

1 **Compare SGLT2I versus non-SGLT2I users in type-2 diabetic mellitus patients on GLP-1**
2 **receptor agonist: A population-based and machine learning causal inference analysis**

3
4 Zhiyao Luo MSc ^{1*}, Oscar Hou-In Chou MSc ^{*2}, Zita Man Wai Ng ³, Cheuk To Skylar Chung ³,
5 Jeffrey Shi Kai Chan MBChB ³, Raymond Ngai Chiu Chan MBChB ³, Lei Lu Ph.D.¹, Tingting Zhu
6 Ph.D. ¹, Bernard Man Yung Cheung MB BChir FRCP ², Tong Liu MD ⁴, Gary Tse MD DM Ph.D.
7 FRCP FFPH ^{4,5,6,^}, Jiandong Zhou Ph.D. ^{7,^}

8
9 ¹Institute of Biomedical Engineering, Department of Engineering Science, University of
10 Oxford, Oxford, United Kingdom

11 ²Division of Clinical Pharmacology and Therapeutics, Department of Medicine, Li Ka Shing
12 Faculty of Medicine, The University of Hong Kong, Hong Kong, China

13 ³Diabetes Research Unit, Cardiovascular Analytics Group, PowerHealth Limited, Hong Kong,
14 China

15 ⁴Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department
16 of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University,
17 Tianjin 300211, China

18 ⁵Kent and Medway Medical School, Canterbury, Kent, CT2 7NT, United Kingdom

19 ⁶School of Nursing and Health Studies, Hong Kong Metropolitan University, Hong Kong,
20 China

21 ⁷Division of Health Science, Warwick Medical School, University of Warwick, Coventry,
22 United Kingdom

23
24
25 * Contributed equally

26 ^ Correspondence to:

27 Gary Tse, MD DM PhD FRCP

28 Kent and Medway Medical School, Canterbury, United Kingdom

29 Tianjin Institute of Cardiology, The Second Hospital of Tianjin Medical University, Tianjin
30 300211, China

31 School of Nursing and Health Studies, Hong Kong Metropolitan University, Hong Kong, China

32 Email: garytse86@gmail.com

33
34 Jiandong Zhou, PhD,

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

36 Email: jiandong.zhou@warwick.ac.uk

39 **Abstract**

40 **Background:** Both sodium-glucose cotransporter-2 (SGLT2) inhibitors and GLP-1 receptor
41 agonists (GLP1a) demonstrated benefits against cardiovascular diseases in type 2 diabetes
42 (T2D). However, the effects of SGLT2I amongst patients already on GLP1a users remain
43 unknown.

44 **Objective:** This real-world study compared the risks of cardiovascular diseases with and
45 without exposure to SGLT2I amongst GLP1a users.

46 **Methods:** This was a retrospective population-based cohort study of patients with type-2
47 diabetes mellitus (T2DM) on GLP1a between 1st January 2015 and 31st December 2020 using
48 a territory-wide registry from Hong Kong. The primary outcomes were new-onset myocardial
49 infarction, atrial fibrillation, heart failure, and stroke/transient ischaemic attack (TIA). The
50 secondary outcome was all-cause mortality. Propensity score matching (1:2 ratio) using the
51 nearest neighbour search was performed. Multivariable Cox regression was used to identify
52 significant associations. The machine learning causal inference analysis was used to estimate
53 the treatment effect.

54 **Results:** This cohort included 2526 T2DM patients on GLP1a (median age: 52.5 years old [SD:
55 10.9]; 57.34 % males). The SGLT2I users and non-SGLT2I users consisted of 1968 patients and
56 558 patients, respectively. After matching, non-SGLT2I users were associated with high risks
57 of myocardial infarction (Hazard ratio [HR]: 2.91; 95% Confidence Interval [CI]: 1.30-6.59) and
58 heart failure (HR: 2.49; 95% CI: 1.22-5.08) compared to non-SGLT2I users after adjusting for
59 demographics, comorbidities, medications, renal function, and glycaemic tests. However,
60 non-SGLT2I users were not associated with the risks of atrial fibrillation (HR: 1.52; 95% CI:
61 0.65-3.53) and stroke/TIA (HR: 1.72; 95% CI: 0.70-4.24). The results remained consistent in
62 the competing risk and the sensitivity analyses.

63 **Conclusions:** SGLT2I non-users was associated with higher risks of myocardial infarction and
 64 heart failure when compared to SGLT2I users after adjustments amongst T2DM patients on
 65 GLP1a. The result remained consistent in the machine learning causal inference analysis.

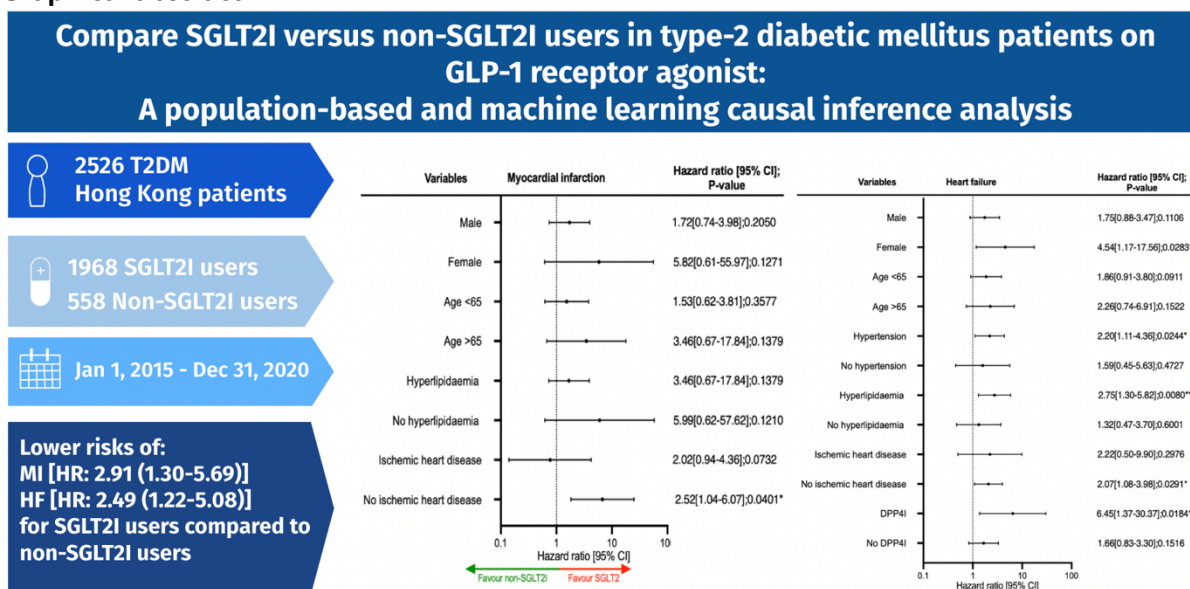
66

67 **Keywords**

68 Sodium-glucose cotransporter-2 inhibitors; glucagon-like peptide-1 receptor agonist;
 69 cardiovascular disease; Myocardial infarction; Stroke; Atrial fibrillation

70

71 **Graphical abstract**



72

73 Introduction

74 Cardiovascular diseases are one of the leading causes of mortality and morbidity
75 worldwide amongst patients with type 2 diabetes mellitus (T2DM) [1]. To reduce the risks of
76 cardiovascular diseases, current clinical guidelines highlight the importance of a multifactorial
77 approach in diabetic management, including the use of anti-diabetic agents [2, 3]. The current
78 literature suggests that novel anti-diabetic medications such as sodium-glucose
79 cotransporter-2 inhibitors (SGLT2i) and glucagon-like peptide receptor agonists (GLP1a)
80 demonstrate promising protective effects against various major adverse cardiovascular
81 events (MACE) and mortality [4-6].

82 Some studies suggested that the cardiovascular benefit of SGLT2i and GLP1a is
83 enhanced when used in conjunction [7]. A real-world cohort study demonstrates that the
84 addition of SGLT2i amongst GLP1a diabetic users showed superior cardioprotective effects
85 compared to the addition of sulfonylurea [8]. It was proposed that the molecular mechanism
86 of the cardioprotective effects of SGLT2i and GLP1a could complement each other. However,
87 the real-world evidence surrounding whether adding SGLT2i would benefit cardiovascular
88 outcomes amongst patients using GLP1a remains limited. Previously, it was suggested that
89 amongst patients on SGLT2i, only 4.4% of the patients had baseline GLP1a uses [9]. The recent
90 development of machine learning (ML) for causal inference may also be applied to better
91 characterize the treatment effects of anti-diabetic agents [10]. Therefore, this present study
92 aims to examine the effect of SGLT2i and GLP1a on cardiovascular outcomes with ML causal
93 inference in a T2DM cohort from Hong Kong.

94

95 Methods

96 This study was approved by the Institutional Review Board of the University of Hong
97 Kong/Hospital Authority Hong Kong West Cluster Institutional Review Board (HKU/HA HKWC
98 IRB) (UW-20-250 and UW 23-339) and complied with the Declaration of Helsinki.

99

100 *Study population*

101 This was a retrospective population-based study of prospectively collected electronic
102 health records using the Clinical Data Analysis and Reporting System (CDARS) by the Hospital
103 Authority (HA) of Hong Kong. The records cover both public hospitals and their associated
104 outpatient clinics as well as ambulatory and day-care facilities in Hong Kong. This system has
105 been used extensively by our teams and other research teams in Hong Kong, including
106 diabetes and medication research.[11, 12] The system contains data on disease diagnosis,
107 laboratory results, past comorbidities, clinical characteristics, and medication prescriptions.
108 T2DM patients who were administered with GLP1a in centres under the HA, between 1st
109 January 2015, to 31st December 2020, were included.

110

111 *Predictors and variables*

112 Patients' demographics include gender and age of initial drug use (baseline), clinical
113 and biochemical data were extracted for the present study. Prior comorbidities were
114 extracted by the *International Classification of Diseases Ninth Edition* (ICD-9) codes
115 (**Supplementary Table 1**). The diabetes duration was calculated by examining the earliest
116 date amongst the first date of (1) diagnosis using ICD-9; (2) HbA1c $\geq 6.5\%$; (3) Fasting glucose
117 ≥ 7.0 mmol/l or Random glucose ≥ 11.1 mmol/l; (4) using anti-diabetic medications. The patients
118 on financial aid were defined as patients on the Comprehensive Social Security Assistance
119 (CSSA) scheme, higher disability allowance, normal disability allowance, waiver, and other

120 financial aid in Hong Kong. The number of hospitalisations in the year prior to the index days
121 was extracted. The Charlson standard comorbidity index was calculated. [13] The baseline
122 laboratory examinations, including the glucose profiles and renal function tests, were
123 extracted. The estimated glomerular filtration rate (eGFR) was calculated using the
124 abbreviated modification of diet in renal disease (MDRD) formula. [14] Furthermore, the
125 time-weighted lipid and glucose profiles after drug initiation were also calculated by the
126 products of the sums of two consecutive measurements and the time interval, then divided
127 by the total time interval, as suggested previously.[15]

128

129 *Study outcomes*

130 The primary outcome of this study was new-onset PAD defined using ICD-9 code
131 enlisted in **Supplementary Table 1**. Mortality data were obtained from the Hong Kong Death
132 Registry, a population-based official government registry with the registered death records of
133 all Hong Kong citizens linked to CDARS. Mortality was recorded using the *International*
134 *Classification of Diseases Tenth Edition* (ICD-10) coding. The as-treat approach was adopted
135 which patients were censored at treatment discontinuation or switching of the comparison
136 medications. The endpoint date of interest for eligible patients was the event presentation
137 date. The endpoint for those without primary outcomes was the mortality date or the end of
138 the study period (31st December 2020).

139

140 *Statistical analysis*

141 Descriptive statistics are used to summarize baseline clinical and biochemical
142 characteristics of patients with and without SGLT2I use. For baseline clinical characteristics,

143 continuous variables were presented as mean (95% confidence interval/standard deviation),
144 and the categorical variables were presented as total numbers (percentage). Propensity score
145 matching generated by logistic regression with a 1:2 ratio for SGLT2I users versus non-users
146 based on demographics, prior anti-diabetic drugs, number of prior anti-diabetic drugs, prior
147 comorbidities, renal function, duration from T2DM diagnosis initial drug exposure, HbA1c and
148 fasting glucose were performed using the nearest neighbour search strategy with a calliper of
149 0.1 (**Figure 1**). Propensity score matching was performed using Stata software (Version 16.0).
150 Baseline characteristics between patients with and without SGLT2I use before and after
151 matching were compared with absolute standardized mean difference (SMD), with $SMD < 0.20$
152 regarded as well-balanced between the two groups.

153 The cumulative incidence curves for the primary outcomes and secondary outcomes
154 were constructed. Cox proportional hazards regression was used to identify significant risk
155 predictors of adverse study outcomes in the matched cohort, with adjustments for
156 demographics, comorbidities, number of prior hospitalisations, medication profile, renal
157 function, glycaemic tests, and the duration of T2DM. The log-log plot was used to verify the
158 proportionality assumption for the proportional Cox regression models. Subgroup analyses
159 were conducted to confirm the association amongst patients with different clinically
160 important predictors accounting to the diabetic and the metabolic profile, as well as the
161 comorbidities and medications associated with the outcome.

162 Cause-specific and sub-distribution hazard models were conducted to consider possible
163 competing risks. Multiple propensity adjustment approaches were used, including propensity
164 score stratification,[16] propensity score with inverse probability of treatment weighting
165 (IPTW) [17] and propensity score with stable inverse probability weighting.[18]. In the
166 sensitivity analysis, patients with less than 3-month follow-up duration, patients with chronic

167 kidney disease (CKD) stage 4/5 (eGFR <30 mL/min/1.73m²), patients on financial aid, and
168 patients at the top or bottom 5% of propensity score matching were excluded to test the
169 robustness of the association. We used the hip fracture as the negative control in the
170 falsification analysis, such that the observed significant association in the falsification analysis
171 should be attributed to bias. The hazard ratio (HR), 95% CI, and P-value were reported.
172 Statistical significance was defined as P-value <0.05. All statistical analyses were performed
173 with RStudio (Version: 1.1.456) and Python (Version: 3.6).

174

175 *Causal Random Forest Model*

176 In clinical research, discerning causal links between treatments and outcomes is
177 essential. While prediction models offer insights into probable outcomes based on various
178 factors, they do not capture intervention causal effects[19], which is crucial distinction when
179 assessing intervention efficacy [20].

180 Causal Random Forests (CRFs) stand at the confluence of machine learning and causal
181 inference, leveraging the robust capabilities of random forests to specifically estimate
182 heterogeneous treatment effects [21]. CRFs adeptly uncover intricate, non-linear dynamics
183 between treatments and outcomes. CRFs manage high-order interactions, pinpointing
184 patient subgroups deriving maximum or minimum benefit from a treatment.

185 Let $Y(t)$ be the potential outcome at time t under treatment A , where A can be 0
186 (no SGLT-2) or 1 (SGLT-2). Our objective is to estimate the Conditional Average Treatment
187 Effect (CATE) for time t :

$$188 \quad CATE(t) = E[Y(t) | A = 1] - E[Y(t) | A = 0]$$

189 For a given patient i , let W_i represent its covariates, A_i its treatment status, and $Y_i(t)$
190 the observed outcome at time t . In CRFs, each tree b in the forest aims to predict the
191 potential outcome differences $Y_i(t)$ for treatment versus control. The splitting criterion in
192 CRFs is based on maximizing the difference in treatment effects between the two child nodes.
193 Specifically, the impurity decrease in a node split is calculated as:

$$\Delta I = \frac{n_{left}}{n} Var_{left}(Y(t)|A = 1) - Var_{left}(Y(t)|A = 0) +$$
$$\frac{n_{right}}{n} Var_{right}(Y(t) | A = 1) - Var_{right}(Y(t) | A = 0)$$

194
195 where n is the total number of observations in the current node, n_{left} and n_{right} are the
196 number of observations in the left and right child nodes, and Var_{left} and Var_{right} are the
197 variances of the potential outcomes in the left and right child nodes, respectively.
198

199 In essence, the CRF optimization process is driven by finding splits that result in nodes
200 with high differences in treatment effects, which, in turn, reveals subgroups with varying
201 responses to treatment. Once the forest is constructed, the CATE for a new observation x at
202 time t is estimated as:

$$\widehat{CATE}(t, x) = \frac{1}{B} \sum_{b=1}^B \frac{1}{n_{b,1}} \left(\sum_{i:A_i=1} Y_i(t) - \frac{1}{n_{b,0}} \sum_{i:A_i=0} Y_i(t) \right)$$

203
204 where B is the number of trees in the forest, $n_{b,1}$ and $n_{b,0}$ are the number of treated and
205 control observations, respectively, in tree b for which x falls in the terminal node.
206

207 Results

208

209 In this territory-wide cohort study of 2526 patients (median age: 52.5 years old [SD:
210 10.9]; 57.34 % males) with T2DM treated with GLP1a between 1st January 2015 and 31st
211 December 2020 in Hong Kong, patients were followed up until 31st December 2020 or until

212 their deaths (**Figure 1**). This study included 1068 patients (77.91%) who used SGLT2Is, and
213 558 patients (22.09%) who did not use SGLT2I (Table 1). Before matching, fewer SGLT2 users
214 had hypertension and were on less medications (metformin, anticoagulants, ACEI/ARB,
215 diuretics).

216 After the propensity score matching, the baseline characteristics and the time-
217 weighted lipid and glucose profiles of the two groups were well-balanced (**Table 1**). The
218 SGLT2I and non-SGLT2I users were comparable after matching with the nearest neighbour
219 search strategy with a calliper of 0.1, and there was no violation of the proportional hazard
220 assumption (**Supplementary Figure 1**). In the matched cohort, 43 patients developed
221 myocardial infarction (2.56%), 26 patients developed atrial fibrillation (1,55%), 59 patients
222 developed heart failure (3.52%), and 30 patients developed stroke/TIA (1.79%). The
223 characteristics of patients by drug use before and after propensity score matching are shown
224 in **Table 1**.

225

226 *Association between SGLT2I and non-SGLT2I users and cardiovascular outcomes*

227 In the matched cohort, 21 SGLT2I users and 22 non-SGLT2I users developed
228 myocardial infarction. After a total follow-up of 9176.7 person-year, the incidence of
229 myocardial infarction was lower amongst SGLT2I users (Incidence rate [IR] per 1000 person-
230 year: 3.4; 95% CI: 2.1-5.2) compared to non-SGLT2I users (IR per 1000 person-year: 7.3; 95%
231 CI: 4.6-11.0) (**Table 2**). Non-SGLT2I users had a higher risk of myocardial infarction after
232 adjustment (HR: 2.91; 95% CI: 1.30-6.50, p=0.0092) compared to SGLT2I users regardless of
233 the demographics, comorbidities, medication profile, renal function, glycaemic tests, number
234 of hospitalisations, and the duration of T2DM (**Table 2**). This was substantiated by the

235 cumulative incidence curves (**Figure 2**). The E-value for the myocardial infarction was 5.27,
236 which suggested that the unobserved confounding variable with at least a 5.26-fold stronger
237 association with myocardial infarction would be needed to explain the significant HR, but
238 weaker confounding could not do so (**Supplementary Figure 4**).

239 Besides, 13 SGLT2I users and 13 non-SGLT2I users developed atrial fibrillation. After a
240 total follow-up of 9220.2 person-year, the incidence of atrial fibrillation was lower amongst
241 SGLT2I users (IR per 1000 person-year: 2.1; 95% CI: 1.1-3.6) compared to non-SGLT2I users
242 (IR per 1000 person-year: 4.3; 95% CI: 2.3-7.3) (**Table 2**). Non-SGLT2I users did not had a
243 significant risk of atrial fibrillation after adjustment (HR: 1.52; 95% CI: 0.65-3.53, p=0.3335)
244 compared to SGLT2I users after adjustments (**Table 2**). This was substantiated by the
245 cumulative incidence curves (**Figure 2**).

246 Furthermore, 31 SGLT2I users and 28 non-SGLT2I users developed heart failure. After
247 a total follow-up of 9146.1 person-year, the incidence of heart failure was lower amongst
248 SGLT2I users (IR per 1000 person-year: 5.1; 95% CI: 3.4-7.2) compared to non-SGLT2I users
249 (IR per 1000 person-year: 9.3; 95% CI: 6.2-13.4) (**Table 2**). Non-SGLT2I users had a significant
250 high risk of heart failure after adjustment (HR: 2.49; 95% CI: 1.22-5.08, p=0.0121) compared
251 to SGLT2I users after adjustments (**Table 2**). This was substantiated by the cumulative
252 incidence curves (**Figure 2**). The E-value for the heart failure was 4.42.

253 Finally, 14 SGLT2I users and 16 non-SGLT2I users developed stroke/TIA. After a total
254 follow-up of 9213.8 person-year, the incidence of stroke/TIA was lower amongst SGLT2I users
255 (IR per 1000 person-year: 2.3; 95% CI: 1.2-3.8) compared to non-SGLT2I users (IR per 1000
256 person-year: 5.2; 95% CI: 3.05-8.5) (**Table 2**). Compared to SGLT2I users, non-SGLT2I users did
257 not have significantly different risk of stroke/TIA after adjustment (HR: 1.72; 95% CI: 0.70-

258 4.24, $p=0.2361$) (**Table 2**). This was substantiated by the cumulative incidence curves (**Figure**
259 **2**).

260

261 *Association between SGLT2I and non-users and all-cause mortality*

262 Overall, 15 SGLT2I users and 13 non-SGLT2I users passed away. After a follow-up of
263 9313.1 person-year, the incidence of all-cause mortality was lower amongst SGLT2I users (IR:
264 2.4; 95% CI: 1.4-4.0) compared to non-users (IR: 4.2; 95% CI: 2.2-7.2) (**Table 2**). Non-SGLT2I
265 users were not associated with higher risks of all-cause mortality after adjustment (HR: 1.66;
266 95% CI: 0.61-4.52, $p=0.3184$) compared to non-SGLT2I users after adjustments (**Table 2**). This
267 was substantiated by the cumulative incidence curves (**Figure 2**).

268

269 *Subgroup analysis*

270 The results of the subgroup analysis for effects of SGLT2I and non-SGLT2I users on the
271 cardiovascular outcomes are shown in **Figure 3**. The result demonstrated that non-SGLT2I
272 users was associated with higher risks of myocardial infarction amongst patients without prior
273 ischaemic heart disease. Besides, non-SGLT2I users were associated with higher risks of atrial
274 fibrillation amongst female, with hypertension, hyperlipidaemia, on insulin, and on dipeptidyl
275 peptidase-4 inhibitor (DPP4I). Furthermore, non-SGLT2I users were associated with higher
276 risks of heart failure amongst female, with hypertension, hyperlipidaemia, without prior
277 ischaemic heart disease, and on DPP4I. Lastly, non-SGLT2I users were associated with higher
278 risks of stroke/TIA amongst patients with hypertension, on insulin, and without DPP4I.

279 The marginal effects analysis demonstrated that SGLT2I use was associated with lower
280 risks of cardiovascular disease regardless of the time-weighted mean fasting glucose level
281 apart from atrial fibrillation. Meanwhile, the differences between SGLT2I users and non-
282 SGLT2I users on myocardial infarction, atrial fibrillation, and stroke/TIA became smaller as the
283 time-weighted mean of HbA1c level increases (**Supplementary Figure 2**).

284

285 *Sensitivity analysis*

286 Sensitivity analyses were performed to confirm the predictability of the models and
287 test the potential effects on bias. The results of the cause-specific hazard models, sub-
288 distribution hazard models and different propensity score approaches demonstrated that
289 different models did not change the point estimates for both the myocardial infarction (all
290 $p < 0.05$) (**Supplementary Table 2**). However, non-SGLT2I users were associated with a higher
291 risk of stroke/TIA compared to SGLT2I users. The association between SGLT2I users and non-
292 SGLT2I users on myocardial infarction and heart failure remained consistent after 1) excluding
293 patients with less than 3-month follow-up duration; 2) Exclude patients with CKD stage 4/5
294 (eGFR < 30), peritoneal dialysis or haemodialysis; 3) Excluding patients on financial aid; 4)
295 Exclude patients at the top or bottom 5% of propensity score matching (**Supplementary Table**
296 **3**).

297

298 *Machine learning causal inference analysis results*

299 The causal inference analysis was performed using Casual Survival Forests model. The
300 result demonstrated that SGLT2I demonstrated a higher estimated treatment effects (ETE)
301 against myocardial infarction (Averaged treatment effect [ATE]: 0.009; 95% CI: 0.005-0.012;

302 p=0.0012) and heart failure (ATE: 0.017; 95% CI: 0.012-0.023; p=0.0007), atrial fibrillation (ATE:
303 0.013; 95% CI: 0.009-0.018; p=0.0041), and stroke/TIA (ATE: 0.029; 95% CI: 0.022-0.032;
304 p<0.0001) (**Figure 4**). Here ATE is calculated by ETE of SGLT2I minus ETE of non-SGLT2I.
305 Furthermore, for the significant outcome myocardial infarction and heart failure in **Table 2**,
306 in the causal inference analysis, SGLT2I users demonstrated a greater estimated treatment
307 effects compared to non-SGLT2I users regardless of sex, hyperlipidaemia, hypertension, prior
308 ischaemic heart disease and insulin (**Supplementary Figure 3**).

309

310 *Falsification analysis*

311 Hip fracture was used as negative control outcome in the falsification analysis for the
312 comparison between SGLT2I and non-SGLT2I users (**Supplementary Table 4**). The results
313 demonstrated that non-SGLT2I users' risk of the hip fracture after adjustment (HR: 0.91; 95%
314 CI: 0.82-1.79, p=0.2311) were not significantly different from SGLT2I users.

315

316 **Discussion**

317 In this territory-wide cohort study, real-world data was used to compare the
318 association between SGLT2I usage and cardiovascular events amongst patients on GLP1a. The
319 results suggest that non-SGLT2I users demonstrated a significantly higher risk of myocardial
320 infarction and heart failure than SGLT2I users. Meanwhile, in the machine learning causal
321 inference analysis, the result demonstrated that amongst patients on GLP1a, SGLT2I had a
322 larger estimated treatment effect compared to non-SGLT2I users on cardiovascular outcomes.

323

324 *Comparison with previous studies*

325 A recent consensus report jointly issued by the American Diabetes Association (ADA)
326 and the European Association for the Study of Diabetes (EASD) has presented an extensive
327 guideline endorsing the utilization of SGLT2I, as well as GLP-1A for individuals with T2DM with
328 increased risk of cardiovascular disease and heart failure [22]. However, limited emphasis is
329 placed on delineating the potential cardioprotective benefits arising from a combined therapy
330 involving SGLT2I and GLP-1A.

331 In our study, we observed a notable increase in the risks of myocardial infarction and
332 heart failure amongst SGLT2I non-users compared to the SGLT2I users in patients on GLP1a.
333 The result remained consistent with a meta-analysis, which demonstrated that SGLT2I was
334 associated with lower risks of heart failure compared to non-SGLT2I users, which was
335 consistent with the clinical trial data [23]. Furthermore, in the meta-analysis including 97 trials,
336 SGLT2I was associated with a lower risk of myocardial infarction compared to control with an
337 HR of 0.86 [24]. SGLT2I has also demonstrated its protective effect against new-onset
338 myocardial infarction compared to DPP4I in a Hong Kong population-based study (HR: 0.81)
339 [25]. Yet, upon conversion, the equivalent HR for myocardial infarction amongst GLP1a
340 patients in this study was 0.48, much lower than the published figures. This might indirectly
341 imply that SGLT2I might exert greater protective effects amongst GLP1a patients.
342 Nevertheless, evidence suggests SGLT2I was not significantly associated with a lower risk of
343 acute coronary artery syndrome compared to non-SGLT2I users amongst elderly T2DM
344 patients [26] and patients with renal failure [27]. These differences could be explained by the
345 baseline atherosclerotic profile and the cardiorenal status; in this study, GLP1a alone, with or
346 without interacting with SGLT2I, might have already improved those factors [28, 29].

347 While our findings corroborate the existing literature concerning the association
348 between SGLT2I and a reduction in MACE, the underlying mechanistic pathways remain an

349 imperative area for ongoing investigation. It was previously suggested that a combination of
350 both SGLT2I and GLP1a could demonstrate superior glycaemic control with a greater
351 reduction in the HbA1c compared to using either alone [30]. The associations between the
352 use of SGLT2I and atrial fibrillation or stroke/TIA were not statistically significant. However,
353 this could be explained by the relatively low case numbers for those outcomes.

354 The promise of the SGLT2I/GLP1a combination therapy in T2DM patients is evident in
355 a post-hoc analysis derived from the EXSCEL trial; the result demonstrated that the
356 combination of SGLT2I and GLP-1A yielded a decreased risk of MACE (HR: 0.68) compared to
357 GLP-1A alone (HR: 0.85) [31]. Meanwhile, a US-based real-world cohort study showcased an
358 overall risk reduction effect resulting from the addition of SGLT2I to the GLP1A regimen on
359 MACE, but not for stroke [8]. This raised the possibility that SGLT2I might play a slightly lesser
360 role in preventing stroke [32, 33].

361 The ML causal inference analysis was designed to distinguish between causal
362 relationships and spurious correlations and would provide actionable information regarding
363 the treatment [34]. The estimated treatment difference was defined as the difference in the
364 predicted outcome for patients treated with that drug versus their outcome if they had not
365 been treated [34]. This approach has been used to compare the effects of antidiabetic drugs
366 amongst poorly controlled diabetes patients [35]. In our ML causal inference analysis, the
367 result demonstrated that SGLT2I demonstrated higher estimated treatment effects against
368 all 4 outcomes we investigated in this study. Besides, the causal ML method demonstrated
369 that differences in outcomes with and without the usage of SGLT2I remained consistent
370 across different subgroups, such that the effects of SGLT2I remained homogenous.

371

372 *Clinical implications*

373 As the prevalence of T2DM continues to rise in Asia [36], the expanding body of
374 literature concerning the potential cardioprotective benefits of antidiabetic drugs holds
375 significant relevance. Under the current reference framework in Hong Kong designates
376 metformin as the first-line therapy for T2DM [37, 38], with second-line antidiabetic drugs to
377 be added when the glucose levels are insufficiently controlled. In our study, the number
378 needed to treat for myocardial infarction and heart failure was 49 and 45, respectively
379 **(Supplementary Table 5)**. While affirming the favourable cardioprotective effects of SGLT2I
380 supplementation for GLP1a-treated T2DM patients, this study also underscores the necessity
381 for further trials and in-depth investigations to ascertain whether these outcomes are solely
382 attributable to SGLT2I or result from synergistic effect with GLP1a.

383

384 *Limitations*

385 However, there are some limitations to the study. The observational nature of this
386 study suggests that data on certain variables, namely alcohol use, BMI, smoking, family
387 history and physical activity, could not be obtained and analysed. In compensation for this,
388 laboratory results and comorbidities related to the cardiovascular outcomes were used to
389 infer possible risk factors. To corroborate, drug cohorts were matched over a variety of
390 medications and diseases, adjusted with regression, and performed sensitivity analyses to
391 reduce the effect of the confounders.

392 The observational design of this study inherently lacks randomization, introducing
393 potential biases. While propensity score matching was employed to mitigate confounding,
394 the possibility of unidentified residual confounding remains. Therefore, a falsification analysis
395 was conducted, and the result did not falsify the association observed. Additionally, the
396 observational nature of the study precludes the establishment of definitive causal

397 relationships. Yet, the incorporation of the machine learning approaches arguably holds
398 promise in unravelling causal relationships, further accentuating the associations between
399 SGLT2I usage and cardiovascular outcomes.

400

401 **Conclusion**

402 This population-based cohort study suggested that non-SGLT2I users were associated
403 with higher risks of myocardial infarction and heart failure compared to SGLT2I users amongst
404 T2DM patients currently on GLP1a. These findings were substantiated by the machine
405 learning causal inference analysis.

406

407

408 **Funding source**

409 The authors received no funding for the research, authorship, and/or publication of this
410 article.

411

412 **Ethical approval statement**

413 This study was approved by the Institutional Review Board of the University of Hong
414 Kong/Hospital Authority Hong Kong West Cluster (HKU/HA HKWC IRB) (UW-20-250), and New
415 Territories East Cluster-Chinese University of Hong Kong (NTEC-UCHK) Clinical Research
416 Ethics Committee (2018.309, 2018.643) and complied with the Declaration of Helsinki.

417

418 **Conflicts of Interest**

419 None.

420

421 **Availability of data and materials**

422 Data are not available, as the data custodians (the Hospital Authority and the Department of
423 Health of Hong Kong SAR) have not given permission for sharing due to patient confidentiality
424 and privacy concerns. Local academic institutions, government departments, or
425 nongovernmental organizations may apply for the access to data through the Hospital
426 Authority's data sharing portal (<https://www3.ha.org.hk/data>).

427

428 **Acknowledgements**

429 None.

430

431 **Guarantor Statement**

432 All authors approved the final version of the manuscript. GT is the guarantor of this work and,
433 as such, had full access to all the data in the study and takes responsibility for the integrity of
434 the data and the accuracy of the data analysis.

435

436 **Author contributions**

437 Data analysis: ZYL, OHIC, JZ

438 Data review: ZYL, OHIC, CTC, GT, JZ

439 Data acquisition: OHIC, ZMWN, GT

440 Data interpretation: OHIC, CTC, JSKC, GT, JZ

441 Critical revision of manuscript: ZYL, RNCC, TTZ, TL, LL, JSKC, GT, JZ

442 Supervision: TTZ, TL, BMYC, GT, JZ

443 Manuscript writing: ZYL, OHIC, CTC, JSKC, JZ

444 Manuscript revision: ZYL, OHIC, CTC, JSKC, JZ

References

1. Raghavan S, Vassy JL, Ho YL, Song RJ, Gagnon DR, Cho K, Wilson PW, Phillips LS: **Diabetes mellitus–related all-cause and cardiovascular mortality in a national cohort of adults**. *Journal of the American Heart Association* 2019, **8**(4):e011295.
2. Association AD: **American Diabetes Association Standards of medical care in diabetes–2017**. *Diabetes care* 2017, **40**(Suppl. 1):S1.
3. Group ESD: **2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD**. *European Heart Journal* 2020, **41**(2):255-323.
4. Kyriakos G, Quiles-Sanchez LV, Garmpi A, Farmaki P, Kyre K, Savvanis S, Antoniou VK, Memi E: **SGLT2 Inhibitors and Cardiovascular Outcomes: Do they Differ or there is a Class Effect? New Insights from the EMPA-REG OUTCOME trial and the CVD-REAL Study**. *Curr Cardiol Rev* 2020, **16**(4):258-265.
5. Giugliano D, Scappaticcio L, Longo M, Caruso P, Maiorino MI, Bellastella G, Ceriello A, Chiodini P, Esposito K: **GLP-1 receptor agonists and cardiorenal outcomes in type 2 diabetes: an updated meta-analysis of eight CVOTs**. *Cardiovascular Diabetology* 2021, **20**(1):189.
6. Lin TK, Lee MC, Cheng YH, Ma T, Chen MC, Yang TY, Jong GP: **The association between SGLT2 inhibitors and new-onset acute coronary syndrome in the elderly: a population-based longitudinal cohort study**. *Diabetol Metab Syndr* 2023, **15**(1):170.
7. Neves JS, Borges-Canha M, Vasques-Nóvoa F, Green JB, Leiter LA, Granger CB, Carvalho D, Leite-Moreira A, Hernandez AF, Del Prato S: **GLP-1 receptor agonist therapy with and without SGLT2 inhibitors in patients with Type 2 diabetes**. *Journal of the American College of Cardiology* 2023, **82**(6):517-525.
8. Dave CV, Kim SC, Goldfine AB, Glynn RJ, Tong A, Paterno E: **Risk of cardiovascular outcomes in patients with type 2 diabetes after addition of SGLT2 inhibitors versus sulfonylureas to baseline GLP-1RA therapy**. *Circulation* 2021, **143**(8):770-779.
9. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondou N, Shaw W, Law G, Desai M, Matthews DR: **Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes**. *N Engl J Med* 2017, **377**(7):644-657.
10. Venkatasubramaniam A, Mateen BA, Shields BM, Hattersley AT, Jones AG, Vollmer SJ, Dennis JM: **Comparison of causal forest and regression-based approaches to evaluate treatment effect heterogeneity: an application for type 2 diabetes precision medicine**. *BMC Med Inform Decis Mak* 2023, **23**(1):110.
11. Zhou J, Zhang G, Chang C, Chou OHI, Lee S, Leung KSK, Wong WT, Liu T, Wai AKC, Cheng SH *et al*: **Metformin versus sulphonylureas for new onset atrial fibrillation and stroke in type 2 diabetes mellitus: a population-based study**. *Acta Diabetol* 2022, **59**(5):697-709.
12. Chou OHI, Zhou J, V Mui J, Satti DI, Chung CT, Tai Loy Lee T, Lee S, Dee EC, Ng K, Cheung BM *et al*: **Lower risks of new-onset acute pancreatitis and pancreatic cancer in sodium glucose cotransporter 2 (SGLT2) inhibitors compared to dipeptidyl peptidase-4 (DPP4) inhibitors: A propensity score-matched study with competing risk analysis**. *Diabetes Epidemiology and Management* 2023, **9**:100115.
13. Charlson ME, Carrozzino D, Guidi J, Patierno C: **Charlson Comorbidity Index: A Critical Review of Clinimetric Properties**. *Psychother Psychosom* 2022, **91**(1):8-35.

14. Soliman AR, Fathy A, Khashab S, Shaheen N: **Comparison of abbreviated modification of diet in renal disease formula (aMDRD) and the Cockcroft-Gault adjusted for body surface (aCG) equations in stable renal transplant patients and living kidney donors.** *Ren Fail* 2013, **35**(1):94-97.
15. Wu M-Z, Chandramouli C, Wong P-F, Chan Y-H, Li H-L, Yu S-Y, Tse Y-K, Ren Q-W, Yu S-Y, Tse H-F *et al*: **Risk of sepsis and pneumonia in patients initiated on SGLT2 inhibitors and DPP-4 inhibitors.** *Diabetes & Metabolism* 2022, **48**(6):101367.
16. Austin PC: **An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.** *Multivariate Behav Res* 2011, **46**(3):399-424.
17. Austin PC, Stuart EA: **Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies.** *Stat Med* 2015, **34**(28):3661-3679.
18. Avagyan V, Vansteelandt S: **Stable inverse probability weighting estimation for longitudinal studies.** *Scandinavian Journal of Statistics* 2021, **48**(3):1046-1067.
19. van Geloven N, Swanson SA, Ramspek CL, Luijken K, van Diepen M, Morris TP, Groenwold RHH, van Houwelingen HC, Putter H, le Cessie S: **Prediction meets causal inference: the role of treatment in clinical prediction models.** *Eur J Epidemiol* 2020, **35**(7):619-630.
20. Muriel A, Hernández-Marrero D, Abaira V: **Longitudinal observational studies and causality.** *Nefrología (English Edition)* 2012, **32**(1):12-14.
21. Wager S, Athey S: **Estimation and Inference of Heterogeneous Treatment Effects using Random Forests.** *Journal of the American Statistical Association* 2018, **113**(523):1228-1242.
22. Davies MJ, Aroda VR, Collins BS, Gabbay RA, Green J, Maruthur NM, Rosas SE, Del Prato S, Mathieu C, Mingrone G *et al*: **Management of hyperglycaemia in type 2 diabetes, 2022. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD).** *Diabetologia* 2022, **65**(12):1925-1966.
23. Ryan PB, Buse JB, Schuemie MJ, DeFalco F, Yuan Z, Stang PE, Berlin JA, Rosenthal N: **Comparative effectiveness of canagliflozin, SGLT2 inhibitors and non-SGLT2 inhibitors on the risk of hospitalization for heart failure and amputation in patients with type 2 diabetes mellitus: A real-world meta-analysis of 4 observational databases (OBSERVE-4D).** *Diabetes Obes Metab* 2018, **20**(11):2585-2597.
24. Zheng SL, Roddick AJ, Aghar-Jaffar R, Shun-Shin MJ, Francis D, Oliver N, Meeran K: **Association Between Use of Sodium-Glucose Cotransporter 2 Inhibitors, Glucagon-like Peptide 1 Agonists, and Dipeptidyl Peptidase 4 Inhibitors With All-Cause Mortality in Patients With Type 2 Diabetes: A Systematic Review and Meta-analysis.** *Jama* 2018, **319**(15):1580-1591.
25. Zhou J, Lee S, Leung KSK, Wai AKC, Liu T, Liu Y, Chang D, Wong WT, Wong ICK, Cheung BM Y *et al*: **Incident heart failure and myocardial infarction in sodium-glucose cotransporter-2 vs. dipeptidyl peptidase-4 inhibitor users.** *ESC Heart Failure* 2022, **9**(2):1388-1399.
26. Lin T-K, Lee M-C, Cheng Y-H, Ma T, Chen M-C, Yang T-Y, Jong G-P: **The association between SGLT2 inhibitors and new-onset acute coronary syndrome in the elderly: a population-based longitudinal cohort study.** *Diabetology & Metabolic Syndrome* 2023, **15**(1):170.

27. Birkeland KI, Bodegard J, Banerjee A, Kim DJ, Norhammar A, Eriksson JW, Thuresson M, Okami S, Ha KH, Kossack N *et al*: **Lower cardiorenal risk with sodium-glucose cotransporter-2 inhibitors versus dipeptidyl peptidase-4 inhibitors in patients with type 2 diabetes without cardiovascular and renal diseases: A large multinational observational study.** *Diabetes Obes Metab* 2021, **23**(1):75-85.
28. Caruso I, Cignarelli A, Sorice GP, Natalicchio A, Perrini S, Laviola L, Giorgino F: **Cardiovascular and Renal Effectiveness of GLP-1 Receptor Agonists vs. Other Glucose-Lowering Drugs in Type 2 Diabetes: A Systematic Review and Meta-Analysis of Real-World Studies.** *Metabolites* 2022, **12**(2).
29. Marx N, Husain M, Lehrke M, Verma S, Sattar N: **GLP-1 Receptor Agonists for the Reduction of Atherosclerotic Cardiovascular Risk in Patients With Type 2 Diabetes.** *Circulation* 2022, **146**(24):1882-1894.
30. Frías JP, Guja C, Hardy E, Ahmed A, Dong F, Öhman P, Jabbour SA: **Exenatide once weekly plus dapagliflozin once daily versus exenatide or dapagliflozin alone in patients with type 2 diabetes inadequately controlled with metformin monotherapy (DURATION-8): a 28 week, multicentre, double-blind, phase 3, randomised controlled trial.** *Lancet Diabetes Endocrinol* 2016, **4**(12):1004-1016.
31. Clegg LE, Penland RC, Bachina S, Boulton DW, Thuresson M, Heerspink HJL, Gustavson S, Sjöström CD, Ruggles JA, Hernandez AF *et al*: **Effects of exenatide and open-label SGLT2 inhibitor treatment, given in parallel or sequentially, on mortality and cardiovascular and renal outcomes in type 2 diabetes: insights from the EXSCEL trial.** *Cardiovasc Diabetol* 2019, **18**(1):138.
32. Sattar N, Lee MMY, Kristensen SL, Branch KRH, Del Prato S, Khurmi NS, Lam CSP, Lopes RD, McMurray JJV, Pratley RE *et al*: **Cardiovascular, mortality, and kidney outcomes with GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis of randomised trials.** *Lancet Diabetes Endocrinol* 2021, **9**(10):653-662.
33. Strain WD, Frenkel O, James MA, Leiter LA, Rasmussen S, Rothwell PM, Sejersten Ripa M, Truelsen TC, Husain M: **Effects of Semaglutide on Stroke Subtypes in Type 2 Diabetes: Post Hoc Analysis of the Randomized SUSTAIN 6 and PIONEER 6.** *Stroke* 2022, **53**(9):2749-2757.
34. Crown WH: **Real-World Evidence, Causal Inference, and Machine Learning.** *Value Health* 2019, **22**(5):587-592.
35. Belthangady C, Giampanis S, Jankovic I, Stedden W, Alves P, Chong S, Knott C, Norgeot B: **Causal deep learning reveals the comparative effectiveness of antihyperglycemic treatments in poorly controlled diabetes.** *Nature Communications* 2022, **13**(1):6921.
36. Yuan H, Li X, Wan G, Sun L, Zhu X, Che F, Yang Z: **Type 2 diabetes epidemic in East Asia: a 35-year systematic trend analysis.** *Oncotarget* 2018, **9**(6):6718-6727.
37. **9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes-2020.** *Diabetes Care* 2020, **43**(Suppl 1):S98-s110.
38. **Hong Kong Reference Framework for Diabetes Care for Adults in Primary Care Settings.** In: *Reference Frameworks*. Health Bureau; 2021.

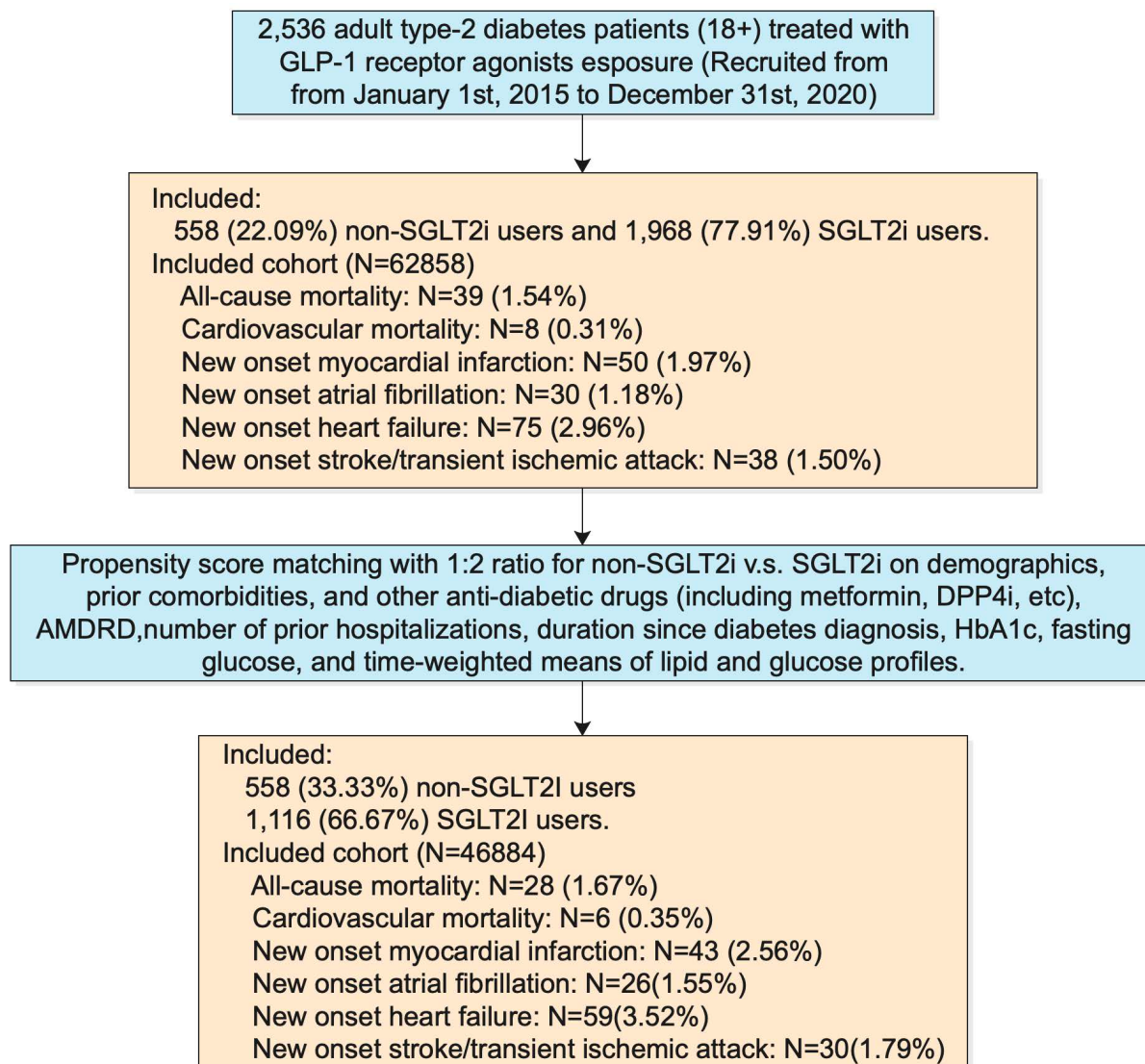
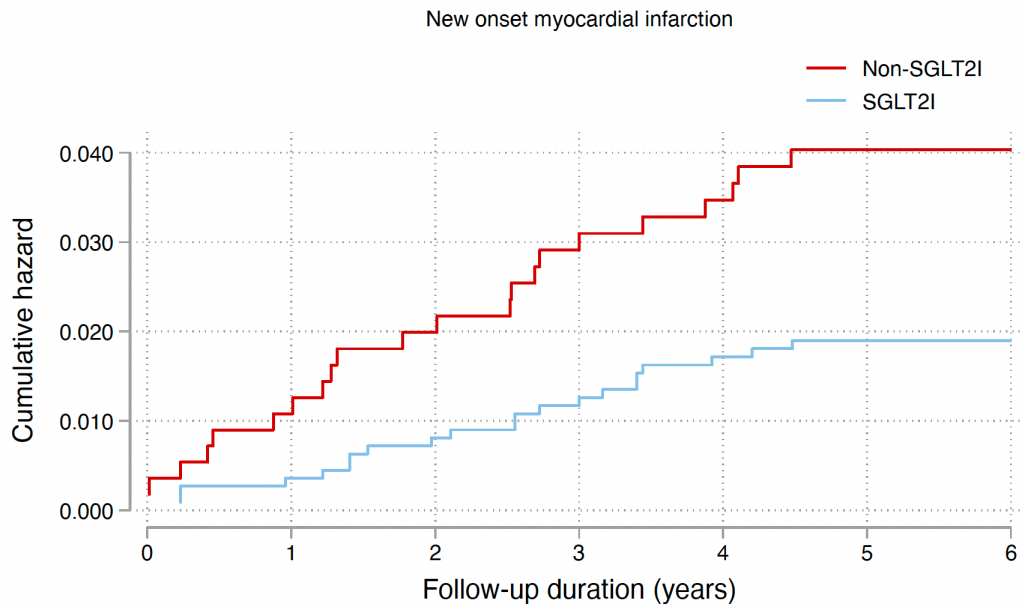


Figure 1. Procedures of data processing for the study cohort

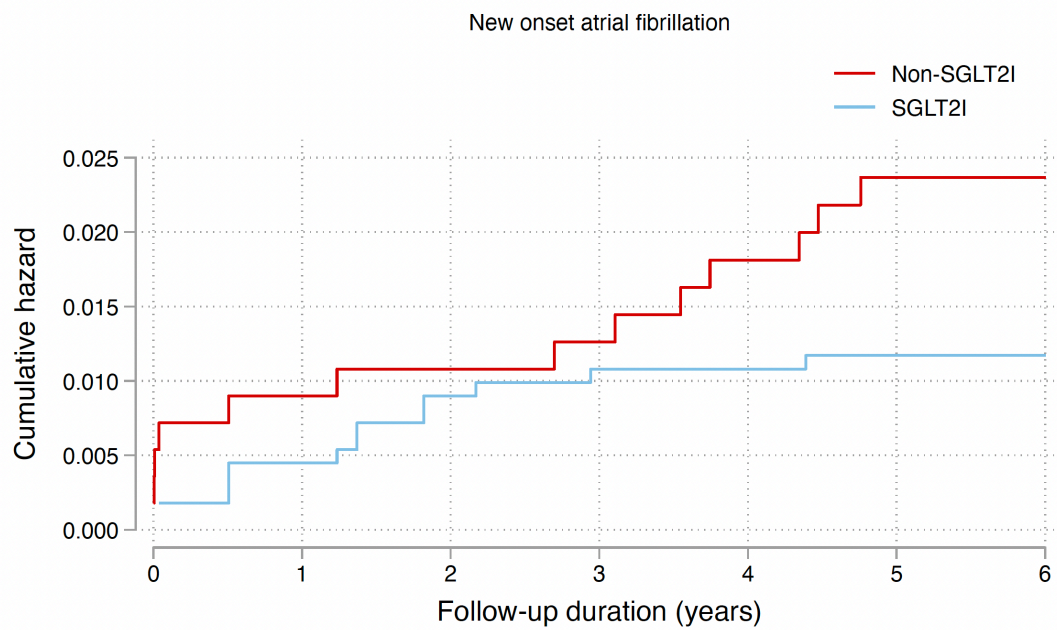
SGLT2i: Sodium-glucose cotransporter-2 inhibitors; MDRD: modification of diet in renal disease.



Number at risk

	0	1	2	3	4	5	6
Non-SGLT2I	558	552	546	539	534	527	3
SGLT2I	1116	1112	1107	1103	1095	1086	2

Cumulative incidence curve by drug use after propensity score matching



Number at risk

	0	1	2	3	4	5	6
Non-SGLT2I	558	553	551	547	543	535	4
SGLT2I	1116	1111	1106	1103	1098	1090	4

Cumulative incidence curve by drug use after propensity score matching

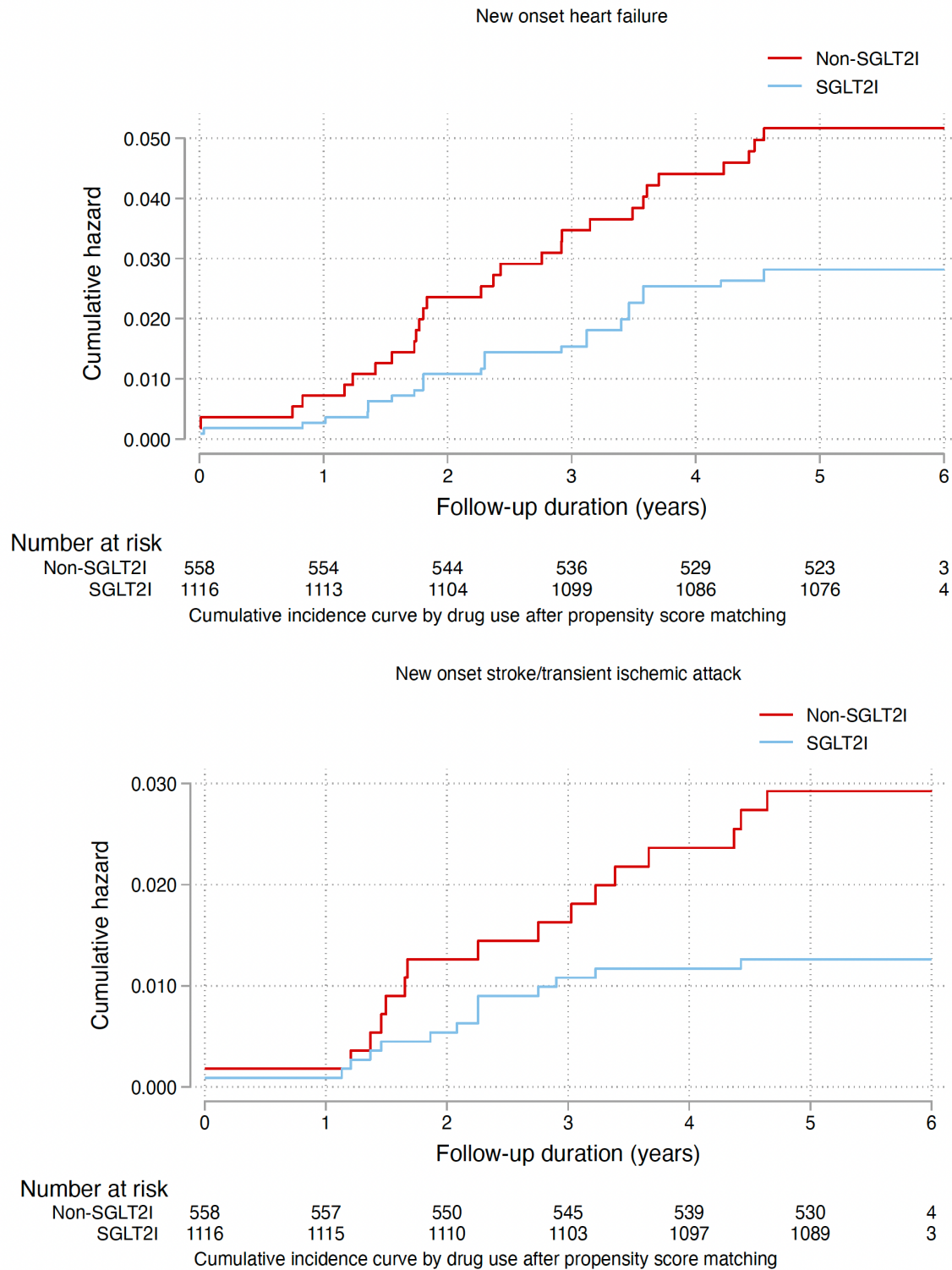
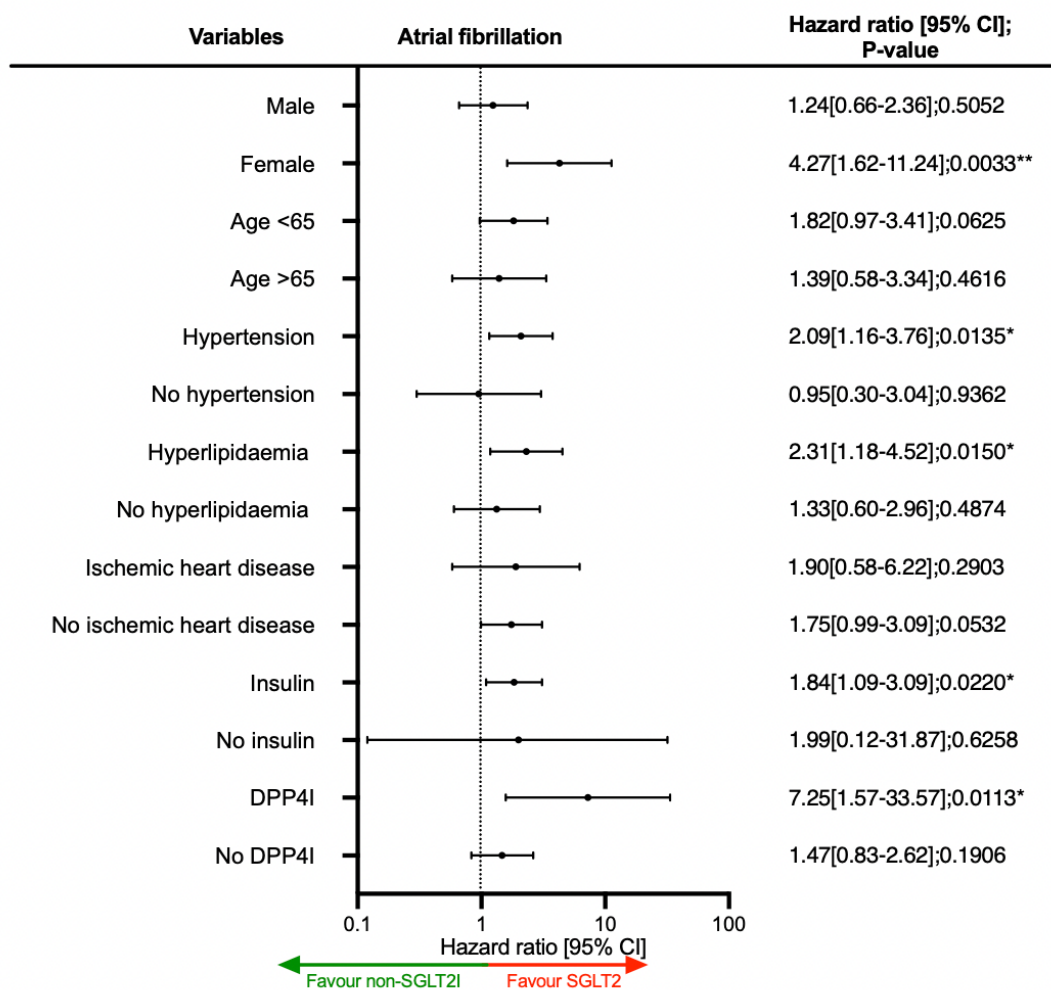
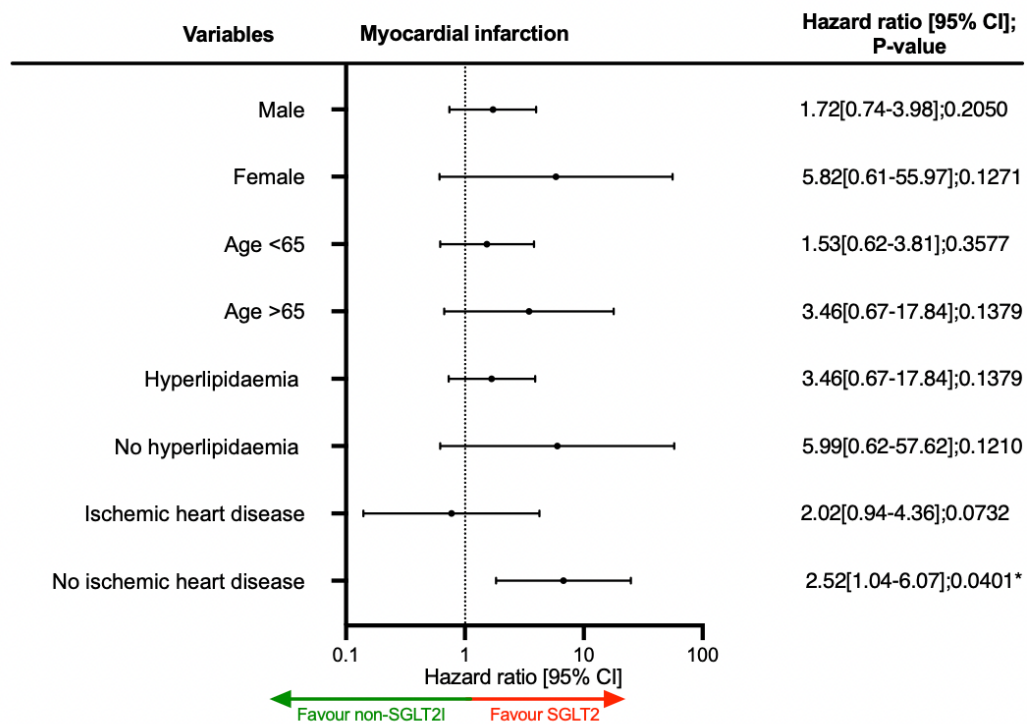


Figure 2. Cumulative incidence curves for new onset cardiovascular outcomes and all-cause mortality stratified by drug exposure effects of SGLT2I users and non-SGLT2I users after propensity score matching (1:2).

SGLT2I: Sodium-glucose cotransporter-2 inhibitors;



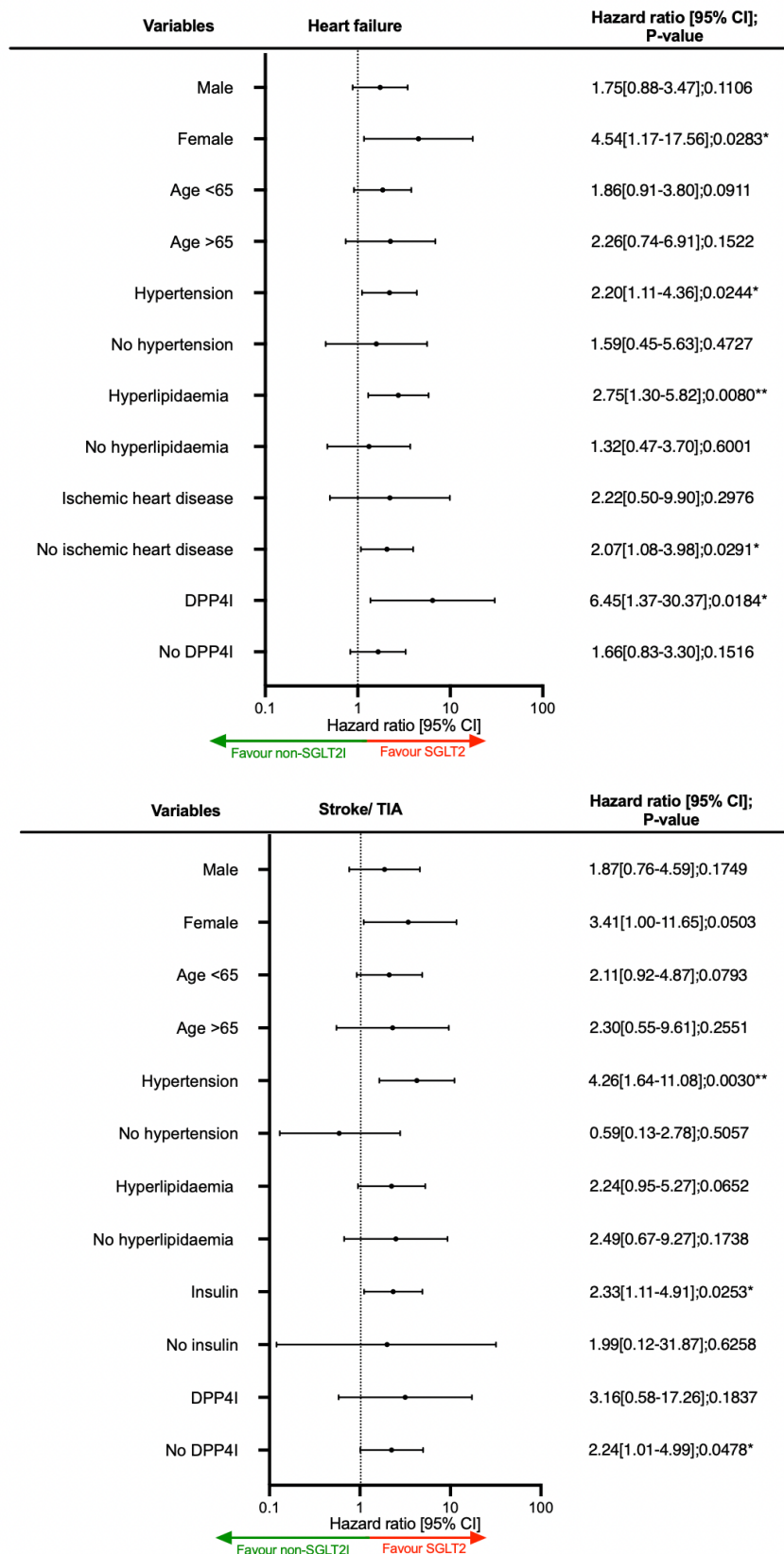
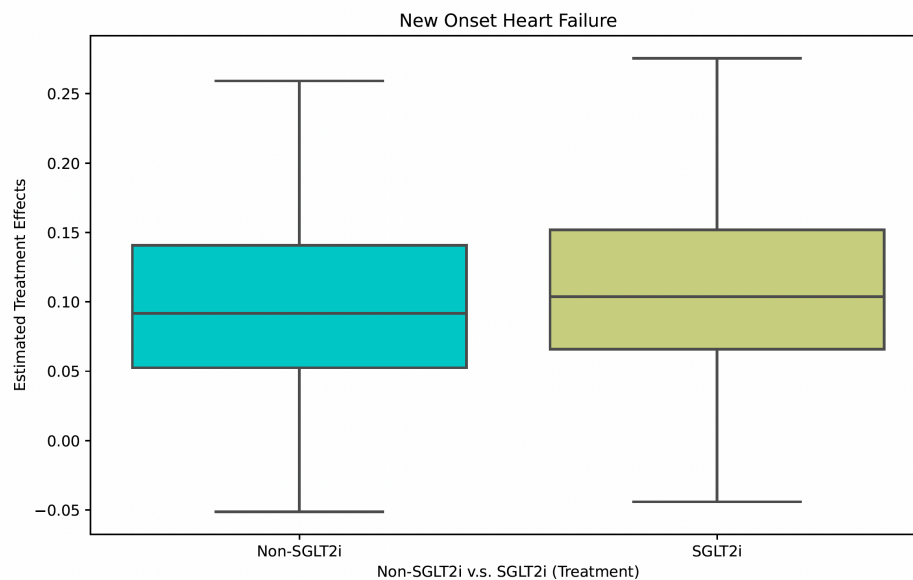
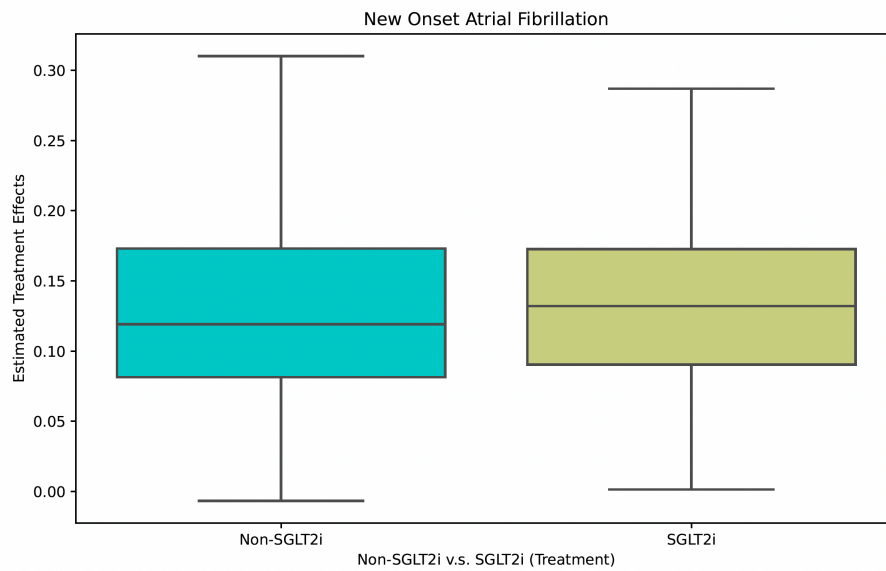
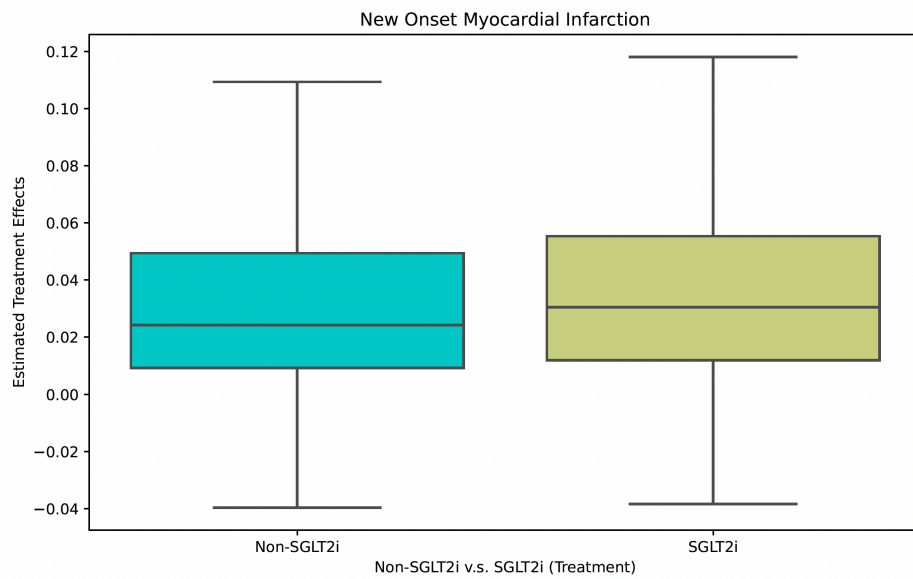


Figure 3. Forests plot of hazard ratios with 95% CI for SGLT2i users versus non-SGLT2i users on cardiovascular outcomes in the matched cohort.

SGLT2i: Sodium-glucose cotransporter-2 inhibitors; DPP4i: Dipeptidyl peptidase-4 inhibitors; TIA: Transient ischaemic attack



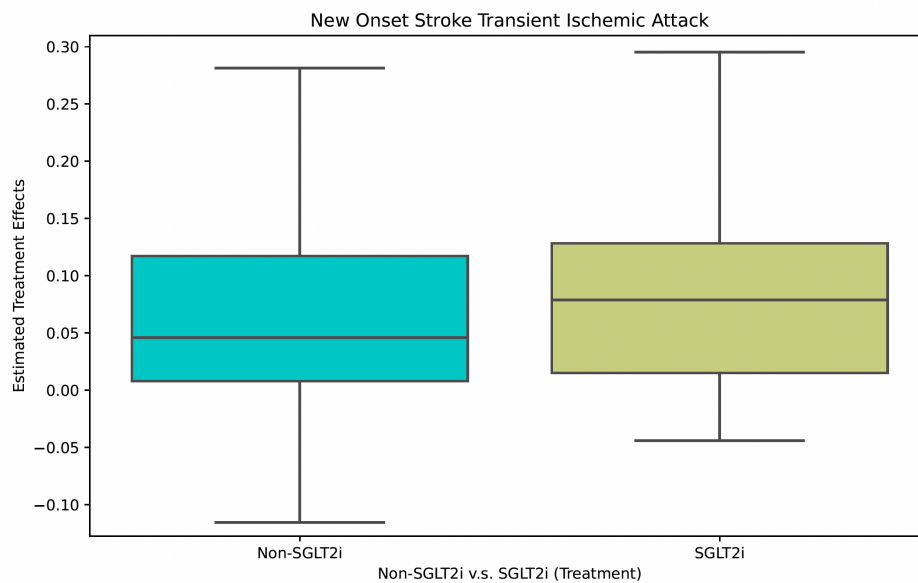


Figure 4. Treatment effect estimated by machine learning causal inference analysis.
SGLT2i: Sodium-glucose cotransporter-2 inhibitors

Table 1. Baseline and clinical characteristics of patients with SGLT2I users versus SGLT2I non-users amongst GLP1A users before and after propensity score matching (1:2).

* for SMD \geq 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.

Characteristics	Before matching				After matching			
	All (N=2526) Mean(SD);N or Count(%)	Non-SGLT2I users (N=558) Mean(SD);N or Count(%)	SGLT2I users (N=1968) Mean(SD);N or Count(%)	SMD	All (N=1674) Mean(SD);N or Count(%)	Non-SGLT2I users (N=558) Mean(SD);N or Count(%)	SGLT2I users (N=1116) Mean(SD);N or Count(%)	SMD
Demographics								
Male gender	1479(58.55%)	320(57.34%)	1159(58.89%)	0.03	978(58.42%)	320(57.34%)	658(58.96%)	0.03
Female gender	1047(41.44%)	238(42.65%)	809(41.10%)	0.03	696(41.57%)	238(42.65%)	458(41.03%)	0.03
Baseline age, years	52.5(10.9);n=2526	53.1(11.4);n=558	52.3(10.7);n=1968	0.07	52.5(10.7);n=1674	53.1(11.4);n=558	52.2(10.4);n=1116	0.09
Past comorbidities								
Charlson standard comorbidity index	1.1(1.1);n=26	1.2(1.1);n=58	1.1(1.1);n=68	0.05	1.1(1.0);n=74	1.2(1.1);n=58	1.1(1.0);n=16	0.1
Financial aid	42(1.66%)	6(1.07%)	36(1.82%)	0.06	18(1.07%)	6(1.07%)	12(1.07%)	<0.01
Duration from earliest diabetes mellitus diagnosis date to baseline date, day	8.1(4.7);n=26	8.07(4.84);n=558	8.05(4.65);n=1968	<0.01	8.2(4.7);n=74	8.1(4.8);n=58	8.2(4.7);n=16	0.03
Number of hospitalizations	1.3(0.9);n=26	1.27(0.81);n=558	1.3(0.96);n=968	0.04	1.2(0.8);n=74	1.3(0.8);n=58	1.2(0.7);n=16	0.07
Diabetic retinopathy	248(9.81%)	48(8.60%)	200(10.16%)	0.05	125(7.46%)	48(8.60%)	77(6.89%)	0.06
Diabetic nephropathy	28(1.10%)	9(1.61%)	19(0.96%)	0.06	25(1.49%)	9(1.61%)	16(1.43%)	0.01
Diabetic neuropathy	77(3.04%)	19(3.40%)	58(2.94%)	0.03	51(3.04%)	19(3.40%)	32(2.86%)	0.03
Hyperlipidaemia	1636(64.76%)	341(61.11%)	1295(65.80%)	0.1	1027(61.35%)	341(61.11%)	686(61.46%)	0.01
Hypertension	1351(53.48%)	360(64.51%)	991(50.35%)	0.29*	1004(59.97%)	360(64.51%)	644(57.70%)	0.14

Ischemic heart disease	236(9.34%)	51(9.13%)	185(9.40%)	0.01	130(7.76%)	51(9.13%)	79(7.07%)	0.08
Liver diseases	52(2.05%)	14(2.50%)	38(1.93%)	0.04	37(2.21%)	14(2.50%)	23(2.06%)	0.03
Renal diseases	221(8.74%)	42(7.52%)	179(9.09%)	0.06	108(6.45%)	42(7.52%)	66(5.91%)	0.06
Prior acute myocardial infarction	59(2.33%)	13(2.32%)	46(2.33%)	<0.01	34(2.03%)	13(2.32%)	21(1.88%)	0.03
Prior heart failure	69(2.73%)	17(3.04%)	52(2.64%)	0.02	48(2.86%)	17(3.04%)	31(2.77%)	0.02
Prior stroke/transient ischemic attack	45(1.78%)	7(1.25%)	38(1.93%)	0.05	21(1.25%)	7(1.25%)	14(1.25%)	<0.01
Prior atrial fibrillation	30(1.18%)	7(1.25%)	23(1.16%)	0.01	21(1.25%)	7(1.25%)	14(1.25%)	<0.01
Cancer	37(1.46%)	7(1.25%)	30(1.52%)	0.02	21(1.25%)	7(1.25%)	14(1.25%)	<0.01
Medications								
Number of anti-diabetic drug classes	4.4(0.6);n=26	4.41(0.59);n=558	4.41(0.63);n=1968	0.01	4.4(0.6);n=74	4.41(0.59);n=558	4.44(0.58);n=1116	0.05
Metformin	2298(90.97%)	543(97.31%)	1755(89.17%)	0.33*	1623(96.95%)	543(97.31%)	1080(96.77%)	0.03
Sulphonylurea	1495(59.18%)	354(63.44%)	1141(57.97%)	0.11	1069(63.85%)	354(63.44%)	715(64.06%)	0.01
Insulin	2040(80.76%)	448(80.28%)	1592(80.89%)	0.02	1345(80.34%)	448(80.28%)	897(80.37%)	<0.01
Acarbose	153(6.05%)	30(5.37%)	123(6.25%)	0.04	76(4.54%)	30(5.37%)	46(4.12%)	0.06
Thiozolidinedone	1193(47.22%)	240(43.01%)	953(48.42%)	0.11	746(44.56%)	240(43.01%)	506(45.34%)	0.05
DPP4i	850(33.65%)	190(34.05%)	660(33.53%)	0.01	489(29.21%)	190(34.05%)	299(26.79%)	0.16
Anticoagulants	1166(46.15%)	318(56.98%)	848(43.08%)	0.28*	868(51.85%)	318(56.98%)	550(49.28%)	0.15
Antiplatelets	511(20.22%)	129(23.11%)	382(19.41%)	0.09	346(20.66%)	129(23.11%)	217(19.44%)	0.09
Lipid-lowering drugs	1173(46.43%)	290(51.97%)	883(44.86%)	0.14	814(48.62%)	290(51.97%)	524(46.95%)	0.1
Statins and fibrates	1694(67.06%)	370(66.30%)	1324(67.27%)	0.02	1135(67.80%)	370(66.30%)	765(68.54%)	0.05
Nitrates	271(10.72%)	57(10.21%)	214(10.87%)	0.02	157(9.37%)	57(10.21%)	100(8.96%)	0.04
ACEI/ARB	1163(46.04%)	303(54.30%)	860(43.69%)	0.21*	834(49.82%)	303(54.30%)	531(47.58%)	0.13

Diuretics for hypertension	981(38.83%)	157(28.13%)	824(41.86%)	0.29*	501(29.92%)	157(28.13%)	344(30.82%)	0.06
Diuretics for heart failure	264(10.45%)	52(9.31%)	212(10.77%)	0.05	138(8.24%)	52(9.31%)	86(7.70%)	0.06
Beta-blockers	639(25.29%)	110(19.71%)	529(26.88%)	0.17	343(20.48%)	110(19.71%)	233(20.87%)	0.03
Calcium channel blockers (dihydropyridine)	874(34.60%)	185(33.15%)	689(35.01%)	0.04	541(32.31%)	185(33.15%)	356(31.89%)	0.03
Calcium channel blocker (non-dihydropyridine)	17(0.67%)	3(0.53%)	14(0.71%)	0.02	9(0.53%)	3(0.53%)	6(0.53%)	<0.01
Calculated biomarkers								
Abbreviated MDRD, mL/min/1.73 m ²	91.7(26.5);n=2232	87.9(28.0);n=465	92.7(26.0);n=1767	0.18	93.3(27.3);n=1364	90.7(30.5);n=465	94.6(25.3);n=899	0.14
Lipid and glucose profiles								
Time weighted mean of triglyceride, mmol/L	2.1(0.9);n=51	2.09(0.99);n=542	2.06(0.94);n=1909	0.03	2.0(0.9);n=1603	2.1(1.0);n=42	2.0(0.9);n=61	0.11
Time weighted mean of low-density lipoprotein, mmol/L	2.2(0.4);n=42	2.15(0.43);n=540	2.15(0.42);n=1902	0.01	2.1(0.4);n=1600	2.2(0.4);n=40	2.1(0.4);n=60	0.05
Time weighted mean of high-density lipoprotein, mmol/L	1.1(0.2);n=51	1.14(0.17);n=542	1.15(0.16);n=1909	0.01	1.1(0.2);n=1603	1.14(0.17);n=542	1.14(0.15);n=1061	0.01
Time weighted mean of total cholesterol, mmol/L	4.3(0.5);n=51	4.28(0.46);n=542	4.26(0.48);n=1909	0.03	4.2(0.4);n=1603	4.3(0.5);n=42	4.2(0.4);n=61	0.13
Time weighted	7.6(1.4);n=52	7.4(1.5);n=48	7.6(1.4);n=1704	0.13	7.5(1.3);n=37	7.4(1.5);n=48	7.6(1.3);n=889	0.1

mean of HbA1C, % Time weighted mean of fasting glucose, mmol/L	8.5(2.2);n=22 90	8.7(2.4);n=4 86	8.5(2.1);n=18 04	0.1	8.6(2.0);n=14 84	8.7(2.4);n=4 86	8.5(1.8);n=99 8	0.0 8
---	---------------------	--------------------	---------------------	-----	---------------------	--------------------	--------------------	----------

Table 2. Incidence rate (IR) per 1000 person-year and multivariate Cox regression models of new onset cardiovascular disease and all-cause mortality in the cohort amongst GLP1A users before and after 1:2 propensity score matching.

* for $p \leq 0.05$, ** for $p \leq 0.01$, *** for $p \leq 0.001$; CI: confidence interval; SGLT2i: sodium glucose cotransporter-2 inhibitor

Adjusted for significant demographics, past comorbidities, duration of diabetes mellitus, and number of prior hospitalizations, number of anti-diabetic drug classes, other anti-diabetic medications, abbreviated MDRD, HbA1c, fasting glucose.

Myocardial infarction	Number of patients	Number of events	Total person-year	Incidence [95% CI] per 1000 person-time	Adjusted hazard ratio [95% CI]	P-value
SGLT2i users	1116	21	6153.1	3.4[2.1-5.2]	1 [Reference]	NA
Non-SGLT2i users	558	22	3023.6	7.3[4.6-11.0]	2.91[1.30-6.50]	0.0092**
Atrial fibrillation	Number of patients	Number of events	Total person-year	Incidence [95% CI] per 1000 person-time	Adjusted hazard ratio [95% CI]	P value
SGLT2i users	1116	13	6162.8	2.1[1.1-3.6]	1 [Reference]	NA
Non-SGLT2i users	558	13	3057.4	4.3[2.3-7.3]	1.52[0.65-3.53]	0.3335
Heart failure	Number of patients	Number of events	Total person-year	Incidence [95% CI] per 1000 person-time	Adjusted hazard ratio [95% CI]	P-value
SGLT2i users	1116	31	6131.3	5.1[3.4-7.2]	1 [Reference]	NA
Non-SGLT2i users	558	28	3014.8	9.3[6.2-13.4]	2.49[1.22-5.08]	0.0121*
Stroke/ transient ischemic attack	Number of patients	Number of events	Total person-year	Incidence [95% CI] per 1000 person-time	Adjusted hazard ratio [95% CI]	P-value
SGLT2i users	1116	14	6165.8	2.3[1.2-3.8]	1 [Reference]	NA
Non-SGLT2i users	558	16	3048	5.2[3.05-8.5]	1.72[0.70-4.24]	0.2361
All-cause mortality	Number of patients	Number of events	Total person-year	Incidence [95% CI] per 1000 person-time	Adjusted hazard ratio [95% CI]	P value
SGLT2i users	1116	15	6215.3	2.4[1.4-4.0]	1 [Reference]	NA
Non-SGLT2i users	558	13	3097.8	4.2[2.2-7.2]	1.66[0.61-4.52]	0.3184