Nasopharyngeal Carcinoma and Head and Neck Cancer in Patients with Type-2 Diabetes Mellitus receiving SGLT2I, DPP4I or GLP1a: A population-based cohort study

Lifang Li PhD 1 *, Oscar Hou In Chou MSc 2*, Kar Kei Mak 3, Yifan Yang PhD 4, Cheuk To Chung 5,6, Guoliang Li MD PhD 7, Catherine Chan MBChB FRCS 8, Wing Tak Wong PhD 9, Tong Liu MD PhD 10, Bernard Man Yung Cheung PhD FRCP 2, Gary Tse MD PhD FRCP FFPH 10, 11, 12#, Jiandong Zhou PhD 13,14#

1 Department of Biostatistics & Health Informatics, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, United Kingdom
2 Division of Clinical Pharmacology and Therapeutics, Department of Medicine, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China
3 School of Medicine, Shanghai Jiao Tong University, Shanghai, China
4 School of Management, Xi’An Jiao Tong University, Xi’an, China
5 Diabetes Research Unit, Cardiovascular Analytics Group, PowerHealth Research Institute, Hong Kong, China
6 Cancer Research Institute, Quest Genomics, Hong Kong, China
7 Department of Cardiovascular Medicine, The First Affiliated Hospital of Xi’an Jiaotong University, Xi’an, China
8 Department of Ear, Nose and Throat, Prince of Wales Hospital, Hong Kong, China
9 School of Life Sciences, Chinese University of Hong Kong, Hong Kong, China
10 Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, Tianjin 300211, China
11 Kent and Medway Medical School, Canterbury, Kent, CT2 7NT, United Kingdom
12 School of Nursing and Health Studies, Hong Kong Metropolitan University, Hong Kong, China
13 Department of Family Medicine and Primary Care, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China
14 Division of Health Science, Warwick Medical School, University of Warwick, Coventry, United Kingdom

*Co-first authors.
# Correspondence to:
Prof. Gary Tse, MD, PhD, FRCP, FFPH
Tianjin Institute of Cardiology, The Second Hospital of Tianjin Medical University, Tianjin 300211, China
Kent and Medway Medical School, Canterbury, United Kingdom
School of Nursing and Health Studies, Hong Kong Metropolitan University, Hong Kong, China

Email: gary.tse@kmms.ac.uk

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Prof. Jiandong Zhou, PhD

Department of Family Medicine and Primary Care, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China

Division of Health Science, Warwick Medical School, University of Warwick, United Kingdom

Email: Jiandong.Zhou@warwick.ac.uk
Abstract

Introduction: Nasopharyngeal carcinoma (NPC) remains endemic in Asian regions, which risk factors were distinct from other head and neck (H&N) cancers. Anti-diabetic drugs has been proposed to reduce the risk of NPC. The associations between sodium glucose cotransporter 2 inhibitors (SGLT2I) versus dipeptidyl peptidase-4 inhibitors (DPP4I) and the risks of NPC and H&N cancer amongst type-2 diabetes mellitus (T2DM) patients remains unknown.

Method: This was a retrospective population-based cohort study including T2DM patients treated with either SGLT2I or DPP4I between 1st January 2015 and 31st December 2019 in Hong Kong. The primary outcome was new-onset NPC and other H&N cancer. The secondary outcome was cancer-related mortality. Propensity score matching (1:1 ratio) was performed using the nearest neighbour search. Multivariable Cox regression was applied to identify significant predictors.

Results: This cohort included 75,884 patients with T2DM, amongst whom 28,778 patients were on SGLT2I and 47,106 patients were on DPP4I. After matching (57556 patients), 106 patients developed NPC (0.18%), and 50 patients developed H&N cancer (0.08%). Compared to DPP4I, SGLT2I was associated with significantly lower risks of NPC (Hazard ratio [HR]: 0.41; 95% Confidence interval [CI]: 0.21-0.81) but not H&N cancer (HR: 1.00; 95% CI: 0.26-3.92) after adjustments. The result remained significant regardless of demographics and obesity. The association remained consistent in different risk models, matching approaches, and the sensitivity analysis.

Conclusion: This study provides real-world evidence that SGLT2I was associated with lower risks of NPC, but not H&N cancer compared to DPP4I after adjustments amongst T2DM patients.
Keywords: Sodium glucose cotransporter 2 inhibitors, dipeptidyl peptidase-4 inhibitors,

Nasopharyngeal cancer, head and neck cancer, cancer, type-2 diabetes
Introduction

Nasopharyngeal carcinoma (NPC) remains as major public health challenge due to their prevalence and mortality rate, especially in many Asian countries, with a reported incidence up to 50 per 100,000 population.\(^1\) This is in contrast to the prevalence of NPC in the Western countries which has a prevalence of less than 1 per 100. In general, head and neck (H&N) cancer are tumours derived from the mucosa of the aerodigestive tract from the nasal cavity to the larynx.\(^2\) One of the most important H&N cancer is NPC. NPC has a distinct epidemiology, aetiology and pathological characteristics compared to other H&N cancer.\(^3\) NPC is endemic in the southern region of China, as well as other parts of east and southeast Asia.\(^4\) Epstein-Barr virus (EBV) infection remained as the centre of the pathogenesis of NPC. Lifestyle and diet habits such as consuming preserved food which is more likely to be practised by Chinese population and smoking may become another potential risk factors for NPC.\(^5\) Since most patients with NPC are diagnosed with late-stage disease, and that patients diagnosed with metastatic NPC have poor prognosis, there remains to be a need for preventing the risks of new-onset NPC.\(^6\)

Anti-diabetic drugs has been described in reducing the risk of NPC. In a Taiwan study, metformin was associated with a significantly lower risks of NPC compared to metformin non-user amongst T2DM patients.\(^7\) Sodium glucose cotransporter 2 inhibitors (SGLT2I) has demonstrated extensive anti-cancer effects across various malignancy, such that SGLT2I was associated with a reduced risks of cancer compared to placebo or other anti-diabetic drugs.\(^8\)\(^-\)\(^10\) EBV-encoded oncprotein latent membrane protein 1 (LMP1) may increase the level of glucose transporter 1 and therefore enhance glucose uptake by the NPC cells and drives the Warburg effect.\(^6,\)\(^11\) As such, SGLT2I might have help prevent NPC by playing its role in the metabolic reprogramming. However, the effects of SGLT2I on NPC on a population scale...
remained uncertain. However, the association between other second-line anti-diabetic drugs such as SGLT2I and dipeptidyl peptidase-4 inhibitors (DPP4I) remained scarce.

Since NPC remains endemic in Hong Kong, using the territory’s electronic health record data allows studying the association between SGLT2I and NPC using a data with large sample size, which is unlikely to be completed using a prospective study or randomised controlled trials. Hence, this study aimed to investigate the role of SGLT2I and DPP4I with new-onset NPC and H&N cancer in a cohort of T2DM patients from a territory-wide population cohort study.
Methods

This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (HKU/HA HKWC IRB) (UW-20-250) and complied with the Declaration of Helsinki.

Study design and population

This was a retrospective, territory-wide cohort study of T2DM patients who received their treatment with SGLT2I or DPP4I between 1st January 2015, and 31st December 2019, in Hong Kong. The follow-up period is continued until 31st December 2020, or until the patients’ demise. A sensitivity analysis that included a cohort of patients on glucagon-like peptide-1 receptor agonist (GLP1a) between 1st January 2015 and 31st December 2019 was conducted to elucidate the relative effects among the second-line oral antidiabetic medications. The patients were identified from the Clinical Data Analysis and Reporting System (CDARS), a territory-wide database that centralizes patient information from individual local hospitals to establish comprehensive medical data, including clinical characteristics, disease diagnosis, laboratory results, and drug treatment details. The system has been used by local teams in Hong Kong to conduct comparative studies. The exclusion criteria of the study involved patients: 1) With prior NPC or H&N cancer 2) Without complete demographics 3) Under 18 years old (Figure 1). This study employed an as-treat approach, including the censoring of patients at the discontinuation of treatment or switching between the comparison medications.

The demographic details of the patients, including gender and age at the initiation of drug use (baseline), clinical diagnoses and laboratory results as well as medications were extracted for this study. International Classification of Diseases Ninth Edition (ICD-9) codes
were used to identify the prior comorbidities. Financial aid patients were defined based on their participation in the Comprehensive Social Security Assistance (CSSA) scheme, higher disability allowance, normal disability allowance, waiver and other financial aid in Hong Kong. Charlson's standard comorbidity index and the duration and frequency of SGLT2I and DPP4I usage were calculated. To calculate the diabetes duration, the earliest date was examined by the first date of (1) diagnosis using ICD-9; (2) Hba1c≥6.5%; (3) Fasting glucose≥7.0 mmol/l or Random glucose 11.1 mmol/l; (4) using antidiabetic drugs. The estimated glomerular filtration rate (eGFR) was computed using the abbreviated modification of diet in renal disease (MDRD) formula.\textsuperscript{16}

**Adverse outcomes**

The primary outcome of this study was new-onset NPC (ICD-9: 147) and H\&N cancer other than NPC (ICD-9: 140-149, 160-161, excluding 147).\textsuperscript{17} Mortality data were obtained from the Hong Kong Death Registry, an official government registry with the registered death records of all Hong Kong citizens linked to CDARS. Mortality was recorded according to the *International Classification of Diseases Tenth Edition* (ICD-10). The endpoint date of interest for eligible patients was the primary outcome presentation date. The endpoint for those without primary outcome was the mortality date or the endpoint of the study, which is 31\textsuperscript{st} December 2020.

**Statistical analysis**

Descriptive statistics are used to outline the baseline clinical and biochemical characteristics of patients using SGLT2I and DPP4I. The continuous variable for baseline
clinical features was expressed as the mean (95% confidence interval/standard deviation [SD])
while the total number (percentage) represented the categorical variables. The two-tailed
Mann-Whitney U test was utilised to make a comparison between the continuous variables,
whereas the two-tailed Chi-square test with Yates’ correction was applied for 2x2 contingency
data. Propensity score matching with 1:1 ratio for SGLT2I use versus DPP4I use based on
demographics, prior comorbidities, number of prior hospitalisation, duration from T2DM
diagnosis initial drug exposure, non-SGLT2I/DPP4I medications, abbreviated MDRD, time-
weighted HbA1c, and time-weighted fasting glucose were performed using the nearest
neighbour search strategy with a calliper of 0.1. Stata software (Version 16.0) was used in
facilitating the propensity score matching procedures.

In this study, we assessed the comparison of the baseline characteristics of patients
before and after consuming SGLT2I and DPP4I with standardized mean difference (SMD),
where SMD<0.10 was considered as indicative of well-balanced between the two groups. A
Cox proportional hazards regression model was performed to identify significant risk
predictors of adverse study outcomes. Adjustments were made for the demographics, past
comorbidities, duration of diabetes mellitus, and number of prior hospitalizations, number of
anti-diabetic drugs, non-SGLT2I/DPP4I medications, abbreviated MDRD, HbA1c, fasting
glucose, lipid profile. To verify the proportionality assumption for the model mentioned above,
the log-log plot was established. The cumulative incidence curves were generated for both
primary and secondary outcomes and the subgroup analysis was undertaken for the
affirmation of the association among patients with pivotal clinical predictors. P Interaction was
applied to examine the differences in the associations between the subgroups.
To account for the potential competing risks in this study, cause-specific and subdistribution hazard models were also carried out. Multiple propensity score methodologies were implemented, encompassing propensity score stratification, inverse probability of treatment weighting (IPTW) and propensity score with stable inverse probability weighting. We also performed a sensitivity analysis to examine the impacts of the various minimum drug duration on the outcome. The contraindication of the use of SGLT2I included patients with CKD stage 4/5 (eGFR <30 mL/min/1.73m^2), peritoneal dialysis or haemodialysis recipients and thus being excluded for the above analysis. Besides, the exclusion criteria were also included: (1) patients who have primary immunodeficiency; (2) Lag time of 6 months, 1 year and 2 years; (3) 10% propensity score at the 2 ends.

To evaluate the relationship and preference amongst the three novel second-line antidiabetic drugs, the three-arm sensitivity analysis involving GLP1a using stabilized IPTW was performed. As part of a falsification analysis to detect residual bias and confounding factors arising from unobserved variables, a negative control outcome was introduced. To be specified, lower limb fractures (ICD-9 codes: 820-829) were utilized in the falsification analysis. Any observed significant association in this context was interpreted as likely stemming from bias rather than a genuine effect. The final results were demonstrated in terms of hazard ratio (HR), 95% CI and P-value. Statistical significance was defined as P-value < 0.05. All statistical analyses were executed with RStudio software (Version: 1.1.456) and Python (Version: 3.6).

Results

Baseline characteristics
In this territory-wide cohort study of 76,147 patients with T2DM treated with SGLT2I/DPP4I between 1st January 2015 and 31st December 2019 in Hong Kong, patients were followed up until 31st December 2020 or their deaths (Figure 1). The following patient groups were excluded: (1) with prior NPC or H&N cancer (N=113); (2) without complete demographics (N=19); (3) under 18 years old (N=108). After exclusion, this study included a total of 75,884 patients with T2DM (mean age: 62.4 years old [Standard deviation: 12.8]; 55.77% males), of whom 28,778 patients used SGLT2Is, and 47,106 patients used DPP4Is (Table 1). Before matching, the SGLT2 users had higher numbers of male patients, were younger, had lower Charlson’s standard comorbidity index, and had more cases of hypertension, use of anti-diabetic drugs, anti-hypertensive drugs and lipid-lowering drugs, and a higher eGFR compared to DPP4I users. The characteristics of patients are shown in Table 1.

After the propensity score matching, baseline characteristics and the time-weighted lipid and glucose profiles of the 2 groups were well-balanced, apart from baseline age (SMD=0.14), number of hospitalizations (SMD=0.12), thiazolidinedione (SMD=0.13), and abbreviated MDRD (SMD=0.11) (Table 1). The DPP4I and SGLT2I cohorts were comparable after matching with nearest neighbour search strategy with calliper of 0.1, and the proportional hazard assumption was confirmed (Supplementary Figure 1). In the matched cohort, 106 patients developed NPC, 50 patients developed H&N cancer, 899 patients passed away due to cancer during the study period (Figure 1).

The association between SGLT2I and the primary outcomes
In the matched cohort, 31 SGLT2I users and 75 DPP4I users developed NPC. After a mean follow-up of 314,709.1 person-year, the incidence of NPC was lower amongst SGLT2I users (Incidence rate [IR] per 100,000 person-year: 19.43; 95% CI: 13.20-27.58) compared to DPP4I users (IR per 100,000 person-year: 48.33; 95% CI: 38.02-60.58) (Table 2). Compared to DPP4I users, SGLT2I users had a 59% lower risk of NPC after adjustment (Hazard ratio [HR]: 0.41; 95% Confidence Interval [CI]: 0.21-0.81) for demographics, past comorbidities, duration of diabetes mellitus, and number of prior hospitalizations, number of anti-diabetic drugs, non-SGLT2I/DPP4I medications, abbreviated MDRD, HbA1c, fasting glucose, lipid profile. (Table 2).

This was substantiated by the cumulative incidence curves stratified by SGLT2I versus DPP4I (Figure 2).

Meanwhile, 20 SGLT2I users and 30 DPP4I users developed H&N cancer. After a mean follow-up of 314,893.5 person-year, the incidence of NPC was lower amongst SGLT2I users (IR per 100,000 person-year: 12.53; 95% CI: 7.66-19.36) compared to DPP4I users (IR per 100,000 person-year: 19.32; 95% CI: 13.03-27.58) (Table 2). SGLT2I users was not associated with H&N cancer (HR: 1.00; 95% CI: 0.26-3.92) compared to users of DPP4I after adjustment (Table 2).

The association between SGLT2I and cancer-related mortality

180 SGLT2I users and 719 DPP4I users died. After a follow-up of 315,010.8 person-year, the incidence of cancer-related mortality was lower amongst SGLT2I users (IR: 112.76; 95% CI: 96.89-130.49) compared to DPP4I users (IR: 462.74; 95% CI: 429.53-497.84) (Table 2). SGLT2I users had a lower risk of cancer-related mortality after adjustment (HR: 0.42; 95% CI: 0.29-0.60) compared to users of DPP4I. This was substantiated by the cumulative incidence curves stratified by SGLT2I versus DPP4I (Figure 2).
Subgroup analysis

The results of the subgroup analysis for effects of SGLT2I and DPP4I on the NPC are shown in Figure 3. The result demonstrated that SGLT2I was associated with lower risks of NPC regardless of sex, age, socioeconomic status, renal disease, and obesity. Besides, the effects of SGLT2I on NPC were not affected by presence of baseline cancer, autoimmune disease, systemic steroid use, or immunosuppressant/chemotherapy use. The marginal effects analysis demonstrated that SGLT2I use was associated with lower risks of NPC regardless the duration of diabetes (Supplementary Figure 2). However, SGLT2I was associated with lower risks of NPC compared to DPP4I only amongst patient with poorer renal function and higher time-weighted mean of HbA1c. Meanwhile, SGLT2I was associated with lower risks of H&N cancer regardless of the time-weighted mean of HbA1c, but only amongst patients with short duration of diabetes, poorer renal function.

Sensitivity analyses

Sensitivity analyses were performed to confirm the predictability of the models. The results of the cause-specific hazard models, sub-distribution hazard models and different propensity score approaches demonstrated that different models did not change association for both the primary and the secondary outcomes (Supplementary Table 2). A three-arm analysis with the inclusion of GLP1a including patients only on (SGLT2I, DPP4I, and GLP1a) using stabilized IPTW was conducted (Supplementary Table 3). While SGLT2I was also associated with higher risks of NPC compared to GLP1a, the results remained insignificant due to the insufficient power in the GLP1a cohort (HR: 1.24; 95% CI: 0.79-1.66). Besides, SGLT2I...
also did not associated with a lower risk of cancer-related mortality compared to GLP1a (HR: 1.09; 95% CI: 0.56-1.95). The sensitivity analysis for 6 months, 1-year and 2-year lag time demonstrated the same pattern. Besides, after excluding patients with extreme 10% propensity score at the 2 end, patients with CKD stage 4/5 (eGFR <30), peritoneal dialysis or haemodialysis, and patients with primary immunodeficiency likewise, the results demonstrated that SGLT2I remained associated with lower risks of NPC compared to DPP4I (Supplementary Table 4).

**Falsification analysis**

Lower limb fracture was used as negative control outcome in the falsification analysis for the comparison between SGLT2I and DPP4I since most of the cases should be traumatic cause (Supplementary Table 9). The results demonstrated that SGLT2I was not associated with lower limb fracture after adjustment (HR: 1.29; 95% CI: 0.95-1.79) compared to DPP4I.

**Discussion**

In this territory-wide cohort study, the associations between SGLT2I, DPP4I and new-onset NPC and H&N were examined using real-world data. Our results firstly demonstrate that SGLT2I use was independently associated with a lowered risk of new-onset NPC than DPP4I usage after adjustments. Secondly, SGLT2I was not associated with lower risks of H&N cancer compared to DPP4I. The results remained consistent across different competing risk and the sensitivity analyses, and the falsification analysis did not falsify the results. To the best of our knowledge, this was the first cohort study to examine the association between SGLT2I and NPC.
Comparison with previous studies

The incidence of NPC in our study for SGLT2I and for DPP4I parallels the values of existing study, with a reported incidence of 15-50 per 100,000 population.¹ NPC is endemic in Southern China, as well as other parts of South-East Asia.²¹ In Hong Kong, a previous study described an incidence of 10.1 per 100,000 in male and 3.3 per 100,000 in female in 2020.²² Previously, SGLT2I was proposed as an potential anti-tumour agents since sodium-glucose cotransporter 2 was over-expressed in multiple tumour models, including head and neck cancer.²³, ²⁴ In this study, we found that SGLT2I was associated with lower risk of NPC compared to DPP4I.

Previously, metformin was reported to be associated with a lower risk of NPC compared to metformin non-user.⁷, ²⁵ ²⁶ However, the effects of the anti-diabetic drugs on NPC might be independent of their ability of glycaemic control. Previously, a meta-analysis identified that T2DM patients might have decreased risk of NPC amongst the Asian population. This could be explained by the pathological phenotypes of undifferentiated NPC. It was found that the prevalence of T2DM in differentiated NPC was higher compared to undifferentiated cases, such that T2DM may only effect on the former pathological subtype.²⁷ T2DM was also reported not to be associated with mortality amongst NPC patients.²⁸ Meanwhile, in a prospective cohort study in Taiwan, diabetic patients who received metformin had a lower risk of developing H&N cancer compared to the matched individuals without metformin. Specifically, the incidence of oropharyngeal cancer and NPC was significantly lower for patients treated with metformin than their counterparts.²⁹ Another study also supported a reduced risk of H&N cancer for metformin, particularly with a history of smoking and heavy
Apart from anti-diabetic medications, a Taiwan case-control study found that the risk of head and neck cancer is reduced by statin. Several studies also demonstrated that patients received statin were associated with increased survival of head and neck cancer. Hence, metabolic drugs might potentially be preventive for head and neck cancer.

Potential underlying mechanisms

SGLT2 inhibitors have recently reported to exhibit possible anti-cancer effects in several malignancies. It was proposed that SGLT2I act to inhibit the active reverse transport of glucose by SGLT-2 located in the proximal renal tubule, which consequently affect the cellular respiration and growth of the tumour cells. Our results suggested that the effects of SGLT2I mainly targeted specifically on NPC but not H&N cancer in general. Besides, SGLT2I may act on the LMP1 induced glucose transporter 1 which reduce glucose uptake by the NPC cells. Meanwhile, DPP4I were hypothesized to diminish the cancer risk by eliminating the degradation of incretins, glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic peptide (GIP) in the cell. Although there are multiple studies regarding both SGLT2 inhibitors and DPP4 inhibitors act on cancer, the mechanism of both medications towards head and neck cancer, especially NPC, remains unclear.

Clinical implications

The secondary protective benefits of SGLT2I and DPP4I on cardiovascular benefits and cancer have received worldwide attention in recent years. While there is substantial evidence supporting the effects of the anti-diabetic medications on cardiovascular diseases, few evidence surrounding the latter. By exploring the association of SGLT2I and DPP4I in HCCNPC
this study added to the growing body of evidence supporting the use of antidiabetic agents in preventing cancer. However, further molecular studies and clinical studies are needed to confirm the causation relationship between SGLT2I and NPC.

**Limitations**

Several limitations should be acknowledged in this study. First of all, the observational nature of this study gives inherent issues such as under-coding, coding errors, and missing data possibly resulting in information bias. Crucial data on predictive variables, for instance, nitrosamine consumption, diet, and family history of NPC were unavailable from CDARS. In addition, the inclusion of patients in CDARS was only limited to those who smoked, consumed alcohol, and were obese to a degree that required medical attention and treatment. To mitigate this limitation, this study incorporated extensive comorbidities related to NPC to indirectly infer the possible risk variables. Furthermore, a falsification analysis was also conducted to minimise the risk of residual confounding. Moreover, the retrospective design allows only for the inference of associations but not the demonstration of causation relationships. Lastly, assessing medication adherence relied solely on prescription refills, lacking direct measurement of drug exposure. This approach may lead to time lag biases and immortal time issues. Hence, it is essential to conduct prospective randomised controlled trials aimed at evaluating the causal links of anti-diabetic medications.

**Conclusion**
In this population-based cohort study, SGLT2I users was associated with lower risks of NPC but not H&N cancer compared to DPP4I users after matching and adjustments compared to DPP4I users in Asia. Future studies are needed about the SGLT2I and its action on NPC.

**Ethical approval statement**

This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (HKU/HA HKWC IRB) (UW-20-250) and complied with the Declaration of Helsinki.

**Availability of data and materials**

The data that support the findings of this study were provided by the Hong Kong Hospital Authority, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Hong Kong Hospital Authority.

**Conflicts of Interest**

None.

**Acknowledgements**
Oscar Hou In Chou and Kar Kei Mak would like to thank Bai Xian Asia Institute (BXAI) for her generosity in providing Asian Future Leaders Scholarship Program scholarship.

**Funding**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Guarantor Statement**

All authors approved the final version of the manuscript. GT 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.

**Author contributions**

LF L, OHI C, JD Z: conception of study, preparation of figures, study design, data contribution, statistical analysis, data interpretation, manuscript drafting, and critical revision of the manuscript.

OHI C, KK M, YF Y, CT C: literature search, data interpretation, literature search, data collection, manuscript drafting.

GL L, C C, WT W, T L, BMY C: critical revision of the manuscript

GT, JD Z: conception of study and literature search, study design, data collection, and critical revision of manuscript, study supervision.
References


SGLT2I: Sodium-glucose cotransporter-2 inhibitors; DPP4I: Dipeptidyl peptidase-4 inhibitors; AMDRD: abbreviated modification of diet in renal disease.
New onset nasopharyngeal cancer

Cumulative incidence curve by drug use after propensity score matching

New onset salivary gland and other head/neck cancer

Cumulative incidence curve by drug use after propensity score matching
Figure 2. Cumulative incidence curves for new onset NPC, other H&N cancer, cancer-related mortality, and all-cause mortality stratified by drug exposure of SGLT2I and DPP4I after propensity score matching (1:1)

SGLT2I: Sodium-glucose cotransporter-2 inhibitors; DPP4I: Dipeptidyl peptidase-4 inhibitors.

NPC: nasopharyngeal carcinoma; H&N: head and neck.
Figure 3. Subgroup analyses for SGLT2I v.s. DPP4I exposure predict new onset NPC in the matched cohort.
* For p ≤ 0.05, ** for p ≤ 0.01, *** for p ≤ 0.001; CI: confidence interval; SGLT2I: Sodium-glucose cotransporter-2 inhibitors; DPP4I: Dipeptidyl peptidase-4 inhibitors; HR: Hazard ratio; NPC: nasopharyngeal carcinoma.
Table 1. Baseline and clinical characteristics of patients with SGLT2I v.s. DPP4I use before and after propensity score matching (1:1).

* for SMD >0.1; SGLT2I: sodium glucose cotransporter-2 inhibitor; DPP4I: dipeptidyl peptidase-4 inhibitor; MDRD: modification of diet in renal disease; ACEI: angiotensin-converting enzyme inhibitors; ARB: angiotensin II receptor blockers.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Before matching</th>
<th>After matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All (N=75884)</td>
<td>SGLT2I users (N=28778)</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male gender</td>
<td>42323 (55.77%)</td>
<td>17127 (59.51%)</td>
</tr>
<tr>
<td>Female gender</td>
<td>33561 (44.22%)</td>
<td>11651 (40.48%)</td>
</tr>
<tr>
<td>Baseline age, years</td>
<td>62.4 (12.8); n=7588</td>
<td>57.9 (11.3); n=28778</td>
</tr>
<tr>
<td>Financial aid</td>
<td>2790 (3.67%)</td>
<td>1167 (4.05%)</td>
</tr>
<tr>
<td>Past comorbidities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charlson's standard comorbidity index</td>
<td>2.0 (1.5); n=7588</td>
<td>1.6 (1.3); n=28778</td>
</tr>
<tr>
<td>Number of hospitalizations</td>
<td>1.2 (0.8); n=7588</td>
<td>1.3 (1.1); n=28778</td>
</tr>
<tr>
<td>Hyperlipidaemia</td>
<td>55285 (72.85%)</td>
<td>21876 (76.01%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>43161 (56.87%)</td>
<td>19179 (66.64%)</td>
</tr>
<tr>
<td>Diabetic retinopathy</td>
<td>5309 (0.69%)</td>
<td>1953 (6.78%)</td>
</tr>
<tr>
<td>Diabetic nephropathy</td>
<td>934 (1.23%)</td>
<td>172 (0.59%)</td>
</tr>
<tr>
<td>Diabetic neuropathy</td>
<td>1842 (2.42%)</td>
<td>680 (2.36%)</td>
</tr>
<tr>
<td>Baseline cancer</td>
<td>1940 (2.55%)</td>
<td>564 (1.95%)</td>
</tr>
<tr>
<td>Renal diseases</td>
<td>6096 (8.03%)</td>
<td>1536 (5.33%)</td>
</tr>
<tr>
<td>Liver diseases</td>
<td>2369 (3.12%)</td>
<td>1053 (3.65%)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>2505 (3.30%)</td>
<td>757 (2.63%)</td>
</tr>
<tr>
<td>Acute myocardial infarction</td>
<td>21582 (2.84%)</td>
<td>955 (3.31%)</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>7860 (10.35%)</td>
<td>3619 (12.57%)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>1979 (2.60%)</td>
<td>656 (2.27%)</td>
</tr>
<tr>
<td>Stroke/transient ischemic attack</td>
<td>2418 (3.18%)</td>
<td>739 (2.56%)</td>
</tr>
<tr>
<td>Autoimmune disease</td>
<td>809 (1.06%)</td>
<td>312 (1.08%)</td>
</tr>
<tr>
<td>HPV infection</td>
<td>128 (0.16%)</td>
<td>45 (0.15%)</td>
</tr>
<tr>
<td>Primary immunodeficiency</td>
<td>492 (0.64%)</td>
<td>105 (0.36%)</td>
</tr>
<tr>
<td>Overweight, obesity and hyperalimentation required treatment</td>
<td>576 (0.75%)</td>
<td>421 (1.46%)</td>
</tr>
<tr>
<td>Alcohol dependence required treatment</td>
<td>169 (0.22%)</td>
<td>42 (1.4%)</td>
</tr>
<tr>
<td>Tobacco dependence required treatment</td>
<td>1066 (1.40%)</td>
<td>163 (0.56%)</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of anti-diabetic drug classes</td>
<td>3.4 (0.9); n=75884</td>
<td>3.5 (1.0); n=28778</td>
</tr>
<tr>
<td>Metformin</td>
<td>68140 (89.79%)</td>
<td>26863 (93.34%)</td>
</tr>
<tr>
<td>Sulphonylurea</td>
<td>58324 (76.85%)</td>
<td>20695 (71.91%)</td>
</tr>
<tr>
<td>Insulin</td>
<td>38694 (50.99%)</td>
<td>14795 (51.41%)</td>
</tr>
<tr>
<td>Acarbose</td>
<td>2043 (2.69%)</td>
<td>1113 (3.86%)</td>
</tr>
<tr>
<td>Thiazolidinedione</td>
<td>15218 (20.05%)</td>
<td>7849 (27.27%)</td>
</tr>
</tbody>
</table>

**Notes:**
- All results are presented as mean(SD); N or Count(%).
- For SMD >0.1, the effect size is considered clinically significant.
- The table includes the count and percentage for each characteristic.
- All statistical tests were two-tailed, and P <0.05 was considered statistically significant.
- The significance level was adjusted to control for multiple comparisons using the Bonferroni correction.
- The sample size for each group is 75884 for SGLT2I users and 47106 for DPP4I users.
- The study was approved by the institutional review board, and all patients provided written informed consent.
<table>
<thead>
<tr>
<th>Category</th>
<th>Value 1 (Mean)</th>
<th>Value 2 (Mean)</th>
<th>Value 3 (Mean)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Glucagon-like peptide-1 receptor agonists</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE/ARB</td>
<td>26842(35.37%)</td>
<td>10601(36.83%)</td>
<td>16241(34.47%)</td>
<td>0.00</td>
</tr>
<tr>
<td>Diuretics</td>
<td>41874(55.18%)</td>
<td>20840(72.41%)</td>
<td>21034(44.65%)</td>
<td>0.00</td>
</tr>
<tr>
<td>Statins and fibrates</td>
<td>40813(53.78%)</td>
<td>18266(63.47%)</td>
<td>22547(47.86%)</td>
<td>0.00</td>
</tr>
<tr>
<td>Systemic steroids</td>
<td>434(0.57%)</td>
<td>306(1.06%)</td>
<td>128(0.27%)</td>
<td>0.00</td>
</tr>
<tr>
<td>Immunosuppressant and chemotherapy</td>
<td>2536(3.34%)</td>
<td>564(1.95%)</td>
<td>1972(4.18%)</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Calculated biomarkers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abbreviated MDRD, mL/min/1.73m²</td>
<td>80.8(28.4);n=63</td>
<td>89.8(24.1);n=24</td>
<td>75.2(29.4);n=38</td>
<td>0.5*</td>
</tr>
<tr>
<td>Time weighted mean of triglyceride, mmol/L</td>
<td>1.8(1.0);n=3885</td>
<td>1.9(1.0);n=1739</td>
<td>1.8(1.0);n=2146</td>
<td>0.00</td>
</tr>
<tr>
<td>Time weighted mean of low-density lipoprotein, mmol/L</td>
<td>2.3(0.6);n=3648</td>
<td>2.3(0.56);n=162</td>
<td>2.34(0.59);n=20</td>
<td>0.00</td>
</tr>
<tr>
<td>Time weighted mean of high-density lipoprotein, mmol/L</td>
<td>1.2(0.2);n=3833</td>
<td>1.16(0.19);n=17</td>
<td>1.18(0.21);n=21</td>
<td>0.00</td>
</tr>
<tr>
<td>Time weighted mean of total cholesterol, mmol/L</td>
<td>4.3(0.6);n=1242</td>
<td>4.4(0.6);n=2147</td>
<td>4.4(0.6);n=3256</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Glucose tests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time weighted mean of HbA1C, %</td>
<td>7.9(1.4);n=4617</td>
<td>7.9(1.4);n=1983</td>
<td>7.8(1.4);n=2633</td>
<td>0.00</td>
</tr>
<tr>
<td>Time weighted mean of fasting glucose, mmol/L</td>
<td>7.8(3.0);n=4283</td>
<td>7.9(2.9);n=1857</td>
<td>7.8(3.0);n=2426</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table 2. Incidence and multivariate Cox regression models of new onset NPC, other H&N cancer and cancer-related mortality in the cohort after 1:1 propensity score matching.

* For p ≤ 0.05, ** for p ≤ 0.01, *** for p ≤ 0.001; CI: confidence interval; SGLT2I: sodium glucose cotransporter-2 inhibitor; DPP4I: dipeptidyl peptidase-4 inhibitor; IQR: Interquartile range.

Adjusted for significant demographics, past comorbidities, duration of diabetes mellitus, and number of prior hospitalizations, number of anti-diabetic drugs, non-SGLT2I/DPP4I medications, abbreviated MDRD, HbA1c, fasting glucose, lipid profile.

<table>
<thead>
<tr>
<th>NPC</th>
<th>Number of patients</th>
<th>Number of events</th>
<th>Total person-year</th>
<th>Follow-up duration, year [IQR]</th>
<th>Incidence per 100,000 person-year [95% CI]</th>
<th>Adjusted hazard ratio [95% CI]</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DPP4I</td>
<td>28778</td>
<td>75</td>
<td>155176.8</td>
<td>5.4[0.7]</td>
<td>48.33[38.02-60.58]</td>
<td>1 [Reference] NA</td>
</tr>
<tr>
<td>Other H&amp;N cancer</td>
<td>SGLT2I</td>
<td>28778</td>
<td>31</td>
<td>195532.3</td>
<td>5.5[0.4]</td>
<td>19.43[13.20-27.58]</td>
<td>0.41[0.21-0.81] 0.0095**</td>
</tr>
<tr>
<td>Cancer-related</td>
<td>Number of patients</td>
<td>Number of events</td>
<td>Total person-year</td>
<td>Follow-up duration, year [IQR]</td>
<td>Incidence per 100,000 person-year [95% CI]</td>
<td>Adjusted hazard ratio [95% CI]</td>
<td>P-value</td>
</tr>
<tr>
<td>mortality</td>
<td>DPP4I</td>
<td>28778</td>
<td>30</td>
<td>155306.8</td>
<td>5.4[0.7]</td>
<td>19.32[13.03-27.58]</td>
<td>1 [Reference] NA</td>
</tr>
<tr>
<td></td>
<td>SGLT2I</td>
<td>28778</td>
<td>20</td>
<td>159586.7</td>
<td>5.5[0.4]</td>
<td>12.53[7.66-19.36]</td>
<td>1.00[0.26-3.92] 0.9985</td>
</tr>
<tr>
<td>Cancer-related</td>
<td>Number of patients</td>
<td>Number of events</td>
<td>Total person-year</td>
<td>Follow-up duration, year [IQR]</td>
<td>Incidence per 100,000 person-year [95% CI]</td>
<td>Adjusted hazard ratio [95% CI]</td>
<td>P-value</td>
</tr>
<tr>
<td>mortality</td>
<td>DPP4I</td>
<td>28778</td>
<td>719</td>
<td>155377.9</td>
<td>5.4[0.7]</td>
<td>462.74[429.53-497.84]</td>
<td>1 [Reference] NA</td>
</tr>
<tr>
<td></td>
<td>SGLT2I</td>
<td>28778</td>
<td>180</td>
<td>159632.9</td>
<td>5.5[0.4]</td>
<td>112.76[96.89-130.49]</td>
<td>0.42[0.29-0.60] &lt;0.0001***</td>
</tr>
</tbody>
</table>