Statistical and Epidemiological Analysis Plan

BI Study No.:	**1245.97**
Title:	Post-authorisation safety study to assess the risk of urinary tract malignancies in relation to empagliflozin exposure in patients with type 2 diabetes: a multi-database European study
Responsible statistician:	[redacted] and [redacted]
Responsible project epidemiologist:	[redacted] and [redacted] (principal investigator).
Date of statistical and epidemiological analysis plan:	26-JUL-2022
Version:	6.0

TABLE OF CONTENTS

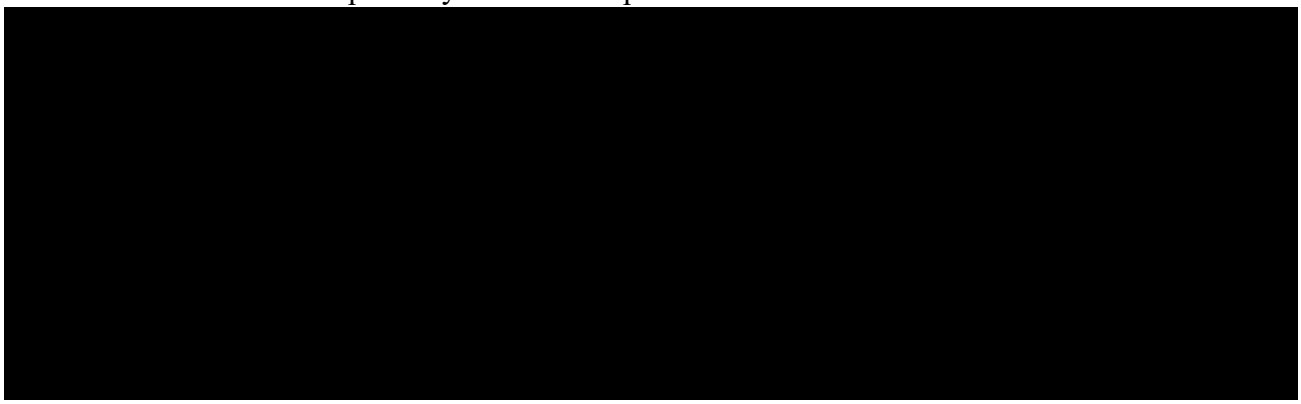
LIST OF ABBREVIATIONS

Term	Definition / description
AT	As-treated analysis
ATC	Anatomical Therapeutic Chemical Classification System
AvoHILMO	Register of Primary Health Care Visits (Finland)
BI	Boehringer Ingelheim International GmbH
BMI	Body mass index
CI	Confidence interval
CLR	Conditional logistic regression
COPD	Chronic obstructive pulmonary disease
CPRD-GOLD	Clinical Practice Research Datalink, General practitioner OnLine Data (United Kingdom)
CPRD Aurum	Clinical Practice Research Datalink, Aurum (United Kingdom)
DPP-4	Dipeptidyl peptidase-4
eGFR	Estimated glomerular filtration rate
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
FCR	Finnish Cancer Registry
FDA	Food and Drug Administration
FPR	Finnish Prescription Register
GLP-1	Glucagon-like peptide-1
GP	General practitioner
GPP	Good Pharmacoepidemiology Practices
GPV	Good Pharmacovigilance Practices
HbA1c	Glycated hemoglobin A1c
HES	Hospital Episode Statistics (United Kingdom)
HF	Heart failure
HILMO	Care Register for Health Care (secondary care; Finland)
HR	Hazard ratio
ICD-10	International classification of diseases

Term	Definition / description
IPTW	Inverse-probability of treatment weights
IR	Incidence rate
ITT	Intention-to-treat analysis
LISA	Longitudinal integration database for health insurance and labour market studies
MDRD	Modification of Diet in Renal Disease study
MI	Myocardial infarction
MSM	Marginal structural model
NCC	Nested case-control
NDR	National diabetes register (Sweden)
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NPR	National patient register (Sweden)
NSAID	Non-steroidal anti-inflammatory drugs
ONS	Office for National Statistics
PASS	Post-safety study
PH	Proportional hazards
PID	Personal identification number
PRAC	Pharmacovigilance Risk Assessment Committee
PS	Propensity score
PSUR	Periodic Safety Update Report
PY	Person-years
RMP	Risk management plan
RR	Relative risk
SCR	Swedish cancer registry
SEAP	Statistical/Epidemiological Analysis Plan
SF	Statistics Finland
SGLT-2	Sodium glucose co-transporter-2
SID	Study identification number
SII	Social Insurance Institution (SII)
SIR	Standardized incidence ratio

Term	Definition / description
SPDR	Swedish prescribed drug register
T2DM	Type 2 diabetes mellitus
THIN	The Health Improvement Network database
THL	National Institute for Health and Welfare (Finland)
UK	United Kingdom
US	United States
UT	Urinary tract
UTI	Urinary tract infection

LIST OF TABLES



1. INTRODUCTION

This statistical analysis plan refers to the final version of the pharmacoepidemiological study protocol entitled “Post-authorisation safety study to assess the risk of urinary tract malignancies in relation to empagliflozin exposure in patients with type 2 diabetes: a multi-database European study”, version 6.0 dated February 10th, 2022 unless otherwise stated.

This is an observational, comparative cohort safety study based on European healthcare databases and includes databases from the UK, Sweden and Finland.

The study will use a new user design to compare new users of empagliflozin to new users of DPP-4 inhibitors in order to assess the risk of urinary tract malignancies.

Individuals with similar treatment and clinical history will be matched by deriving propensity scores conditional on factors potentially affecting the treatment and outcome.

In the primary analysis, propensity-score (PS) matched cohorts will be derived from the study base by matching patients who initiate empagliflozin with patients who initiate DPP-4 inhibitors at index date with a variable matching ratio of up to 1:3 (1:1 matching in a sensitivity analysis) using greedy matching methods.

In addition, patients using metformin in free or fixed dose combination with empagliflozin or with a DPP-4 inhibitor will be identified and the primary analysis will be repeated within this subset of patients.

2. RESEARCH QUESTION AND OBJECTIVES

2.1. RESEARCH QUESTION

The aim of the study is to assess the risk of urinary tract malignancies in patients initiating empagliflozin compared to patients initiating a DPP-4 inhibitor.

2.2. OBJECTIVES

Primary objectives

i. To estimate the adjusted hazard ratio and incidence rates of all urinary tract cancers (bladder, renal and other urinary tract cancers) in patients initiating empagliflozin compared to patients initiating a DPP-4 inhibitor.

ii. To estimate the adjusted hazard ratios and incidence rates of bladder cancer in patients initiating empagliflozin compared to patients initiating a DPP-4 inhibitor.

iii. To estimate the adjusted hazard ratios and incidence rates of renal cancer in patients initiating empagliflozin compared to patients initiating a DPP-4 inhibitor.

Secondary objectives

i. To estimate the adjusted hazard ratio and incidence rates of each primary outcome (all urinary tract cancers, bladder cancer, renal cancer) with respect to:

   a. Increasing cumulative dosage to empagliflozin.
   b. Dosage of empagliflozin prescribed per use (10mg vs. 25mg).
   c. Time since first prescribed/dispensed dose of empagliflozin.

3. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

No.	Date	Section of the SEAP	Amendment or update	Reason
1	13-October-2017	Interim reports	Use of patient level data for the 2nd and 3rd interim reports instead of aggregated data.	PRAC assessment comments on the the 1st interim report.
2	29-NOV-2019	Multiple	Inclusion of new study country, Finland.	To increase of the size of the patient population
3	29-NOV-2019	Multiple	Using DPP-4 inhibitors as the primary comparison group and other SGLT-2 inhibitors as the secondary comparison group.	Comparison with other SGLT-2 inhibitors is expected to be underpowered due to the low uptake of SGLT-2 inhibitors other than empagliflozin. The number of patients initiating SGLT-2 inhibitors other than empagliflozin and verifying the inclusion and exclusion criteria in both study countries between August 1st, 2014 and November 30th, 2018 was 10,528 subjects. In contrast, the corresponding number of eligible patients initiating DPP-4 inhibitors was 3.8 times higher than the number of empagliflozin patients with 67,511 subjects.
4	29-NOV-2019	Multiple	Extending the study period until 31 December 2020.	Increasing the study sample size and accounting for the long latency period related to risk of malignancies
5	29-NOV-2019	Multiple	Investigation of patients counts in CPRD Aurum.	Based on the results of the PRAC feasibility evaluation, the number of persons available in CPRD-Aurum will be monitored.

6	13-APR-2022	Multiple	Removal of SGLT-2 inhibitors as a secondary comparator group	Low uptake of SGLT-2 inhibitors other than empagliflozin was observed in the 5th interim report. The number of patients initiating SGLT-2 inhibitors other than empagliflozin and verifying the inclusion and exclusion criteria during the study period was lower than the number of similar patients initiating empagliflozin in Sweden and Finland. It would therefore not be possible to carry out even 1:1 matching for empaglifozin users.
7	13-APR-2022	6.5	Change of the PS matching ratio to up to 1:3 mathcing in the main analysis and to 1:1 matching in a sensitivity analysis.	Up to 1:3 matching will allow an optimal use of available DPP-4 initiators. Using as many matches as possible in ech country will increase study precision and statistical power.
8	13-APR-2022	6.7.2	The description of the Crude, base and fullly adjusted models was aligned with the protocol. Additional criteria for selecting covariates for adjustement were introduced.	The incidence rates observed in the 5th interim report suggest there will be few events for outcomes such as renal cancer. With few events, it will not be possible to build a model including all selected confounders. The chosen method limits the covariates of the model to confounders which are not balanced in the matched sample.

9	13-APR-2022	Multiple	Pooled analysis of individual level data are replaced with aggregate level data meta-analyses. Pooled IRs will be estimated by pooling aggregate level data. Pooled HRs will be estimated using meta-analysis methods.	Individual level data from Finland and Sweden cannot be pooled together due to new regulations in the storage and use of individual level data in Sweden and Finland.
10	13-APR-2022	Multiple	Covariate selection criteria were added to mitigate the effect of missing data as well as potential Cox model convergence failures due to small number of events.	Observed incidence rates in interim report 5 suggests that there will be very few events observed for some outcomes such as renal cancer. In addition, stratified analyses that were to be carried out on pooled individual data will now be carried out at the single country level. It is therefore necessary to plan for possible convergence failure for some Cox proportional hazards models.
[redacted]				

12	13-APR-2022	Multiple	Specified algortihms for inclusion and exclusion criteria Redefined the AT and ITT exposure definitions as two time-fixed variables Adjusted the definitions of time at risk, latency period and treatment discontinuation. Specified algorithms for drug exposure definitions. Removed stratification of metformin use from stratified analyses. (Analyses in study groups have covered subgroups by metformin use) Changed some covariates from time-dependent to time-fixed	Modifications made to clarify the expected analyses for programmers.
13	13-APR-2022	Multiple	Sections and tables shells relating to analyses only relevant for interim reports were removed.	This SAP will be used for the final analysis only.
14	13-APR-2022	6.8.3	Added sensitivity analysis for grace period	The sensitivity analysis will assess the effect of lengthening the grace period to 90 days.
15	26 JUL 2022	Multiple	PS variable “database” added for UK CPRD Additional sensitivity analyses added for UK CPRD	The CPRD Aurum database was added to the study upon PRAC request.

16	26 JUL 2022	Multiple	Extending the study period until 31 December 2021 for UK CPRD-GOLD and Aurum.	Increasing the study sample size and accounting for the long latency period related to risk of malignancies

4. RESEARCH METHODS

4.1. STUDY DESIGN

This is a non-interventional, comparative, cohort safety study based on European healthcare databases and includes databases from the UK, Sweden and Finland. The databases for this study are constructed from linked prescription, hospital, general practitioner, cancer and death registration records.

The study will use an “incident users” design and compare new users of empagliflozin to new users of DPP-4 inhibitors. The index date will be defined as the date on which each identified new user receives the index prescription for empagliflozin, or a DPP-4 inhibitor.

To minimize channelling bias, individuals will be matched with similar treatment and clinical history at index date by deriving propensity scores (PS) conditional on factors affecting the treatment and outcome [1], [2]. Propensity-score-matched cohorts will be derived from the study base by matching patients who initiate empagliflozin with patients who initiate DPP-4 inhibitors at index date with the highest possible matching ratio (1:1, 1:2, or 1:3 using Greedy matching methods).

Depending on sample size, combination exposures with metformin will be accounted for through separate PS, stratification, or adjustment. Additional analyses will explore the presenceof detection/diagnostic bias as screening for urinary tract signs or symptoms may be more frequently among empagliflozin users. The potential for diagnostic bias will be evaluated and addressed through consideration of the frequency of urine dipstick testing (as captured through albuminuria tests) and stage of cancer at the time of diagnosis (as more early-stage cancers in one group would be suggestive of diagnostic bias).

The primary, secondary, and further analyses utilize a cohort design which will allow direct estimation of the IRs and aHRs of multiple outcomes of interest among new users of empagliflozin compared with new users of a DPP-4 inhibitor. The covariate information will be assessed during the time preceding treatment initiation (varying length look-back (pre-index) period) and during follow-up and will include all historical information available for each patient up until occurrence of outcome or censoring. Follow-up will start 6 months after the index date (12 months in sensitivity analyses) to account for an empirical induction/promotion period.

In the context of data sources such as the CPRD and the Swedish and Finnish national registers, the use of a cohort design has more advantages than limitations compared with the use of a nested case-control (NCC) design—see the appendix discussion in Schneeweiss (2010) [1] and Patorno et al. (2014) [2]. Thus, an observational cohort is constructed to study the risk of urinary tract malignancies among the two treatment groups. Several sensitivity analyses will be conducted including analyses using a NCC design (given the rarity of the outcomes) and analyses using marginal structural models (MSM) to control for time-varying confounderstowards switching or discontinuation of primary exposure.

The country-level datasets will be analysed separately. Country-level effect size estimates (IRs, aHRs, etc) will be pooled using suitable meta-analysis methods to compute overall estimates.

4.2. SETTING

4.2.1. STUDY POPULATIONS

The broad study population, in each of the study countries (see data sources in Table 1), will include individuals who have purchased at least one prescription of empagliflozin or of a DPP-4 inhibitor between 1st August, 2014 and 31st December, 2020 (31st December 2021 for UK-CPRD).

The Swedish prescribed drug register (SPDR) and the Finnish prescription register (FPR) will be used to identify the source population in Sweden and Finland. The CPRD-GOLD and CPRD Aurum data will be used to identify the source population in the UK [3]. Patients belonging to a practice having moved from GOLD to Aurum will be present in data extracted from both CPRD-GOLD and Aurum. These patients will be identified using a migrating practices file and their data will be removed from the extracted CPRD-GOLD data.Data from GOLD and Aurum will be combined and treated as a single data source.

ATC codes for empagliflozin and DPP-4 inhibitors are provided in Annex II.

4.2.2. STUDY PERIOD

The study will include individuals who have purchased at least one prescription of empagliflozin or of DPP-4 inhibitors from 1st August 2014 through 31st December 2020 (31st December 2021 for UK-CPRD data). The start of the study period corresponds to the time of drug licensing (second half of 2014 in all three countries) and subsequent introduction into clinical practice. The maximum follow-up time for an individual will be six years.

4.2.3. BASELINE AND LOOK-BACK PERIOD

To characterize the empagliflozin and DPP-4 inhibitor cohorts at the time of study drug initiation, all information available during the look-back (pre-index) time period will be collected.

The look-back time period is defined as the time period ending on the day before the index date (Table 1). Since all cohort members are required by inclusion criteria to have at least 1 year of data before the index date (baseline period), the look-back period will include at least 365 days during which covariates can be evaluated. For some study participants, data on covariates might be available beyond 1 year prior to the index date; in these cases, all available information will be considered for covariate classification related to diabetes, diabetes medications, and concomitant chronic conditions. For concomitant medications for diseases other than diabetes, the look-back time period will be limited to 180 days prior to the index date.

Table 1 Data sources with look-back periods and total data collection periods.

Data source	Data holder	Look-back period	Total data collection period
United Kingdom (UK)			
Clinical Practice Research Datalink, General practitioner OnLine Data (CPRD-GOLD)	Clinical Practice Research Datalink (CPRD)	1st Jan 1987– index date	1st Jan 1987 – 31st Dec 2021
Clinical Practice Research Datalink, Aurum (CPRD Aurum)	Clinical Practice Research Datalink (CPRD)	1st Jan 1995 – index date	1st Jan 1995 – 31st Dec 2021
Hospital Episode Statistics (HES)	Clinical Practice Research Datalink (CPRD)	1st Jan 1997 – index date	1st Jan 1997 – 31st Dec 2021
Sweden			
Swedish prescribed drug register (SPDR)[1]	National Board of Health and Welfare	1st Jul 2005 – index date	1st Jul 2005 – 31st Dec 2020
National patient register (NPR), both inpatient[2] and outpatient[3] registers	National Board of Health and Welfare	Inpatient: 1st Jan 1997 – index date Outpatient: 1st Jan 2001 – index date	Inpatient: 1st Jan 1997 – 31st Dec 2020 Outpatient: 1st Jan 2001 – 31st Dec 2020
Swedish cancer registry (SCR)[4]	National Board of Health and Welfare	1st Jan 1958 – index date	1st Jan 1958 – 31st Dec 2020
Cause of death register[5]	National Board of Health and Welfare	Not applicable	1st Jan 2014 – 31st Dec 2020
Total population register[6]	Statistics Sweden	1st Jan 2013 – index date	1st Jan 2013 – 31st Dec 2020
Longitudinal integration database for health insurance and labor market studies (LISA)[6]	Statistics Sweden	1st Jan 2013 – index date	1st Jan 2013 – 31st Dec 2020
National diabetes register (NDR)[7]	Centre of Registers, Region Västra Götaland	1st Jan 2012 – index date	1st Jan 2012 – 31st Dec 2020
Finland			

Finnish Prescription Register (FPR)[8]	Social Insurance Institution (SII)	1st Jan 1994 – index date	1st Jan 1994 – 31st Dec 2020
Finnish Registry for Reimbursed Medications[9]	Social Insurance Institution (SII)	1st Jan 1996 – index date	1st Jan 1996 – 31st Dec 2020
Care Register for Health Care (HILMO)[10]	National Institute for Health and Welfare (THL)	1st Jan 1996 – index date	1st Jan 1996 – 31st Dec 2020
Register of Primary Health Care Visits (AvoHILMO)[11]	National Institute for Health and Welfare (THL)	1st Jan 2011 – index date	1st Jan 2011 – 31st Dec 2020
Finnish cancer registry (FCR)[12]	Cancer Society of Finland	1st Jan 1953 – index date	1st Jan 1953 – 31st Dec 2020
Finnish Cause of Death Register[5]	Statistics Finland	Not applicable	1st Jan 2014 – 31st Dec 2020
Four regional EMR databases (Helsinki, Espoo, Vantaa, and Helsinki University Hospital [HUS] area)[6]	Cities of Helsinki, Espoo, and Vantaa, and Helsinki University Hospital	1st Jan 2013 – index date	1st Jan 2013 – 31st Dec 2020
Population register centre[6]	Population register centre	1st Jan 2013 – index date	1st Jan 2013 – 31st Dec 2020

[1] Data since the register was established in July 2005.
[2] Data since ICD-10 codes have been used (1997-).
[3] Data since the register was established (2001-).
[4] Data since the registry was established (1958-).
[5] Data since the start of the study period (2014-). However, data will also be extracted since one year before the start of the study period (2013-), as in few cases observations of dispensed study drugs defining the population may occur after death.
[6] Data since 1 year before the start of the study period (2013-).
[7] Data since 2 years before the start of the study period (2012-).
[8] Data since the register was established (1994-).
[9] Data since the register was established (1996-).
[10] Data since ICD-10 codes have been used (1996-).
[11] Data since the register was established (2011-).
[12] Data since the register was established (1953-).

4.2.4. INCLUSION AND EXCLUSION CRITERIA

The following inclusion/exclusion criteria will be applied to the source population in order to identify the study population in Sweden, Finland and the UK.

Table 2 Inclusion criteria.

Criteria	Definition	Data source

Diagnosis of type 2 diabetes	A diagnosis of type 2-diabetes mellitus (T2DM) identified using the country-specific algorithms detailed below. Relevant ICD-10 codes and READ diagnosis codes are provided in Annex I, and relevant ATC codes are provided in Annex II. **All available data prior to index date [2] inclusive.**	UK: CPRD-GOLD and Aurum Sweden: NDR, National Patient Register (NPR), and SPDR Finland: avoHILMO, FPR, and Reimbursement register
Age over 18 years at index date	Patient age over 18 years. **Evaluated at index date[2].**	UK: CPRD-GOLD and Aurum Sweden: SPDR Finland: FPR and Population register centre
At least one year of database membership	Difference between index date and database membership start ≥ 365 days. **Evaluated at index date[2]**	UK: CPRD-GOLD and Aurum Sweden: Total population register[1] Finland Population register centre
Inclusion criteria verified	All of the above.	

[1] The national patient register will be used if data on immigration/emigration are unavailable from the Total population register.
[2] The index date is defined as the date of first purchase/prescription of empagliflozin or a DPP-4 inhibitor during the study period (from August 1, 2014 until December 31, 2020, 2021 for CPRD).

The following country-specific algorithms will be used for the identification of a diagnosis of T2DM at or prior to index date date:

UK

A diagnosis of T2DM will be identified using relevant diagnosis codes (listed in Annex I) in CPRD-GOLD (clinical and therapy data) or CPRD Aurum (observation and problem data).

Sweden

Patients with a diagnosis of T2DM will be identified in following steps:

- Search for a diagnosis of T2DM in NDR data (diabetes registry records) using the corresponding variables that indicate the registered type of diabetes.

- If no T2DM diagnosis is present in NDR data, search for a diagnosis in NPR data (inpatient and outpatient records) using the relevant ICD-10 diagnosis codes listed in Annex I.

- If no T2DM diagnosis is present in NDR or NPR, search for a previous (before index date) prescription of metformin in SPDR data using the ATC codes listed in Annex II.

Finland

Patients with a diagnosis of T2DM will be identified in following steps:

- Search for a diagnosis of T2DM in the reimbursement register using ICD-10 diagnosis codes listed in Annex I.

- If no T2DM diagnosis is present in the reimbursement register, search for a main diagnosis in AvoHILMO data using relevant ICD-10 diagnosis codes listed in Annex I.

- If no T2DM diagnosis is found in the reimbursement register or in AvoHILMO, search for a previous (before index date) prescription of metformin using the ATC codes listed in Annex II.

Exclusion criteria

Table 3 Exclusion criteria.

Criteria	Definition	Data source
History of cancer prior to index date	A registered diagnosis of any cancer (excluding non-melanoma skin cancer) (ICD-10, ICD-7, ICD-O-3 and READ diagnosis codes in Annex I) **All available medical history prior to index date inclusive.**	UK: CPRD-GOLD and Aurum Sweden: Swedish cancer registry (SCR) Finland: Finnish cancer registry (FCR)
Diabetes diagnosis other than type 2	A diagnosis of type 1 diabetes mellitus (T1DM) or other non-type 2 diabetes identified using the country-specific identification algorithms described below. Relevant ICD-10 codes and READ diagnosis codes are provided in Annex I. **All available medical history prior to index date inclusive.**	UK: CPRD-GOLD and Aurum Sweden: NDR and NPR Finland: Reimbursment registry

SGLT-2 inhibitor or DPP-4 inhibitor use any time prior to index date	Any prescription of a SGLT-2 inhibitor drug (including empagliflozin) or a DPP-4 inhibitor drug, alone or in combination. (ATC codes in Annex II) **All available prescription history prior to index date exclusive.**	UK: CPRD-GOLD and Aurum Sweden: SPDR Finland: FPR
Use of fixed or free dose combinations of SGLT-2 inhibitors with DPP-4 inhibitors or empaglifozin.	Any prescription record of a SGLT-2 inhibitor drug in combination with a DPP-4 inhibitor drug. (ATC codes in Annex II) **All available prescription history prior to index date inclusive.**	UK: CPRD-GOLD and Aurum Sweden: SPDR Finland: FPR
End-stage renal disease diagnosis or receipt of renal dialysis any time prior to index date	Any registered record with a diagnosis or procedures of end-stage renal disease, receipt of renal dialysis or eGFR ≤ 15 (ml/min/1.73 m^2). (ICD-10 codes, procedure codes and READ diagnosis codes in Annex I) **All available medical history prior to index date inclusive.**	UK: CPRD-GOLD and Aurum Sweden: NPR and NDR Finland: HILMO and AvoHILMO
Patient excluded from the study population	Any of the above exclusion criteria verified prior to enrolment.	

The following country-specific algorithms will be used to identify patients with a diagnosis of any diabetes other than T2DM for exclusion criteria:

UK

Patients with a diagnosis of any diabetes other than T2DM will be identified in CPRD-GOLD (clinical and therapy data) or CPRD Aurum (observation and problem data) using relevant diagnosis codes (listed in Annex I).

Sweden

Patients with a diagnosis of any diabetes other than T2DM will be identified in following steps:

- Search for a diagnosis of any diabetes other than T2DM in NDR data using the corresponding variables that indicate the registered type of diabetes.

- For patients without a diagnosis of diabetes other than T2DM in NDR, search for a main diagnosis of T1DM in NPR data (inpatient and outpatient records) using the

relevant ICD-10 diagnosis codes listed in Annex I. A T1DM diagnosis will be confirmed and the patient is excluded if the latest event with a main diagnosis of T1DM is not succeeded by an event with a main diagnosis of T2DM within the look-back period.

Finland

Patients with a diagnosis of any diabetes other than T2DM will be identified in the reimbursement register using relevant ICD-10 diagnosis codes listed in Annex I.

5. VARIABLES

5.1. EXPOSURES

5.1.1. STUDY MEDICATIONS

Main Cohort:

In each country, after identifying the study cohort, patients will be grouped in two exposure groups:

- New users of empagliflozin: Patients prescribed empagliflozin at index date.
- New users of DPP-4 inhibitors: Patients prescribed any DPP-4 inhibitor at index date.

Sub-cohort of users of empagliflozin in combination with metformin

Patients using metformin in free or fixed dose combination will be identified in the main study cohort and grouped in two exposure groups for a separate analysis:

- Patients prescribed empagliflozin in free or fixed dose combination with metformin at index date.
- Patients prescribed any DPP-4 inhibitor in free or fixed dose combination with metformin at index date.

5.1.2. INDEX DATE

The index date is defined as the date of first dispensation/prescription of empagliflozin or of a DPP-4 inhibitor during the study period (from 1st August, 2014 until 31st December, 2020 – 2021 for UK-CPRD).

5.1.3. EXPOSURE TIME

Time at risk will start on index date and continue until first treatment discontinuation (see definition in Section 5.1.6). Continuous exposure will be defined as successive exposure periods for the study drugs separated by gaps of 30 days at most. Duration of exposure

episodes for individual prescriptions of study drugs will be derived based on package size, unit strength and daily dose.

To account for an empirical induction/promotion period of cancer, a latency period of 6 months (180 days) after the index date is applied. Patients with less than 6 months (180 days) of follow-up time will be excluded. In addition, patients with any diagnosis of any cancer within the latency period will be excluded. In a sensitivity analysis, the latency period will be extended to 12 months (365 days).

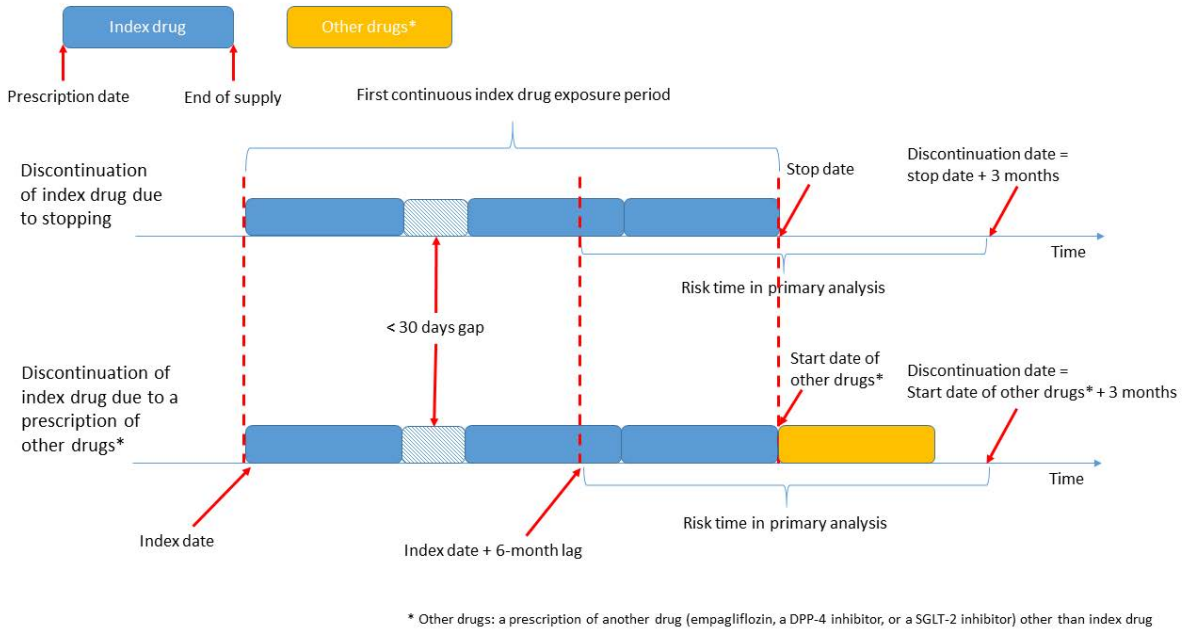


Figure 1 Definition of continous exposure period of index drug

For the primary AT analyses, the time at risk will begin after the latency period and end at urinary tract (bladder, renal or other) cancer or at any of the censoring events (listed in Table 4), whichever occurs first. In addition, treatment discontinuation of the index drug (defined in Section 5.1.6) is an additional censoring variable.

For the sensitivity ITT analyses, the time at risk will begin after the latency period and end at urinary tract (bladder, renal or other) cancer or any of the censoring events excluding treatment discontinuation of the index drug.

Patients with censored time at risk will not be able to re-enter the study cohort at a later time point. The time at risk will not be censored if oral or injectable glucose-lowering drugs other than the index drugs and SGLT-2 inhibitors are prescribed in addition to empagliflozin or a DPP-4 inhibitor after the index date.

Table 4 Censoring events.

Censoring event	Definition	Data source
All-cause mortality	Censoring of follow-up time occurs at the date of death regardless of the cause. **Applied in all analyses.**	UK: CPRD-GOLD and Aurum Sweden: Cause of death register Finland: Cause of death register
Emigration from the study country	Censoring of follow-up time occurs at the date of emigration from the study country. **Applied in all analyses.**	UK: CPRD-GOLD and Aurum Sweden: Total population register[1] Finland: Population register centre
Patient transfer	Censoring of follow-up time occurs at the date of patient transfer. **Applied in all analyses.**	UK: CPRD-GOLD and Aurum Sweden: NA Finland: NA
Incidence of any cancer during follow-up	A registered diagnosis of any cancer (excluding non-melanoma skin cancer) during follow-up. (ICD-10, ICD-7, ICD-O-3 and READ diagnosis codes in Annex I) **Applied in all analyses.**	UK: CPRD-GOLD and Aurum Sweden: Swedish cancer registry (SCR) Finland: FCR
Diabetes diagnosis other than type 2	A diagnosis of type 1 diabetes or other non-type 2 diabetes during follow-up identified using the same algorithm as used for the exclusion criteria (Section 4.2.4) Relevant ICD-10 codes and READ diagnosis codes are provided in Annex I. **Applied in all analyses.**	UK: CPRD and Aurum Sweden: NDR, NPR Finland: reimbursement registry
End of study period	Censoring of follow-up time occurs at end of study date i.e. 31 December 2020 (2021 for UK-CPRD). **Applied in all analyses.**	NA

Treatment discontinuation (AT exposure- main analysis)	Censoring of follow-up time occurs 90 days after - last day of medication (i.e. stop date) - new prescription of the other study drug or of a SGLT-2 inhibitor **Applied only in analyses using the AT exposure definition.**	UK: CPRD-GOLD and Aurum Sweden: SPDR Finland: FPR
Treatment discontinuation (AT exposure- sensitivity analysis)	Censoring of follow-up time occurs 180 days after - last day of medication (i.e. stop date) - new prescription of the other study drug or of a SGLT-2 inhibitor **Applied only in the sensitivity analysis using the AT exposure definition.**	UK: CPRD-GOLD and Aurum Sweden: SPDR Finland: FPR

[1] The National patient register will be used if data on immigration/emigration are unavailable from the Total population register.

5.1.4. Days' supply and duration of exposure

The estimated days' supply of each prescription/dispensing will be used to determine exposure episodes. Continuous exposure periods will then consist of consecutive exposure episodes separated by a gap of at most 30 days (Section 5.1.6).

The construction of exposure episodes is based on the assumptions that the total prescribed/dispensed dose is consumed as 1 unit per day, 1 defined daily dosage (DDD) per day, or 1 prescribed daily dose (PDD) per day. The construction of exposure episodes using PDD exihibited better performance and led to fewer outcome misclassifications, compared to the alternative approaches [4]. In this study, the length of each exposure episode (days' supply) will be calculated at the prescription level as the total prescribed/dispensed doses divided by the daily dose.

The total doses will be calculated as the total number of product units per prescription, using corresponding variables including the package size, unit strength and prescribed number of packages. The daily dose in product units will be determined based on the PDD (or recommended daily dose) and the unit strength in the prescribed packages. For study drugs, the daily doses will be ascertained from recommended dosages and pill strength. In case of missing information, the days' supply will be estimated as the median days' supply of the same drug prescription. The available variables will be extracted from the relevant databases, as described below.

UK CPRD-GOLD and Aurum

The total doses will be estimated as the total number of prescribed units using the follow variables provided by CPRD:

- Number of individual product packs prescribed
- Pack size
- Unit strength of the product

For study drugs and metformin, daily dose and days' supply will be derived from pill strength and dosage recommendations. For other drugs, the daily dose will be ascertained based on the prescribed daily dose (if available) or defined daily dose and the unit strength of the dispensed packages.

Sweden

The total doses will be calculated as the total number of dispensed units using the follow variables:

- Size of dispensed package
- Number of dispensed packages
- Drug strength

For study drugs and metformin, daily dose and days' supply will be derived from pill strength and dosage recommendations as described in Annex IV. For other drugs, the daily dose will be ascertained based on the defined daily dose and the unit strength of the dispensed packages.

Finland

The prescription information are available as purchase records, including

- ATC codes
- Number of packages prescribed
- Nordic Article Number (VNR).

VNRs will be mapped to retrieve information on the package size and unit strength of the products, using the publically available database from Finnish Medicines Agency (FIMEA). For study drugs and metformin, daily dose and days' supply will be derived from pill strength and dosage recommendations as described in Annex IV. For other drugs, the daily dose will be ascertained based on the defined daily dose and the unit strength of the dispensed packages.

5.1.5. Dose

Dose will be the current daily dose as described in Table 6.

The current daily dose will be derived from pill strength and dosage recommendations as described in Section 5.1.5.

5.1.6. Definition of continuous treatment and discontinuation

Continuous treatment

Continuous treatment is defined as having consecutive treatment episodes separated by a grace period of 30 days (Figure 1). In a sensitivity analysis, the grace period is extended to 90 days.

In case of overlaps between treatment episodes (i.e., a new prescription/dispensation during an ongoing exposure period), the overlapping days are added to the end of the following treatment episode, resulting in a longer treatment episode to take into account possible stockpiling.

Treatment discontinuation

Treatment discontinuation (see Table 4) is defined either by

- The start of a new prescription of any new drug from the following drug groups: empagliflozin, DPP-4 inhibitors, or SGLT-2 inhibitors), or of a fixed-dose combination of a SGLT-2 inhibitor with a DPP-4 inhibitor (identified from the ATC class A10BD)

- The end date of the first continuous exposure period of the index drug (stop date) plus a 3-month (90 days) period following it.

The discontinuation date is defined as the stop date/start date of a new drug plus 90 days.

5.1.7. Primary Exposure definition

The primary definition of exposure, i.e. as-treated exposure (AT), is defined as a continuous exposure to empagliflozin or to a DPP-4 inhibitor, starting from index date. Continuous exposure and treatment discontinuation are defined in Section 5.1.6.

In the Swedish dataset, drug exposure will be identified from SPDR. In Finland, the FPR will be used. In the UK datasets, drug exposure will be identified from CPRD data. The study drugs will be identified using ATC codes (med codes for CPRD) listed in Annex II.

Table 5 As-treated exposure definition

Exposure	Definition	Details	Data source
Exposure to index drugs (as treated)	Exposure status for empagliflozin (alone or in combination with metformin) or DPP-4 inhibitor (including combinations) based on purchased amounts and daily defined dose. ATC codes for empagliflozin and DPP-4 inhibitors in Annex II. **Evaluated at index date**	Time-fixed Categorical: Empagliflozin DPP-4 inhibitor	UK: CPRD Sweden: SPDR Finland: FPR

5.1.8. SECONDARY EXPOSURE DEFINTIONS

Exposure definitions based on dosage or duration

Table 6 Exposure definitions based on dosage or duration

Exposure	Definition	Details	Data source
Cumulative dose of empagliflozin use in grams	Time-updating cumulative sum of empagliflozin dose consumption based on the daily prescribed dose and purchased amounts since the start of follow-up. ATC codes for empagliflozin in Annex II. **Updating from index date onwards.***	Time-dependent Categorical[1]: Never 0 - <cumulative dose Q1 ≥ Q1 to < Q2 ≥ Q2 to < Q3 ≥ Q3	UK: CPRD Sweden: SPDR Finland: FPR
Daily dose of empagliflozin	The dosage of empagliflozin prescribed per daily use (10mg vs. 25mg). ATC codes for empagliflozin in Annex II. **Updating from index date onwards.***	Time-dependent Categorical: Never 10 mg 25 mg	UK: CPRD Sweden: SPDR Finland: FPR

Cumulative exposure duration to empagliflozin in months	Time-updating cumulative sum of duration since the first use of empagliflozin (alone or in combination with metformin), in months. ATC codes for empagliflozin in Annex II. **Updating from index date onwards.***	Time-dependent Categorical[2]: Never 0 - < cumulative duration Q1 ≥ Q1 to < Q2 ≥ Q2 to < Q3 ≥ Q3	UK: CPRD Sweden: SPDR Finland: FPR

*The last observed value will be carried forward for 3 months (90 days) to account for the 90-day delay between end of treatment and treatment discontinuation as defined in Section 5.1.6.

[1] The categories are defined based on the quartiles of overall cumulative dose (grams) of the study population

[2] The categories are defined based on the quartiles of overall cumulative exposure duration (months) of the study population

Intention-to-treat exposure definition

The intention to treat exposure definition assumes that individuals do not discontinue the treatment of an index drug.

Table 7 Intention-to-treat exposure definition

Exposure	Definition	Details	Data source
Exposure to index drugs (Intention to treat)	Purchase of empagliflozin (alone or in combination with metformin) or any DPP-4 inhibitors (including combinations). ATC codes for empagliflozin and DPP-4 inhibitors in Annex II. **Evaluated at index date.**	Time-fixed Categorical: Empagliflozin DPP-4 inhibitor	UK: CPRD Sweden: SPDR Finland: FPR

5.2. OUTCOMES

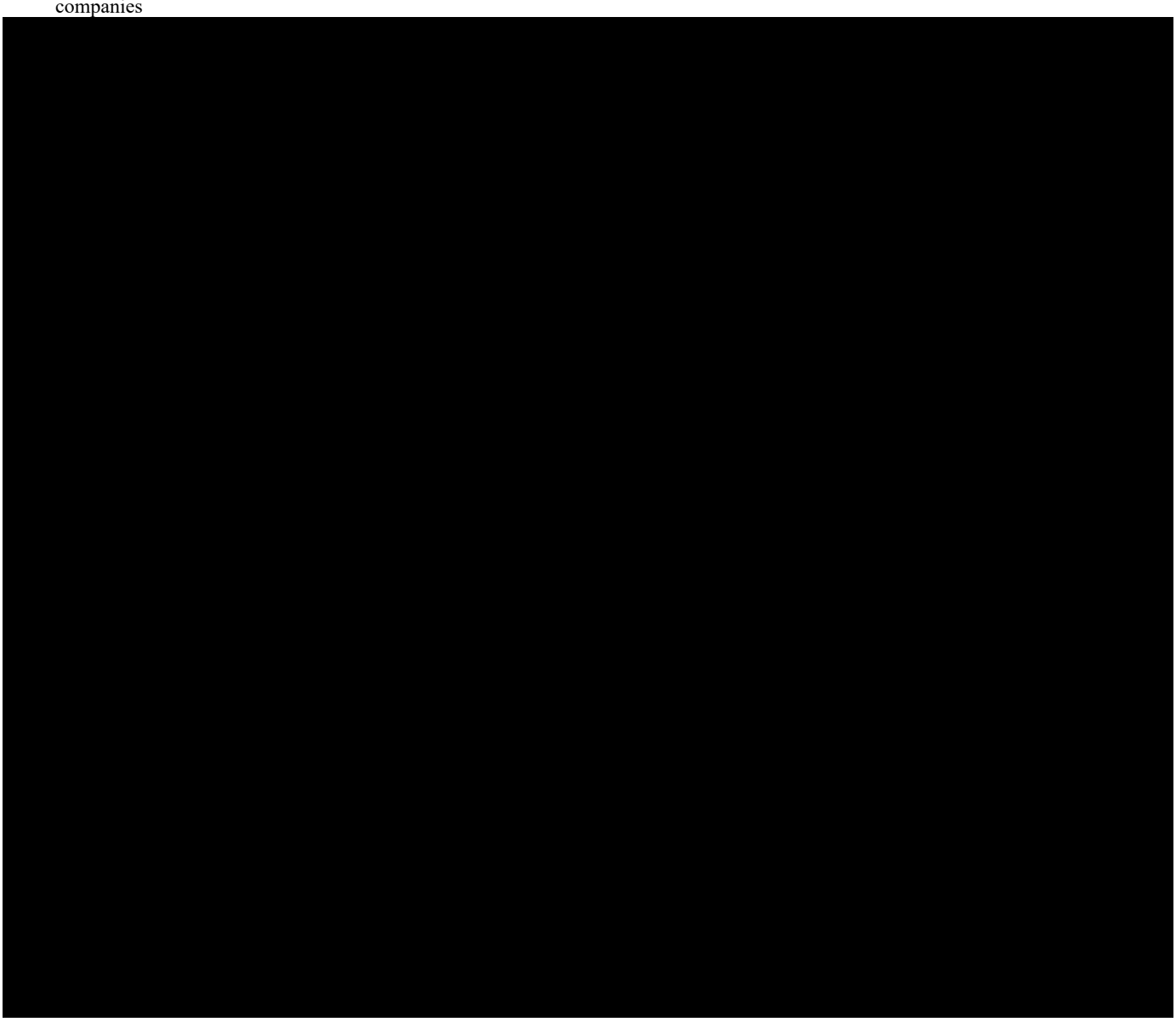
5.2.1. Primary outcomes

The outcomes of interest for this study are urinary tract cancers, bladder cancer, and renal cancer. These outcomes will be identified through Clinical Practice Research Datalink's General Practitioner OnLine Data (CPRD-GOLD and CPRD Aurum) data in the UK, the Swedish cancer registry (SCR) in Sweden and Finnish cancer registry (FCR) in Finland.

Table 8 Main outcomes

Outcome	Definition	Details	Data source
Urinary tract cancer (malignant or carcinoma in situ)	First record of urinary tract cancer (ICD-10, ICD-O-3 and READ diagnosis codes in Annex I). **During follow-up.**	Time to first event during follow-up.	UK: CPRD-GOLD and Aurum Sweden: SCR Finland: FCR
Bladder cancer (malignant and carcinoma in situ)	First record of bladder cancer (ICD-10, ICD-O-3 and READ diagnosis codes in Annex I). **During follow-up.**	Time to first event during follow-up.	UK: CPRD-GOLD and Aurum Sweden: SCR Finland: FCR
Renal cancer (malignant)	First record of renal cancer (ICD-10, ICD-O-3 and READ diagnosis codes in Annex I). **During follow-up.**	Time to first event during follow-up.	UK: CPRD-GOLD and Aurum Sweden: SCR Finland: FCR

The date of diagnosis of the first incidence of urinary tract cancers as specified in CPRD-GOLD and CPRD Aurum, in the SCR or in the FCR after the entry into the study will be used as the primary outcome date.
In the Swedish and Finnish datasets, urinary tract cancer events will be identified from the SCR and FCR using ICD-10 and ICD-O-3 codes (Annex I).
In CPRD, urinary tract cancer events will be identified in CPRD-GOLD using READ codes and in CPRD Aurum using READ and SNOMED CT codes (Annex I).



5.3. COVARIATES

5.3.1. Demographics and other baseline variables

Table 11 Demographics and other baseline variables.

Variable	Definition	Details	Data source
Age at index	Patient age at index. **Evaluated at index date.**	Time-fixed Continuous Categorical: 18-49 50-64 65-69 70-74 75-79 80-84 85 and over	UK: CPRD Sweden: SPDR Finland: FPR, Population register centre
Sex	Patient sex. **Evaluated at index date.**	Time-fixed Categorical: Female Male	UK: CPRD Sweden: SPDR Finland: Population register centre
Highest level of education	Patient highest level of education **Evaluated for the calendar year of the index date.**	Time-fixed Categorical: Compulsory Upper secondary Post secondary	UK: N/A Sweden: Longitudinal integration database for health insurance and labor market studies (LISA) Finland: Statistics Finland

Disposable income, individual at index	Patient available income. **Evaluated for the calendar year of the index date.**	Time-fixed Continuous Categorical:[1] < Q1 ≥ Q1 – < Q2 ≥ Q2 – < Q3 ≥ Q3	UK: N/A Sweden: LISA Finland: N/A
Disposable income, family at index	Patient family income. **Evaluated for the calendar year of the index date.**	Time-fixed Continuous Categorical:[1] < Q1 ≥ Q1 – < Q2 ≥ Q2 – < Q3 ≥ Q3	UK: N/A Sweden: LISA Finland: N/A
Socioeconomic summary variable	Socioeconomic status based on area of residence. **Evaluated at index date.**	Time-fixed Categorical:	UK: CPRD Sweden: N/A Finland: N/A
CPRD database	Source of CPRD data **Evaluated at index date.**	Time-fixed Categorical: GOLD Aurum	UK: CPRD Sweden: N/A Finland: N/A
Calendar year of index date	Calendar year of index date. **Evaluated at index date.**	Time-fixed Categorical: 2014 2015 2016 2017 2018 2019 2020 2021	UK: CPRD Sweden: SPDR Finland: FPR

Duration of look-back period (in years)	Time difference between index date and database membership start. **Evaluated at index date.**	Time-fixed Continuous Categorical: 1 2 – 5 6 – 9 10 and over	UK: CPRD Sweden: Total population register[2] Finland: Population register centre
BMI	Patient BMI value at index. **Evaluated at index date using closest value.**	Time-fixed Continuous Categorical:[2] <20 ≥20-24.9 25-29.9 >30	UK: CPRD Sweden: NDR Finland: AvoHILMO, EMR databases
Smoking status	Patient smoking status at index date. **Evaluated at index date using closest value.**	Time-fixed Categorical: Yes[3] No	UK: CPRD Sweden: NDR Finland: AvoHILMO, EMR databases
Alcohol use status	Patient alcohol use status at index date. **Evaluated at index date using closest value.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NA Finland: NA
HbA1c value (in mmol/mol)	Time-dependent HbA1c measurement value in mmol/mol. **Initialized at index using closest value to index date and updating from index date onwards.**	Time-dependent Continuous Categorical: < 48 (6.5%) ≥48 (6.5 %)-<64 (8 %) ≥64 (8%)	UK: CPRD Sweden: NDR Finland: EMR databases

Number of albuminuria tests	The number of albuminuria tests received by the patient prior to index date. **Evaluate at index using all available history prior to index date.**	Time-fixed Continuous	UK: CPRD Sweden: NDR Finland: EMR databases

[1] The national patient register will be used if data on immigration/emigration are unavailable from the Total population register.
[2] Categories under-weight (<20), normal weight (20-24.9), overweight (25-29.9), and obese (>30) may be adjusted based on the observed distribution of the data.
[3] In Sweden, a smoker is a patient smoking one or more cigarettes per day or a pipe daily or who had stopped smoking within the past 3 months. If data allows, categories current, never, past will be used.

5.3.2. Concomitant diseases

Table 12 Diabetic complications variables.

Variable	Definition	Details	Data source
Time since first diabetes diagnosis at index date	Time difference between [index date] and [first type 2 diabetes diagnosis date]. ICD-10 and READ diagnosis codes for type 2 diabetes in Annex I. **Evaluated at index date.**	Time-fixed Continuous Categorical: < Q1 ≥ Q1-< Q2 ≥ Q2-< Q3 ≥ Q4	UK: CPRD Sweden: NDR Finland: HILMO, AvoHILMO
Duration of treated diabetes at index date	Time difference between [index date] and [first antidiabetic treatment prescription/purchase date] in the data. ATC codes for antidiabetic treatment in Annex II. **Evaluated at index date.**	Time-fixed Continuous	UK: CPRD Sweden: SPDR Finland: FPR
Diabetic complications at index date	The variable is evaluated at index date and indicates the number of recorded diagnoses of diabetic complications using all available history prior to index date. ICD-10, READ diagnosis codes and corresponding NDR variables for diabetic complications in Annex I. **Evaluated at index using all available history prior to index date.**	Time-fixed Continous Categorical: 0 1-2 >2	UK: CPRD Sweden: NPR NDR Finland: HILMO, AvoHILMO

Diabetic retinopathy or diabetic maculopathy at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of diabetic retinopathy or diabetic maculopathy using all available history prior to index date. ICD-10, READ diagnosis codes and corresponding NDR variables for diabetic retinopathy or diabetic maculopathy in Annex I. **Evaluated at index using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR NDR Finland: HILMO, AvoHILMO
Diabetic nephropathy at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of diabetic nephropathy using all available history prior to index date. ICD-10 and READ diagnosis codes for diabetic nephropathy in Annex I. **Evaluated at index using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
Diabetic neuropathy at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of diabetic neuropathy using all available history prior to index date. ICD-10 and READ diagnosis codes for diabetic neuropathy in Annex I. **Evaluated at index using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
Peripheral vascular disease at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of peripheral vascular disease using all available history prior to index date. ICD-10 and READ diagnosis codes for peripheral vascular disease in Annex I. **Evaluated at index using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO

Diabetic lower limb severe complications at index date	The variable is initialized at index date and indicates a register record with a diagnosis of diabetic lower limb severe complications using all available history prior to index date. ICD-10 and READ diagnosis codes for diabetic lower limb severe complications in Annex I. **Evaluated at index using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO

[1] Categories to be defined based on the observed distribution of the data.

Table 13 Other non-diabetic concomitant disease variables.

Variable	Definition	Details	Data source
Kidney or genitourinary stones	The variable is initialized at index date and indicates a register record with a diagnosis of kidney or genitourinary stones using all available history. If no diagnosis prior to index date, it is time-updated to indicate any diagnosis of kidney or genitourinary stones after index date. ICD-10 and READ diagnosis codes for kidney or genitourinary stones in Annex I. **Initialized at index and updating from index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
Estimated glomerular filtration rate (eGFR) at index date	eGFR estimate derived from serum creatinine measurement using the CKD-Epi equation: eGFR (mL/min/1.73 m^2) = 141 x $\min(S_{cr}/K,1)^a$ x $\max(S_{cr}/K,1)^{-1.209}$ x 0.993^{Age} x 1.018 [if female]. a = -0.329 (females) or -0.411 (males). **Evaluated at index using closest value to index date.**	Time-fixed Continuous Categorical <15 (ml/min/1.73 m2) (kidney failure) 15-29 (ml/min/1.73 m2) (severely decreased) 30-44 (ml/min/1.73 m2) (Moderately to severely decreased) 45-59 (ml/min/1.73 m2) (Mildly to moderately decreased) 60-89 (ml/min/1.73 m2) (Mildly decreased) ≥90 (ml/min/1.73 m2) (Normal)	UK: CPRD Sweden: NDR Finland: EMR databases

Renal impairment at index date	The variable is evaluated at index date and indicates an eGFR < 60 (ml/min/1.73 m2) in all available history prior to index date. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical Yes No	UK: CPRD Sweden: NDR Finland: EMR databases
UTI or pyelonephritis	The variable is initialized at index date and indicates a register record with a diagnosis of UTI or pyelonephritis using all available history. If no diagnosis prior to index date, it is time-updated to indicate any diagnosis of UTI or pyelonephritis after index date. ICD-10 and READ diagnosis codes for UTI or pyelonephritis in Annex I. **Initialized at index and updating from index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
Liver disease at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of liver disease using all available history prior to index date. ICD-10 and READ diagnosis codes for liver disease in Annex I. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
ICU admission at index date	The variable is evaluated at index date and indicates a record of ICU admission in all available history prior to index date. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NA Finland: HILMO, AvoHILMO
Pancreatitis at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of pancreatitis using all available history prior to index date. ICD-10 and READ diagnosis codes for pancreatitis in Annex I. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO

Congestive heart failure at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of congestive heart failure using all available history prior to index date. ICD-10 and READ diagnosis codes for congestive heart failure in Annex I. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
Systolic blood pressure at index date	Systolic blood pressure measurement value in mm Hg at index date. **Evaluated at index using closest value to index date.**	Time-fixed Continuous Categorical[1]: < 120 ≥120-<139 ≥140-<159 ≥160	UK: CPRD Sweden: NDR Finland: EMR databases
Diastolic blood pressure at index date	Diastolic blood pressure measurement value in mm Hg at index date. **Evaluated at index using closest value to index date.**	Time-fixed Continuous Categorical:[2] < 80 ≥80-89 ≥90-99 100 or higher	UK: CPRD Sweden: NDR Finland: EMR databases
Hypertension at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of hypertension, or two records within 6 months (180 days) of systolic blood pressure ≥ 140 and diastolic blood pressure ≥ 90 using all available history prior to index date. ICD-10 and READ diagnosis codes for hypertension in Annex I. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical:[2] Yes No	UK: CPRD Sweden: NPR, NDR Finland: HILMO, AvoHILMO, EMR databases

Stroke at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of stroke using all available history prior to index date. ICD-10 and READ diagnosis codes for stroke in Annex 1. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
Myocardial infarction at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of myocardial infarction using all available history prior to index date. ICD-10 and READ diagnosis codes for myocardial infarction in Annex I. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
Autoimmune disease at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of autoimmune disease using all available history prior to index date. ICD-10 and READ diagnosis codes for autoimmune disease in Annex I. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO
COPD at index date	The variable is evaluated at index date and indicates a register record with a diagnosis of COPD using all available history prior to index date. ICD-10 and READ diagnosis codes for COPD in Annex I. **Evaluated at index date using all available history prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: NPR Finland: HILMO, AvoHILMO

[1] Modification of Diet in Renal Disease (MDRD) study http://clinchem.aaccjnls.org/content/53/4/766.long

[2] http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/AboutHighBloodPressure/Understanding-Blood-Pressure-Readings_UCM_301764_Article.jsp#

5.3.3. Concomitant medications

Table 14 Other diabetic medication exposure definitions

Exposure	Definition	Details	Data source
Concomitant use of insulin at index date	Variable indicating the concomitant use of insulin with study drugs at index date, using data within the 180-day period prior to index date. ATC codes for insulins in Annex II. **Evaluated at index using data within the 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Ever exposed to insulin	Variable indicating any use of insulin using data from 180 days prior to index date onwards. ATC codes for insulins in Annex II. **Updating from 180 days prior to index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Metformin exposure	Time updating current exposure status to metformin based on purchased amounts and daily defined or prescribed dose, using data from 180 days prior to index date onwards. ATC codes for metformin in Annex II. **Updating from 180 days prior to index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Ever exposed to metformin	Variable indicating any use of metformin using data from 180 days prior to index date onwards. ATC codes for metformin in Annex II. **Updating from 180 days prior to index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Ever exposed to GLP-1 drugs	Variable indicating any use of GLP-1 drugs using data from 180 days prior to index date onwards. ATC codes for GLP-1 drugs in Annex II. **Updating from 180 days prior to index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR

Ever exposed to thiazolidinediones	Variable indicating any use of thiazolidinediones using data from 180 days prior to index date onwards. ATC codes for thiazolidinediones in Annex II. **Updating from 180 days prior to index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Ever exposed to sulfonylureas	Variable indicating any use of sulfonylureas using data from 180 days prior to index date onwards. ATC codes for sulfonylureas in Annex II. **Updating from 180 days prior to index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Ever exposed to other glucose lowering medication	Variable indicating any use of oral glucose lowering drugs other than the study drugs using data from 180 days prior to index date onwards. ATC codes for other glucose lowering drugs in Annex II. **Updating from 180 days prior to index date onwards.**	Time-dependent Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Treatment complexity at index	Treatment complexity at index date is categorized by the number of drug classes in concomitant use including study drugs, using data from 1ate80 days prior to index date onwards. **Evaluated at index using data from 180 days prior to index date onwards.**	Time-fixed Categorical: Monotherapy Dual combination therapy Triple combination therapy or higher	UK: CPRD Sweden: SPDR Finland: FPR

Table 15 Other non-diabetic concomitant medication use variables

Variable	Definition	Details	Data source
Antihypertensives/diuretics exposure at index date	Exposure status at index date of any antihypertensives/diuretics (including combinations) based on purchased amounts and daily defined dose using data within 180-day period prior to index date. ATC codes for antihypertensives/diuretics in Annex II. **Evaluated at index date using data within 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Non-steroidal anti-inflammatory drugs (NSAIDs) exposure at index date	Exposure status at index date of any NSAID (including combinations) based on purchased amounts and daily defined dose using data within 180-day period prior to index date. ATC codes for NSAID in Annex II. **Evaluated at index date using data within 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Oral steroids exposure at index date	Exposure status at index date of any oral steroid (including combinations) based on purchased amounts and daily defined dose using data within 180-day period prior to index date. ATC codes for oral steroids in Annex II. **Evaluated at index date using data within 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Statins, fibrates exposure at index date	Exposure status at index date of any statin or fibrate (including combinations) based on purchased amounts and daily defined dose using data within 180-day period prior to index date. ATC codes for statins, fibrates in Annex II. **Evaluated at index date using data within 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR

Lipid modifying agents exposure at index date	Exposure status at index date of any lipid-modifying agent (including combinations) based on purchased amounts and daily defined dose using data within 180-day period prior to index date. ATC codes for lipid modifying agent in Annex II. **Evaluated at index date using data within 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Zoledronic acid exposure at index date	Exposure status at index date of any zoledronic acid (including combinations) based on purchased amounts and daily defined dose using data within 180-day period prior to index date. ATC codes for zoledronic acid in Annex II. **Evaluated at index date using data within 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR
Antibiotics exposure at index date	Exposure status at index date of any antibiotic (including combinations) based on purchased amounts, daily defined dose and dose per time estimation using data within 180-day period prior to index date. ATC codes for antibiotic in Annex II. **Evaluated at index date using data within 180-day period prior to index date.**	Time-fixed Categorical: Yes No	UK: CPRD Sweden: SPDR Finland: FPR

6. PLANNED ANALYSIS

6.1. SUMMARY OF STUDY POPULATION IDENTIFICATION

The identification process used to identify the study populations and subgroups will be summarised using population flowcharts including the following information:

- The total number of patients in the source population.
- The number and proportion of patients not meeting inclusion criteria.
- The number and proportion of patients meeting inclusion criteria.
- The number and proportion of patients excluded after applying exclusion criteria.

- The number and proportion of patients included in the study after applying inclusion and exclusion criteria [total study population].

- The number and proportion of patients using empagliflozin (out of total study population) at index date.

- The number and proportion of patients using DPP-4 inhibitors (out of total study population) at index date.

For each of the study countries, a population flowchart similar to Figure 2 will be presented.

6.2. POPULATION DESCRIPTION BEFORE MATCHING

The study population and exposure groups identified in each of the study countries will be described at index date using the following baseline characteristics:

- Demographics and other baseline characteristics at index date Table 11.

- Diabetic complications Table 12 (status at index date).

- Non-diabetic concomitant disease Table 13 (status at index date).

- Diabetic concomitant medication use Table 14 (status at index date).

- Non-diabetic concomitant medication use Table 15 (status at index date).

All continuous variables will be described using standard statistical measures, i.e. range, mean, standard deviation, median, first and third quartile. All categorical variables will be summarized with absolute and relative frequencies.

A summary table will be presented for each of the study countries separately (2 tables). A table template is given in Table 17.

6.3. METHODS ADDRESSING BIAS

Potential biases will be addressed as follows:

- Selection bias and immortal time bias is adressed thru the study design. It include a new user cohort design, a selection of comparator drugs with the same indication and prescribed to patients with potentially similar T2D disease severity, and a definition of exposure and study entry based on information before or at index date.

- Channeling bias is adressed by using propensity score matching to construct balanced exposure groups. In addition, the year of the index date (date of first prescription of the study group) is used as an exact matching criteria.

- Bias due to definitions of exposure, definition of follow-up times and other relevant covariates will be addressed through sensitivity analyses.

- Diagnostic bias will be evaluated and addressed in sensitivity analyses through consideration of the frequency of urine dipstick testing (as captured through albuminuria tests) and stage of cancer at the time of diagnosis (as more early-stage cancers in one group would be suggestive of diagnostic bias).

6.4. PROPENSITY SCORE MATCHED COHORTS

Matched full cohort

Propensity score matched cohorts will be derived in each country by matching each initiator of empagliflozin at index date with initiators of DPP-4 inhibitors at index date with a variable matching ratio up to 1:3 (1:1 in a sensitivity analysis) using greedy matching methods (cf. Section 6.5). Free or fixed dose combination use of metformin at index date will be used as an exact matching variable.

Matched metformin-user only cohort

Using the propensity scores computed using the entire cohort, a matched metformin-user only cohort will be derived by matching within a subgroup including only patients using metformin in free or fixed dose combination with the study drugs at index date. Each initiator of empagliflozin at index date will be matched with initiators of DPP-4 inhibitors at index date with a variable matching ratio up to 1:3 using greedy matching methods (cf. Section 6.5).

6.5. IMPLEMENTATION OF THE PROPENSITY SCORE MATCHING

6.5.1. Estimation of the propensity score

Propensity score model specification

The propensity score will be estimated using a binary logistic regression model. For all individuals, propensity scores for the comparison empagliflozin vs. DPP-4 inhibitors will be computed at the index date as the probability to receive empagliflozin, conditional on treatment and clinical history up to index date within each stratum.

The resulting propensity score model will be used to create two matched cohorts in each country:

- Matched full cohort, including patients initiating empagliflozin or a DPP-4 inhibitor at index date.

- Matched metformin-user only cohort, including patients using metformin in free or fixed dose combination with empagliflozin or a DPP-4 inhibitor at index date.

Variables considered for inclusion in the propensity score model

The following variables will be evaluated at index date and considered for inclusion in the propensity score model:

- Demographics and other baseline characteristics at index date Table 11.

- Diabetic complications Table 12 (status at index date).

- Non-diabetic concomitant disease Table 13 (status at index date).

- Diabetic concomitant medication use Table 14 (status at index date).

- Non-diabetic concomitant medication use Table 15 (status at index date).

6.5.2. Matching method

For matching the full cohort, concomitant use of metformin at index date will be an exact matching variable.

Matching will be carried out in several rounds to achieve the highest possible matching ratio given the available data in each country (up to 1:3 matching). In the first round of matching, for each empagliflozin initiator i_E with index date T_{i_E}, all individuals initiating DPP-4 inhibitors within 3 months (90 days) before or after T_{i_E} and with the same status of concomitant metformin use at T_{i_E} will be eligible as comparators. For UK CPRD data, only individuals from the same database (GOLD or Aurum) as the empagliflozin initiator will be considered as eligible comparators. Of these, the individual with the closest propensity score(s) to that of i_E within a predefined propensity score interval will be selected (Greedy matching method). In case of ties, the selection will be made randomly. After each round of selection, the selected comparators are then removed from the pool of possible comparators for further selection (sampling without replacement) [5].

The predefined propensity score interval will be calculated based on the standard deviation of the logit of propensity score δ multiplied by a coefficient k. The value of the coefficient k will be initially set to 0.2, i.e. using a caliper of 0.2 standard deviations (i.e. 0.2δ). The coefficient k will then be calibrated to ensure at least 99% of empagliflozin initiators have matches from initiators of DPP-4 inihibitors after the first round of propensity score matching [6]. The same calibrated matching caliper $k\delta$ used in the first matching round will be reused in the second and third matching rounds.

Matching by time-period, based on index dates, enables better control of time-varying confounders including changing prescription patterns for empagliflozin and potential confounding. In case there is a strong time bias in the number of matches per empagliflozin user (e.g., initiators in early years have more matches than initiators from later years), a relaxation of the criteria for matching within time-period will be considered. In addition, the criteria for calibrating coefficient k may be modified if the data does not allow matching of 99% of empagliflozin initiators.

6.6. ASSESSING THE QUALITY OF THE MATCHING

6.6.1. Propensity score distribution before and after matching and common support region

For each propensity score set (all patient cohort and concomitant metformin users cohort), the propensity score distribution will be plotted for empagliflozin users and DPP-4 inihibitor users before matching (Figure 3). This will be used for visual inspection of the common support region.

The propensity score distribution in non-matched empagliflozin users, matched empagliflozin users, and matched DPP-4 inhibitor users will be plotted for each cohort in each country. Results will be presented in 2 plots similar to Figure 5, one for each of the two propensity score matched cohorts.

6.6.2. Assessment of covariate balance before and after matching

The covariates balance between treatment groups before and after matching will be evaluated using:

- Standardized differences (percentage).
- χ^2 test for equality of proportions in dichotomous variables (p-value).
- t-test for equality of means in continuous variables (p-value).

For continuous variables, the standardized difference is defined as

$$d = \frac{(\bar{x}_{empagliflozin} - \bar{x}_{DPP-4\ inihibitor})}{\sqrt{\frac{s^2_{empagliflozin} + s^2_{DPP-4\ inihibitor}}{2}}} \times 100$$

Where $\bar{x}_{empagliflozin}$ and $\bar{x}_{DPP-4\ inihibitor}$ denote the sample mean of the covariate in empagliflozin users at index date and comparison drug users at index date, respectively, while $s^2_{empagliflozin}$ and $s^2_{DPP-4\ inihibitor}$ denote the sample variance of the covariate in empagliflozin users at index date and comparison drug users at index date, respectively.

For dichotomous variables, the standardized difference is defined as

$$d = \frac{(\hat{p}_{empagliflozin} - \hat{p}_{DPP-4\ inihibitor})}{\sqrt{\frac{\hat{p}_{empagliflozin}(1-\hat{p}_{empagliflozin}) + \hat{p}_{DPP-4\ inihibitor}(1-\hat{p}_{DPP-4\ inihibitor})}{2}}} \times 100$$

Where $\hat{p}_{empagliflozin}$ and $\hat{p}_{DPP-4\ inihibitor}$ denote the prevalence of the dichotomous variable in empagliflozin users at index date and DPP-4 inhibitor users at index date, respectively [7]

.

If a categorical variable is defined with more than two categories (i.e. not dichotomous), each category will be treated as a dichotomous variable when computing the standardized differences.

The balance assessment will be performed with regard to the following covariates:

- Demographics and other baseline characteristics at index date Table 11.

- Diabetic complications Table 12 (status at index date).

- Non-diabetic concomitant disease Table 13 (status at index date).

- Diabetic concomitant medication use Table 14 (status at index date).

- Non-diabetic concomitant medication use Table 15 (status at index date).

Balance in baseline characteristics before and after matching between empagliflozin users at index and the DPP-4 inihibitor users at index will be reported in each country, separately.

All continuous variables will be described using standard statistical measures, i.e. range, mean, standard deviation, median, first and third quartile. All categorical variables will be summarized with absolute and relative frequencies. The standardized differences and p-values for the statistical tests will be also reported.

Results will be reported in 4 tables similar to Table 18.

Standardized difference < 10% will be used as indicator of satisfactory distribution balance after matching [8]. Variables with standardized difference larger than 10% might still exist. The unbalanced variables will be considered as potential confounders and included in the adjustment of Cox's PH model, as described in the comparative analysis (Section 6.7.2). In case a substantial amount of variables remain unbalanced, the PS matching procedure will be re-evaluated by, for example, reducing the caliper width.

6.7. AS-TREATED ANALYSIS

6.7.1. Descriptive analysis

Crude incidence rates (IRs) for the outcomes defined in Table 8 will be estimated within each exposure category as defined in Table 5 using Poisson regression.

In addition, stratified IRs will be derived for the following variables:

- Demographics and other baseline characteristics at index date Table 11.

- Diabetic complications Table 12.

- Non-diabetic concomitant disease Table 13.

- Diabetic concomitant medication use Table 14.

- Non-diabetic concomitant medication use Table 15.

Crude IRs will be presented along with 95% CIs. Other descriptive summaries such as number of events and person-time for calculating the IR will be included. Analysis will be performed for each country. Results will be reported as in in Table 22.

Kaplan-Meier plots,with 95% CIs, for event-free survival probabilities will be presented for each country for the outcomes defined in Table 8 with exposures as defined in Table 5**.** Results will be reported as the figure template shown in Figure 6.

Meta-analyses of IRs will be carried out using aggregated data across countries as described in Section 6.7.3.1 and reported as shown in Table 22.

6.7.2. Comparative analysis

Hazard ratios (HRs) for the outcomes defined in Table 8 will be estimated for PS matched cohorts using Cox's PH model with time-varying covariates, if model pre-requisites are fulfilled. HRs will be presented along with 95% CIs. Proportional hazards assumptions for time-fixed exposure variables will be assessed for each model by plotting scaled Schoenfeld residuals.

The comparative analyses will be performed separately for the matched full cohort and matched metformin-user cohort.

Crude model

Crude HRs are calculated using a Cox PH model with as-treated exposure (as defined in Section 5.1.7) as the only included variable.

HRs estimated using the crude model will be presented along with 95% CIs. A table template for the result is given in Table 23. Analysis will be performed separately for each country.

Crude HRs with 95% CIs will be collected from all study countries to carry out meta-analyses as described in Section 6.7.3.2. Results will be reported as in Figure 8.

Base model

Based on the crude model, the base model will be further adjusted with the unbalanced covariates at index date (standard difference ≥ 10%) among the variables included in the propensity score model (Section 6.6). The inclusion of adjusting variables will be considered based on data availablity as described in Section 10.

HRs estimated using the base model will be presented along with 95% CIs. A table template for the result is given in Table 24. Analysis will be performed separately for each country. HRs with 95% CIs estimated using base model will be collected from all study countries to carry out meta-analyses as described in Section 6.7.3.2. Results will be reported as in Figure 8.

Adjusted model

The base model will be further adjusted with selected time-dependent variables. The following time-dependent variables will be initially considered for inclusion (as detailed in Section 9) in the adjusted Cox's PH model:

- HbA1c value (as defined in Table 11)
- Kidney or genitourinary stones (as defined in Table 13)
- UTI or pyelonephritis (as defined in Table 13)
- Ever exposed to insulin (as defined in Table 14)
- Ever exposed to metformin (as defined in Table 14)

- Ever exposed to other glucose lowering medication (as defined in Table 14)
- Ever exposed to GLP-1 drugs (as defined in Table 14)
- Ever exposed to thiazolidinediones (as defined in Table 14)
- Ever exposed to sulfonylureas (as defined in Table 14)

HRs estimated using the adjusted model with time-dependent covariates will be presented along with 95% CIs. A table template for the result is given in Table 25. Analysis will be performed and model adjustment will be applied separately for each country.

Adjusted HRs with 95% CIs will be collected from all study countries to carry out meta-analyses as described in Section 6.7.3.2. Results will be reported as in Figure 8.

Handling of model convergence failure

For outcomes with a small number of events, converge failure of multivariable Cox PH regression model may happen if there are no events for a particular combination of covariate values. Thus, a minimum of 5 outcome events per category of a variable is required for it to be included in the Cox model. In case of persisting model convergence failure while meeting this minimum requirement, the included covariates will be removed sequentially based on the rank of their PS standardised differences. The covariate with the smallest standardised difference will be removed iteratively until the model converges. In case of ties in selection, the choice is made randomly. The removed covariates will be reported along with the HR estimates.

6.7.3. META-ANALYSES

Meta-analyses will be carried out when there is no relevant heterogeneity of variable definitions and effect size of outcomes with respect to primary exposures among country-specific analyses. Country-level effect-sizes will be pooled using meta-analysis models to compute pooled effect-size estimates.

6.7.3.1. META-ANALYSIS OF DESCRIPTIVE ANALYSES

Pooled incidence rates

For each outcome, the aggregated number of events and person-time at risk will be collected from all study countries to estimate the pooled IRs and their 95% CIs.

The pooled IRs for all K countries will be estimated as

$$IR_{pooled} = \frac{\sum_{k=1}^{K} x_k}{\sum_{k=1}^{K} t_k}$$

where x_k and t_k are the number of events and the person-time at risk in each country k.

95% CIs will be estimated using Poisson regression and reported similarly using Table 22.

6.7.3.2. META-ANALYSIS OF COMPARATIVE ANALYSIS

Meta-analyses will combine the estimates of adjusted HRs across countries using the inverse-variance method. Log-transformed HRs and their standard errors will be collected to calculate pooled HR estimates using random-effects models. Pooled HRs and their 95% CI will be reported using forest plots. In a sensitivity analysis, pooled HRs will be estimated using fixed-effect models.

Between-study heterogeneity

Prior to conducting meta-analyses, heterogeneity across the countries in terms of study design and conduct will be assessed. Specifically, the variability in population characteristics (including data availability), in pre- and post- PS matching balance in baseline characteristics between cohorts, as well as in observed incidence rates will be examined. If needed, changes in study population selection criteria or other analysis parameters will be considered to reduce such heterogeneity.

Statistical heterogeneity between country-level estimates will be estimated using Higgins' and Thompson's I^2 [9] and the moment-based estimate of τ^2.

Both I^2 and τ^2 are derived using Cochran's Q statistic defined as

$$Q = \sum_{k=1}^{K} w_k \left(\hat{\theta}_k - \frac{\sum_{k=1}^{K} w_k \hat{\theta}_k}{\sum_{k=1}^{K} w_k}\right)^2; w_k = \frac{1}{s_k^2}$$

where $\hat{\theta}_k$ denotes the log-transformed per-country HR estimate,s_k^2 its variance, w_k the corresponding weight from the country-level study k calculated as the inverse of variance s_k^2.

<u>I^2 statistic</u>

I^2 is the percentage of variability in the treatment estimates which is attributable to heterogeneity between studies rather than to sampling error. I^2 is formulated as

$$I^2 = max\left\{0, 1 - \frac{K-1}{Q}\right\}$$

where Q is Cochran's Q statistic and K is the number of the per-country studies.

I^2 and its 95% CIs are used to assess the importance of between study heterogeneity as follows:

- 0% to 40%: might not be important
- 30% to 60%: may represent moderate heterogeneity
- 50% to 90%: may represent substantial heterogeneity
- 75% to 100%: considerable heterogeneity

I^2 is sensitive to the precision of the included studies, with higher I^2 associated with increased precision in individual studies.

τ^2 statistic

The τ^2 statistic represents the variance of the distribution of the true effect-sizes. τ^2 will be calculated using the DerSimonian and Laird estimate:

$$\hat{\tau}^2 = max\left\{0,\ \frac{Q-(K-1)}{\sum_{k=1}^{K} w_k - \frac{\sum_{k=1}^{K} w_k^2}{\sum_{k=1}^{K} w_k}}\right\};\ w_k = \frac{1}{s_k^2}$$

where w_k denotes the corresponding weight from the per-country study k, and s_k^2 is the variance of HR estimate of per-country study k.

τ^2 is insensitive to the number of studies or to their precision but can be difficult to interpret. It is however expressed at the same scale as square of the effect size; hence its 95% CI represents the estimated range of true effect sizes.

Random-effects model

The random-effects model considers the observed per-country effect sizes as randomly sampled from a distribution of true effect sizes. The pooled effect size estimated by the model represents the mean of this distribution of true effect sizes.

In addition to the sampling error (squared as variance s_k^2), the random-effects model takes into account the variance of the distribution of the true effects, which is estimated as between-study heterogeneity tau-squared τ^2. The weight w_k for each country k is calculated with τ^2 as an additional scaling factor, resulting in the adjusted random-effects weights w_k^*.

$$w_k^* = \frac{1}{s_k^2 + \tau^2}$$

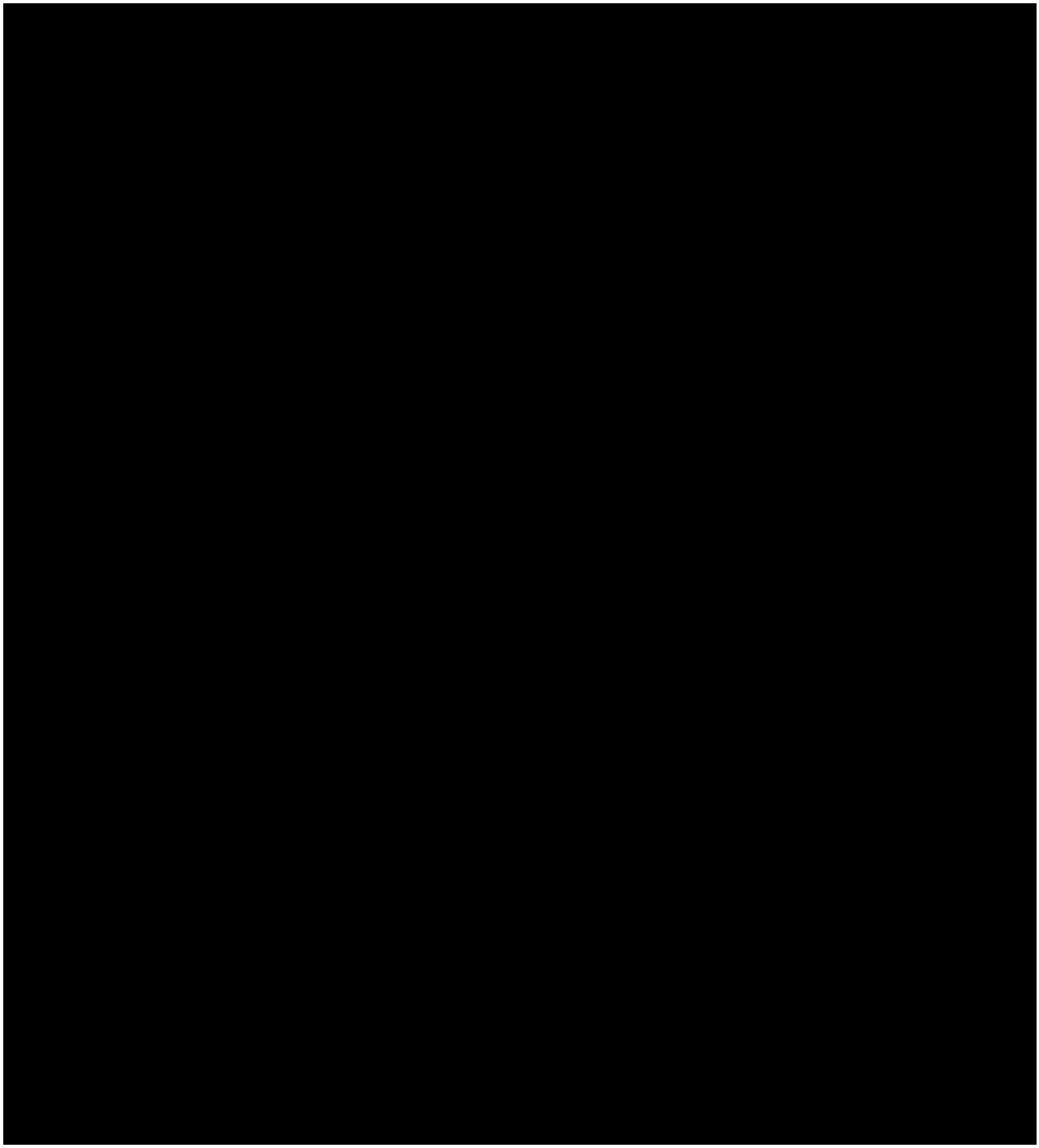
The pooled HR estimate $\hat{\theta}$ is calculated using adjusted weights w_k^*.

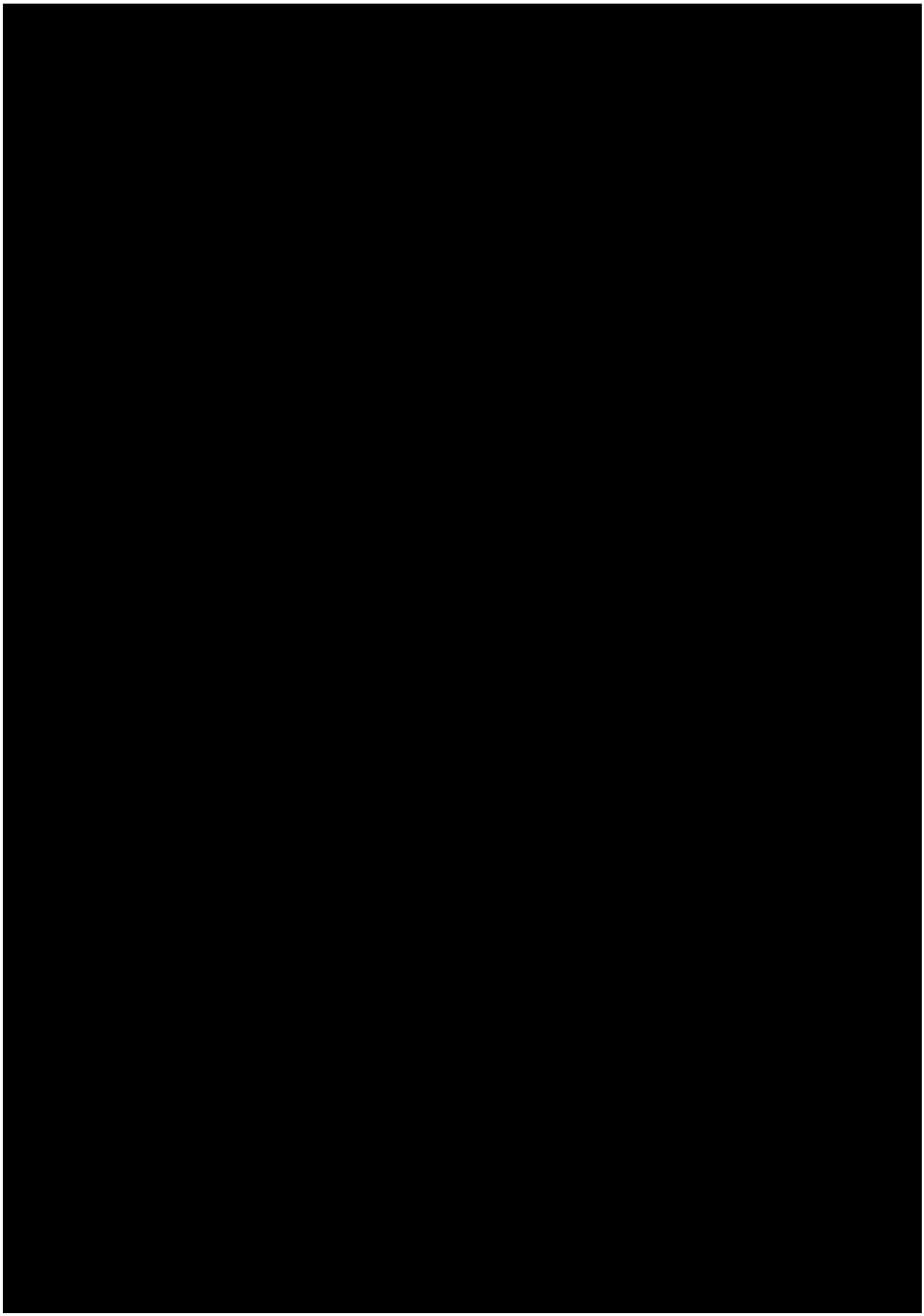
$$\hat{\theta} = \frac{\sum_{k=1}^{K} \hat{\theta}_k w_k^*}{\sum_{k=1}^{K} w_k^*}$$

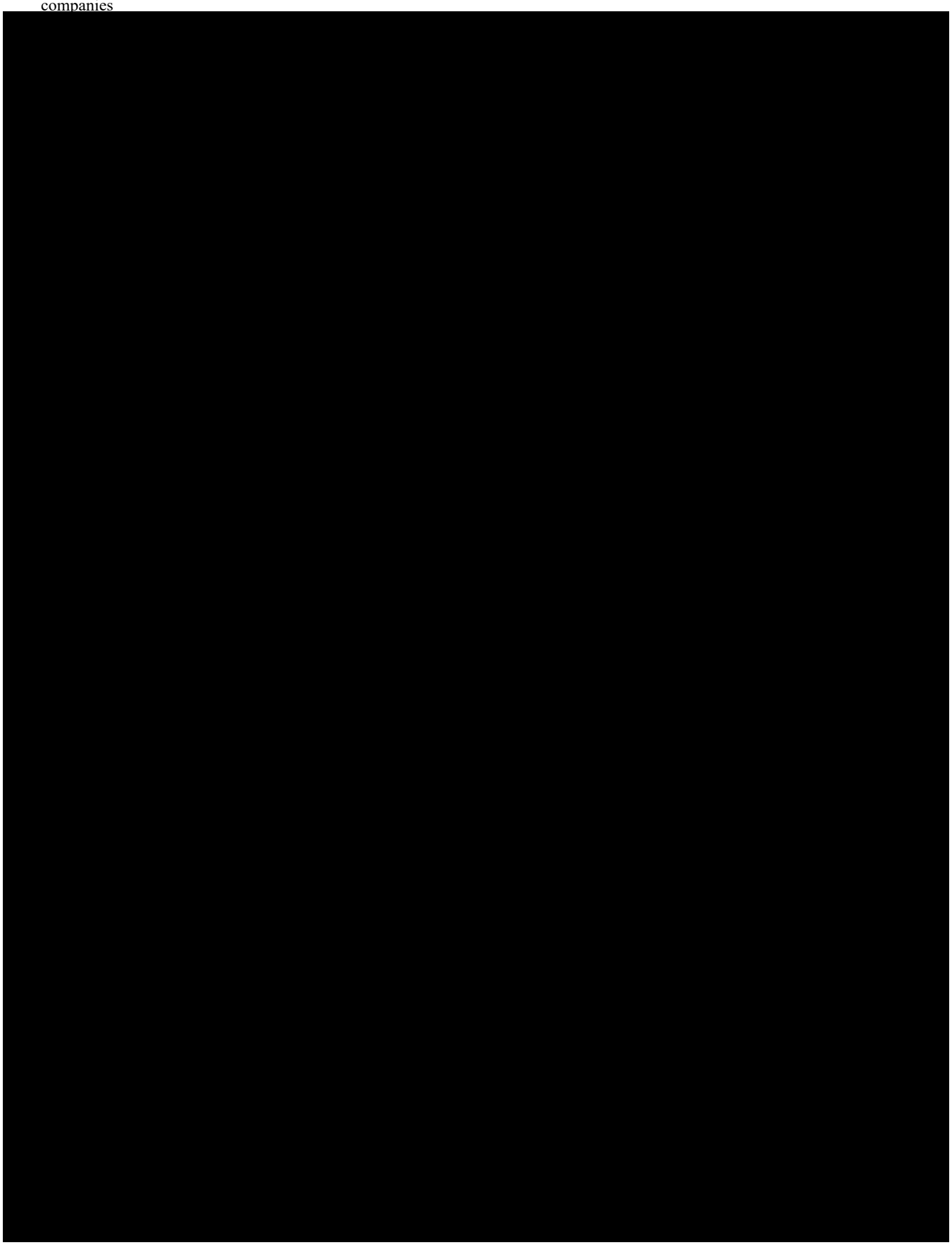
For each meta-analysis, pooled and country-level results will be presented in forest plots similar to Figure 8 showing the following items:

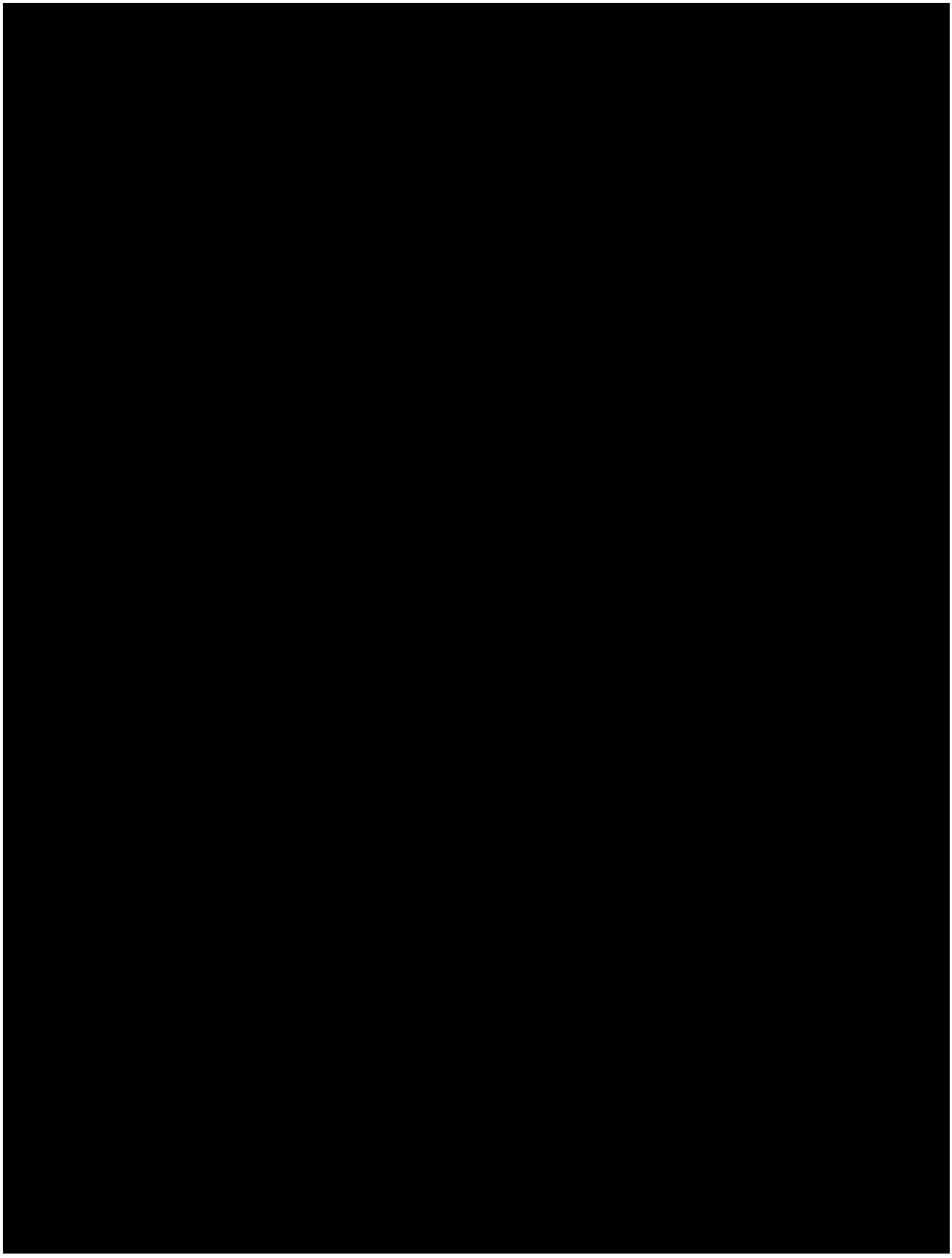
- Country identifier
- Number of events (IR/1000 person years) per study group
- HR and 95% CI for each country (on log scale)
- The weight (%) assigned to each country
- The estimated pooled HR and its 95% CI using a random-effect models (on log scale)

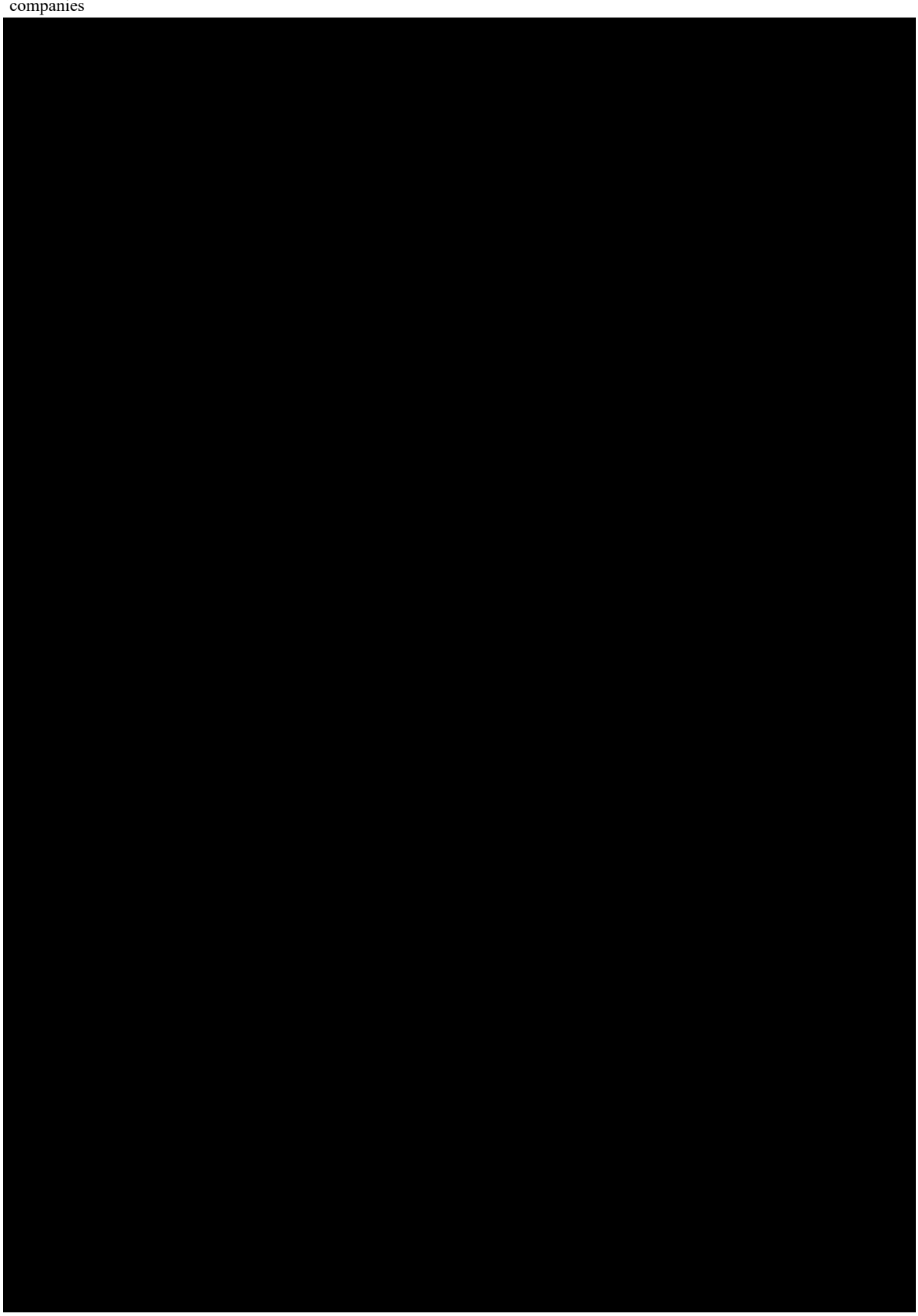
- Total number of patients per study group
- Average follow-up time in person years
- Between-study heterogeneity τ^2 and its 95% CI
- The I^2 statistics and its 95% CI

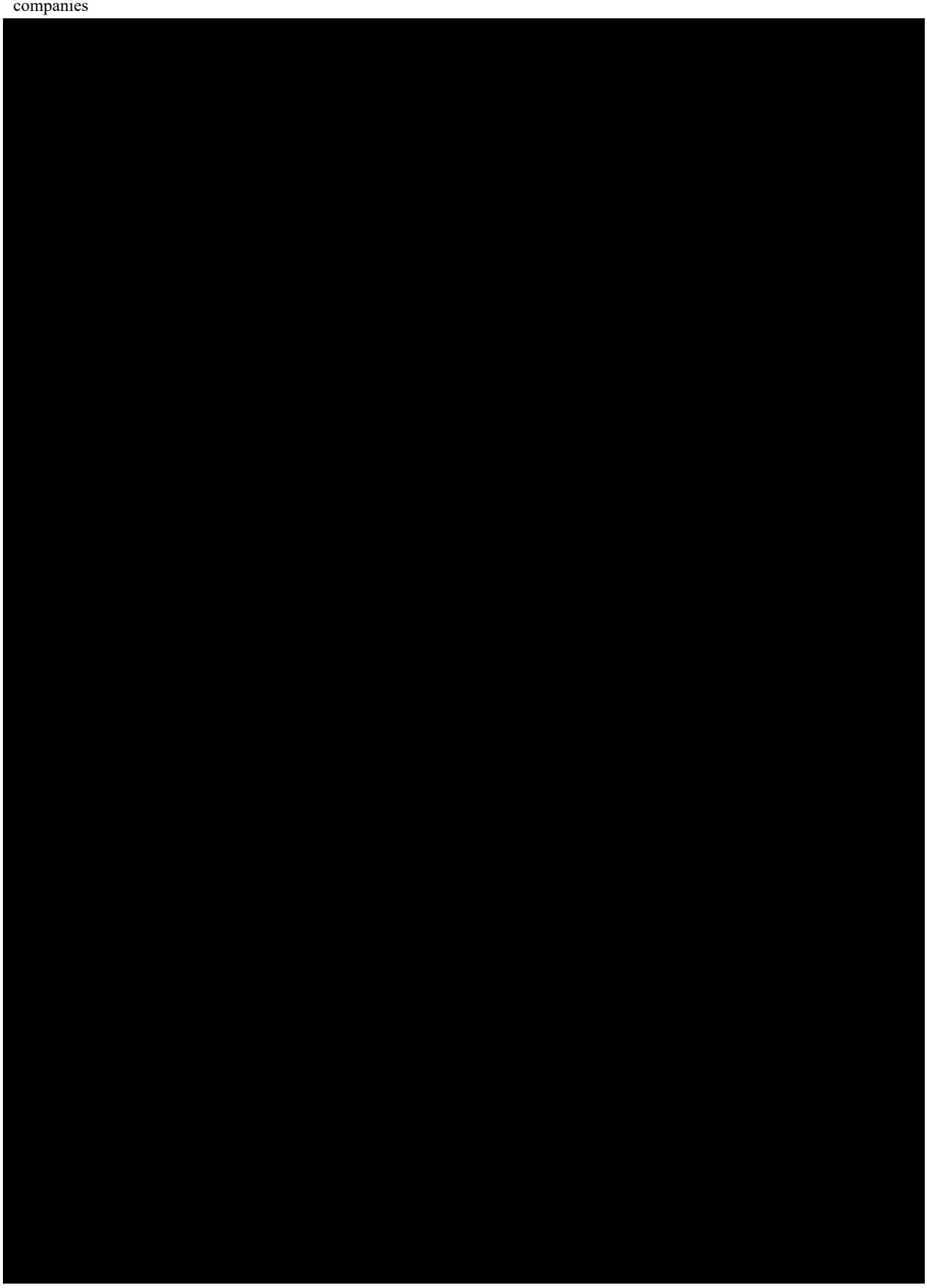


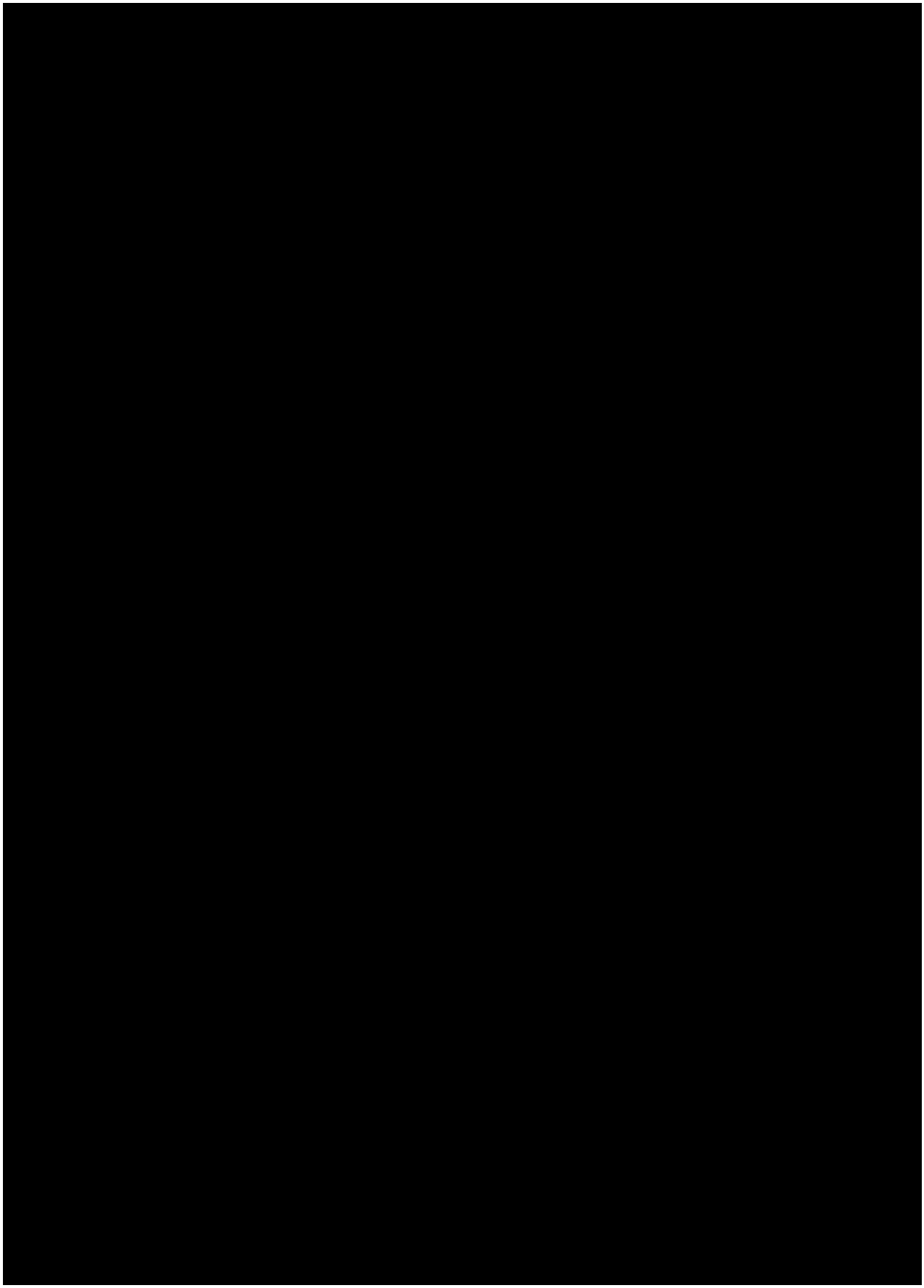


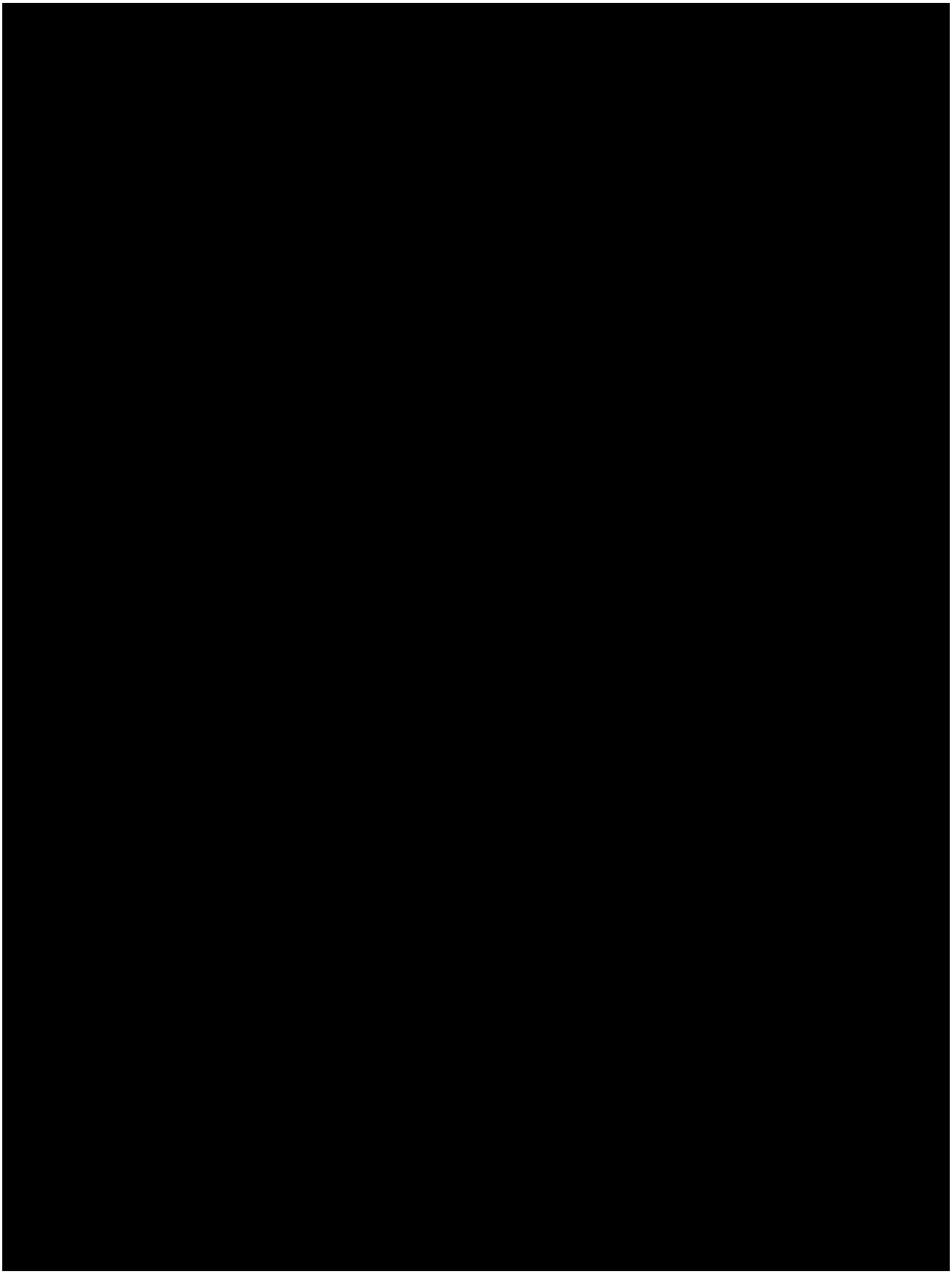


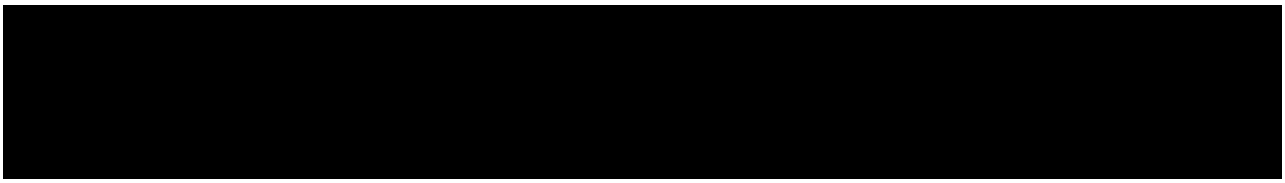












6.10. LIST OF PLANNED ANALYSIS

Table 16 List of planned analyses.

Analysis	Details	Exposure definition	Outcomes	Template
Population identification				
Summary of the process used to identify the study populations and study treatment groups.	Country specific analysis (3 figures, one for each country)	N/A	N/A	Figure 2
Population description before matching				
Description of patient baseline characteristics for the study population and the study treatment groups before matching.	Country specific analysis (3 tables, one for each country)	N/A	N/A	Table 17
Assessment of covariate balance before matching				
Balance assessment for baseline characteristics between empagliflozin users at index and DPP-4 inhibitors users at index before matching.	Country specific analysis (3 tables, one for each country)	N/A	N/A	Table 18
Checking overlap in propensity score distribution				
Propensity score distribution plot for empagliflozin users at index and DPP-4 inhibitors users at index.	Country specific analysis (3 figures, one for each country)	N/A	N/A	Figure 3
Assessing the quality of the matching				
Balance assessment for baseline characteristics between empagliflozin users at index and DPP-4 inhibitors users at index after matching.	Country specific analysis (3 tables, one for each country)	N/A	N/A	Table 18
Standardized differences plot before and after PS matching for empagliflozin users at index vs. DPP-4 inhibitors users at index	Country specific analysis (3 figures, one for each country)	N/A	N/A	Figure 4

Analysis	Details	Exposure definition	Outcomes	Template
Propensity score distribution in non-matched empagliflozin users at index date matched empagliflozin users at index and matched DPP-4 inhibitors users at index.	Country specific analysis (3 figures, one for each country)	N/A	N/A	Figure 5
Propensity score assessment distribution	Country specific analysis (3 tables, one for each country)	N/A	N/A	Table 19
Primary objectives: Descriptive analysis				
Stratified incidence rates	Using matched cohorts Country specific analysis (9 tables, 3 outcomes x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 22
Meta-analysis of incidence rates	Using pooled number of outcome events and person-time from all study countries (3 tables, 3 outcomes)	As treated exposure definition	Main outcomes-Table 8	Table 22
Kaplan-Meier survival probability plots	Using matched cohorts Country specific analysis (9 figures, 3 outcomes x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Figure 6
Primary objectives: Comparative analysis				
Hazard ratio estimates using crude model	Using matched cohorts Country specific analysis (18 tables, 2 cohorts x 3 outcomes x 3 countries	As treated exposure definition	Main outcomes-Table 8	Table 23

Analysis	Details	Exposure definition	Outcomes	Template
Meta-analysis of crude HR estimates	Using adjusted HR estimates and 95% CIs from all study countries (2 cohorts: 6 figures, 3 outcomes)	As treated exposure definition	Main outcomes-Table 8	Figure 8
Hazard ratio estimates using the base model	Using matched cohorts Country specific analysis (18 tables, 2 cohorts x 3 outcomes x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 23
Meta-analysis of HR estimates using the base model	Using adjusted HR estimates and 95% CIs from all study countries (2 cohorts: 6 figures, 3 outcomes)	As treated exposure definition	Main outcomes-Table 8	Figure 8
Hazard ratio estimates using the adjusted model	Using matched cohorts Country specific analysis (18 tables, 2 cohorts x 3 outcomes x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 25
Meta-analysis of adjusted HR estimates	Using adjusted HR estimates and 95% CIs from all study countries (2 cohorts: 6 figures, 3 outcomes)	As treated exposure definition	Main outcomes-Table 8	Figure 8
Stratified analysis				
Hazard ratio estimates using adjusted model	Using matched cohorts Stratified by treatment complexity at index. (27 tables, 3 outcomes x 3 strata by treatment complexity at index x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 25

Analysis	Details	Exposure definition	Outcomes	Template
Hazard ratio estimates using adjusted model	Using matched cohorts Stratified by PS quartiles at index. (36 tables, 3 outcomes x 4 strata by propensity score quartiles x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 25
Hazard ratio estimates using adjusted model	Using matched cohorts Stratified by age category at index. (63 tables, 3 outcomes x 7 strata by age category x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 25
Hazard ratio estimates using adjusted model	Using matched cohorts Stratified by gender. (18 tables, 3 outcomes x 2 strata by gender x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 25
Hazard ratio estimates using adjusted model	Using matched cohorts Stratified by insulin use at index. (18 tables, 3 outcomes x 2 strata by insulin use at index x 3 countries)	As treated exposure definition	Main outcomes-Table 8	Table 25
Meta-analyses for adjusted HR estimates	Using adjusted HR estimates and 95% CIs from all study countries. For each stratification variable listed above. (54 figures, 3 outcomes x (3+4+7+2+2) strata)	As treated exposure definition	Main outcomes-Table 8	Figure 8
Secondary objectives: Descriptive analysis				

Analysis	Details	Exposure definition	Outcomes	Template
Stratified incidence rates	Using matched cohorts Country specific analysis (9 tables, 3 outcomes x 3 countries)	Cumulative exposure dose of empagliflozin	Main outcomes-Table 8	Table 30
Stratified incidence rates	Using empagliflozin users subgroup Country specific analysis (9 tables, 3 outcomes x 3 countries)	Empagliflozin daily dose	Main outcomes-Table 8	Table 31
Stratified incidence rates	Using empagliflozin users subgroup Country specific analysis (9 tables, 3 outcomes x 3 countries)	Cumulative exposure duration to empagliflozin	Main outcomes-Table 8	Table 32
Meta-analysis of incidence rates	Using pooled number of outcome events and person-time from all study countries (9 tables, 3 outcomes x 3 secondary exposure definitions)	The secondary exposure definitions listed above	Main outcomes-Table 8	Table 30-Table 32
Secondary objectives: Comparative analysis				
Hazard ratio estimates using adjusted model	Using matched cohorts Country specific analysis (9 tables, 3 outcomes x 3 countries)	Cumulative exposure dose of empagliflozin	Main outcomes-Table 8	Table 26
Hazard ratio estimates using adjusted model	Using matched cohorts Country specific analysis (9 tables, 3 outcomes x 3 countries)	Empagliflozin daily dose	Main outcomes-Table 8	Table 27

Analysis	Details	Exposure definition	Outcomes	Template
Hazard ratio estimates using adjusted model	Using matched cohorts Country specific analysis (9 tables, 3 outcomes x 3 countries)	Cumulative exposure duration to empagliflozin	Main outcomes-Table 8	Table 28
Meta-analysis of adjusted HR estimates	Using adjusted HR estimates with 95% CIs from all study countries (9 figures, 3 outcomes x 3 secondary exposure	3 secondary exposure definitions listed above	Main outcomes-Table 8	Figure 8
[redacted]				
Sensitivity analysis: NCC study design				
Balance assessment for baseline characteristics between matched cases and controls.	Country specific analysis (3 tables, one for each country)	N/A	N/A	Table 33
Odds ratio using conditional logistic regression	Nested case-control study design. (3 tables, 1 outcome (all cancer cases) x 3 countries)	As treated exposure definition	All cancer cases defined in Table 8	Table 34
Meta-analysis for odds ratio estimates	Using odds ratio estimates with 95% CIs from all study countries (1 figure, 1 outcome (all cancer cases))	As treated exposure definition	All cancer cases defined Table 8	Figure 8

Analysis	Details	Exposure definition	Outcomes	Template
Sensitivity analysis: Adjusting for time-dependent confounding				
Hazard ratio estimates using marginal structural model	Using matched cohorts MSM using IPTW. (3 tables, 1 outcome x 3 countries)	As treated exposure definition	Urinary tract cancer as defined in Table 8	Table 35
Meta-analysis of adjusted HR estimates using MSM	Using adjusted HR estimates with 95% CIs from all study countries (1 figure, 1 outcome)	As treated exposure definition	Urinary tract cancer as defined in Table 8	Figure 8
Sensitivity analysis: Exposures				
Hazard ratio estimates using adjusted model	Using matched cohorts (3 tables, 1 outcome x 3 countries)	Intention to treat exposure definition	Urinary tract cancer as defined in Table 8	Table 25
Meta-analysis of adjusted HR estimates	Using adjusted HR estimates with 95% CIs from all study countries (1 figure, 1 outcome)	Intention to treat exposure definition	Urinary tract cancer as defined in Table 8	Figure 8
Sensitivity analysis: Grace period				
Hazard ratio estimates using adjusted model	Using matched cohorts Allowed gaps between exposure episodes extended to 90 days (3 tables, 1 outcome x 3 countries)	Intention to treat exposure definition	Urinary tract cancer as defined in Table 8	Table 25
Meta-analysis of adjusted HR estimates	Using adjusted HR estimates with 95% CIs from all study countries (1 figure, 1 outcome)	Intention to treat exposure definition	Urinary tract cancer as defined in Table 8	Figure 8

Analysis	Details	Exposure definition	Outcomes	Template
Sensitivity analysis: Outcomes				
Crude incidence rates	Using matched cohorts Country specific analysis (9 tables, 3 outcomes x 3 countries)	As treated exposure definition	Further outcomes- Table 9	Table 29
Hazard ratio estimates using adjusted model	Using matched cohorts (3 tables, 3 outcomes)	As treated exposure definition	Further outcomes- Table 9	Table 25
TNM staging at diagnosis	Using matched cohorts (9 tables, 3 outcomes x 3 countries)	As treated exposure definition	Further outcomes- Table 10	Table 21
Albuminaria tests each year	Using matched cohorts (3 tables, 3 countries)	As treated exposure definition	NA	Table 36
Sensitivity analysis: Extended latency period				
Hazard ratio estimates using adjusted model	Using matched cohorts Duration of latency period is extended to 12 months. (3 tables, 1 outcome x 3 countries)	As treated exposure definition	Urinary tract cancer as defined in Table 8	Table 25
Meta-analysis of adjusted HR estimates	Using adjusted HR estimates with 95% CIs from all study countries (1 figure, 1 outcome)	As treated exposure definition	Urinary tract cancer as defined in Table 8	Figure 8
Hazard ratio estimates using adjusted model	Using matched cohorts Discontinuation is 6 months after last day of medication or switch. (3 tables, 1 outcome x 3 countries)	As treated exposure definition	Urinary tract cancer as defined in Table 8	Table 25
Meta-analysis of adjusted HR estimates (discontinuation is 6 months after last day of medication or switch)	Using adjusted HR estimates with 95% CIs from all study countries (1 figure, 1 outcome)	As treated exposure definition	Urinary tract cancer as defined in Table 8	Figure 8

Analysis	Details	Exposure definition	Outcomes	Template
Sensitivity analysis: Inclusion of the Aurum CPRD database				
Description of patient baseline characteristics for the UK, stratified by CPRD database	CPRD specific analysis (1 table)	N/A	N/A	Table 37
Stratified incidence rates	Using UK matched cohorts (3 tables, 3 outcomes x 1 country)	As treated exposure definition	Main outcomes-Table 8	Table 38
Hazard ratio estimates using crude model after adding CPRD database as a stratification variable	Using UK matched cohorts (3 tables, 3 outcomes x 1 country)	As treated exposure definition	Main outcomes-Table 8	Table 39
Hazard ratio estimates using the base model after adding CPRD database as a stratification variable	Using matched cohorts Only for the UK (3 tables, 3 outcomes x 1 table)	As treated exposure definition	Main outcomes-Table 8	Table 40
Hazard ratio estimates using the adjusted model after adding CPRD database as a stratification variable	Using matched cohorts Only for the UK (3 tables, 3 outcomes x 1 table)	As treated exposure definition	Main outcomes-Table 8	Table 41
Hazard ratio estimates using the crude model, stratified by UK-CPRD databases	Using UK matched cohorts (3 tables, 3 outcomes x 1 country)	As treated exposure definition	Main outcomes-Table 8	Table 42
Hazard ratio estimates using the base model, stratified by UK-CPRD databases	Using UK matched cohorts (3 tables, 3 outcomes x 1 country)	As treated exposure definition	Main outcomes-Table 8	Table 43
Hazard ratio estimates using the base model, stratified by UK-CPRD databases	Using UK matched cohorts (3 tables, 3 outcomes x 1 country)	As treated exposure definition	Main outcomes-Table 8	Table 44
Sensitivity analysis: Meta-analysis using fixed-effect model				

Analysis	Details	Exposure definition	Outcomes	Template
Meta-analysis of adjusted HR estimates using a fixed-effect model	Using adjusted HR estimates and 95% CIs from all study countries (3 figures, 3 outcomes)	As treated exposure definition	Main outcomes-Table 8	Figure 9

[redacted]

7. DATA MANAGEMENT

[redacted] will collect electronic patient records and comply with procedures that include checking electronic files, maintaining security and data confidentiality, following analyses plans, and performing quality checks for all programs.

[redacted] will maintain any patient-identifying information securely on site according to up-to-date standard operating procedures. [redacted] will also maintain appropriate data storage, and archiving procedures will be followed with periodic backup of files.

8. STATISTICAL SOFTWARE

R language (http://www.r-project.org) will be used in statistical analysis for creating tabulations and graphics as well as in all statistical modelling. R language is described in more details in report "R: Regulatory Compliance and Validation Issues: A Guidance Document for the Use of R in Regulated Clinical Trial Environments" (www.r-project.org/doc/R-FDA.pdf, read 21 December 2016)[12].

9. MISSING DATA

A high frequency of missing values is not expected for most variables, with the possible exception of lifestyle and biological measurement variables. For the medical history conditions / comorbidities to be collected for inclusion in the propensity score, the absence of a code for a condition will be interpreted as an absence of the event.

If a variable is missing for only some of the patients, a missing data category will be added and utilized in the analysis. If a variable is missing for more than 50% of patients, removal of the variable from the relevant model or analysis (propensity score, multivariable Cox's regression, NCC matching) will be considered.

10. QUALITY CONTROL

The data analysis will be conducted as specified in the SEAP. The principal investigator, the co-investigators and the sponsors of the study must approve all revisions to the SEAP. All major changes to the SEAP shall be properly documented as SEAP amendments and when necessary such SEAP amendments will be implemented in the study protocol and delivered to relevant ethics committees and register holders.

[redacted] will be the registered holder for the retrieved Swedish, Finnish and UK data, also with the responsibility for destroying the data after the study, when legally possible. A quality control will be performed on the retrieved register data, including the inclusion of compulsory variables, such as SID and main diagnosis, in the delivered register data. If data are missing or incorrect, the dataset is sent back to the register holder for correction.

All processes from data management leading to dissemination of study results will undergo quality control checks for programs, result tables and written text, according to [redacted]'s Standard Operating Procedures and project-specific instructions RWI_WI_EPI0010 "Non-Interventional Study and Low Intervention Clinical Trial Report Development", RW_PI_OP-401.01_003_01 "Programming and statistical analysis", and RW_PI_OP-406.01_003_01 "Programming validation". A detailed audit trail of all documents (programs, result tables, reports) along with quality control processes will be maintained.

All programs for data management and data analyses written by study statistician(s) will be self-documenting with comments about the data handling process, the population selection and the analysis performed by the program. The documentation will be sufficient for another statistician to be able to repeat the program. A statistician other than the one who writes the program will carry out quality control checks of these programs. The programming QC will be performed for all programs related to the database, its individual datasets from different registers and statistical analyses. The QC will be documented with the following information:

- Name of the program
- Purpose of the program
- Name of the programmer
- Validation status and date of validation
- Updates and dates of updates
- Name of the QC statistician
- QC findings or comments
- Date of QC

The statistician responsible for the QC of a program will ensure that:

- The correct dataset and population has been used
- The analysis is carried out according to the SEAP
- The analysis results are consistent
- The output is in line with the table shells in the SEAP
- The analysis can be interpreted in both statistical and clinical terms.

All QC findings by the QC statistician will be documented, and corrections are made to the programs accordingly. All changes are documented as updates to the original program.

The raw dataset and statistical programs used for generating the data included in the final study report will be kept in electronic format and be available for auditing and inspection.

As the register holder of the study register [redacted] is in charge of archiving and deleting the data. Secure archives will be maintained for the orderly storage and retrieval of all study related material. An index shall be prepared to identify the archived contents and their location. Access to the archives will be controlled and limited to authorised personnel only. Access to the study data cannot be given to any third parties, neither the study data can be used to other purposes than prescribed in the study protocol. All requests to use the study data for other purposes than mentioned in the study protocol must be subjected to appropriate data permit processes.

11. REFERENCES

[1] S. Schneeweiss, "A basic study design for expedited safety signal evaluation based on electronic healthcare data," *Pharmacoepidemiol Drug Saf*, vol. 19, no. 8, pp. 858–868, Aug. 2010, doi: 10.1002/pds.1926.

[2] E. Patorno *et al.*, "Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations," *Diabetologia*, vol. 57, no. 11, pp. 2237–2250, Nov. 2014, doi: 10.1007/s00125-014-3364-z.

[3] CPRD, "Clinical Practice Research Datalink," 2021. https://www.cprd.com/ (accessed Nov. 04, 2021).

[4] L. Pazzagli *et al.*, "Methods for constructing treatment episodes and impact on exposure-outcome associations," *Eur J Clin Pharmacol*, vol. 76, no. 2, pp. 267–275, Feb. 2020, doi: 10.1007/s00228-019-02780-4.

[5] J. A. Rassen, A. A. Shelat, J. Myers, R. J. Glynn, K. J. Rothman, and S. Schneeweiss, "One-to-many propensity score matching in cohort studies," *Pharmacoepidemiol Drug Saf*, vol. 21 Suppl 2, pp. 69–80, May 2012, doi: 10.1002/pds.3263.

[6] P. C. Austin, "Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies," *Pharm Stat*, vol. 10, no. 2, pp. 150–161, Apr. 2011, doi: 10.1002/pst.433.

[7] P. C. Austin, "Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples," *Statistics in Medicine*, vol. 28, no. 25, pp. 3083–3107, 2009, doi: 10.1002/sim.3697.

[8] S.-L. T. Normand *et al.*, "Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores," *Journal of Clinical Epidemiology*, vol. 54, no. 4, pp. 387–398, Apr. 2001, doi: 10.1016/S0895-4356(00)00321-8.

[9] J. P. T. Higgins and S. G. Thompson, "Quantifying heterogeneity in a meta-analysis," *Statistics in Medicine*, vol. 21, no. 11, pp. 1539–1558, Jun. 2002, doi: 10.1002/sim.1186.

[10] V. Essebag, R. W. Platt, M. Abrahamowicz, and L. Pilote, "Comparison of nested case-control and survival analysis methodologies for analysis of time-dependent exposure," *BMC Medical Research Methodology*, vol. 5, no. 1, p. 5, Jan. 2005, doi: 10.1186/1471-2288-5-5.

[11] S. R. Cole and M. A. Hernán, "Constructing Inverse Probability Weights for Marginal Structural Models," *Am J Epidemiol*, vol. 168, no. 6, pp. 656–664, Sep. 2008, doi: 10.1093/aje/kwn164.

[12] "R: Regulatory Compliance and Validation Issues: A Guidance Document for the Use of R in Regulated Clinical Trial Environments." www.r-project.org/doc/R-FDA.pdf

Annex I

Diagnosis and operation codes.

The list of diagnosis and operation codes will be provided in separate documents.

- Annex 1.1 - Diagnosis and operation codes for CPRD Gold

Annex 1.1.zip

- Annex 1.2 - Diagnosis and operation codes for CPRD Aurum

Annex 1.2.zip

- Annex 1.3 - Diagnosis and operation codes for Finland and Sweden

Annex 1.3.xlsx

Annex II

ATC codes used to define treatments

The corresponding READ codes used to define treatments for CPRD will be provided in a separate document.

- Annex 2.1 - READ codes for treatment definition for CPRD Gold

Annex 2.1.zip

- Annex 2.2 - READ codes for treatment definition for CPRD Aurum

Annex 2.2.zip

Drug Group	ATC codes
Empagliflozin alone or in combination with metformin; excluding fixed dose combinations with drugs other than empagliflozin	
Empagliflozin	A10BK03 (previously A10BX12)
metformin and empagliflozin	A10BD20
Other SGLT-2 inhibitors; excluding empagliflozin and excluding fixed-dose combinations with DPP-4 inhibitors	
Dapagliflozin	A10BK01 (previously: A10BX09)
Canagliflozin	A10BK02 (previously: A10BX11)
Ertugliflozin	A10BK04
Ipragliflozin	A10BK05
metformin and dapagliflozin	A10BD15
metformin and canagliflozin	A10BD16
Metformin and ertugliflozin	A10BD23
DPP-4 inhibitors; excluding fixed-dose combinations with SGLT-2 inhibitors	
Sitagliptin	A10BH01
Vildagliptin	A10BH02
Saxagliptin	A10BH03
Alogliptin	A10BH04
Linagliptin	A10BH05
Evogliptin	A10BH07
Gemigliptin	A10BH06
sitagliptin and simvastatin	A10BH51
gemigliptin and rosuvastatin	A10BH52
metformin and sitagliptin	A10BD07
metformin and vildagliptin	A10BD08
pioglitazone and alogliptin	A10BD09
metformin and saxagliptin	A10BD10
metformin and linagliptin	A10BD11
pioglitazone and sitagliptin	A10BD12
metformin and alogliptin	A10BD13
metformin and gemigliptin	A10BD18
metformin and evogliptin	A10BD22
Fixed dose combinations of SGLT-2 inhibitors with DPP-4 inhibitors	
linagliptin and empagliflozin	A10BD19
saxagliptin and dapagliflozin	A10BD21
Sitagliptin and ertugliflozin	A10BD24
metformin, saxagliptin and dapagliflozin	A10BD25
Insulin	
Insulins and analogues	A10A
Metformin (alone or in fixed dose combinations)	
Metformin	A10BA02
metformin and sulfonylureas	A10BD02
metformin and rosiglitazone	A10BD03
metformin and pioglitazone	A10BD05
metformin and sitagliptin	A10BD07
metformin and vildagliptin	A10BD08

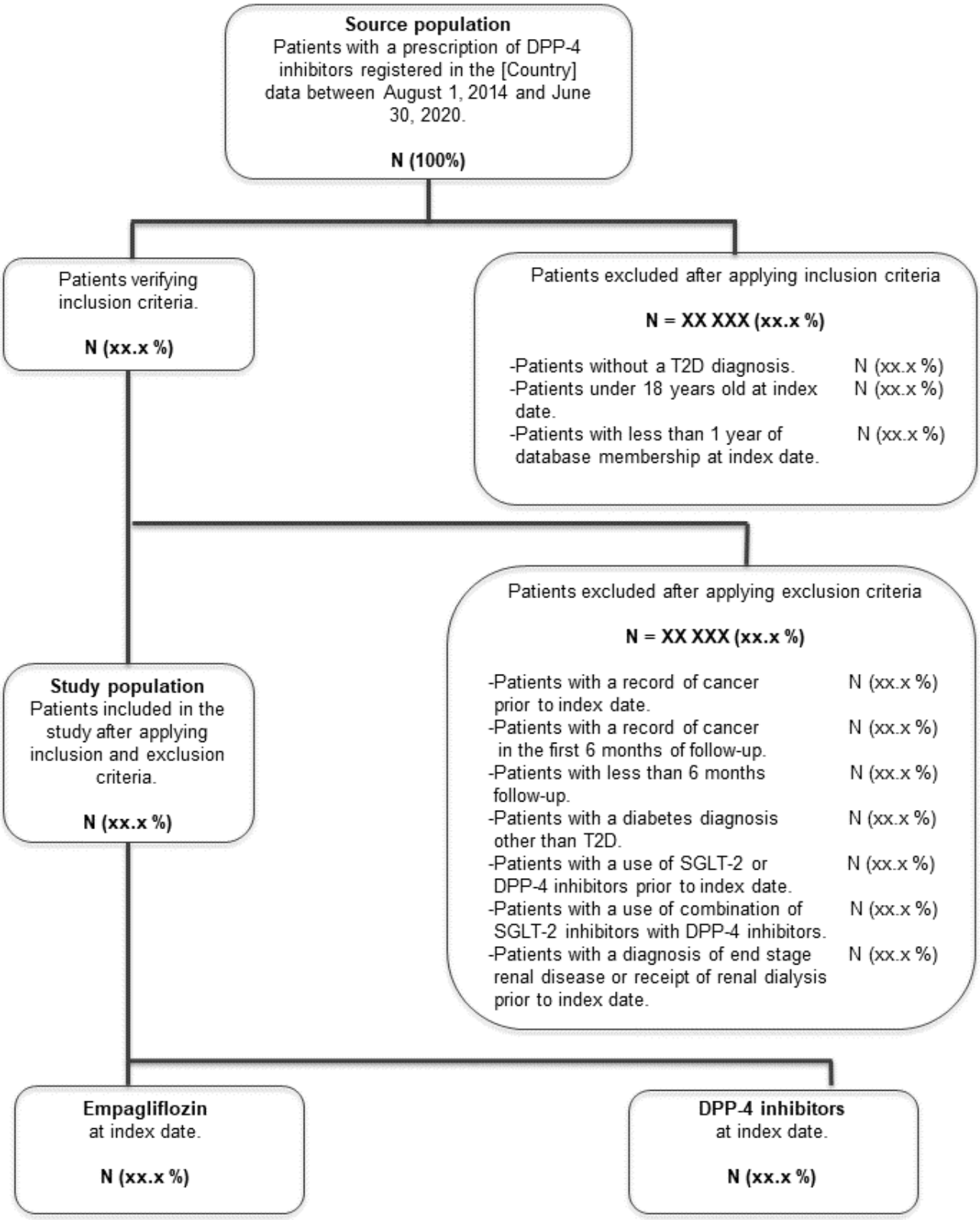
metformin and saxagliptin	A10BD10
metformin and linagliptin	A10BD11
metformin and alogliptin	A10BD13
metformin and repaglinide	A10BD14
metformin and dapagliflozin	A10BD15
metformin and canagliflozin	A10BD16
metformin and acarbose	A10BD17
metformin and gemigliptin	A10BD18
metformin and empagliflozin	A10BD20
metformin and evogliptin	A10BD22
metformin and ertugliflozin	A10BD23
metformin, saxagliptin and dapagliflozin	A10BD25
Other glucose lowering medication	
exenatide	A10BX04
liraglutide	A10BX07
lixisenatide	A10BX10
albiglutide	A10BX13
repaglinide	A10BX02
nateglinide	A10BX03
mitiglinide	A10BX08
metformin and repaglinide	A10BD14
pramlintide	A10BX05
benfluorex	A10BX06
guar gum	A10BX01
GLP-1 drugs	
Glucagon-like peptide-1 receptor (GLP-1) analogues	A10BJ
Thiazolidinediones	
Thiazolidinediones	A10BG
glimepiride and rosiglitazone	A10BD04
metformin and rosiglitazone	A10BD03
glimepiride and pioglitazone	A10BD06
metformin and pioglitazone	A10BD05
pioglitazone and alogliptin	A10BD09
pioglitazone and sitagliptin	A10BD12
Sulfonylureas	
Sulfonylureas	A10BB
phenformin and sulfonylreas	A10BD01
metformin and sulfonylreas	A10BD02
glimepiride and rosiglitazone	A10BD04
glimepiride and pioglitazone	A10BD06
glymidine	A10BC01
tolbutamide	V04CA01
Antihypertensives/diuretics	
Antihypertensives	C02
Diuretics	C03
Beta blocking agents	C07
Calcium channel blockers	C08
Agents acting on the renin-angiotensin system	C09
minoxidil	D11AX01
clonidine	N02CX02
clonidine	S01EA04

guanethidine	S01EX01
theobromin	R03DA07
theobromine, combinations	R03DA57
diazoxide	V03AH01
Non-steroidal anti-inflammatory drugs (NSAIDs)	
Antiinflammatory and antirheumatic products, non-steroids	M01A
Oral steroids	
Corticosteroids for systemic use, plain	H02A
Corticosteroids for systemic use, combinations	H02B
Statins, fibrates	
HMG CoA reductase inhibitors	C10AA
Fibrates	C10AB
HMG CoA reductase inhibitors in combination with other lipid modifying agents	C10BA
HMG CoA reductase inhibitors, other combinations	C10BX
sitagliptin and simvastatin	A10BH51
Lipid modifying agents	
Lipid modifying agents	C10
sitagliptin and simvastatin	A10BH51
nicotinic acid	C04AC01
nicotinyl alcohol (pyridylcarbinol)	C04AC02
Zoledronic acid	
Zoledronic acid	M05BA08
zoledronic acid, calcium and colecalciferol, sequential	M05BB08
Antibiotics	
Antibacterials for systemic use	J01

Annex III

Tables and figures templates

Figure 2 Flowchart of the study population and study subgroups identification in [Study country].



Abbreviations: DPP-4, dipeptidyl peptidase-4; T2DM, Type 2 diabetes mellitus.

Definitions: Index date, the date of first purchase/prescription of empagliflozin or DPP-4 between August 1, 2014 and June 30, 2020.

Table 17 Patients baseline characteristics in each of the study groups and in the total study population in [Study country].

	Empagliflozin at index date **N**	**DPP-4 inhibitors at index date** **N**	**[Country] Study population (Total)**
Age at index (years)			
18-49	N (x.xx%)	N (x.xx%)	N (x.xx%)
50-64	N (x.xx%)	N (x.xx%)	N (x.xx%)
65-69	N (x.xx%)	N (x.xx%)	N (x.xx%)
70-74	N (x.xx%)	N (x.xx%)	N (x.xx%)
75-79	N (x.xx%)	N (x.xx%)	N (x.xx%)
80-84	N (x.xx%)	N (x.xx%)	N (x.xx%)
85 and over	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Sex			
Male	N (x.xx%)	N (x.xx%)	N (x.xx%)
Female	N (x.xx%)	N (x.xx%)	N (x.xx%)
Highest level of education			
Compulsory	N (x.xx%)	N (x.xx%)	N (x.xx%)
Upper secondary	N (x.xx%)	N (x.xx%)	N (x.xx%)
Post secondary	N (x.xx%)	N (x.xx%)	N (x.xx%)
Disposable income, Individual			
< Q1	N (x.xx%)	N (x.xx%)	N (x.xx%)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
Q1 – Q2	N (x.xx%)	N (x.xx%)	N (x.xx%)
Q2 – Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Disposable income, family			
< Q1	N (x.xx%)	N (x.xx%)	N (x.xx%)
Q1 – Q2	N (x.xx%)	N (x.xx%)	N (x.xx%)
Q2 – Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Marital status			
Registered partnership (incl. married)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Widowed	N (x.xx%)	N (x.xx%)	N (x.xx%)
Separated	N (x.xx%)	N (x.xx%)	N (x.xx%)
Single	N (x.xx%)	N (x.xx%)	N (x.xx%)
Calendar year of index date			
2014	N (x.xx%)	N (x.xx%)	N (x.xx%)
2015	N (x.xx%)	N (x.xx%)	N (x.xx%)
2016	N (x.xx%)	N (x.xx%)	N (x.xx%)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
2017	N (x.xx%)	N (x.xx%)	N (x.xx%)
2018	N (x.xx%)	N (x.xx%)	N (x.xx%)
2019	N (x.xx%)	N (x.xx%)	N (x.xx%)
2020	N (x.xx%)	N (x.xx%)	N (x.xx%)
2021	N (x.xx%)	N (x.xx%)	N (x.xx%)
Duration of look-back period			
1	N (x.xx%)	N (x.xx%)	N (x.xx%)
2-5	N (x.xx%)	N (x.xx%)	N (x.xx%)
6-9	N (x.xx%)	N (x.xx%)	N (x.xx%)
10 and over	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Time since first diabetes diagnosis at index date			
<Q1	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ (x<Q2	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ (x<Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ (	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)

	Empagliflozin at index date **N**	**DPP-4 inhibitors at index date** **N**	**[Country] Study population (Total)**
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Duration of treated diabetes at index date			
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Diabetes treatment complexity at index date			
Monotherapy	N (x.xx%)	N (x.xx%)	N (x.xx%)
Dual combination therapy	N (x.xx%)	N (x.xx%)	N (x.xx%)
Triple combination therapy	N (x.xx%)	N (x.xx%)	N (x.xx%)
Insulin use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Insulin use at index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Metformin use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Metformin use at index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Oral glucose lowering drugs use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Oral glucose lowering drugs use at index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Sulfonylureas use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Sulfonylureas use at index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Thiazolidinediones use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Thiazolidinediones use at index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
GLP-1 use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
GLP-1 use at index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic complications prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic retinopathy or diabetic maculopathy prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic nephropathy prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic neuropathy prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
History of Diabetic lower limb severe complications prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of peripheral vascular disease prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of kidney or genitourinary stones prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of UTI or pyelonephritis prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of liver disease prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of pancreatitis prior to index date			

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of congestive heart failure prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of hypertension prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of stroke prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of myocardial infarction prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of autoimmune disease prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of COPD prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Any ICU admission prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Antihypertensives/ diuretics use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
NSAIDs use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Oral steroids use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Statins/fibrates use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Lipid modifying agents use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Zoledronic acid use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Antibiotic use prior to index date			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
BMI (closest value prior to index date)			
<20	N (x.xx%)	N (x.xx%)	N (x.xx%)
20-24.9	N (x.xx%)	N (x.xx%)	N (x.xx%)
25-29.9	N (x.xx%)	N (x.xx%)	N (x.xx%)
>30	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Smoking status (closest value prior to index date)			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Alcohol use (closest value prior to index date)			

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
HbA1c in mmol/mol(closest value prior to index date)			
< 48 (6.5%) mmol/mol	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ 48 (6.5 %) - <64 mol/mol (8 %)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ 64 mmol/mol (8 %)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
eGFR in (mL/min/1.73 m2)			
<15 (ml/min/1.73 m2) (kidney failure)	N (x.xx%)	N (x.xx%)	N (x.xx%)
15-29 (ml/min/1.73 m2) (severely decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)
30-44 (ml/min/1.73 m2) (Moderately to severely decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)
45-59 (ml/min/1.73 m2) (Mildly to moderately decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)
60-89 (ml/min/1.73 m2) (Mildly decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥90 (ml/min/1.73 m2) (Normal)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
mean (± sd) median (Q1,Q3)	mean (± sd) median (Q1,Q3)	mean (± sd) median (Q1,Q3)	mean (± sd) median (Q1,Q3)
Renal impairment			
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)
Systolic blood pressure			
< 120	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥120-<139	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥140-<159	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥160	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Diastolic blood pressure			
< 80	N (x.xx%)	N (x.xx%)	N (x.xx%)
80-89	N (x.xx%)	N (x.xx%)	N (x.xx%)
90-99	N (x.xx%)	N (x.xx%)	N (x.xx%)
100 or higher	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)

	Empagliflozin at index date N	DPP-4 inhibitors at index date N	[Country] Study population (Total)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)

Table 18 Covariate balance assessment [before/after] matching comparing empagliflozin group to DPP-4 inhibitors group – [Study country].

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
Age at index (years)				
18-49	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
50-64	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
65-69	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
70-74	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
75-79	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
80-84	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
85 and over	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Sex				
Male	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Female	N (x.xx%)	N (x.xx%)		
Highest level of education				
Compulsory	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Upper secondary	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Post secondary	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Disposable income, individual				
< Q1	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
Q1 – Q2	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Q2 – Q3	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
≥ Q3	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Disposable income, family				
< Q1	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Q1 – Q2	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Q2 – Q3	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
≥ Q3	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Marital status				
Registered partnership (incl. married)	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Widowed	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Separated	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Single	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Calendar year of index date				
2014	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
2015	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
2016	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
2017	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
2018	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
2019	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
2020	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
2021	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Duration of look-back period				
1	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
2-5	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
6-9	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
10 and over	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Time since first diabetes diagnosis at index date				
<Q1	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
≥Q1-<Q2	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
≥Q2-<Q3	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
≥Q4	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff

	Empagliflozin users at index date [before/after] matching N	**DPP-4 inhibitors users at index date [before/after] matching N**	**P-Value**	**Standardized differences**
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Duration of treated diabetes at index date				
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Diabetes treatment complexity at index date				
Monotherapy	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Dual combination therapy	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Triple combination therapy	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Insulin use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Metformin use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Oral glucose lowering drugs use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Sulfonylureas use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Thiazolidinediones use at index date				

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
GLP-1 use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of diabetic complications prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of diabetic retinopathy or diabetic maculopathy prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of diabetic nephropathy prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of diabetic neuropathy prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
History of Diabetic lower limb severe complications prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of peripheral vascular disease prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of kidney or genitourinary stones prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of UTI or pyelonephritis prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of liver disease prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of pancreatitis prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
History of congestive heart failure prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of hypertension prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of stroke prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of myocardial infarction prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of autoimmune disease prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
History of COPD prior to index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Any ICU admission prior to index date				

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Antihypertensives/ diuretics use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
NSAIDs use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Oral steroids use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Statins/fibrates use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Lipid modifying agents use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Zoledronic acid use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
Antibiotic use at index date				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
BMI (closest value prior to index date)				
<20	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
20-24.9	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
25-29.9	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
>30	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Smoking status (closest value prior to index date)				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		
Alcohol use (closest value prior to index date)				
Yes	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
No	N (x.xx%)	N (x.xx%)		

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
HbA1c mmol/mol (closest value prior to index date)				
< 48 (6.5%) mmol/mol	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
≥ 48 (6.5 %) - <64 mol/mol (8 %)	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
≥ 64 mmol/mol (8 %)	N (x.xx%)	N (x.xx%)	χ2 test p-value	std diff
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
Number of albuminuria tests at index date				
Range (min,max)	(min, max)	(min, max)		
mean (± sd)	mean (± sd)	mean (± sd)	t-test p-value	std diff
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)		
eGFR in (mL/min/1.73 m2)				
<15 (ml/min/1.73 m2) (kidney failure)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
15-29 (ml/min/1.73 m2) (severely decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
30-44 (ml/min/1.73 m2) (Moderately to severely decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
45-59 (ml/min/1.73 m2) (Mildly to moderately decreased)				
60-89 (ml/min/1.73 m2) (Mildly decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ ((ml/min/1.73 m2) (Normal)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)

	Empagliflozin users at index date [before/after] matching N	DPP-4 inhibitors users at index date [before/after] matching N	P-Value	Standardized differences
mean (± sd) median (Q1,Q3)	mean (± sd) median (Q1,Q3)	mean (± sd) median (Q1,Q3)	mean (± sd) median (Q1,Q3)	mean (± sd) median (Q1,Q3)
Renal impairment				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Systolic blood pressure				
< 120	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥120-<139	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥140-<159	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥160	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Diastolic blood pressure				
< 80	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
80-89	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
90-99	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
100 or higher	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)

Figure 3 Propensity score distribution in empagliflozin users at index and DPP-4 inhibitors users before matching – [Study country], [Cohort].

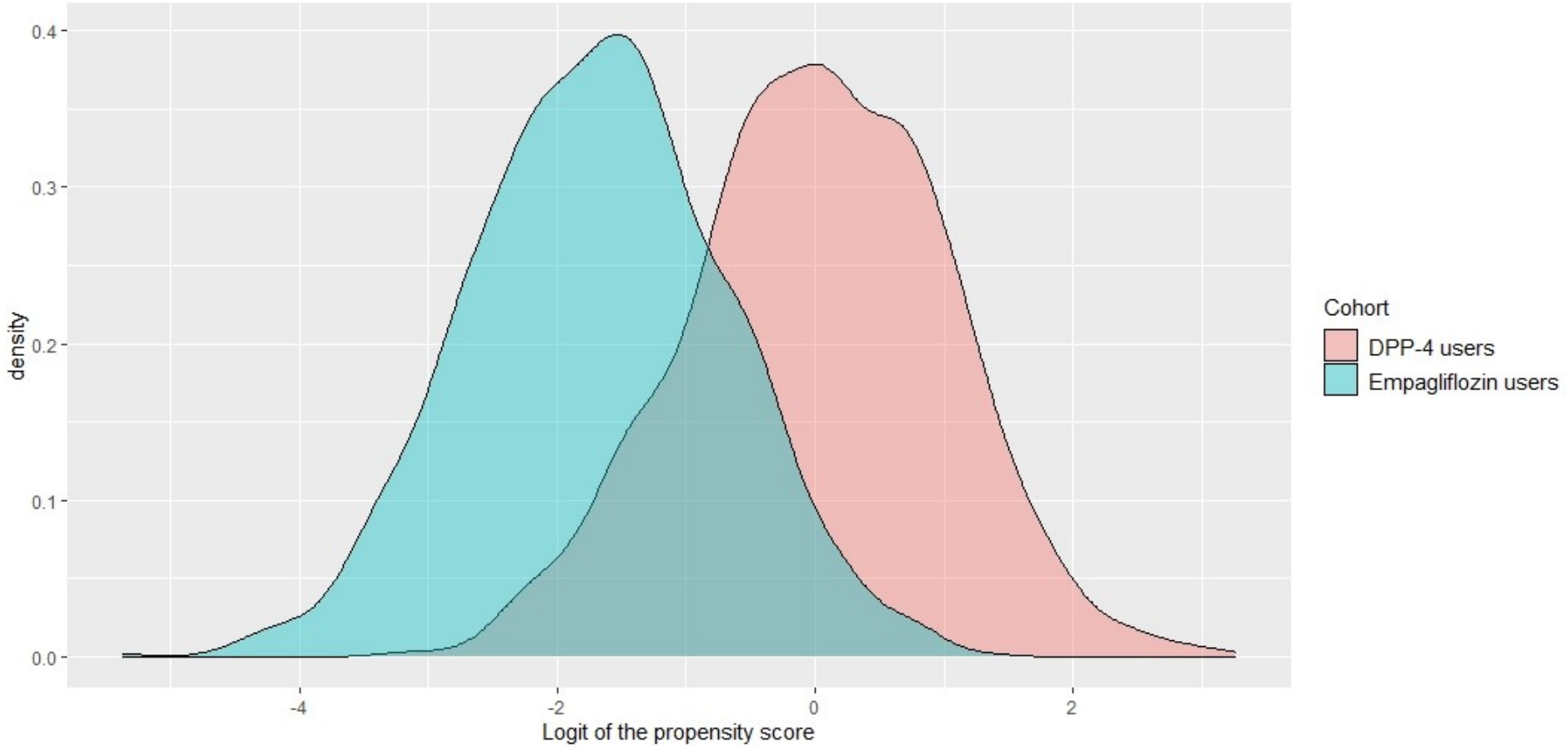


Figure 4 Standardized differences for baseline covariates comparing patients using empagliflozin at index and DPP-4 inhibitors users at index date before and after matching – [Study country].

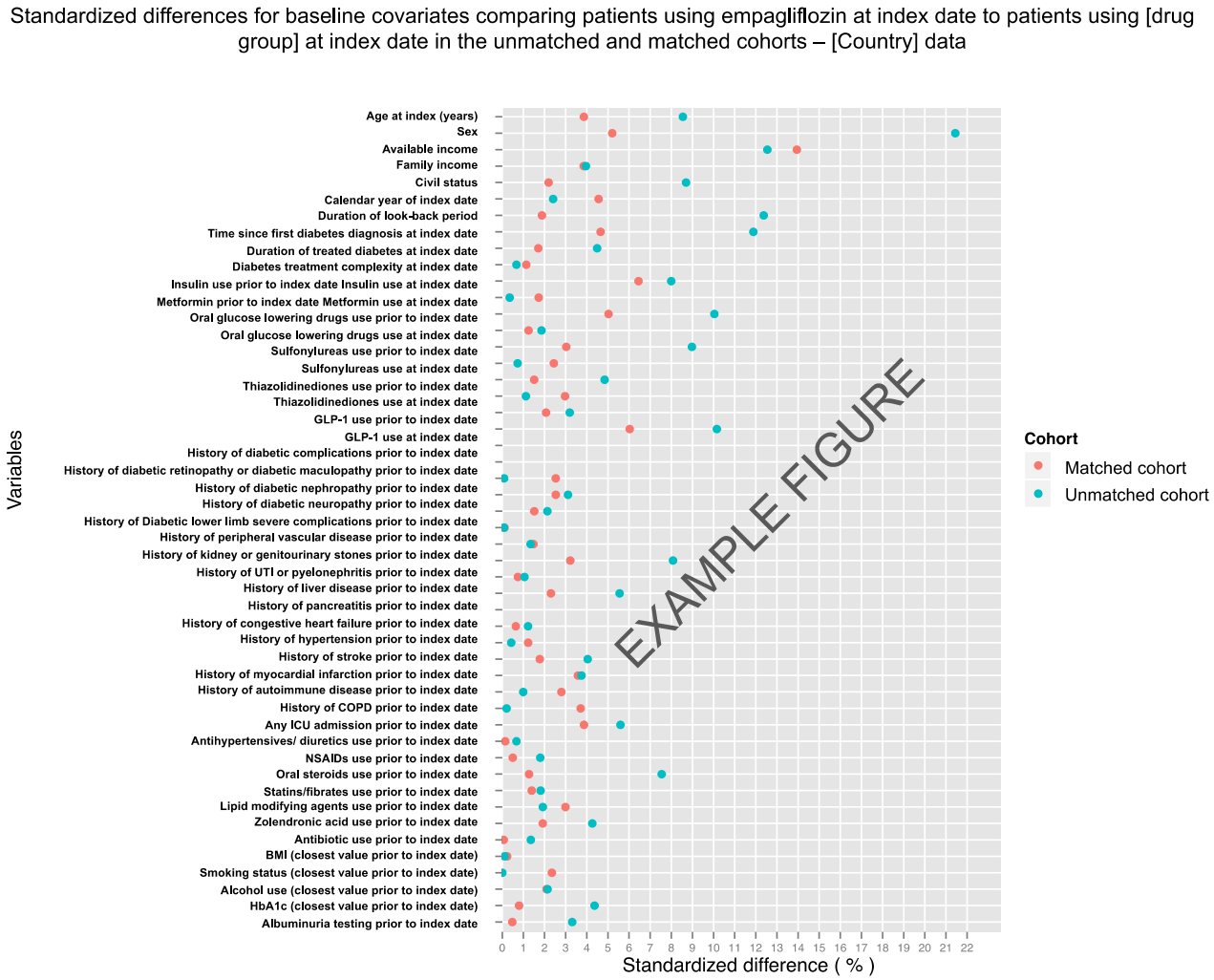


Figure 5 Propensity score distribution in non-matched empagliflozin users at index date, matched empagliflozin users at index and matched DPP-4 inhibitors users at index – [Study country], [Cohort].

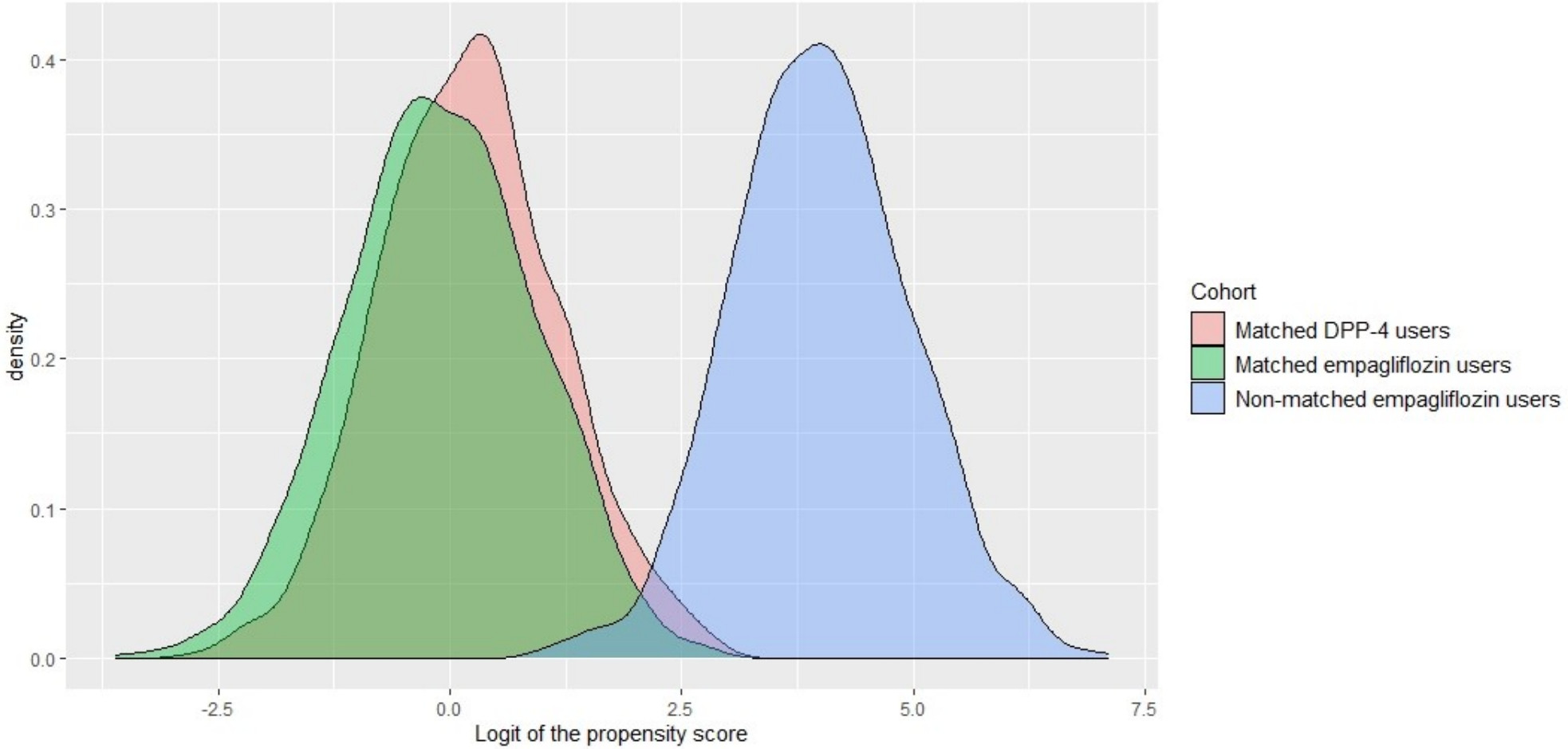


Table 19 Propensity score distribution before and after propensity score matching of empagliflozin users and DPP-4 inhibitors users by decile in [country]

Propensity score	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9
Before matching									
Empagliflozin users	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
DPP-4 inhibitors users	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
After matching									
Matched empagliflozin users	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx
Matched DPP-4 inhibitors users	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Table 20 Number of patients with incidence of cancer within 12 months after index date

	Empagliflozin users at index date	DPP-4 inhibitors users at index date
Number of cancers within 12 months after index date.	N (x.xx%)	N (x.xx%)

Table 21 TNM staging of [Study outcome] at diagnosis during follow-up after exclusion of cancers occurring within 6 months after index date.

	[Study outcome] cases in empagliflozin users at index date	[Study outcome] cases in DPP-4 inhibitors users at index date
TNM staging (T)		
T0	N (x.xx%)	N (x.xx%)
T1	N (x.xx%)	N (x.xx%)
T1a	N (x.xx%)	N (x.xx%)
T1b	N (x.xx%)	N (x.xx%)
T1c	N (x.xx%)	N (x.xx%)
T2	N (x.xx%)	N (x.xx%)
T3	N (x.xx%)	N (x.xx%)
T4	N (x.xx%)	N (x.xx%)
TX	N (x.xx%)	N (x.xx%)
NA	N (x.xx%)	N (x.xx%)
TNM staging (N)		
N0	N (x.xx%)	N (x.xx%)
N1	N (x.xx%)	N (x.xx%)
NX	N (x.xx%)	N (x.xx%)
NA	N (x.xx%)	N (x.xx%)
TNM staging (M)		
M0	N (x.xx%)	N (x.xx%)
M1	N (x.xx%)	N (x.xx%)
MX	N (x.xx%)	N (x.xx%)
NA	N (x.xx%)	N (x.xx%)

Table 22 Incidence rates per 1000 patient years of [Outcome name] with 95% confidence interval in the [Study country/pooled] matched cohort(s). The number of patients in each strata and the total number of person years are also presented.

Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Absolute incidence						
Total	N	N	N	x.xx	x.xx	x.xx
Exposure category (As treated)						
Empagliflozin	N	N	N	x.xx	x.xx	x.xx
DPP-4 inhibitors	N	N	N	x.xx	x.xx	x.xx
Age at index (years)						
18-49	N	N	N	x.xx	x.xx	x.xx
50-64	N	N	N	x.xx	x.xx	x.xx
65-69	N	N	N	x.xx	x.xx	x.xx
70-74	N	N	N	x.xx	x.xx	x.xx
75-79	N	N	N	x.xx	x.xx	x.xx
80-84	N	N	N	x.xx	x.xx	x.xx
85 and over	N	N	N	x.xx	x.xx	x.xx
Sex						
Male	N	N	N	x.xx	x.xx	x.xx
Female	N	N	N	x.xx	x.xx	x.xx
Highest level of education						
Compulsory	N	N	N	x.xx	x.xx	x.xx
Upper secondary	N	N	N	x.xx	x.xx	x.xx
Post secondary	N	N	N	x.xx	x.xx	x.xx
Disposable income, individual						
< Q1	N	N	N	x.xx	x.xx	x.xx
Q1 – Q2	N	N	N	x.xx	x.xx	x.xx
Q2 – Q3	N	N	N	x.xx	x.xx	x.xx
≥ Q3	N	N	N	x.xx	x.xx	x.xx
Disposable income, family						
< Q1	N	N	N	x.xx	x.xx	x.xx
Q1 – Q2	N	N	N	x.xx	x.xx	x.xx
Q2 – Q3	N	N	N	x.xx	x.xx	x.xx
≥ Q3	N	N	N	x.xx	x.xx	x.xx

Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Marital status						
Registered partnership (incl. married)	N	N	N	x.xx	x.xx	x.xx
Widowed	N	N	N	x.xx	x.xx	x.xx
Separated	N	N	N	x.xx	x.xx	x.xx
Single	N	N	N	x.xx	x.xx	x.xx
Calendar year of index date						
2014	N	N	N	x.xx	x.xx	x.xx
2015	N	N	N	x.xx	x.xx	x.xx
2016	N	N	N	x.xx	x.xx	x.xx
2017	N	N	N	x.xx	x.xx	x.xx
2018	N	N	N	x.xx	x.xx	x.xx
2019	N	N	N	x.xx	x.xx	x.xx
2020	N	N	N	x.xx	x.xx	x.xx
Diabetes treatment complexity at index date						
Monotherapy	N	N	N	x.xx	x.xx	x.xx
Dual combination therapy	N	N	N	x.xx	x.xx	x.xx
Triple combination therapy	N	N	N	x.xx	x.xx	x.xx
Insulin use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Metformin use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Oral glucose lowering drugs use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Sulfonylureas use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Thiazolidinediones use						
Yes	N	N	N	x.xx	x.xx	x.xx

Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
No	N	N	N	x.xx	x.xx	x.xx
GLP-1 use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Diabetic complications						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Diabetic retinopathy or diabetic maculopathy						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Diabetic nephropathy						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Diabetic neuropathy						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Diabetic lower limb severe complications						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Peripheral vascular disease						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Kidney or genitourinary stones						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
UTI or pyelonephritis						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Liver disease						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Pancreatitis						

Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Congestive heart failure						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Hypertension						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Stroke						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Myocardial infarction						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Autoimmune disease						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
COPD						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Any ICU admission prior to index date						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Antihypertensives/ diuretics use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
NSAIDs use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Oral steroids						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx

Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Statins/fibrates use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Lipid modifying agents use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Zoledronic acid use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Antibiotic use						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
BMI (closest value prior to index date)						
<20	N	N	N	x.xx	x.xx	x.xx
20-24.9	N	N	N	x.xx	x.xx	x.xx
25-29.9	N	N	N	x.xx	x.xx	x.xx
>30	N	N	N	x.xx	x.xx	x.xx
Smoking status (closest value prior to index date)						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Alcohol use (closest value prior to index date)						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
HbA1c (closest value prior to index date)						
< 48 (6.5%) mmol/mol	N	N	N	x.xx	x.xx	x.xx
≥ 48 (6.5 %) - <64 mol/mol (8 %)	N	N	N	x.xx	x.xx	x.xx
≥ 64 mmol/mol (8 %)	N	N	N	x.xx	x.xx	x.xx
eGFR in (mL/min/1.73 m2)						
<15 (ml/min/1.73 m2) (kidney failure)	N	N	N	x.xx	x.xx	x.xx

Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
15-29 (ml/min/1.73 m2) (severely decreased)	N	N	N	x.xx	x.xx	x.xx
30-44 (ml/min/1.73 m2) (Moderately to severely decreased)	N	N	N	x.xx	x.xx	x.xx
45-59 (ml/min/1.73 m2) (Mildly to moderately decreased)	N	N	N	x.xx	x.xx	x.xx
60-89 (ml/min/1.73 m2) (Mildly decreased)	N	N	N	x.xx	x.xx	x.xx
≥90 (ml/min/1.73 m2) (Normal)	N	N	N	x.xx	x.xx	x.xx
Renal impairment						
Yes	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx
Systolic blood pressure						
< 120	N	N	N	x.xx	x.xx	x.xx
≥120-<139	N	N	N	x.xx	x.xx	x.xx
≥140-<159	N	N	N	x.xx	x.xx	x.xx
≥160	N	N	N	x.xx	x.xx	x.xx
Diastolic blood pressure						
< 80	N	N	N	x.xx	x.xx	x.xx
80-89	N	N	N	x.xx	x.xx	x.xx
90-99	N	N	N	x.xx	x.xx	x.xx
100 or higher	N	N	N	x.xx	x.xx	x.xx

Figure 6 Kaplan-Meier survival probability estimates for [Outcome name] in the [Study country] matched cohort stratified by exposure using the as treated definition.

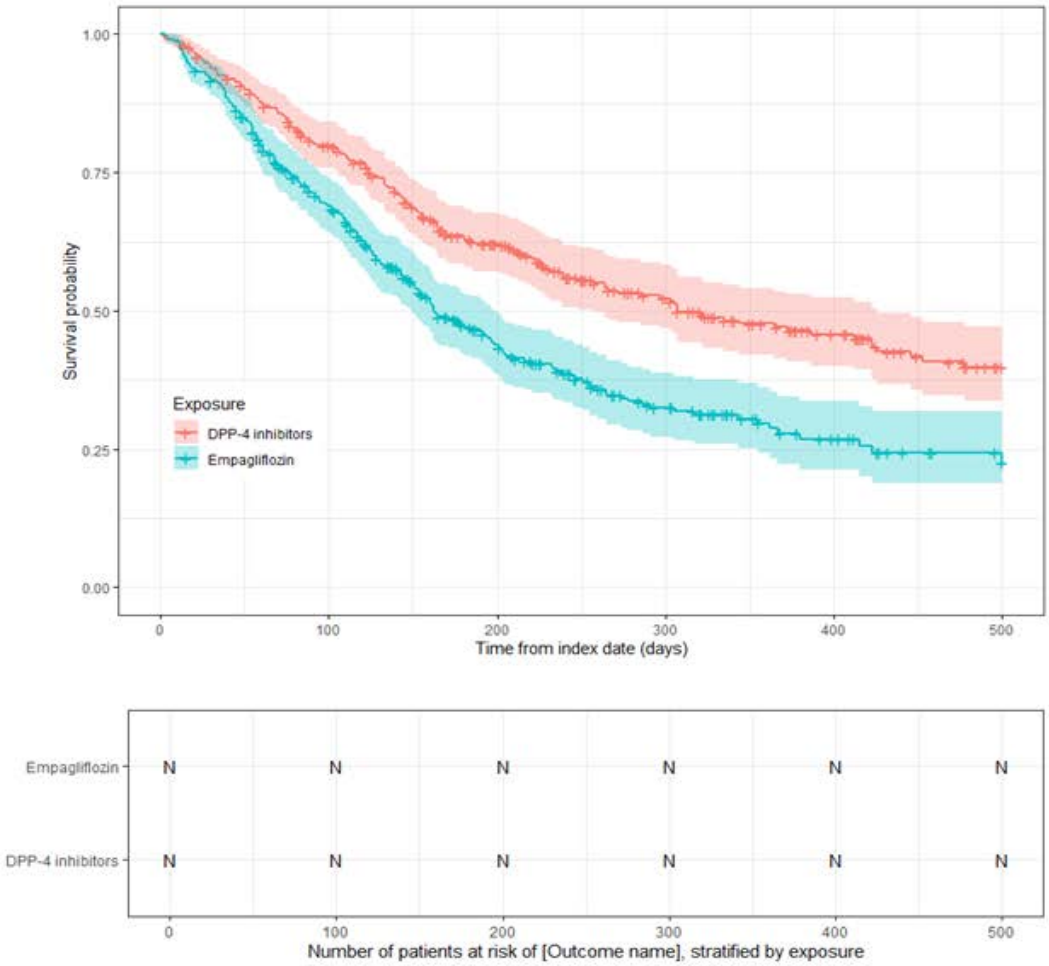


Table 23 Crude hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the [Study country] matched cohort. The table shows the hazard ratios (HRs) using the crude model with the corresponding 95% CI and p-value–As treated

Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure				
DPP-4 inhibitors	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx

Table 24 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the [Study country] matched cohort. The table shows the hazard ratios (HRs) using the base model with the corresponding 95% CI and p-value–As treated exposure definition.

Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure				
DPP-4 inhibitors	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx

Table 25 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the [Study country] matched cohort. The table shows the hazard ratios (HRs) using the adjusted model with the corresponding 95% CI and p-value–As treated exposure definition. [Stratifying variable if any]

Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure				
DPP-4 inhibitors	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx

Table 26 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the [Study country] matched cohort. The table shows the hazard ratios (HRs) using the adjusted model with the corresponding 95% CI and p-value–Cumulative dose of empagliflozin exposure definition.

Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
Cumulative dose of empagliflozin				
Never exposed	reference	reference	reference	reference
< Q1	x.xxx	x.xxx	x.xxx	x.xxx
Q1 – Q2	x.xxx	x.xxx	x.xxx	x.xxx
Q2 – Q3	x.xxx	x.xxx	x.xxx	x.xxx
≥ Q3	x.xxx	x.xxx	x.xxx	x.xxx

Table 27 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the [Study country] matched cohort. The table shows the hazard ratios (HRs) using the adjusted model with the corresponding 95% CI and p-value–daily dose of empagliflozin exposure definition.

Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
Daily dose of empagliflozin				
Never exposed	reference	reference	reference	reference
10 mg	x.xxx	x.xxx	x.xxx	x.xxx
25 mg	x.xxx	x.xxx	x.xxx	x.xxx

Table 28 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the [Study country] matched cohort. The table shows the hazard ratios (HRs) using the adjusted model with the corresponding 95% CI and p-value–Cumulative exposure duration to empagliflozin.

Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
Cumulative duration of empagliflozin				
Never exposed	reference	reference	reference	reference
< Q1	x.xxx	x.xxx	x.xxx	x.xxx
Q1 – Q2	x.xxx	x.xxx	x.xxx	x.xxx
Q2 – Q3	x.xxx	x.xxx	x.xxx	x.xxx
≥ Q3	x.xxx	x.xxx	x.xxx	x.xxx

Table 29 Incidence rates per 1000 patient years of [Outcome name] with 95% confidence interval in the [Study country] matched cohort. The number of patients in each strata and the total number of person years are also presented.

Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
AT exposure						
DPP-4 inhibitors	N	N	N	x.xx	x.xx	x.xx
Empagliflozin	N	N	N	x.xx	x.xx	x.xx

Table 30 Incidence rates per 1000 patient years of [Outcome name] with 95% confidence interval in the [Study country] matched cohort. The number of patients in each strata and the total number of person years are also presented.

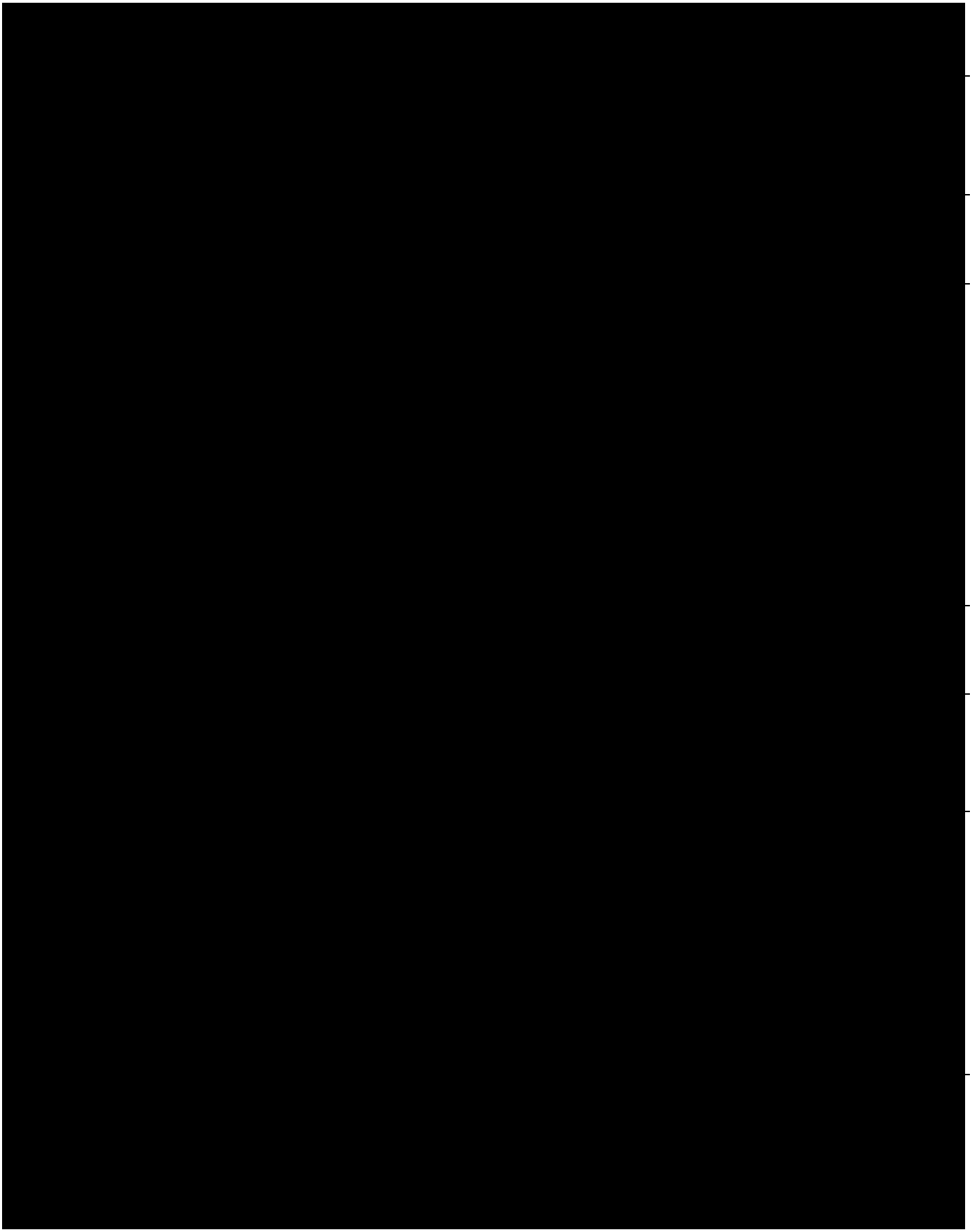
Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Cumulative dose of empagliflozin						
Never exposed	N	N	N	x.xx	x.xx	x.xx
< Q1	N	N	N	x.xx	x.xx	x.xx
Q1 – Q2	N	N	N	x.xx	x.xx	x.xx
Q2 – Q3	N	N	N	x.xx	x.xx	x.xx
≥ Q3	N	N	N	x.xx	x.xx	x.xx

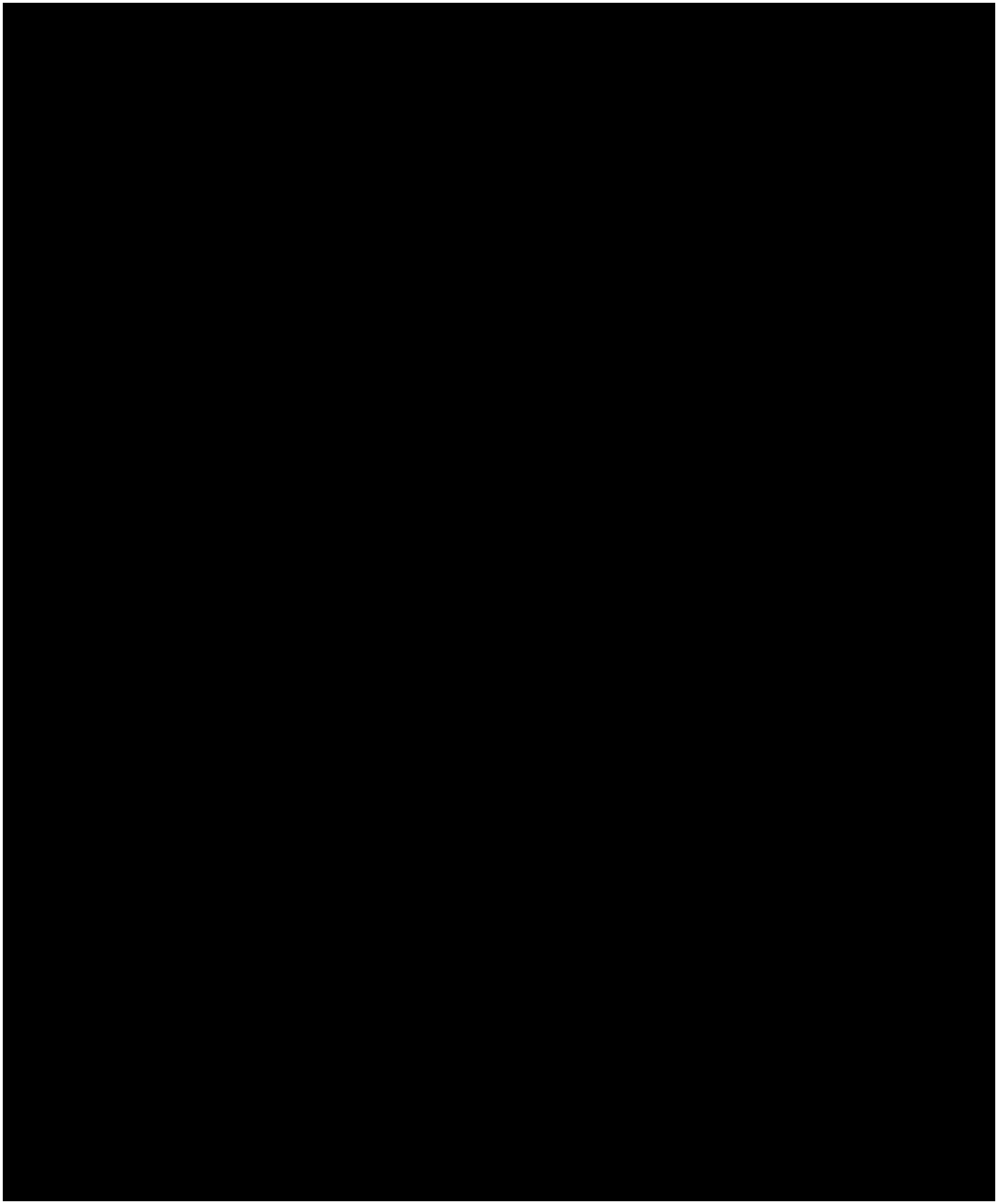
Table 31 Incidence rates per 1000 patient years of [Outcome name] with 95% confidence interval in the [Study country] matched cohort. The number of patients in each strata and the total number of person years are also presented.

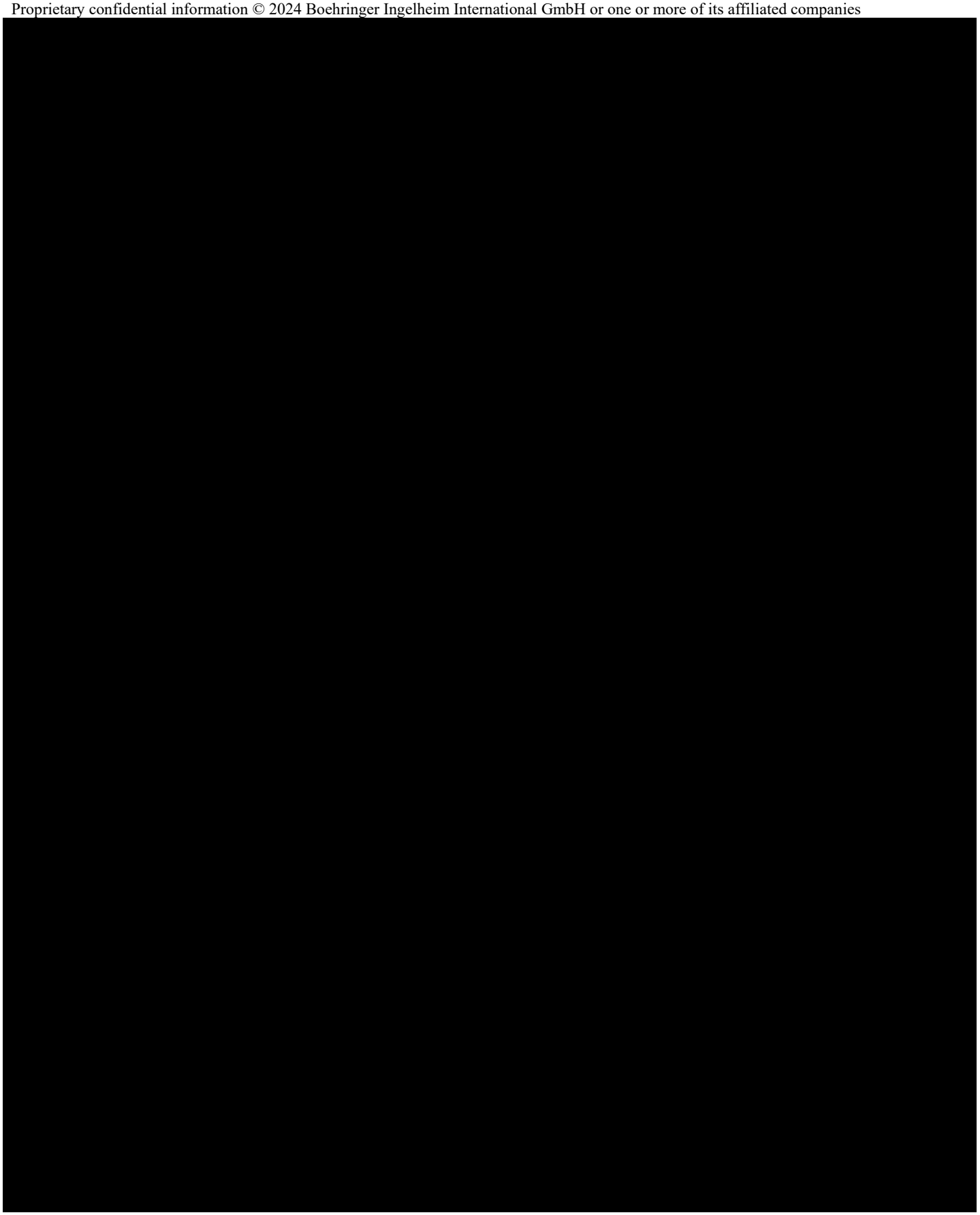
Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Daily dose of empagliflozin						
Never exposed	N	N	N	x.xx	x.xx	x.xx
10mg	N	N	N	x.xx	x.xx	x.xx
25mg	N	N	N	x.xx	x.xx	x.xx

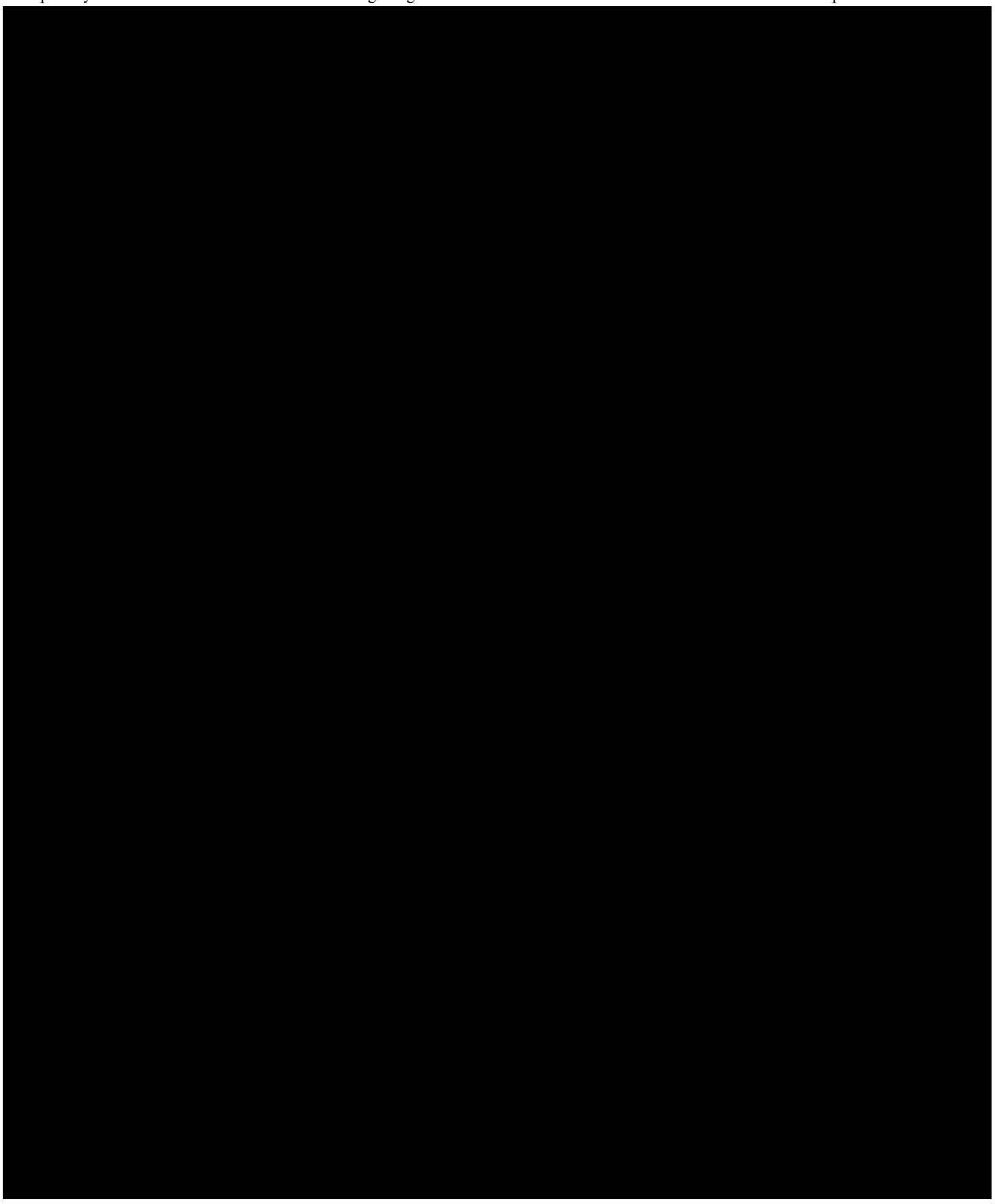
Table 32 Incidence rates per 1000 patient years of [Outcome name] with 95% confidence interval in the [Study country] matched cohort. The number of patients in each strata and the total number of person years are also presented.

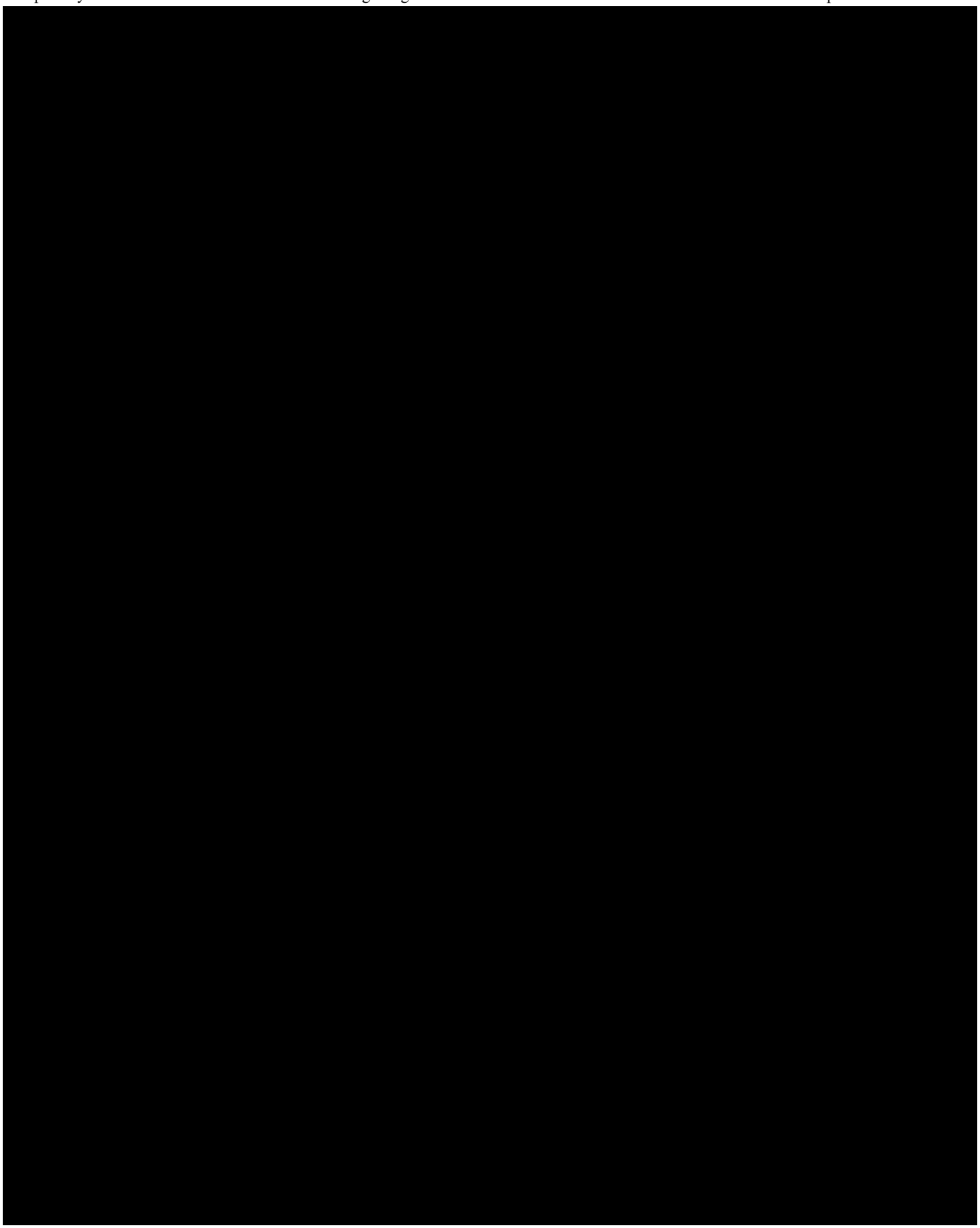
Variable and strata	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Cumulative exposure duration to empagliflozin						
Never exposed	N	N	N	x.xx	x.xx	x.xx
< Q1	N	N	N	x.xx	x.xx	x.xx
Q1 – Q2	N	N	N	x.xx	x.xx	x.xx
Q2 – Q3	N	N	N	x.xx	x.xx	x.xx
≥ Q3						

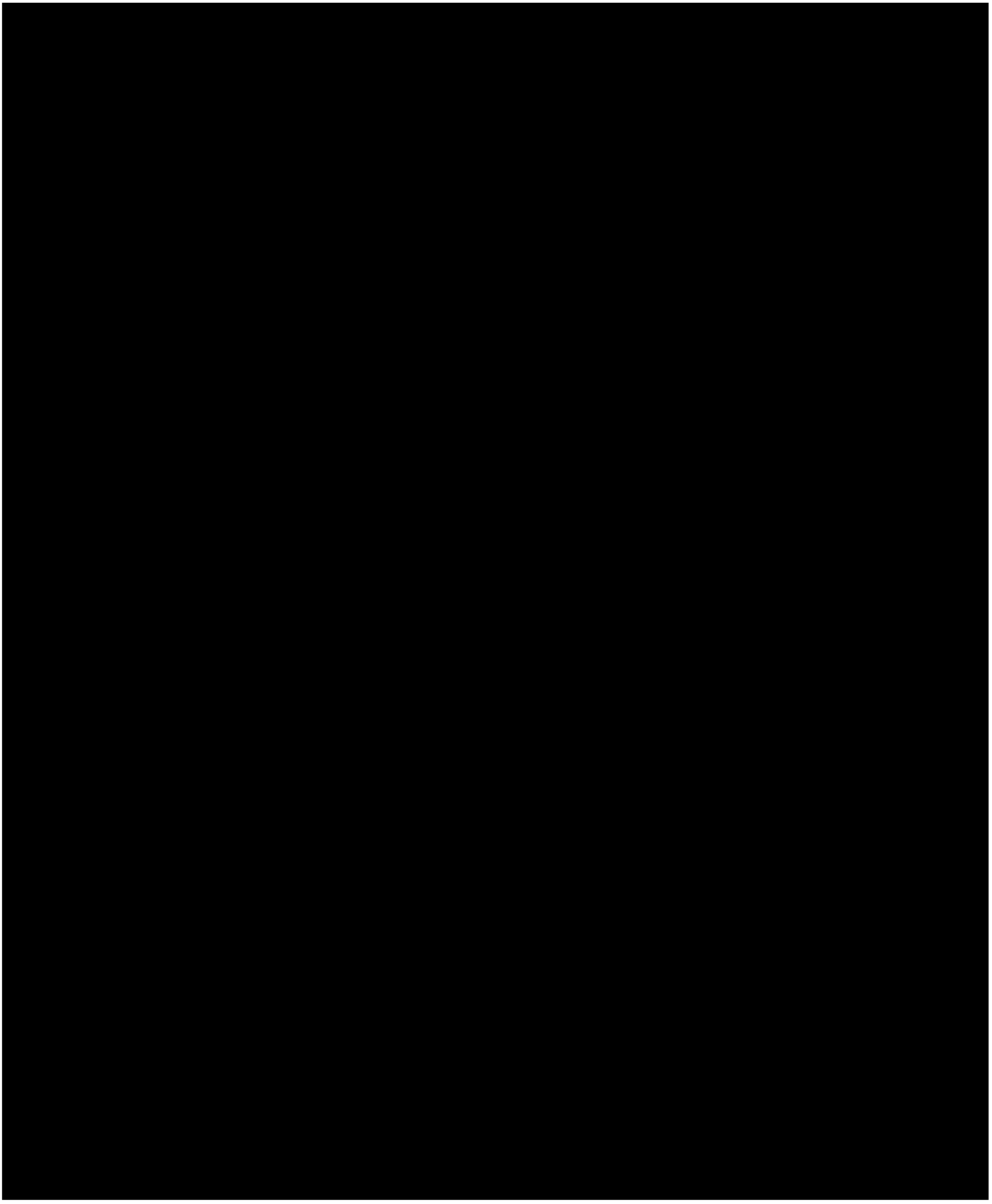


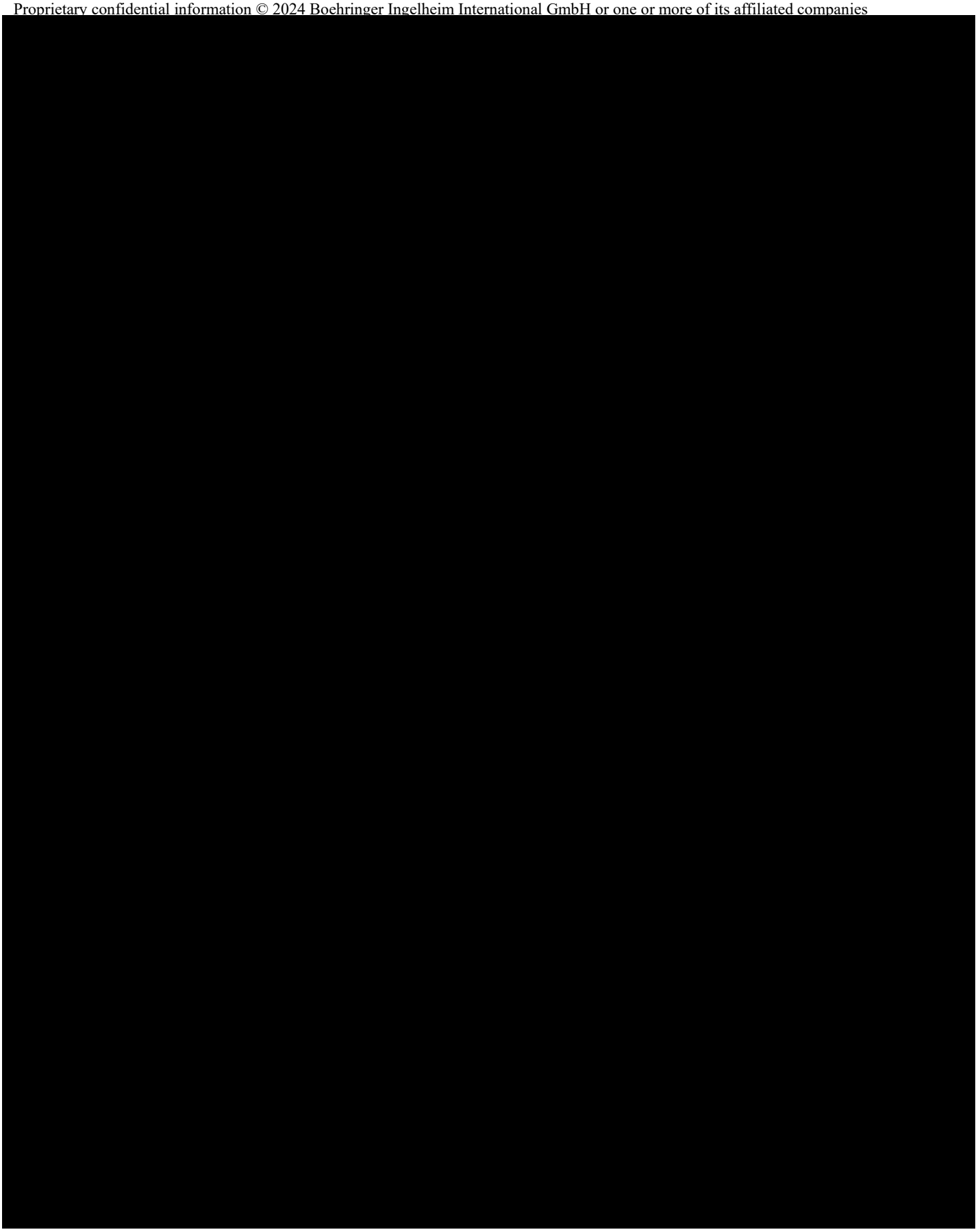












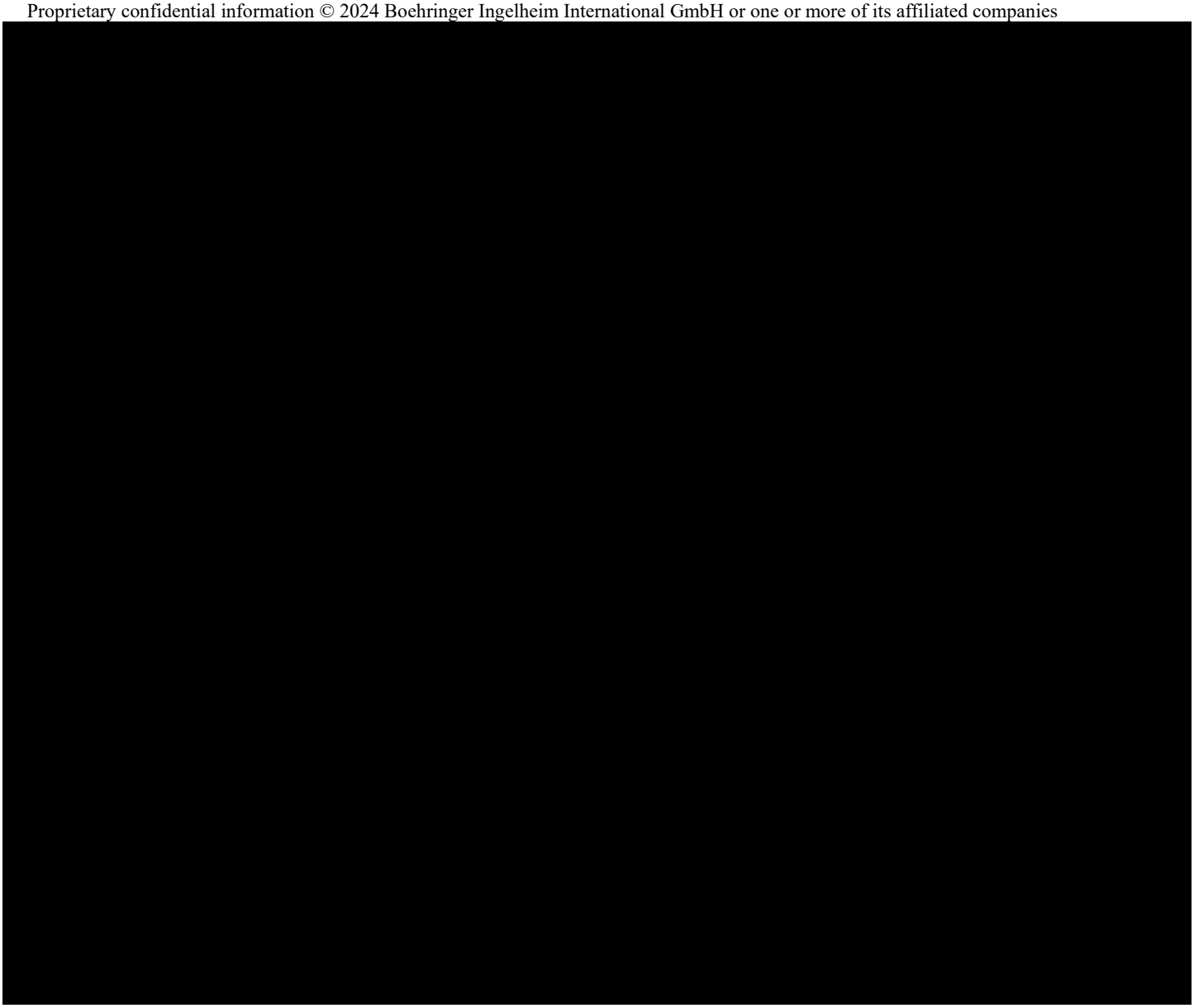


Figure 7. Distribution of the stabilized weights on log scale by time from index date for [drug group]

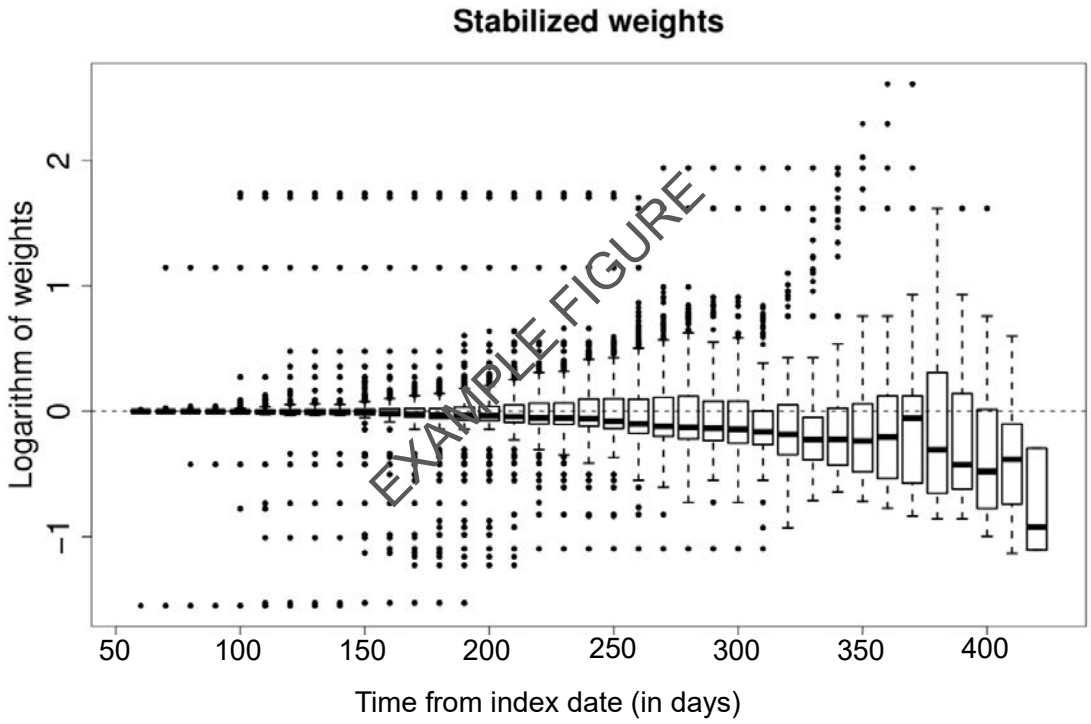


Table 36. Number of yearly albuminaria tests stratified by exposure group - [Study country]

	Empagliflozin	DPP-4 inhibitors
No. of tests		
0	N (x.xx%)	N (x.xx%)
1	N (x.xx%)	N (x.xx%)
2	N (x.xx%)	N (x.xx%)
More than 2	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)

Table 37 Patients baseline characteristics in each of the study groups in the UK, stratified by CPRD database.

	GOLD		Aurum	
	Empagliflozin N =	**DPP-4 inhibitors N =**	**Empagliflozin N =**	**DPP-4 inhibitors N =**
Age at index (years)				
18-49	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
50-64	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
65-69	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
70-74	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
75-79	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
80-84	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
85 and over	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Sex				
Male	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Female	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Highest level of education				
Compulsory	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Upper secondary	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Post secondary	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Disposable income, Individual				
< Q1	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Q1 – Q2	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Q2 – Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Disposable income, family				
< Q1	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Q1 – Q2	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Q2 – Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)

Range (min,max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)
Marital status				
Registered partnership (incl. married)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Widowed	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Separated	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Single	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Calendar year of index date				
2014	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2015	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2016	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2017	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2018	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2019	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2020	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2021	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Duration of look-back period				
1	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
2-5	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
6-9	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
10 and over	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)
Time since first diabetes diagnosis at index date				
<Q1	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ (x<Q2	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ (x<Q3	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ (	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)

Duration of treated diabetes at index date				
Range (min,max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)
Diabetes treatment complexity at index date				
Monotherapy	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Dual combination therapy	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Triple combination therapy	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Insulin use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Insulin use at index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Metformin use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Metformin use at index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Oral glucose lowering drugs use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Oral glucose lowering drugs use at index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)

Sulfonylureas use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Sulfonylureas use at index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Thiazolidinediones use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Thiazolidinediones use at index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
GLP-1 use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
GLP-1 use at index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic complications prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic retinopathy or diabetic maculopathy prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic nephropathy prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)

No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of diabetic neuropathy prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of Diabetic lower limb severe complications prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of peripheral vascular disease prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of kidney or genitourinary stones prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of UTI or pyelonephritis prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of liver disease prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of pancreatitis prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)

History of congestive heart failure prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of hypertension prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of stroke prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of myocardial infarction prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of autoimmune disease prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
History of COPD prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Any ICU admission prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Antihypertensives/ diuretics use prior to index date				

Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
NSAIDs use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Oral steroids use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Statins/fibrates use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Lipid modifying agents use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Zoledronic acid use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Antibiotic use prior to index date				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
BMI (closest value prior to index date)				
<20	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
20-24.9	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
25-29.9	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
>30	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Smoking status (closest value prior to index date)				

Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Alcohol use (closest value prior to index date)				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
HbA1c in mmol/mol(closest value prior to index date)				
< 48 (6.5%) mmol/mol	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ 48 (6.5 %) - <64 mol/mol (8 %)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥ 64 mmol/mol (8 %)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)	(min, max) mean (± sd) median (Q1,Q3)
eGFR in (mL/min/1.73 m2)				
<15 (ml/min/1.73 m2) (kidney failure)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
15-29 (ml/min/1.73 m2) (severely decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
30-44 (ml/min/1.73 m2) (Moderately to severely decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
45-59 (ml/min/1.73 m2) (Mildly to moderately decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
60-89 (ml/min/1.73 m2) (Mildly decreased)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)

≥90 (ml/min/1.73 m2) (Normal)	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Renal impairment				
Yes	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
No	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Systolic blood pressure				
< 120	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥120-<139	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥140-<159	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
≥160	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)
Diastolic blood pressure				
< 80	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
80-89	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
90-99	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
100 or higher	N (x.xx%)	N (x.xx%)	N (x.xx%)	N (x.xx%)
Range (min,max)	(min, max)	(min, max)	(min, max)	(min, max)
mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)	mean (± sd)
median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)	median (Q1,Q3)

Table 38 Incidence rates per 1000 patient years of [Outcome name] with 95% confidence interval in the UK-CPRD population stratified by CPRD databases (GOLD and Aurum) . The number of patients in each strata and the total number of person years are also presented.

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Absolute incidence												
Total	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Exposure category (As treated)												
Empagliflozin	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
DPP-4 inhibitors	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Age at index (years)												
18-49	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
50-64	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
65-69	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
70-74	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
75-79	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
80-84	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
85 and over	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Sex												
Male	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Female	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Highest level of education												
Compulsory	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Upper secondary	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Post secondary	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Disposable income, individual												
< Q1	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Q1 – Q2	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Q2 – Q3	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥ Q3	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Disposable income, family												
< Q1	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Q1 – Q2	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Q2 – Q3	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥ Q3	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Marital status												
Registered partnership (incl. married)	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Widowed	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Separated	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Single	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Calendar year of index date												
2014	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
2015	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
2016	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
2017	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
2018	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
2019	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
2020	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
2021	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Diabetes treatment complexity at index date												
Monotherapy	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Dual combination therapy	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Triple combination therapy	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Insulin use												

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Metformin use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Oral glucose lowering drugs use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Sulfonylureas use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Thiazolidinediones use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
GLP-1 use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Diabetic complications												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Diabetic retinopathy or diabetic maculopathy												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Diabetic nephropathy												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Diabetic neuropathy												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Diabetic lower limb severe complications												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Peripheral vascular disease												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Kidney or genitourinary stones												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
UTI or pyelonephritis												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Liver disease												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Pancreatitis												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Congestive heart failure												

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Hypertension												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Stroke												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Myocardial infarction												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Autoimmune disease												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
COPD												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Any ICU admission prior to index date												

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Antihypertensives/ diuretics use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
NSAIDs use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Oral steroids												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Statins/fibrates use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Lipid modifying agents use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Zoledronic acid use												

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Antibiotic use												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
BMI (closest value prior to index date)												
<20	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
20-24.9	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
25-29.9	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
>30	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Smoking status (closest value prior to index date)												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Alcohol use (closest value prior to index date)												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
HbA1c (closest value prior to index date)												
< 48 (6.5%) mmol/mol	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥ 48 (6.5 %) - <64 mol/mol (8 %)	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥ 64 mmol/mol (8 %)	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
eGFR in (mL/min/1.73 m2)												
<15	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
15-29	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
30-44	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
45-59	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
60-89	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥90	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Renal impairment												
Yes	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
No	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Systolic blood pressure												

Variable and strata	GOLD						Aurum					
	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI	N Patients	N events	Person years	Rate	Lower 95% CI	Upper 95% CI
< 120	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥120-<139	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥140-<159	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
≥160	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
Diastolic blood pressure												
< 80	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
80-89	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
90-99	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx
100 or higher	N	N	N	x.xx	x.xx	x.xx	N	N	N	x.xx	x.xx	x.xx

Table 39 Crude hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the UK matched cohort, stratified by databases. The table shows the hazard ratios (HRs) using the crude model with the corresponding 95% CI and p-value–As treated

	GOLD				Aurum			
	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure								
DPP-4 inhibitors	reference	reference	reference	reference	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Table 40 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the UK matched cohort, stratified by databases. The table shows the hazard ratios (HRs) using the base model with the corresponding 95% CI and p-value–As treated

	GOLD				Aurum			
	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure								
DPP-4 inhibitors	reference	reference	reference	reference	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Table 41 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the UK matched cohort, stratified by databases. The table shows the hazard ratios (HRs) using the adjusted model with the corresponding 95% CI and p-value–As treated

	GOLD				Aurum			
	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure								
DPP-4 inhibitors	reference	reference	reference	reference	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx	x.xxx

Table 42 Crude hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the UK matched cohort after adding CPRD database as a stratification variable. The table shows the hazard ratios (HRs) using the crude model with the corresponding 95% CI and p-value–As treated

	Country- UK			
	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure				
DPP-4 inhibitors	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx

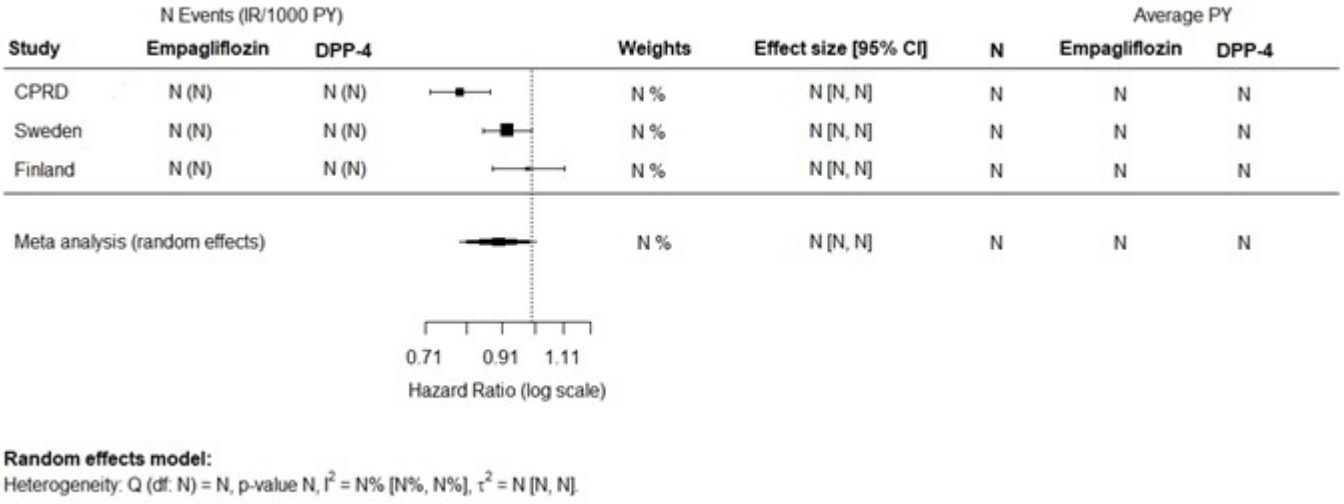
Table 43 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the UK matched cohort after adding CPRD database as a stratification variable. The table shows the hazard ratios (HRs) using the base model with the corresponding 95% CI and p-value–As treated exposure definition.

Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure				
DPP-4 inhibitors	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx

Table 44 Adjusted hazard ratios for [Outcome Name] incidence estimated using the Cox proportional hazards model in the [Study country] matched cohort after adding CPRD database as a stratification variable. The table shows the hazard ratios (HRs) using the adjusted model with the corresponding 95% CI and p-value–As treated exposure definition.

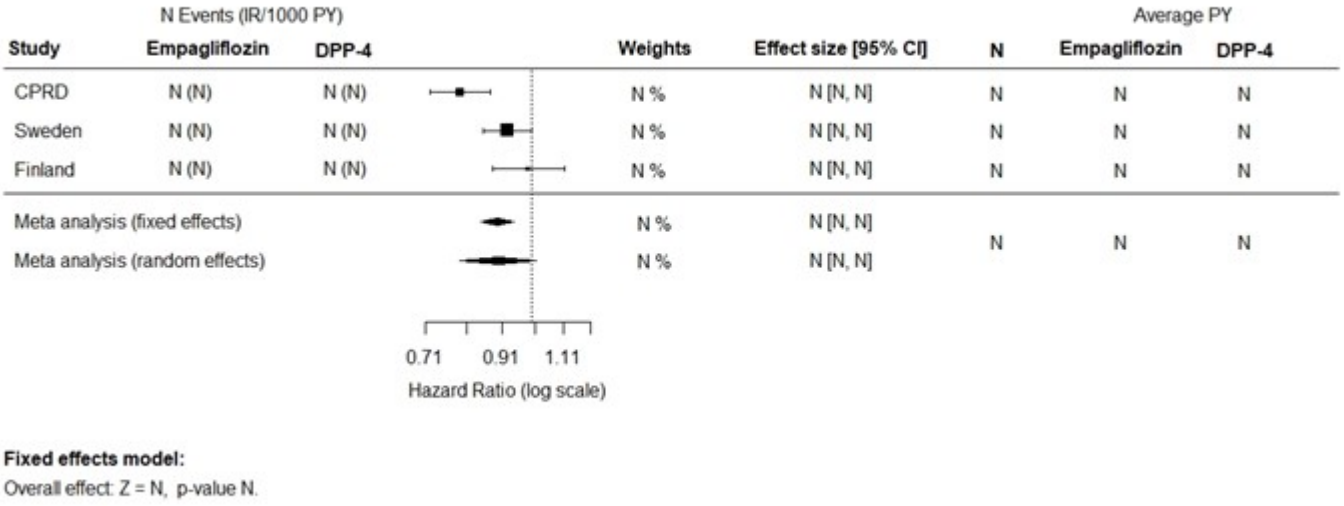
Variables	Hazard ratio	Lower 95% CI	Upper 95% CI	P-value
AT exposure				
DPP-4 inhibitors	reference	reference	reference	reference
Empagliflozin	x.xxx	x.xxx	x.xxx	x.xxx

Figure 8. Meta-analysis of [hazard/odds] ratios of [outcome] for the [analysis type] using a random-effects model



Random effects model:
Heterogeneity: Q (df: N) = N, p-value N, I^2 = N% [N%, N%], τ^2 = N [N, N].

Figure 9. Meta-analysis of hazard ratios of [outcome] for the as-treated analysis of empaglifozin users vs DPP-4 users using a fixed-effect model (sensitivity analysis)



Fixed effects model:
Overall effect: Z = N, p-value N.

Figure 10. Average number of diabetes-related healthcare visits per patient recorded in the source population for study years 2015-[end of study year]

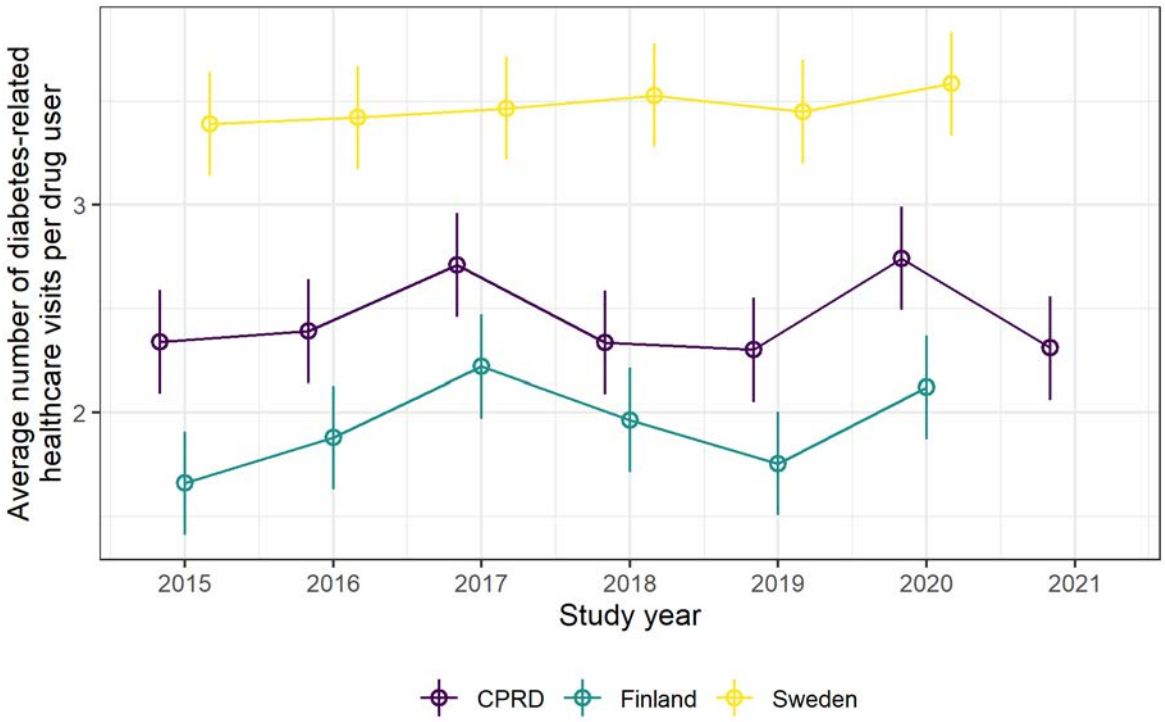


Figure 11. Number of monthly incident users in [Year] for empaglifozin and DPP-4 inhibitors for [Study country]

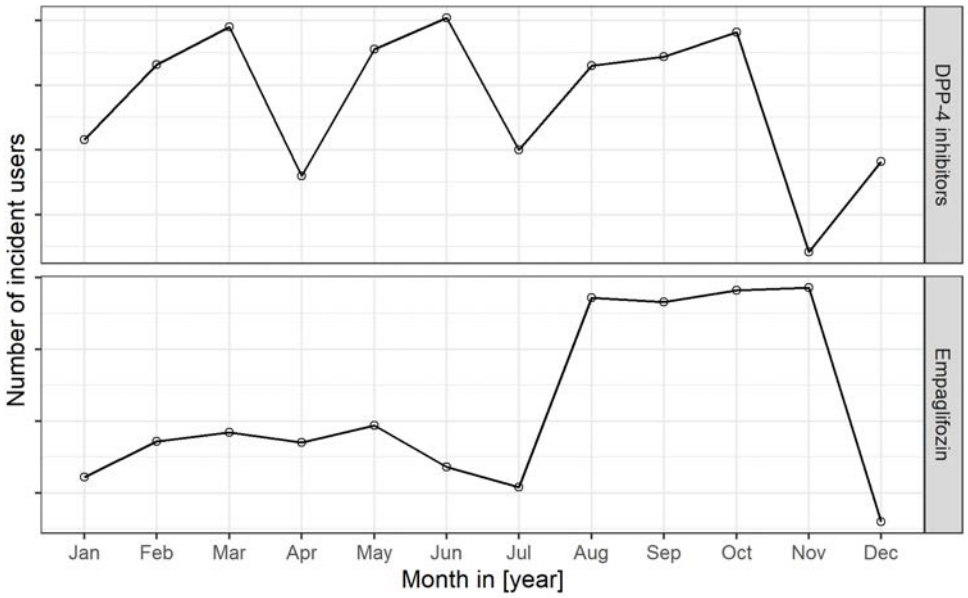


Figure 12. Number of quarterly incident users in [year] for empaglifozin and DPP-4 inhibitors for [Study country]

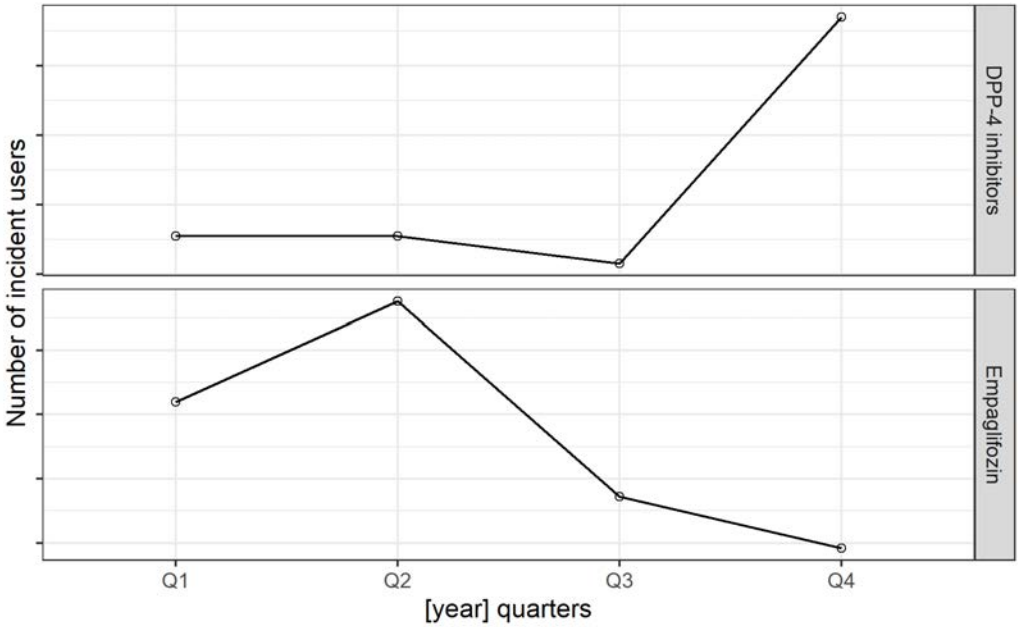


Figure 13. Number of incident users of each study drug group for each study year in [Study country].

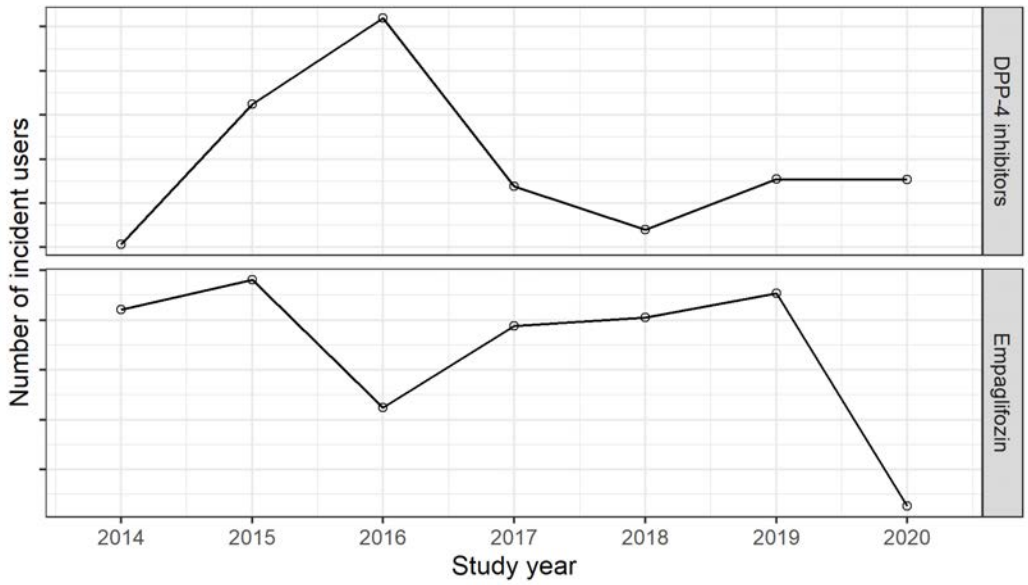


Figure 14. Average number (95% CIs) of prescriptions per patient in the source population in [Study country] per quarter from January 2015 to December [End of study year]

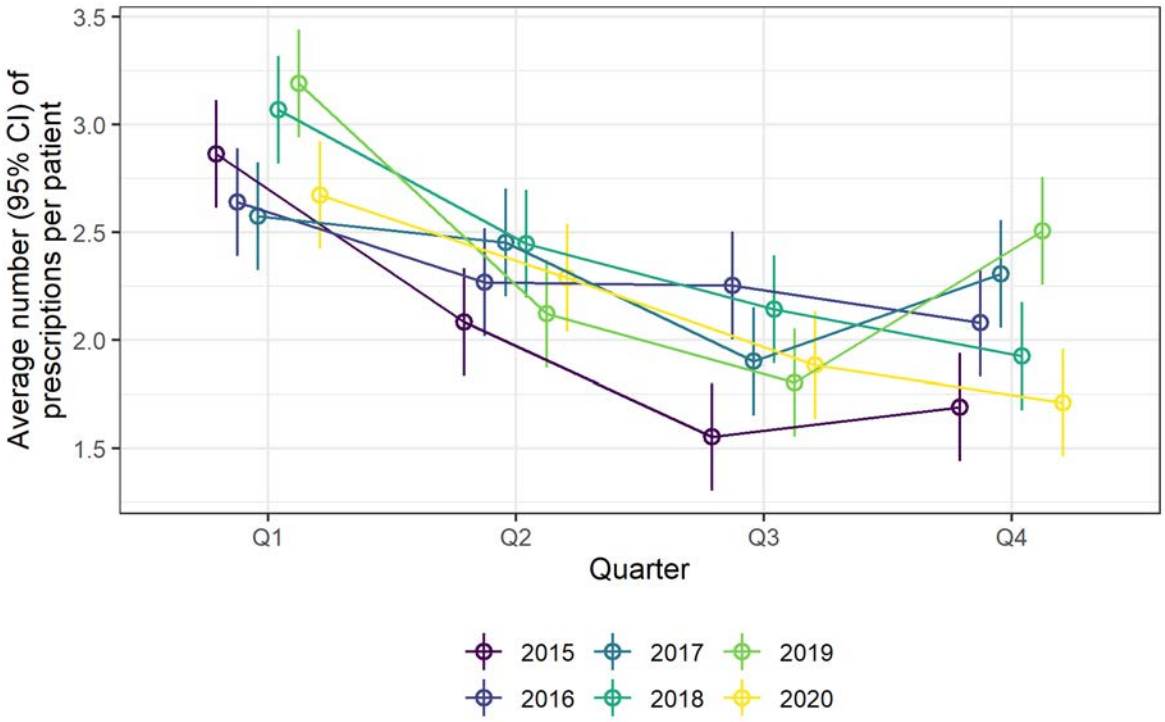


Table 45. Incidence rate of urinary tract cancer in the source population for all study years from 2015 to end of study year.

Country and study year	N of patients	N of events	Person years	Incidence rate	Lower 95% CI	Upper 95% CI
[Study country]						
2015	N	N	N	x.xx	x.xx	x.xx
2016	N	N	N	x.xx	x.xx	x.xx
2017	N	N	N	x.xx	x.xx	x.xx
2018	N	N	N	x.xx	x.xx	x.xx
2019	N	N	N	x.xx	x.xx	x.xx
2020	N	N	N	x.xx	x.xx	x.xx
2021	N	N	N	x.xx	x.xx	x.xx

Annex IV

List of drug dosages for empagliflozin, DPP-4 inhibitors and metformin

List of drug dosages extracted from registries in Finland and Sweden:

Annex 4 - List of drug dosages.xlsx