Boehringer Ingelheim

Protocol for observational studies based on existing data

Document Number:	**c03270726-04**
BI Study Number:	1245.96
BI Investigational Product(s):	Jardiance® (empagliflozin) Synjardy® (empagliflozin/metformin)
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, genital infections, and diabetic ketoacidosis among patients treated with empagliflozin compared to patients treated with other SGLT2 inhibitors or DPP-4 inhibitors
Brief lay title:	Post-authorisation safety study in patients with type 2 diabetes to assess the risk of liver injury, kidney injury, urinary tract and genital infections, and diabetic ketoacidosis in patients treated with empagliflozin, compared to other SGLT2 inhibitors or DPP-4 inhibitors
Protocol version identifier:	4.0
Date of last version of protocol:	21 Oct 2015
PASS:	Yes
EU PAS Register number:	ENCEPP/SDPP/13413; http://www.encepp.eu/encepp/viewResource.htm?id=13414
Active substance:	A10BX12 Empagliflozin A10BD20 Empagliflozin/metformin
Medicinal product:	Jardiance Synjardy
Product reference:	EMEA/H/C/002677 EMEA/H/C/003770

Procedure number:	EMEA/H/C/002677/MEA/002
Joint PASS:	No
Research question and objectives:	To estimate, among patients with type 2 diabetes mellitus (T2D), the risk of acute liver injury, the risk of acute kidney injury and chronic kidney disease, the risk of severe complications of urinary tract infections, the risk of genital infections, and the risk of diabetic ketoacidosis among patients treated with empagliflozin compared with patients treated with other sodium-glucose cotransporter 2 (SGLT2) inhibitors and compared with patients treated with dipeptidyl peptidase-4 (DPP-4) inhibitors
Country(-ies) of study:	United Kingdom
Authors:	Manel Pladevall-Vila, Cristina Rebordosa [redacted]
Marketing authorisation holder(s):	Boehringer Ingelheim International GmbH [redacted]
MAH contact person:	[redacted]
EU-QPPV:	[redacted]
Signature of EU-QPPV:	The signature of the EU-QPPV is provided electronically
Date:	10 Jun 2016

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

ADQI	Acute Dialysis Quality Initiative
AKI	Acute Kidney Injury
ALI	Acute Liver Injury
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Transaminase
ATC	Anatomical Therapeutic Chemical
BI	Boehringer Ingelheim International GmbH
CI	Confidence Interval
CKD	Chronic Kidney Disease
CKD-EPI	CKD Epidemiology Collaboration
CPR	Central Personal Registration (number), Denmark
CPRD	Clinical Practice Research Datalink
CPT	Current Procedural Terminology (codes)
DKA	Diabetic Ketoacidosis
DPP-4	Dipeptidyl Peptidase-4
ED	Emergency Department
eGFR	Estimated Glomerular Filtration Rate
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
FDA	United States Food and Drug Administration
GLDs	Glucose-Lowering Drugs
GLP-1	Glucagon-Like Peptide-1
GP	General Practitioner
HCPCS	Healthcare Common Procedure Coding System
HES	Hospital Episode Statistics
IACS	Instituto Aragonés de Ciencias de la Salud (Aragón Institute of Health Sciences) in Aragon, Spain

ICD-10	International Statistical Classification of Diseases and Related Health Problems, 10th Revision
ICD-9	International Classification of Diseases, 9th Revision
ICD-9-CM	International Classification of Diseases, 9th Revision, Clinical Modification
IRB	Institutional Review Board
IRR	Incidence Rate Ratio
ISAC	Independent Scientific Advisory Committee
MedDRA	Medical Dictionary for Drug Regulatory Activities
NICE	National Institute for Health and Care Excellence
OQA	Office of Quality Assurance (RTI Health Solutions)
PaCO2	Partial Pressure of Carbon Dioxide
PASS	Post-Authorisation Safety Study
RR	Relative Risk
RTI-HS	RTI Health Solutions
SGLT2	Sodium-Glucose Cotransporter 2
SIDIAP	Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària (Information System for the Advancement of Research in Primary Care), Catalonia, Spain
SIRS	Systemic Inflammatory Response Syndrome
STD	Sexually Transmitted Disease
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
T1D	Type 1 Diabetes Mellitus
T2D	Type 2 Diabetes Mellitus
THIN	The Health Improvement Network
UK	United Kingdom
UKPDS	UK Prospective Diabetes Study
ULN	Upper Limit of the Normal Range
US	United States
UTI	Urinary Tract Infection

3. RESPONSIBLE PARTIES

The study investigators at RTI Health Solutions (RTI-HS) share responsibility with Boehringer Ingelheim International GmbH (BI) for the design of the study. The investigators are responsible for conducting the study in a manner that meets regulatory standards, conducting analyses, and preparing scientific reports. The study shall be conducted as described in the approved protocol. The authors will not develop or implement any deviation or change to the protocol without prior review by BI.
The financial sponsor of this study is BI. The sponsor is responsible to assure study progress. BI is also responsible for communicating with the European Medicines Agency (EMA) (called 'the agency') about the study protocol, the progress of the study, and study results.

4. ABSTRACT

Name of company: Boehringer Ingelheim			
Name of finished medicinal product: Jardiance Synjardy			
Name of active ingredient: A10BX12 Empagliflozin A10BD20 Empagliflozin/metformin			
Protocol date: 05 Feb 2015	**Study number:** 1245.96	**Version/Revision:** 4.0	**Version/Revision date:** 10 Jun 2016
Title of study:	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, genital infections, and diabetic ketoacidosis among patients treated with empagliflozin compared to patients treated with other SGLT2 inhibitors or DPP-4 inhibitors		
Rationale and back-ground:	As part of the risk management plan, Boehringer Ingelheim International GmbH (BI) has committed to submit a protocol for a post-authorisation safety study (PASS) to evaluate the liver and renal safety of empagliflozin. The study will also evaluate the risks of severe complications of urinary tract infections and genital infections. This updated version of the protocol adds diabetic ketoacidosis (DKA) as an additional safety topic in line with BI's commitment within the Article 20 referral, started by the European Medicines Agency (EMA) in June 2015 with the aim to evaluate the risk of DKA from sodium-glucose cotransporter 2 (SGLT2) inhibitors.		
Research question and objectives:	To estimate, among patients with type 2 diabetes mellitus (T2D), the risk of acute liver injury, the risk of acute kidney injury and chronic kidney disease, the risk of severe complications of urinary tract infections, the risk of genital infections, and the risk of diabetic ketoacidosis among patients treated with empagliflozin compared with patients treated with other SGLT2 inhibitors and compared with patients treated with dipeptidyl peptidase-4 (DPP-4) inhibitors.		
Study design:	This will be an observational cohort study using existing data (records from routine medical care). The study will use a "new users" design and compare new users of empagliflozin to new users of SGLT2 inhibitors other than empagliflozin and to new users of DPP-4 inhibitors. Propensity scores based on information prior to the index date will be used to account for potential confounding. The index date will be defined as the date on which each identified new user receives the index prescription for empagliflozin, another SGLT2 inhibitor, or a DPP-4 inhibitor.		
Population:	The study population will include all eligible patients with T2D initiating treatment with empagliflozin, with another SGLT2 inhibitor, or with a DPP-4 inhibitor. Eligible patients will be included if they are aged 18 or more years and have at least 12 months of continuous registration in the Clinical Practice Research Datalink		

	(CPRD) in the United Kingdom (UK). Each member of the empagliflozin-exposed population must also have at least one prescription for empagliflozin, with or without other glucose-lowering drugs (GLDs), and no prior prescriptions of empagliflozin, other SGLT2 inhibitors, or DPP-4 inhibitors during the available pre-index period. Each member of the population exposed to SGLT2 inhibitors other than empagliflozin must have at least one prescription for an SGLT2 inhibitor other than empagliflozin, with or without other GLDs, and no prior prescriptions of these other SGLT2 inhibitors, empagliflozin, or DPP-4 inhibitors during the available pre-index period. Each member of the population exposed to a DPP-4 inhibitor must have at least one prescription for a DPP-4 inhibitor, with or without other GLDs, and no prior prescriptions of a DPP-4 inhibitor, empagliflozin, or other SGLT2 inhibitor during the available pre-index period. The analyses will be stratified by treatment complexity (mono vs. dual vs. triple therapy), and by calendar time of treatment initiation. Prior or current use of insulin is allowed; an additional stratified analysis by insulin use is planned. Only patients with T2D will be included in the study. Patients with type 1 diabetes will be excluded. Different exclusion criteria will be applied according to each of the five outcomes of interest, which will result in five different cohorts (see Section 9.2.7.1). Follow-up will start the day after the index date and, for each specified outcome, will continue until the occurrence of the study outcome, end of study data (date of death, date of transfer out of the practice, date of end of registration, date of study end), the date during follow-up that specific exclusion criteria are met, the end date of the first continuous treatment of the index drug (empagliflozin or other SGLT2 inhibitor or DPP-4 inhibitor) plus a defined grace period (30 days after the end of the last prescription's days' supply in main analyses), or the date in which a new treatment episode starts with any of the other index drugs.
Variables:	*Primary Outcomes*: • Acute liver injury • Acute kidney injury • Hospitalisation due to severe complications of urinary tract infection (UTI) • Genital infections • Hospitalisation due to diabetic ketoacidosis *Secondary Outcomes*: • Acute liver injury in a subset of patients without predisposing factors • Chronic kidney disease • Outpatient severe complications of UTI • Genital infections resulting in hospitalisation or systemic treatment Validation of identified cases will be implemented for all five primary outcomes. If the number of cases identified for any of the five outcomes is 100 or fewer, all cases will be validated for those outcomes. Otherwise, a random sample of 100 cases will be selected for validation for each outcome with more than 100 cases identified. Cases identified in the CPRD and Hospital Episode Statistics (HES) will be validated by review of computerised clinical information and general

	practitioner (GP) questionnaires. For the individual outcomes where the initial algorithm does not demonstrate sufficient predictive value, after algorithm refinement up to 100 additional cases (if available) will be validated. *Exposures (index drugs)*: • Empagliflozin (and fixed-dose combinations with metformin) • Other SGLT2 inhibitors: dapagliflozin, canagliflozin (and fixed-dose combinations of these drugs with metformin) • DPP-4 inhibitors: sitagliptin, saxagliptin, linagliptin, vildagliptin, alogliptin (and fixed-dose combinations of these drugs with metformin) Current use of the index drugs will be defined from the date of prescription of empagliflozin, other SGLT2 inhibitor, or DPP-4 inhibitor to the end of supply for that prescription plus a period of 30 days. Recent use will be defined from the end of current use (30 days after end of supply) and ends 90 days later (which is 120 days after end of supply). Discontinuation of the index drug will be defined as no further prescription 120 or more days after the end of the last prescription's days' supply. Patients who discontinue an index drug and then subsequently initiate another index drug will not be allowed to re-enter the study.
Data sources:	The UK CPRD and HES are proposed as the data sources for the present study. In the UK, nearly all residents are registered in a general medical practice that uses electronic medical records. The CPRD contains diagnostic and prescribing information recorded by general practitioners as part of their routine clinical practice in the UK. The database currently contains data for over 13.2 million patients with research-quality data from 680 UK practices; 5.69 million of these patients are active (still registered with a general practice that is contributing data to the CPRD). Patients registered are representative of the whole UK population in terms of age and sex. Danish and Spanish databases in Europe and a commercial data source in the United States (US) are proposed as additional data sources in case an insufficient number of users and events are available in the CPRD for rare events such as acute liver injury. These proposed databases have not been yet contacted to determine their interest in study participation. Detailed information on prescriptions written by the general practitioners, including prescribed dose and duration, is routinely recorded in the database. Additional diagnostic and treatment information can be found in letters from specialists, hospitals, and other sources. Because general practitioners (GPs) serve as the gatekeepers for all medical services, any visit to a specialist or hospital requires communication back to the GP, who enters that information into the medical record. Identifying patients with both CPRD and HES data enables access to hospital discharge diagnosis and procedural coding. Linkage to the HES is currently limited to approximately 40% of study patients. Cause-of-death information is available for the same set of practices. If patient numbers accumulating in the CPRD are below the expected counts or an unexpectedly low event rate indicates insufficient study power for certain rare outcomes (e.g., liver injury), additional data sources or countries may be added (see Section 9.6).
Study size:	The study size will be driven by the uptake of empagliflozin following the approval and launch of empagliflozin for the treatment of T2D to improve glycaemic control in adults in the UK. The study size required to detect an incidence rate ratio (IRR) of 3 is about 35,000 person-years of empagliflozin use for liver injury and about

	4,000 person-years of empagliflozin use for kidney injury. It is expected that by the end of 2018, approximately 5,000-6,000 empagliflozin-treated patients will accumulate in the CPRD database, and by 2020 this number will reach approximately 9,000; however, the number of patients available for analysis may be lower. The number of empagliflozin users and crude incidence rates of the events of interest will be assessed annually starting June 2016, although only the number of users will be assessed in this first interim report. If the number of events for certain outcomes is not adequate to answer the scientific questions of interest, the addition of data sources from other countries (e.g., Denmark, Spain, and/or the US) will be considered. The study period may be prolonged to up to 5 years after launch.
Data analysis:	The following estimates and comparisons will be generated: • Crude and adjusted incidence rates of each of the outcomes among empagliflozin new users, among other SGLT2 inhibitor new users, and among DPP-4 inhibitor new users. Incidence rates will be reported as point estimates (in cases per 1,000 person-years) and 95% confidence intervals (CIs). • Summary IRRs, after adjusting (via stratified Mantel-Haenszel analysis) for propensity score deciles and calendar year, among empagliflozin new users versus other SGLT2 inhibitor new users and versus DPP-4 inhibitor new users. The adjusted IRRs for each of the five outcomes will be the main effect estimates of interest. Adjusted incidence rates and IRRs will be calculated by weighting the incidence or IRR of each stratum (defined by propensity score decile and calendar year) by the amount of empagliflozin person-time within the stratum. An additional analysis will further stratify the IRRs by categories of insulin use at the index date. Additional sensitivity analyses will be performed to evaluate potential for bias and confounding.
Milestones:	Launch of empagliflozin in the UK occurred on 01 August 2014. Protocol version 1.0 was submitted to the EMA for review in February 2015, version 2.0 in June 2015, and version 3.0 in October 2015. The EMA approved protocol version 3.0 on 17 December 2015. The study start will depend on the dates of protocol approval and market uptake and will take into account the 3-month lag time needed to access the CPRD data. The number of users in each treatment cohort will be monitored and reported annually in three interim reports. The three interim reports will include data up to 19, 24, and 36 months after empagliflozin use starts being captured in CPRD and are expected to be produced June each year. Crude incidence rates of acute liver and kidney injury outcomes (overall, not stratified by treatment) will be generated in all interim reports except the first one. Based on the available patient numbers and the event rates observed in CPRD, a decision will be made to proceed with adjusted, treatment-stratified analyses for none, all, or only certain (more frequent) study outcomes at the interim report stage. Use of other country data sources (Denmark, Spain, and/or the US) will be considered during the course of the study for certain (more rare) study outcomes if patient numbers and/or event rates observed at the interim reports indicate that an analysis in the CPRD may not be feasible. Final report will be produced using data through 36 months from the start of use of empagliflozin in the UK as reflected in the database (currently planned study period

	August 2014–August 2017). Taking into account the database lag, the analysis will be performed in January 2018 and final report will be produced by July 2018. Based on the available number of patients, the study period could be extended to August 2019, in which case final report would be expected in July 2020.

5. AMENDMENTS AND UPDATES

Protocol version 4.0 is an amendment of version 3.0 to add diabetic ketoacidosis as an additional safety endpoint.

Version Number	Date	Section of study protocol	Amendment or update	Reason
4.0	27 May 2016	Cover page and abstract	Addition of Synjardy (empagliflozin/ metformin)	Synjardy came on the market in August 2013, after the approval of protocol version 3.0
4.0	27 May 2016	Section 7.1	Added data on diabetic ketoacidosis in clinical trials and safety studies	New safety outcome added to the study
4.0	27 May 2016	Section 7.7	Added data on the epidemiology of diabetic ketoacidosis	New safety outcome added to the study
4.0	27 May 2016	Section 8	Added the evaluation of the risk of diabetic ketoacidosis as a primary objective	New safety outcome added to the study
4.0	27 May 2016	Section 9.3.2 and Annex 5	Added definition of the diabetic ketoacidosis outcome	New safety outcome added to the study
4.0	27 May 2016	Section 9.6	Changed chi-square sample size calculation method for Poisson method	The Poisson method is considered more appropriate for rare events, and confidence intervals are also going to be estimated using Poisson method
4.0	27 May 2016	Annex 6	Added covariates to be considered for inclusion in the propensity score model for diabetic ketoacidosis	New safety outcome added to the study

6. MILESTONES

Milestone	Planned Date
Protocol endorsed by the EMA	December 2015
Start of data collection	15 March 2016
End of data collection	31 December 2017
First interim report	Will be based on data available 19 months after use of empagliflozin is first captured in the CPRD Report expected in June 2016
Second interim report	Will be based on data available 24 months after use of empagliflozin is first captured in the CPRD Report expected in June 2017
Third interim report	Will be based on data available 36 months after use of empagliflozin is first captured in the CPRD Report expected in June 2018
Registration in the EU PAS Register1	10 May 2016
Final report of study results	Expected in July 2018. In case of insufficient number of patients / events, study period may be expanded: final report then would be expected July 2020

CPRD = Clinical Practice Research Datalink; EMA = European Medicines Agency.

Note: Approvals by data protection, data custodian, ethics, and scientific review bodies are pending. Timelines may be impacted by approvals of these bodies, duration of contract reviews, and availability of data and staff at research institutions once contracts and approvals are finalised.

1. ENCEPP/SDPP/13413; http://www.encepp.eu/encepp/viewResource.htm?id=13414.

7. RATIONALE AND BACKGROUND

Jardiance (empagliflozin), a highly potent and selective inhibitor of the sodium-glucose cotransporter 2 (SGLT2), was approved in Europe in May 2014 for the treatment of type 2 diabetes mellitus (T2D) to improve glycaemic control in adults. SGLT2 is highly expressed in the kidney; as the predominant glucose transporter, it is responsible for the reabsorption of glucose from the glomerular filtrate back into the circulation. Empagliflozin improves glycaemic control in patients with T2D by reducing renal glucose reabsorption [R14-4617].

The recommended starting dose is 10 mg empagliflozin once daily for monotherapy and add-on combination therapy with other glucose-lowering medicinal products including insulin. In patients tolerating empagliflozin 10 mg once daily who have an estimated glomerular filtration rate (eGFR) ≥ 60 mL/min/1.73 m2 and need tighter glycaemic control, the dose can be increased to 25 mg once daily. The maximum daily dose is 25 mg [R14-4617].

The overall frequency of treatment-emergent adverse events observed in the empagliflozin clinical trials and safety studies was comparable between treatment groups, between 70% and 74% [P14-17456].

As part of the risk management plan, Boehringer Ingelheim International GmbH (BI) has committed to conduct a post-authorisation safety study (PASS) to evaluate the liver and renal safety of empagliflozin as well as the incidence of urinary tract and genital infections.

For the available information on renal and liver safety of empagliflozin, please refer to Section 7.1.

The study will also evaluate the risks of (1) severe complications of urinary tract infections (UTIs) and (2) genital infections. The rationale for looking at these risks is the fact that inhibition of SGLT2 in patients with T2D leads to excess glucose excretion in the urine [R14-4617], which, together with hyperglycaemia, may be the main cause of increased susceptibility of patients with diabetes to urinary and genital infections. Although this mechanism is not completely understood, it is known that increased glucose levels in genitourinary tissues enhance yeast adhesion and growth; thus, by providing a favourable growth environment for otherwise commensal microorganisms, glycosuria could potentially increase the risk for UTIs, vulvovaginitis, and balanitis. Moreover, hyperglycaemia not only impairs various aspects of host defense, including neutrophils and complement proteins, but also promotes the virulence of infecting organisms in patients with diabetes [P14-02878, R14-5237].

In addition, this version of the protocol (version 4.0) adds diabetic ketoacidosis (DKA) as a safety topic of interest. The rationale for assessing this risk is that cases of diabetic ketoacidosis have occurred in patients taking SGLT2 inhibitors for T2D, and a number of these cases have been atypical, with patients not having blood sugar levels as high as expected or even with levels in the normal range [P15-08785]
Atypical DKA reported in patients with T2D treated with SGLT2 has a different origin than that observed in patients with type 1 diabetes mellitus (T1D). Full-dose SGLT2 inhibition

induces a rapid increase in urinary glucose excretion, ranging from 50 to 100 g/day, lasting slightly more than 24 hours. Concomitant insulin intensification therapy was a common factor in most if not all cases. Depending on body size, glomerular filtration rate, and degree of hyperglycaemia, SGLT2-induced glucose loss can make up a substantial fraction of daily carbohydrate availability [P15-08785]. The different pathophysiology of DKA versus atypical DKA induced by SGLT2 inhibitors is that (1) in the latter, insulin deficiency and insulin resistance are milder, with less glucose overproduction and underutilization, which in most cases led to significant reductions in total daily insulin requirements, consequently leading to hypoinsulinaemia, and (2) renal glucose clearance (i.e., the ratio of glycosuria vs. glycaemia) is twice as large in atypical DKA as in DKA. Ketoacidosis follows the same sequence of events in both presentations of DKA. However, in SGLT2-treated patients with T2D, the lower insulin-to-glucagon ratio stimulates lipolysis (increase in free fatty acids) and enhances lipid oxidation at the expense of carbohydrate oxidation. At low glucose concentrations, non-oxidative glucose disposal falls. The augmented free fatty acids delivered to the liver result in mild stimulation of ketogenesis, while fasting and mean post-meal beta-hydroxybutyrate levels increase; conversely, plasma lactate levels decrease as an expression of reduced carbohydrate utilization [P15-08785].

This protocol describes a cohort study, to be conducted among patients with T2D, comparing the incidence of the five outcomes of interest in patients initiating empagliflozin compared to patients initiating other SGLT2 inhibitors and compared to patients initiating a DPP-4 (dipeptidyl peptidase-4) inhibitor. For each outcome, different inclusion/exclusion criteria will be applied, resulting in five slightly different study cohorts to be assessed for the risk of each outcome.

7.1 Data on the Five Outcomes of Interest in Empagliflozin Clinical Trials and Safety Studies

According to the Jardiance assessment report, published 20 March 2014 and including data across empagliflozin clinical trials and safety studies, as of 31 August 2012, there were 3,522 patients in the placebo group, with a total duration of exposure of 2,758.1 patient-years; 3,630 patients in the empagliflozin 10-mg group, with a total duration of exposure of 3,258.2 patient-years; and 4,602 patients in the empagliflozin 25-mg group, with a total duration of exposure of 4448.1 patient-years [P14-17456].

The frequency of hepatic injury (Standard MedDRA Query) was low and similar for all treatment groups (see Table 1) [P14-17456]. However, the number of serious hepatic events was higher in patients treated with empagliflozin than in the placebo group, with 19 of the 22 cases of serious liver enzyme elevations occurring in one of the empagliflozin groups. There was also some imbalance seen for elevations of alanine aminotransferase (ALT) and/or aspartate transaminase (AST) (≥ 5 x the upper limit of the normal range [ULN], ≥ 10 x ULN, and ≥ 20 x ULN), when comparing empagliflozin and placebo users. A total of 7 patients (5 during treatment with empagliflozin, 1 after treatment with empagliflozin, and 1 during the screening period) had laboratory values consistent with Hy's Law of drug-induced liver injury [R14-5256], although none of them were finally qualified as drug-induced liver injury due to plausible alternative hypotheses.

The frequency of decreased renal function was low and similar for all treatment groups (see Table 1) [P14-17456]. The most common adverse event was renal impairment: 0.5% among the placebo group and 0.7% each among the empagliflozin 10-mg and 25-mg groups.

The frequency of UTIs was similar for all treatment groups (see Table 1) [P14-17456]. The incidence rate of UTI was 10.9 per 100 patient-years among the placebo group, 10.5 per 100 patient-years among the empagliflozin 10-mg group, and 9.6 per 100 patient-years among the empagliflozin 25-mg group.

The frequency of genital infections was consistently higher among empagliflozin groups, compared to placebo (see Table 1) [P14-17456]. The proportion of women with genital infections was 2-fold higher among women than among men. Empagliflozin groups had a 4-fold higher rate of genital infections than comparator groups irrespective of sex [P14-17456].
In a retrospective analysis of randomized phase 2 and 3 empagliflozin trials (13,000 participants with T2D), there were eight events consistent with DKA, with no imbalance observed between patients treated with empagliflozin 10 mg (two events), empagliflozin 25 mg (one event), and placebo (five events). In the cardiovascular outcome trial EMPA-REG, with approximately 7,000 patients, the frequency of reported blinded events of DKA was less than 0.1% [P15-08785].

Table 1. Frequency of adverse events of interest by treatment group in the empagliflozin clinical trials and safety studies – all patients

Adverse event	**Frequency (%) of adverse events by treatment group**		
	Placebo (N = 3,522)	**Empagliflozin 10 mg** (N = 3,630)	**Empagliflozin 25 mg** (N = 4,602)
Hepatic injury	1.5	1.2	1.4
Decreased renal function	1.0	1.1	1.3
Urinary tract infections	8.1	8.9	8.8
Genital infections	1.0	4.4	4.7
Diabetic ketoacidosisa	5	2	1

a Source for diabetic ketoacidosis: Rosenstock and Ferrannini [P15-08785]. The total number of patients in each treatment group does not include the diabetic ketoacidosis outcome, for which these numbers were not reported.
Source for adverse events except for diabetic ketoacidosis: Jardiance, European Public Assessment Report [P14-17456].

7.2 EPIDEMIOLOGY OF ACUTE LIVER INJURY

7.2.1 Epidemiology of acute liver injury in general population

The incidence of acute liver injury (ALI) and drug-induced liver injury in the general population has been poorly investigated. Data from several observational studies suggest that

the annual incidence of drug-induced ALI in the general population ranges from 0.7 cases per 100,000 persons (95% confidence interval [CI], 0.6-0.9) [P03-00488] to 13.9 cases per 100,000 persons [P02-05969].

The study with the lowest incidence was performed in a hospital network surveillance system in Spain between 1993 and 1998 and identified 107 cases of ALI with a total follow-up of 14.6 x 106 person-years [P03-00488]. The study with the highest incidence was conducted in France using intensive surveillance of cases in a well-defined geographic region between 1997 and 2000 and identified 34 cases of ALI [P02-05969]. Other cohort studies reported an incidence rate (per 100,000 persons) of drug-induced ALI of 2.3 (n = 77 cases) in a hospital outpatient hepatology clinic in Sweden between 1995 and 2005 [P06-11008] and 3.4 (N = 461 cases) in a Spanish regional registry between 1994 and 2004 [P05-08822]. Similarly, in a study conducted in the United Kingdom (UK) Clinical Practice Research Datalink (CPRD) between 1994 and 1999, 128 cases of ALI were identified, and the estimated incidence rate was 2.4 per 100,000 person-years (95% CI, 2.0-2.8) [P04-07683].

7.2.2 Epidemiology of acute liver injury in the diabetes population

Even fewer data are available on the incidence of ALI among patients with diabetes. The incidence rate of ALI among patients with diabetes without risk factors for liver disease in the UK CPRD between 1994 and 1998 was estimated to be 14.2 cases per 100,000 person-years (n = 14 cases) [P03-03701]. Among users of oral glucose-lowering drugs (GLDs) the incidence rate was 22.0 cases per 100,000 person-years (n = 9 cases), and among users of insulin the incidence rate was 13.2 per 100,000 person-years (n = 3 cases). Among patients with diabetes, the adjusted relative risk of ALI among users of GLDs or insulin, compared with non-users of these medications, was 2.8 (95% CI, 0.6-12.5). In the same study, the incidence rate of ALI was estimated to be 8.8 per 100,000 person-years in the general population without diabetes [P03-03701].

In a cohort study performed in the United States (US) Veterans Affairs database between 1985 and 2000, 173,643 hospitalised veterans with type 1 and type 2 diabetes, identified through ICD-9 (*International Classification of Diseases, 9th Revision*) code 250, were matched to 650,620 hospitalised veterans without diabetes. The incidence rate of ALI was 2.31 per 10,000 person-years (n = 346 cases after 1,494,995 person-years of follow-up) among patients with diabetes, and 1.44 per 10,000 person-years (n = 942 cases after 6,556,350 person-years of follow-up) among patients without diabetes, with an adjusted hazard ratio of 1.44 (95% CI, 1.26-1.63). Chronic liver disease and increasing age increased the risk of ALI. Congestive heart failure was more frequent among patients with diabetes and ALI (31%) than patients without diabetes and ALI (22%). The 6-week mortality after hospitalisation with ALI was 60% among patients with diabetes and 63% among patients without diabetes [R12-3632].
For the study size and power calculations (see Section 9.6), the estimates of the ALI incidence in the population with diabetes were used as reported by the two studies above [R12-3632, P03-03701]. Only one of these studies was performed in CPRD; however, as this study used data from 1994-1998 and was in a highly selected population (diabetes patients without risk factors for liver disease), both available studies were considered for the study size calculations.

7.3 EPIDEMIOLOGY OF ACUTE KIDNEY INJURY

Acute kidney injury (AKI) is characterised by an abrupt decline in renal function. The definition of acute renal failure in epidemiologic studies has been based on absolute increases of serum creatinine from normal values (1.7 or 2 x ULN) or change from baseline (20% to 50%) or both [P02-05226, P05-04032, P14-17372, P14-17373, P14-17376, P14-17377].

7.3.1 Epidemiology of acute kidney injury in general population

Two studies have used large administrative and/or claims databases to examine secular trends in the epidemiology of AKI among the general population in the US. The first study was based on a total of 5,403,015 hospital discharges for AKI based on 5% of Medicare claims and reported that the incidence of AKI rose from 14.6 to 36.4 per 1,000 discharges between 1992 and 2001 [R11-5329]. The second study, using the Nationwide Inpatient Sample, analysed a total of 1,083,745 AKI discharges (320,370 of which required dialysis) and reported that AKI incidence rose from 4 to 21 per 1,000 discharges between 1988 and 2002. For AKI requiring dialysis, the percentage of annual discharges increased from 0.3 in 1988 to 2 per 1,000 discharges in 2002. For AKI requiring dialysis, the incidence rose from 4 per 100,000 population in 1988 to 27 per 100,000 population in 2002 [R14-5285].

7.3.2 Epidemiology of acute kidney injury in the diabetes population

Among patients with diabetes, a study performed in the UK CPRD between 2003 and 2007 reported an incidence rate of 198 per 100,000 person-years among patients with T2D and 27 per 100,000 patient-years among those without T2D. The adjusted incidence rate ratio (IRR) of acute kidney injury was 2.5 (95% CI, 2.2-2.7) among patients with T2D compared to patients with no T2D [R11-5319]. Although the manuscript does not provide the incidence of hospitalisation for AKI, it could be estimated to be 65% of the above incidence rates given that this percentage of patients with AKI was identified through a referral or hospitalisation for AKI.

Scarce data have been published on the risk of AKI associated with exposure to antidiabetic medications. A nested case-control study performed in the UK CPRD between 1997 and approximately 2004 (end of study period not reported) reported a relative risk of AKI of 2.5 (95% CI, 0.8-7.8) among patients with current exposure to insulin (n = 5) compared to unexposed patients (n = 98). Similarly, the relative risk of AKI was 1.3 (95% CI, 0.5-3.1) among patients exposed to other antidiabetic drugs (n = 11), compared to unexposed patients (n = 88) [P05-04032].

For the study size and power calculations, the estimates of the AKI incidence in the population with diabetes were used as reported in the UK CPRD study [R11-5319]. In addition, due to sparse data available, estimates reported in general population [R14-5285] have been used as well; see Section 9.6.

7.4 Epidemiology of chronic kidney disease

In a cross-sectional analysis of random samples from the nationally representative Health Survey for England that took place in England in 2003 and in 2009-2010, there were 13,896 adults aged 16 years or older participating. Among 305 participants in 2003 and 322 participants in 2009-2010 with doctor-diagnosed diabetes, the prevalence of chronic kidney disease (CKD) was 17.3% [R15-3134].

In a retrospective, longitudinal study assessing adults with prevalent or incident CKD (identified using estimated glomerular filtration rate readings and/or Read codes) in the CPRD in 2010, the prevalence of stage 3–5 CKD in 2010 was estimated to be 5.9% (n = 165,942), and the prevalence of mildly impaired eGFR was 21.2% (n = 602,437), being the denominator for prevalence calculations the total CPRD population (n = 2,836,476). The prevalence of diabetes was 19.2% among patients with CKD stage 3-5, 13.2 among patients with mildly impaired eGFR and 6.1% in the general population [R15-3139].

The UK Prospective Diabetes Study (UKPDS) [R03-0585] was a clinical trial designed to evaluate the effects of improved blood glucose control and/or blood pressure control on the incidence of complications in patients with hypertension and newly diagnosed type 2 diabetes. The study included 5,102 participants 25-65 years of age (mean, 53 years), 60% male; 4,031 without albuminuria and 5,032 with normal plasma creatinine at diagnosis. All patients were followed until the trial ended in 1997. The renal outcomes assessed were two measures of albuminuria (micro- and macroalbuminuria) and two measures of renal impairment (reduced creatinine clearance and doubling of baseline plasma creatinine). A reduced glomerular filtration rate was defined as an estimated creatinine clearance ≤ 60 mL/min per 1.73 m2. Altogether, 1,544 of 4,031 patients (38%) developed albuminuria and 1,449 of 5,032 (29%) developed renal impairment over a median of 15 years after diagnosis of type 2 diabetes [R12-1479].

In the same UKPDS study, the absolute risk of renal failure was 1.4 events per 1,000 patient-years among patients with tight control of the diabetes, and 2.3 events per 1,000 patient-years among patients with less tight control of diabetes. The absolute risk of death from renal failure was 0.3 event per 1,000 patient-years among patients with tight control of the diabetes, and 1 event per 1,000 patient-years among patients with less tight control of diabetes [R03-0585].

7.5 EPIDEMIOLOGY OF SEVERE COMPLICATIONS OF URINARY TRACT INFECTIONS

Untreated UTIs can lead to acute or chronic kidney infections (pyelonephritis), which could permanently damage the kidneys, and lead to urosepsis [R14-5259, R14-5286]. People with T2D are at higher risk for infections than people without diabetes [R12-1100], and women with T2D have a 2-fold higher risk of culture-confirmed UTIs than women without diabetes [R12-1083, R14-5236, R14-5283].

7.5.1 Epidemiology of severe UTI in general population

In a US hospital-based study of the general population, using the 1997 Healthcare Cost and Utilization Project Nationwide Inpatient Sample Release 6, the incidence of hospital discharges for acute pyelonephritis was estimated to be 11.7 per 10,000 persons-years in females (n = 160,848), and 2.4 per 10,000 persons-years in males (n = 30,718) [R14-5258]. Similarly, in another US study, 4,887 enrolees of Group Health Cooperative, based in Seattle, Washington (US), received a diagnosis of acute pyelonephritis from 1997 through 2001. The annual rates of pyelonephritis among women were 12 to 13 outpatient cases per 10,000 population and 3 to 4 inpatient cases per 10,000 population; among males, the annual rates of pyelonephritis were 2 to 3 outpatient cases per 10,000 population and 1 to 2 inpatient cases per 10,000 population [P14-17370].

7.5.2 Epidemiology of severe UTI in the diabetes population

In a review of hospital charts of pyelonephritis (n = 838 cases, 35 among patients with diabetes) and UTI (n = 976 cases, 11 among patients with diabetes) conducted in Canada between 1991 and 1992, the rate of hospitalisations due to pyelonephritis ranged from 66 to 144 per 10,000 person-years among women with diabetes and from 24.7 to 34.3 per 10,000 person-years among men with diabetes, compared with 6.0 to 11.2 per 10,000 person-years among females without diabetes and 1.7 to 10 per 10,000 person-years among males without diabetes [R14-5251].

Another cohort study performed in Denmark evaluated 10,063 individuals from the Danish general population who were participants in The Copenhagen City Heart Study and estimated the risk of hospitalisation caused by any infectious disease between 1991 and 2000. The study reported a total of 314 hospitalisations due to UTI and an incidence rate for pyelonephritis of 158 per 10,000 person-years among patients with diabetes and 41 per 10,000 person-years among patients from the general population [R12-1080].

A cohort study of the computerised medical database of the University Medical Center Utrecht General Practitioners Research Network selected all patients aged 45 years or older with diabetes between 1995 and 2003; the estimated annual incidence rate of pyelonephritis among patients with diabetes was 3 per 1,000 person-years (n = 16 cases), while the rate of outpatient diagnosis of UTI was 101 per 1,000 person-years (n = 2,000 cases) [R12-1105].

In a more recent study performed in the UK CPRD for the period 1990 to 2007, the incidence of outpatient diagnosis of UTI per 1,000 person-years was 46.9 (95% CI, 45.8-48.1) among patients with diabetes and 29.9 (95% CI, 28.9-30.8) among patients without diabetes. The study reported that, over the 1-year follow-up period, 5,967 UTI events were observed among patients with diabetes, and 3,708 UTI events were observed among patients without diabetes. The relative risk of UTI was 1.46 (95% CI, 1.40-1.53) among newly diagnosed patients with diabetes compared with patients without diabetes. Similarly, the relative risk of UTI was 2.08 (95% CI, 1.93-2.24) among patients with previously diagnosed diabetes compared with patients without diabetes. The risk was higher among females, older patients, patients with previous diagnosis of diabetes, worst level of diabetes control, and a recent history of UTI [R12-5226].

7.5.3 Epidemiology of sepsis in general population and in diabetes patients

The reported incidence rates of sepsis vary by study design and methods. A review of the variations in the incidence and mortality of severe sepsis in the US used four methods/definitions and found that the annual incidence of sepsis ranged from 3 to 10.3 per 1,000 patient-years, depending on the method used [R14-5287].

Among patients with diabetes, the incidence rate of sepsis was evaluated in a cohort study comparing all people with diabetes in Ontario, Canada, between 1999 and 2000, to a matched cohort without diabetes (513,749 patients in each group). The rate of sepsis among the diabetes cohort was 5.4 per 1,000 patient-years (number of cases not reported), and the risk ratio of sepsis versus the non-diabetes cohort was 2.45 (95% CI, 2.23-2.68) [R10-6632].

For the study size and power calculations, the incidence estimates of the acute pyelonephritis in the population with diabetes were drawn from the study in Netherlands [R12-1105]; estimates from the study performed in UK CPRD [R12-5226] could not be used as it had evaluated all forms of UTI including mild and not resulting in hospitalisation. Incidence rate of UTI hospitalisations reported in Denmark [R12-1080] has been used for the power calculations as well; see Section 9.6.

7.6 EPIDEMIOLOGY OF GENITAL INFECTIONS

Individuals with T2D are at higher risk for infections, including genital infections, than patients without diabetes [R10-6632, R12-1080, R12-1088, R12-1100, R12-3639, R14-5248], and T2D is a risk factor for vaginal infections and balanitis [R12-2432, R14-5237, R14-5260]. Moreover, patients with diabetes have up to a 4-fold greater risk of dying due to infectious disease than patients without diabetes [R10-6632, R12-1080]. Women with diabetes have increased rates of asymptomatic vaginal carriage of *Candida* species and increased frequency of symptomatic infections [R12-2432, R14-5254].
While assessing the incidence or prevalence of genital infections using existing databases it is important to remember that the rates may be underestimated, as some of these cases may be self-limiting or may be treated with over-the-counter medications only.

An observational cohort study in the UK evaluated the risk of vaginitis and balanitis among patients with T2D and patients without diabetes in the UK CPRD database using electronic medical records between 1990 and 2007 [R12-3639]. Among females with T2D, there were 1,243 cases of vaginitis: 87.5% recorded as vaginitis/candidiasis, 8.6% vulvitis, 2.2% bacterial vaginosis, and 1.7% related sexually transmitted diseases (STDs). The incidence rate of vaginitis was 21 per 1,000 person-years among patients with T2D and 10.3 per 1,000 person-years among females without diabetes (adjusted relative risk [RR], 1.81; 95% CI, 1.64-2.00). Similarly, there were 592 cases of balanitis among patients with T2D classified as 85.4% balanitis/candidiasis, 8.1% penile/*Candida* infection, and 6.5% other (no STDs). The incidence rate of balanitis was 8.4 per 1,000 person-years among males with T2D and 2.5 per 1,000 person-years among males without diabetes (RR, 2.85; 95% CI, 2.39-3.39). The incidence rate of vaginitis and balanitis decreased with increasing age for males and females with T2D and females without diabetes, but not for males without diabetes (see Table 2). Among patients with T2D, previous history of genital infections predisposed to genital

infections, with an adjusted RR of 6.99 (95% CI, 5.97-8.18) for vaginitis and an adjusted RR of 11.22 (95% CI, 9.02-13.97) for balanitis. Also, worse control of diabetes, assessed by measuring glycated haemoglobin (HbA1c), was associated with an increased risk of genital infections among both treated and untreated patients with T2D (see Table 2).

Table 2. Incidence of vaginitis and balanitis among patients with T2D and patients without diabetes in the UK CPRD, 1990-2007

	Vaginitis (females)		**Balanitis (males)**	
	T2D (N = 62,537, 59,279 patient-years)	**No diabetes** (N = 62,700, 57,844 patient-years)	**T2D** (N = 73,383, 70,088 patient-years)	**No diabetes** (N = 73,220, 67,676 patient-years)
All ages	21 (19.8-22.1)	10.3 (9.5-11.1)	8.4 (7.8-9.1)	2.5 (2.1-2.9)
Age 18-39	~531	~26	~18	~2
Age 40-49	~42	~20	~16	~2
Age 50-59	~24	~11	~9	~2
Age 60-69	~19	~10	~7	~3
Age 70+	~12	~8	~5	~3
Not treated fair control	11.7 (10.0-13.5)		4.6 (3.6-5.6)	
Not treated poor control	38.2 (31.6-44.8)		20.7 (16.6-24.8)	
Treated fair control	15.0 (12.4-17.7)		4.4 (3.1-5.7)	
Not treated poor control	23.9 (20.8-27.0)		11.5 (9.6-13.5)	

Note: Incidence rates (95% CIs) are per 1,000 person-years.

1. Incidence rates with an ~ symbol are approximate estimates derived from the published graph [R12-3639].

Another cohort study conducted in Canada (1999-2000) study compared administrative data from all people with diabetes (n = 513,749) in Ontario with a matched cohort of 513,749 of people without diabetes. The rate of genital infections among male patients with diabetes was 1,340 per 100,000 patients. The rate of genital infections among female patients with diabetes was 234 per 100,000 patients [R10-6632].

For the study size and power calculations, the incidence estimates of the genital infections in men and women with diabetes were drawn from the study performed in UK CPRD [R12-3639], as it is most relevant for the study setting; see Section 9.6.

7.7 EPIDEMIOLOGY OF DIABETIC KETOACIDOSIS

DKA is the most common hyperglycemic emergency in patients with diabetes mellitus, leading to more than 100,000 hospital admissions each year in the US and comprising 4% to 9% of all hospital discharge summaries among patients with diabetes mellitus. In 20% to 30% of the cases of DKA, DKA may be the initial manifestation of diabetes, particularly for T1D [R16-1981]. Although DKA occurs most frequently in patients with T1D, it can also occur in patients with T2D [R15-2053]. The incidence of DKA is difficult to establish, ranging from 4.6 to 8 episodes per 1,000 patients with diabetes in population-based studies from the US [R16-1373]. In England, according to the National Diabetic Audit, the 5-year prevalence of recorded DKA was over 12% in people with T1D and less than 1% in people with T2D. DKA occurred in 3.9% of the people with T1D and 0.48% of all people with diabetes (T1D and T2D) during 2009-2010; the occurrence of DKA in people with T2D was not reported [R16-1374].
The incidence of DKA among patients with T2D is not well established. A review of the medical records of all adult patients admitted to the medical intensive care unit of a medical center in an Arizona (US) identified 226 patients with DKA: 47% had T1D, 26% had T2D, and 27% had DKA as the initial manifestation of diabetes (type not specified) [R16-1375]. A retrospective population-based study performed in Sweden on data from 1997 through 2000 was performed to determine the occurrence of DKA in adult patients with T1D and T2D [R14-3272]. All adult patients with severe hyperglycaemia or suspected DKA admitted to The Umeå University Hospital with a diagnosis code of DKA (ICD-10 codes: E10.0, E10.1, E11.0 and E11.1) were included. The average annual incidence rate for DKA was 5.9 per 100,000 adult inhabitants. The annual incidence rate for DKA in patients with diabetes was 50 per 100,000 adult patients with T2D and 1,585 per 100,000 adult patients with T1D. A total of 25 patients developed DKA: 8 (32%) had T2D, while 17 (68%) had T1D [R14-3272].
In a similar study performed in an Arizona (US) hospital, 226 patients with DKA were identified: 47% had T1D, 26% had T2D, and 27% had DKA as the initial manifestation of diabetes (type not yet specified) [R16-1375].
The incidence of DKA among patients with T2D using SGLT2 inhibitors in the context of routine clinical practice is not well understood. In January 2016, one case of DKA was reported in a patient using empagliflozin as part of routine clinical practice [P15-11541].

7.8 EPIDEMIOLOGY OF DIABETES AND ANTIDIABETIC TREATMENT PATTERNS IN THE UK

The prevalence of diabetes has increased in the UK from 2.8% in 1996 to 4.3% in 2005, and the incidence has increased from 2.7 per 1,000 person-years in 1996 to 4.4 per 1,000 person-years in 2005. During the period 1996-2005, a change in oral GLD use has occurred, predominantly from sulfonylureas to metformin [R11-5320]. Moreover, since 2005-2006, the use of thiazolidinediones has decreased due to concerns about cardiovascular safety, which led to suspension of the rosiglitazone marketing authorisation in the European Union in 2010 [R12-1620]. Together with the introduction in the market of DPP-4 inhibitors, this has changed the selection of second-line treatment regimens, as shown in two studies performed in the UK. One was a cohort study performed in the CPRD from 2000 to 2010, which found that the combination of metformin and DPP-4-inhibitors represented 0.7% of all second-line regimens in 2007, but DPP-4 inhibitors were prescribed in 20.2% of all second-line regimens

in 2010 [R14-5249]. On the other hand, the combination of metformin and thiazolidinediones (pioglitazone or rosiglitazone) represented 34% of the all second-line regimens in 2007 but only 9.8% in 2010 [R14-5249]. The other study was performed in The Health Improvement Network (THIN) database, where the annual incidence of prescriptions of thiazolidinediones decreased from 1.2 per 1,000 person-years in 2007 to 0.8 per 1,000 person-years in 2009, at the same time that "other glucose-lowering drugs," including DPP-4 inhibitors, increased from 0.2 per 1,000 person-years to 1.1 per 1,000 person-years [R14-5244].

8. RESEARCH QUESTION AND OBJECTIVES

The primary research question is to evaluate whether, among patients with T2D, initiation of empagliflozin changes the adjusted incidence of five outcomes compared with initiation of an SGLT2 inhibitor other than empagliflozin and compared with initiation of a DPP-4 inhibitor (the two comparator groups).
The five primary objectives of the study are as follows:

- To estimate the adjusted incidence rate ratio of acute liver injury by comparing patients initiating empagliflozin with patients initiating other SGLT2 inhibitors and with patients initiating a DPP-4 inhibitor.
- To estimate the adjusted incidence rate ratio of acute kidney injury by comparing patients initiating empagliflozin with patients initiating other SGLT2 inhibitors and with patients initiating a DPP-4 inhibitor.
- To estimate the adjusted incidence rate ratio of severe complications of UTIs (pyelonephritis and urosepsis) resulting in hospitalisation by comparing patients initiating empagliflozin with patients initiating other SGLT2 inhibitors and with patients initiating a DPP-4 inhibitor.
- To estimate the adjusted incidence rate ratio of genital infections by comparing patients initiating empagliflozin with patients initiating other SGLT2 inhibitors and with patients initiating a DPP-4 inhibitor.
- To estimate the adjusted incidence rate ratio of diabetic ketoacidosis resulting in hospitalisation by comparing patients initiating empagliflozin with patients initiating other SGLT2 inhibitors and with patients initiating a DPP-4 inhibitor.

The secondary objectives of the study are as follows:

- To estimate the adjusted incidence rate ratio of the secondary outcomes defined for acute liver injury, renal injury, severe complications of UTIs, and genital infections (see Section 9.3.2) by comparing patients initiating empagliflozin with patients initiating other SGLT2 inhibitors and with patients initiating a DPP-4 inhibitor.
- To estimate adjusted incidence rates for each of the five primary outcomes of interest stratified by categories of insulin use at the index date, age, sex, and other variables of interest such as diabetes control or prior history of UTI/genital infections, among patients initiating empagliflozin, other SGLT2 inhibitors, or a DPP-4 inhibitor.
- To estimate adjusted incidence rate ratios for each of the five primary outcomes of interest stratified by categories of insulin use at the index date by comparing patients initiating empagliflozin with patients initiating other SGLT2 inhibitors and with patients initiating a DPP-4 inhibitor.

9. RESEARCH METHODS

9.1 STUDY DESIGN

An observational cohort study will be conducted in the Clinical Practice Research Datalink (CPRD) in the United Kingdom (UK). The study will use a "new users" or "incident users" design and will compare new users of empagliflozin to two comparison groups, new users of other SGLT2 inhibitors and new users of DPP-4 inhibitors.

The new-user design avoids comparing a population predominantly composed of first-time users of a newly marketed drug such as empagliflozin with a population of prevalent users of an older drug who may have stayed on the comparator treatment for a longer time and be less susceptible to the events of interest. To avoid the inclusion of prevalent users, patients starting empagliflozin will be required to have no exposure during the available pre-index period to empagliflozin, another SGLT2 inhibitor drug, or a DPP-4 inhibitor. Patients starting another SGLT2 inhibitor will be required to have no exposure during the available pre-index period to an SGLT2 inhibitor, empagliflozin, or a DPP-4 inhibitor. Similarly, patients starting a DPP-4 inhibitor will be required to have no exposure during the available pre-index period to a DPP-4 inhibitor, empagliflozin, or another SGLT2 inhibitor [R13-1120, R14-4378].

Empagliflozin is usually a second- or third-line treatment for T2D; thus, it is expected that few patients with T2D initiating empagliflozin will be treatment naïve. For the majority of patients, empagliflozin will be added to an existing treatment (e.g., added to metformin), or patients will be switched to empagliflozin (e.g., from metformin plus an oral GLD other than the study drugs to metformin plus empagliflozin) due to disease progression, treatment failure, or side effects that may be related to study outcome. In this scenario, additional analysis will be done to achieve a fair comparison between switchers starting empagliflozin and switchers starting other SGLT2 inhibitors or DPP-4 inhibitors, by comparing patients starting third-line therapy with empagliflozin as an add-on drug to GLDs other than the study drugs with patients starting third-line add-on therapy with another SGLT2 inhibitor or with patients starting third-line therapy with a DPP-4 inhibitor [R13-1120]. In the same way, patients starting a combination of empagliflozin and metformin (whether a fixed-dose or free combination) will be compared with patients starting a combination of another SGLT2 inhibitor and metformin or with patients starting a combination of a DPP-4 inhibitor and metformin.

A cohort design will allow direct estimation of the absolute rates, rate differences, and relative risk or hazard ratios of multiple outcomes of interest among new users of empagliflozin compared with new users of other SGLT2 inhibitors and compared with new users of a DPP-4 inhibitor. A cohort study design will also allow accurate chronologic confounder assessment and assessment of the outcomes at multiple time points. The covariate information will be assessed during the time preceding treatment initiation and will include all historical information available for each patient. Follow-up will start the day after treatment initiation. In the context of a data source such as the CPRD, the use of a cohort design has more advantages than limitations compared with the use of a nested case-control

design—see the appendix discussion in Schneeweiss (2010) [R13-1120] and Patorno et al. (2014) [R14-4378].

Other SGLT2 inhibitors have been selected as a comparator group as patients with T2D treated with SGLT2 inhibitors other than empagliflozin are expected to constitute the most comparable patient group due to similar indications and target population. SGLT2 inhibitors were also suggested as an appropriate comparator per a European Medicines Agency request. DPP-4 inhibitors have been selected as an additional comparator group for several reasons. First, the National Institute for Health and Care Excellence (NICE) appraisal of dapagliflozin (an SGLT2 inhibitor) recommended that dapagliflozin should be used as described for DPP-4 inhibitors. The NICE Evidence Review Group considered that, overall, "DPP-4 inhibitors are the key comparators for dapagliflozin in both the dual therapy and triple therapy settings" [R13-5134]. Second, DPP-4-inhibitors, SGLT2 inhibitors, and thiazolidinediones (pioglitazone) have similar indications and target population, while dual therapy with GLP-1 (glucagon-like peptide-1) analogues has a restricted target population [P14-17374]. Finally, the use of thiazolidinediones has decreased in recent years, given increasing concerns about their safety, and at the same time, use of DPP-4 inhibitors increased, making second-line regimens with DPP-4-inhibitors the most common second-line regimens after metformin with sulfonylurea (see Section 9.4) [R14-5244, R14-5249].

Propensity scores will be estimated for each cohort member based on information prior to the index date. Propensity scores will incorporate measured potential predictors of therapy as independent variables and exposure group status (empagliflozin group vs. other SGLT2 inhibitors as a group and vs. the DPP-4 inhibitor group) as the outcome. Propensity scores will be used to minimise confounding.

9.2 SETTING

9.2.1 Study population

Empagliflozin is expected to be prescribed mainly by general practitioners (GPs) and specialists, and most of the follow-up prescriptions (for chronic treatment) will also be issued by GPs or primary care physicians. Thus, the selected study populations will be adult patients with T2D identified using data on general practice diagnoses and prescriptions written by GPs available in the UK CPRD during the study period. For ascertainment of hospitalisation-related study outcomes (UTI, DKA) linkage with Hospital Episode Statistics (HES) will be necessary to obtain more complete information on the study outcomes, but will be available only for a subset of available patients.
The study population will include eligible patients with T2D initiating treatment with empagliflozin, initiating other SGLT2 inhibitors, or initiating a DPP-4 inhibitor.

9.2.2 Study period

The study period will start 01 August 2014 (the empagliflozin launch date in the UK). The study end date is currently planned at 01 August 2017, but it may be extended up to August 2019, depending on the number of empagliflozin new users and study outcomes accrued in the CPRD at the planned interim reports. A sufficient number of more common safety

outcomes (UTI and genital infections) may also accumulate at an earlier time point (e.g., by the 24-month interim report, expected in mid-2017). The study period may be prolonged to up to 5 years after launch (August 2019), in which case final report would be expected in July 2020. If data obtained at the 24-month interim report indicates that insufficient new users and/or outcomes will accumulate in the CPRD for certain outcomes, additional data sources might be used to meet the study size requirements for these outcomes.

9.2.3 Index prescription definition

The index prescription will be the first prescription for empagliflozin (Anatomical Therapeutic Chemical [ATC] code A10BX12), other SGLT2 inhibitor comparator (dapagliflozin, ATC code A10BX09, or canagliflozin, ATC code A10BX11), or a comparator DPP-4 inhibitor (ATC code A10BH) during the study period for each new user available in the CPRD. Index prescriptions of the study drugs include the single study drugs or fixed-dose combinations of the study drugs with metformin when available.

9.2.4 Index date

The index date will be defined as the date on which each identified new user receives the index prescription for empagliflozin, other SGLT2 inhibitor comparator, or DPP-4 inhibitor comparator.

9.2.5 Baseline and lookback period

To characterise the empagliflozin, other SGLT2 inhibitor, and DPP-4 inhibitor cohorts at the time of study drug initiation, all information available during the lookback (pre-index) time period will be collected. The lookback time period is defined as the time period ending on the index date. Since all cohort members are required by inclusion criteria to have at least 12 months of data before the index date (baseline period), the lookback period will include at least 365 days during which covariates can be evaluated. For some of the cohort members, more data on covariates might be available beyond 365 days, and all available information will be considered for covariate classification related to diabetes, diabetes medications, and concomitant chronic conditions. Nevertheless, for comedications (i.e., for diseases other than diabetes) the lookback time period will be limited to 180 days prior to the index date. For acute conditions considered for exclusion criteria that will lead to exclusion if occur within 180 days prior to index date, information on the events occurring prior to that time point will be collected to be used as covariates, using all available data.
If the distribution of the duration of lookback time period is different among empagliflozin, other SGLT2 inhibitor, and DPP-4 inhibitor groups, categories of lookback time will be created using indicator variables. Those indicator variables will then be used as covariates in the multivariable regression models for outcome prediction, and for propensity score development, to control for possible differences in availability of information between the empagliflozin and comparator cohorts.

9.2.6 Inclusion criteria

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

- Be aged 18 or more years at the index date (date of initiation of empagliflozin or comparator).
- Have at least 12 months of continuous registration prior to the index date in a practice contributing up-to-standard data to the CPRD.
- Have T2D ever before the index date:
- A list of Read codes used to identify patients with T2D in the CPRD is included in Annex 3. The final algorithm to identify patients with T2D, which might include medication codes, will be described in the statistical epidemiological analysis plan.

The empagliflozin-exposed population must also meet the following criteria:

- Have at least one prescription for empagliflozin or fixed-dose combination of empagliflozin with metformin, with or without treatment with another GLD.
- Have no prior prescriptions of SGLT2 inhibitors (including empagliflozin) alone or in fixed-dose combination during the available pre-index period.
- Have no prior prescriptions of a DPP-4 inhibitor alone or in fixed-dose combination during the available pre-index period.

The population exposed to SGLT2 inhibitors other than empagliflozin must meet the following criteria:

- Have at least one prescription for an SGLT2 inhibitor other than empagliflozin or a fixed-dose combination of an SGLT2 other than empagliflozin with metformin with or without treatment with another GLD.
- Have no prior prescriptions of any SGLT2 inhibitor (including empagliflozin) or fixed-dose combination of any SGLT2 inhibitor with metformin during the available pre-index period.
- Have no prior prescriptions of a DPP-4 inhibitor or fixed-dose combination of a DPP-4 inhibitor with metformin during the available pre-index period.

The population exposed to a DPP-4 inhibitor must meet the following criteria:

- Have 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.
- Have no prior prescriptions of a DPP-4 inhibitor or fixed-dose combination of a DPP-4 inhibitor with metformin during the available pre-index period.
- Have no prior prescriptions of any SGLT2 inhibitor (including empagliflozin) or fixed-dose combination of any SGLT2 inhibitor with metformin during the available pre-index period.

9.2.7 Exclusion criteria

Patients with a confirmed diagnosis of T1D before the index date will not be included in the study. To identify patients with T1D, a combination of diagnosis and drug prescription codes

will be used. A list of Read codes used to identify patients with T1D in the CPRD is included in Annex 4.
Patients prescribed combinations of SGLT2 inhibitors with DPP-4 inhibitors (as fixed-dose combinations or as non–fixed-dose combinations of the two individual medications prescribed on the same date) will be excluded.

9.2.7.1 Exclusion criteria by outcome of interest

Different exclusion criteria will be applied to generate five sets of cohorts for the analysis of the five outcomes of interest.
For analysis of the ALI outcome, patients will be excluded if they meet *any* of the following criteria:

1. A diagnosis of any of the following acute conditions within 6 months before or at the index date - Annex 4 contains Read codes (Annex Table 4-1) and ICD - 10 (International Statistical Classification of Diseases and Related Health Problems, 10th Revision) codes (Annex Table 4-2) for these exclusion conditions:
   - ALI
   - Acute infectious hepatitis
   - Acute cholelithiasis and cholecystitis
   - Intra- or extrahepatic biliary obstruction
   - Acute pancreatic disease
   - Decompensated congestive heart failure (i.e., hospitalisation)
2. Pregnancy at baseline, because pregnancy can be associated with an increased risk of hepatic injury. Specific liver disorders associated with pregnancy include preeclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hyperemesis gravidarum [P15-00346]. Pregnancy 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.

For the analysis of the secondary outcome “ALI in patients with no predisposing conditions” the following set of exclusion criteria will be applied in addition to the ones applied for the primary ALI outcome:

1. A diagnosis of ALI is recorded any time before or at the index date (i.e., during the available lookback time). Annex 4 contains Read codes (Annex Table 4-1) and ICD-10 (*International Statistical Classification of Diseases and Related Health Problems, 10th Revision*) codes (Annex Table 4-2) for these exclusion conditions.
2. A diagnosis of the following chronic conditions recorded at any time before or at the index date—Annex 4 contains Read codes (Annex Table 4-1) and ICD-10 codes (Annex Table 4-2):

- Chronic liver disease
- 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
- Congestive heart failure

For the analysis of 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. Annex 4 contains Read codes (Annex Table 4-3) and ICD-10 codes (Annex Table 4-4) for AKI.
- A diagnosis of chronic kidney disease is recorded on or any time before the index date (i.e., during the available lookback time). Annex 4 contains Read codes (Annex Table 4-5) and ICD-10 codes (Annex Table 4-6).

For the analysis of the AKI secondary outcome, patients will be excluded if they meet the following criterion:

- A diagnosis of chronic kidney disease is recorded on or any time before the index date (i.e., during the available lookback time). Annex 4 contains Read codes (Annex Table 4-5) and ICD-10 codes (Annex Table 4-6).

For analysis of the UTI outcomes (both primary and secondary), patients will be excluded if they meet *any* of the following criteria:

- The patient experienced chronic or acute pyelonephritis within the 6 months before the index date (see Annex Table 4-7 for Read codes and Annex Table 4-8 for ICD-10 codes).

9.2.8 Follow-up of subjects

Follow-up will start the day after the index date, which will be the date of the first prescription for empagliflozin, another SGLT2 inhibitor, or a DPP-4 inhibitor.
For the analysis of each outcome, follow-up time will end at whichever of the following dates occurs first:

- The date of the outcome event; a diagnosis of ALI, AKI or chronic kidney disease, severe complication of UTI, genital infection, or DKA.
- The date of death.

- The date of study end.
- The date of transfer out of the practice or end of registration in the practice.
- The date that specific exclusion criteria are met (see exclusion criteria in Section 9.2.7). For both primary and secondary ALI outcomes, all exclusion criteria, including those used only for the secondary outcome, will serve as censoring events.
- The end date of the first continuous treatment of the index drug (empagliflozin or other SGLT2 inhibitor or DPP-4 inhibitor) plus a defined grace period (see also Section 9.3.1.1):
    - Main analysis: 30 days after the end of the last prescription's days' supply
    - Sensitivity analysis: 90 days after the end of the last prescription's days' supply.
- The date in which a new treatment episode starts with any of the other index drugs.

Patients with censored follow-up time will not be able to re-enter study cohort at a later time point. Follow-up will *not* be censored if oral or injectable glucose-lowering drugs other than the index drugs are prescribed in addition to empagliflozin, another SGLT2 inhibitor, or a DPP-4 inhibitor after the index date.

9.3 VARIABLES

9.3.1 Exposures

For this study, eligible patients will be identified from GP prescriptions for the study medications of interest listed in the CPRD.

Empagliflozin alone (Jardiance, ATC code A10BX12) or in fixed-dose combination with metformin hydrochloride (Synjardy, ATC code A10BD20) will be the study drug of interest. The oral GLDs currently suggested as comparators are SGLT2 inhibitors other than empagliflozin and DPP-4 inhibitors (see Sections 9.1 and9.4). Other comparator classes could be considered, depending on the prescribing practices and market uptake. Currently, the following SGLT2 inhibitors and DPP-4 inhibitors are marketed in the UK and will be comparator medications for study purposes:

SGLT2 inhibitors other than empagliflozin:

- Forxiga (dapagliflozin: ATC code A10BX09)
- Invokana (canagliflozin: ATC code A10BX11)
- Xigduo (dapagliflozin and metformin hydrochloride: ATC code A10BD15)
- Vokanamet (canagliflozin and metformin hydrochloride: ATC code A10BD16)

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 and metformin hydrochloride: ATC code A10BD07)
- Eucreas (vildagliptin and metformin hydrochloride: ATC code A10BD08)
- Kombiglyze (saxagliptin and metformin hydrochloride: ATC code A10BD10)
- Vipdomet (alogliptin and metformin hydrochloride: ATC code A10BD13)
- Jentadueto (linagliptin and metformin hydrochloride: ATC code A10BD11)

If additional SGLT2 inhibitor or DPP-4 inhibitor drugs are marketed in the future in the UK during the study period, they will also be considered to be members of the respective comparator group.

In clinical practice, new users of index drugs (empagliflozin, other SGLT2 inhibitors, and DPP-4 inhibitors) may be prescribed the drug in the context of (1) adding it to an existing GLD regimen as double or triple therapy or (2) switching from one GLD to the study drug as monotherapy or combination therapy. There will be relatively few patients with T2D starting SGLT2 inhibitors or DPP-4 inhibitors and naïve to other GLDs since monotherapy indication for these drugs is restricted to patients intolerant to metformin and sulfonylureas. Thus, new users of the study drugs would usually be (1) switching from monotherapy with another GLD to monotherapy with a study drug, (2) switching from dual or triple therapy with another GLD to dual or triple therapy with a study drug and other GLDs, or (3) adding a study drug to therapy with one or two other GLDs to become patients on dual or triple therapy. This approach is consistent with the new-user design described previously in Section 9.1 [R13-1120]. Information on whether patients received prior GLD therapy or if they were "added on" or "switched to" empagliflozin or a DPP-4 inhibitor at the time of inclusion in the study will be collected. Patients will be classified according to their treatment complexity as receiving mono vs. dual vs. triple therapy.

9.3.1.1 Exposure and time at risk

For this study, it will be assumed that the risk of ALI, AKI, hospitalisation for UTI/DKA, and the risk of genital infections related to empagliflozin, other SGLT2 inhibitors, or DPP-4 inhibitors (the index drugs) increases at the beginning of therapy, is maintained at an increased level for the duration of treatment, and decreases gradually to the background risk once treatment is stopped.

Only use at the first continuous treatment will be considered, defined as having consecutive prescriptions separated by 30 days or less. Therefore, the risk/exposure time window for each new user of an index drug—empagliflozin, other SGLT2 inhibitor, or DPP-4 inhibitor—will be categorised into two mutually exclusive categories of risk, as follows (Figure 1):

- Current use (current time at risk): The risk/exposure window for current use starts on the date of the prescription and ends 30 days after end of supply. The main analysis is based on current use. This time at risk will be used for comparisons and estimation of incidence rate ratios in the study main analysis.
- Recent use (recent time at risk): The risk/exposure window for recent use starts at the end of current use (30 days after end of supply) and ends 90 days later (which is 120 days after end of supply). The time at risk for recent use will be used in an additional analysis (see Section 9.8.3).

Overlapping time at risk from current use for consecutive prescriptions of the index medication will be concatenated, with the overlapping time counted only once. For consecutive prescriptions of the index medication separated by gaps of 30 days or less, time at risk from current use will include the gaps between prescriptions. Discontinuation of the index drug will be defined as no further prescription 120 or more days after the end of the last prescription's days' supply.

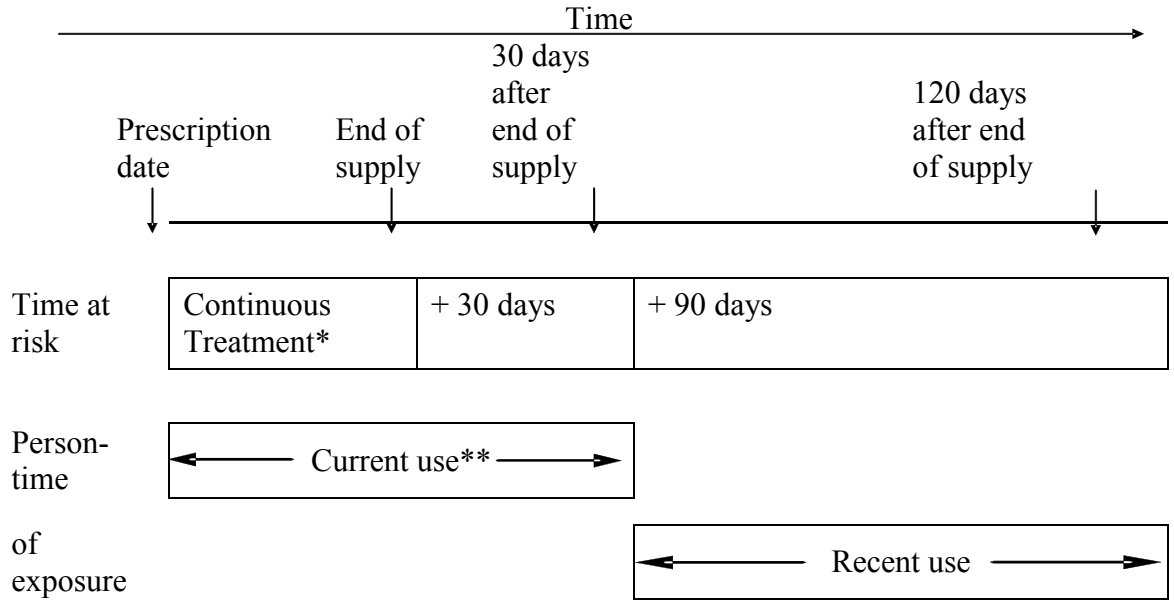

Figure 1. Exposure Definition

*A continuous treatment period of time at risk is defined as a prescription or multiple prescriptions separated by gaps of days' supply no longer than 30 days.

** Sensitivity analyses include a change in the current use definition that ends 90 days after the end of supply instead of 30 days.

Since most oral GLD prescriptions are supplied for 30 days, for the majority of cohort members, the risk window defining current use will end 60 days after the last prescription date. By adding 30 days to the end of the days' supply, a delayed increase in risk for ALI, AKI, hospitalisation for UTI/DKA, or genital infection after termination of the index drugs can be detected.
In a sensitivity analysis, the risk/exposure window for current use starts on the date of the prescription and ends 90 days after the end of supply. This sensitivity assessment will allow

exploration of any further potential delay in effect. A 90-day period was selected because this time is long enough to account for non-adherence and extended use of the discontinued index drug and a delay in effect.
Dose will be the dosage at index date. When the dose is missing, the dose will be estimated from the available recorded information (e.g., strength, number of units, amount of drug prescribed).

9.3.2 Study outcomes

The five primary outcomes of interest for this study are ALI, AKI, hospitalisation for UTI/DKA, and occurrence of genital infection. Secondary outcomes are chronic kidney disease (CKD), outpatient cases of severe complications of UTI, and genital infections resulting in hospitalization. Validation of identified cases to confirm diagnosis and date of the event will be implemented for all outcomes, and each of the outcome-specific sections that follow includes a case definition to be used for validation purposes and the algorithm of codes that will be used for case identification. The final section describes the common validation process that will be followed for all five outcomes.
There is scarce literature on the validity of diagnosis code algorithms used to identify genital infections and hospitalisations due to severe complications of UTI or to DKA. For this reason, initial testing of potential algorithms will be implemented in small samples of around 100 identified cases. Then, if the lower bound of the 95% CI for the positive predictive value of the coding algorithm for these outcomes is found to be below 50%, the coding algorithm will be updated for that outcome to achieve a higher positive predictive value; if necessary, another sample of medical records to validate the revised algorithm will be drawn from the database. The formal validation process will occur after a final algorithm for case identification is selected (see Section 9.3.2.6).

9.3.2.1 Acute liver injury, primary outcome

- Outcome type: primary
- Secondary outcomes within this section: yes (ALI in patients with no predisposing conditions)
- Further outcomes within this section: none
- Outcome name: incidence of acute liver injury
- Time frame: up to 5 years
- Safety issue: no

Drug-induced acute liver injury is defined as an adverse hepatic reaction that is unexpected on the basis of the pharmacological action of the drug administered [R14-1933]. Acute liver injury has been defined in terms of an elevation in the serum concentration of ALT or AST, conjugated bilirubin, or alkaline phosphatase (ALP). It has been considered that elevations of ALT/AST are indicators of liver injury, whereas increases of conjugated bilirubin are measures of overall liver function. Liver injury alone may not lead to clinically significant liver damage, whereas impaired liver function is a marker of severe drug-induced

hepatotoxicity. Thus, a combined elevation of ALT or AST and conjugated bilirubin without evidence of intra- or extra-biliary obstruction (i.e., no significant elevation of ALP) could be used to define potentially clinically significant elevations of serum liver enzyme levels [P06-02059]. The concept of combining markers of liver injury and function evolved from the observation of Hyman Zimmerman [R05-1093] that "drug-induced hepatocellular jaundice is a serious lesion." Zimmerman noted that the combination of pure hepatocellular injury (ALT elevation without much ALP elevation) and jaundice among patients with drug-induced liver injury had a poor prognosis, with a mortality of 10% to 15% [R14-5256, R05-1093, P14-17375, P09-12413]. This observation is referred to as "Hy's Law" by the FDA and has been used by the FDA over the years to assess the potential for a drug to cause severe liver injury—that is, irreversible liver failure that is fatal or requires liver transplantation [P09-12413, P14-17375, R14-5256].
According to guidance from the FDA, acute liver injury is defined as an elevation of the serum concentration of ALT or AST of at least 3 times the ULN and a contemporaneous (i.e., within 30 days of the ALT or AST increase) elevation of bilirubin concentration of at least 2 times the ULN in the absence of intra- or extrahepatic bilirubin obstruction or Gilbert's syndrome (Table 3).

Table 3. Clinical criteria for acute liver injury

Endpoint	**Definition**
Acute liver injury1	All the following criteria : • ALT or AST ≥ 3 × ULN • Bilirubin ≥ 2 × ULN • No intra- or extrahepatic bilirubin obstruction or Gilbert's syndrome

ALT = alanine aminotransferase; ULN = upper limit of normal range.
1. Sources: FDA (2009) [P09-12413], Temple (2001) [P14-17375], Temple (2006) [R14-5256].

Validation case definition

In this study, a case of acute liver injury is defined as any person with a recorded diagnosis compatible with acute liver injury who meets the criteria recommended by the FDA of having an elevation of the serum concentration of ALT or AST of at least 3-fold the ULN and a contemporaneous elevation of bilirubin concentration of at least 2-fold the ULN in the absence of intra- or extrahepatic bilirubin obstruction or Gilbert's syndrome (Table 3). Cases of liver injury can include patients requiring and not requiring hospitalisation for acute liver injury.

Case identification

Potential cases of liver injury will be identified by diagnosis codes suggestive of acute liver injury. In the CPRD, potential cases will be identified using Read codes and liver function test results. In HES data, potential cases will be identified by ICD-10 codes. Annex 5

contains a preliminary list of Read and biochemistry test codes (Annex Table 5-1) and ICD-10 codes (Annex Table 5-2).

9.3.2.1.1 Acute liver injury in patients with no predisposing conditions, secondary outcome

- Outcome type: secondary
- Secondary outcome: none
- Further outcomes within this section: none
- Outcome name: incidence of acute liver injury in patients with no predisposing conditions
- Time frame: up to 5 years
- Safety issue: no

For the assessment of this secondary outcome, additional exclusion criteria will be applied as described in Section 9.2.7.1.

9.3.2.2 Acute kidney injury, primary outcome

- Outcome type: primary
- Secondary outcomes within this section: Yes (CKD)
- Further outcomes within this section: none
- Outcome name: acute kidney injury
- Time frame: up to 5 years
- Safety issue: no

Multiple definitions of acute kidney injury have been used in clinical practice and research. In 2002, an international consensus conference of the Acute Dialysis Quality Initiative (ADQI) proposed the RIFLE criteria for defining and classifying acute kidney injury [R14-5242]. The RIFLE acronym indicates Risk of renal dysfunction, Injury to the kidney, Failure of kidney function, Loss of kidney function, and End-stage kidney disease. The RIFLE criteria are based on three levels of renal dysfunction (risk, injury, and failure) and two clinical outcomes (loss of function and end-stage renal disease). The RIFLE classification includes separate criteria for serum creatinine and urine output. The serum creatinine criterion is based on changes from baseline values. An increase of 1.5 times baseline indicates risk of kidney dysfunction, 2 times baseline indicates injury to the kidney, and 3 times baseline indicates failure of kidney function.

The definition of acute kidney injury in epidemiologic studies has been based on absolute increases of serum creatinine from normal values (1.7 or 2 times the ULN) or changes from baseline (20% to 50%) or both [P02-05226, P05-04032, P14-17372, P14-17376, P14-17377].

Validation case definition

In this study, a case of acute kidney injury is defined as any person with acute kidney injury according to the following criteria:

- At least a 2-fold increase in serum creatinine from the lowest baseline value recorded at any time before the index date, and the value is above the ULN; or
- An increase in serum creatinine to at least 2 times the ULN in the absence of a recorded baseline value; and
- Absence of a recorded diagnosis of chronic renal failure at any time before the index date.

Case identification

Potential cases of outpatient and hospitalised acute kidney injury will be identified by diagnosis codes suggestive of renal failure. In the CPRD, potential cases will be identified using Read codes. In the HES, potential cases will be identified by ICD-10 codes. Preliminary lists of codes are presented in Annex Table 5-3 (Read codes) and Annex Table 5-4 (ICD-10 codes).

9.3.2.2.1 Chronic kidney disease, secondary outcome

- Outcome type: secondary
- Secondary outcome: none
- Further outcomes within this section: none
- Outcome name: chronic kidney disease
- Time frame: up to 5 years
- Safety issue: no

CKD will be defined based on the estimated glomerular filtration rate (eGFR), as calculated using the 2009 CKD-EPI equation as described below [R13-4387; R12-1392; R15-5270].

141 × min(SCr/k, 1)α × max(SCr/k, 1)-1.209 × 0.993Age [× 1.018 if female] [× 1.159 if black]

SCr is serum creatinine (in mg/dl), 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.

CKD will be differentiated from acute kidney injury by requesting confirmation of the first abnormal test result.

Validation case definition

In this study, a case of chronic kidney disease is defined as any person with chronic kidney disease according to the following criteria that are met *after* the index date:

- Estimated GFR < 60 ml/min/1.73 m2

  AND

- Estimated GFR < 60 ml/min/1.73 m2 is confirmed in a separate test result performed at least 3 months after the initial post-index date at which the abnormal result was identified.

  AND

- Absence of a recorded diagnosis of chronic renal failure or an estimated GFR <60 ml/min/1.73 m2 at any time before the index date.

Case identification

Potential cases of chronic kidney disease will be identified by diagnosis codes suggestive of chronic renal failure. In the CPRD, potential cases will be identified using Read codes [R15-3136]. In the HES, potential cases will be identified by ICD-10 codes [R15-3138]. Preliminary lists of codes are presented in Annex Table 5-5(Read codes) and Annex Table 5-6(ICD-10 codes). All codes selected and displayed in the tables had positive predictive values larger than 70% [R15-3136, R15-3138].

9.3.2.3 Hospitalisation due to severe complications of urinary tract infections, primary outcome

- Outcome type: primary
- Secondary outcome within this section: Yes (outpatient cases of severe complications of UTI)
- Further outcomes within this section: none
- Outcome name: hospitalisation due to severe complications of urinary tract infections
- Time frame: up to 5 years
- Safety issue: no

Validation case definition

To define **pyelonephritis**, the criteria established by Patkar et al. (2009) [R14-5253] will be used.

1. At least two of the following will have to be present:

- History of fever or documented fever > 38.0°C or 104.0°F
- Dysuric complaints
- Flank pain/costovertebral angle tenderness
- Leukocytosis (white blood cell count > 12,000/cubic mm)
- Abnormal urine (cloudy, frank pus or blood in urine, foul smell)

AND (2 OR 3)

2. Any one of the following:
   - Computed tomography, magnetic resonance imaging, or ultrasonography findings consistent with renal inflammation
   - Computed tomography, magnetic resonance imaging, or ultrasonography findings consistent with renal abscess
   - Computed tomography, magnetic resonance imaging, or ultrasonography findings consistent with hydronephrosis

   OR

3. Any one of the following:
   - Blood cultures and urine cultures positive for the same organism
   - Blood cultures positive for Gram-negative organisms, *Enterococcus* species, or *Staphylococcus saprophyticus*
   - Urine culture positive for more than 10^5 Gram-negative organisms (e.g., *Escherichia coli*), *Enterococcus* species, or *S. saprophyticus*
   - Urine culture positive for fewer than 10^5 of any organism AND patient treated for at least 7 days with antibiotics

Urosepsis is clinically defined as sepsis caused by infection of the urinary tract and/or male genital organs (e.g., prostate). Patients are affected by microorganisms capable of inducing inflammation within the urinary and male genital tract. The following criteria from Wagenlehner et al. (2008) [R14-5286] need to be met for the diagnosis of urosepsis:

1. Diagnosis of infection of the urinary tract
   AND

2. One of the following criteria:
   - Proof of bacteraemia
   - Clinical suspicion of sepsis

AND

3. Two or more of following, which indicate the presence of systemic inflammatory response syndrome (SIRS):
    - Body temperature ≥ 38°C or ≤ 36°C
    - Tachycardia (≥ 90 beats per minute)
    - Tachypnoea (≥ 20 breaths per minute)
    - Respiratory alkalosis (PaCO2 ≤ 32 mm Hg)*
    - Leucocytes ≥ 12,000 per µL or ≤ 4,000 per µL or band forms > 10%

The case definition allows a patient to meet the criteria for both pyelonephritis and urosepsis, but the case will be counted only once for analysis purposes.

Case identification

The outcome hospitalisation due to severe complications of UTIs comprises the following conditions:

- Hospitalisation, emergency department (ED) visit, or a GP record of hospitalisation or referral to hospitalisation for pyelonephritis (Read codes in Annex Table 5-5, ICD-10 codes in Annex Table 5-6)
- Hospitalisation or ED visit for urosepsis. Combination of diagnosis code for UTI and a diagnosis code for sepsis within 1 week (Read codes for sepsis in Annex Table 5-5, ICD-10 codes for sepsis in Annex Table 5-6, Read codes for UTI in Annex Table 5-7, ICD-10 codes for UTI in Annex Table 5-8).

Initially, potential cases of severe complications of UTI will be identified via GP mentions of hospitalisations or ED visits associated with the required diagnoses. Because of CPRD limitations, linkage to HES data will also be used to identify relevant hospitalisations using ICD-10 discharge codes with relevant diagnoses not listed in the CPRD and to confirm or ascertain the correct hospitalisation date for any relevant hospitalisation listed in the CPRD (see Annex Table 5-8 for a list of the ICD-10 discharge codes of interest). Potential cases will be those with hospitalisations for the listed diagnoses dated in the CPRD during follow-up and up to 120 days after the end of follow-up. Linkage to the HES is currently limited to approximately 40% of study patients [P14-17413].

9.3.2.3.1 Outpatient severe complications of urinary tract infection, secondary outcome

- Outcome type: secondary
- Secondary outcome: none
- Further outcomes within this section: none
- Outcome name: Outpatient severe complications of UTI

* PaCO2 = partial pressure of carbon dioxide.

- Time frame: up to 5 years
- Safety issue: no

Outpatient cases of UTI will have to meet the same pyelonephritis definition used for the previous outcome (Section 9.3.2.3). Although unlikely to occur, if an outpatient case meets the previous definition of sepsis and not the one of pyelonephritis, the case will be included as a case for this outcome.

Case identification

The outcome outpatient severe complication of UTI comprises the following conditions:

- Outpatient visit for pyelonephritis (Read codes in Annex Table 5-5, ICD-10 codes in Annex Table 5-6)
- Outpatient visit for urosepsis. Combination of diagnosis code for UTI and a diagnosis code for sepsis within 1 week (Read codes for sepsis in Annex Table 5-5, ICD-10 codes for sepsis in Annex Table 5-6; Read codes for UTI in Annex Table 5-7, ICD-10 codes for UTI in Annex Table 5-8).

Initially, potential cases of outpatient severe complications of UTI will be identified via GP mentions of visits associated with the required diagnoses.

9.3.2.4 Genital infections primary outcome

- Outcome type: primary
- Secondary outcome: yes (severe genital infections)
- Further outcomes within this section: none
- Outcome name: incidence of genital infections
- Time frame: up to 5 years
- Safety Issue: yes

Diabetes is a risk factor for vaginal infections and balanitis, especially those produced by *Candida* species, with studies suggesting that the rate of vulvovaginal candidiasis and balanitis is higher among patients with diabetes [R14-5237, R12-3639, R12-2432, R14-5260, R14-5243].

This study will include only non-sexually transmitted genital infections. For women, bacterial vaginosis and vulvovaginal candidiasis will be included. For men, only non-sexually transmitted cases of balanitis will be included, which comprise *Candida* balanitis and aerobic balanitis.

Clinical diagnosis of bacterial vaginosis is presumptive and based on the presence of typical symptoms of vulvovaginitis, elevated pH (> 4.7 or > 4.5), and the presence of clue cells in saline wet mount or Gram stain of vaginal discharge. Diagnosis is enhanced by fishy odour of vaginal discharge after addition of 1-2 drops of 10% potassium hydroxide. Cultures are not useful. *Gardnerella vaginalis* is commonly found in women with bacterial vaginosis. Other organisms associated with bacterial vaginosis include *Prevotella* species, *Mycoplasma hominis*, and *Mobiluncus* species [R14-5247, R14-5252].

Clinical diagnosis of vulvovaginal candidiasis is presumptive if there are typical symptoms of vulvovaginitis and microscopic identification of yeast forms or hyphae in Gram stain or potassium hydroxide wet-mount preparations of vaginal discharge. Diagnosis is definitive by positive culture for *C. albicans* (or other *Candida* species) in symptomatic women [R14-5247, R14-5252].
Balanitis is defined as an inflammation of the penis that often involves the prepuce (balanoposthitis). Clinical diagnosis of balanitis is established based on clinical symptoms followed by culture confirmation; in the case of *Candida* balanitis, the isolation of yeast is definitive proof of fungal infection. There is a wide variety of causes and predisposing factors such as not being circumcised, neutropenia, or diabetes [R14-5248, R14-5243]. Balanitis can be due to several microorganisms, but *Candida* is the most common organism in patients with diabetes [R12-3639].
Non-sexually transmitted balanitis may also be caused by anaerobic and aerobic organisms, (such as *Gardnerella vaginalis* and Group B *Streptococcus*). Typical symptoms of *Candida* balanitis include burning and itching of the penis with generalised erythaema of the glans and/or prepuce (which may have a dry, glazed appearance) and with erosions, papules, and white discharge. In patients with diabetes, the presentation may be more severe, with oedema and fissuring of the foreskin, which may become non-retractile. Symptoms of *Gardnerella* balanitis are milder and include irritation of the prepuce and glans penis, macular erythaema, and fishy subpreputial discharge. Symptoms of balanitis due to Group B *Streptococcus* include non-specific erythaema with or without exudate [R14-5243].

Validation case definition

Vulvovaginitis: all potential cases (outpatient and hospitalised) of non-sexually transmitted vulvovaginitis will be included in the study and classified as vulvovaginal candidiasis, bacterial vaginosis, or non-specific vulvovaginitis or vulvitis.
Based on the clinical diagnosis mentioned above [R14-5247], either of the following criteria will need to be met for the diagnosis of vulvovaginal candidiasis:

1. A specific diagnosis of vulvovaginitis due to Candida AND any of the following
    a. Treatment with antifungals
    b. Yeast in Gram stain
    c. Culture positive for *Candida*
2. Non-specific diagnosis of vulvovaginitis OR symptoms of vulvovaginitis AND two or more of the following:
    a. Treatment with antifungals
    b. Yeast in Gram stain
    c. Culture positive for *Candida*

Based on the clinical diagnosis mentioned above [R14-5247], either of the following criteria will need to be met for the diagnosis of bacterial vaginosis:

1. A specific diagnosis of vulvovaginitis due to bacterial vaginosis AND any of the following:
    a. Treatment with metronidazole
    b. Clue cells in saline wet mount or Gram stain
    c. Culture positive for *Gardnerella*

d. pH > 4.7

2. Non-specific diagnosis of vulvovaginitis OR symptoms of vulvovaginitis AND two or more of the following:
    a. Treatment with metronidazole
    b. Clue cells in saline wet mount or Gram stain
    c. Culture positive for *Gardnerella*
    d. pH > 4.7

Cases with diagnosis or symptoms of vulvovaginitis or vulvitis that do not fulfil the above criteria will be classified as "non-specific vulvovaginitis or vulvitis" or excluded from the analysis if the event is suggestive of an alternative diagnosis such as a sexually transmitted genital infection.

*Balanitis***:** all potential cases of non-sexually transmitted balanitis (outpatient and hospitalised) will be included in the study and classified as *Candida* balanitis, aerobic balanitis, or non-specific balanitis.

Based on the clinical diagnosis mentioned above [R14-5247], either of the following criteria will need to be met for the diagnosis of *Candida* balanitis:

1. A specific diagnosis of balanitis due to Candida AND any of the following:
    a. Treatment with antifungals
    b. Yeast in Gram stain
    c. Culture positive for *Candida*
2. Non-specific diagnosis of balanitis OR symptoms of balanitis AND two or more of the following:
    a. Treatment with antifungals
    b. Yeast in Gram stain
    c. Culture positive for *Candida*

Based on the clinical diagnosis mentioned above [R14-5247], either of the following criteria will need to be met for the diagnosis of aerobic balanitis:

1. Non-specific diagnosis of balanitis AND culture positive for a non-sexually transmitted microorganism such as *Gardnerella* or Group B Streptococcus.
2. Symptoms of balanitis AND culture positive for a non-sexually transmitted microorganism such as *Gardnerella* or Group B *Streptococcus*.

Cases with diagnosis or symptoms of balanitis that do not fulfil the above criteria will be classified as "non-specific balanitis" or excluded from the analysis if the event is suggestive of an alternative diagnosis such as a sexually transmitted genital infection.

Case identification

Vulvovaginitis: potential cases of vulvovaginitis will be identified by the presence of an outpatient diagnosis or a hospitalisation or ED visit for vulvovaginitis, specifically, codes for specific diagnosis of bacterial vaginosis or vulvovaginal candidiasis, codes for non-specific diagnosis of vulvovaginitis, codes for specific microbiology results for *Candida* or *Gardnerella*, or diagnosis codes for non-specific positive microbiology results (see Read codes in Annex Table 5-9 and ICD-10 codes in Annex Table 5-13). Cases will be excluded if there is a diagnosis code with a specific diagnosis for a sexually transmitted infection within 30 days before or after the index infection date (see Annex Table 5-10). During the validation

process, cases will be classified into definite, probable, or possible cases, or no case, based on the combination of these and additional codes such as those with the results of laboratory tests (such as nucleic acid testing for STDs), treatment (such as antifungals), or microbiology results showing STD or negative infection.
Balanitis: potential cases of balanitis will be identified by the presence of an outpatient diagnosis of or a hospitalisation or ED visit for balanitis, including *Candida* balanitis, aerobic balanitis, and non-specific balanitis (see Read codes in Annex Table 5-13 and ICD-10 codes in Annex Table 5-15). Cases will be excluded if there is a diagnosis code with a specific diagnosis for a sexually transmitted infection within 30 days before or after the index infection date (see Annex Table 5-14). During the validation process, cases will be classified into definite, probable, or possible cases, or no case, based on the combination of these and additional codes such as those with the results of laboratory tests (such as detection of *Neisseria gonorrhoeae* nucleic acid or antibody titre), treatment (such as antifungals), or microbiology results showing STD or negative infection.

9.3.2.4.1 Severe genital infections: secondary outcome

- Outcome type: secondary
- Secondary outcome: none
- Further outcomes within this section: none
- Outcome name: incidence of severe genital infections
- Time frame: up to 5 years
- Safety issue: yes

Among the cases of genital infections identified in the primary outcome, those that resulted in hospitalisation will be classified as severe genital infections. To ensure that genital infection was the reason for hospitalisation and not a nosocomial infection, only primary discharge diagnoses will be used to identify genital infections (see ICD-10 codes in Annex Table 5-15). Those genital infections identified through an outpatient diagnosis and that required systemic treatment with antifungals or antibiotics (as opposed to topical or vaginal treatment) will also be classified as severe genital infections. Systemic treatment will be considered if prescribed on the date of the genital infection diagnosis or within 30 days before or after the genital infection diagnosis date. The UK recommended systemic treatment for vulvovaginitis includes metronidazole and tinidazole for bacterial vaginosis and fluconazole and itraconazole for vulvovaginal candidiasis [R14-5247]. The systemic treatment for balanitis recommended by the European Union is fluconazole for *Candida* balanitis, erythromycin or amoxicillin/clavulanic acid for aerobic balanitis, and metronidazole or amoxicillin/clavulanic acid for anaerobic balanitis [R15-0067].

9.3.2.5 Hospitalisation due to diabetic ketoacidosis: primary outcome

- Outcome type: primary
- Secondary outcomes within this section: no
- Further outcomes within this section: none

- Outcome name: incidence of diabetic ketoacidosis
- Time frame: up to 5 years
- Safety issue: no

Diabetic ketoacidosis is defined by the biochemical triad of ketosis, hyperglycaemia, and acidosis. DKA usually occurs as a consequence of absolute or relative insulin deficiency that is accompanied by an increase in counter-regulatory hormones (i.e., glucagon, cortisol, growth hormone, catecholamines). This type of hormonal imbalance enhances hepatic gluconeogenesis and glycogenolysis, resulting in severe hyperglycaemia. Enhanced lipolysis increases serum free fatty acids that are then metabolised as an alternative energy source in the process of ketogenesis. This results in accumulation of large quantities of ketone bodies and subsequent metabolic acidosis. Ketones include acetone, 3-beta-hydroxybutyrate, and acetoacetate. The predominant ketone in DKA is 3-beta-hydroxybutyrate. DKA has been considered to be indicative, or even diagnostic, of T1D, but increasingly cases of ketone-prone T2D are being recognised [R16-1372, R16-1373]. However, the initial treatment is the same. Several mechanisms are responsible for fluid depletion in DKA. These include osmotic diuresis due to hyperglycaemia; vomiting (commonly associated with DKA); and eventually, inability to take in fluid due to a diminished level of consciousness. Electrolyte shifts and depletion are in part related to the osmotic diuresis. Hyperkalaemia and hypokalaemia need particular attention. Serious complications of DKA and its treatment are hypokalaemia and hyperkalaemia, hypoglycaemia, and cerebral oedema [R16-1372, R16-1373] .

Validation case definition

In this study, a case of DKA is defined as any person with a recorded diagnosis compatible with DKA who meets the following criteria recommended by the Joint British Diabetes Societies [R16-1372] and the American Diabetes Association [R16-1371, R15-2058]:

- Ketonaemia > 1.5 mmol/L or ketonuria (+)
- pH less than 7.3. In the absence of pH values: anion gap > 10 mEq/L or venous bicarbonate (HCO3) below 15 mmol/L

Although blood glucose levels are usually a parameter used to define DKA, this criterion will not be used because atypical DKA can occur among users of SGLT2 inhibitors.
In this study, cases of DKA will include only patients requiring hospitalisation for DKA because some authorities contended that true DKA will always require hospitalisation or emergency department (ED) admission for treatment; otherwise, this study would include patients with milder forms of ketosis and unconfirmed DKA events [R16-1371, R16-1372].

Case identification

The outcome “hospitalisation due to DKA” comprises hospitalisation, ED visit, or a GP record of hospitalisation or referral to hospitalisation for DKA. Potential cases of DKA will be identified by diagnosis codes suggestive of DKA. In the CPRD, potential cases will be identified using Read codes and biochemistry test results, together with codes for hospitalisation. In HES data, potential cases will be identified by ICD-10 codes in the primary or secondary discharge position. Annex 5 contains a preliminary list of Read codes (Annex Table 5-16) and ICD-10 codes (Annex Table 5-17).

9.3.2.6 Case validation process common to all five primary outcomes

The goal of case validation is to confirm the diagnosis and date of onset of the identified cases of the five outcomes according to the clinical definitions described in the previous sections. All potential cases of ALI, all cases of AKI/CKD, and all cases of DKA will be validated, provided the total number of potential cases does not exceed 100 for each outcome. Otherwise, 100 cases will be selected at random and only those cases will be validated. Because many more cases of severe complications of UTI and genital infections are expected than cases of ALI and AKI/CKD or DKA, subsets of 100 cases of severe complications of UTI and 100 cases of genital infection in the CPRD and HES will be validated. For the individual outcomes where the initial algorithm does not demonstrate sufficient predictive value, up to 100 additional cases (if available) will be validated. Validation will follow a two-step process:

Review of the computerised clinical information and general practitioner questionnaires

In the first step, any potential case of a primary outcome identified from the CPRD will be evaluated further through review of the patient's computerised clinical information.
In a second step, for patients for whom available electronic data do not allow confirmation, including those patients for whom no HES linkage is available, patient computerised clinical information (patient profiles) will be reviewed.
However, as is almost always the case for outcome validation efforts, it is unlikely that access to additional validation information can be obtained for 100% of code-identified cases.
The review of patient computerised clinical information will be conducted by epidemiologists with clinical expertise. The relevant clinical information from these sources will be abstracted using a standardised abstraction form. As data are available, information to be abstracted will include admission and discharge diagnoses and dates, presence or absence of signs, symptoms, imaging results, and laboratory test results. Final confirmation of cases will be conducted independently by medical epidemiologists who will be blinded to medication exposure. Difficult cases will be evaluated by consensus between the validation physicians.
For all outcomes, all cases (whether or not they have been selected for the validation process) will be included in the analysis. However, the positive predictive value estimates available from the validation will help during interpretation of the results. Furthermore, given that 100 cases will be validated per outcome and the low expected incidence of ALI, AKI/CKD, and DKA, the majority of cases for those three outcomes will most likely be selected for validation.

9.3.3 Covariates

Exclusion diagnoses will be identified based on recorded GP diagnoses during the lookback period.
Variables potentially associated with the five outcomes of interest—such as sociodemographic variables including age, sex, socioeconomic status, body mass index, smoking or alcohol consumption; concomitant medications; and duration of lookback period (see Annex 6)—will be identified for all cohort members prior to and including the index date. In the CPRD, severity of T2D will be assessed by HbA1c values, diagnosis codes and duration since the first diagnosis, when available. These and other variables that can differ by

exposure group will be considered for inclusion as predictors in the logistic regression models used to generate the propensity scores. Propensity scores will be used in the analysis to quantify the probability of receiving empagliflozin at the index date for new users of both empagliflozin and other SGLT2 inhibitors. Additional propensity scores will be used in the analysis to quantify the probability of receiving empagliflozin at the index date for new users of both empagliflozin and DPP-4 inhibitors.

At the index date, cohort members will also be classified by indicator variables on the calendar time of cohort entry (by quarter) whether the index treatment (empagliflozin, other SGLT2 inhibitor, or DPP-4 inhibitor) was added to existing medication (adding on), or if the index treatment was initiated as a replacement for another GLD (switching from the existing GLD to empagliflozin, other SGLT2 inhibitor, or DPP-4 inhibitor), and whether this treatment was received as monotherapy or as dual or triple therapy. A variable indicating whether or not patients were receiving insulin at the index date will also be created. Patients will be classified as having "switched to" empagliflozin, other SGLT2 inhibitor, or a DPP-4 inhibitor if no prescriptions for any previously prescribed GLD are recorded after the index date. To calculate the end of previous GLD therapy, days' supply will be used, and 30 days will be assumed if days' supply is missing. At the index date, patients will be classified as "adding on" empagliflozin, other SGLT2 inhibitor, or DPP-4 inhibitor if at least one prescription of the previous GLD treatment is recorded after the index date.

The approaches to handling concomitant GLDs (including oral and injectable treatments such as insulin) in the analyses are summarised in Table 4.

Table 4. Approaches to handling concomitant glucose-lowering drugs

Timing and Type of GLD Prescription	Analysis Approach
At the index date, any GLD taken during baseline that is not the new prescription and is not eligible to be a study exposure	Include in propensity score
At the index date, any drug combination that includes the study drugs and metformin (i.e., empagliflozin plus metformin or any other SGLT2 inhibitor plus metformin or a DPP-4 inhibitor plus metformin)	Use an indicator variable to allow for stratification of patients according to whether or not at cohort entry study drugs were part of a combination with metformin. The indicator variable can be used also if appropriate in multivariable marginal structural Cox proportional hazard models (see Section 9.8.4).
Designation whether the GLD initiated at the index date is an add-on to current medication or a switch to a different medication	Include in propensity score and indicator variable to be used if appropriate in multivariable marginal structural Cox proportional hazard models (see Section 9.8.4).
Insulin use at the index date	Include in propensity score and conduct additional stratified analyses by insulin use at index date (Yes or No)
Insulin and/or other GLDs that have been used in the past (or not) and have been added during follow-up	If appropriate (see Section 9.8.4), use of multivariable marginal structural Cox proportional hazard models

1. GLD = glucose-lowering drug.

9.4 DATA SOURCES

The proposed study design requires data sources that longitudinally capture inpatient, ED, and outpatient diagnoses and procedures; capture prescription information; and allow validation of the five outcomes of interest. A data source meeting all those study requirements is the UK CPRD (website: cprd.com/), which is proposed as the study data source.

In the UK, nearly all residents are registered in a general medical practice that uses electronic medical records. Some of those records are available for research purposes in the CPRD. The CPRD contains diagnostic and prescribing information recorded by GPs as part of their routine clinical practice in the UK. The database currently contains data for over 13.2 million patients with research-quality data from 680 UK practices; 5.69 million of these patients are active (still registered with a contributing GP practice) [R14-5257]. Patients registered are representative of the whole UK population in terms of age and sex. A large proportion of patients are linkable to central mortality records. A large and growing proportion, currently

approximately 40% of the patients, can also be linked to hospitalisation records (hospital discharge diagnoses and procedures are coded using ICD-10 codes) via the patient's National Health Service number, sex, date of birth, and postal code. Updated, valid, linked CPRD data are available through the CPRD Division of the UK Medicines and Health Care Products Regulatory Agency.
Detailed information on prescriptions written by GPs, including prescribed dose and duration, is routinely recorded in the data source. Read codes are used for diagnoses, and Gemscript codes are used for medications. Additional diagnostic and treatment information can be found in free-text fields, letters from specialists and hospitals, and other sources. Because GPs serve as the gatekeepers for all medical services, any visit to a specialist or hospital requires communication back to the GP, who might enter that information into the medical record.
The CPRD contains information on lifestyle factors with a variable proportion of missing values. Although information on race is not available, other user characteristics of interest are likely to be captured. For example, data on body weight and height, smoking, and alcohol use were available for approximately 70% of patients in the CPRD [R14-5279]. In contrast, the pharmaceutical exposures and comorbidities are expected to be based on outpatient prescriptions and to be complete. The diagnosis of T2D, after excluding individuals with diagnosis codes for T1D, has been validated in the CPRD and found to have a high positive predictive value: 98.6% [R14-5280].
The UK is an ideal setting for population-based studies of diabetes because diabetes care is largely coordinated by the GP, and metabolic parameters, cardiovascular risk factors, diabetes comorbidities, and disease outcomes are collected electronically. Furthermore, clinical guidelines in the UK facilitate consistency in patterns of care [R14-5255]. In general, the validity of the former General Practice Research Database, upon which the CPRD was founded, as a reliable data source for drug safety studies in numerous therapeutic areas is well established [R11-2162, R99-1044]. However, all of the outcomes of interest in this study have not been well validated in the CPRD. For example, the validity of the diagnosis codes for severe complications of UTI or pyelonephritis has not yet been determined and does not appear in a recent systematic review of the validated outcomes in the CPRD. For ALI and AKI, the positive predictive value of the codes analysed in one study seems to be below 50% [R11-5210]. On the other hand, the combined infectious and parasitic endpoints studied in the CPRD have a median proportion of cases confirmed of 93%; for the combined genitourinary system endpoints, the median proportion of cases confirmed is 91% [R11-5211].
Access to de-identified, patient-level data from the CPRD is available following approval of the study protocol by the Independent Scientific Advisory Committee. A more detailed description of data available from the CPRD is shown in Annex 7.

9.5 POTENTIAL ADDITIONAL DATA SOURCES

As described in Section 9.6, the number of users and events in the CPRD might not provide sufficient power for the rarest outcome, ALI. In that situation, additional data sources may be considered. Those data sources should have an adequate number of empagliflozin users and should allow validation of the ALI outcome. Although in version 1.0 of this protocol the national registers in Sweden (patient register, prescription register, causes of death register, etc.) were considered an option for obtaining additional exposed patients, currently, validation of cases through medical record abstraction is not possible using Swedish data sources.

Potential additional data sources in Europe include the Danish national registers, the EpiChron database at IACS (Instituto Aragonés de Ciencias de la Salud [Aragón Institute of Health Sciences]) in Aragon, Spain, and the SIDIAP (Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària [Information System for the Advancement of Research in Primary Care]) database in Catalonia, Spain. A commercial database in the United States is also proposed as an additional data source. None of these databases has been contacted for participation in the study and further feasibility assessments are required.
In the Danish national registers, and through collaboration with academic centres, validation of hospitalised cases might be feasible. However, validation of outpatient cases of ALI would be feasible only for cases identified in hospital clinics. Moreover, given the total population of Denmark and the expected number of empagliflozin users in the country, additional data sources might still be required. In Europe, two data sources in Spain, EpiChron and SIDIAP, are based on GP records, have a structure similar to that of the CPRD, and would allow validation of cases of ALI because liver function tests and access to hospital records or GP questionnaires might be available.
As an additional option to the combination of available data sources in Denmark and Spain, a commercial data source in the US could be used. For example, one of the large administrative insurance claims databases with linkage to electronic health records should allow identification of sufficient numbers of empagliflozin users and cases of ALI while at the same time allowing the research team to validate a large proportion of the identified cases. In the following sections we describe in detail the proposed potential alternative data sources in Europe and the US.

9.5.1 Additional data sources in Denmark and Spain

Key characteristics of the study data sources proposed in Spain and Denmark are described in Annex 8 with information on the population of the countries where the data sources are located. A brief description of each proposed new study data source follows.

9.5.1.1 Danish data sources

The Danish health care system provides universal coverage to all Danish residents (5.6 million inhabitants; https://www.sundhed.dk/service/english/an-ehealth-nation/healthcare-in-dk/). Health care coverage includes visits to general practitioners and specialists, hospital admissions, and outpatient visits. The costs of medicines are partially covered by the Danish health system. The centralised Civil Registration System in Denmark allows for personal identification of each person in the entire Danish population and for the possibility of linkage to all Danish registries containing civil registration numbers, such as the Danish National Patient Register, Danish National Prescription Registry, prescription databases of the Central Denmark Region, and the Danish Register of Causes of Death. Data collected in these registries are available for research purposes. The process requires collaboration with a local university or investigator affiliated with a research institute to access the data and ethics committee notification or approval to handle data [R15-3135, R15-3137]. All applications have to be submitted in Danish.
Denmark's primary health care sector, which includes GPs, specialists, and dentists, generates about 96% of the prescription sales, most of which are reimbursable and are

dispensed by community pharmacies. Each dispensing record contains information on the patient, drug, and prescriber. Dispensing records retain the patient's universal personal identifier, allowing for individual-level linkage to all Danish registries and medical databases. Three national registries—Danish National Patient Register, Danish National Prescription Registry, and Danish National Database of Reimbursed Prescriptions—will be of particular interest. The Danish National Civil Registration System will be used to obtain information on death and migration status.

9.5.1.1.1 Danish National Patient Register

The registry includes data on all hospital admissions since 1 January 1977 and on outpatient clinic and emergency department visits since 1995 [R15-3137]. Hospital discharge diagnoses and information on surgical procedures, in-hospital deaths, and some selected drugs are recorded. After 1993, hospital discharge diagnoses are coded using ICD-10 codes.

9.5.1.1.2 Danish National Prescription Registry

The registry provides patient-level data on drug prescriptions dispensed by pharmacies since 1994 [R15-3141] and collects data on reimbursed and unreimbursed drugs.

9.5.1.1.3 The Danish National Database of Reimbursed Prescriptions

This data source encompasses the reimbursement records of all reimbursed drugs sold in community pharmacies and hospital-based outpatient pharmacies in Denmark since 2004 [R15-3140]. On average, approximately 3.5 million users are recorded in the database each year. Individuals are identified by the unique central personal registration (CPR) number assigned to all persons born in or immigrating to Denmark. This new data source avoids restrictions imposed on data use at the Danish National Prescription Registry. Most importantly, CPR numbers are reversibly encrypted, which allows re-identification of medication users. These features are very important for validation purposes.

9.5.1.1.4 Strengths and limitations of the Danish data sources

- Data from national registers include all age ranges in the population.
- At the national level, all dispensed prescriptions, regardless of reimbursement, are available.
- Source medical records can be accessed for selected projects and with special approvals for studies conducted in the Danish data sources.
- Most potential confounders can be obtained from the national databases, although only hospital-based diagnoses are available.

9.5.1.2 Spanish data sources

9.5.1.2.1 EpiChron database

A group of researchers at IACS in Aragon has linked the electronic medical and administrative databases in the region into a single data source (EpiChron). This database contain administrative, clinical, and drug information from outpatient clinics (primary care centres), emergency departments, hospitals, and pharmacies. From 2010 onwards, data are available for 1.3 million patients covered by all outpatient practices from the Aragon public health system. The following types of data are available: administrative and clinical information from outpatient clinics (primary care centres), emergency department diagnoses and care, hospital procedures and discharge diagnoses, and pharmacy prescription data. Studies are conducted in collaboration with the Institute of Public Health and Health Services Research; ethics committee approval is needed for the study.

9.5.1.2.2 SIDIAP database

The SIDIAP database in Catalonia, Spain, is a primary care database set up by the Institute of Research in Primary Care (Institut D'Investigació en Atenció Primària) and Catalan Institute of Health (Institut Català de la Salut). The database collects information from 279 primary health care centres and includes more than 5.8 million patients, about 78% of the Catalan population covered by the Catalan Institute of Health [R15-3142]. Data from health care visits are recorded in the electronic medical records. Linkage by an individual's national security number provides the potential to access information from different data sources, including demographic information from the Catalan Health Services database, electronic primary care clinical and laboratory test records, drugs dispensed in community pharmacies, and other available disease or procedural registries.
Information on pharmacy-dispensed drugs is available since 2005. Additional data available are the date and value of clinical variables, prescriptions issued, dispensed prescriptions (since 2005), and laboratory results (since 2006). The database can be linked to the Catalonian death registry, which includes date and cause of death of all residents [R15-3133]. All research projects applying to use SIDIAP data are assessed by an institutional review board (IRB) and the SIDIAP scientific review committee.

9.5.1.2.3 Strengths and limitations of the databases available in Spain for this study

- Broad coverage and representation of the general population covered by the Catalan Health System. The database in Aragón covers the whole population in the region.
- Empagliflozin prescriptions are reimbursed in Spain.
- The Aragón and Catalan health systems have universal coverage of reimbursed drugs and other health care of the population assigned to the primary health care centre.
- Linkage to hospital discharge codes is possible in a subset of the population.
- Access to pharmacy-dispensed prescriptions. Prescriptions dispensed in a hospital setting or purchased over the counter are not captured in the database.
- Information on relevant confounders is captured in both databases.

- Lifestyle habits (e.g., smoking) are captured in SIDIAP and to a lesser extent in IACS.
- Laboratory ambulatory test results are available in SIDIAP and IACS.
- Access to medical records for validation of events is available in IACS.
- Fewer studies have been published that either use one of the Spanish data sources or evaluate the validity of diagnoses in SIDIAP or IACS than for the CPRD.

9.5.2 Additional data source in the US: example, a large administrative insurance claims databases with linkage to electronic health records

Many of the commercial, administrative insurance claims databases available in the US now have linkages to electronic health records. The proportion of patients with linked data and the period of time where linked data are available vary by data source. If necessary, further feasibility assessments will be required before a specific US data source can be selected. However, in general, data sources of this type contain fully adjudicated paid claims with dates of service for all non-capitated ambulatory, emergency department, inpatient, and outpatient encounters (including administrative claims for laboratory tests) for large numbers of members with eligibility at the time of service, as well as claims for outpatient dispensings of prescription pharmaceuticals from pharmacies. Some data sources have the ability to redact or abstract inpatient and outpatient medical records for the health plan members, identify and contact providers and members for survey research through vendor relationships, and link data to national vital records. Most of these data sources are geographically representative within the US and have been used as a data source in multiple studies related to safety outcomes and validation. Health plans contributing data to these data sources typically include several different lines of business such as health maintenance organisations, point-of-service plans, preferred provider organisations, and indemnity plans.
Data on patient enrollment, medical care (professional and facility claims), outpatient prescription drug use, laboratory test results, and health care utilisation may be tracked for patients in the database. Diagnoses and procedures are identified by ICD-9-CM (*International Classification of Diseases, 9th Revision, Clinical Modification*), Current Procedural Terminology (CPT), and Healthcare Common Procedure Coding System (HCPCS) codes for both outpatient visits and inpatient stays. Drug claims are captured by National Drug Codes, which can be translated to broader, more meaningful classification systems such as Generic Product Identifier codes. Standard Logical Observation Identifiers Names and Codes are used to define specific laboratory test result data. Physician, specialist, and emergency department visits, as well as hospital stays, are captured through CPT codes, uniform billing (UB-92) revenue codes (e.g., room and board), and place-of-service codes. Information on physician specialty is also retained.

9.5.2.1 Strengths and limitations of administrative insurance claims data linked to electronic health records

- A large percentage of the US population is covered in these data sources.
- Patients aged 65 years or older are underrepresented in these data sources.

- Access to medical records for all cases requiring validation will not be possible. Medical record retrieval rates have been low for some studies.
- These are claims-based data sources with limited clinical information.

9.6 STUDY SIZE

The study size will be driven by the uptake of empagliflozin following approval and launch of empagliflozin for the treatment of T2D to improve glycaemic control in adults in the UK. In the latest available data from CPRD, with a cut-off date of 30 November 2014, the number of prescriptions of DPP-4 inhibitors in the CPRD was 790,139 prescriptions among a total of 43,509 patients. In the same data-cut, the number of prescriptions of SGLT2 inhibitors was 16,110 prescriptions among a total of 2,960 patients. Most of the prescriptions were for dapagliflozin, and only one prescription was for empagliflozin.
The required study size to detect an IRR of 3 among empagliflozin new users compared with DPP-4 inhibitor new users, with a comparator:empagliflozin ratio of 10:1 and a power of 80%, would be about 30,000 person-years of empagliflozin for liver injury and about 3,200 person-years for kidney injury. For all other outcomes, the number of empagliflozin new-user person-years required to detect an IRR of 3 would be less than 1,700 (Table 5). The study size required to detect an IRR of 3 among empagliflozin new users compared with other SGLT2 new users would be higher than the number presented above for the comparison with DPP-4 inhibitors, since reaching a comparator:empagliflozin ratio of 10:1 before study end is unlikely. It is expected that by the end of 2018, approximately 5,000-6,000 empagliflozin-treated patients will accumulate in the CPRD database, and by 2020 this number will reach approximately 9,000; however, the number of patients available for analysis may be lower. Occurrence of the events of interest in the CPRD will be monitored at prespecified interim report time points.
The empagliflozin UK launch occurred in August 2014. Delay in reimbursement and formulary entry may negatively affect the projected uptake. Therefore, it may be possible that due to slow patient accrual, the total numbers may be insufficient to address rare outcomes such as hepatic impairment. Based on early experience, the data source may be modified, i.e., if patient numbers accumulating in the CPRD are significantly below the expected counts, or an unexpectedly low event rate indicates insufficient study power for certain rare outcomes (e.g., ALI), an additional analysis using one or two additional data sources in countries with better market uptake of empagliflozin will be implemented via a protocol amendment. These potential additional sources are discussed in section 9.5.

Table 5. Number of empagliflozin-exposed person-years needed to detect an IRR of 1.5, 2, 3, or 4 with a power of 80% and alpha = 0.05 (using a two-sided test for the ratio of two Poisson rates)

Outcome	Background incidence rates1	Comparator:empagliflozin ratio, 10:1				Comparator:empagliflozin ratio, 20:1			
		IRR, 1.5	IRR, 2	IRR, 3	IRR, 4	IRR, 1.5	IRR, 2	IRR, 3	IRR, 4
Acute liver injury	0.14a	309,102	92,008	30,027	16,360	293,301	86,858	28,101	15,200
	0.23b	188,149	56,005	18,277	9,958	178,531	52,870	17,105	9,252
Acute kidney injury	1.29c,d	33,546	9,985	3,259	1,776	31,831	9,427	3,050	1,650
	1.98c	21,856	6,506	2,123	1,157	20,739	6,142	1,987	1,075
	2.88e	15,026	4,473	1,460	795	14,258	4,222	1,366	739
Acute pyelonephritis	3f	14,425	4,294	1,401	764	13,687	4,053	1,311	709
UTI leading to hospitalisation	15.8g	2,739	815	266	145	2,599	770	249	135
Vaginitis2	21h	2,061	613	200	109	1,955	579	187	101
Balanitis3	8.4h	5,152	1,534	501	273	4,888	1,448	468	253
Diabetic ketoacidosis in T2D	0.5i	86,549	25,762	8,408	4,581	82,124	24,320	7,868	4,256

IRR = incidence rate ratio; UTI = urinary tract infection.

1 Incidence rates per 1,000 patient-years.

2 Patient-years restricted to females.

3 Patient-years restricted to males.

Sources: a Huerta et al. (2002) [P03-03701]; b El-Serag and Everhart [R12-3632]; Girman et al. (2012) [R11-5319]; d Estimate calculated as the 65% of the incidence reported by the study, 2 per 1,000 person-years. Only 65% of those were based on hospital records; e Waikar et al. (2006) [R14-5285] (the results reported are on patients without T2D); f Venmans et al. (2009) [R12-1105]; g Benfield et al. (2007) [R12-1080]; h Hirji et al. (2012) [R12-3639]; i Wang et al (2008) [R14-3272].

9.7 DATA MANAGEMENT

All conversion of the original data to analysis variables will be performed using SAS software version 9.2 or higher (SAS Institute, Inc., Cary, North Carolina). Data management for CPRD data will be carried out in accordance with RTI Health Solutions (RTI-HS) standard operating procedures. Routine procedures include checking electronic files, maintaining security and data confidentiality, following the statistical epidemiological analysis plan, and performing quality-control checks of all programs.
Security processes will be in place to ensure the safety of all systems and data. Every effort will be made to ensure that data are kept secure so that they cannot be accessed by anyone except select study staff. Appropriate data storage and archiving procedures will be followed (i.e., storage on CD-ROM and DVD), with periodic backup of files to tape. A more complete description of the data management procedures will be included in the statistical epidemiological analysis plan.

9.8 DATA ANALYSIS

The final approach to data analysis will be presented in a separate statistical epidemiological analysis plan, to be developed prior to data collection.

9.8.1 Propensity score approach

Decisions to begin a specific GLD are influenced by demographic, medical, and clinical factors, and those same factors might be associated with the outcomes of interest. In the context of this study where the expected number of patients meeting the case definition is small for some of the outcomes (e.g., ALI), the number of covariates that can be used in a regression model predicting those outcomes is limited [R08-1486, R08-1494]. To overcome this problem, the set of confounding variables will be summarised into a single summary confounder score, a propensity score. The propensity score is the predicted probability of being assigned to a particular treatment conditional on a set of observed covariates. Because the models predict not the probability of experiencing the outcome but the probability of being treated with empagliflozin in this study, many more variables can be used in the predicting regression model [P12-04844, R14-5241, R14-5284, R14-5389]. Two sets of propensity scores, one for the comparison of empagliflozin versus other SGLT2 inhibitors and one for the comparison of empagliflozin versus DPP-4 inhibitors, will be generated. Furthermore, given the different inclusion/exclusion criteria used for each of the five primary outcomes, the two sets of propensity scores 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).
As a first step, a propensity score is estimated for each cohort member at the index date, based on the values of the observed covariates. In this study, propensity scores will be estimated by conducting multivariable logistic regression modelling and incorporating measured potential predictors of therapy as independent variables and exposure group status (empagliflozin group vs. other SGLT2 inhibitors group, and vs. DPP-4 inhibitor group) as the outcome. “Combination with metformin” status will be taken into account by indicator

variables that will be used in the analysis to stratify by whether or not study drugs were used as combinations with metformin at cohort entry.
The selection of variables to be included in the propensity score modelling (see Section 9.3.3) will be based on examination of exposure group differences in the distribution of each covariate and within categories of insulin use at the index date. Ideally, the included covariates should be associated with the outcomes of interest. Simulation studies show that variables that are unrelated to the exposure but are related to the outcome should always be included in the estimation of propensity scores [R12-1913]. Inclusion of these variables increases the precision of the estimated effect of exposure without increasing bias. In contrast, inclusion of variables that are related to the exposure but not to the outcome can decrease precision of the estimated effect of exposure without decreasing bias. Potential associations will be evaluated from the literature and, when needed, from bivariate associations within the data.
The variables listed in Annex 6 are potential candidates for inclusion in the propensity score model: Annex Table 6-1 (acute liver injury), Annex Table 6-2 (acute kidney injury), Annex Table 6-3 (urinary tract infection), Annex Table 6-4 (genital infection), and Annex Table 6-5 (diabetic ketoacidosis). Those variables will be assessed at the index date and during the previous lookback time period. Duration of the lookback period will be categorised, with indicator variables to be used for propensity score development and potential adjustment in multivariable regression models (see Section 9.7.4).
Prescription patterns change over time, and the confounding influence of the determinants of the prescription may also change. To allow for changing prescription patterns for empagliflozin from the time it is first available through the date of receipt of the data, the propensity score models will be developed for patients within each calendar year, based on index dates. Fitting separate models by annual periods will enable better control of time-varying confounders.
Descriptive analyses of covariates at baseline will include means, standard deviations, and medians/interquartile ranges, when appropriate, for continuous variables and percentages for categorical variables. Results will be stratified on index year–specific propensity score deciles and will be conducted annually. Strata will be formed within each year of data. Next, selective removal of cases, known as “trimming” [P12-04844], which occurs at both ends of the propensity score range, will be implemented. At the low end of the range, all patients, exposed or unexposed, with a propensity score below the 2.5 percentile value of the distribution of scores in the exposed group will be excluded. At the upper end of the range, we will exclude all patients, exposed and unexposed, with scores greater than the 97.5 percentile of scores among the comparator patients. Trimming will be performed separately for each outcome-specific cohort and, within each cohort, for each index year–specific set of propensity scores.
After trimming is completed, data will be stratified into deciles of propensity scores based on the distribution among empagliflozin new users. If the resulting decile strata are too small, deciles within index years will be combined. Additional details on how propensity scores will be developed and used will be provided in the statistical epidemiological analysis plan.
The obtained propensity score can be used in the analysis in different ways. One option is to stratify study patients into ranges of the propensity score and conduct a stratified analysis to control confounding. Another option is to match patients by propensity score. A third option is to use the propensity score in a multivariable regression model (see Section 9.7.4) predicting the outcome. That same model includes terms for the exposure and the propensity

score and might also include a few other covariates. Those covariates may or may not have been included in the model predicting the propensity score or may be modelled in the propensity score differently from their modelling in the final outcome model to obtain additional control of such covariates (for example, a variable may be categorised in the propensity score model and included as a continuous term in the final outcome model).
The propensity score methodology to be applied in this study will be outlined in the statistical epidemiological analysis plan.

9.8.2 Primary and secondary objectives: estimate adjusted incidence rate ratios and compare adjusted incidence rates for each of the study outcomes

Adjusted incidence rates of acute liver injury, AKI/CKD, severe complications of UTI, and incidence rates of genital infections among empagliflozin new users, among other SGLT2 inhibitor new users, and among DPP-4 inhibitor new users will be estimated and compared. Incidence rates will be reported as point estimates (in cases per 1,000 person-years) and 95% CIs. Ascertainment during follow-up will allow estimation of the number of new cases for each of the five primary outcomes. Current use person-time for each patient will be allocated as the time between the date of the first prescription for either empagliflozin, another SGLT2 inhibitor, or DPP-4 inhibitor and the end of time at risk (see Section 9.3.1.1 for time at risk definitions). The total person-time of observation among individuals at risk will then be calculated.

9.8.2.1 Main analysis

For each of the five primary outcomes (primary objective) and four secondary outcomes (secondary objective) of interest, estimation of adjusted current use IRRs with 95% CIs will be considered the main analysis of interest. Adjustment will be implemented by using propensity score methodology, e.g., by stratifying for propensity score deciles and calendar year among empagliflozin new users versus other SGLT2 inhibitor new users and versus DPP-4 inhibitor new users. More details on the analysis methods will be included in the statistical epidemiological analysis plan.
Although incidence rate ratios will be estimated using two different comparison groups and for five different outcomes, no Bonferroni type I error adjustment for multiple comparisons is planned for this study [R14-1393, R14-2938].
Crude IRRs will facilitate comparison with the adjusted IRRs to provide an indication of the degree of confounding. However, crude IRRs cannot be used for any comparisons given that this is a non-randomised, observational study and crude IRRs are expected to be biased due to channelling. Secondary objectives are to estimate adjusted incidence rates of each of the four outcomes among empagliflozin, other SGLT2 inhibitor, and DPP-4 inhibitor new users and estimate adjusted IRRs stratified by insulin use at baseline.

9.8.2.2 Secondary outcome analyses

The following adjusted incidence rates for each of the five primary and four secondary outcomes of interest among empagliflozin new users, among other SGLT2 inhibitor new users, and among DPP-4 inhibitor new users will be estimated:

- Kaplan-Meier estimates will be reported graphically to describe the occurrence of events during the follow-up.
- Adjusted incidence rates, stratified by categories of insulin use at the index date.
- Adjusted incidence rates, stratified by age, sex, and other variables of interest such as diabetes control or prior history of UTIs/genital infections.

Because the patients with T2D treated with insulin may have diabetes in a more advanced stage, may receive a different GLD treatment, and may be at different risk of the outcomes of interest, the incidence rate for each of the five outcomes of interest will be estimated stratified by insulin use at the index date in each cohort. Sensitivity analysis will include the combined analysis of patients with T2D treated with and without insulin, and tests of interaction will be performed to evaluate whether these estimates are biased. Incidence rates will be reported as point estimates (in cases per 1,000 person-years) and 95% CIs. The following estimates and comparisons will be generated:

- Summary incidence rate ratios after adjusting, e.g. by stratifying for propensity score deciles and calendar year, overall and by categories of insulin use at index date, comparing empagliflozin new users with other SGLT2 inhibitor new users and with DPP-4 inhibitor new users.

9.8.3 Duration, dose, and recent use effects analysis

- Adjusted IRRs by categories of duration of exposure will be estimated among empagliflozin new users versus other SGLT2 inhibitor new users and versus DPP-4 inhibitor new users. For example, patients with at least 1 year of continuous empagliflozin exposure will be compared with patients with at least 1 year of continuous other SGLT2 inhibitor exposure and separately with patients with at least 1 year of continuous DPP-4 inhibitor exposure; patients with less than 1 year of exposure to empagliflozin will be compared with patients with less than 1 year of exposure to other SGLT2 inhibitor and separately with patients with less than 1 year of exposure to DPP-4 inhibitors. Categories of duration will be defined based on available data.
- Adjusted IRRs by exposure dose categories will be estimated among empagliflozin new users versus other SGLT2 inhibitor new users and versus DPP-4 inhibitor new users.
- Additional analyses will estimate adjusted IRRs using recent use (recent time at risk) instead of current use (current time at risk).

9.8.4 Identification of potential confounders and effect modifiers during follow-up

The main strategy to address confounding is the use of propensity scores based on information prior to the index date. The degree to which we can pursue analyses of confounders and effect modifiers, especially during follow-up, is contingent on the number of events. If the number of events is sufficient, we plan to conduct the following analyses. Other variables (both fixed and time-dependent) will be classified during follow-up time, and analyses stratifying exposure time by level of these variables will be performed to explore potential effect modification and confounding by these variables on the adjusted incidence rates and adjusted IRRs for the five outcomes of interest. Some potential stratification

variables will include index year / quarter, HbA1c level, each specific concomitant GLD medication class added during follow-up, and prescription for a medication associated with the outcomes (e.g., ALI) during follow-up. Also, variables for which close balance was not achieved within propensity score strata may be examined further.
We will calculate the following statistics:

- Adjusted incidence rates by exposure category stratified by potential effect modifiers or confounders
- Adjusted incidence rate ratios stratified by potential effect modifiers or confounders

Use of insulin and other GLDs during follow-up will be examined. If differences exist in treatment intensification between the empagliflozin and other SGLT2 inhibitor cohorts, or between the empagliflozin and DPP-4 inhibitor cohorts, marginal structural Cox proportional hazard models will be used to account for time-dependent confounders that can also be on the causal pathways between exposures and outcomes.

9.8.5 Interim reports to monitor accrual of empagliflozin users and the event rates of acute liver injury and acute kidney injury

Accrual of empagliflozin users will be monitored and reported annually in three interim reports. The interim reports, which will include data up to 19, 24, and 36 months after use of empagliflozin is first captured in the CPRD, are expected to be released in June of each year. Crude incidence rates of acute liver and kidney injury outcomes (overall, not stratified by treatment) will be generated in the second and third interim reports. The number of events for acute liver and kidney injury is expected to be low. After each analysis, the available power to estimate the association between empagliflozin use and these two rare outcomes will be examined. If there is insufficient power because, given the current event rates, the number of new users of empagliflozin accrued up to that point is too low to yield acceptable precision, a decision will be made about extending the study population by including a country or countries with better market uptake of empagliflozin. In that case, a protocol amendment will be implemented to reflect inclusion of the additional data source(s).

9.8.6 Imputation of missing values

In the CPRD, no high frequency of missing values is expected for most variables, with the possible exception of lifestyle variables. If missing data are common for lifestyle variables, multiple imputation methods will be used to replace missing values during propensity score generation and multivariable analysis. Additional details on when and how multiple imputation methods will be used will be provided in the statistical epidemiological analysis plan. We propose to use multiple imputation methods because in most cases they allow for better bias correction than alternative methods and are more efficient than the complete-case approach. The complete-case approach can be very costly of information in a body of high-dimensional data, since the proportion of complete cases will decline with the increase in the number of variables [R07-2456, R14-5281].
Of note, for the medical history conditions / comorbidities to be collected for inclusion in the propensity score, the absence of a code for a condition will be interpreted as an absence of the event.

9.8.7 Sensitivity analyses

The following sensitivity analyses will be conducted:

- In the main analysis (see Figure 1) current use ends 30 days after end of supply but in this sensitivity analysis current use will end 90 days after end of supply. This change will be applied to all exposure groups: empagliflozin new users, other SGLT2 inhibitor new users, and DPP-4 inhibitor new users. New adjusted IRRs will then be estimated for empagliflozin users versus other SGLT2 inhibitor new users and versus DPP-4 inhibitor users.
- Conduct analyses including only validated cases.
- For hospitalised cases of AKI, hospitalised cases of severe complications of UTI, and hospitalisations due to DKA, conduct analyses excluding hospital-acquired cases of AKI, UTI, and ketoacidosis. The rationale for this analysis considers that including hospital-acquired cases of UTI and AKI does not allow appropriate control of potential confounders (e.g., medications, medical procedures) that occur within the hospitalisation. Nosocomial infections are particularly prone to confounding by patient characteristics and procedures conducted during the hospital stay. Moreover, hospitalisations with these conditions might present opportunities for diagnostic suspicion bias and confounding by association with the main reason for hospitalisation [P12-13528].

  To exclude hospital-acquired cases of AKI, in addition to discharge diagnosis codes, diagnosis codes or test results for one of the following events will have to be present within 7 days before the hospital admission:

  - Emergency department visit with a recorded diagnosis code for acute kidney injury
  - Outpatient visit with a recorded diagnosis code for acute kidney injury
  - A serum creatinine laboratory test result

  To rule out nosocomial infection, at least one of the following criteria will have to be present within 7 days prior to the hospital admission:

  - GP prescription for an antimicrobial agent
  - GP visit with a recorded diagnosis code for UTI or pyelonephritis
  - Emergency department visit with a recorded diagnosis code for UTI or pyelonephritis
- For cases of chronic kidney injury, conduct analyses to explore the time between exposure and chronic kidney injury using varying time windows to explore the lag time and minimal cumulative exposure to the outcome. The time window will be defined in the statistical epidemiological analysis plan. Exploratory analyses using different definitions of chronic kidney disease will also be conducted to compare definitions and to facilitate comparison to other studies. The definitions will be defined in the statistical epidemiological analysis plan.
- Compare characteristics of validated cases with those of cases that were not validated.

- Assess the potential effect of unmeasured confounders on the association between empagliflozin use and, for example, ALI, by using the method described by Lash et al. [R14-5373]. More details and examples of how this method will be used will be provided in the statistical epidemiological analysis plan.
- Estimate summary adjusted IRRs using propensity score methods not used in the main analysis (e.g., if propensity score stratification is the main analysis, sensitivity analyses will be conducted in propensity score–matched and propensity score-adjusted cohorts).
- Estimate summary adjusted IRRs using intention-to-treat analysis, carrying forward the initial exposure status and disregarding changes in treatment status during all available follow-up time.

9.8.8 Further analysis

If additional data sources are used because of insufficient accrual of cases when using the CPRD as the only data source, techniques to pool the data will be applied to combine IRR estimates across data sources, if appropriate.

9.9 QUALITY CONTROL

Standard operating procedures will be used to guide the conduct of the study. These procedures include internal quality audits, rules for secure and confidential data storage, methods to maintain and archive project documents, quality-control procedures for programming, standards for writing analysis plans, and requirements for senior scientific review.
All programming written by one study analyst will be reviewed independently by a different analyst, with oversight by a senior statistician. All key study documents, such as the analysis plan, abstraction forms, and study reports, will undergo quality-control review, senior scientific review, and editorial review.
An independent Office of Quality Assurance (OQA) will perform audits and assessments that involve various aspects of the project, including but not limited to education and training documentation, data entry and data transfer procedures and documentation, and IRB documentation. Such audits will be conducted by the OQA according to established criteria in standard operating procedures and other applicable procedures. Standard procedures will be in place to restore files in the event of a hardware or software failure.
A quality-assurance audit of this study may be conducted by the sponsor or the sponsor's designees.

9.10 LIMITATIONS OF THE RESEARCH METHODS

There are several methodological challenges when conducting epidemiologic studies to evaluate the association between glucose-lowering medications and outcomes of interest among patients with diabetes. These challenges include, but are not limited to, changes in treatment in response to advancing diabetes or due to adverse effects of specific drugs and time-varying risk of an outcome depending on duration of exposure [P12-13528, P14-17457, R14-4378].

9.10.1 Confounding

Although use of propensity scores will facilitate the control of measured confounders, unmeasured and unidentified confounders could still introduce bias if they are differentially distributed among the exposed and comparator groups and are related to the outcome. As an example, use of over-the-counter medications will remain unmeasured in this study.
Confounding by indication or severity, also known as channelling bias, is a common bias in pharmacoepidemiology. Patients starting treatment with a newly marketed drug might have more severe disease than patients not taking the medication either because of self-selection or because of physician preference. They may also have a less severe form of the disease if physicians prefer to test new drugs with a less familiar safety profile in less severely affected patients. New medications may also be prescribed differentially by physicians who are "early adopters" of new technologies and who systematically treat more severely affected patients with the new medications. The use of propensity scores and other SGLT2 inhibitors and DPP-4 inhibitors as the comparator groups—relative to other GLDs such as sulfonylureas (both SGLT2 inhibitors and DPP-4 inhibitors were recently introduced in the market)—reduces the risk of this type of bias, but residual confounding could still operate. For example, empagliflozin could be preferentially prescribed to patients with more severe diabetes or for whom other GLDs have failed. Empagliflozin could also be more likely prescribed to patients with fewer risk factors for severe complications of UTI. Similarly, the decision for hospitalisation could be affected by the perceived side effects profile. These channelling patterns could bias the risk estimate towards or away from the null.
Insulin use among patients with T2D is a well-known indicator of disease severity. Although this study will control for use of insulin at baseline, complete control of the potential confounding by severity associated with insulin use might not be achieved. Moreover, different use of insulin during follow-up could still confound the association between empagliflozin and the outcomes of interest. Use of insulin during follow-up could be related to both the severity of the disease and the lack of metabolic control achieved by empagliflozin or the comparators: other SGLT2 inhibitor drugs or DPP-4 inhibitor drugs. Thus, insulin use could be a consequence of the exposure and at the same time influence metabolic control and therefore the outcomes of interest. In this setting, glycated haemoglobin might be a time-dependent confounder on the causal pathways between exposures and outcome. In this study, differences in insulin use during follow-up between the empagliflozin and the other SGLT2 inhibitor cohorts and between the empagliflozin and the DPP-4 inhibitor cohorts will be examined. If the differences in insulin use are relevant, marginal structural Cox proportional hazard models, which can account for time-dependent confounders on the causal pathways between exposures and outcome, will be used.
The occurrence of the outcomes of interest may also vary over time according to the duration or cumulative dose of drug exposure. To account for these time-varying hazards, a stratified analysis estimating the IRRs of the outcomes of interest by categories of cumulative duration of exposure will be conducted.
In the planned sensitivity analyses, the potential effect of unmeasured confounders on the association between empagliflozin use and each of the five outcomes of interest will be evaluated. These analyses will enable estimation of the degree of possible bias by assuming different and plausible values for those confounders.

9.10.2 Other biases

Misclassification bias can occur when study patients are not correctly assigned to the outcome and/or exposure. Because prescribing records will be used, misclassification of exposure is unlikely. However, analyses will not control for non-adherence to the study drugs. Moreover, misclassification of new users could happen if free samples of empagliflozin or DPP-4 inhibitor are provided to patients for different periods of time. Misclassification of the outcome will be reduced by the validation process that has been planned for each of the five outcomes of interest.
Empagliflozin will be compared with other SGLT2 inhibitors as a group and to DPP-4 inhibitors as a group. Information about the risk of the outcomes of interest among patients using specific SGLT2 inhibitors or specific DPP-4 inhibitors is scarce; therefore, the comparator groups will include specific drugs within each drug class that may or may not have a differential risk for some of the outcomes of interest. This could potentially bias empagliflozin risk estimates for this study.
The main analysis in this study will be performed as treated; cohort members will be censored after discontinuation of the index oral GLD (120 days after the end the days' supply of the last prescription). However, if drug discontinuation predicts future outcomes of interest, then an informative censoring bias may occur because we are removing outcomes from their appropriate exposure category. In this study, the risk time window for current use will terminate after the days' supply of the last prescription has elapsed plus 30 days. A sensitivity analysis with varying latency periods after drug discontinuation will evaluate the potential for informative censoring bias.
In addition, another sensitivity analysis will be performed by conducting an intention-to-treat analysis, which will carry forward the initial exposure status and disregard changes in treatment status over time.

9.10.3 Limitation due to study size

As described in more detail in Section 9.5, the CPRD is probably the richest data source to conduct this study since it captures lifestyle factors not available in other data sources. It is expected that by the end of 2020, around 9,000 users of empagliflozin will have been accrued in the CPRD. However, the number of patients available for analysis might be lower. Therefore, it might take longer than expected until sufficient power to address the rare outcomes is reached, especially for the acute liver and acute kidney injury outcomes. With 9,000 accumulated person-years, an IRR between 2 and 3 could be detected with an 80% power for the AKI, UTI, and genital infection outcomes (see Table 5).
However, the occurrence of the events of interest will be monitored at 12, 24, and 36 months after use of empagliflozin is first captured in the CPRD. Based on postlaunch uptake in the CPRD, the study population may be modified, i.e., if patient numbers accumulating in the CPRD are significantly below the expected counts, or an unexpectedly low event rate indicates insufficient study power for certain rare outcomes (e.g., acute liver injury). In those circumstances, expanding the study population by including a country or countries with better market uptake of empagliflozin will be implemented via a protocol amendment. At this time, data sources in Denmark, Spain, and the US are considered the most appropriate to complement the CPRD, if necessary (see Section 9.6). However, this will be reassessed based on the market experience in different European Union countries and the US. A US data

source may be used if empagliflozin exposure in other European data sources is also too low to appropriately address rare outcomes.

9.10.4 Generalisability

Use of the CPRD, a population-based database, provides data entered by primary care practitioners in a routine clinical care setting. Therefore, the study results can be generalised to similar patients with T2D in other geographic settings, including most industrialised countries.

10. PROTECTION OF HUMAN SUBJECTS

This is a non-interventional study using an existing database (secondary data) and does not pose any risks for patients. All data collected in the study will be de-identified with no breach of confidentiality with regard to personal identifiers or health information. RTI-HS will apply for an independent ethics committee review according to local regulations in the UK; in addition, RTI-HS will obtain approval from the RTI International† IRB.
Data protection and privacy regulations will be observed in collecting, forwarding, processing, and storing data from study participants.

10.1 RTI INTERNATIONAL

RTI International holds a Federal-Wide Assurance from the Department of Health and Human Services Office for Human Research Protections that allows the organisation to review and approve human subjects protocols through its IRB committees. RTI International currently has three IRB committees available to review research protocols. One IRB committee is constituted to review medical research and has two members who are physicians. These IRBs have been audited by the US Food and Drug Administration and are fully compliant with applicable regulatory requirements.

10.2 CPRD

RTI-HS will submit the final study protocol for approval to the Independent Scientific Advisory Committee (ISAC) (http://www.cprd.com/ISAC). The CPRD has obtained ethical approval from a Multicentre Research Ethics Committee for all observational research using CPRD data without patient involvement; however, ISAC may recommend that the Multicentre Research Ethics Committee review the study documentation if any ethical issues arise.

10.3 OTHER GOOD RESEARCH PRACTICE

This study adheres to the *Guidelines for Good Pharmacoepidemiology Practices (GPP)* of the International Society for Pharmacoepidemiology [R11-4318] and has been designed in line with the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) *Guide on Methodological Standards in Pharmacoepidemiology* [R14-5282]. The *ENCePP Checklist for Study Protocols* [R13-1395] is included in Annex 2.
The study is a PASS and will comply with the definition of the non-interventional (observational) study referred to in the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use tripartite guideline *Pharmacovigilance Planning E2E* [R11-2259] and provided in the European Medicines Agency (EMA) *Guideline on Good Pharmacovigilance Practices (GVP) Module VIII: Post-Authorisation Safety Studies* [R13-5420] and with the 2012 European Union

† RTI Health Solutions is a business unit of RTI International, a private, not-for-profit research organization.

pharmacovigilance legislation, adopted 19 June 2012 [R14-5246]. Following EMA regulations, the study protocol has been registered at the EU PAS Register, and an abstract of results will be registered at the EU PAS Register (registration number: ENCEPP/SDPP/13413; http://www.encepp.eu/encepp/viewResource.htm?id=13414).

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Based on current guidelines from the International Society for Pharmacoepidemiology [R11-4318] and the EMA [R13-1970], non-interventional studies such as the one described in this protocol, conducted using medical chart reviews or electronic claims and health care records, do not require expedited reporting of suspected adverse events/reactions. Specifically, as stated in section VI.C.1.2.1 of *Guideline on Good Pharmacovigilance Practices (GVP), Module VI – Management and Reporting of Adverse Reactions to Medicinal Products*, for non-interventional study designs, which are based on use of secondary data, reporting of adverse reactions is not required.
The data generated in the course of the study will be monitored by the BI responsible person. When an observation is identified that may qualify as a special safety issue or that may have implications for the benefit-risk balance of empagliflozin, appropriate BI functions will be notified according to BI standard operating procedures.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Study milestones will be agreed with the European Medicines Agency. The study progress will be reported by BI in regulatory communications in line with the risk management plan, Periodic Safety Update Reports, and other regulatory milestones and requirements. Study reports will be prepared using a template following the *Guideline on Good Pharmacovigilance Practices (GVP), Module VIII*, Section B.6.3 [R13-5420]. The planned periodic interim reports at 12, 24, and 36 months will be reported within the earliest corresponding Periodic Safety Update Report.
Section V of *Guidelines for Good Pharmacoepidemiology Practices (GPP)* [R11-4318] contends that "there is an ethical obligation to disseminate findings of potential scientific or public health importance"; for example, results pertaining to the safety of a marketed medication. Moreover, a well-developed publication strategy is encouraged in the *Guideline on Good Pharmacovigilance Practices, Module VIII*, Section B.7 [R13-5420].
RTI-HS reserves the right to submit the results from any of the study analyses for publication and commits that at least the final results will be published. Any publications will follow guidelines, including those for authorship, established by the International Committee of Medical Journal Editors [R13-5418]. When reporting results of this study, the appropriate STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist will be followed [R13-2485].

13. REFERENCES

13.1 PUBLISHED REFERENCES

[R14-1933]	Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, et al. Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther. 2011 Jun; 89 (6): 806-15.
[R15-3134]	Aitken GR, Roderick PJ, Fraser S, Mindell JS, O'Donoghue D, Day J, et al. Change in prevalence of chronic kidney disease in England over time: comparison of nationally representative cross-sectional surveys from 2003 to 2010. BMJ Open 2014; 4 (9): e005480.
[P05-08822]	Andrade RJ, Lucena MI, Fernández MC, Pelaez G, Pachkoria K, García-Ruiz E, et al. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. Gastroenterology 2005 Aug; 129 (2): 512-21.
[R14-5241]	Arbogast PG, Ray WA. Performance of disease risk scores, propensity scores, and traditional multivariable outcome regression in the presence of multiple confounders. Am J Epidemiol 2011; 174 (5): 613-20.
[R07-2456]	Arnold AM, Kronmal RA. Multiple Imputation of Baseline Data in the Cardiovascular Health Study. Am J Epidemiol 2003; 157 (1): 74-84.
[R15-2053]	Barski L, Nevzorov R, Harman-Boehm I, Jotkowitz A, Rabaev E, Zektser M, et al. Comparison of diabetic ketoacidosis in patients with type-1 and type-2 diabetes mellitus. Am J Med Sci. 2013 Apr;345(4):326-30.
[R14-5242]	Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P, Acute Dialysis Quality Initiative w. Acute renal failure - definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. Crit Care 2004; 8 (4): R204-12.
[R12-1080]	Benfield T, Jensen JS, Nordestgaard BG. Influence of diabetes and hyperglycaemia on infectious disease hospitalisation and outcome. Diabetologia 2007; 50 (3): 549-54.
[R14-4617]	Boehringer Ingelheim International GmbH. Jardiance (empagliflozin) summary of product characteristics. 2014. Available at: website: ema.europa.eu/ema/index.jsp?curl = pages/medicines/human/medicines/002677/human_med_001764.jsp&mid = WC0b01ac058001d124. Accessed 09 October 2014.

[R15-3133]	Bolíbar B, Fina Avilés F, Morros R, Garcia-Gil Mdel M, Hermosilla E, Ramos R, et al. [SIDIAP database: electronic clinical records in primary care as a source of information for epidemiologic research]. Med Clin (Barc) 2012; 138 (14): 617-21.
[R14-5237]	Bohannon NJ. Treatment of vulvovaginal candidiasis in patients with diabetes. Diabetes Care 1998; 21 (3): 451-6.
[R12-1083]	Boyko EJ, Fihn SD, Scholes D, Abraham L, Monsey B. Risk of urinary tract infection and asymptomatic bacteriuria among diabetic and nondiabetic postmenopausal women. Am J Epidemiol 2005; 161 (6): 557-64.
[R14-5236]	Boyko EJ, Fihn SD, Scholes D, Chen CL, Normand EH, Yarbro P. Diabetes and the risk of acute urinary tract infection among postmenopausal women. Diabetes Care 2002; 25 (10): 1778-83.
[R08-1486]	Braitman LE, Rosenbaum PR. Rare outcomes, common treatments: analytic strategies using propensity scores. Ann Intern Med. 2002; 137 (8): 693-5.
[R12-1913]	Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. Am J Epidemiol. 2006 Jun 15; 163 (12): 1149-56.
[R14-5283]	Brown JS, Wessells H, Chancellor MB, Howards SS, Stamm WE, Stapleton AE, et al. Urologic complications of diabetes. Diabetes Care 2005; 28 (1): 177-85.
[R08-1494]	Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. Am J Epidemiol 2003; 158 (3): 280-7.
[P14-17370]	Czaja CA, Scholes D, Hooton TM, Stamm WE. Population-based epidemiologic analysis of acute pyelonephritis. Clin Infect Dis 2007; 45 (3): 273-80.
[R14-5284]	D'Agostino RB, Jr. Propensity scores in cardiovascular research. Circulation 2007; 115 (17): 2340-3.
[R15-3135]	Danish Data Protection Agency. Introduction to the Danish Data Protection Agency. 2014. website: http://www.datatilsynet.dk/english/ (access date: 16 June 2015).
[R15-3137]	Danish Health and Medicines Authority. Epidemiology. 2013. website: http://www.ssi.dk/English/RandD/Research%20areas/Epidemiology.aspx (access date: 16 June 2015).
[P04-07683]	de Abajo FJ, Montero D, Madurga M, García Rodríguez LA. Acute and clinically relevant drug-induced liver injury: a population based

	case-control study. Br J Clin Pharmacol. 2004 Jul; 58 (1): 71-80.
[R12-2432]	de Leon EM, Jacober SJ, Sobel JD, Foxman B. Prevalence and risk factors for vaginal Candida colonization in women with type 1 and type 2 diabetes. BMC Infect Dis 2002; 2: 1.
[P06-11008]	De Valle MB, Av Klinteberg V, Alem N, Olsson R, Björnsson E. Drug-induced liver injury in a Swedish University hospital out-patient hepatology clinic. Aliment Pharmacol Ther 2006; 24 (8): 1187-95.
[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): 1138-49.
[R14-5243]	Edwards S. Balanitis and balanoposthitis: a review. Genitourin Med. 1996 Jun; 72 (3): 155-9.
[R15-0067]	Edwards S, Bunker C, Ziller F, van der Meijden WI. 2013 European guideline for the management of balanoposthitis. Int J STD AIDS. 2014 May 14; 25 (9): 615-26.
[R12-3632]	El-Serag HB, Everhart JE. Diabetes increases the risk of acute hepatic failure. Gastroenterology. 2002 Jun; 122 (7): 1822-8.
[P14-17456]	EMA. Assessment report: Jardiance. EMA/CHMP/137741/2014. London: European Medicines Agency, Committee for Medicinal Products for Human Use; 20 March 2014. Available at: website: ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002677/WC500168594.pdf. Accessed 09 October 2014.
[R12-1620]	EMA. European Medicines Agency recommends suspension of Avandia, Avandamet and Avaglim (press release). 23 September 2010. Available at: website: ema.europa.eu/ema/index.jsp?curl = pages/news_and_events/news/2010/09/news_detail_001119.jsp&mid = WC0b01ac058004d5c1. Accessed 09 October 2014.
[R13-1970]	EMA. Guideline on good pharmacovigilance practices (GVP). Module VI – Management and reporting of adverse reactions to medicinal products. European Medicines Agency; 22 June 2012. Available at: website: emea.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129135.pdf. Accessed 7 February 2013.
[R13-5420]	EMA. Guideline on good pharmacovigilance practices (GVP). Module VIII – Post-authorisation safety studies. European Medicines Agency; 9 July 2012. Available at: website: emea.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129137.pdf. Accessed 7 February 2013.
[R14-5282]	ENCePP. Guide on methodological standards in

	pharmacoepidemiology (revision 3). EMA/95098/2010 Rev.3. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance; 2014. Available at: website: encepp.eu/standards_and_guidances/methodologicalGuide.shtml. Accessed 14 July 2014.
[R13-1395]	ENCePP. ENCePP checklist for study protocols (revision 2). European Network of Centres for Pharmacoepidemiology and Pharmacovigilance; 14 January 2013. Available at: website: encepp.eu/standards_and_guidances/index.shtml. Accessed 7 February 2013.
[R15-5707]	Erondu N, Desai M, Ways K, Meininger G. Diabetic Ketoacidosis and Related Events in the Canagliflozin Type 2 Diabetes Clinical Program. Diabetes Care. 2015 Sep;38(9):1680-6.
[R14-5246]	European Commission. Commission implementing regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council. 2012. Available at: website: eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri = OJ:L:2012:159:0005:0025:EN:PDF. Accessed 25 February 2014.
[P16-03420]	European Commission. SGLT2 inhibitors and diabetic ketoacidosis. Notification to the PRAC/EMA Secretariat of a referral under Article 20 of Regulation (EC) 726/2004. 10 June 2015. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Referrals_document/SGLT2_inhibitors__20/Procedure_started/WC500187925.pdf. Accessed 18 March 2016.
[R14-5313]	Eurostat. Population data for Europe. 2014. website: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&language=en&pcode=tps00001&plugin=1 (access date: 09 October 2014).
[R14-5247]	Faculty of Sexual and Reproductive Healthcare. Management of vaginal discharge in non-genitourinary medicine settings. London (UK): Faculty of Sexual and Reproductive Healthcare, British Association for Sexual Health and HIV; 28 February 2012. Available at: website: guideline.gov/content.aspx?id = 36683. Accessed 09 October 2014.
[R14-5260]	Fakjian N, Hunter S, Cole GW, Miller J. An argument for circumcision. Prevention of balanitis in the adult. Arch Dermatol 1990; 126 (8): 1046-7.
[P09-12413]	FDA. Guidance for industry. Drug-induced liver injury: premarketing clinical evaluation. Silver Spring (MD): US Department of Health and Human Services, Food and Drug Administration, Center for Drug

	Evaluation and Research (CDER), Center for Biologics Evaluation and Research(CBER); July 2009. website: http://www.fda.gov/downloads/Drugs/.../Guidances/UCM174090.pdf (access date: 26 November 2014).
[R16-1981]	Fishbein H, Palumbo P. Acute metabolic complications in diabetes. Diabetes in America. 1995;2:283-92.
[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.
[R14-5259]	Foxman B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am J Med 2002; 113 Suppl 1A: 5S-13S.
[R14-5258]	Foxman B, Klemstine KL, Brown PD. Acute pyelonephritis in US hospitals in 1997: hospitalization and in-hospital mortality. Ann Epidemiol. 2003 Feb; 13 (2): 144-50.
[R14-5287]	Gaieski DF, Edwards JM, Kallan MJ, Carr BG. Benchmarking the incidence and mortality of severe sepsis in the United States. Crit Care Med 2013; 41 (5): 1167-74.
[P14-17413]	Gallagher AM, Smeeth L, Seabroke S, Leufkens HG, van Staa TP. Risk of death and cardiovascular outcomes with thiazolidinediones: a study with the general practice research database and secondary care data. PLoS One 2011; 6 (12): e28157.
[P14-17376]	García-Rodríguez LA, Massó-González EL, Wallander MA, Johansson S. The safety of rosuvastatin in comparison with other statins in over 100,000 statin users in UK primary care. Pharmacoepidemiol Drug Saf 2008; 17 (10): 943-52.
[P14-17377]	García Rodríguez LA, González-Pérez A, Stang MR, Wallander MA, Johansson S. The safety of rosuvastatin in comparison with other statins in over 25,000 statin users in the Saskatchewan Health Databases. Pharmacoepidemiol Drug Saf 2008; 17 (10): 953-61.
[P14-02878]	Geerlings S, Fonseca V, Castro-Diaz D, List J, Parikh S. Genital and urinary tract infections in diabetes: impact of pharmacologically-induced glucosuria. Diabetes Res Clin Pract 2014; 103 (3): 373-81.
[R14-5279]	Gelfand JM, Margolis DJ, Dattani H. The UK General Practice Research Database. In: Strom BL, editor. Pharmacoepidemiology, 4th ed: John Wiley & Sons; 2005. p. 337-46.
[R11-5319]	Girman CJ, Kou TD, Brodovicz K, Alexander CM, O'Neill EA, Engel S, et al. Risk of acute renal failure in patients with Type 2 diabetes mellitus. Diabet Med 2012; 29 (5): 614-21.

[P12-04844]	Glynn RJ, Schneeweiss S, Sturmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. Basic Clin Pharmacol Toxicol 2006; 98 (3): 253-9.
[P14-17373]	Goettsch WG, Heintjes EM, Kastelein JJ, Rabelink TJ, Johansson S, Herings RM. Results from a rosuvastatin historical cohort study in more than 45,000 Dutch statin users, a PHARMO study. Pharmacoepidemiol Drug Saf 2006; 15 (7): 435-43.
[R11-5320]	Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. J Epidemiol Community Health 2009; 63 (4): 332-6.
[R16-1371]	Gosmanov AR, Gosmanova EO, Dillard-Cannon E. Management of adult diabetic ketoacidosis. Diabetes Metab Syndr Obes. 2014;7:255-64.
[P02-05226]	Griffin MR, Yared A, Ray WA. Nonsteroidal antiinflammatory drugs and acute renal failure in elderly persons. Am J Epidemiol 2000; 151 (5): 488-96.
[P15-00346]	Hay JE. Liver disease in pregnancy. Hepatology. 2008 Mar; 47 (3): 1067-76.
[R11-5211]	Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. Br J Clin Pharmacol 2010; 69 (1): 4-14.
[R12-3639]	Hirji I, Andersson SW, Guo Z, Hammar N, Gomez-Caminero A. Incidence of genital infection among patients with type 2 diabetes in the UK General Practice Research Database. J Diabetes Complications 2012; 26 (6): 501-5.
[R12-5226]	Hirji I, Guo Z, Andersson SW, Hammar N, Gomez-Caminero A. Incidence of urinary tract infection among patients with type 2 diabetes in the UK General Practice Research Database (GPRD). J Diabetes Complications 2012; 26 (6): 513-6.
[P03-03701]	Huerta C, Zhao SZ, García Rodríguez LA. Risk of acute liver injury in patients with diabetes. Pharmacotherapy. 2002 Sep; 22(9):1091-6.
[P05-04032]	Huerta C, Castellsague J, Varas-Lorenzo C, García Rodríguez LA. Nonsteroidal anti-inflammatory drugs and risk of ARF in the general population. Am J Kidney Dis 2005; 45 (3): 531-9.
[P03-00488]	Ibáñez L, Pérez E, Vidal X, Laporte JR, Grup d'Estudi Multicènteric d'Hepatotoxicitat Aguda de Barcelona. Prospective surveillance of acute serious liver disease unrelated to infectious, obstructive, or metabolic diseases: epidemiological and clinical features, and exposure to drugs. J Hepatol 2002; 37 (5): 592-600.
[R11-2259]	ICH. Pharmacovigilance planning. E2E. International Conference on

	Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use; 2004. Available at: website: ich.org/products/guidelines/efficacy/efficacy-
[R13-5418]	ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. International Committee of Medical Journal Editors; August 2013. Available at: website: icmje.org/urm_main.html. Accessed 25 February 2014.
[R14-5257]	ISAC. Annual report Jan 2013-Dec 2013. Independent Scientific Advisory Committee for Medicines and Healthcare Products Regulatory Agency (MHRA) for database research; 2013. Available at: website: mhra.gov.uk/home/groups/pl-a/documents/committeedocument/con448379.pdf. Accessed 14 October 2014.
[R11-4318]	ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Revision 2. International Society for Pharmacoepidemiology; 2007. Available at: website: pharmacoepi.org/resources/guidelines_08027.cfm. Accessed 07 February 2013.
[R15-3139]	Jameson K, Jick S, Hagberg KW, Ambegaonkar B, Giles A, O'Donoghue D. Prevalence and management of chronic kidney disease in primary care patients in the UK. Int J Clin Pract 2014; 68 (9): 1110-21.
[R99-1044]	Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. BMJ 1991; 302 (6779): 766-8.
[R11-2162]	Jick SS, Kaye JA, Vasilakis-Scaramozza C, García Rodríguez LA, Ruigómez A, Meier CR, et al. Validity of the general practice research database. Pharmacotherapy 2003; 23 (5): 686-9.
[R15-3140]	Johannesdottir SA, Horváth-Puhó E, Ehrenstein V, Schmidt M, Pedersen L, Sørensen HT. Existing data sources for clinical epidemiology: the Danish National Database of Reimbursed Prescriptions. Clin Epidemiol 2012; 4: 303-13.
[R16-1372]	Joint British Diabetes Societies, Inpatient Care Group. The management of diabetic ketoacidosis in adults. 2013. Available at: https://www.diabetes.org.uk/Documents/About%20Us/What%20we%20say/Management-of-DKA-241013.pdf. Accessed 16 March 2016.
[R12-1088]	Joshi N, Caputo GM, Weitekamp MR, Karchmer AW. Infections in patients with diabetes mellitus. N Engl J Med 1999; 341 (25): 1906-12.
[R11-5210]	Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. Br J

	Gen Pract 2010; 60 (572): e128-36.
[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.
[R15-3141]	Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. Scand J Public Health 2011; 39 (7 Suppl): 38-41.
[R15-2058]	Kitabchi AE, Umpierrez GE, Murphy MB, Kreisberg RA. Hyperglycemic crises in adult patients with diabetes: a consensus statement from the American Diabetes Association. Diabetes Care. 2006 Dec;29(12):2739-48.
[R14-5373]	Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Dordrecht ; New York: Springer; 2009.
[R14-5244]	Leal I, Romio SA, Schuemie M, Oteri A, Sturkenboom M, Trifiro G. Prescribing pattern of glucose lowering drugs in the United Kingdom in the last decade: a focus on the effects of safety warnings about rosiglitazone. Br J Clin Pharmacol 2013; 75 (3): 861-8.
[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.
[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.
[R14-5248]	Lisboa C, Santos A, Dias C, Azevedo F, Pina-Vaz C, Rodrigues A. Candida balanitis: risk factors. J Eur Acad Dermatol Venereol 2010; 24 (7): 820-6.
[P12-13528]	Lo Re V, 3rd, Haynes K, Ming EE, Wood Ives J, Horne LN, Fortier K, et al. Safety of saxagliptin: rationale for and design of a series of postmarketing observational studies. Pharmacoepidemiol Drug Saf 2012; 21 (11): 1202-15.
[P14-16383]	Macedo AF, Douglas I, Smeeth L, Forbes H, Ebrahim S. Statins and the risk of type 2 diabetes mellitus: cohort study using the UK clinical practice pesearch datalink. BMC Cardiovasc Disord. 2014; 14: 85.
[R14-5249]	Morgan CL, Poole CD, Evans M, Barnett AH, Jenkins-Jones S, Currie CJ. What next after metformin? A retrospective evaluation of the outcome of second-line, glucose-lowering therapies in people with type 2 diabetes. J Clin Endocrinol Metab 2012; 97 (12): 4605-12.

[R12-1100]	Muller LM, Gorter KJ, Hak E, Goudzwaard WL, Schellevis FG, Hoepelman AI, et al. Increased risk of common infections in patients with type 1 and type 2 diabetes mellitus. Clin Infect Dis 2005; 41 (3): 281-8.
[P06-02059]	Navarro VJ, Senior JR. Drug-related hepatotoxicity. N Engl J Med. 2006 Feb 16; 354 (7): 731-9.
[P14-17374]	NICE. Type 2 diabetes: the management of type 2 diabetes. NICE clinical guideline 87. National Institute for Health and Care Excellence; 2014. Available at: website: nice.org.uk/guidance/cg87/resources/guidance-type-2-diabetes-pdf. Accessed 09 October 2014.
[R13-5134]	NICE. Dapagliflozin in combination therapy for treating type 2 diabetes. NICE technology appraisal guidance 288. National Institute for Health and Care Excellence; 2013. Available at: website: nice.org.uk/guidance/ta288/resources/guidance-dapagliflozin-in-combination-therapy-for-treating-type2-diabetes-pdf. Accessed 09 October 2014.
[R14-5251]	Nicolle LE, Friesen D, Harding GK, Roos LL. Hospitalization for acute pyelonephritis in Manitoba, Canada, during the period from 1989 to 1992; impact of diabetes, pregnancy, and aboriginal origin. Clin Infect Dis 1996; 22 (6): 1051-6.
[R14-5252]	Patel DA, Burnett NM, Curtis KM. Reproductive tract infections. Reproductive Health Epidemiology Series—Module 3. Atlanta (GA): Centers for Disease Control and Prevention; 2003. Available at: website: cdc.gov/reproductivehealth/ProductsPubs/PDFs/Epi_Module_03A_Tag508.pdf. Accessed 09 October 2014.
[R14-5253]	Patkar NM, Curtis JR, Teng GG, Allison JJ, Saag M, Martin C, et al. Administrative codes combined with medical records based criteria accurately identified bacterial infections among rheumatoid arthritis patients. J Clin Epidemiol. 2009 Mar; 62 (3): 321-7, 7 e1-7.
[R14-4378]	Patorno E, Patrick AR, Garry EM, Schneeweiss S, Gillet VG, Bartels DB, et al. Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations. Diabetologia 2014; 57 (11): 2237-50.
[R14-5254]	Peer AK, Hoosen AA, Seedat MA, van den Ende J, Omar MA. Vaginal yeast infections in diabetic women. S Afr Med J 1993; 83 (10): 727-9.
[P14-17372]	Pérez Gutthann S, García Rodríguez LA, Raiford DS, Duque Oliart A, Ris Romeu J. Nonsteroidal anti-inflammatory drugs and the risk of hospitalization for acute renal failure. Arch Intern Med 1996; 156

	(21): 2433-9.
[P14-17457]	Rawson NS. Review of the quality of observational studies of the association between rosiglitazone and acute myocardial infarction. J Popul Ther Clin Pharmacol 2014; 21 (2): e214-32.
[R12-1479]	Retnakaran R, Cull CA, Thorne KI, Adler AI, Holman RR, Group US. Risk factors for renal dysfunction in type 2 diabetes: U.K. Prospective Diabetes Study 74. Diabetes 2006; 55 (6): 1832-9.
[P15-11541]	Roach P, Skierczynski P. Euglycemic diabetic ketoacidosis in a patient with type 2 diabetes after treatment with empagliflozin. Diabetes Care. 2016 Jan;39(1):e3.
[P15-08785]	Rosenstock J, Ferrannini E. Euglycemic diabetic ketoacidosis: a predictable, detectable, and preventable safety concern with SGLT2 inhibitors. Diabetes Care. 2015 Sep;38(9):1638-42.
[R14-1393]	Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990 Jan; 1 (1): 43-6.
[R14-2938]	Rothman KJ. Six persistent research misconceptions. J Gen Intern Med. 2014 Jul; 29 (7): 1060-4.
[R14-5255]	Rubino A, McQuay LJ, Gough SC, Kvasz M, Tennis P. Delayed initiation of subcutaneous insulin therapy after failure of oral glucose-lowering agents in patients with type 2 diabetes: a population-based analysis in the UK. Diabet Med 2007; 24 (12): 1412-8.
[R16-1373]	Savage MW, Dhatariya KK, Kilvert A, Rayman G, Rees JA, Courtney CH, et al. Joint British Diabetes Societies guideline for the management of diabetic ketoacidosis. Diabet Med. 2011 May;28(5):508-15.
[R13-1120]	Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. Pharmacoepidemiol Drug Saf 2010; 19 (8): 858-68.
[R14-5389]	Segal JB, Griswold M, Achy-Brou A, Herbert R, Bass EB, Dy SM, et al. Using propensity scores subclassification to estimate effects of longitudinal treatments: an example using a new diabetes medication. Med Care 2007; 45 (10 Supl 2): S149-S57.
[P02-05969]	Sgro C, Clinard F, Ouazir K, Chanay H, Allard C, Guilleminet C, et al. Incidence of drug-induced hepatic injuries: a French population-based study. Hepatology 2002; 36 (2): 451-5.
[R10-6632]	Shah BR, Hux JE. Quantifying the risk of infectious diseases for people with diabetes. Diabetes Care 2003; 26 (2): 510-3.
[R15-3142]	SIDIAP. Database website. General details. 2014. website: http://www.sidiap.org/index.php?option=com_content&view=article&id=61&Itemid=11&lang=en (access date: 3 June 2014).

[R13-3589]	Southern DA, Quan H, Ghali WA. Comparison of the Elixhauser and Charlson/Deyo methods of comorbidity measurement in administrative data. Med Care 2004; 42 (4): 355-60.
[R14-5281]	Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009; 338: b2393.
[P14-17375]	Temple RJ. Hepatotoxicity through the years: impact on the FDA (presentation). 2001. website: http://www.fda.gov/downloads/Drugs/ScienceResearch/ResearchAreas/ucm122149.pdf (access date: 09 October 2014).
[R14-5256]	Temple R. Hy's law: predicting serious hepatotoxicity. Pharmacoepidemiol Drug Saf 2006; 15 (4): 241-3.
[R16-1374]	The NHS Information Centre. National Diabetes Audit executive summary 2009-2010. 2011. Available at: http://www.hscic.gov.uk/catalogue/PUB02573/nati-diab-audi-09-10-exec-summ.pdf. Accessed 17 March 2016.
[R03-0585]	UK Prospective Diabetes Study Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. UK Prospective Diabetes Study Group. BMJ 1998; 317 (7160): 703-13.
[R14-5280]	Van Staa T-P, Abenhaim L. The quality of information recorded on a UK database of primary care records: a study of hospitalizations due to hypoglycemia and other conditions. Pharmacoepidemiol Drug Saf 1994; 3, 15-21.
[R12-1105]	Venmans LM, Hak E, Gorter KJ, Rutten GE. Incidence and antibiotic prescription rates for common infections in patients with diabetes in primary care over the years 1995 to 2003. Int J Infect Dis 2009; 13 (6): e344-51.
[R13-2485]	von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008; 61 (4): 344-9.
[R14-5286]	Wagenlehner FM, Pilatz A, Naber KG, Weidner W. Therapeutic challenges of urosepsis. Eur J Clin Invest 2008; 38 Suppl 2: 45-9.
[R14-5285]	Waikar SS, Curhan GC, Wald R, McCarthy EP, Chertow GM. Declining mortality in patients with acute renal failure, 1988 to 2002. J Am Soc Nephrol 2006; 17 (4): 1143-50.
[R14-3272]	Wang ZH, Kihl-Selstam E, Eriksson JW. Ketoacidosis occurs in both type 1 and type 2 diabetes—a population-based study from northern

	Sweden. Diabet Med. 2008 Jul;25(7):867-70.
[R16-1375]	Westphal SA. The occurrence of diabetic ketoacidosis in non-insulin-dependent diabetes and newly diagnosed diabetic adults. Am J Med. 1996 Jul;101(1):19-24.
[R11-5329]	Xue JL, Daniels F, Star RA, Kimmel PL, Eggers PW, Molitoris BA, et al. Incidence and mortality of acute renal failure in Medicare beneficiaries, 1992 to 2001. J Am Soc Nephrol 2006; 17 (4): 1135-42.
[R05-1093]	Zimmerman HJ. Drug-induced liver disease. In: Hepatotoxicity, the adverse effects of drugs and other chemicals on the liver, 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 1999. p. 428-33.

13.2 UNPUBLISHED REFERENCES

None.

Annex 1. LIST OF STAND-ALONE DOCUMENTS

None.

Annex 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

Doc.Ref. EMEA/540136/2009

ENCePP Checklist for Study Protocols (Revision 2, amended)

Adopted by the ENCePP Steering Group on 14/01/2013

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the page number(s) of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). Note, the Checklist is a supporting document and does not replace the format of the protocol for PASS as recommended in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

Study title:

Post-authorisation safety study in patients with T2D to assess the risk of acute liver injury, hospitalisation for acute kidney injury, hospitalisation for urinary tract infection, and the risk of genital infections, treated with empagliflozin compared to patients treated with other SGLT2 inhibitors and with DPP-4 inhibitors

Study reference number:

ENCEPP/SDPP/13413; http://www.encepp.eu/encepp/viewResource.htm?id=13414

Section 1: Milestones	**Yes**	**No**	**N/A**	**Page Number(s)**

1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection‡	☒	☐	☐	16
1.1.2 End of data collection§	☒	☐	☐	16
1.1.3 Study progress report(s)	☐	☒	☐	
1.1.4 Interim progress report(s)	☒	☐	☐	16
1.1.5 Registration in the EU PAS Register	☒	☐	☐	16
1.1.6 Final report of study results	☒	☐	☐	16

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
2.1 Does the formulation of the research question and objectives clearly explain:				
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	17-18
2.1.2 The objectives of the study?	☒	☐	☐	25
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	30
2.1.4 Which formal hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Page Number(s)**
3.1 Is the study design described? (e.g. cohort, case-control, randomised controlled trial, new or alternative design)	☒	☐	☐	29-30
3.2 Does the protocol specify the primary and secondary (if applicable) endpoint(s) to be investigated?	☒	☐	☐	38-50
3.3 Does the protocol describe the measure(s) of effect? (e.g. relative risk, odds ratio, deaths per 1000 person-years, absolute risk, excess risk, incidence rate ratio, hazard ratio, number needed to harm (NNH) per year)	☒	☐	☐	60-66

Comments:

‡ Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.

§ Date from which the analytical dataset is completely available.

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Page Number(s)**
4.1 Is the source population described?	☒	☐	☐	30
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period?	☒	☐	☐	30-31
4.2.2 Age and sex?	☒	☐	☐	31
4.2.3 Country of origin?	☒	☐	☐	30
4.2.4 Disease/indication?	☒	☐	☐	30
4.2.5 Co-morbidity?	☐	☒	☐	
4.2.6 Seasonality?	☐	☒	☐	
4.3 Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	31-34

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
5.1 Does the protocol describe how exposure is defined and measured? (e.g. operational details for defining and categorising exposure)	☒	☐	☐	35-38
5.2 Does the protocol discuss the validity of exposure measurement? (e.g. precision, accuracy, prospective ascertainment, exposure information recorded before the outcome occurred, use of validation sub-study)	☒	☐	☐	35-38
5.3 Is exposure classified according to time windows? (e.g. current user, former user, non-use)	☒	☐	☐	35-38
5.4 Is exposure classified based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☒	☐	
5.5 Does the protocol specify whether a dose-dependent or duration-dependent response is measured?	☒	☐	☐	62-63

Comments:

Section 6: Endpoint definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
6.1 Does the protocol describe how the endpoints are defined and measured?	☒	☐	☐	38-50
6.2 Does the protocol discuss the validity of endpoint measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, prospective or retrospective ascertainment, use of validation sub-study)	☒	☐	☐	38-50

Comments:

Section 7: Confounders and effect modifiers	**Yes**	**No**	**N/A**	**Page Number(s)**
7.1 Does the protocol address known confounders? (e.g. collection of data on known confounders, methods of controlling for known confounders)	☒	☐	☐	50-52, 60-62, and 63
7.2 Does the protocol address known effect modifiers? (e.g. collection of data on known effect modifiers, anticipated direction of effect)	☒	☐	☐	50-52, 60-62, and 63

Comments:

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
8.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview, etc.)	☒	☐	☐	52-58 and Annex 7
8.1.2 Endpoints? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics, etc.)	☒	☐	☐	52-58 and Annex 7
8.1.3 Covariates?	☒	☐	☐	52-58 and Annex 7
8.2 Does the protocol describe the information available from the data source(s) on:				
8.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	52-58 and Annex 7
8.2.2 Endpoints? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	52-58 and Annex 7
8.2.3 Covariates? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, life style, etc.)	☒	☐	☐	52-58 and Annex 7
8.3 Is a coding system described for:				
8.3.1 Diseases? (e.g. International Classification of Diseases (ICD)-10)	☒	☐	☐	38-50, 52-58, and Annex 4
8.3.2 Endpoints? (e.g. Medical Dictionary for Regulatory Activities(MedDRA) for adverse events)	☒	☐	☐	38-50, 52-58, and Annex 5
8.3.3 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	52-58
8.4 Is the linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	52-58

Comments:

Section 9: Study size and power	**Yes**	**No**	**N/A**	**Page Number(s)**
9.1 Is sample size and/or statistical power calculated?	☒	☐	☐	58-60

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Page Number(s)**
10.1 Does the plan include measurement of excess risks?	☒	☐	☐	60-66
10.2 Is the choice of statistical techniques described?	☒	☐	☐	60-66
10.3 Are descriptive analyses included?	☒	☐	☐	60-66
10.4 Are stratified analyses included?	☒	☐	☐	60-66
10.5 Does the plan describe the methods for adjusting for confounding?	☒	☐	☐	60-66
10.6 Does the plan describe methods addressing effect modification?	☒	☐	☐	60-66

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Page Number(s)**
11.1 Is information provided on the management of missing data?	☒	☐	☐	64-65
11.2 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	66-66
11.3 Are methods of quality assurance described?	☒	☐	☐	66-66
11.4 Does the protocol describe possible quality issues related to the data source(s)?	☒	☐	☐	66-70
11.5 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Page Number(s)**
12.1 Does the protocol discuss:				
12.1.1 Selection biases?	☒	☐	☐	66-70
12.1.2 Information biases? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	☐	☒	☐	

12.2 Does the protocol discuss study feasibility? (e.g. sample size, anticipated exposure, duration of follow-up in a cohort study, patient recruitment)	☒	☐	☐	58-60
12.3 Does the protocol address other limitations?	☒	☐	☐	66-70

Comments:

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.1 Have requirements of Ethics Committee/ Institutional Review Board approval been described?	☒	☐	☐	70-72
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☒	☐	
13.3 Have data protection requirements been described?	☒	☐	☐	70-72

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Page Number(s)**
14.1 Does the protocol include a section to document future amendments and deviations?	☒	☐	☐	15

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Page Number(s)**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	73
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	73

Comments:

Name of the main author of the protocol: Manel Pladevall-Vila and Cristina Rebordosa

Date: 21 June 2016

Signature: _____

Annex 3. Read codes to identify type 2 diabetes

Read code	Read term
66A3.00	Diabetic on diet only
66A4.00	Diabetic on oral treatment
66AH000	Conversion to insulin
66AH100	Conversion to insulin in secondary care
66AH200	Conversion to insulin by diabetes specialist nurse
66Ao.00	Diabetes type 2 review
66At100	Type II diabetic dietary review
66At111	Type 2 diabetic dietary review
66AV.00	Diabetic on insulin and oral treatment
8I2P.00	Sulphonylureas contraindicated
8I2S.00	Glitazones contraindicated
C100112	Non-insulin dependent diabetes mellitus
C107400	NIDDM with peripheral circulatory disorder
C109.00	Non-insulin dependent diabetes mellitus
C109.11	NIDDM - Non-insulin dependent diabetes mellitus
C109.12	Type 2 diabetes mellitus
C109.13	Type II diabetes mellitus
C109000	Non-insulin-dependent diabetes mellitus with renal comps
C109011	Type II diabetes mellitus with renal complications
C109012	Type 2 diabetes mellitus with renal complications
C109100	Non-insulin-dependent diabetes mellitus with ophthalm comps
C109111	Type II diabetes mellitus with ophthalmic complications
C109112	Type 2 diabetes mellitus with ophthalmic complications
C109200	Non-insulin-dependent diabetes mellitus with neuro comps
C109211	Type II diabetes mellitus with neurological complications
C109212	Type 2 diabetes mellitus with neurological complications
C109300	Non-insulin-dependent diabetes mellitus with multiple comps
C109312	Type 2 diabetes mellitus with multiple complications
C109400	Non-insulin dependent diabetes mellitus with ulcer
C109411	Type II diabetes mellitus with ulcer
C109412	Type 2 diabetes mellitus with ulcer
C109500	Non-insulin dependent diabetes mellitus with gangrene

C109511	Type II diabetes mellitus with gangrene
C109512	Type 2 diabetes mellitus with gangrene
C109600	Non-insulin-dependent diabetes mellitus with retinopathy
C109611	Type II diabetes mellitus with retinopathy
C109612	Type 2 diabetes mellitus with retinopathy
C109700	Non-insulin dependent diabetes mellitus - poor control
C109711	Type II diabetes mellitus - poor control
C109712	Type 2 diabetes mellitus - poor control
C109900	Non-insulin-dependent diabetes mellitus without complication
C109911	Type II diabetes mellitus without complication
C109912	Type 2 diabetes mellitus without complication
C109A00	Non-insulin dependent diabetes mellitus with mononeuropathy
C109A11	Type II diabetes mellitus with mononeuropathy
C109B00	Non-insulin dependent diabetes mellitus with polyneuropathy
C109B11	Type II diabetes mellitus with polyneuropathy
C109B12	Type 2 diabetes mellitus with polyneuropathy
C109C00	Non-insulin dependent diabetes mellitus with nephropathy
C109C11	Type II diabetes mellitus with nephropathy
C109C12	Type 2 diabetes mellitus with nephropathy
C109D00	Non-insulin dependent diabetes mellitus with hypoglyca coma
C109D11	Type II diabetes mellitus with hypoglycaemic coma
C109D12	Type 2 diabetes mellitus with hypoglycaemic coma
C109E00	Non-insulin depend diabetes mellitus with diabetic cataract
C109E11	Type II diabetes mellitus with diabetic cataract
C109E12	Type 2 diabetes mellitus with diabetic cataract
C109F00	Non-insulin-dependent d m with peripheral angiopath
C109F11	Type II diabetes mellitus with peripheral angiopathy
C109F12	Type 2 diabetes mellitus with peripheral angiopathy
C109G00	Non-insulin dependent diabetes mellitus with arthropathy
C109G11	Type II diabetes mellitus with arthropathy
C109G12	Type 2 diabetes mellitus with arthropathy
C109H00	Non-insulin dependent d m with neuropathic arthropathy
C109H11	Type II diabetes mellitus with neuropathic arthropathy
C109H12	Type 2 diabetes mellitus with neuropathic arthropathy
C109J00	Insulin treated Type 2 diabetes mellitus

C109J11	Insulin treated non-insulin dependent diabetes mellitus
C109J12	Insulin treated Type II diabetes mellitus
C109K00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus
C10D.11	Maturity onset diabetes in youth type 2
C10F.00	Type 2 diabetes mellitus
C10F.11	Type II diabetes mellitus
C10F000	Type 2 diabetes mellitus with renal complications
C10F011	Type II diabetes mellitus with renal complications
C10F100	Type 2 diabetes mellitus with ophthalmic complications
C10F111	Type II diabetes mellitus with ophthalmic complications
C10F200	Type 2 diabetes mellitus with neurological complications
C10F211	Type II diabetes mellitus with neurological complications
C10F300	Type 2 diabetes mellitus with multiple complications
C10F311	Type II diabetes mellitus with multiple complications
C10F400	Type 2 diabetes mellitus with ulcer
C10F411	Type II diabetes mellitus with ulcer
C10F500	Type 2 diabetes mellitus with gangrene
C10F511	Type II diabetes mellitus with gangrene
C10F600	Type 2 diabetes mellitus with retinopathy
C10F611	Type II diabetes mellitus with retinopathy
C10F700	Type 2 diabetes mellitus - poor control
C10F711	Type II diabetes mellitus - poor control
C10F900	Type 2 diabetes mellitus without complication
C10F911	Type II diabetes mellitus without complication
C10FA00	Type 2 diabetes mellitus with mononeuropathy
C10FA11	Type II diabetes mellitus with mononeuropathy
C10FB00	Type 2 diabetes mellitus with polyneuropathy
C10FB11	Type II diabetes mellitus with polyneuropathy
C10FC00	Type 2 diabetes mellitus with nephropathy
C10FC11	Type II diabetes mellitus with nephropathy
C10FD00	Type 2 diabetes mellitus with hypoglycaemic coma
C10FD11	Type II diabetes mellitus with hypoglycaemic coma
C10FE00	Type 2 diabetes mellitus with diabetic cataract
C10FE11	Type II diabetes mellitus with diabetic cataract
C10FF00	Type 2 diabetes mellitus with peripheral angiopathy

C10FF11	Type II diabetes mellitus with peripheral angiopathy
C10FG00	Type 2 diabetes mellitus with arthropathy
C10FG11	Type II diabetes mellitus with arthropathy
C10FH00	Type 2 diabetes mellitus with neuropathic arthropathy
C10FH11	Type II diabetes mellitus with neuropathic arthropathy
C10FJ00	Insulin treated Type 2 diabetes mellitus
C10FJ11	Insulin treated Type II diabetes mellitus
C10FK00	Hyperosmolar non-ketotic state in type 2 diabetes mellitus
C10FK11	Hyperosmolar non-ketotic state in type II diabetes mellitus
C10FL00	Type 2 diabetes mellitus with persistent proteinuria
C10FL11	Type II diabetes mellitus with persistent proteinuria
C10FM00	Type 2 diabetes mellitus with persistent microalbuminuria
C10FM11	Type II diabetes mellitus with persistent microalbuminuria
C10FN00	Type 2 diabetes mellitus with ketoacidosis
C10FN11	Type II diabetes mellitus with ketoacidosis
C10FP00	Type 2 diabetes mellitus with ketoacidotic coma
C10FP11	Type II diabetes mellitus with ketoacidotic coma
C10FQ00	Type 2 diabetes mellitus with exudative maculopathy
C10FR00	Type 2 diabetes mellitus with gastroparesis
C10P100	Type II diabetes mellitus in remission
L180600	Pre-existing diabetes mellitus, non-insulin-dependent
TJ23200	Adverse reaction to chlorpropamide
TJ23300	Adverse reaction to glibenclamide
TJ23400	Adverse reaction to gliclazide
TJ23500	Adverse reaction to glipizide
TJ23800	Adverse reaction to tolazamide
TJ23900	Adverse reaction to tolbutamide
TJ23A00	Adverse reaction to metformin hydrochloride
U602314	[X] Adverse reaction to chlorpropamide
U602315	[X] Adverse reaction to glibenclamide
U602316	[X] Adverse reaction to gliclazide
U602317	[X] Adverse reaction to glipzide
U602318	[X] Adverse reaction to gliquidone
U60231A	[X] Adverse reaction to tolazamide
U60231B	[X] Adverse reaction to tolbutamide

U60231C	[X] Adverse reaction to metformin hydrochloride
ZC2CA00	Dietary advice for type II diabetes
ZV6DB00	[V]Admitted for conversion to insulin
1434.00	H/O: diabetes mellitus
3881.00	Education score - diabetes
3882.00	Diabetes well being questionnaire
3883.00	Diabetes treatment satisfaction questionnaire
7276.00	Pan retinal photocoagulation for diabetes
9360.00	Patient held diabetic record issued
13AB.00	Diabetic lipid lowering diet
13AC.00	Diabetic weight reducing diet
13B1.00	Diabetic diet
14F4.00	H/O: Admission in last year for diabetes foot problem
14P3.00	H/O: insulin therapy
1IA..00	No evidence of diabetic nephropathy
2BBF.00	Retinal abnormality - diabetes related
2BBk.00	O/E - right eye stable treated prolif diabetic retinopathy
2BBL.00	O/E - diabetic maculopathy present both eyes
2BBl.00	O/E - left eye stable treated prolif diabetic retinopathy
2BBM.00	O/E - diabetic maculopathy absent both eyes
2BBo.00	O/E - sight threatening diabetic retinopathy
2BBP.00	O/E - right eye background diabetic retinopathy
2BBQ.00	O/E - left eye background diabetic retinopathy
2BBR.00	O/E - right eye preproliferative diabetic retinopathy
2BBr.00	Impaired vision due to diabetic retinopathy
2BBS.00	O/E - left eye preproliferative diabetic retinopathy
2BBT.00	O/E - right eye proliferative diabetic retinopathy
2BBV.00	O/E - left eye proliferative diabetic retinopathy
2BBW.00	O/E - right eye diabetic maculopathy
2BBX.00	O/E - left eye diabetic maculopathy
2G51000	Foot abnormality - diabetes related
2G5A.00	O/E - Right diabetic foot at risk
2G5B.00	O/E - Left diabetic foot at risk
2G5C.00	Foot abnormality - diabetes related
2G5d.00	O/E - Left diabetic foot at increased risk

2G5E.00	O/E - Right diabetic foot at low risk
2G5e.00	O/E - Right diabetic foot at increased risk
2G5F.00	O/E - Right diabetic foot at moderate risk
2G5G.00	O/E - Right diabetic foot at high risk
2G5H.00	O/E - Right diabetic foot - ulcerated
2G5I.00	O/E - Left diabetic foot at low risk
2G5J.00	O/E - Left diabetic foot at moderate risk
2G5K.00	O/E - Left diabetic foot at high risk
2G5L.00	O/E - Left diabetic foot - ulcerated
2G5V.00	O/E - right chronic diabetic foot ulcer
2G5W.00	O/E - left chronic diabetic foot ulcer
42c..00	HbA1 - diabetic control
42W..00	Hb. A1C - diabetic control
42WZ.00	Hb. A1C - diabetic control NOS
661M400	Diabetes self-management plan agreed
661N400	Diabetes self-management plan review
66A..00	Diabetic monitoring
66A1.00	Initial diabetic assessment
66A2.00	Follow-up diabetic assessment
66A5.00	Diabetic on insulin
66A6.00	Last hypo. attack
66A7.00	Frequency of hypo. attacks
66A7000	Frequency of hospital treated hypoglycaemia
66A7100	Frequency of GP or paramedic treated hypoglycaemia
66A8.00	Has seen dietician - diabetes
66A9.00	Understands diet - diabetes
66Aa.00	Diabetic diet - poor compliance
66AA.11	Injection sites - diabetic
66Ab.00	Diabetic foot examination
66Ac.00	Diabetic peripheral neuropathy screening
66AD.00	Fundoscopy - diabetic check
66Ad.00	Hypoglycaemic attack requiring 3rd party assistance
66Ae.00	HbA1c target
66Ae000	HbA1c target level - IFCC standardised
66Af.00	Patient diabetes education review

66AG.00	Diabetic drug side effects
66Ag.00	Insulin needles changed daily
66AH.00	Diabetic treatment changed
66Ah.00	Insulin needles changed for each injection
66AI.00	Diabetic - good control
66Ai.00	Diabetic 6 month review
66AJ.00	Diabetic - poor control
66Aj.00	Insulin needles changed less than once a day
66AJ.11	Unstable diabetes
66AJ000	Chronic hyperglycaemia
66AJ300	Recurrent severe hypos
66AJz00	Diabetic - poor control NOS
66AK.00	Diabetic - cooperative patient
66Ak.00	Diabetic monitoring - lower risk albumin excretion
66AL.00	Diabetic-uncooperative patient
66Al.00	Diabetic monitoring - higher risk albumin excretion
66AM.00	Diabetic - follow-up default
66Am.00	Insulin dose changed
66AN.00	Date diabetic treatment start
66AO.00	Date diabetic treatment stopp.
66AP.00	Diabetes: practice programme
66Ap.00	Insulin treatment initiated
66AQ.00	Diabetes: shared care programme
66Aq.00	Diabetic foot screen
66AR.00	Diabetes management plan given
66AS.00	Diabetic annual review
66As.00	Diabetic on subcutaneous treatment
66AS000	Diabetes Year of Care annual review
66AT.00	Annual diabetic blood test
66At.00	Diabetic dietary review
66AU.00	Diabetes care by hospital only
66Au.00	Diabetic erectile dysfunction review
66Av.00	Diabetic assessment of erectile dysfunction
66AW.00	Diabetic foot risk assessment
66Aw.00	Insulin dose

66Ax.00	Checking accuracy of blood glucose meter
66AY.00	Diabetic diet - good compliance
66AZ.00	Diabetic monitoring NOS
66o..00	Further diabetic monitoring
671F000	Insulin alert pat information booklet information discussed
679c.00	Insulin administration education
679L.00	Health education - diabetes
679L000	Education in self management of diabetes
679L200	Education about diabetes and driving
679L211	Advice about diabetes and driving
679R.00	Patient offered diabetes structured education programme
67D8.00	Provision of diabetes clinical summary
67H9.00	Education about lifestyle for risk of diabetes
67IJ100	Pre-conception advice for diabetes mellitus
68A7.00	Diabetic retinopathy screening
68A9.00	Diabetic retinopathy screening offered
68AB.00	Diabetic digital retinopathy screening offered
7L19800	Subcutaneous injection of insulin
889A.00	Diab mellit insulin-glucose infus acute myocardial infarct
8A12.00	Diabetic crisis monitoring
8A13.00	Diabetic stabilisation
8B3l.00	Diabetes medication review
8BAi.00	Insulin passport completed
8BAm.00	Insulin passport checked
8BL2.00	Patient on maximal tolerated therapy for diabetes
8CA4100	Pt advised re diabetic diet
8CMW700	Diabetes clinical pathway
8CP2.00	Transition of diabetes care options discussed
8CR2.00	Diabetes clinical management plan
8CS0.00	Diabetes care plan agreed
8H2J.00	Admit diabetic emergency
8H3O.00	Non-urgent diabetic admission
8H4e.00	Referral to diabetes special interest general practitioner
8H4F.00	Referral to diabetologist
8H7C.00	Refer, diabetic liaison nurse

8H7f.00	Referral to diabetes nurse
8H7r.00	Refer to diabetic foot screener
8HBG.00	Diabetic retinopathy 12 month review
8HBH.00	Diabetic retinopathy 6 month review
8Hg4.00	Discharged from care of diabetes specialist nurse
8HgC.00	Discharged from diabetes shared care programme
8HHy.00	Referral to diabetic register
8Hj0.00	Referral to diabetes structured education programme
8Hj3.00	Referral to DAFNE diabetes structured education programme
8Hj4.00	Referral to DESMOND diabetes structured education programme
8Hj5.00	Referral to XPERT diabetes structured education programme
8HKE.00	Diabetology D.V. requested
8Hl1.00	Referral for diabetic retinopathy screening
8Hl4.00	Referral to community diabetes specialist nurse
8Hlc.00	Referral to community diabetes service
8HLE.00	Diabetology D.V. done
8HME.00	Listed for Diabetology admissn
8HTE100	Referral to community diabetes clinic
8HTi.00	Referral to multidisciplinary diabetic clinic
8HTk.00	Referral to diabetic eye clinic
8HVU.00	Private referral to diabetologist
8I3W.00	Diabetic foot examination declined
8I3X.00	Diabetic retinopathy screening refused
8IAs.00	Diabetic dietary review declined
8IEa.00	Referral to DAFNE diabetes structured educn prog declined
93C4.00	Patient consent given for addition to diabetic register
9b92000	Diabetic medicine
9h4..00	Exception reporting: diabetes quality indicators
9h41.00	Excepted from diabetes qual indicators: Patient unsuitable
9h42.00	Excepted from diabetes quality indicators: Informed dissent
9h43.00	Excepted from diabetes qual indicators: service unavailable
9kL..00	Insulin initiation - enhanced services administration
9m0..00	Diabetic retinopathy screening administrative status
9M00.00	Informed consent for diabetes national audit
9m00.00	Eligible for diabetic retinopathy screening

9M10.00	Informed dissent for diabetes national audit
9N0m.00	Seen in diabetic nurse consultant clinic
9N0n.00	Seen in community diabetes specialist clinic
9N0o.00	Seen in community diabetic specialist nurse clinic
9N1i.00	Seen in diabetic foot clinic
9N1o.00	Seen in multidisciplinary diabetic clinic
9N1Q.00	Seen in diabetic clinic
9N1v.00	Seen in diabetic eye clinic
9N2i.00	Seen by diabetic liaison nurse
9N4I.00	DNA - Did not attend diabetic clinic
9N4p.00	Did not attend diabetic retinopathy clinic
9NiA.00	Did not attend diabetes structured education programme
9NiC.00	Did not attend DAFNE diabetes structured education programme
9NM0.00	Attending diabetes clinic
9NN9.00	Under care of diabetes specialist nurse
9NND.00	Under care of diabetic foot screener
9OL..00	Diabetes monitoring admin.
9OL..11	Diabetes clinic administration
9OL1.00	Attends diabetes monitoring
9OL2.00	Refuses diabetes monitoring
9OL3.00	Diabetes monitoring default
9OL4.00	Diabetes monitoring 1st letter
9OL5.00	Diabetes monitoring 2nd letter
9OL6.00	Diabetes monitoring 3rd letter
9OL7.00	Diabetes monitor.verbal invite
9OL8.00	Diabetes monitor.phone invite
9OL9.00	Diabetes monitoring deleted
9OLA.00	Diabetes monitor. check done
9OLA.11	Diabetes monitored
9OLB.00	Attended diabetes structured education programme
9OLD.00	Diabetic patient unsuitable for digital retinal photography
9OLF.00	Diabetes structured education programme completed
9OLG.00	Attended XPERT diabetes structured education programme
9OLH.00	Attended DAFNE diabetes structured education programme
9OLJ.00	DAFNE diabetes structured education programme completed

9OLK.00	DESMOND diabetes structured education programme completed
9OLL.00	XPERT diabetes structured education programme completed
9OLM.00	Diabetes structured education programme declined
9OLN.00	Diabetes monitor invitation by SMS (short message service)
9OLZ.00	Diabetes monitoring admin.NOS
9Oy..00	Diabetes screening administration
9Oy0.00	Diabetes screening invitation
9Oy0000	Diabetic foot screening invitation
9Oy0200	Diabetic foot screening invitation first letter
9Oy0300	Diabetic foot screening invitation second letter
9Oy0400	Diabetic foot screening invitation third letter
C10..00	Diabetes mellitus
C100.00	Diabetes mellitus with no mention of complication
C100100	Diabetes mellitus, adult onset, no mention of complication
C100111	Maturity onset diabetes
C100z00	Diabetes mellitus NOS with no mention of complication
C101.00	Diabetes mellitus with ketoacidosis
C101100	Diabetes mellitus, adult onset, with ketoacidosis
C101y00	Other specified diabetes mellitus with ketoacidosis
C101z00	Diabetes mellitus NOS with ketoacidosis
C102.00	Diabetes mellitus with hyperosmolar coma
C102100	Diabetes mellitus, adult onset, with hyperosmolar coma
C102z00	Diabetes mellitus NOS with hyperosmolar coma
C103.00	Diabetes mellitus with ketoacidotic coma
C103100	Diabetes mellitus, adult onset, with ketoacidotic coma
C103y00	Other specified diabetes mellitus with coma
C103z00	Diabetes mellitus NOS with ketoacidotic coma
C104.00	Diabetes mellitus with renal manifestation
C104.11	Diabetic nephropathy
C104100	Diabetes mellitus, adult onset, with renal manifestation
C104y00	Other specified diabetes mellitus with renal complications
C104z00	Diabetes mellitus with nephropathy NOS
C105.00	Diabetes mellitus with ophthalmic manifestation
C105100	Diabetes mellitus, adult onset, + ophthalmic manifestation
C105y00	Other specified diabetes mellitus with ophthalmic complicatn

C105z00	Diabetes mellitus NOS with ophthalmic manifestation
C106.00	Diabetes mellitus with neurological manifestation
C106.11	Diabetic amyotrophy
C106.12	Diabetes mellitus with neuropathy
C106.13	Diabetes mellitus with polyneuropathy
C106100	Diabetes mellitus, adult onset, + neurological manifestation
C106y00	Other specified diabetes mellitus with neurological comps
C106z00	Diabetes mellitus NOS with neurological manifestation
C107.00	Diabetes mellitus with peripheral circulatory disorder
C107.11	Diabetes mellitus with gangrene
C107.12	Diabetes with gangrene
C107100	Diabetes mellitus, adult, + peripheral circulatory disorder
C107200	Diabetes mellitus, adult with gangrene
C107z00	Diabetes mellitus NOS with peripheral circulatory disorder
C108y00	Other specified diabetes mellitus with multiple comps
C108z00	Unspecified diabetes mellitus with multiple complications
C10A.00	Malnutrition-related diabetes mellitus
C10A000	Malnutrition-related diabetes mellitus with coma
C10A100	Malnutrition-related diabetes mellitus with ketoacidosis
C10A500	Malnutritn-relat diabetes melitus wth periph circul complctn
C10M.00	Lipoatrophic diabetes mellitus
C10y.00	Diabetes mellitus with other specified manifestation
C10y100	Diabetes mellitus, adult, + other specified manifestation
C10yy00	Other specified diabetes mellitus with other spec comps
C10yz00	Diabetes mellitus NOS with other specified manifestation
C10z.00	Diabetes mellitus with unspecified complication
C10z100	Diabetes mellitus, adult onset, + unspecified complication
C10zy00	Other specified diabetes mellitus with unspecified comps
C10zz00	Diabetes mellitus NOS with unspecified complication
Cyu2.00	[X]Diabetes mellitus
Cyu2000	[X]Other specified diabetes mellitus
Cyu2300	[X]Unspecified diabetes mellitus with renal complications
F171100	Autonomic neuropathy due to diabetes
F345000	Diabetic mononeuritis multiplex
F35z000	Diabetic mononeuritis NOS

F372.00	Polyneuropathy in diabetes
F372.11	Diabetic polyneuropathy
F372.12	Diabetic neuropathy
F372000	Acute painful diabetic neuropathy
F372100	Chronic painful diabetic neuropathy
F372200	Asymptomatic diabetic neuropathy
F381300	Myasthenic syndrome due to diabetic amyotrophy
F381311	Diabetic amyotrophy
F3y0.00	Diabetic mononeuropathy
F420.00	Diabetic retinopathy
F420000	Background diabetic retinopathy
F420100	Proliferative diabetic retinopathy
F420200	Preproliferative diabetic retinopathy
F420300	Advanced diabetic maculopathy
F420400	Diabetic maculopathy
F420500	Advanced diabetic retinal disease
F420600	Non proliferative diabetic retinopathy
F420700	High risk proliferative diabetic retinopathy
F420800	High risk non proliferative diabetic retinopathy
F420z00	Diabetic retinopathy NOS
F440700	Diabetic iritis
F464000	Diabetic cataract
G73y000	Diabetic peripheral angiopathy
K01x100	Nephrotic syndrome in diabetes mellitus
K08yA00	Proteinuric diabetic nephropathy
K08yA11	Clinical diabetic nephropathy
K27y700	Erectile dysfunction due to diabetes mellitus
Kyu0300	[X]Glomerular disorders in diabetes mellitus
L180500	Pre-existing diabetes mellitus, insulin-dependent
L180700	Pre-existing malnutrition-related diabetes mellitus
M037200	Cellulitis in diabetic foot
M21yC00	Insulin lipohypertrophy
M21yC11	Insulin site lipohypertrophy
M271000	Ischaemic ulcer diabetic foot
M271100	Neuropathic diabetic ulcer - foot

M271200	Mixed diabetic ulcer - foot
N030000	Diabetic cheiroarthropathy
N030011	Diabetic cheiropathy
N030100	Diabetic Charcot arthropathy
R054200	[D]Gangrene of toe in diabetic
R054300	[D]Widespread diabetic foot gangrene
TJ23.00	Adverse reaction to insulins and antidiabetic agents
TJ23000	Adverse reaction to insulins
TJ23z00	Adverse reaction to insulins and antidiabetic agents NOS
U602300	[X]Insul/oral hypoglyc drugs caus adverse eff therapeut use
U602311	[X] Adverse reaction to insulins and antidiabetic agents
U602312	[X] Adverse reaction to insulins
U60231E	[X] Adverse reaction to insulins and antidiabetic agents NOS
ZC2C800	Dietary advice for diabetes mellitus
ZL22500	Under care of diabetic liaison nurse
ZL62500	Referral to diabetes nurse
ZL62600	Referral to diabetic liaison nurse
ZLA2500	Seen by diabetic liaison nurse
ZLD7500	Discharge by diabetic liaison nurse
ZRB4.00	Diabetes clinic satisfaction questionnaire
ZRB4.11	CSQ - Diabetes clinic satisfaction questionnaire
ZRB5.00	Diabetes treatment satisfaction questionnaire
ZRB5.11	DTSQ - Diabetes treatment satisfaction questionnaire
ZRB6.00	Diabetes wellbeing questionnaire
ZRB6.11	DWBQ - Diabetes wellbeing questionnaire
ZRBa.00	Education score - diabetes
ZV65312	[V]Dietary counselling in diabetes mellitus
ZV6DA00	[V]Admitted for commencement of insulin

Annex 4. CODES TO BE USED FOR EXCLUSION CRITERIA

Annex Table 4-1. Liver injury exclusion criteria: Read codes

Read code	Read term
History of acute liver injury	
J633000	Toxic hepatitis
J635.00	Toxic liver disease
J635000	Toxic liver disease with cholestasis
J635100	Toxic liver disease with hepatic necrosis
J635200	Toxic liver disease with acute hepatitis
J635300	Toxic liver disease with chronic persistent hepatitis
J635400	Toxic liver disease with chronic lobular hepatitis
J635500	Toxic liver disease with chronic active hepatitis
J635600	Toxic liver disease with fibrosis and cirrhosis of liver
J635700	Acute hepatic failure due to drugs
J635X00	Toxic liver disease, unspecified
Jyu7600	[X]Toxic liver disease, unspecified
J60..00	Acute and subacute liver necrosis
J600.00	Acute necrosis of liver
J600000	Acute hepatic failure
J600011	Acute liver failure
J600100	Acute hepatitis - non-infective
J600z00	Acute necrosis of liver NOS
J601.00	Subacute necrosis of liver
J601000	Subacute hepatic failure
J601100	Subacute hepatitis - non-infective
J601z00	Subacute necrosis of liver NOS
J60z.00	Acute and subacute liver necrosis NOS
J622.00	Hepatic coma
J622.11	Encephalopathy - hepatic
Liver transplant	
7800	Transplantation of liver
7800000	Orthotopic transplantation of liver
7800100	Heterotopic transplantation of liver
7800111	Auxiliary liver transplant

7800112	Piggy back liver transplant
7800200	Replacement of previous liver transplant
7800400	Orthotopic transplantation of whole liver
7800500	Orthotopic transplantation of liver NEC
7800y00	Other specified transplantation of liver
7800z00	Transplantation of liver NOS
8LH..00	Liver transplant planned
SP08600	Liver transplant failure and rejection
TB00200	Liver transplant with complication, without blame
ZV42700	[V]Liver transplanted
ZV7C000	[V]Assessment for liver transplant
Chronic liver disease and alcoholism	
136T.00	Harmful alcohol use
136W.00	Alcohol misuse
7L1f.00	Compensation for liver failure
7L1fz00	Compensation for liver failure NOS
8HlW.00	Referral to liver unit
9kR..00	Chronic hepatitis annual review - enhanced services admin
C150500	Alcohol-induced pseudo-Cushing's syndrome
C310311	Glycogenosis of liver and muscle
C310400	Glycogenosis with hepatic cirrhosis
C32y511	Hepatic familial steatosis
C350012	Pigmentary cirrhosis of liver
E01..00	Alcoholic psychoses
E010.00	Alcohol withdrawal delirium
E011000	Korsakov's alcoholic psychosis
E012.00	Other alcoholic dementia
E012.11	Alcoholic dementia NOS
E012000	Chronic alcoholic brain syndrome
E23..00	Alcohol dependence syndrome
E23..11	Alcoholism
E231.00	Chronic alcoholism
E231000	Unspecified chronic alcoholism
E231100	Continuous chronic alcoholism
E231300	Chronic alcoholism in remission

E231z00	Chronic alcoholism NOS
E23z.00	Alcohol dependence syndrome NOS
Eu10211	[X]Alcohol addiction
Eu10212	[X]Chronic alcoholism
F11x000	Cerebral degeneration due to alcoholism
F11x011	Alcoholic encephalopathy
F375.00	Alcoholic polyneuropathy
F394100	Alcoholic myopathy
G555.00	Alcoholic cardiomyopathy
G81..00	Portal vein thrombosis
G820.11	Hepatic vein thrombosis
G852200	Oesophageal varices in cirrhosis of the liver
G852300	Oesophageal varices in alcoholic cirrhosis of the liver
G852300	Oesophageal varices in alcoholic cirrhosis of the liver
J153.00	Alcoholic gastritis
J6...00	Liver, biliary, pancreas + gastrointestinal diseases NEC
J61..00	Cirrhosis and chronic liver disease
J610.00	Alcoholic fatty liver
J610.00	Alcoholic fatty liver
J611.00	Acute alcoholic hepatitis
J611.00	Acute alcoholic hepatitis
J612.00	Alcoholic cirrhosis of liver
J612.00	Alcoholic cirrhosis of liver
J612000	Alcoholic fibrosis and sclerosis of liver
J612000	Alcoholic fibrosis and sclerosis of liver
J613.00	Alcoholic liver damage unspecified
J613.00	Alcoholic liver damage unspecified
J613000	Alcoholic hepatic failure
J613000	Alcoholic hepatic failure
J614.00	Chronic hepatitis
J614000	Chronic persistent hepatitis
J614100	Chronic active hepatitis
J614111	Autoimmune chronic active hepatitis
J614200	Chronic aggressive hepatitis
J614300	Recurrent hepatitis

J614400	Chronic lobular hepatitis
J614y00	Chronic hepatitis unspecified
J614z00	Chronic hepatitis NOS
J615.00	Cirrhosis - non-alcoholic
J615z11	Macronodular cirrhosis of liver
J615z12	Cryptogenic cirrhosis of liver
J615z13	Cirrhosis of liver NOS
J615z15	Hepatic fibrosis
J616.00	Biliary cirrhosis
J616000	Primary biliary cirrhosis
J616100	Secondary biliary cirrhosis
J616200	Biliary cirrhosis of children
J616z00	Biliary cirrhosis NOS
J617.00	Alcoholic hepatitis
J617.00	Alcoholic hepatitis
J617000	Chronic alcoholic hepatitis
J617000	Chronic alcoholic hepatitis
J61y.00	Other non-alcoholic chronic liver disease
J61y.00	Other non-alcoholic chronic liver disease
J61y200	Hepatosplenomegaly
J61y400	Hepatic fibrosis
J61y500	Hepatic sclerosis
J61y600	Hepatic fibrosis with hepatic sclerosis
J61yz00	Other non-alcoholic chronic liver disease NOS
J61z.00	Chronic liver disease NOS
J62..00	Liver abscess and sequelae of chronic liver disease
J620.00	Liver abscess - excluding amoebic liver abscess
J620000	Liver abscess due to portal pyaemia
J620000	Liver abscess due to portal pyaemia
J620100	Liver abscess due to cholangitis
J620200	Liver abscess via hepatic artery
J620300	Liver abscess via umbilicus
J620z00	Liver abscess NOS
J621.00	Portal pyaemia
J623.00	Portal hypertension

J624.00	Hepatorenal syndrome
J625.00	[X] Hepatic failure
J625.11	[X] Liver failure
J62y.00	Other sequelae of chronic liver disease
J62y.11	Hepatic failure NOS
J62y.12	Liver failure NOS
J62y.13	Hepatic failure
J63..00	Other liver disorders
J630.00	Chronic passive liver congestion
J634.00	Hepatic infarction
J638.00	Peliosis hepatis
J63y.00	Other specified liver disorder
J63y000	Hepatoptosis
J63z.00	Liver disorder NOS
Jyu7.00	[X]Diseases of the liver
Jyu7100	[X]Other and unspecified cirrhosis of liver
Jyu7300	[X]Other specified diseases of liver
PB6yz00	Other liver or biliary system anomalies NOS
PB6z.00	Liver or biliary system anomalies NOS
R091.00	[D]Hepatomegaly
R091z00	[D]Hepatomegaly NOS
R148.00	[D]Abnormal liver function test
SP14300	Hepatorenal syndrome as a complication of care
Infectious hepatitis	
2126700	Hepatitis C resolved
141F.00	History of viral hepatitis
14b0.00	History of one hepatitis B vaccination
14b1.00	History of two hepatitis B vaccinations
14b2.00	History of three hepatitis B vaccinations
14C5.00	H/O: liver disease
14i..00	H/O hepatitis C antiviral drug therapy
43X3.00	Hepatitis C antibody test positive
44D2.00	Liver function tests abnormal
44G2.00	Liver enzymes abnormal
4JQD.00	Hepatitis C viral ribonucleic acid PCR positive

4JQD.11	Hepatitis C PCR positive
4JQF.00	Hepatitis C antigen positive
65Q7.00	Viral hepatitis carrier
9kV..00	Hepatitis C screening positive - enhanced services admin
9kV..11	Hepatitis C screening positive
9kZ..00	Hepatitis B screening positive - enhanced services admin
9kZ..11	Hepatitis B screening positive
A053.00	Amoebic liver abscess
A17y400	Tuberculosis of liver
A392200	Actinomycosis of liver
A70..00	Viral hepatitis
A700.00	Viral hepatitis A with coma
A701.00	Viral (infectious) hepatitis A
A701.11	Infective hepatitis
A702.00	Viral hepatitis B with coma
A703.00	Viral (serum) hepatitis B
A704.00	Other specified viral hepatitis with coma
A704000	Viral hepatitis C with coma
A704z00	Other specified viral hepatitis with hepatic coma NOS
A705.00	Other specified viral hepatitis without coma
A705000	Viral hepatitis C without mention of hepatic coma
A705100	Acute delta-(super)infection of hepatitis B carrier
A705200	Acute hepatitis E
A705400	Hepatitis non A non B
A705z00	Other specified viral hepatitis without mention of coma NOS
A707.00	Chronic viral hepatitis
A707000	Chronic viral hepatitis B with delta-agent
A707100	Chronic viral hepatitis B without delta-agent
A707200	Chronic viral hepatitis C
A707X00	Chronic viral hepatitis, unspecified
A709.00	Viral hepatitis without hepatic coma
A70G.00	Acute hepatitis C
A70z.00	Unspecified viral hepatitis
A70z000	Hepatitis C
A70z100	Acute viral hepatitis NOS

A72x000	Mumps hepatitis
A785200	Cytomegaloviral hepatitis
A916100	Secondary syphilitic hepatitis
A953.00	Syphilis of liver
A98yy11	Gonococcal hepatitis
A98yy13	Gonococcal perihepatitis
AB50300	Blastomycosis liver
AC10.11	Cat liver fluke infection
AC11.11	Chinese liver fluke disease
AC13.11	Liver flukes NOS
AC13.12	Sheep liver fluke infection
AC20.00	Liver echinococcus granulosus
AC25.00	Liver echinococcus multilocularis
AC2y.00	Liver echinococcus unspecified
AC8y100	Capillaria hepatica
AD05.00	Toxoplasma hepatitis
AE23.00	Sequelae of viral hepatitis
AyuB.00	[X]Viral hepatitis
AyuB000	[X]Other specified acute viral hepatitis
AyuJ900	[X]Sequelae of viral hepatitis
J631.00	Hepatitis in viral diseases EC
J631000	Hepatitis in coxsackie virus
J631100	Hepatitis in cytomegalic inclusion virus
J631200	Hepatitis in infectious mononucleosis
J631400	Hepatitis in yellow fever
J631500	Hepatitis in other viral disease
J631600	Hepatitis + adenovirus
J631z00	Hepatitis in viral diseases EC NOS
J632.00	Hepatitis in other infectious diseases EC
J632000	Hepatitis in malaria
J632200	Hepatitis in secondary syphilis
J632300	Hepatitis in toxoplasmosis
J632z00	Hepatitis in infectious diseases EC NOS
J633.00	Hepatitis unspecified
J633z00	Hepatitis unspecified NOS

J639.00	Hepatic granulomas in berylliosis
J63A.00	Hepatic granulomas in sarcoidosis
Jyu7400	[X]Liver disorders in infectious and parasitic diseases CE
Q409.00	Congenital viral hepatitis
Q409000	Congenital hepatitis A infection
Q409100	Congenital hepatitis B infection
Q409z00	Congenital viral hepatitis NOS
ZV02600	[V]Viral hepatitis carrier
ZV02612	[V]Hepatitis Australia antigen carrier
ZV02B00	[V]Hepatitis B carrier
ZV02C00	[V]Hepatitis C carrier
Chronic disease involving the liver or causing hyperbilirubinaemia	
C350000	Haemochromatosis
C351000	Hepatolenticular degeneration (Wilson's disease)
C374200	Gilbert's syndrome
C376100	Alpha-1-antitrypsin hepatitis
C376200	Alpha-1-antitrypsin deficiency
G820.00	Budd - Chiari syndrome (hepatic vein thrombosis)
Biliary disease	
J620100	Liver abscess due to cholangitis
J651.00	Other cholecystitis
J651000	Chronic cholecystitis
J661100	Chronic cholangitis
J661200	Recurrent cholangitis
PB6y100	Congenital hepatomegaly
PB6y600	Atrophy of left lobe of liver
PB6y900	Liver hyperplasia
PB6yw00	Other congenital anomaly of liver
PB6yw11	Liver hamartoma
PB6yw12	Abnormal liver lobulation
PB6yy00	Other congenital anomaly of hepatic or bile ducts
Pyu5E00	[X]Other congenital malformations of liver
Q48yz11	Congenital hepatic fibrosis
ZV11300	[V]Personal history of alcoholism
Pancreatic disease	

14CH.00	History of chronic pancreatitis
1J0K.00	Suspected pancreatic cancer
A723.00	Mumps pancreatitis
A785100	Cytomegaloviral pancreatitis
B17..00	Malignant neoplasm of pancreas
B170.00	Malignant neoplasm of head of pancreas
B171.00	Malignant neoplasm of body of pancreas
B172.00	Malignant neoplasm of tail of pancreas
B173.00	Malignant neoplasm of pancreatic duct
B175.00	Malignant neoplasm, overlapping lesion of pancreas
B17y.00	Malignant neoplasm of other specified sites of pancreas
B17y000	Malignant neoplasm of ectopic pancreatic tissue
B17yz00	Malignant neoplasm of specified site of pancreas NOS
B17z.00	Malignant neoplasm of pancreas NOS
B716.00	Benign neoplasm of pancreas, excluding islets of Langerhans
B716000	Benign neoplasm of head of pancreas
B716100	Benign neoplasm of body of pancreas
B716200	Benign neoplasm of tail of pancreas
B716z00	Benign neoplasm of pancreas, excluding islets NOS
B717011	Endocrine tumour of pancreas
B80z000	Carcinoma in situ of pancreas
B905100	Neoplasm of uncertain behaviour of pancreas
BB5B.00	[M]Pancreatic adenomas and carcinomas
BB5Bz00	[M]Pancreatic adenoma or carcinoma NOS
J671.00	Chronic pancreatitis
J671000	Alcohol-induced chronic pancreatitis
J671100	Gallstone chronic pancreatitis
J672.00	Cyst and pseudocyst of pancreas
J672000	Pancreatic cyst
J672100	Pseudocyst of pancreas
J672z00	Pancreatic cyst and pseudocyst NOS
J67y000	Atrophy of pancreas
J67y200	Fibrosis of pancreas
J67y300	Aseptic necrosis of pancreas
J67y400	Fat necrosis of pancreas

J67y500	Pancreatic insufficiency
J67y600	Exocrine pancreatic insufficiency
J67yz00	Other diseases of pancreas NOS
J67z.00	Diseases of pancreas NOS
J694.00	Pancreatic steatorrhoea
Jyu8400	[X]Other chronic pancreatitis
Hepatobiliary and pancreatic neoplasms	
B15..00	Malignant neoplasm of liver and intrahepatic bile ducts
B150.00	Primary malignant neoplasm of liver
B150000	Primary carcinoma of liver
B150100	Hepatoblastoma of liver
B150200	Primary angiosarcoma of liver
B150300	Hepatocellular carcinoma
B150z00	Primary malignant neoplasm of liver NOS
B151.00	Malignant neoplasm of intrahepatic bile ducts
B151200	Malignant neoplasm of intrahepatic biliary passages
B151400	Malignant neoplasm of intrahepatic gall duct
B151z00	Malignant neoplasm of intrahepatic bile ducts NOS
B152.00	Malignant neoplasm of liver unspecified
B153.00	Secondary malignant neoplasm of liver
B15z.00	Malignant neoplasm of liver and intrahepatic bile ducts NOS
B16..00	Malignant neoplasm gallbladder and extrahepatic bile ducts
B161.00	Malignant neoplasm of extrahepatic bile ducts
B161100	Malignant neoplasm of hepatic duct
B161z00	Malignant neoplasm of extrahepatic bile ducts NOS
B16y.00	Malignant neoplasm other gallbladder/extrahepatic bile duct
B16z.00	Malignant neoplasm gallbladder/extrahepatic bile ducts NOS
B577.00	Secondary malignant neoplasm of liver
B577.11	Liver metastases
B713000	Benign neoplasm of hepatic flexure of colon
B715.00	Benign neoplasm of liver and biliary ducts
B715000	Benign neoplasm of liver
B715100	Benign neoplasm of intrahepatic bile ducts
B715400	Benign neoplasm of hepatic duct
B803000	Carcinoma in situ of hepatic flexure of colon

B808.00	Carcinoma in situ of liver and biliary system
B808000	Carcinoma in situ of liver
B808100	Carcinoma in situ of intrahepatic bile ducts
B808200	Carcinoma in situ of hepatic duct
B808z00	Carcinoma in situ of liver or biliary system NOS
B903.00	Neoplasm of uncertain behaviour of liver and biliary passage
B903000	Neoplasm of uncertain behaviour of liver
B903200	Neoplasm of uncertain behaviour of hepatic duct
BB5D.00	[M]Hepatobiliary tract adenomas and carcinomas
BB5D400	[M]Liver cell adenoma
BB5D411	[M]Hepatocellular adenoma
BB5D412	[M]Hepatoma, benign
BB5D500	[M]Hepatocellular carcinoma NOS
BB5D511	[M]Hepatoma NOS
BB5D512	[M]Hepatoma, malignant
BB5D513	[M]Liver cell carcinoma
BB5D700	[M]Combined hepatocellular carcinoma and cholangiocarcinoma
BB5D800	[M]Hepatocellular carcinoma, fibrolamellar
BB5Dz00	[M]Hepatobiliary adenoma or carcinoma NOS
BBL8.00	[M]Hepatoblastoma
BBL8.11	[M]Embryonal hepatoma
Byu1100	[X]Other specified carcinomas of liver
C310200	Hepatorenal glycogenosis
Heart failure	
1O1..00	Heart failure confirmed
388D.00	New York Heart Assoc classification heart failure symptoms
661M500	Heart failure self-management plan agreed
662f.00	New York Heart Association classification - class I
662g.00	New York Heart Association classification - class II
662h.00	New York Heart Association classification - class III
662i.00	New York Heart Association classification - class IV
662p.00	Heart failure 6 month review
662T.00	Congestive heart failure monitoring
662W.00	Heart failure annual review
8H2S.00	Admit heart failure emergency

8HBE.00	Heart failure follow-up
9N0k.00	Seen in heart failure clinic
9Or..00	Heart failure monitoring administration
9Or3.00	Heart failure monitoring first letter
G58..00	Heart failure
G580.00	Congestive heart failure
G580.12	Right heart failure
G580000	Acute congestive heart failure
G580100	Chronic congestive heart failure
G580400	Congestive heart failure due to valvular disease
G582.00	Acute heart failure
G583.00	Heart failure with normal ejection fraction
G583.11	HFNEF - heart failure with normal ejection fraction
G58z.00	Heart failure NOS
ZRad.00	New York Heart Assoc classification heart failure symptoms

H/O = history of; NEC = not elsewhere classified; NOS = not otherwise specified.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 4-2. Liver injury exclusion criteria: ICD-10 codes

Read code	**Description**
History of acute liver injury	
K71	Toxic liver disease
K72	Hepatic failure, not elsewhere classified (includes acute and subacute hepatic failure, chronic hepatic failure, and hepatic failure, unspecified)
Liver transplant	
Z94.4	Liver transplant status
T86.4	Liver transplant failure and rejection
Chronic liver disease and alcoholism	
K70	Alcoholic liver disease
K73	Chronic hepatitis, not elsewhere classified
K74	Fibrosis and cirrhosis of liver
K75	Other inflammatory liver disease
K76	Other disease of liver
K77.x	Liver disorders in disease classified elsewhere

F10.1-F10.9	Alcohol dependence syndrome/Mental and behavioural disorders due to use of alcohol: dependence syndrome
I85.01	Oesophageal varices with bleeding
I85.9	Oesophageal varices without mention of bleeding
I98.2 I98.3	Oesophageal varices in diseases classified elsewhere/Secondary oesophageal varices with bleeding
K29.2	Alcoholic gastritis
Z65.8	Personal history of alcoholism/Other specified problems related to psychosocial circumstances
K85.2	Alcohol-induced acute pancreatitis
E24.4	Alcohol-induced pseudo-Cushing's syndrome
G31.2	Degeneration of nervous system due to alcohol
G62.1, G72.1	Alcoholic polyneuropathy; Alcoholic myopathy
I42.6	Alcoholic cardiomyopathy
Z50.2	Alcohol rehabilitation
Z71.4	Alcohol abuse counselling and surveillance
Infectious hepatitis	
B15—B19, B26.8, A51.4, Z22.5, B25.1, (B00.8+K77.0), B58.1, B65.x, A52.7	Acute infectious hepatitis; Mumps hepatitis; Secondary syphilitic hepatitis; Other symptomatic late syphilis; Schistosomiasis; Toxoplasma hepatitis; Carrier or suspected carrier of viral hepatitis; Personal history of hepatitis/Other secondary hepatitis; Cytomegaloviral hepatitis; Herpes viral hepatitis
Chronic disease involving the liver or causing hyperbilirubinaemia	
E83.1	Haemochromatosis
E83.0	Wilson's disease
E88.0	Deficit of alpha-1-antitrypsin
I82.0	Budd-Chiari syndrome
E80.4, E80.5, E80.6, E80.7	Disorders of bilirubin excretion (Gilbert syndrome)
Biliary disease	
K80	Cholelithiasis
K81	Cholecystitis
K82	Other diseases of gallbladder
K83	Other diseases of biliary tract
K87.0	Disorders of gallbladder and biliary tract in diseases classified elsewhere
Pancreatic diseases	
K85	Acute pancreatitis

K86	Other diseases of pancreas
K87.1	Disorders of pancreas in diseases classified elsewhere
Hepatobiliary and pancreatic neoplasms	
C22	Malignant neoplasm of liver and intrahepatic bile ducts
C23	Malignant neoplasm of gallbladder
C24	Malignant neoplasm of other and unspecified parts of biliary tract
Heart failure	
I50.x, I13.0, I13.2, I11.0, I09.81,	Heart failure, hypertensive heart, and chronic kidney disease with heart failure

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: website: apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 4-3. Kidney injury exclusion criteria: Read codes

Read code	Description
Acute kidney disease	
G500400	Acute pericarditis - uraemic
K00..00	Acute glomerulonephritis
K00..11	Acute nephritis
K000.00	Acute proliferative glomerulonephritis
K001.00	Acute nephritis with lesions of necrotising glomerulitis
K00y000	Acute glomerulonephritis in diseases EC
K00y100	Acute exudative nephritis
K00y200	Acute focal nephritis
K00y300	Acute diffuse nephritis
K00z.00	Acute glomerulonephritis NOS
K04..00	Acute renal failure
K040.00	Acute renal tubular necrosis
K041.00	Acute renal cortical necrosis
K042.00	Acute renal medullary necrosis
K043.00	Acute drug-induced renal failure
K044.00	Acute renal failure due to urinary obstruction
K04y.00	Other acute renal failure
K04z.00	Acute renal failure NOS
Kyu2000	[X]Other acute renal failure
SK05.00	Renal failure following crush syndrome
SK05.11	Renal failure after crushing
SK08.00	Acute renal failure due to rhabdomyolysis
SP15400	Renal failure as a complication of care

SP15411	Kidney failure as a complication of care
SP15412	Post operative renal failure
Chronic kidney disease	
1Z12.00	Chronic kidney disease stage 3
1Z13.00	Chronic kidney disease stage 4
1Z14.00	Chronic kidney disease stage 5
1Z15.00	Chronic kidney disease stage 3A
1Z16.00	Chronic kidney disease stage 3B
1Z1B.00	Chronic kidney disease stage 3 with proteinuria
1Z1B.11	CKD stage 3 with proteinuria
1Z1C.00	Chronic kidney disease stage 3 without proteinuria
1Z1C.11	CKD stage 3 without proteinuria
1Z1D.00	Chronic kidney disease stage 3A with proteinuria
1Z1D.11	CKD stage 3A with proteinuria
1Z1E.00	Chronic kidney disease stage 3A without proteinuria
1Z1E.11	CKD stage 3A without proteinuria
1Z1F.00	Chronic kidney disease stage 3B with proteinuria
1Z1F.11	CKD stage 3B with proteinuria
1Z1G.00	Chronic kidney disease stage 3B without proteinuria
1Z1G.11	CKD stage 3B without proteinuria
1Z1H.00	Chronic kidney disease stage 4 with proteinuria
1Z1H.11	CKD stage 4 with proteinuria
1Z1J.00	Chronic kidney disease stage 4 without proteinuria
1Z1J.11	CKD stage 4 without proteinuria
1Z1K.00	Chronic kidney disease stage 5 with proteinuria
1Z1K.11	CKD stage 5 with proteinuria
1Z1L.00	Chronic kidney disease stage 5 without proteinuria
1Z1L.11	CKD stage 5 without proteinuria
7L1A000	Renal dialysis
C341.00	Gouty nephropathy
C341z00	Gouty nephropathy NOS
C354711	Renal calcinosis
C373600	Nephropathic amyloidosis
D310100	Henoch-Schonlein nephritis
G222.00	Hypertensive renal disease with renal failure
K0...00	Nephritis, nephrosis and nephrotic syndrome
K00..12	Bright's disease

K00y.00	Other acute glomerulonephritis
K00yz00	Other acute glomerulonephritis NOS
K01..00	Nephrotic syndrome
K010.00	Nephrotic syndrome with proliferative glomerulonephritis
K011.00	Nephrotic syndrome with membranous glomerulonephritis
K012.00	Nephrotic syndrome+membranoproliferative glomerulonephritis
K013.12	Steroid sensitive nephrotic syndrome
K015.00	Nephrotic syndrome, focal and segmental glomerular lesions
K016.00	Nephrotic syndrome, diffuse membranous glomerulonephritis
K017.00	Nephrotic syn difus mesangial prolifertiv glomerulonephritis
K018.00	Nephrotic syn,difus endocapilary proliftv glomerulonephritis
K019.00	Nephrotic syn,diffuse mesangiocapillary glomerulonephritis
K01A.00	Nephrotic syndrome, dense deposit disease
K01B.00	Nephrotic syndrome, diffuse crescentic glomerulonephritis
K01w.00	Congenital nephrotic syndrome
K01w000	Finnish nephrosis syndrome
K01x000	Nephrotic syndrome in amyloidosis
K01x100	Nephrotic syndrome in diabetes mellitus
K01x111	Kimmelstiel - Wilson disease
K01x200	Nephrotic syndrome in malaria
K01x300	Nephrotic syndrome in polyarteritis nodosa
K01x400	Nephrotic syndrome in systemic lupus erythematosus
K01x411	Lupus nephritis
K01y.00	Nephrotic syndrome with other pathological kidney lesions
K01z.00	Nephrotic syndrome NOS
K02..00	Chronic glomerulonephritis
K02..11	Nephritis - chronic
K02..12	Nephropathy - chronic
K020.00	Chronic proliferative glomerulonephritis
K021.00	Chronic membranous glomerulonephritis
K022.00	Chronic membranoproliferative glomerulonephritis
K023.00	Chronic rapidly progressive glomerulonephritis
K02y.00	Other chronic glomerulonephritis
K02y000	Chronic glomerulonephritis + diseases EC
K02y200	Chronic focal glomerulonephritis
K02y300	Chronic diffuse glomerulonephritis
K02yz00	Other chronic glomerulonephritis NOS

K02z.00	Chronic glomerulonephritis NOS
K030.00	Proliferative nephritis unspecified
K031.00	Membranous nephritis unspecified
K032.00	Membranoproliferative nephritis unspecified
K032000	Focal membranoproliferative glomerulonephritis
K032300	Anaphylactoid glomerulonephritis
K032400	Familial glomerulonephritis in Alport's syndrome
K032500	Other familial glomerulonephritis
K032600	Berger's IgA or IgG nephropathy
K032y00	Nephritis unsp+OS membranoprolif glomerulonephritis lesion
K032y11	Hypocomplementaemic persistent glomerulonephritis NEC
K032y13	Mesangioproliferative glomerulonephritis NEC
K032y14	Mesangiocapillary glomerulonephritis NEC
K032y15	Mixed membranous and proliferative glomerulonephritis NEC
K032z00	Nephritis unsp+membranoprolif glomerulonephritis lesion NOS
K033.00	Rapidly progressive nephritis unspecified
K035.00	Renal medullary necrosis unspecified
K03T.00	Tubulo-interstit nephritis, not specif as acute or chron
K03U.00	Unspecif nephr synd, diff concentric glomerulonephritis
K03V.00	Unspecified nephritic syndrome, dense deposit disease
K03W.00	Unsp nephrit synd, diff endocap prolif glomerulonephritis
K03X.00	Unsp nephrit synd, diff mesang prolif glomerulonephritis
K03y.00	Other nephritis and nephrosis unspecified
K03y000	Other nephritis and nephrosis in diseases EC
K03z.00	Unspecified glomerulonephritis NOS
K042.11	Necrotising renal papillitis
K05..12	End-stage renal failure
K050.00	End-stage renal failure
K080.00	Renal osteodystrophy
K080000	Phosphate-losing tubular disorders
K080100	Renal dwarfism
K080200	Renal infantilism
K080300	Renal rickets
K080z00	Renal osteodystrophy NOS
K081.00	Nephrogenic diabetes insipidus
K08y.00	Other impaired renal function disorder
K08y000	Hypokalaemic nephropathy

K08y300	Renal function impairment with growth failure
K08yz00	Other impaired renal function disorder NOS
K091.00	Bilateral small kidneys
K09z.00	Small kidneys unspecified
K0A..00	Glomerular disease
K0A2100	Recur+persist haematuria, focal+segmental glomerular lesions
K0A2200	Recur+persist haematuria difus membranous glomerulonephritis
K0A2300	Recur+persist haemuria df mesangial prolif glomerulnephritis
K0A2500	Recur+persist hmuria df mesangiocapilary glomerulonephritis
K0A2600	Recurrent and persistent haematuria, dense deposit disease
K0A2700	Recur+persist haematuria difus crescentic glomerulonephritis
K0A2800	IgA nephropathy
K0A3.00	Chronic nephritic syndrome
K0A3000	Chronic nephritic syndrome, minor glomerular abnormality
K0A3100	Chronic nephritic syndrm focal+segmental glomerular lesions
K0A3200	Chron nephritic syndrom difuse membranous glomerulonephritis
K0A3300	Chron neph syn difus mesangial prolifrtiv glomerulonephritis
K0A3500	Chronic neph syn difus mesangiocapillary glomerulonephritis
K0A3600	Chronic nephritic syndrome, dense deposit disease
K0A3700	Chronic nephritic syn diffuse crescentic glomerulonephritis
K0A5100	Hereditary nephropathy NEC,focal+segmnt glomerular lesion
K0A5200	Hereditry nephropathy NEC,difus membran glomerulnephritis
K0A5300	Heredtry nephrpthy NEC difus mesangial prolif glomnephrit
K0A5600	Hereditary nephropathy, NEC, dense deposit disease
K0A5X00	Hereditary nephropathy, unspecif morphological changes
K0A7.00	Glom disordr in blood diseas+disordr invlvg imun mechansm
K0B..00	Renal tubulo-interstitial disorders in diseases EC
K0B1.00	Renal tubulo-interstitial disorder/ neoplastic diseases
K0B2.00	Ren tub-interst disordr/blood dis+disordr inv immune mech
K0B4.00	Ren tub-interstitl disordr/systemc connectv tiss disorder
K0B4000	Renal tubulo-interstitial disorder in SLE
K0B5.00	Renal tubulo-interstitial disordrs in transplant rejectn
K0B6.00	Balkan nephropathy
K0C..00	Drug/heavy-metal-induced tubulo-interstitial and tub conditn
K0C0.00	Analgesic nephropathy
K0C1.00	Nephropathy induced by other drugs meds and biologl substncs
K0C2.00	Nephropathy induced by unspec drug medicament or biol subs

K0C4.00	Toxic nephropathy, not elsewhere classified
K0D..00	End-stage renal disease
K0E..00	Acute-on-chronic renal failure
K0y..00	Other specified nephritis, nephrosis or nephrotic syndrome
K0y0.00	Late syphilis of kidney
K0z..00	Nephritis, nephrosis and nephrotic syndrome NOS
Kyu0900	[X]Unsp nephrit synd, diff mesang prolif glomerulonephritis
Kyu1.00	[X]Renal tubulo-interstitial diseases
Kyu1000	[X]Other chronic tubulo-interstitial nephritis
Kyu1400	[X]Nephropathy induced by other drugs+biological substances
Kyu1C00	[X]Renal tubulo-interstitial disorders/transplant rejection
Kyu4000	[X]Other disorders resulting/impaired renal tubular function
PD03000	Bilateral renal hypoplasia
PD03011	Potter's syndrome
PD04.00	Dysplasia of kidney
PD04000	Bilateral renal dysplasia
PD04011	Bilateral renal dysgenesis
PD1..00	Congenital cystic kidney disease
PD1..11	Congenital cystic renal disease
PD1..12	Fibrocystic kidney
PD1..13	Polycystic kidney
PD1..14	Sponge kidney
PD10.00	Congenital renal cyst, single
PD11.00	Polycystic kidney disease
PD11100	Polycystic kidneys, adult type
PD11z00	Polycystic kidney disease NOS
PD11z11	Cystic kidney disease NEC
PD12.00	Medullary cystic disease
PD12000	Medullary cystic disease, juvenile type
PD12011	Nephronophthisis
PD12100	Medullary cystic disease, adult type
PD12111	Medullary sponge kidney
PD12y00	Medullary cystic disease OS
PD12z00	Medullary cystic disease NOS
PD13.00	Multicystic renal dysplasia
PD13.11	Multicystic kidney
PD1y.00	Other specified congenital cystic kidney disease

PD1y000	Fibrocystic kidney disease
PD1y011	Fibrocystic renal degeneration
PD1yz00	Other congenital cystic kidney disease NOS
PD1z.00	Congenital cystic kidney disease NOS
Q48y000	Congenital renal failure
TB00100	Kidney transplant with complication, without blame
TB11.00	Kidney dialysis with complication, without blame
Other	
14S2.00	H/O: kidney recipient
14V2.00	H/O: renal dialysis
14V2.11	H/O: kidney dialysis
7B00.00	Transplantation of kidney
7B00000	Autotransplant of kidney
7B00100	Transplantation of kidney from live donor
7B00111	Allotransplantation of kidney from live donor
7B00200	Transplantation of kidney from cadaver
7B00211	Allotransplantation of kidney from cadaver
7B00300	Allotransplantation of kidney from cadaver, heart-beating
7B00400	Allotransplantation kidney from cadaver, heart non-beating
7B00500	Allotransplantation of kidney from cadaver NEC
7B00y00	Other specified transplantation of kidney
7B00z00	Transplantation of kidney NOS
7B01511	Excision of rejected transplanted kidney
7B06300	Exploration of renal transplant
7B0F.00	Interventions associated with transplantation of kidney
7B0F100	Pre-transplantation of kidney work-up, recipient
7B0F300	Post-transplantation of kidney examination, recipient
7B0Fy00	OS interventions associated with transplantation of kidney
7B0Fz00	Interventions associated with transplantation of kidney NOS
7L1A011	Thomas intravascular shunt for dialysis
7L1A100	Peritoneal dialysis
7L1A200	Haemodialysis NEC
7L1A300	Haemofiltration
7L1A400	Automated peritoneal dialysis
7L1A500	Continuous ambulatory peritoneal dialysis
7L1A600	Peritoneal dialysis NEC
7L1A700	Haemoperfusion

7L1Az00	Compensation for renal failure NOS
7L1A.00	Compensation for renal failure
7L1A.11	Dialysis for renal failure
7L1Ay00	Other specified compensation for renal failure
7L1B.00	Placement ambulatory apparatus compensation renal failure
7L1B.11	Placement ambulatory dialysis apparatus - compens renal fail
7L1By00	Placement ambulatory apparatus- compensate renal failure OS
7L1C.00	Placement other apparatus for compensation for renal failure
7L1Cz00	Placement other apparatus- compensate for renal failure NOS
8L50.00	Renal transplant planned
D215.00	Anaemia secondary to renal failure
D215000	Anaemia secondary to chronic renal failure
SP08011	Det.ren.func.after ren.transpl
SP08300	Kidney transplant failure and rejection
TA02000	Accid cut,puncture,perf,h'ge - kidney dialysis
TA22000	Failure of sterile precautions during kidney dialysis
TB00111	Renal transplant with complication, without blame
TB11.11	Renal dialysis with complication, without blame
ZV42000	[V]Kidney transplanted
ZV45100	[V]Renal dialysis status
ZV56.00	[V] Aftercare involving intermittent dialysis
ZV56000	[V] Aftercare involving extracorporeal dialysis
ZV56011	[V] Aftercare involving renal dialysis NOS
ZV56100	[V] Preparatory care for dialysis
ZV56y00	[V] Other specified aftercare involving intermittent dialysi
ZV56y11	[V] Aftercare involving peritoneal dialysis
ZV56z00	[V] Unspecified aftercare involving intermittent dialysis
ZVu3G00	[X] Other dialysis

NEC = not elsewhere classified; NOS = not otherwise specified; O/E = on examination; OS = otherwise specified.

Sources: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 4-4. Kidney injury exclusion criteria: ICD-10 codes

ICD-10 code	**ICD-10 term**
N00	Acute nephritic syndrome
N10	Acute tubulo-interstitial nephritis
N12	Tubulo-interstitial nephritis, not specified as acute or chronic
N14	Drug- and heavy-metal-induced tubule-interstitial and tubular conditions

N17	Acute kidney injury
N19	Unspecified kidney failure
T86.1	Kidney transplant failure and rejection
Y84.1	Kidney dialysis
Z49	Care involving dialysis
Z94.0	Kidney transplant status
Z99.2	Dependence on renal dialysis

Source:International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: website: apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 4-5. Prior chronic kidney disease exclusion criteria: Read codes

Read code	**Description**
1Z12.00	Chronic kidney disease stage 3
1Z13.00	Chronic kidney disease stage 4
1Z14.00	Chronic kidney disease stage 5
1Z15.00	Chronic kidney disease stage 3A
1Z16.00	Chronic kidney disease stage 3B
1Z1B.00	Chronic kidney disease stage 3 with proteinuria
1Z1C.00	Chronic kidney disease stage 3 without proteinuria
1Z1D.00	Chronic kidney disease stage 3A with proteinuria
1Z1E.00	Chronic kidney disease stage 3A without proteinuria
1Z1F.00	Chronic kidney disease stage 3B with proteinuria
1Z1G.00	Chronic kidney disease stage 3B without proteinuria
1Z1H.00	Chronic kidney disease stage 4 with proteinuria
1Z1J.00	Chronic kidney disease stage 4 without proteinuria
1Z1K.00	Chronic kidney disease stage 5 with proteinuria
1Z1L.00	Chronic kidney disease stage 5 without proteinuria
66i..00	Chronic kidney disease monitoring
68D1.11	Chronic kidney disease screening
6AA..00	Chronic kidney disease annual review
7B00.00	Transplantation of kidney
7B00000	Autotransplant of kidney
7B00100	Transplantation of kidney from live donor
7B00111	Allotransplantation of kidney from live donor
7B00200	Transplantation of kidney from cadaver
7B00211	Allotransplantation of kidney from cadaver

7B00300	Allotransplantation of kidney from cadaver, heart-beating
7B00400	Allotransplantation kidney from cadaver, heart non-beating
7B00y00	Other specified transplantation of kidney
7B00z00	Transplantation of kidney NOS
7B01.11	Total excision of kidney
7B01300	Heminephrectomy for horseshoe kidney
7B01311	Excision of half of horseshoe kidney
7B01511	Excision of rejected transplanted kidney
7B02.11	Partial excision of kidney
7B02000	Heminephrectomy for duplex kidney
7B02100	Division of isthmus of horseshoe kidney
7B03300	Rovsing's operation for polycystic kidney
7B0F100	Pre-transplantation of kidney work-up, recipient
B4A..00	Malig neop of kidney and other unspecified urinary organs
B4A0.00	Malignant neoplasm of kidney parenchyma
B4Az.00	Malignant neoplasm of kidney or urinary organs NOS
B580.00	Secondary malignant neoplasm of kidney
K01y.00	Nephrotic syndrome with other pathological kidney lesions
K05..13	Chronic kidney disease
K053.00	Chronic kidney disease stage 3
K054.00	Chronic kidney disease stage 4
K055.00	Chronic kidney disease stage 5
PD1..13	Polycystic kidney
PD1..14	Sponge kidney
PD11.00	Polycystic kidney disease
PD11000	Polycystic kidneys, infantile type
PD11011	Autosomal recessive polycystic kidney disease
PD11100	Polycystic kidneys, adult type
PD11111	Autosomal dominant polycystic kidney disease
PD11z00	Polycystic kidney disease NOS
PD11z11	Cystic kidney disease NEC
PD12111	Medullary sponge kidney
PD13.11	Multicystic kidney
PD1y.00	Other specified congenital cystic kidney disease
PD1y000	Fibrocystic kidney disease

PD1yz00	Other congenital cystic kidney disease NOS
PD1z.00	Congenital cystic kidney disease NOS
SP08300	Kidney transplant failure and rejection
SP08P00	Stenosis of vein of transplanted kidney
TB00100	Kidney transplant with complication, without blame
TB11.00	Kidney dialysis with complication, without blame
U612200	[X]Failure sterile precautions dur kidney dialys/other perf
Z4B4.00	Polycystic kidney disease counselling
ZV10512	[V]Personal history of malignant neoplasm of kidney
ZV10513	[V]Personal history of malignant neoplasm of kidney
ZV42000	[V]Kidney transplanted

H/O = history of; NEC = not elsewhere classified; NOS = not otherwise specified.
Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 4-6. Prior chronic kidney disease exclusion criteria: ICD-10 codes

Read code	**Description**
N18	Chronic kidney disease
Q61	Cystic kidney disease
Z94.0	Kidney transplant status
C64	Malignant neoplasm of kidney, except renal pelvis
C65	Malignant neoplasm of renal pelvis

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: website: apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 4-7. Prior chronic pyelonephritis exclusion criteria: Read codes

Read code	**Description**
A160100	Tuberculous pyelitis
K100.00	Chronic pyelonephritis
K100000	Chronic pyelonephritis without medullary necrosis
K100100	Chronic pyelonephritis with medullary necrosis
K100200	Chronic pyelitis
K100300	Chronic pyonephrosis
K100400	Nonobstructive reflux-associated chronic pyelonephritis
K100500	Chronic obstructive pyelonephritis
K100z00	Chronic pyelonephritis NOS
K104.00	Xanthogranulomatous pyelonephritis

NOS = not otherwise specified.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 4-8. Prior chronic pyelonephritis exclusion criteria: ICD-10 codes

Read code	Description
N11	Chronic tubulo-interstitial nephritis (including N11.0Nonobstructive reflux-associated chronic pyelonephritis, N11.1Chronic obstructive pyelonephritis, N11.8Other chronic tubulo-interstitial nephritis, and N11.9Chronic tubulo-interstitial nephritis, unspecified)

Source:International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: website: apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 4-9. Type 1 Diabetes Mellitus exclusion criteria: Read codes

Read code	Read term
66AJ100	Brittle diabetes
66An.00	Diabetes type 1 review
7L10000	Continuous subcutaneous infusion of insulin
7L10011	Subcutaneous infusion with insulin pump
C100000	Diabetes mellitus, juvenile type, no mention of complication
C100011	Insulin dependent diabetes mellitus
C101000	Diabetes mellitus, juvenile type, with ketoacidosis
C102000	Diabetes mellitus, juvenile type, with hyperosmolar coma
C103000	Diabetes mellitus, juvenile type, with ketoacidotic coma
C104000	Diabetes mellitus, juvenile type, with renal manifestation
C105000	Diabetes mellitus, juvenile type, + ophthalmic manifestation
C106000	Diabetes mellitus, juvenile, + neurological manifestation
C107000	Diabetes mellitus, juvenile +peripheral circulatory disorder
C107300	IDDM with peripheral circulatory disorder
C108.00	Insulin dependent diabetes mellitus
C108.11	IDDM-Insulin dependent diabetes mellitus
C108.12	Type 1 diabetes mellitus
C108.13	Type I diabetes mellitus
C108000	Insulin-dependent diabetes mellitus with renal complications
C108011	Type I diabetes mellitus with renal complications
C108012	Type 1 diabetes mellitus with renal complications
C108100	Insulin-dependent diabetes mellitus with ophthalmic comps
C108112	Type 1 diabetes mellitus with ophthalmic complications
C108200	Insulin-dependent diabetes mellitus with neurological comps

Read code	Read term
C108211	Type I diabetes mellitus with neurological complications
C108212	Type 1 diabetes mellitus with neurological complications
C108300	Insulin dependent diabetes mellitus with multiple complicatn
C108311	Type I diabetes mellitus with multiple complications
C108400	Unstable insulin dependent diabetes mellitus
C108411	Unstable type I diabetes mellitus
C108412	Unstable type 1 diabetes mellitus
C108500	Insulin dependent diabetes mellitus with ulcer
C108511	Type I diabetes mellitus with ulcer
C108512	Type 1 diabetes mellitus with ulcer
C108600	Insulin dependent diabetes mellitus with gangrene
C108700	Insulin dependent diabetes mellitus with retinopathy
C108711	Type I diabetes mellitus with retinopathy
C108712	Type 1 diabetes mellitus with retinopathy
C108800	Insulin dependent diabetes mellitus - poor control
C108811	Type I diabetes mellitus - poor control
C108812	Type 1 diabetes mellitus - poor control
C108900	Insulin dependent diabetes maturity onset
C108911	Type I diabetes mellitus maturity onset
C108912	Type 1 diabetes mellitus maturity onset
C108A00	Insulin-dependent diabetes without complication
C108A11	Type I diabetes mellitus without complication
C108B00	Insulin dependent diabetes mellitus with mononeuropathy
C108B11	Type I diabetes mellitus with mononeuropathy
C108C00	Insulin dependent diabetes mellitus with polyneuropathy
C108D00	Insulin dependent diabetes mellitus with nephropathy
C108D11	Type I diabetes mellitus with nephropathy
C108E00	Insulin dependent diabetes mellitus with hypoglycaemic coma
C108E11	Type I diabetes mellitus with hypoglycaemic coma
C108E12	Type 1 diabetes mellitus with hypoglycaemic coma
C108F00	Insulin dependent diabetes mellitus with diabetic cataract
C108F11	Type I diabetes mellitus with diabetic cataract
C108G00	Insulin dependent diab mell with peripheral angiopathy
C108H00	Insulin dependent diabetes mellitus with arthropathy

Read code	Read term
C108H11	Type I diabetes mellitus with arthropathy
C108J00	Insulin dependent diab mell with neuropathic arthropathy
C108J11	Type I diabetes mellitus with neuropathic arthropathy
C108J12	Type 1 diabetes mellitus with neuropathic arthropathy
C10E.00	Type 1 diabetes mellitus
C10E.11	Type I diabetes mellitus
C10E.12	Insulin dependent diabetes mellitus
C10E000	Type 1 diabetes mellitus with renal complications
C10E011	Type I diabetes mellitus with renal complications
C10E012	Insulin-dependent diabetes mellitus with renal complications
C10E100	Type 1 diabetes mellitus with ophthalmic complications
C10E111	Type I diabetes mellitus with ophthalmic complications
C10E112	Insulin-dependent diabetes mellitus with ophthalmic comps
C10E200	Type 1 diabetes mellitus with neurological complications
C10E212	Insulin-dependent diabetes mellitus with neurological comps
C10E300	Type 1 diabetes mellitus with multiple complications
C10E311	Type I diabetes mellitus with multiple complications
C10E312	Insulin dependent diabetes mellitus with multiple complicat
C10E400	Unstable type 1 diabetes mellitus
C10E411	Unstable type I diabetes mellitus
C10E412	Unstable insulin dependent diabetes mellitus
C10E500	Type 1 diabetes mellitus with ulcer
C10E511	Type I diabetes mellitus with ulcer
C10E512	Insulin dependent diabetes mellitus with ulcer
C10E600	Type 1 diabetes mellitus with gangrene
C10E611	Type I diabetes mellitus with gangrene
C10E612	Insulin dependent diabetes mellitus with gangrene
C10E700	Type 1 diabetes mellitus with retinopathy
C10E711	Type I diabetes mellitus with retinopathy
C10E712	Insulin dependent diabetes mellitus with retinopathy
C10E800	Type 1 diabetes mellitus - poor control
C10E811	Type I diabetes mellitus - poor control
C10E812	Insulin dependent diabetes mellitus - poor control
C10E900	Type 1 diabetes mellitus maturity onset

Read code	Read term
C10E911	Type I diabetes mellitus maturity onset
C10E912	Insulin dependent diabetes maturity onset
C10EA00	Type 1 diabetes mellitus without complication
C10EA11	Type I diabetes mellitus without complication
C10EA12	Insulin-dependent diabetes without complication
C10EB00	Type 1 diabetes mellitus with mononeuropathy
C10EC00	Type 1 diabetes mellitus with polyneuropathy
C10EC11	Type I diabetes mellitus with polyneuropathy
C10EC12	Insulin dependent diabetes mellitus with polyneuropathy
C10ED00	Type 1 diabetes mellitus with nephropathy
C10ED12	Insulin dependent diabetes mellitus with nephropathy
C10EE00	Type 1 diabetes mellitus with hypoglycaemic coma
C10EE12	Insulin dependent diabetes mellitus with hypoglycaemic coma
C10EF00	Type 1 diabetes mellitus with diabetic cataract
C10EF12	Insulin dependent diabetes mellitus with diabetic cataract
C10EG00	Type 1 diabetes mellitus with peripheral angiopathy
C10EH00	Type 1 diabetes mellitus with arthropathy
C10EJ00	Type 1 diabetes mellitus with neuropathic arthropathy
C10EK00	Type 1 diabetes mellitus with persistent proteinuria
C10EL00	Type 1 diabetes mellitus with persistent microalbuminuria
C10EL11	Type I diabetes mellitus with persistent microalbuminuria
C10EM00	Type 1 diabetes mellitus with ketoacidosis
C10EM11	Type I diabetes mellitus with ketoacidosis
C10EN00	Type 1 diabetes mellitus with ketoacidotic coma
C10EN11	Type I diabetes mellitus with ketoacidotic coma
C10EP00	Type 1 diabetes mellitus with exudative maculopathy
C10EP11	Type I diabetes mellitus with exudative maculopathy
C10EQ00	Type 1 diabetes mellitus with gastroparesis
C10EQ11	Type I diabetes mellitus with gastroparesis
C10J.00	Insulin autoimmune syndrome
C10P000	Type I diabetes mellitus in remission
C10P011	Type 1 diabetes mellitus in remission
C10z000	Diabetes mellitus, juvenile type, + unspecified complication
ZC2C900	Dietary advice for type I diabetes

Read code	Read term
ZC2C911	Diet advice for insulin-dependent diabetes
ZRbH.00	Perceived control of insulin-dependent diabetes

Annex 5. CODES TO DEFINE STUDY OUTCOMES

Annex Table 5-1. Acute liver injury and biochemistry tests Read codes

Read code/ biochemistry test (enttype)	**Description**
Enttype = 155	Alanine aminotransferase
Enttype = 156	Aspartate aminotransferase
Enttype = 158	Bilirubin
1675	Yellow/jaundiced colour
1675.11	Jaundice - symptom
2274	O/E - jaundiced colour
2274.11	O/E - jaundiced
7806	Therapeutic endoscopic operations on liver using laparoscope
7807	Diagnostic endoscopic examination of liver using laparoscope
7800111	Auxiliary liver transplant
7800112	Piggy back liver transplant
7800500	Orthotopic transplantation of liver NEC
7804200	Open wedge biopsy of lesion of liver
7805211	Exploration of liver transplant
7807000	Diagnostic laparoscopic examination and biopsy liver lesion
7807100	Laparoscopic ultrasound examination liver biop lesion liver
7807200	Laparoscopic ultrasound examination of liver NEC
44CU.00	Plasma alkaline phosphatase level
44D2.00	Liver function tests abnormal
44E.00	Serum bilirubin level
44E2.00	Serum bilirubin raised
44E6.00	Serum bilirubin borderline
44G2.00	Liver enzymes abnormal
44G3100	ALT/SGPT level abnormal
44H5100	AST/SGOT level abnormal
44H5200	AST/SGOT level raised
46R5.11	Bilirubin in urine
7800z00	Transplantation of liver NOS
7807y00	Diagnostic laparoscopic examination of liver OS
7807z00	Diagnostic laparoscopic examination of liver NOS

Read code/ biochemistry test (enttype)	Description
780A.00	Diagnostic percutaneous operations on liver
780A000	Percutaneous transvascular biopsy of lesion of liver
780A100	Percutaneous biopsy of lesion of liver NEC
780A111	Menghini needle biopsy of liver
780A112	Needle biopsy of liver NEC
780A113	Sheeba needle biopsy of liver
780Az00	Diagnostic percutaneous operation on liver NOS
780B000	Biopsy of liver NEC
780B011	Biopsy of lesion of liver NEC
780F000	Endoscopic ultrasound examination liver biopsy lesion liver
9N0v.00	Seen in liver clinic
J60.00	Acute and subacute liver necrosis
J600.00	Acute necrosis of liver
J600000	Acute hepatic failure
J600011	Acute liver failure
J600100	Acute hepatitis – non-infective
J600200	Acute yellow atrophy
J600z00	Acute necrosis of liver NOS
J601.00	Subacute necrosis of liver
J601000	Subacute hepatic failure
J601100	Subacute hepatitis – non-infective
J601200	Subacute yellow atrophy
J601z00	Subacute necrosis of liver NOS
J60z.00	Acute and subacute liver necrosis NOS
J622.00	Hepatic coma
J622.11	Encephalopathy - hepatic
J625.00	[X] Hepatic failure
J625.11	[X] Liver failure
J62y.11	Hepatic failure NOS
J62y.12	Liver failure NOS
J62y.13	Hepatic failure
J63.00	Other liver disorders
J633.00	Hepatitis unspecified

Read code/ biochemistry test (enttype)	Description
J633000	Toxic hepatitis
J633z00	Hepatitis unspecified NOS
J635.00	Toxic liver disease
J635000	Toxic liver disease with cholestasis
J635100	Toxic liver disease with hepatic necrosis
J635200	Toxic liver disease with acute hepatitis
J635700	Acute hepatic failure due to drugs
J635X00	Toxic liver disease, unspecified
J636.00	Central haemorrhagic necrosis of liver
J63y.00	Other specified liver disorder
J63y100	Non-specific reactive hepatitis
J63yz00	Other specified liver disorder NOS
J63z.00	Liver disorder NOS
J66y600	Obstructive jaundice NOS
R024.00	[D] Jaundice (not of newborn)
R024000	[D] Cholemia NOS
R024100	[D] Icterus NOS
R024111	[D] Jaundice
R024z00	[D] Jaundice (not of newborn) NOS
R104000	[D] Transaminase or lactic acid dehydrogenase raised
R104013	[D] Transaminase raised
R104200	[D] Alkaline phosphatase raised
R148.00	[D] Abnormal liver function test
R148.11	[D] LFTs abnormal
R148z00	[D] Abnormal liver function test NOS
ZV42700	[V] Liver transplanted
ZV7C000	[V] Assessment for liver transplant

ALT = alanine aminotransferase; AST = aspartate aminotransferase; LFT = liver function tests; NEC = not elsewhere classified; NOS = not otherwise specified; O/E = on examination; OS = otherwise specified.

Sources: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 5-2. Acute liver injury ICD-10 codes

ICD-10 code	ICD-10 term
K71.1	Toxic liver disease with hepatic necrosis
K71.2	Toxic liver disease with acute hepatitis
K71.6	Toxic liver disease with hepatitis, not elsewhere classified
K71.9	Toxic liver disease, unspecified
K72.0	Acute and subacute hepatic failure
K72.9	Hepatic failure, unspecified
K76.8	Other specified diseases of liver
K76.9	Liver disease, unspecified
R17	Unspecified jaundice, excludes neonatal
Z94.4	Liver transplant

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: http://apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 5-3. Acute kidney injury Read codes

Read code	Description
K00..00	Acute glomerulonephritis
K00..11	Acute nephritis
K000.00	Acute proliferative glomerulonephritis
K001.00	Acute nephritis with lesions of necrotising glomerulitis
K00y100	Acute exudative nephritis
K00y200	Acute focal nephritis
K00y300	Acute diffuse nephritis
K00z.00	Acute glomerulonephritis NOS
K04..00	Acute renal failure
K040.00	Acute renal tubular necrosis
K041.00	Acute renal cortical necrosis
K042.00	Acute renal medullary necrosis
K042.11	Necrotising renal papillitis
K043.00	Acute drug-induced renal failure
K044.00	Acute renal failure due to urinary obstruction
K04y.00	Other acute renal failure
K04z.00	Acute renal failure NOS
14V2.00	H/O: renal dialysis
14V2.11	H/O: kidney dialysis

Read code	Description
7L1A.00	Compensation for renal failure
7L1A.11	Dialysis for renal failure
7L1A000	Renal dialysis
7L1A011	Thomas intravascular shunt for dialysis
7L1Ay00	Other specified compensation for renal failure
7L1Az00	Compensation for renal failure NOS
ZV45100	[V] Renal dialysis status
ZV56.00	[V] Aftercare involving intermittent dialysis
ZV56000	[V] Aftercare involving extracorporeal dialysis
ZV56011	[V] Aftercare involving renal dialysis NOS
ZV56100	[V] Preparatory care for dialysis
ZV56y00	[V] Other specified aftercare involving intermittent dialysis
ZV56y11	[V] Aftercare involving peritoneal dialysis
ZV56z00	[V] Unspecified aftercare involving intermittent dialysis
ZVu3G00	[X] Other dialysis
K0D..00	End-stage renal disease
7L1A200	Haemodialysis NEC
7L1A300	Haemofiltration
7L1A700	Haemoperfusion
K06..00	Renal failure unspecified
K06..11	Uraemia NOS
K060.00	Renal impairment
K060.11	Impaired renal function
Kyu2000	[X] Other acute renal failure
G500400	Acute pericarditis - uraemic
7L1A100	Peritoneal dialysis
7L1A400	Automated peritoneal dialysis
7L1A500	Continuous ambulatory peritoneal dialysis
7L1A600	Peritoneal dialysis NEC
14S2.00	H/O: kidney recipient
7B00.00	Transplantation of kidney
7B00000	Autotransplant of kidney
7B00100	Transplantation of kidney from live donor
7B00111	Allotransplantation of kidney from live donor

Read code	Description
7B00200	Transplantation of kidney from cadaver
7B00211	Allotransplantation of kidney from cadaver
7B00300	Allotransplantation of kidney from cadaver, heart-beating
7B00400	Allotransplantation kidney from cadaver, heart non-beating
7B00500	Allotransplantation of kidney from cadaver NEC
7B00y00	Other specified transplantation of kidney
7B00z00	Transplantation of kidney NOS
ZV42000	[V] Kidney transplanted

H/O = history of; NEC = not elsewhere classified; NOS = not otherwise specified.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 5-4. Acute kidney injury ICD-10 codes

ICD-10 code	ICD-10 term
N00	Acute nephritic syndrome
N10	Acute tubulo-interstitial nephritis
N12	Tubulo-interstitial nephritis, not specified as acute or chronic
N14	Drug- and heavy-metal-induced tubule-interstitial and tubular conditions
N17	Acute renal failure
N19	Unspecified kidney failure
T86.1	Kidney transplant failure and rejection
Y84.1	Kidney dialysis
Z49	Care involving dialysis
Z94.0	Kidney transplant status
Z99.2	Dependence on renal dialysis

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: website: apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 5-5. Chronic kidney disease Read codes

Read code	Read term
1Z12.00	Chronic kidney disease stage 3
K05..00	Chronic renal failure
1Z13.00	Chronic kidney disease stage 4
1Z1..00	Chronic renal impairment
9hE0.00	Except chronic kidney disease qual indic: Patient unsuitable*

K060.00	Renal impairment
66i..00	Chronic kidney disease monitoring
K06..00	Renal failure unspecified
1Z14.00	Chronic kidney disease stage 5
9Ot0.00	Chronic kidney disease monitoring first letter
K08..00	Impaired renal function disorder
6AA..00	Chronic kidney disease annual review
K060.11	Impaired renal function
1Z1C.00	Chronic kidney disease stage 3 without proteinuria
9Ot..00	Chronic kidney disease monitoring administration
9hE1.00	Exc chronic kidney disease quality indicators: Inform dissen*
K050.00	End-stage renal failure
7L1A200	Haemodialysis NEC
7L1A.11	Dialysis for renal failure
9hE..00	Exception reporting: chronic kidney disease quality indicato*
7A60100	Creation of arteriovenous fistula NEC
7L1A100	Peritoneal dialysis
K08z.00	Impaired renal function disorder NOS
K06..11	Uraemia NOS
14V2.00	H/O: renal dialysis
4519.00	Deteriorating renal function
D215000	Anaemia secondary to chronic renal failure
D215.00	Anaemia secondary to renal failure
7L1B000	Insertion of ambulatory peritoneal dialysis catheter
1Z1B.00	Chronic kidney disease stage 3 with proteinuria
1Z15.00	Chronic kidney disease stage 3A
8L50.00	Renal transplant planned
7L1A000	Renal dialysis
K05..11	Chronic uraemia
K0D..00	End-stage renal disease
G22..11	Nephrosclerosis
Read code	Description
ZV45100	[V]Renal dialysis status
7L1B100	Removal of ambulatory peritoneal dialysis catheter
1Z16.00	Chronic kidney disease stage 3B

1Z1E.00	Chronic kidney disease stage 3A without proteinuria
1Z1J.00	Chronic kidney disease stage 4 without proteinuria
1Z1H.00	Chronic kidney disease stage 4 with proteinuria
SP08300	Kidney transplant failure and rejection
1Z1G.00	Chronic kidney disease stage 3B without proteinuria
SP05613	[X] Peritoneal dialysis associated peritonitis

Source: Denburg et al. (2011) [R15-3136[R15-3136][R15-3136][R15-3136]].

Annex Table 5-6.Chronic kidney disease ICD-10 codes

ICD-10 code	ICD-10 term
E132	Other specified diabetes mellitus with incipient diabetes nephropathy adequately or inadequately controlled by insulin, diet, or oral agents
I12	Hypertensive renal disease
I13	Hypertensive renal and heart disease
N08	Glomerular disorders in diseases classified elsewhere
N18	Chronic renal failure

Source: Fleet et al. (2013) [R15-3138].

Annex Table 5-7. Pyelonephritis and sepsis Read codes

Read code	Description
Pyelonephritis	
K101.00	Acute pyelonephritis
K10y000	Pyelonephritis unspecified
K101z00	Acute pyelonephritis NOS
K100600	Calculous pyelonephritis
K10yz00	Unspecified pyelonephritis NOS
K10y.00	Pyelonephritis and pyonephrosis unspecified
K10y000	Pyelonephritis unspecified
K10y100	Pyelitis unspecified
K10y200	Pyonephrosis unspecified
K10y400	Pyelitis in diseases EC
K101000	Acute pyelonephritis without medullary necrosis
K101200	Acute pyelitis
K101300	Acute pyonephrosis
K10y300	Pyelonephritis in diseases elsewhere classified
K103.00	Pyeloureteritis cystica

Read code	Description
K106.00	Candida pyelonephritis
Sepsis	
A38z.11	Sepsis
A3C..00	Sepsis
K190600	Urosepsis
J666.00	Biliary sepsis
L090y00	Sepsis NOS following abortion/ectopic/molar pregnancy
Q404z00	Umbilical sepsis NOS
L40..11	Sepsis - puerperal
A3Cy.00	Other specified sepsis
A3Cz.00	Sepsis NOS
A3C3.00	Sepsis due to Gram-negative bacteria
A3C0100	Sepsis due to Streptococcus group B
AB2y500	Candidal sepsis
A3C0300	Sepsis due to Streptococcus pneumoniae
A3C1.00	Sepsis due to Staphylococcus
A3C2.11	Sepsis due to anaerobes
A270611	Listerial sepsis
A3C1000	Sepsis due to Staphylococcus aureus
A3C0000	Sepsis due to Streptococcus group A
A3C0.00	Sepsis due to Streptococcus
A023.00	Salmonella sepsis
A3C0y00	Other streptococcal sepsis
A396.00	Sepsis due to Actinomyces
A3C2.00	Sepsis due to anaerobic bacteria
A3C3y00	Sepsis due to other Gram-negative organisms
A270600	Sepsis due to Listeria monocytogenes
A3C0z00	Streptococcal sepsis, unspecified
AB2y511	Sepsis due to Candida

NOS = not otherwise specified.

Source: Medical and product dictionary browsers, version 1.3. London: General Practice Research Database (now the CPRD); March 2010. Accessed 15 October 2014..

Annex Table 5-8. Pyelonephritis and sepsis ICD-10 codes

ICD-10 code	Description
A40	Streptococcal sepsis
A41	Other sepsis
N10	Acute tubule-interstitial nephritis
N13.6	Pyonephrosis
N20	Calculus of kidney and ureter

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: http://apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 5-9. Urinary tract infections Read codes

Read code	Description
K190z00	Urinary tract infection, site not specified NOS
K15..00	Cystitis
K190.00	Urinary tract infection, site not specified
1AG..00	Recurrent urinary tract infections
K150.00	Acute cystitis
K190.11	Recurrent urinary tract infection
K17y000	Urethritis unspecified
A994.00	Non-specific urethritis
K190300	Recurrent urinary tract infection
K190500	Urinary tract infection
K15z.00	Cystitis NOS
L166800	Urinary tract infection complicating pregnancy
L166.00	Genitourinary tract infections in pregnancy
K17..00	Urethritis due to non-venereal causes
K171.00	Postmenopausal atrophic urethritis
L166z11	UTI - urinary tract infection in pregnancy
K190400	Chronic urinary tract infection
L166600	Urinary tract infection following delivery
L166300	Genitourinary tract infection in pregnancy - not delivered
L166.11	Cystitis of pregnancy
L166z00	Genitourinary tract infection in pregnancy NOS
K17z.00	Urethritis due to non-venereal cause NOS
K17y.00	Other urethritis

Read code	Description
L166000	Genitourinary tract infection in pregnancy unspecified
K17yz00	Other urethritis NOS
Kyu5500	[X] Other urethritis
L166100	Genitourinary tract infection in pregnancy - delivered
Kyu5100	[X] Other cystitis
K21..11	Prostatitis and other inflammatory diseases of prostate
K211.00	Chronic prostatitis
K210.00	Acute prostatitis
K21z.00	Prostatitis NOS

NOS = not otherwise specified.

Source: Medical and product dictionary browsers, version 1.3. London: General Practice Research Database (now the CPRD); March 2010. Accessed 15 October 2014..

Annex Table 5-10. Urinary tract infections ICD-10 codes

ICD-10 code	Description
N30	Cystitis
N34	Urethritis and urethral syndrome
N37.0	Urethritis in diseases classified elsewhere
N39.0	Urinary tract infection, site not specified
N41	Inflammatory diseases of prostate
O23.2	Infections of urethra in pregnancy
O23.4	Unspecified infection of urinary tract in pregnancy
O23.5	Infections of the genital tract in pregnancy
O23.9	Other and unspecified genitourinary tract infections in pregnancy
O86.2	Urinary tract infection following delivery

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: website: apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 5-11. List of diagnoses and corresponding Read codes and terms to define vulvovaginitis

Read code	Description
"Specific diagnosis" of bacterial vaginosis or *Candida* vulvovaginitis	
A3By700	Gardnerella vaginalis
AB21.00	Candidal vulvovaginitis
AB21.11	Monilial vulvovaginitis
AB21000	Candidiasis of vulva

Read code	Description
AB21100	Candidiasis of vagina
AB21111	Vaginal trush
AB22.00	Other urogenital candidiasis
AB21z00	Candidal vulvovaginitis NOS
K421900	Bacterial vaginitis
K421911	Bacterial vaginosis
Specific microbiology results	
4JK2300	HVS culture - *Gardenella vaginalis*
4JK2400	High vaginal swab: fungal organism isolated
4J74.11	Fungus on microscopy
4KE0.00	Clue cells present
4KE.00	Clue cells
Non-specific positive microbiology results	
4JK2500	High vaginal swab: white cells seen
4JK7.00	Vaginal swab culture positive
4JK2000	High vaginal swab culture positive
4KA2.00	Vaginal vault smear-inadequate
4KA4.00	Vaginal vault smear abnormal
Non-specific diagnosis of vulvovaginitis	
K421.00	Vaginitis and vulvovaginitis
K421000	Vaginitis unspecified
K421100	Vulvitis unspecified
K421200	Vulvovaginitis unspecified
K421400	Vaginitis in diseases EC
K421500	Vulvitis in diseases EC
K421600	Vulvovaginitis in diseases EC
K421A00	Acute vulvitis
K421z00	Vaginitis and vulvovaginitis NOS

EC = elsewhere classified; NOS = not otherwise specified.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 5-12. Preliminary list of diagnoses and corresponding Read codes and terms suggestive of sexually transmitted infection to be used to exclude vulvovaginitis cases

Read code	Read term
4JK2200	HVS culture - trichomonas vaginalis
43U8.00	Chlamydia test positive
K5A3.11	Senile (atrophic) vaginitis
A78A.00	Chlamydial infection
A98z.11	Gonorrhoea
K420900	Chlamydia cervicitis
A541100	Herpetic vulvovaginitis
AD10111	Trichomonal vaginitis
AD10100	Trichomonal vulvovaginitis
Ayu4L00	[X]Vulval warts
43U1.00	Chlamydia antigen ELISA positive
A78A000	Chlamydial infection of lower genitourinary tract
A78AX00	Chlamydial infection of genitourinary tract, unspecified
K5A5.00	Perimenopausal atrophic vaginitis
A78A500	Chlamydial infection of genital organs NEC
A541200	Herpetic ulceration of vulva
4JQA.00	Gonorrhoea test positive
A980.00	Acute gonorrhoea of lower genitourinary tract
Ayu6200	[X]Chlamydial infection, unspecified
A980z00	Acute Gonorrhoea of lower genitourinary tract NOS
Ayu4D00	[X]Sexually transmitted chlamydial infection of other sites
A980200	Acute gonococcal vulvovaginitis
Ayu4K00	[X]Chlamydial infection of genitourinary tract, unspecified
A78A300	Chlamydial inf of pelviperitoneum oth genitourinary organs
Kyu8500	[X]Vaginits,vulvits+vulvovaginitis/infect+parasitc diseas CE
A982200	Chronic gonococcal vulvovaginitis
A982.00	Chronic gonorrhoea lower genitourinary tract
Ayu6100	[X]Other chlamydial diseases
A981z00	Acute gonorrhoea upper genitourinary tract NOS
A981.00	Acute gonorrhoea of upper genitourinary tract
A913500	Secondary syphilis of vulva
Kyu8400	[X]Ulceration of vulva in infectious+parasitic diseases CE

Read code	Read term
A983.00	Chronic gonorrhoea of upper genitourinary tract
A982z00	Chronic gonorrhoea of lower genitourinary tract NOS
K421300	Postirradiation vaginitis
K5A3.00	Postmenopausal atrophic vaginitis
A980200	Acute gonococcal vulvovaginitis
Ayu4K00	[X]Chlamydial infection of genitourinary tract, unspecified
A78A300	Chlamydial inf of pelviperitoneum oth genitourinary organs
Kyu8500	[X]Vaginits,vulvits+vulvovaginitis/infect+parasitc diseas CE
A982200	Chronic gonococcal vulvovaginitis

CE = classified elsewhere, EC = elsewhere classified; NOS = not otherwise specified.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 5-13. List of diagnoses and corresponding Read codes and terms to define balanitis

Read code	Description
K272.11	Infection of penis
K271z00	Balanoposthitis NOS
K271100	Posthitis
K271000	Balanitis
K271.11	Balanitis
K271.00	Balanoposthitis
AB22000	Candidal balanitis
K272300	Cellulitis of penis
K272200	Penile carbuncle
2663.11	O/E - discharge - penis
Kyu6A00	[X]Balanitis in diseases classified elsewhere
AB22011	Penile candidiasis (thrush)
K272100	Penile boil
K272000	Penile abscess
4JK8000	Penile swab culture positive

NOS = not otherwise specified; O/E = on examination.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014..

Annex Table 5-14 Preliminary list of diagnoses and corresponding Read codes and terms suggestive of sexually transmitted infection that will be used to exclude balanitis cases

Read code	Description
K271200	Zoon's balanitis
A78A000	Chlamydial infection of lower genitourinary tract
A78AX00	Chlamydial infection of genitourinary tract, unspecified
A78A500	Chlamydial infection of genital organs NEC
A541300	Herpetic infection of penis
K274.11	Balanitis xerotica obliterans
A060.00	Balantidiasis
A980.00	Acute gonorrhoea of lower genitourinary tract
A980z00	Acute gonorrhoea of lower genitourinary tract NOS
A05y100	Amoebic balanitis
A982z00	Chronic gonorrhoea of lower genitourinary tract NOS
A781212	Penile warts
A78A.00	Chlamydial infection
A98z.11	Gonorrhoea
4JQA.00	Gonorrhoea test positive
Ayu4D00	[X]Sexually transmitted chlamydial infection of other sites
A78A300	Chlamydial inf of pelviperitoneum oth genitourinary organs
Ayu4K00	[X]Chlamydial infection of genitourinary tract, unspecified

NEC = not elsewhere classified; NOS = not otherwise specified; O/E = on examination.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, April 2014. Accessed 15 October 2014.

Annex Table 5-15 . Genital infections ICD-10 codes

ICD-10 code	Description
Vulvovaginitis	
B37.3	Candidiasis of vulva and vagina
B37.4	Candidiasis of other urogenital sites
N77.1	Vaginitis, vulvitis and vulvovaginitis in infectious and parasitic diseases classified elsewhere
N76	Other inflammation of vagina and vulva (includes vaginitis, vulvitis, etc…) Additional codes (B95-B98) are used to identify infectious agent.
Balanitis	
N48.1	Balanoposthitis (additional codes (B95-B98) are used to identify infectious agent)
N51.2	Balanitis in diseases classified elsewhere

ICD-10 code	Description
B37.4	Candidiasis of other urogenital sites
Exclude case of vulvovaginitis or balanitis if any of the following is present 30 days before or after the index diagnosis code:	
A06.8	Amoebic infection of other sites
A50-A64	Infections with a predominantly sexual mode of transmission (includes syphilis, gonococci, chlamydia, trichomonas, herpetical, and other)

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: website: apps.who.int/classifications/icd10/browse/2010/en. Accessed 15 October 2014.

Annex Table 5-16. Preliminary list of diagnoses and corresponding Read codes and terms suggestive of diabetic ketoacidosis

Read code	Description
C362700	Ketoacidaemia NEC
C362600	Metabolic ketoacidaemia
C10FP11	Type II diabetes mellitus with ketoacidotic coma
C10FP00	Type 2 diabetes mellitus with ketoacidotic coma
C10FN11	Type II diabetes mellitus with ketoacidosis
C10FN00	Type 2 diabetes mellitus with ketoacidosis
C10A100	Malnutrition-related diabetes mellitus with ketoacidosis
C103z00	Diabetes mellitus NOS with ketoacidotic coma
C103100	Diabetes mellitus, adult onset, with ketoacidotic coma
C103.00	Diabetes mellitus with ketoacidotic coma
C101z00	Diabetes mellitus NOS with ketoacidosis
C101y00	Other specified diabetes mellitus with ketoacidosis
C101100	Diabetes mellitus, adult onset, with ketoacidosis
C101.00	Diabetes mellitus with ketoacidosis

NEC = not elsewhere classified; NOS = not otherwise specified.

Source: CPRD Medical and Product Dictionary Browsers, version 1.4.0. Clinical Practice Research Datalink, Medicines and Health Products Regulatory Agency, United Kingdom, February 2016. Accessed 16 March 2016.

Annex Table 5-17. Diabetic ketoacidosis ICD-10 codes

ICD-10 code	Description
E11.1	Type 2 diabetes mellitus, with ketoacidosis
E12.1	Malnutrition-related diabetes mellitus, with ketoacidosis
E13.1	Other specified diabetes mellitus, with ketoacidosis
E14.1	Unspecified diabetes mellitus, with ketoacidosis

E11.0	Type 2 diabetes mellitus, with coma
E12.0	Malnutrition-related diabetes mellitus, with coma
E13.0	Other specified diabetes mellitus, with coma
E14.0	Unspecified diabetes mellitus, with coma

Source: International Statistical Classification of Diseases and Related Health Problems 10th Revision. Available at: http://apps.who.int/classifications/icd10/browse/2010/en. Accessed 16 March 2016.

Annex 6. COVARIATES TO BE CONSIDERED FOR INCLUSION IN THE PROPENSITY SCORE MODEL, BY OUTCOME

Acute Liver Injury

Annex Table 6-1. Acute liver injury outcome: variables of interest to be collected for propensity score development

Demographic or lifestyle	
Age	Smoking history
Sex	Alcohol consumption
Calendar year of index date	History of alcohol abuse
Duration of lookback time	Socioeconomic status: Index of multiple socioeconomic deprivation, quintiles: first least deprived, fifth most deprived
Body mass index > 30 or obesity surgery	

Medications		
Drugs associated with liver injury1		
Acarbose	Estrogens	Phenytoin
Acetaminophen (prescription)	Fluoxetine	Pyrazinamide
Allopurinol	Flutamide	Rifampicin
Amiodarone	HAART drugs	Risperidone
Amitriptyline	Irbesartan	Sertraline
Amoxicillin + clavulanic acid	Isoniazid	Statins
Anabolic steroids	Ketoconazole	Sulfonamides
Azathioprine	Lisinopril	Terbinafine
Baclofen	Losartan	Tetracyclines
Bupropion	Methotrexate	Trazodone
Captopril	Mirtazapine	Trazodone
Carbamazepine	Nitrofurantoin	Tricyclics
Chlorpromazine	NSAIDs	Trimethoprim-sulfamethoxazole
Clindamycin	Omeprazole	Trovafloxacin
Clopidogrel	Oral contraceptives	Valproic acid
Cyproheptadine	Paroxetine	Verapamil
Enalapril	Phenobarbital	
Erythromycins	Phenothiazines	

Other medications		
Cardiovascular system drugs	Antiepileptics	Methotrexate
Lipid-modifying agents	Drugs for asthma and obstructive airways disease	Cyclosporin
Other antirheumatic agents	Systemic corticosteroids	Other immunosuppressants excluding systemic tacrolimus
Hormone-replacement therapy	Systemic tacrolimus	Systemic antivirals
Insulins	Azathioprine	Other antimicrobials
Other oral antidiabetic drugs (including specification of add-on or switch)		
Medical comorbidities		
Ischaemic heart disease	Diffuse diseases of connective tissue	Pancreatitis
Hypertensive disease	Rheumatoid arthritis	Urinary infections (chronic or recurring)
Heart failure	Osteoarthrosis	Immunosuppressive diseases such as human immunodeficiency virus infection/AIDS
Peripheral vascular disease	Polymyalgia rheumatica	Being hospitalised, especially for a serious condition that requires intensive care
Other cardiovascular disease	Renal insufficiency	Length of hospitalisation
Cerebrovascular disease	Other malignancies	Chronic disease score2
Hyperlipidaemia	Dementia	
Autoimmune disease	Peptic ulcer disease	
Asthma	Colon polyps	
Chronic obstructive pulmonary disease, emphysema, respiratory insufficiency	Crohn's disease	
	Ulcerative colitis	
Indicators of diabetes severity		
Renal insufficiency or diabetic nephropathy	Coronary heart disease	
Retinopathy	Cerebrovascular disease	
Neuropathy	Amputations	
Peripheral vascular disease	Time since first diagnosis of type 2 diabetes mellitus	

HAART = highly active antiretroviral therapy; NSAIDs = non-steroidal anti-inflammatory drugs.

1. Source: Navarro and Senior, 2006.
2. For example, scores like the ones developed by Elixhauser et al. [R13-5418] or Charlson/Deyo [R13-3589] to be specified in analysis plan.

Hospitalisation for Acute Kidney Injury

Annex Table 6-2. Acute kidney injury outcome: variables of interest to be collected for propensity score development

Demographic or lifestyle	Medications
Age Sex Calendar year of index date Duration of lookback time Body mass index > 30 or obesity surgery Smoking history History of alcohol abuse Socioeconomic status: index of multiple socioeconomic deprivation, quintiles–first least deprived to fifth most deprived	Antihypertensives, diuretics including angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers, beta blockers, calcium-channel blockers, other antihypertensives, antiarrhythmics, digoxin, nitrates NSAIDs (non-steroidal anti-inflammatory drugs) Statins, fibrates Oral steroids Zoledronic acid Lipid-modifying agents Other: acetaminophen, antibiotics (penicillins, sulfa),, anticonvulsants, antifungals, antituberculars, chemotherapeutic agents, methotrexate, aspirin and other antiplatelets (e.g., clopidogrel, ticlopidine, prasugrel), systemic antivirals, anticoagulants Concomitant antidiabetics (including specification of add-on or switch)

Medical comorbidities		Indicators of diabetes severity
Prior history of acute kidney injury > 6 months before index date Being hospitalised, especially for a serious condition that requires intensive care Length of hospitalisation High blood pressure Heart failure Chronic renal disease or renal dialysis Liver disease Peripheral artery disease Chronic disease score1 Other cardiovascular disease Autoimmune disease Chronic obstructive pulmonary disease, emphysema, respiratory insufficiency Diffuse diseases of connective tissue	Rheumatoid arthritis Osteoarthrosis Polymyalgia rheumatica Urinary infections (chronic or recurring) Kidney stones Crohn's disease Ulcerative colitis Pancreatitis Immunosuppressive diseases such as HIV/AIDS Peptic ulcer disease Dementia Asthma	Renal insufficiency or diabetic nephropathy, peripheral Retinopathy Neuropathy Peripheral vascular disease Coronary heart disease Cerebrovascular disease Amputations Time since first diagnosis of type 2 diabetes mellitus, if available

1. For example, scores like the ones developed by Elixhauser et al. [R13-5418] or Charlson/Deyo [R13-3589], to be specified in analysis plan.

Hospitalisation for Urinary Tract Infection

Annex Table 6-3. Urinary tract infection outcome: variables of interest to be collected for propensity score development

Demographic or lifestyle	Medications
Age Sex Calendar year of index date Duration of lookback time Body mass index > 30 or obesity surgery Smoking history History of alcohol abuse Socioeconomic status: Index of multiple socioeconomic deprivation, quintiles—first least deprived to fifth most deprived	Antihypertensives/diuretics including angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers, beta blockers, calcium-channel blockers, other antihypertensives, antiarrhythmics, digoxin, nitrates NSAIDs (non-steroidal anti-inflammatory drugs) Statins, fibrates Oral steroids Zoledronic acid Lipid-modifying agents Other: acetaminophen, antibiotics (e.g., penicillins, sulfa), anticonvulsants, antifungals, antituberculars, chemotherapeutic agents, methotrexate, aspirin and other antiplatelets (e.g., clopidogrel, ticlopidine, prasugrel), anticoagulants, systemic antivirals Concomitant antidiabetics (including specification of add-on or switch)

Medical comorbidities		Indicators of diabetes severity
Prior history of UTI leading to hospitalisation or acute pyelonephritis > 6 months before index date Being hospitalised, especially for a serious condition that requires intensive care Length of hospitalisation Kidney diseases Kidney and genitourinary stones and disease Pregnancy High blood pressure Heart failure Liver disease Other cardiovascular disease Autoimmune disease Chronic obstructive pulmonary disease, emphysema, respiratory insufficiency Diffuse diseases of connective tissue	Rheumatoid arthritis Osteoarthrosis Polymyalgia rheumatica Urinary infections (chronic or recurring) Crohn's disease Ulcerative colitis Pancreatitis Immunosuppressive diseases such as HIV/AIDS Peptic ulcer disease Dementia Asthma Chronic disease score1	Renal insufficiency or diabetic nephropathy, peripheral Retinopathy Neuropathy Peripheral vascular disease Coronary heart disease Cerebrovascular disease Amputations Time since first diagnosis of type 2 diabetes mellitus, (CPRD only)

HIV = human immunodeficiency virus.

1. For example, scores like the ones developed by Elixhauser et al. (1998) [R13-3591] or Charlson/Deyo [R13-3589], to be specified in analysis plan.

Genital infections

Annex Table 6-4. Genital infections outcome: variables of interest to be collected for propensity score development

Demographic or lifestyle	Medications
Age Sex Calendar year of index date Duration of lookback time Body mass index > 30 or obesity surgery Smoking history History of alcohol abuse Socioeconomic status: Index of multiple socioeconomic deprivation, quintiles—first least deprived to fifth most deprived	Antihypertensives/diuretics including angiotensin-converting enzyme inhibitors/angiotensin II receptor blockers, beta blockers, calcium-channel blockers, other antihypertensives, antiarrhythmics, digoxin, nitrates NSAIDs (non-steroidal anti-inflammatory drugs) Statins, fibrates Oral steroids Zoledronic acid Lipid-modifying agents Other: acetaminophen, antibiotics (e.g., penicillins, sulfa), anticonvulsants, antifungals, antituberculars, chemotherapeutic agents, methotrexate, aspirin and other antiplatelets (e.g., clopidogrel, ticlopidine, prasugrel), anticoagulants, systemic antivirals Concomitant antidiabetics (including specification of add-on or switch)

Medical comorbidities		Indicators of diabetes severity
Prior history of genital infections Being hospitalised, especially for a serious condition that requires intensive care Length of hospitalisation Kidney diseases Kidney and genitourinary stones and disease Pregnancy High blood pressure Heart failure Liver disease Other cardiovascular disease Autoimmune disease Chronic obstructive pulmonary disease, emphysema, respiratory insufficiency Diffuse diseases of connective tissue	Rheumatoid arthritis Osteoarthrosis Polymyalgia rheumatica Urinary infections (chronic or recurring) Crohn's disease Ulcerative colitis Pancreatitis Immunosuppressive diseases such as HIV/AIDS Peptic ulcer disease Dementia Asthma Chronic disease score1	Renal insufficiency or diabetic nephropathy, peripheral Retinopathy Neuropathy Peripheral vascular disease Coronary heart disease Cerebrovascular disease Amputations Time since first diagnosis of type 2 diabetes mellitus, (CPRD only)

HIV = human immunodeficiency virus.

1. For example, scores like the ones developed by Elixhauser et al. (1998) [R13-3591] or Charlson/Deyo [R13-3589], to be specified in analysis plan.

Annex Table 6-5. Diabetic ketoacidosis outcome: variables of interest to be collected for propensity score development

Demographic or lifestyle	**Medications**
Age	Clozapine or olanzapine
Sex	Lithium
Calendar year of index date	Terbutaline
Duration of lookback time	Corticosteroids
Body mass index > 30 or obesity surgery	Thiazides
Smoking history	Pentamidine
History of alcohol abuse	Concomitant non-insulin antidiabetics (including specification of add-on or switch)
Socioeconomic status: Index of multiple socioeconomic deprivation, quintiles—first least deprived to fifth most deprived	Insulin therapy, change in insulin dose
Cocaine	
Medical comorbidities	**Indicators of diabetes severity**
Prior history of DKA	Renal insufficiency or diabetic nephropathy, peripheral
Acute illness: infections (such as UTI, gastroenteritis, urosepsis, influenza), recent surgery, or trauma	Retinopathy
Thyroid problems (e.g., thyroid storm and thyrotoxicosis)	Neuropathy
Myocardial ischemia/infarction	Peripheral vascular disease
Pancreatitis	Coronary heart disease
Psychological stress	Cerebrovascular disease
Reduced caloric or fluid intake	Amputations
Hypovolemia	Time since first diagnosis of type 2 diabetes mellitus
Alcohol intake	
Hypoxemia	
Acute renal failure	
Heart Failure	
Cerebrovascular accident	
Chronic disease score1	

UTI = urinary tract infection.

1. For example, scores like the ones developed by Elixhauser et al. (1998) [R13-3591] or Charlson/Deyo [R13-3589], to be specified in the analysis plan.

Annex 7. OVERVIEW OF CHARACTERISTICS OF THE CLINICAL PRACTICE RESEARCH DATALINK

Annex Table 7-1. Characteristics of the Clinical Practice Research Datalink

Characteristic	**United Kingdom (, 62,435,709)1**
Database type	Primary care electronic medical records of patients enrolled in practices contributing to the CPRD. Linkage to hospital data (Hospital Episode Statistics [HES]), mother-child data, practice-level socioeconomic data, death certificates (Office for National Statistics), cancer and cardiovascular disease registries, and others is possible. Linkage is available for a proportion of the practices.
Database population (n)	5.69 million2
Population covered, description	Most UK residents are registered with a GP. Patients registered with practices that contribute to the CPRD are included. Prisoners and members of the armed forces are not included. The homeless are underrepresented.
Proportion of the country's population covered	10.9%
Representativeness of patients and practices	Age and sex of patients are representative of the UK population
Demographics	
Lifestyle risk factors	Yes, but missing data. Marital status is updated, but there is no information on marital status at the time of a past event
Geographic location	First digits of physician's practice postal code
Medication information	
Source	All prescriptions issued by GPs. Repeat prescriptions may be implemented. There is a sequence number to know whether the prescription is new. The presence of a repeat prescription does not ensure that the prescription was picked up (or filled).
Drug dictionary codes/ therapeutic classification	Multilex/British National Formulary
Unique product code	Yes
Prescribed/dispensed drugs	GP prescriptions issued
Date drug prescribed/ dispensed	Yes, date the drug was prescribed
Dose	Yes, but it is not a mandatory field. The dose is a text code and requires some handling to be transformed into a number. This transformation may be performed by the researcher or by the CPRD.
Duration	There is a field to record duration, but it is highly incomplete. Duration can be derived from the number of prescriptions.

Clinical indication	There is no field for indication. The user needs to assess diagnoses on the prescription date or obtain free-text data. Prior diagnosis can be used as a proxy.
Inpatient medications	No
Specialist-prescribed medications	Only if the GP decided to include these in the medical record. GPs typically issue repeat prescriptions; there is a higher risk for not capturing the first specialist-initiated prescriptions than subsequent ones.
Computerised free-text comments available	Yes
Diagnoses and procedures	
Coding system	Read
Outpatient visits	Yes, as entered by the GP
Hospitalisation data	Partial linkage to HES (ICD-10 codes); Read codes as recorded by GPs
Specialist visits	Information from referral letters
Emergency department visits	As entered by the GP
Time period covered	Since 1987
Updates	Quarterly
Approximate time lag	6-12 weeks
Access to medical records	GPs can be sent questionnaires via the CPRD for validation; also partial linkage to HES
Data transfer	Yes, third-party approval for standard data and linked databases. Data set will be delivered for analysis
Approval process	ISAC approval of short protocol

CPRD = Clinical Practice Research Datalink; GP = general practitioner or general practice; HES = Hospital Episode Statistics; ICD-10 = *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*; ISAC = Independent Scientific Advisory Committee; UK = United Kingdom.

1. In thousands. Population data from Eurostat [R14-5313].
2. Based on "active" patients (i.e., patients having consulted a GP at least once in the year) [R14-5257].

Annex 8. Key Features of Data sources in Denmark and Spain

Annex Table 8-1. Characteristics of the data sources available in Denmark and Spain

Database feature	Danish National Patient and Prescription Registers	EpiChron	SIDIAP (Information System for the Advancement of Research in Primary Care)
Population of country1	Denmark: 5,627,235	Spain: 46,512,199	Spain: 46,512,199
Database type	National health record databases, link to other national databases through a unique personal identification number	Primary health care electronic medical record database; link to hospital discharge data and pharmacy data	Primary health care electronic medical record database, link to hospital discharge data, pharmacy data, and mortality data
Data on medications and type of prescriptions	All (reimbursed and non-reimbursed) pharmacy-dispensed prescriptions; in regional databases, only reimbursed prescriptions	Reimbursed pharmacy-dispensed prescriptions	Reimbursed pharmacy-dispensed prescriptions and electronically prescribed drugs
Drug dictionary codes/therapeutic classification	ATC	ATC	ATC
Disease and procedure coding system(s)	ICD-10-CM	Primary health care, ICPC; hospital, ICD-9-CM	ICD-10-CM
Laboratory (requests, results)	No	Yes	Yes
Data availability	Since 1994	Partial since 2005; complete 2010 through 2013	2006 to Dec 2014
Approximate time lag (updates per year)	1 year (1 per year)	1 year (1 per year)	1 year (2 per year)
Access to hospital medical records	Yes	Yes	No

ATC = Anatomical Therapeutic Chemical (classification system); GP = general practitioner; ICD-10-CM = *International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Clinical Modification;* ICD-9-CM = *International Classification of Diseases, 9th Revision, Clinical Modification;* ICPC = International Classification of Primary Care; SIDIAP = Information System for the Advancement of Research in Primary Care (Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària).

1. Population data from Eurostat (2014) [R14-5313].