Statistical and Epidemiological Analysis Plan

BI Trial No.:	1245.96
Title:	Post-authorisation safety study in patients with type 2 diabetes mellitus to assess the risk of acute liver injury, acute kidney injury and chronic kidney disease, severe complications of urinary tract infection, and the risk of genital infections, and diabetic ketoacidosis among patients treated with empagliflozin compared to patients treated with DPP-4 inhibitors
Responsible trial statistician:	[redacted] and [redacted]
Responsible project epidemiologist:	[redacted] and [redacted]
Date of statistical and epidemiological analysis plan:	16 Sep 2019
Version:	3.0
Annotated SEAP, date and version:	15Jul2022, version Final

1. TABLE OF CONTENTS

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ADIN	Action, Decision, Issue, Notification (log at [redacted])
AKI	Acute Kidney Injury
ALI	Acute Liver Injury
ALI1	ALI primary endpoint, Acute Liver Injury Without Predisposing Conditions (analysis table or figure prefix)
ALI2	ALI secondary endpoint, Acute Liver Injury With or Without Predisposing Conditions (analysis table or figure prefix)
ALT	Alanine Aminotransferase
ALP	Alkaline Phosphatase
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical (classification system)
BI	Boehringer Ingelheim International GmbH
BMI	Body Mass Index
CI	Confidence Interval
CKD	Chronic Kidney Disease
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
COPD	Chronic Obstructive Pulmonary Disease
CPRD	Clinical Practice Research Datalink
DK	Denmark
DKA	Diabetic Ketoacidosis
dm+d	Dictionary of Medicines and Devices
DPP-4	Dipeptidyl Peptidase-4
ED	Emergency Department
eGFR	Estimated Glomerular Filtration Rate
EMA	European Medicines Agency
FHSA	Family Health Services Authority
GFR	Glomerular Filtration Rate
GI	Genital Infections
GIF	Genital Infections in Females
GIFH	Severe Genital Infection in Females
GIM	Genital Infections in Males
GIMH	Severe Genital Infection in Males
GLDs	Glucose-Lowering Drugs
GOLD	General Practitioner Online Database of CPRD
GP	General Practitioner
GPI	Generic Product Identifier
HAART	Highly Active Antiretroviral Therapy

Term	Definition / description
HbA1c	Glycated Haemoglobin
HCO3	Bicarbonate
HES	Hospital Episode Statistics
HIRD®	HealthCore Integrated Research Database
HIV	Human Immunodeficiency Virus
HTX	Hepatotoxic
ICD-10	*International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*
ICD-10-CM	*International Classification of Diseases, 10th Revision, Clinical Modification*
ICD-9	*International Classification of Diseases, 9th Revision*
ICD-9-CM	*International Classification of Diseases, 9th Revision, Clinical Modification*
IR	Incidence Rate
IRR	Incidence Rate Ratio
ITT	Intention to Treat
LOINC	Logical Observations Identifiers Names and Codes
MAR	Missing at Random
MCAR	Missing Completely at Random
MNAR	Missing Not at Random
NHS	National Health Service
NIS	Non-interventional Study
NSAID	Non-steroidal Anti-Inflammatory Drug
OPCS4	*Office of Population Censuses and Surveys, Fourth Revision*
PASS	Post-authorisation Safety Study
PCOS	Polycystic Ovarian Disease
PPV	Positive Predictive Value
PRAC	Pharmacovigilance Risk Assessment Committee
QC	Quality Control
[redacted]	
SEAP	Statistical Epidemiological Analysis Plan
SES	Socioeconomic Status
SGLT2	Sodium-Glucose Cotransporter 2
SOP	Standard Operating Procedure
STI	Sexually Transmitted Infection
T1D	Type 1 Diabetes Mellitus
T2D	Type 2 Diabetes Mellitus
UK	United Kingdom
ULN	Upper Limit of the Normal Range
US	United States
UTI	Urinary Tract Infection

3. RESPONSIBLE PARTIES

Non-interventional study (NIS) Statistician (statistical epidemiological analysis plan [SEAP] author): [redacted] and [redacted] from [redacted]

Annotated SEAP reviewers are as follows:

[redacted]:

- [redacted]

[redacted]

- [redacted]

[redacted] in the US:

- [redacted]

Boehringer Ingelheim International GmbH (BI):

- [redacted], Team Metabolism
- [redacted] Team Metabolism

4. INTRODUCTION

This statistical epidemiologic analysis plan (SEAP) provides a detailed description of the plan for conducting the analyses for a post-authorisation safety study (PASS) for empagliflozin. As part of the risk management plan, Boehringer Ingelheim International GmbH (BI) has committed to perform a PASS to estimate, among patients with type 2 diabetes mellitus (T2D), the risk of acute liver injury (ALI), the risk of acute kidney injury (AKI) and chronic kidney disease (CKD), the risk of severe complications of urinary tract infection (UTI), the risk of genital infections (GI), and the risk of diabetic ketoacidosis (DKA) among patients treated with empagliflozin compared with patients treated with dipeptidyl peptidase-4 (DPP-4) inhibitors.

The primary objective of the study is as follows:

- To estimate the **adjusted incidence rate ratios (IRRs)** of the **primary outcomes** (ALI1 in patients with no predisposing conditions, AKI, severe complications of UTI, GI, and DKA) by comparing patients initiating empagliflozin with patients initiating a DPP-4 inhibitor, among patients with T2D and after applying outcome-specific exclusion criteria.

The secondary objectives of the study are as follows:

- To estimate the **adjusted IRRs** of the **secondary outcomes** (ALI2 in patients with or without predisposing conditions, CKD, and severe GI) by comparing patients initiating empagliflozin with patients initiating a DPP-4 inhibitor, among patients with T2D and after applying outcome-specific exclusion criteria.
- To estimate the **adjusted IRRs for each of the primary outcomes** (ALI in patients with no predisposing conditions, AKI, severe complications of UTI, GI, and DKA) **and secondary outcomes** (ALI2 in patients with or without predisposing conditions, CKD, and severe GI)—**stratified** by categories of insulin use at the index date, age, sex, and other variables of interest such as diabetes control—by comparing patients initiating empagliflozin with patients initiating a DPP-4 inhibitor, among patients with T2D and after applying outcome-specific exclusion criteria.
- To estimate the **adjusted incidence rates** (IRs) for each of the **primary outcomes** (ALI1 in patients with no predisposing conditions, AKI, severe complications of UTI, GI, and DKA) **and secondary outcomes** (ALI in patients with or without predisposing conditions, CKD, and severe GI)—**overall and stratified** by categories of insulin use at the index date, age, sex, and other variables of interest such as diabetes control—among patients with T2D initiating empagliflozin or a DPP-4 inhibitor and after applying outcome-specific exclusion criteria.

5. AMENDMENTS AND UPDATES

A list of changes in the planned analysis in SEAP version 3.0 versus SEAP version 1.0 and 2.0 is presented below, by SEAP section.

Date	Section	Change	Reason for change
26 Sep 2019 (vs. SEAP version 2.0)	Across SEAP	Added clarifications from protocol version 7.0 versus version 6.0	Version 7.0 of the protocol was approved after 01 Aug 2019, minor changes and clarifications were translated into the updated SEAP version 3.0
01 Aug 2019 (vs. SEAP version 1.0)	Across SEAP	Eliminated SGLT2 comparator	Streamlining endorsed by the Pharmacovigilance Risk Assessment Committee (of the EMA)
01 Aug 2019 (vs. SEAP version 1.0)	Across SEAP	Added the Danish population registers and the HealthCore US database	Per EMA feedback on the second interim report
			To increase study number of users of empagliflozin in order to achieve desired study size for the rarest outcomes
01 Aug 2019 (vs. SEAP version 1.0)	5	Updated study objectives	To delete secondary outcome for UTI
01 Aug 2019 (vs. SEAP version 1.0)	7.2	Extended the study period	Per EMA feedback on the second interim report
			To increase study number of users of empagliflozin in order to achieve desired study size for the rarest outcomes
01 Aug 2019 (vs. SEAP version 1.0)	7.3	Updated study population to modify the definition of new users to not have a prescription or dispensing of the study medications during the 12 months prior to the index date (previously defined as no prior use ever)	To allow inclusion of more users of empagliflozin in the study, the main limitation of which is the accrual of patients exposed to the drug of interest. This definition has been used in other EMA-approved PASS, and it is considered appropriate because the comparators are indicated for the same target population and same stage of the disease, and a 12-month wash-out period is considered enough for the acute outcomes being evaluated.
01 Aug 2019 (vs. SEAP version 1.0)	8.2	Updated outcomes definition	To include ED visits in addition to those outcomes that require hospitalisation
01 Aug 2019 (vs. SEAP version 1.0)	8.2	Modified the ALI primary outcome definitions to be hospitalisation or referral to a specialist for ALI in patients with no predisposing conditions	To have a primary outcome that can be assessed in the three data sources included in the study. Outpatient ALI cannot be assessed in Denmark due to lack of primary care data but will be assessed in a sensitivity analysis in CPRD and HIRD

Date	Section	Change	Reason for change
01 Aug 2019 (vs. SEAP version 1.0)	8.2	Modified ALI secondary outcome definition to be hospitalisation or referral to a specialist for ALI in patients with and without predisposing conditions	To have a secondary outcome that can be assessed in the three data sources included in the study
01 Aug 2019 (vs. SEAP version 1.0)	8.2	Modified the AKI primary outcome definitions to be hospitalisation or referral to a specialist for AKI	To have a primary outcome that can be assessed in the three data sources included in the study. Outpatient AKI cannot be assessed in Denmark due to lack of primary care data but will be assessed in a sensitivity analysis in CPRD and HIRD
01 Aug 2019 (vs. SEAP version 1.0)	8.2	Deleted UTI secondary outcome and modified primary UTI outcome to include outpatient diagnosis	Based on PRAC feedback on versions 1 and 2 of the protocol
01 Aug 2019 (vs. SEAP version 1.0)	8.2	Added description of cases of elevated liver enzymes identified through laboratory results	To gain more knowledge about the potential hepatotoxicity of empagliflozin
01 Aug 2019 (vs. SEAP version 1.0)	11.5.3	Provided additional details on potential methods to deal with propensity score deciles with zero events	To suggest alternative methods of estimating the adjusted treatment effect from propensity scores (in the situation of zero margins in propensity score strata)
01 Aug 2019 (vs. SEAP version 1.0)	11.6.3	Added analysis using positive predictive value	To correct for outcome misclassification
01 Aug 2019 (vs. SEAP version 1.0)	11.6.5	Added description of the study outcomes by type of DPP-4 inhibitor	To provide an indication regarding whether the distribution of study outcomes in each individual type of DPP-4 inhibitor is different from the expected
01 Aug 2019 (vs. SEAP version 1.0)	11.6.8	Provided more detail on meta-analysis techniques	To describe in more detail the meta-analytic techniques and clarify that a random-effects model will be used as the main analysis and the fixed-effects model will be used in a sensitivity analysis
05 October 2016 (vs. Protocol version 4.0, updated in later Protocol amendments)	8.3	The list of covariates to be included in each outcome-specific propensity score model has been updated.	To harmonise when needed the list of covariables to be evaluated for each outcome

Date	Section	Change	Reason for change
05 October 2016 (vs. Protocol version 4.0, updated in later Protocol amendments)	11.5.3	Instead of doing a propensity score model for each calendar year, a full interaction propensity score model will be generated where calendar year will be included as a covariate along with its interaction with each of the other covariates.	This full interaction propensity score model will allow the influence of each covariate in predicting treatment to vary across calendar years (thus accounting for potential channelling bias) and has an advantage of efficiency in providing one overall comprehensive model compared with generating separate propensity score models by calendar year.

AKI = acute kidney injury; ALI = acute liver injury; CPRD = Clinical Practice Research Datalink; DPP-4 = dipeptidyl peptidase 4; ED = emergency department; EMA = European Medicines Agency; HIRD = HealthCore Integrated Research Database; PASS = post-authorisation safety study; PRAC = Pharmacovigilance Risk Assessment Committee; SEAP = statistical epidemiological analysis plan; SGLT2 = sodium-glucose cotransporter 2; US = United States; UTI = urinary tract infection.

6. IMPORTANT PROTOCOL VIOLATIONS

None.

7. RESEARCH METHODS

7.1 STUDY DESIGN

This is an observational cohort study using existing data from routine medical care in the Clinical Practice Research Datalink (CPRD) database in the United Kingdom (UK), the Danish Population Registries in Denmark, and the HealthCore Integrated Research Database (HIRD®) in the US.

The study uses a "new users" design and compares new users of empagliflozin with new users of DPP-4 inhibitors. The comparator group has the same indication as that of the empagliflozin group, and comparator medications were potentially prescribed to patients with T2D with similar characteristics; that is, similar T2D severity and underlying diabetes-related comorbidities.

7.2 SETTING

The data sources included in the study will be CPRD in the UK (including both GOLD [General Practitioner Online Database] and Aurum) and, for the evaluation of the rarest outcomes, also the Danish Population Registries in Denmark and HIRD in the US. For further details on these data sources, please see Section 9 ~~7.2~~ and the study protocol.

The study period will start on 01 August 2014, the date of empagliflozin launch in the UK, US, and Denmark. The study end date is planned for 01 August 2019 (01 August 2019 in CPRD and DK) or up to the latest available date (31 July 2019 in HIRD) at the time of data extraction for the final report.

7.3 STUDY POPULATION

The study population will include adults with T2D who are new users of empagliflozin or of a DPP-4 inhibitor. New users will be defined as those having a prescription/dispensing of a study medication during the study period and who have not received a prescription or dispensing of any of the study medications during the 12 months before the index date. The date of the first prescription/dispensing (prescriptions in CPRD and dispensing in the Danish Population Registries and HIRD) of a study medication of interest within the study period which meet all the inclusion criteria will be the index date. Patients will be allowed to enter the other cohort (i.e., the empagliflozin cohort if they first entered the DPP-4 inhibitor cohort or the DPP-4 inhibitor cohort if they first entered the empagliflozin cohort) if they fulfil the inclusion/exclusion criteria.

7.3.1 Inclusion criteria

All patients will be required to meet *all* of the following criteria:

1. Was aged 18 or more years at the index date.
2. Had at least 12 months of available data before the index date.

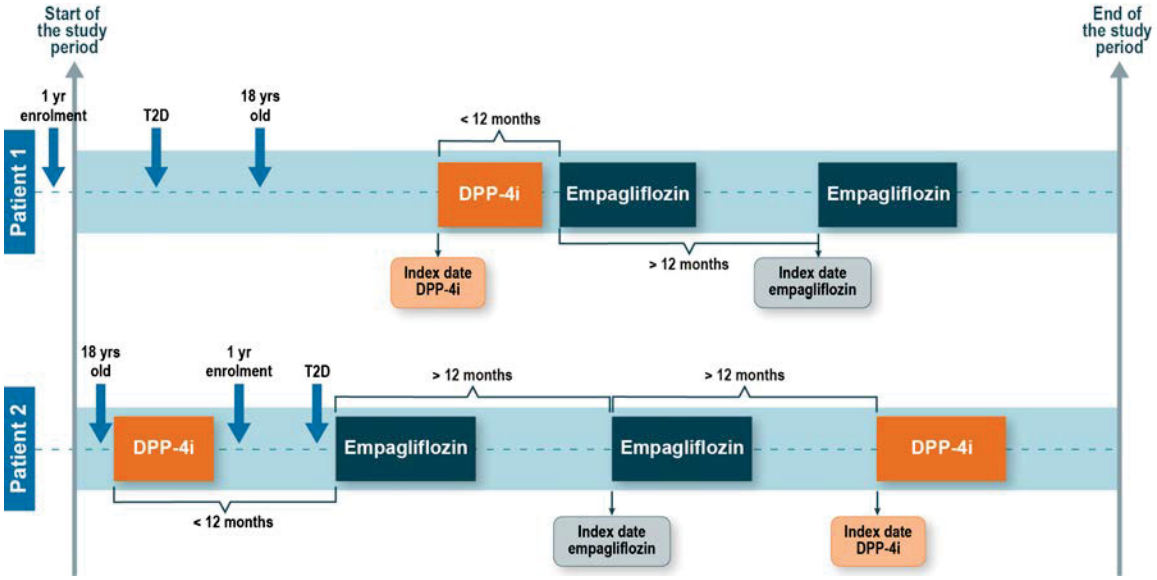
- **In CPRD**: at least 12 months of continuous registration before the index date in a practice contributing up-to-standard data to CPRD (up-to-standard date variable was not available for Aurum at the time of data extraction).
- **In Denmark**: at least 12 months of residency in Denmark (i.e., patient must have not immigrated during the previous 1 year).
- **In HIRD**: 12 months of continuous enrolment prior to the index date with complete pharmacy and medical coverage in a health insurance plan.

3. Had T2D ever before or on the index date, as described in Section 7.3.3.
4. The empagliflozin-exposed population also met the following criteria:
    - Had at least one prescription for empagliflozin or fixed-dose combination of empagliflozin with metformin, with or without treatment with another glucose-lowering drug (GLD).
    - Had no prior prescriptions for any sodium-glucose cotransporter 2 (SGLT2) inhibitor (including empagliflozin) or a DPP-4 inhibitor alone or in fixed-dose combination during the previous 12 months.
5. The population exposed to a DPP-4 inhibitor must meet the following criteria:
    - Had at least one prescription for a DPP-4 inhibitor or a fixed-dose combination of a DPP-4 inhibitor with metformin, with or without treatment with other GLDs.
    - Had no prior prescriptions for a DPP-4 inhibitor or any SGLT2 inhibitor (including empagliflozin) alone or in fixed-dose combination during the previous 12 months.

The first prescription/dispensing of empagliflozin or DPP-4 inhibitor that meets all the above criteria will be the index prescription, and the date of the prescription/dispensing will be the index date. Note that the first prescription within the study period may not have met all the inclusion criteria, but a subsequent prescription/dispensing may have met them (see Figure 1).

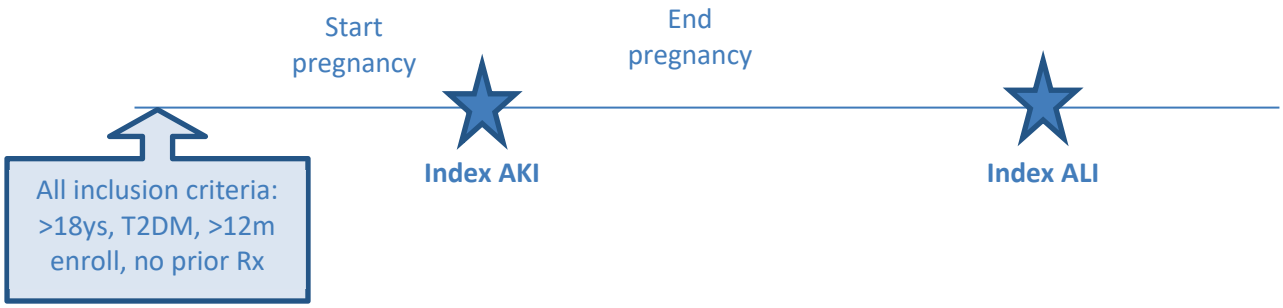
For all analyses (i.e., main, secondary, and sensitivity analyses), a patient can enter both cohorts, but only once if they fulfil inclusion and exclusion criteria, including no prior use of any of the study medications during the previous 12 months (i.e., patients can switch from their initial study drug cohort, for example from empagliflozin to the DPP4i cohort, but not back to the empagliflozin cohort).

Figure 1. Eligibility criteria and index date definition



DPP-4i = dipeptidyl peptidase-4 inhibitor; T2D = type 2 diabetes mellitus.

As a further clarification, note that index date will be determined independently for each outcome cohort. See example figure below:



7.3.2 Exclusion criteria

1. Patients with a diagnosis of type 1 diabetes ever before (see lookback periods in each database in Section 7.5) or on the index date will not be included in the study (see operational definition in Section 7.3.3).
2. Patients prescribed/dispensed combinations of SGLT2 inhibitors (including empagliflozin) with DPP-4 inhibitors on the index date for either empagliflozin or DPP-4 inhibitors will be excluded (fixed-dose combinations or non–fixed-dose combinations of the two individual medications).

    - Even though this criterion is under the exclusion criteria in the protocol, rather than exclusion, prescriptions/dispensings of SGLT2 inhibitors and DPP-4 inhibitors occurring on the same date will not qualify as potential index dates; this does not preclude future prescriptions/dispensings of either medication from being selected if they fulfil the inclusion/exclusion criteria.

- Although unlikely, if there is prescription/dispensing for both empagliflozin and another SGLT2 inhibitor on the same index date, this will be operationalised as if the patient had prior use of SGLT2 inhibitors, i.e. that prescription would not fulfil inclusion/exclusion criteria.

7.3.3 Operational definitions (inclusion and exclusion criteria) for type 2 diabetes mellitus and type 1 diabetes mellitus

7.3.3.1 Type 1 and 2 diabetes mellitus in CPRD

Operationally, patients with T2D will be selected if they met *one* of the following criteria (Table 1):

- T2D diagnosis AND no type 1 diabetes mellitus (T1D) diagnosis—irrespective of the number of non-insulin GLD types prescribed, and "diabetes (type not specified)" diagnosis
- "Diabetes (type not specified)" diagnosis AND no T1D diagnosis (irrespective of the number of non-insulin GLD types prescribed) AND no T2D diagnosis
- T2D diagnosis AND only one T1D diagnosis AND more than one non-insulin GLD (or one fixed-dose combination of non-insulin GLDs)
- "Diabetes (type not specified)" diagnosis AND only one T1D diagnosis AND more than one non-insulin GLD (or one fixed-dose combination of non-insulin GLDs)

Operationally, patients with T1D diagnosis will be excluded if they met one of the following criteria (Table 1):

- One T1D diagnosis AND one non-insulin GLD, irrespective of insulin prescriptions/ dispensings and diagnosis of T2D or "diabetes (type not specified)"
- Had more than one T1D diagnosis (irrespective of any other variable)

Table 1. Operational definitions for T2D and T1D inclusion and exclusion criteria

Criterion	Include/exclude	T2D	T1D	Non-insulin GLD	"Diabetes (type not specified)"
1	Include	≥ 1	0	Irrespective	Irrespective
2	Include	0	0	Irrespective	≥ 1
3	Include	≥ 1	1	> 1*	Irrespective
4	Include	0	1	> 1*	≥ 1
5	Exclude	Irrespective	1	1	Irrespective
6	Exclude	Irrespective	> 1	Irrespective	Irrespective
7	Exclude	0	0	≥ 1	0
8	Exclude	0	1	> 1	0

*Or one fixed-dose combination of non-insulin GLDs.
GLD = glucose-lowering drug; T1D = type 1 diabetes; T2D = type 2 diabetes.

In Table 1, criteria 1 and 2 include no conflicting diagnoses. Criteria 3 and 4 use conflicting diagnoses but are based on the assumption that the use of at least two different non-insulin GLDs confirms the diagnosis of T2D. Criterion 5 is based on the assumption that one

diagnosis of T1D in the presence of only one type of non-insulin GLD, irrespective of the patient having a diagnosis code for T2D or for diabetes (type unspecified), cannot differentiate between T1D and T2D. Criterion 6 is based on the assumption that two T1D diagnoses are in conflict with any other criteria or variable that could potentially define T2D. Patients with one T1D code will be included if they had at least 1 T2D code or a diabetes (type unspecified) code and used more than one type of GLD (non-insulin type). Thus, criteria will be independent of insulin prescription/dispensing; for example, patients with T1D codes (even without insulin prescriptions/dispensings) will be excluded (and not considered T2D) if they had more than one T1D code, or if they did not have any concomitant T2D code or diabetes (type unspecified) code, even if they had prescriptions/ dispensings for more than one type of non-insulin GLD. Operational definitions were adapted from Holden et al. [R13-3433]; in the current study, patients with T1D and no insulin prescription/dispensing will be excluded if they fulfil criterion 5 or 6.

In CPRD, diagnoses of T2D, T1D, diabetes (type not specified), and non-insulin GLDs will be considered if they occurred ever before or at the index date using Read codes and Gemscript in GOLD and SNOMED/EMIS codes and Dictionary of Medicines and Devices (dm+d) codes in Aurum.

The following drug classes are included under non-insulin GLDs:

- Alpha glucosidase inhibitors
- Biguanides
- Combinations (will count as 2 GLDs):
    - DPP4/metformin
    - SGLT2/metformin
- Thiazolidinediones/metformin
- Dipeptidyl Peptidase-4 inhibitors (named as DPP4 in the code list)
- GLP-1 receptor agonists
- Meglitinides
- Sodium-Glucose Cotransporter 2 inhibitors (SGLT2 in the code list)
- Sulphonylureas
- Thiazolidinediones

All codes for diagnoses and medications are listed in Annex 1.

7.3.3.2 Type 1 and 2 diabetes mellitus in Denmark

In Denmark, patients with T2D will be identified based on a combination of community pharmacy prescription/dispensing data and/or hospital codes [R18-0269, R18-1213]. The classification of patients as having T1D or T2D based on register date only is not accurate, and the approach chosen is to identify persons with T1D with reasonably high specificity and classify the rest as having T2D. Thus, T2D will be equivalent to “cannot be classified as T1D with reasonable certainty,” and hence any register-based classification in Denmark should be used cautiously. For example, patients with severe new T2D may have been treated with insulin in the beginning (but likely will eventually start non-insulin GLDs) and may have

been incorrectly coded as T1D in the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10) system.

Patients with incident T2D will be defined as those with either of the following:

- A first prescription for any GLD (Anatomical Therapeutic Chemical [ATC] code A10), OR
- A first inpatient or outpatient hospital contact for diabetes of any type (type 1, type 2, or non-specified)(see list of codes in Annex 1)

Among the population with incident T2D, patients who are likely to have incident T1D will be excluded. Patients with incident T1D will be defined as follows:

- Under age 30 years at the time of their first diabetes record (prescription or diagnosis), AND
- Positively and exclusively received prescriptions for insulin as GLD within first 3 months of their first diabetes record [R18-0269].

Note: Aarhus decided not to use PCOS, and it is not planned to be used in CPRD or HIRD so the text related to PCOS has been deleted.

7.3.3.3 Type 1 and 2 diabetes in HIRD

In HIRD, patients with T2D will be identified applying the same criteria as the CPRD definition (Table 1). Using all available data before or on the index date, T1D, T2D, and other specified diabetes (i.e., diabetes (type not specified)) will be based on *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) and *International Classification of Diseases, 10th Revision, Clinical Modification* (ICD-10-CM) diagnosis codes, and dispensings for GLDs will be based on Generic Product Identifier [GPI] codes, as follows:

1. T2D: medical visit with ICD-9-CM/ICD-10-CM diagnosis codes (ICD-9-CM 250.X0 or 250.X2, or ICD-10-CM E11 (Annex 1) medical visit with an ICD-9-CM/ICD-10-CM diagnosis codes ICD-9-CM 250.X1 or 250.X3, or ICD-10-CM E10 (Annex 1).
2. Diabetes (type not specified): medical visit with ICD-10-CM E13 (Annex 1). Because ICD-9-CM codes for T2D include unspecified diabetes type, these codes will not be used in the definition.
3. Non-insulin GLD drug classes: dispensings with GPI codes 2715x-2720x, 2725x-2728x, 2750x-2755x, or 2760x-2799x.

7.3.4 Outcome-specific exclusion criteria

The three- or four-letter prefixes for table and figure specification of primary outcomes will be as follows:

- ALI1: hospitalisation, ED visit, or specialist visit for acute liver injury among patients without predisposing conditions – in all data sources,
- AKI: hospitalisation, ED visit, or specialist visit for acute kidney injury – in all data sources

- UTI: severe complications of UTI (inpatient and outpatient)—only in CPRD,
- GIM: GI in males (inpatient and outpatient)—only in CPRD,
- GIF: GI in females (inpatient and outpatient)—only in CPRD, and
- DKA: hospitalisation or ED visit due to diabetic ketoacidosis – in all data sources.

The prefixes for table and figure specification of secondary outcomes will be as follows:

- ALI2: hospitalisation, ED visit, or specialist visit for acute liver injury with or without predisposing conditions – in all data sources,
- CKD: chronic kidney disease (inpatient and outpatient)—only in CPRD,
- GIMH: severe genital infection in males that results in hospitalisation or ED visit or that required systemic treatment with specific antifungals or antibiotics—only in CPRD, and
- GIFH: severe genital infection in females that results in hospitalisation or ED visit or that required systemic treatment with specific antifungals or antibiotics—only in CPRD.

Different exclusion criteria will be applied to generate sets of outcome-specific populations for the analysis of the outcomes of interest.

For the analysis of GI, severe GI, and DKA, no additional exclusion criteria will be applied. Prior diagnosis of GIM/GIF, GIMH/GIFH, DKA, or fulfilling the laboratory or validation criteria for GIM/GIF or DKA prior to or at the index date will not constitute exclusion criteria.

7.3.4.1 ALI outcome-specific exclusion criteria—in all data sources

For analysis of the ALI1 primary outcome, patients will be excluded if they meet *any* of the following criteria:

1. A diagnosis of any of the following acute conditions any time before or at the index date (see Annex 2 for codes):
   - ALI
   - Acute infectious hepatitis
   - Acute cholelithiasis and cholecystitis
   - Acute intra- or extrahepatic biliary obstruction
   - Acute pancreatic disease
   - Decompensated congestive heart failure (i.e., hospitalisation)
2. Pregnancy at the index date will be identified through diagnosis codes compatible with initiation and/or termination of pregnancy, and duration of pregnancy will be estimated through specific time windows set up around the date of diagnosis.

   - **In CPRD**, see Annex 5.
   - **In HIRD**, ICD-9-CM and ICD-10-CM codes for outcome of delivery/normal delivery will be used to identify the date of delivery and the prior 40-week (280 days) pregnancy period. For early pregnancy termination codes, a 20-week (140 days) pregnancy period will be defined. We will also use ICD-9-CM and ICD-10-CM codes

for occurrences of pregnancy medical encounters to identify pregnancy status at specific points in time.

- **In Denmark,** pregnancy will be defined using an algorithm based on ICD-10 codes. Normal duration of pregnancy: 37-42 gestational weeks (259-294 days). However, **the first 14 days cover a period from last ovulation until true pregnancy**, thus, the **true** pregnant period for normal pregnancy duration: 245-280 days. See further details in **Annex 5**.
  - Date of ICD-10 diagnosis will identify date of delivery/abortion.
  - For ICD-10 codes for delivery/outcome of delivery without any specification of gestational week at delivery, a 280-day pregnancy period will be applied.
  - For post term delivery specific ICD-10 codes, a 294-day pregnancy period will be applied.
  - For preterm specific delivery codes, a 245-day pregnancy period will be applied.
  - For early pregnancy termination codes, either a 140-day pregnancy period (for codes specified as after gestational week 12 or without specification) or a 70-day pregnancy period (for codes specified as prior to gestational week 12) will be applied. ICD-10 codes for conditions related to pregnancy but without a specification regarding gestational week will possibly also be used.
  - We will also use ICD-10 codes for conditions related to pregnancy but without a specification of specific pregnancy time point. For these codes, we will apply a pregnancy period defined as: End pregnancy = indate + 280 days and start pregnancy = outdate – 280. Thus, the potential pregnancy period will be up to 560 days if the hospital contact only last 1 day and shorter for contacts with longer duration.
  - In case of overlapping pregnancy periods, ICD-10 codes of delivery/abortion will be ranked above ICD-10 codes for conditions related to pregnancy but without a specification of specific pregnancy time point.

3. A diagnosis of the following chronic conditions recorded at any time before or at the index date (see variable definitions in Section 8.3):

   - Chronic liver disease (including transplant)
   - Chronic alcoholism
   - Chronic infectious hepatitis
   - Chronic disease involving the liver or causing hyperbilirubinaemia
   - Chronic cholelithiasis and cholecystitis
   - Intra- or extrahepatic biliary obstruction
   - Chronic pancreatic disease
   - Primary or secondary hepatic, biliary, or pancreatic cancer
   - Decompensated congestive heart failure (i.e., inpatient hospitalisation): this has been modified to clarify that the exclusion is limited to those with hospitalisation for heart failure

For the analysis of the ALI2 secondary outcome "~~ALI~~ALI in patients with or without predisposing conditions," only outcome-specific exclusion criteria 1 and 2 will be applied.

For outcome-specific exclusion criterion 1, the time window will be 6 months before or at the index date instead of any time before or at the index date (outcome-specific exclusion criterion 3 – does not apply to ALI2 secondary endpoint). Note that for most of the conditions in outcome-specific exclusion criteria 1 and 3, the status of acute or chronic cannot be determined through the codes used to record these conditions. For this reason, the lookback period of 6 months before the index date will be applied when identifying the excluding conditions that apply to the secondary outcome cohort, and a lookback period of *any time* before or on the index date will be applied when identifying the excluding conditions that apply to the primary outcome cohort. For example, ICD-10: K81 Cholecystitis will be considered an exclusion condition for the secondary outcome only if it occurred in the 6 months before or on the index date, and it will be an exclusion criteria for the primary outcome if it occurs at *any time* before or on the index date. See the list of codes in the Inclusion_Exclusion excel, where codes to be evaluated as potential exclusion criteria for ALI1 primary outcome are all those identified in a column with a 1 or a 2 and the time period is any time before index date, while codes to be used for the exclusion of patient to the secondary ALI2 cohort are those identified in a column with a 2 and the time period for ascertainment is 6 months before index date..

Use of laboratory data as a proxy for predisposing conditions:

In **Denmark**, because there are no primary care data and some of these conditions may not be seen in the hospital setting, there may be underrecording of these conditions. Given that nationwide laboratory data will be available (no longer nationwide at the time of the analysis, see next page) , electronically available laboratory values before the or on index date will be used to identify some of the conditions described as exclusion criteria. Note that after discussions only eGFR will be used for censoring-purposes. Details on identification of these conditions through laboratory values will be included in the annotated SEAP. In Denmark, laboratory values will also be used to exclude patients with acute conditions for ALI2, and acute and chronic conditions for ALI1. See Tables in page 31 and 32.

In **CPRD** and **HIRD**, although availability of electronic laboratory data is incomplete (no in-hospital laboratory data in CPRD and laboratory data for only a subset in HIRD), electronically available laboratory values before or on the index date will be used to identify some of the conditions described as exclusion criteria.

In **all data sources**, *chronic liver injury* will be defined using laboratory criteria recorded within 1 year before or at the index date for exclusion purposes, but not for censoring. C*hronic liver injury* will be defined as fulfilling laboratory criteria for ALI at least two times separated by more than 180 days. For laboratory test related to the time window after index date see validation plan.

- The laboratory values between the first and second time that the patient fulfils the criteria for chronic liver injury will not have to fulfil the criteria of *chronic liver injury*.

- To qualify for chronic liver injury in the two different times, the criterion met does not need to be the same both times, i.e. it could be criteria (A) first time and criteria (C) the second time.

- Laboratory criteria for ALI are specified in Section 8.2.1.1., briefly:

    - (**A**) ALT or AST ≥ 5 × ULN **or**
      (**B**) ALT or AST ≥ 3 × ULN *AND* total bilirubin ≥ 2 × ULN **or**
      (**C**) ALP ≥ 2 × ULN

    - ALT will be used as first option, whenever available. If both ALT and AST are available on the same date, only ALT will be used. If ALT is missing, then AST will used, except if the patient has a record for a muscle disease within 30 days before or after the AST laboratory test. The list and all codes of muscle diseases to be used in each database is available in the code list for inclusion and exclusion criteria (Annex 2).

    - In **CPRD GOLD** and **Aurum**,
        - ULN can be missing, and these vary by laboratory. Based on patient profile review we have seen that patients usually have many ALT AST measurements all with the same ULN, form the same lab, and occasionally one ULN is missing. In this scenario, we will impute the ULN as the mode of the ULN within the same patient lab results.
        - After inspection of data values, very few laboratory results were extreme outliers. Because exclusions are based on two measures and outcomes are identified based on diagnosis, no data cleaning will be performed in CPRD. In addition, if outliers within <p0,01 or >p.99.9% were set to missing, there would not be patients with ALT or AST measures x 20 ULN and those are described in Table 13.

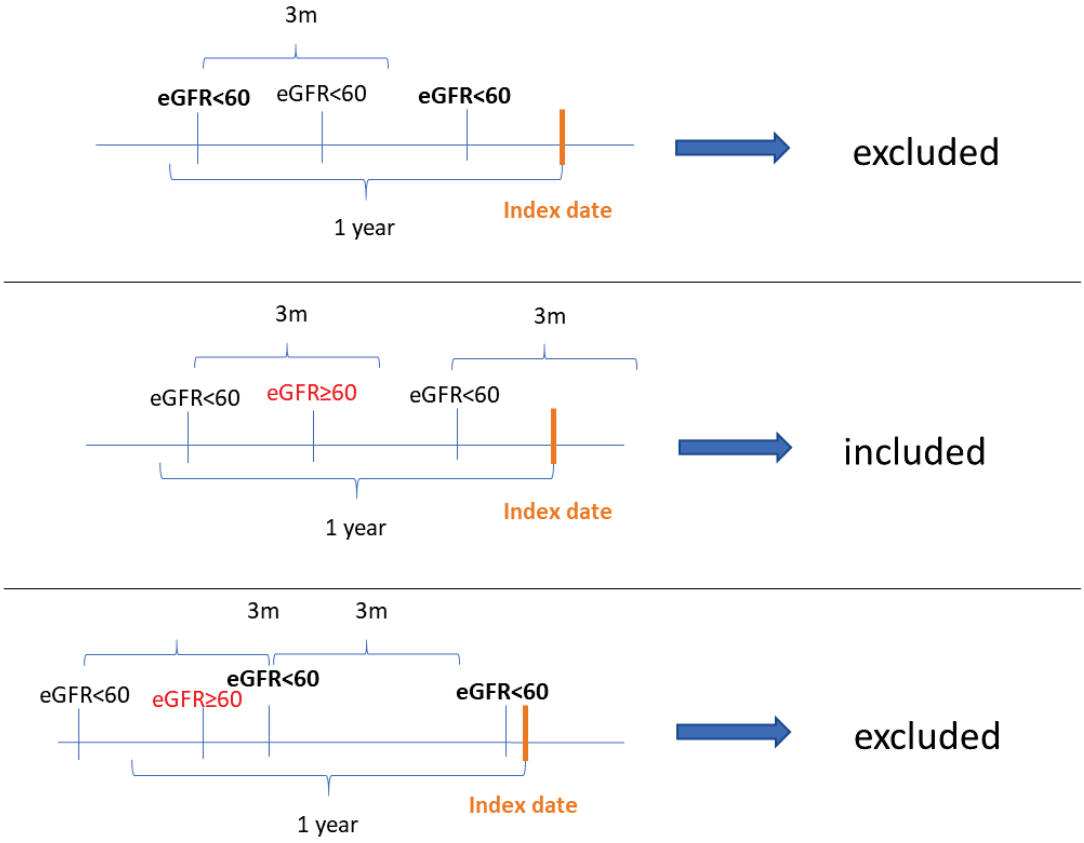
In Denmark, complete laboratory data is available from 2015 onwards for the whole of country. Prior to 2015 it is possible that in some regions of Denmark the data is partially complete. For this reason, a sensitivity analysis will be performed to investigate the available data by region and look at differences in the amount of lookback baseline time based on exposure. Note from DK: In Denmark, we have laboratory data from 2 different registers; regional Labka at [redacted] (covering 2 out of 5 regions = North and Central Denmark back to approximately 2000), and the new nationwide Lab_F (covering 4 out of 5 regions (not Central Denmark) from approx. 2015 onwards). The regional Labka database covering North and Central Denmark has presently been closed for further data achievement due to legislative discussions. This means we have the data that we have now but will not receive more data from this regional Labka database, unless legislative discussions turn out favourably. This means we cannot achieve the remaining 2019 lab-data for the fifth Danish region = Central Denmark. Also, the criteria about platelets and ALP has been added to our research project after we first applied for lab data in spring 2019, thus, we can only assess platelets and ALP for 4 out of the 5 Danish Regions (not Central Denmark) in the entire study period. Of note, the approval of use of nationwide laboratory data from Lab_F for research purposes is new and the completeness of these lab-data is still not known. Thus, from 2015 4 out of 5 Danish regions deliver lab data to the nationwide LAB_F database, but we do not know if there may be incomplete intra-regional lab data in some regions also after 2015. We are about to find out in the validation process. Based on further inspection of available lab data it was agreed **that age-platelet would not be used** for exclusion, censoring or

validation, although it may be of interest to describe it among cases around event date (final decision will be specified in the validation plan).

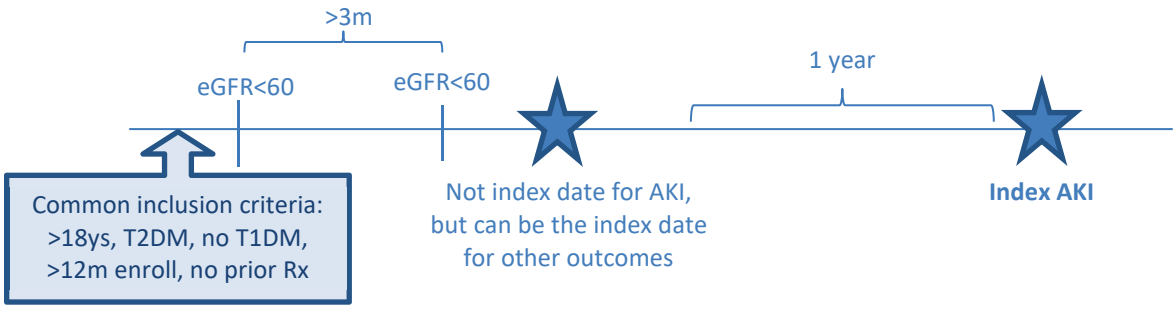
7.3.4.2 AKI outcome-specific exclusion criteria – in all data sources

For analysis of the AKI outcome, patients will be excluded if they meet *any* of the following criteria:

- A diagnosis of AKI is recorded within 6 months before or at the index date (see variable definitions in Section 8.2.1.2)
- A diagnosis of CKD is recorded any time before or at the index date (i.e., during the available lookback time) (see variable definitions in Section 8.3)
- In **all data sources**, if laboratory data are available, the patient fulfils laboratory criteria for CKD before or at the index date. (Electronic lab data for eGFR lab tests are not available in HIRD.) Laboratory criteria within 1 year before or at the index date in CPRD and within 1 year before or at the index date in Denmark for CKD are as follows:
  - Estimated glomerular filtration rate (GFR) < 60 mL/min/1.73 m^2, AND
  - Another Estimated GFR < 60 mL/min/1.73 m^2 performed at least 3 months and a maximum of 365 days apart.
    If there are other records for eGFR test between the above tests, these will also have to be < 60. If after an eGFR<60 there is one eGFR≥60, then the process will restart, i.e., for the two measures of eGFR<60 separated more than 90 days all measures should also be <60 or there should be have no measures between those. See Figure below for further clarifications. If a patient only has one eGFR or no eGFR measures, that patients will be considered as not having CKD, i.e. that patient will be included.
- No data cleaning will be performed in CPRD. After inspection of data values, very few laboratory results were extreme outliers. Because exclusions are based on two measures of eGFR or SrCreat and also CKD outcome sensitivity is based on two measures, there is less risk that the diagnosis is identified or confirmed through one outlier laboratory result



As described in Section 7.3.1., the index date will be determined independently for each outcome cohort. For example, for AKI, at each prescription, inclusion and exclusion criteria will be assessed to determine the first prescription that fulfils all the criteria, i.e. the index prescription that determines the index date. The second AKI event in the figure below will have to have: no CKD ever before, and no two eGFR values<60 separated 90 to 365 days a part. And this is evaluated within the 1 year prior to the second Rx.



7.3.4.3 CKD outcome-specific exclusion criteria—in CPRD

For analysis of the CKD secondary outcome, patients will be excluded if they meet any of the following criteria:

- A diagnosis of CKD is recorded any time before or on the index date, i.e., during the available lookback time (see variable definitions in Section 8.3).
- The patient fulfils laboratory criteria for CKD before or at the index date (see Section 7.3.4.2).

7.3.4.4 Severe complications of UTI-specific exclusion criteria—in CPRD

For analysis of the severe complications of UTI outcome, patients will be excluded if they experienced chronic or acute pyelonephritis within the 6 months before or at the index date (see variable definitions in Section 8.3).

Prior diagnosis of UTI or fulfilling the validation criteria for severe complications of UTI outcome prior or at the index date will not constitute an exclusion criterion.

7.4 FOLLOW-UP

Start of follow-up: Follow-up will start the day after the index date (i.e., the date of the first prescription/dispensing for empagliflozin or for a DPP-4 inhibitor plus one day).

In CPRD only, start of follow-up for the two sensitivity analyses studying the CKD outcome will be defined as (1) starting follow-up after a minimum lag time of 3 months after start of exposure and (2) starting follow-up after a minimum lag time of 6 months after start of exposure.

End of follow-up will be defined for each of the outcome-specific cohorts. For the analysis of each outcome, each patient's follow-up time in a given cohort and in a given exposure category will end at the earliest of the following dates:

1. The date of the first outcome event (outcome-specific): a diagnosis of ALI1, ALI2, AKI, ~~or~~ CKD, severe complications of UTI, male/female GI, DKA, or male/female severe GI.
2. The date of death. Note that in HIRD, date of death will be based on discharge status for medical visits, recognizing that it is limited in sensitivity. Because the health plan is often terminated retroactively for decedents, they will be captured for censoring based on date of end of health plan eligibility.
3. The date of study end.
4. The date of transfer out of the practice or last collection date of a practice in CPRD, emigration date in Denmark, or end of health plan eligibility in HIRD.
5. The date ~~before~~ an outcome-specific exclusion criterion is met (see exclusion criteria in Section 7.3.4). Exclusion criteria will be specific for primary and secondary outcomes. Presence of outcome-specific exclusion criteria constitute criterion for not considering a diagnosis of the outcome as an identified outcome (i.e. not a potential outcome).
    a. An exception would be for ALI2 secondary outcome, that is, "among patients with or without predisposing conditions," for whom follow-up will not be censored for any of the ALI-specific exclusion criteria (see Section 7.3.4.1).
    b. For the ALI1 primary outcome, a diagnosis of any of the excluding conditions, except ALI, at any time between the index date and the event date will constitute a censoring criterion. If ALI1 excluding conditions occur at event date, it will constitute criterion for not considering the ALI diagnosis as an identified outcome.
        i. **The text above has been deleted so that in all data sources:** the presence of laboratory values that fulfil ALI criteria will be assessed as part of the case validation (not as part of case identification) and will not be used as censoring criteria (see Section 7.4). Electronic laboratory data for ALI will be used for the identification and description of cases of isolated elevated liver enzymes as described in Section 8.2.2.1.1.

    c. For AKI, a diagnosis of CKD anytime between the index date and the event date will constitute a censoring criterion, and AKI diagnosis will not be counted as an identified outcome.
        i. When using electronic laboratory values to identify CKD, the time period any time prior to the outcome date will be used. (Electronic lab

data for eGFR lab tests are not available in HIRD.) A patient will be censored at the date of the second eGFR test <60, where the first eGFR test should occur between 3 months and 1 year (i.e. 90 days and 365 days) before the second test. The first test can occur before or after the index date. If there are other records for eGFR test between the above tests, these will also have to be < 60. The process will be the same as described for CKD as an exclusion criterion in Section 7.3.4.2.

6. The end date of the first continuous treatment episode of the index drug (empagliflozin or DPP-4 inhibitor) plus a grace period, defined as follows (see also Section 8.1.2):
    a. Current use analysis (main analysis): 30 days after the end of days' supply of the last prescription/dispensing of the first treatment episode. Clarification notes:
        i. If a patient is censored during the current use time period, only include the exposure time until the censor date in the analysis. Subsequently, this patient would not have a recent use time period.
        ii. Having a prescription for SGLT2 (including empagliflozin) for DPP4i cohort, or a prescription for DPP4i or SGLT2 for empagliflozin cohort, is a censoring event for current use analysis (it's not a censoring event for ITT analysis)
    b. Recent use analysis: 90 days after the end of current use, i.e., 120 days after the end of the days' supply of the last prescription/dispensing of the first treatment episode. Clarification notes:
        i. If a patient has a censoring event during the recent use time period, the recent use time window would only include the days between the (current use end date + 1) to censor date. Thus, the recent use period will start the day after the end of current use period and will end the date of the censoring event.
        ii. If a patient whose study drug is empagliflozin gets an empagliflozin prescription during the recent use time window (e.g., on the 60th day of the recent use period), the empagliflozin rx during recent use would be considered a censoring event for this analysis, since patients are not allowed to enter the cohort if they do not have at least 12 months without empagliflozin, while recent use period ends 120 days after the end of supply.
        iii. Having a prescription for empagliflozin or other SGLT2i for DPP4i cohort, or a prescription for DPP4i or SGLT2 for empagliflozin cohort is a censoring event for recent use analysis (it's not a censoring event for ITT analysis)
    c. Sensitivity analysis extending the grace period: 90 days after the end of the days' supply of the last prescription/dispensing of the first treatment episode (see analysis in Section 11.6.2) This would be redefining current use with gaps and grace periods of 90 days instead of 30 days.
    d. Intention-to-treat sensitivity analysis: follow-up will **not** be censored at the end of the first continuous treatment of the index drug (continuous use or

treatment is defined as consecutive prescriptions/dispensings separated by 30 days or less. See further details of the definition of continuous treatment in Section 8.1.2.), i.e., the criterion of ending follow-up with the end date of the first continuous treatment episode of the index drug will not be applied for this analysis (see analysis in Section 11.6.2)

i. As further clarification: no censoring will be done for medication stop or change.

e. Within each exposure cohort, the end date will be the date of a prescription or dispensing of the other study medication (empagliflozin or DPP-4 inhibitor) or other SGLT2 inhibitor (even if these are not study medications). This criterion will not be applied for the intention-to-treat sensitivity analysis (see analysis in Section 11.6.2). Follow-up or current use will end the day of the prescription for this other medication. Switching will start the day after the other prescription. If a patient has an outcome at the day after a switching-redemption if would not be counted in the main analysis, only under the ITT analysis.

Patients will be able to enter another study cohort only once (not the original cohort, i.e. patients can switch from their initial study drug cohort, empagliflozin for example, to the DPP4i cohort, but not back to the empagliflozin cohort) if they fulfil inclusion and exclusion criteria, including no prior use of any of the study medications during the previous 12 months. Follow-up will *not* be censored if oral or injectable GLD (other than the index study medications) are prescribed in addition to empagliflozin or a DPP-4 inhibitor after the index date. All censoring criteria, except otherwise specified, will be applied in the intention-to-treat sensitivity analysis. In the ITT analysis, if a patient enters both cohorts, all follow-up after each index date is counted in each cohort, so the time on the second drug would also be counted in the time FU of the first drug.

Summary of exclusion and censoring criteria due to predisposing conditions:

	ALI1 Primary Outcome – in all data sources	
	For diagnosis of conditions: Any Time Before or at Index Date **For Lab values: – 1 Year (365 Days) Before or at Index Date**	**During follow-up (i.e. from Day after Index Date to Event Date (including event date)[a]**
Acute conditions (diagnosis codes)[a]	Exclude	Censor[a]
Acute conditions (lab criteria) – **Aarhus only** Fulfils **1** time the lab criteria for ALI during the 180 days before index date: (**A**) ALT or AST[a] ≥ 5 × ULN **or** (**B**) ALT or AST[a] ≥ 3 × ULN *AND* total bilirubin ≥ 2 × ULN **or** (**C**) ALP ≥ 2 × ULN	Exclude (**Aarhus only**)	Do not censor, see validation plan
Pregnancy	Exclude, only if pregnant at index date	Censor
Chronic predisposing conditions (diagnosis codes)	Exclude	Censor
Fulfils **2** times the lab criteria for ALI separated at least 180 days each, and both occurring during the 365 days before index date: (**A**) ALT or AST[a] ≥ 5 × ULN **or** (**B**) ALT or AST[a] ≥ 3 × ULN *AND* total bilirubin ≥ 2 × ULN **or** (**C**) ALP ≥ 2 × ULN	Exclude, if the 2 times are before index date (**all data sources**)	Do not censor, see validation plan

[a] Acute conditions referred to as "ALI" in the code list for exclusions will be assessed during follow as a potential case (since the codes are those in the outcome code list) rather than censored. These lab-criteria will not be used for censoring or for case-identification (in main analyses) – only as exclusion diagnosis and for validation purposes.

[b]AST value will be used only if ALT is not available.

	ALI2 Secondary Outcome – in all data sources	
	6 months (180 Days) Before or at Index Date	**During follow-up (i.e. from day after Index Date to Event Date, including event date)[a]**
Acute conditions (diagnosis codes) [a]	Exclude	Do not censor [a]
Acute conditions (lab criteria) – **Aarhus only** Fulfils **1** time the lab criteria for ALI during the 180 days before index date: (**A**) ALT or AST[a] ≥ 5 × ULN **or** (**B**) ALT or AST[a] ≥ 3 × ULN *AND* total bilirubin ≥ 2 × ULN **or** (**C**) ALP ≥ 2 × ULN	Exclude (**Aarhus only**)	Do not censor, see validation plan
Pregnancy	Exclude, only if pregnant at index date	Do not censor
Chronic predisposing conditions	Not applicable, do not exclude [b]	Do not censor
Fulfils 2 times the lab criteria for ALI separated at least 180 days each, and both occurring during the 365 days before index date: (**A**) ALT or ASTc ≥ 5 × ULN **or** (**B**) ALT or ASTc ≥ 3 × ULN *AND* total bilirubin ≥ 2 × ULN **or** (**C**) ALP ≥ 2 × ULN	Do not exclude, even if the 2 times are before index date (see validation plan)	Do not censor, see validation plan

[a] Acute conditions referred to as "ALI" in the code list for exclusions will be assessed during follow as a potential case (since the codes are those in the outcome code list) rather than censored.

[b] Note that there are codes that are shared between the list of acute and chronic conditions, and that only the duration of the lookback defines chronicity. E.g. viral hepatitis without mention of acute or chronic will be considered acute if recorded in the 6 months before index date. Thus, assessment of chronic conditions in the evaluation of the secondary endpoint is not applicable.

[c] AST value will be used only if ALT is not available.

	Acute Kidney Injury – in all data sources	
	Baseline: – 1 Year (365 Days) Before or at Index Date[a]	**During follow-up (i.e. from Day after Index Date to Event Date, including event date)[a]**
AKI diagnosis	Exclude if 6 months prior to index date	Do not censor, asses as potential case
CKD diagnosis	Exclude if any time prior to index date	Censor
Presence of 2 recorded eGFR < 60 mL/min/1.73 m^2 of 3 to 12 months apart[b]	Exclude[c]	Censor[d]

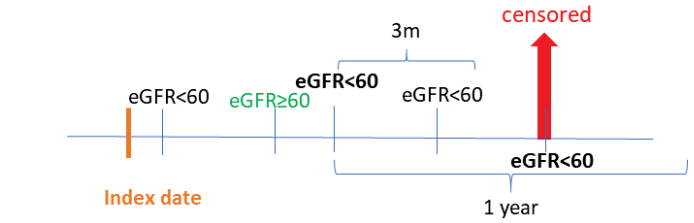
[a] Baseline period for validation purposes is 1 year before or on the index date.

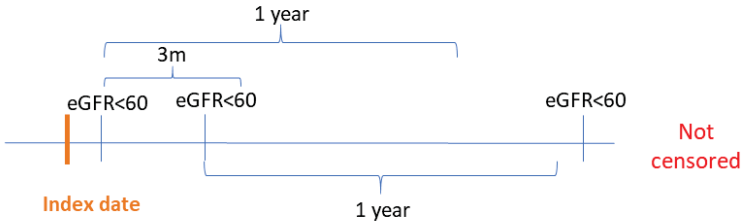
[b] Within 1 year before index date. The second eGFR measure that is <60 (when the patient is considered chronic) needs to occur from index date to event date (including event date), but the first occur before index date.

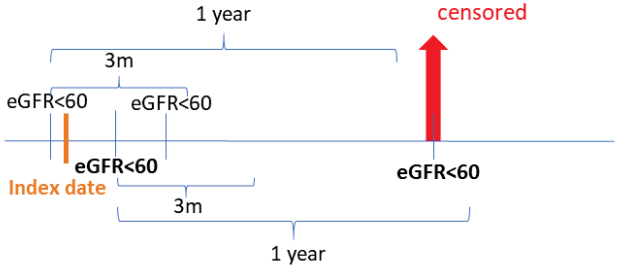
[c] The exclusion/censoring criteria for eGFR is used only for CPRD and Denmark. Lab results for eGFR are not available in HIRD.

[d] If the 2 during follow-up eGFR<60 are separated more than 12m, we do not censor.

Figure below shows examples of potential eGFR scenarios during follow-up.







	Chronic Kidney Disease -- in CPRD only	
	Baseline: – 1 Year (365 Days) Before or at Index Date[a]	**During Follow-up (i.e. from Day after Index Date to and at Event Date)**
CKD diagnosis	Exclude if any time prior to index date	Do not censor, asses as potential case
Presence of 2 recoded eGFR < 60 mL/min/1.73 m^2 of 3 to 12 months apart [b]	Exclude	Do not censor, asses as potential case

[a] Baseline period for validation purposes is 1 year before or on the index date.

[b] Within 1 year before index date

7.5 BASELINE, TIME WINDOWS, AND CALCULATED VISITS

In CPRD, the baseline or lookback time period is defined as the period of continuous registration from the practice up-to-standard date (CPRD specific) to the index date. The earliest date of the lookback period will be the later of the practice up-to-standard date and the current patient registration date. The practice up-to-standard date in CPRD is the date at which data from the practice are deemed to be of research quality. As the up-to-standard date variable was not available for CPRD Aurum at the time of data extraction, only the current patient registration date into the practice will be used.

In Denmark, the baseline time period is defined as the period from the start of ICD-10 coding system use in the registries (1994) to the index date. For laboratory data, this baseline period will start in 2000 for data from the Central Denmark Region and from 2015 for nationwide data. The baseline period for validation purposes will be specified in the validation plan.

- For lab data, the look back period is only 1 year prior to index date, i.e. earliest possible date used for lab-data is August 01 2013.

- Northern + Central Denmark region: maybe 2000 onwards, though platelets and ALP are presently not available for central Denmark region at all (post-hoc application for platelets and ALP was not possible in Central Denmark).

- Nationwide data: maybe 2015 onwards, though platelets and ALP not available for central Denmark region at all.

- Based on further inspection of available lab data it was agreed **that age-platelet would not be used** for exclusion, censoring or validation, although it may be of interest to describe it among cases around event date (final decision will be specified in the validation plan).

In HIRD, all available data prior to and including the index date will be used for the baseline or lookback time period. The baseline period is defined as the period of continuous enrolment prior to and including the index date. The earliest date of the lookback period can be 01 January 2006. The baseline period that will be used for validation purposes will be specified in the validation plan.

For medications and when specified for some conditions, the baseline period will be limited to the 6 months before and including the index date.

Due to the observational nature of this study, no mandated visits are required. Therefore, calculated study visits are not applicable.

8. VARIABLES

8.1 EXPOSURES

8.1.1 Study medications

For this study, eligible patients will be identified from prescriptions/dispensings for the following study medications of interest (index medications):

Empagliflozin:

- Jardiance (empagliflozin, ATC code A10BK03)
- Synjardy (empagliflozin/metformin hydrochloride, ATC code A10BD20) (Synjardy XR not used during the study period. IN DK: all Synjardy have same ATC, will check product number to confirm)

DPP-4 inhibitors:

- Januvia (sitagliptin: ATC code A10BH01)
- Galvus (vildagliptin: ATC code A10BH02)
- Onglyza (saxagliptin: ATC code A10BH03)
- Vipidia (alogliptin: ATC code A10BH04)
- Trajenta (linagliptin: ATC code A10BH05)
- Janumet (sitagliptin/metformin hydrochloride: ATC code A10BD07)
- Eucreas (vildagliptin/metformin hydrochloride: ATC code A10BD08)
- Kombiglyze (saxagliptin/metformin hydrochloride: ATC code A10BD10)
- Vipdomet (alogliptin/metformin hydrochloride: ATC code A10BD13)
- Jentadueto (linagliptin/metformin hydrochloride: ATC code A10BD11)

If additional DPP-4 inhibitor drugs are marketed in the future in the UK, Denmark, or the US during the study period, they will be added to the comparator group.

Fixed-dose combinations of SGLT2 inhibitors with DPP-4 inhibitors, such as Glyxambi (empagliflozin/linagliptin), will not be included in either study cohort, but will be included in the lists of SGLT2 inhibitors with DPP-4 inhibitors to be considered to define use prior to index date.

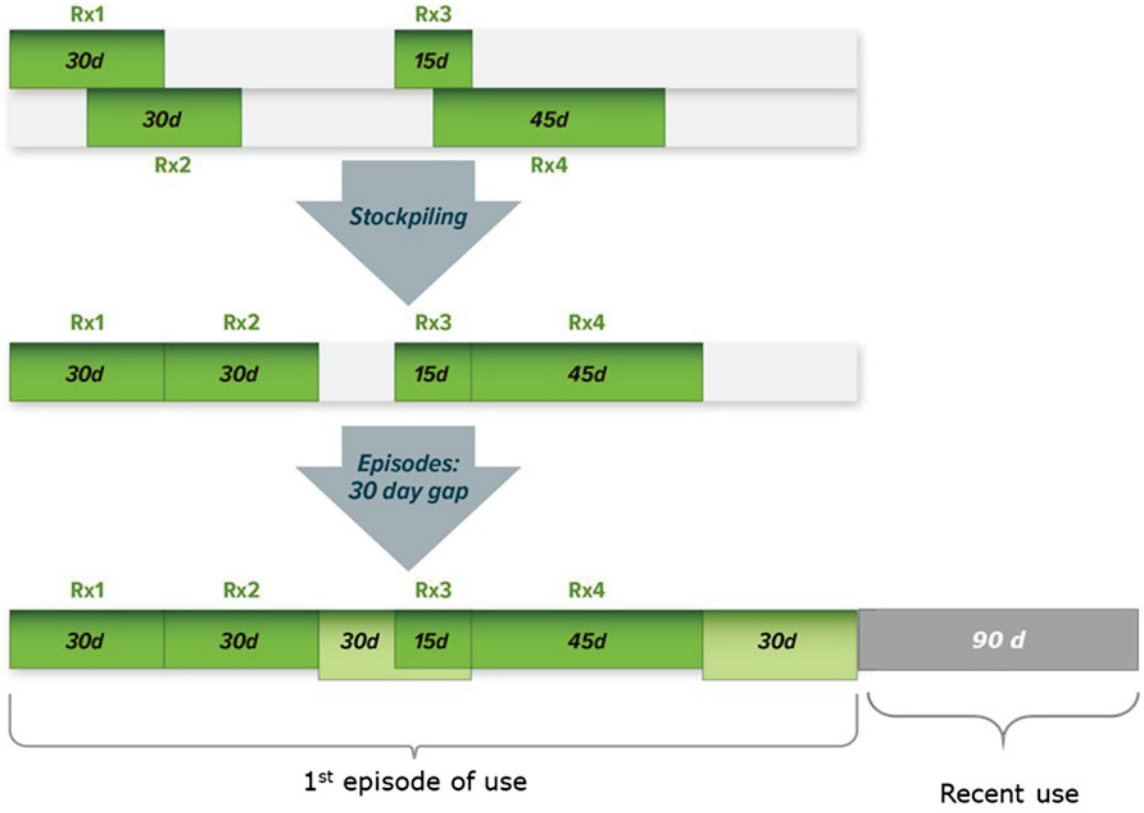
8.1.2 Exposure time

The time at risk or risk/exposure time window for each new user of a study medication will be categorised into two mutually exclusive categories of risk, as follows:

- *Current use* (current time at risk): the risk/exposure window for the first episode of **continuous use**, defined as consecutive prescriptions/dispensings separated by 30 days or less. Current use will start on the day after* the date of the first prescription/dispensing during the study period, concatenate days' supply for consecutive prescriptions/ dispensings (as long as the gap between the end of days' supply of one prescription/ dispensing and the start of the next one is no more than 30 days) and end 30 days after the

end of the last prescription's/dispensing's days' supply. Overlapping time at risk from current use for consecutive prescriptions/dispensings of the study medication will be concatenated, with the overlapping time added at the end of the concatenated prescription/dispensing (stockpiling). Note that in CPRD and HealthCore days' supply have been imputed for missing values and outliers, therefore, already accounted for an upper limit of days' supply. In Denmark, based on descriptive data on stockpiling, an upper limit of 180 days was applied for stockpiling. For consecutive prescriptions/ dispensings of the study medication separated by gaps of 30 days or less, time at risk from current use will include the gaps between prescriptions/dispensings. The main analysis is based on current use during the first continuous episode of treatment. End of day's supply will be estimated according to prescription/dispensing instructions in CPRD or based on available information on the duration of dispensings (e.g., days' supply, number of packages bought, strength, and number of pills) in Denmark and HIRD. See Figure below.
*Current use will start on the day after the date of the first prescription (index date), thus, assuming a patient redeems a prescription of pills for 30 days use at day 1, then day 2 would be *current use* start, day 31 is last day of current use, day 61 would be last day of the 30 days grace period, i.e. last day of the *current use* (if there are no further prescriptions), and day 62 would be the discontinuation (and start of recent use). An event happening at the day of prescription redemption (day 1) would not count, and events happening at day 62 or later would also not count for main analysis. So *current use* period is from the start of the day 2 (the day after index date) until the end of day 61."
Note that for all prescriptions of medications being evaluated either as current use of study meds or switching drugs, i.e. empagliflozin, DPP4i, SGLT2i, start of each prescription would be the day after the prescription.

- *Recent use* (recent time at risk): the risk/exposure window for recent use will start at the end of current use (30 days after the end of supply) and end 90 days later (which is 120 days after the end of supply). Recent use will be defined from the end of the first episode of current use (30 days after the end of supply) through 90 days later (which is 120 days after the end of supply).



Exposure time definitions will be modified in sensitivity analyses, as follows:

- Extended grace period: current use grace period of 90 days after the end of supply (instead of 30 days), i.e. continuous use is defined as consecutive prescriptions or dispensings separated by 90 days or less (instead of 30 days)
- Intention-to-treat analysis: all follow-up will be considered to be exposed time
- For the CKD outcome (CPRD only), a sensitivity analysis will be performed with current use starting at the following time points:
  - At least 3 months after the index date
  - At least 6 months after the index date

8.1.3 Day's supply and duration of exposure

Episodes of continuous use will be defined based on the estimated days' supply of each prescription/dispensing. Day's supply will be defined as follows:

- **In CPRD** as (1) recoded in CPRD; (2) calculated from numeric daily dose, quantity, and pack type when available; or (3) if not available or if values are implausible, imputed using the mode within each drug substance, stratum of age category (<40,40-49, 50-59, 60-69, 70-79, or 80+ years), sex (male, female), and calendar period (categorised as <2014, 2014-2015, 2016-2017, and 2018-2019), as possible.
  Distribution of days of supply above percentile 99 and below percentile 1 will be checked, and their values will be set to missing if considered implausible.

When days' supply are missing, mode value for the same drug substance, age category, sex, and calendar period :
Otherwise, mode by drug substance, age category, and sex
Otherwise, mode by drug substance and age category
Otherwise, mode by drug substance
Otherwise, mode by cohort (Empagliflozin/DPP4i)

See also assumptions described in section 8.1.4, in case quantity is available but not numeric daily dose (ndd): "When ndd is missing, a value of 1 will be assumed for the study drugs (empa/DPP4i) not in combination with metformin, and 2 will be assumed for the study drugs in fixed-dose combination with metformin (empa/metf, DPP4i/metf). Vildagliptin is an exception, for which we will assume that each patient uses 2 pills each day."
For insulin, if days´s supply is missing, then we will assume a duration of 28 days (the mode of days between prescription or refill frequency). Quantity and daily_dose will not be used as after examining the data it mostly results in implausible values for this drug.

- **In Denmark**, calculated from pack sizes, and refill dates. For mono-preparation formulas, we will assume that each patient takes one pill each day for both Empa and DPP4is. Vildagliptin [A10BH02] is an exception, for which we will assume that each patient uses 2 pills each day. For fixed-combinations with metformin, we will assume that each patient uses 2 pills each day.
- **In HIRD**, days' supply is a data element in the pharmacy claims database and is available as a separate variable. When days' supply values are missing or implausible (e.g., > 150 days), days' supply will be imputed using the mode within each drug substance, stratum of age category, sex, and calendar period, as possible. Implausibility will be decided after a close examination of drug use patterns in the data.

Duration of exposure will include only the first episode of continuous use, i.e. current use. Duration of use will be categorised as short or long based on the available data on duration of exposure for empagliflozin, which will be applied to the comparator medications. Exposure and outcome events will be accounted in one or the other cohort, i.e., cohorts can't overlap as time is censored if the patient switches cohorts, except for the for ITT analysis where the patient can be accounted in more than one cohort.

8.1.4 Dose

Dose will be the daily dose at the index date.

In all data sources, the dose will be estimated from the available recorded information (e.g., strength, number of units or dose number, amount of drug prescribed, dose frequency), if possible or will be imputed using the mode within each stratum of drug substance, age category, sex, and calendar period. Evaluation of different doses of empagliflozin will be performed if variation in the dose used by the empagliflozin cohort is observed and an adequate number of outcome events occur within the different daily dose categories. Dose

will be categorised as low or high according to the values of daily dose below or above the recommended dosages. The dose thresholds that will be used to categorise empagliflozin and DPP-4 inhibitors (including those in fixed-dose combination with metformin) will be as follows:

- Empagliflozin: low if ≤ 10 mg/day, and high if > 10 mg/day (the recommended starting dose for Jardiance is 10 mg once a day, which can be increased if necessary to 25 mg daily in suitable patients, and the recommended starting dose in the Synjardy fixed-dose combination of empagliflozin and metformin is 5 mg of empagliflozin twice a day, which can be increased if necessary to 12.5 mg twice a day in suitable patients; thus, the dose will be low if 10 mg/day or less and high if greater than 10 mg/day)
- Sitagliptin: low if ≤ 50 mg/day and high if > 50 mg/day. Sitagliptin dose categories will be defined once all the data on the distribution of sitagliptin dosages in all data sources has been examined.
- Saxagliptin: low if ≤ 2.5 mg/day and high if > 2.5 mg/day
- Alogliptin: low if ≤ 12.5 mg/day and high if > 12.5 mg/day
- Linagliptin: low if ≤ 2.5 mg/day and high if > 2.5 mg/day*
- Vildagliptin: low if ≤ 50 mg/day and high if > 50 mg/day

* Trajenta dose <5mg will be considered as high dose as this medication is most likely prescript as 5 mg.

Dose will be the daily dose at the index date and will be defined as follows for each data source:

- **In CPRD**: The daily dose prescribed will be estimated for each drug separately (e.g., empagliflozin, and empagliflozin/metformin) as the strength multiplied by number of daily doses (*n.d.d.*). For fixed-dose combinations, daily dose will be estimated for the active substance other than metformin.
  When ndd is missing, a value of 1 will be assumed for the study drugs (empagliflozin/DPP4i) not in combination with metformin, and 2 will be assumed for the study drugs in fixed-dose combination with metformin (empagliflozin/metformin, DPP4i/metformin).
  Vildagliptin is an exception, for which we will assume that each patient uses 2 pills each day. If a patient has more than 1 prescription at the index date, the one with the lowest daily dose (ndd*strength) will be considered the index prescription.
- **In Denmark**: The daily dose prescribed will be estimated for each drug separately (e.g., empagliflozin and empagliflozin/metformin) as the strength assuming that each patient uses 1 pill each day for monopreparation formulas (except vildagliptin for which we assume that each patient uses 2 pills each day) and 2 pills each day for fixed-combinations with metformin. For fixed-dose combinations, daily dose will be estimated for the active substance in the combination that is other than metformin. In Denmark, based on this definition, there will be no patients in "intermediate dose categories", i.e. dose categories other than those determined by the strength.
- **In HIRD**: The prescribed daily dose at the index date will be estimated for each drug separately (e.g., empagliflozin and empagliflozin/metformin) as the empagliflozin strength multiplied by the quantity dispensed, divided by the days' supply.

As indicated in the footnote of Analysis Table 3, (applies also for dose thresholds for Analysis Table 12.6):

- When there are two prescriptions/dispensings at index date for the same drug substance but with different strength, the lowest dose will be reported.

- When there are two prescriptions/dispensings at index date for different drug substance used as single drug, i.e. not in fixed-dose combination, both will be reported. This is very infrequent and due to concerns on small cell counts and deductive disclosure in Denmark will report the one with the lexicographically first ATC code, and this will be described in the report.

- When there are two prescriptions/dispensings at index date, one as a single drug substance (e.g. Jardiance) and the other one as a fixed-dose combination (e.g. Synjardy), the dose of the drug used as single drug substance will be reported.

As indicated in the footnote of Analysis Table S7, if a patient had fills for more than one DPP4-i on their index date these were counted in each relevant row.

HIRD annotation/clarification: we only reported maximum one fill per patient per therapy-outcome cohort. When there are 2 fills at index for different drug of same type (i.e., empa single entity, empa + metformin, metformin, DPP4i):

1 – keep fill with lowest daily dose

2 – if multiples remain (e.g., same daily dose but some other difference, keep single entity over +metformin)

3 – if multiples remain keep fill with longest days supply, highest quantity dispensed, and lowest drug strength dispensed

Denmark:

DK: 12.6 Prescriptions for same DPP4i at index-date: Regarding the comment to 12.6 "please consider/use highest dose". Previous (in table 3) I used the lowest dose. This is now updated, so for future versions of table 3 I will be using the highest dose, if prescriptions with different strength is redeemed for the same drug at index-date. And this is true for table 12.6 as well.

- Regarding different types of DPP4i. Here we have already discussed:
  - In case of both combination drug (fixed dose compared to metformin) and single drug (free dose) → I select single drug.
  - In the very seldom case (n<5) of multiple DPP4-type drugs → I choose a semi-random (I choose the lexicographic first name).

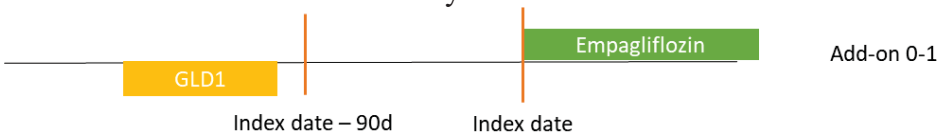
  (This has been described in mail 6/10-2020, and decided at meeting 23/10-2020).

8.1.5 Mono/dual/triple therapy, add-on, and switching

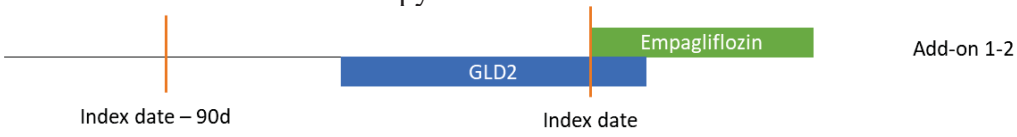
8.1.5.1 In CPRD and HIRD:

New users of the index medications will also be classified as patients augmenting their therapy (adding medications to baseline therapy) or switching medications without looking for prescriptions/dispensing after the index date. Subgroups defined by antidiabetic medications used concomitantly with the index exposure, and taking into account prior GLD therapy (including insulin),(not including index date), are as follows (summary figure added at the end):

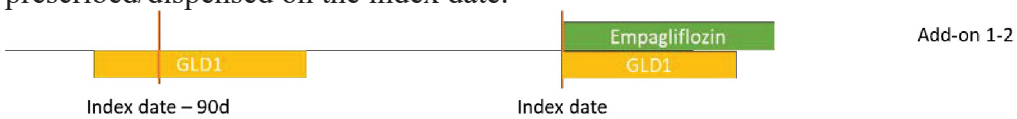
1. **Add-on** to naïve to become monotherapy **"Add 0-1":** patients who start monotherapy index medication with no other prescription/dispensing for other GLD and have no prior use, i.e. no prescription/dispensing, of GLDs in the 90 days before the index date. Relatively few patients with T2D are expected to start DPP-4 inhibitors and be naive to other GLDs, as the monotherapy indication for these drugs is restricted to patients intolerant to metformin and sulfonylureas.



2. **Add-on** to monotherapy to become dual therapy **"Add 1-2"**: adding an index medication to monotherapy with one other GLDs to become patients on dual therapy:
   a. Patients who start a monotherapy index medication and

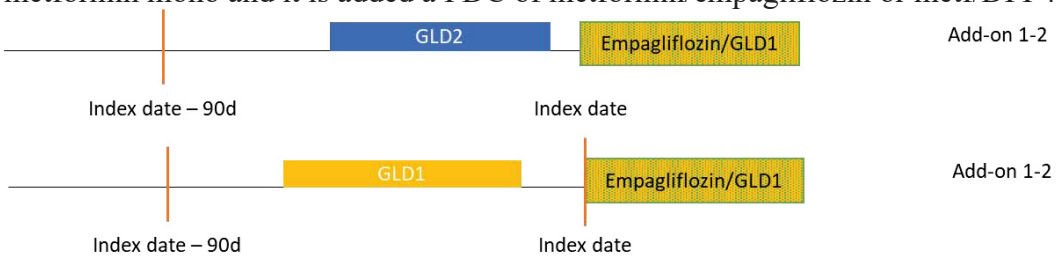


   - have a prescription/dispensing of a different GLD before the index date with end of days' supply after the index date.
   - have a prescription/dispensing of a different GLD in the 90 days before the index date with end of days' supply before the index date and have the other GLD medication (irrespective of being the same as the prior one or a different one) prescribed/dispensed on the index date.



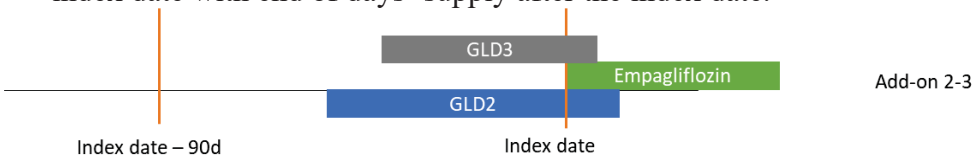
   b. Patients who start fixed-dose combinations of index medications with no other GLD and have one other GLD prescribed/dispensed before the index date and with end of days' supply in the 90 days before the index date (irrespective of being the same as the one at index on FDC or a different one). It means that the patient used 1 drug and stops that one and starts a FDC. However, it could be that the patient was on

metformin mono and it is added a FDC of metformin/empagliflozin or metf/DPP4i.

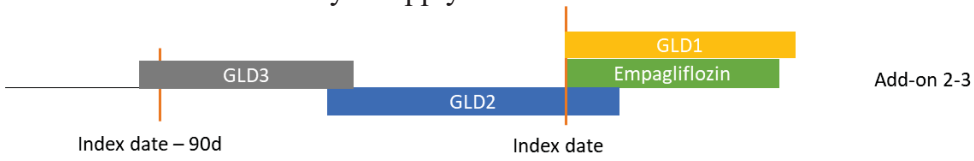


3. **Add-on** to dual therapy to become triple therapy "Add 2-3": adding a study drug to therapy with two other GLDs to become patients on triple therapy:

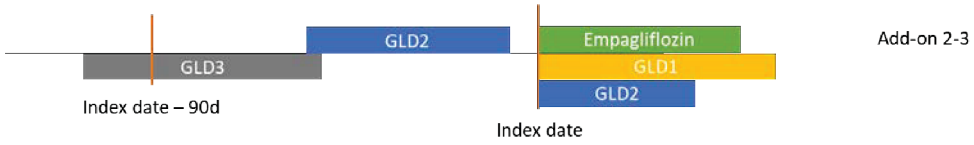
   a. Patients who start monotherapy index medication and have either:
   - prescriptions/dispensings for two (exactly 2, not 3 or more) other GLDs before the index date with end of days' supply after the index date.



   - One prescription/dispensing for one other GLDs on the index date and have prescription/dispensing of two GLDs with end of days' supply in the 90 days before index date and one prescription/dispensing for another GLDs before the index date with end of days' supply after the index date.



   - two prescriptions/dispensings for other GLDs on the index date and have prescription/dispensing of two GLDs with end of days' supply in the 90 days before the index date.

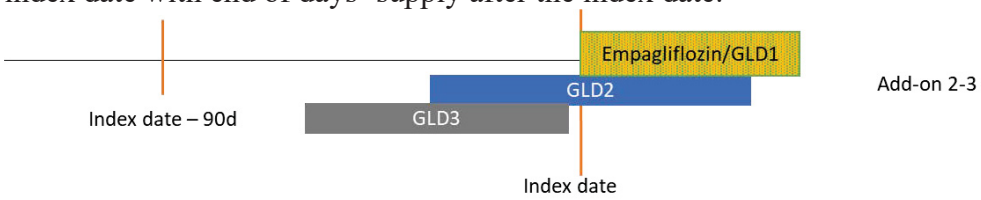


   b. Patients who start fixed-dose combinations of index medications with either:

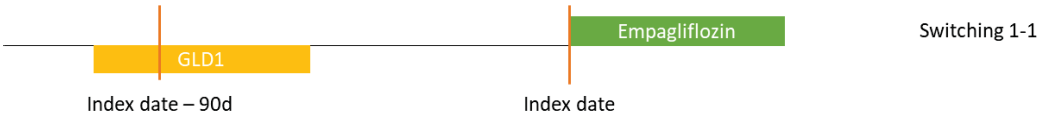
   - Two other GLDs prescriptions/dispensings with end of days' supply in the 90 days before the index date and one other for another GLD prescription/dispensing on the index date.

   - One other GLD prescription/dispensing with end of days' supply in the 90 days before the index date and 1 prescription/dispensing for another GLD before the

index date with end of days' supply after the index date.

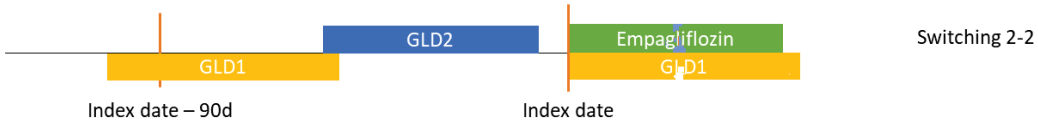


4. Switching from monotherapy to monotherapy "Switch 1-1": switching from monotherapy with another GLD to monotherapy with a study medication:
   a. Patients who start monotherapy with index medication with no other prescriptions/dispensing for another GLD on the index date and have ~~oneano~~other GLD prescription/dispensing with end of days' supply in the 90 days before the index date.

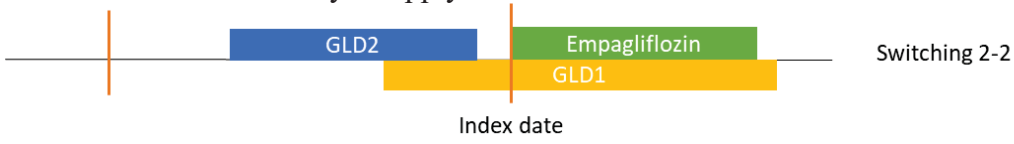


5. Switching from dual to dual therapy "Switch 2-2": switching from dual therapy with another GLD to dual therapy with a study drug and another GLD:

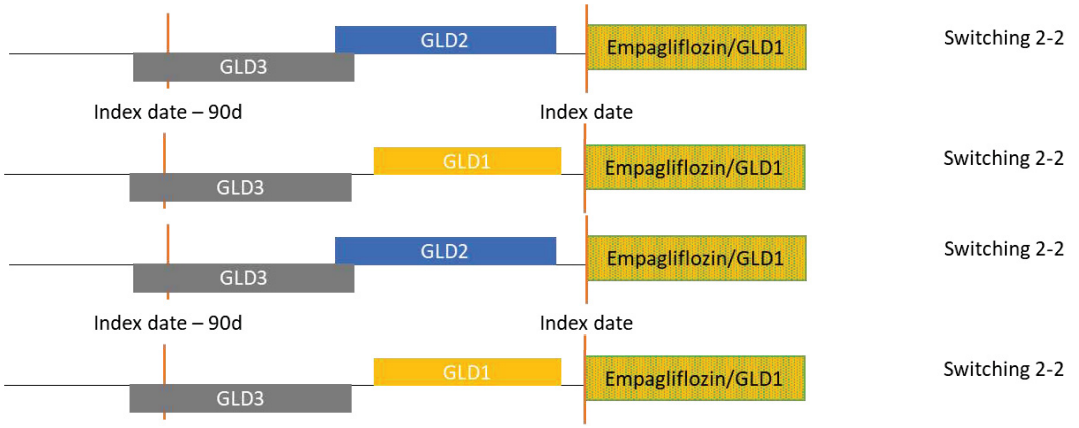
   a. Patients who start monotherapy index medication and have:
      - One other GLD prescription/dispensing at index date and two other GLDs prescriptions/dispensings with end of days' supply in the 90 days before the index date.



      - One other GLDs prescriptions/dispensings with end of days' supply in the 90 days before the index date and one other GLDs prescriptions/dispensings before the index date with end of days' supply after the index date.

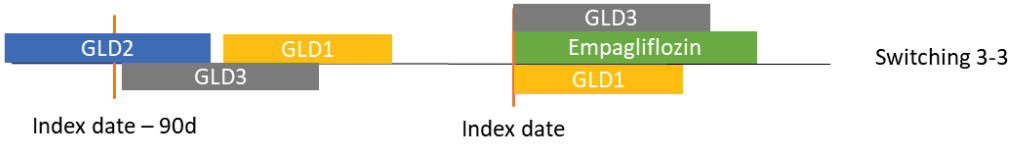


   b. Patients who start a fixed-dose combination index medication and have two other GLDs prescriptions/dispensings with end of days' supply in the 90 days before the index date (either with one of the two drugs being equal to the other GLD in the fixed-dose combination or not).

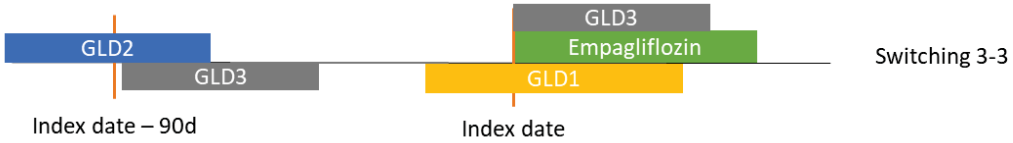


6. Switching from triple to triple therapy "Switch 3-3": switching from triple therapy with another GLD to triple therapy with a study drug and 2 other GLDs:

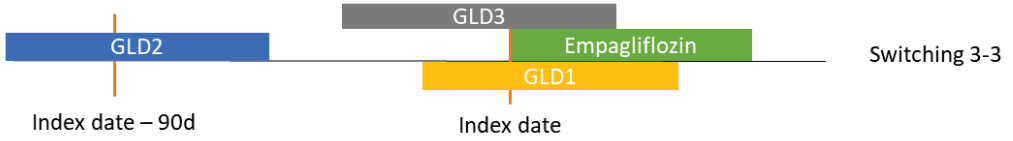
   a. Patients who start monotherapy index medication and have:
   - two prescriptions/dispensings for other GLDs at the index date and have three prescriptions/dispensings of other GLDs with end of days' supply in the in the 90 days before the index date.



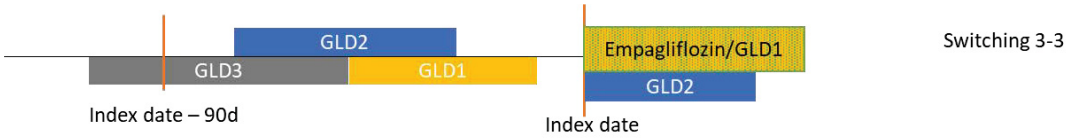
   - Two prescriptions/dispensings of other GLDs with end of days' supply in the in the 90 days before the index date and one prescriptions/dispensings before the index date with end of days' supply after the index date



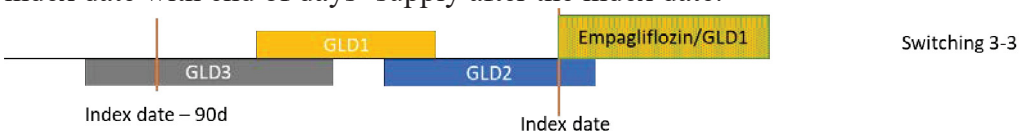
   - One prescriptions/dispensings of other GLDs with end of days' supply in the in the 90 days before the index date and two prescriptions/dispensings of other GLDs (other than the index medication) before the index date with end of days' supply after the index date



   b. Patients who start a fixed-dose combinations of index medication and have:
   - One other GLD at the index date and have ~~prior use of~~ three prescriptions/dispensings for 3 GLDs with end of days' supply in the 90 days before the index date.

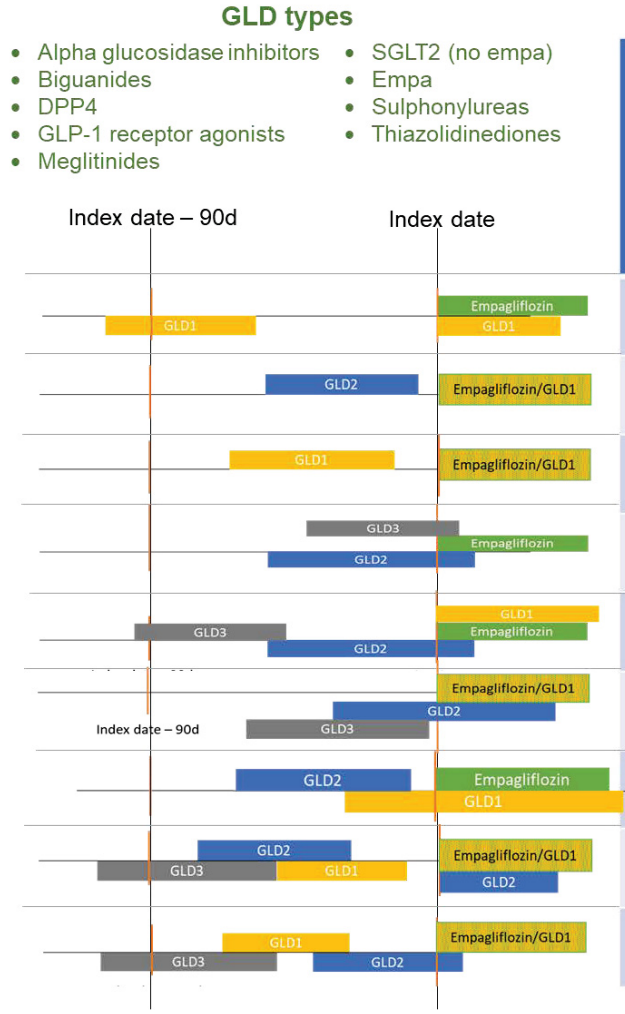


- Two prescriptions/dispensings of other GLD with end of days' supply in the in the 90 days before the index date and one prescriptions/dispensings before the index date with end of days' supply after the index date.



7. A group called "Other add-on", "Other switching", and "Reduced" will be created for patients not classified in the categories described above and listed in the table, e.g., switching from triple to dual or monotherapy or from dual to monotherapy. Depending on the number of patients in this group, these may be evaluated for risk of the outcomes of interest. Includes for example patients with 4 medications or patients reducing treatment.

Figure summarizing the classification:



Number of GLD classes with a Rx ending within 90 prior index (E)	Number of GLD classes with a Rx starting within 90 prior index and ending after index (S)	Number of GLD classes with a Rx on the index date (I)	Classification E + S → S + I If E+S < S+I → Add-on If E+S = S+I → Switching
1	0	2	Add-on 1-2
1	0	2	Add-on 1-2
1	0	2	Add-on 1-2
0	2	1	Add-on 2-3
1	1	2	Add-on 2-3
1	1	2	Add-on 2-3
1	1	1	Switching 2-2
3	0	3	Switching 3-3
2	1	2	Switching 3-3

In addition, patients will be considered users of insulin at the index date if (1) they receive a prescription/dispensing of insulin at the index date or (2) they have a prescription/dispensing of insulin before the index date with an end of days' supply that overlaps the index date for at least 1 day. The insulin use indicator variable will be used as a stratification variable in the analyses.

End of each GLD prescription (including Insulins) will be based on days' supply as defined in each database. When these are missing, we will impute them following the same process described in section 8.1.3:

- **In CPRD** as (1) recoded in CPRD; (2) calculated from numeric daily dose, and quantity when available; or (3) if not available or if values are implausible, imputed using the mode within each drug substance, stratum of age, sex, and calendar period as possible. Distribution of days of supply above percentile 99 and below percentile 1 will be checked (after step 2), and their values will be set to missing if considered implausible. In CPRD, if a non-insulin GLD prescription has missing ndd, this will be assumed to be 2 per day if it is in fixed-dose combination with metformin, and 1 per day if it's a single ingredient prescription. Vildagliptin is an exception, for which we will assume that each patient uses 2 pills each day. In CPRD, for insulin, if days´s supply is missing, then we will assume a duration of 28 days (the mode of days between prescriptions), quantity and daily_dose will not be used as they mostly results in implausible values.
- **In Denmark**, calculated from pack sizes, and refill dates.
- **In HIRD**, days' supply is a data element in the pharmacy claims database and is available as a separate variable. When days' supply values are missing or implausible (e.g., > 150 days), days' supply will be imputed using the mode within each drug substance, stratum of age category, sex, and calendar period, as possible. Implausibility will be decided after a close examination of drug use patterns in the data.

The following information will be needed for Analysis Table 2 in the SEAP Table Shells:

- Number of new users of fixed-dose and free-dose <u>combinations with metformin</u>:
    - Number of new users of fixed-dose combinations with metformin
    - Number of new users of free-dose combinations with metformin

  Combination with metformin at the index date will be defined as (1) receiving a prescription/dispensing of a fixed-dose drug combination with the drug of interest at the index date or (2) they have a prescription/dispensing of metformin before the index date with an end of days' supply that overlaps the index date for at least 1 day. If the patient has, for instance, empagliflozin and metformin/sulfonylureas at the index date, this patient will be reported in the table as being with free dose combination with metformin.

- Number of new users who were switchers:
    - Switchers mono-mono therapy: 1-1
    - Switchers dual-dual therapy: 2-2
    - Switchers triple-triple therapy: 3-3
    - Other switching patterns
- Number of new users with "augmentation therapy"
    - Add-on to naive to become monotherapy: 0-1
    - Add-on to become dual therapy: 0-2, 1-2
    - Add-on to become triple therapy: 0-3, 1-3, 2-3
    - Other Add-on patterns
- Reduced number of therapies
- Number of new users on monotherapy: x-1 (with x being any number, 0-1, 1-1, 2-1, 3-1)
- Number of new users on dual therapy: x-2 (0-2, 1-2, 2-2, 3-2)
- Number of new users on triple therapy: x-3 (0-3, 1-3, 2-3, 3-3, 4-3,…)
- Number of new users with concomitant use of insulin
- Number of prescriptions of the study medication during the index episode
- Duration of index episode (days)

8.1.5.2 In Denmark

The duration of drug use in Denmark requires assumptions on number of pills per day. In clinical practice, the number of pills and dose (e.g., mg) per day is variable for some drugs such as metformin or sulfonylureas and assuming a specific number of pills per day may lead to under or overestimation of the duration of use. The same may occur if DDDs are used since these are fixed and patients may be taking different amount per day. Thus, based on Aarhus assessment it was considered that deriving duration of use of all GLDs based on information from variables DDD and number of pills is not reliable.

The Danish register has available the date of prescription of a drug and investigators in Denmark proposed to look into the number of prescriptions before and after index date to determine "Switchers" or "Add-Ons".

DK will assess the number of GLD classes redeemed during

- A) the time period 90 days before index date, and
- B) the time period from index date (included) to 90 days after.

8.2 OUTCOMES

8.2.1 Primary outcomes

Study outcomes will be identified based on diagnoses, tests, and procedures recorded in the outpatient visits (Read codes in CPRD GOLD and SNOMED/EMIS codes in CPRD Aurum) or hospitalisations (*International Statistical Classification of Diseases and Related Health Problems, 10th Revision* [ICD-10], *Office of Population Censuses and Surveys, Fourth Revision* [OPCS4] codes in Hospital Episode Statistics [HES], and ICD-10-CM, and ICD-9-CM codes in HIRD), during follow-up. Note that, unless otherwise specified, for hospitalisations, all diagnoses, i.e., primary and secondary discharge diagnoses, will be considered. Absence of any record (or claim in HIRD) with a diagnosis code for a condition will be taken to mean absence of the condition. Only the first outcome event during follow-up for each outcome will be considered.

For outcomes identified through hospital data, the date of the event will be the admission date, as opposed to the discharge date. In CPRD GOLD/Aurum, for outcomes identified through Read/SNOMED diagnosis code + records for hospitalisation, ED visit, or referral, the date of the event will be the date of the Read/SNOMED diagnosis code in GOLD/Aurum.

Note that:

- ALI1, ALI2, and AKI outcomes are identified through primary or secondary discharge codes, referrals/specialist visits (hospital outpatient clinic visits in DK) and ED visits. Note that the secondary diagnosis refers not the second diagnosis but all those not identified as primary, e.g. in CPRD these may be up to 15.
- DKA outcome is identified via only hospital primary or secondary discharge diagnosis and ED visits, but not outpatient specialist or referral data.
- Severe UTI: identification of pyelonephritis is only through primary discharge diagnosis
- CKD and Genital infections are ascertained through all type of codes, including primary and secondary discharge diagnosis, referrals, ED visits, and outpatient primary care codes.

If the outcome event occurs the same date of an outcome-specific censoring reason it will not be considered an identified outcome for study purposes.

Validation of identified cases other than through electronic algorithms will be implemented for all five primary outcomes, all three secondary outcomes, and in CPRD and HIRD it will also be implemented for the two sensitivity analysis outcomes (outpatient ALI and AKI). If the number of cases identified for any of the outcomes is 150 or fewer, all cases will be validated for those outcomes. Otherwise, a sample of 150 cases will be selected for validation for each outcome with more than 150 cases identified. Stratified random sampling will be used to sample cases evenly across the empagliflozin and comparator cohorts, as feasible. Moreover, to maximise the efficiency of the validation process, the cases from the pool of cases on current use will be selected preferentially over the other cases. Details on validation will be described in the validation plan stand-alone document (Annex 4). Briefly, validation will be performed as follows:

- **In CPRD**, potential cases will be validated by review of computerised clinical information, including outpatient laboratory data, and questionnaires sent to general practitioners (GPs).
- **In Denmark**, potential cases will be validated by computerised laboratory data in registers (both primary and secondary care test results, i.e., laboratory results from samples taken by GPs and other primary care providers outside hospitals and lab results from samples taken in the hospital [emergency department (ED), hospital specialist outpatient clinics, or inpatient admissions]).
- **In HIRD**, potential cases will be validated by computerised clinical information, including partial outpatient laboratory data and/or by abstraction of information from provider or hospital medical records, including extraction of inpatient laboratory data when possible.

In all data sources: for all the outcomes that undergo validation, and based on results from validation, a sensitivity analysis will repeat the main analysis but include only confirmed cases.. The protocol states that the if less than 70% of the cases are validated, then the analysis will correct for the estimated PPVs (see Section 11.6.3, Analysis Table S2). However, because PPVs may be low for some of these outcomes with less than 70% of the cases validated, and the analysis among validated cases may be required a posteriori, it has been decided a priori that both Analysis Table S1.2 (including confirmed cases only) and Analysis Table S5 (correcting IRR by the PPVs) will be done for all the outcomes [R18-1561] (see the description of analyses in Section 11.6.3).

Although initially potential cases occurring during the first 72 hours of hospitalisation were going to be excluded, based on discussions during the development of Validation Plan it was decided that cases of ALI, AKI and DKA occurring during the first 72h of hospitalisation would also be included given that evaluation of these 72 h can only be done accurately in HIRD, and only in a subset of cases from CPRD that have GP Qxs responses, but cannot be done among cases validated through electronic laboratory values in CPRD and in Denmark.

In addition, for DKA the study aims to describe, if possible, the number of cases that occurred after surgery. Because it is not possible, in most data sources, to disentangle if the event occurred before or after surgery, the study would describe the *number of cases that had recent surgery*. **Recent surgery** will be defined as that occurring within 30 days before the event date or within the same hospitalisation (see Validation Plan for further details).

8.2.1.1 Acute liver injury: hospitalisation, ED visit, or specialist visit for ALI1 in patients without predisposing conditions—in all data sources

8.2.1.1.1 Case definition and identification

Potential cases of ALI1 will be identified by the following process:

- **In Denmark,** potential cases will be identified through primary or secondary hospital discharge ICD-10 codes, ED visit ICD-10 codes, or hospital outpatient clinic ICD-10 codes suggestive of ALI1.
- **In HIRD**, potential cases will be identified as patients who had one or more claims for a hospital visit, ED visit, or gastroenterology specialist office visit with an ICD-9-CM/ICD-10-CM diagnosis code for ALI1.
- **In CPRD:**
  - **In HES**, potential cases will be identified through primary or secondary hospital discharge ICD-10 codes suggestive of ALI1.
  - **In GOLD,**
    - During non-linkable person-time**,** potential cases of ALI1 will be identified through primary care codes (Read codes) suggestive of ALI1 associated with a record for hospitalisation within 30 days of the code suggestive of ALI1,
    - During all available person-time, i.e. irrespective of linkage with HES, potential cases of ALI1 will be identified through primary care codes (Read codes) suggestive of ALI1 associated with a record for specific specialist referral within 30 days of the code suggestive of ALI1.
  - ED visits, or specific specialist referral within 30 days of the code suggestive of ~~ALI~~ALI1. Hospitalisations, ED visits, and referrals will be defined as follows:
    - A referral to a internal medicine specialist gastroenterologist, or liver specialist will be identified through Read codes for referral to these specialists, through National Health Service (NHS) classification of specialities recorded as gastroenterology (*nhsspec* = 17) (referral to internal medicine specialist is not available in as lookups for FJSA or NHS).
    - A hospitalisation or ED visit: identified through Read codes for hospitalisation/ED visit or referral for hospitalisation/ED visit, a referral for hospitalisation recorded in the referral data set (inpatient = 1), or a consultation type indicating that the patient has been hospitalised or seen in the ED (*constype* = 11, 12, 18, 23, 47, or 48, or an *nhsspec* = 14).
  - **In Aurum,** potential cases of ALI will be identified through primary care codes (SNOMED/EMIS codes) suggestive of ~~ALI~~ALI1 associated with a record for hospitalisation, ED visits, or specific specialist referral within 30 days of the code suggestive of ALI1. Hospitalisations, ED visits, and referrals will be defined as follows:
    - A referral to an internal medicine specialist, gastroenterologist, or liver specialist will be identified through codes for referral to these specialists in the *medcodeid*

variable. Identification through the variables *refsourceorgid*, *reftargetorgid* in the referral data set was not available at the time of the data extraction (Sept 2019).

- A hospitalisation or ED visit will be identified through any of the following:
  - SNOMED/EMIS codes for hospitalisation/ED visit or referral for hospitalisation/ED visit in *medcodeid* variable in the observation data set or in variable *consmedcodeid* from the consultation data set.
  - Referral for hospitalisation recorded in the variable *refservicetypeid = 2* in the referral data set. Identification through variables *refsourceorgid* or *reftargetorgid* variables in the referral data set was not available at the time of the data extraction (Sept 2019).
  - Consultation type indicating that the patient was hospitalised or seen in the ED using the *conssourceid* variable. Identification of consultation type through the *cprdconstype* variable was not available at the time of the data extraction (Sept 2019).

The date of the event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.

The list of all codes to be used in each database is available in Annex 2.

In all data sources, electronic laboratory data will **not** be used to identify potential cases of ALI1 (nor ALI2) (except in the analysis of elevated isolated liver enzymes) (see Section 8.2.2.1.1). However, for validation of potential cases, electronic laboratory data will be used, whenever possible, to assess whether the patients fulfil the laboratory criteria for ALI and identified cases can be confirmed without needing to undergo further validation through GP questionnaires or chart abstraction. Evaluation of electronic laboratory data for confirmation purposes and to identify chronic liver injury (see CLI - see Section 7.3.4.1.page 23 of this document) in ALI1, will take place on a time window from 90 days before to 90 days after the outcome event date. This time window was defined to maximize the likelihood of identifying tests results that can allow evaluation for confirmation purposes of the identified cases (see details in the data validation plan).

- In **Denmark**, electronic assessment of whether the patients fulfil the laboratory criteria for ALI will be performed using available electronic laboratory data using the centralized Lab F database, without resorting to validation through GP questionnaires or chart abstraction.

- In **CPRD and HIRD**, because laboratory data are incomplete in these two data sources (e.g., no in-hospital laboratory data in HES; only outpatient laboratory data available for a ~30% subset of patients in HIRD):

  - If this electronic assessment does not confirm a case, the case will not be classified as a non-case; instead, the potential case will undergo additional validation procedures. Additional laboratory data will be obtained during validation, i.e., through GP questionnaires in CPRD and through chart abstraction in HIRD.

        - In CPRD, for cases identified in the outpatient setting (primary care and referrals), no additional laboratory data will be obtained during validation through GP questionnaires or chart abstraction, because no additional data is expected to be provided by the GP or abstracted (other than the already available outpatient electronic laboratory test results).

    - If with the available laboratory results from the outpatient setting the potential case can be confirmed, there will be no need to obtain medical records or completed GP questionnaires.

    - In all data sources, the baseline period will be set at 12 months for GP questionnaires and record abstraction.

The following specific laboratory data are available in each database and will be used for validation:

- **In Denmark**, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and bilirubin values will be defined as recorded in population-based laboratory databases, which cover tests made either in primary care (ordered by GPs or privately practising specialist physicians) or during hospital care (ED visits, hospital outpatient clinic visits, or inpatient admissions).
- **In HIRD**, outpatient laboratory data are available for a subset of patients. However, given that no in-hospital laboratory data are available in claims, no cases will be ruled out based on laboratory criteria in identifying potential cases of ALI, ALT, AST, ALP, and bilirubin values will be defined using Logical Observations Identifiers Names and Codes (LOINC).
- **In CPRD GOLD**, ALT, AST, ALP, and bilirubin values will be defined as recorded in data2 of entity types 155, 156, 153, and 158, respectively, and will be used to validate cases of ALI1 and ALI2.
- **In CPRD Aurum**, ALT, AST, ALP, and bilirubin values will be defined as recorded in the variable *value* with corresponding units (*numunitid*) for specific SNOMED/EMIS codes referring to these tests; values will be used to validate cases of ALI.
- In **CPRD GOLD** and **Aurum**, ULN can be missing, and these vary by laboratory. Based on patient profile review we have seen that patients usually have many ALT AST measurements all with the same ULN, form the same lab, and occasionally one ULN is missing. In this scenario, we will impute the ULN as the mode of the ULN within the same patient lab results.
  After inspection of data values, very few laboratory results were extreme outliers. Because exclusions are based on two measures and outcomes are identified based on diagnosis, no data cleaning will be performed in CPRD. In addition, if outliers within <p0,01 or >p.99.9% were set to missing, there would not be patients with ALT or AST measures x 20 ULN and those are described in Table 13.

Note that for "(B) ALT or AST ≥ 3 × ULN AND total bilirubin ≥ 2 × ULN", ALT/AST and bilirubin should be measured on the same calendar day.

Validation criteria are specified in the validation plan.

In all data sources for the assessment of this primary outcome, predisposing exclusion criteria will be applied to the cohorts as described in Section 7.3.4. Follow-up will be censored at the occurrence of any exclusion criteria for this cohort, e.g., any predisposing condition such as an ICD-10 diagnosis of chronic liver disease among patients in the ALI1 cohort, as described in Section 7.4. Use of laboratory values during follow-up for validation purposes is described in the validation plan.

8.2.1.1.2 ALI sensitivity analysis (CPRD and HIRD)

An ALI sensitivity analysis will be performed in the CPRD and HIRD, the data sources where primary care data are available. Cases will include previously identified cases in the primary outcome as well as patients who only received an ALI diagnosis in the primary care setting. The exclusion and censoring conditions will be the same as those for the primary outcome.

- **In CPRD**, includes all cases identified in the primary outcome, and also includes outpatient cases of ALI identified in the primary care or outpatient setting, i.e. codes for ALI without an association with a hospitalisation, ED visit, or specialist visit. The exclusion and censoring conditions will be the same as those for the primary outcome. The date of the event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.
- **In the HIRD**, cases will include all previously identified cases in the primary outcome as well as those patients who had one or more claims for a primary care provider (PCP) office visit with an ALI-related ICD-9-CM/ICD-10-CM diagnosis code and no claims for ALI-related hospitalisations, ED visits, or gastroenterology specialist visits (no hepatology visits recording available). PCPs includes the provider specialties of family practice, internal medicine, general practice, physician assistant, osteopathic medicine, paediatric medicine, and geriatric medicine. The date of the event will be the admission date in the hospital, the date of the ED visit, the date of the specialist visit, or the date of the diagnosis in the PCP office visit. The exclusion and censoring conditions will be the same as those for the primary outcome.

8.2.1.2 Acute kidney injury: hospitalisation, ED visit, or specialist visit for AKI—in all data sources

8.2.1.2.1 Case definition and identification

Potential cases of hospitalisation, ED visit, or specialist visit for AKI will be identified by the following procedure:

- **In Denmark,** potential cases will be identified through primary or secondary hospital discharge ICD-10 codes, ED visit ICD-10 codes, or hospital outpatient clinic ICD-10 codes suggestive of AKI.
- **In HIRD,** potential cases will be identified as patients who had one or more claims for a hospital visit, ED visit, or specialist office visit with an ICD-9-CM/ICD-10-CM diagnosis

code suggestive of AKI. Specialist visits will be those office visits with a provider whose speciality is nephrology.

- **In CPRD**:
  - **In HES**, potential cases will be identified through primary or secondary hospital discharge ICD-10 codes suggestive of AKI. No ED visit information is available in HES.
  - **In GOLD**, potential cases will be identified using primary care codes (Read codes) suggestive of AKI and associated with codes for hospitalisation, ED visit, or specialist visit. Hospitalisations, ED visits, and referrals will be defined as follows:
    - A referral to a nephrologist specialist will be identified through Read codes for referral to these specialists, through NHS speciality recorded as nephrology (*nhsspec* = 33).
    - A hospitalisation or ED visit: as defined in Section 8.2.1.1.1 for ALI.
  - **In Aurum**, potential cases will be identified using primary care codes (SNOMED/EMIS codes) suggestive of AKI associated with codes for hospitalisation, ED visit, or specialist visit within 30 days of the code suggestive of AKI. Hospitalisations, ED visits, and referrals will be defined as follows:
    - A referral to an internal medicine specialist or nephrologist will be identified through codes for referral to these specialists in the *medcodeid* variable. Identification through the variables *refsourceorgid*, *reftargetorgid* in referral data set was not available at the time of the data extraction (Sept 2019).
    - A hospitalisation or ED visit will be identified through any of the following:
      - SNOMED/EMIS codes for hospitalisation/ED visit or referral for hospitalisation/ED visit in the variable *medcodeid* in the observation data set or in the variable *consmedcodeid* from the consultation data set
      - Referral for hospitalisation recorded in the variable *conssourceid* in the consultation data set. Identification through the variables *refsourceorgid* or *reftargetorgid* in the referral data set was not available at the time of the data extraction (Sept 2019).
      - Consultation type indicating that the patient was hospitalised or seen in the ED using the *conssourceid* variable. Identification of consultation type through the *cprdconstype* variable was not available at the time of the data extraction (Sept 2019).

    The date of the event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.

The list of all codes to be used in each database is available in Annex 2.

In all data sources, electronic laboratory data will not be used to identify potential cases of AKI. However, electronic assessment of whether the patients fulfil the laboratory criteria for AKI will be performed before validation, whenever possible, using available electronic laboratory data without resorting to validation through GP questionnaires or chart abstraction.

- In Denmark, electronic assessment of whether the patients fulfil the laboratory criteria for AKI will be performed using available electronic laboratory data using the centralized Lab F database, without resorting to validation through GP questionnaires or chart abstraction.

- In CPRD and HIRD, because laboratory data are incomplete in these two data sources (e.g., no in-hospital laboratory data in HES, and laboratory data for only a subset of outpatients in HIRD), if this electronic assessment does not confirm an identified case, the identified case will not be classified as a non-case; instead, the identified case will undergo additional validation procedures. Additional laboratory data will be obtained during validation, i.e., through GP questionnaires in CPRD and through chart abstraction in HIRD. On the other hand, if with the available laboratory results from the outpatient setting the potential case can be confirmed, even if there was a hospitalization/ED visit/specialist office visit, there will be no need to obtain medical records or completed GP questionnaires. Similarly, for outpatient/primary care potential cases, outpatient laboratory results may be enough. In CPRD and HIRD the baseline period will be set at 12 months for GP questionnaires and record abstraction.

Exclusion conditions will be assessed via programming algorithms in all databases as described in Section 7.3.4. Laboratory values will be used to identify and exclude patients with CKD prior to the index date and for the assessment of AKI. In addition, laboratory values will be used to identify patients with undiagnosed CKD developed during follow-up and that will constitute a censoring criterion.

- **In Denmark**, serum creatinine values will be defined as recorded in population-based laboratory databases, covering tests made either in primary care (ordered by GPs or privately practising specialist physicians) or during hospital care (ED visits, hospital outpatient clinic visits, or inpatient admissions). eGFR will be used as exclusion, validation (to confirm absence of CKD) and censoring criteria, whereas creatinine increase from baseline serum creatinine level is used only for validation purposes – not censoring, except when creatinine is used to estimate eGFR, if applicable.
- **In HIRD**, serum creatinine values will be defined as recorded in outpatient laboratory data for the subset of patients for whom the data are available and defined using LOINC. Because race and ethnicity are not available in HIRD, serum creatinine will not be used to estimate eGFR. eGFR will be used as exclusion, validation(to confirm absence of CKD) and censoring criteria, (except in the HIRD), whereas creatinine increase from baseline serum creatinine level is used only for validation purposes – not censoring.
- **In CPRD GOLD**, serum creatinine values will be defined as recorded in data2 of entity type 165 and used to validate cases of AKI.
- **In CPRD Aurum**, serum creatinine values will be defined as recorded in the variable.
    - No data cleaning will be performed in CPRD After inspection of data values, very few laboratory results were extreme outliers. Because exclusions are based on two measures of eGFR or SrCreat, and CKD outcome sensitivity analysis is based on two measures, there is less risk that the diagnosis is identified or confirmed through one outlier laboratory result.

Validation criteria are specified in the validation plan.

8.2.1.2.2 AKI sensitivity analysis (CPRD and HIRD)

- **In CPRD**, includes all cases identified in the primary outcome, and also includes outpatient cases of AKI identified in the primary care or outpatient setting, i.e. codes for AKI without an association with a hospitalisation, ED visit, or specialist visit. The exclusion and censoring conditions will be the same as those for the primary outcome. The date of the event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.
- **In the HIRD,** cases will include all previously identified cases as well as those patients who had one or more claims for a PCP office visit with an AKI-related ICD-9-CM/ICD-10-CM diagnosis code and no claims for AKI-related hospitalisations, ED visits, or specialist (nephrology) visits. PCPs includes the provider specialties of family practice, internal medicine, general practice, physician assistant, osteopathic medicine, paediatric medicine, and geriatric medicine. The exclusion and censoring conditions will be the same as those for the primary outcome.

8.2.1.3 Diabetic ketoacidosis: hospitalisation or ED visit due to DKA—in all data sources

In Denmark and HIRD and linkable practices from CPRD, cases of DKA will be identified through hospital and ED discharge codes suggestive of DKA (see ICD-10 codes in Annex 2).

The outcome “hospitalisation due to DKA” comprises hospitalisation, ED visit, or an outpatient record of hospitalisation for DKA. Potential cases of DKA will be identified by diagnosis codes suggestive of DKA.

- **In Denmark,** potential cases will be identified through primary or secondary hospital discharge ICD-10 codes or ED visit ICD-10 codes (inpatient, but not outpatient hospital visits) suggestive of DKA.
- **In HIRD**, potential cases will be identified through hospital claims or ED visit claims with an ICD-9-CM/ICD-10-CM code suggestive of DKA.
- **In CPRD:**
  - **In HES**, potential cases will be identified through hospital discharge ICD-10 codes in the primary or secondary discharge position.
  - **In GOLD/Aurum,** potential cases will be identified using primary care codes, together with codes for hospitalisation or ED visit within 30 days. A hospitalisation or ED visit will be identified as defined for ALI1 or ALI2 (Section 8.2.1.1).
  - The date of the event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.

In all data sources, electronic laboratory data will not be used to identify potential cases of DKA. However, whenever possible, to confirm some of the cases that would not undergo

further validation, electronic assessment of whether the patients fulfil the laboratory criteria for DKA will be performed before validation, using available electronic laboratory data without resorting to validation through GP questionnaires or chart abstraction. The following variables will be defined in each data source:

- **In Denmark,** laboratory values will be defined as recorded in population-based laboratory databases, covering tests made either in primary care (ordered by GPs or privately practising specialist physicians) or during hospital care (ED visits, hospital outpatient clinic visits, or inpatient admissions). Validation criteria are specified in the validation plan. However, dipstick results cannot be validly assessed in Denmark.
- **In HIRD,** the outpatient laboratory database does not include the relevant laboratory values for identifying DKA. Laboratory criteria will be assessed through chart abstraction.
- **In CPRD GOLD** (for additional details see DVP)**:**
    - Ketonaemia > 1.5 mmol/L: identified through recorded data values associated with Read codes for blood ketone levels (see Annex 2) and categorised as present if data2 > 1.5 mmol/L (after transforming to correct units) or if there is a Read code for ketonaemia.
    - Ketonuria: identified through recorded data values associated with entity type 432 "Urine dipstick for ketones" and/or associated with Read codes for urine ketone levels (see Annex 2) and categorised as positive if *value* is, e.g., +, ++, +++ for positive, present, or high, respectively, or if the Read code indicates ketonuria.
    - Acidaemia: identified through either Read codes for acidaemia, or a serum pH < 7.3. In the absence of serum PH values or diagnosis of acidaemia then: a serum anion gap > 10 mEq/L, or venous HCO_3 below 15 mmol/L:
        - Read codes for acidaemia are available in Annex 2.
        - pH (from a blood sample): identified through recorded data values associated with a Read code for pH sample and categorised as acidaemia if *data2* < 7.3 or from a Read code with column ph_acid=1
        - Anion gap: identified through recorded data values associated with entity type 447 "Anion gap" and/or associated with Read codes for anion gap levels and categorised as acidaemia if *data2* > 10 mEq/L (after transforming to correct units)
        - Venous HCO_3 (bicarbonate): identified through recorded data values associated with entity type 391 "Serum bicarbonate" and and categorised as acidaemia if *data2* < 15 mmol/L (after transforming to correct units) or Read code in tab bicarbonate with column bicarbonate_abnromal = 1
- **In CPRD Aurum** (for additional details see data validation plan)**:**
    - Ketonaemia: identified through recorded data values associated with SNOMED/EMIS codes for blood ketone levels or through SNOMED/EMIS codes identifying ketonaemia (see Annex 2) and categorised as present if *value* > 1.5 mmol/L (after transforming to correct units) or if there is a SNOMED/EMIS code for ketonaemia.
    - Ketonuria: identified through SNOMED/EMIS codes for urine ketone levels or ketonuria (see Annex 2) and categorised as positive if *value* is, e.g., +, ++, +++ for positive, present, or high, respectively, or if the code indicates ketonuria.

- Acidaemia: identified through either SNOMED/EMIS codes for acidaemia, a serum pH < 7.3, a serum anion gap > 10 mEq/L, or venous HCO_3 (bicarbonate) below 15 mmol/L:
  - SNOMED/EMIS codes for acidaemia are available in Annex 2
  - pH (from a blood sample): identified through recorded data values associated with a SNOMED/EMIS code for pH sample and categorised as acidaemia if *value* < 7.3
  - Anion gap: identified through recorded data values associated with SNOMED/EMIS codes for anion gap levels and categorised as acidaemia if *value* > 10 mEq/L (after transforming to correct units)
  - Venous HCO_3 (bicarbonate): identified through recorded data values associated with SNOMED/EMIS codes for serum HCO_3 and categorised as acidaemia if *value* < 15 mmol/L (after transforming to correct units)

The list of all codes to be used in each database is available in Annex 2.

Validation criteria are specified in the validation plan.

8.2.1.4 Severe complications of urinary tract infections (inpatient and outpatient)—only in CPRD

Potential cases of severe complications of UTIs will be identified by any of the following conditions:

- **In CPRD GOLD,** irrespective of records for hospitalisation, ED visit, or referrals:
  - GP record/outpatient visit for pyelonephritis or urosepsis: any Read code for pyelonephritis or urosepsis (only one Read code, K190600, is available for urosepsis)
  - In the outpatient setting: a combination of diagnosis code for UTI and a diagnosis code for sepsis within 1 week
- **In CPRD Aurum,** irrespective of records for hospitalisation, ED visit, or referrals:
  - GP record/outpatient visit for pyelonephritis or urosepsis: any SNOMED/EMIS code for pyelonephritis or urosepsis (only two SNOMED codes are available for urosepsis)
  - In the outpatient setting: a combination of diagnosis code for UTI and a diagnosis code for sepsis within 1 week
- ~~**In HES**:~~
- **In HES**:(no hospital secondary discharge diagnosis will be used as opposed to the previous outcomes)
  - A primary discharge ICD-10 code for pyelonephritis (there are no ICD-10 codes for urosepsis)
  - A primary discharge ICD-10 code for sepsis and a secondary discharge code for UTI
  - A primary discharge ICD-10 code for sepsis and a Read or SNOMED/EMIS code for UTI within 1 week before or after the hospitalisation for sepsis

The date of the outcome event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.

Lists of Read, SNOMED/EMIS, and ICD-10 codes are presented in Annex 2.

Validation criteria are specified in the validation plan.

8.2.1.5 Genital infections (inpatient and outpatient)—only in CPRD

The outcome GI comprises vulvovaginitis in females and balanitis in males.

Complications and severe consequences of GI will be described among confirmed severe GI cases. (See identification of these complications in the Validation Plan).

Vulvovaginitis

Among all females, potential cases of vulvovaginitis will be identified through the following codes in GOLD /Aurum:

- A code **specific vulvovaginitis**:
  - bacterial vaginosis (Code identified in Outcomes.xls, Genital infections tab, column Type of GI = "specific vulvovaginitis Gardnerella") or for specific microbiology results indicating *Gardnerella* (Type of GI= "culture positive for Gardnerella confirmatory"),
  - vulvovaginal candidiasis (Type of GI= "specific vulvovaginitis candida"), or fungus (Type of GI= "culture positive for candida confirmatory" or "yeast in gram confirmatory")
- A code for **non-specific vulvovaginitis** (Type of GI= "nonspecific vulvovaginitis"), or for non-specific positive microbiology results (Type of GI= "nonspecific positive micro")

Among female patients linkable to HES, the information in HES will be used to identify additional potential cases of vulvovaginitis through the following codes:

- A primary or secondary discharge ICD-10 code for:
  - vulvovaginal candidiasis (which is a specific vulvovaginitis) (Type of GI= "specific vulvovaginitis candida) or
  - non-specific vulvovaginitis (Type of GI= "nonspecific vulvovaginitis")

For validation purposes, these following variables will need to be generated separately although using information from both primary care and HES, given different validation criteria.

- Specific vulvovaginitis.
  - vulvovaginal candidiasis
  - bacterial vaginosis
- Nonspecific vulvovaginitis

Identification on whether the case has been identified in GOLD/Aurum or HES will be needed to generate the secondary endpoint severe genital infections.

All potential cases of vulvovaginitis will require absence of a diagnosis code for a sexually transmitted infection (STI) within 30 days before or after the genital infection date.

The date of the event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.

Lists of Read, SNOMED/EMIS, and ICD-10 codes are presented in Annex 2.

Validation criteria are specified in the validation plan.

Balanitis

Among all males, potential cases of balanitis will be identified through a code in GOLD/Aurum or through a primary or secondary discharge code in HES for balanitis or balanoposthitis, including:

- specific *Candida* balanitis (Type of GI= "specific balanitis candida"),
- aerobic balanitis (NO codes present), and
- non-specific balanitis (Type of GI= "nonspecific balanitis").

For validation purposes, these two variables will need to be generated separately given different validation criteria. Identification on whether the case has been identified in GOLD/Aurum or HES will be needed to generate the secondary endpoint severe genital infections.

All potential cases of balanitis will require the absence of a specific diagnosis code for an STI within 30 days before or after the genital infection date.

The date of the event will be the admission date in HES or the date of the diagnosis from GOLD/Aurum.

Lists of Read, SNOMED/EMIS, and ICD-10 codes are presented in Annex 2.

Validation criteria are specified in the validation plan.

8.2.2 Secondary outcomes

8.2.2.1 Acute liver injury: hospitalisation, ED visit, or specialist visit for ALI2 in patients with or without predisposing conditions—in all data sources

ALI is both a primary outcome and a secondary outcome. For the assessment of ALI2 as a secondary outcome, different exclusion criteria and censoring will be applied as described in Section 7.3.4.1 and Section 7.4. Validation criteria are specified in the validation plan.

8.2.2.1.1 Description of cases of elevated liver enzymes

In addition to the evaluation of the ALI1 and ALI2 endpoints, this study will describe if patients in the ALI2 cohort, (which includes ALI1 cohort *per se*, and results are also provided for those patients without predisposing conditions- see Table 13) with elevated liver enzymes

irrespective of whether the patients had a diagnosis or symptom code suggestive of ALI during current use and recent use (presented together and separately) as defined in Section 8.1.2. Cases of elevated liver enzymes will be classified in the highest of the following categories:

- ALT and/or AST ≥ 3 × ULN and < 5 x ULN
- ALT and/or AST ≥ 5 × ULN and < 10 x ULN
- ALT and/or AST ≥ 10 × ULN and < 20 x ULN
- ALT and/or AST ≥ 20 × ULN

Elevated enzymes will be reported during either current, recent, and current+recent exposure periods. Three variables may be created as follows: ELE_current, ELE_recent, ELE_current+recent, where Elevated liver enzymes (ELE) will be categorised as [1 = (3-5); 2 = (5-10); 3 = (10-20); 4= ">=20" x ULN AST/ALT]. A patient may have elevated liver enzyme during both current and recent use. In this case, the period during which the elevation is the highest would be the one considered. For example:

- If a patient has ELE_current = 2, & ELE_recent = 1, then consider ELE_current = 2, & ELE_recent = ., ELE_current+recent=2
- If a patient has ELE_current = 2, & ELE_recent = 3, then consider ELE_current = ., & ELE_recent = 3, ELE_current+recent=3
- If a patient has ELE_current = 2, & ELE_recent = 2, then consider ELE_current = 2, & ELE_recent = 2, ELE_current+recent=2

The number and proportion of patients with elevated liver enzymes identified in each data source will be described overall and stratified by age and sex, by treatment group, by ALI diagnosis (as defined above in this section 8.2.2.1), and by prior history of predisposing conditions (see analyses described in Section 11.6.1). Identification of these tests has been described in Section 8.2.1.1.

ALT will be used as first option, whenever available. If both ALT and AST are available on the same date, only ALT will be used. If ALT is missing, then AST will used, except if the patient has a record for a muscle disease within 30 days before or after the AST laboratory test. The list and all codes of muscle diseases to be used in each database is available in the code list for inclusion and exclusion criteria (Annex 2).

The validation plan will also include another analysis describing patients that *had a diagnosis or symptom code* suggestive of ALI2 but that did not fulfil the validation criteria (see further details in the validation plan).

8.2.2.2 Chronic kidney disease (inpatient and outpatient)—only in CPRD

Potential cases of CKD will be identified by CKD diagnosis codes. Lists of Read, SNOMED/EMIS, ICD-10, and OPCS4 codes are presented in Annex 2.

- **In CPRD GOLD and Aurum**, potential cases of CKD will be identified through Read and SNOMED/EMIS codes for CKD.
- **In HES**, additional potential cases of CKD will be identified through primary or secondary discharge ICD-10 codes for CKD or OPCS4 procedure codes for dialysis and transplant.
- All identified potential cases will be differentiated from AKI and confirmed as CKD by requesting confirmation of two abnormal eGFR test results, one within at least 3 months (90 days) before or after the CKD record, and another one between 91 and 365 days before the confirmatory test (See further details in the DVP): An estimated glomerular filtration rate (eGFR) below 60 mL/min/1.73 m^2 will be identified through either recorded data values associated with specific codes or estimated eGFR using serum creatinine as follows:
    - Recorded data values associated with eGFR Read or SNOMED/EMIS codes:
        - Estimated using the CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) equation [R13-4387, R12-1392, R15-5270].
        - Estimated GFR (mL/min/1.73 m^2) = 141 × min(SCr/k, 1)α × max(SCr/k, 1) - 1.209 × 0.993*Age [× 1.018 if female] [× 1.159 if black], where SCr is serum creatinine (in mg/dL) and is the recorded value of serum creatinine, k is 0.7 for females and 0.9 for males, α is −0.329 for females and −0.411 for males, min is the minimum of SCr/k or 1, and max is the maximum of SCr/k or 1.
        - As race is not available for all patients in CPRD, the race coefficient will not be applied if race is missing; this may result in underestimation of eGFR in black patients.

Lists of Read, SNOMED/EMIS, ICD-10, and OPCS4 codes are presented in Annex 2.

Additional details on the validation criteria are specified in the validation plan.

All potential cases of CKD identified will be classified as definite cases or non-cases based on laboratory results; the 2009 CKD-EPI formula will also be used to calculate eGFR during the time windows of interest (see validation plan). Further details on case validation will be provided in the validation plan document.

A sensitivity analysis will be performed by identifying patients with CKD based on eGFR below 60 mL/min/1.73 m^2, through either recorded data values associated with Read or SNOMED/EMIS codes or estimated eGFR using serum creatinine. In this analysis, all eGFR data values recorded through Read or SNOMED/EMIS codes during the whole follow-up time and all serum creatinine laboratory results will be used to define eGFR; a patient will be considered to have CKD if there are at least two eGFR values below 60 mL/min/1.73 m^2 separated by 90 days or more and less than 365 days. The date of the second reduced eGFR will be the outcome event date. (See section 3.3.8 and 4.5.6 in the Validation Plan for additional details on validation criteria).

8.2.2.3 Severe genital infections—only in CPRD

Among the cases of GI identified in the primary outcome, those that resulted in hospitalisation or ED visit or that required systemic treatment with specific antifungals or antibiotics (as opposed to topical or vaginal treatment) will be classified as severe GI in males and females separately. Hospitalisations and ED visits recorded in CPRD will be defined as for the primary DKA outcome.

Vulvovaginitis—severe genital infection in females

Among all female patients, for periods of no HES coverage for patients whose records are linkable to HES or for patients whose records are not linkable to HES, potential cases of vulvovaginitis will be identified using the information from CPRD GOLD/Aurum as follows Cases identified in the primary outcome together with a code for hospitalisation or a prescription for antibiotics or antifungals as follows:

- GP record for hospitalisation within 30 days before or after a Read or SNOMED/EMIS code specific for bacterial vaginosis or *Candida* vulvovaginitis as defined in the primary endpoint,
- GP record for hospitalisation within 30 days before or after a Read or SNOMED/EMIS code for non-specific diagnosis of vulvovaginitis as defined in the primary endpoint
- Read or SNOMED/EMIS code specific for bacterial vaginosis or *Candida* vulvovaginitis as defined in the primary endpoint fungus and a prescription/dispensing for systemic antifungals or antibiotics* within 30 days before or after this Read or SNOMED/EMIS code, or
- Read or SNOMED/EMIS code for a non-specific diagnosis of vulvovaginitis as defined in the primary endpoint and a prescription/dispensing for systemic antifungals or antibiotics* within 30 days before or after this Read or SNOMED/EMIS code

In summary: any event identified in the primary outcome with a hospitalisation within 30 days or treated with antifungals or antibiotics within 30 days. *Systemic antibiotics and antifungals are identified in the Outcome.xls, tab GI_meds, column [(antifungals=1 or column antibiotics= 1), and column systemic= 1]

Among patients linkable to HES, all potential cases of vulvovaginitis identified in the primary outcome using HES primary or secondary discharge diagnosis (see Section 8.2) will be considered potential cases of severe GI in females.

All potential cases of severe GIs will require the absence of a specific diagnosis code for an STI within 30 days before or after the genital infection date.

Lists of Read, SNOMED/EMIS, and ICD-10 codes are presented in Annex 2.

Validation criteria are specified in the validation plan.

Balanitis—severe genital infection in males

Among all male patients, for periods of no HES coverage for patients whose records are linkable to HES or patients whose records are not linkable to HES, potential cases of balanitis will be identified using information from CPRD GOLD/Aurum with a GP record for hospitalisation or a prescription/dispensing for systemic antifungals or antibiotics* within 30 days before or after a Read or SNOMED/EMIS code for balanitis (as defined in the primary outcome).Among patients linkable to HES, all potential cases of balanitis identified in the primary outcome using HES (see Section 8.2) will be considered potential cases of severe GI in males. *Systemic antibiotics and antifungals are identified in the Outcome.xls, tab GI_meds, column [(antifungals=1 or column antibiotics= 1), and column systemic= 1].

All potential cases of severe GIs will require the absence of a specific diagnosis code for an STI within 30 days before or after the genital infection date.

Lists of Read, SNOMED/EMIS, and ICD-10 codes are presented in Annex 2.

Validation criteria are specified in the validation plan.

8.3 COVARIATES

Other variables will be used as (1) descriptors of the study population, (2) exclusion criteria and censoring criteria, (3) potential confounding or effect modifier variables, or (4) stratification variables. Variables will be defined based on all prior information history unless otherwise specified. See lists of other variables in Table 2 through Table 4.

Table 2. Variables that will be used as cohort exclusion/censoring indicator variables

Condition	ALI1[a]	ALI2	AKI	DKA	CKD	UTI	GI
Type 1 diabetes	√	√	√	√	√	√	√
Co-prescription/dispensing of SGLT2 and DPP-4 at index date[b]	√	√	√	√	√	√	√
ALI	√	√ 6m	—	—	—	—	—
Acute infectious hepatitis	√	√ 6m	—	—	—	—	—
Acute cholelithiasis and cholecystitis	√	√ 6m	—	—	—	—	—
Intra- or extrahepatic biliary obstruction	√	√ 6m	—	—	—	—	—
Acute pancreatic disease	√	√ 6m	—	—	—	—	—
Decompensated congestive heart failure (i.e., hospitalisation)	√	√ 6m	—	—	—	—	—
Pregnancy at baseline	√	√	—	—	—	—	—
Chronic liver disease	√	—	—	—	—	—	—
Chronic alcoholism	√	—	—	—	—	—	—
Chronic infectious hepatitis	√	—	—	—	—	—	—

Condition	ALI1[a]	ALI2	AKI	DKA	CKD	UTI	GI
Chronic disease involving the liver or causing hyperbilirubinaemia	√	—	—	—	—	—	—
Chronic cholelithiasis and cholecystitis	√	—	—	—	—	—	—
Intra- or extrahepatic biliary obstruction	√	—	—	—	—	—	—
Chronic pancreatic disease	√	—	—	—	—	—	—
Primary or secondary hepatic, biliary, or pancreatic cancer	√	—	—	—	—	—	—
Congestive heart failure	√	—	—	—	—	—	—
AKI	—	—	√ 6m	—	—	—	—
CKD	—	—	√	—	√	—	—
Chronic pyelonephritis	—	—	—	—	—	√ 6m	—
Acute pyelonephritis	—	—	—	—	—	√ 6m	—

AKI = acute kidney injury; ALI = acute liver injury; CKD = chronic kidney disease; DKA = diabetic ketoacidosis; DPP-4 = dipeptidyl peptidase-4; GI = genital infection; SGLT2 = sodium-glucose cotransporter 2; UTI = urinary tract infection.

Key: "√" means that the variable will be used for that outcome and assessed using information any time before or on the index date- "—" means that it will not be used. "6m" means that the variable will be used using the information during the 6 months before and including the index date.

[a] The primary ALI outcome, which has different exclusion criteria from the secondary outcome.

[b] The index date is the date of the first prescription/dispensing of the study medication of interest within the study period.

The list of all codes to be used in each database is available in Annex 3.

Table 3. Variables that will be used as potential confounders or effects modifiers

	ALI1 and ALI2	AKI and CKD	DKA	UTI (CPRD only)	GI (CPRD only)
Lifestyle variables					
Age	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Alcohol consumption	CPRD, DK n.a.	CPRD, DK n.a.	CPRD, DK n.a.	CPRD	CPRD
Body mass index > 30, obesity diagnosis, or obesity surgery	CPRD, DK	CPRD, DK	CPRD, DK	CPRD	CPRD
Calendar year of index date[a]	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Duration of lookback period	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
History of alcohol abuse	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Sex	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Smoking history	CPRD, DK n.a.	CPRD, DK n.a.	CPRD, DK n.a.	CPRD	CPRD

	ALI1 and ALI2	**AKI and CKD**	**DKA**	**UTI** (CPRD only)	**GI** (CPRD only)
Socioeconomic status	CPRD, DK	CPRD, DK	CPRD, DK	CPRD	CPRD
Indicators of T2D severity					
Glycated haemoglobin in the 12 months before or on the index date	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Amputations	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Cerebrovascular disease	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Coronary heart disease	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD,
Neurologic complications	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD,	CPRD,
Peripheral circulatory complications	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Renal complications	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Ophthalmic complications	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Other diabetes complications	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Time since first diagnosis of T2D	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Medical history					
Chronic disease score: Charlson Comorbidity Index	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Heart failure	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
High blood pressure/hypertensive disease	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Asthma	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Autoimmune disease	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Being hospitalised, requiring intensive care (ICU) in the last 12 months prior to index date	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Chronic obstructive pulmonary disease, emphysema, respiratory insufficiency	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Dementia	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Autoimmune and connective tissue diseases	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Rheumatoid arthritis	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Polymyalgia rheumatica	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD

	ALI1 and ALI2	**AKI and CKD**	**DKA**	**UTI** (CPRD only)	**GI** (CPRD only)
Other autoimmune and connective tissue diseases	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Immunosuppressive diseases such as HIV/AIDS, transplants, malignant neoplasms of lymphoid, haematopoietic and related tissue, and other immunodeficiency disorders	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Kidney and genitourinary stones	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Length of hospitalisation in the ICU in the last 12 months prior to index date	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Liver and biliary diseases	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Osteoarthritis and Osteoarthrosis	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Gastro-oesophageal reflux, gastroesophagitis, and peptic ulcer disease	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Chron's disease	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Ulcerative colitis	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Colon polyps	CPRD, DK, HIRD	—	—	—	
Pancreatic diseases	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Pancreatitis	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Hyperlipidaemia	CPRD, DK, HIRD	—	—	—	
~~Other~~ Malignancies	CPRD, DK, HIRD	—	—	—	—
Chronic renal disease or renal dialysis	—	CPRD, DK, HIRD	—	—	—
Prior history of AKI > 6 months before the index date[a]	—	CPRD, DK, HIRD	—	—	—
Urinary infections	—	CPRD, DK, HIRD	—	CPRD	CPRD
Pregnancy	—	—	—	CPRD	CPRD
Prior history of UTI leading to hospitalisation > 6 months before the index date[a*]	—	—	—	CPRD$^{>6m}$	—
Prior history of acute pyelonephritis > 6 months before the index date[a]	—	—	—	CPRD$^{>6m}$	—
Prior history of genital infections	—	—	—	—	CPRD

	ALI1 and ALI2	**AKI and CKD**	**DKA**	**UTI** (CPRD only)	**GI** (CPRD only)
Recent myocardial ischaemia/infarction (6 months prior to index date)	—	—	CPRD^{6m}, DK^{6m}, HIRD^{6m}	—	—
Prior history of DKA	—	—	CPRD, DK, HIRD	—	—
Psychological stress	—	—	CPRD, DK, HIRD	—	—
Thyroid problems (e.g., thyroid storm and thyrotoxicosis)	—	—	CPRD, DK, HIRD	—	—
Medications					
Omeprazole	CPRD^{HTX}, DK, HIRD	—	—	—	—
Drugs used in diabetes	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Insulins	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Other oral antidiabetic drugs	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Acarbose	CPRD^{HTX}, DK, HIRD	—	—	—	—
Anabolic steroids	CPRD^{HTX}, DK, HIRD	—	—	—	—
Antithrombotic agents	—	CPRD, DK, HIRD	—	CPRD	CPRD
Platelet aggregation inhibitors	—	CPRD, DK, HIRD	—	CPRD	CPRD
Clopidogrel	CPRD^{HTX}, DK, HIRD	—	—	—	—
Cardiovascular system drugs	CPRD, DK, HIRD	—	—	—	—
Digoxin	—	CPRD, DK, HIRD	—	CPRD	CPRD
Antiarrhythmic, class I and III	—	CPRD, DK, HIRD	—	CPRD	CPRD
Amiodarone	CPRD^{HTX}, DK, HIRD	—	—	—	—
Nitrates	—	CPRD, DK, HIRD	—	CPRD	CPRD
Antihypertensives	—	CPRD, DK, HIRD	—	CPRD	CPRD
Other antihypertensive	—	CPRD, DK, HIRD	—	CPRD	CPRD
Diuretics	—	CPRD, DK, HIRD	—	CPRD	CPRD
Thiazides	—	—	CPRD, DK, HIRD	—	—

	ALI1 and ALI2	**AKI and CKD**	**DKA**	**UTI** (CPRD only)	**GI** (CPRD only)
Beta blocking agents	—	CPRD, DK, HIRD	—	CPRD	CPRD
Calcium channel blockers	—	CPRD, DK, HIRD	—	CPRD	CPRD
Verapamil	CPRD^{HTX}, DK, HIRD	—	—	—	—
Agents acting on the renin-angiotensin system	—	CPRD, DK, HIRD	—	CPRD	CPRD
Captopril	CPRD^{HTX}, DK, HIRD	—	—	—	—
Enalapril	CPRD^{HTX}, DK, HIRD	—	—	—	—
Lisinopril	CPRD^{HTX}, DK, HIRD	—	—	—	—
Irbesartan	CPRD^{HTX}, DK, HIRD	—	—	—	—
Losartan	CPRD^{HTX}, DK, HIRD	—	—	—	—
Lipid modifying agents	—	CPRD, DK, HIRD	—	CPRD	CPRD
Statins	CPRD, DK, HIRD	CPRD, DK, HIRD	—	CPRD	CPRD
Fibrates	CPRD, DK, HIRD	CPRD, DK, HIRD	—	CPRD	CPRD
Terbinafine	CPRD^{HTX}, DK, HIRD	—	—	—	—
Flutamide	CPRD^{HTX}, DK, HIRD	—	—	—	—
Estrogens	CPRD^{HTX}, DK, HIRD	—	—	—	—
Systemic glucocorticosteroids	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD
Antibacterial for systemic use	—	CPRD, DK, HIRD	—	CPRD	CPRD
Tetracycline	CPRD^{HTX}, DK, HIRD	—	—	—	—
Amoxicillin + clavulanic acid	CPRD^{HTX}, DK, HIRD	—	—	—	—
Sulfonamides	CPRD^{HTX} DK, HIRD	—	—	—	—
Trimethoprim-sulfamethoxazole	CPRD^{HTX}, DK, HIRD	—	—	—	—
Clindamycin	CPRD^{HTX}, DK, HIRD	—	—	—	—
Erythromycins	CPRD^{HTX}, DK, HIRD	—	—	—	—

	ALI1 and ALI2	**AKI and CKD**	**DKA**	**UTI** (CPRD only)	**GI** (CPRD only)
Trovafloxacin	CPRD^{HTX}, DK, HIRD	—	—	—	—
Nitrofurantoin	CPRD^{HTX}, DK, HIRD	—	—	—	—
Antimycotics for systemic use	—	CPRD, DK, HIRD	—	CPRD	CPRD
Ketoconazole	CPRD^{HTX}, DK, HIRD	—	—	—	—
Posaconazole	—	CPRD, DK, HIRD	—	—	—
Terbinafine (for systemic use)	CPRD^{HTX}, DK, HIRD	—	—	—	—
Drugs for treatment of tuberculosis	CPRD^{HTX}, DK, HIRD	CPRD, DK, HIRD	—	CPRD	CPRD
Isoniazid	CPRD^{HTX}, DK, HIRD	—	—	—	—
Pyrazinamide	CPRD^{HTX}, DK, HIRD	—	—	—	—
Rifampicin	CPRD^{HTX}, DK, HIRD	—	—	—	—
Direct acting antivirals	—	CPRD, DK, HIRD	—	CPRD	CPRD
HAART drugs	CPRD^{HTX}, DK, HIRD	—	—	—	—
Antineoplastic agents	—	CPRD, DK, HIRD	—	CPRD	CPRD
Immunosuppressant	CPRD^{HTX}, DK, HIRD	—	—	—	—
Azathioprine	CPRD^{HTX}, DK, HIRD	—	—	—	—
Methotrexate	CPRD^{HTX}, DK, HIRD	CPRD, DK, HIRD	—	CPRD	CPRD
Ciclosporine	CPRD, DK, HIRD	—	—	—	—
Other immunosuppressant excluding systemic tacrolimus	CPRD, DK, HIRD	—	—	—	—
Systemic tacrolimus	CPRD, DK, HIRD	—	—	—	—
Anti-inflammatory and antirheumatic products	CPRD^{HTX}, DK, HIRD	CPRD, DK, HIRD	—	CPRD	CPRD
NSAIDs	CPRD^{HTX}, DK, HIRD	CPRD, DK, HIRD	—	CPRD	CPRD
Other antirheumatic agents	CPRD, DK, HIRD	—	—	—	—
Baclofen	CPRD^{HTX}, DK^{HTX}, HIRD	—	—	—	—

	ALI1 and ALI2	**AKI and CKD**	**DKA**	**UTI** (CPRD only)	**GI** (CPRD only)
Allopurinol	CPRD^{HTX}, DK, HIRD	—	—	—	—
Zoledronic acid	—	CPRD, DK, HIRD	—	CPRD	CPRD
Acetaminophen (prescription/ dispensing)	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	CPRD, DK, HIRD	—	CPRD	CPRD
Antiepileptic	—	CPRD, DK, HIRD	—	CPRD	CPRD
Carbamazepine	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Phenobarbital	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Phenytoin	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Valproic acid	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Antipsychotics	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	
Chlorpromazine	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Phenothiazines	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Clozapine	—	—	CPRD, DK, HIRD	—	—
Olanzapine	—	—	CPRD, DK, HIRD	—	—
Lithium	—	—	CPRD, DK, HIRD	—	—
Antidepressants	—	—	—	—	—
Amitriptyline	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Bupropion	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—
Fluoxetine	CPRD^{HTX}, DK^{HTX}, HIRD^{HTX}	—	—	—	—

	ALI1 and ALI2	**AKI and CKD**	**DKA**	**UTI** (CPRD only)	**GI** (CPRD only)
Mirtazapine	$CPRD^{HTX}$, DK^{HTX}, $HIRD^{HTX}$	—	—	—	—
Paroxetine	$CPRD^{HTX}$, DK^{HTX}, $HIRD^{HTX}$	—	—	—	—
Risperidone	$CPRD^{HTX}$, DK^{HTX}, $HIRD^{HTX}$	—	—	—	—
Sertraline	$CPRD^{HTX}$, DK^{HTX}, $HIRD^{HTX}$	—	—	—	—
Trazodone	$CPRD^{HTX}$, DK^{HTX}, $HIRD^{HTX}$	—	—	—	—
Tricyclics	$CPRD^{HTX}$, DK^{HTX}, $HIRD^{HTX}$	—	—	—	—
Pentamidine	—	—	CPRD, DK, HIRD	—	—
Terbutaline	—	—	CPRD, DK, HIRD	—	—
Cyproheptadine	$CPRD^{HTX}$, DK^{HTX}, $HIRD^{HTX}$	—	—	—	—

AKI = acute kidney injury; ALI = acute liver injury; CKD = chronic kidney disease; CPRD = Clinical Practice Research Datalink; DK = Denmark; DKA = diabetic ketoacidosis; GI = genital infection; HAART = highly active antiretroviral therapy; HIRD = HealthCore Integrated Research Database; HIV = human immunodeficiency virus; HTX = hepatotoxic (see key); NSAID = non-steroidal anti-inflammatory drug; T2D = type 2 diabetes mellitus; UTI = urinary tract infection.

Key: "CPRD, DK, HIRD": the variable is available in the mentioned database and will be used for that outcome and assessed using information any time before or on the index date; "—" variable will not be used; "6m": use of the information during the 6 months before or on the index date for this variable; ">6m": use of the information more than 6 months after the index date for this variable; "HTX": medications to be combined in a single variable indicating use of hepatotoxic medications; "n.a.": a proxy of the variable is planned to be used, if available, but the variable itself is not available in the database.

[a] The index date is the date of the first prescription/dispensing of the study medication of interest within the study period.

Note: see MAP files, table shells, and dataset template for an updated list of variables and cohorts.

The list of all codes to be used in each database is available in Annex 3.

[*] Prior history of UTI leading to hospitalisation > 6 months before the index date will be defined as having a code for UTI on more than 6 months before the index date or having a code (Readcode or Snomed) for UTI and a code for hospitalitsation within 7 days.

Table 4. Variables that will be used as stratification variables

	ALI1 and ALI2	**AKI**	**DKA**	**CKD** (CPRD only)	**UTI** (CPRD only)	**GI** (CPRD only)
Calendar year	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD	CPRD
Add-on vs. switching status[a]	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD	CPRD
Mono, dual, and triple therapy[b]	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD	CPRD
Insulin	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD	CPRD
Age	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD	CPRD
Sex	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD	CPRD
Diabetes control	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD, DK, HIRD	CPRD	CPRD	CPRD

AKI = acute kidney injury; ALI1 = acute liver injury primary endpoint; ALI2 = acute liver injury secondary endpoint; CKD = chronic kidney disease; CPRD = Clinical Practice Research Datalink; DK = Denmark; DKA = diabetic ketoacidosis; GI = genital infection; HIRD = HealthCore Integrated Research Database; UTI = urinary tract infection.

Key: "CPRD, DK, HIRD" means that the variable is available in the mentioned database and will be used for that outcome and assessed using information any time before and including the index date.

[a] Whether patients switch to vs. add on the index medication as defined in Section 8.1.5.

[b] Whether new users by starting the study medication become new users of monotherapy, dual therapy, or triple therapy as defined in Section 8.1.5.

The list of all codes to be used in each database is available in Annex 3.

8.3.1 Demographic and lifestyle variables

Age in years: defined as year of the index date minus year of birth and categorised as 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, or 80+ years. (The index date is the date of the first prescription/dispensing of the study medication of interest within the study period.) In the subgroup analysis, the age categories will be ≤ 65 years and > 65 years.

Sex: as recorded in the database and categorised as male or female.

Smoking status

- **In CPRD GOLD**: smoking status will be categorised as current smoker, former smoker, never smoker, or missing. The definition will be based on the available information closest to the index date from the *10 years before or at the index date*. Smoking status will be derived from entity type 4 *OR* from Read codes:
  - Using entity type 4: smoking status will be assigned based on the following hierarchy: Current (data1 = 1); Former (data1 = 3); Non-smoker (data1 = 2) (i.e., if a patient is classified as former and current at a visit, then the patient is treated as a current smoker at that date).
  - Using Read codes: as classified by categories using Read codes for "Smoking" provided in Annex 3.
  - If patients have no data in entity type 4 and none of the Read codes, smoking will be classified as missing. Patients that would be classified as never smokers by their most recently available smoking data (i.e., data closest to the index date) but that have prior data that would classify them as current or former smokers will be classified as former smokers. If, on the same day, there is an entity type and a Read code resulting in a different classification, the classification from entity type will be kept. If there are two Read codes on the same day with different classifications, the "worst case" classification will be kept—i.e., if a patient is classified on the same day as both former and current smoker, the patient will be considered a current smoker.
- **In CPRD Aurum**: smoking status will be categorised as current smoker, former smoker, never smoker, or missing. The definition will be based on the available information closest to the index date from the *10 years before or at the index date*. Smoking status will be derived from SNOMED/EMIS codes and values, as follows:
  - Using SNOMED/EMIS codes: As classified by categories using codes in Annex 3.
  - If the patient does not have any recorded code listed, smoking will be classified as missing.
  - If two or more SNOMED/EMIS codes occur on the same day with different classifications, the "worst case" classification will be kept—i.e., if a patient is classified on the same day as both a former and current smoker, the patient will be considered a current smoker.
  - If a patient is classified as a never smoker by his or her most recently available smoking data (i.e., data closest to index date) and has prior data classifying him or her as a current or former smoker, he or she will be classified as a former smoker.
- **In Denmark**, smoking status is not available. Smoking will be defined as a dichotomous variable, i.e. (History of) Smoking: yes/no. In DK, we will be able to determine whether a person fulfil proxy criteria for a history of smoking, but not whether a person is active smoker on the index date. Proxies of smoking ~~(i.e., current use of~~ will include hospital diagnoses related to smoking history; e.g. ICD-10 codes: F17 (psychiatric conditions related to use of tobacco), DZ720E (daily smoking), asthma/chronic obstructive pulmonary disease (COPD), use of medication for COPD and medications ~~6 months before the index date, and/or previous hospital diagnoses of COPD, and/or previous hospital diagnoses of disorders due to use of tobacco)~~used as support while stopping

tobacco smoking (e.g. varenicline). Proxies of smoking will be evaluated as other covariates to potentially be included in the analysis.

- **In HIRD**, smoking status is not available using claims data. Because the ICD-10 codes for tobacco-related disorders are not used regularly by providers and are suggested to have low sensitivity, we had removed this condition as a proxy. Clinical data on smoking in the HIRD are also not available for a large portion of the population. Therefore, we will not have proxies for smoking status in HIRD. We will be assessing COPD as a variable under Medical History as specified in Table 3 above.

Alcohol use

- **In CPRD GOLD**, alcohol use will be categorised as never or former drinker, low-moderate drinker (1-6 units/day), heavy or very heavy drinker (7+ units/day), drinker with amount unknown, or missing. The definition will be based on the available information on the drinking status closest to the index date from the *10 years before or at the index date*, and the alcohol amount drunk will be derived from entity type 5 *OR* based on Read codes:
  - Using entity type 5:
    - Drinking status will be defined as recorded in entity type 5 (recoded in units per week): drinking status will be assigned using the following hierarchy Current (data1 = 1); Former (data1 = 3); Non-drinker (data1 = 2).
    - Alcohol amount: the maximum reported value in data2 will be assigned.
    - If patients have no data in entity type 5 and none of the codes for "Alcohol" provided in a separate Microsoft Excel file (StudyMedicalCodes.xlsx), drinking status will be classified as missing.
    - Alcohol use will be classified as follows:
      - 0 = Non-drinker, if drinking status = Non-drinker
      - 0 = Former drinker, if drinking status = Former
      - 1 = Low to moderate drinker, if drinking status = Current and alcohol amount = 1 to 6 units/day
      - 2 = Heavy or very heavy drinker, if drinking status = Current and alcohol amount = 7+ units/day
      - 8 = Drinker amount unknown, if drinking status = Current and alcohol amount = missing
  - Using Read codes: based on the presence of any Read code for "Alcohol" provided in a separate Microsoft Excel file (StudyMedicalCodes.xlsx) and classified according to the following categories: non-drinker/former; drinker/low to moderate drinker (1-6 units/day); drinker/heavy to very heavy drinker (7+ units/day); drinker–unknown quantity
  - If, on the same day, there is an entity and a Read code resulting in a different classification, the classification from entity will be kept. Unless the medical code gives information on quantity, the medical code classification will be kept. If there are two Read codes on the same day with different classifications, the "worst case" classification will be kept—i.e., if a patient is classified on the same day as both a

heavy drinker and a low to moderate, the patient will be considered a heavy drinker. If the last available record is drinker with unknown quantity, this record will be used regardless of any prior information on quantity.
- Codes indicating history of alcohol abuse or chronic alcohol use-related disorders have been classified under heavy drinker in CPRD and will not be reported as a separate variable.

- **In CPRD Aurum**, alcohol use will be categorised as never or former drinker, low-moderate drinker (1-6 units/day), heavy or very heavy drinker (7+ units/day), drinker with amount unknown, or missing. The definition will be based on the available information on the drinking status closest to the index date from the *10 years before or at the index date*, and the alcohol amount drunk will be derived from SNOMED/EMIS codes and values, as follows:
  - Based on the presence of any code related to alcohol consumption that is listed in Annex 3, patients will be classified according to the following categories: never or former drinker, low to moderate drinker (1-6 units/day), heavy to very heavy drinker (7+ units/day), or drinker–unknown quantity.
  - If the quantity of alcohol it is unknown from the code, then the value and numunitid will be used to derive it as follows:
    - If the units indicate low or high (column value_unk in tab alcohol_units) then assign this as the amount.
    - If the numunitid corresponds to a value in the column value_unk =’unk’ from tab alcohol_units, then the amount of drinks will be calculated as follows:
      - If the numunitid corresponds to a value in the column value_div not missing, then the amount = value / value_div.
      - If the numunitid corresponds to a value in the column value_mult not missing, then the amount = value * value_mult.
      - Then the category will be assigned as low to moderate drinker (1-6 units/day), heavy to very heavy drinker (7+ units/day) depending on this calculated amount.
  - If two or more SNOMED/EMIS codes occur on the same day with different classifications, the “worst case” classification will be kept—i.e., if a patient is classified on the same day as both a heavy drinker and a low to moderate drinker, the patient will be considered a heavy drinker.
  - If the last available record is drinker with unknown quantity, this record will be used regardless of any prior information on quantity.
  - Codes indicating history of alcohol abuse or chronic alcohol use-related disorders have been classified under heavy drinker in CPRD and will not be reported as a separate variable.

- **In Denmark**, alcohol use and overuse will be based on proxies (i.e., previous hospital diagnoses of problems related to alcohol use, alcohol dependence, previous alcohol intoxication, alcoholic organ disease, and/or previous use of disulfiram for alcoholism treatment).

- **In HIRD**, alcohol use is not available using claims data. However, diagnosis and procedure codes can be used to identify patients with a history of alcohol abuse/dependence, to indicate severe alcohol use.

Body mass index (kg/m^2)/obesity

- **In CPRD GOLD**, body mass index (BMI) will be categorised as underweight (BMI < 20), normal (BMI 20 to < 25), overweight (BMI 25 to < 30), obese (BMI ≥ 30), or missing. The definition of BMI/obesity will be based on the available information closest to the index date during the *3 years before or at the index date*, which will be derived from entity type 13 and/or entity type 14 or from Read codes:
  - BMI as recorded in entity type 13, data3 (BMI = maximum recorded BMI measure taken that day) *OR* calculated BMI based on the data recorded in variables "weight" (entity type = 13) and "height" (entity type = 14):
    - Using the most recent available information during the 3 years before or on the index date for "weight" as recorded in entity type = 13 data1 (weight = maximum recorded weight for that day) *AND*
    - Using the most recent available information ever for "height," as recorded in entity type = 14 data1 (height = maximum recorded height for that day) (since height will remain relatively constant through time)
    - Formula for BMI calculation: weight/height^2
  - Using Read codes: based on the Read Codes for "BMI and Obesity" provided in a separate Microsoft Excel file (StudyMedicalCodes.xlsx) and that are classified as underweight, normal weight, overweight, or obese.
  - If patients have no data in "recorded BMI (in entity type 13)," or "calculated BMI," or "BMI and Obesity" codes, then BMI/Obesity = unknown. If a patient has two measures on the same day, the maximum will be considered. If a patient has a Read code on the same day as the entity, the entity measure will be included. When describing percentage of patients with obesity, these will be identified as those in the category "Obese."
- **In CPRD Aurum**, BMI will be categorised as underweight (BMI < 20), normal (BMI 20 to < 25), overweight (BMI 25 to < 30), obese (BMI ≥ 30), or missing. The definition of BMI/obesity will be based on the available information closest to the index date during the *3 years before or at the index date*, which will be derived from SNOMED/EMIS codes for BMI and its related values or calculated from weight and height, using codes, value, and unit, as follows:
  - <u>BMI as a numeric variable</u> will be derived. All needed values will be kept at each record of the observation data set:
    - BMI: only records within 3 years before or at the index date will be used.
      - To obtain numeric BMI records of BMI will be identified from SNOMED/EMIS codes listed in Annex 3 and:
        - BMI will be assigned as variable value.

- If records indicate 'Overweight', 'Obese', 'Underweight', or 'Normal' and numunitid in (108, 157 , 568, 657, 907, 1856 , 1859, 2050 , 10176 , 13294) then BMI = value.
- If the resulting BMI value is < 10 kg/m^2 or > 80 kg/m^2, it will be set to missing.

- Height: all records of height at any time before or on the index date will be used.
  - Records of height will be identified from SNOMED/EMIS codes listed in Annex 3 specific for height and if variable value is 100 < value < 200 then it will be transformed to meters dividing by 100.

- If records indicate 'Height_Weight', 'Overweight', 'Obese', 'Underweight', or 'Normal' and numunitid in (122, 173, 408, 432) then variable value will be considered height if 1 < value <2 or if variable value is 100 < value < 200 then it will be transformed to meters dividing by 100.Weight: only records of weight within 3 years before or at the index date will be used.
  - Records of weight will be identified from SNOMED/EMIS codes listed in Annex 3 specific for weight and if and 40 < value < 180 then weight = value.

- If records indicate 'Height_Weight', 'Overweight', 'Obese', 'Underweight', or 'Normal' and numunitid in (156, 827, 1778, 1788, 2514, 3283, 3906) and 40 < value < 180 then weight = value.Once we have all individual records for each patient over time, we will use the values of height and weight closest to index date, and calculate BMI as weight/$height^2$. These values do not need to be recorded on the same date. BMI will be calculated as a numeric (continuous) variable and will be classified into four groups: underweight (< 20), normal (20 to < 25), overweight (25 to < 30), obese (≥ 30), or missing. The last value of BMI before or on the index date will be kept.

- BMI as a categorical variable from SNOMED/EMIS codes:

  - Based on the codes in Annex 3, which are classified as underweight (< 20), normal (20 to < 25), overweight (25 to < 30), obese (≥ 30), or missing
  - Then, if a patient has two or more different classifications on the same date:
    - If the patient has two or more numeric different results, the highest BMI will be considered the actual BMI.
    - If the patient has contradictory numeric value and SNOMED/EMIS code for BMI, the numeric result will be considered the actual BMI.
    - Finally, if the patient has two or more SNOMED/EMIS codes classifying BMI in different categories on the same date, the “worst case” scenario will be considered, e.g., a patient with codes for Normal and Overweight will be classified as Overweight.

  - Finally, the date closest to the index date within 3 years before or at the index date will be chosen as the final classification. If numeric BMI is available, the value will be categorised as underweight (< 20), normal (20 to < 25), overweight (25 to < 30), obese (≥ 30).
  - Dichotomous BMI variable: When describing the percentage of patients with obesity, patients will be classified as obese (those in the category “obese”) or not

obese; patients with missing BMI data will be included as not obese in this classification.

- **In Denmark**, presence of obesity will be categorised as yes/no. The definition will be based on previous hospital diagnoses of obesity, hospital treatment including surgical treatment for obesity, and previous use of drug treatment for obesity.
- **In HIRD**, BMI (or a proxy) is not available using claims data at this time.

Socioeconomic status

- **In CPRD**, socioeconomic status (SES) will be defined as reported in CPRD linkable database for the Index of Multiple Deprivation and categorised as quintiles of the Index of Multiple Deprivation, at a patient level if available or practice level if not.
- **In Denmark**, proxies of SES will be based on marital status (married, divorced, widowed, never married) and place of residence (~~rural, urban,~~based on the municipality of residence: rural or urban).
- **In HIRD**, SES (or a proxy) is not available using claims data at this time.

Other variables

Calendar year of the index date (cohort entry): defined as the year when the first prescription/ dispensing that qualifies the patient to enter the study occurred; categorised as 2014, 2015, 2016, 2017, 2018, and 2019. Some calendar years may be grouped if data are sparse within individual calendar years.

Duration of lookback period: defined as the number of years that a patient has been enrolled in the database, estimated as the time from date of first enrolment in the database until the index date. In HIRD, there may be patients with intermittent enrolment prior to the required baseline period. Therefore, the total time enrolled prior to the follow-up will be the sum of the duration of the periods of enrolment at baseline.

8.3.2 Medical conditions

Medical conditions will be defined based on diagnoses recorded in the outpatient visits (Read codes in CPRD GOLD, SNOMED/EMIS codes in CPRD Aurum, ICD-10 codes in Denmark, and ICD-9-CM/ICD-10-CM in HIRD) or hospitalisations (ICD-10 and procedure codes in HES and Denmark, and ICD-9-CM/ICD-10-CM diagnosis codes and procedure codes HIRD) at any time before or at the index date, unless otherwise specified (see Table 2 through Table 4). Read codes to identify comorbidities in CPRD GOLD, SNOMED/EMIS codes in CPRD Aurum, ICD-10 and procedure codes in HES and Denmark, and ICD-10-CM codes in HIRD are provided in Annex 3. Because in Denmark only hospital data are available, proxies via procedures, medication, and other will be included. Absence of a diagnosis code or proxies for a condition will be regarded as absence of the condition.

Programming note regarding calculation of Charlson Comorbidity Index: for solid tumors, only the highest of malignant neoplasm or metastatic solid tumor should be included. Likewise, for liver, the highest of mild or moderate liver disease should be included.

In CPRD, ICU will be defined both from codes in GOLD/Aurum and HES. In GOLD/Aurum any code in the tab "Intensive_care_unit", but length of ICU stay cannot be calculated. In HES, the ICU and its length of stay will be defined from episodes with variables *mainspef* or *tretspef* = 192, and the length of duration will be the sum of variable epidur from different episodes within the last 12 months prior or at the index date.

8.3.3 Variables related to severity of type 2 diabetes

Diabetes control

HbA1c as a continuous variable will be described in mmol/mol. Categories defining diabetes control will be:

- HbA1c ≤ 53 mmol/mol or ≤7%, n (%)
- HbA1c > 53 mmol/mol and ≤ 63,9 mmol/mol or >7% and ≤8%, n (%)
- HbA1c > 63,9 mmol/mol and ≤74,9 mmol/mol or > 8% and ≤9%, n (%)
- HbA1c > 74,9 mmol/mol or >9%, n (%)
- HbA1c missing, n (%)

- **In CPRD GOLD**, diabetes control will be defined as poor or appropriate according to values of glycated haemoglobin (HbA1c). HbA1c values will be defined through laboratory results recorded under entity type 275 "HbA1c - diabetic control" or Read codes in excel code list within the 12 months before or on the index date. When more than one value exists, the value recorded at the date closest to and before or on the index date will be used. If, in the same date, a value was recorded through a Read code and, also, a value was recorded through entity type, entity type value will be used.
  Categorization of reported HbA1c:
  - If data2/value <2 set to missing
  - If data2/value between 2 & 20, consider reported as %
  - If data2/value between >20 & <= 200 consider reported as mmol/mol
  - If data2/value >200 set to missing

  If no valid values for HbA1c available, define control through new column in Excel indicating control (if possible)
  To estimate mean (SD) transform % to mmol/mol using the following formulae:
  [HbA1c in mmol/mol] = (10.93 * [HbA1c in %]) – 23.50
- **In CPRD Aurum**, diabetes control will be defined as poor or appropriate according to values of glycated haemoglobin (HbA1c). HbA1c values will be defined through laboratory results recorded under variable value for observations with SNOMED/EMIS codes listed in Annex 3 within the 12 months before or at the index date. When more than one value exists, the value recorded at the date closest to and before the index date will be used.
- **In Denmark**, diabetes control will be defined based on latest measured HbA1c within the 12 months before or on the index date recorded in the LAB_F and LABKA databases.

When more than one value exists, the value recorded at the date closest to and before or on the index date will be used.

- **In HIRD**, diabetes will be defined as controlled (appropriately controlled) or not controlled (poorly controlled) as assessed in the subset of patients with outpatient laboratory data (around 33% of the patients in HIRD) based on the most recent laboratory result for HbA1c within the 12 months before or at the index date recorded in the in-house laboratory database. When more than one value exists, the value recorded at the date closest to and before or at the index date will be used. Diabetes will be considered to be controlled if the HbA1c is ≤ 7% or 53 mmol/mol and uncontrolled if the HbA1c is > 7% or > 53 mmol/mol.

Time since first diagnosis of T2D: defined as the time (in months~~years~~) between the first diagnosis of T2D recorded in the database and the index date. Categories in years will be: < 2, [2-5), [5-10), ≥10 years.

T2D-related medical conditions will be defined from diagnoses recorded at outpatient visits (Read codes in CPRD GOLD, SNOMED/EMIS codes in CPRD Aurum) or hospitalisations (ICD-10 codes or procedure codes in Denmark and HES, and ICD-9-CM/ICD-10-CM codes in HIRD) at any time before or at the index date, unless otherwise specified in Table 4. Because in Denmark only hospital data are available, proxies via procedures, medication, and other will be included. T2D-related medical conditions of interest in this study are as follows:

- Renal complications
- Ophthalmic complications
- Neurologic complications
- Peripheral circulatory complications
- Coronary heart disease
    - Recent myocardial ischaemia/infarction (in the 6 months before index date)
- Cerebrovascular disease
- Amputations
- Other diabetes complications

Lists of Read, SNOMED/EMIS, procedure, and ICD-10 codes are presented in Annex 3.

8.3.4 Medications

Medications at baseline: use of medications will be defined as any prescription/dispensing received for medications in the 180 days before or at the index date, unless otherwise specified. Medications to be assessed are provided in Table 3. A single variable for use of any medication associated with ALI in the 180 days before or at the index date will be created. The medications associated with ALI are identified in Table 3 (marked as hepatotoxic [HTX]). The Gemscript codes used in CPRD GOLD, the dm+d codes used in CPRD Aurum, the ATC codes used in Denmark, and GPI codes used in HIRD for medications are included in Annex 3.

One summary variable indicating use of hepatotoxic medications will be created combining medications, indicated in Table 3 with a “HTX” superscript.

9. Data Sources

See Section 7.2.

10. DATA MANAGEMENT AND SOFTWARE/TOOLS

10.1 SOFTWARE/TOOLS

All analyses will be performed in SAS version 9.4 or higher (2013).

10.2 HANDLING OF MISSING VALUES

We will investigate the degree of missingness in the data by examining the frequency of missing values for each variable. It is expected that there will be no missing data for age, sex, disease diagnoses, and medications; however, a varying degree of missingness is expected to be seen in lifestyle variables in applicable data sources such as the Index of Multiple Deprivation, smoking, alcohol consumption, and BMI and in variables that require laboratory results to be defined, but not for lab data used for validation, except that in CPRD, when ULN is missing one can impute as the ULN of the same test performed in the same patient given that patients tend to do all the labs testing in the same lab, thus the ULN are the same. We found that this was true during patient profile review. To estimate the propensity score, we will assess the pattern of missingness for relevant variables with a significant amount of missing data (e.g., > 10%). We need to understand the mechanism of missingness because some imputation methods may introduce bias when assumptions do not hold. The strategy will be as follows:

- If data are assumed to be missing completely at random (MCAR), then no imputation will be performed, and missing data will be ignored in the propensity score models;
- If data are assumed to be missing at random (MAR), then we will impute missing values of covariates included in the propensity score model five times, and each patient's propensity score will be his/her averaged propensity score across the five imputed data sets;
- If data are assumed to be missing not at random (MNAR), then a causal pathway of the missingness mechanism would need to be developed in conjunction with clinical input so that the correct selection model could be specified. Because MNAR variables are unmeasured confounders, the potential effects of this missing data mechanism on the estimation of the treatment effect under a variety of assumptions is inherently included in the sensitivity analyses outlined in Section 11.6.4.

If the amount of missing is not significant (e.g., less 10%), for categorical variables, indicator category will be created for missing and included in the propensity score. As no continues variables will be used as a predictor of the propensity score, no needs are required. "*when data are missing not at random and the covariate with missing value is strongly associated to its missing indicator, the missing indicator variable in a propensity model may yield smaller bias than the model without it*" Choi, J., Dekkers, O. M., & le Cessie, S. (2019).

10.3 HANDLING OF SMALL CELL COUNTS AND INCONSISTENCIES IN DATA AND OUTLIERS

Small cell counts: the possibility of unintentional (deductive) disclosure arises when cells with small numbers of patients are quoted. When reporting the data, the policy is that no cell should contain less than 5 events in CPRD and Denamrk, and less than 10 in HelathCore. When distributed outside the exact number of events will be replaced by "≤5" or "≤10" as appropriate, and if needed other cells such as person-years will be masked to avoid back calculation.

- In CPRD: [redacted] will be able to share small cell counts (≤5), with BI and RPs, but not outside. A waiver may be possible to share with regulatory authorities, however this requires providing an additional report to be shared outside the regulatory bodies and/or indications for each single cell. Given the complexity, the risks of unintentional disclosure by regulatory authorities, and that we will mask in other data sources, we plan to follow same approach and mask small cells for CPRD too when sharing with regulatory authorities.
- In Denmark: Aarhus is NOT allowed to share with [redacted] BI, or outside their organisation the exact numbers for cells with small numbers of counts(≤5), not even for the PS-deciles, i.e. Aarhus will not be able to share numbers of events and person-years for PS-deciles with less than 5 events. Will still be able to share the IR and IRR with CI.
- In HIRD: upon signing a Data Sharing Agreement (DSA), HealthCore will be able to share exact numbers (including small cell counts, i.e. ≤10) with [redacted] in order to reviewing results and to perform pooled analysis. HealthCore is not able to share exact numbers outside [redacted] (not even with BI).

If data are considered to be outliers, they will be treated as missing values, except otherwise specified (see lab values in Section 8.2). Approaches to handling missing values are described in the previous section. Outliers will likely be found in laboratory results data or exposure data, depending on each data source.

11. PLANNED ANALYSIS

11.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic and lifestyle variables are described in Section 8.3. Each study outcome has its own set of associated demographic and lifestyle variables to be used for describing the study population and for consideration as covariates in each outcome-specific propensity score model (see details on the descriptive and propensity score analyses in Sections 11.5 and 11.6). A list of demographic and lifestyle variables by study outcome is provided in Table 3.

11.2 CONCOMITANT DISEASES AND MEDICATION

Variables related to concomitant diseases and medications are described in Section 8.3. Each study outcome has a specific set of medical comorbidities of interest. Concomitant diseases and medication variables will be used for the following purposes: (1) descriptors of the study population, (2) exclusion and censoring criteria, (3) potential confounding or effect modifying variables to be evaluated for inclusion in each outcome-specific propensity score model, and (4) stratification variables (see details on each analysis in Sections 11.5 and 11.6). Lists of medical comorbidities and medications of interest by study outcome are provided in Table 3 and Table 4.

11.3 METHODS ADDRESSING BIAS

Potential biases will be addressed as follows:

- Implementing a study design that minimises selection bias and immortal time bias. Features of the study design 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 the or at index date.
- Performing sensitivity analysis with different definitions of exposure—as treated, intention to treat (ITT), different latency periods (grace period of current use 90 days after the end of supply instead of 30 days), and different lag times—to evaluate the effects of potential exposure misclassification.
- Performing outcome validation to minimise outcome misclassification.
- Performing subgroup analysis, e.g., patients in mono-, dual, or triple therapy and patients that are naive to, switch to, or augment T2D therapy.
- Estimating propensity score with interaction terms by calendar period to adjust for potential channelling bias. Calendar periods will be calendar years, i.e. 2014, 2015, 2016, 2017, 2018, and 2019. Calendar periods may also be combined to allow estimation of propensity scores when there are low number of users.
- In the context of this study, where the expected number of events is small for some of the outcomes (e.g., ALI), the number of covariates that could be used in a regression model predicting those outcomes is limited. To address this limitation, the set of confounding variables will be summarised into a single summary confounder score, a propensity score.
    - A propensity score will be generated representing the probability of being in the empagliflozin cohort relative to the DPP-4 inhibitor cohort. Given the different inclusion/exclusion criteria used for each outcome of interest, the propensity score will be cohort specific to the outcome being analysed (i.e., propensity scores will be calculated for only those patients included in the specific outcome analysis).

11.4 METHODS ADDRESSING CONFOUNDING/EFFECT MEASURE MODIFICATION

Methods to address confounding and effect measure modification have been described in Section 11.3.

11.5 MAIN ANALYSES

Table and figure shells of planned analyses described in this section are provided in a separate document. Any *P* values that may be generated during the course of analysis will be interpreted as descriptive statistics rather than inferential statistics.

Due to different inclusion criteria for the outcomes of interest, the analyses described in the following sections will be performed separately for each outcome. For each table or figure number specified in the following sections, a three- or four-letter prefix will be added in the table shells to identify the outcome to which the table or figure applies. For example, if a table is specified in the text as **Analysis Table 1**, then this table in the analysis would be named Table ~~ALI~~ALI1-1 for the ALI1 outcome, Table AKI-1 for the AKI outcome, Table UTI-1 for hospitalisation due to UTI, and so on. The three- or four-letter prefixes for table and figure specification of primary outcomes will be as follows:

- ALI1 for acute liver injury without predisposing conditions,
- AKI for acute kidney injury,
- UTI for hospitalisation due to severe complications of UTI (CPRD only),
- GIM for GI in males (CPRD only),
- GIF for GI in females (CPRD only), and
- DKA for hospitalisation due to diabetic ketoacidosis.

The prefixes for table and figure specification of secondary outcomes will be as follows:

- ALI2 for acute liver injury with or without predisposing conditions,
- CKD for chronic kidney disease (CPRD only),
- GIMH for severe genital infection in males (CPRD only), and
- GIFH for severe genital infection in females (CPRD only).

The prefixes for table and figure specification of sensitivity analysis outcomes will be as follows:

- ALI1S (CPRD and HIRD): for ALI1 sensitivity analysis, i.e. all cases of ALI1 identified in the primary outcome + patients who only received an ALI diagnosis in the primary care setting.
- AKIS (CPRD and HIRD): for AKI sensitivity analysis, i.e.:
    - all cases of AKI identified in the primary outcome +
    - CPRD: outpatient cases of AKI identified in the primary care or outpatient setting
    - HIRD: patients who had one or more claims for a PCP office visit with an AKI-related ICD-9-CM/ICD-10-CM diagnosis code and no claims for AKI-related hospitalisations, ED visits, or specialist (nephrology) visits.
- CKDS (CPRD only): for CKD sensitivity analysis, i.e. cases of CKD identified based on eGFR below 60 mL/min/1.73 m2, through either recorded data values associated with Read or SNOMED/EMIS codes or estimated eGFR using serum creatinine.

11.5.1 Cohort attrition and study therapy users

The study cohort assembly process is outlined in **Analysis Table 1**. New users of empagliflozin or a DPP-4 inhibitor will be identified as separate study cohorts. The number of prescriptions/dispensings and patients eligible to be included in the analysis will be provided, starting with any prescription/dispensing of the cohort therapy (e.g., empagliflozin or DPP-4 inhibitor) during the study period and upon the application of each individual entry criterion independently of the other criteria, as follows:

- Age 18 or older on the date of the cohort therapy prescription
- At least 12 months of available data before the date of the cohort therapy prescription
- Indication of T2D at any time before or on the date of the cohort therapy prescription
- No previous prescription of any SGLT2 inhibitor (including empagliflozin) or DPP-4 inhibitor alone or in fixed-dose combination within 12 months prior to and including the date of the cohort therapy prescription
- No indication of T1D at any time before or on the date of the cohort therapy prescription

Note for programmers: application of each individual entry criterion independently means that for each criterion (each row in the table), the number of patients remaining in the sample after applying only that criterion to the total population (numbers in row 1) will be reported. That is, start with the numbers in row 1, then subtract the patients excluded after applying each criterion as if it were the only criterion.

The number of prescriptions and patients meeting all of the above criteria will then be tabulated for each study cohort. Among these prescriptions and patients, the number meeting the outcome-specific inclusion/exclusion criteria (Section 7.3.4) will then be tabulated.

Additional criteria specific to certain outcomes of interest, as specified in Section 7.3.4, will be applied to create some outcome-specific cohorts for analysis. No additional criteria will be applied to select cohorts that will be used to evaluate GI, severe GI, and DKA; the main cohorts will be used for these analyses.

Characteristics of new therapy use will be presented in **Analysis Table 2**. Descriptive statistics will be presented for the following characteristics of new use: total number of new users; number of new users of study index medications that are in fixed-dose and free-dose combination with metformin (i.e., taken with metformin but under a separate product); number of new users that are switchers and subgroups of switchers; number of new users overall that have augmented their therapy (i.e., add-on users) and subgroups of add-on new users; number of new users on monotherapy, dual therapy, or triple therapy (see definitions in Section 8.1.5); daily dose; use of concomitant insulin; number of prescriptions/dispensings at the index treatment episode; and the duration of the index treatment episode. A description of daily dose of each study medication at the index date will be presented in **Analysis Table 3**.

11.5.2 Descriptive analyses of study population

For continuous variables, descriptive statistics will include the number of non-missing values, mean, standard deviation, median, first quartile, third quartile, minimum value, and

maximum value (percentile 1, and percentile 99 to avoid small cell counts). For categorical variables, descriptive statistics will include frequencies and percentages. Descriptive statistics will be generated for the patient demographic and lifestyle variables defined in Section 8.3, at the time of the index date, for each outcome-specific population (**Analysis Table 4**).

Similar tables with descriptive statistics will be generated for patient comorbidities, including indicators of diabetes severity, at the index date (**Analysis Table 5**) and patient medications at the index date (**Analysis Table 6**). Variables to be included in these descriptive tables are defined in Section 8.3. Additionally, the standardised bias of the DPP-4 inhibitor cohort vs. the empagliflozin cohort will be calculated for each demographic or lifestyle variable to assess the initial balance of measured covariates [R16-4915] before propensity score trimming. The standardised bias for categorical variables will be computed as the difference in proportions of each level of the covariate (DPP-4 inhibitor minus empagliflozin) divided by the standard deviation of the DPP-4 inhibitor cohort. For continuous variables, the standardised bias will be computed as the difference in means (DPP-4 inhibitor minus empagliflozin) divided by the standard deviation of the DPP-4 inhibitor cohort. Potential imbalance in each covariate would be indicated initially if the standardised bias (SB) is more than 0.25 or less than -0.25. However, if a certain variable has a standardised bias greater than 0.25 or less than -0.25 and the difference in that variable between cohorts is deemed to have no clinical significance, then the +/-0.25 standardised bias criterion may be relaxed for that particular variable. SB across deciles is a stricter measure of assessing balance, and it may be that balance is achieved across deciles even if there is one decile with SB>0,25, i.e. some of the differences in the individual deciles may go in opposite directions and cancel each other on average.

For categorical variables:

$$Standardized\ Bias = \frac{\hat{p}_{DPP4} - \hat{p}_{Empa}}{\sqrt{\hat{p}_{DPP4}(1 - \hat{p}_{DPP4})}}$$

For continuous variables (SD = Standard Deviation):

$$Standardized\ Bias = \frac{\hat{\mu}_{DPP4} - \hat{\mu}_{Empa}}{SD_{DPP4}}$$

Average SB will be computed as "SB_mean: an average of the values of the SB from the 10 deciles, not weighted and including the sign". These will be estimated in CPRD and Denmark where there bigger differences have been observed across deciles, while balance was achieved in all deciles in HIRD.

11.5.2.1 Unadjusted estimates and 95% confidence intervals

Unadjusted or crude IRs of the primary and secondary outcomes of interest will be computed as the number of events divided by the total person-years at risk in the empagliflozin and DPP-4 inhibitor cohorts overall and for subgroups with and without insulin use at the index date. Unadjusted IRs will be estimated overall and stratified by age and sex (**Analysis Table**

7). The 95% confidence intervals (CIs) around IR point estimates will be computed using the relationship between the Poisson distribution and chi-square distributions as described in Dobson et al. [R16-4917]: Note that formulas have been updated.

$$IR = \frac{Number\ of\ cases\ during\ time\ at\ risk}{Total\ time\ at\ risk\ (person\ years)} = \frac{E}{Y}$$

$$IR_{LB} = max\left(0, \left(\frac{X^2_{2E,0.025}}{2}\right)/Y\right), IR_{UB} = \left(\frac{X^2_{2(E+1),0.975}}{2}\right)/Y$$

In the above equation, *IR* is the IR for each outcome in each study cohort, *E* is the number of events, *Y* is the total number of person-years at risk, ~~*IR*~~$_L$*IR*$_{LB}$ is the lower confidence limit, and ~~*IR*~~$_U$*IR*$_{UB}$ is the upper confidence limit.

Additionally, unadjusted IRRs will be generated comparing empagliflozin with DPP-4 inhibitor (IR for empagliflozin cohort divided by IR for DPP-4 inhibitor cohort) for each of the eight primary and secondary outcomes of interest. The 95% CIs around the IRR point estimate will be computed using the Poisson distribution and test-based methods defined in Sahai and Kurshid [R19-2327]:

$$IRR_{LB} = \left(\frac{Y_{Comp}}{Y_{Empa}}\right)\left(\frac{E_{Empa}}{E_{Comp}+1}\right)\frac{1}{F_{0.025,2(E_{Comp}+1),2E_{Empa}}}$$

$$IRR_{UB} = \left(\frac{Y_{Comp}}{Y_{Empa}}\right)\left(\frac{E_{Empa}+1}{E_{Comp}}\right)F_{0.025,2(E_{Empa}+1),2E_{Comp}}$$

In the above equations, *IRR* is the IR ratio of empagliflozin vs. DPP-4 inhibitor, E_{Empa} is the number of events observed in the empagliflozin cohort, Y_{Empa} is the total person-years at risk in the empagliflozin cohort, E_{Comp} is the number of events in the DPP-4 inhibitor cohort, Y_{Comp} is the total person-time at risk in the DPP-4 inhibitor cohort, IRR_{LB} is the lower confidence limit, and IRR_{UB} is the upper confidence limit. Where $F_{0.025}$ represents the upper 2.5th percentile of the F distribution with the corresponding degrees of freedom. In SAS, the code would correspond to finv(0.975,…,…).

Cumulative incidence function plots by cohort will be displayed graphically to describe the cumulative IRs of each outcome of interest during the current use period in follow-up. These figures will plot the cumulative incidence of the outcome (while accounting for censoring and competing risks) vs. time using separate lines for the empagliflozin and DPP-4 inhibitor study cohorts (**Analysis Figure 1**).

11.5.3 Main analysis for each primary and secondary outcome: propensity score analyses

Cases to be included in the analysis will be those that have been identified through electronic algorithms whether or not they have been validated in the three different data sources.

The following analyses will be performed for all primary and secondary outcomes unless otherwise indicated.

11.5.3.1 Propensity score

A propensity score model will be generated for each outcome population. All variables related to the specific outcome will be used for the estimation of propensity score, regardless of whether they are related to the exposure, as suggested by Brookhart et al. [R12-1913]. A screening process for candidate variable inclusion in the propensity score model will be performed. For each outcome of interest, the initial list of candidate predictor variables will be the relevant patient demographic/lifestyle characteristics, comorbidities, medications, and indicators of diabetes severity variables presented in Table 3, except for the variables that are exclusion criteria for the specific outcome of interest.

For each baseline covariate of interest, a univariable Poisson regression model with log-time offset will be generated by calendar period (i.e. natural years) of index date where the outcome during current time at risk will be modelled as a function of the baseline covariate only (**Analysis Table 8**). The selection of the variables as described in the following sentences was modified according based on team discussions and following the text in the next paragraph. If the resulting point estimate of the IRR meets a certain threshold (e.g., > 1.25 or < 0.8) within any given calendar period, then the baseline covariate will be selected to be in the propensity score model. Certain variables that do not meet these thresholds may also be included in the propensity score model if they are deemed clinically important. Additionally, if not enough events are observed to perform univariable Poisson regression models, the list of variables to be included in the propensity score model may be determined by clinical judgement or from the literature. Variables with a low prevalence (< 5% in both groups) will not be considered for inclusion in the propensity score model unless they are considered to be clinically important risk factors for the outcome. This will be based on previous knowledge and on the results of the association of each variable with the outcome and exposure. Alternatively, rather than univariable Poisson models, certain core variables can be forced into the model to create a “basic” Poisson model. The effect of each additional covariate on top of these “basic” variables can be assessed for its association with the outcome. If a research partner decides to use this covariate screening strategy, the list of “basic” variables should be made explicit.

Based on discussions during development of the analysis, it was agreed that as a first step of PS building and selection of the variables, all variables developed for each outcome would be included in the propensity score except the following: 1) higher level variables for which its components have already been included in the PS, e.g: “Autoimmune and connective tissue diseases will not be included since “Rheumatoid arthritis”, “Polymyalgia rheumatica”, and “Other autoimmune and connective tissue diseases” are already included in the PS, 2) hepatotoxic and nephrotoxic medications have been collapsed into two variables and included in the PS for ALI (hepatotoxic medications) and for AKI and CKD (nephrotoxic medications), 3) some medications such as immunosuppressants have been included despite some specific components such as methotrexate is under the variable hepatotoxic medications have been included. Also, the PS does not include variables with very low or very high OR, as they are covariates that predict treatment status almost perfectly. Consider excluding

variables with very low or very high OR (i.e., or < 0.10 and or > 10), as they may cause limited common support. A document "MAP_variables for PS.xls" was created and discussed and contains the details of the variables included.

For CPRD:

Cocaine and Zoledronic acid were excluded from DKA and from GIM, GIF and UTI PS models respectively.

For each outcome-specific population, among patients in the empagliflozin and DPP-4 inhibitor study cohorts, propensity scores will be estimated through logistic regression models where the dependent variable will be the index medication (1 = empagliflozin initiator, 0 = DPP-4 inhibitor initiator) and all variables identified as described above will be the independent variables. To account for potential channelling bias, a full interaction propensity score multivariable model (interactions of each variable with year of index date) will be generated where calendar year will also be included as a covariate, along with its interaction with each of the other covariates. This full interaction propensity score multivariable model will allow the influence of each covariate in predicting treatment to vary across calendar years (thus, accounting for potential channelling bias) and has an advantage of efficiency in providing one overall comprehensive model compared with generating separate propensity score models by calendar year [R16-4007, P12-08895]. The final multivariable model (full interaction PS model) will be used to compute a propensity score for each patient in the empagliflozin and DPP-4 inhibitor study cohorts. Additionally, interactions between specific covariates may be added to the propensity score model if deemed appropriate at the level of the individual data source. Odds ratios and 95% CIs will be tabulated for the final multivariable propensity score model (**Analysis Table 9**).

PS and calendar year:

- **In CPRD**: during development of the PS analysis, no difference was observed between models including and not including interaction with calendar year. Further exploration of calendar year variable found that it was not to be associated with the outcome (Table 8), and strongly associated with exposure (predicts exposure and separates PS curves) (Table 9). Repeating Table 5 and 6 by calendar year, showed that no differences in the distribution of baseline variables within a cohort by calendar years, and also no difference in the distribution between cohorts. PS model with insulin and advanced therapy interactions did not modify the PS distributions. The two variables provided complementary information, i.e., not all patients in triple therapy were users of insulin at index date. PS distributions including interaction term with insulin or advanced therapy were similar to those without such interaction terms. Because adding calendar year interaction term increases the degrees of freedom and does not improve the PS model, and based on the parsimony principle, it was decided not to add any interaction for these variables, but still keep them in the model.
- In HIRD there were no issues with PS distributions, resulting in good overlap of scores between the two therapy cohorts. In general, based on distribution of baseline comorbidity and medication use variables (Table 5 and 6), it seems that differences between the two cohorts are smaller in HIRD than in Europe. This might explain why

HIRD does not see the 2 bumps in the PS curves that are observed in the European data. An important factor is that age differences between therapy cohorts are not as large in HIRD as in the EU databases. Several variations of PS models were run to see results of including/excluding or adding interactions with: calendar year, insulin use, advanced therapy, and "Chronic renal disease or renal dialysis." The association between history of CKD ("Chronic renal disease or renal dialysis") and exposure is not that strong (OR around 0.7 in the PS models), and with the outcomes the OR is between 0.64 and 2.29 but for most years it is above 1. Good overlap in propensity scores for the two therapy cohorts is noted in Figure 2 after removing calendar year from the PS models. Plan: run the PS models with all variables in the MAP file as covariates, except for year of index date, without any interactions.

- In DK, PS score was explored by calendar year, CKD, and other and main drivers of nonoverlapping bimodal curves was use of insulin and other GLDs. A variable named 3 or more GLDs. DK results for the outcome DKA showed that this variable was clearly a confounder, i.e. IRR estimates of models including this variable versus crude models including age and resulted in > 10% change in the IRR coefficient. 3 or more GLDs was used to stratify the population and to do 2 PS models which lead to PS distributions with considerable overlap. The IRR estimates can be combined via meta-analysis if they are similar.

The distribution of propensity scores will be presented graphically as two stacked histograms: one histogram for empagliflozin patients and one histogram for DPP-4 inhibitor patients (**Analysis Figure 2**). This figure will allow the visual assessment of overlap of propensity score distributions between empagliflozin and DPP-4 inhibitor cohorts. If limited overlap is observed, then treatment groups would have limited comparability, and propensity score models would be regenerated after removing variables from the original model that were not related to the outcome or were too highly correlated with the exposure. If limited overlap in propensity score distributions persists, then the analytic approach and feasibility of comparison between study cohorts may be re-evaluated. Only the results of the final model will be presented.

The distribution of propensity scores for each study cohort and degree of overlap will be examined to determine the best method to trim extreme values further while minimising the exclusion of patients [R15-0905, R16-4914]. For the propensity score model, a propensity score–trimmed cohort will be created by trimming extreme values of propensity scores. Patients with propensity scores less than the 2.5th percentile of the empagliflozin cohort or greater than the 97.5th percentile of the DPP-4 inhibitor cohort will be excluded. If there is a high percentage of the cohort (e.g. 30% or more) excluded using 2.5%th percentile, trimming will be done for patients with propensity score less than percentile 1st of the empagliflozin cohort and greater than the 99th percentile of the DPP-4 inhibitor cohort. Balance should be obtained. As of May 2021: CPRD is doing 1% trimming, HIRD is doing 2.5%, and Aarhus is evaluating the two options to see if results are comparable and how many patients are trimmed, since balance is not obtained either way. The propensity score–trimmed cohorts will be used in all subsequent treatment comparisons.

Within the propensity score–trimmed cohort, the sample will be divided into deciles based on the propensity score distribution of the empagliflozin cohort to create 10 mutually exclusive strata for both the empagliflozin and DPP-4 inhibitor cohorts. If too few patients are available to form these 10 strata, then fewer strata (e.g., quintiles rather than deciles) or a matching-weights approach may be used.

Within each propensity score decile, standardised bias will be used to assess the balance of the covariates included in the propensity score model between the empagliflozin and DPP-4 inhibitor cohorts (**Analysis Table 10**). The standardised bias between DPP-4 inhibitor patients and empagliflozin patients within each propensity score decile will be computed as explained above. If substantial imbalance remains (e.g., standardised bias greater than the prespecified cut-off value of 0.25), then the logistic regression models to generate the propensity scores will be refined to achieve a greater degree of covariate balance. This refinement may take the form of adding non-linear and/or interaction terms to the logistic regression model. If important covariates remain imbalanced, then they may be included as predictor variables in the outcome model in addition to propensity score decile. Stricter trimming can be considered if imbalance is found in the extreme deciles.

11.5.3.2 Crude and adjusted estimates after trimming

Crude or unadjusted IRs will be generated for empagliflozin patients and DPP-4 inhibitor patients, along with the unadjusted IRRs (empagliflozin vs. DPP-4 inhibitors) for the overall propensity score–trimmed sample (**Analysis Table 11**). The 95% CIs around point estimates will be derived as detailed in Section 11.5.2.1.

Incidence rates adjusted for propensity score decile will be generated for empagliflozin and DPP-4 inhibitor patients along with adjusted IRRs (empagliflozin vs. DPP-4 inhibitors) overall (non-stratified by any other variable; see study objectives in Section 4). These estimates with corresponding 95% CIs will be generated through the application of a Poisson regression model where the outcome is modelled as a function of treatment cohort (empagliflozin or DPP-4 inhibitors) and propensity score decile (specified as a categorical variable) with the log of time of exposure (in years) as the offset [R16-4919]. If any variable remains unbalanced after trimming, it will be added as an independent variable in the Poisson regression model.

Protocol: "*After trimming is completed, data will be stratified into deciles of propensity scores based on the distribution among empagliflozin new users. However, if the resulting decile strata are too small (i.e., have zero events in both treatment groups), deciles may be combined (e.g., to form quintiles), or alternative methods of producing summary estimates in the presence of sparse data will be explored, such as matching weights, which has been demonstrated to yield the smallest amount of bias in health care database studies with rare binary outcomes [ref: P18-11001]. The propensity scores obtained after trimming will be used in the analysis to control for measured confounding variables simultaneously while reducing the loss in degrees of freedom that would have been experienced with a multivariable model, which is particularly important in studies with rare events. The propensity score methodology to be applied in this study will be further detailed in the statistical epidemiological analysis plan*."

Issues with convergence were encountered when there were analyses with small numbers of events. PS weighting was evaluated as an alternative to PS deciles, but results did not show an improvement and although values were obtained from the analysis, these were on the unreasonable scale (e.g.: ULN of the 95% exp(35080).

When there are Poisson regressions with warning messages such as: "The negative of the Hessian is not positive definite. The convergence is questionable… … the validity of the model fit is questionable…", generally due to low numbers: it was agreed that this would be noted in the report (maybe with a footnote in the table).

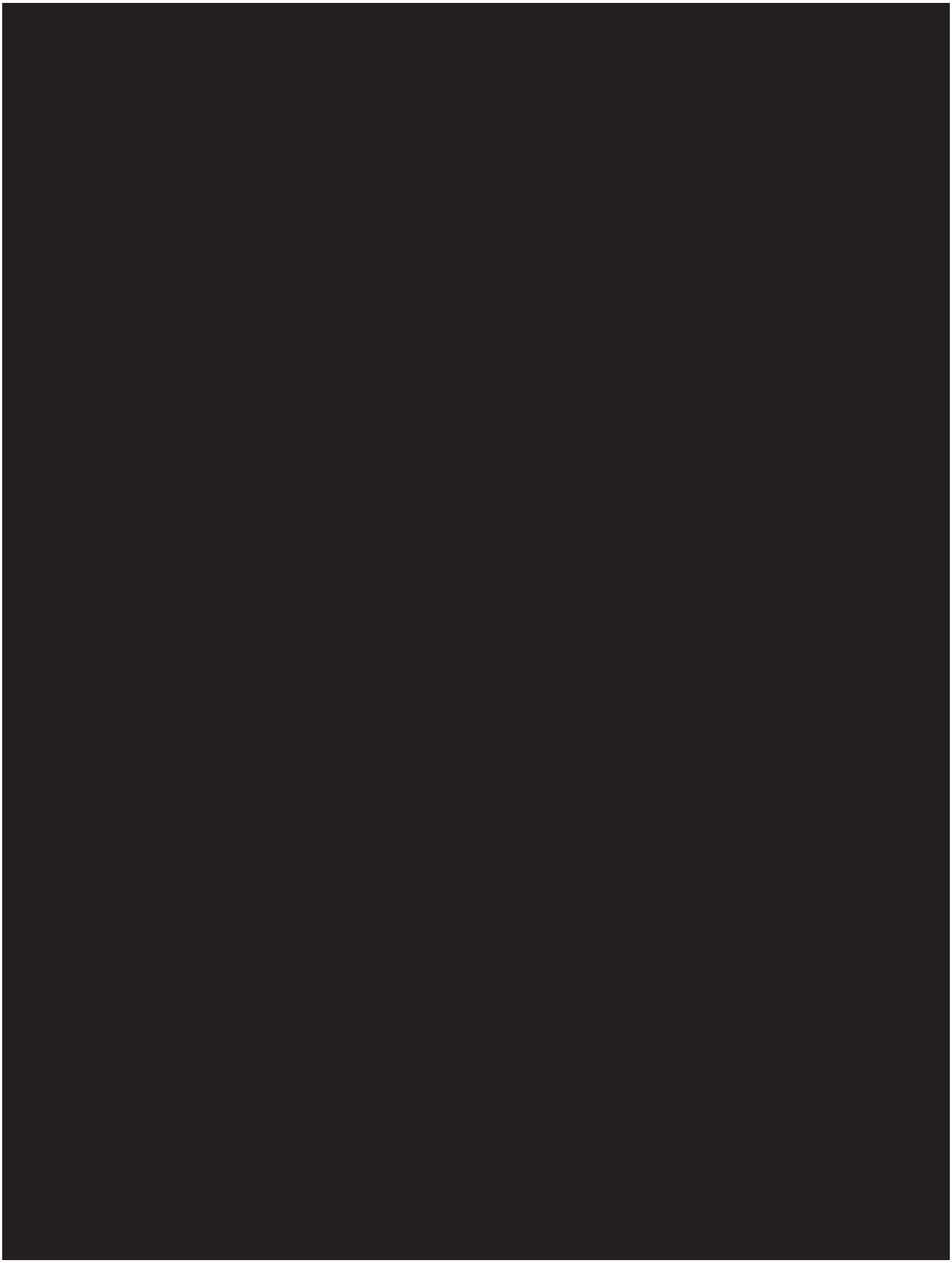
A summary of the main analysis that will be performed is presented in Table 5.

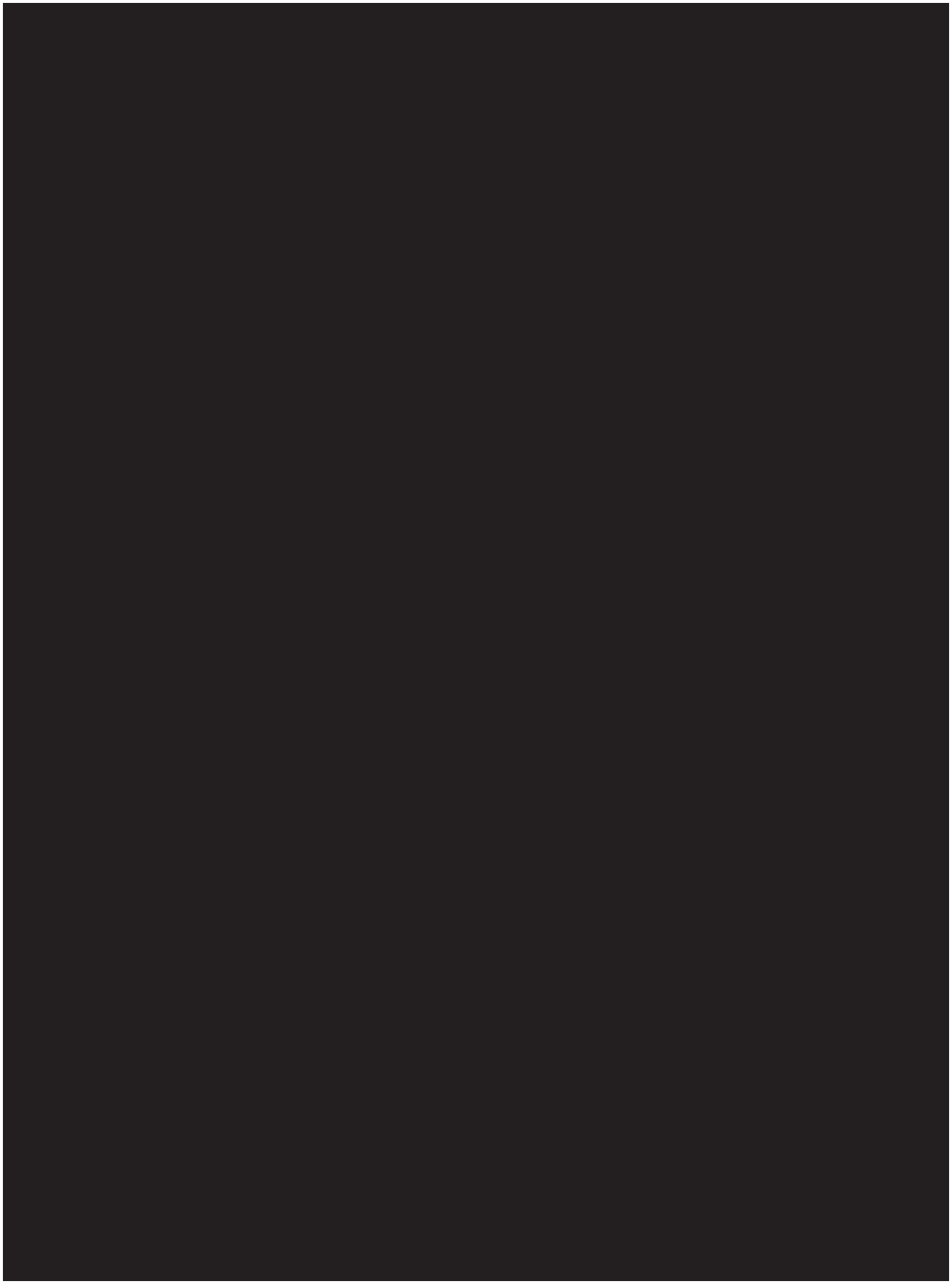
Table 5. Corresponding analysis tables and figures for analyses in all data sources providing the comparison of empagliflozin with DPP-4 inhibitors

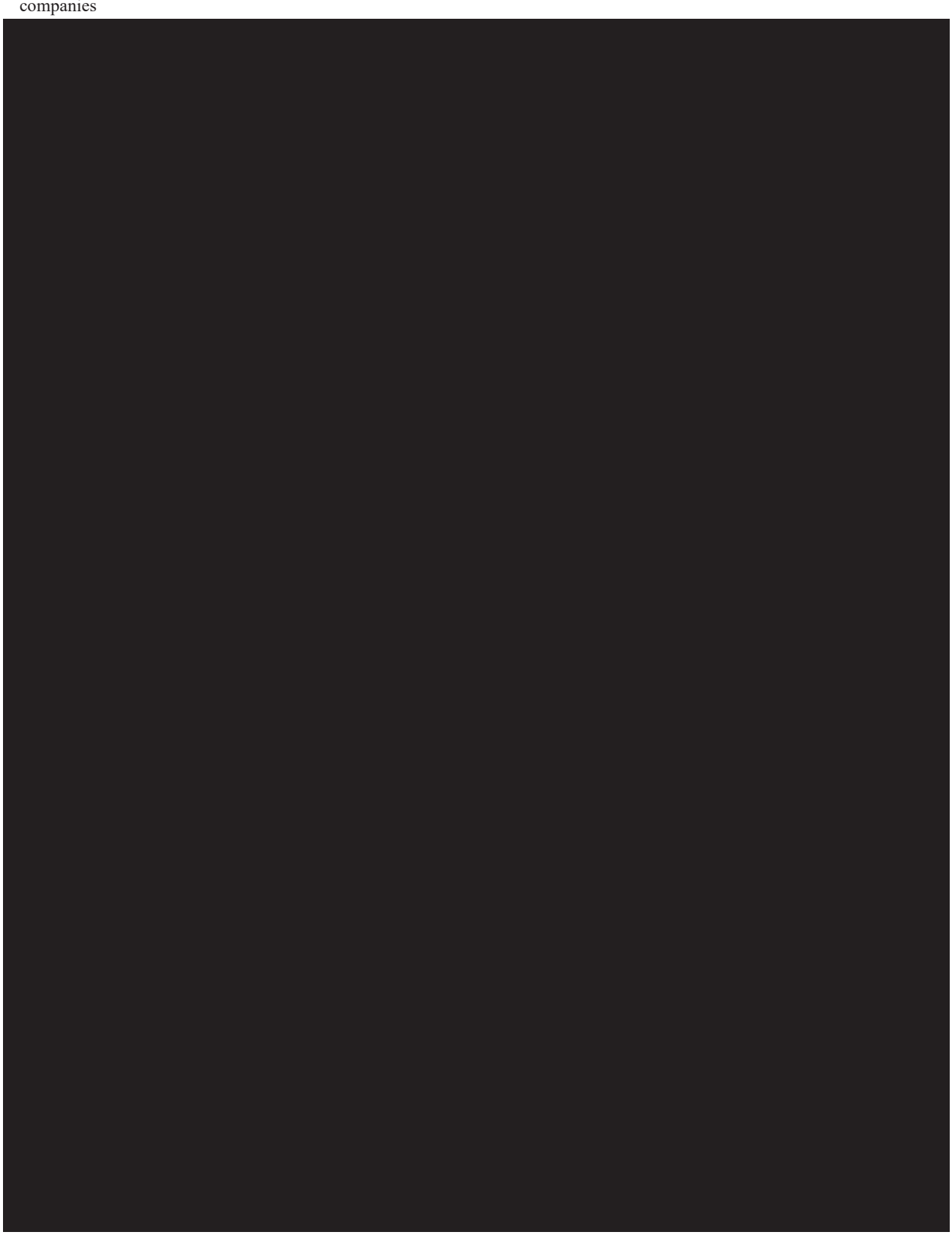
Outcomes	**Type of analyses**	**Comparison of empagliflozin with DPP-4 inhibitors**
All[a]	Odds ratio predicting empagliflozin use	Analysis Table 9
All	Distribution of propensity score	Analysis Figure 2
All	Standardised bias by propensity score decile	Analysis Tables 10
All	Unadjusted incidence rate and IRR overall and adjusted incidence rate and IRR	Analysis Table 11

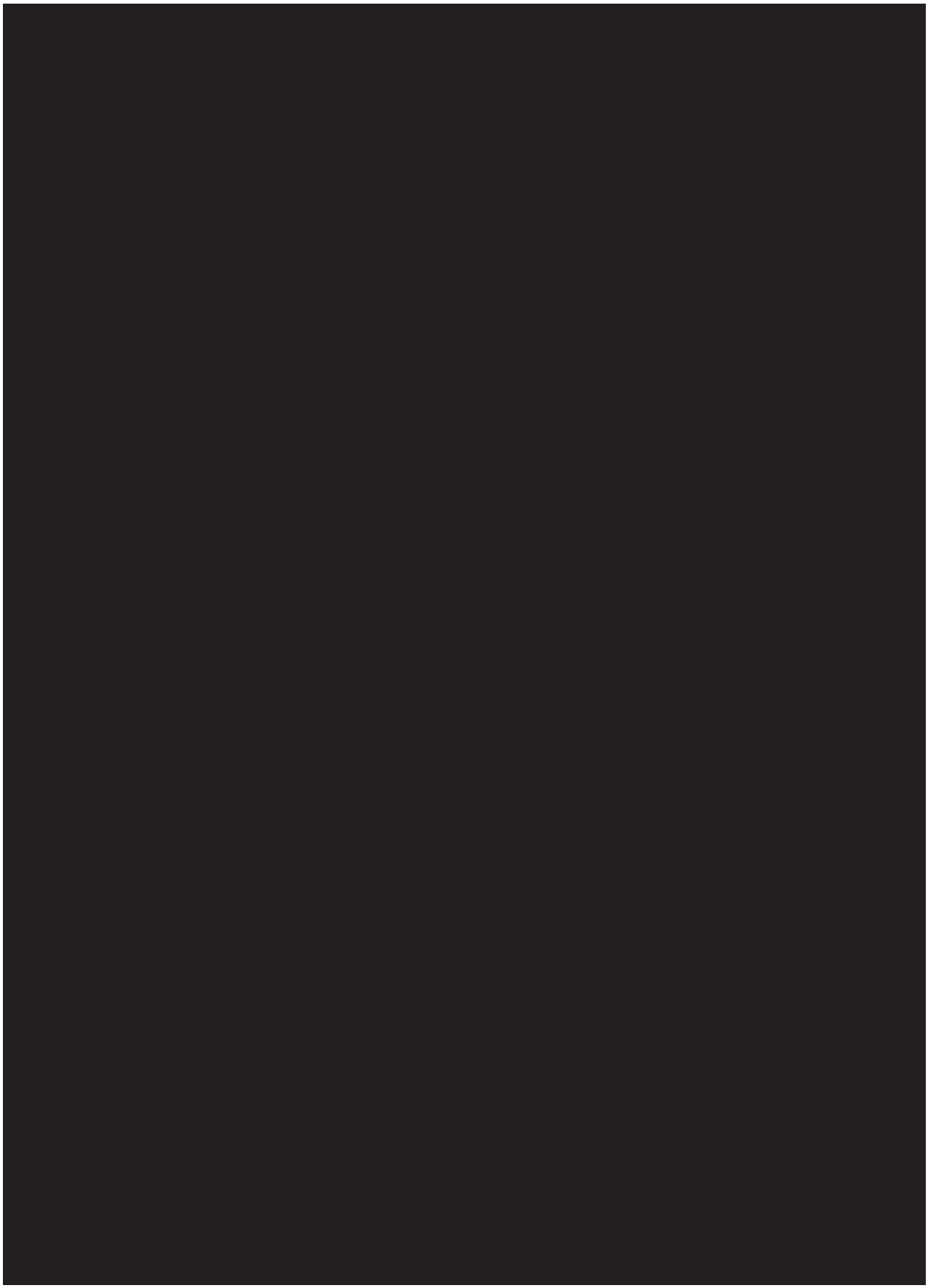
DPP-4 = dipeptidyl peptidase-4; IRR = incidence rate ratio.
[a] All outcomes as they apply to each data source.

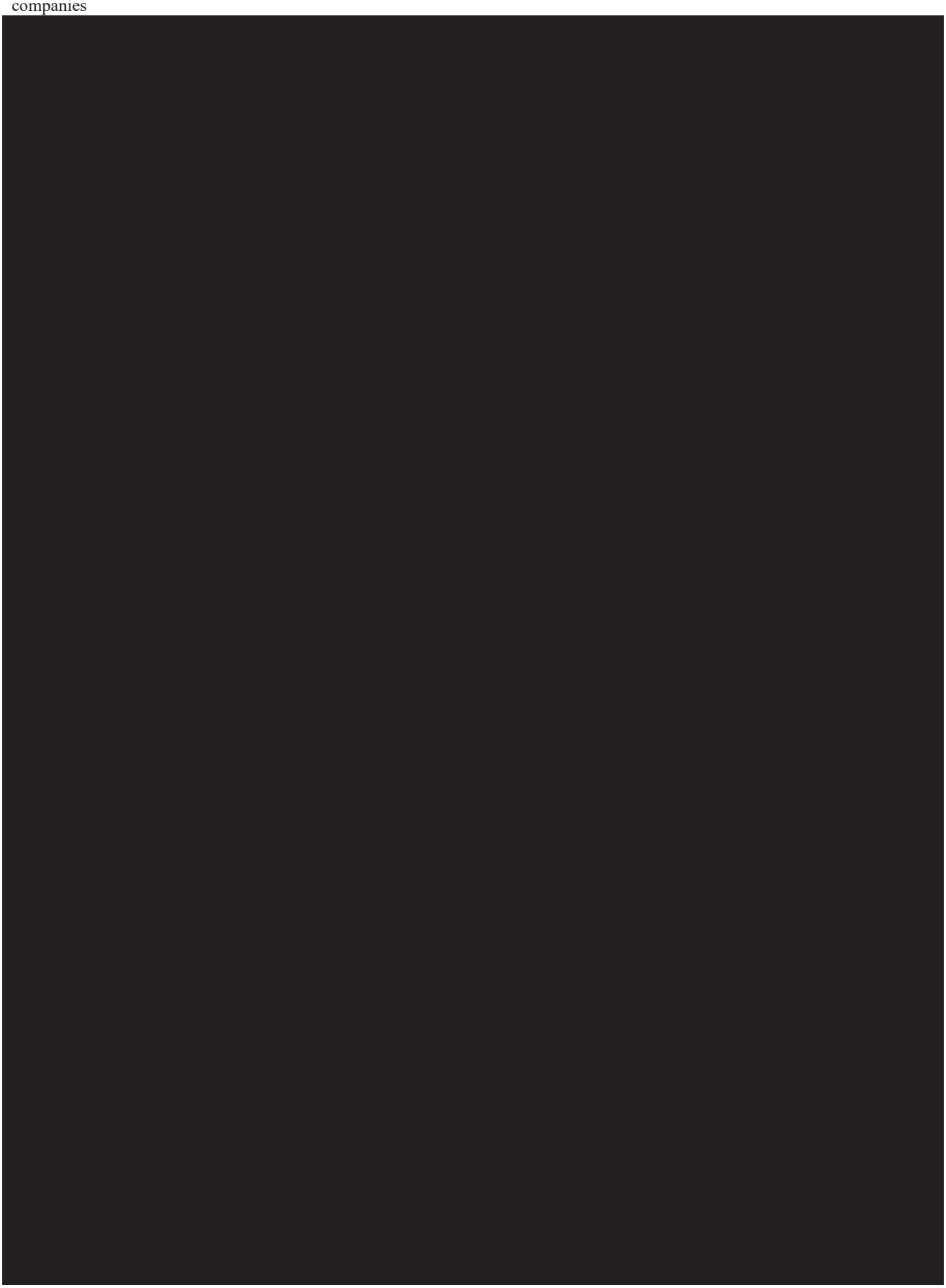
Note: when number of events is lower than allowed to be reported in each database (i.e., 1-4 in Denmark and CPRD and 1-10 in HIRD), the number of events and person-years will not be reported, but IR and IRR will be reported. This applies to all tables of results.

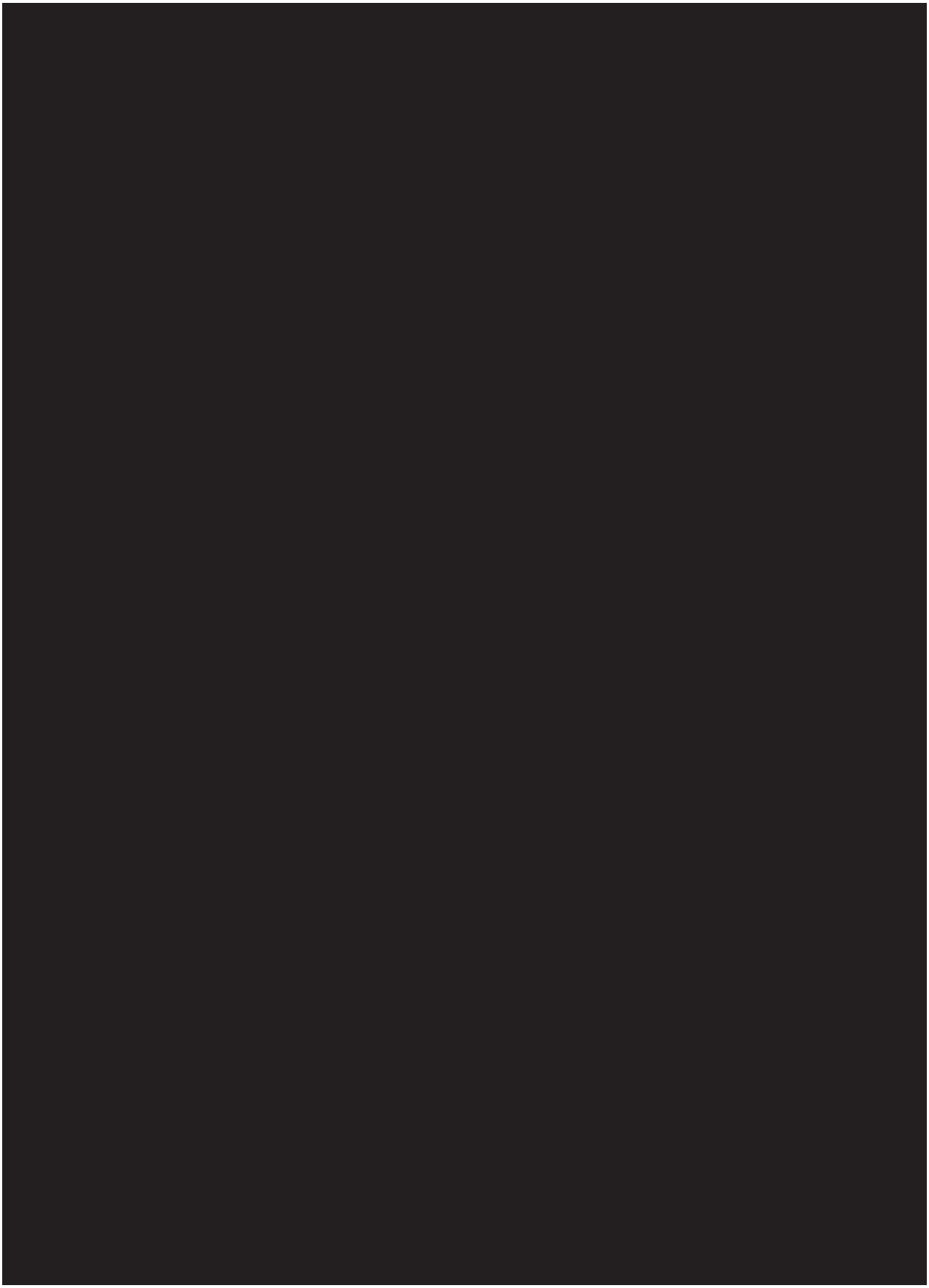












11.7 SAFETY ANALYSIS

Covered in prior sections.

12. QUALITY CONTROL

Standard operating procedures (SOPs) at each research centre will be used to guide the conduct of the study. These SOPs include rules for secure and confidential data storage and quality-control procedures for programming, standards for writing analysis plans, and requirements for senior scientific review.

For work conducted at [redacted] quality control (QC) of the data management and analysis activities of this study will be performed according to [redacted] SOPs. All data transfers from the research partners to [redacted] will be verified by [redacted] according to SOPs for electronic file acquisition and checking practices. Programming written by one study analyst will be independently reviewed by a different analyst.

Analysis data sets and programme output will be checked for accuracy and integrity according to SOPs that include the following steps:

- Checking program logs for errors and warnings
- Checking output for errors and inconsistencies
- Running quality-control programs to verify that specifications were implemented correctly and that any output generated accurately reflects the data
- Checking all results tables for accuracy

For work conducted at the [redacted] applicable SOPs will guide the conduct of the study. These procedures include internal quality review, rules for secure and confidential data storage, methods to maintain and archive project documents, and quality-control procedures for programming, as well as standards for writing analysis plans, and requirements for senior scientific review, as applicable. Epidemiologists and statisticians (analysts) at the [redacted]

[redacted] execute QC at multiple levels of the research process per internal SOPs. Quality-control measures are implemented to help prevent errors in code and output and to ensure the possibility of later verification of code and reproduction of results.

For work conducted at [redacted] the study will be tracked at various levels to help ensure that all aspects including project delivery, infrastructure, quality processes, resource management, and financial issues are addressed. To help ensure the highest level of quality on every project, HealthCore has established several layers of quality assurance throughout the project life cycle.

- *Role-Based Control Checks:* Each member of the team is responsible for performing thorough quality-control checks on their work; in addition, the Principal Investigator and Research Project Manager are also accountable for the quality of all deliverables.
- *Quality Check Points:* Regularly scheduled, centralised "checkpoints" have been implemented during the data management cycle to help ensure accurate translation of programming requests.
- *Quality Assurance Standards:* Standard review procedures have been developed and are applied throughout the project life cycle.
- *Automation:* HealthCore has developed standard definitions of many variables and disease states and has developed programs to apply these standards as needed on projects. These standards help ensure consistency, repeatability, and accuracy for each project.

HealthCore's research team documents study progress and scientific and quality review of all study activities and deliverables (e.g., protocol, data management, data analysis, reports, manuscripts) in an ADIN (Action, Decision, Issue, Notification) log and in a QC Log. The ADIN Log provides documentation of study progress, action items, issues/issue resolution, and notifications and is updated weekly during internal project team meetings. In addition, the QC Log documents the quality-control measures performed for each study activity during the conduct of the study.

All programming required for study database extraction and creation of the analytic data sets from HIRD will be performed in accordance with HealthCore Programming Standards. HealthCore Programming Standards is a set of documents describing data extraction methods that are referenced in HealthCore SOPs and provide a guideline for basic, frequently used terms and definitions and respective coding information to maintain operational consistency. Data validation will occur throughout the data management and analysis process. Data quality checks include, but are not limited to, programming checks by an individual who is not the main programmer for the study, internal data set consistency, and checks to ensure that protocol criteria were met. If validation checks are not satisfied, then an examination of the problem will be performed on the data set or data sets in question and the problem will be resolved.

13. REFERENCES

R15-0905 Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. Stat Med. 2014 Mar 30;33(7):1242-58.

R18-1561 Brenner H, Gefeller O. Use of the positive predictive value to correct for disease misclassification in epidemiologic studies. Am J Epidemiol. 1993 Dec 1;138(11):1007-15.

R12-1913 Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. Am J Epidemiol. 2006 Jun 15;163(12):1149-56.

R15-3136 Denburg MR, Haynes K, Shults J, Lewis JD, Leonard MB. Validation of The Health Improvement Network (THIN) database for epidemiologic studies of chronic kidney disease. Pharmacoepidemiol Drug Saf 2011; 20 (11): 1138-1149.

R16-4917 Dobson AJ, Kuulasmaa K, Eberle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. Stat Med. 1991 Mar;10(3):457-62.

R15-3138 Fleet JL, Dixon SN, Shariff SZ, Quinn RR, Nash DM, Harel Z, et al. Detecting chronic kidney disease in population-based administrative databases using an algorithm of hospital encounter and physician claim codes. BMC Nephrol 2013; 14: 81.

R16-4007 Girman CJ, Gokhale M, Kou TD, Brodovicz KG, Wyss R, Stürmer T. Assessing the impact of propensity score estimation and implementation on covariate balance and confounding control within and across important subgroups in comparative effectiveness research. Med Care. 2014 Mar;52(3):280-7.

R16-4915 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods. 2010 Sep;15(3):234-249.

R13-3433 Holden SH, Barnett AH, Peters JR, Jenkins-Jones S, Poole CD, Morgan CL, et al. The incidence of type 2 diabetes in the United Kingdom from 1991 to 2010. Diabetes Obes Metab. 2013 Sep;15(9):844-52.

R13-4387 Kidney Disease: Improving Global Outcomes (KDIGO) Blood Pressure Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int Suppl 2013; 3 (1), 1-150.

R14-5373 Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Dordrecht; New York: Springer; 2009.

R16-4914 Lee BK, Lessler J, Stuart EA. Weight trimming and propensity score weighting. PLoS One. 2011 Mar 31;6(3):e18174.

R15-5270 Levey AS, Becker C, Inker LA. Glomerular filtration rate and albuminuria for detection and staging of acute and chronic kidney disease in adults: a systematic review. *JAMA* 2015; 313(8):837-846.

R12-1392 Levey AS, Stevens LA, Schmid CH, Zhang Y, Castro AF, Feldman HI, Kusek JW, Eggers P, Lente F van, Greene T, Coresh J, Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI). A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; 150 (9): 604-612.

R18-0269 Mor A, Dekkers OM, Nielsen JS, Beck-Nielsen H, Sørensen HT, Thomsen RW. Impact of glycemic control on risk of infections in patients with type 2 diabetes: a population-based cohort study. Am J Epidemiol. 2017 Jul 15;186(2):227-36.

R16-4920 Parsons LS. Performing a 1:N case-control match on propensity score. Proceedings from SAS Users Group International (SUGI) 29 Conference; Paper 165-29; 2004.

P12-08895 Rassen JA, Glynn RJ, Rothman KJ, Setoguchi S, Schneeweiss S. Applying propensity scores estimated in a full cohort to adjust for confounding in subgroup analyses. Pharmacoepidemiol Drug Saf. 2012 Jul;21(7):697-709.

R18-1213 Roberto G, Leal I, Sattar N, Loomis AK, Avillach P, Egger P, et al. Identifying cases of type 2 diabetes in heterogeneous data sources: strategy from the EMIF Project. PLoS One. 2016 Aug 31; 11 (8): e0160648.

R19-2327 Sahai H, Khurshid A. Statistics in epidemiology: methods, techniques, and applications. Boca Raton, Florida: CRC Press LLC; 1996.

R18-3667 Solberg LI, Engebretson KI, Sperl-Hillen JM, Hroscikoski MC, O'Connor PJ. Are claims data accurate enough to identify patients for performance measures or quality improvement? The case of diabetes, heart disease, and depression. Am J Med Qual 2006;21:238-245.

R18-3576 Southern DA, Roberts B, Edwards A, Dean S, Norton P, Svenson LW, et al. Validity of administrative data claim-based methods for identifying individuals with diabetes at a population level. Can J Public Health 2010;101:61-64.

R18-3569 Spittal MJ, Pirkis J, Gurrin LC. Meta-analysis of incidence rate data in the presence of zero events. BMC Med Res Methodol. 2015 Apr 30;15:42.

R16-4919 Wyss R, Girman CJ, LoCasale RJ, Brookhart AM, Stürmer T. Variable selection for propensity score models when estimating treatment effects on multiple outcomes: a simulation study. Pharmacoepidemiol Drug Saf. 2013 Jan;22(1):77-85.

R16-3197 Zhong VW, Pfaff ER, Beavers DP, Thomas J, Jaacks LM, Bowlby DA, et al. Use of administrative and electronic health record data for development of automated algorithms for childhood diabetes case ascertainment and type classification: the SEARCH for Diabetes in Youth Study. Pediatr Diabetes 2014; 15: 573-84.

Choi, J., Dekkers, O. M., & le Cessie, S. (2019). A comparison of different methods to handle missing data in the context of propensity score analysis. European Journal of Epidemiology, 34(1), 23–36. https://doi.org/10.1007/s10654-018-0447-z

14. Appendices

ANNEX 1. CODES TO DEFINE DIABETES AND EXCLUSION CRITERIA

See separate stand-alone document.

ANNEX 2. CODES TO DEFINE STUDY OUTCOMES

See separate stand-alone document.

ANNEX 3. CODES TO DEFINE STUDY COVARIABLES

See separate stand-alone document.

ANNEX 4. VALIDATION PLAN

See separate stand-alone document.

ANNEX 5. OPERATIONAL DEFINITIONS FOR SELECTED VARIABLES

Pregnancy in Denmark:

Normal duration of pregnancy: 37-42 gestational weeks (259-294 days).
However, the first 14 days cover a period from last ovulation until pregnancy, thus, the true pregnant period: 244-280 days.
Regarding abortion and definitions:
Prior to gestational week 22+0: abortion
After gestational week 22+0: stillborn
Among females aged 15 to 50 years (both inclusive):
For codes flagged as 1 in the excel (codes for labour/delivery): end_preg=date of diagnosis (in-date). start_preg = end_preg – 280 days
For codes flagged as 2 in the excel (codes for preterm delivery, i.e. prior to gestational week 37): end_date=date of diagnosis (in-date). Start-preg = end_preg – 244 days
For codes flagged as 3 in the excel (codes for postterm delivery, i.e. after gestational week 42): end_date=date of diagnosis (in-date). Start-preg = end_preg – 294 days
For codes flagged as 4 in the excel (codes for abortion, either without specification of period or specified after gestational week 12): end_date=date of diagnosis (in-date). Start-preg = end_preg – 140 days
For codes flagged as 5 in the excel (codes for abortion, specified prior to gestational week 12): end_date=date of diagnosis (in-date). Start-preg = end_preg – 70 days
For codes flagged as 6 in the excel (codes for conditions related to pregnancy but without a specification regarding period during pregnancy):
For these codes, we will apply a period of 40 weeks = 280 days around the code:
end_date=date of diagnosis (in-date) + 280 days and start_date = (out-date) - 280 days
Thus, for codes flagged as 6 in the excel the potential pregnancy period will be up to 560 days. We will therefore apply a ranking system.
Ranking:
The definitions above, will lead to a number of pregnancy periods. A given person may be pregnant more than once. Pregnancy periods identified with different codes may be overlapping, thus we will apply a ranking system for overlapping pregnancy periods in order to determine, which pregnancy period that will be used to identify pregnancy as exclusion and censoring criteria.
In theory, codes flagged as 4/5 (abortion) should not have overlapping pregnancy periods with codes flagged as 1/2/3 (delivery).
But in ~10% of episodes there are overlap.
Also, in order to determine the best ranking algorithm, we have looked into the proportion of codes flagged as 6 that are not followed by either:
i) a code flagged as 1/2/3 within 280 days after
ii) a code flagged as 4/5 within 140 days after
And this proportion is overall small (less than 4%), however larger in the latest part of the study period (i.e. pregnant women who has not given birth yet)
In case of overlapping pregnancy episodes:
Rank 1: Codes flagged as 2 and 3 should be ranked 1 (assuming that there will be no overlapping pregnancy periods identified with codes flagged as 2 and 3 codes flagged with 1 are not specific for delivery within normal pregnancy range, thus pre/postterm delivery codes are ranked higher
Rank 2: Codes flagged as 1, 4, 5
Rank 3: Codes flagged as 6

Of note, we must bear in mind that in the latest part of the study period more patients will be coded with 6 than either 1-5 because all the pregnancies aren't followed (up to) 40 weeks. Thus for the latest part of the study period the pregnancy episodes will be expected to last longer time (up to 280+280 days)

If pregnancy episodes do not overlap, all episodes will count as individual pregnancy-episodes.

Pregnancy in CPRD:

PREGNANCY ALGORTIM IN CPRD GOLD

Identification of pregnancy start and end date based on entity types and codes. This algorithm is based on the paper by Minassian et al 2019 PDS.

Process:

1. 1st Identify end pregnancy date
2. 2nd Identify start pregnancy date for some when end pregnancy date is not possible
3. 3rd Create pregnancy period: if both available for a patient use those, if one or the other missing then:
    - If only end_preg available then start_preg=end_preg+294 days
    - If only start_preg available then end_preg=start_preg-294 days
4. 4th Identify if index date is within the pregnancy period.
5. 5th Among patients for whom we have not been able to identify start_preg or end_preg, first list or tab to see N, then evaluate the need to do patients profile review of those.
6. 6th use Read codes and ICD10 OPCS4 codes

TABLE 2 Entity codes used to derive pregnancy start and end dates

Entity Code	Description	Estimation of Pregnancy Date
To estimate the start of pregnancy		
60 61 154	Ante-natal booking Ante-natal consultation Alpha fetoprotein	Event date minus the number of weeks specified in the relevant data field[a] (allowing a maximum of 42 wk).
119 120	Gestation—maternity outcome Gestational age of baby	Estimated date of delivery (derived by the algorithm) minus the number of weeks specified in data 1 (allowing a maximum of 42 wk).
129	Pregnancy dates	Expected delivery date (in data 2) minus 280 d.
284	Maternity ultra sound scan	Expected delivery date (in data 8) minus 280 d. If data 8 is missing, use event date minus the number of weeks specified in data 2 (allowing a maximum of 42 wk).
To estimate the end of pregnancy		
35 80 84	Hearing (6 wk) Muscle tone for 6 wk (CHS) Vision CHS 6 wk	Event date minus 42 d.
63[b]	CHS examination	If CHS stage (in data 2) = birth, use event date. If CHS stage = 6 wk, use event date minus 42 d.
69 150	Postnatal examination Postnatal visit	Event date minus the number of days or weeks specified in data 2 (allowing a maximum of 56 d or 8 wk).
78 93 112 115 119 120 126 128 144 145	Stages of labour Delivery details CHS Apgar score at 1 min Delivery details (CHS) Gestation—maternity outcome Gestational age of baby Maternity infant details Perineum Maternity outcome placenta CHS Apgar score at 5 min	Event date
100	Perinatal problems	Event date minus 7 d.
114[c]	Pregnancy outcome	Discharge date (in data 1) minus 2 d.

Abbreviation: CHS, Child Health Surveillance.

[a]Data 1 (entity codes 60 and 61), data 8 (entity code 154).

[b]Used only if CHS stage specifies birth or 6 weeks.

[c]Used to estimate either a delivery date or an early pregnancy loss date (depending on the outcome specified in the record). All other entity codes are used to estimate delivery dates only.

USING READ CODES

- Based on the review of the crosstab between read codes and entity types, for some "easy" read codes associated with entity types that are not in the list above we propose to follow the following algorithms:

- In the tab Pregnancy Read from the excel 4366XXX, for read codes identified in the column classification=1, end_preg=eventdate, and =evendate-294days

- In the tab Pregnancy Read from the excel 4366XXX, for read codes identified in the column classification=2, end_preg=eventdate, and start_preg=evendate-140days

- In the tab Pregnancy Read from the excel 4366XXX, for read codes identified in the column classification=0,3,4,5,6, we do not apply this algorithm

USING ICD10 and OPCS4 codes in HES

- In the tab Pregnancy ICD-10 OPCS4 from the excel 4366XXX, for codes identified in the column classification=1, end_preg=epistart, and start_preg= epistart-294days (or ovdate for procedures)

Among females with age between 18 and 50:

1st for codes flagged as 1 in the excel (these will be codes for delivery, labour, birth….),: end_preg=EV, and START_PREG=EV-280 days (rank =1)

2nd for all the rest of codes, if they have maternity file and AV is not missing, then END_PREG=AV+196 days, and START_PREG=AV-84 days (rank =2)

3rd If they do not have maternity file or AV is missing, for those codes flagged with a 2, use formula in column "formula in CPRD", e.g. END_PREG=EV, START_PREG=EV-140 (rank =3)

4th Else, if they do not have maternity file or AV is missing, for those codes flagged with 3 (rank =4):

- **In Aurum***: do profile (for exclusion).
- In GOLD: use entity types as:
    - END DATE :
        - If entype=35, 80, 84, end_preg=(eventdate-42 days)
        - If entype=63, if data2=1 then end_preg=eventdate, if data2=2, then end_preg=eventdate-42days
        - If entype=69, end_preg=eventdate-7*data2 (if data2≤56days)
        - If enttype = 150, end_preg=eventdate- data2 (if data2 ≤56days)
        - If entype=78, 93, 112, 115, 119, 120, 126, 128, 144, 145, end_preg=eventdate
        - If entype=129, end_preg= data2
        - If entype=100, end_preg=(eventdate-7days)
        - If entype=114, end_preg= (data1-2 days)
- START DATE
    - For some entity type end date cannot be estimated, so we can apply an algorithm using estimated start_preg+42 weeks (that is 294 days). Start date will be as follows:

- If enttype=60 or 61, start_preg=(eventdate -data1*7)
- If enttype=154, start_preg=(eventdate- data8*7)
- If enttype=284, start_preg=(data8-280days), if data8=missing, start_preg=(eventdate- data2*7) (if data2 =≤29480day s)

- Some patients may have a record for an enttype 60, or 61, several times and the derived start_preg and end_preg vary. In these cases, the patient will be considered pregnant if cohort entry date is within start and end of pregnancy for any of the dates for the same patient (worst case scenario)

5th In GOLD, if code flagged as 3 and entity type is not under any of those indicated above, then do profile (rank=5)

If, for a potential index date, the patient has different rows with pregnancy details (with possibly conflicting dates or missing dates), the row with lowest rank will be considered. If more than 1 row with same rank and different dates, she will be considered pregnant if for any of the dates the patient is found pregnant at the potential index date.

For the follow-up, no patient profile review will be done, if start or end cannot be identified for a pregnancy code, the date of the observation will be used as censoring date.

DEFINE PREGNANCY FOR COVARIATE DEFINITION, if start and end of pregnancy is missing (rank5 GOLD, rank 4 Aurum) will be considered as pregnant at index date if the patient has at least 1 code for pregnancy in the 294 days before index date and at least 1 code in the 294 days after index date. Otherwise the patient will be considered as not pregnant at index date.

Programmatically deleted code :"*9NZ..00 Patient encounter data NOS*", MedCoID *15291000000114*. Under pregnancy codes but not related to pregnancy as shown in patient profiles.

15. HISTORY TABLE

Version No.	Date (dd Mmm yyyy)	Author	Sections changed	Brief description of change