized in prior 9 months, hospitalization within 45 days of the cohort entry date, number of physician visits, emergency department visits and costs of facility and pharmacy, etc.)

Predictors of exenatide initiation will be empirically identified by listing the 200 most prevalent drug classes dispensed to exenatide initiators along with the 200 most prevalent diagnoses (at the 3-digit ICD-9-CM level) and 200 most prevalent procedures. The prevalence of these variables will be tabulated by exposure.

The cohorts will be further characterized with respect to a number of healthcare utilization variables, including total and drug-specific costs, the number of unique ICD-9-CM diagnoses, and the number of unique drugs dispensed.

3.10 Nested Case-control Study

If there is a sufficient number of chart-confirmed cases available (Section 3.11.1), a nested case-control analysis will be conducted for pancreatic cancer and for thyroid cancer separately to account for potential confounders that are captured poorly in the claims data. The cases will consist of all chart-confirmed cases of pancreatic and thyroid cancer, separately, from the previous validation study (01 June 2005 - 31 July 2010) and the extension validation study (01 August 2010 – 31 December 2015). The validation studies occurred (or will occur) among the patient-identifiable subsets of the ORD. Controls will consist of sampled person-days from the at-risk follow-up experience of exenatide and OAD cohort members. Controls will be matched to cases within calendar time blocks. The boundaries of the calendar block for each risk set will be determined according to the observed distribution of cases across calendar time. Both cases and controls will be required to meet the inclusion and exclusion criteria below.

- Inclusion criteria, case-control analysis
  - Cases: a chart-confirmed diagnosis of pancreatic or thyroid cancer during the period of 01 June 2005 through 31 December 2015
  - Controls: a random sample of the source cohorts and no chart-confirmed diagnoses of pancreatic or thyroid cancer (separately applied for each outcome) at the time of the case occurrence

- Exclusion criteria, case-control analysis

- Controls: a diagnosis of benign neoplasm of the thyroid or pancreas, separately applied for each outcome, preceding the given risk set time.
- Cases and controls: no charts available for review, or a history of cancer excluding non-melanoma skin cancer

We will identify up to 4 controls for each case. Controls will be randomly sampled cohort members at-risk during the risk set time block and matched on the number of baseline visits, age, and sex. Additionally, we will frequency match cases and controls within each risk-set to balance the distribution of charts from hospitals versus outpatient facilities. Different types of facilities may record different types of information about patients. The frequency matching will mitigate differences in covariate detail in medical records between cases and controls.

Controls sampled within risk-sets (time blocks) are matched to cases on sampling time, thus representing time at-risk for the outcome when the case occurred. Operationally, the analyst will assign cases an index date equal to the chart-confirmed cancer date. Controls will be assigned an index date corresponding to the date of the first office visit within the same risk-set block. We will explore the feasibility of matching cases to the controls on the type of office visit. The exposure variable will represent patients' exposure status at cohort entry so that the case-control effect estimate will correspond to the intent-to-treat analog planned for the cohort analysis.

Data will be abstracted from the charts of cases (up to 40 charts for each outcome) and controls (up to 160 charts) for up to 20 clinical characteristics that may represent unmeasured confounders in the underlying cohort analysis or that may inform the multiple imputation analysis described below. The variables include:

- Race/ethnicity
- Height and weight/body mass index
- Smoking
- Alcohol use
- Blood pressure
- Family history of cancers including pancreatic and thyroid cancers
- Personal history of medical conditions, separately for pancreatic cancer and thyroid cancer (e.g., pancreatic and thyroid diseases, gallstone, cholecystectomy, non-alcoholic fatty liver disease)
- Various lab results (e.g., A1C, C-reactive protein)
- Exposure to ionizing radiation and low-iodine diet (for thyroid cancer only)

3.10.1 **Decision Rule to Determine Feasibility of Case-control Analysis**

The feasibility of the case-control analysis with respect to sample size will be determined empirically according to the number of confirmed cases in the original validation study, the number of algorithm-identified cases in the extension period, and the expected number of controls. The estimates of number of cases from the extension period will include consideration of the positive predictive value of the algorithm in previous work [28]. For new cases and controls, the expected number will be reduced by 30% to account for incomplete medical record retrieval. The case-control study will proceed if the number of expected cases (i.e., n=80 for each outcome) and controls provides a statistical power of 80% for an odds ratio (OR) equal to

2. In the event that the number of cases exceeds 80, all confirmed cases (separately for pancreatic cancer and thyroid cancer) will be included in the nested case-control analysis and covariates will be abstracted from the medical records. If sufficent statistical power is unavailable, the case-control analysis will not be conducted. Instead, covariate information will be abstracted from the medical records of cases, and this information will be used to inform the range of prevalence of key confounders (e.g., smoking and obesity) within the sensitivity analyses for the cohort study.

3.11 Medical Record Abstraction and Adjudication

Although the same restricted algorithms applied in the previous study will be used to identify the study outcomes, medical records will be sought to confirm the diagnosis of the potential cases of pancreatic cancer and malignant and benign thyroid neoplasms that occurred during the extension period. Medical records are available for a subset of the ORD where investigators may access protected health information of enrollees, which is govered by several ethics approval processes.

Adjudication will occur via the methods described in the study protocol of 11 April 2012 of the BO15 study conducted by Optum for Eli Lilly and Company. One hundred fifteen medical records will be sought for the potential cases of pancreatic cancer or each thyroid neoplasm occurring after one year following drug initiation. Charts will be sought for each potential case identified by the relaxed algorithms as defined in the final report of 25 July 2013. To increase the total number of medical records available to evaluate the algorithms, cases will include patients who are matched and unmatched (i.e., excluded from the primary analyses) based on the propensity score. Medical record information related to case diagnoses will be abstracted for the 9 months preceding the date of the claim that identifed the potential case through the 2 months following the same date. Two adjudication panels will be formed with each comprised of 2 adjudicators: one with expertise in the clinical diagnoses of the outcome of interest (pancreatic cancer or thyroid neoplasm) and one with expertise in general oncology. The adjudication elements applied in the previous work will be applied in these analyses. Discrepant adjudications will be resolved by mutual consensus among the adjudicators. An Optum senior scientist/clinician will serve as a final arbiter in the event that the adjudicators cannot resolve any discrepancies. All of the reviewers will be blinded to exposure status.

Covariates (Section 3.11) will be abstracted that are captured poorly in the claims data for the potential cases with medical records available. The maximum number of charts to be sought for covariate abstraction will be 80 for cases (a portion of the 115 charts; no covariates will be abstracted for potential cases of benign thyroid neoplasm) and 160 for controls. Historically, Optum has obtained 70-85% of medical records requested. There were 55 chart-confirmed pancreatic and thyroid cancer cases in the original validation study and their covariates will be abstracted from the existing charts. No new records will be sought for these cases.

4 METHODS OF ANALYSIS

4.1 Cohort Formation

The number of patients who meet the inclusion and exclusion criteria will be documented for each database (Shell Figure 1 and 2). The number of patients retained and excluded after matching on propensity scores will be presented (Shell Figure 3) for each database and in the combined database. The number of patients in the analytic cohorts by duration of follow-up (<6 months, 6-12 months, and >12 months) will also be presented.

4.2 Description of Baseline Characteristics

Baseline characteristics including demographics, medical history (including individual malignancies), prescription drug history, and health care services from the 9-month baseline period (inclusive of cohort entry) will be tabulated for the exenatide and OAD cohorts in the ORD and in the Impact Database, separately. Continuous variables will be summarized by mean and standard deviation or median and interquartile range. Categorical variables will be summarized by frequency and percentage.

4.3 Medical Record Adjudication

The number of charts sought and received will be summarized. A measure of the sensitivity of case finding for each outcome will be calculated as the fraction of cases identified by the restricted algorithm among the chart-confirmed cases. Because the cases for chart confirmation were identified from a more sensitive “relaxed” algorithm, the chart-confirmed set provides a reasonable denominator for estimating the sensitivity of the restricted algorithm. Positive predictive values for each outcome will be calculated as the number of chart-confirmed cases divided by the total number of cases identified by the restricted algorithm for which charts are available. The sensitivities and positive predictive values in the validation studies of the previous period (01 June 2005 - 31 July 2010) and extension period (01 August 2010 - 31 December 2015) as well as the combined period (01 June 2005 - 31 December 2015) will be presented.

4.4 Common Data Model

To combine the ORD and Impact Database, a common data model will be created using an approach similar to that used within the FDA’s Mini-Sentinel program [29,30] and the Observational Medical Outcomes Partnership [31]. Because both databases are owned and operated by Optum, they exist in similar formats within Optum’s research database. In the creation of study analytic files, the data management and analyses (up to the point of the pooled analyses) will be run in parallel. The same specifications will be followed for variable definitions in each dataset to form a common data model at the patient level. No aggregate data are required for this analysis because both datasets reside within the same firewalls at Optum, and patient-level data can be pooled directly.

Table 1. Common Data Model

Common Data	Descriptions
Unique person identifier	Patient ID
Demographics	Age, gender, region, cohort entry year
Exposure	Exenatide and OAD
	Duration of exenatide and OAD use
	Cumulative duration of exenatide use and non-exenatide use
	Cumulative dose of exenatide use and non-exenatide use

Common Data	Descriptions
Covariates	Covariates to be used for propensity score estimation
Propensity score	Propensity score calculated given the covariate patterns
Matching ratio	1:1 and 1:2
Events	Pancreatic cancer, thyroid cancer, benign thyroid neoplasm, MTC, non-MTC
Person-time	Person years
Data source identifier	ORD and Impact Database

As part of Optum's concurrent management of the ORD and the Impact Database, overlapping patients are de-duplicated. The source of each patient's data will be indicated by a data source identifier.

4.5 Time-fixed Analysis

An "intent-to-treat" or "as-matched" analysis will be conducted that holds the original exposure assignment constant from the date of accrual through the end of follow-up. At-risk person-time will be accrued from one year post drug initiation until the earliest occurrence of an outcome (pancreatic cancer or thyroid neoplasm), health plan disenrollment (a gap of >32 days in membership), or 31 December 2015.

Follow-up time, starting one-year post drug initiation, will be summed and characterized as exposed to exenatide or OADs, in totality and stratified by duration of follow-up. The categories for duration of follow-up will be the same as the previous study: 1 to <2 years, 2 years to <3 years, and ≥3 years. The person-time for each type of newly diagnosed thyroid neoplasm will be calculated for the 2 cohorts by subgroup of thyroid neoplasm (e.g., benign tumor, MTC, and non-MTC neoplasms).

The IR of each outcome occurring at least one year following drug initiation will be calculated as the number of events divided by the sum of corresponding person-years at-risk in each cohort. Similar IR estimates of pancreatic and thyroid cancer will be calculated in the subgroup of patients with and without concurrent use of insulin. Concurrent use of insulin and exenatide will be defined as the use of insulin within 32 days before or after cohort entry. Any insulin use outside of 32 days will be defined as non-concurrent use of insulin.

Kaplan-Meier plots will be used to depict the cumulative probability of event-free time. Cox proportional hazards regression models will be used to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) of newly diagnosed pancreatic cancer and thyroid cancer among exenatide initiators compared with OAD initiators, in totality and by duration of follow-up. In the event that the propensity score matching does not fully balance potential confounders, the Cox models will adjust for these covariates. Any covariate that is residually imbalanced in any stratum will be added to all models up to the limit of 10 outcomes per covariate.

The IRs and HRs of pancreatic and thyroid cancers will be estimated in the combined dataset, in the ORD only, and in the Impact Database only. Person-level data will be directly pooled using Cox regression (time-fixed analyses) or Poisson regression models (cumulative exposure analyses). The models of pooled data will account for database of origin by conditioning on the indicator of database membership.

The database-specific estimates will be evaluated to determine the suitability of the data for pooling. Interaction terms in the primary regression models of the database membership indicator and exposure at baseline will be tested to assess the heterogeneity of estimates across databases. A p-value smaller than 0.05 will indicate a potential departure from homogeneity of effects across databases. The evaluation of heterogeneity will also be based on clinical rationale, the distributions of patient characteristics in the 2 datasets, and consideration of the similarities of the 2 data sources [33]. If the results across databases are clearly heterogeneous, the interpretation of the pooled data may be deemphasized [32]. The results of both the pooled analysis as well as the database specific estimates will appear in the report. Estimation for subgroups of thyroid neoplasm and subgroups of concurrent users of insulin will occur in the combined dataset only.

4.6 Analysis of Cumulative Exposure

To estimate the cumulative effect of exenatide on the outcomes, an analysis will be conducted of the cumulative duration and cumulative dose of exenatide use. The general classification schemes described in Section 3.8.2 will apply. Among the exenatide cohort, follow-up time of each person will be assigned to corresponding categories of cumulative duration (non-use, 0 to <1 year, 1 to <2 year, 2 to <3 years, and ≥3 years) and to appropriate categories of cumulative dose (non-use, 0-1,499 mcg, 1,500-5,999 mcg, and 6000+ mcg).

Multivariable Poisson regression models will be used to estimate crude and adjusted relative rates (RRs), and 95% CIs for each outcome comparing different categories of cumulative duration and cumulative dose of exenatide use to non-use of exenatide (principally current use of OADs). Indicators of cumulative duration and dose will be time-dependent covariates in the model. To address potential residual confounding introduced when patients switch drug regimens, indicators of OAD use will also be time-dependent covariates in the Poisson regression models. Lastly, age will be included as a time-dependent covariate, ascertained at each apparent change to OAD exposure.

Importantly, the time-dependent adjustment variables (age and OAD use) are assumed to not be mediators or colliders; otherwise, the estimates could be biased. We expect that age fulfills this assumption. While age is associated with exenatide use and pancreatic or thyroid cancer, it is not on the causal pathway between exenatide and those outcomes. Age is not a collider because both exenatide and cancers do not directly cause increased age. The assumption is stronger for OAD exposure, but following the precedent of this project, OAD use will be assumed to be a non-mediator and non-collider. To evaluate the possibility that OADs are a mediator, each OAD use indicator during follow-up will be introduced into the Poisson model sequentially and results from each step will be reported. Changes in the effect estimates toward the null may indicate confounding or OAD-mediation (i.e., that conditioning on the mediator blocks a portion of the causal path between exenatide exposure and the outcome). Regarding OADs as colliders, the use of exenatide may cause concomitant use of OADs, but the development of cancers is less likely to cause concomitant use of OADs (with the possible exception of protopathic effects of pancreatic cancer). Therefore, collider bias (i.e., selection bias) is not anticipated.

The cumulative Poisson regression models are represented by the following equation:

$$\log\left[E\left[d_{jk}\middle|x_{jk}\right]\right] = \log\left[n_{jk}\right] + \alpha_j + \beta_1 x_{jk1} + \beta_2 x_{jk2} + \cdots + \beta_q x_{jkq}$$

n_{jk} is the number of person-years of follow-up observed among patients in the j[th] stratum who are in the k[th] exposure category.
d_{jk} is the number of events observed in these n_{jk} person-years of follow-up.
$x_{jk1}, x_{jk2}, \ldots, x_{jkq}$ are explanatory variables (e.g. age, OAD dynamic status, or time-fixed covariates) that describe the k[th] exposure group of patients in stratum j.
$\alpha_{1,\ldots,}\alpha_j$ are unknown nuisance parameters.
$\beta_1, \beta_2, \ldots, \beta_q$ represent cumulative exposures of interest.

The following table is an example of the data format for the Poisson regression models with time-varying covariates:

Table 2. Data Format for the Poisson Regression Models with Time-varying Covariates

ID	Cumulative duration of exenatide	Initiating OADs[ǂ]	Person years	Outcome	Age	MET	SU	TZD	Non-SU	PA	Alpha-GI	Insulin
1	<1 year *	0	0	0	59	1	1	0	0	0	0	0
1	1 to <2 years	0	0.5	1	59.5	0	1	0	0	0	0	0
1	1 to <2 years	0	0.5	1	59.5	0	0	1	0	0	0	0
2	<1 year*	0	0	0	45	0	1	0	0	0	0	0
2	1 to <2 years	0	1	0	46	0	1	0	0	0	0	0
2	2 to <3 years	0	4	0	50	0	0	1	0	0	0	0
3	<1 year*	0	0	0	64	1	0	0	0	0	0	0
3	1 to <2 years	0	1	0	65	0	0	0	0	0	0	1
3	2 to <3 years	0	1	0	66	0	0	0	0	0	0	1
3	>=3 years	0	0.5	1	66.5	0	0	0	0	0	0	1

MET=Metformin; SU=Sulfonylureas; TZD=Thiazolidinediones; Non-SU=Non-sulfonylurea Secretagogues; PA=Pramlintide Acetate; Alpha-GI=Alpha-Glucosidase Inhibitors.
*The person time within the first year following drug initiation will be set to zero.
[ǂ]To keep the data format in a simple fashion, the OAD use is categorized to unexposed group.

A trend test for RRs across different categories of cumulative duration and dose will be examined to explore the potential dose-response relationship.

These analyses will be conducted in the combined dataset, in the ORD only, and in the Impact Database only.

Appendix IV includes a summary of the statistical analyses.

4.7 Sensitivity Analysis and Detection Bias Evaluation

A sensitivity analysis will be conducted to evaluate potential bias from excluding events that occur within the first year of drug initiation by repeating the analysis with exclusion of events occurring within the first 6 months after study entry (i.e., including events that occur 6 to 12 months after initiation). Cases from the first 6 months of follow-up will be assumed to reflect outcomes that cannot be affected by recent antidiabetic drug initiation, although these rates may be shown for comparative purposes. Any changes in estimates will be observed by including cases from the 6 to 12 months of follow-up. Specifically, the duration of follow-up will be stratified in the same fashion as in the previous study (>6 months to <1 year, 1 to <2 years, 2 to <3 years, and ≥3 years).

In order to verify the assumption that the use of OADs has not affected the study outcomes, the following sensitivity analyses will be conducted. First, we will compare exenatide with OAD within comparable categories of cumulative dose and duration (e.g., 1-2 years of exenatide use versus 1-2 years of OAD use). For the analysis of cumulative dose, the categories of OAD exposure will be defined according to the distribution of cumulative dose (e.g., tertile or quartile). Exenatide initiators will be grouped in the same way based on the distribution of cumulative dose (e.g., tertile or quartile) of exenatide use. For the analysis of cumulative duration, the cumulative duration of each initiated OAD will be measured by summing the days supply of each dispensing over time. Both OAD and exenatide initiators will be categorized using the categories of cumulative duration that have been defined in the study protocol (e.g., 1-2 years, 2-3 years, ≥3 years). Second, we will compare higher with lower cumulative dosage and duration categories of OAD exposure defined according to the distribution of cumulative dose and duration.

Covariate information will largely be inferred from claims for medical services. However, some important covariates (e.g. smoking and obesity) are captured poorly by claims data, and the inability to adjust for them could result in bias. In order to evaluate the impact of these variables on the observed estimates, a sensitivity analysis will be incorporated to assess the effect of an unmeasured confounder across a range of plausible prevalences and associations with both exposure and outcome using the method provided by Schneeweiss et al [34]. The data from Amylin's pancreatitis project will be used where covariates (e.g. smoking and obesity) were ascertained from chart review. The current project and Amylin's pancreatitis project have similar study populations and the same study exposure, so that the prevalence of these covariates can be estimated from the Amylin project [35]. However, the potential for unmeasured confounding still remains even outside such assumed characteristics. This sensitivity analysis will supplement the nested case-control analysis in that the covariate information for the case-control study will correspond to the follow-up experience of patients, whereas this sensitivity analysis will use baseline values of covariates.

Detection bias can occur when a disease outcome is subject to differential diagnosis (overdiagnosis or underdiagnosis) induced by exposure or an uncontrolled correlate of exposure. The result can be an exaggerated or attenuated association [36]. In order to detect the presence of detection bias in this study, the frequency of physician visits and diagnostic procedures related to the thyroid or the pancreas during the baseline period and the follow-up period will be tabulated. If there is a signal indicative of a potential higher rate of detection in the exenatide cohort, the impact of this potential detection bias on the observed RR estimates will be evaluated using the method presented by Greenland and Neutra [36].

These sensitivity analyses will be performed in the combined dataset only.

4.8 Nested Case-control Analysis

In the nested case-control analysis, the RRs (as ORs) of the malignant outcomes comparing the exenatide and OAD comparator cohorts will be estimated. Baseline characteristics including demographics, claims-based medical history and prescription drug history, and health care services from the 9 months prior to the case-control index date and from medical records will be tabulated for the cases—separately for pancreatic and thyroid cancer—and corresponding controls (Shell Tables 11.1, 11.2, 12.1 and 12.2). The claims-based characteristics that are included in propensity score models will be described for cases of pancreatic cancer or thyroid cancer, and their controls. The chart-based covariates will include common biometrics (e.g., BMI, blood pressure) and disease-specific variables (e.g. pancreas-related diseases for pancreatic cancer and thyroid-related diseases for thyroid cancer). If there is any difference in the characteristics ascertained from claims data and medical records, the data abstracted from the medical records will be used. The chart-derived covariates will be described among the controls by exenatide and OAD use, because the control series is representative of the source population. This description will provide information about confounding in the cohort analysis by displaying any differences in covariate distributions between the exenatide and OAD comparison groups.

Missing data obtained from medical records is expected and may vary by covariate. In our study of exenatide and acute pancreatitis study, which had a similar design, 48% of patients had missing information on overweight/obesity, 12% on smoking status, and 15% on alcohol use. To account for missing data, we will apply a multiple imputation technique based on Markov Chain Monte Carlo methodology [37]. Should there be computational limitations of the Markov Chain approach, we will adopt a chained equation approach. Both methods assume data are missing at random (MAR) and improve the validity and/or efficiency of analyses of missing data when the assumption holds. Given that the MAR is not testable in the source dataset, we will include a description of the proportion of missing values, the proportion of missing information, and patterns of missing variables. This information can inform whether the MAR assumption is plausible. Also, Collins, Schafer, and Kam (2001)[38] demonstrated that in many realistic cases, a departure of MAR assumption has only a minor impact on estimates and standard errors. Note that the MAR assumption of multiple imputation is less stringent than the missing completely at random assumption of a complete-case analysis (including only patients with complete data)[39]. Multiple imputation also appropriately provides standard errors that account for statistical uncertainty from missing data.

In this study, imputation will be based on patterns of non-missing values of those variables with missing data, the variables most predictive of exenatide initiation in the propensity score modeling, exposure status, and demographics. The imputation process will involve 10 imputations and will create 10 complete analytic datasets. Uncertainty is accounted for by creating 10 imputation sets and observing the variability across the 10 datasets. Each dataset will be analyzed separately, accounting for the matched design and incorporating potential confounders derived from claims data and medical records using conditional logistic regression modeling. By using the risk set sampling approach, the generated ORs will estimate the RRs. The ORs and covariance matrices of each analytic dataset will then be combined to produce estimates and CIs that incorporate missing-data uncertainty [38]. The results based on the imputed data will be compared with the results based on the raw data.

Appendix IV includes a summary of the nested case-control statistical analysis.

5 SAMPLE SIZE AND STUDY POWER

5.1 Cohort study

The FDA requested that this extension study be conducted. As a result, this work has been initiated as requested without a formal evaluation of sample size and statistical power. The number of patients available for analysis through 11 May 2013 was previously quantified, and this information is presented in the remainder of this section.

In the combined ORD and Impact Database dataset, there were 81,788 individuals who had at least one dispensing of exenatide BID or once weekly from 01 June 2005 through 11 May 2013 and who had at least 9 months continuous enrollment prior to the first dispensing (Table 2).

Based on previous work, Optum estimates that 83% of the 81,788 exenatide initiators will meet the inclusion and exclusion criteria, so that 67,884 individuals will remain eligible. Six percent are expected to have a dispensing of DPP-4 inhibitors/GLP-1 agonists prior to or on the first dispensing of exenatide, reducing the number of individuals to 63,811. It is expected that 79% (50,410) of exenatide initiators will be matched to up to 2 OAD initiators after excluding patients with extreme values of propensity scores. Among the remaining 50,410 individuals, 30% will be excluded from the analysis after restricting to those with at least one year of follow-up, giving a final cohort of 35,287 exenatide initiators. The average follow-up time per person is about 1.7 years estimated from the 25 July 2013 report, giving a total of approximately 60,000 person-years.

To estimate statistical power, Optum assumed that the IR of pancreatic cancer is 0.2 per 1,000 person-years and the IR of thyroid cancer is 0.3 per 1,000 person-years in the OAD cohort and that the OAD cohort will be 1.5 times larger than the exenatide cohort (estimated from the previous report). Figure 4 shows the potential power for pancreatic cancer analysis to detect a range of RRs at 60,000 person-years of exenatide exposure in a future study with accrual through May 2013 using the ORD and Impact Database. The future study would have 89% power to detect a RR=2.5 and 98% power for a RR=3, but only 24% power for a RR=1.5 overall. Figure 5 shows the potential power for the thyroid cancer analysis to detect a range of RRs with 60,000 person-years of exenatide exposure. The study with this much exenatide exposure would be feasible in the future and have an estimated 79% power to detect a RR=2.0 and 98% power for a RR=2.5, but only 33% power for a RR=1.5.

The final sample size is expected to be slightly larger with the study period extending to 31 December 2015.

Table 2. Exenatide* Initiators in Optum Research Database and Impact Database from 01 June 2005 to 11 May 2013

Calendar Year	Optum Research Database		Impact Database		Total	
	N	%	N	%	N	%
2005	3,965	7.3	3,358	12.3	7,323	9.0
2006	13,768	25.3	8,060	29.4	21,828	26.7
2007	11,111	20.4	6,631	24.2	17,742	21.7
2008	9,110	16.7	3,639	13.3	12,749	15.6

Calendar Year	Optum Research Database		Impact Database		Total	
	N	%	N	%	N	%
2009	5,594	10.3	2,168	7.9	7,762	9.5
2010	3,140	5.8	1,171	4.3	4,311	5.3
2011	2,649	4.9	1,008	3.7	3,657	4.5
2012	3,723	6.8	1,343	4.9	5,066	6.2
2013**	1,350	2.5	-	-	1,350	1.7
Total	54,410		27,378		81,788	

* Includes both exenatide BID and exenatide once weekly.

** Data available until 11 May 2013 in the ORD but unavailable in the Impact Database.

Figure 4. Study Power for Pancreatic Cancer among All Study Patients

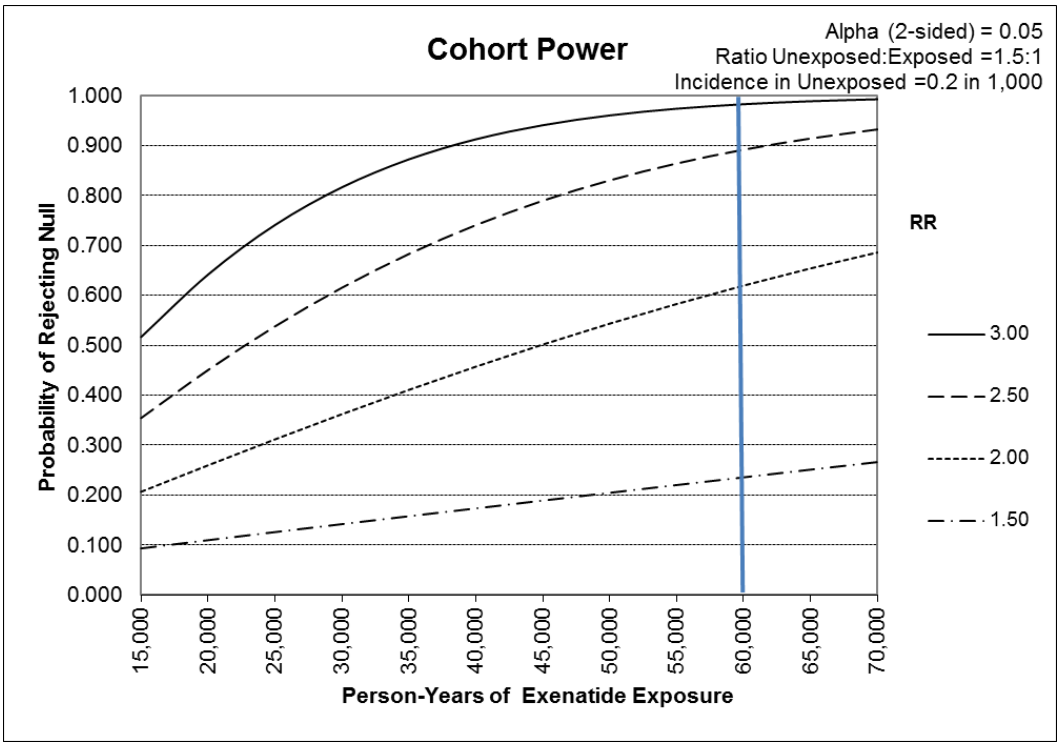
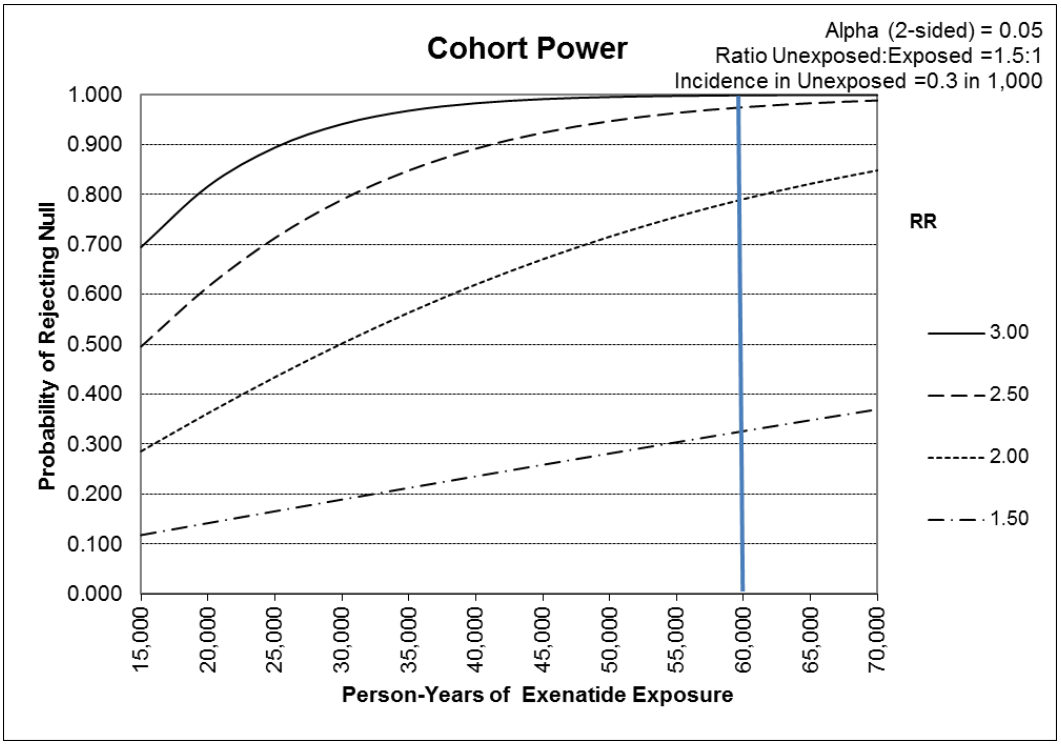


Figure 5. Study Power for Thyroid Cancer among All Study Patients

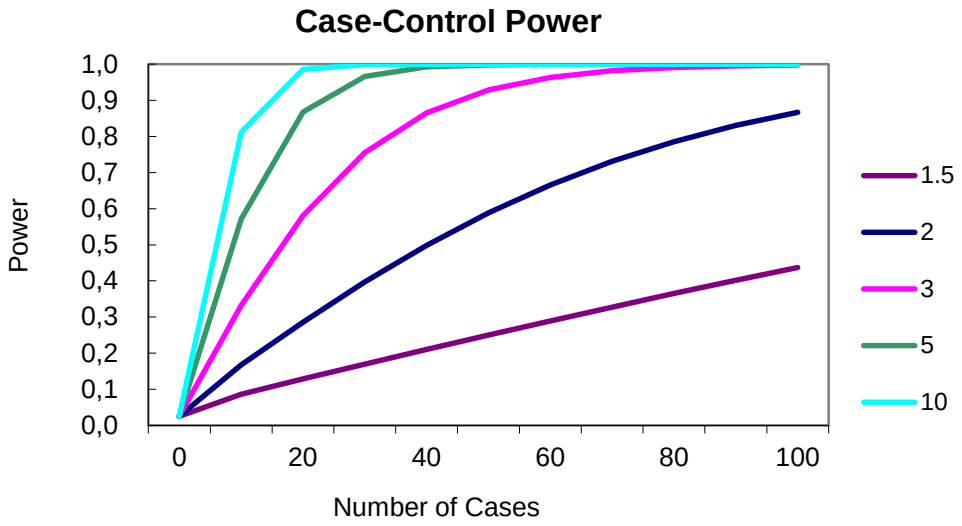


5.2 Nested Case-control Study

In the previous validation study (01 June 2005 - 31 July 2010), there were 26 pancreatic cancer cases and 29 thyroid cancer cases confirmed by medical record review. With this extension, there may be 40 pancreatic cancer cases and 40 thyroid cancer cases confirmed by charts total. With 40 confirmed cases of pancreatic cancer (or thyroid cancer) between 01 June 2005 and 31 December 2015 and 4 controls for each case, the study power is estimated to be 50% for an OR=2 and 87% for an OR=3. To reach a 80% statistical power for an OR=2, 80 confirmed cases would be necessary for each outcome. This calcuation is based on the following assumptions:

- A prevalence of exposure of 36% among controls. This percentage was the observed prevalence of exposure in the source cohorts in the previous iteration of this study, in which one exenatide initiator was cohort-matched to up to 2 comparators.
- Four controls will be sampled for each case;
- 95% confidence (z-alpha=1.96) with a two-sided test.

Figure 6. Study Power for the Nested Case-control study



6 LIMITATIONS

6.1 General Limitations of the Claims Data

This study is based on an analysis of automated medical and prescription claims, potentially supplemented by information abstracted from the medical record. While claims data are extremely valuable for the efficient and effective examination of health care outcomes, treatment patterns, health care resource utilization, and costs, all claims databases have certain inherent limitations because the claims are collected for the purpose of payment and not research. Presence of a claim for a filled prescription does not indicate that the medication was consumed or that it was taken as prescribed. Medications filled over the counter or provided as samples by the physician will not be observed in the claims data. Presence of a diagnosis code on a medical claim is not indicative of the positive presence of disease, as the diagnosis code may be incorrectly coded or included as rule-out criteria rather than actual disease. Since the claims are used to justify the service and not to clinically describe a patient, there often exist discrepancies between diagnoses associated with claims and actual clinical diagnoses, including comorbidities. This discrepancy could be differential with respect to drug exposure if physicians who treat patients with exenatide monitored their patients differently and followed up on abdominal pain reports from their patients differently. This form of detection bias would be expected to be less of a problem for the more severe forms of the outcomes (such as hospitalized cases) and for confirmed outcomes (through medical record review), but it could be substantial for minor manifestations. If physicians were more likely to warn patients on exenatide about abdominal pain rather than those on OADs, such patients might contact their physicians more readily with complaints. Similarly, physicians might be more inclined to evaluate (and attach a provisional diagnosis to) mild abdominal pain reported by a patient on exenatide than the same minimal pain in patients on OADs. This type of detection bias will likely bias the results upward, i.e., leading to a spurious association or overestimation of the association of exenatide and thyroid and/or pancreatic cancer. While this bias may lessen over

time, the presence and magnitude of this detection bias will be evaluated in the as-matched analysis described in Section 4.7.

6.2 Short Average Duration of Enrollment in Claims Data

For patients in the ORD and Impact Databases, like nearly all commercial health insurance claims databases in the US, duration of follow-up can be limited due to individuals changing health insurance plans. Within the ORD, patients are enrolled for an average of 2 years. For patients on an antidiabetic drug with 9 months of continuous enrollment, the average duration of enrollment increases to approximately 5 years; a subset of the population will remain enrolled for a long period, and this subset's later contribution of person-time can be evaluated, albeit with some caveats about the pattern of loss to follow-up in the study. Because cancer outcomes tend to have long latency periods, in a modification of the as-matched analysis, person-time will be categorized according to length of follow-up. Person-time that occurs later in follow-up is more likely to give rise to pancreatic and thyroid cancer, allowing empirical assessment of the latency period of the outcomes. Given the long latency period, follow-up time in this study period may not be sufficient to allow for complete observation of the associated study outcomes. Currently, however, no databases are available with a large number of exenatide users and long duration of follow-up. In order to mitigate this limitation, the current study will extend the study period to accrue additional exenatide users and to increase the observation period for those accrued in the original study.

6.3 Misclassification of Exposure and Outcomes Arising in the Analysis

In this study, an intent-to-treat analysis will be applied that has the advantage of preserving the randomization-like features of the propensity score matching but risks the misclassification of exposure. Such misclassification might provide conservative estimates of effect as subjects switch exposure status throughout the course of the follow-up; however, the attribution of remote outcomes to the baseline exposure status that occurs with this analysis is appropriate for cancer outcomes with a long latency period. To account for cumulative exposure to exenatide on the risk of study outcomes and reduce potential misclassification of exposure, a time-dependent exposure analysis will be used to examine the time-dependent exposure on the risk of study outcomes. However, the comparison drugs will be pooled together and considered as fixed in the analysis, which may introduce misclassification of person-time allocation in the comparison group if the risk of study outcomes changes with the cumulative exposure of OADs.

The identification of outcomes will be derived by defined algorithms, which may misclassify the true cases and non-cases. However, compared with the outcomes identified solely based on diagnosis codes within the claims data, this algorithm approach is expected to increase the specificity of outcome ascertainment, reducing the misclassification of outcomes in estimated RRs.

7 STUDY MANAGEMENT

7.1 Quality Assurance

The study will follow Optum Epidemiology's internal standard operating procedures (SOPs) that are consistent with the International Society for Pharmacoepidemiology's Guidelines for Good Pharmacoepidemiology Practices (http://www.pharmacoepi.org). In particular, the SOPs in place at Optum Epidemiology prescribe that processes and deliverables are documented, reviewed, and validated in sufficient detail to allow for subsequent re-examination or replication.

8 REFERENCES

1. Byetta and Bydureon [package insert]. Wilmington, DE: AstraZeneca Pharmaceuticals LP; August 2014 and February 2014.

2. Cure P, Pileggi A, Alejandro R et al. Exenatide and rare adverse events. N Engl J Med 2008; 358: 1969–1972.

3. Denker PS, Dimarco PE.Exenatide (exendin-4)-induced pancreatitis: a case report. Diabetes Care 2006; 29: 471.

4. Dore DD, Bloomgren GL, Wenten M, Hoffman C, Clifford CR, Quinn SG et al. A cohort study of acute pancreatitis in relation to exenatide use. Diabetes Obes Metab 2011.

5. Garg R, Chen W, Pendergrass M. Acute pancreatitis in type 2 diabetes treated with exenatide or sitagliptin: a retrospective observational pharmacy claims analysis. Diabetes Care 2010; 33(11):2349-2354.

6. Singh S, Chang HY, Richards TM, Weiner JP, Clark JM, Segal JB. Glucagonlike peptide 1-based therapies and risk of hospitalization for acute pancreatitis in type 2 diabetes mellitus: a population-based matched case-control study. JAMA Intern Med. 2013 Apr 8;173(7):534-9.

7. Funch D, Gydesen H, Tornøe K, Major-Pedersen A, Chan KA. A prospective, claims-based assessment of the risk of pancreatitis and pancreatic cancer with liraglutide compared to other antidiabetic drugs. Diabetes Obes Metab. 2014 Mar;16(3):273-5.

8. Gokhale M, Sturmer T, Pate V et al. Incretin-Based Drugs and Comparative Pancreatic Cancer Risk among Older Adults. Pharmacoepidemiology and Drug Safety 2013; 22: (SUPPL. 1) 1–521.

9. Parks M, Rosebraugh C. Weighing risks and benefits of liraglutide--the FDA's review of a new antidiabetic therapy. N Engl J Med 2010; 362(9):774-777.

10. Bulchandani D, Nachnani JS, Herndon B, et al. Effect of exendin (exenatide)--GLP 1 receptor agonist on the thyroid and parathyroid gland in a rat model. Eur J Pharmacol. 2012 Sep 15;691(1-3):292-6.

11. Lamari Y, Boissard C, Moukhtar MS, Jullienne A, Rosselin G, Garel JM. Expression of glucagon-like peptide 1 receptor in a murine C cell line: regulation of calcitonin gene by glucagon-like peptide 1. FEBS Lett 1996; 393(2-3):248-252.

12. Bjerre KL, Madsen LW, Andersen S, Almholt K, de Boer AS, Drucker DJ 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.

13. “Victoza (liraglutide injection): Human relevance of rodent thyroid C-cell tumors,” 2009, http://www.fda.gov/downloads/AdvisoryCommittees/Committees%20MeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM151129.pdf. Accessed 14 November 2014.

14. Victoza (prescribing information). Bagsvaerd, Denmark: Novo Nordisk; 2010. http://www.novo-pi.com/victoza.pdf. Accessed 14 November 2014.

15. Körner M1, Stöckli M, Waser B, Reubi JC. GLP-1 receptor expression in human tumors and human normal tissues: potential for in vivo targeting. J Nucl Med. 2007 May;48(5):736-43.

16. Boess F, Bertinetti-Lapatki C, Zoffmann S, George C, Pfister T, Roth A, Lee SM, Thasler WE, Singer T, Suter L. Effect of GLP1R agonists taspoglutide and liraglutide on primary thyroid C-cells from rodent and man. J Mol Endocrinol. 2013 Jun 1;50(3):325-36.

17. Liang C, Everage, N, et al. Incidence of Pancreatic Malignancy and Thyroid Neoplasm in Type 2 Diabetes Mellitus Patients who Initiate Exenatide Compared to Other Antihyperglycemic Drugs. Revised Final Report. Prepared for Eli Lilly and Company. 25 July 2013.

18. Final Protocol of Extension Exenatide Study, 26 March 2014.

19. Liang C, et al. Incidence of Pancreatic Malignancy and Thyroid Neoplasm in Type 2 Diabetes Mellitus Patients who Initiate Exenatide Compared to Other Antihyperglycemic Drugs. Revised Final Report. Prepared for Eli Lilly and Company, 25 July 2013. Submitted 29 July 2013 to NDA 021-773, Seq. No. 0332.

20. Liang C, et al. Incidence of Pancreatic Malignancy and Thyroid Neoplasm in Type 2 Diabetes Mellitus Patients who Initiate Exenatide Compared to Other Antihyperglycemic Drugs. Brief Report for Additional Analyses Requested by the FDA. Final Version. Prepared for AstraZeneca Pharmaceuticals LP, 20 September 2013. Submitted 30 Sept 2013 to NDA 21-773, Seq. No. 0335.

21. Liang C, Seeger JD, Dore DD. Implications of Immortal Time When Outcomes Are Non-Fatal. Pharmacoepidemiol Drug Saf 2012;21:120.

22. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika (1983) 70 (1): 41-55.

23. Sturmer, T., et al., Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. Am J Epidemiol, 2010. 172(7): p. 843-54.

24. Data Matching – Optimal and Greedy. Available at http://ncss.wpengine.netdna-cdn.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Data_Matching-Optimal_and_Greedy.pdf. Accessed 06 March 2014.

25. Parsons, L. S. (2001). Reducing bias in a propensity score matched-pair sample using greedy matching techniques. In SAS SUGI 26, Paper 214-26. Available at http://www2.sas.com/proceedings/sugi26/p214-26.pdf. Accessed 06 March 2014.

26. 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.

27. Breslow, N. E. and N. E. Day (1987). "Statistical methods in cancer research. Volume II--The design and analysis of cohort studies." IARC Sci Publ(82): 1-406.

28. Liang C, Clifford RC, Turnbull BR et al. Algorithms to Identify Pancreatic Cancer and Thyroid Neoplasm from Health Insurance Claims Data. 29th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE). Aug. 25–28, 2013 | Montreal.

29. Curtis LH et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. Pharmacoepidemiol Drug Saf. 2012 Jan;21 Suppl 1:23-31.

30. Defining and Evaluating Possible Database Models to Implement the FDA Sentinel Initiative. Accessible at http://www.fdcreports.com/~/media/Images/Publications/Archive/The%20Gray%20Sheet/35/020/01350200013/sentinel_database_models_05_09.pdf. Accessed 06 March 2014.

31. Observational Medical Outcomes Partnership. Website: http://omop.org/CDM. Accessed 06 March 2014.

32. Glasziou PP, Sanders SL. Investigating causes of heterogeneity in systematic reviews .Stat Med 2002; 21: 1503-11.

33. Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. Ann Intern Med 1997; 127 (9): 820-826.

34. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf 2006; 15(5):291-303.

35. Dore DD, et al. A cohort study of acute pancreatitis in relation to exenatide use. Diabetes Obes Metab. 2011;13(6):559-66.

36. Greenland S, Neutra R. An analysis of detection bias and proposed corrections in the study of estrogens and endometrial cancer. J Chronic Dis 1981; 34(9-10):433-438.

37. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res 1999; 1(8):3-15.

38. Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing-data procedures. *Psychological Methods, 6,* 330–351.

39. Rubin, D.B. (1987) Multiple Imputation for Nonresponse in Surveys. J. Wiley & Sons, New York.

9 APPENDICES

9.1 Appendix I. Other Antidiabetic Drugs (OADs) Considered for Cohort Entry, Excluding Dipeptidyl Peptidase-4 / Glucagon-like Peptide-1 (DPP-4/GLP-1) Agonist

METFORMIN
MEFORMIN HCL
METFORMIN/CAFF/AA7/HRB125/CHOL
METFORMIN/AA COMB.#7/HC#125/CH
SULFONYLUREAS
ACETOHEXAMIDE
TOLBUTAMIDE
CHLORPROPAMIDE
TOLAZAMIDE
GLYBURIDE
GLIPIZIDE
GLIMEPIRIDE
THIAZOLIDINEDIONES
TROGLITAZONE
ROSIGLITAZONE MALEATE
PIOGLITAZONE HCL
AMYLINOMIMETICS
PRAMLINTIDE ACETATE
NONSULFONYLUREA SECRETAGOGUES
REPAGLIINIDE
NATEGLINIDE
ALPHA–GLUCOSIDASE INHIBITORS
ACARBOSE
MIGLITOL
INSULIN

9.2 **Appendix II.** Diagnosis Codes for Pancreatic Cancer and Thyroid Neoplasm

The International Classification of Diseases, 9th Revision (ICD–9) diagnosis codes for identification of pancreatic cancer and thyroid neoplasm.

ICD–9	**Description**
	Pancreatic Cancer
157.X	Malignant neoplasm of pancreas
157.0	Malignant neoplasm of head of pancreas
157.1	Malignant neoplasm of body of pancreas
157.2	Malignant neoplasm of tail of pancreas
157.3	Malignant neoplasm of pancreatic duct
157.4	Malignant neoplasm of islets of Langerhans
157.8	Malignant neoplasm of other specified sites of pancreas
157.9	Malignant neoplasm of pancreas, part unspecified
	Thyroid Neoplasm
193	Malignant neoplasm of thyroid gland
226	Benign neoplasm of thyroid glands

9.3 Appendix III. Predefined Algorithms for Pancreatic Cancer and Thyroid Neoplasm

1. **Case Algorithm for Pancreatic Cancer**
    a. Any in– or out– patient diagnosis codes of pancreatic cancer, and
    b. Without a diagnosis of benign pancreatic neoplasm within 60 days after the diagnosis of pancreatic cancer, and
    c. With **one or more** pancreas surgery, chemotherapy or radiation therapy within 180 days after the diagnosis of pancreatic cancer, ***and***
    d. Without a diagnosis of other cancers (see list below) within 60 days before or after the diagnosis of pancreatic cancer

ICD-9 code	***Cancer***
150.xx	***Malignant neoplasm of esophagus***
151.xx	***Malignant neoplasm of stomach***
152.xx	***Malignant neoplasm of small intestine, including duodenum***
153.xx	***Malignant neoplasm colon***
154.xx	***Malignant neoplasm of rectum, rectosigmoid junction, and anus***
155.xx	***Malignant neoplasm of liver and intrahepatic bile ducts***
156.xx	***Malignant neoplasm of gallbladder and extrahepatic bile ducts***
158.xx	***Malignant neoplasm of retroperitoneum and peritoneum***
159.xx	***Malignant neoplasm of other and ill-defined sites within the digestive organs and peritoneum***
162.xx	***Malignant neoplasm of trachea, bronchus, and lung***
165.xx	***Malignant neoplasm of other and ill-defined sites within the respiratory system and intrathoracic organs***
171.5	***Malignant neoplasm of connective and other soft tissue (Abdomen)***
188.xx	***Malignant neoplasm of bladder***
195.2	***Malignant neoplasm of other and ill-defined sites (Abdomen)***

2. **Case Algorithm for Thyroid Cancer**
    a. Any in– or out– patient diagnosis codes of thyroid cancer, and
    b. Without a diagnosis of benign thyroid neoplasm within 60 days after the diagnosis of thyroid cancer, and
    c. With **one or more** of thyroid surgery, chemotherapy, radioiodine therapy or radiation therapy within 180 days after the diagnosis of thyroid cancer

3. **Case Algorithm for Medullary Thyroid Cancer (MTC)**
    a. Any in– or out– patient diagnosis codes of thyroid cancer, and
    b. Without a diagnosis of benign thyroid neoplasm within 60 days after the diagnosis of thyroid cancer, and
    c. With **2 or more** of thyroid surgery, chemotherapy, radioiodine therapy or radiation therapy **plus** thyroid hormone replacement therapy within 180 days after the diagnosis of thyroid cancer, and
    d. With one or more claims evidence of **serum calcitonin** levels within 180 days after thyroid surgery or thyroid cancer diagnosis

    NOTE: A relaxed algorithm was used in the final analysis of MTC. That algorithm included either 3.a + 3.b + 3.d or 3.a + 3.c + 3.d.

4. **Case Algorithm for Benign Thyroid Neoplasm**
    a. Any in– or out– patient diagnosis codes of benign thyroid neoplasm, and
    b. Without a diagnosis of thyroid cancer within 60 days after the diagnosis of benign thyroid neoplasm, and
    c. With biopsy claims within 90 days before the diagnosis of benign thyroid neoplasm

Please note that the date of diagnosis above refers to the date of first claim for the diagnosis.

9.4 Appendix IV. Summary of Statistical Analysis

Outcome	Exposure	Estimates	Database
COHORT STUDY			
Time-fixed analysis			
Algorithm-identified outcomes* • Pancreatic cancer • Thyroid cancer	Exenatide initiation vs OADs initiation • Overall • Stratified by duration of use	**Confounding control:** * Propensity score matching * Adjust for imbalance characteristics, overall and within stratification **Cox proportional hazards model:** IRs and HRs	* Combined dataset * ORD only * Impact Database only
	With and without concurrent use of insulin	IRs	* Combined dataset only
Algorithm-identified outcomes* • MTC • Non-MTC • Benign thyroid neoplasm	Exenatide initiation and OADs initiation	IRs	* Combined dataset only
Cumulative exposure analysis			
Algorithm-identified outcomes* • Pancreatic cancer • Thyroid cancer	Cumulative exenatide use vs non-use • Cumulative dose (Category 1, 2…) vs non-use • Cumulative duration (Category 1, 2…) vs non-use	**Confounding control:** * Propensity score matching * Adjust for time-dependent OADs **Poisson regression model:** IRs and RRs with trend test	* Combined dataset * ORD only * Impact Database only

Cases	Controls	Exposure	Estimates	Database
CASE-CONTROL STUDY				
Chart-confirmed cases • Pancreatic cancer • Thyroid cancer	**Non-cases:** • Randomly selected from source cohorts and chart obtained	Exposure status at baseline as defined by drug initiation that qualified patient for cohort entry	ORs from conditional logistic regression model that approximate RRs from cohort study. Estimates adjusted for claims-based and chart-based covariates with multiple imputation as needed.	* ORD only, identifiable subset

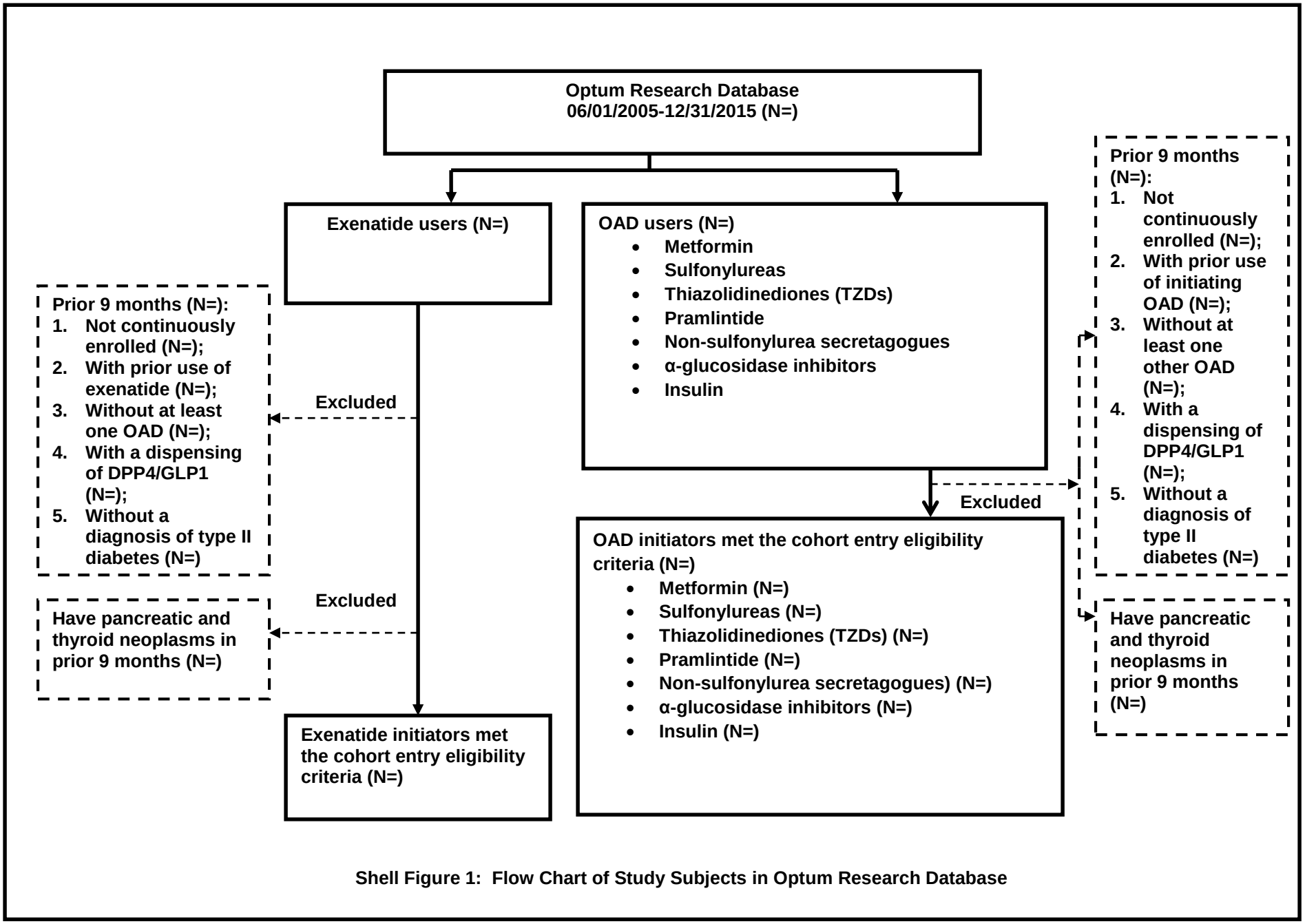
Abbreviations: ORD=Optum Research Database; OADs=Other Antidiabetic Drugs; MTC=Medullary Thyroid Cancer; IRs=Incidence Rates; HRs=Hazard Ratios; RRs=Rate Ratios; ORs=Odds Ratios.

* Outcomes identified using the restricted algorithms in Appendix III.

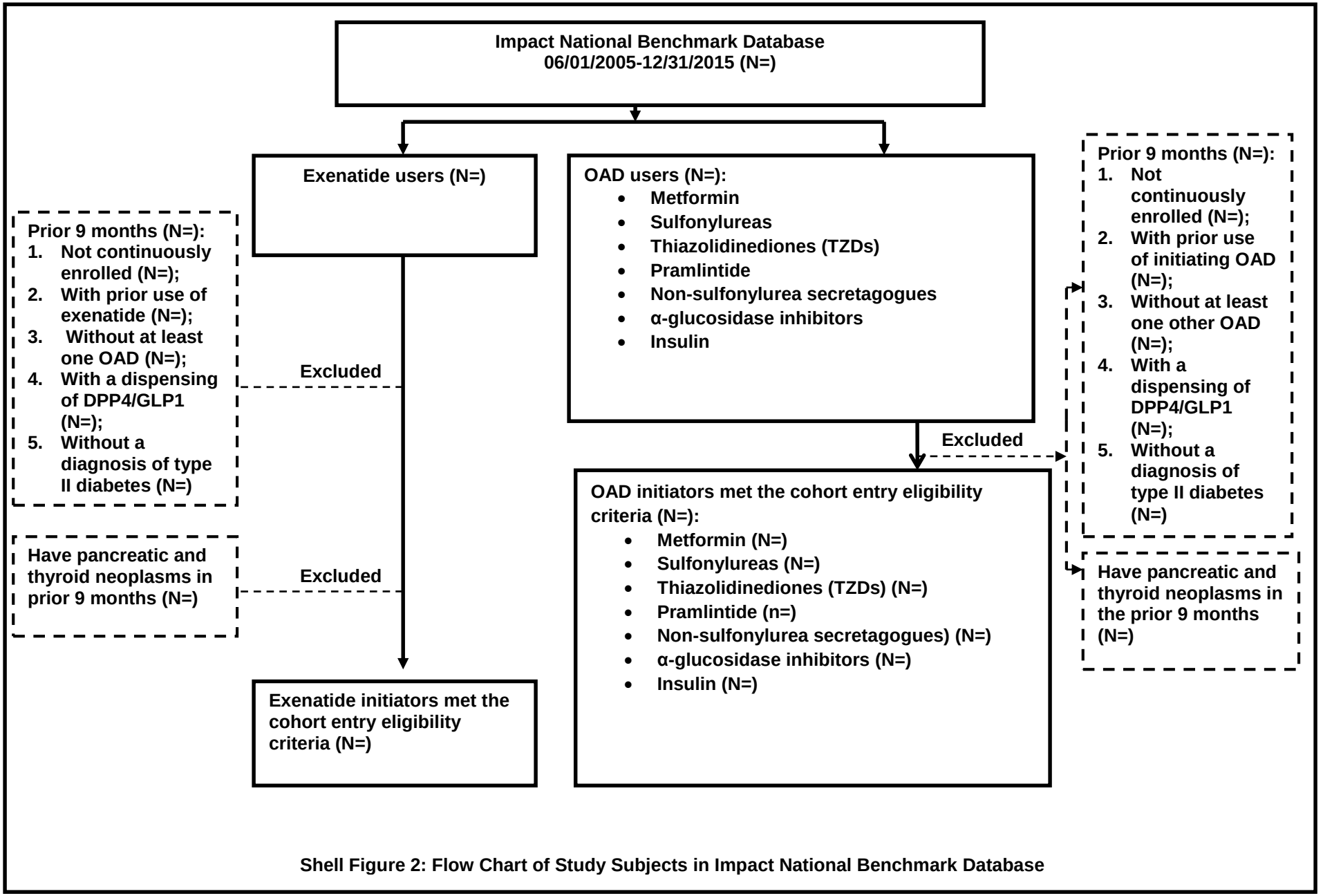
10 SHELL FIGURES*

1. Shell Figure 1, 2, and 3: Flow Charts of Study Subjects
2. Shell Figure 4: Distribution of Propensity Score by Study Cohorts Prior Matching (by database)
3. Shell Figure 5: Distribution of Propensity Score by Study Cohorts After Matching (by matching ratio and database)
4. Shell Figure 6: Kaplan-Meier Curves of Time to Pancreatic Cancer by Cohort in Combined Database
5. Shell Figure 7: Kaplan-Meier Curves of Time to Thyroid Cancer by Cohort in Combined Database
6. Shell Figure 8: Evaluation of the Confounding Caused by Smoking Needed to Explain the Apparent Relative Risk of Pancreatic Cancer, Combined Database
7. Shell Figure 9: Evaluation of the Confounding Caused by Smoking Needed to Explain the Apparent Relative Risk of Thyroid Cancer, Combined Database
8. Shell Figure 10: Evaluation of the Confounding Caused by Obesity Needed to Explain the Apparent Relative Risk of Pancreatic Cancer, Combined Database
9. Shell Figure 11: Evaluation of the Confounding Caused by Obesity Needed to Explain the Apparent Relative Risk of Thyroid Cancer, Combined Database

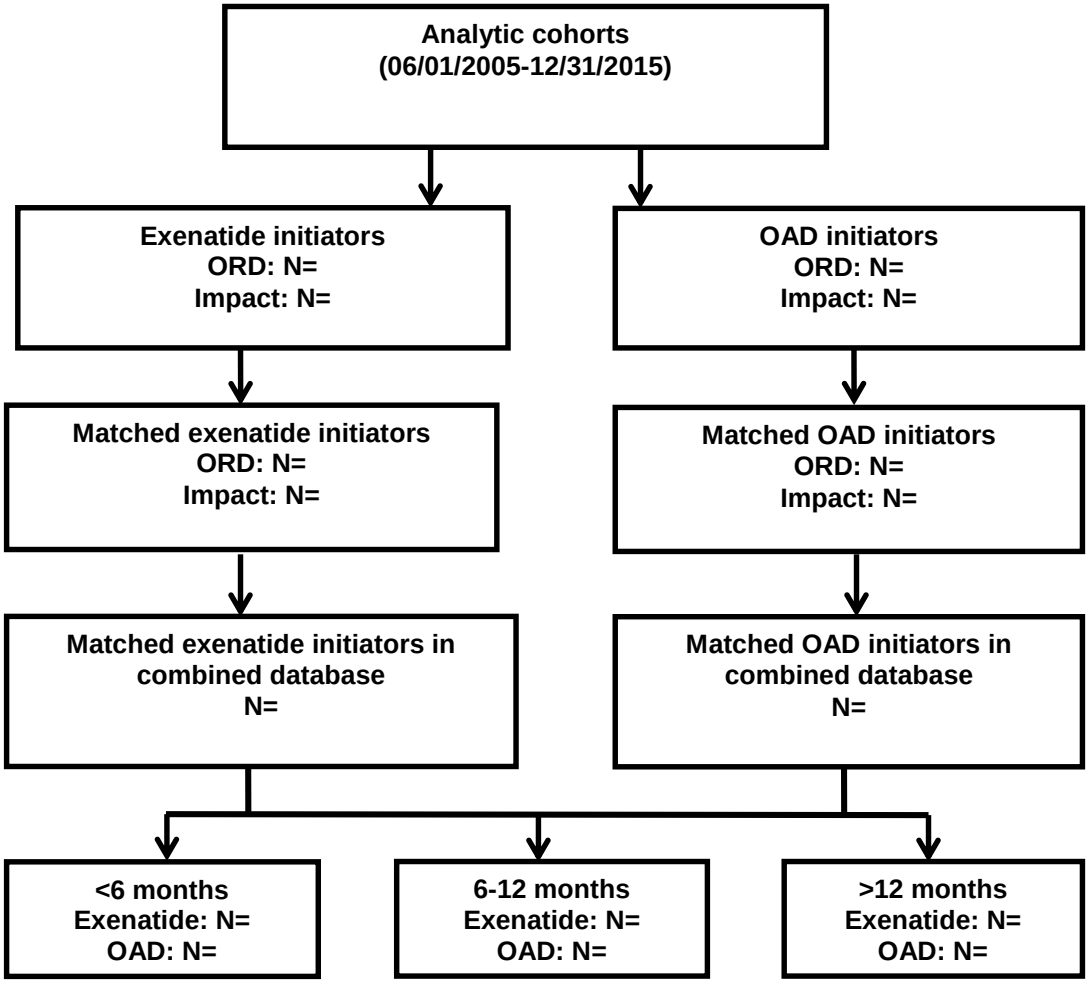
*Note: Shell Figures 4-11 are not shown in the protocol because the information underlying these figures is not available at this time. The figures will be presented in the report.



Shell Figure 1: Flow Chart of Study Subjects in Optum Research Database



Shell Figure 2: Flow Chart of Study Subjects in Impact National Benchmark Database



Shell Figure 3: Flow Chart for Combined Initiators in Analytic File

11 SHELL TABLES