Risk of Thyroid Cancer Associated with Use of Liraglutide and Other Antidiabetic Drugs in a US Commercially Insured Population

Donnie Funch, Kathleen Mortimer, Najat J Ziyadeh, John D Seeger, Li Zhou, Eva Ng, Douglas Ross, Atheline Major-Pedersen, Heidrun Bosch-Traberg, Helge Gydesen, David D Dore

Optum Epidemiology, Boston, MA, USA; Massachusetts General Hospital, Thyroid Associates, Boston, MA, USA; Harvard Medical School, Department of Medicine, Boston, MA, USA; Global Safety, Novo Nordisk A/S, Copenhagen, Denmark; Global Development, Novo Nordisk A/S, Copenhagen, Denmark; Epidemiology, Novo Nordisk A/S, Copenhagen, Denmark; Department of Health Services, Policy & Practice, Brown University School of Public Health, Providence, RI, USA

Background: Quantify association between the glucagon-like peptide-1 receptor agonist liraglutide and risk of thyroid cancer (TC) compared to other antidiabetics.

Patients and Methods: Initiators of liraglutide, exenatide, metformin, pioglitazone or groups of dipeptidyl peptidase-4 inhibitors or sulfonylureas were identified in a US health plan (2010–2014) and followed for a median of 17 months. Thyroid cancer cases during follow-up were identified via a validated algorithm. Incidence rates of TC among liraglutide and comparators were assessed using relative risks estimated within propensity score-matched cohorts using intention to treat (ITT) and time on drug analyses. Latency effects and potential surveillance bias were evaluated.

Results: Relative risks from ITT analyses ranged from 1.00 (95% confidence interval (CI) 0.56–1.79) versus metformin to 1.70 (95% CI 1.03–2.81) versus all comparators excluding exenatide. Effect estimates from latency analyses were slightly attenuated. Time on drug analyses suggested no increased risk for either longer duration or higher cumulative dose of liraglutide. Medical record review found 85% were papillary or a follicular variant of papillary or both; 46% were microcarcinomas (≤10 millimeters), which were more prevalent in the liraglutide cohort (67% versus 43% in all comparators).

Conclusion: Relative risks were elevated for several comparisons, which should be interpreted cautiously because of potential residual confounding and surveillance bias. Liraglutide cases had smaller thyroid nodules and shorter time-to-diagnosis, suggesting increased surveillance for TC among liraglutide initiators, especially shortly after the drug´s approval. After adjusting the primary analyses (ITT) for latency, no significant elevated risk of TC was observed among liraglutide initiators.

Keywords: glucagon-like peptide-1 receptor agonist, type 2 diabetes, administrative claims, intention-to-treat, time-on-drug

Introduction

Type 2 diabetes mellitus (T2DM) is the most common type of diabetes, accounting for approximately 90% of all cases worldwide, and incretin-based therapies are widely used by patients with T2DM. Liraglutide is a once-daily glucagon-like peptide-1 receptor agonist (GLP-1 RA) approved for use in the United States (US) in 2010 to improve glycemic control in adults with T2DM as an adjunct to diet and exercise. On liraglutide’s approval, the Food and Drug Administration (FDA) required a five-year epidemiological study to compare the incidence of thyroid cancer (TC) between initiators of liraglutide and initiators of other
antidiabetic drugs (ADs).\textsuperscript{1} The GLP-1 receptor is expressed in all types of thyroid cells including normal, premalignant or malignant tissues. This research requirement arose from studies showing that GLP-1 RAs, including liraglutide, caused thyroid C-cell hyperplasia and C-cell tumors in carcinogenicity studies in rodents.\textsuperscript{2} However, animal experiments with monkeys did not show increased proliferation of C-cells in thyroid glands after chronic administration of liraglutide, and longitudinal data from clinical trials have not demonstrated a causal association between GLP-1 analogs and thyroid C-cell pathology.\textsuperscript{2,3}

The incidence of TC has been increasing on average 4.5% annually over the past decade,\textsuperscript{4} and specifically, medullary thyroid carcinoma (MTC), the subtype that originates from C cells, accounts for 1–2% of TCs in the US.\textsuperscript{5} The FDA label contains a warning regarding MTCs for GLP-1 RAs; however, the applicability of rodent studies to the occurrence of MTC or TC generally in humans remains uncertain, and the rarity and difficulty of identifying MTC without pathology reports complicates its study.\textsuperscript{6,7}

Rates of spontaneous reports of TC from 2004 through 2009 were elevated with exenatide, another GLP-1 RA.\textsuperscript{8,9} Although spontaneous reporting systems enable identification of a range of potential safety signals, biased reporting is common and often profound; therefore, evidence on GLP-1 RA and TC has been sparse and inconclusive.\textsuperscript{9,10}

In this paper, results of the primary analyses of the FDA post-marketing requirement are presented where the relationship between liraglutide and TC relative to several specific drug or drug class (ie, SU or DPP-4i) during the preceding 6-month baseline period. The operational definition of an initiator included patients who were naïve to AD treatments, those who switched to cohort-defining therapy from another specific AD or drug class, and those who started cohort-defining therapy as add-on treatment to existing therapy. Patients with claims for International Classification of Diseases, 9th Revision (ICD-9) diagnosis codes 193 (Malignant neoplasm of thyroid gland) or V10.87 (Personal history of malignant neoplasm of thyroid) during baseline were excluded to allow for identification of incident cancers during follow-up.

Accrual occurred from 01 February 2010 to 30 November 2014. Due to changing prescribing patterns and composition of ADs in the US market, cohorts were assembled quarterly for the first 3 years (2010–2012) and annually thereafter (2013–2014), for 14 total calendar blocks. Patients were assigned to the cohort of the first drug initiated during a calendar accrual block. Throughout

**Methods**

The patients included in this prospective cohort study were drawn from the Optum Research Database (ORD), a proprietary research database that contains the eligibility data, medical claims, and pharmacy claims from a large, commercial health plan affiliated with Optum. The individuals covered by this health plan are geographically diverse across the US and comprise approximately 4% of the US population.\textsuperscript{13}

**Data Source**

The ORD is a proprietary research database containing claims and enrollment data dating back to 1993 for members of a large US health plan. For 2014, data relating to approximately 12.7 million individuals with both medical and pharmacy benefit coverage and an average of 2.5 years of enrollment were available. The medical and pharmacy claims for these individuals forms a longitudinal record of reimbursed medical services, irrespective of treatment site, along with detailed information on drug dispensings (including drug name, dosage form, strength, fill date, days of supply, and financial information) and on medical claims (including multiple diagnosis codes, procedures, site of service, provider specialty, and paid amounts) collected from available health-care sites (ie, inpatient hospital, outpatient hospital, emergency room, physician’s office, etc.) for virtually all types of provided services, allowing for longitudinal tracking of patients. Access to a subset of patients’ medical charts (ie, approximately 35% of all patients within the ORD) allows for confirmation of outcomes identified through claims.

**Study Population**

Study ADs included liraglutide and comparator medications (exenatide, metformin, pioglitazone, a sulfonylurea [SU; glyburide, glipizide, gimepiride] or a dipeptidyl peptidase-4 inhibitor [DPP-4i; sitagliptin, saxagliptin, linagliptin]). Patients aged 18–89 years who initiated liraglutide or another AD following at least 6 months of continuous health plan enrollment with complete medical and pharmacy benefits. Drug initiation was defined by pharmacy claims for ADs, with no dispensing of that specific drug or drug class (ie, SU or DPP-4i) during the preceding 6-month baseline period. The operational definition of an initiator included patients who were naïve to AD treatments, those who switched to cohort-defining therapy from another specific AD or drug class, and those who started cohort-defining therapy as add-on treatment to existing therapy. Patients with claims for International Classification of Diseases, 9th Revision (ICD-9) diagnosis codes 193 (Malignant neoplasm of thyroid gland) or V10.87 (Personal history of malignant neoplasm of thyroid) during baseline were excluded to allow for identification of incident cancers during follow-up.
the study, however, patients were permitted to enter more than one cohort if they met eligibility criteria with initia-
tions in later accrual blocks. Follow-up began the day after
first eligible initiation and continued until the earliest of
disenrollment from health plan, TC occurrence, or end of
study period (31 December 2014). A National Death Index
search was conducted to identify deaths due to TC, but
because it is rarely fatal, results were sparse; therefore,
date of death was not used to censor.

Propensity Score Matching
Treatment of (T2DM) consists of a stepwise escalation of
therapy along the progressive continuum of the disease
with metformin widely accepted as the first-line
therapy.14,15 This treatment strategy may create a
selection bias within observational studies, and in order
to mitigate this potential bias, baseline demographics,
comorbidities, medications, and healthcare utilization cov-
ariates in addition to the 150 most frequently occurring
diagnoses, procedures, and drug dispensings in the ORD
were balanced using propensity scores (PS), which were
estimated separately for 8 comparison groups (5 compar-
tor drugs as well as all comparators, all comparators minus
exenatide [to remove any GLP-1 RA effect], and all com-
parators minus exenatide and DPP-4i [to remove any in-
cretnin effect from the all comparator cohort]).16–18

A priori variables correlated with exposures and likely
to be correlated with TC were forced into the model
including known risk factors (benign thyroid disease, over-
weight/obese, markers for diabetes severity, and variables
representing baseline AD use [number of AD dispensings,
use of insulin, individual ADs]). For the remaining base-
line and Top 150 variables, a stepwise selection technique
was used with an entry p-value of ≤0.20 and an exit
p-value of ≥0.30. Candidate variables for inclusion in PS
models are in Supplementary Appendix 1.

To account for changes in prescribing patterns and
availability of ADs on the market, PS estimation and
matching were performed by calendar periods and separ-
ately for each comparator. Initiators of liraglutide were
PS-matched 1:1 to initiators in each of the comparison
groups using a standard greedy matching algorithm.19–21
Patients who did not match during one accrual time block,
but initiated any study drug in a later accrual block, were
eligible for matching, with PS estimated based on updated
variable information for the later initiation. Patients could
enter into multiple matched cohorts but only once for each
drug cohort pair.

Identification of TC
Thyroid cancer cases were identified during follow-up
using an algorithm developed within this population and
validated through medical record review.22 The algorithm
had a positive predictive value (PPV) of 0.91 (95% con-
fidence interval [CI] 0.81–0.96) requiring both
a thyroidectomy and at least 2 separate diagnoses for
malignant neoplasm of thyroid gland (ICD-9 193) within
90 days of the thyroidectomy. Date of diagnosis was the
first TC claim associated with thyroidectomy.

Analysis
Baseline distributions of covariates were calculated for
liraglutide initiators and all combined comparators.
Because of the complexity of drug use patterns and the
different mechanisms of cancer occurrence, several analy-
tic methods were undertaken to quantify the risks associ-
ated with initiating liraglutide, recency of use of
liraglutide, and cumulative exposure history.

As the primary analysis, an intent to treat (ITT) analy-
sis quantified the risk associated with starting treatment
by attributing follow-up time in the cohort to the drug used at
cohort entry and accounting for potential risk after drug
discontinuation, especially relevant when investigating
cancer outcomes. Incidence rates (IR; TC cases per
100,000 person-years [py]) were estimated as number of
patients with TC divided by py at-risk. Poisson regression
models computed estimated IR ratios (RRs) and 95% CIs
for liraglutide versus each comparator. Generalized esti-
mating equations (GEE) with exchangeable correlation
structures and robust variance estimators accounted for
the paired nature of the data created by matching, except
where data were too sparse to support GEE.23 In a latency
analysis, person-time and TC cases from the first year of
follow-up were excluded to account for the probable
latency period of any exposure effect on TC.

Two time on drug (TOD) analyses evaluated the effect
of recency of use and cumulative exposure to liraglutide.
Based on dispensing date and days’ supply, each person-
day was assigned to “current”, “recent”, and “past” use
categories independently for each study drug. A “grace
period” of 31 days was added to account for medication
non-adherence and the uncertainty surrounding duration of
biologic effect for these drugs. Current use was defined as
the initiation date through end of days’ supply plus the 31-
day grace period and continued as long as new dispensings
were observed before the grace period ended. If the grace
period expired, it was assumed the patient discontinued that drug, and the subsequent 31 days were considered “recent” use. Person-time occurring after “recent” use was categorized as “past”, which persisted unless the patient re-started the same treatment, thereby reentering the current use category. Because many patients were on more than one AD and switched therapies regularly, those in matched pairs may have had exposure to both liraglutide and the comparator during follow-up. If TC was diagnosed while the patient was currently on liraglutide and had past use of the comparator, for example, the TC case (and corresponding person-time) counted toward both the incidence estimate for current liraglutide use and past comparator use. Poisson regression estimated RRs and 95% CIs for current, recent, and past use of liraglutide versus the same category for comparators, adjusted by the logit of the PS to address confounding.

For the second TOD analysis, each patient’s cumulative person-time exposed and unexposed to liraglutide was quantified. If a patient initiated another study drug, subsequent person-time was categorized as “liraglutide-unexposed” and continued until end of follow-up or TC diagnosis. All person-time was categorized as “liraglutide-exposed” time once a dispensing for liraglutide was observed. Rate ratios were estimated using Poisson regression modeling within observed categories of cumulative time (<6, 6–18, >18 months) relative to all “liraglutide-unexposed” time, and models were adjusted for the logit of the PS.

Concerns about increased screening for TC among liraglutide initiators were evaluated by tabulating frequencies of TC screening and diagnostic tests during follow-up. In addition, time-to-diagnosis by drug cohort and calendar years were plotted to assess whether screening intensity was higher soon after liraglutide’s approval. For the subset of cases confirmed via medical records, tumor histology and count of microcarcinomas were summarized by exposure to test the assumption that a higher proportion of papillary microcarcinomas may indicate more vigilant screening.

Privacy and Confidentiality
Approval of the study protocol and a waiver of patient authorization from the New England Institutional Review Board and affiliated Privacy Board was obtained. All analyses were conducted using appropriately de-identified data.

Analyses were conducted using SAS Version 9.4.

Results
Before matching, there were 27,287 eligible liraglutide initiators and 362,809 comparators. The size of matched cohorts varied considerably due to patient characteristics and prevalence of drugs on the market. Follow-up times ranged from one day to 58.5 months (Table 1, Section A). Median lengths of follow-up were similar within matched cohorts, with some variation across matched sets (eg, medians for the matched liraglutide: pioglitazone cohorts were over 2 months longer than those for the liraglutide: sulfonylurea matched cohorts). Median time to TC was shorter for liraglutide than most comparator cohorts (Table 1, Section B; Supplementary Appendix Figures S-1A–S-1H). Composition of these all comparator groups by specific study drug is available in Supplementary Appendix Table S-1. Propensity score matching balanced baseline characteristics of liraglutide initiators and all comparators (Supplementary Appendix Table S-2) and within all other comparator pairs (data not shown).

ITT and TOD Analyses
Results for primary ITT analyses are displayed in Figure 1. Across drug comparisons, RRs ranged from 1.00 (95% CI 0.56–1.79) for the comparison with metformin to approximately 1.70 for each of the following comparisons: all comparators (RR 1.66, 95% CI 1.00–2.75), all comparators except exenatide (RR 1.70, 95% CI 1.03–2.81), and exenatide (RR 1.66, 95% CI 0.73–3.79). In the latency analysis, RRs ranged from 0.54 (95% CI 0.20–1.47) to 1.62 (95% CI 0.76–3.43).

In TOD analyses, RRs among current users ranged from 0.92 (95% CI 0.41–2.05) versus pioglitazone to 2.16 (95% CI 0.70–6.70) versus exenatide. Recent use RRs were higher than current but were imprecise due to the small amount of person-time. Past-time RRs followed no consistent pattern. Adjusted RRs were 1.75 (95% CI 0.96–3.19) for <6 months exposure, 1.85 (95% CI 0.93–3.65) for 6–18 months, and 2.10 (95% CI 0.89–4.96) for >18 months exposure.

Across all ITT and TOD analyses, the only comparison that excluded the null was the ITT analysis for all comparators except exenatide (RR 1.70, 95% CI 1.03–2.81).

Surveillance Bias Assessments
The percentage of initiators with thyroid-related assays, imaging procedures or diagnostic procedures, or presence
Table 1 Amount of Follow-Up Time Available [Section A] and Time to Thyroid Cancer Diagnosis (in Days) [Section B] by Matched Drug Pairs

<table>
<thead>
<tr>
<th>Study Drug</th>
<th>Section A</th>
<th></th>
<th></th>
<th></th>
<th>Section B</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Follow-Up Time in Days</td>
<td></td>
<td></td>
<td></td>
<td>Time to Thyroid Cancer in Days for All Thyroid Cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Median</td>
<td>IQR</td>
<td>Median</td>
<td>IQR</td>
<td>Cases</td>
<td>Median</td>
<td>IQR</td>
</tr>
<tr>
<td>All Comparator Drugs Except Exenatide</td>
<td>34,707</td>
<td>526</td>
<td>234–958</td>
<td>528</td>
<td>235–945</td>
<td>41</td>
<td>311</td>
<td>150–616</td>
</tr>
<tr>
<td>All Comparator Drugs Except Exenatide And DPP-4i</td>
<td>32,935</td>
<td>523</td>
<td>233–951</td>
<td>519</td>
<td>238–939</td>
<td>35</td>
<td>311</td>
<td>131–649</td>
</tr>
<tr>
<td>Exenatide</td>
<td>17,589</td>
<td>536</td>
<td>244–937</td>
<td>543</td>
<td>244–937</td>
<td>15</td>
<td>310</td>
<td>206–556</td>
</tr>
<tr>
<td>DPP-4i (Saxagliptin/Sitagliptin/Linagliptin)</td>
<td>25,579</td>
<td>519</td>
<td>234–932</td>
<td>520</td>
<td>229–922</td>
<td>24</td>
<td>257</td>
<td>88–385</td>
</tr>
<tr>
<td>Metformin</td>
<td>24,747</td>
<td>504</td>
<td>224–917</td>
<td>510</td>
<td>238–914</td>
<td>23</td>
<td>263</td>
<td>83–370</td>
</tr>
<tr>
<td>Pioglitazone</td>
<td>18,210</td>
<td>573</td>
<td>250–1049</td>
<td>560</td>
<td>246–1033</td>
<td>18</td>
<td>355</td>
<td>150–888</td>
</tr>
</tbody>
</table>

Notes: Patients were allowed to initiate multiple times during the study period and to match only once into each drug cohort pair. Follow-up began on the day following initiation and continued until end of enrollment or end of study period (12/31/2014), whichever came first. Time is calculated as number of days from index date to first claim of ICD-9 code 193. Case assignment is based on initiation at time of matched cohort entry (ITT analysis).

Abbreviations: DPP-4i, dipeptidyl peptidase-4 inhibitor; IQR, interquartile range; ITT, intention to treat.

Discussion

This 5-year prospective cohort study was designed to fulfill the FDA’s post-marketing requirement to quantify the risk of TC associated with liraglutide. Cohorts were constructed within an administrative claims database, and outcomes were validated through medical record review. Analyses addressed plausible sources of bias and compared therapies with different mechanisms of action on the occurrence of TC. While RRs were elevated for papillary microcarcinomas and evidence of increased TC on surveillance among the liraglutide cohort, especially shortly after its approval, limited and conflicting data are available on the association between GLP-1 RAs and TC. Published works referred primarily to an analysis of the FDA Adverse Event Reporting System database, which has limitations common to many administrative databases, including lack of exposure verification, use of diagnostic codes in place of true diagnoses, and limited ability to identify new TC cases. Studies that have measured the incidence of TC in patients using GLP-1 RAs are uncommon, and a recent analysis of FDA databases comparing Liraglutide (Victoza®) to placebo found no increased risk of TC in participants treated with liraglutide (https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/208259s012_final.pdf).

Medical records for 150 patients identified with potential TC were obtained and adjudicated by a clinician. Of those, 22 cases were confirmed, and pathology information was available (Table 2). There were no cases of MTC in contrast to 3 of 7 cases (42.9%) for all comparators. Limited and conflicting data are available on the association between GLP-1 RAs and MTC. Published works refer primarily to an analysis of the FDA Adverse Event Reporting System database, which has limitations common to many administrative databases, including lack of exposure verification, use of diagnostic codes in place of true diagnoses, and limited ability to identify new TC cases. Studies that have measured the incidence of MTC in patients using GLP-1 RAs are uncommon. A recent analysis of FDA databases comparing Liraglutide (Victoza®) to placebo found no increased risk of MTC in participants treated with liraglutide (https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/208259s012_final.pdf).

In 2010 and 2011, when liraglutide was new to the market, time to TC diagnosis was substantially shorter among liraglutide initiators (Figure 2). This difference attenuated over calendar time; specifically, among patients who initiated in 2010, 12 of 17 (70.6%) cases of TC among liraglutide initiators occurred within the first 300 days of follow-up (approximate median follow-up time), compared with 67% among all comparators (67% versus 43%) (49.8% versus 48.5%, respectively) (Supplementary Appendix Table S-3).

Analyses addressed plausible sources of bias and compared therapies with different mechanisms of action on the occurrence of TC. While RRs were elevated for papillary microcarcinomas and evidence of increased TC on surveillance among the liraglutide cohort, especially shortly after its approval, limited and conflicting data are available on the association between GLP-1 RAs and TC. Published works referred primarily to an analysis of the FDA Adverse Event Reporting System database, which has limitations common to many administrative databases, including lack of exposure verification, use of diagnostic codes in place of true diagnoses, and limited ability to identify new TC cases. Studies that have measured the incidence of TC in patients using GLP-1 RAs are uncommon, and a recent analysis of FDA databases comparing Liraglutide (Victoza®) to placebo found no increased risk of TC in participants treated with liraglutide (https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/208259s012_final.pdf).
Reporting System data, which found higher TC reporting with exenatide (odds ratio [OR] 4.73). A more robust analysis found a significant increase in risk with exenatide (OR > 3.0), although the report acknowledged limitations of spontaneous report data that affect both studies and represent plausible explanations for the findings. These limitations include heightened exenatide reporting arising from public awareness of concerns regarding TC and...
GLP-1 RAs and selective prescribing of exenatide due to its weight-lowering properties in obese patients who are at greater risk for TC. A meta-analysis of 25 longitudinal studies (22 randomized controlled trials, 3 retrospective cohort analyses, 775,602 patients) concluded that liraglutide was not associated with an increased risk of TC (liraglutide OR 1.54, 95% CI 0.40–6.02).

Randomized trials cannot easily characterize rare outcomes. Even the relatively large Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results (LEADER) trial could not achieve sufficient power to analyze TCs (5 TC cases with liraglutide and 3 with placebo). Intention to treat, TOD, and latency analyses were conducted in this study, and results were generally consistent with no increase in risk of TC among liraglutide initiators. The study findings support existing evidence that surveillance bias is likely when the AD is new to market. Among liraglutide-exposed cases, there was a larger

Figure 2 Proportion of thyroid cancer cases diagnosed since cohort entry by year of initiation.
This study has a number of strengths. The database source used is a commercial US health insurance plan, which allowed for a well-powered and generalizable analysis. The age distribution of the ORD is similar to that of the US population through age 64. Approximately 4.5% of members in the database are aged 65 and older, compared with approximately 13% of the US 2014 Census Population. Because it is one of the largest administrative healthcare databases in the US, the ORD supports surveillance on cohorts exposed to liraglutide and the study of rare outcomes. Moreover, while some healthcare products or services may be paid for out-of-pocket, the majority of services used for diagnosis and treatment of TC are likely to be reimbursed by the health plan and therefore, captured in the ORD. These results are representative of the US working population under age 65 and contribute external validity to this area of drug safety.

The 5-year study time period is sufficient to cover not only the time around initial drug approval (when provider uncertainty regarding the new therapy may result in shifting patterns and other temporal changes by applying the matching models throughout calendar time), but also considerate time after initiation drug, are reported. Latency analyses using claims data where markers for confounders or patients' characteristics are missing or incomplete. Additionally, TC cases were identified using an ORD-validated algorithm with a high PPV, rather than relying on a single ICD-9 code. Similar to randomized clinical trials, the results of ITT analyses, where all cases and person-time are ascribed to patients' initiation drug, are reported. Latency analyses using claims data where markers for confounders and discontinuation of treatment bring about changes in exposure category. In the TOD analysis, time off drug could be informative about long-term residual effects of studied medications. Alternatively, time off drug could provide information about possible baseline imbalance in risk factors such as obesity, diabetes, and other medical conditions.

### Table 2: Histology of Confirmed Thyroid Cancer Cases by Exposure (N=72 Cases)

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Follicular</th>
<th>Papillary</th>
<th>Papillary with Follicular Variant</th>
<th>F/P</th>
<th>P/PF</th>
<th>F/P/PF</th>
<th>Papillary with Other Variant</th>
<th>PF/PO</th>
<th>Medullary</th>
<th>Unknown</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liraglutide</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Exenatide</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>DPP-4 Inhibitors (Saxagliptin/Saxagliptin)</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Linagliptin</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Metformin</td>
<td>2</td>
<td>23</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>Sulfonamide (Glyburide/Glipizide/Glimepiride)</td>
<td>2</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Pioglitazone</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Grand Total</td>
<td>5</td>
<td>46</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>72</td>
</tr>
</tbody>
</table>

**Note:** Exposure is based on first initiation.

**Abbreviations:** DPP-4i, dipeptidyl peptidase-4 inhibitor; F/P, follicular and papillary; P/PF, papillary and papillary with follicular variant; F/P/PF, follicular, papillary, and papillary with follicular variant; PF/PO, papillary with follicular variant and papillary with other variant.
between study groups. The consistency of the findings from each of these multiple approaches and analyses yielded similar results, adding weight to the conclusions.

This study also had several limitations. Evaluation of single-therapy is limited because, while all patients were initiators of an AD, many were already on other therapies, some of which continued during the study period; even many naïve initiators were likely to add or switch AD therapies during follow-up. While TC outcomes were identified using an algorithm with a high PPV in the validation sub-study, only patients who had undergone a thyroidectomy were classified as cases. In the presence of different distributions of tumor size by exposure, this may be a source of differential misclassification, although the apparent sensitivity of the algorithm was high and the algorithm performed similarly for both microcarcinomas and larger tumors. Inadequate access to medical records reduced the number of cases with pathology data to the validation subset, restricting the ability to identify cancer and not actual use of AD therapies, although this approach may be a source of differential misclassification, although the apparent sensitivity of the algorithm was high and the algorithm performed similarly for both microcarcinomas and larger tumors. Inadequate access to medical records reduced the number of cases with pathology data to the validation subset, restricting the ability to identify cancer type (e.g., medullary) and classify cases by tumor size.

Exposure assessment was based on pharmacy claims and not actual use of AD therapies, although this approach has been validated and widely used.\textsuperscript{36,37} This potential discrepancy between the study definition of drug exposure and actual drug exposure is more likely in the ITT analysis where a single dispensing is adequate to establish an exposure category versus that of the TOD analysis where persons who repeatedly refill their medication (and therefore are more likely to be taking it) contribute more person-time to follow-up. Actual use of the drug is inferred from dispensings and may result in some misclassification of exposure, which is expected to be non-differential with respect to TC.

Median follow-up time was approximately 17 months, which may be too short to identify all potential cases since TC clinical development may require years. The broad inclusion of baseline variables in the PS directly addresses many known risk factors for TC by explicitly including them in the PS and thereby balancing them across the liraglutide and comparator cohorts. In addition, many of the risk factors for TC that were not explicitly included in the PS had proxy variables present among claims data that were included. Only a few variables that potentially represent unmeasured and therefore, unaddressed, confounding variables (at least among the known risk factors for TC) remain unaccounted for, most notably obesity. Further, the potential confounding effects of such variables are limited in the context of a PS-matched study, since such confounding would need to operate independently of the many variables that were included (and are therefore balanced between cohorts).\textsuperscript{20} Study results may be a consequence of uncontrolled confounding associated with the inability to adjust adequately for surveillance bias.

Liraglutide initiators are markedly different from initiators of other ADs, chiefly metformin. Propensity score-matched comparisons were made, for example, between liraglutide initiators and those metformin initiators who resemble liraglutide initiators. Nevertheless, for the liraglutide-metformin cohort in particular, because of the number of metformin users in the ORD, it was possible to identify sufficient PS-matched metformin initiators. Comparisons between certain drug cohorts should be made thoughtfully because remaining unmeasured differences in risk factors are more probable with some comparators.

In summary, improved study designs are warranted to ascertain any true causal associations for newly marketed drugs, like liraglutide, in the presence of surveillance bias. A number of analyses were conducted and were generally consistent with no increase in risk of TC among liraglutide

<table>
<thead>
<tr>
<th>Table 3 Proportion of Papillary Thyroid Microcarcinomas by Exposure for All Chart-Adjudicated Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of Cases Exposed</strong></td>
</tr>
<tr>
<td>All Comparator Drugs</td>
</tr>
<tr>
<td>Liraglutide</td>
</tr>
<tr>
<td>Exenatide</td>
</tr>
<tr>
<td>DPP-4 Inhibitors (Sitagliptin/Saxagliptin/ Linagliptin)</td>
</tr>
<tr>
<td>Metformin</td>
</tr>
<tr>
<td>Sulfonylureas (Glyburide/Glipizide/Glimepiride)</td>
</tr>
<tr>
<td>Pioglitazone</td>
</tr>
</tbody>
</table>

**Abbreviations:** PTMC, papillary thyroid microcarcinoma; mm, millimeters; CI, confidence interval; DPP, dipeptidyl peptidase-4.
initiators. Sensitivity analyses support the interpretation that the observed elevated point estimates likely resulted from surveillance bias.

**Data Sharing Statement**
Research data are proprietary and cannot be shared.

**Acknowledgments**
Findings from this study were presented at the 34th International Conference on Pharmacoepidemiology and Therapeutic Risk Management in Prague, Czech Republic as a poster presentation: Kathleen Mortimer, Donnie Funch, Li Zhou, Eva Ng, Atheline Major-Pedersen, Heidrun Bosch-Traberg, Helge Gydesen, David Dore. Propensity Score (PS) Matched Approaches to Estimate the Association Between Liraglutide Use and Thyroid Cancer (TC). Publication #1072. https://pharmae coopi.org/pub?id=F17BF46A-D4E3-F6F2-827A- D6862969DD8B. The authors would like to acknowledge the contributions of Nicole Brooks for her project oversight and Tamar Aroyan and Danielle Cupka for their technical assistance in preparing this manuscript.

**Author Contributions**
All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

**Funding**
This study was funded by a research contract with Novo Nordisk A/S.

**Disclosure**
All Optum-affiliated researchers (Funch, Mortimer, Ziyadeh, Seeger, Zhou, Ng, Dore) received support from this contract. As part of this contract, Optum Epidemiology retains the right to publish all results. Funch, Ziyadeh, and Seeger were employed by Optum and hold stock/stock options in the parent company of Optum (UnitedHealth Group, Inc.). Seeger was also involved in a research contract with Boehringer-Ingelheim, during the conduct of the study. Dore was an employee and shareholder of UnitedHealth Group. Ross received payment from Novo Nordisk for case adjudication and for consulting services provided to Optum Epidemiology. He also reports that UBC Corp, which manages the safety monitoring for the Medullary Thyroid Cancer Consortium, pays him personal consultation fees for his time. Major-Pedersen, Bosch-Traberg, and Gydesen are employees of Novo Nordisk A/S. The authors report no other conflicts of interest in this work.

**References**


