This manuscript was accepted for publication in Diabetes Care on 1 May 2025. The final version of the paper is available on the Diabetes Care website at https://doi.org/10.2337/dc25-0154.

Risk of thyroid tumors with GLP-1 receptor agonists: a retrospective cohort study

Daniel R Morales, PhD,[1] Fan Bu, PhD,[2] Benjamin Viernes, PhD,[3,4] Scott L DuVall, PhD,[3,4] Michael E Matheny, MD,[5,6] Katherine R Simon, AA,[5,6] Thomas Falconer, MS,[7] Lauren R Richter, MD,[7,8] Anna Ostropolets, MD, PhD,[7] Wallis CY Lau, PhD,[9,10,11,12] Kenneth KC Man, PhD,[9,10,11,12] Shounak Chattopadhyay, PhD,[13] Nestoras Mathioudakis, MD,[14] Evan Minty, MD,[15] Akihiko Nishimura, PhD,[16] Feng Sun, PhD,[17] Can Yin, MS,[18] Sarah L Seager, BS,[18] Yi Chai, PhD,[9] Jin J Zhou, PhD,[13,19] Yuan Lu, ScD,[20] Carlen Reyes, MD, PhD,[21] Andrea Pistillo, MSc,[21] Talita Duarte-Salles, PhD,[21] Clair Blacketer, MPH,[22] Martijn J Schuemie, PhD,[13,23] Patrick B Ryan, PhD,[7,23] Harlan M Krumholz, MD,[20,24,25] George Hripcsak, MD,[7] Rohan Khera, MD,[20,24,26] Marc A Suchard, MD, PhD,[3,13,27,28]

1. Division of Population Health and Genomics, School of Medicine, University of Dundee, Dundee, United Kingdom
2. Department of Biostatistics, University of Michigan at Ann Arbor, Ann Arbor, MI, USA
3. Veterans Affairs Informatics and Computing Infrastructure, United States Department of Veterans Affairs, Salt Lake City, UT, USA
4. Department of Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA
5. Tennessee Valley Healthcare System, Veterans Affairs Medical Center, Nashville, TN, USA
6. Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA
7. Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, USA
8. Division of Pediatric Endocrinology, Diabetes, & Metabolism, Department of Pediatrics, Columbia University Irving Medical Center, New York, NY, USA
9. Centre for Safe Medication Practice and Research, Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, Hong Kong
10. Laboratory of Data Discovery for Health (D24H), Hong Kong Science Park, Hong Kong, China
11. Research Department of Practice and Policy, School of Pharmacy, University College London, London, United Kingdom
12. Centre for Medicines Optimisation Research and Education, University College London Hospitals NHS Foundation Trust, London, United Kingdom
13. Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA, USA
14. Division of Endocrinology, Diabetes, and Metabolism, Johns Hopkins University School of Medicine, Baltimore, MD, USA
15. Faculty of Medicine, O'Brien Institute for Public Health, University of Calgary, Calgary, AB, Canada

16. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA
17. Department of Epidemiology and Biostatistics, Peking University, Beijing, China
18. DTA&AI, Real World Solutions, IQVIA, Durham, NC, USA
19. Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA
20. Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University, New Haven, CT, USA
21. Real-World Epidemiology Research Group, Fundaci Institut Universitari per a la recerca a l'Atenci Primria de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain
22. Observational Health Data Analytics, Janssen Research and Development, LLC, Titusville, NJ, USA
23. Epidemiology, Office of the Chief Medical Officer, Johnson & Johnson, Titusville, NJ, USA
24. Center for Outcomes Research and Evaluation (CORE), Yale New Haven Hospital, New Haven, CT, USA
25. Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA
26. Section of Health Informatics, Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA
27. Department of Biomathematics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA
28. Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Running Title: Thyroid tumors and GLP-1 receptor agonists

Corresponding author

Marc A Suchard, Departments of Biomathematics, Biostatistics and Human Genetics 6558 Gonda Building, 695 Charles E. Young Drive, South Los Angeles, CA 90095-1766 (310) 825-7442; msuchard@ucla.edu; @suchard_group

Twitter summary

Do GLP-1 receptor agonists increase thyroid tumor risk? In this federated new user comparator cohort study, GLP-1 receptor agonists were not associated with an increased risk of thyroid tumors.

Word count: 3279

Number of tables: 3 Number of figures: 1

Number of supplementary digital content files: 1

ABSTRACT

Objective: To assess the association between glucagon-like peptide-1 receptor agonists (GLP-1RA) use and risk of incident thyroid tumors.

Research Design and Methods: Retrospective new user active comparator cohort study using international administrative claims and electronic health record databases. Participants included T2DM patients with prior metformin therapy initiating GLP-1RA versus new users of sodium-glucose transport protein 2 inhibitors (SGLT2I), dipeptidyl peptidase IV inhibitors (DPP4I) and sulfonylureas (SU). The outcome was incident thyroid tumor and thyroid malignancy. Propensity score (PS) matching, and stratification were used for confounding adjustment with intention-to-treat and on-treatment strategy. Cox regression was used to estimate hazard ratios (HR) pooled using random-effects meta-analysis. Unmeasured confounding was evaluated using negative outcomes, with calibration of the HR.

Results: 460,032 users of GLP-1RA, 717,792 users of SGLT2I, 2,055,583 users of DPP4I and 1,119,868 users of SU were included. Only US cohorts passed study diagnostics. Thyroid tumor incidence ranged from 0.88 to 1.03 per 1000 person years in GLP-1RA cohorts. GLP-1RA exposure was not associated with an increased risk of thyroid tumors compared with SGLT2I, DPP4Is or SUs (Meta-analysis: GLP-1RA vs SGLT2I range HR 0.83 (0.57-1.27) to HR 0.95 (0.85–1.06); GLP-1RA vs SU range HR 0.95 (0.75-1.20) to HR 1.03 (0.87-1.23); GLP-1RA vs DPP4I range HR 0.78 (0.60-1.01) to HR 0.93 (0.83-1.04)). Analysis using thyroid malignancy, and including a 1-year lag period produced similar conclusions.

Conclusion: In T2DM patients initiating second-line treatments, we observed no increased risk of thyroid tumors with GLP-1RA exposure.

ARTICLE HIGHLIGHTS

- **Why did we undertake this study?**

Safety concerns have been raised over the risk of thyroid tumors with GLP-1 receptor agonist therapy in pre-clinical models, with different international regulatory advice.

- **What is the specific question(s) we wanted to answer?**

What is the risk of incident thyroid tumor in people with T2DM initiating GLP-1 receptor agonists compared to DPP4Is, SGLT2Is and SUs?

- **What did we find?**

In this new user active comparator cohort design, no increased risk of thyroid tumors was observed with GLP-1 receptor agonist therapy in people with T2DM.

- **What are the implications of our findings?**

The findings do not suggest that the increased risk of thyroid tumors observed in pre-clinical models translates to humans.

INTRODUCTION

Glucagon-like peptide-1 receptor agonists (GLP-1RA) are incretin hormone mimetics indicated for the treatment of type 2 diabetes mellitus (T2DM) and obesity (semaglutide and liraglutide) that stimulate glucose-dependent pancreatic insulin secretion and suppress glucagon secretion through direct activation of the GLP-1 receptor.[1] Evidence suggests that GLP-1RA reduce cardiovascular risk and reduce the progression of renal disease in patients with T2DM.[2-6] As a result, prescribing of GLP-1RA as second-line T2DM treatment has increased since their introduction, particularly in the US and are also used first-line in some patients.[7,8]

Rodent studies have reported that GLP-1RA may cause thyroid C-cell hyperplasia, adenomas and carcinomas, such as medullary thyroid cancer, in a dose-dependent and time-dependent manner.[9-11] However, the level of GLP-1 receptor expression in thyroid tissue is species-specific with higher levels of receptor expression reported in rodents compared with humans.[9] GLP1-RA exposure in rodent studies are also administered at high dose. It therefore remains unclear whether such findings are relevant to humans as no risk has been reported in existing clinical trials.

GLP-1RA including liraglutide, dulaglutide, exenatide and semaglutide have been contraindicated by the US Food and Drug Administration (FDA) in patients with a personal or family history of medullary thyroid cancer and multiple endocrine neoplasia type 2 (MEN2).[12-15] US GLP-1RA product information also contains warnings raising awareness about the symptoms of thyroid tumors. In contrast, no similar contraindications or warnings exist in product information in Europe where this potential risk is monitored as part of routine pharmacovigilance.

A nested case-control study using French national health insurance data examined the association between GLP-1RA and thyroid cancer in people with T2DM treated with second-line therapy.[16] That study reported a 58% increased risk of thyroid cancer with exposure to GLP-1RA for one to three years, and a 36% increased risk with exposure longer than three years. However, several commentaries have raised concerns around the study design and the potential for residual unmeasured confounding, with a call to provide further information to confirm or refute these findings.[17-20]

The Large-scale Evidence Generation and Evaluation across a Network of Databases for Type 2 Diabetes Mellitus (LEGEND-T2DM) initiative aims to generate reliable evidence on the effects of second-line T2DM glucose-lowering agents using observational healthcare data by executing a series of comparative observational studies specifically aimed at minimizing bias and increasing reproducibility.[21] The aim of this analysis was to determine whether T2DM patients initiating GLP-1RA have a differential risk of incident thyroid tumors compared to patients initiating other second-line agents consisting of sodium-glucose transport protein 2 inhibitors (SGLT2I), dipeptidyl peptidase IV inhibitors (DPP4I) or sulfonylureas (SU).

METHODS

Study design and data sources

We conducted an active comparator, new-user cohort design [22,23] in patients identified in routinely collected electronic health records (EHRs) and claims data from six national-level and four health-system datasets from the US (US Open Claims, Optum EHR, Optum DOD, CCAE, MDCD, MDCR, VA), Germany (IQVIA Germany), Spain (SIDIAP), and the United Kingdom (IMRD UK). All LEGEND-T2DM data sources were previously standardized to the Observational Health Data Sciences and Informatics (OHDSI)'s Observational Medical Outcomes Partnership (OMOP) common data model (CDM) version 5, which mapped international coding systems into standard vocabulary concepts. The use of the OMOP CDM allows a federated analysis of data and standardised analytics without patient-level data sharing. All data partners received institutional approval or exemption for their participation. Details of data sources are presented in Supplementary Table 1 and the study package is available at https://github.com/ohdsi-studies/LegendT2dm. This package provides end-to-end reproducible study code under our pre-specified protocol [21] and is built upon the open-source Health Analytics Data-to-Evidence Suite (HADES) software library [24] using the R statistical programming language.

Study population

The study population included all patients in a data source who met the inclusion criteria for the second-line T2DM agent exposure cohorts. Three cohorts were created from the pairwise comparisons using people initiating GLP-1RA as the target cohort and those initiating DPP4I, SU and SGLT2I as the comparator cohorts, respectively. These cohorts consisted of patients with T2DM, and prior metformin monotherapy, who initiated treatment with one of four drug classes: GLP-1RA, sodium-glucose transport protein 2 inhibitors (SGLT2I), dipeptidyl peptidase IV inhibitors (DPP4I) and sulfonylureas (SU). The index date for each patient entering cohort was defined by the first exposure to any drug ingredient in the four drug classes. The inclusion criteria consisted of:

- T2DM diagnosis and no type 1 or secondary diabetes mellitus diagnoses before the index date
- At least 1 year of observation time before the index date
- No prior drug exposure to a comparator second-line or other antihyperglycemic agent (i.e. thiazolidinediones, acarbose, acetohexamide, bromocriptine, glibornuride, miglitol and nateglinide)
- < 30 days of insulin exposure before the index date
- > 90 days of continuous metformin exposure before the index date

A schematic of the LEGEND-T2DM new-user cohort design is summarised in Figure 1.

Exposures

Exposures consisted of the 22 drug ingredients that comprise several classes of T2DM second-line therapy:GLP-1RA, DPP4I, SGT2I and SU (Supplement Section 9.1 for included drug substances and codes).

Outcome

The outcome was thyroid tumors defined as an incident record for any type of neoplasm of the thyroid gland. This outcome definition included both malignant and benign thyroid tumors (Supplement Section 9.2). Patients who experienced the outcome before their index date were excluded from analysis. A post-hoc analysis was performed for malignant thyroid tumor.

Analysis

To adjust for measured confounding and to ensure balance between target and comparator cohorts, large-scale propensity score models were constructed for each comparison and data source using a consistent data-driven process through regularised regression.[25,26] This process engineers a large set of predefined baseline patient characteristics (>90,000 variables) that included demographics, comorbidities, concomitant medication use, and healthcare utilization to provide improved covariate balance between target and comparator cohorts and reduced residual systematic bias of effect estimates than smaller models.[26,27] We used L_1 regularization through the high-performance Cyclops package [28] in HADES with 10-fold cross-validation to choose the appropriate strength of regularization across covariates. All covariates were identified within the 365 days before and including the index date. Given the subcutaneous route of administration of GLP-1RA compared with other drugs administered orally, device codes representing needles and associated health management encounters have been excluded from propensity score construction. We then used the resulting propensity score estimates for variable-ratio matching and 10-partition stratification of patients in the target and comparator cohorts.

Cox proportional hazards models were used to estimate hazard ratios (HRs) for the risk of thyroid tumors in each data source, and estimates across non-overlapping data sources were meta-analysed using a random-effects meta-analysis.[29] Potential residual bias from unmeasured confounding was evaluated by using approximately 100 negative outcome control experiments, where the null hypothesis of no effect was believed to be true. We first identified candidates for these control outcomes through a data-rich algorithm that identifies prevalent condition concept occurrences lacking evidence of association with exposures in published literature, drug-product labelling, and spontaneous reports. These candidates were then screened by clinical review to select the final set.[30] The empirical null distributions from these negative outcome control experiments were used to calibrate the HR estimate, 95% CI, and p-value for each study.[31,32]

Blinded to the results, study diagnostics were first evaluated for all comparisons. These diagnostics consisted of:

- Minimum detectable risk ratio (MDRR) as a typical proxy for power, with event counts <5 undisclosed for governance reasons,
- Propensity score distributions to evaluate empirical equipoise and population generalizability,
- Plots of standardised mean differences to evaluate cohort balance before and after PS-adjustment across the extensive range of patient characteristics,
- Negative control calibration plots to assess residual bias, and
- Kaplan-Meier plots to assess proportional hazard assumptions.

Comparisons were deemed acceptable if, after propensity score adjustment, they returned an MDRR <10, an absolute standardised mean differences between characteristics <0.15, an empirical equipoise >0.25[33,34, and no observed crossing of estimated survival probability 95% CI bands. If the comparisons failed diagnostics related to covariate imbalance or systematic error (suggesting potential residual confounding) they were not included in the meta-analysis but are still reported for completeness. Patient follow-up was censored at the end of their time-at-risk (TAR) or of data source observation period. Each comparison was executed using two different TAR definitions reflecting different causal contrasts [35]:

- Intent-to-treat TAR where follow-up was censored at the end of observation.
- On-treatment TAR where follow-up was censored at treatment discontinuation or at escalation with other T2DM agents.

Additional analysis

We performed two post-hoc analyses in. First, we specifically examined malignant thyroid tumour as an outcome. Second, we incorporated a 1-year exposure lag period before outcomes could be counted to account for potential latency period in cancer development for the ITT analyses.

Ethics

All data partners received institutional review board approval or waiver in accordance with their institutional governance guidelines.

RESULTS

The cumulative number of T2DM patients included in each cohort varied by the analytical approach and therapeutic class. Unadjusted incidence rates and hazard ratio estimates for each individual data source are reported in Supplement Sections 5.1 to 5.3. Baseline patient characteristics for new users of GLP-1RA, SGLT2I, DPP4I and SU for each data source before and after propensity score matching and stratification are shown in Section 6 of the Supplement (Tables 14 to 42). For example, Table 1 summarise baseline patient characteristics for GLP1RA and SGLT2I new-users in

the Open Claims data source and covariate balance after propensity score matching and stratification. The number of patients included in the meta-analysis across databases ranged from: 316,587 to 460,032 initiating GLP-1RA; 713,801 to 717,792 initiating SGLT2I; 1,990,074 to 2,055,583 initiating DPP4I; and 1,104,270 to 1,119,868 initiating SU. A total of 51 of 120 pairwise comparisons passed diagnostics and were considered acceptable for inclusion in the meta-analyses (a study diagnostic summary is reported in Supplement Section 7 and full details of covariate balance, systematic error and time at risk in Section 8). All data sources included in the meta-analysis were from the US with the Open Claims database contributing approximately 85% of patients.

GLP-1RA vs SGLT2I

The aggregate crude incidence of thyroid tumors ranged from 0.88 (95%CI 0.82-0.94) to 1.00 (95%CI 0.89-1.13) per 1000 person years (pys) in people initiating GLP-1RA and 0.82 (95%CI 0.79-0.86) to 0.94 (95%CI 0.86-1.02) per 1000 pys in people initiating SGLT2I (Table 2). Cohort balance and systematic error diagnostics across data sources the GLP-1RA vs SGLT2I cohorts are shown in Supplement Section 8.1. There was no statistically significant increased risk of thyroid tumors for GLP-1RA exposure compared to SGLT2I exposure for any of the analyses (Table 3).

GLP-1RA vs DPP4I

The aggregate crude incidence of thyroid tumors ranged from 0.96 (95%CI 0.90-1.03) to 1.03 (95%CI 0.91-1.15) per 1000 pys in people initiating GLP-1RA and 0.94 (95%CI 0.91-0.97) to 1.13 (95%CI 1.06-1.20) per 1000 pys in people initiating DPP4I (Table 2). Cohort balance and systematic error diagnostics across data sources the GLP-1RA vs DPP4I cohorts are shown in Supplement Section 8.2. There was no statistically significant increased risk of thyroid tumors for GLP-1RA exposure compared to DPP4I exposure for any of the analyses (Table 3).

GLP-1RA vs SU

The aggregate crude incidence of thyroid tumors ranged from 0.88 (95%CI 0.82-0.94) to 0.95 (95%CI 0.84-1.08) per 1000 pys in people initiating GLP-1RA and 0.72 (95%CI 0.70-0.73) to 0.81 (95%CI 0.77-0.85) per 1000 pys in people initiating SU (Table 2). Cohort balance and systematic error diagnostics across data sources the GLP-1RA vs SU cohorts are shown in Supplement Section 8.3. There was no statistically significant increased risk of thyroid tumors for GLP-1RA exposure compared to SU exposure for any of the analyses (Table 3).

Additional analysis

Post-hoc analysis specifically measuring hazard ratios for malignant thyroid tumor in databases available at the time did not alter the findings of the main analysis (Supplementary Tables 5-7). The same was observed when a one-year lag exposure period was included following cohort entry (Supplementary Tables 8-10).

DISCUSSION

Using a federated network of healthcare databases, we examined the association between GLP-1RA exposure and incident thyroid tumors in a large number of T2DM patients initiating different second-line treatments. We observed no statistically significant increased risk of thyroid tumors among new users of GLP-1RA compared to new users of SGLT2I, DPP4I or SU therapy, with findings consistent across several analytical approaches.

Evidence suggests that GLP-1 receptor expression in thyroid tissue is species-dependent.[9] Studies in mice and rats have suggested that the formation of medullary thyroid tumors is a potential safety concern of GLP-1RA use in humans. One rodent study investigating GLP-1 receptor proliferative action in thyroid C-cells reported that GLP-1RA caused receptor activation on rodent thyroid C-cells, and that long-term activation (up to 2 years) was associated with increased levels of C-cell proliferation and benign and malignant tumors formation in rats and mice.[9] Intriguingly, the incidence of thyroid adenomas and malignant carcinomas was different across rodent species treated with liraglutide, with rats having a higher incidence of both thyroid adenomas and carcinomas whilst in mice an increase of adenomas was observed. However, no similar C-cell proliferative effects were observed in nonhuman primates.[9] Another study reported diffuse C-cell hyperplasia and an increase in the number of adenomas and carcinomas in rats treated with dulaglutide over 93 weeks whilst a further study of rats reported an increase in C-cell hyperplasia with exenatide.[10,11] However, those studies used doses that exceed equivalent doses in humans and may not be representative. A meta-analysis of 45 clinical trials including 52,600 patients with GLP-1RA use also reported no statistically significant increased risk of thyroid cancer or thyroid masses.[36] However, the authors similarly acknowledged limitations in sample size and imprecise effect estimates.

The nested case-control study by Bezin, *et al.* demonstrated a 46% increased risk of thyroid carcinoma with current GLP-1RA use and a 58% increased risk with one to three years of cumulative GLP-1RA use. This association remained elevated with cumulative use of longer than three years. That study examined exenatide, liraglutide, semaglutide, albiglutide, and dulaglutide, which were the GLP-1RA available in France during the study period. Whilst that study adjusted for important potential confounders, the extent of confounding adjustment was potentially limited with exclusion of patients if the number of dispensations were less than 3 (or 2 if at least one was described a large package) within a 1-year period between 2006 and 2018. The association was also compared with DPP4I use with a smaller association still observed seen with cumulative DPP4I use of longer than three years. Comparison with SGLT2I was not performed because SGLT2I were not marketed in France during the study period. Several commentaries raised caution over the interpretation of those results highlighting the potential for further unmeasured confounding, uncertainty around outcome validity, and not demonstrating further increasing risk with increasing cumulative exposure.[17-20] Our study differs in that we applied an active comparator new user cohort study design across multiple databases, and had more robust approaches to account for observed and unobserved confounding. We did not examine differences in dosing which remains a limitation of our study.

A recent cohort study by Bea, *et al.* analysed claims data from the Korean National Health Insurance Service from 21,722 new users of GLP-1RA with T2DM compared to new users of SGLT2I.[37] That study adjusted for a larger number of potential confounding variables and observed no increased risk of incident thyroid cancer with GLP-1RA (primary analysis HR 0.98, 95% CI 0.62-1.53). Whilst our results are more consistent to the findings of Bea, *et al.* our study differs in that our outcome include both benign and malignant thyroid tumors that could limit our study's direct comparability. However, our outcome is relevant because safety concerns are based upon the potential extrapolation of effects in pre-clinical models, which show an increase in both benign and malignant thyroid tumors, with the former being more frequent.[9,10]

Strengths of our study include using the LEGEND approach aimed at minimizing bias and increasing reproducibility and the use of an active comparator new user design across multiple data sources. The large network of databases increased precision and provided the ability to examine consistency of results. Large-scale propensity score models were used for improved confounding adjustment and implemented in a standardised way across each database. We examined propensity score diagnostics blinded to the treatment effect results to check model performance and thereby determine whether treatment effects were at risk of bias.

There were several limitations to our study. Despite the use of large-scale propensity score adjustment, 69 of 120 cohort comparisons failed study diagnostics, predominantly relating to covariate imbalance and potential residual confounding, which affected all European databases. The results of these analyses are reported for completeness in the supplement but should be treated with caution. For this reason, only estimates from cohorts that did pass study diagnostics have been included in the meta-analysis that were all from US databases. However, a recently reported Scandinavian cohort study also found no evidence of an increased risk of thyroid cancer with GLP-1RA use.[38] Furthermore, in the cohorts that passed study diagnostics, residual unobserved confounding remains possible, although examination of negative control outcome experiment plots and their use in calibrating the effect estimates reduces this possibility. Our retrospective cohort study was not able to include tirzepatide within the class level exposure due to it being only recently available. We did not adjust for post-baseline risk factors for adherence in our on-treatment analysis due to the tightly defined cohorts using several active comparators producing consistent results. We did not condition on filling a second prescription to avoid immortal time bias. However, 30-days (the usual time of second prescriptions) typically fell lower than the 10th percentile of patients' time-at-risk. DPP4Is also enhance levels of GLP1 but to a lower extent than GLP1-Ras.[39] Whilst using DPP4Is as a control may reduce the size of any potential risk difference, we observed similar and consistent results using SGLT2I and SU as comparators suggesting this had negligible impact

We used a composite outcome for our primary analysis consisting of benign and malignant thyroid tumors given potential GLP-1RA effects on both seen in animal models. We also examined the association with malignant thyroid tumors We identified thyroid tumors using clinical codes and we do not have pathological confirmation of the diagnosis or include procedures in our definition. We excluded people with a prior history of thyroid tumors only and people with metastases to the

thyroid may have been included. However, this presentation is rare.[40] Whilst the ITT analysis allows the observance of outcomes that occur after the patient has stopped taking the drug, a cumulative exposure analysis was not specifically performed to directly compare our results with those from Bezin et al. Cohort follow-up began at drug initiation and cases occurring shortly after initiation may not be causally related although this may be non-differential given the active comparator is also a second-line agent. Like other observational studies, drug exposure is based upon prescription and dispensing information only and that treatment was accurately characterised. The study assumes data are valid and that an individual does not have the condition if not recorded. We included only individuals whose baseline inclusion criteria could be characterised using all features, except for demographics, to indicate the presence or absence of health records in a given time-period. Death data was not comprehensively recorded in several included US databases precluding a competing risk analysis. However, this is unlikely to affect the findings due to robust assessment of cohort diagnostics, consistent findings and the lack of significant known mortality differences between the comparators. Bias may be introduced if differential screening for thyroid tumors prior to therapy occurred. However, this risk is minimised through the use of large-scale confounding adjustment and results using a one-year lag exposure period did not change our conclusion.

Thyroid tumors have been identified as an important safety concern included in risk minimisation plans for GLP-1RA use following drug approval. The study by Bezin, *et al.* suggests an association may exist in humans, leading to a review of this potential safety concern in Europe. However, whilst acknowledging the differences between studies, we did not observe an association between GLP-1RA and thyroid tumors. Instead, our findings more closely align with the conclusions by Bea, *et al.*, suggesting no further warnings are warranted. Despite these findings, the risk of thyroid tumors in people prescribed GLP-1RA is likely to remain as an important potential concern for the foreseeable future until further information can be generated through ongoing post-approval safety studies (EUPAS registration EUPAS45172, EUPAS32646, EUPAS11850) and routine pharmacovigilance monitoring.

Conclusion

In this federated network study involving patients with T2DM initiating different second-line antihyperglycemic treatments, GLP-1RA exposure was not associated with an increased risk of thyroid tumors compared with SGLT2I, DPP4I or SU therapy.

DECLARATION OF INTERESTS

RK is an Associate Editor at JAMA and received support from the National Heart, Lung, and Blood Institute of the National Institutes of Health (under award K23 HL153775) and the Doris Duke Charitable Foundation (under award, 2022060). He also receives research support, through Yale, from Bristol-Myers Squibb, Novo Nordisk, and BridgeBio. He is a coinventor of U.S. Provisional Patent Applications 63/619,241, 63/606,203, 63/177,117, 63/346,610, 63/428,569, and 63/484,426, unrelated to current work. He is also a co-founder of Evidence2Health and Ensight-AI, both representing precision health platforms to improve evidence-based cardiovascular care. SLD reports grants from Alnylam Pharmaceuticals, Inc., AstraZeneca Pharmaceuticals LP, Biodesix, Inc, Celgene Corporation, Cerner Enviza, GSK PLC, Janssen Pharmaceuticals, Inc., Novartis International AG, Parexel International Corporation through the University of Utah or Western Institute for Veteran Research outside the submitted work. KKCM receives support from C W Maplethorpe Fellowship, National Institute of Health Research, UK, European Commission Framwork Horizon 2020, Hong Kong Research Grant Council, Innovation and Technology Commission of the Hong Kong Special Administration Region Government, outside the submitted work. In the past three years, HMK received expenses and/or personal fees from UnitedHealth, Element Science, Aetna, Reality Labs, Tesseract/4Catalyst, F-Prime, the Siegfried and Jensen Law Firm, Arnold and Porter Law Firm, and Martin/Baughman Law Firm. He is a co-founder of Refactor Health and HugoHealth, and is associated with contracts, through Yale New Haven Hospital, from the Centers for Medicare & Medicaid Services and through Yale University from Johnson & Johnson. Other authors declare no conflicts of interest

ACKNOWLEDGMENTS

This study was partially funded through the National Institutes of Health grants K23 HL153775, R01 LM006910, R01 HG006139 and R01 HL169954, and the US Department of Veterans Affairs under the research priority to "Put VA Data to Work for Veterans" (VA ORD 22-D4V). The funders had no role in the design and conduct of the protocol; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. DRM is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors contributed to the acquisition, interpretation of the data and critically revised the manuscript. RK and MAS conceived and designed the study. FB, RK, LSD, KL, JJZ, AA and MAS conducted the statistical analysis and DRZ drafted the manuscript. MAS is the study guarantor.

DATA AND RESOURCE AVAILABILITY

All data results and aggregate data are included in the article. This was a federated analysis and no raw data was transferred, which remains with the original data holders, which were used under license for the current study and therefore are not publicly available.

REFERENCES

1. Drucker DJ. Mechanisms of Action and Therapeutic Application of Glucagon-like Peptide-1. Cell Metab. 2018;27(4):740-756

2. Zinman B, Wanner C, Lachin JM, et al.. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med 2015;373:2117–28.

3. Wiviott SD, Raz I, Sabatine MS, et al.. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. reply. N Engl J Med 2019;380:1881–2.

4. Neal B, Perkovic V, Matthews DR, et al.. Canagliflozin and cardiovascular and renal events in type 2 diabetes. N Engl J Med 2017;377:2099.

5. Marso SP, Bain SC, Consoli A, et al.. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med 2016;375:1834–44.

6. Marso SP, Daniels GH, Brown-Frandsen K, et al.. Liraglutide and cardiovascular outcomes in type 2 diabetes. N Engl J Med 2016;375:311–22.

7. Khera R, Dhingra LS, Aminorroaya A, Li K, Zhou JJ, Arshad F, Blacketer C, Bowring MG, Bu F, Cook M, Dorr DA, Duarte-Salles T, DuVall SL, Falconer T, French TE, Hanchrow EE, Horban S, Lau WC, Li J, Liu Y, Lu Y, Man KK, Matheny ME, Mathioudakis N, McLemore MF, Minty E, Morales DR, Nagy P, Nishimura A, Ostropolets A, Pistillo A, Posada JD, Pratt N, Reyes C, Ross JS, Seager S, Shah N, Simon K, Wan EY, Yang J, Yin C, You SC, Schuemie MJ, Ryan PB, Hripcsak G, Krumholz H, Suchard MA. Multinational patterns of second line antihyperglycaemic drug initiation across cardiovascular risk groups: federated pharmacoepidemiological evaluation in LEGEND-T2DM. BMJ Med. 2023;2(1):e000651

8. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, Collins BS, Hilliard ME, Isaacs D, Johnson EL, Kahan S, Khunti K, Leon J, Lyons SK, Perry ML, Prahalad P, Pratley RE, Seley JJ, Stanton RC, Gabbay RA, on behalf of the American Diabetes Association. 9. Pharmacologic Approaches to Glycemic Treatment: Standards of Care in Diabetes-2023. Diabetes Care. 2023;46(Suppl 1):S140-S157.

9. Bjerre Knudsen L, 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):1473–86.

10. Byrd RA, Sorden SD, Ryan T, Pienkowski T, LaRock R, Quander R, et al. Chronic Toxicity and Carcinogenicity Studies of the Long-Acting GLP-1 Receptor Agonist Dulaglutide in Rodents. Endocrinology. 2015;156(7):2417–28.

11. Bulchandani D, Nachnani JS, Herndon B, Molteni A, Pathan MH, Quinn T, et al. Effect of exendin (exenatide)--GLP 1 receptor agonist on the thyroid and parathyroid gland in a rat model. Eur J Pharmacol. 2012;691(1–3):292–6.

12. Highlights of prescribing information VICTOZA. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2019/022341s031lbl.pdf. Accessed 21/10/2023.

13. Highlights of prescribing information TRILUCITY. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/125469s036lbl.pdf. Accessed 21/10/2023.

14. Highlights of prescribing information BYDUREON. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/022200s026lbl.pdf. Accessed 21/10/2023.

15. Highlights of prescribing information OZEMPIC. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/209637lbl.pdf. Accessed 21/10/2023.

16. Bezin J, Gouverneur A, Pénichon M, Mathieu C, Garrel R, Hillaire-Buys D, Pariente A, Faillie JL. GLP-1 Receptor Agonists and the Risk of Thyroid Cancer. Diabetes Care. 2023;46(2):384-390.

17. Endo M, Roth MY, Tylee TS, DeSantis A, Hirsch IB. Comment on Bezin et al. GLP-1 Receptor Agonists and the Risk of Thyroid Cancer. Diabetes Care 2023;46:384-390. Diabetes Care. 2023;46(5):e118.

18. Goldenberg RM, Jain AB. Comment on Bezin et al. GLP-1 Receptor Agonists and the Risk of Thyroid Cancer. Diabetes Care 2023;46:384-390. Diabetes Care. 2023;46(5):e117.

19. Smits MM, van Raalte DH. Comment on Bezin et al. GLP-1 Receptor Agonists and the Risk of Thyroid Cancer. Diabetes Care 2023;46:384-390. Diabetes Care. 2023;46(5):e120.

20. Mañas-Martinez AB, Gimeno-Orna JA. Comment on Bezin et al. GLP-1 Receptor Agonists and the Risk of Thyroid Cancer. Diabetes Care 2023;46:384-390. Diabetes Care. 2023;46(5):e119.

21. Khera R, Schuemie MJ, Lu Y, Ostropolets A, Chen R, Hripcsak G, Ryan PB, Krumholz HM, Suchard MA. Large-scale evidence generation and evaluation across a network of databases for type 2 diabetes mellitus (LEGEND-T2DM): a protocol for a series of multinational, real-world comparative cardiovascular effectiveness and safety studies. BMJ Open. 2022;12(6):e057977.22.

Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. Nature Reviews Rheumatology 2015;11:437–41.

23. Johnson ES, Bartman BA, Briesacher BA, et al. The incident user design in comparative effectiveness research. Pharmacoepidemiology and Drug Safety 2013;22:1–6.

24. Schuemie M, Reps J, Black A, Defalco F, Evans L, Fridgeirsson E, Gilbert JP, Knoll C, Lavallee M, Rao GA, Rijnbeek P, Sadowski K, Sena A, Swerdel J, Williams RD, Suchard M. Health-Analytics Data to Evidence Suite (HADES): Open-Source Software for Observational Research. Stud Health Technol Inform. 2024;310:966-970.

25. Rosenbaum PR Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70: 41-55.

26. Tian Y Schuemie MJ Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. Int J Epidemiol. 2018; 47: 2005-2014.

27. Schneeweiss S, Rassen JA, Glynn RJ, et al. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology 2009;20:512.

28. Suchard MA, Simpson SE, Zorych I, Ryan P, Madigan D. Massive parallelization of serial inference algorithms for a complex generalized linear model. ACM Trans Model Comput Simul. 2013;23:10.1145/2414416.2414791.29. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials 1986;7:177–88.

30. Voss EA, Boyce RD, Ryan PB, et al. Accuracy of an automated knowledge base for identifying drug adverse reactions. Journal of Biomedical Informatics 2017;66:72–81.

31. Schuemie MJ, Hripcsak G, Ryan PB, et al. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. Proceedings of the National Academy of Sciences of the United States of America 2018;115:2571–7.

32. Schuemie MJ Ryan PB DuMouchel W Suchard MA Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Stat Med. 2014; 33: 209-218.

33. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Statistics in Medicine 2009;28:3083–107.

34. Walker A, Patrick A, Lauer M, Hornbrook M, Marin M, Platt R, Roger V, Stang P, Schneeweiss S. A tool for assessing the feasibility of comparative effectiveness research. Comparative Effectiveness Research. 2013;3:11-20.

35. Hernán MA, Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. Clinical Trials 2012;9:48–55.

36. Hu W, Song R, Cheng R, Liu C, Guo R, Tang W, Zhang J, Zhao Q, Li X, Liu J. Use of GLP-1 Receptor Agonists and Occurrence of Thyroid Disorders: a Meta-Analysis of Randomized Controlled Trials. Front Endocrinol (Lausanne). 2022;13:927859.

37. Bea S, Son H, Bae JH, Cho SW, Shin JY, Cho YM. Risk of thyroid cancer associated with glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in patients with type 2 diabetes: A population-based cohort study. Diabetes Obes Metab. 2023 Sep 21. doi: 10.1111/dom.15292.

38. Pasternak B, Wintzell V, Hviid A, Eliasson B, Gudbjörnsdottir S, Jonasson C, Hveem K, Svanström H, Melbye M, Ueda P. Glucagon-like peptide 1 receptor agonist use and risk of thyroid cancer: Scandinavian cohort study. BMJ. 2024 Apr 10;385:e078225. doi: 10.1136/bmj-2023-078225. PMID: 38683947; PMCID: PMC11004669.

39. Gilbert MP, Pratley RE. GLP-1 Analogs and DPP-4 Inhibitors in Type 2 Diabetes Therapy: Review of Head-to-Head Clinical Trials. Front Endocrinol (Lausanne). 2020 Apr 3;11:178. doi: 10.3389/fendo.2020.00178.

40. Orlandi AM, Alcaraz G, Bielski L, Brenta G, Jozami LC, Cavallo A, Guerra J, Zund S; Thyroid Department of Sociedad Argentina de Endocrinología y Metabolismo. Thyroid gland: a rare site of metastasis. Endocrine. 2024 Jan 15. doi: 10.1007/s12020-023-03626-x. Epub ahead of print.

Table 1. Baseline patient characteristics for GLP-1RA (T) and SGLT2I (C) new-users in the Open Claims data source. We report proportion of initiators satisfying selected baseline characteristics and the standardized difference of population proportions (StdDiff) before and after propensity score adjustment.

	Before adjustment			After matching			After stratification		
Characteristic	T (%)	C (%)	StdDiff	T (%)	C (%)	StdDiff	T (%)	C (%)	StdDiff
Age group									
15 - 19	0.2	0.1	0.03	0.2	0.1	0.01	0.2	0.1	0.01
20 - 24	0.8	0.4	0.05	0.6	0.6	0.00	0.6	0.5	0.01
25 - 29	1.4	0.7	0.07	1.1	1.1	0.00	1.1	1.0	0.01
30 - 34	2.8	1.7	0.07	2.2	2.3	0.00	2.1	2.1	0.00
35 - 39	5.1	3.5	0.08	4.3	4.4	0.00	4.1	4.1	0.00
40 - 44	8.3	6.2	0.08	7.5	7.3	0.01	7.1	7.0	0.00
45 - 49	11.9	9.9	0.06	11.3	11.0	0.01	10.8	10.5	0.01
50 - 54	15.4	13.8	0.05	15.0	15.0	0.00	14.4	14.4	0.00
55 - 59	16.4	16.6	0.00	16.9	16.8	0.00	16.9	16.8	0.00
60 - 64	15.0	16.7	-0.05	15.8	15.8	0.00	15.9	16.1	0.00
65 - 69	11.1	13.1	-0.06	12.0	12.1	0.00	12.4	12.3	0.00
70 - 74	6.7	8.9	-0.08	7.5	7.7	-0.01	7.9	8.2	-0.01
75 - 79	3.3	5.3	-0.10	3.7	3.9	-0.01	4.4	4.6	-0.01
80 - 84	1.4	2.9	-0.10	1.7	1.9	-0.01	2.2	2.4	-0.01
85 - 89	0.1	0.2	-0.03	0.1	0.2	-0.01	0.1	0.2	-0.02
Gender: female	59.8	45.6	0.29	50.9	51.0	0.00	50.9	51.0	0.00
Medical history: General									
Acute respiratory disease	14.6	13.4	0.03	13.5	13.9	-0.01	13.9	13.9	0.00
Attention deficit hyperactivity disorder	1.2	0.6	0.06	0.8	0.8	0.00	0.8	0.8	0.00
Chronic liver disease	0.8	1.0	-0.03	0.8	0.9	-0.01	0.8	0.9	-0.01
Chronic obstructive lung disease	4.1	4.4	-0.01	4.0	4.2	-0.01	4.2	4.4	-0.01
Crohn's disease	0.2	0.2	0.00	0.1	0.2	-0.01	0.1	0.2	0.00
Dementia	0.3	0.4	-0.02	0.3	0.4	-0.01	0.4	0.4	0.00
Depressive disorder	10.3	7.0	0.12	8.5	8.8	-0.01	8.4	8.3	0.00
Gastroesophageal reflux disease	9.6	8.2	0.05	8.7	9.0	-0.01	8.7	8.9	0.00
Gastrointestinal hemorrhage	1.1	1.2	0.00	1.1	1.0	0.00	1.1	1.1	0.00
Human immunodeficiency virus infection	0.3	0.3	0.01	0.3	0.3	0.00	0.3	0.3	0.00
Hyperlipidemia	39.9	44.4	-0.09	40.7	40.8	0.00	42.7	42.8	0.00
Hypertensive disorder	46.7	49.0	-0.05	46.6	46.7	0.00	48.3	48.2	0.00
Lesion of liver	0.6	0.6	0.00	0.5	0.6	-0.01	0.6	0.6	0.00
Obesity	20.3	12.8	0.20	15.9	16.3	-0.01	16.0	15.5	0.01
Osteoarthritis	15.7	13.5	0.06	14.3	14.7	-0.01	14.3	14.4	0.00
Pneumonia	2.2	2.0	0.01	2.0	2.1	-0.01	2.0	2.1	-0.01
Psoriasis	1.3	1.0	0.03	1.2	1.1	0.01	1.2	1.1	0.01
Renal impairment	4.3	5.2	-0.04	4.4	4.6	-0.01	4.9	5.0	0.00
Rheumatoid arthritis	1.1	0.9	0.02	1.0	1.0	-0.01	1.0	1.0	0.00
Schizophrenia	0.3	0.3	0.00	0.3	0.3	0.00	0.3	0.3	0.00
Urinary tract infectious disease	4.7	3.9	0.04	4.1	4.4	-0.01	4.1	4.3	-0.01
Viral hepatitis C	0.2	0.3	-0.01	0.2	0.3	-0.01	0.2	0.3	0.00
Visual system disorder	16.1	16.8	-0.02	15.9	16.4	-0.01	16.4	16.7	-0.01
Medical history: Cardiovascular disease									
Atrial fibrillation	2.5	3.7	-0.07	2.7	2.9	-0.01	3.1	3.3	-0.01
Cerebrovascular disease	1.8	2.5	-0.04	2.0	2.0	0.00	2.2	2.3	0.00
Coronary arteriosclerosis	5.5	9.1	-0.14	6.1	6.4	-0.01	7.5	7.8	-0.01
Heart disease	13.4	18.3	-0.14	14.0	14.6	-0.02	16.0	16.7	-0.02
Heart failure	2.6	4.7	-0.11	2.8	3.1	-0.02	3.4	4.0	-0.03
Ischemic heart disease	2.6	4.5	-0.10	2.9	3.1	-0.01	3.6	3.8	-0.01
Peripheral vascular disease	3.6	4.3	-0.04	3.6	3.9	-0.01	3.9	4.1	-0.01
Pulmonary embolism	0.5	0.4	0.01	0.5	0.5	0.00	0.5	0.5	0.00
Venous thrombosis	0.6	0.7	0.00	0.6	0.7	0.00	0.7	0.7	0.00
Medical history: Neoplasms									
Hematologic neoplasm	0.5	0.5	-0.01	0.5	0.5	0.00	0.5	0.5	0.00
Malignant lymphoma	0.3	0.3	-0.01	0.3	0.3	0.00	0.3	0.3	0.00
Malignant neoplasm of anorectum	0.1	0.1	-0.01	0.1	0.1	-0.01	0.1	0.1	-0.01
Malignant neoplastic disease	4.4	4.9	-0.02	4.4	4.6	-0.01	4.7	4.8	-0.01
Malignant tumor of breast	1.0	0.9	0.01	1.0	1.0	0.00	1.0	1.0	-0.01
Malignant tumor of colon	0.2	0.2	-0.01	0.2	0.2	0.00	0.2	0.2	0.00
Malignant tumor of lung	0.1	0.2	-0.01	0.1	0.1	0.00	0.1	0.1	0.00
Malignant tumor of urinary bladder	0.1	0.2	-0.01	0.2	0.1	0.00	0.2	0.2	0.01
Primary malignant neoplasm of prostate	0.6	0.9	-0.04	0.8	0.8	0.00	0.8	0.8	0.00
Medication use									
Agents acting on the renin-angiotensin system	62.0	68.4	-0.13	64.8	64.7	0.00	66.1	66.2	0.00

Antibacterials for systemic use	58.0	54.4	0.07	55.9	56.7	-0.02	55.6	56.0	-0.01
Antidepressants	39.2	28.9	0.22	34.8	35.5	-0.01	33.2	33.2	0.00
Antiepileptics	22.8	17.9	0.12	20.6	20.9	-0.01	20.1	20.0	0.00
Antiinflammatory and antirheumatic products	38.3	35.9	0.05	37.0	37.2	0.00	36.9	37.1	0.00
Antineoplastic agents	5.6	4.5	0.05	5.0	5.1	0.00	5.0	5.1	0.00
Antipsoriatics	1.4	0.9	0.04	1.2	1.2	0.00	1.2	1.1	0.00
Antithrombotic agents	13.8	18.5	-0.13	14.7	15.2	-0.01	16.5	16.9	-0.01
Beta blocking agents	28.6	33.6	-0.11	29.3	30.3	-0.02	31.3	32.0	-0.01
Calcium channel blockers	22.6	24.6	-0.05	23.3	23.7	-0.01	23.8	24.0	0.00
Diuretics	40.8	39.7	0.02	39.8	40.6	-0.02	39.9	40.3	-0.01
Drugs for acid related disorders	33.8	31.5	0.05	32.4	33.0	-0.01	32.3	32.7	-0.01
Drugs for acid related disorders	33.8	31.5	0.05	32.4	33.0	-0.01	32.3	32.7	-0.01
Drugs for obstructive airway diseases	41.9	36.8	0.10	39.4	40.1	-0.01	38.9	39.0	0.00
Immunosuppressants	3.3	2.7	0.04	3.0	3.0	0.00	3.0	2.9	0.00
Lipid modifying agents	64.9	72.3	-0.16	68.3	68.8	-0.01	69.4	69.9	-0.01
Opioids	24.9	23.1	0.04	23.6	24.1	-0.01	23.8	23.8	0.00
Psycholeptics	26.1	21.6	0.11	23.6	24.4	-0.02	23.3	23.5	-0.01
Psychostimulants	4.8	2.9	0.10	3.7	3.8	0.00	3.7	3.6	0.01

Please refer to the supplementary material for tables of characteristics related to other databases and exposures.

Table 2. Number of patients, person time, outcomes and crude incidence of thyroid tumors in each cohort and type of analysis.

	Patients		Exposure (per 1000 pys)		Outcomes		Incidence (per 1000 pys)	
	T	C	T	C	T	C	T (95%CI)	C (95%CI)
GLP-1RA vs SGLT2I								
PS matching on-treatment	366,899	713,801	242.7	555.1	223	520	0.92 (0.80 - 1.05)	0.94 (0.86 - 1.02)
PS stratification on-treatment	446,146	713,873	279.8	555.2	280	520	1.00 (0.89 - 1.13)	0.94 (0.86 - 1.02)
PS matching ITT	369,051	717,707	855.1	2093.6	751	1,725	0.88 (0.82 - 0.94)	0.82 (0.79 - 0.86)
PS stratification ITT	448,528	717,792	958.5	2093.7	900	1,725	0.94 (0.88 - 1.00)	0.82 (0.79 - 0.86)
GLP-1RA vs DPP4I								
PS matching on-treatment	316,587	1,104,270	224.4	968.6	218	1,091	0.97 (0.85 - 1.11)	1.13 (1.06 - 1.20)
PS stratification on-treatment	457,288	1,107,973	285.7	972.0	293	1,098	1.03 (0.91 - 1.15)	1.13 (1.06 - 1.20)
PS matching ITT	320,991	1,133,559	876.3	4767.9	845	4,514	0.96 (0.90 - 1.03)	0.95 (0.92 - 0.97)
PS stratification ITT	460,032	1,119,868	1011.0	4734.0	1,000	4,446	0.99 (0.93 - 1.05)	0.94 (0.91 - 0.97)
GLP-1RA vs SU								
PS matching on-treatment	353,356	1,990,074	233.6	1936.6	206	1,571	0.88 (0.77 - 1.01)	0.81 (0.77 - 0.85)
PS stratification on-treatment	425,525	2,055,583	266.2	2001.5	253	1,627	0.95 (0.84 - 1.08)	0.81 (0.77 - 0.85)
PS matching ITT	357,988	2,042,127	873.7	8520.9	769	6,210	0.88 (0.82 - 0.94)	0.73 (0.71 - 0.75)
PS stratification ITT	425,525	2,055,583	939.5	8691.8	857	6,230	0.91 (0.85 - 0.98)	0.72 (0.70 - 0.73)

ITT=intention to treat. PS=propensity score. T=target cohort. C=comparator cohort. CI=confidence interval. Pys=person years. GLP-1RA = Albiglutide, Dulaglutide, Exenatide, Liraglutide, Lixisenatide, Semaglutide (Tirzepatide was not included due to it being only recently available and the cohort was retrospective). DPP4I = Alogliptin, Linagliptin, Saxagliptin, Sitagliptin, Vildagliptin. SGLT2I = Canagliflozin, Dapagliflozin, Empagliflozin, Ertugliflozin. SU = Chlorpropamide, Glimepiride, Glipizide, Gliquidone, Glyburide, Tolazamide, Tolbutamide.

Table 3. Calibrated and uncalibrated meta-analysed hazard ratios for the association with GLP-1RA exposure and thyroid tumors across the databases passing study diagnostics.

	Uncalibrated		Calibrated	
	HR (95% CI)	**P-value**	**HR (95% CI)**	**P-value**
GLP-1RA vs SGLT2I				
Unadjusted on-treatment	1.05 (0.96 - 1.14)	0.29	0.96 (0.60 - 1.52)	0.86
PS matching on-treatment	0.86 (0.59 – 1.24)	0.41	0.83 (0.57 – 1.27)	0.33
PS stratification on-treatment	0.90 (0.77 – 1.05)	0.18	0.88 (0.75 – 1.03)	0.13
Unadjusted ITT	1.12 (1.07 - 1.17)	<0.01	1.02 (0.77 - 1.36)	0.88
PS matching ITT	0.92 (0.77 – 1.10)	0.37	0.89 (0.74 – 1.07)	0.22
PS stratification ITT	0.97 (0.89 – 1.06)	0.55	0.95 (0.85 – 1.06)	0.35
GLP-1RA vs DPP4I				
Unadjusted on-treatment	0.95 (0.76 - 1.18)	0.63	0.98 (0.44 - 2.17)	0.93
PS matching on-treatment	0.79 (0.65 - 0.97)	0.02	0.78 (0.60 - 1.01)	0.06
PS stratification on-treatment	0.81 (0.70 - 0.95)	<0.01	0.83 (0.67 - 1.03)	0.10
Unadjusted ITT	1.09 (0.95 - 1.25)	0.21	1.08 (0.65 - 1.79)	0.79
PS matching ITT	0.93 (0.85 - 1.02)	0.12	0.92 (0.79 - 1.06)	0.24
PS stratification ITT	0.94 (0.87 - 1.01)	0.11	0.93 (0.83 - 1.04)	0.22
GLP-1RA vs Sulfonylureas				
Unadjusted on-treatment	0.49 (0.42 - 0.57)	<0.01	0.48 (0.19 - 1.21)	0.12
PS matching on-treatment	0.94 (0.78 - 1.15)	0.56	0.95 (0.75 - 1.20)	0.68
PS stratification on-treatment	0.93 (0.80 - 1.09)	0.37	0.94 (0.73 - 1.21)	0.64
Unadjusted ITT	0.56 (0.51 - 0.63)	<0.01	0.55 0.27 - 1.13)	0.11
PS matching ITT	1.04 (0.95 - 1.14)	0.42	1.03 (0.87 - 1.23)	0.72
PS stratification ITT	1.00 (0.92 - 1.09)	0.96	1.02 (0.84 - 1.24)	0.86

ITT=intention to treat. PS=propensity score.HR=hazard ratio. CI=confidence interval. GLP-1RA = Albiglutide, Dulaglutide, Exenatide, Liraglutide, Lixisenatide, Semaglutide (Tirzepatide was not included due to it being only recently available and the cohort was retrospective). DPP4I = Alogliptin, Linagliptin, Saxagliptin, Sitagliptin, Vildagliptin. SGLT2I = Canagliflozin, Dapagliflozin, Empagliflozin, Ertugliflozin. SU = Chlorpropamide, Glimepiride, Glipizide, Gliquidone, Glyburide, Tolazamide, Tolbutamide.

FIGURE LEGEND

Figure 1. Schematic of new-user cohort design to study thyroid tumor with second-line treatments for T2DM.