Protocol Number: MB102-104 ST

Date: 28 February 2017

Observational Study Protocol MB102-104 ST

COMPARISON OF THE RISK OF ACUTE LIVER INJURY BETWEEN PATIENTS WITH TYPE 2 DIABETES EXPOSED TO DAPAGLIFLOZIN AND THOSE EXPOSED TO OTHER ANTIDIABETIC TREATMENTS

AZ Study Director
Robert J. LoCasale, PhD, MS
[redacted]

External Investigators
Catherine Johannes, PhD;
RTI Health Solutions Principal Investigator
[redacted]

Stephan Lanes; Principal Epidemiologist, Safety and Epidemiology
HealthCore, Inc.
[redacted]

AstraZeneca Pharmaceuticals LP
[redacted]

SYNOPSIS

Observational Study Protocol MB102-104 ST

Protocol Title: Comparison of the Risk of Acute Liver Injury Between Patients With Type 2 Diabetes Exposed to Dapagliflozin and Those Exposed to Other Antidiabetic Treatments

Department: AstraZeneca Epidemiology

Objectives

Primary Objective: To compare, by insulin use at the index date, the incidence of hospitalization for acute liver injury (ALI) among patients with type 2 diabetes mellitus who are new users of dapagliflozin with those who are new users of antidiabetic drugs (ADs) in classes other than sodium-glucose cotransporter 2 (SGLT2) inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

Secondary Objective #1: To compare, by insulin use at the index date, baseline patient characteristics of patients with type 2 diabetes mellitus who are new users of dapagliflozin with those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy and to identify important prognostic variables that should be balanced between the study groups.

Secondary Objective #2: To examine potential risk factors for ALI if new users of dapagliflozin are found to be at greater risk for this outcome than new users of other ADs.

Study Design: This will be a cohort study that will be conducted with data from the Clinical Practice Research Datalink (CPRD) in the United Kingdom (UK) and the HealthCore Integrated Research Database (HIRD^{SM}) and Medicare databases in the United States of America (US). The study will compare the incidence of hospitalizations for ALI among new users of dapagliflozin with that among patients who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy. The planned study duration is 5 years; however, duration will depend on the market uptake of dapagliflozin.

Study Population: Patients will be eligible for inclusion in this study if they meet *all* of the following criteria: (1) receive newly prescribed dapagliflozin (with or without other ADs) or a newly prescribed AD (with or without other ADs) in a class other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy; (2) do not have evidence of type 1 diabetes; (3) are aged 18 years or older at the index date for CPRD patients, 18-64 years for HIRD^{SM} patients, or 65 years or older for Medicare patients; and (4) have been enrolled in the data source for at least 180 days before the first prescription or dispensing for dapagliflozin or comparator AD. Patients with a previous diagnosis of ALI; liver, biliary, or pancreatic disease; hepatobiliary or pancreatic neoplasm; or congestive heart failure will be excluded. In the HIRD^{SM}, all available eligible comparator patients will be used in the analysis. In the CPRD and Medicare, it may be necessary to select a subsample of AD users before propensity score estimation if all available data cannot be used. If used, the subsample will be identified by frequency matching eligible comparator patients to dapagliflozin patients by 5-year age groups, sex, geographic region, and calendar year of the index date in at least a 6:1 ratio in the CPRD and a 15:1 ratio in Medicare.

Data Collection Methods

Data Sources: The CPRD contains electronic medical records including outpatient diagnoses and prescriptions from general practitioner (GP) practices in the UK and mentions of diagnoses associated with hospitalizations. The HIRD^{SM} contains health insurance claims from the largest commercially insured population in the US. Medicare data include health insurance claims from the federally sponsored health insurance program for individuals in the US aged 65 years or older and individuals with permanent disabilities.

Exposures: New use of dapagliflozin will be defined to begin on the date of first dapagliflozin prescription or dispensing recorded in the data source. New use of an AD in a class eligible for inclusion will be defined to begin on the date of first prescription or dispensing for these medications in the data source.

Outcomes: The primary outcome in this study will be hospital admission for ALI. In the CPRD, additional clinical information will be obtained for a subset of potential cases via a questionnaire sent to the GP. In the HIRD^{SM} and Medicare cohorts, individuals with health insurance claims for hospitalization with ALI will be identified as potential cases. For a subset of these cases, the hospital medical records will be abstracted or photocopied and redacted to confirm the diagnosis and diagnosis date of ALI. If the electronic algorithm for

identifying potential cases is associated with a lower 95% confidence interval for the positive predictive value less than 0.80, then the case ascertainment algorithm will be modified.

Follow-up: Follow-up will begin on the day after the index date. The index date is the day a patient is first prescribed or dispensed dapagliflozin or a comparator AD. Follow-up time for a given exposure will continue until hospitalization for ALI; death; the end of study data or study period; initiation of an SGLT2 inhibitor other than dapagliflozin; or the end of the risk window for the index AD, defined as 30 days after the estimated discontinuation of dapagliflozin or the index comparator AD; whichever occurs first. Sensitivity analyses will involve each of the following variations, one at a time: (1) the risk window will be extended to 90 days after the estimated discontinuation of the index AD; (2) follow-up will end if new use of any study AD is added; and (3) comparator cohorts will be limited to new users of any study drug class.

Data Analyses: Descriptive statistics will be generated to compare baseline characteristics (e.g., demographic information, comorbidities, and medication use at the index date) between dapagliflozin initiators versus comparator AD initiators, by insulin use at index date. Propensity scores will be estimated by logistic regression analyses, incorporating measured potential predictors of therapy and calendar year of the index date as independent variables in the regression model and exposure group (dapagliflozin group vs. comparator group) as the outcome. Duration of lookback time and timing of information on key covariates will be included in the model. Incidence rates of ALI will be determined in each cohort. Propensity score–stratified analysis will be used to estimate unadjusted and adjusted incidence rate ratios (with 95% confidence intervals) of the outcome of interest in dapagliflozin initiators versus other AD initiators. If necessary, further adjustment for therapies added during follow-up will be performed via time-dependent Cox proportional hazards regression. Analyses will be conducted in each data source, and a pooled estimate will be calculated if deemed appropriate.

Sample Size/Power: An estimate of the occurrence of ALI in the general diabetes population treated with oral antidiabetics is 22 per 100,000 person-years. It is currently projected that at the end of 5 years there will be a total of 91,927 person-years of exposure to dapagliflozin across the three data sources (CPRD: 3,600 person-years; HIRD^{SM}: 42,473 person-years; and Medicare: 45,854 person-years). This includes approximately 73,542 person-years among those not on insulin at the index date and 18,385 person-years among those on insulin at the index date. Using these estimates and assuming that the true dapagliflozin:comparator incidence rate ratio is 1.0 and there are 73,000 person-years of dapagliflozin follow-up among new users of dapagliflozin not on insulin at the index date, there will be an 85% probability that the upper 95% confidence limit of the observed incidence rate ratio will be less than 2.3.

Limitations/Strengths: The UK is an ideal setting for population-based studies of diabetes because diabetes care is largely coordinated by the GP, and information about metabolic parameters, cardiovascular risk factors, diabetes comorbidities, and disease outcomes are collected electronically. In addition, this population-based data source provides data entered by the GP practices without any awareness of the hypothesis or of an ongoing study. Furthermore, clinical guidelines in the UK facilitate consistency in patterns of care. However there are still limitations. In the CPRD, there may be inaccuracies in the recorded dates of hospitalization and information about care by specialists may be missing. Patient-specific variability in the availability of laboratory data and medical records, misclassification of exposures and outcome, and limitations in the numbers of available of study patients are the major potential limitations of this study.

In the HIRD^{SM} and Medicare cohorts, the health insurance claims databases include claims for all medical services for cohort members during the study period. The Medicare data cover a very large proportion of US residents aged 65 years or older, and the HIRD^{SM} covers a large proportion of the US population younger than 65 years of age. Information on potentially important confounders such as high body mass index and smoking is virtually nonexistent unless treatment for either is detectable through claims. Therefore, an evaluation of the impact of missing confounders is planned.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Dapagliflozin (BMS-512148) is a highly potent, selective, and reversible inhibitor of human renal sodium-glucose cotransporter 2 (SGLT2), the major transporter responsible for renal glucose reabsorption. Dapagliflozin is the first approved drug in its class. It lowers plasma glucose by inhibiting the renal reabsorption of glucose, thereby promoting urinary excretion of glucose, making it a member of an emerging therapeutic class in the treatment of type 2 diabetes mellitus (T2DM) (Bristol-Myers Squibb [BMS] and AstraZeneca [AZ], 2011).

Epidemiology of Acute Liver Injury

Very few studies have estimated the incidence of acute liver injury (ALI) in the general population. In a study conducted in the United States of America (US), the incidence of acute hepatitis from any cause was 8.4 cases per 100,000 persons per year (Carson et al., 1993). Alcoholic cirrhosis, nonalcoholic nonbiliary cirrhosis, infectious hepatitis, chronic hepatitis, and toxic hepatitis have been reported as the most frequent parenchymal liver diseases potentially causing liver injury and failure (Almdal and Sørensen, 1991).

Drug-Induced Acute Liver Injury

Drug-induced ALI is one of the most common forms of drug toxicity, and drug-induced liver injury has been the most frequent single cause of safety-related drug marketing withdrawals in the last several decades (US Food and Drug Administration [FDA], 2009).

The incidence of drug-induced liver injury is difficult to assess as its clinical ascertainment is usually based in the exclusion of other causes of ALI. In most epidemiological studies on drug-induced ALI, patients with other potential causes or risk factors for liver injury (e.g., chronic liver disease) are usually excluded from the study population.

Very few studies have estimated the incidence of drug-induced ALI in the general population. A summary of the main characteristics and results from these studies is presented in Appendix 5. In these studies, the annual incidence of drug-induced ALI in the general population ranged from 0.7 cases per 100,000 persons (95% confidence interval [CI], 0.6-0.9) (Ibáñez et al., 2002) to 13.9 cases per 100,000 persons (95% CI, 9.7-19.5) (Sgro et al., 2002). The study with the highest incidence was conducted in France using prospective intensive surveillance of cases in a well-defined geographical region (Sgro et al., 2002). In a study conducted in the United Kingdom (UK), the incidence rate was 2.4 per 100,000 person-years (95% CI, 2.0-2.8) (de Abajo et al., 2004).

Incidence of Acute Liver Injury in Patients with Diabetes

Very few studies have examined the risk of ALI in patients with diabetes. In a cohort study using the General Practice Research Database (now the Clinical Practice Research Datalink [CPRD])

in the UK, the incidence of ALI in patients with diabetes was 14.2 cases per 100,000 person-years (Huerta et al., 2002). The incidence rate in users of oral antidiabetic drugs (ADs) was 22.0 cases per 100,000 person-years, and the rate in users of insulin was 13.8 per 100,000 person-years. The incidence in the general population without diabetes was 8.8 per 100,000 person-years. The study excluded patients with a history of liver, biliary, or pancreatic disease and did not report on potential risk factors for ALI. Among patients with diabetes, the adjusted relative risk comparing users of ADs (oral ADs and/or insulin) and nonusers of these medications was 2.8 (95% CI, 0.6-12.5).

Challenges in the Identification of Acute Liver Injury

The diagnostic criteria for ALI are based on elevations of serum enzymes such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), and alkaline phosphatase (AP) and/or elevations of total or conjugated bilirubin. Overall, the definitive ascertainment of ALI in populations enrolled in administrative databases requires access to laboratory values to confirm potential cases identified with diagnosis codes. The existing studies evaluating ALI in patients with diabetes identified potential cases using diagnostic codes related to liver injury and validated these cases by reviewing medical records (Chan et al., 2003; Huerta et al., 2002).

1.1 Study Rationale

There were no meaningful, consistent changes from baseline in mean liver function test values across studies in the clinical development program for dapagliflozin, and there were no clinically meaningful differences in liver function test values between the dapagliflozin and placebo groups. There were also no cases of severe drug-induced liver injury, defined as fatal or requiring liver transplantation. In both dapagliflozin and control groups, 5.7% of patients had elevated values for liver tests based on laboratory values and/or reported adverse events (AEs) of hepatic disorder in the All Phase 2b and 3 Pool, the data source used for safety analyses at the time of filing, comprising all clinical data from the phase 2b and 3 data on dapagliflozin and placebo. There was no imbalance in the proportion of patients with laboratory values for ALT or AST greater than 3 times the upper limit of normal (ULN) and concomitant or subsequent total bilirubin (TB) greater than 2 times ULN up to the 4-month safety update. Five (0.1%) patients treated with dapagliflozin had ALT or AST values greater than 3 times ULN and concomitant or subsequent TB greater than 2 times ULN, versus three (0.2%) treated with the comparator (two with placebo and one with glipizide). All had possible underlying causes of these elevations (BMS and AZ, 2011).

This post-authorization safety study is being conducted as part of the BMS/AZ Dapagliflozin Risk Management Plan to monitor the safety of dapagliflozin in real-world use. Although there was no overall signal of liver toxicity from dapagliflozin use, based on hepatic enzyme

monitoring during the dapagliflozin clinical program of studies, one case of possible drug-induced liver injury was experienced by a patient in the dapagliflozin arm. Therefore, there is interest in evaluating whether there is increased risk of ALI among dapagliflozin users compared with users of other ADs. This protocol describes a cohort study to be conducted in the CPRD (UK) and two US data sources: the HealthCore Integrated Research Database (HIRD^{SM}) and Centers for Medicare and Medicaid Services (CMS) Medicare databases. The study will compare the incidence of ALI among new users of dapagliflozin with the incidence of ALI among those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

1.2 Research Question

What is the risk of hospitalization for ALI for patients with T2DM who are new users of dapagliflozin compared with those who are new users of other ADs?

2 STUDY OBJECTIVES

2.1 Primary Objective

Primary Objective: To compare, by insulin use at the index date, the incidence of hospitalization for ALI among patients with T2DM who are new users of dapagliflozin to those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.

2.2 Secondary Objectives

Secondary Objective #1: To compare, by insulin use at the index date, baseline patient characteristics between patients with T2DM who are new users of dapagliflozin with those who are new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy and to identify important prognostic variables that should be balanced between the study groups.

Secondary Objective #2: To examine potential risk factors for ALI if patients taking dapagliflozin are found to be at greater risk for this outcome than patients newly starting other ADs.

2.3 Exploratory Objective

Not applicable.

3 STUDY DESIGN

3.1 Overview of Study Design

This cohort study will compare the incidence of hospitalization for ALI among new users of dapagliflozin with that among new users of ADs in classes other than SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy. As currently designed, the source study populations are the UK and the US. A cohort design will allow for direct estimation of the incidence and risk of the outcome of interest associated with dapagliflozin. Further, the cohort design permits assessment of the outcome at multiple time points. The planned study duration is 5 years; however, actual duration will depend on the market uptake of dapagliflozin.

The index date for each patient will be defined as the date the patient is newly prescribed or dispensed either dapagliflozin (single-entity dapagliflozin or the fixed-dose combination of dapagliflozin and metformin) or an eligible comparator AD after the beginning of the study observation period, according to the time of approval of dapagliflozin in each country. The lookback time, all available data before the index date, will be used to evaluate patient characteristics among the exposure groups and the potential for confounding. The follow-up time, which begins the day after the index date, will be used to evaluate the incidence of ALI and will be used to identify other confounders, effect modifiers, and exposures not controlled at baseline or that change during follow-up. Propensity scores will be estimated at each planned data cut and will be adjusted by calendar year. Multiple logistic regression models will be used to compute propensity scores, and adjusted incidence rate ratios (IRRs) will be derived by propensity score stratification. Other analytic methods that do not control for intermediate variables to evaluate and adjust for added concomitant AD during follow-up will be considered.

3.2 Study Population

During the conduct of the study, patients will be identified at selected intervals. Study populations of patients with T2DM will be identified using data on general practitioner (GP) diagnoses and prescriptions in the CPRD in the UK and health insurance claims for outpatient medication dispensings in the HIRD^{SM} and CMS Medicare databases of the US. These patients will be new users of dapagliflozin or other selected ADs, as detailed in Section 3.2.1, Inclusion Criteria.

3.2.1 Inclusion Criteria

Patients are eligible for the study if they meet *all* of the following criteria:

- Were newly prescribed or dispensed dapagliflozin (with or without other ADs) or are newly prescribed or dispensed an AD (with or without other ADs) in a class other than

SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy (see Appendix 4)

- Were enrolled in the data source for at least 180 days before the first prescription or dispensing for dapagliflozin or other AD qualifying for the comparator group
- Were aged (at the index date)
  - 18 years or older in the CPRD
  - 18-64 years in the HIRD^{SM} or
  - 65 years or older in Medicare and were participants only in the fee-for-service program (i.e., were not in a managed care program); were enrolled in Parts A, B, and D of the Medicare program for at least 180 days before entering the study (follow-up will be censored if Part D coverage was discontinued); had a residence in a US state or district of Columbia; and Medicare eligibility was not due to end-stage renal disease.

Our rationale for comparing new users of dapagliflozin with new users of ADs *in a class other than SGLT2 inhibitors* is to ensure that we do not miss potentially important associations that are due to the SGLT2 class after more such compounds become available. Analysis results in each cohort will be evaluated separately by insulin use at the index date because insulin use is clinically considered to be associated with a longer history of diabetes.

Our rationale for not including new users of metformin monotherapy or sulfonylurea monotherapy is that patients diagnosed with T2DM are likely to be prescribed these medications early in the course of the disease—e.g., following guidelines of the UK National Institute for Health and Clinical Excellence (2015) and the American Diabetes Association (2014)—whereas dapagliflozin is expected to be used after initial treatment with these therapies. In addition, the long history of availability of sulfonylureas and metformin could result in misclassification of new use when these medications are started again after an interruption of treatment.

Further, in clinical practice, patients may be newly prescribed dapagliflozin or another AD with or without other ADs already prescribed as part of their regimen (i.e., patients may have new AD medications added on or they may switch agents). Therefore, we plan to include patients, regardless of whether or not they are taking other ADs at the time they are newly prescribed either dapagliflozin or an eligible comparator AD. In addition, we will collect information on whether patients received prior AD therapy and/or if they were "added on" or "switched to" dapagliflozin or other ADs at the time of inclusion in the study.

3.2.2 Exclusion Criteria

The clinical diagnosis of drug-induced liver injury is usually accomplished by the exclusion of other causes of ALI. Therefore, most epidemiological studies on drug-induced ALI, including those conducted in patients with diabetes, excluded those cases that had other known causes of liver injury (e.g., chronic liver disease, viral hepatitis, hepatocarcinoma) (Appendix 6). Exclusion

of these cases from the numerator could result in biased effect estimates if the same criteria are not applied to the denominator, that is, to all cohort members or person-time at risk (Rothman and Ray, 2002). Therefore, patients will be excluded if they meet *any* of the following criteria:

- The study outcome of ALI was experienced by a patient at any time before or at the index date (i.e., during the available lookback time). *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) codes to be mapped to Read codes for these exclusion conditions are provided in Appendix 1.
- A diagnosis recorded at any time before or at the index date for any of the following chronic conditions and, for the acute conditions, up to 2 years before or at the index date (Appendix 1):
  - Chronic liver disease
  - Chronic alcoholism
  - Chronic or acute infectious hepatitis
  - Chronic disease involving the liver or causing hyperbilirubinemia
  - Cholelithiasis and cholecystitis
  - Intra- or extrahepatic biliary obstruction
  - Pancreatic disease
  - Primary or secondary hepatic, biliary, or pancreatic cancer
  - Congestive heart failure
- The patient was prescribed an SGLT2 inhibitor other than dapagliflozin on or before the index date.
- The patient initiated metformin or sulfonylurea as AD monotherapy at the index date.
- The patient initiated insulin monotherapy at the index date.
- The patient had evidence of type 1 diabetes before cohort entry or first recorded AD is insulin monotherapy.

3.2.3 Selection of Patients

Eligible study patients will be selected from the study data sources separately. All eligible patients in each data source who meet inclusion and exclusion criteria and are new users of dapagliflozin will be selected for inclusion. In the HIRD^{SM}, all available eligible comparator patients will be used in the analysis. In the CPRD and Medicare, it may be necessary to select a subsample of AD users before propensity score estimation if it is not feasible to use all available data. If used, the subsample will be identified by frequency matching eligible comparator patients with eligible comparator episodes to dapagliflozin new users by 5-year age groups, sex, geographic region, and calendar year of the index date in at least a 6:1 ratio in the CPRD and at least a 15:1 ratio in Medicare. The rationale for selecting at least 6 comparator AD new users for each dapagliflozin new user in CPRD and at least 15 comparator AD new users in Medicare is to ensure that we have sufficient numbers of patients to develop the propensity score and to conduct secondary analyses, as needed.

Patients will enter the study cohort based on the first new use of a qualifying study medication after all inclusion criteria are met. However, inclusion in the study cohort as a new user of dapagliflozin or one of the comparator AD medications does not preclude the patient from being included as a new user of another study medication if the criteria for inclusion in the study cohort are met.

3.2.4 Follow-up of Patients

Follow-up will begin on the day after, but not including, the index date, the day a patient is newly prescribed or dispensed dapagliflozin or a comparator AD. Since ALI could occur soon after exposure to a medication, we assume that patients prescribed dapagliflozin or comparator AD could be at risk for the study outcome the day after the index date.

Follow-up time in a given exposure category will continue until hospitalization for ALI, death, the end of study data or study period, or the end of the risk window for the index AD, whichever occurs first. Discontinuation will be defined as no further prescription 30 or more days after the end of the days' supply of the last consecutive prescription in the index exposure episode (assigned to be 30 days if days' supply is missing). Follow-up will be censored at the addition of a non-dapagliflozin SGLT2 inhibitor in either group. Follow-up will *not* be censored if other ADs are prescribed in addition to dapagliflozin or the comparator AD after the index date. If a patient develops type 1 diabetes during follow-up (physician diagnosis in the CPRD or fulfillment of a claims definition in the HIRD^{SM} or Medicare), follow-up time will be censored at the date of diagnosis.

3.2.5 Exposure and Time at Risk

We assume that the potential increased risk of ALI related to drug injury could occur at the beginning of therapy, be maintained at an increased level during the duration of treatment, and decrease gradually to the background risk after stopping treatment. According to the report of an international consensus meeting, ALI occurring within 30 days of stopping therapy may be compatible with drug-induced liver injury (Bénichou, 1990). Therefore the time window of risk relating to use of the index drug will start the day after the index date and will end 30 days after the end of the last prescription's days' supply (assigned to be 30 days if missing) for the index drug.

For most patients, the risk window will end 60 days after the start of the last prescription (assuming the last prescription was a 30-day supply) for the index AD. Adding 30 days after the estimated end of supply will capture a potential effect after stopping therapy and any delay in the detection of and hospitalization for ALI. For sensitivity analyses, the risk window for the index drug will extend to 90 days after end of the last prescription's days' supply of the index AD (see

Appendix 2). This sensitivity assessment will allow exploration of any further potential delay in effect. We selected 90 days because this period is long enough to account for noncompliant and extended use of the discontinued index exposure and a delay in effect. If a comparator initiator starts on dapagliflozin, that patient will switch to the dapagliflozin group. See Appendix 2 for a description of how switching study exposure groups and the assignment of person-time will be handled under various AD initiation scenarios. If any patient discontinues the index AD and starts on another AD that qualifies as a study exposure, some person-time at risk will pertain both to the tailing off of the first AD and to the startup of the second AD. Person-time during follow-up with combined exposure to dapagliflozin and a comparator AD will be grouped into a combined exposure category and analyzed separately from the dapagliflozin-only exposed person-time.

3.3 Data Source/Data Collection Process

This study requires data sources that longitudinally capture inpatient and outpatient diagnoses and procedures, capture prescription and dispensing information, and allow validation of data source listings of ALI. This study will be conducted using three sources of longitudinal data: CPRD in the UK and the HIRD^{SM} and Medicare databases in the US. A summary of the available data fields and other characteristics of the data in each data source is provided in Appendix 3.

3.3.1 Clinical Practice Research Datalink – UK

In the UK, GPs are the gatekeepers for the health care of the patients registered with them. In practices that contribute information to the CPRD, common software is used to create the electronic medical record that GPs keep for the clinical follow-up of their patients (http://www.cprd.com/intro.asp). The CPRD includes information on patient demographics, lifestyle factors (admittedly not complete for all patients), outpatient diagnoses (documented to be complete), additional clinical information (completeness dependent upon the type of information), referrals, prescriptions issued by GPs (complete), and other information that GPs consider important for clinical care (e.g., results from complementary exams, procedures, hospitalizations, and reports from specialists—these generally require validation). In addition, validation of outcomes can be implemented by surveying the GP. Approximately 65% of the English practices contributing to the CPRD have consented to have their patient information linked, by a trusted third party, to other health care data sets (including Hospital Episode Statistics [HES] data) via the patient's National Health Service number, sex, date of birth, and postal code. English practices represent approximately 75% of all practices contributing to the CPRD; therefore, approximately half of the total CPRD practices can be linked to HES data if needed. Previous experience with linkage of CPRD and HES data suggests that the

hospitalization dates listed in the CPRD may be up to 120 days later than the actual hospital admission date. The CPRD has information on 5.1 million individuals (active contributors), which represents approximately 8% of the UK population.

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 (Gelfand et al., 2005). In contrast, the pharmaceutical exposures and comorbidities are expected to be based on outpatient prescriptions and to be complete. In particular, the diagnosis of T2DM, after excluding individuals with type 1 diabetes diagnoses, has been validated in the CPRD and found to have a high positive predictive value: 98.6% (95% CI, 92.2%-99.7%, calculated using Episheet [Rothman, 2012]) (Van Staa and Abenhaim, 1994).

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 (Rubino et al., 2007). The prevalence of diabetes has increased in the UK from 2.8% in 1996 to 4.3% in 2005. The incidence of diabetes 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 AD use has occurred, predominantly from sulfonylureas to metformin (González et al., 2009).

3.3.2 HealthCore Integrated Research Database – US

HealthCore, Inc., (hereafter, HealthCore) is a wholly owned subsidiary of Anthem, Inc., which is the largest health benefits company in the US in terms of medical membership. Anthem is an independent licensee of the Blue Cross and Blue Shield Association and serves its members as the Blue Cross licensee in 14 states and through UniCare. Anthem is also the parent of Health Management Corporation, a preventive health and disease management company.

The HIRD^{SM} contains fully adjudicated paid claims from the largest commercially insured population in the US, with dates of service for all noncapitated ambulatory, emergency department, inpatient, and outpatient encounters (including administrative claims for laboratory tests) for members with eligibility at the time of service. It also includes claims for outpatient dispensings of prescription pharmaceuticals from pharmacies. The full HIRD^{SM} dates back to 01 January 2006. The majority of data can be accessed from that time period through the most recent update. Data are updated monthly, with an approximate 3-month time lag for up to 95% full capture of paid medical claims. The lag for pharmacy data is shorter, with approximately 99% paid within 30 days. As of January 2014, the HIRD^{SM} contained claims information for

approximately 35.8 million lives available for research. In addition, HealthCore has the ability to redact or abstract inpatient and outpatient medical records for the health plan members represented in the HIRD^{SM}, identify and contact providers and members for survey research through vendor relationships, and link data to national vital records. The HIRD^{SM} enables rapid access to US population-based health care data resources representing all major geographic regions and health care settings and varied clinical indications that permit long-term longitudinal patient follow-up. The specific geographic regions represented in the HIRD^{SM} are the Northeast, Mid-Atlantic, Southeast, Midwest, Central, and West regions of the US. The HIRD^{SM} has been used as a data source in multiple studies related to safety outcomes and validation.

Health plans contributing data to the HIRD^{SM} include several different lines of business such as health maintenance organizations, point-of-service plans, preferred provider organizations, and indemnity plans.

Data on patient enrollment, medical care (professional and facility claims), outpatient prescription drug use, laboratory test result data, and health care utilization may be tracked for patients in the database. Diagnoses and procedures are identified by ICD-9-CM or ICD-10-CM,[1] 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 (NDCs), 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 in the database.

Patients aged 65 years or older will be excluded from this data source to avoid any duplication with Medicare data. In addition, patients in the HIRD^{SM} will be censored during follow-up the day before their 65th birthday.

3.3.3 Medicare – US

Medicare is a federally sponsored health insurance program in the US that offers health coverage to 47 million people, including 39 million people aged 65 years or older and 8 million nonelderly people with a permanent disability (Cubanski et al., 2010). Most adults become eligible for Medicare when they reach 65 years of age, although younger adults can qualify if they are permanently disabled. Medicare beneficiaries make up approximately 15% of the total US population and include more than 98% of the US population aged 65 years or older (Research

[1] ICD-9-CM = *International Classification of Diseases, 9th Revision, Clinical Modification;*
ICD-10-CM = *International Classification of Diseases, 10th Revision, Clinical Modification.*

Data Assistance Center, 2013). In 2012, 11.2 million people aged 65 years or older had been diagnosed with diabetes (Centers for Disease Control and Prevention, 2014), and most would have Medicare coverage. Therefore, Medicare data are particularly useful for the current study.

Medicare consists of Part A, which is hospitalization insurance; Part B, which covers physician services and outpatient care; and Part D, which is outpatient prescription drug coverage. Parts B and D are optional, and enrollees must pay a monthly premium for this coverage. Part D coverage has been available since 2006 and is purchased by beneficiaries through private insurance companies approved by Medicare. As of 2010, about 60% of Medicare beneficiaries were enrolled in Part D.

Analytic files on claims contain information collected by Medicare to pay for health care services provided to Medicare beneficiaries. Data are available for each claim type: institutional (inpatient, outpatient, skilled nursing facility, hospice, or home health agency) and noninstitutional (physician and durable medical equipment providers). Similar to the HIRD^{SM}, diagnoses and procedures are identified by ICD-9-CM or ICD-10-CM, CPT, and HCPCS codes for both outpatient visits and inpatient stays. Additionally, the Part D data claims file contains information on prescription drug fills, including product codes (NDCs), quantity dispensed, and days' supply.

There is currently a 2-year lag in accessing Medicare Part D data. Generally, Medicare releases Part D data each January. Therefore, if the first interim analysis occurs in January 2018, Medicare data would be available through the end of 2015.

3.4 Definitions of Study Variables

3.4.1 Outcomes/Endpoint Variables

In each data source, electronically available diagnosis data will be utilized to screen for potential cases of ALI. After all potential cases are identified, a validation effort will be implemented to obtain more detail to confirm cases. Depending upon the total number of potential cases in each data source, the validation process may be initiated for all potential cases. If the number of potential cases is relatively large, e.g., more than 125 potential cases in any single data source, validation of a potential algorithm will be implemented for a sample of potential cases. Resampling and validation may be required if the algorithm requires modification (for criteria, see Section 3.4.1.2, Case Validation via Medical Record Review).

3.4.1.1 Electronic Case Identification

ALI will be identified as follows:

- Hospitalization for ALI (see Read and ICD-9-CM codes in Appendix 1; ICD-10[1] codes will be included in the statistical analysis plan)

  OR
- Specialist visit for ALI (CPRD only).

Potential cases of ALI will be identified as follows:

- CPRD: via GP mentions of hospitalizations associated with the required diagnoses or referral to a specialist. Due to limitations of the CPRD, we will also link to the HES data to identify hospitalizations with relevant diagnoses not listed in the CPRD (see Appendix 1 for Read codes) and to confirm or ascertain the correct hospitalization date for any relevant hospitalization listed in the CPRD. Therefore, potential cases will be those with hospitalizations for the included 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 55% of study patients (Gallagher et al., 2011).
- HIRD^{SM} and Medicare: via claims for hospitalization stays associated with the required diagnoses (see Appendix 1 for ICD-9-CM codes).

3.4.1.2 Case Validation via Medical Record Review

For each data source, the validation process will be detailed in the validation plan, to be developed in the future. In summary, the medical records for up to 125 potential cases will be requested and abstracted to assess whether algorithms based on codes accurately identify the event of interest (acute liver injury). The medical record data will be reviewed by endpoint adjudicators with relevant clinical expertise. We aim to identify an electronic algorithm of codes that results in a positive predictive value greater than 80% for medical record–confirmed outcomes (See Section 3.4.1.3 for the specific case definition for confirmed cases).

If the lower bound of the 95% CI for the positive predictive value of the coding algorithm for ALI is found to be below 80% in one or more of the data sources, we will update the coding algorithm for that outcome to achieve a higher positive predictive value; if necessary, we will draw another sample of medical records to validate the revised algorithm. The algorithm development process will be described in the validation plan.

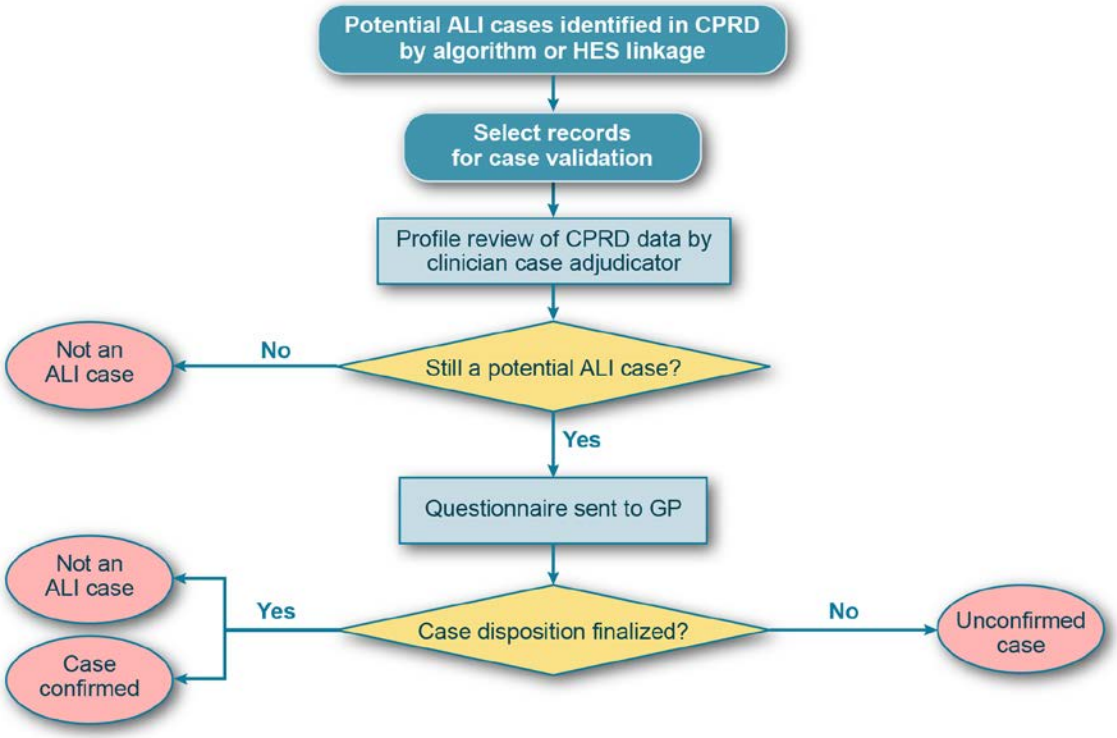
CPRD

Presence of the clinical criteria (see Section 3.4.1.3, Case Definition via Medical Record Review) in each potential case of ALI (up to 125 cases) will be evaluated further through review of the patient profiles generated from information recorded in the CPRD. For patients for whom

[1] ICD-10 = *International Statistical Classification of Diseases and Related Health Problems, 10th Revision.*

available clinical data in the CPRD cannot rule out ALI, including those for whom no HES linkage is available, we will further attempt to validate the ALI diagnosis by collecting the relevant clinical information through a questionnaire to the GP. The process for case validation is shown in Figure 1.

Figure 1: Case Validation in the Clinical Practice Research Datalink

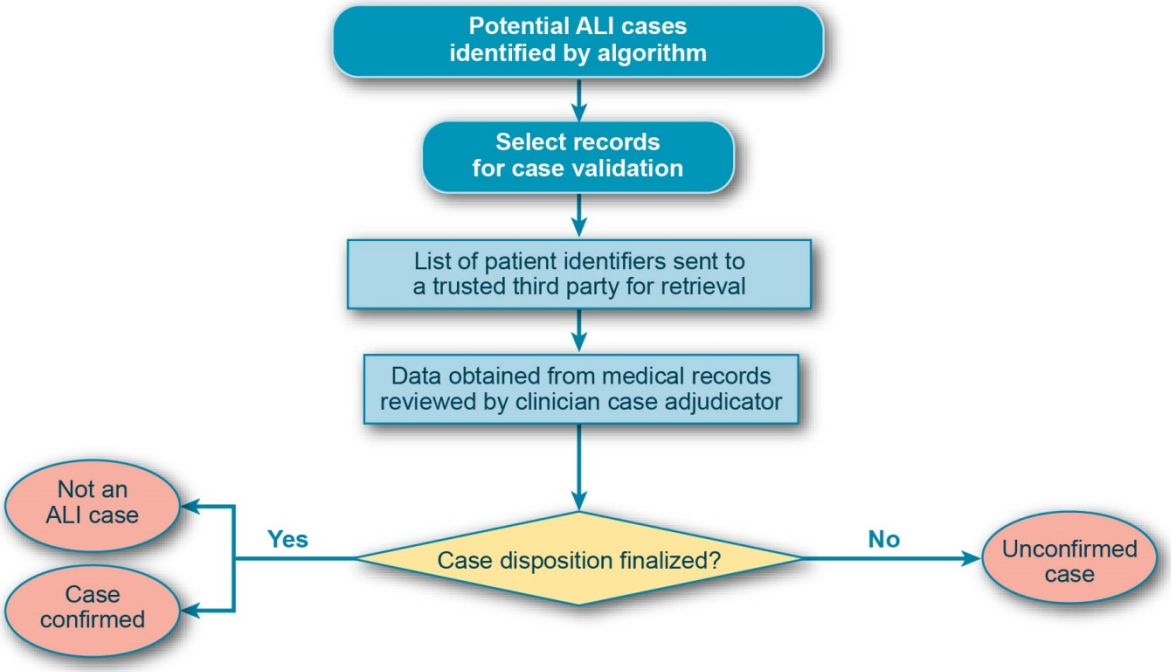


ALI = acute liver injury; CPRD = Clinical Practice Research Datalink; GP = general practitioner; HES = Hospital Episode Statistics.

Medicare/ HIRD[SM]

The process for case validation in Medicare data and the HIRD[SM] is shown in Figure 2. Patient identifiers (name, date of birth, and social security number) are included in the Medicare files and can be used for further data abstraction. Additionally, individual and institutional providers have a unique identification number that is used to identify specific providers. Patients meeting the outcome definition of ALI based on diagnosis codes associated with hospital claims will be identified, and information about the individual provider for these patients will be collected. Relevant potential cases will be identified from each cohort, and a list will be sent to separate individual trusted third parties for Medicare and for HIRD[SM]. Each third party will contact the individual provider to obtain to obtain the required information from relevant medical records. For patients in the Medicare data, details from the medical record will be obtained by record abstraction. For patients in the HIRD[SM], redacted copies of the medical record for the hospitalization of interest will be obtained. Structured forms for abstraction in Medicare and for guiding the copying of relevant records for HIRD[SM] patients will be used to collect the relevant information to confirm the outcome (forms will be provided as part of the study report). Final confirmation of cases will be conducted independently by endpoint adjudicators who will be blinded to exposure to medications and will be identified by RTI Health Solutions (RTI-HS).

Figure 2: Case Validation in Medicare and HIRD[SM] Data



ALI = acute liver injury. HIRD[SM] = HealthCore Integrated Research Database.

3.4.1.3 Case Definition via Medical Record Review

The diagnostic criteria for ALI recommended in current guidelines and publications are summarized in Appendix 6. The diagnostic criteria for ALI were first defined in 1989 in an international consensus meeting involving a panel of 12 European and American experts (Bénichou, 1990). The diagnostic criteria were revised in 2001 before a conference sponsored by the US Food and Drug Administration (FDA) Center for Drug Evaluation and Research, the Pharmaceutical Research and Manufacturers of America, and the American Association for the Study of Liver Diseases (AASLD) (FDA Working Group, 2000; Navarro and Senior, 2006). ALI was defined as an ALT level of more than 3 times the ULN, or an AP level of more than 2 times the ULN, or a TB level of more than 2 times the ULN associated with any elevation of the ALT or AP (Navarro and Senior, 2006). The diagnostic criteria for ALI used in observational studies in patients with diabetes are summarized in Appendix 6, Table 6-2. One study evaluated ALI (Huerta et al., 2002), using the criteria proposed by an international consensus meeting (Bénichou, 1990).

For this study, confirmed cases of ALI will be a hospitalization for ALI meeting the criteria proposed by the FDA Working Group (FDA Working Group, 2000; Navarro and Senior, 2006). Detailed definitions of the study endpoint follow).

An ALI case will be any patient from the study population meeting *all* the following criteria:

- Recorded hospitalization for ALI with the following increases of liver enzymes:
  - Elevation of ALT greater than 3 times the ULN, OR
  - Any increase of ALT and AP and an increase of TB greater than 2 times the ULN, OR
  - Elevation of AP greater than 2 times the ULN
- Elevation of liver enzymes detailed above detected within 26 weeks before the date of hospitalization or within the first 48 hours of hospitalization
- No evidence of previous chronic liver disease or any of the exclusion criteria detailed in Section 3.2.2.

3.4.2 Exposure/Independent Variables of Interest

For this study, we plan to identify the study medications of interest among eligible patients from GP prescriptions in the CPRD and from outpatient pharmacy claims in the HIRD^{SM} and Medicare data. New use of dapagliflozin will be defined as the date of first dapagliflozin prescription (CPRD) or dispensing (HIRD^{SM} and Medicare data) in the data source. New use of an AD in a class other than SGLT2 inhibitor, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy will be defined as the date of first prescription or dispensing for these medications in the data source during the study period. As illustrated in Appendix 2, more than one treatment episode within the person-time of a patient can be selected as comparator

exposures if a qualifying drug is initiated at a point in time after the first eligible treatment episode ends and is a different drug than the first. Potential comparators are eligible to enter the pool of patients from which comparators can be selected multiple times (i.e., if they qualify with drug A and then later switch to drug B, which also qualifies as a new comparator drug, they can enter both times). Follow-up will not be censored with addition of any other antidiabetic drug (other than a non-dapagliflozin SGLT2 inhibitor) during the index risk window.

The index medication is the dapagliflozin or comparator AD exposure that qualifies the patient to enter the study. When the episodes (time windows at risk) of study medications overlap, only one of medications will be used as the index medication. If the overlapping episodes include dapagliflozin, then dapagliflozin will be used as the index medication. If the overlapping episodes do not include dapagliflozin, then the first episode will be used as the index medication. If the episodes of two or more comparator AD episodes start at the same date, we will randomly choose one as the index medication, and the exposure will be classified as index combined therapy.

Initiators of the ADs listed in Appendix 4 will be included in the comparison group.

3.4.3 Other Covariates/Control Variables

Information to characterize the cohorts at the time of study drug initiation will be collected from the period ending on the index date, using all available information in the data source. Because all patients in the study are required to have at least 180 days of data before the index date, there will be a minimum of 180 days of data from which to evaluate covariate values. For some patients, more information will be available, and all information will be considered to reduce misclassification of covariate information. During development of propensity scores, use of indicator variables for the duration of lookback time and timing of information on key covariates will address possible differential availability of information on covariates by exposure group, as well as control for associations that vary by time of recorded information. Exclusion diagnoses will be identified based on recorded GP diagnoses (CPRD) or, in the US (HIRD^{SM} and Medicare), claims diagnoses during the lookback period. Data on likely predictors of ALI (Table 1) will be identified for all patients prior to and including the index date. Although severity of T2DM may be a predictor for ALI, indicators for severity, e.g., glycated hemoglobin (HbA1c) value, will be available in the CPRD and for approximately 30% of patients in the HIRD^{SM}. All of these variables and additional variables that could potentially differ by exposure (e.g., history of urinary tract infections; see Appendix 7) will be included in a logistic regression model that will be used to generate propensity scores for the final analysis. The propensity scores will quantify the probability of receiving dapagliflozin at the time of the index date.

Table 1: Variables of Interest to be Collected for Propensity Score Development

Demographic or Lifestyle	
Age	Smoking history (in CPRD only)
Sex	Alcohol consumption (in CPRD only)
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 (in CPRD only)
Body mass index > 30 (in CPRD only) or obesity surgery	
Geographic region of residence	

Medications		
Drugs with a known association with liver injury[a]		
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
Duloxetine	Phenobarbital	
Enalapril	Phenothiazines	
Erythromycins		
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)		Opioids

Medical Comorbidities[b]		
Ischemic heart disease	Chronic obstructive pulmonary disease, emphysema, respiratory insufficiency	Renal insufficiency
Hypertensive disease	Diffuse diseases of connective tissue	All malignancies other than non-melanoma skin cancer
Peripheral vascular disease	Rheumatoid arthritis	Dementia
Other cardiovascular disease	Osteoarthrosis	Peptic ulcer disease
Cerebrovascular disease	Polymyalgia rheumatica	Being hospitalized, especially for a serious condition that requires intensive care
Hyperlipidemia	Pancreatitis	Length of hospitalization
Autoimmune disease	Immunosuppressive diseases, such as HIV/AIDS	Pregnancy in the 180 days before and including the index date
Asthma		Hospitalization for a serious condition that required intensive care in the 180 days before the index date
Kidney and genitourinary stones		
Urinary infections, chronic or recurring		
Colon polyps		
Crohn's disease		
Ulcerative colitis		

Indicators of Diabetes Severity	
Renal insufficiency or diabetic nephropathy	Coronary heart disease
Retinopathy	Cerebrovascular disease
Peripheral neuropathy	Amputations
Peripheral vascular disease	Time since first diagnosis of type 2 diabetes mellitus, (CPRD only)

CPRD = Clinical Practice Research Datalink; HAART = highly active antiretroviral therapy; NSAIDs = nonsteroidal antiinflammatory drugs.

[a] Source: Navarro and Senior, 2006.

[b] Additional medical comorbidities that may be considered can be found in Appendix 7.

Other variables to include in the propensity score models will be indicators of health care utilization in the 180 days before but not including the index date. These variables are number of outpatient visits, number of hospitalizations, number of emergency department visits, and number of specialty care visits.

Another variable to be examined will be whether patients, at the time of the index date, "added on" dapagliflozin or new use of another AD, that is, whether prescriptions or dispensings for any AD that the patient was receiving in the 90 days before the prescription or dispensing for dapagliflozin or the newly initiated AD are continued in the 90 days after the index date (i.e., at least one more prescription is recorded). Patients will be classified as having "switched to" dapagliflozin or new use of another AD if there is a prescription or dispensing for the medication recorded in the 90 days before the index date and no prescriptions or dispensings for that medication are recorded after the index date. Days' supply will be used to determine the calculated end of previous AD therapy or will be assumed to be 30 days if missing.

To explore the impact of addition of ADs other than the index exposure, we will determine the rates of new AD use during the person-time of the index exposure episode among dapagliflozin initiators and comparator AD initiators (see Section 4.1.3). The additional drug could be any AD

that was not a part of the initial treatment episode, including insulin. To be considered exposed to an added AD during follow-up (i.e., added after the index date), one prescription or dispensing for the medication must be added to the regimen during follow-up (recorded as yes/no for whether an AD was prescribed). Variables other than the addition of other ADs, both fixed and time-dependent, that could represent possible confounders and effect modifiers will be identified, if deemed necessary and feasible, and classified during follow-up time. The degree to which we can pursue analyses of these variables is contingent on the number of events within each outcome.

The approaches to handling concomitant ADs in the analyses are summarized in Table 2.

Table 2: Approaches to Handling Concomitant Antidiabetic Drugs

Timing and Type of Antidiabetic Drug Dispensing or Prescription	Analysis Approach
At index date, any AD taken during baseline that is not the new prescription and is not eligible to be a study exposure	Include in propensity score
At index date, any AD taken during baseline that is not the new prescription and is eligible to be a study exposure	Include in propensity score
Designation whether the AD initiated at the index date is an add-on to current medication or a switch to a different medication	Include in propensity score
Insulin at the index date	Conduct descriptive and comparative analyses separately by insulin use at index date (Yes or No)
ADs that have been used in the past added during follow-up	Consider use of time-varying covariate methods
Potential bias from new use, by exposure status, of a drug class	Sensitivity analysis excluding all individuals not new to the index drug class—e.g., a new user of saxagliptin (comparator group) will be included in the sensitivity analysis only if he/she was not treated with any DPP-4 inhibitor during the baseline time period
SGLT2 inhibitor other than dapagliflozin added during follow-up	End follow-up and censor

AD = antidiabetic drug; SGLT2 = sodium-glucose cotransporter 2.

4 STATISTICAL ANALYSIS

4.1 Statistical Analysis Methods

In the following analyses, descriptive analyses (i.e., secondary objective #1) will be done before other analyses. Specifics of variable definitions will be described in the statistical analysis plan, to be developed after finalization of the protocol.

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 and Medicare data will be carried out in accordance with RTI-HS standard operating procedures. Routine procedures include checking electronic files, maintaining security and data confidentiality, following the statistical analysis plan, and performing quality-control checks of all programs. Data extraction programming for creating the study population from the HIRD^{SM} and creating the analytic file will be performed in accordance with HealthCore Programming Standards. The HealthCore Programming Standards are a set of documents describing data extraction and data development methods that are referenced in HealthCore standard operating procedures.

4.1.1 Propensity Score Approach

Demographic, medical, and clinical factors that may be associated with the decision to begin therapy with any particular AD may also be associated with the outcome. However, the number of outcomes will likely be small, limiting the number of variables that could be included in a regression model that predicts these outcomes (Cepeda et al., 2003). To address this difficulty, we will summarize the set of confounding variables, using a propensity score, based on the literature on potential confounding variables associated with the outcome. The propensity score is the predicted probability of being assigned to a particular treatment conditional on a set of observed covariates (Braitman and Rosenbaum, 2002; D'Agostino et al., 1998; Perkins et al., 2000).

Within each data source, propensity scores will be estimated by conducting logistic regression modeling and incorporating measured potential predictors of therapy as independent variables and exposure group status (dapagliflozin group vs. comparator group) as the outcome. The variables listed in Table 1 and Appendix 7, if available in the data source, will be assessed on the index date or, for chronic conditions, before the exposure index date, and will be considered for inclusion in the propensity score model.

If matching is performed, the propensity score models will be fitted after matching. The propensity score models will be developed for patients within each data source and will adjust for calendar year of the index date as a continuous variable to allow for changing prescription patterns for dapagliflozin from the time it is first available through the date of receipt of the data. As prescription patterns change, the confounding influence of the determinants of the prescription may also change. During development of propensity scores, indicator variables for the duration of lookback time and timing of information on key covariates will be included. Descriptive analyses will be stratified by data source and propensity score deciles and will be conducted at each scheduled data cut.

For data sources for which it is not feasible to obtain data on all eligible comparator cohort members in the data source, propensity scores will be obtained after selection of the comparator cohort through matching. For all analyses, we will exclude patients who have estimated propensity scores outside the range that is common to both exposed and comparator cohorts. This process is known as "trimming." Trimming occurs at both ends of the propensity score scale. At the bottom end, we will exclude all patients, exposed or unexposed, who have a propensity score below the 2.5 percentile value of the distribution of scores among the exposed group. At the upper end, we will exclude all patients, exposed and unexposed, with scores greater than the 97.5 percentile of scores among the comparator patients. This trimming will be performed separately for each index year–specific set of propensity scores.

Within each propensity score analysis, after trimming, the data will be stratified into deciles of propensity scores based on the distribution among dapagliflozin initiators. Within each of these 10 propensity score–based strata, we will investigate the extent to which covariates are balanced between the two treatment groups by use of the absolute standardized difference to assess the balance of measured baseline covariates between the dapagliflozin group and the comparator AD group before and after propensity stratification (Austin, 2009). Any imbalance will be addressed by either revising the propensity score model or by making adjustments in the final outcome model (Braitman and Rosenbaum, 2002; Perkins et al., 2000). We will report the number of patients trimmed from the analysis because of nonoverlap of propensity scores. If using deciles to create strata results in strata that are too small, it may be necessary to combine adjacent deciles.

4.1.2 Primary Objective

The incidence rate of ALI among dapagliflozin initiators and among the comparator group will be estimated. The number of new hospitalized cases of ALI during follow-up will be determined using the validated algorithm. Person-time for each patient will be determined as the time between the date of first prescription or dispensing for either dapagliflozin or comparator ADs and end of time at risk. The total person-time of observation among individuals at risk will then be calculated. The incidence rate of ALI will be estimated by insulin use at the index date in each cohort. Within each data source, crude and adjusted rates will be calculated as the number of new cases of ALI during the observation period divided by the total person-time of observation among individuals at risk. Incidence rates will be reported as point estimates (in cases per 1,000 person-years) and 95% CIs.

The following incidences and comparisons will be generated:

- Crude incidence, by categories of insulin use at the index date, among dapagliflozin initiators versus among the entire comparator group.

- Summary IRRs after adjusting by propensity score decile, calendar year, and data source, estimated by categories of insulin use at the index date, among dapagliflozin initiators versus the entire comparator group.

The adjusted IRRs will be the primary endpoint. In CPRD and Medicare data, crude IRRs will facilitate comparison with the adjusted ratios to provide an indication of the degree of confounding. Follow-up time that includes exposure to more than one index AD, e.g., first 30 days after patients switch from dapagliflozin exposure group and selection into the comparator group, will be assigned to a separate category of multiple exposures.

Adjusted incidence rates and IRRs will be calculated by standardizing to the dapagliflozin person-time distribution—weighting the incidence or IRR in each stratum (defined by propensity score decile, calendar year, and data source) by the amount of dapagliflozin person-time within the stratum. More details on the analysis methods will be included in the statistical analysis plan.

4.1.3 Secondary Objectives

4.1.3.1 Secondary Objective #1—Compare Baseline Patient Characteristics

Descriptive statistics, by insulin use at the index date, will be generated for each data source to compare baseline characteristics between dapagliflozin initiators and comparator AD initiators. Categorical variables will be summarized by frequencies and proportions, and continuous variables will be summarized by means and standard deviations or medians and interquartile ranges. The following variables will be characterized:

- Age stratified by sex
- Body mass index (CPRD only)
- Duration of lookback time prior to index date
- Prescription or dispensing of each specific non-dapagliflozin AD at index date
- Switch versus add-on initiation of study exposure (dapagliflozin or other AD) at index date
- History of diabetic nephropathy diagnosis
- History of cardiovascular disease diagnosis
- History of hypertension diagnosis
- HbA1c at index date (most recent measurement on or before the index date) (CPRD only)
- History of hyperlipidemia diagnosis or treatment
- History of chronic obstructive pulmonary disease or asthma diagnosis
- History of malignancy diagnosis
- History of peripheral vascular disease diagnosis
- History of vasculitis, scleroderma, or lupus diagnosis
- History of trauma diagnosis in last 6 months

Results of the descriptive analyses will be used to inform subsequent analyses.

4.1.3.2 Secondary Objective #2—Identify Confounders and Effect Modifiers

The degree to which we can pursue analyses of confounders and effect modifiers 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 relative incidence rates and rate ratios for ALI. Some potential stratification variables will include index year, HbA1c levels (when available in the CPRD and HIRD^{SM}), whether patients were "added on" or "switched to" dapagliflozin or other ADs, each specific concomitant AD medication class added during follow-up, and prescription for a medication associated with ALI during follow-up. Also, variables for which close balance was not achieved within propensity score strata may be further examined.

We will calculate the following statistics:

- Incidence rates by exposure category stratified by potential effect modifiers or confounders and data source
- Incidence rate ratios stratified by potential effect modifiers or confounders

To explore the impact of differential addition of ADs or factors associated with the need for an additional AD by exposure, we will compare rates of new AD use among dapagliflozin initiators and comparator AD initiators. If necessary, further adjustment for therapies added during follow-up will be performed via time-dependent Cox proportional hazards regression. Techniques to pool the data will be applied to combine IRR estimates across data sources if appropriate (see Section 4.1.7).

4.1.4 Imputation of Missing Values

We expect that relatively few key variables will have notable missing values. Variables such as smoking and alcohol consumption levels, body mass index, and HbA1c will be unavailable or have high levels of missing values. The pharmaceutical exposures and comorbidities are expected to be based on outpatient prescriptions and dispensings and to be complete. If there are considerable missing data for lifestyle covariates, multiple imputation will be used to fill in missing values for the propensity score creation and multivariable analyses. The decision to use multiple imputation will depend on the strength of the association between the variable and treatment and the extent of missing data. Based on information from the observations with nonmissing values, we will impute five simulated versions of the data set. The imputed data sets will be used for creation of propensity scores and in the multivariable analyses, with the results being combined appropriately to generate final point estimates and CIs. In theory, this should

give point estimates with equal or less bias than those that would be obtained if we had limited the sample to those with complete data, and it should give greater precision because of the larger number of patients that will be included using this method as opposed to restricting the analysis to observations with complete data. The specific approach will be detailed in the statistical analysis plan.

We have selected the multiple imputation approach because existing methods for imputation penalize the standard errors when imputing data and multiple imputation allows for better bias correction than most alternatives, including the complete-case approach, for many, although not all, applications. 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.

4.1.5 Exploratory Objectives

Not applicable.

4.1.6 Sensitivity Analyses

The following sensitivity analyses will be conducted:

- Estimate IRRs for dapagliflozin compared with comparator group by applying the risk window extended from 30 to 90 days in the exposed follow-up time for dapagliflozin initiators and for comparator initiators.
- Estimate the crude incidence rate without stratification by insulin categories at the index date, among dapagliflozin initiators versus among the entire comparator group.
- Estimate summary IRRs after adjusting for propensity score decile, calendar year, and data source, without stratification by insulin categories at the index date, among dapagliflozin initiators versus the entire comparator group.
- Estimate summary IRRs after adjusting for propensity score decile, calendar year, and data source, without stratification by insulin categories at the index date, among dapagliflozin initiators versus initiators of the most commonly used classes of ADs (each class separately).
- Assess the effect of unmeasured confounders, one at a time, on the association between dapagliflozin use and ALI by assuming a plausible range of values for the prevalences of each of the unmeasured confounders among the dapagliflozin group and the comparator group and risk ratio for the association between each of the unmeasured confounders and the outcome of interest (Lash et al., 2009). For example, race or ethnicity and use of specific over-the-counter medications can be risk factors for ALI and are likely to be unmeasured. Based on the available literature, we can assume a reasonable range of prevalence values for a given unmeasured confounder and the outcome of interest. However, at this time, there is no reason to expect differential distribution of race or over-the-counter medication use among dapagliflozin initiators versus comparator initiators.

- Estimate summary IRRs after adjusting for propensity score, calendar year, and data source after excluding individuals not new to the index AD drug class.

4.1.7 Pooled Analysis

The results of two or more data sources with similar results for the primary outcome (i.e., plausibly differing only from sampling variability) will be pooled. The pooled analysis will be designed to estimate the effect of the exposure while controlling for confounding using the data stratified on propensity scores. The data source will be retained as a stratification variable, so the effect within each data source can be estimated. Mantel-Haenszel techniques will be used to pool the data from each data source and calculate overall adjusted incidence rate ratios. This analysis is designed to estimate the effect of the exposure while controlling for confounding by using the data source–specific propensity score stratification.

If residual confounding is suspected in any of the data sources (for example, due to lack of information on one or more confounding variables), external adjustment can be used to assess the impact of unidentified confounders and to reduce the amount of such confounding (see Section 4.1.6) (Lash et al., 2009).

4.2 Power/Sample Size

The observed study size will depend upon the market uptake of dapagliflozin in the US and UK. We estimate a total of 91,927 person-years of follow-up among new users of dapagliflozin across all three data sources. The derivation of this total follows.

Currently, we estimate that in the CPRD, approximately 3,600 person-years of follow-up will be available among all new users of dapagliflozin during the 5 years of the study; including approximately 2,900 person-years among those not on insulin at the index date and 700 person-years among those on insulin at the index date. These estimates are based on the following assumptions: number of patients aged 18 years and older with a newly prescribed AD in the CPRD per year is 23,970 (CPRD data as of 31 December 2011); the proportion of new users starting dapagliflozin among patients who meet inclusion criteria will be 1% during year 1, 2% during year 2, 3% during year 3, 4% during year 4, and 5% during year 5 of the study; on average, each new user will contribute 12 months of person-time; and approximately 20% of new dapagliflozin users will be on insulin at the index date (derived from Hall et al., 2012).

In the US data sources, we estimate that there will be approximately 88,327 person-years of follow-up available among all new users of dapagliflozin (HIRD^{SM}: 42,473 person-years; Medicare: 45,854 person-years). This exposure would include approximately 70,662 person-years among those not on insulin at the index date and 17,665 person-years among those on insulin at the index date. These estimates are based on the following assumptions: (1) the age distribution among oral AD users is 61.9% aged 18-64 years and 37.7% aged 65 years or older

(Boccuzzi et al., 2001), (2) 34.15% coverage rate for Medicare Part D among Medicare beneficiaries, (3) the HIRD^{SM} covers 6% of the US population aged 64 years or younger, (4) each new dapagliflozin user will contribute 12 months of exposed person-time, and (5) approximately 20% of new dapagliflozin users will be on insulin at the index date (Hall et al., 2012).

To provide precision estimates in relation to the total projected study size, we first estimated the background ALI incidence rates based on the literature. The incidence of ALI associated with referral or hospitalization in patients with diabetes using oral hypoglycemic agents in the UK has been estimated to be 22 cases per 100,000 person-years (Huerta et al., 2002). If this is the rate in in the total sample cohort, we would expect to observe approximately 101 events overall. Table 3 shows the expected number of ALI events among all cohort members by study data source.

Table 3: Estimated Number of Events for Acute Liver Injury Among All Cohort Members

	HIRD^{SM}	Medicare	CPRD	Total
Total sample cohort[a]	212,366	229,270	18,000	459,636
Average length of follow-up	1 year	1 year	1 year	1 year
Exposed person-years	42,473	45,854	3,600	91,927
Rate of acute liver injury (per 100,000 person-years)	22	22	22	
Estimated number of events	47	50	4	101

CPRD = Clinical Practice Research Datalink; HIRD^{SM} = HealthCore Integrated Research Database.
[a] Assumes 34.15% coverage rate for Medicare enrollment criteria and 6% population coverage of HIRD^{SM}.

Using the expected background incidence rate of 22 per 100,000 person-years, we estimated the precision of the study under various scenarios. Table 4 shows the probability that the upper 95% confidence limit around the observed IRR will be less than the specified IRRs for dapagliflozin users not on insulin, assuming that the true IRR is 1.0 and a 1:4 dapagliflozin:comparator person-year ratio. For example, a study size of 73,000 person-years of follow-up among new users of dapagliflozin not on insulin at the index date will provide 85% probability that the upper 95% confidence limit of the IRR will be less than 2.3.

Table 4: Probability That the Upper 95% Confidence Limit of the IRR is Below the Specified Value, Assuming IRR in Population = 1.0

Dapagliflozin-Exposed Person-years	Upper 95% Confidence Limit of IRR for Dapagliflozin Versus Comparator		
	1.5	1.8	2.0
50,000	0.22	0.41	0.54
60,000	0.26	0.48	0.62
73,000	0.31	0.56	0.70
85,000	0.35	0.62	0.76
90,000	0.36	0.65	0.79

IRR = incidence rate ratio.

Note: Assuming 22 per 100,000 person-years is the rate of ALI associated with hospitalization among patients not exposed to dapagliflozin, a 1:4 dapagliflozin:comparator person-year ratio, and population incidence rate ratio (IRR) = 1.0. This table was calculated using Episheet (Rothman, 2012).

In studies of medication safety, it is desirable to reduce the uncertainty around the relative risks associated with treatments of interest. However, the desired precision is often difficult to obtain because of the low frequency of events and the low numbers of patients taking the medication. Nevertheless, useful results can be obtained. For example, results from a smaller-than-ideal study size can achieve the following accomplishments:

- Reduce the uncertainty around the frequency of adverse events of interest compared with results from spontaneous reporting
- Detect a large relative risk associated with dapagliflozin use, if it exists
- Establish the absolute risk of ALI associated with dapagliflozin, a measure of the potential public health impact related to dapagliflozin among patients with diabetes

4.3 Milestones

Descriptive and, when appropriate, comparative analyses are planned for this study. The descriptive analyses, which include comparison of baseline characteristics for each cohort and total case counts will be performed after dapagliflozin has been on the market in the US for approximately 30 months. Execution of an interim comparative analysis will depend upon a sufficient number of dapagliflozin users and events. The final analysis will be conducted after dapagliflozin has been available for 60 months. The interim comparative analysis will be performed if at least two outcome events are observed in the entire study cohort (dapagliflozin and comparator AD cohorts combined). The proposed timeline for analyses is shown in Table 5.

Table 5: Milestones

Report	Time From Dapagliflozin Availability to Patients in the US to Data Cut (Anticipated Month/Year)[a]
Interim descriptive analysis	24 months (January 2016) Includes only CPRD and HIRD^{SM}
Interim comparative analysis	48 months (January 2018) Includes all data sources
Final analysis	60 months (January 2019) Includes all data sources

CPRD = Clinical Practice Research Datalink; HIRD^{SM} = HealthCore Integrated Research Database; US = United States of America.

[a] Due to database lags, which are typically 4-6 months in the HIRD^{SM} and CPRD, the 24-month report will likely include data from the first 18 months of dapagliflozin use in the HIRD^{SM} and CPRD, the 48-month report will include data through the first 42 months of dapagliflozin use in the HIRD^{SM} and CPRD and 18 months of dapagliflozin use in Medicare, and the 60-month report will include data through the first 54 months of dapagliflozin use in the CPRD and HIRD^{SM} and 30 months of dapagliflozin use in Medicare.

5 STUDY LIMITATIONS/STRENGTHS

5.1 Confounding

All potential confounding variables for which there are data will be controlled to the extent possible, primarily through the use of propensity scores. Potential unidentified confounding conditions are those that cannot be identified in this manner and that are differentially distributed between the exposure groups, related to the outcome, and largely uncorrelated with the measurable characteristics.

In the CPRD, the lack of specialist prescriptions may result in lack of information on the early prescriptions for some medications, before prescribing is transferred to the GP. Some variables, such as use of over-the-counter medication, will remain unmeasured in all data sources.

Confounding by indication (or channeling bias) is a common bias in observational pharmacoepidemiology studies whereby the indication for therapy may be associated with both treatment and outcome. Since patients who receive a particular drug therapy typically have more severe disease or a perceived higher risk (due to self-selection or physician preference) compared with patients not on the medication, selection of treatment can be confounded with clinical and nonclinical patient factors that may be related to the outcomes of interest. New medications may be prescribed differentially to healthier patients who physicians believe could tolerate a product with a lesser-known safety profile, or to patients who have more severe disease, have failed previous treatment regimens, or have contraindications to other drugs (e.g., thiazolidinediones are not recommended for use in patients with heart failure). New medications may also be prescribed differentially by physicians who are “early adopters” of new technologies. As much

as possible, such considerations are taken into account by the propensity score, but some aspects may remain unmeasured, and could result in residual confounding. Specifically, dapagliflozin could be preferentially prescribed to patients with more severe diabetes or who have failed other therapies. In the CPRD, HbA1c levels are likely to be available for most patients with T2DM, so such bias may be assessed and adjusted for in the analysis. In the HIRD^{SM}, this variable is available for approximately 30% of individuals, but this variable is not available in Medicare data. Comparisons of adjusted IRRs from the CPRD with those assessed in HIRD^{SM} and Medicare data can facilitate assessment of such uncontrolled confounding in the claims-based data sources. Dapagliflozin could also be preferentially prescribed to patients with fewer risk factors for ALI. These channeling patterns could bias the hazard ratio toward or away from the null.

To assess the effect of unmeasured confounders on the association between dapagliflozin use and ALI, we will conduct sensitivity analyses to estimate the degree of possible bias that might be present by assuming a plausible range of values for those confounders.

5.2 Other Biases

Misclassification bias can result if study patients are not categorized correctly with regard to exposure or outcome. We expect minimal misclassification with respect to exposure, since this will be determined from prescribing records. However, actual adherence to instructions for taking dapagliflozin or other ADs cannot be confirmed. Further, misclassification as to whether the patient is a new user could exist if providers supplied samples of dapagliflozin or comparator ADs for varying duration to patients, at no cost, and with no record in the data source. Because of the newness of dapagliflozin, we expect little misclassification of dapagliflozin initiators. However, initiators of older ADs will be more likely to be misclassified as new users if they used the medication of interest before the patients' data were included in the data sources.

Classification of type 2 versus type 1 diabetes mellitus may also be a source of misclassification. Potential patients with evidence of type 1 diabetes mellitus (T1DM) are to be excluded. However, with the repeated health care that individuals with T1DM or T2DM require, we anticipate that accuracy of classification of diabetes type will be improved from the relative frequency of the use of these two diagnoses in individual patients.

Misclassification of the outcome will be reduced by evaluating ALI that is associated with hospitalization or visit to a specialist by validating all or a sample of potential cases identified through the review of all clinical information that can be obtained to confirm the diagnosis of ALI and liver enzyme levels. Misclassification of whether ALI resulted in a hospitalization or was hospital-acquired is possible and could affect study results. The overall clinical picture that the clinical adjudicator will form from the available information should clarify whether the ALI

developed in the hospital or not. However, by obtaining liver enzyme levels before the hospitalization or specialist visit of interest or within the first 48 hours of hospitalization, we expect to identify the timing of onset relative to the hospitalization date.

5.3 Study Size

The ability to meet the sample size projections depends upon the uptake of dapagliflozin as well as the validity of assumptions used to calculate the sample size estimates. It is currently projected that at the end of 5 years, the study will have 91,927person-years of exposure to dapagliflozin; approximately 73,542 among those not on insulin at the index date and 18,385 among those on insulin at the index date.

As discussed in the study size section, with 73,000 person-years among those taking dapagliflozin and not on insulin, we will have an 85% probability that the upper 95% confidence limit of the IRR will be less than 2.3 (assuming that the true IRR is 1.0 and a 1:4 dapagliflozin:comparator person-year ratio).

If the uptake of dapagliflozin is less than expected, we still expect to have good precision. For example, if the use of dapagliflozin among patients not taking insulin is one-quarter of that expected, then the final study size from this data source will be 18,385person-years, with a 70% probability that the upper 95% confidence limit will be less than 4.0 under the null hypothesis of no association.

5.4 Generalizability

Use of the CPRD and US claims data sources provides data entered or submitted by pharmacies, general medical practices, and US clinics and hospitals without any awareness of the hypothesis of studies that may use these data. Study results from the CPRD data will be generalizable to patients with T2DM in the UK meeting the inclusion and exclusion criteria. Results from the Medicare data will be generalizable to US patients with T2DM aged 65 years or older and not in a residential care facility. Results from HIRD^{SM} data are generalizable to patients with claims-identified T2DM among the employable US population.

6 STUDY CONDUCT

This study will be conducted in accordance with International Society for Pharmacoepidemiology (ISPE) *Guidelines for Good Pharmacoepidemiology Practices* (ISPE, 2015) and applicable regulatory requirements. As with all research at RTI International that involves human patients or data on human patients, RTI-HS will request review of the protocol

by the RTI International[1] institutional review board (IRB), and we anticipate that the IRB will agree to exemption because we will not have any patient identifiers.

6.1 Ethics Committee Review and Informed Consent

6.1.1 Ethics Committee Review

6.1.1.1 CPRD

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

6.1.1.2 HIRD^{SM}

This component of the overall study is designed as an analysis based on medical and pharmacy claims data from a large insured population in the US. There is no active enrollment or active follow-up of study patients, and no data will be collected directly from individuals.

HealthCore maintains Data Sharing Agreements and Business Associate Agreements with all covered entities who provide data to the HIRD^{SM}. HealthCore's access, use, and disclosure of protected health information (PHI) are in compliance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule (45 CFR Part 160 and Subparts A and E of Part 164). HealthCore does not access, use, or disclose identifiable PHI unless under a specific waiver of authorization (e.g., a HIPAA Waiver of Authorization from an IRB). HealthCore accesses the data in a manner that complies with federal and state laws and regulations, including those related to the privacy and security of individually identifiable health information.

As PHI must be accessed in order to acquire medical records to validate electronic case-finding algorithms, a HIPAA Waiver of Authorization will be applied for from an IRB. HealthCore will submit the protocol to a central IRB for review and approval. Approval is typically provided within 2 to 3 weeks of submission. Once IRB approval is obtained, HealthCore's vendor will proceed with the conduct of medical record acquisition. If changes to the protocol are required, HealthCore will submit an amendment to the IRB. As the IRB is independent, HealthCore cannot control the approval or whether there are conditions for the approval.

Notwithstanding receipt of approval from a central IRB, in some instances, individual institutions may require approval from their local IRB, which would require a separate protocol

[1] RTI Health Solutions is a business unit of RTI International, a not-for-profit research organization.

submission and, in some cases, additional fees. In these cases, HealthCore, RTI-HS, and AZ will need to agree whether or not to proceed with chart acquisition from these institutions.

HealthCore will provide the vendor only the minimum amount of patient information that is necessary to execute the medical record acquisition. HealthCore uses only vendors that follow federal and state laws and regulations, including but not limited to privacy and security rules such as HIPAA.

At no time during the conduct of this study will HealthCore provide patient- or provider-identifying information to RTI-HS, BMS, or AZ. Only aggregated data will be reported to RTI-HS, BMS, and AZ.

6.1.1.3 Medicare

For use of Medicare data, the CMS requires that IRB review and approval be obtained before use of Medicare data for research can be approved. This protocol will be reviewed by the RTI International IRB before applying to use Medicare data and will undergo a continuing IRB review at least once per year.

Under the Privacy Rule (under 45 CFR 164.512), CMS may disclose protected health information for research without documentation of individual authorization only if an IRB or a CMS Privacy Board has approved a waiver of research. Such a waiver must be provided to CMS.

Data requests for research identifiable data must be reviewed by the CMS Privacy Board to ensure that any study patient's privacy is protected and the need for identifiable data is justified.

6.2 Responsibilities Within the Study

The study shall be conducted as described in this approved protocol. All revisions to the protocol must be discussed with, and be prepared by AZ.

6.2.1 Sponsor Roles and Responsibilities

The sponsor, AZ, is responsible for providing reasonable resources for study implementation and to assure study progress. They are also responsible for communicating with regulatory agencies about the study protocol, the progress of the study, and study findings.

6.2.2 Investigator Roles and Responsibilities

The study investigators at RTI-HS and HealthCore share responsibility with AZ for the design of the study. The investigators at RTI-HS are responsible for conducting the CPRD and Medicare components in a manner that meets regulatory standards, conducting analyses, and preparing scientific reports. The investigators at HealthCore are responsible for conducting the HIRD^{SM} component in a manner that meets regulatory standards, conducting analyses, and preparing

scientific reports. The study will 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 AZ.

6.3 Confidentiality of Study Data

The confidentiality of records that could identify patients within the individual data sources must be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).

Data that could directly identify the patient will not be collected in the "study database."

6.4 Quality Control

Experienced US-based RTI-HS programmers will perform all analyses for the CPRD and Medicare data. To ensure the integrity and quality of the study results, RTI-HS will follow the programming validation life cycle process for all analyses. This includes quality-checking programs, logs, and output for accuracy according to relevant standard operating procedures. All programs will be independently reviewed by a second programmer/analyst.

This project will be guided by a written plan to ensure that all collaborators conduct quality-control checks of all aspects of data manipulation and analysis and preparation of study deliverables. The plan will specify that all collaborators will establish and maintain adequate documentation of performance of major tasks. The RTI-HS Office of Quality Assurance will conduct periodic audits during the study period to ensure that such documentation meets the necessary standards, especially the completion of these quality-control checks, according to the plan.

HealthCore's quality system is organized around the Quality Manual, the quality checks with the project life cycle, and standard operating procedures. HealthCore performs internal audits to endure adherence to the quality system according to a formal procedure and has procedures for retention of PHI and project data. The study will be tracked at various levels to help ensure that all aspects including project delivery, infrastructure, quality processes, resource management, and financial issues are addressed. To help ensure the highest level of quality on every project, HealthCore has established multiple layers of quality assurance throughout the project life cycle:

- Role-Based Control Checks: Each member of the team is responsible for performing thorough quality assurance checks on his or her work. In addition, the Project Director, in collaboration with the Lead Epidemiologist, is also accountable for quality of all deliverables.
- Quality Check Points: Centralized "check points" have been implemented during the data management cycle to help ensure accurate translation of programming requests.

- Quality Assurance Standards: Standard review procedures have been developed and are applied throughout the project life cycle.
- Automation: HealthCore has developed standard definitions of many variables and disease states and developed programs to apply these standards as needed on projects. These standards help ensure consistency, repeatability, and accuracy for each project.

HealthCore's research team documents the progress and scientific and quality review of all study activities and deliverables (e.g., protocol, reports, and manuscripts) in a project log. The project log provides documentation of the major study tasks related to a specific study activity performed by the research team, to develop and execute the requirements of the protocol or other guiding document for a HealthCore research project. In addition, the project log documents the quality assurance measures performed for each study activity during the conduct of the research project. Also, any research team and/or sponsor interaction resulting in a change to study specifications (e.g., protocol, study database, variables in the analytic files) is described in the project log. This is necessary to ensure that such communications are appropriately documented, that the most up-to-date versions of relevant documents are readily identifiable, and that affected documents are clearly tracked in the project log.

6.5 Database Retention and Archiving of Study Documents

Each investigator must retain all study records and source documents for the maximum period required by applicable regulations and guidelines, or institution procedures, or for the period specified by the sponsor, whichever is longer. Investigators must contact the study sponsor prior to destroying any records associated with the study.

The location of database and supporting documentation will be outlined in the final observational study report.

6.6 Registration of Study on Public Website

The study will be registered in the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) EU PAS Register (ENCePP, 2016) and ClinicalTrials.gov before the first data cut. The research team and study sponsor adhere to the general principles of transparency and independence in the ENCePP code of conduct (ENCePP, 2014).

6.7 Plans for Disseminating and Communicating Study Results

In accordance with the *Guidelines for Good Pharmacoepidemiology Practices* (ISPE, 2015), there is an ethical obligation to disseminate findings of potential scientific or public health importance, e.g., results pertaining to the safety of a marketed medication. The Consolidated Standards of Reporting Trials (CONSORT) statement refers to randomized studies, but also provides useful guidance applicable to reporting results of nonrandomized studies (Moher et al.,

2001). A well-developed publication strategy is encouraged in the *Guideline on Good Pharmacovigilance Practices*, module VIII, Section B.7 (European Medicines Agency, 2016).

Reports will be provided after each of the analyses, i.e., the descriptive analysis and the comparative analyses. RTI-HS personnel will work with HealthCore to submit the results from any of these analyses for publication and commit that they will, at least, publish the final results. The authorship of publications shall be in accordance with standards as described in the *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals* by the International Committee of Medical Journal Editors (2016).

7 ADVERSE EVENT REPORTING

7.1 Adverse Event Definitions

An AE is any untoward medical occurrence in a patient or clinical investigation patient administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding, for example), symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the medicinal product.

A nonserious adverse event is any AE that is not classified as serious.

A serious AE (SAE) is any untoward medical occurrence that at any dose:

- Results in death
- Is life-threatening (defined as an event in which the patient was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe)
- Requires inpatient hospitalization or causes prolongation of existing hospitalization (see Note below)
- Results in persistent or significant disability/incapacity
- Is a congenital anomaly/birth defect
- Is an important medical event (defined as a medical event(s) that may not be immediately life-threatening or result in death or hospitalization but, based upon appropriate medical and scientific judgment, may jeopardize the patient or may require intervention [e.g., medical, surgical] to prevent one of the other serious outcomes listed in the definition above.) Examples of such events include, but are not limited to, intensive treatment in an emergency room or at home for allergic bronchospasm; blood dyscrasias or convulsions that do not result in hospitalization.)

Suspected transmission of an infectious agent, pathogenic or nonpathogenic, via the AZ product under study is an SAE.

An ***overdose*** is defined as the accidental or intentional administration of any dose of a product that is considered both excessive and medically important. Although pregnancy, overdose and cancer are not always serious by regulatory definition, these events are handled as SAEs.

Note: The following hospitalizations are not considered SAEs:

- A visit to the emergency room or other hospital department < 24 hours that does not result in admission (unless considered an important medical or life-threatening event).
- Elective surgery, planned prior to signing consent.
- Routine health assessment requiring admission for baseline/trending of health status (e.g., routine colonoscopy).
- Medical/surgical admission other than to remedy ill health and planned prior to entry into the study.
- Admission encountered for another life circumstance that carries no bearing on health status and requires no medical/surgical intervention (e.g., lack of housing, economic inadequacy, caregiver respite, family circumstances, administrative reasons).

7.2 Adverse Event Collection and Reporting

All AEs collected will be reported in aggregate in the final study report.

8 GLOSSARY OF TERMS AND LIST OF ABBREVIATIONS

8.1 Glossary of Terms

Not applicable.

8.2 List of Abbreviations

Term	Definition
AASLD	American Association for the Study of Liver Diseases
AD	antidiabetic drug
AE	adverse event
ALI	acute liver injury
ALT	alanine aminotransferase
AP	alkaline phosphatase
AST	aspartate aminotransferase
AZ	AstraZeneca Pharmaceuticals LP
BMS	Bristol-Myers Squibb
CB	conjugated bilirubin
CI	confidence interval
CMS	Centers for Medicare and Medicaid Services
CONSORT	Consolidated Standards of Reporting Trials
CPRD	Clinical Practice Research Datalink
CPT	Current Procedural Terminology
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union

Term	Definition
FDA	Food and Drug Administration
GP	general practitioner
GPRD	General Practice Research Database
HAART	highly active antiretroviral therapy
HbA1c	hemoglobin A1c (glycated hemoglobin)
HCPCS	Healthcare Common Procedure Coding System
HES	Hospital Episode Statistics database (UK)
HIPAA	Health Insurance Portability and Accountability Act
HIRD^{SM}	HealthCore Integrated Research Database
ICD-9-CM	International Classification of Diseases, 9th Revision, Clinical Modification
IRB	institutional review board
IRR	incidence rate ratio
ISAC	Independent Scientific Advisory Committee
ISPE	International Society for Pharmacoepidemiology
LFT	liver function test
MREC	Multicentre Research Ethics Committee
NDC	National Drug Code
NEC	not elsewhere classified
NOS	not otherwise specified
NSAID	nonsteroidal antiinflammatory drugs
OS	otherwise specified
PHI	protected health information
QC	quality control
RTI-HS	RTI Health Solutions
SAE	serious adverse event
SGLT2	sodium-glucose cotransporter 2
SGOT	another designation for aspartate aminotransferase (AST)
SGPT	another designation for alanine aminotransferase (ALT)
T1DM	type 1 diabetes mellitus
T2DM	type 2 diabetes mellitus
TB	total bilirubin
UK	United Kingdom
ULN	upper limit of normal
US	United States of America

9 REFERENCES

Almdal TP, Sørensen TI. Incidence of parenchymal liver diseases in Denmark, 1981 to 1985: analysis of hospitalization registry data. The Danish Association for the Study of the Liver. Hepatology. 1991;13:650-5.

American Diabetes Association. Standards of medical care in diabetes—2014. Diabetes Care. 2014 Jan;37 Suppl 1:S14-80.

Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009;28(25):3083-107.

Bénichou C. Criteria of drug-induced liver disorders. Report of an international consensus meeting. J Hepatol. 1990;11:272-6.

Boccuzzi SJ, Wogen J, Fox J, Sung JCY, Shah AB, Kim J. Utilization of oral hypoglycemic agents in a drug-insured U.S. population. Diabetes Care. 2001;24:1411-15.

Braitman LE, Rosenbaum PR. Rare outcomes, common treatments: analytic strategies using propensity scores. Ann Intern Med. 2002;137:693-5.

Bristol-Myers Squibb (BMS) and AstraZeneca (AZ). Background document: dapagliflozin. BMS-512148. NDA 202293. 13 June 2011. Available at: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM262996.pdf. Accessed 03 October 2012.

Carson JL, Strom BL, Duff A, Gupta A, Das K. Safety of nonsteroidal anti-inflammatory drugs with respect to acute liver disease. Arch Intern Med. 1993;153:1331-6.

Centers for Disease Control and Prevention. National diabetes statistics report, 2014. Available at: https://www.cdc.gov/diabetes/pdfs/data/2014-report-estimates-of-diabetes-and-its-burden-in-the-united-states.pdf. Accessed 22 February 2017.

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;158:280-7.

Chan KA, Truman A, Gurwitz JH, Hurley JS, Martinson B, Platt R, et al. A cohort study of the incidence of serious acute liver injury in diabetic patients treated with hypoglycemic agents. Arch Intern Med. 2003;163:728-34.

Cubanski J, Huang J, Damico A, Jacobson G, Neuman T. Medicare chartbook, 4th edition. Menlo Park, CA: The Henry J. Kaiser Family Foundation; 2010. Available at: http://kff.org/medicare/report/medicare-chartbook-2010/. Accessed 17 May 2013.

D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med. 1998;17:2265-81.

de Abajo FJ, Montero D, Madurga M, Garcia Rodriguez LA. Acute and clinically relevant drug-induced liver injury: a population based case-control study. Br J Clin Pharmacol. 2004;58:71-80.

European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module VIII – Post-authorisation safety studies (Rev 2). 04 August 2016. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129137.pdf Accessed 24 February 2017.

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. The ENCePP code of conduct for scientific independence and transparency in the conduct of pharmacoepidemiological and pharmacovigilance studies. Revision 3. 21 February 2014. Available at: http://www.encepp.eu/code_of_conduct/documents/ENCePPCodeofConduct_Rev3.pdf. Accessed 17 June 2014.

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). The European Union Electronic Register of Post-Authorisation Studies (EU PAS Register). 15 July 2016. Available at: http://www.encepp.eu/encepp_studies/indexRegister.shtml.Accessed 24 February 2017.

Food and Drug Administration. FDA Working Group. P*h*RMA/FDA/AASLD drug-induced hepatotoxicity white paper: postmarketing considerations. November 2000. Available at: http://www.fda.gov/downloads/Drugs/ScienceResearch/ResearchAreas/ucm091462.pdf. Accessed 15 August 2012.

Food and Drug Administration. Guidance for Industry. Drug-induced liver injury: premarketing clinical evaluation. 2009. Available at http://www.fda.gov/downloads/Drugs/.../Guidances/UCM174090.pdf. Accessed 04 June 2014.

Gallagher AM, Smeeth L, Seabroke S, Huber GML, 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:e28157.

Gelfand JM, Margolis DJ, Dattani H. The UK General Practice Research Database. In: Strom BL, editor Pharmacoepidemiology, 4th edition. John Wiley & Sons; 2005. p. 337-46.

González EL, Johansson S, Wallander MA, Rodríguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996–2005. J Epidemiol Community Health. 2009;63:332-6.

Hall GC, McMahon AD, Carroll D, Home PD. Macrovascular and microvascular outcomes after beginning of insulin versus additional oral glucose-lowering therapy in people with type 2 diabetes: an observational study. Pharmacoepidemiol Drug Saf. 2012;21:305-13.

Huerta C, Zhao SZ, García Rodríguez LA. Risk of acute liver injury in patients with diabetes. Pharmacotherapy. 2002;22:1091-6.

Ibáñez L, Perez E, Vidal X, Laporte JR. 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;37:592-600.

International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. December 2016. Available at: http://www.icmje.org/urm_main.html. Accessed 31 January 2017.

International Society for Pharmacoepidemiology (ISPE). Guidelines for good pharmacoepidemiology practices (GPP). Revision 3. June 2015. Available at: http://www.pharmacoepi.org/resources/guidelines_08027.cfm. Accessed 31 January 2017.

Lash TL, Fox MP, Fink AK. Applying Quantitative Bias Analysis to Epidemiologic Data. Springer Dordrecht Heidelberg London New York. 2009.

Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet. 2001;357:1191-4.

National Institute for Health and Clinical Excellence. Type 2 diabetes in adults: management. 2015. Available at: https://www.nice.org.uk/guidance/ng28. Accessed 22 February 2017.

Navarro VJ, Senior JR. Drug-related hepatotoxicity. N Engl J Med 2006;354:731-9.

Perkins SM, Tu W, Underhill MG, Zhou XH, Murray MD. The use of propensity scores in pharmacoepidemiologic research. Pharmacoepidemiol Drug Saf. 2000;9:93-101.

Research Data Assistance Center. Strengths and limitations of CMS administrative data in research. 2013. Available at: http://www.resdac.org/resconnect/articles/156. Accessed 17 May 2013.

Rothman KJ. Episheet: spreadsheets for the analysis of epidemiologic data (software). 04 October 2012. Available at: http://krothman.hostbyet2.com/Episheet.xls. Accessed 24 February 2017.

Rothman KJ, Ray W. Should cases with a 'known' cause of their disease be excluded from study? (commentary). Pharmacoepidemiol Drug Saf. 2002;11:11-4.

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;24:1412-8.

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;36:451-5.

Temple R. Hepatotoxicity through the years: impact on the FDA (presentation). 28 March 2001. Available at: http://www.fda.gov/downloads/Drugs/ScienceResearch/ResearchAreas/ucm122149.pdf. Accessed 27 July 2012.

Temple R. Hy's law: predicting serious hepatotoxicity. Pharmacoepidemiol Drug Saf 2006;15:241-3.

Van Staa TP, 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.

APPENDIX 1. DIAGNOSIS CODES

Table 1-1: Electronic Algorithm ICD-9-CM Codes to Use in Screening for Acute Liver Injury in the US Data Sources

Code	Type	Description
572.2	ICD-9	HEPATIC COMA
570	ICD-9	ACUTE NECROSIS OF LIVER
572.4	ICD-9	HEPATORENAL SYNDROME
573.3	ICD-9	UNSPECIFIED HEPATITIS
573.8	ICD-9	OTHER SPECIFIED DISORDERS OF LIVER
996.82	ICD-9	COMPL LIVER TRANSPLANT
V42.7	ICD-9	LIVER TRANSPLANT STATUS
782.4	ICD-9	JAUNDICE, UNSPECIFIED, NOT OF NEWBORN
50.5	ICD-9-PX	LIVER TRANSPLANT
47133	CPT	DONOR HEPATECTOMY; CADAVER
47135	CPT	LIVER ALLOTRANSPLANTATION; ORTHOTOPIC; PARTIAL OR WHOLE
47136	CPT	LIVER ALLOTRANSPLANTATION; HETEROTOPIC; PARTIAL OR WHOLE
47143	CPT	PREP CADAVER DONOR ALLOTRANSPLANTATION; NO SPLIT
47144	CPT	PREP CADAVER DONOR ALLOTRANSPLANTATION; TRISEGMENT SPLIT
47145	CPT	PREP CADAVER DONOR ALLOTRANSPLANTATION; LOBE SPLIT
47146	CPT	RECONSTRUCTION LIVER GRAFT PRE-ALLOTRANSPLANTATION; CADAVER OR LIVE DONOR; VENOUS ANASTOMOSIS
47147	CPT	RECONSTRUCTION LIVER GRAFT PRE-ALLOTRANSPLANTATION; CADAVER OR LIVE DONOR; ARTERIAL ANASTOMOSIS

CPT = Current Procedural Terminology (code); ICD-9 = *International Classification of Diseases, 9th Revision*; US = United States of America.

Table 1-2: Electronic Algorithm Read Codes to Use in Acute Liver Injury in the CPRD

Code	Description
1675.00	Yellow/jaundiced color
1675.11	Jaundice - symptom
2274.00	O/E - jaundiced color
2274.11	O/E - jaundiced
7806.00	Therapeutic endoscopic operations on liver using laparoscope
7807.00	Diagnostic endoscopic examination of liver using laparoscope
7800111	Auxiliary liver transplant
7800112	Piggy back liver transplant

Code	Description
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
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
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 - noninfective

Code	Description
J600200	Acute yellow atrophy
J600z00	Acute necrosis of liver NOS
J601.00	Subacute necrosis of liver
J601000	Subacute hepatic failure
J601100	Subacute hepatitis - noninfective
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
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	Nonspecific 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

Code	Description
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
67P4200	Discussion about liver transplantation
7800.00	Transplantation of liver
7800000	Orthotopic transplantation of liver
7800100	Heterotopic transplantation of liver
7800400	Orthotopic transplantation of whole liver
7800y00	Other specified transplantation of liver
7806y00	Therapeutic laparoscopic operation on liver OS
7806z00	Therapeutic laparoscopic operation on liver NOS
780Ay00	Other specified diagnostic percutaneous operation on liver
7L1f.00	Compensation for liver failure
7L1fz00	Compensation for liver failure NOS
8A7..00	Liver function monitoring
8HlW.00	Referral to liver unit
8LH..00	Liver transplant planned

ALT = alanine aminotransferase; AST = aspartate aminotransferase; CPRD = Clinical Practice Research Datalink; LFT = liver function tests; NEC = not elsewhere classified; NOS = not otherwise specified; OS = otherwise specified; SGOT = another designation for AST; SGPT = another designation for ALT.

Source: Medical and product dictionary browsers, version 3.0. London: General Practice Research Database (now the Clinical Practice Research Datalink); September 2015.

Table 1-3: Exclusion Criteria: ICD-9-CM Codes to be Mapped to Read Codes and ICD-10 Codes

ICD-9-CM Code	ICD-9-CM Description
History of acute liver injury	
570	Acute and subacute necrosis of liver
572.2	Hepatic coma
573.3	Hepatitis unspecified
V42.7	Liver transplant
50.5 procedure	Liver transplant
Chronic liver disease and alcoholism	
571	Chronic liver disease and nonalcoholic cirrhosis
571.0	Alcoholic fatty liver
571.1	Acute alcoholic hepatitis
571.2	Alcoholic cirrhosis of liver
571.3	Alcoholic liver damage
571.4x	Chronic hepatitis
571.5	Cirrhosis of liver without mention of alcohol
571.6	Biliary cirrhosis
571.8	Other chronic nonalcoholic liver disease
571.9	Unspecified chronic liver disease without mention of alcohol
572.0	Abscess of liver
572.4	Hepatorenal syndrome
572.1	Portal pyemia
572.3	Portal hypertension
572.8	Other sequelae of chronic liver disease
573.0	Chronic passive congestion of the liver
573.4	Hepatic infarction
573.8	Other specific disorders of the liver
573.9	Unspecified disorder of liver
291.x	Alcoholic psychoses
303.x	Alcohol dependence syndrome
357.5	Alcoholic polyneuropathy
425.5	Alcoholic cardiomyopathy
V11.3	Alcoholism
Infectious hepatitis	
070.x	Viral hepatitis

ICD-9-CM Code	ICD-9-CM Description
072.3	Mumps pancreatitis
072.71	Mumps hepatitis
091.62	Secondary syphilitic hepatitis
130.5	Hepatitis due to toxoplasmosis
573.1	Hepatitis in viral disease classified elsewhere
573.2	Hepatitis in other infectious diseases classified elsewhere
V02.6x	Viral hepatitis
Chronic disease involving the liver or causing hyperbilirubinemia	
275.x	Hemochromatosis
275.1	Wilson's disease
277.6	Deficit of alpha-1-antitrypsin
453.0	Budd-Chiari syndrome
277.4	Gilbert's disease
Biliary disease	
574.x	Cholelithiasis (with and without cholecystitis)
575.x	Other disorders of gallbladder (cholecystitis)
576.x	Other disorders of biliary tract (cholangitis)
Pancreatic disease	
577.x	Disease of pancreas
Hepatobiliary and pancreatic neoplasms	
155.x-157.x, 197.7	Malignant neoplasms of liver and intrahepatic bile ducts, gallbladder and extrahepatic bile ducts, pancreas. Secondary neoplasm of the liver
Congestive heart failure	
428.x	Heart Failure

Source: ICD-9-CM = *International Classification of Diseases, 9th Revision, Clinical Modification.*

APPENDIX 2. ASSESSMENT OF INCIDENCE DURING CURRENT USE AND SWITCHING USE AND ESTIMATION OF PERSON-TIME DURING RECENT USE

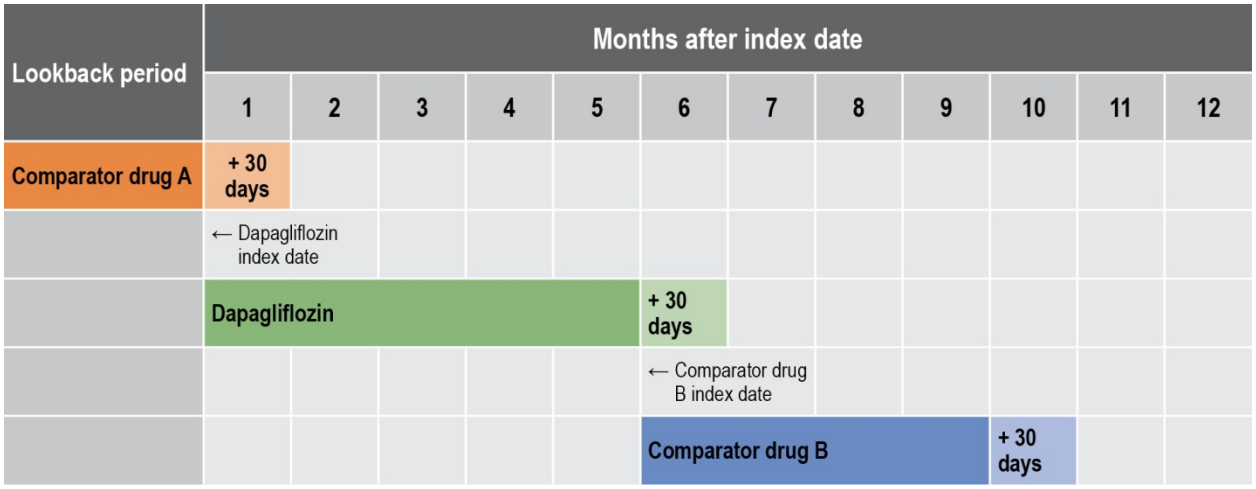
Example 1 through Example 6 on the following pages focus on assessment of incidence during dapagliflozin use compared with (1) use of other comparator antidiabetic drugs (ADs) as a group and (2) use of specific ADs.

- For calculation of crude incidence and incidence rate ratios, the following comparisons can be made:
  - Incidence of ALI during any dapagliflozin person-time can be compared with incidence during any comparator-exposed person-time.
  - Incidence of ALI during combined dapagliflozin-plus-comparator–exposed person-time can be compared with comparator-exposed person-time (see Example 1).
- For specific medication comparisons, e.g., comparing dapagliflozin to pioglitazone, the person-time and events during dapagliflozin person-time excluding any time overlapping with pioglitazone person-time can be compared with any non-dapagliflozin person-time in the pioglitazone group.
- For adjusted incidence, (1) the propensity score will be used to adjust for concomitant medications at the index date, but only if not newly initiated at the time they were started, and (2) stratification or a multivariable model will be used to adjust for concomitant medications added during follow-up or switched to at the end of follow-up.

Example1. Drug A was a medication newly initiated before the study period and is not an excluded AD. Drug A is not eligible to be a comparator drug as it was not newly initiated in the observation period; therefore, the risk window for drug A is not evaluated, but use of drug A is controlled for in development of the propensity score.

The switch is from drug A, a comparator drug that was newly initiated but not during the observation period, to dapagliflozin and then from dapagliflozin to drug B, a newly initiated drug from the comparator group. At the initiation of drug B, the patient is eligible to be sampled for the comparator group.

Figure 2-1: Switch From a Potential Comparator Drug Not in Exclusion Criteria to Dapagliflozin and Then to a Different, New Comparator Drug



Risk window related to dapagliflozin = months 1-6; dapagliflozin categorized as “add-on” to drug A at the index date.

Risk window is further categorized into the following mutually exclusive categories:

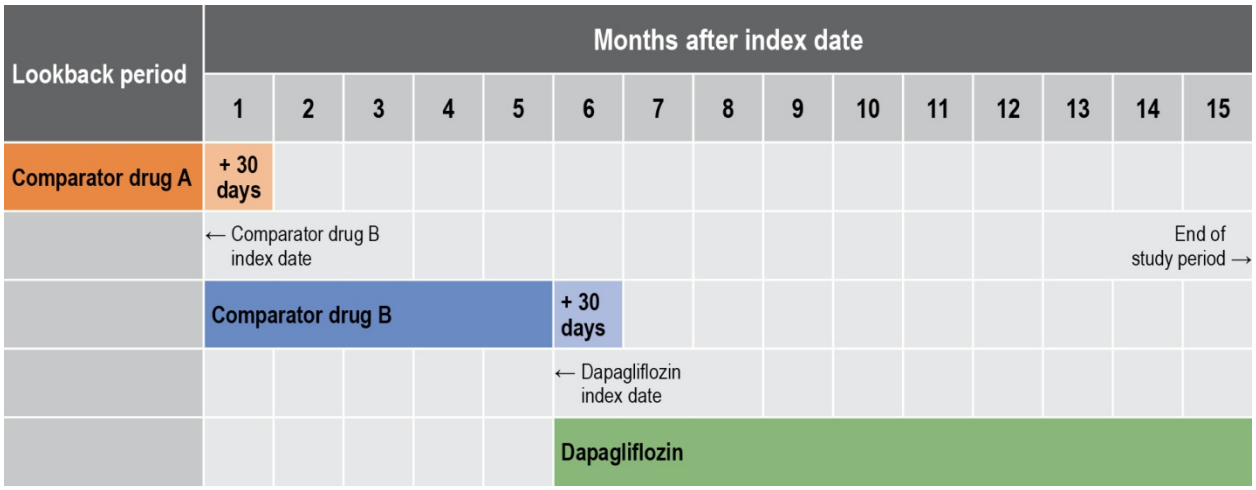
- Single exposure during follow-up: drug A: 0 months; dapagliflozin: months 2-5; drug B: months 7-10
- Multiple exposure during follow-up: months 1 and 6. Person-time in these months will be grouped into a “combined” exposure category and analyzed separately from the dapagliflozin-only exposure time and the comparator drug B–only exposure time.

Example 2. Drug A is not eligible to be a comparator drug; therefore, the risk window for drug A is not evaluated, but use of drug A is controlled for in development of the propensity score.

Drug B is newly initiated and is eligible to be selected for the comparator cohort. If selected, follow-up starts at the beginning of month 1. Drug B risk window = months 1-6. Drug B is categorized as “add-on” to drug A at the index date.

When the patient switches to dapagliflozin, person-time will be counted as “combined” exposure in month 6 and dapagliflozin-only in months 7-15. Person-time in months 1 and 6 will be grouped into a “combined” exposure category and analyzed separately from the dapagliflozin-only exposure time and the comparator drug B–only exposure time.

Figure 2-2: Switch From Comparator to Dapagliflozin



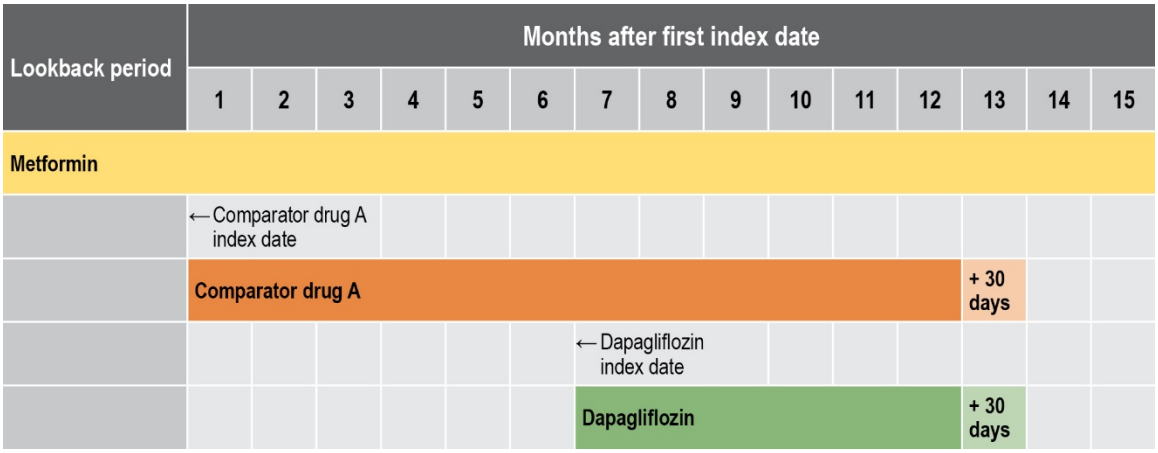
Example 3. Comparator drug A is initiated as an add-on to metformin, then dapagliflozin is added in month 7 of follow-up. If dapagliflozin is added to a comparator drug treatment episode during follow-up, the person-time with combined exposure to dapagliflozin and comparator drug A (months 7-12) will be analyzed in a separate category of "combined exposure."

At the time of the addition of dapagliflozin (start of month 7), the patient is exposed to comparator drug A and dapagliflozin (and metformin, exposure to which will be accounted for in development of the propensity score).

Risk window for drug A only = months 1-6.

Combined (dapagliflozin and comparator) exposure risk window = months 7-13.

Figure 2-3: Addition of Dapagliflozin to Comparator During Follow-up



Example 4. Comparator drug A is initiated as an add-on to metformin, then drug A is switched to comparator drug B, another newly initiated AD.

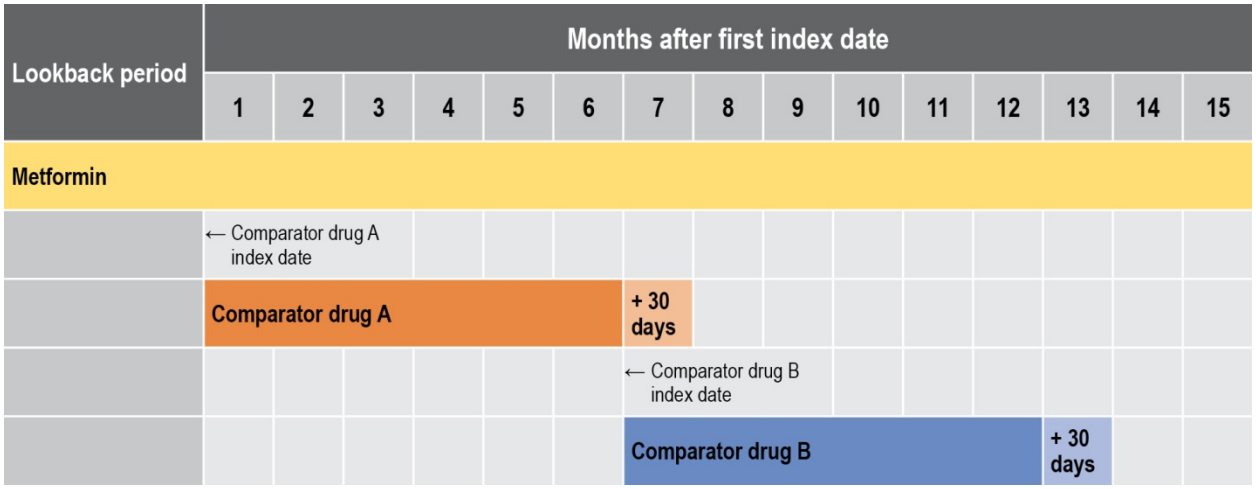
At the time of the switch (start of month 7), the patient is eligible to be selected into the comparator group (again).

Risk window for drug A only = months 1-6.

Risk window for drug A plus drug B (combined exposure) = month 7

Risk window for drug B, if selected into comparator group = months 8-13.

Figure 2-4: Add on Drug A to Metformin, Then Switch to Drug B



Example 5. Sensitivity analysis with 60-day extended risk window

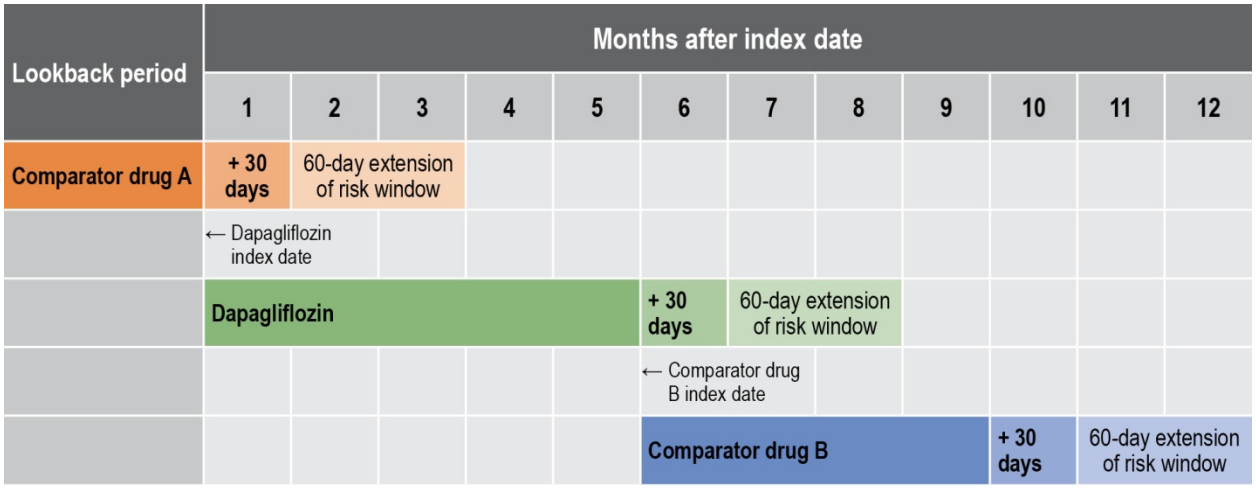
Drug A is not eligible to be a comparator drug; therefore, the risk window for drug A is not evaluated, but use of drug A is controlled for in development of propensity score.

Risk window for dapagliflozin only = months 1-5.

Drug B is newly initiated by the patient, so patient is eligible to be selected for the comparator cohort; however, patient is not selected, so months 9-12 do not contribute to person-time exposure to drug B.

Risk window for dapagliflozin plus drug B = months 6-8.

Figure 2-5: Sensitivity Analysis With Extended Risk Window



Example 6. Sensitivity analysis with exposure limited to only the index exposure episodes

A sensitivity analysis will be performed that includes only the index exposure episodes. In Figure 2-6 and Figure 2-7, the follow-up time for the sensitivity analysis will be censored at the end of month 6 for each type of index exposure.

Risk window for dapagliflozin plus drug B (combined exposure) = month 6.

Figure 2-6: Sensitivity Analysis Including Only the Index Exposure Episode, Dapagliflozin

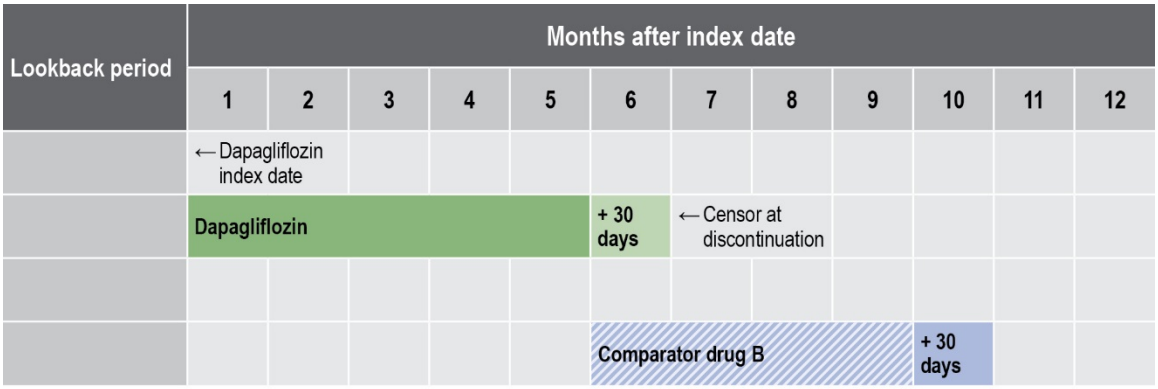
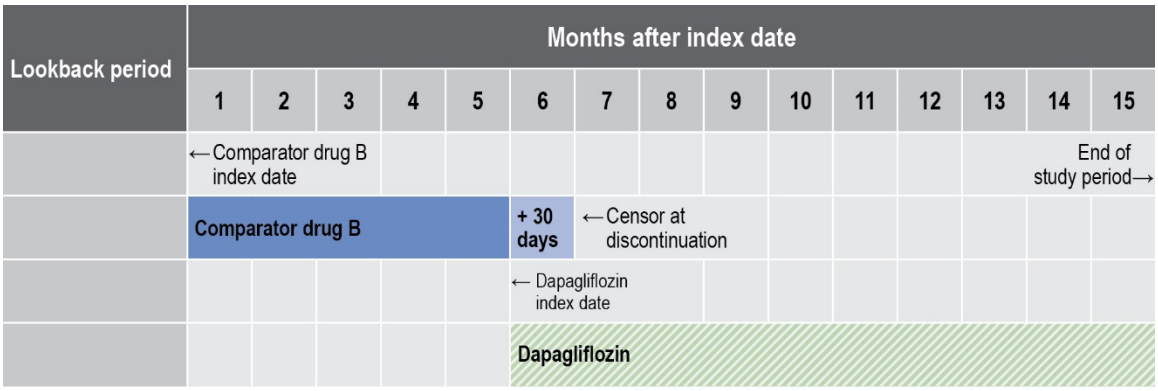


Figure 2-7: Sensitivity Analysis Including Only the Index Exposure Episode, Comparator



APPENDIX 3. OVERVIEW OF DATA SOURCE CHARACTERISTICS

Characteristic	United Kingdom (Population, 62,435,709) Clinical Practice Research Datalink (CPRD)	United States HIRD[SM], Aged 64 Years or Younger	United States Medicare, Aged 65 Years or Older
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.	Health insurance claims of patients enrolled in Anthem-affiliated health plans	Health insurance claims of patients enrolled in Medicare health insurance program; health insurance claims include pharmacy dispensings for those with Part D, hospital and outpatient claims, and procedure claims; medical record review is an option through a trusted third party
Database population (n)	5.1 million	35.8 million lives since 2006[a]	47 million
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.	United States residents who are enrolled in a covered employer-sponsored health plan	Federally sponsored health insurance program that offers health coverage to 47 million people, including 39 million people aged 65 years or older and 8 million nonelderly people with a permanent disability
Proportion of the country's population covered	8%	8%	60% of individuals aged 65 years or older and with Part D coverage
Representativeness of patients and practices	Age and sex of patients are representative of the UK population	Representative of all major geographic regions and health care settings for commercially insured US population aged less than 65 years	Considered to be representative of US population aged 65 years or older
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	None	None
Geographic location	First digits of physician's practice postal code	US	US

Characteristic	United Kingdom (Population, 62,435,709) Clinical Practice Research Datalink (CPRD)	United States HIRD℠, Aged 64 Years or Younger	United States Medicare, Aged 65 Years or Older
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 (of filled).	All claims submitted to Anthem, Inc.	All pharmacy claims submitted to Medicare
Drug dictionary codes/ therapeutic classification	Multilex/British National Formulary	National Drug Codes for outpatient prescriptions	National Drug Codes for outpatient prescriptions
Unique product code	Yes	Yes	Yes
Prescribed/dispensed drugs	GP prescriptions issued	Dispensed drugs at outpatient pharmacies	Dispensed drugs at outpatient pharmacies
Date drug prescribed/ dispensed	Yes, date the drug was prescribed	Yes, date dispensed	Yes, date dispensed
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.	Yes	Yes
Duration	There is a field to record duration, but it is highly incomplete. Duration can be derived from the number of prescriptions.	Days' supply	Days' supply is provided
Clinical indication	There is no field for indication. The user needs to assess diagnoses on the prescription. Prior diagnosis can be used as a proxy.	There is no field for indication. The user needs to assess diagnoses on the dispensing date. Prior diagnosis can be used as a proxy.	There is no field for indication. The user needs to assess diagnoses on the dispensing date. Prior diagnosis can be used as a proxy.
Inpatient medications	No	No	No

Characteristic	United Kingdom (Population, 62,435,709) Clinical Practice Research Datalink (CPRD)	United States HIRD[SM], Aged 64 Years or Younger	United States Medicare, Aged 65 Years or Older
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.	Available for many but not all outpatient pharmacy claims	Yes, if dispensed in outpatient setting
Diagnoses and procedures			
Coding system	Read	ICD-9-CM, ICD-10-CM, CPT, HCPCS	ICD-9-CM, ICD-10-CM, CPT, HCPCS
Outpatient visits	Yes, as entered by the GP	Yes, one or more diagnoses on submitted claim	Yes, one or more diagnoses on submitted claim
Hospitalization data	Partial linkage to HES; as recorded by GPs	Yes, one or more diagnoses on submitted claim	Yes, one or more diagnoses on submitted claim
Specialist visits	Information from referral letters	Yes, one or more diagnoses on submitted claim	Yes, one or more diagnoses on submitted claim
Emergency room visits	As entered by the GP	Yes, one or more diagnoses on submitted claim	Yes, one or more diagnoses on submitted claim
Time period covered	Since 1987	Since 2006	Medicare Part D available since 2006
Updates	Quarterly	Monthly	Yearly
Approximate time lag	6-12 weeks	3-4 months	Up to 24 months
Access to medical records	GPs can be sent questionnaires via the CPRD for validation; also partial linkage to HES	Only through trusted third party	Only through trusted third party
Data transfer	Yes, third-party approval for standard data and linked databases. Data set will be delivered for analysis	No, data remain with at HIRD[SM]	Yes, after CMS protocol review
Approval process	ISAC approval of short protocol	Central institutional review board, possibly local institutional review board(s)	Local institutional review board, CMS Privacy Board

CMS = Centers for Medicare and Medicaid Services; CPRD = Clinical Practice Research Datalink; CPT = Current Procedural Terminology; GP = general practitioner; HCPCS = Healthcare Common Procedure Coding System; HES = Hospital Episode Statistics; HIRD[SM] = HealthCore Integrated Research

Database; ICD-9-CM = *International Classification of Diseases, 9th Revision, Clinical Modification*; ISAC = Independent Scientific Advisory Committee; UK = United Kingdom; US = United States of America.

[a] HIRD[SM] patients of any age enrolled in a qualifying health plan at any time between January 2006 and January 2014.

APPENDIX 4. ANTIDIABETIC DRUGS ELIGIBLE FOR INCLUSION IN THE COMPARATOR GROUP

Blood Glucose–Lowering Drugs (Excluding Insulin) by ATC Subgroup	Active Substance
A10BA, Biguanides	Metformin
A10BB, Sulfonamides, urea	Glibenclamide/glyburide
	Tolbutamide
	Gliclazide
	Glimepiride
	Carbutamide
	Chlorpropamide
	Tolazamide
	Glipizide
	Gliquidone
	Glyclopyramide
	Acetohexamide
A10BD, Combinations	Metformin/glibenclamide
	Metformin/rosiglitazone
	Rosiglitazone/glimepiride
	Pioglitazone/metformin hydrochloride
	Pioglitazone/glimepiride
	Sitagliptin/metformin hydrochloride
	Vildagliptin/metformin hydrochloride
	Pioglitazone/alogliptin
A10BF, Alpha glucosidase inhibitors	Acarbose
	Voglibose
	Miglitol
A10BG, Thiazolidinediones	Pioglitazone
A10BH, DPP-4 (dipeptidyl peptidase-4) inhibitors	Sitagliptin
	Vildagliptin
	Saxagliptin
	Linagliptin
	Alogliptin

Blood Glucose–Lowering Drugs (Excluding Insulin) by ATC Subgroup	Active Substance
A10BH, DPP-4 Combinations	Alogliptin/metformin
	Linagliptin/metformin
	Saxagliptin/metformin
A10BX, Other	Repaglinide
	Nateglinide
	Mitiglinide
	Exenatide
	Liraglutide
	Albiglutide
	Dulaglutide
	Lixisenatide

ATC = Anatomical Therapeutic Chemical (classification system).

Source: World Health Organization Collaborating Centre for Drug Statistics Methodology. ATC/DDD index 2015. Available at: http://www.whocc.no/atc_ddd_index/. Accessed 30 October 2015.

APPENDIX 5. INCIDENCE OF DRUG-INDUCED ACUTE LIVER INJURY IN THE GENERAL POPULATION

Reference	Study Population	Study Design	Exclusions	Number of Cases	Annual Incidence per 100,000 Population (95% CI)	Endpoint Definition	Endpoint Validation
De Valle et al., 2006	Sweden, hospital outpatient hepatology clinic	Retrospective assessment of cases	None	1,164	2.3	Criteria of International Consensus Meeting[a]	Yes
Andrade et al., 2005	Spain, regional registry	Prospective surveillance; outpatient and inpatient cases	None	461	3.4 (1.3-5.5)	Hepatotoxicity Criteria of International Consensus Meeting[a]	Yes
de Abajo et al., 2004	UK, GPRD	Retrospective cohort	History of hepatobiliary disease, malignancy, alcohol-related disorders	128	Incidence rate per 100,000 person-years: 2.4 (2.0-2.8)	Criteria of International Consensus Meeting[a]	Yes
Carson et al., 1993	US, Medicaid Michigan and Florida	Retrospective cohort	History of liver disease	107	2.2 (2.0-2.4)	ALT or AST or bilirubin or AP > 2 the control value	Yes
Sgro et al., 2002	France, Nevers area	Prospective intensive surveillance	None	34	13.9 (9.7-19.5)	Criteria of International Consensus Meeting[a]	Yes
Ibáñez et al., 2002	Spain, Hospital network	Prospective surveillance	None	86	0.7 (0.6-0.9)	Criteria of International Consensus Meeting[a]	Yes

ALT = alanine aminotransferase; AP = alkaline phosphatase; AST = aspartate aminotransferase; CI = confidence interval; GPRD = General Practice Research Database (now the Clinical Practice Research Datalink [CPRD]); UK = United Kingdom; US = United States.

[a] Bénichou, 1990.

APPENDIX 6. DIAGNOSTIC CRITERIA FOR ACUTE LIVER INJURY

Table 6-1: Diagnostic Criteria for Acute Liver Injury

	International Consensus Meeting 1990; Bénichou 1990	FDA Working Group, 2000; Navarro and Senior, 2006	Temple, 2001; 2006 Hy's law
Acute liver injury			
ALT	> 2 x ULN or	> 3 x ULN or	> 3 x ULN and
CB	> 2 x ULN or		≥ 2 x ULN[a]
AST, AP, TB	All increased and one > 2 x ULN		
ALT, AP, TB		All increased and TB > 2 x ULN or	
AP		> 2 x ULN	

ALT = alanine aminotransferase; AP = alkaline phosphatase; AST = aspartate aminotransferase; CB = conjugated bilirubin; FDA = US Food and Drug Administration; TB = total bilirubin; ULN = upper limit of normal.

[a] And no evidence of intra-or extrahepatic bilirubin obstruction (elevated alkaline phosphatase) or Gilbert's syndrome.

Table 6-2: Diagnostic Criteria for Acute Liver Injury Used in Observational Studies of Patients With Diabetes

	Lo Re III, 2012 (saxagliptin)	Chan, 2003	Graham, 2003	Huerta, 2002
Acute liver injury	Not evaluated	Not evaluated	Not evaluated	Criteria of international consensus meeting (Bénichou, 1990)

Table 6-3: Exclusions in Observational Studies of Patients With Diabetes and Acute Liver Injury

Exclusion Diagnosis	Huerta, 2002 GPRD, UK	Chan, 2003 5 Health Maintenance Organizations, US	Graham, 2003 UnitedHealth Group, US
Liver disease	X[a]	X[b]	X
Abnormal liver function tests	X[c]		
Gallbladder diseases	X[d]	X[e]	
Pancreatic diseases	X		
Alcoholism	X	X	
Congestive heart failure	X	X	
Pregnancy	X		
Cancer	X		
Primary or secondary hepatic or biliary tract cancer		X	X
Liver transplantation		X	
Shock (hypovolemic, cardiogenic, septic)		X	

GPRD = General Practice Research Database (now the CPRD); UK = United Kingdom; US = United States.

[a] Fulminant hepatitis, hepatitis, liver necrosis, other liver disorders, hepatocellular damage, liver biopsy, enlarged liver, pale stools, jaundice, drug-induced jaundice.

[b] Viral hepatitis, disorders of copper metabolism, necrosis of liver, chronic liver disease, cirrhosis, hepatic coma, hepatorenal syndrome, unspecified hepatitis, hepatic infarction, other specified liver disorders, unspecified liver disorders, alcoholic hepatitis.

[c] Bilirubin serum level abnormal, abnormal liver function test, biochemical liver dysfunction, abnormal liver enzymes, liver enzymes raised, abnormal hepatic function, alanine aminotransferase raised, aspartate aminotransferase raised.

[d] Obstructive jaundice, cholelithiasis, other gallbladder diseases.

[e] Biliary tract obstruction.

APPENDIX 7. COVARIATES TO BE CONSIDERED FOR INCLUSION IN THE PROPENSITY SCORE MODEL

Additional variables that are risk factors specific to ALI are listed in Section 3.4.3, Table 1.

Medical Comorbidities
Urinary infections (chronic or recurring)
Kidney stones
Bladder stones
Colon polyps
Crohn's disease
Ulcerative colitis
Pancreatitis
Immunosuppressive diseases such as human immunodeficiency virus infection/AIDS