nature portfolio

Peer Review File

Glucagon-like Peptide-1 Receptor Agonists and Risk of Tuberculosis in Type 2 Diabetes

Corresponding Author: Dr Kuang-Ming Liao

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript reports a retrospective analysis of a large database, using a propensity score-matched cohorts approach, and showing an association of GLP-1 RA, when compared to other blood glucose lowering-treatments, with tuberculosis, in patients with type 2 diabetes.
The observation is novel and potentially interesting, but the study shows some relevant issues, which may affect the interpretation of results. With the data reported in the present form, the Conclusion could be a little overenthusiastic.

Major points:

1. Covariates did not include some characteristics which could have biased the results, such as:
a. Geographical area/countries: the epidemiology of tuberculosis is widely heterogeneous across countries; in many countries with a higher incidence/prevalence of TB, GLP-1RA were introduced later, and have a lower market penetration. Despite this background, country of origin was not considered among covariates for propensity score calculation and/or for matching of cohorts.
b. Sociodemographic status: in many countries, GLP-1RA are either bought by patients out of pocket, or paid for by insurance plans, with a coverage which varies on the basis of the cost of the insurance plan. As a consequence, in many countries the use of GLP-1RA is more common in affluent fractions of the population, that have a lower risk of tubercolosis. Family/personal income or some other measure of socioeconomic status must be included among covariates.

2. Some important covariates, such as HbA1c and BMI, were analyzed after a categorization (I.e., BMI>25, and HbA1c>9), which is largely arbitrary. This procedure could have produced a distortion (meaning that mean BMi and HbA1c could have been higher in the GLP-1RA cohort even after matching for BMI>25 and HbA1c>9). This is likely to bias the results, particularly for BMI, since the incidence of tuberculosis increases with malnourishment, which could be considered as a relative contraindication to the use of GLP-1 RA. I suggest to use A1c and BMI as continuous variables, possibly introducing a further category for BMI<20.

3. The category BMI>25 is problematic in a multinational cohort: 25 kg/m2 is the usual threshold for overweight in Whites, but not in Asians.

4. The length of observation should be reported (at least as mean in each cohort for all the comparisons).

5. The Discussion is not well focused. The elaboration on immune mechanisms involved in risk of tuberculosis (Lines 274-323) is substantially out of the scope of the paper, and it could be synthesized in a few line. The extensive description of animal studies (lines 324-368) is more typical of a narrative review, and it could also be considerably shortened. On the other hand, the section of limitations is insufficient. The authors should stress more clearly the possible effect of unaccounted confounders (see previous points). In addition, they should make clearer that we know little about the completeness and reliability of diagnosis of TB in this database, underscoring the possibility of misclassifications.

Minor points:
A. Introduction, line 59-60: I do not understand whether the reported prevalence of tuberculosis refers to 2013 or 2023.
B. Introduction, line 60-61: the authors state that in most recent years TB is more frequent than COVID-19 as a cause of

death, but they do not cite their source.
C. Introduction, line 61-62: specify the years to which reported data are referred.
D. Please provide some data on the reliability of diagnosis of type 2 diabetes in this database.
E. Figure 2: although I understand the small error determined by the fact that one in 4 years lasts 366 instead of 365 days, I would use years instead of days on the X axis, for better readability. Please add number of patients on observation at each year in each group.
F. Figure 3. Report p for interaction (for comparisons across subgroups). Numbers already reported in Figures do not need to be repeated in the text of the Results section, where a brief description would be easier to read.
G. Discussion, line 365-367: I am not sure that the fact that the TriNetX platform includes a predefined analysis for propensity score-.matched cohorts represents an actual strength of the study: the application of this analysis model to any database of appropriate size and detail would have worked equally well.
H. Conclusion, line 392: I am not certain that ay observational study can “demonstrate” an association: I would rather say that it “suggests” an association.
I. Conclusion, line 398-399: beside mechanistic and observational studies, I would make a string quest for analyses of events of TB reported as adverse events in randomized trials.

Reviewer #2

(Remarks to the Author)
GENERAL COMMENTS:
I thank the authors for the opportunity to review their manuscript. In this population-based cohort study, the authors use the TriNetX platform to investigate the association between GLP-1RAs and other antidiabetic drugs and the risk of tuberculosis in pairwise comparisons. The study has several strengths, including a large sample size. However, it also has some limitations and other aspects that require clarifications; these issues are described in the Specific Comments below.

SPECIFIC COMMENTS:
1. The analytical approach and interpretation are highly focused on p-values and the presence or absence of statistical significance. Rather than focusing on statistical significance alone, the interpretation should focus on the magnitude of the effect estimates and the limits of the corresponding 95% confidence intervals.

2. Along those lines, p-values should be removed throughout the manuscript, tables, and figures. The one exception is Figure 3, where the existing p-values should be replaced by p-values testing the interactions.

3. Greater clarity is required regarding cohort construction. In Figure 1, 3.8 million patients are excluded. The figure should be revised such that it reports the number excluded at each exclusion criterion. In addition, it is unclear why patients with prior use of the study drugs before the index date were excluded from the overall cohort. Presumably, this exclusion was applied to the individual pairwise comparisons.

4. The study would benefit from a figure to clarify the measurement windows used (https://pubmed.ncbi.nlm.nih.gov/30856654/). For example, in the Methods, it states that “individuals with any previous or concurrent exposure to these medications before the index period were excluded”. This appears to refer to any previous use. Elsewhere, it suggests that a one-year lookback window was used.

5. Could the authors please confirm whether GLP-1RA users were identified first (followed by users of the comparator drugs) or if users were assigned the first of the two drug classes used in the study period? The former could result in a time-related bias.

6. One of the comparisons involves GLP-1 RAs vs metformin. Given that metformin is typically used as first line therapy whereas GLP-1 RAs are used as second or third line therapy, the clinical relevance of this comparison is unclear.

7. Clinical variables should be modeled continuously using splines rather than categorized in the propensity score.

8. The Results are very repetitive with respect to the contents of the tables and figures.

9. It would be important to explicitly define the exposure definition used. It appears as though the authors used an approach analogous to an intention-to-treat. The inclusion of an on treatment analysis in which patients are censored upon switch or discontinuation would be helpful as a sensitivity analysis.

10. The manuscript would benefit from the inclusion of a negative control analysis.

Reviewer #3

(Remarks to the Author)
This study addresses an interesting question. There have been quite some data on the association between metformin and lower TB risk and better TB outcomes, but less is known about possible effects of other anti-diabetes drugs, especially GLP-1 receptor antagonists which have shown immunomodulatory effects.
The authors use electronic health record data from a international research network including 153 million people to examine if GLP-1 receptor antagonists are associated with lower risk of TB.

they include individuals who started GLP-1 ra's and other DM drugs over a 8-year period. propensity scoring was used to compare risk associated with different DM drugs.
the results show a lower TB risk associated with GLP1 ra compared to other DM drugs.

1. for propensity score matching general demographics and DM variables and immunosuppression are used, but no other variables that reflect TB risk (smoking, alcohol, socioeconomic factors, undernutrition, possible presence of TB contacts). the authors mention this limitation in their discussion. I still wonder if the authors are able to add som covariable (even if only present for the incident TB cases) to show there is low risk of confounding by TB exposure or other TB risk factors.
2. duration of DM may be another factor. have the authors included this in their model and did this change their result?
3. an importat question is if DM medication reduces risk of TB infection, or if it reduces TB disease breakthrough among those who are TST/IGRA positive. does this database hold data (even in a subset) on TST/IGRA testing to make thius ditinction?
4. have the authors looked at dose for the various DM drugs? could there be some kind of dose-effect?
5. the authors do not report levels of dysglycemia and presence of dyslipidemia, both of which have previously been linked to Tb risk. can the authors show how HBa1c level links to TB risk? and same for level of dyslipidemia? do these interact with the association with GLP-1 ra's?
6. same, how does BMI interact with medication risk?
7. under medication no mention is made of statins, which may also lower TB risk
8. can the authors report on TB severity (X-ray extent, bacterial load, level of CRP or anemia, BMI) to show if GLP-1 ra also link to milder disease?
9. I think the discussion on the effects of DM on host immune response should either be improved / much shortened or left out. there are good reviews on this topic (eg SSekematta P. 2023 and otehrs) and there are some fallacies in this section. For instance, TNF production is not consistently down in DM, rather macrophages even have shown to produce higher levels in some studies. I would similarly much shorten the section on metabolic factors in the discussion.
10. This is a large epidemiological study and I think the discussion should focus on internal and external validity of the epidemiological associations (and brief reference on immunomodulatory effects of GLP-1 ra's). I miss reference ot important epidemiological studies on TB risk like from Julia Critchley and Matthew magee (PMID 39040177, 32061457, 37963989, 30377249, 30104296).
11. the labels for Fig 2C should be reversed

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
All the issues raises by this reviewer have been properly addressed.

Reviewer #2

(Remarks to the Author)
I thank the authors for the opportunity to re-review their manuscript. They have addressed most comments raised, and the manuscript is substantially improved as a result. I do have two additional comments.

1. In Figure 2, there is a relatively early splitting of the curves, which may be indicative of a time-related bias. Do the authors have any explanation for the observed early effects?

2. In response to my previous comment, the authors added a sensitivity analysis using an alternative exposure definition in which patients were required to have at least one prescription of the index drug during the first 1 to 5 years following the index date and to have no exposure to comparator antidiabetic drug classes during the follow-up period. By looking into the future to establish eligibility, this analysis has introduced immortal time bias. Patients should be censored at discontinuation or crossover rather than excluded.

Reviewer #3

(Remarks to the Author)
I am happy with the revisions and have no further comments

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
I thank the authors for the opportunity to re-review their manuscript. I remain concerned by the inclusion of sensitivity

analysis using an alternative exposure definition. While I understand that the authors are limited by the analyses that can be performed using their platform, this analysis remains biased. Rather than including a biased sensitivity analysis as part of the manuscript, it would be more appropriate to explain why they were unable to conduct such an analysis in the limitations section.

Reviewer #1 (Remarks to the Author):

The manuscript reports a retrospective analysis of a large database, using a propensity score-matched cohorts approach, and showing an association of GLP-1 RA, when compared to other blood glucose lowering-treatments, with tuberculosis, in patients with type 2 diabetes. The observation is novel and potentially interesting, but the study shows some relevant issues, which may affect the interpretation of results. With the data reported in the present form, the Conclusion could be a little overenthusiastic.

Reply: Thank you for your thoughtful review and for recognizing the novelty and potential clinical significance of our findings. We appreciate your concerns regarding the interpretation of our results. In response to your feedback and the specific issues you have identified in your detailed comments, we have carefully revised our manuscript to address these limitations. We have also revised our conclusion to present our findings more cautiously, acknowledging the observational nature of our study and the need for confirmatory research. We believe these revisions have strengthened the manuscript and provided a more balanced interpretation of our results.

Major points:

1. Covariates did not include some characteristics which could have biased the results, such as:

a. Geographical area/countries: the epidemiology of tuberculosis is widely heterogeneous across countries; in many countries with a higher incidence/prevalence of TB, GLP-1RA were introduced later, and have a lower market penetration. Despite this background, country of origin was not considered among covariates for propensity score calculation and/or for matching of cohorts.

Reply: We thank the reviewer for highlighting the potential impact of geographical heterogeneity in tuberculosis epidemiology and differential adoption of GLP-1RAs across countries. We agree that country- or region-level differences could introduce system-level confounding if not addressed. To evaluate the robustness of our findings within a more geographically homogeneous setting, we performed a sensitivity analysis restricted to the US network in TriNetX (Table S7), where drug availability and prescribing practices are more comparable and where country-level TB epidemiology is relatively uniform. In this US-restricted analysis, GLP-1RA use remained associated with a lower risk of tuberculosis compared with sulfonylureas, metformin, and DPP-4i, with effect estimates consistent in direction with the primary analysis (HR 0.67 and

0.70, respectively). The association was attenuated, and was not statistically significant compared with SGLT2i in the US-restricted analysis. This attenuation may reflect lower event rates in the US setting, reduced statistical power, and/or greater clinical comparability between newer drug classes, leading to smaller differences in TB risk.

Methods

Sensitivity analyses were conducted to assess the robustness of the primary findings under alternative exposure definitions and geographic restrictions. ... Second, to address potential geographic heterogeneity in TB epidemiology and drug adoption patterns, analyses were repeated in a cohort restricted to the US network within TriNetX.

Results

In the US network cohort, GLP-1RA use remained associated with a lower risk of TB compared with SU (HR, 0.67; 95% CI, 0.60–0.76), metformin (HR, 0.70; 95% CI, 0.60–0.81), and DPP-4is (HR, 0.86; 95% CI, 0.76–0.97). However, this association was attenuated and not statistically significant when compared with SGLT2is (HR, 0.94; 95% CI, 0.83–1.06) (Table S7).

Table S7. Hazard ratio of tuberculosis between GLP-1RA and other comparator groups in the US network.

Comparator	Events (n)	Rate (‰)	Events (n)	Rate (‰)	HR (95% CI)	E-value (95% LCL)
GLP-1RA (n = 669,255) vs. SU (n = 669,255)	453	0.68	879	1.31	0.67 (0.60,0.76)	2.4 (2.0)
GLP-1RA (n = 428,580) vs. Metformin (n = 428,580)	277	0.65	485	1.13	0.70 (0.60,0.81)	2.2 (1.8)
GLP-1RA (n =517,546) vs. DPP-4i (n = 517,546)	427	0.82	648	1.25	0.86 (0.76,0.97)	1.6 (1.2)
GLP-1RA (n = 624,817) vs. SGLT2i (n = 624,817)	498	0.80	504	0.81	0.94 (0.83,1.06)	1.3 (1.0)

Discussion

Forth, we acknowledge that TB epidemiology varies substantially across countries, and GLP-1RA adoption differs by region, which could introduce geographical confounding that was not explicitly adjusted for in our PSM. To address this limitation, we performed a sensitivity analysis restricted to the US network in TriNetX (Table S7), where TB epidemiology and drug availability are more uniform. In this US-restricted analysis, GLP-1RA use remained associated with a lower risk of tuberculosis compared with SU, metformin and DPP-4i, though the association was attenuated and not statistically significant when compared with SGLT2is. This attenuation may reflect lower event rates in the US setting, reduced statistical power, and/or greater clinical comparability between newer drug classes, leading to smaller differences in TB risk.

b. Sociodemographic status: in many countries, GLP-1RA are either bought by patients out of pocket, or paid for by insurance plans, with a coverage which varies on the basis of the cost of the insurance plan. As a consequence, in many countries the use of GLP-1RA is more common in affluent fractions of the population, that have a lower risk of tubercolosis. Family/personal income or some other measure of socioeconomic status must be included among covariates.

Reply: We thank the reviewer for this important comment regarding potential confounding by socioeconomic status. We agree that access to GLP-1RAs may be influenced by patients' socioeconomic conditions and insurance coverage, which could in turn be associated with TB risk. To address this concern, we expanded the set of covariates included in the propensity score model to incorporate available proxies of socioeconomic and lifestyle status recorded in the TriNetX database. These variables included housing and economic circumstances, employment and unemployment status, education and literacy, occupational exposure to risk factors, alcohol-related disorders, and nicotine dependence. After propensity score matching, these variables were well balanced between treatment groups, with all standardized mean differences below 0.1 (Table 1). We acknowledge that TriNetX does not provide direct measures of personal or household income, insurance tier, or out-of-pocket expenditure, and therefore residual confounding related to unmeasured socioeconomic factors cannot be fully excluded. This limitation has been explicitly acknowledged in the revised manuscript.

Methods

To address potential confounding related to socioeconomic position and lifestyle, available proxy variables recorded in the TriNetX database were incorporated into the propensity score model. These included housing and economic circumstances, employment and unemployment status, education and literacy, occupational exposure to risk factors, alcohol-related disorders, and nicotine dependence.

Results

Before PSM, considerable differences in baseline characteristics were observed between GLP-1RA users and those receiving SU, metformin, DPP-4i, or SGLT2i (Table S4). After PSM, baseline variables, including age, sex, race, BMI, HbA1c, socioeconomic and lifestyle proxy variables, comorbidities, and relevant

medications, were well balanced across groups within each cohort, with all SMDs below 0.10 (Table 1).

Discussion

Second, although we adjusted for several available proxies of socioeconomic and lifestyle status, direct measures of income, insurance coverage, or out-of-pocket payment were not available in the TriNetX database. Residual confounding related to unmeasured socioeconomic factors therefore cannot be excluded.

2. Some important covariates, such as HbA1c and BMI, were analyzed after a categorization (I.e., BMI>25, and HbA1c>9), which is largely arbitrary. This procedure could have produced a distortion (meaning that mean BMi and HbA1c could have been higher in the GLP-1RA cohort even after matching for BMI>25 and HbA1c>9). This is likely to bias the results, particularly for BMI, since the incidence of tuberculosis increases with malnourishment, which could be considered as a relative contraindication to the use of GLP-1 RA. I suggest to use A1c and BMI as continuous variables, possibly introducing a further category for BMI<20.

Reply: We thank the reviewer for this important methodological comment. We agree that categorization of continuous variables such as BMI and HbA1c using arbitrary cut-offs may inadequately capture residual differences and potentially introduce bias. In response, we revised the propensity score model to incorporate BMI and HbA1c as continuous variables (modeled as mean with standard deviation). In addition, to specifically address the concern regarding malnutrition and its association with tuberculosis risk, we introduced an additional BMI category (<20 kg/m²) and explicitly adjusted for malnutrition as a baseline comorbidity. After propensity score matching using these revised specifications, BMI and HbA1c distributions were well balanced between treatment groups, with standardized mean differences below 0.10 across all comparisons except GLP-1RA vs SU cohorts (Table 1). These revisions mitigate the potential distortion raised by the reviewer and further support the robustness of the findings.

Methods

Propensity score matching (PSM) was performed using a comprehensive set of baseline characteristics to enhance comparability between treatment groups. Demographic variables included age (mean and standard deviation), sex, and race

(White, Black or African American, Asian, other, and unknown). Clinical variables comprised body mass index (BMI, modeled as mean with standard deviation and categorized as <25 and ≥25 kg/m²) and hemoglobin A1c (HbA1c, modeled as mean with standard deviation and categorized as ≥9%).

Results

Before PSM, considerable differences in baseline characteristics were observed between GLP-1RA users and those receiving SU, metformin, DPP-4i, or SGLT2i (Table S4). After PSM, most of the baseline variables, including age, sex, race, BMI, HbA1c, socioeconomic and lifestyle proxy variables, comorbidities, and relevant medications, were well balanced across groups within each cohort, with all SMDs below 0.10 (Table 1).

3. The category BMI>25 is problematic in a multinational cohort: 25 kg/m2 is the usual threshold for overweight in Whites, but not in Asians.

Reply: We thank the reviewer for raising this important point regarding the appropriateness of BMI cut-offs in a multinational cohort. We agree that a BMI threshold of 25 kg/m² reflects the conventional definition of overweight primarily in White populations and may not optimally capture adiposity-related risk in Asian populations. To address this concern, BMI was incorporated into the propensity score model primarily as a continuous variable (modeled as mean with standard deviation), rather than relying solely on categorical cut-offs. Furthermore, individuals identified as Asian accounted for approximately 4–5% of the overall study population, and their distribution was well balanced between treatment groups after matching. Given this relatively small proportion and the use of continuous BMI modeling, the potential impact of applying a uniform BMI cut-off on the overall effect estimates is likely limited. Nevertheless, we acknowledge that ethnicity-specific BMI thresholds could not be fully explored within the constraints of the current database, and this limitation has been acknowledged in the revised manuscript.

Discussion

Finally, although ethnicity-specific BMI thresholds may be particularly relevant for Asian populations, individuals classified as Asian comprised a small proportion of the study cohort, and BMI was primarily modeled as a continuous variable to mitigate potential bias related to population-specific cut-offs.

4. The length of observation should be reported (at least as mean in each cohort for all the comparisons).

Reply: We thank the reviewer for this important comment. We agree that reporting the duration of follow-up is essential for interpreting time-to-event analyses. Accordingly, we have added the length of observation for each comparison group. Given the right-skewed distribution of follow-up time in real-world data, follow-up duration was summarized using the median number of days. The median follow-up times were as follows: GLP-1RA versus SGLT2 inhibitors, 868 vs. 772 days; GLP-1RA versus sulfonylureas, 778 vs. 1,789 days; GLP-1RA versus dipeptidyl peptidase-4 inhibitors, 784 vs. 1,595 days; and GLP-1RA versus metformin, 767 vs. 1,442 days.

Results

The median duration of follow-up differed across comparison groups, with median follow-up times of 868 vs. 772 days for GLP-1RA versus SGLT2is, 778 vs. 1,789 days for GLP-1RA versus SU, 784 vs. 1,595 days for GLP-1RA versus DPP-4is, and 767 vs. 1,442 days for GLP-1RA versus metformin.

5. The Discussion is not well focused. The elaboration on immune mechanisms involved in risk of tuberculosis (Lines 274-323) is substantially out of the scope of the paper, and it could be synthesized in a few line. The extensive description of animal studies (lines 324-368) is more typical of a narrative review, and it could also be considerably shortened. On the other hand, the section of limitations is insufficient. The authors should stress more clearly the possible effect of unaccounted confounders (see previous points). In addition, they should make clearer that we know little about the completeness and reliability of diagnosis of TB in this database, underscoring the possibility of misclassifications.

Reply: We appreciate your thoughtful feedback on improving the focus and balance of our Discussion section. We have substantially revised this section in response to these concerns.

Discussion

Diabetes and Increased Risk of Tuberculosis: Immunological and Metabolic Mechanisms

Patients with diabetes mellitus have a markedly increased risk of developing active

tuberculosis due to combined defects in adaptive and innate immunity and the effects of chronic hyperglycemia [20, 21]. Diabetes is associated with reduced T-lymphocyte number and function, particularly impaired Th1 responses and production of cytokines and macrophage activation also abnormal [20]. Innate immune defenses are also compromised, as diabetic macrophages and neutrophils exhibit reduced phagocytosis, chemotaxis, and reactive oxygen species production, limiting intracellular killing of M. tuberculosis [22, 23]. Additionally, impaired monocyte migration—unimproved by insulin therapy—further weakens granuloma formation [24]. Chronic hyperglycemia exacerbates these defects by impairing the respiratory burst and promoting oxidative stress and advanced glycation end-product formation, resulting in dysfunctional immune signaling and sustained susceptibility to TB [22-25].

The role of GLP-1RA and infection

GLP-1 receptors (GLP-1R) are expressed on multiple immune cells, including macrophages, monocytes, and B and T lymphocytes, indicating a role in immune regulation [24]. Disruption of GLP-1R signaling impairs lymphocyte proliferation, while GLP-1R activation increases intracellular cAMP, promoting peripheral T-cell proliferation and regulating Treg activity, thereby supporting adaptive immunity and immune tolerance [26].

Preclinical studies suggest that GLP-1RA exert immunomodulatory effects in diabetes-related immune dysfunction [27]. In obese, high-fat diet–induced diabetic mice, GLP-1RA outperformed insulin in restoring host defense against infection by enhancing neutrophil phagocytosis, migration, and bacterial clearance [28]. Prolonged GLP-1RA treatment further improved infection resistance by correcting hyperglycemia, increasing neutrophil counts, and restoring innate immune competence [29].

Limitations

Despite these strengths, several limitations must be acknowledged. First, the database lacked information on several clinically relevant TB variables, including contact history, latent infection status, radiographic findings, and disease severity, limiting our ability to assess whether GLP-1RA effects differ across TB exposure patterns or disease phenotypes. Additionally, the TriNetX database does not capture medication dosage or adherence data, preventing assessment of dose-response relationships and the potential impact of treatment compliance on the observed associations. Future studies incorporating these clinical variables are warranted. Second, although we adjusted for several available proxies of socioeconomic and lifestyle status, direct measures of income, insurance coverage, or out-of-pocket

payment were not available in the TriNetX database. Residual confounding related to unmeasured socioeconomic factors therefore cannot be excluded. Third, as with all retrospective observational studies, residual confounding remains possible despite extensive covariate adjustment and propensity score matching. In addition, the TriNetX platform does not support flexible non-linear modeling approaches, such as spline-based transformations, within the propensity score framework, and potential non-linear associations between continuous clinical variables and tuberculosis risk may not be fully captured. Forth, we acknowledge that TB epidemiology varies substantially across countries, and GLP-1RA adoption differs by region, which could introduce geographical confounding that was not explicitly adjusted for in our PSM. To address this limitation, we performed a sensitivity analysis restricted to the US network in TriNetX (Table S7), where TB epidemiology and drug availability are more uniform. In this US-restricted analysis, GLP-1RA use remained associated with a lower risk of tuberculosis compared with SU, metformin and DPP-4i, though the association was attenuated and not statistically significant when compared with SGLT2is. This attenuation may reflect lower event rates in the US setting, reduced statistical power, and/or greater clinical comparability between newer drug classes, leading to smaller differences in TB risk. Fifth, although ethnicity-specific BMI thresholds may be particularly relevant for Asian populations, individuals classified as Asian comprised a small proportion of the study cohort, and BMI was primarily modeled as a continuous variable to mitigate potential bias related to population-specific cut-offs. Finally, our study is also limited knowledge regarding the completeness and reliability of TB diagnoses within this database, which raises concerns about potential misclassifications. The accuracy of the diagnosis may not be fully verifiable, and incomplete or inaccurate diagnostic data could influence the overall findings, potentially affecting the conclusions drawn from the study.

Minor points:

A. Introduction, line 59-60: I do not understand whether the reported prevalence of tuberculosis refers to 2013 or 2023.

Reply: We apologize for the ambiguity. The reported prevalence refers to 2023. This has been corrected in the revised manuscript.

Introduction
In 2023, the global burden of tuberculosis (TB) rose to an estimated 10.8 million

cases (134 per 100,000 population)

B. Introduction, line 60-61: the authors state that in most recent years TB is more frequent than COVID-19 as a cause of death, but they do not cite their source.

Reply: We thank the reviewer for identifying this omission. The appropriate reference has now been added.

Introduction
TB has reemerged as the leading cause of death from an infectious disease globally, surpassing COVID-19 [2].

C. Introduction, line 61-62: specify the years to which reported data are referred.

Reply: Thank you for this suggestion. We have specified that the data refer to 2023.

Introduction
8.2 million new TB cases were reported in 2023,

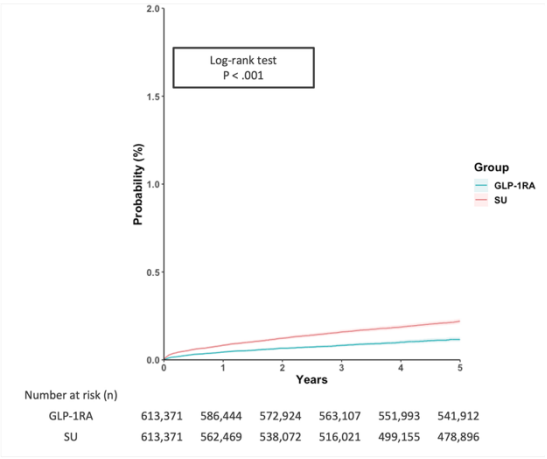
D. Please provide some data on the reliability of diagnosis of type 2 diabetes in this database.

Reply: We appreciate the reviewer's inquiry regarding diagnostic validity. The TriNetX database relies on ICD-10 codes documented in routine clinical practice across healthcare organizations. While TriNetX itself does not publish specific validation studies for type 2 diabetes diagnoses, several previous studies using similar administrative health databases for type 2 diabetes based on diagnostic codes. Additionally, our requirement for documented antidiabetic medication use provides an additional layer of diagnostic specificity, as it confirms active diabetes treatment rather than relying solely on diagnostic codes

E. Figure 2: although I understand the small error determined by the fact that one in 4 years lasts 366 instead of 365 days, I would use years instead of days on the X axis, for better readability. Please add number of patients on observation at each year in each group.

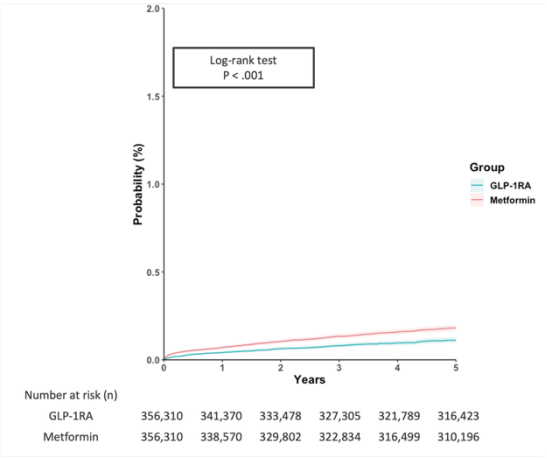
Reply: We thank the reviewer for this helpful suggestion. We have revised Figure 2 to improve readability by presenting follow-up time on the x-axis in years rather than days. In addition, we have added the number of patients at risk at each yearly interval for both treatment groups below the Kaplan–Meier curves. These modifications have been incorporated into the revised figure.

Fig 2A GLP-1RA vs. Sulfonylureas



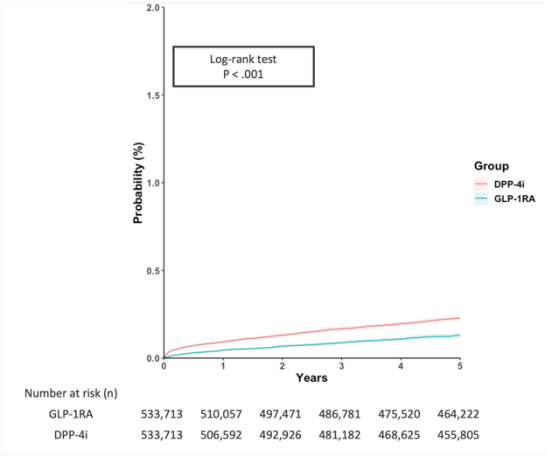
Number at risk (n)						
GLP-1RA	613,371	586,444	572,924	563,107	551,993	541,912
SU	613,371	562,469	538,072	516,021	499,155	478,896

Fig 2B GLP-1RA vs. Metformin



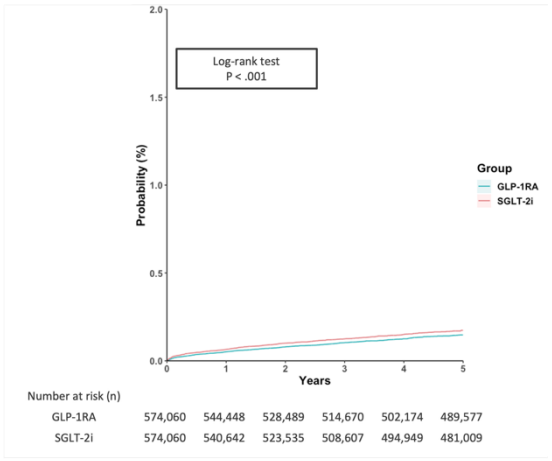
Number at risk (n)						
GLP-1RA	356,310	341,370	333,478	327,305	321,789	316,423
Metformin	356,310	338,570	329,802	322,834	316,499	310,196

Fig 2C GLP-1RA vs. DPP-4i



Number at risk (n)						
GLP-1RA	533,713	510,057	497,471	486,781	475,520	464,222
DPP-4i	533,713	506,592	492,926	481,182	468,625	455,805

Fig 2D GLP-1RA vs. SGLT2is



F. Figure 3. Report p for interaction (for comparisons across subgroups). Numbers already reported in Figures do not need to be repeated in the text of the Results section, where a brief description would be easier to read.

Reply: Thank you for this suggestion. We have revised Figure 3 to include interaction p-values for subgroup comparisons. We have also removed redundant numbers from the Results text and replaced them with brief descriptions to enhance readability.

Fig 3A

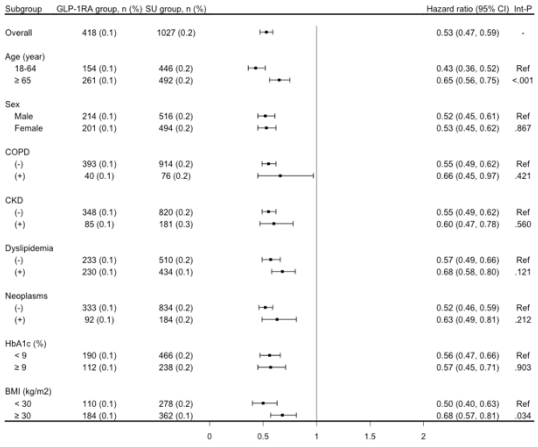


Fig 3B

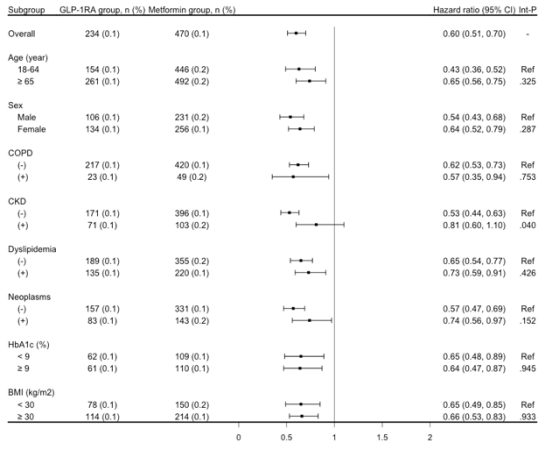


Fig 3C

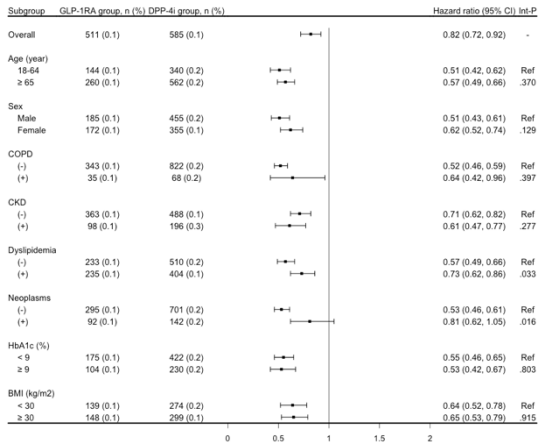
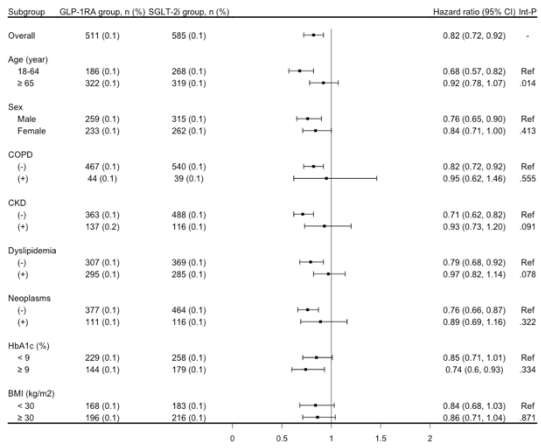


Fig 3D



G. Discussion, line 365-367: I am not sure that the fact that the TriNetX platform includes a predefined analysis for propensity score matched cohorts represents an actual strength of the study: the application of this analysis model to any database of appropriate size and detail would have worked equally well.

Reply: Thank you for your suggestion. We remove this sentence according.

H. Conclusion, line 392: I am not certain that ay observational study can “demonstrate” an association: I would rather say that it “suggests” an association.

Reply: Thank you for your comment. The sentence was revised accordingly.

Conclusion
this study provides the first large-scale evidence from multiple nationwide medical databases suggesting that the use of GLP-1RAs is associated with a significantly reduced risk of tuberculosis among individuals with T2D

I. Conclusion, line 398-399: beside mechanistic and observational studies, I would make a string quest for analyses of events of TB reported as adverse events in randomized trials.

Reply: Thank you for your suggestion. The sentence was revised accordingly.

Conclusion
These findings suggest that beyond their established metabolic benefits, GLP-1RAs may confer additional advantages in lowering infection risk. To validate these observations, future research should prioritize mechanistic studies, prospective clinical investigations, and systematic analyses of tuberculosis events reported as adverse events in randomized controlled trials.

Reply: Thank you for your instructive comment.

Reviewer #2 (Remarks to the Author):

GENERAL COMMENTS:

I thank the authors for the opportunity to review their manuscript. In this population-based cohort study, the authors use the TriNetX platform to investigate the association between GLP-1RAs and other antidiabetic drugs and the risk of tuberculosis in pairwise comparisons. The study has several strengths, including a large sample size. However, it also has some limitations and other aspects that require clarifications; these issues are described in the Specific Comments below.

Reply: Thank you for taking the time to review our manuscript and for your constructive feedback. We appreciate your recognition of the study's strengths, particularly the large sample size afforded by the TriNetX platform. We have carefully considered all of the limitations and clarifications you identified in your specific comments and have made comprehensive revisions to address each point. We believe these revisions have significantly strengthened the manuscript and improved the clarity and interpretation of our findings. Please find our detailed point-by-point responses to your specific comments below.

SPECIFIC COMMENTS:

1. The analytical approach and interpretation are highly focused on p-values and the presence or absence of statistical significance. Rather than focusing on statistical significance alone, the interpretation should focus on the magnitude of the effect estimates and the limits of the corresponding 95% confidence intervals.

Reply: We appreciate your valuable feedback. In response to your comment, we have revised our analytical approach and interpretation throughout the manuscript. We now place greater emphasis on effect sizes and their corresponding 95% confidence intervals, rather than focusing primarily on p-values and statistical significance. We present and interpret the magnitude of effect estimates, highlighting their clinical and practical relevance. P-values are now included only when necessary and appropriate. Additionally, for the subgroup analyses, we have included p-values for interaction to evaluate potential effect modification across subgroups.

2. Along those lines, p-values should be removed throughout the manuscript, tables, and figures. The one exception is Figure 3, where the existing p-values should be replaced by p-values testing the interactions.

Reply: Thank you for this clarification. We have removed p-values throughout the entire manuscript, including all tables and figures, in accordance with your recommendation. The only exception is Figure 3, where we have replaced the previously reported p-values with p-values for interaction testing to evaluate effect modification across subgroups.

Fig 3A

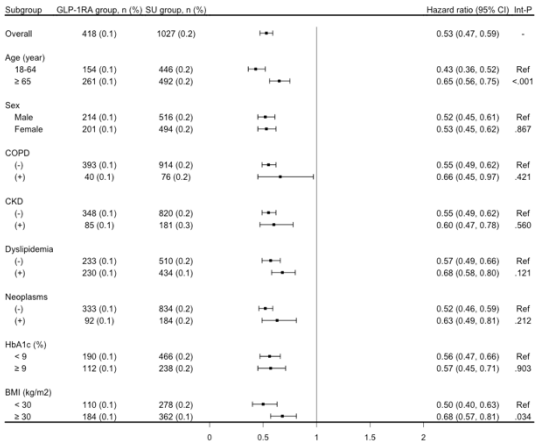


Fig 3B

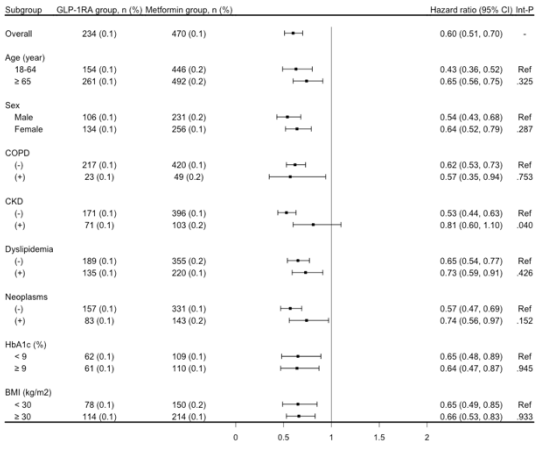


Fig 3C

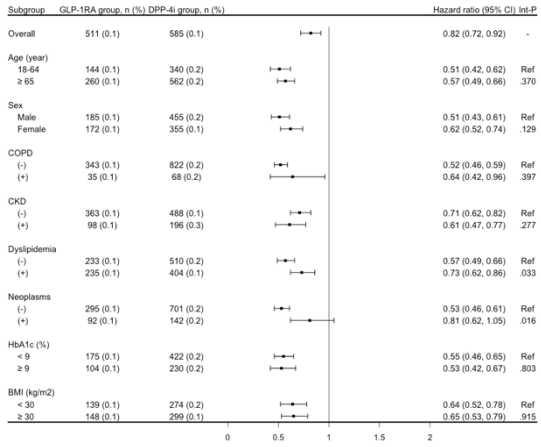
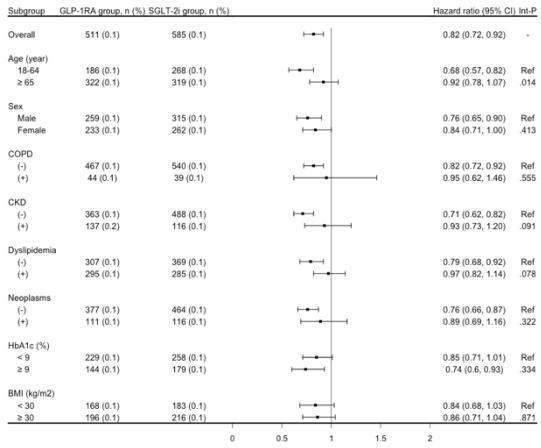
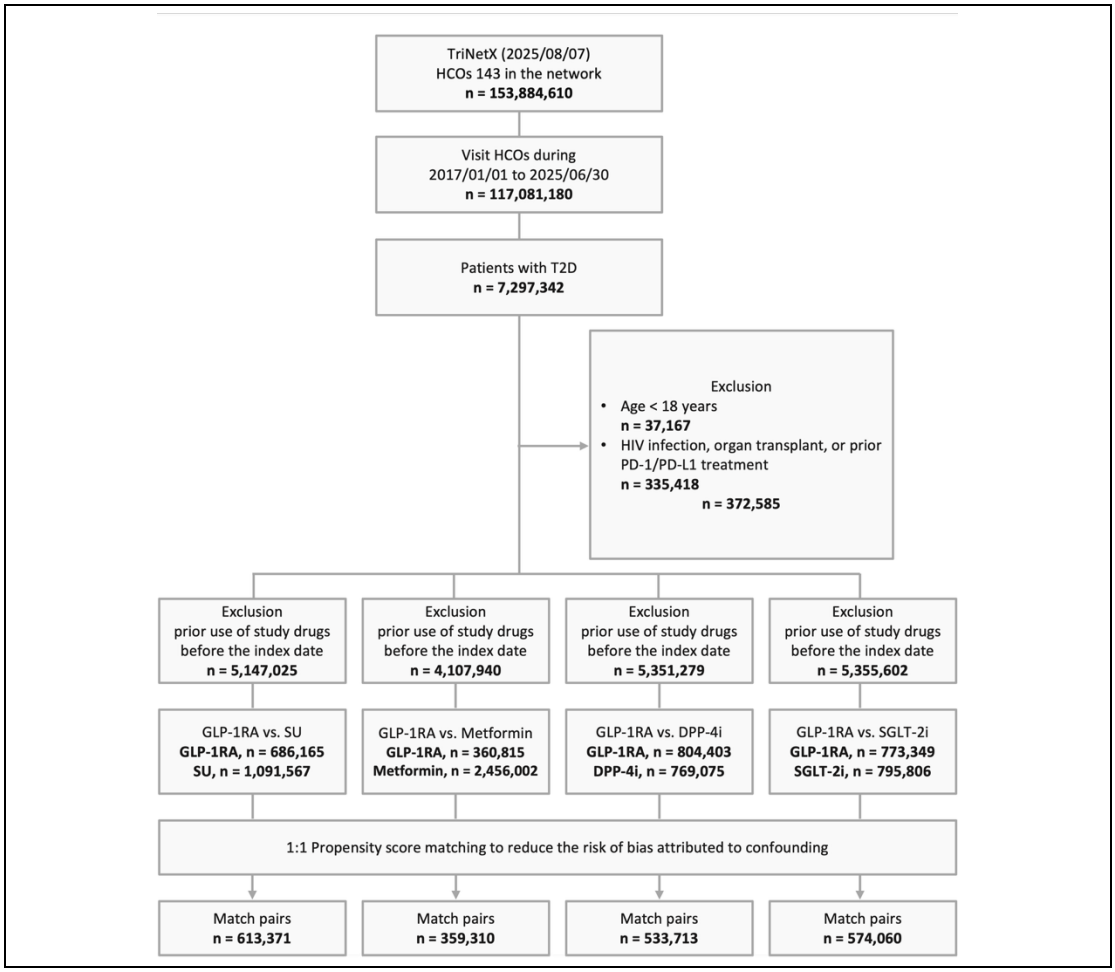


Fig 3D



3. Greater clarity is required regarding cohort construction. In Figure 1, 3.8 million patients are excluded. The figure should be revised such that it reports the number excluded at each exclusion criterion. In addition, it is unclear why patients with prior use of the study drugs before the index date were excluded from the overall cohort. Presumably, this exclusion was applied to the individual pairwise comparisons.

Reply: We thank the reviewer for this helpful comment regarding cohort construction. We have revised Figure 1 to clearly report the number of patients excluded at each exclusion criterion. We also clarify that the exclusion of patients with prior use of the study drugs was not applied to the overall cohort, but rather implemented separately within each pairwise comparison to enforce a new-user design. Specifically, for each comparison (GLP-1RA vs sulfonylureas, metformin, DPP-4is, and SGLT2is), patients with any prior use of the index or comparator drug class before the index date were excluded. The figure and legend have been revised accordingly to improve transparency and avoid ambiguity regarding cohort construction.



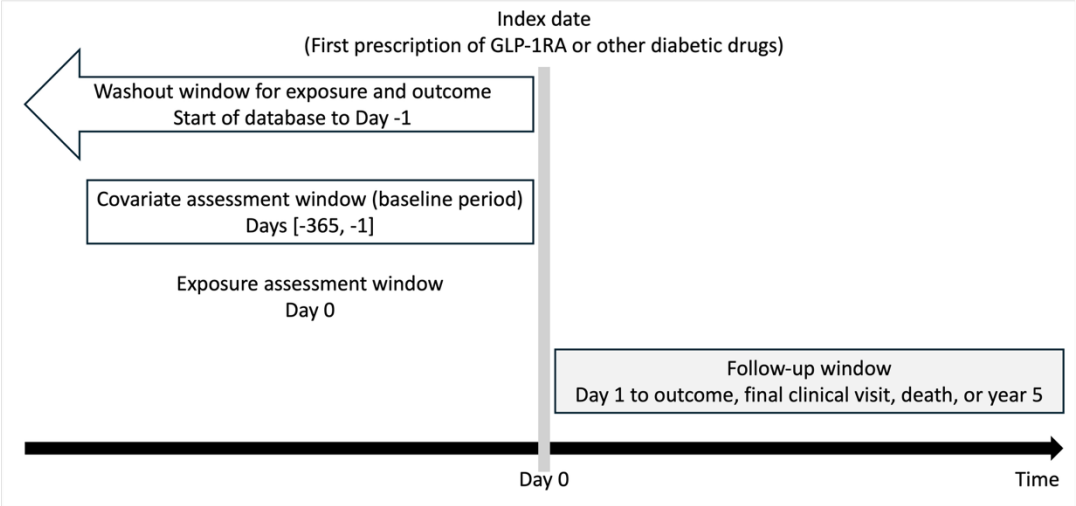
4. The study would benefit from a figure to clarify the measurement windows used (https://pubmed.ncbi.nlm.nih.gov/30856654/). For example, in the Methods, it states that “individuals with any previous or concurrent exposure to these medications before the index period were excluded”. This appears to refer to any previous use. Elsewhere, it suggests that a one-year lookback window was used.

Reply: Thank you for this helpful suggestion. We agree that the measurement windows should be explicitly clarified. In our study, baseline covariates were assessed during the 12-month lookback period (day −365 to day −1). However, the new-user requirement for medication exposure was stricter: to avoid prevalent-user bias, we excluded individuals with any record of prior exposure to the index or comparator drug classes at any time before the index date. We have revised the Methods to clearly distinguish the covariate lookback window from the exposure washout definition, and we have added/updated a schematic figure to illustrate these windows.

Methods

Baseline covariates were assessed during the 12-month lookback window (day −365

to day −1).



5. Could the authors please confirm whether GLP-1RA users were identified first (followed by users of the comparator drugs) or if users were assigned the first of the two drug classes used in the study period? The former could result in a time-related bias.

Reply: We thank the reviewer for this important clarification. In this study, treatment groups were defined using an intention-to-treat, new-user design. All eligible individuals were classified according to the first antidiabetic drug class initiated during the study period, which served as the index exposure. Specifically, patients were assigned to the GLP-1RA or comparator group based on which drug class was prescribed first, rather than identifying GLP-1RA users first and subsequently selecting comparator users. The index date was defined as the date of the first prescription of the assigned drug class, and follow-up began thereafter.

Methods

Patients were assigned to treatment groups according to the first antidiabetic drug class initiated during the study period (intention-to-treat approach), and the index date was defined as the date of first prescription.

6. One of the comparisons involves GLP-1 RAs vs metformin. Given that metformin is typically used as first line therapy whereas GLP-1 RAs are used as second or third line therapy, the clinical relevance of this comparison is unclear.

Reply: We thank the reviewer for this important point. We agree that metformin is

most used as first-line therapy, whereas GLP-1RAs are typically initiated as second- or later-line agents. Therefore, we did not intend the GLP-1RA versus metformin comparison to represent a direct head-to-head alternative at the same treatment line. Rather, this comparison was included as a benchmark/anchor comparator to evaluate the robustness and clinical interpretability of the association in a widely used reference population and to reflect real-world treatment sequencing. To mitigate confounding related to disease severity and treatment stage, we performed propensity score matching on a broad set of baseline characteristics, including continuous BMI and HbA1c, diabetes-related complications, comorbidities, and concomitant medications. In addition, we performed multiple active-comparator analyses (vs sulfonylureas, DPP-4is, and SGLT2is), with the more clinically comparable comparisons being those with newer second-line agents (DPP-4is and SGLT2is).

7. Clinical variables should be modeled continuously using splines rather than categorized in the propensity score.

Reply: We thank the reviewer for this important methodological suggestion. We agree that modeling continuous clinical variables using flexible approaches such as splines can better capture non-linear relationships and is desirable in principle. However, within the TriNetX analytics platform, the propensity score is implemented using a predefined logistic regression framework, which does not allow user-specified non-linear transformations or spline-based modeling of covariates. Within these constraints, we incorporated key clinical variables, including BMI and HbA1c, as continuous variables (mean and standard deviation) in the propensity score model, rather than relying solely on categorical thresholds. We have acknowledged the inability to implement spline-based propensity score modeling as a limitation of the database and analytical platform in the revised manuscript.

Methods

Propensity score matching (PSM) was performed using a comprehensive set of baseline characteristics to enhance comparability between treatment groups. Demographic variables included age (mean and standard deviation), sex, and race (White, Black or African American, Asian, other, and unknown). Clinical variables comprised body mass index (BMI, modeled as mean with standard deviation and categorized as <20 and ≥25 kg/m^2) and hemoglobin A1c (HbA1c, modeled as mean with standard deviation and categorized as ≥9%). Comorbidities included

hypertension, dyslipidemia, malnutrition, chronic lower respiratory diseases, chronic ischemic heart disease, neoplasms, chronic kidney disease, cerebrovascular diseases, rheumatoid arthritis with rheumatoid factor, arthropathic psoriasis, dependence on renal dialysis, systemic lupus erythematosus, Sjögren syndrome, and pneumoconiosis due to other dust containing silica. Diabetes-related complications were also included, comprising kidney complications, ophthalmic complications, neurological complications, and circulatory complications. To address potential confounding related to socioeconomic position and lifestyle, available proxy variables recorded in the TriNetX database were incorporated into the propensity score model. These included housing and economic circumstances, employment and unemployment status, education and literacy, occupational exposure to risk factors, alcohol-related disorders, and nicotine dependence. Baseline medications encompassed HMG-CoA reductase inhibitors, systemic corticosteroids, inhaled corticosteroids, tumor necrosis factor-alpha inhibitors, and interleukin inhibitors. Detailed definitions and coding algorithms for all covariates are provided in Table S2.

Results

Before PSM, considerable differences in baseline characteristics were observed between GLP-1RA users and those receiving SU, metformin, DPP-4i, or SGLT2i (Table S4). After PSM, baseline variables, including age, sex, race, BMI, HbA1c, socioeconomic and lifestyle proxy variables, comorbidities, and relevant medications, were well balanced across groups within each cohort, with all SMDs below 0.10 (Table 1).

Discussion

Third, as with all retrospective observational studies, residual confounding remains possible despite extensive covariate adjustment and propensity score matching. In addition, the TriNetX platform does not support flexible non-linear modeling approaches, such as spline-based transformations, within the propensity score framework, and potential non-linear associations between continuous clinical variables and tuberculosis risk may not be fully captured.

8. The Results are very repetitive with respect to the contents of the tables and figures.

Reply: Thank you for pointing this out. We have substantially revised the Results section to eliminate redundancy. The text now provides a narrative summary of key

findings and patterns rather than repeating numerical values already presented in the tables and figures.

9. It would be important to explicitly define the exposure definition used. It appears as though the authors used an approach analogous to an intention-to-treat. The inclusion of an on treatment analysis in which patients are censored upon switch or discontinuation would be helpful as a sensitivity analysis.

Reply: We thank the reviewer for this important comment. We agree that explicit clarification of the exposure definition is essential. Our primary analysis was conducted using an ITT framework, in which patients were classified according to the initially prescribed antidiabetic therapy at the index date, regardless of subsequent treatment changes. To address the reviewer's suggestion and to evaluate the robustness of our findings under an alternative exposure definition, we performed an additional exposure-restricted sensitivity analysis. In this analysis, patients were required to have at least one prescription of the index drug during the first 1 to 5 years following the index date and to have no exposure to comparator antidiabetic drug classes during the follow-up period. This approach minimizes treatment crossover and exposure misclassification and approximates an on-treatment (as-treated-like) framework within the constraints of the TriNetX platform. We acknowledge that a fully time-varying on-treatment analysis with censoring at treatment discontinuation or switching is not feasible in TriNetX due to the lack of detailed information on exact treatment duration and discontinuation dates. Nevertheless, the results of this exposure-restricted analysis were consistent with those of the primary ITT analysis, supporting the robustness of our findings. The exposure definition and sensitivity analysis have now been explicitly described in the Methods section.

Methods

Sensitivity analyses were conducted to assess the robustness of the primary findings under alternative exposure definitions and geographic restrictions. First, an exposure-restricted sensitivity analysis was performed to approximate an on-treatment framework, in which patients were required to have at least one prescription of the index drug during the first one to five years after the index date and no exposure to comparator antidiabetic drug classes during follow-up.

Results

In the exposure-restricted sensitivity analysis, the association between GLP-1RA use and reduced TB risk remained consistent across all comparisons. Compared with SU, metformin, DPP-4is, and SGLT2is, GLP-1RA use was associated with a lower risk of TB, with HRs ranging from 0.56 to 0.70 (Table S8).

Table S8. Sensitivity analysis using an exposure-restricted approach.

Comparator	Events (n)	Rate (‰)	Events (n)	Rate (‰)	HR (95% CI)	E-value (95% LCL)
GLP-1RA (n = 212,044) vs. SU (n = 212,044)	196	0.92	416	1.96	0.57 (0.48,0.67)	2.9 (2.4)
GLP-1RA (n = 88,456) vs. Metformin (n = 88,456)	75	0.85	138	1.56	0.67 (0.51,0.89)	2.4 (1.5)
GLP-1RA (n =163,520) vs. DPP-4i (n = 163,520)	163	1.00	340	2.08	0.56 (0.46,0.67)	3.0 (2.4)
GLP-1RA (n = 145,249) vs. SGLT2i (n = 145,249)	168	1.16	230	1.58	0.70 (0.58,0.86)	2.2 (1.6)

10. The manuscript would benefit from the inclusion of a negative control analysis.

Reply: We thank the reviewer for this suggestion. We conducted a negative control outcome analysis using skin cancer, an outcome not expected to be causally affected by GLP-1 receptor agonists. Across all comparisons, the hazard ratios were close to 1.0 and the confidence intervals included 1.0 (Table S9), supporting that the primary findings are unlikely to be explained solely by systematic residual confounding.

Methods

Additionally, a negative control outcome analysis using skin cancer was conducted to evaluate the presence of residual confounding unrelated to TB risk.

Results

In negative control outcome analyses using skin cancer, no associations were observed between GLP-1RA use and the control outcome across all comparisons (Table S9).

Table S9. Negative control outcome analysis using skin cancer.

Comparator	HR (95% CI)
GLP-1RA vs. SU	0.99 (0.96,1.02)
GLP-1RA vs. Metformin	0.97 (0.93,1.01)
GLP-1RA vs. DPP-4i	1.02 (0.99,1.06)
GLP-1RA vs. SGLT2i	1.03 (0.92,1.16)

Reviewer #3 (Remarks to the Author):

This study addresses an interesting question. There have been quite some data on the association between metformin and lower TB risk and better TB outcomes, but less is known about possible effects of other anti-diabetes drugs, especially GLP-1 receptor antagonists which have shown immunomodulatory effects. The authors use electronic health record data from a international research network including 153 million people to examine if GLP-1 receptor antagonists are associated with lower risk of TB. they include individuals who started GLP-1 ra's and other DM drugs over a 8-year period. propensity scoring was used to compare risk associated with different DM drugs. the results show a lower TB risk associated with GLP1 ra compared to other DM drugs.

Reply: Thank you for this thoughtful summary of our study and for recognizing the clinical significance of our research question. We agree that while the protective effects of metformin on tuberculosis have been increasingly documented, the potential role of GLP-1RAs remains underexplored despite their known immunomodulatory properties. We believe our findings contribute valuable evidence to this gap by leveraging a large, international dataset and robust propensity score methods to demonstrate the association between GLP-1RA use and reduced TB risk compared to other antidiabetic medications. We appreciate your recognition of the importance of investigating effects beyond metformin in the diabetes-TB interface.

1. for propensity score matching general demographics and DM variables and immunosuppression are used, but no other variables that reflect TB risk (smoking, alcohol, socioeconomic factors, undernutrition, possible presence of TB contacts). the authors mention this limitation in their discussion. I still wonder if the authors are able to add som covariable (even if only present for the incident TB cases) to show there is low risk of confounding by TB exposure or other TB risk factors.

Reply: We thank you for this thoughtful comment regarding potential confounding by TB-related risk factors. While direct information on TB exposure history or household contact is not available in the TriNetX database, we have incorporated a broad range of established proxy variables reflecting TB risk into the propensity score matching. Specifically, in addition to general demographics and diabetes-related characteristics, we included nicotine dependence (smoking), alcohol-related disorders, malnutrition, socioeconomic indicators (housing and economic circumstances, employment/unemployment, education and literacy), chronic respiratory disease, and a comprehensive set of immunosuppressive conditions and medications (including systemic and inhaled corticosteroids, TNF-α inhibitors, interleukin inhibitors, chronic

kidney disease, malignancy, and autoimmune diseases). After propensity score matching, these TB risk-related proxy variables were well balanced between treatment groups, with standardized mean differences below 0.10 across all comparisons (Table 1), suggesting a low likelihood of substantial confounding by measured TB risk factors. Nevertheless, we acknowledge that unmeasured factors such as direct TB contact or community-level exposure cannot be fully accounted for in observational data and have retained this point as a limitation in the Discussion.

Methods

Propensity score matching (PSM) was performed using a comprehensive set of baseline characteristics to enhance comparability between treatment groups. Demographic variables included age (mean and standard deviation), sex, and race (White, Black or African American, Asian, other, and unknown). Clinical variables comprised body mass index (BMI, modeled as mean with standard deviation and categorized as <20 and ≥25 kg/m²) and hemoglobin A1c (HbA1c, modeled as mean with standard deviation and categorized as ≥9%). Comorbidities included hypertension, dyslipidemia, malnutrition, chronic lower respiratory diseases, chronic ischemic heart disease, neoplasms, chronic kidney disease, cerebrovascular diseases, rheumatoid arthritis with rheumatoid factor, arthropathic psoriasis, dependence on renal dialysis, systemic lupus erythematosus, Sjögren syndrome, and pneumoconiosis due to other dust containing silica. Diabetes-related complications were also included, comprising kidney complications, ophthalmic complications, neurological complications, and circulatory complications. To address potential confounding related to socioeconomic position and lifestyle, available proxy variables recorded in the TriNetX database were incorporated into the propensity score model. These included housing and economic circumstances, employment and unemployment status, education and literacy, occupational exposure to risk factors, alcohol-related disorders, and nicotine dependence. Baseline medications encompassed HMG-CoA reductase inhibitors, systemic corticosteroids, inhaled corticosteroids, tumor necrosis factor-alpha inhibitors, and interleukin inhibitors. Detailed definitions and coding algorithms for all covariates are provided in Table S2

Discussion

First, the database lacked information on several clinically relevant TB variables, including contact history, latent infection status, radiographic findings, and disease severity, limiting our ability to assess whether GLP-1RA effects differ across TB exposure patterns or disease phenotypes.

Second, although we adjusted for several available proxies of socioeconomic and lifestyle status, direct measures of income, insurance coverage, or out-of-pocket payment were not available in the TriNetX database. Residual confounding related to unmeasured socioeconomic factors therefore cannot be excluded.

Forth, we acknowledge that TB epidemiology varies substantially across countries, and GLP-1RA adoption differs by region, which could introduce geographical confounding that was not explicitly adjusted for in our PSM. To address this limitation, we performed a sensitivity analysis restricted to the US network in TriNetX (Table S7), where TB epidemiology and drug availability are more uniform. In this US-restricted analysis, GLP-1RA use remained associated with a lower risk of tuberculosis compared with SU, metformin and DPP-4i, though the association was attenuated and not statistically significant when compared with SGLT2is. This attenuation may reflect lower event rates in the US setting, reduced statistical power, and/or greater clinical comparability between newer drug classes, leading to smaller differences in TB risk.

2. duration of DM may be another factor. have the authors included this in their model and did this change their result?

Reply: Thank you for this valuable suggestion. While exact diabetes duration was not directly available in our database, we have included diabetes-related complications and HbA1c levels as proxies for disease severity and duration, which may indirectly capture the effects of longer disease course. We have clarified this approach in the Methods section.

Methods

Propensity score matching (PSM) was performed using a comprehensive set of baseline characteristics to enhance comparability between treatment groups. … Clinical variables comprised body mass index (BMI, modeled as mean with standard deviation and categorized as <20 and ≥25 kg/m²) and hemoglobin A1c (HbA1c, modeled as mean with standard deviation and categorized as ≥9%). …

Diabetes-related complications were also included, comprising kidney complications, ophthalmic complications, neurological complications, and circulatory complications.

3. an importat question is if DM medication reduces risk of TB infection, or if it reduces TB disease breakthrough among those who are TST/IGRA positive. does

this database hold data (even in a subset) on TST/IGRA testing to make thius ditinction?

Reply: Thank you for highlighting this important mechanistic distinction. Unfortunately, TST/IGRA testing data were not systematically available in our database, precluding us from distinguishing between effects on TB infection versus progression to active disease. This represents an important avenue for future prospective studies with more detailed immunological assessments, and we have added this point to our limitations.

Limitations

First, the database lacked information on several clinically relevant TB variables, including contact history, latent infection status, radiographic findings, and disease severity, limiting our ability to assess whether GLP-1RA effects differ across TB exposure patterns or disease phenotypes.

4. have the authors looked at dose for the various DM drugs? could there be some kind of dose-effect?

Reply: Thank you for this suggestion. We agree that examining dose-response relationships would strengthen causal inference. However, detailed dosing information was insufficiently standardized across the participating databases to conduct reliable dose-response analyses. We have added this limitation to the Discussion and suggest this as a priority for future studies.

Limitations

Additionally, the TriNetX database does not capture medication dosage or adherence data, preventing assessment of dose-response relationships and the potential impact of treatment compliance on the observed associations. Future studies incorporating these clinical variables are warranted.

5. the authors do not report levels of dysglycemia and presence of dyslipidemia, both of which have previously been linked to Tb risk. can the authors show how HBa1c level links to TB risk? and same for level of dyslipidemia? do these interact with the association with GLP-1 ra's?

Reply: Thank you for this insightful comment. We have now conducted stratified analyses according to HbA1c levels and dyslipidemia status to examine potential effect modification. Interaction p-values were calculated to formally assess whether these metabolic factors modify the association between GLP-1 RAs and TB risk. These results

are now presented in Figure 3 and discussed in the Results section.

6. same, how does BMI interact with medication risk?

Reply: Thank you for this suggestion. We have performed stratified analyses by BMI categories and calculated interaction p-values to assess whether BMI modifies the protective association of GLP-1 RAs with TB risk. These findings are now included in Figure 3.

7. under medication no mention is made of statins, which may also lower TB risk

Reply: Thank you for noting this omission. Given the potential immunomodulatory effects of statins, we have now included statin use as a covariate in our propensity score matching model. The baseline characteristics table has been updated to show statin use before and after matching.

8. can the authors report on TB severity (X-ray extent, bacterial load, level of CRP or anemia, BMI) to show if GLP-1 ra also link to milder disease?

Reply: Thank you for this important question. We agree that examining disease severity would provide valuable insights into whether GLP-1RAs influence not only TB incidence but also disease phenotype. Unfortunately, detailed clinical data on radiographic extent, bacterial load, inflammatory markers, or BMI at TB diagnosis were not systematically captured in our database. We have added this as a limitation and suggest that future studies incorporate such granular clinical data.

Limitations

First, the database lacked information on several clinically relevant TB variables, including contact history, latent infection status, radiographic findings, and disease severity, limiting our ability to assess whether GLP-1RA effects differ across TB exposure patterns or disease phenotypes.

9. I think the discussion on the effects of DM on host immune response should either be improved / much shortened or left out. there are good reviews on this topic (eg SSekematta P. 2023 and otehrs) and there are some fallacies in this section. For instance, TNF production is not consistently down in DM, rather macrophages even

have shown to produce higher levels in some studies. I would similarly much shorten the section on metabolic factors in the discussion.

Reply: Thank you for this critical feedback and we have substantially revised and shortened the sections on immune response and metabolic factors, correcting the oversimplified statements about TNF production and other immune mediators.

Discussion

Diabetes and Increased Risk of Tuberculosis: Immunological and Metabolic Mechanisms

Patients with diabetes mellitus have a markedly increased risk of developing active tuberculosis due to combined defects in adaptive and innate immunity and the effects of chronic hyperglycemia [20, 21]. Diabetes is associated with reduced T-lymphocyte number and function, particularly impaired Th1 responses and production of cytokines and macrophage activation also abnormal [20]. Innate immune defenses are also compromised, as diabetic macrophages and neutrophils exhibit reduced phagocytosis, chemotaxis, and reactive oxygen species production, limiting intracellular killing of M. tuberculosis [22, 23]. Additionally, impaired monocyte migration—unimproved by insulin therapy—further weakens granuloma formation [24]. Chronic hyperglycemia exacerbates these defects by impairing the respiratory burst and promoting oxidative stress and advanced glycation end-product formation, resulting in dysfunctional immune signaling and sustained susceptibility to TB [22-25].

The role of GLP-1RA and infection

GLP-1 receptors (GLP-1R) are expressed on multiple immune cells, including macrophages, monocytes, and B and T lymphocytes, indicating a role in immune regulation [24]. Disruption of GLP-1R signaling impairs lymphocyte proliferation, while GLP-1R activation increases intracellular cAMP, promoting peripheral T-cell proliferation and regulating Treg activity, thereby supporting adaptive immunity and immune tolerance [26].

Preclinical studies suggest that GLP-1RA exert immunomodulatory effects in diabetes-related immune dysfunction [27]. In obese, high-fat diet–induced diabetic mice, GLP-1RA outperformed insulin in restoring host defense against infection by enhancing neutrophil phagocytosis, migration, and bacterial clearance [28]. Prolonged GLP-1RA treatment further improved infection resistance by correcting hyperglycemia, increasing neutrophil counts, and restoring innate immune competence [29].

10. This is a large epidemiological study and I think the discussion should focus on internal and external validity of the epidemiological associations (and brief reference on immunomodulatory effects of GLP-1 ra's). I miss reference ot important epidemiological studies on TB risk like from Julia Critchley and Matthew magee (PMID 39040177, 32061457, 37963989, 30377249, 30104296).

Reply: Thank you for this excellent guidance. We have refocused the Discussion to emphasize the internal and external validity of our epidemiological findings, with appropriate attention to potential biases, confounding, and generalizability. The important epidemiological studies you referenced (PMIDs: 39040177, 32061457, 37963989, 30377249, 30104296) have now been cited and discussed in relation to our findings.

Discussion

The risk of long-term infections increases as HbA1c levels rise. When compared to individuals without T2D, those with T2D - whether optimally controlled or poorly controlled - showed significantly higher risks of hospitalization due to infection [30]. Furthermore, the relationship between TB and T2D appears to be bidirectional. The likelihood of developing TB is heightened among individuals with T2D, while the risk of developing T2D is also elevated in those with a history of TB disease [31]. This suggests that T2D requires more intensive management, and the use of GLP-1RA may not only help in controlling T2D but could potentially reduce the risk of developing TB as well. The prevalence of a positive test for Mycobacterium tuberculosis varies according to an individual's diabetes status, even when stratified by race and ethnicity. Notably, the Hispanic and Asian populations have contributed significantly to the previously observed positive correlation between diabetes and a positive test for Mycobacterium tuberculosis infection. As such, Hispanic and Asian adults with diabetes could particularly benefit from proactive screening and treatment for tuberculosis infection [32]. In addition, African individuals with T2D also exhibit a high prevalence of TB infection, particularly among older adults or those with poorly controlled T2D [33]. These populations require more intensive monitoring and intervention to manage the risk of TB.

11. the labels for Fig 2C should be reversed

Reply: Thank you for catching this error. The labels in Figure 2C have been corrected.

Reviewer #1 (Remarks to the Author):

All the issues raises by this reviewer have been properly addressed.

Reply: We sincerely appreciate your careful review and constructive feedback throughout this revision process. Your suggestions have significantly strengthened the manuscript Thank you for your instructive suggestion.

Reviewer #2 (Remarks to the Author):

I thank the authors for the opportunity to re-review their manuscript. They have addressed most comments raised, and the manuscript is substantially improved as a result. I do have two additional comments.

1. In Figure 2, there is a relatively early splitting of the curves, which may be indicative of a time-related bias. Do the authors have any explanation for the observed early effects?

Reply: Thank you for this insightful observation. We agree that the early divergence of the survival curves in Figure 2 warrants careful consideration. We have thoroughly investigated this finding and believe it reflects a true biological effect rather than a time-related bias, for the following reasons:

First, we conducted a proportional hazards test (Schoenfeld residuals) to assess whether the assumption of proportionality was violated. The test results ($p > 0.05$) indicated no significant deviation from proportional hazards, suggesting that the early separation is consistent with a sustained protective effect rather than a time-dependent artifact.

Second, the early divergence may be biologically plausible given the mechanism of action of GLP-1RAs. These agents have demonstrated immune-modulatory effects and anti-inflammatory properties that could theoretically provide early protection against tuberculosis infection or progression.

Discussion we acknowledge that the relatively early divergence of survival curves observed in Figure 2 requires careful interpretation. While proportional hazards testing did not indicate significant violations ($p > 0.05$), and the early separation may reflect biologically plausible immune-modulatory effects of GLP-1RAs, we cannot entirely exclude the possibility of unmeasured time-related confounding factors that may

have contributed to this pattern.

2. In response to my previous comment, the authors added a sensitivity analysis using an alternative exposure definition in which patients were required to have at least one prescription of the index drug during the first 1 to 5 years following the index date and to have no exposure to comparator antidiabetic drug classes during the follow-up period. By looking into the future to establish eligibility, this analysis has introduced immortal time bias. Patients should be censored at discontinuation or crossover rather than excluded.

Reply: We sincerely thank the reviewer for this critical methodological observation. We fully agree that the sensitivity analysis requiring at least one prescription during years 1-5 while excluding patients with comparator exposure inadvertently introduced immortal time bias by conditioning eligibility on future exposure information. This approach effectively required patients to survive and remain in the study long enough to fulfill these criteria, creating an immortal period that violates fundamental principles of survival analysis. We acknowledge that the appropriate approach would be to censor patients at discontinuation or crossover rather than exclude them based on future medication patterns. Unfortunately, the TriNetX platform does not currently support time-dependent exposure modeling with appropriate censoring at discontinuation or crossover, which prevented us from implementing this preferred analytical strategy. In response to this concern, we have added a clear acknowledgment of this limitation in the manuscript's Limitations section. We explicitly note that this sensitivity analysis introduced immortal time bias and should be interpreted with considerable caution. We emphasize that our primary findings rely on the intention-to-treat active comparator new-user design, which appropriately mitigates immortal time bias by defining exposure at baseline and following all patients from the index date regardless of subsequent treatment changes. We appreciate the reviewer's expertise in bringing this issue to our attention, as it strengthens the transparency and rigor of our reporting. Future studies with access to platforms supporting time-dependent covariate analysis would be better positioned to evaluate sustained exposure effects without introducing such bias.

Discussion
we note that one of our sensitivity analyses, which required patients to have at least

one prescription of the index drug during years 1-5 following the index date while having no exposure to comparator drugs during follow-up, inadvertently introduced immortal time bias by conditioning on future exposure information. Unfortunately, the TriNetX platform does not currently support time-dependent exposure modeling with appropriate censoring at discontinuation or crossover, preventing us from addressing this limitation within the sensitivity analysis. Although our primary analysis employed an active comparator new-user design to mitigate immortal time bias, future investigations using platforms that support time-dependent covariate analysis are warranted to properly evaluate sustained exposure effects.

Reviewer #3 (Remarks to the Author):

I am happy with the revisions and have no further comments

Reply: We appreciate your thorough review and are grateful for your approval of the revisions.

1. Please refrain from using words such as new/novel/first, when referring to the scientific findings. Please also remove exaggerated language such as 'extremely'/'outstanding. Remove the following sentences:
   "To the best of our knowledge, this is the first large-scale investigation utilizing multiple nationwide medical databases to evaluate the relationship between GLP-1RAs and the risk of developing TB in individuals with diabetes."
   Reply: We thank the editor for this suggestion. The requested revisions have been made. The sentence was revised as “This large-scale investigation utilized multiple nationwide medical databases to evaluate the relationship between GLP-1RAs and the risk of developing TB in individuals with diabetes.”
   "In conclusion, this study provides the first large-scale evidence from multiple nationwide medical databases suggesting that the use of GLP-1RAs is associated with a significantly reduced risk of tuberculosis among individuals with T2D."
   Reply: We thank the editor for this suggestion. The requested revisions have been made. The sentence was revised as “In conclusion, this study provides evidence from multiple nationwide medical databases suggesting that the use of GLP-1RAs is associated with a significantly reduced risk of tuberculosis among individuals with T2D.”
2. Please define "HMG CoA reductase inhibitors" in the legend for Table 1. All abbreviations in the figures and tables must be defined.
   Reply: We have revised HMG-CoA reductase inhibitors as statin and reviewed all figures and tables to ensure that all abbreviations are defined in their respective legends.
3. Please remove or rename the Conclusion heading,
   as the main text should only include the sections Abstract, Introduction, Results, optional Discussion and optional Methods. We also allow a combined
   Results and Discussion section.
   Reply: We have removed the “Conclusion” heading from the main text as requested.
4. Please add that you used TriNetX in the abstract; for example line 29-30 could be: "We conducted a multicenter cohort study using TriNetX electronic health records from 143 healthcare organizations across multiple countries between 2017–2025."

Reply: We have revised the abstract accordingly and now explicitly state the use of the TriNetX electronic health record network in the Methods section of the abstract.

5. Table 3 is mentioned in the text but seems to be missing from the actually tables.
   Reply: We thank the editor for identifying this error. The citation should have referred to Table 2 rather than Table 3. This has been corrected throughout the manuscript.
6. Please fix your citations at the beginning of line 323. Please proofread your manuscript.
   Reply: The citation error has been corrected. In addition, we carefully proofread the manuscript and corrected minor typographical and formatting inconsistencies.
7. Wherever statistics have been derived (e.g. error bars, box plots, statistical significance) the legend needs to provide and define the n number (i.e. the sample size used to derive statistics) as a precise value (not a range). Please note that this information is missing in the legend(s) of all of your figures. In addition, All error bars need to be defined in the legends (e.g. SD, SEM) together with a measure of centre (e.g. mean, median). For example, the legends should state something along the lines of “Data are presented as mean values +/- SEM” as appropriate. All box plots need to be defined in the legends in terms of minima, maxima, centre, bounds of box and whiskers and percentile. This information is missing from your figure legends.
   Reply: We have revised the relevant figure legends to include the requested statistical information.
8. Please ensure that all Figures are cited in numerical order. Please note that Production team will place figures after their first citation in the manuscript.
   Reply: We have reviewed the manuscript and corrected the order of figure citations to ensure they appear sequentially in the text.
9. All relevant figure legends should include the template text below: 'Source data are provided as a Source Data file.' A reference to the source data file should be added in the 'Data Availability' section, using the text: "Source data are provided with this paper."
   Reply: We have added the statement “Source data are provided as a Source Data file” to the relevant figure legends and included the statement “Source data are provided with this paper” in the Data Availability sectio.
10. Please remove keywords section from the article file.

Reply: The keywords section has been removed from the manuscript.

11. Wherever statistics have been derived (e.g. error bars, box plots, statistical significance) the legend needs to provide and define the n number (i.e. the sample size used to derive statistics) as a precise value (not a range). Please define how many replicates were performed and whether they are biological (derived from different experimental units or subjects) or technical (multiple contemporary measurements from the same experimental unit or subject). Samples should be unambiguously described, including a clear definitionn of the unit of study. In studies using model organisms, cell lines, primary cell cultures, plants or micro-organisms, the unit of study is the smallest object that could be randomly and independently assigned to an intervention. The groups being compared, including control groups, should be clearly defined. If no control group has been used, the rationale for this should be stated. Please ensure there are enough details about sample collection to distinguish between independent data points and technical replicates- splitting a biological sample into 3 tubes or wells receiving the same treatment does not
    constitute independent replication. We strongly discourage deriving statistics from technical replicates or less than 3 biological replicates, unless there is a clear scientific justification for why providing this information is important. Conflating technical and biological variability, e.g., by pooling technically replicates samples across independent experiments is strongly discouraged.
    Please note that this information is missing in the legend(s) of figure(s) 3a, 3b, 3c, 3d.
    Reply: All the relevant information was added.
12. All error bars need to be defined in the legends (e.g. SD, SEM) together with a measure of centre (e.g. mean, median). For example, the legends should state something along the lines of “Data are presented as mean values +/- SEM” as appropriate. All box plots need to be defined in the legends in terms of minima, maxima, centre, bounds of box and whiskers and percentile. Please note that the error bars/error bands need to be defined in the legend(s) of figure(s) 2a, 2b, 2c, 2d Please note that the measure of centre for the error bars/error bands needs to be defined in the legend(s) of figure(s) 3a, 3b, 3c, 3d.
    Reply: All the relevant information was added.
13. To ensure figures and tables are standalone, please ensure all discipline-specific acronyms are defined in each caption where they are used.

Reply: We reviewed all figure and table legends and ensured that all abbreviations and discipline-specific acronyms are defined upon first use within each caption to facilitate standalone interpretation.

14. Please include your STROBE checklist in the Supplementary Information file.

Reply: The STROBE checklist has been included in the Supplementary Information file as requested.