16 February 2024

Final study report

Title: Association between exposure to GLP-1 receptor agonists and risk of suicide-related and self-harm-related events – including **POST-HOC ANALYSES**

Administrative details of the data analysis	
Substance(s)	GLP-1 receptor agonists
Condition/ADR(s)	Suicide-related and self-harm-related events (suicidal ideation, self-injurious ideation, self-injury/self-harm, suicide attempt, completed suicide)
Short title of topic	Suicide-related and self-harm-related events and GLP-1 receptor agonists
RWE study team	Andrei Barbulescu, María Clara Restrepo-Méndez and Robert Flynn
Reviewer	Daniel Leaver Morales

Table of Contents

1. Brief description of the study (for publication in HMA-EMA Catalogue of RWD studies)

A cohort study aimed at investigating a potentially increased risk of suicide- and self-harm- related events among patients treated with GLP-1 receptor agonists.

2. Rationale and background

Suicide-related events, including ideation, attempt, and completed suicide, are an important drug safety issue. Therapies involved as risk factors for suicidal outcomes include antidepressants (Barbui et al., 2009; Fergusson et al., 2005), anticonvulsants (Mula & Hesdorffer, 2011) interferon (Fragoso et al., 2010; Lucaciu & Dumitrascu, 2015) and hormonal contraceptives (Skovlund et al., 2018), among others. However, the strength of evidence for some of these associations remains uncertain.

Currently, there is an ongoing signal of suicidal ideation and self-injurious ideation after exposure to liraglutide and semaglutide. These substances are glucagon-like peptide-1 (GLP-1) receptor agonists which are known as incretin mimetics because they act by increasing insulin release from the pancreas in response to food. Liraglutide is authorized for weight management in adult and adolescent patients aged ≥12 years and for the treatment of adults, adolescents, and children >10 years of age with insufficiently controlled type 2 diabetes mellitus as an adjunct to diet and exercise.[1] Semaglutide is authorized for the treatment of adult patients with insufficiently controlled type 2 diabetes mellitus (T2DM) as an adjunct to diet and exercise and for weight loss and weight maintenance in adults who have a body mass index (BMI) ≥30 kg/m² (obesity) or a BMI between 27 kg/m² and <30 kg/m² (overweight).[2] Suicidal behaviour is not currently listed as a side effect in the EU product information for any GLP-1 receptor agonists. However, the FDA prescribing information for liraglutide includes a warning regarding suicidal behaviour and ideation.[3]

Weight control failure has been related to suicidal ideation, particularly among overweight and obese adolescents.(Ju et al., 2016) Obesity and diabetes are associated with multiple comorbidities (Guh et al., 2009; Zheng et al., 2018), premature mortality (Flegal et al., 2013; Zheng et al., 2018), and impaired health-related quality of life (Garner et al., 2012; Zheng et al., 2018). Studies suggest that diabetic patients and obese individuals have a higher prevalence of neuropsychiatric disorders, such as depression, compared to the general population (Khaledi et al., 2019; Luppino et al., 2010; Roy & Lloyd, 2012; Scott et al., 2008; Zhao et al., 2009). It has been indicated that incretin-based therapies may have neuropsychiatric effects given the presence of glucagon-like peptide-1 (GLP-1) receptors in the central nervous system. (Gamble et al., 2018)

To support the evaluation of the signal, a real-world data study was proposed to assess whether there is an association between exposure to GLP-1 receptor agonists and increased risk of suicide-related and self-harm-related events among patients with T2DM when compared to patients who are prescribed alternative treatment (i.e., sodium-glucose cotransporter 2 (SGLT-2) inhibitors). Of note, it is of interest for the signal evaluation to investigate the effect of the entire class of GLP-1 receptor agonists.

[1] EMA. Victoza | European Medicines Agency (europa.eu)
[2] EMA. Ozempic | European Medicines Agency (europa.eu)
[3] FDA. SAXENDA (liraglutide).

3. Research question and objectives

Is the use of GLP-1 receptor agonists associated with an increased risk of suicide-related and self-harm-related events when compared to patients who are prescribed SGLT-2 inhibitors among T2DM patients?

4. Research methods

4.1. Study design

Comparative cohort design (active comparator, new user design).

4.2. Data sources

The following database was used: IQVIA™ Medical Research Data (IMRD) UK. A brief description of this database is provided in **Annex 1**.

4.3. Setting and study population

The population eligible for the study consisted of patients:

- registered with a GP-practice covered by IMRD,
- who initiated treatment with GLP-1 receptor agonists or SGLT-2 inhibitors (new users) during the study period,
- who had not used GLP-1 receptor agonists or SGLT-2 inhibitors before index-date (i.e., the baseline of the study, representing the date of index treatment initiation and the start of follow-up in the study),
- with at least one year of recorded medical history prior to index-date,
- with at least one diagnosis of T2DM recorded before index-date,
- with no recorded history of suicide-related or self-harm-related events prior to index-date,
- who have been treated with biguanides (e.g., metformin, which is considered first-line anti-diabetic treatment according to NICE guidelines (NICE, 2022)) before index-date

No exclusion was applied according to age.

4.4. Study period

The study period covered from the first use of SGLT-2 inhibitors in the database (i.e., **01 March 2013**) onwards to the most recent data available (**01 June 2023**). Although the first use of GLP-1 receptor agonists dates to August 2009 in the IMRD, we opted to include only the period in which both treatments were available, thus patients would have had the option to receive either treatment.

4.5. Variables

Exposure: Exposures of interest consisted of GLP-1 receptor agonists (**target group**) and SGLT-2 inhibitors (**comparator group**).

In the UK, NICE clinical guidelines (NICE, 2022) recommend use of SGLT-2 inhibitors after metformin and at least one other oral antidiabetic. GLP-1 receptor agonists are only recommended when triple therapy with metformin and two other oral antidiabetics was not effective, contraindicated or not tolerated. GLP-1 receptor agonists should replace one of the combination components, and it is indicated to T2DM adults:

- who have a BMI of ≥35 Kg/m^2 or other medical problems associated with obesity, or
- who have a BMI <35 Kg/m^2 for whom insulin therapy is not an option or weight loss would benefit other obesity related comorbidities.

New users of GLP-1 receptor agonists or SGLT-2 inhibitors were identified based on the date of first prescription in the database (**index-date**).

Exposures were identified through keyword searches of prescriptions recorded in the electronic health record. Detailed list of codes is provided in **Annex 2**.

Products containing GLP-1 receptor agonists or SGLT-2 inhibitors were combined under each corresponding treatment arm. The GLP-1 receptor agonists arm included the following substances (ordered by frequency of prescribing): liraglutide (alone and in combination with insulin), semaglutide, dulaglutide, exenatide and lixisenatide. The SGLT-2 inhibitor arm included the following substances (also ordered by frequency of prescribing): dapagliflozin (alone or in combination with metformin or saxagliptin), empagliflozin (alone or in combination with metformin or linagliptin), canagliflozin (alone or in combination with metformin) and ertugliflozin. **Supplementary Table 1** shows the frequency of individuals who initiated each drug under each study treatment arm.

Outcome: The primary outcome consisted of a composite endpoint which includes the first recorded occurrence of any of the following events:

> (a) suicide-related events which encompassed codes related to suicidal ideation, attempted suicide, intentional overdose/self-poisoning, completed suicide.
>
> (b) self-harm-related events which covered codes in which intention of self-harm is clearly specified (e.g., deliberate self-harm/self-injury, self-inflicted harm, self-injurious behaviour) to differentiate from unintentional/accidental harm.

The complete list of clinical terms and READ codes used to identify these events is provided in **Annex 2**.

Potential confounding factors:

Analyses accounted for the following baseline covariates measured before index-date:

- Sociodemographic: age, sex, and deprivation index.
- History of:
    - obesity
    - depression (including depressive episode in bipolar disorder and cyclothymia)
    - treatment for depression
    - anxiety

- treatment for anxiety
- schizophrenia/psychosis (including anti-psychotic treatment)
- dementia
- malignancy
- cardiovascular diseases (myocardial infarction, cerebrovascular disease, congestive heart failure)
- kidney disease
- smoking
- alcohol and substance abuse

- Prior anti-diabetic treatments before index-date other than biguanide (metformin)

Definitions and codes used to generate these covariates are described in **Annex 2**. The choice of covariates took into consideration known risk-factors for the outcome, which may potentially be associated with the choice of exposure.

4.6. Statistical analysis

4.6.1. Brief summary of the analysis method (for publication in HMA-EMA Catalogue of RWD studies)

In the main analysis, the estimated treatment effect was the comparison of the risk of suicide-related and self-harm-related events, over the maximum available follow-up (~10 years), between patients who initiated GLP-1 agonists and patients who initiated SGLT-2 inhibitors at baseline, regardless of subsequent changes in treatment (i.e., **intention-to-treat analysis** (**ITT**)).

4.6.2. Descriptive analysis

Descriptive analyses were performed to describe the study cohorts at baseline in terms of demographic characteristics, lifestyle factors, comorbidities, and history of treatment with anti-diabetic and psychiatric medication.

4.6.3. Main statistical analysis

Inverse probability of treatment weighting (IPTW)

Inverse probability of treatment weighting (IPTW) was used to render the assignment of study treatments independent of the baseline measured covariates, thus minimising the potential confounding effect of these covariates. Using IPTW, the average treatment effect (ATE) in the entire study population can be estimated, assuming that all important confounding variables have been accounted for. For each observation in the study population, the IPTW is the inverse of the probability of receiving the observed treatment conditional on all variables considered sufficient for confounding adjustment (i.e., the propensity score (PS) for patients treated with GLP-1 receptor agonists and 1-PS for patients treated with SGLT-2 inhibitors). In order to stabilize the weights (i.e., less extreme weights, that are closer to the mean weight of one), the numerator of one is replaced by the marginal probability of receiving the observed treatment in the study population (i.e., the proportion of

observations in the study population with the respective treatment).(Hernan & Robins, 2006) All analyses are then conducted in the re-weighted population without additional confounding adjustment on the assumption of no unmeasured confounding. To account for weighting the population (essentially multiplying observations by the weight coefficient) robust standard error estimators were used as recommended. (Hernán M & Robins J, 2020)

The distributions of baseline covariates before and after weighting were compared between the two treatment arms by calculating and plotting standardized mean differences (SMD), with a SMD of <0.1 used to determine appropriate covariate balance. For each variable, the SMD is the difference in mean (for continuous variables) or proportion (for binary variables) between the two treatment arms, divided by the squared mean variance of the variable in the two treatment arms.(Austin, 2011)

Intention-to-treat (ITT) analysis

The ITT approach chosen for the main analysis involves following patients from the date of study treatment initiation (index date) until the earliest of any of the following: first outcome event date, transfer out date from the general practice, date of death or the end of the study. Additionally, outcome events were attributed to the baseline treatment regardless of treatment change. Moreover, it is assumed that intercurrent events that may have occurred during follow-up are independent of the risk of the outcome (i.e., the risk of experiencing the outcome among individuals remaining in the analysis over the course of follow-up is representative of the risk among censored individuals).

Incidence rates (IRs)

IRs were calculated as the number of events occurring during follow-up divided by the total person-time in each treatment-arm. IRs are presented as number of events per 100 person-years.

Cumulative incidence (Incidence proportion)

Survival (i.e., the proportion of the patients included at baseline who have not yet experienced a suicide-related or a self-harm-related event) over the course of follow-up was estimated by treatment arm using the product-limit method.(Bland & Altman, 1998) The cumulative incidence was calculated as the complement of survival (i.e., the proportion of the patients included at baseline who have experienced a suicide-related or a self-harm-related event) at each follow-up time, and was presented as number of events per 100 patients.

Cox proportional hazards model

Hazard ratios of suicide-related and self-harm-related events associated with treatment of interest (GLP-1 receptor agonists) *versus* comparator (SGLT-2 inhibitors) were estimated using a Cox proportional hazards model.

More details on the procedures used for analysis can be found in **Annex 3**.

4.6.4. Sensitivity analysis

The following sensitivity analyses were performed to test the validity of the underlying assumptions and the robustness of the study findings:

- **Restricting to only suicide-related events:** The component part comprising only suicide-related events (a) of the primary outcome was assessed separately. The limited sample size of the component part comprising only self-harm-related events (b) precluded conducting a separate analysis for it.
- **Applying asymmetric PS trimming** (Stürmer trimming method): Patients with extreme PS, among whom unmeasured confounding may be more likely, were excluded from the study.

Observations with a PS lower than the 5th percentile of the PS distribution observed among patients treated with GLP-1 receptor agonists or higher than the 95th percentile of the PS distribution observed among patients treated with SGLT-2 inhibitors were excluded from the analysis.(Sturmer et al., 2010)

- **Applying "on-treatment" (OT) approach:** In this analysis patients were followed only as long as they continued the baseline treatment (i.e., censoring follow-up at baseline treatment discontinuation). Assuming that treatment discontinuation is independent of the risk of suicide-related and self-harm-related events, this analysis estimates the treatment effect had patients remained on the baseline treatment for the entire follow-up.

  The discontinuation of baseline treatment was defined following these steps:

  i. For each patient, all prescriptions corresponding to the index treatment class were extracted (e.g., for GLP-1 receptor agonist initiators, all GLP-1 receptor agonist prescriptions, for the index drug or another GLP-1 receptor agonist, were extracted).
     - The end of treatment was set to 180 days after the last prescription followed by the first gap longer than 180 days in the sequence of prescriptions.

  ii. For each patient, all prescriptions corresponding to the alternative treatment class were extracted (e.g., for GLP-1 receptor agonist initiators, all SGLT-2 inhibitor prescriptions were extracted).
     - The date of crossing-over to the alternative treatment was set to the date of the first prescription from the alternative treatment arm.

  Thus, we considered a patient "on-treatment" from index-date to the identified end of index treatment or to the date of crossing over to the alternative treatment arm. Thus, patients were censored at the earliest of the following events: end of index-treatment, crossing over to the alternative treatment arm, first outcome event date, transfer out date, date of death or at the end of the study period.

4.6.5. Sample size

The sample size was driven by the availability of individuals with exposures and outcomes within each database and no *a priori* sample size requirement was stipulated.

Analyses were done using SAS Enterprise Guide 7.1 software.

4.7. Quality control

The study was conducted according to the ENCePP code of conduct (European Medicines Agency 2018).

Standard operating procedures or internal process guidance were adhered to for the conduct of the study. These procedures included rules for secure and confidential data storage, quality-control procedures for all aspects of the study from protocol development to the reporting of the results.

All documents underwent at least one round a review by an experienced reviewer, while the results from the statistical analysis were reviewed by another experienced analyst.

The quality control of the data is the responsibility of the data holder.

5. Deviation from the protocol

- We could not include ethnicity as predictor variable for the propensity score because of the number of individuals with missing data (n=1548).
- We could not conduct the following sensitivity analyses because sample size constraints, i.e., the number of events were too low in these subsamples to conduct meaningful analyses:
    - Restricting the study population to a subsample of "low-risk" patients, i.e., individuals without history of psychiatric conditions (depression, anxiety and psychosis).
    - Assessing self-harm-related events only as a separate endpoint.
- We performed post hoc analyses described in Section 7 as a result of the discussion of the findings of the main analysis during the PRAC plenary held on 29 November 2023.

6. Results

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) have been removed prior to publication of this report. Additional cells have been redacted (events/patients typically being rounded up to the nearest 10) if needed in order to ensure that the aforementioned low cell counts cannot be re-identified. This may include both events/patients and follow-up times.

6.1. Descriptive results

Overall, 38,336 patients who were prescribed GLP-1 receptor agonists or SGLT-1 inhibitors from 1 March 2013 to 01 June 2023 were identified. After excluding patients who did not meet the selection criteria specified in Section 5.3, or who initiated both study treatments on the same day or had missing data for deprivation index (**Figure 1**), 6027 patients prescribed GLP-1 receptor agonists and 20855 patients prescribed SGLT-1 inhibitors were included in the study (**Figure 1**).

The GLP-1 receptor agonist and SGLT-2 inhibitor cohorts were followed for a median of 3.5 (IRQ 1.5-6.2) and 2.0 (IRQ 0.8-4.3) years, respectively. A total of 300 incident cases of suicide-related and self-harm-related events were identified during the study period. The clinical terms which most contributed to the composite endpoint definition were suicidal thoughts (28%), thoughts of suicide or self-harm (26%), suicidal ideation (15%), intentional drug overdose (5%), and thoughts of deliberate self-harm (4%) (**Supplementary Table 2**).

Table 1 shows the baseline characteristics for both treatment groups. Before IPTW, patients starting GLP-1 receptor agonists were more likely to be female, obese, had a higher prevalence of insulin, sulphonylureas or glitazone use history, and of psychiatric conditions (such as anxiety or depression) or renal disease, and were less likely to have suffered cardiovascular disease (such as congestive heart failure or cerebrovascular disease) compared with those starting SGLT-2 inhibitors.

The distribution of PS and inverse probability of treatment weights by treatment arm are shown in **Supplementary Figures 1** and **2**. The standardized mean differences between treatment arms for each baseline covariate before and after IPTW are shown in **Supplementary Figures 3a** and **3b**,

respectively. The treatment groups were well balanced across all covariates after IPTW, with all SMDs <0.1 (**Table 1, Supplementary Figure 3a** and **3b**).

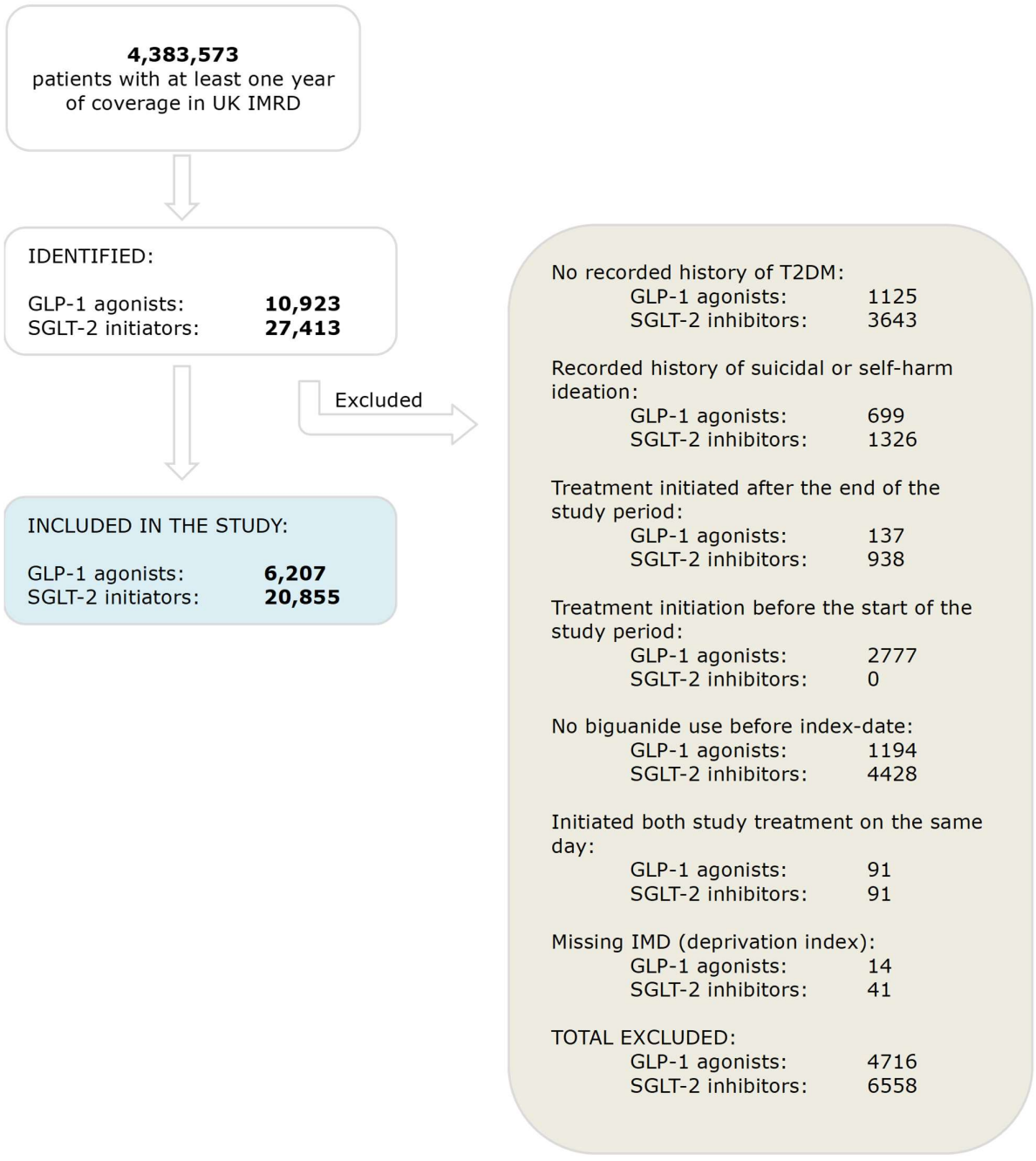


Figure 1. Flowchart for cohort selection.

Table 1. Baseline characteristics of the study population

	Before IPTW (unadjusted)		After IPTW (adjusted)	
	GLP-1 agonists	SGLT-2 inhibitors	GLP-1 agonists	SGLT-2 inhibitors
Number of individuals	**6207**	**20855**	**6207**	**20855**
Calendar year at index-data, median (min-max)	2019 (2013-2023)	2021 (2013-2023)	2019 (2013-2023)	2020 (2013-2023)
Age at index-date (years), median (min-max)	58.4 (10.2-96.6)	61.0 (16.3-101.1)	60.0 (10.2-96.6)	60.4 (16.3-101.1)
Years between first T2DM diag. and index-date, median (min-max)	7.9 (0.0-55.1)	8.1 (0.0-72.3)	8.1 (0.0-55.1)	8.1 (0.0-72.3)
Multiple deprivation index, median (min-max)	4.0 (1.0-10.0)	4.0 (1.0-10.0)	4.0 (1.0-10.0)	4.0 (1.0-10.0)
Male, n (%)	3274 (52.7)	12874 (61.7)	3513 (58.6)	12442 (59.5)
Insulin use history, n (%)	1604 (25.8)	2692 (12.9)	1029 (17.2)	3399 (16.2)
Sulphonylureas use history, n (%)	3709 (59.8)	10101 (48.4)	3203 (53.4)	10735 (51.3)
Glitazone use history, n (%)	1059 (17.1)	2390 (11.5)	826 (13.8)	2715 (13.0)
DPP-4 inhibitor use history, n (%)	3141 (50.6)	9839 (47.2)	3056 (51.0)	10084 (48.2)
Glinide use history, n (%)	72 (1.2)	201 (1.0)	74 (1.2)	216 (1.0)
Other antidiabetic drug use history, n (%)	163 (2.6)	281 (1.3)	108 (1.8)	351 (1.7)
Obesity history (within 2 y before index-date), n (%)	5472 (88.2)	12815 (61.4)	4151 (69.2)	14163 (67.7)
Obesity treatment history (within 2 y before index-date), n (%)	276 (4.4)	288 (1.4)	133 (2.2)	454 (2.2)
Anxiety history, n (%)	1503 (24.2)	3905 (18.7)	1229 (20.5)	4211 (20.1)
Anxiety treatment history, n (%)	1600 (25.8)	4841 (23.2)	1465 (24.4)	5013 (24.0)
Depression history, n (%)	2617 (42.2)	7370 (35.3)	2259 (37.7)	7759 (37.1)
Depression treatment history, n (%)	3687 (59.4)	10024 (48.1)	3131 (52.2)	10647 (50.9)
Schizophrenia history, n (%)	132 (2.1)	393 (1.9)	116 (1.9)	414 (2.0)
Dementia history, n (%)	36 (0.6)	166 (0.8)	57 (0.9)	157 (0.8)
Stress syndromes history, n (%)	253 (4.1)	669 (3.2)	216 (3.6)	718 (3.4)
Renal disease history, n (%)	1074 (17.3)	2842 (13.6)	876 (14.6)	3035 (14.5)
Myocardial infarction history (within 2 y before index-date), n (%)	113 (1.8)	444 (2.1)	134 (2.2)	431 (2.1)
Congestive heart failure history, n (%)	259 (4.2)	1378 (6.6)	329 (5.5)	1257 (6.0)
Cerebrovascular disease history, n (%)	245 (3.9)	980 (4.7)	287 (4.8)	957 (4.6)
Cancer history, n (%)	567 (9.1)	2053 (9.8)	599 (10.0)	2027 (9.7)

Substance abuse history, n (%)	278 (4.5)	1001 (4.8)	256 (4.3)	984 (4.7)
Current smoker, n (%)	1147 (18.5)	3817 (18.3)	1067 (17.8)	3833 (18.3)

IPTW: inverse probability of treatment weighting.

6.2. Main results

Intention-to-treat (ITT) analysis

Incidence rates (IRs)

Table 2 shows the unadjusted and IPTW adjusted incidence rates of the composite endpoint (incl. both suicide-related and self-harm-related events) by treatment arm. The unadjusted IR of the composite endpoint (incl. both suicide-related and self-harm-related events) was higher in the GLP-1 receptor agonist cohort than in the SGLT-2 inhibitor cohorts (IR: 0.50 [95%CI: 0.42–0.60] vs 0.30 [95%CI: 0.26-0.35] events per 100 person-years, respectively). However, this difference was attenuated after IPTW adjustment (IR: 0.36 [95%CI: 0.30–0.44] vs 0.33 [95%CI: 0.29–0.39]).

Cumulative incidence (Incidence proportion)

Figure 2 shows the unadjusted and IPTW adjusted cumulative incidence of the composite endpoint (incl. both suicide-related and self-harm-related events) by treatment arm. The unadjusted cumulative incidence curves of the composite endpoint diverged soon after treatment initiation and throughout the follow-up period (**Figure 2a**). At 8.8 years of the follow-up, around the time of the last outcome events observed in each treatment arm, the unadjusted cumulative incidence of the composite endpoint was 3.9% (95%CI: 3.1-4.8%) for the GLP-1 receptor agonist cohort vs 3.1% (95% CI: 1.8-4.4%) for the SGLT-2 inhibitor cohort. After the IPTW adjustment, the cumulative incidence curves overlapped throughout the follow-up period (**Figure 2b**).

Hazard ratios

Table 2 shows the unadjusted and IPTW adjusted hazard ratios of the composite endpoint (incl. both suicide-related and self-harm-related events). In the unadjusted analysis, the use of GLP-1 receptor agonists was associated with a 71% increased risk of suicide-related and self-harm-related events (HR: 1.71 [95%CI: 1.35-2.16]) compared with the use of SGLT-2 inhibitors. However, the HR was substantially attenuated after the IPTW adjustment (HR: 1.10 [95%CI: 0.86-1.41]).

Sensitivity analysis

Only suicide-related events: The composite endpoint consisted largely of suicide-related events, e.g., 115 events out of 125 unadjusted number of events in the GLP-1 receptor agonist cohort. When restricting the outcome definition to only suicide-related events, the unadjusted and IPTW adjusted incidence rates, hazard ratios (**Table 3**) and cumulative incidence curves (**Figure 3**) were consistent with those for the composite endpoint in the ITT analysis (unadjusted HR: 1.67 [95%CI: 1.30-2.13] and IPTW adjusted HR: 1.08 [0.83-1.39] for GLP-1 receptor agonists). The differences in incidence rate, and cumulative incidence between the two treatment cohorts as well as the hazard ratio were also attenuated after the IPTW adjustment.

Asymmetric propensity score trimming: When excluding patients with extreme PS from each treatment cohort, unadjusted and IPTW adjusted results were consistent with those for the main analysis, i.e., for the composite endpoint in the ITT analysis (**Supplementary Table 3**, and **Supplementary Figure 4**).

On-treatment analysis: The results based on the OT analytical approach and using the composite outcome were consistent with those for the main ITT analysis of the composite endpoint (**Table 2** and **Figure 4**), the ITT analysis of suicide-related events only (**Table 3** and **Figure 5**), and with the results from the trimmed study population analysis (**Supplementary Table 3** and **Supplementary Figure 5**), respectively. It is worth noting that the follow-up time was shorter under OT for both treatment arms

compared to the ITT analysis. The GLP-1 receptor agonist and SGLT-2 inhibitor cohorts were followed for a median of 1.3 (IRQ: 0.6-2.3) and 1.0 (IRQ: 0.5-2.5) years, respectively under OT (vs median follow-up: 3.5 [IRQ: 1.5-6.2] and 2.0 [IRQ: 0.8-4.3] years, respectively, under ITT). The percentage of patients censored by discontinuing baseline treatment or crossing over was 57% for the GLP-1 receptor agonist cohort and 39% for the SGLT-1 inhibitor cohort.

Table 2. Unadjusted and adjusted incidence rates per 100 person- years and hazard ratios of **composite endpoint** (incl. both suicide-related and self-harm-related events) by treatment arm and analytical approach

	ITT									OT							
Treatment arm	**N**	**Follow-up (person-years)**	**n events**	**IR**	**95% CI**		**HR**	**95% CI**		**Follow-up (person-years)**	**n events**	**IR**	**95% CI**		**HR**	**95% CI**	
Unadjusted analysis																	
GLP-1 receptor agonist	6207	24812.3	125	**0.50**	0.42	0.60	**1.71**	1.35	2.16	12445.4	60	**0.48**	0.37	0.62	**1.65**	1.20	2.26
SGLT-2 inhibitors	20855	57547.4	175	**0.30**	0.26	0.35	**1.00**	[Reference]		37364.5	110	**0.29**	0.24	0.35	**1.00**	[Reference]	
IPTW adjusted analysis																	
GLP-1 receptor agonist	6103	22487.4	81.8	**0.36**	0.30	0.44	**1.10**	0.86	1.41	11672.0	41.1	**0.35**	0.27	0.46	**1.11**	0.80	1.56
SGLT-2 inhibitors	20923	59805.2	200.0	**0.33**	0.29	0.39	**1.00**	[Reference]		38156.4	120.4	**0.32**	0.26	0.38	**1.00**	[Reference]	

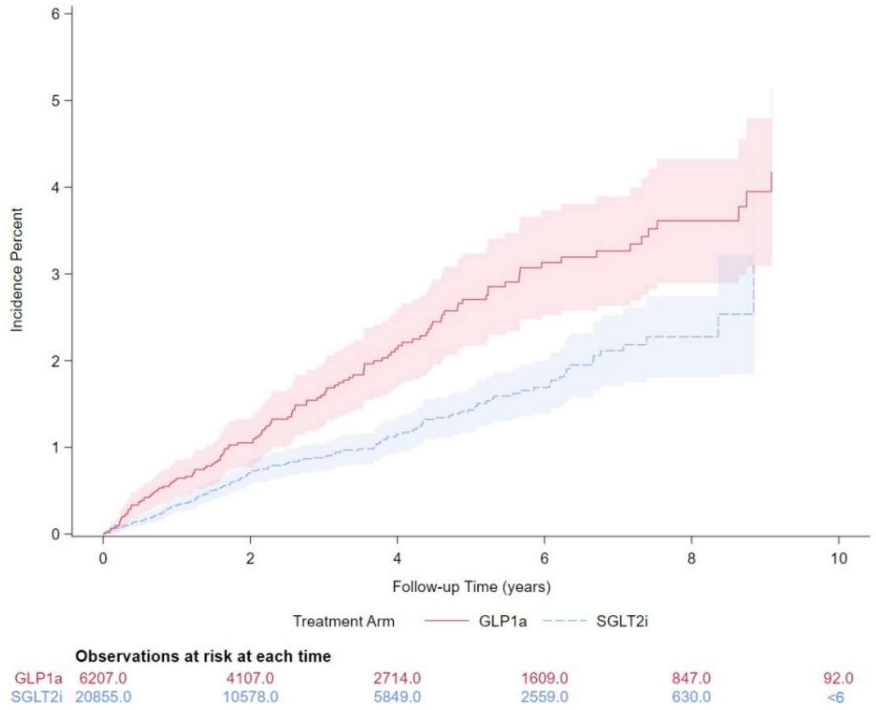
ITT: intention-to-treat analytical approach, OT: On-treatment analytical approach; IR: incidence rate; 95% CI: 95% confidence interval; HR: Hazard ratio; IPTW: inverse probability of treatment weighting.

Table 3. Unadjusted and adjusted incidence rates per 100 person- years and hazard ratios of **only suicide-related events** by treatment arm and analytical approach

	ITT									OT							
Treatment arm	**N**	**Follow-up (person-years)**	**n events**	**IR**	**95% CI**		**HR**	**95% CI**		**Follow-up (person-years)**	**n events**	**IR**	**95% CI**		**HR**	**95% CI**	
Unadjusted analysis																	
GLP-1 receptor agonist	6207	24847.0	115	**0.46**	0.39	0.56	**1.67**	1.30	2.13	12450.8	55	**0.44**	0.34	0.58	**1.59**	1.14	2.21
SGLT-2 inhibitors	20855	57580.3	164	**0.28**	0.24	0.33	**1.00**	[Reference]		37380.9	104	**0.28**	0.23	0.34	**1.00**	[Reference]	
IPTW adjusted analysis																	
GLP-1 receptor agonist	6103	22506.8	75.9	**0.34**	0.28	0.41	**1.08**	0.83	1.39	11675.6	38.4	**0.33**	0.25	0.44	**1.09**	0.77	1.55
SGLT-2 inhibitors	20923	59843.5	187.6	**0.31**	0.27	0.37	**1.00**	[Reference]		38174.2	114.5	**0.30**	0.25	0.37	**1.00**	[Reference]	

ITT: intention-to-treat analytical approach, OT: On-treatment analytical approach; IR: incidence rate; 95% CI: 95% confidence interval; HR: Hazard ratio; IPTW: inverse probability of treatment weighting.

(a) Unadjusted



(b) IPTW Adjusted

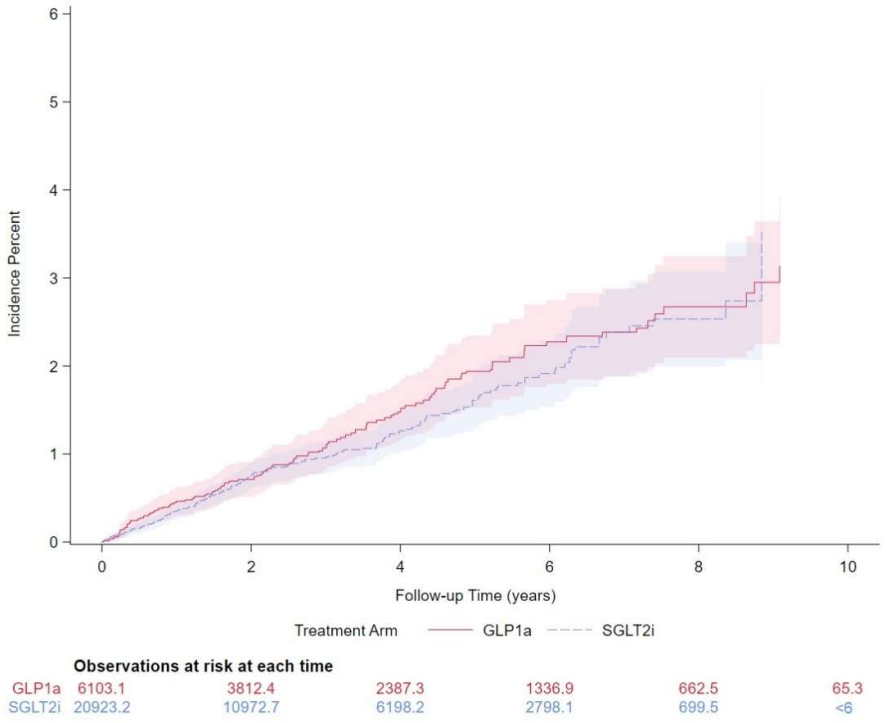
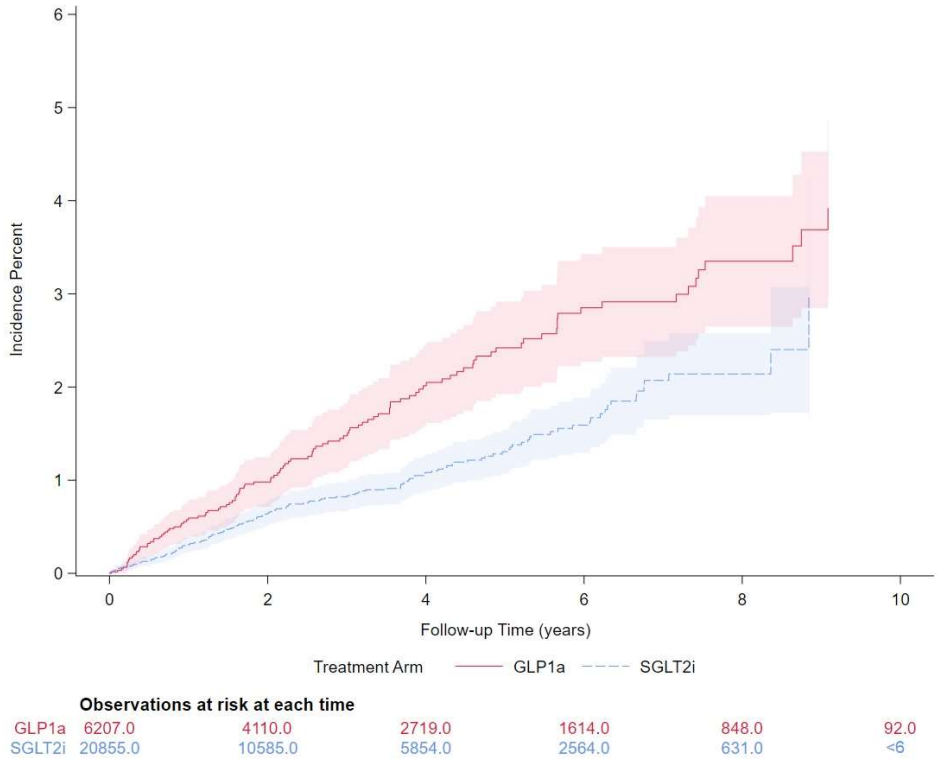


Figure 2. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **ITT** analytical approach

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

(a) Unadjusted



(b) IPTW Adjusted

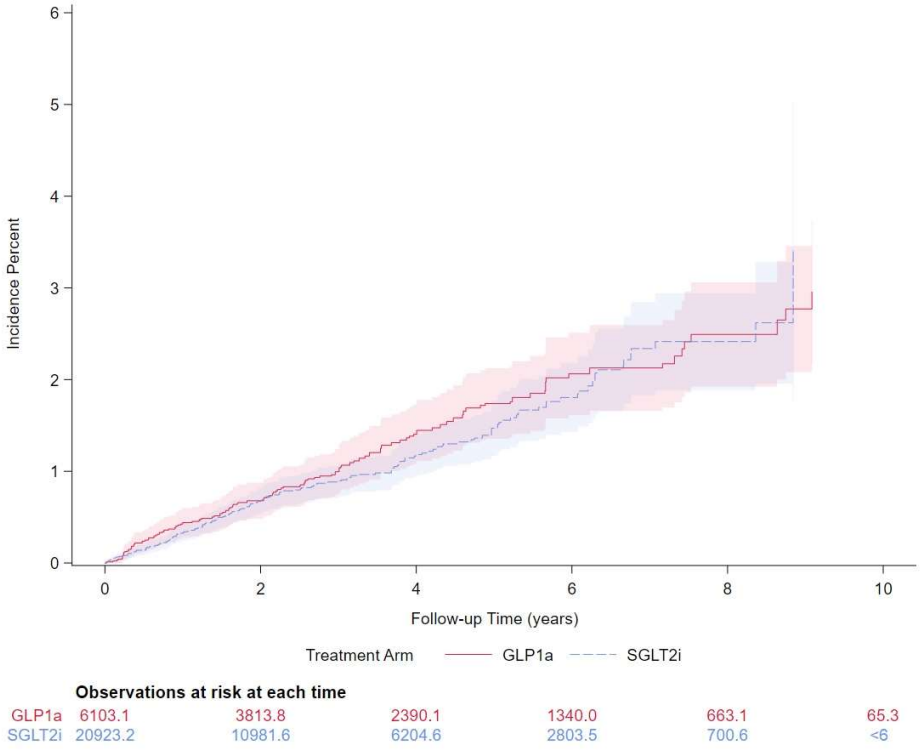
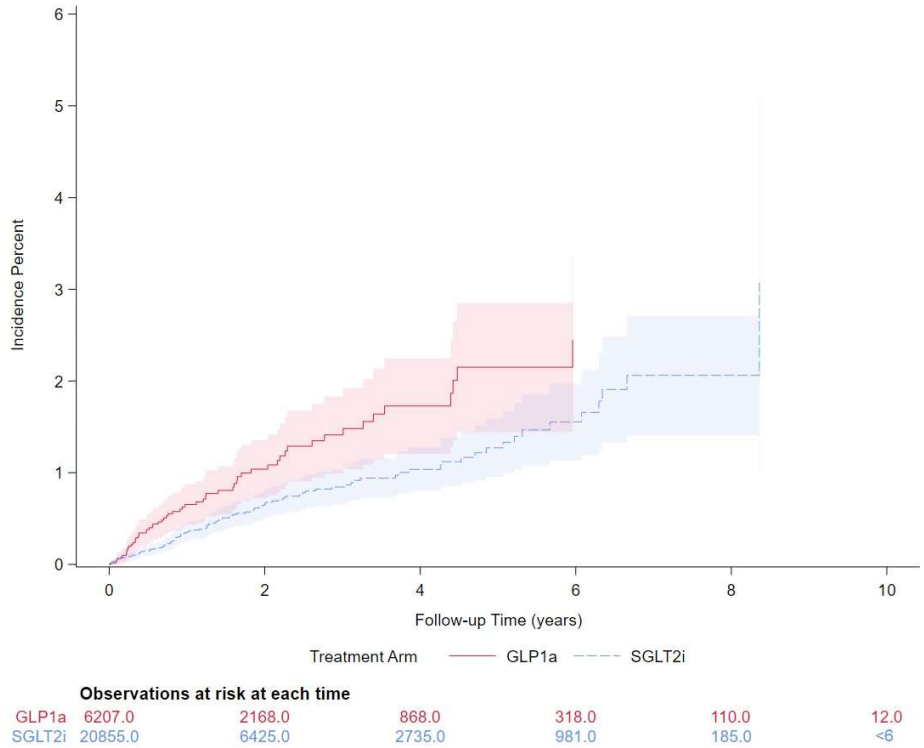


Figure 3. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **only suicide-related events** based on **ITT** analytical approach

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

(a) Unadjusted



(b) IPTW Adjusted

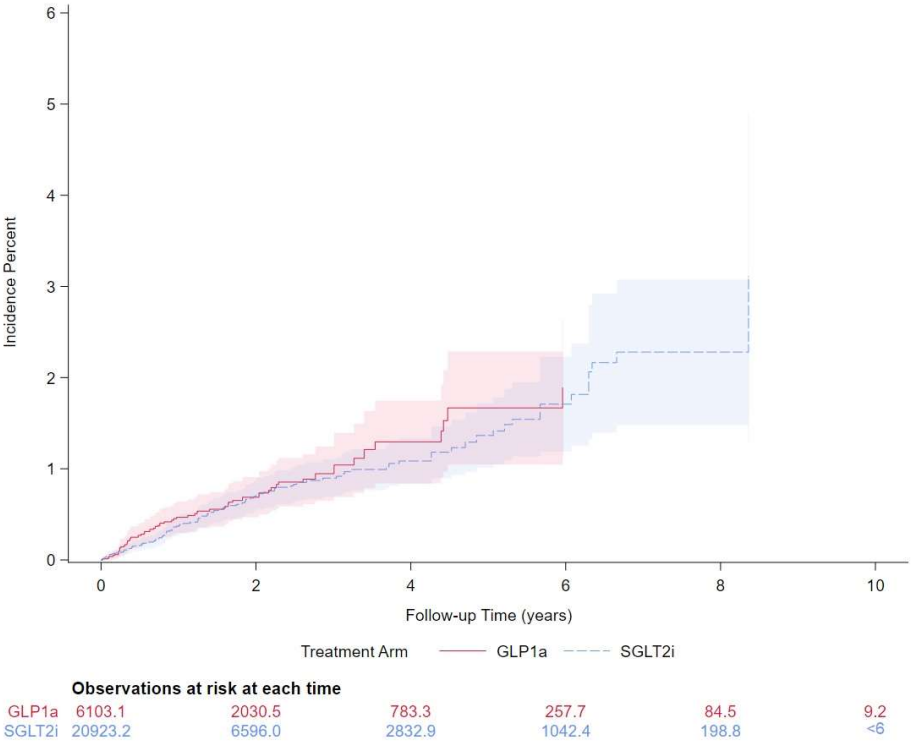
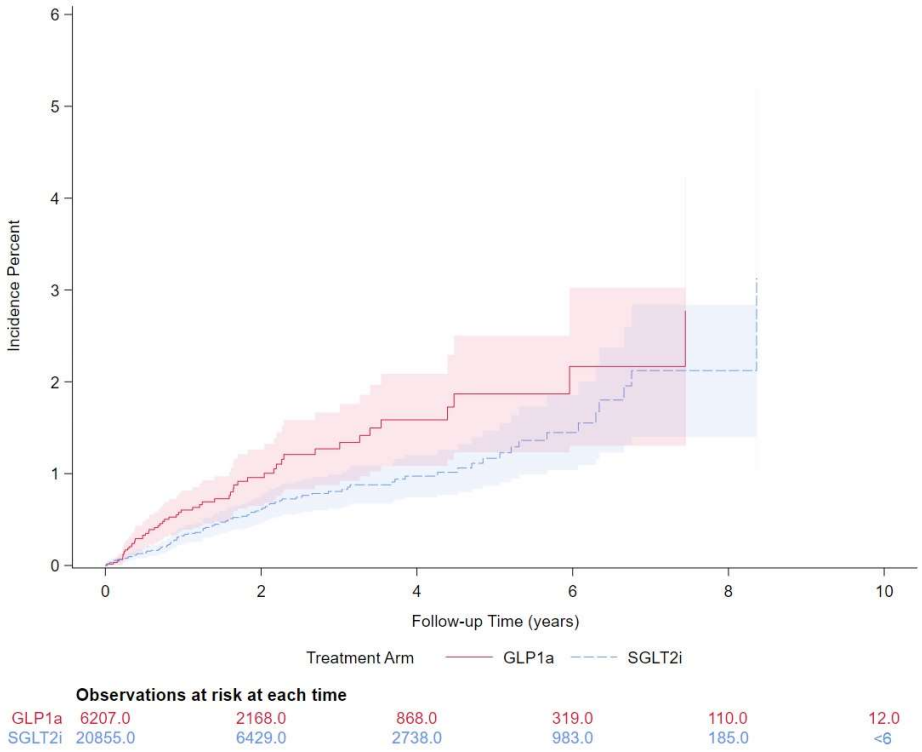


Figure 4. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **OT** analytical approach

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

(a) Unadjusted



(b) IPTW Adjusted

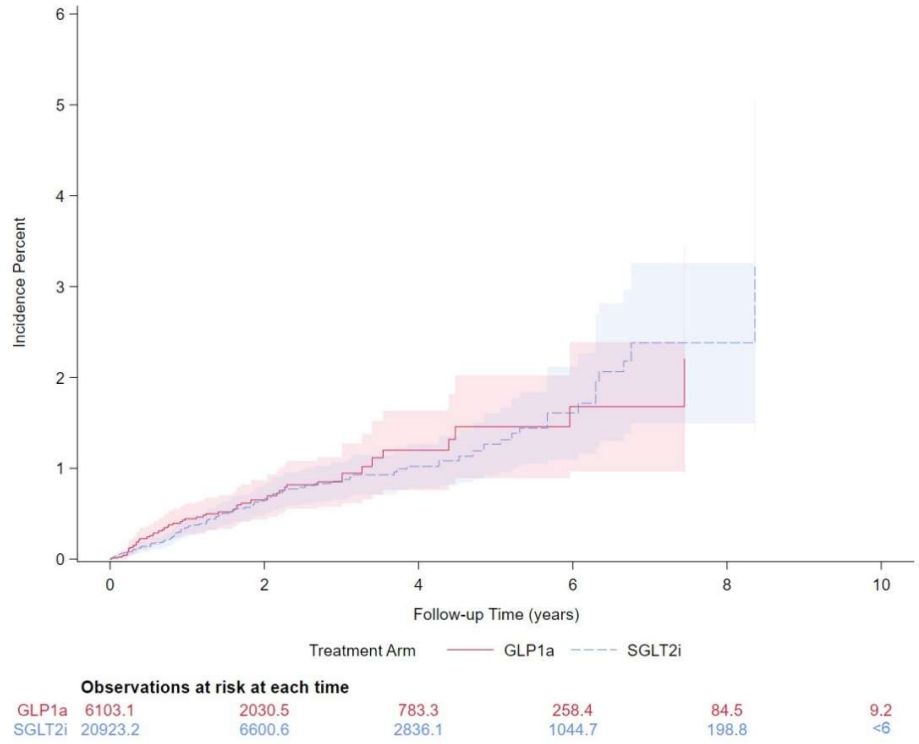


Figure 5. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **only suicide-related events** based on **OT** analytical approach.

7. Post-hoc analyses

7.1. Post-hoc analyses: Methods

Following review and discussion by PRAC, some additional explorative analyses not originally described in the protocol were suggested and accepted by the Rapporteur.

- **Restricting the study population to T2DM patients with obesity:** To explore if the effect of GLP-1 receptor agonists on the risk of suicide and self-harm might be modified by obesity, we conducted the analyses in a subset of T2DM patients with obesity. Thus, the study population was restricted to patients with at least one diagnosis of obesity within 2 years before index-date.

  Inverse probability of treatment weights were estimated in this subpopulation using the same variables (Section 5.5) and models (Section 5.6.3) as in the main analysis.

  Incidence rates, hazard ratios and cumulative incidence proportions were estimated as described for the main analysis (Section 5.6.3).

- **Restricting the study population to those with history of depression:** Improved glycaemic control and decreased weight have been reported to improve psychological and emotional well-being and health perceptions as suggested by some studies which showed reductions in the depression rating scale scores of T2DM patients treated with GLP-1 receptor agonists compared to those treated with control treatments (Chen et al., 2024). Therefore, to explore potential effect modification by this strong confounding factor, we conducted analyses restricting the study population to a subset of patients with at least one diagnosis of depression before index-date.

  Inverse probability of treatment weights were estimated in this subpopulation using the same variables (Section 5.5) and models (Section 5.6.3) as in the main analysis.

  Incidence rates, hazard ratios and cumulative incidence proportions were estimated as described for the main analysis (Section 5.6.3).

- **Use of shorter gap periods (i.e., 90 days) as a cut-off for treatment discontinuation:** To explore if the choice of prescription gap could influence the OT results, this was reduced from 180 days (in the main analysis) to 90 days. The entire baseline study population and the IPTWs calculated at index-date in the main analysis (Section 5.6.3) were used in this analysis.

  "On-treatment" follow-up was modified by considering the end of the index treatment as 90 days after the date of the last observed prescription followed by the first gap longer than 90 days in the sequence of prescriptions, and then censoring it at this time.

  Incidence rates, hazard ratios and cumulative incidence proportions were estimated as described for the main analysis (Section 5.6.3).

- **Adjustment for informative censoring at treatment discontinuation in the "On Treatment" analysis:** In the "on-treatment" analyses described in the Methods (Section 5.6.4 above), patients were censored when they discontinued their index treatment (i.e., either GLP-1 receptor agonists or SGLT-2 inhibitors) or when they switched to the opposite treatment arm. In this analysis, it was observed that treatment discontinuation/switch was potentially dependent on the treatment patients initiated at index-date resulting in informative censoring (Section 7.2, sensitivity analysis). Implicit in the "on-treatment"

analysis is an assumption that discontinuation/censoring is independent of the studied outcome, so that this differential treatment selection should not introduce bias.

In this post hoc analysis, we therefore adjusted for potential informative censoring, by reweighting patient observations using the inverse of the probability of being censored at treatment discontinuation/switch conditional on baseline confounders, measured during follow-up. (Hernan et al., 2000; Robins et al., 2000) To conduct this analysis, we performed the following steps:

i. The follow-up of each patient was divided into 90-day consecutive episodes, each receiving a **time indicator** which takes integer values from 1 to 42 (corresponding to the maximum follow-up of approximately 10 years). Additionally, each episode contained binary indicators for the outcome event, treatment discontinuation/switch or loss to follow-up taking place during the episode. Either one of these censoring indicators takes the value 1 in the last episode of a person's follow-up.

ii. The baseline confounders included in IPTW (Section 5.5) were time-updated before the start of each 90-day episode during follow-up. For anti-diabetic treatment (other than GLP-1 receptor agonist, SGLT-2 inhibitors and biguanide), use within 180 days before each 90-day episode (corresponding to recent use) was considered. For the other variables the same time windows were kept as defined at baseline.

iii. The probability of being censored via treatment discontinuation/switch was estimated conditionally on the **time indicator** and the **time-updated** confounders (Section 5.5) using pooled logistic regression applied on the dataset with several observations per patient. In this model, the time indicator was treated as a continuous variable and using polynomial terms to reduce model dimensionality. In addition, interaction terms between the time indicator and time-updated confounders were included. This analysis was stratified by treatment arm. The conditional probability of remaining uncensored in the dataset at each time was predicted as the conditional probability of not being censored at the previous time and represents the denominator of the inverse probability of censoring weights (IPCW) component at each time.

iv. To stabilize the weights, the numerator of the IPCW component was estimated using a similar procedure but omitting the adjustment by confounding factors from the model, such as to model the probability of being censored via treatment discontinuation/switch conditional **only** on **time indicator** and stratified by treatment arm.

v. Finally, the IPCWs were calculated for each episode by multiplying weight components from all previous episodes of an individual up to and including the current episode.

To adjust for both confounding bias at baseline and potential selection bias during follow-up, the baseline IPTWs (Section 5.6.3) were multiplied by the time-updated IPCWs. Extreme overall inverse probability weights were truncated to the 1^{st} or 99^{th} percentile values of the distribution.

Odds ratios, interpretable as hazard ratios, were estimated using a pooled logistic regression model of the binary outcome as a function of treatment and time (modelled as a continuous variable, the same as in the IPCW models). No interaction between treatment

and time was modelled to emulate the Cox proportional hazards model of the main analysis, which estimates one hazard ratio over the entire follow-up time.

In the unadjusted model, each observation had a weight of one while in the adjusted analyses observations were re-weighted. First, the baseline set of IPTWs were used to compare baseline confounding adjusted odds ratios obtained via the pooled logistic regression with the baseline confounding adjusted hazard ratios obtained via Cox proportional hazards modelling in the main analysis (Section 5.6.3 and Table 2). Then, the weights incorporating both baseline confounding adjustment as well as follow-up selection bias adjustment were used.

Cumulative incidence proportions were calculated using an estimator similar to the Kaplan-Meier estimator used in the main analysis, but discretizing time in 90-day episodes. At each time (i.e., in each 90-day episode), the probability of not experiencing the outcome event was calculated among patients at risk at that time. Survival up to each time was then calculated as the product of all such probabilities up to and including that time. Cumulative incidence, expressed per 100 individuals, was calculated as *(1-survival) * 100*.

7.2. Post-hoc analyses: Results

Restricting the study population to T2DM patients with obesity: This subpopulation of T2DM patients who were also obese included 5,472 patients initiating GLP-1 receptor agonists and 12,815 patients initiating SGLT-1 inhibitors. A total of 254 incident cases of suicide-related and self-harm-related events were identified in this subpopulation.

Supplementary Table 4 shows the baseline characteristics for both treatment groups, which were similar to those described for the entire population. The distribution of PS and inverse probability of treatment weights by treatment arm are shown in **Supplementary Figures 6** and **7**. After IPTW, the treatment groups were well balanced across all covariates, with all SMDs <0.1. The standardized mean differences between treatment arms for each baseline covariate before and after IPTW are shown in **Supplementary Figure 8**.

In the ITT analysis, the unadjusted risk of suicide-related and self-harm related events was higher in GLP-1 receptor agonists users compared with SGLT-2 inhibitor users (0.55 vs 0.35 events per 100 person-years), corresponding to a HR of 1.66 [95%CI: 1.30-2.13]). The HR dropped to 1.33 (95%CI: 1.02-1.72) after adjusting for confounding (**Supplementary Table 5**). Similarly, the contrast observed between the unadjusted cumulative incidence curves of the two treatment arms was attenuated after the IPTW adjustment (**Supplementary Figure 9**).

These results were consistent with those for the OT analysis (**Supplementary Table 5 and Figure 9**).

Restricting the study population to those with history of depression: This subpopulation of T2DM patients who had a history of depression included 2,617 patients initiating GLP-1 receptor agonists and 7,370 patients initiating SGLT-1 inhibitors. A total of 220 incident cases of suicide-related and self-harm-related events were identified in this subpopulation.

Supplementary Table 6 shows the baseline characteristics for both treatment groups. The distribution of PS and inverse probability of treatment weights by treatment arm are shown in **Supplementary Figures 11** and **12**. The treatment groups were well balanced across all covariates

after IPTW with all SMDs <0.1. The standardized mean differences between treatment arms for each baseline covariate before and after IPTW are shown in **Supplementary Figure 13**.

In the ITT analysis, the unadjusted risk of suicide-related and self-harm related events was higher in GLP-1 receptor agonists users compared with SGLT-2 inhibitor users (0.96 vs 0.65 events per 100 person-years), corresponding to a HR of 1.52 [95%CI: 1.16-2.00]). However, the HR dropped to 1.14 (95%CI: 0.86-1.52) after adjusting for confounding (**Supplementary Table 7**). Similarly, the contrast observed between the unadjusted cumulative incidence curves of the two treatment arms was attenuated after the IPTW adjustment (**Supplementary Figure 14**).

These results were consistent with those for the OT analysis (**Supplementary Table 7** and **Figure 15**) and the main analysis (**Table 2 and Figure 2**).

Use of shorter gap periods (i.e., 90 days) as a cut-off for treatment discontinuation including the entire study population: In the OT analysis where a 90-day period was considered to define treatment discontinuation, the unadjusted risk of suicide-related and self-harm related events was higher in GLP-1 receptor agonists users compared with SGLT-2 inhibitor users (0.53 vs 0.29 events per 100 person-years), corresponding to a HR of 1.81 [95%CI: 1.28-2.56]). This HR dropped to 1.24 (95%CI: 0.86-1.79) after adjusting for confounding (**Supplementary Table 8**). Similarly, the contrast observed between the unadjusted cumulative incidence curves of the two treatment arms was attenuated after the IPTW adjustment (**Supplementary Figure 16**).

Consistent with the main OT analysis in which a 180-day prescription gap was used as a cut-off for treatment discontinuation, the current OT analysis using a shorter gap showed a marginally numerically higher risk of suicide-related and self-harm related events in the GLP-1 receptor agonists arm (HR of 1.24 [95%CI 0.86-1.79] when using 90-day gap vs 1.11 [95%CI 0.80-1.56] when using 180-day gap).

Adjustment for informative censoring at treatment discontinuation in the "On Treatment" analysis including the entire study population: **Supplementary Figure 17** shows the distribution of the combined inverse probability weights (IPW), that is the product of the estimated inverse probability-of-treatment weights (IPTWs) and the inverse probability-of-censoring weights (IPCWs). Overall, mean weights for each follow-up episode (time indicator) were close the value of 1, except for the end of follow-up, when few patients remained in the study.

Supplementary Table 9 shows the unadjusted, IPTW adjusted and IPW adjusted (time-updated) pooled logistic odds ratios of the composite endpoint (incl. both suicide-related and self-harm-related events). In the unadjusted analysis, the use of GLP-1 receptor agonists was associated with a 67% increased risk of suicide-related and self-harm-related events (OR: 1.67 [95%CI: 1.22-2.29]) compared with the use of SGLT-2 inhibitors. However, this risk was attenuated after adjusting for baseline confounding via IPTW (OR: 1.12 [95%CI: 0.80-1.57]) and for time-varying confounding and selection bias via IPW (OR: 1.16 [95%CI: 0.83-1.63]).

Supplementary Figure 18 shows the cumulative incidence curves of the composite endpoint (including both suicide-related and self-harm-related events) over the follow-up discretized into approximately forty 90-day episodes. The contrast observed between the unadjusted cumulative incidence curves of the two treatment arms was attenuated after adjustment for baseline confounding and for time-varying confounding and selection bias. These results were consistent with the OT analysis of the main analysis.

8. Discussion

- In the main original analysis, prior to confounding adjustment, we observed a consistent association between GLP-1 receptor agonist initiation and suicide- and self-harm-related events in patients with T2DM, compared to SGLT-2 inhibitor initiation. However, following adjustment for baseline confounders using an IPTW approach, no significant association was observed in either of the ITT or OT treatment strategies. This pattern of results remained robust to the different sensitivity analyses performed.
- In additional post-hoc analyses, we repeated the main analysis restricting first to patients with obese T2DM only and secondly to people with T2DM and a prior history of depression. In the analysis of obese patients with T2DM, a higher risk of suicide- and self-harm-related events in the GLP-1 receptor agonists users was still observed following IPTW adjustment in the ITT analysis, with a similar sized but non-significant association in the OT analysis. No association was observed in either ITT or OT analysis in the cohort of T2DM patients with a prior history of depression.
- The additional analysis accounting for **informative censoring** during follow-up in the OT was similar to that observed in the main OT analysis suggesting no significant selection bias was present.
- Our findings align with two previous studies which assessed the association between use of GLP-1 receptor agonists and mental health-related outcomes. A cohort study conducted using the UK-based CPRD compared incidence rates of new-onset depression or self-harm in patients initiating incretin-based therapies with that of sulfonylureas and other glucose-lowering agents. This study used a new-user active comparator design with High-Dimensional Propensity Scores to control for confounding. Researchers found no strong evidence of an increased risk of a new diagnosis of depression or episode of self-harm in new users of GLP-1 receptor agonists (n=501) compared to new users of sulfonylureas (n=16409), thiazolidinediones (n=2011) or insulin (n=2745). (Gamble et al., 2018) Another study, a *post hoc* analysis, evaluated pooled neuropsychiatric safety data from the five randomized, double-blind, and placebo-controlled trials (n=5325) with liraglutide 3.0 mg used for weight management. The overall incidence of depression, anxiety and insomnia adverse events was similarly low (≤3.6%) in both liraglutide 3.0 mg and placebo treatment groups, with a small numerical imbalance not favouring liraglutide in insomnia events and in suicidality ideation/behaviour. However, there were no between-treatment imbalances in depression or suicidality indicators when prospectively assessed in the phase 3a programme, which excluded individuals with a history of severe psychiatric disorders or suicidal ideation/behaviour. (O'Neil et al., 2017) In addition, a recent cohort study investigated the association between use of semaglutide and risk suicidal ideation comparing to use of non-GLP-1 receptor agonist anti-obesity medications (including orlistat, phentermine, topiramate, bupropion, naltrexone) as well as anti-diabetes medications (excluding GLP-1 receptor agonists). (Wang et al., 2024) The study population included 240,618 patients who were overweight or obese and 1,589,855 patients with T2DM. Propensity score matching was used to control for confounding. In patients with overweight or obesity, semaglutide compared with non-GLP1R agonist anti-obesity medications was associated with lower risk for incident (HR: 0.27; 95% CI: 0.20–0.36) and recurrent (HR: 0.44; 95% CI: 0.32–0.60) suicidal ideation. Similar findings were replicated in patients with T2DM. Authors concluded that these results do not support higher risks of suicidal ideation with semaglutide compared with non-GLP1R agonist anti-obesity or anti-diabetes medications.

8.1. Strengths and limitations

- A strength of our study is the use of new-user active comparator design, an approach that limits prevalent user bias. To reduce the possibility of major unmeasured confounding, an active comparator (SGLT-2 inhibitors), typically prescribed at the same disease stage as the target treatment (GPL-1 receptor inhibitors), was used. Additionally, IPTW was used to control for several potential confounding factors.
- The entire class of GLP-1 receptor agonists was combined (dulaglutide, exenatide, liraglutide, lixisenatide, semaglutide and teduglutide) and therefore it is uncertain how generalisable our findings are to specific drug substances. Moreover, the two individual GLP-1 receptor agonists which were the subject of the safety signal (liraglutide and semaglutide) could not be studied separately due to sample size constraints (Supplementary Table 1). However, it is worth noting that investigating the entire class of GLP-1 receptor agonists was also of interest for the signal evaluation.
- Exposure is misclassified in ITT analyses if a large proportion of study participants change the baseline treatment early during follow up. We found that 57% of GLP-1 receptor agonist initiators and 39% of SGLT-2-inhibitor initiators interrupted their baseline treatment over the course of ITT follow-up (i.e., before experiencing an outcome or censoring event). The median proportion of the ITT follow-up spent on the baseline treatment, among those who interrupted it, was 30 to 40%.
- An alternative to ITT is to follow study participants only as long as they remain on the treatment received at the start of the study. Thus, in such "on treatment" (OT) analyses, exposure is correctly classified. However, if censoring of patients at baseline treatment discontinuation is informative of the outcome, this could introduce selection bias. Therefore, an OT sensitivity analysis was conducted. Besides adjusting for baseline confounding as in the main ITT analysis, time varying selection bias was also adjusted for (i.e., accounting for informative censoring) in the OT analysis. This additional adjustment had only a reduced influence on effect estimates suggesting that treatment discontinuation might not be strongly associated with the outcome.
- A certain degree of outcome misclassification cannot be excluded. Accurate and complete ascertainment of suicidal attempts and completed suicide in electronic health records presents challenges (Hall, 2009; Salmeron et al., 2013; Swain et al., 2019; Wijlaars et al., 2013). For example, suicide attempts (or suicidal ideation) may not always result in health care encounters. Absence of recorded clinical care is likely to be a major source of under ascertainment (i.e., false negatives). In addition, health care practitioners must accurately describe the harm as intentional or accidental, which may not always be clear; misclassification of intent can result in both false-positives and false-negatives, although false-negatives are expected to occur more frequently (Swain et al., 2019). Of note, there was a limited number of self-harm events, and the study was not powered to detect clinically relevant differences across exposure groups for this component of the composite endpoint.
- The possibility of the outcome being differentially misclassified if, for instance, GLP-1 receptor agonist users were more closely screen for mental health conditions cannot be ruled out. However, all patients with T2DM requiring second line treatment are regularly monitored in UK primary care suggesting this may be unlikely, particularly considering that the information about an increased risk of suicide/self-harm is not yet present in the SmPCs of either liraglutide or semaglutide containing products.

- To our knowledge, the accuracy of coding suicide- or self-harm- related events (such as suicidal ideation, self-injury ideation, intentional self-injury/self-harm, attempted suicide) has not been assessed for the UK IMRD database. A study which assessed the validity of suicide recording in the UK THIN database found that the cause of death is incomplete even when free text entries are reviewed, therefore, any study of fatal, acute conditions would underestimate the rate of an event. (Hall, 2009)
- Finally, our study is also subject to the common limitations of observational studies including the potential for residual confounding and informative censoring. Although 26 potential confounders were adjusted for, not all relevant potential confounders (such as severity of T2DM, severity of depression, and patient-level socioeconomic status) could be captured.

9. Conclusion

Our study results do not support a clinically relevant increased risk of suicide-/self-harm- related events associated with GLP-1 receptor agonists versus SGLT-2 receptor inhibitors in T2DM patients. However, further research will be needed to replicate and confirm our findings and strengthen the available evidence. Studies of individuals for whom GLP-1 receptor agonists are indicated for weight management (i.e., obese patients) are of particular interest.

10. References

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res*, *46*(3), 399-424. https://doi.org/10.1080/00273171.2011.568786

Barbui, C., Esposito, E., & Cipriani, A. (2009). Selective serotonin reuptake inhibitors and risk of suicide: a systematic review of observational studies. *CMAJ*, *180*(3), 291-297. https://doi.org/10.1503/cmaj.081514

Bland, J. M., & Altman, D. G. (1998). Survival probabilities (the Kaplan-Meier method). *BMJ*, *317*(7172), 1572. https://doi.org/10.1136/bmj.317.7172.1572

Chen, X., Zhao, P., Wang, W., Guo, L., & Pan, Q. (2024). The Antidepressant Effects of GLP-1 Receptor Agonists: A Systematic Review and Meta-Analysis. *Am J Geriatr Psychiatry*, *32*(1), 117-127. https://doi.org/10.1016/j.jagp.2023.08.010

Fergusson, D., Doucette, S., Glass, K. C., Shapiro, S., Healy, D., Hebert, P., & Hutton, B. (2005). Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials. *BMJ*, *330*(7488), 396. https://doi.org/10.1136/bmj.330.7488.396

Flegal, K. M., Kit, B. K., Orpana, H., & Graubard, B. I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *JAMA*, *309*(1), 71-82. https://doi.org/10.1001/jama.2012.113905

Fragoso, Y. D., Frota, E. R., Lopes, J. S., Noal, J. S., Giacomo, M. C., Gomes, S., Goncalves, M. V., da Gama, P. D., & Finkelsztejn, A. (2010). Severe depression, suicide attempts, and ideation during the use of interferon beta by patients with multiple sclerosis. *Clin Neuropharmacol*, *33*(6), 312-316. https://doi.org/10.1097/WNF.0b013e3181f8d513

Gamble, J. M., Chibrikov, E., Midodzi, W. K., Twells, L. K., & Majumdar, S. R. (2018). Examining the risk of depression or self-harm associated with incretin-based therapies used to manage hyperglycaemia in patients with type 2 diabetes: a cohort study using the UK Clinical Practice Research Datalink. *BMJ Open*, *8*(10), e023830. https://doi.org/10.1136/bmjopen-2018-023830

Garner, R. E., Feeny, D. H., Thompson, A., Bernier, J., McFarland, B. H., Huguet, N., Kaplan, M. S., Orpana, H., Ross, N. A., & Blanchard, C. (2012). Bodyweight, gender, and quality of life: a population-based longitudinal study. *Qual Life Res*, *21*(5), 813-825. https://doi.org/10.1007/s11136-011-9989-1

Guh, D. P., Zhang, W., Bansback, N., Amarsi, Z., Birmingham, C. L., & Anis, A. H. (2009). The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC Public Health*, *9*, 88. https://doi.org/10.1186/1471-2458-9-88

Hall, G. C. (2009). Validation of death and suicide recording on the THIN UK primary care database. *Pharmacoepidemiol Drug Saf*, *18*(2), 120-131. https://doi.org/10.1002/pds.1686

Hernán M, & Robins J. (2020). *Estimating IP weights via modeling. In: Causal Inference: What If*. Chapman & Hall/CRC: p 160.

Hernan, M. A., Brumback, B., & Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, *11*(5), 561-570. https://doi.org/10.1097/00001648-200009000-00012

Hernan, M. A., & Robins, J. M. (2006). Estimating causal effects from epidemiological data. *J Epidemiol Community Health*, *60*(7), 578-586. https://doi.org/10.1136/jech.2004.029496

Ju, Y. J., Han, K. T., Lee, T. H., Kim, W., Park, J. H., & Park, E. C. (2016). Association between weight control failure and suicidal ideation in overweight and obese adults: a cross-sectional study. *BMC Public Health*, *16*, 259. https://doi.org/10.1186/s12889-016-2940-1

Khaledi, M., Haghighatdoost, F., Feizi, A., & Aminorroaya, A. (2019). The prevalence of comorbid depression in patients with type 2 diabetes: an updated systematic review and meta-analysis on huge number of observational studies. *Acta Diabetol*, *56*(6), 631-650. https://doi.org/10.1007/s00592-019-01295-9

Lucaciu, L. A., & Dumitrascu, D. L. (2015). Depression and suicide ideation in chronic hepatitis C patients untreated and treated with interferon: prevalence, prevention, and treatment. *Ann Gastroenterol*, *28*(4), 440-447. https://www.ncbi.nlm.nih.gov/pubmed/26424594

Luppino, F. S., de Wit, L. M., Bouvy, P. F., Stijnen, T., Cuijpers, P., Penninx, B. W., & Zitman, F. G. (2010). Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry*, *67*(3), 220-229. https://doi.org/10.1001/archgenpsychiatry.2010.2

Mula, M., & Hesdorffer, D. C. (2011). Suicidal behavior and antiepileptic drugs in epilepsy: analysis of the emerging evidence. *Drug Healthc Patient Saf*, *3*, 15-20. https://doi.org/10.2147/DHPS.S13070

NICE. (2022). *NICE guideline. Type 2 diabetes in adults: management* Retrieved from https://www.nice.org.uk/guidance/ng28/resources/type-2-diabetes-in-adults-management-pdf-1837338615493

O'Neil, P. M., Aroda, V. R., Astrup, A., Kushner, R., Lau, D. C. W., Wadden, T. A., Brett, J., Cancino, A. P., Wilding, J. P. H., Satiety, Clinical Adiposity - Liraglutide Evidence in individuals, w., & without diabetes study, g. (2017). Neuropsychiatric safety with liraglutide 3.0 mg for weight management: Results from randomized controlled phase 2 and 3a trials. *Diabetes Obes Metab*, *19*(11), 1529-1536. https://doi.org/10.1111/dom.12963

Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, *11*(5), 550-560. https://doi.org/10.1097/00001648-200009000-00011

Roy, T., & Lloyd, C. E. (2012). Epidemiology of depression and diabetes: a systematic review. *J Affect Disord*, *142 Suppl*, S8-21. https://doi.org/10.1016/S0165-0327(12)70004-6

Salmeron, D., Cirera, L., Ballesta, M., & Navarro-Mateu, F. (2013). Time trends and geographical variations in mortality due to suicide and causes of undetermined intent in Spain, 1991-2008. *J Public Health (Oxf)*, *35*(2), 237-245. https://doi.org/10.1093/pubmed/fds103

Scott, K. M., McGee, M. A., Wells, J. E., & Oakley Browne, M. A. (2008). Obesity and mental disorders in the adult general population. *J Psychosom Res*, *64*(1), 97-105. https://doi.org/10.1016/j.jpsychores.2007.09.006

Skovlund, C. W., Morch, L. S., Kessing, L. V., Lange, T., & Lidegaard, O. (2018). Association of Hormonal Contraception With Suicide Attempts and Suicides. *Am J Psychiatry*, *175*(4), 336-342. https://doi.org/10.1176/appi.ajp.2017.17060616

Sturmer, T., Rothman, K. J., Avorn, J., & Glynn, R. J. (2010). Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. *Am J Epidemiol*, *172*(7), 843-854. https://doi.org/10.1093/aje/kwq198

Swain, R. S., Taylor, L. G., Braver, E. R., Liu, W., Pinheiro, S. P., & Mosholder, A. D. (2019). A systematic review of validated suicide outcome classification in observational studies. *Int J Epidemiol*, *48*(5), 1636-1649. https://doi.org/10.1093/ije/dyz038

Wang, W., Volkow, N. D., Berger, N. A., Davis, P. B., Kaelber, D. C., & Xu, R. (2024). Association of semaglutide with risk of suicidal ideation in a real-world cohort [Article]. *Nature Medicine*, *30*(1), 168-176. https://doi.org/10.1038/s41591-023-02672-2

Wijlaars, L. P., Nazareth, I., Whitaker, H. J., Evans, S. J., & Petersen, I. (2013). Suicide-related events in young people following prescription of SSRIs and other antidepressants: a self-controlled case series analysis. *BMJ Open*, *3*(9), e003247. https://doi.org/10.1136/bmjopen-2013-003247

Zhao, G., Ford, E. S., Dhingra, S., Li, C., Strine, T. W., & Mokdad, A. H. (2009). Depression and anxiety among US adults: associations with body mass index. *Int J Obes (Lond)*, *33*(2), 257-266. https://doi.org/10.1038/ijo.2008.268

Zheng, Y., Ley, S. H., & Hu, F. B. (2018). Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol*, *14*(2), 88-98. https://doi.org/10.1038/nrendo.2017.151

Annex 1 - Information on Databases and Healthcare systems included

IQVIA™ Medical Research Data (IMRD) UK

IQVIA™ Medical Research Data (IMRD) UK is a primary care database from the UK. GPs play a gatekeeper role in the healthcare system in the UK, as they are responsible for delivering primary health care and specialist referrals. Over 98% of the UK-resident population is registered with a GP, so that GP patient records are broadly representative of the UK population in general. Patients are affiliated to a practice, which centralizes the medical information from GPs, specialist referrals, hospitalizations, and tests.

Annex 2 - Codelists

2.1 List of codes to identify study population

20230807--FA_IMRD__T2DM_codes_trim

2.2 List of codes to identify antidiabetic treatment exposure

Class	Code (EPHMRA ATC)
Insulins (including combinations)	A10C, A10D
Sulphonylureas (including combinations)	A10H, A10J2, A10K2
Thiazolidinediones (including combinations)	A10K
Biguanides (including combinations)	A10J
DPP-4 inhibitors (including combinations)	A10N, A10P5
SGLT-2 inhibitors (including combinations)	A10P
GLP-1 receptor agonists (including combinations)	A10S, A10C9
Glinide	A10M
Other (acarbose, guar gum)	A10X, A10L

2.3 List of codes to identify outcome

20230915_UK_IMRD_suicide_selfinjury_r

Note: PHQ9 score - thoughts of suicide or self harm" was only used where the value associated with the code was between 1 and 3. The code is derived from a questionnaire aiming to assess severity of depression. "PHQ9 score - thoughts of suicide or self harm" is one of 9 questions with the following answer options: "0 - Not at all", "1 - Several days", "2 - More than half the days", "3 - Nearly every day". Zero was the answer in approximately 80% of cases.

2.4 List of codes to identify and define covariates

2.4.1 Obesity diagnosis

List of codes to identify obesity diagnoses:


obes_bin.csv

List of codes to identify BMI measurements:


obes_th.csv

Definition:

A patient is considered to have had a recent history of obesity if at least one code from the list of obesity diagnoses above was recorded within 2 years before index-date or if they had a BMI measurement with value between 30 and 80 within 2 years before index-date.

2.4.2 Obesity treatment

List of codes to identify obesity treatments:


obes_rx_bin.csv

Definition:

A patient is considered to have had a recent obesity treatment if at least one code from the list of obesity treatments above was recorded within 2 years before index-date.

2.4.3 Cancer

Lists of codes to identify cancer diagnoses and chemotherapy or radiotherapy procedures:


cancer_bin.csv


chimio_bin.csv


radio_bin.csv

Definition:

A patient is considered to have had a history of cancer if at least one code from any of the lists above was recorded any time before index-date.

2.4.4 Renal disease

List of codes to identify renal disease diagnoses or dialysis procedures:


renal_bin.csv

Definition:

A patient is considered to have had a history of renal disease if at least one code from the list above was recorded any time before index-date.

2.4.5 Myocardial infarction

List of codes to identify myocardial infarction diagnoses:

MI_bin.csv

Definition:

A patient is considered to have had a recent myocardial infarction if at least one code from the list above was recorded within 2 years before index-date.

2.4.6 Congestive heart failure

List of codes to identify congestive heart failure diagnoses:

CHF_bin.csv

Definition:

A patient is considered to have had a history of congestive heart failure if at least one code from the list above was recorded any time before index-date.

2.4.7 Cerebrovascular disease

List of codes to identify cerebrovascular disease diagnoses:

stroke_bin.csv

Definition:

A patient is considered to have had a history of cerebrovascular disease if at least one code from the list above was recorded any time before index-date.

2.4.8 Anxiety diagnosis

List of codes to identify anxiety diagnoses:

anxi_bin.csv

Definition:

A patient is considered to have had a history of anxiety if at least one code from the list above was recorded any time before index-date.

2.4.9 Anxiety treatment

List of codes to identify anxiolytic treatments:

anxi_rx_bin.csv

Definition:

A patient is considered to have had a history of anxiolytic medication use if at least one code from the list above was recorded any time before index-date.

2.4.10 Depression diagnosis

List of codes to identify depression diagnoses:

depres_bin.csv

Definition:

A patient is considered to have had a history of depression if at least one code from the list above was recorded any time before index-date.

2.4.11 Depression treatment

List of codes to identify anti-depressive medication:

depres_rx_bin.csv

Definition:

A patient is considered to have had a history of anti-depressive medication use if at least one code from the list above was recorded any time before index-date.

2.4.12 Schizophrenia/Psychosis diagnosis

List of codes to identify schizophrenia diagnoses:

psycho_bin.csv

Definition:

A patient is considered to have had a history of schizophrenia diagnosis if at least one code from the list above was recorded any time before index-date.

2.4.13 Stress related and adjustment disorders

List of codes to identify stress related diagnosis:

stress_bin.csv

Definition:

A patient is considered to have had a history of stress disorders if at least one code from the list above was recorded any time before index-date.

2.4.14 Dementia diagnosis

List of codes to identify stress related diagnosis:

dement_bin.csv

Definition:

A patient is considered to have had a history of stress of dementia if at least one code from the list above was recorded any time before index-date.

2.4.15 Substance (including alcohol) abuse

List of codes to identify substance abuse diagnoses:

subst_abus_bin.csv

List of codes to identify answer to individual AUDIT score questions:

AUDIT_q_th.csv

List of codes to identify short AUDIT score versions:

AUDIT_c_th.csv

List of codes to identify full AUDIT score:

AUDIT_f_th.csv

List of codes to identify weekly quantitative alcohol consumption measurements:

alc_cons_w_th.csv

List of codes to identify daily quantitative alcohol consumption measurements:

alc_cons_d_th.csv

Definition:

A patient is considered to have had a recent history of substance abuse if at least one code from the list of substance abuse diagnoses above was recorded within 2 years before index-date or if they had any of the following measurements within the same time window: (1) a score of 4 for the answer to an individual AUDIT question; (2) a score between 9 and 12 for answers to short AUDIT versions; (3) a score between 20 and 40 for answers to the full AUDIT version; (4) an alcohol consumption value between 35 unit, for females, or 50 units, for males, and 500 units per week or (5) an alcohol consumption value between 6 units, for females, or 8 units, for males, and 100 units per day.

2.4.16 Smoking

List of codes to identify recent smoking:

smoking_th.csv

Definition:

A patient was considered a recent smoker if at least one code from the list above was recorded within 2 years before index-date, accompanied by either a missing value or a value higher than 0 (as some codes were accompanied by the number of cigarettes smoked per time unit).

Annex 3 – Statistical Analyses

3.1 Estimation of propensity scores

Propensity scores were estimated using logistic regression. Continuous variables were modelled as third-degree polynomials (i.e., including main variable plus square and cubic transformations) to allow non-linear relationships between continuous co-variates and exposure.

```
proc logistic data=input_data_set noprint;
      model trt_arm(ref="0")= <list of covariates>
      output out=PS_output predicted=_ps_ ;
run;
```

3.2 Incidence rate estimation (with weighting adjusted 95% confidence intervals)

```
proc genmod data=input_data_set;
class patid trt_arm;
model &outcome_event.=trt_arm /dist=poisson link=log offset=&ln_FU.;
lsmeans trt_arm /ilink cl;
weight &adjust.;
repeated subject=patid / type=ind;
ods output LSMeans=&pop._&outc._&adjust._&analysis._IR_2;
run;
```

<weight &adjust.;> option used to allow re-weighting the population using the inverse probability of treatment weights.

<repeated subject=patid / type=ind;> option used to account for weighting the population in the standard error estimation.

3.2 Cumulative incidence estimation

```
proc phreg data=input_data_set covs(aggregate);
model &FU.*&event.(0)=;
strata trt_arm;
baseline out=output_survival survival=_all_ /method=pl ;
weight &adjust.;
id patid;
run;
```

<model &FU.*&event.(0)=;> empty model indicated the non-parametric survival estimator.
<baseline…/method=pl ;> indicates the use of the product limit estimator.

<weight &adjust.;> option used to allow re-weighting the population using the inverse probability of treatment weights.

<covs(aggregate);> <id patid;> option used to account for weighting the population in the standard error estimation.

Annex 4 – Supplementary material

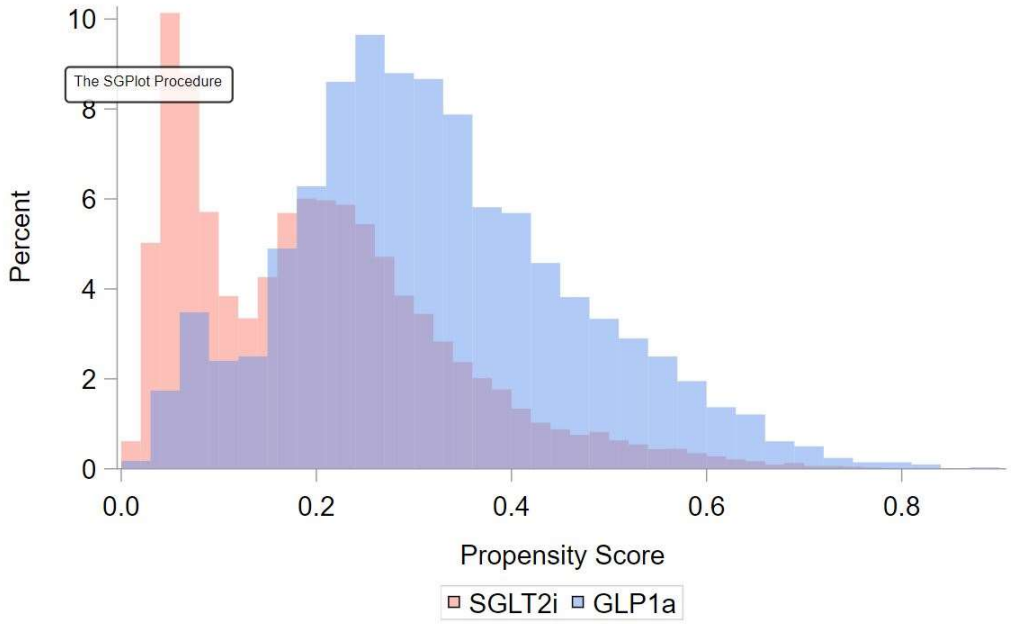
Supplementary Table 1 Individual drugs included under each treatment arm

GLP-1 receptor agonists		
Substance	**No. of individuals initiating each drug**	**% of individuals initiating each drug**
LIRAGLUTIDE	1922	30.97
SEMAGLUTIDE	1638	26.39
DULAGLUTIDE	1619	26.08
EXENATIDE	572	9.22
LIXISENATIDE	383	6.17
INSULIN DEGLUDEC/LIRAGLUTIDE	73	1.18
SGLT-2 inhibitors		
Substance	**No. of individuals initiating each drug**	**% of individuals initiating each drug**
DAPAGLIFLOZIN	10527	50.48
EMPAGLIFLOZIN	7410	35.53
CANAGLIFLOZIN	2412	11.57
METFORMIN/EMPAGLIFLOZIN	232	1.11
METFORMIN/DAPAGLIFLOZIN	160	0.77
ERTUGLIFLOZIN	68	0.33
METFORMIN/CANAGLIFLOZIN	23	0.11
SAXAGLIPTIN/DAPAGLIFLOZIN	14	0.07
LINAGLIPTIN/EMPAGLIFLOZIN	9	0.04

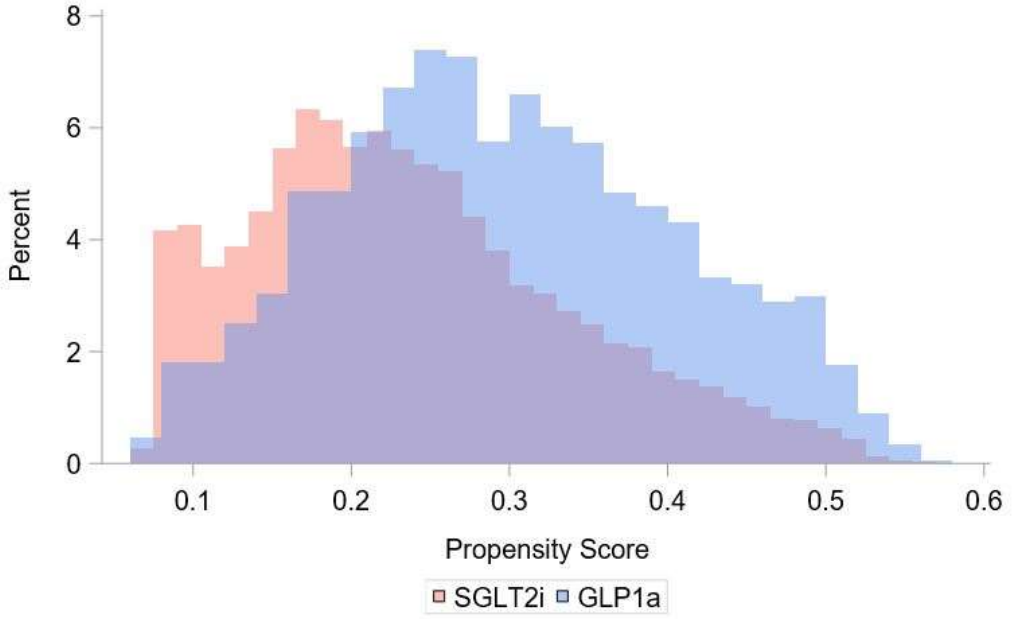
Supplementary Table 2. Frequency of clinical terms corresponding to identified outcome events in the main ITT analysis

Clinical terms included in outcome definition	N	Percent
Suicidal thoughts	85	28.3
PHQ9 score - thoughts of suicide or self-harm	77	25.7
Suicidal ideation	46	15.3
Intentional drug overdose	14	4.7
Thoughts of deliberate self-harm	12	4.0
Moderate suicide risk	9	3.0
Deliberate drug overdose / other poisoning	9	3.0
Suicidal	8	2.7
Suicide attempt	7	2.3
Intentionally harming self	6	2.0
At risk of DSH - deliberate self-harm	<6	<2.0
Suicidal intent	<6	<2.0
Suicide risk	<6	<2.0
Self-harm	<6	<2.0
Self-injurious behaviour	<6	<2.0
Attempted suicide - hanging	<6	<2.0
Expresses suicidal/homicidal thoughts	<6	<2.0
Feeling suicidal	<6	<2.0
Moderate risk of deliberate self-harm	<6	<2.0
Planning suicide	<6	<2.0
Samaritans advisory service	<6	<2.0
Suicidal - symptom	<6	<2.0
Threatening suicide	<6	<2.0
Attempted suicide	<6	<2.0
Intentional self-poison/exposure to other/unspecified drug/medicament	<6	<2.0
Intentional self-harm by sharp object	<6	<2.0
Total	300	100.0

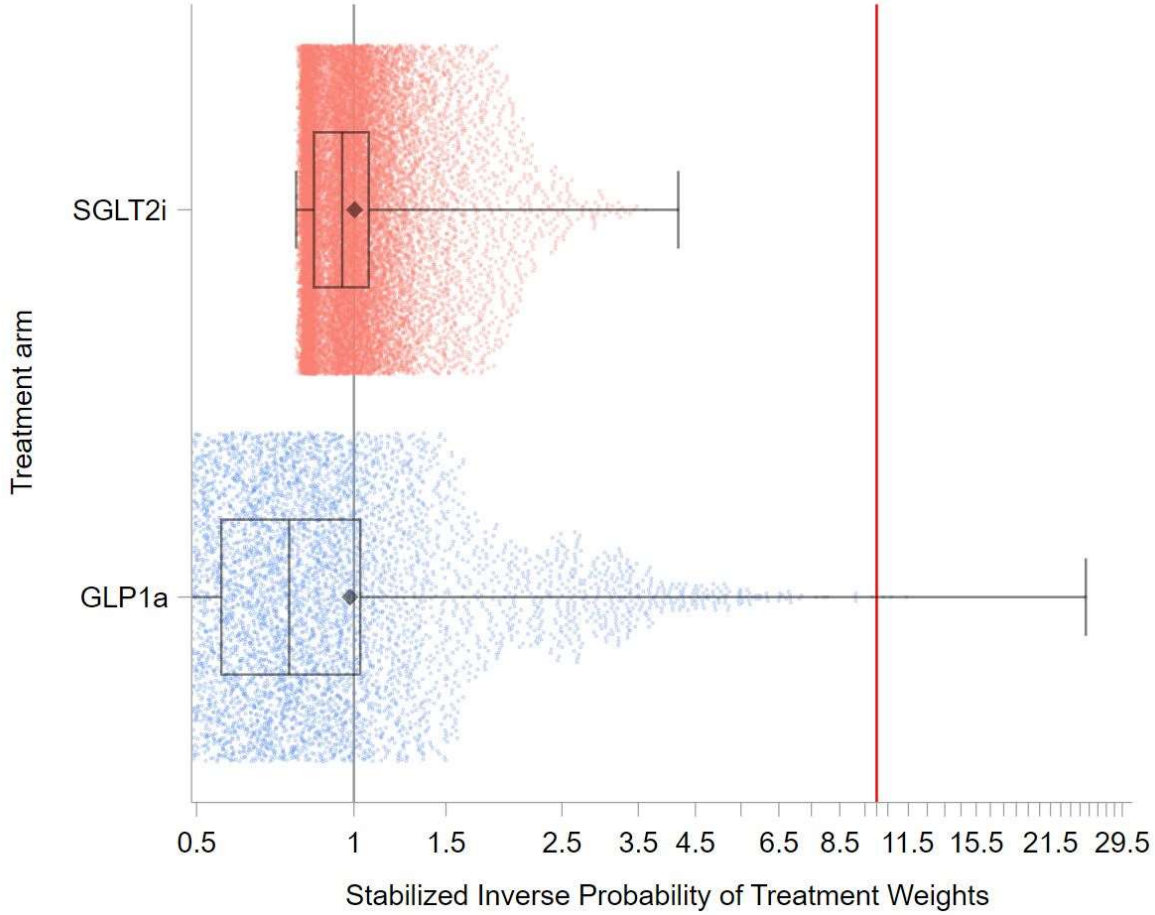
Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.



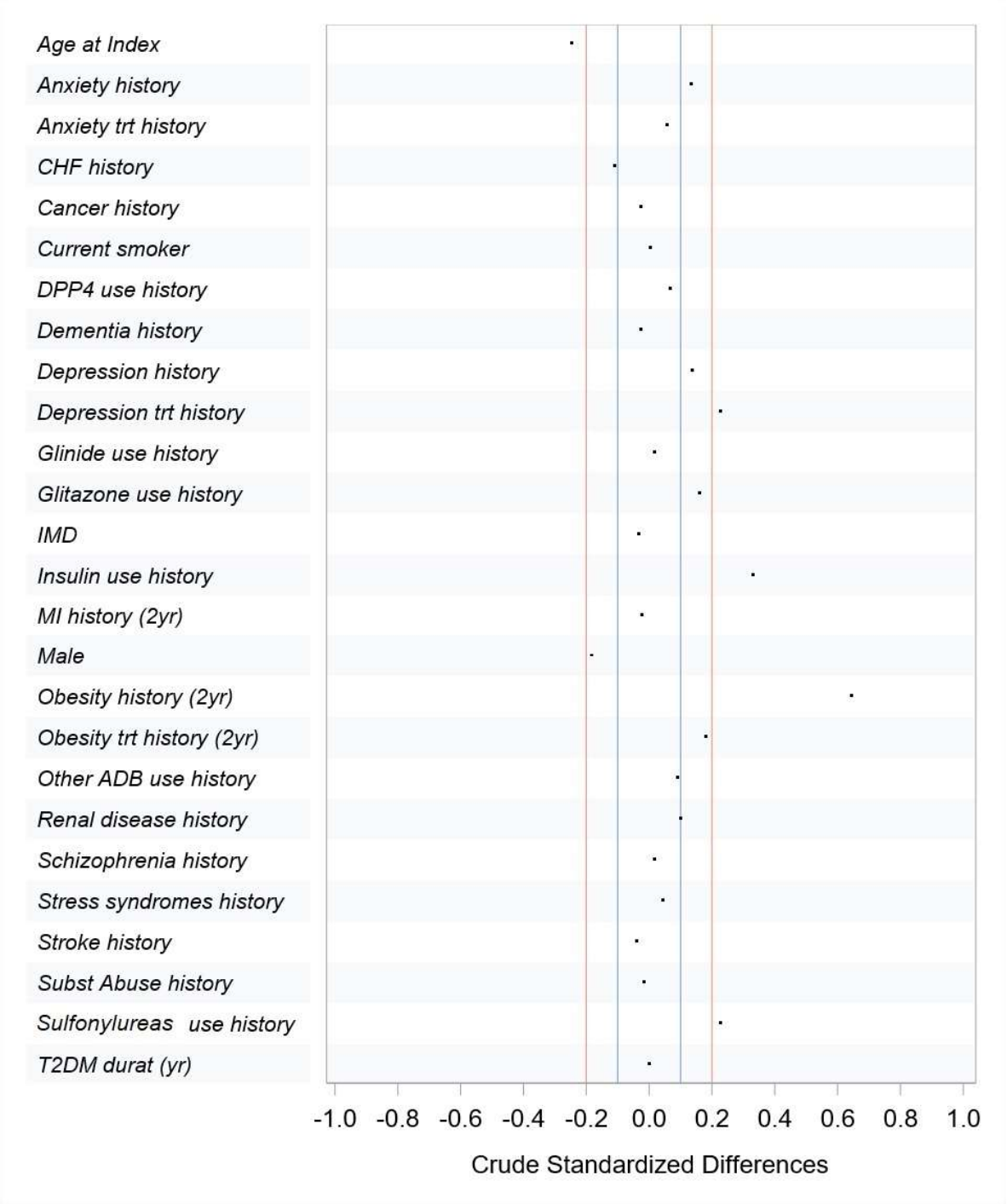
Supplementary Figure 1a. Propensity score distribution by treatment arm in the entire study population.



Supplementary Figure1b. Propensity score distribution by treatment arm in the study population trimmed according to the Stürmer method.



Supplementary Figure 2. Distribution of inverse probability of treatment weights by treatment arm. The red vertical line indicates a value of 10.



Supplementary Figure 3a. Love plot of standardized mean differences between treatment arms for each baseline covariate **before** inverse probability of treatment weighting (IPTW). The blue vertical line indicates a standardized mean difference of 0.1 and the red line a difference of 0.2 (in either direction).

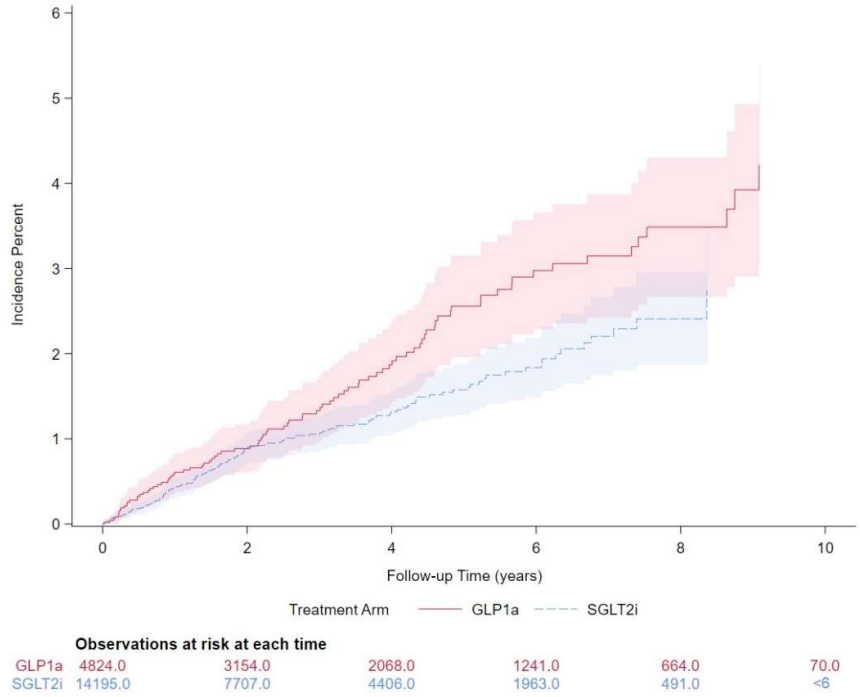
Supplementary Figure 3b. Love plot of standardized mean differences between treatment arms for each baseline covariate **after** inverse probability of treatment weighting (IPTW). The blue vertical line indicates a standardized mean difference of 0.1 and the red line a difference of 0.2 (in either direction).

Supplementary Table 3. Unadjusted and adjusted incidence rates per 100 person- years and hazard ratios of composite endpoint (incl. both suicide-related and self-harm-related events) after Stürmer trimming by treatment arm and analytical approach

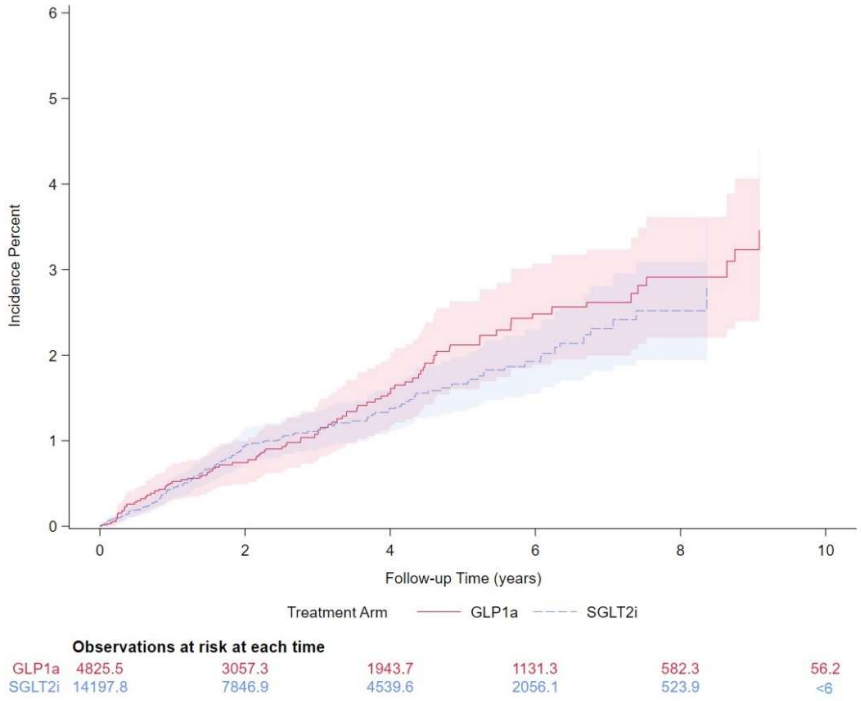
	ITT									OT							
Treatment arm	**N**	**Follow-up (person-years)**	**n events**	**IR**	**95% CI**		**HR**	**95% CI**		**Follow-up (person-years)**	**n events**	**IR**	**95% CI**		**HR**	**95% CI**	
Unadjusted analysis																	
GLP-1 receptor agonist	4824	19108.9	90	0.47	0.38	0.58	**1.42**	1.09	1.85	9734.8	42	0.43	0.32	0.58	**1.25**	0.87	1.81
SGLT-2 inhibitors	14195	41808.1	143	0.34	0.29	0.40	**1.00**	[Reference]		26710.7	92	0.34	0.28	0.42	**1.00**	[Reference]	
IPTW adjusted analysis																	
GLP-1 receptor agonist	4825	18253.3	71.6	0.39	0.32	0.49	**1.13**	0.86	1.48	9476.5	35.7	0.38	0.27	0.52	**1.04**	0.71	1.53
SGLT-2 inhibitors	14198	42658.6	152.2	0.36	0.30	0.42	**1.00**	[Reference]		27000.6	96.8	0.36	0.29	0.44	**1.00**	[Reference]	

ITT: intention-to-treat analytical approach, OT: On-treatment analytical approach; IR: incidence rate; 95% CI: 95% confidence interval; HR: Hazard ratio; IPTW: inverse probability of treatment weighting.

(a) Unadjusted



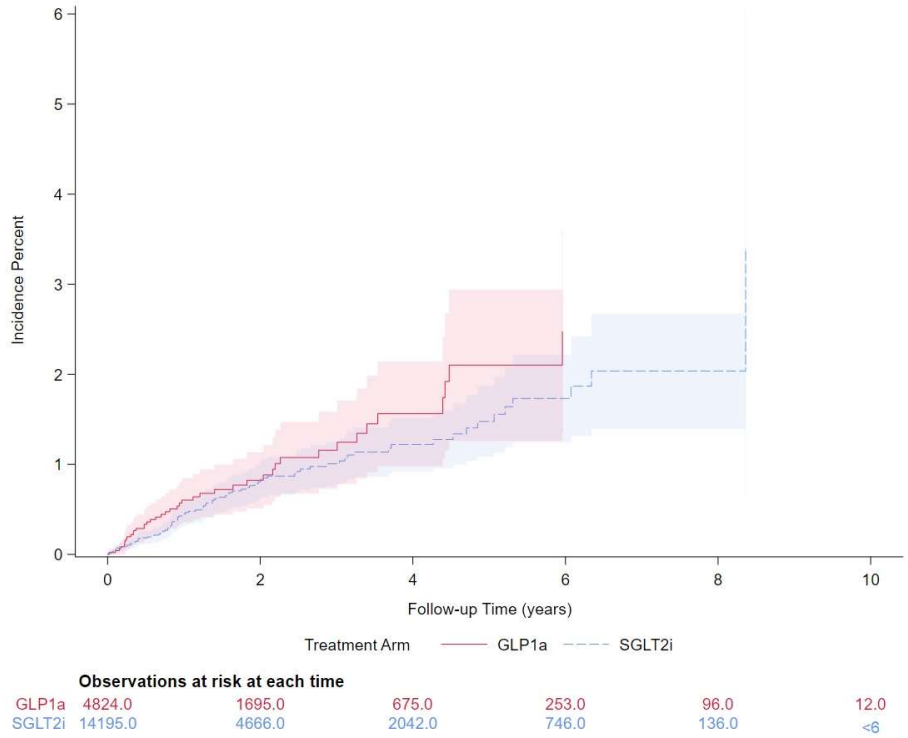
(b) IPTW Adjusted



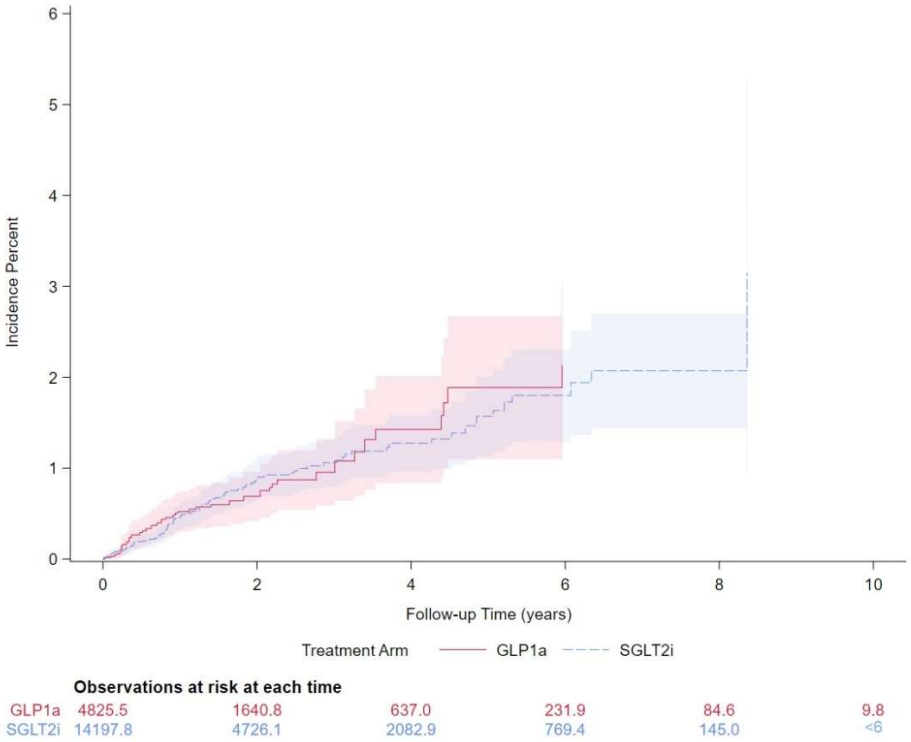
Supplementary Figure 4. Unadjusted and IPTW adjusted cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **ITT** analytical approach and after **Stürmer trimming**.

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

(a) Unadjusted



(b) IPTW Adjusted



Supplementary Figure 5. Unadjusted and IPTW adjusted cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **OT** analytical approach and after **Stürmer trimming.**

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

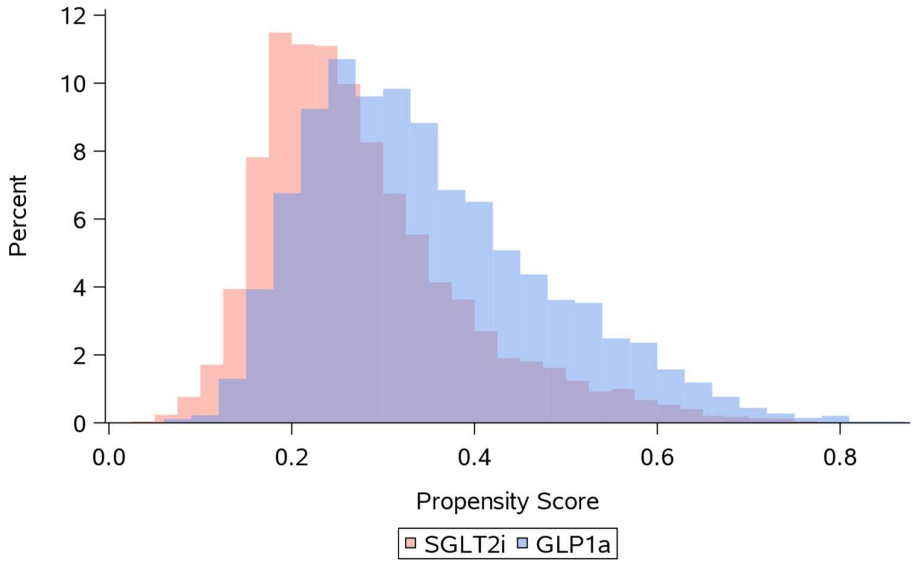
Post-hoc analyses

Restricting the study population to T2DM patients with obesity

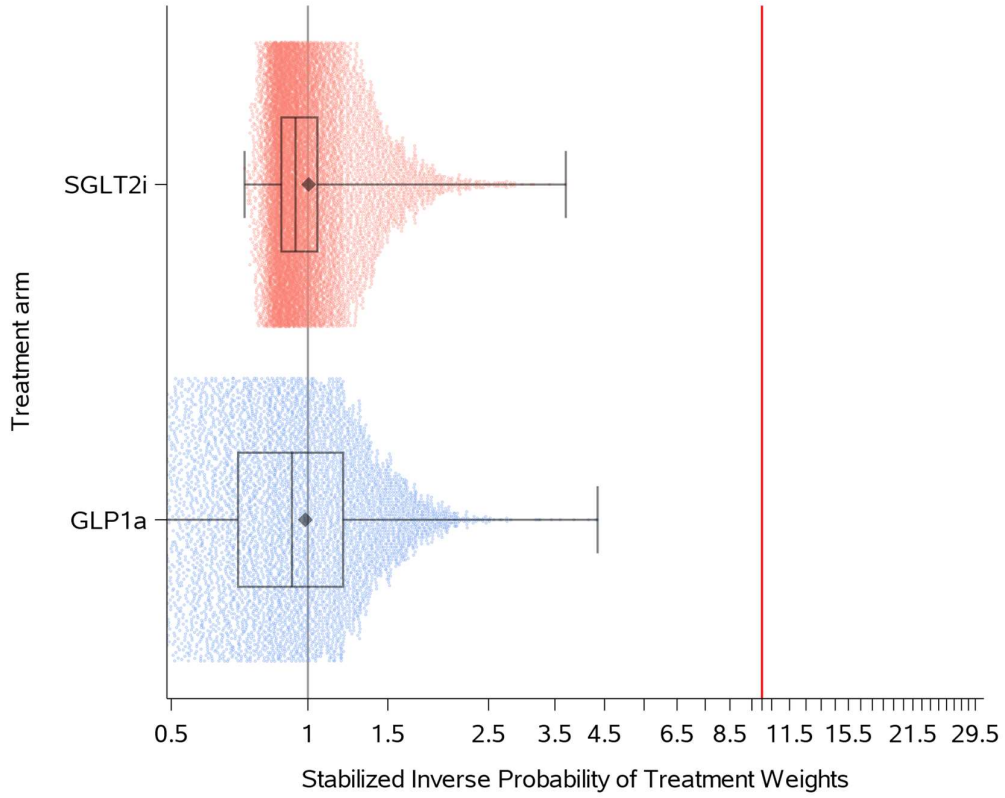
Supplementary Table 4. Baseline characteristics of the subpopulation of T2DM patients with obesity

	Before IPTW (unadjusted)	
	GLP-1 agonists	SGLT-2 inhibitors
Number of individuals	***5472***	***12815***
Calendar year at index-data, median (min-max)	2019 (2013-2023)	2020 (2013-2023)
Age at index-date (years), median (min-max)	58.0 (10.2-91.9)	59.6 (18.2-95.9)
Years between first T2DM diag. and index-date, median (min-max)	7.5 (0.0-50.1)	7.0 (0.0-63.4)
Multiple deprivation index, median (min-max)	4.0 (1.0-10.0)	4.0 (1.0-10.0)
Male, n (%)	2823 (51.6)	7399 (57.7)
Insulin use history, n (%)	1357 (24.8)	1458 (11.4)
Sulphonylureas use history, n (%)	3174 (58.0)	5564 (43.4)
Glitazone use history, n (%)	905 (16.5)	1357 (10.6)
DPP-4 inhibitor use history, n (%)	2673 (48.8)	5623 (43.9)
Glinide use history, n (%)	57 (1.0)	94 (0.7)
Other antidiabetic drug use history, n (%)	136 (2.5)	149 (1.2)
Obesity treatment history (within 2 y before index-date), n (%)	272 (5.0)	275 (2.1)
Anxiety history, n (%)	1360 (24.9)	2622 (20.5)
Anxiety treatment history, n (%)	1423 (26.0)	3140 (24.5)
Depression history, n (%)	2347 (42.9)	4824 (37.6)
Depression treatment history, n (%)	3279 (59.9)	6530 (51.0)
Schizophrenia history, n (%)	122 (2.2)	261 (2.0)
Dementia history, n (%)	26 (0.5)	74 (0.6)
Stress syndromes history, n (%)	224 (4.1)	457 (3.6)
Renal disease history, n (%)	909 (16.6)	1531 (11.9)
Myocardial infarction history (within 2 y before index-date), n (%)	93 (1.7)	236 (1.8)
Congestive heart failure history, n (%)	231 (4.2)	807 (6.3)
Cerebrovascular disease history, n (%)	202 (3.7)	532 (4.2)
Cancer history, n (%)	469 (8.6)	1171 (9.1)
Substance abuse history, n (%)	260 (4.8)	648 (5.1)
Current smoker, n (%)	1027 (18.8)	2327 (18.2)

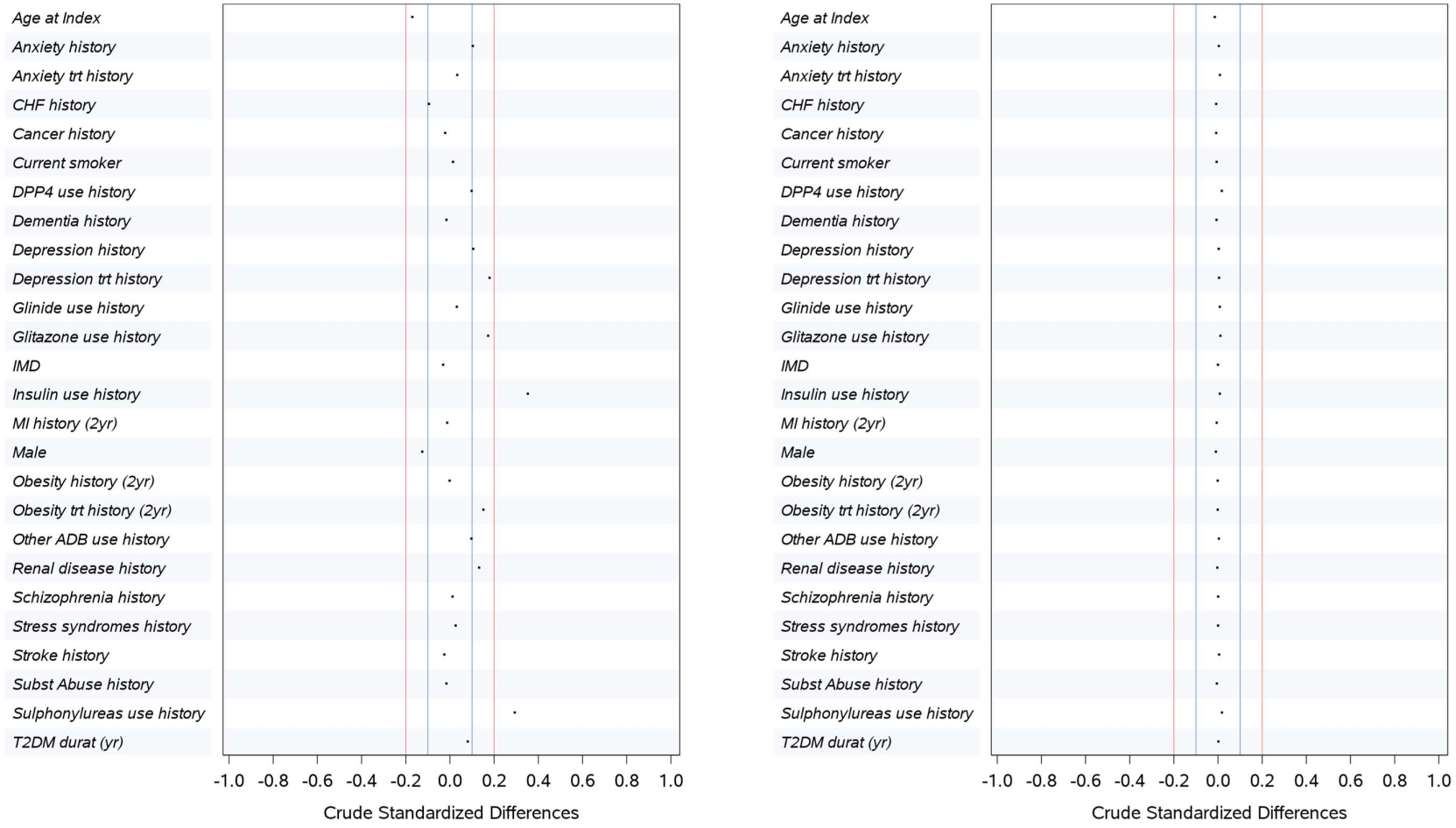
IPTW: inverse probability of treatment weighting.



Supplementary Figure 6. Propensity score distribution by treatment arm in the subpopulation of T2DM patients with obesity.



Supplementary Figure 7. Distribution of inverse probability of treatment weights by treatment arm in the subpopulation of T2DM with obesity. The red vertical line indicates a value of 10.



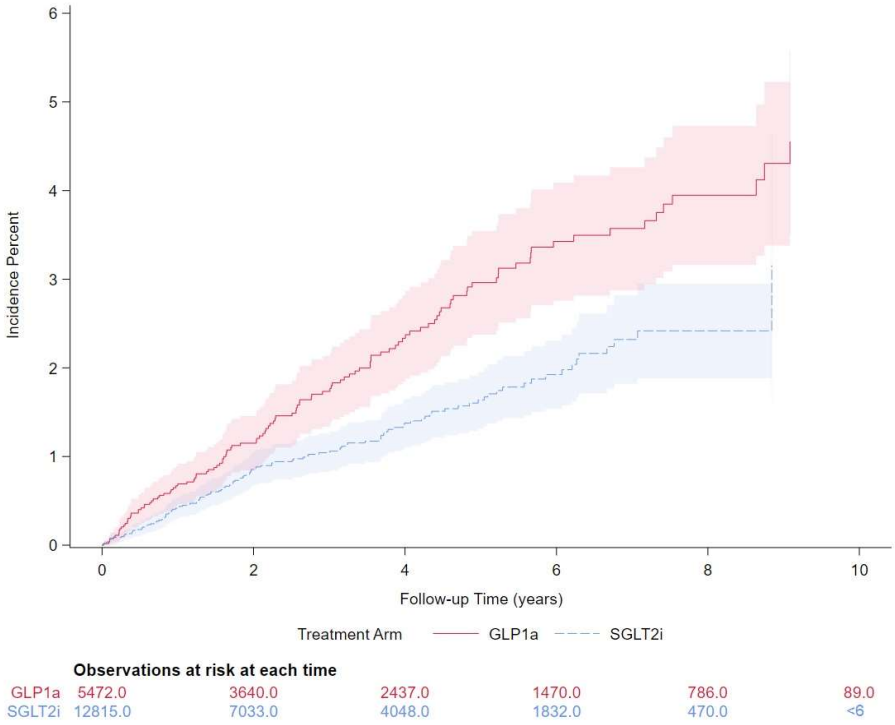
Supplementary Figure 8. Love plot of standardized mean differences between treatment arms for each baseline covariate before (left) and after (right) inverse probability of treatment weighting (IPTW) in the subpopulation of T2DM patients with obesity (n=18287). The blue vertical line indicates a standardized mean difference of 0.1 and the red line a difference of 0.2 (in either direction).

Supplementary Table 5. Unadjusted and adjusted incidence rates per 100 person- years and hazard ratios of composite endpoint (incl. both suicide-related and self-harm-related events) by treatment arm and analytical approach in the subpopulation of T2DM patients with obesity

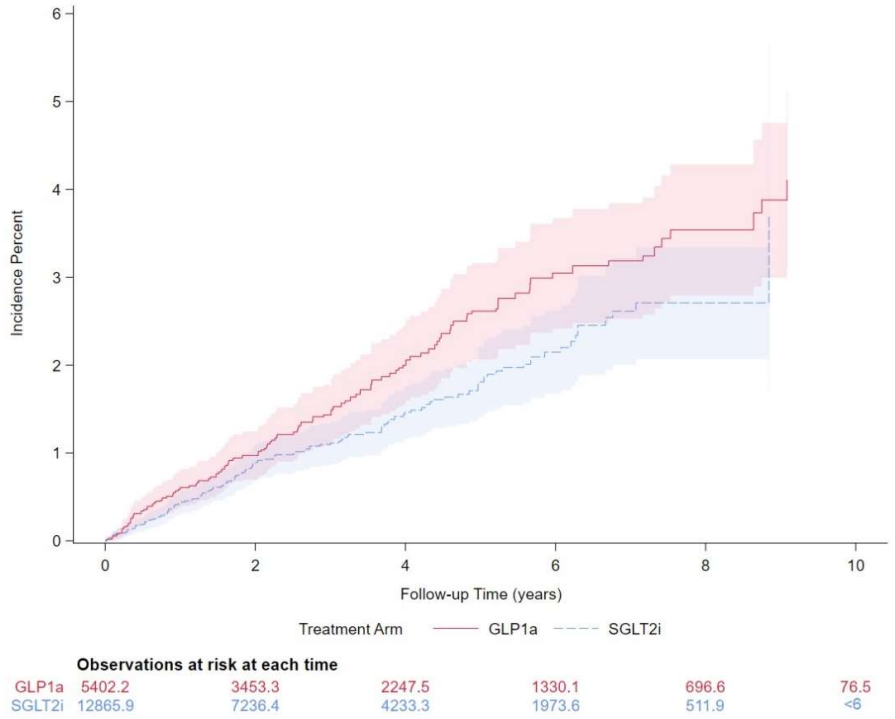
Treatment arm	ITT									OT							
	N	Follow-up (person-years)	n events	IR	95% CI		HR	95% CI		Follow-up (person-years)	n events	IR	95% CI		HR	95% CI	
Unadjusted analysis																	
GLP-1 receptor agonist	5472	22203.5	122	**0.55**	0.46	0.66	**1.66**	1.29	2.13	11069.2	58	**0.52**	0.40	0.68	**1.56**	1.11	2.19
SGLT-2 inhibitors	12815	38129.7	132	**0.35**	0.29	0.41	**1.00**	[Reference]		23917.6	81	**0.34**	0.27	0.42	**1.00**	[Reference]	
IPTW adjusted analysis																	
GLP-1 receptor agonist	5472	20882.5	101.2	**0.48**	0.40	0.58	**1.33**	1.02	1.72	10687.9	49.7	**0.46**	0.35	0.61	**1.31**	0.92	1.86
SGLT-2 inhibitors	12815	39365.5	147.2	**0.37**	0.31	0.45	**1.00**	[Reference]		24330.1	86.8	**0.36**	0.29	0.45	**1.00**	[Reference]	

ITT: intention-to-treat analytical approach, OT: On-treatment analytical approach; IR: incidence rate; 95% CI: 95% confidence interval; HR: Hazard ratio; IPTW: inverse probability of treatment weighting.

(a) Unadjusted



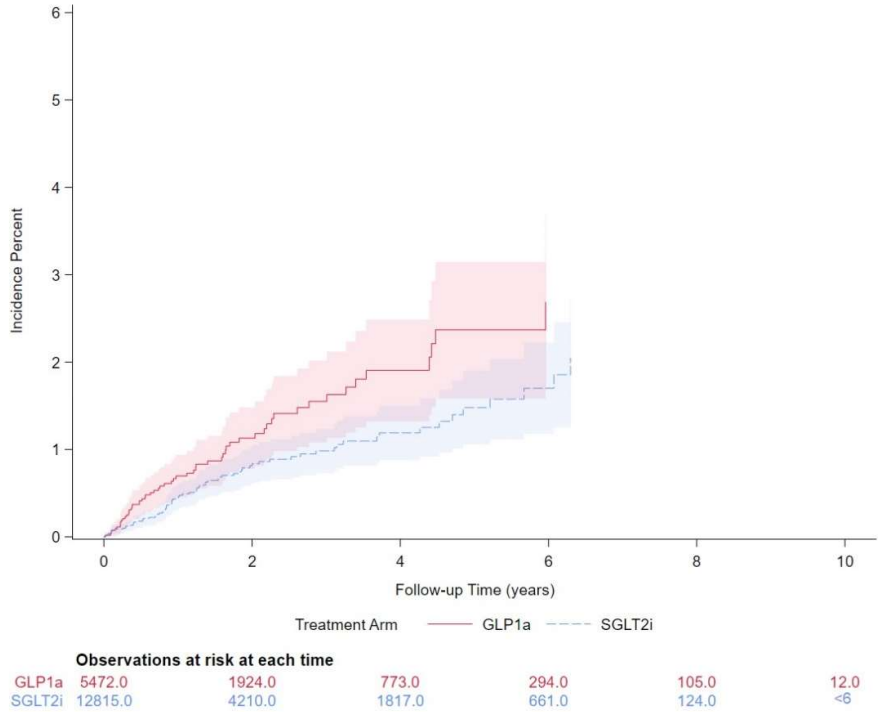
(b) IPTW Adjusted



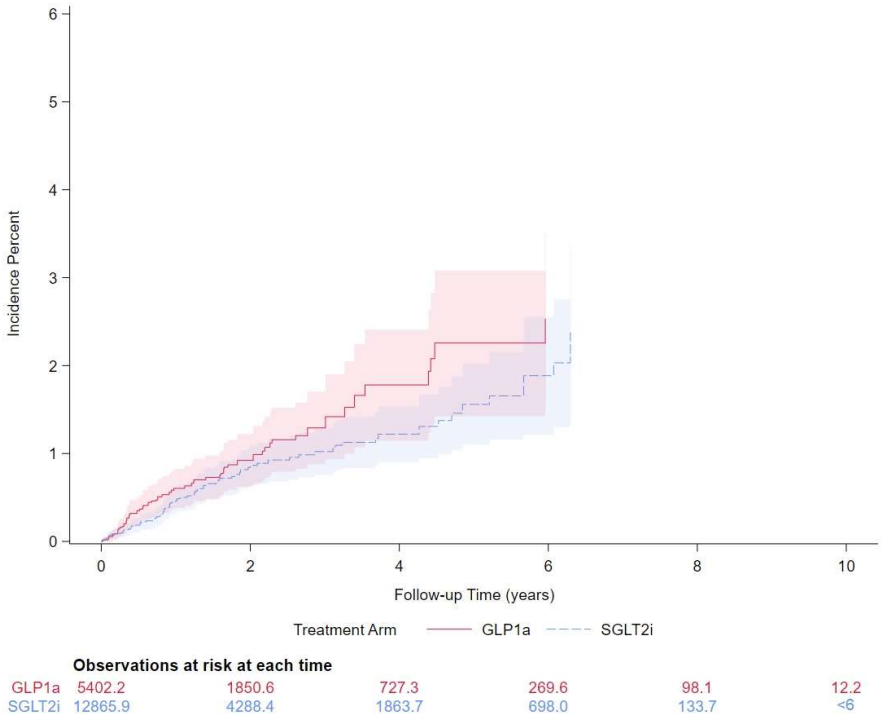
Supplementary Figure 9. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **ITT** analytical approach in the subpopulation of T2DM with obesity.

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

(a) Unadjusted



(b) IPTW Adjusted



Supplementary Figure 10. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **OT** analytical approach in the subpopulation of T2DM with obesity.

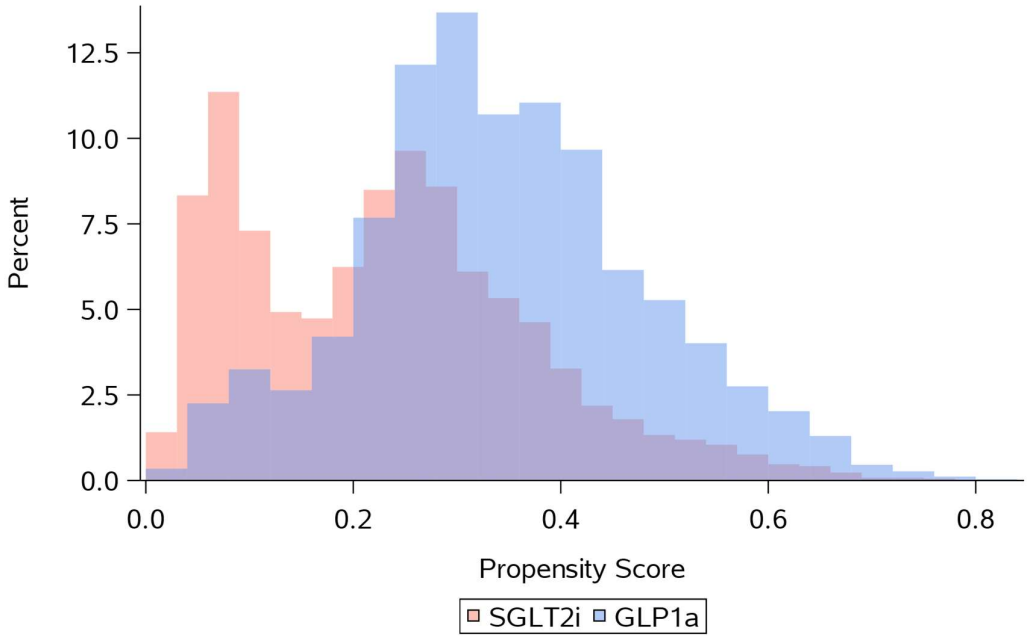
Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

Restricting the study population to those with history of depression

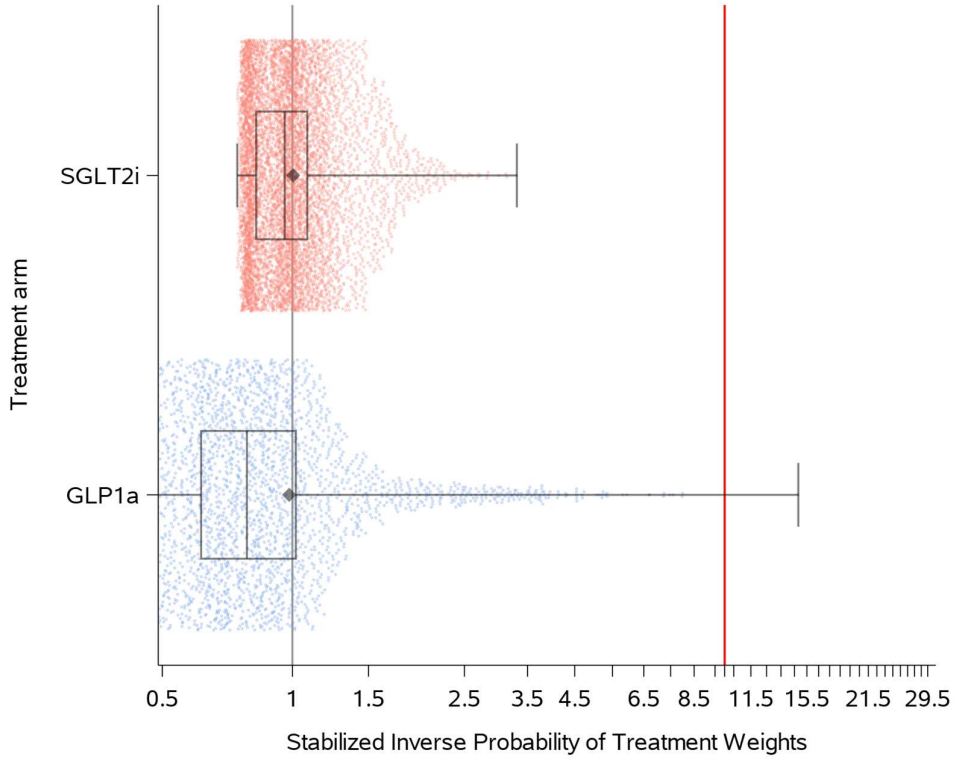
Supplementary Table 6. Baseline characteristics of the subpopulation of T2DM patients with history of depression

	Before IPTW (unadjusted)	
	GLP-1 agonists	SGLT-2 inhibitors
Number of individuals	***2617***	***7370***
Calendar year at index-data, median (min-max)	2019 (2013-2023)	2021 (2013-2023)
Age at index-date (years), median (min-max)	57.1 (18.4-88.2)	60.4 (16.3-100.2)
Years between first T2DM diag. and index-date, median (min-max)	7.6 (0.0-51.8)	8.0 (0.0-72.3)
Multiple deprivation index, median (min-max)	4.0 (1.0-10.0)	4.0 (1.0-10.0)
Male, n (%)	1146 (43.8)	3818 (51.8)
Insulin use history, n (%)	707 (27.0)	1087 (14.7)
Sulphonylureas use history, n (%)	1508 (57.6)	3550 (48.2)
Glitazone use history, n (%)	393 (15.0)	845 (11.5)
DPP-4 inhibitor use history, n (%)	1274 (48.7)	3437 (46.6)
Glinide use history, n (%)	26 (1.0)	83 (1.1)
Other antidiabetic drug use history, n (%)	68 (2.6)	108 (1.5)
Obesity history (within 2 y before index-date), n (%)	2347 (89.7)	4824 (65.5)
Obesity treatment history (within 2 y before index-date), n (%)	136 (5.2)	146 (2.0)
Anxiety history, n (%)	1126 (43.0)	2872 (39.0)
Anxiety treatment history, n (%)	932 (35.6)	2499 (33.9)
Depression treatment history, n (%)	2284 (87.3)	5799 (78.7)
Schizophrenia history, n (%)	94 (3.6)	269 (3.6)
Dementia history, n (%)	16 (0.6)	88 (1.2)
Stress syndromes history, n (%)	162 (6.2)	422 (5.7)
Renal disease history, n (%)	430 (16.4)	1065 (14.5)
Myocardial infarction history (within 2 y before index-date), n (%)	38 (1.5)	151 (2.0)
Congestive heart failure history, n (%)	100 (3.8)	493 (6.7)
Cerebrovascular disease history, n (%)	114 (4.4)	429 (5.8)
Cancer history, n (%)	233 (8.9)	737 (10.0)
Substance abuse history, n (%)	158 (6.0)	435 (5.9)
Current smoker, n (%)	577 (22.0)	1635 (22.2)

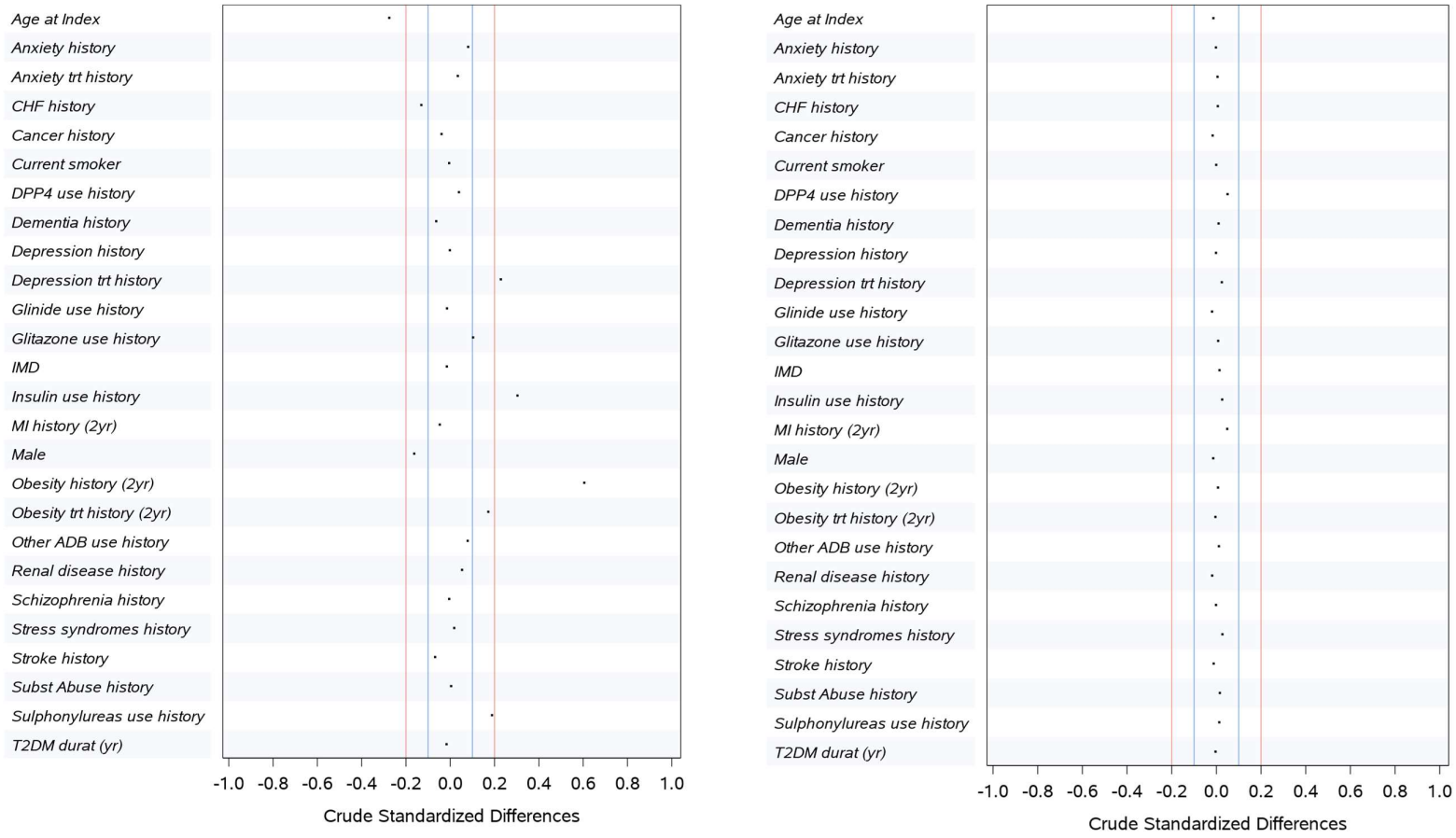
IPTW: inverse probability of treatment weighting.



Supplementary Figure 11. Propensity score distribution by treatment arm in the subpopulation of T2DM patients with history of depression.



Supplementary Figure 12. Propensity score distribution by treatment arm in the subpopulation of T2DM patients with history of depression.



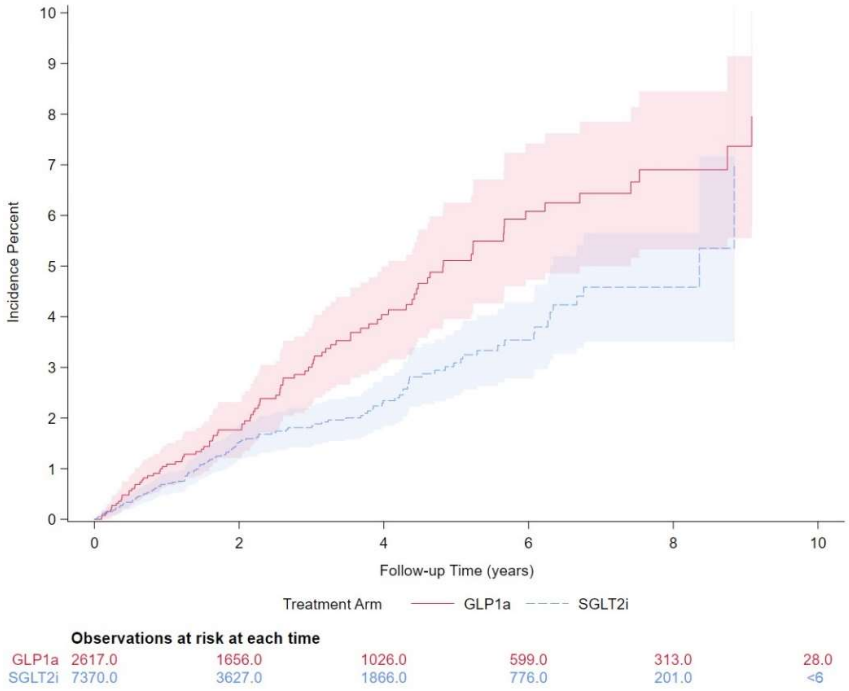
Supplementary Figure 13. Love plot of standardized mean differences between treatment arms for each baseline covariate before (left) and after (right) inverse probability of treatment weighting (IPTW) in the subpopulation of T2DM patients with history of depression. The blue vertical line indicates a standardized mean difference of 0.1 and the red line a difference of 0.2 (in either direction).

Supplementary Table 7. Unadjusted and adjusted incidence rates per 100 person- years and hazard ratios of composite endpoint (incl. both suicide-related and self-harm-related events) by treatment arm and analytical approach in the subpopulation of T2DM patients with history of depression

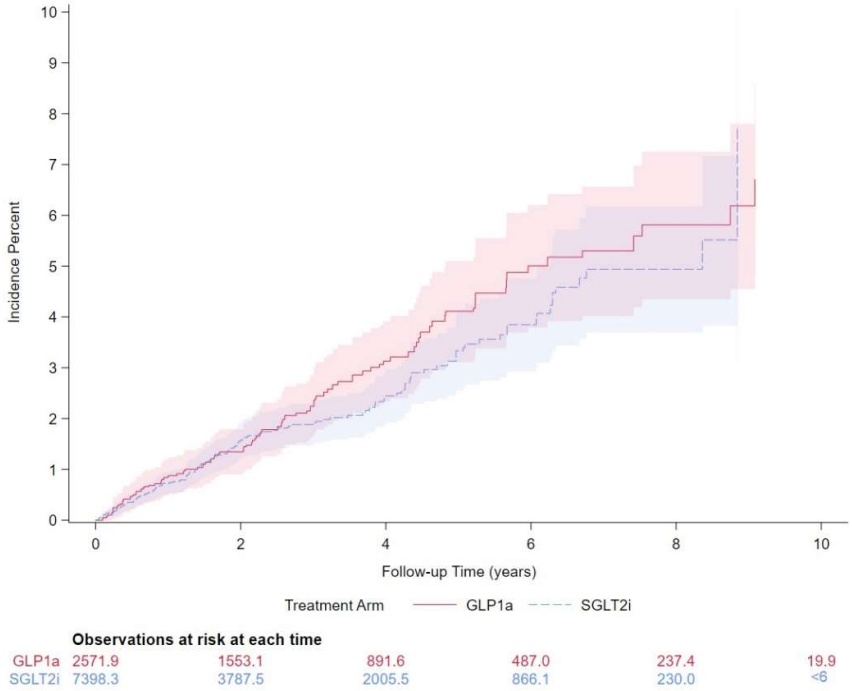
Treatment arm	ITT									OT							
	N	Follow-up (person-years)	n events	IR	95% CI		HR	95% CI		Follow-up (person-years)	n events	IR	95% CI		HR	95% CI	
Unadjusted analysis																	
GLP-1 receptor agonist	2617	9776.7	94	**0.96**	0.79	1.18	**1.52**	1.16	2.00	5016.4	43	**0.86**	0.64	1.16	**1.31**	0.90	1.91
SGLT-2 inhibitors	7370	19464.7	126	**0.65**	0.54	0.77	**1.00**	[Reference]		12153.9	80	**0.66**	0.53	0.82	**1.00**	[Reference]	
IPTW adjusted analysis																	
GLP-1 receptor agonist	2617	8864.8	68.9	**0.78**	0.63	0.96	**1.14**	0.86	1.52	4745.9	33.3	**0.70**	0.51	0.97	**1.02**	0.69	1.51
SGLT-2 inhibitors	7370	20360.1	138.9	**0.68**	0.57	0.82	**1.00**	[Reference]		12442.7	85.9	**0.69**	0.55	0.86	**1.00**	[Reference]	

ITT: intention-to-treat analytical approach, OT: On-treatment analytical approach; IR: incidence rate; 95% CI: 95% confidence interval; HR: Hazard ratio; IPTW: inverse probability of treatment weighting.

(a) Unadjusted



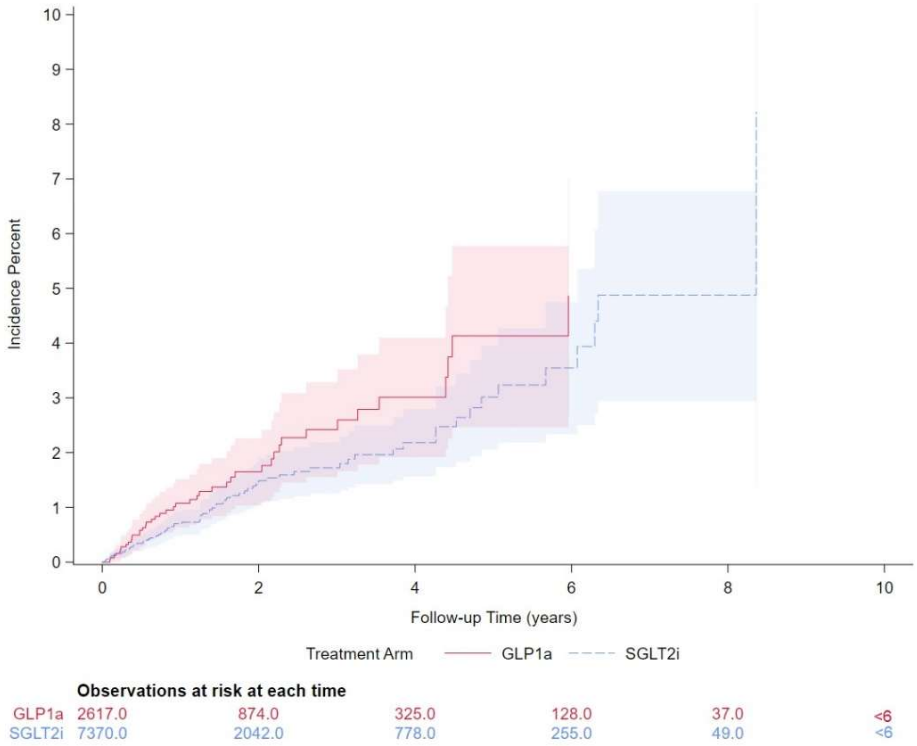
(b) IPTW Adjusted



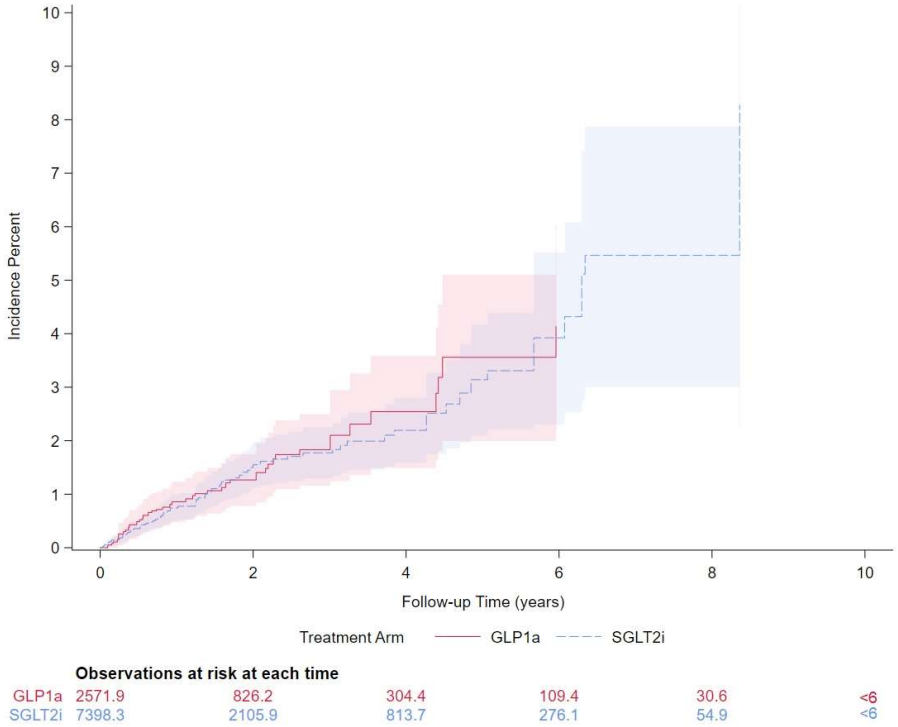
Supplementary Figure 14. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **ITT** analytical approach in the subpopulation of T2DM patients with history of depression.

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

(a) Unadjusted



(b) IPTW Adjusted



Supplementary Figure 15. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **OT** analytical approach in the subpopulation of T2DM patients with history of depression.

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.

Exploring shorter gap periods (i.e., 90 days) as a cut-off for treatment discontinuation

Supplementary Table 8. Unadjusted and adjusted incidence rates per 100 person- years and hazard ratios of **composite endpoint** (incl. both suicide-related and self-harm-related events) by treatment arm and gap period to define treatment discontinuation

Treatment arm	OT main analysis: 180 days*									OT additional analysis: 90 days**							
	N	Follow-up (person-years)	n events	IR	95% CI		HR	95% CI		Follow-up (person-years)	n events	IR	95% CI		HR	95% CI	
Unadjusted analysis																	
GLP-1 receptor agonist	6207	12445.4	60	**0.48**	0.37	0.62	**1.65**	1.20	2.26	9784.6	52	**0.53**	0.40	0.70	**1.81**	1.28	2.56
SGLT-2 inhibitors	20855	37364.5	110	**0.29**	0.24	0.35	**1.00**	[Reference]		29537.1	87	**0.29**	0.24	0.36	**1.00**	[Reference]	
IPTW adjusted analysis																	
GLP-1 receptor agonist	6103	11672.0	41.1	**0.35**	0.27	0.46	**1.11**	0.80	1.56	9246.3	36.0	**0.39**	0.29	0.52	**1.24**	0.86	1.79
SGLT-2 inhibitors	20923	38156.4	120.4	**0.32**	0.26	0.38	**1.00**	[Reference]		30160.7	94.1	**0.31**	0.25	0.39	**1.00**	[Reference]	

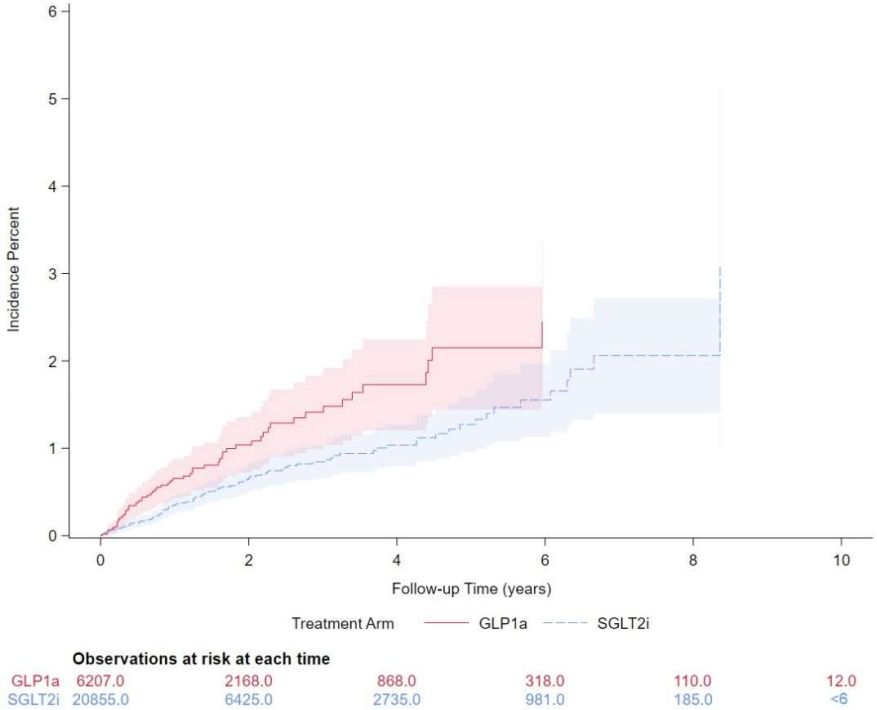
OT: On-treatment analytical approach; IR: incidence rate; 95% CI: 95% confidence interval; HR: Hazard ratio; IPTW: inverse probability of treatment weighting.

* In the main analysis (left), "on-treatment" follow-up considered the end of the baseline treatment 180 days after the last prescription followed by a gap longer than 180 days in the sequence of prescriptions and censoring it at this time.

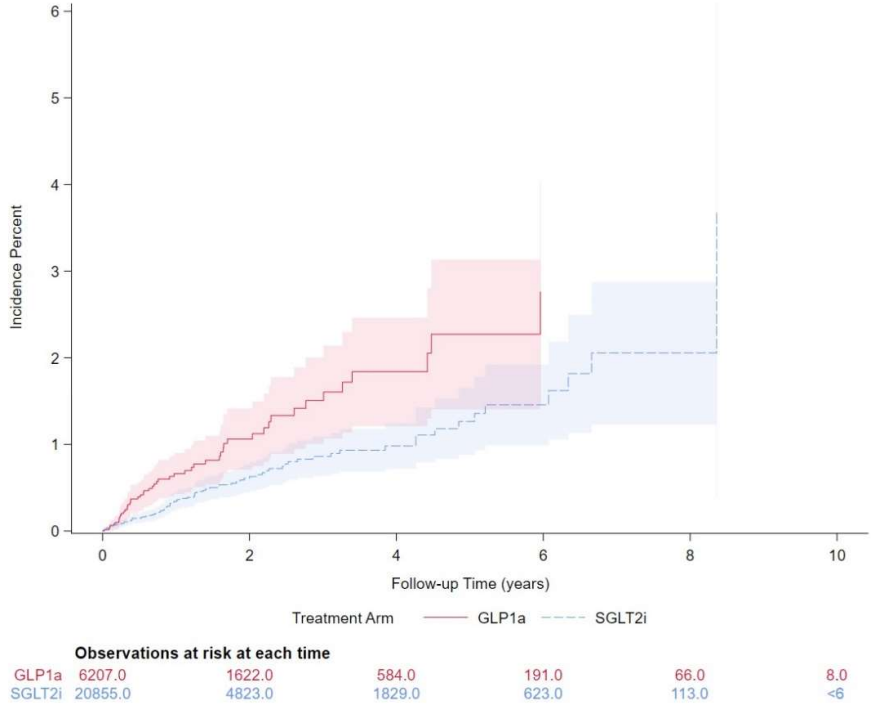
** In the additional analysis (right), "on-treatment" follow-up was modified by considering the end of the baseline treatment 90 days after the last prescription followed by a gap longer than 90 days in the sequence of prescriptions and censoring it at this time.

(a) Unadjusted

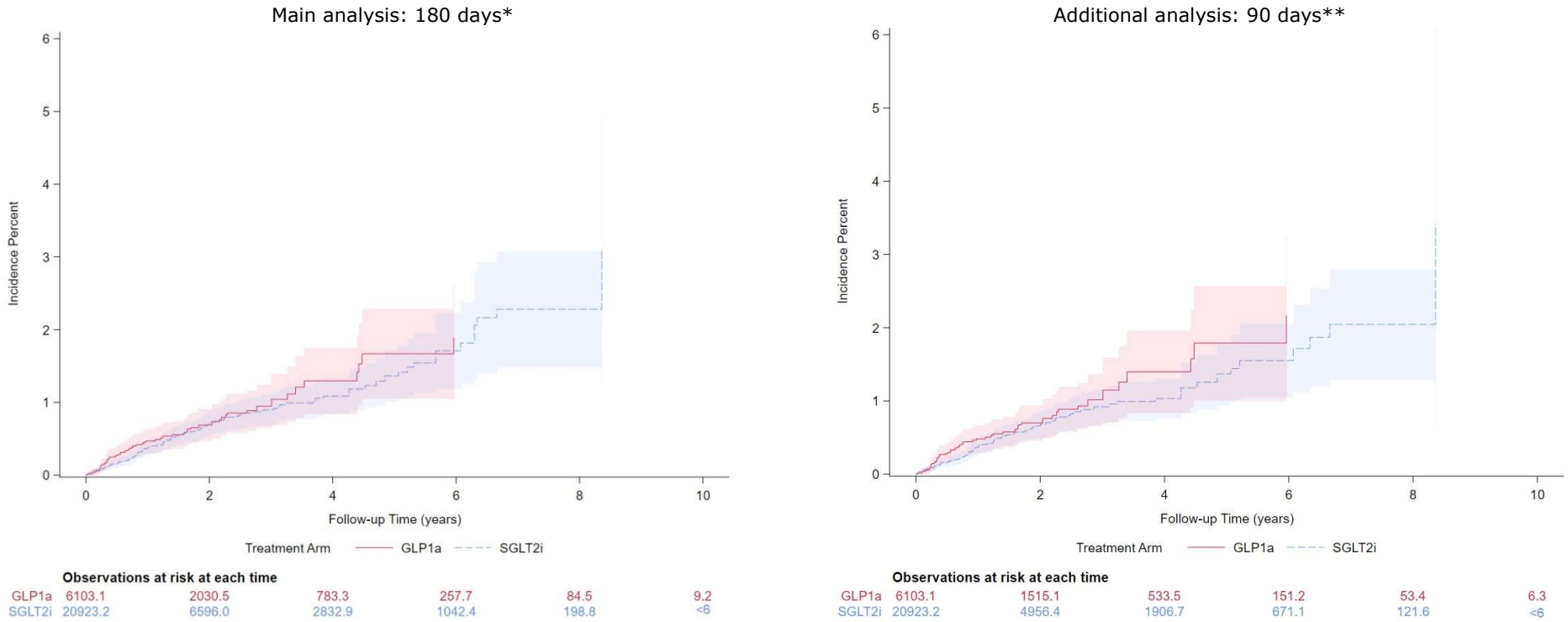
Main analysis: 180 days*



Additional analysis: 90 days**



(b) IPTW Adjusted

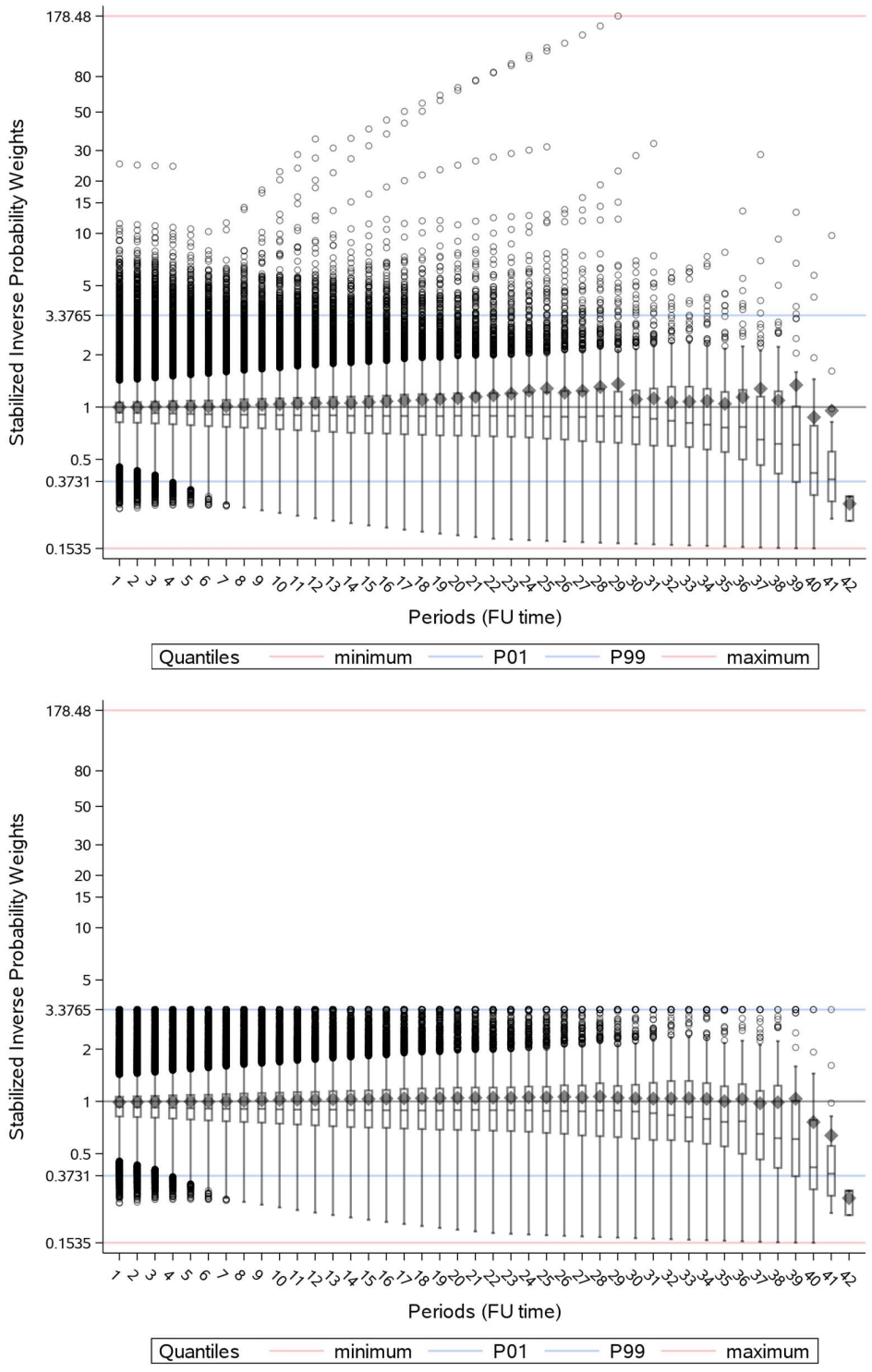


Supplementary Figure 16. Unadjusted (a) and IPTW adjusted (b) cumulative incidence (% and 95% CIs) of **composite endpoint** (incl. both suicide-related and self-harm-related events) based on **OT** analytical approach in the entire population of T2DM patients.

* In the main analysis (left), "on-treatment" follow-up considered the end of the baseline treatment 180 days after the last prescription followed by a gap longer than 180 days in the sequence of prescriptions and censoring it at this time.

** In the additional analysis (right), "on-treatment" follow-up was modified by considering the end of the baseline treatment 90 days after the last prescription followed by a gap longer than 90 days in the sequence of prescriptions and censoring it at this time.

Note: In accordance with IMRD database rules on the management of low cell counts, cells with low numbers (<6) in the IMRD database have been removed prior to publication of this report.



Supplementary Figure 17. Distribution of the inverse probability weights (IPWs) which is the product of the estimated inverse probability-of-treatment weights (IPTWs) and the inverse probability-of-censoring weights (IPCWs) (**top**) and after truncating extreme overall inverse probability weights

(**bottom**) to the 1st or 99th of the distribution. Box plots illustrate the distribution of weights within each episode. Diamonds represent the mean weight for each episode.

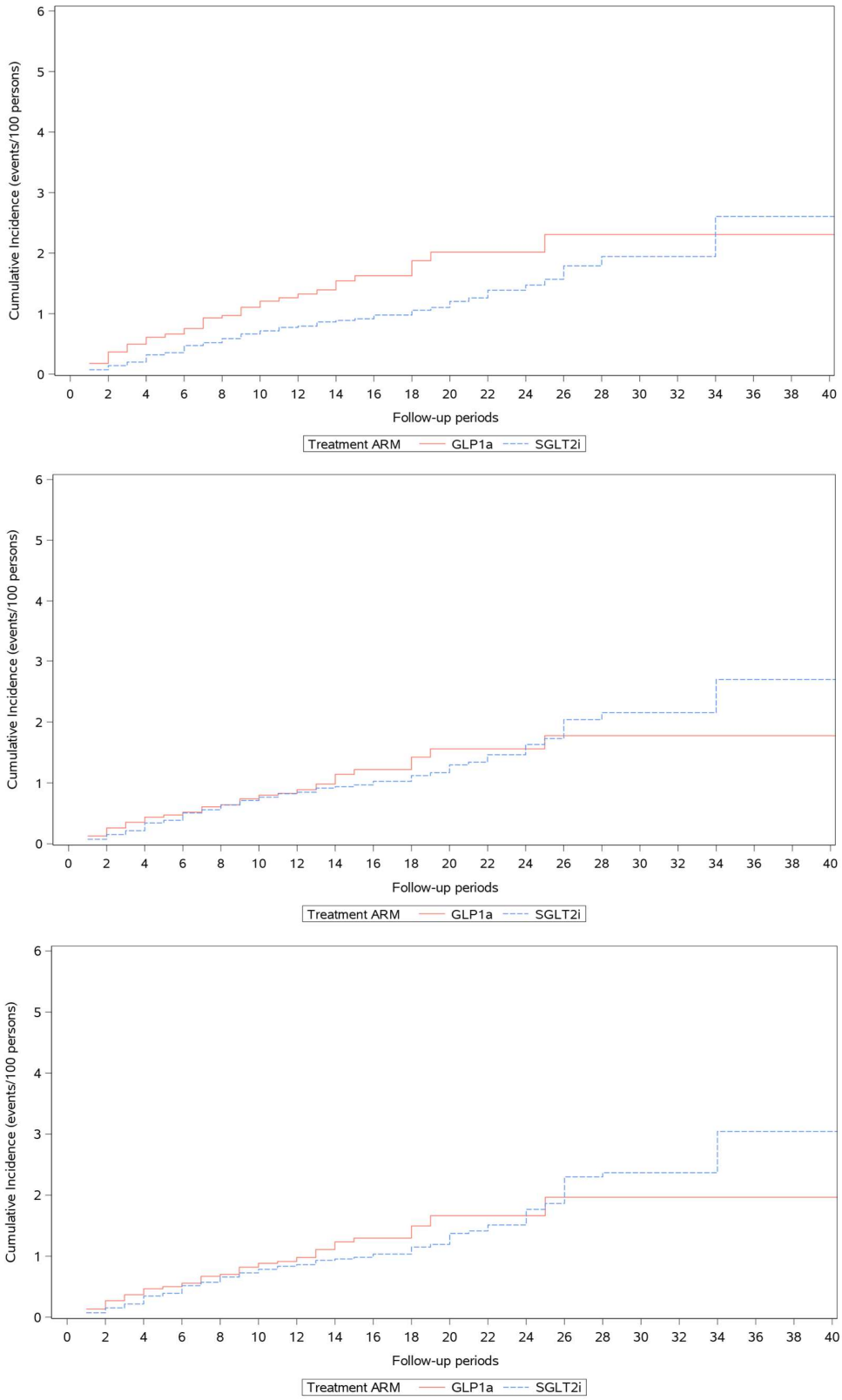
Supplementary Table 9. Unadjusted, IPTW adjusted and IPW adjusted (time-updated) pooled logistic odd ratios of composite endpoint (incl. both suicide-related and self-harm-related events) based on OT analytical approach and the entire population of T2DM patients.

Analysis	Treatment arm	OR	95% CI
Unadjusted	GLP-1 receptor agonist	1.67	1.22 2.29
	SGLT-2 inhibitors	1.00	[Reference]
IPTW adjusted (Baseline)	GLP-1 receptor agonist	1.12	0.80 1.57
	SGLT-2 inhibitors	1.00	[Reference]
IPW adjusted (Time-updated)	GLP-1 receptor agonist	1.16	0.83 1.63
	SGLT-2 inhibitors	1.00	[Reference]

IPTW: inverse probability of treatment weighting.

IPW: inverse probability of weighting (i.e., the product of treatment and censoring weights).

OR: Odds ratio; 95% CI: 95% Confidence intervals.



Supplementary Figure 18. Unadjusted (**top**), IPTW adjusted (**middle**) and IPW (time-updated) adjusted (**bottom**) cumulative incidence (% and 95% CIs) of composite endpoint (incl. both suicide-related and self-harm-related events) based on OT analytical approach in the entire population of

T2DM patients. An estimator similar to the Kaplan-Meier estimator used in the main analysis was applied, but discretizing time in 90-day episodes.

IPTW: inverse probability of treatment weighting. IPW: inverse probability weighting, i.e., the product of treatment and censoring weights.