Boehringer Ingelheim

Protocol for non-interventional studies based on existing data

Document Number:	c26457869-02
BI Study Number:	1245.195
BI Investigational Product(s):	Jardiance ® (empagliflozin)
Title:	**Multi-country non-interventional study on the effectiveness and safety of Empagliflozin in adult patients with type 2 diabetes in Europe and Asia**
Protocol version identifier:	2.0
Date of the last version of the protocol:	29 November 2018
PASS:	No
EU PAS register number:	EUPAS27606
Active substance:	A10BK03 empagliflozin
Medicinal product:	Jardiance
Product reference:	Europe: EMEA/H/C/002677 Asia: NA
Procedure number:	N/A
Joint PASS:	No
Research question and objectives:	The main objective of this study is to compare effectiveness outcomes between the incident users of empagliflozin or any SGLT-2 inhibitor and the incident users of any DPP-4 inhibitor. The effectiveness outcomes include, where available, cardiovascular events, renal effectiveness outcomes, and cardiovascular and all-cause mortality. Additional analyses include safety outcomes. Health care resources utilization and cost of care outcomes will also be evaluated.
Countries of study:	Denmark, Finland, Germany, Israel, Japan, Norway, South Korea, Spain, Sweden, Taiwan, United Kingdom. Other countries may be included later.

Author:	[redacted]
Marketing authorization holder:	[redacted]
MAH contact person:	[redacted]
In case of PASS, add: **<EU-QPPV:>**	N/A
In case of PASS, add: **<Signature of EU-QPPV:>**	N/A
Date:	11 October 2019

Page 1 of 132

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

AT	As-treated
ATC	Anatomical Therapeutic Chemical classification
BI	Boehringer Ingelheim International GmbH
BNP	B-type natriuretic peptide
BUN	Blood urea nitrogen
CABG	Coronary artery bypass grafting
CHF	Congestive heart failure
CI	Confidence interval
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
COPD	Chronic obstructive pulmonary disease
CV	Cardiovascular
DPP-4	Dipeptidyl peptidase-4
eGFR	Estimated glomerular filtration rate
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
ESRD	End-stage renal disease
GP	Grace period
HbA1c	Glycated hemoglobin A1c
HCRU	Healthcare resource utilization
HR	Hazard ratio
ICD-9	International Classification of Diseases, 9th revision
ICD-10	International Classification of Diseases, 10th revision
ITT	Intention to treat
KDIGO	Kidney Disease Improving Global Outcomes
MA	Marketing authorization
MACE	Major adverse cardiovascular event
MI	Myocardial infarction
NT	N-terminal
PS	Propensity score
PTCA	Percutaneous transluminal coronary angioplasty
SAP	Statistical analysis plan
SCr	Serum creatinine
SD	Standard deviation

SGLT-2	Sodium-glucose cotransporter-2
STD.	Standardized difference
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus

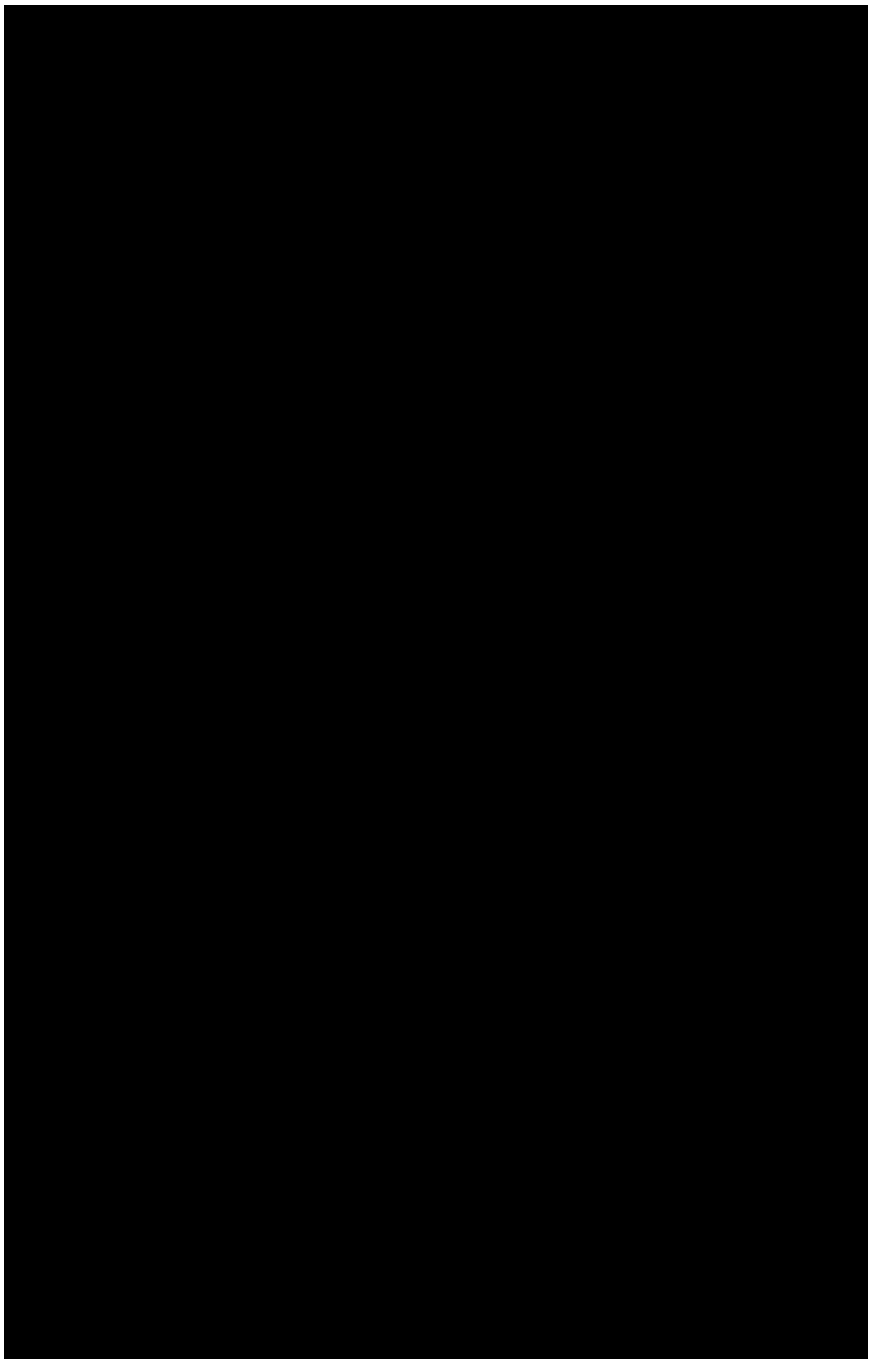
3. RESPONSIBLE PARTIES

Global Principal Investigators

[redacted]

Local Principal Investigators and Participating Institutions

[redacted]



Sponsor
Boehringer Ingelheim International GmbH

In the master protocol, the study investigators at [redacted] share responsibility with Boehringer Ingelheim International GmbH (BI) for the design of this multi-country study. The investigators are responsible for conducting the multi-country study in a manner that meets regulatory standards, conducting analyses, and preparing scientific reports. The multi-country study shall be conducted as described in this approved master protocol. The authors will not develop or implement any deviation or change to the master protocol without prior review by BI.

The financial sponsor of this study is BI. The sponsor is responsible for ensuring the progress of the study. BI is also responsible for communicating with regulatory agencies about the master protocol, the progress of the study, and study results.

4. ABSTRACT

Name of company: Boehringer Ingelheim			
Name of the finished medicinal product: Jardiance®			
Name of active ingredient: Empagliflozin (A10BK03)			
Protocol date: 29 November 2018	**Study number:** 1245.195	**Version/Revision:** 2.0	**Version/Revision date:** 11 October 2019
Title of study:	Multi-country non-interventional study on the effectiveness and safety of Empagliflozin in adult patients with type 2 diabetes in Europe and Asia		
Rationale and background:	Type 2 diabetes mellitus (T2DM) is associated with high cardiovascular morbidity and mortality. Certain novel pharmacological treatments of T2DM, such as sodium-glucose cotransporter-2 (SGLT-2) inhibitors, have shown protective effects on cardiovascular (CV) events, CV mortality and all-cause mortality, in addition to glucose-lowering effects. The beneficial effects on CV events and mortality, as well as on renal events, have been observed in clinical trials and in observational studies. However, no observational study has focused specifically on CV events and mortality associated with empagliflozin use in Europe or Asia. Previous studies investigating safety outcomes have shown diverse results. Further studies are needed to confirm the association between use of empagliflozin or any SGLT-2 inhibitor and effectiveness and safety events in patient populations with T2DM representing real-world practice.		
Research question and objectives:	The overall objective of this study is to examine effectiveness, safety, health care resource utilization (HCRU), and cost of care outcomes associated with the use of empagliflozin or any SGLT-2 inhibitors, compared with the use of dipeptidyl peptidase-4 (DPP-4) inhibitors, among patients with T2DM. The primary objectives related to effectiveness are to examine the risk of hospitalization for heart failure, all-cause mortality and two composite outcomes: one including hospitalization for heart failure and all-cause mortality and another including myocardial infarction (MI), stroke, and all-cause mortality. The secondary objectives related to cardiovascular effectiveness are to examine the risk of cardiovascular mortality, coronary revascularization procedures, and two composite outcomes: one including hospitalization for heart failure and cardiovascular mortality and another including MI, stroke, and cardiovascular mortality (i.e. 3-point major adverse cardiovascular events [MACE]). Further secondary objectives related to renal effectiveness are to		

Name of company: Boehringer Ingelheim			
Name of the finished medicinal product: Jardiance®			
Name of active ingredient: Empagliflozin (A10BK03)			
Protocol date: 29 November 2018	**Study number:** 1245.195	**Version/Revision:** 2.0	**Version/Revision date:** 11 October 2019
	examine the risk of end-stage renal disease (ESRD), estimated glomerular filtration rate (eGFR) decline, progression to micro- or macroalbuminuria, and a composite outcome, including an eGFR decline and a progression to micro- or macroalbuminuria. The secondary objectives related to safety are to compare the risk of bone fractures, diabetic ketoacidosis, severe hypoglycemia, lower-limb amputation, and acute kidney injury requiring dialysis. Further secondary objectives are to examine HCRU and cost of care. The objectives related to effectiveness will be investigated by comparing patients initiating treatment with empagliflozin to patients initiating treatment with any DPP-4 inhibitor, as well as patients initiating treatment with any SGLT-2 inhibitor and patients initiating treatment with any DPP-4 inhibitor. The secondary objectives on safety outcomes, HCRU, and cost of care will be investigated for empagliflozin and DPP-4 inhibitor exposure only.		
Study design:	This will be a non-interventional, multi-country cohort study using secondary data sources in European and Asian countries. This master protocol is an overarching protocol for the multi-country study. Individual-level anonymous datasets will be analyzed in the included countries, using localized protocols adapted from this master protocol. The analyses of effectiveness outcomes will include exposure to empagliflozin and SGLT-2 inhibitors, on the other hand, safety outcomes, HCRU, and cost of care will be investigated for empagliflozin only. The comparator drugs for all outcomes are DPP-4 inhibitors. To meet the objectives of this multi-country study, aggregate-level results obtained from the countries will be combined: effectiveness and safety results will be pooled in a meta-analysis, while HCRU and cost of care results will be presented descriptively.		
Population:	This study will include adults (≥18 years) with T2DM in Denmark, Finland, Germany, Israel, Japan, Norway, South Korea, Spain, Sweden, Taiwan, and the United Kingdom. The study population will consist of three sub-cohorts: patients initiating empagliflozin use, patients initiating		

Name of company: Boehringer Ingelheim			
Name of the finished medicinal product: Jardiance®			
Name of active ingredient: Empagliflozin (A10BK03)			
Protocol date: 29 November 2018	**Study number:** 1245.195	**Version/Revision:** 2.0	**Version/Revision date:** 11 October 2019
	any SGLT-2 inhibitor use, and patients initiating any DPP-4 inhibitor use.		
Variables:	The main exposures in this multi-country study are incident use of empagliflozin (Anatomical Therapeutic Chemical [ATC]: A10BK03, A10BD20) and incident use of any SGLT-2 inhibitor (A10BK). The comparator will be incident use of any DPP-4 inhibitor (A10BH). Incident use of empagliflozin and any SGLT-2 inhibitor will be defined as the first dispensation or any other record of the drug use after the drugs received marketing authorization, and not having any dispensation or any other record of the drug use of any SGLT-2 inhibitor or DPP-4 inhibitor during the preceding 12 months. Incident use of any DPP-4 inhibitor during the corresponding study period of empagliflozin or any SGLT-2 inhibitor will be defined as first dispensation or any other record of the drug use after 12 months with no dispensation or any other record of any DPP-4 inhibitor or SGLT-2 inhibitor use. Exposure periods will be defined based on available data on dispensations or any other records of the drug use in each data source. A grace period of 100% will be utilized to account for administration uncertainty. Variables related to the primary and secondary effectiveness outcomes will be cardiovascular events (MI, stroke, hospitalization for heart failure, coronary revascularization procedure), all-cause and cardiovascular mortality, diagnosis of ESRD, serum creatinine measurements, and albuminuria measurements. Additionally, diagnoses associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters, and dispensation or any other record of the high-ceiling diuretics (loop diuretics) will be used for secondary definition of heart failure. Also, the composite outcomes will be based on these variables. Variables related to the safety outcomes will be diagnoses of bone fractures, diabetic ketoacidosis, severe hypoglycemia, lower-limb amputation, and acute kidney injury requiring dialysis. Other secondary outcomes related to HCRU will be the use of inpatient or outpatient services and drug use. The cost of care will be defined as the costs associated with the use of these resources.		

Name of company: Boehringer Ingelheim			
Name of the finished medicinal product: Jardiance®			
Name of active ingredient: Empagliflozin (A10BK03)			
Protocol date: 29 November 2018	**Study number:** 1245.195	**Version/Revision:** 2.0	**Version/Revision date:** 11 October 2019

	Other variables collected from the data sources, according to data availability, will include covariates at baseline on sociodemographic characteristics, related to lifestyle, diabetes complications, other comorbidities, laboratory values, prior/concomitant use of other antidiabetic drugs, prior use of other drugs, healthcare resource utilization covariates, and cost covariates.
Data sources:	This study will utilize nationwide healthcare registers, regional quality registers, regional high-quality medical health records, and other health claims data available in Denmark, Finland, Germany, Israel, Japan, Norway, South Korea, Spain, Sweden, Taiwan, and United Kingdom.
Study size:	With a baseline outcome rate of 1.5% per year (e.g. heart failure), the analyses are sufficiently powered (>80%) to detect 30% decrease in the baseline rate with 20,000 patients per sub-cohort, i.e., 40,000 patients in total, and 0.5 years of follow-up per patient. The anticipated number of patients varies per country and in total is expected to be over 150,000.
Data analysis:	In the analyses performed in each country, propensity score (PS) matching will be used to reduce confounding in the comparative analyses across study sub-cohorts. Pairwise PS models between i) empagliflozin sub-cohort and DPP-4 sub-cohort and ii) SGLT-2 sub-cohort and DPP-4 sub-cohort will be estimated using logistic regression including appropriate covariates. Standardized differences will be used to assess the success of matching, and variables with standardized difference 10% or greater will also be adjusted for in the outcome analyses. For all included patients, descriptive statistics will be generated separately in each sub-cohort. Continuous covariates will be described by the mean, standard deviation, median, $25^{th,}$ and 75^{th} percentiles, minimum and maximum. Categorical covariates and continuous covariates that are also categorized will be described by proportion and frequency in each category. The primary and secondary effectiveness outcomes, as well as the safety outcomes, will be analyzed and compared across sub-cohorts by incidence rates, cumulative-incidence plots, and Cox proportional hazards models.

Name of company: Boehringer Ingelheim			
Name of the finished medicinal product: Jardiance®			
Name of active ingredient: Empagliflozin (A10BK03)			
Protocol date: 29 November 2018	**Study number:** 1245.195	**Version/Revision:** 2.0	**Version/Revision date:** 11 October 2019
	However, the change of eGFR value in time will be described by moving averages and compared across sub-cohorts by linear regression models. For the secondary outcomes on HCRU and cost of care, the number of outpatient and inpatient healthcare visits, dispensations or other records of drug use, and amount of costs will be determined during the follow-up (per years of follow-up, and per member per month). Comparison across sub-cohorts will be done using generalized linear regression models. [redacted] In analyses combining the results from each country, meta-analyses will be performed for the primary and secondary effectiveness outcomes as well as safety outcomes to combine individual country-level results. Subgroups of countries (e.g. Nordic, European or Asian countries) will also be included in separate meta-analyses. Country-level results on HCRU and cost of care will be presented descriptively.		
Milestones:	In this multi-country study, the start of data collection is in Q1 2019, and the end of data collection in Q1 2021. The first interim data report with combined results from the selected countries is planned to be available in Q4 2019. At least one additional interim data report and analyses will be conducted as needed before completing the final report with all study results from the countries in Q1 2021.		

5. AMENDMENTS AND UPDATES

Amendment Number	Date	Section of the study protocol	Amendment or update	Reason
1	13 September 2019	Multiple	New renal and safety outcomes included	To demonstrate an association of the study drug and the relevant safety outcomes.
		Multiple	Update of the index date	To align with the US study protocol.
		Multiple	Update of the inclusion/exclusion criteria	To improve the precision and comprehensiveness of the study group selection.
		Multiple	Safety outcomes and, healthcare resource utilization and cost analyses will be performed exclusively for empagliflozin as the exposure group and DPP-4 inhibitors as the comparison group	To have drug-specific safety outcomes and healthcare utilization and costs
		Multiple	Other minor amendments	To implement updates holistically.

6. MILESTONES

The following table estimates timelines for the milestones of this multi-country study, combining results from each country; country-level timelines for the country-specific studies will be detailed in localized protocols.

Milestone	Planned Date
Start of data collection[1]	Q1 2019
End of data collection[2]	Q1 2021
Interim data reports[3]	Q4 2019 and Q1 2021
Registration in the EU PAS register	Q4 2018
Final report of study results[4]	Q1 2021

[1] In any country; [2] In the last country; [3] At least two interim reports will be performed, including combined results of the countries for which results were available between Q3 2019 and Q4 2020; [4] Including combined results from all countries.

7. RATIONALE AND BACKGROUND

The number of patients with type 2 diabetes mellitus (T2DM) is increasing all over the world [R15-1988]. In Europe, the number of patients with diabetes is estimated to be 55 million, and it is estimated to increase to 67 million by 2030. The corresponding increase in the Western Pacific is from 77 million to 113 million. T2DM constitutes 90% of these numbers.

Cardiovascular (CV) diseases are considerable comorbidities among patients with T2DM [R18-3510]. Up to one-third of patients with T2DM experience adverse CV events, such as atherosclerosis, coronary heart disease, heart failure, angina, myocardial infarction (MI), and stroke. Further, CV events are the cause of death for 20% of patients with T2DM [R18-3508]. The high CV morbidity also contributes to increasing healthcare costs [R18-3509]. A systematic review found that in comparison to T2DM patients who do not have CV diseases, CV events contribute to a 112% increase in healthcare costs, with stroke being the event with the highest increase in costs among patients with T2DM [R18-3509]. CV events contribute to up to 50% of healthcare costs in the treatment of T2DM. The main cost drivers are hospitalization, drug, and outpatient care costs.

Sodium-glucose cotransporter-2 (SGLT-2) inhibitors represent one of the most recent pharmacological treatment options of T2DM [P17-01701]. These drugs have been available since 2012. The globally available SGLT-2 inhibitors include canagliflozin, dapagliflozin, empagliflozin, and ertugliflozin. Additional SGLT-2 inhibitors (e.g. ipragliflozin, tofogliflozin) are available also in Asia [P17-01701]. Generally, SGLT-2 inhibitors are not the first-line treatment, instead, they are often used mainly later in the disease course of

T2DM [P18-10828]. In addition to the glucose-lowering effects, these drugs also have beneficial effects on weight, blood pressure, and potentially also lipids [P17-01701].

In large CV outcome trials, empagliflozin and canagliflozin use decreased the risk of adverse CV events among patients with T2DM compared to the placebo group [P18-04466]. Both drugs showed to reduce the risk of CV mortality, non-fatal MI, and non-fatal stroke (a composite outcome referred to as 3-point major adverse cardiovascular events [MACE]) by 14%. Further, empagliflozin use reduced the risk of MI by 13%, while the risk of CV or all-cause mortality and hospitalization for heart failure was reduced by 32-38%. Also, renal outcomes have been investigated: dapagliflozin and canagliflozin decreased the risk of eGFR decline, and canagliflozin also decreased the risk of progression of albuminuria [P18-02804, P19-04731]. However, the patients included in these trials had high CV morbidity and thus the results might not be generalizable to with a broader population of patients with T2DM including those with fewer CV risk factors.

Observational studies have found an association between the use of dapagliflozin or any SGLT-2 inhibitor and a decreased risk of CV events and mortality compared to other glucose-lowering drugs among broader populations of T2DM patients [R18-3505, P17-05895, P18-06283, P17-08964]. Dapagliflozin/any SGLT-2 inhibitor use has been associated with a 40-50% decrease in the risk of all-cause mortality [R18-3505] and a 30-40% decrease in the risk of heart failure [R18-3505, P17-05895, P18-06283, P17-08964]. Dapagliflozin/any SGLT-2 inhibitor use has also been associated with a lower risk of 3-point MACE. The previous results on the risk of MI, stroke, and CV mortality associated with dapagliflozin/any SGLT-2 inhibitor use have been directionally consistent, suggesting decreased risks, but the results of the studies have been inconsistent in detecting significant associations [R18-3505, P18-06283, P17-08964]. The proportion of empagliflozin use in the studies investigating any SGLT-2 inhibitor has been <10% [P17-05895, P18-06283, P17-08964] and, therefore, the results may not be generalizable to empagliflozin users. Preliminary results from a study being performed in the United States of America found that empagliflozin was associated with a 49% decrease in heart failure risk. To our best knowledge, no study has investigated the CV benefits associated with empagliflozin use in a real-world setting in Europe or in Asia.

Previous studies, including clinical trials and observational studies, have shown diverse results regarding empagliflozin, dapagliflozin, canagliflozin or any SGLT-2 inhibitor use and the risk of safety events among patients with T2DM. Risk of diabetic ketoacidosis is similar between empagliflozin or canagliflozin users and placebo group [P18-02804, P18-01920]. However, diabetic ketoacidosis was more frequent among dapagliflozin users in comparison to the placebo group (hazard ratio [HR] 2.18; 95% CI 1.10 to 4.30) [P19-04731]. Moreover, in a register-based cohort study, the associated risk of diabetic ketoacidosis was increased (HR, 2.14; CI, 1.01 to 4.52) among patients with T2DM using SGLT2 inhibitors, compared with glucagon-like peptide 1 receptor agonist users [P18-10986]. Further, dapagliflozin use

has also been found to decrease the risk of acute kidney injuries in a clinical trial (HR, 0.69; 95% CI, 0.55 to 0.87) [P19-04731] while no difference was observed between users of canagliflozin and placebo [P18-02804]. Further, no association was found in an observational study that compared SGLT-2 inhibitor use with glucagon-like peptide 1 receptor agonist drug use (HR, 0.69; CI 0.45 to 1.05) [P18-10986].

The frequency of fractures was similar between users of empagliflozin and placebo (3.9% vs. 3.8%, respectively) [P18-01920] and dapagliflozin and placebo (5.3% vs. 5.1%, respectively) [P19-04731]. On the contrary, the frequency of fractures was higher among canagliflozin users in comparison to placebo (HR 1.26; 95% CI 1.04 to 1.52) [P18-02804]. SGLT-2 inhibitor use was not associated with an increased risk of bone fracture at a register-based cohort study, compared with glucagon-like peptide 1 receptor agonist users (HR, 1.11; CI 0.93 to 1.33) [P18-10986].

There was no difference in the frequency of amputations between dapagliflozin and placebo group (1.4% vs. 1.3%) [P19-04731]. However, canagliflozin use has shown a greater risk of amputations (HR, 1.97; 95% CI, 1.41 to 2.75) compared with placebo [P18-02804]. In addition, a register-based cohort study found an association with an increased risk of lower limb amputation (HR, 2.32; CI, 1.37 to 3.91) among patients with T2DM using SGLT2 inhibitors, compared with glucagon-like peptide 1 receptor agonist users [P18-10986].

Further real-world evidence studies are needed to examine the association between empagliflozin use and adverse CV events as well as mortality. It is especially important to account for confounding by indication in observational studies. The previous observational studies investigating the association between SGLT-2 inhibitors and adverse CV events and mortality compared new users of SGLT-2 inhibitors with new users of other glucose-lowering drugs. This may have resulted in residual confounding and, thus, biased results, since SGLT-2 inhibitors are not the first-line treatment of T2DM. Immortal time bias, resulting from a comparison of first-line treatment and drugs used mainly later in the disease course, may have contributed to the major differences between SGLT-2 inhibitors and other glucose-lowering drugs. Further, comparison with other glucose-lowering drugs means that patients with the first-line treatment of T2DM, but also patients with a more severe stage of the disease, might have been included in the comparison group. Alternatively, SGLT-2 inhibitor use can be compared with drugs that are used in the treatment of a similar line in the treatment pathway of T2DM [P18-10828]. For example, dapagliflozin use has been compared with dipeptidyl peptidase-4 (DPP-4) inhibitor use [R18-3505], and the study found that dapagliflozin use was associated with a decreased risk of hospitalization for heart failure, all-cause mortality, and 3-point MACE, but not of MI, stroke, and CV mortality. Further, the specific inclusion and exclusion criteria in the clinical trials [P18-02804, P19-04731, P18-01920] limit the generalizability of the results and more studies are needed to examine the association between empagliflozin/any SGLT-2 inhibitor use and renal events as well as safety outcomes.

This study will be an expansion of the non-interventional study being conducted in the United States of America (EMPRISE, 1245.92) with additional renal effectiveness outcomes. New users of empagliflozin and new users of any SGLT-2 inhibitor will be compared with new users of DPP-4 inhibitors in several European and Asian countries. The effectiveness outcomes of interest include non-fatal and fatal CV events, all-cause and CV mortality. Further effectiveness outcomes are coronary revascularization procedures, end-stage renal disease (ESRD), estimated glomerular filtration rate (eGFR) decline, a progression from normoalbuminuria to micro- or macroalbuminuria, and derived cardiovascular and renal composite outcomes. In addition, safety outcomes, i.e. risk of bone fractures, diabetic ketoacidosis, severe hypoglycemia, lower-limb amputation, and acute kidney injury requiring dialysis, will be examined in this study. Furthermore, healthcare resource utilization (HCRU) and cost of care outcomes will be investigated.

8. RESEARCH QUESTION AND OBJECTIVES

The overall objective of this study is to examine the effectiveness, safety, HCRU, and cost of care outcomes in patients with T2DM initiating empagliflozin or any SGLT-2 inhibitor use in comparison to patients with T2DM initiating any DPP-4 inhibitor use after propensity score (PS) matching.

The specific objectives of this study are related to effectiveness (primary objectives 1-4, secondary objectives 1-8), safety (secondary objectives 9-13), HCRU (secondary objective 14), and cost of care (secondary objective 15). The comparisons for objectives in sections 8.1-8.3 will be performed between patients initiating empagliflozin or any SGLT-2 inhibitor use and patients initiating any DPP-4 inhibitor use. The comparisons in sections 8.4-8.5 will be performed between patients initiating empagliflozin use and patients initiating any DPP-4 inhibitor use only.

8.1 PRIMARY OBJECTIVES: EFFECTIVENESS

Primary objective 1: Compare the risk of hospitalization for heart failure.

Primary objective 2: Compare the risk of all-cause mortality.

Primary objective 3: Compare the risk of a composite outcome including hospitalization for heart failure and all-cause mortality.

Primary objective 4: Compare the risk of a composite outcome including myocardial infarction (MI), stroke, and all-cause mortality, as well as its individual components.

8.2 SECONDARY OBJECTIVES ON CARDIOVASCULAR EFFECTIVENESS

Secondary objective 1: Compare cardiovascular (CV) mortality.

Secondary objective 2: Compare the risk of a composite outcome including hospitalization for heart failure and cardiovascular (CV) mortality.

Secondary objective 3: Compare the risk of 3-point major adverse cardiovascular (CV) events (MACE), defined as a composite outcome including myocardial infarction (MI), stroke, and cardiovascular (CV) mortality

Secondary objective 4: Compare the risk of coronary revascularization procedure.

8.3 SECONDARY OBJECTIVES ON RENAL EFFECTIVENESS

Secondary objective 5: Compare the risk of end-stage renal disease (ESRD).

Secondary objective 6: Compare the risk of estimated glomerular filtration rate (eGFR) decline.

Secondary objective 7: Compare the risk of progression from normoalbuminuria to micro- or macroalbuminuria.

Secondary objective 8: Compare the risk of composite outcome including eGFR decline and progression from normoalbuminuria to micro- or macroalbuminuria.

8.4 SECONDARY OBJECTIVES ON SAFETY

Secondary objective 9: Compare the risk of bone fractures.

Secondary objective 10: Compare the risk of diabetic ketoacidosis.

Secondary objective 11: Compare the risk of severe hypoglycemia.

Secondary objective 12: Compare the risk of lower-limb amputation.

Secondary objective 13: Compare the risk of acute kidney injury requiring dialysis.

8.5 SECONDARY OBJECTIVES ON HEALTH CARE RESOURCE UTILIZATION (HCRU) AND COST OF CARE

Secondary objective 14: Compare healthcare resource utilization (HCRU) during the follow-up period.

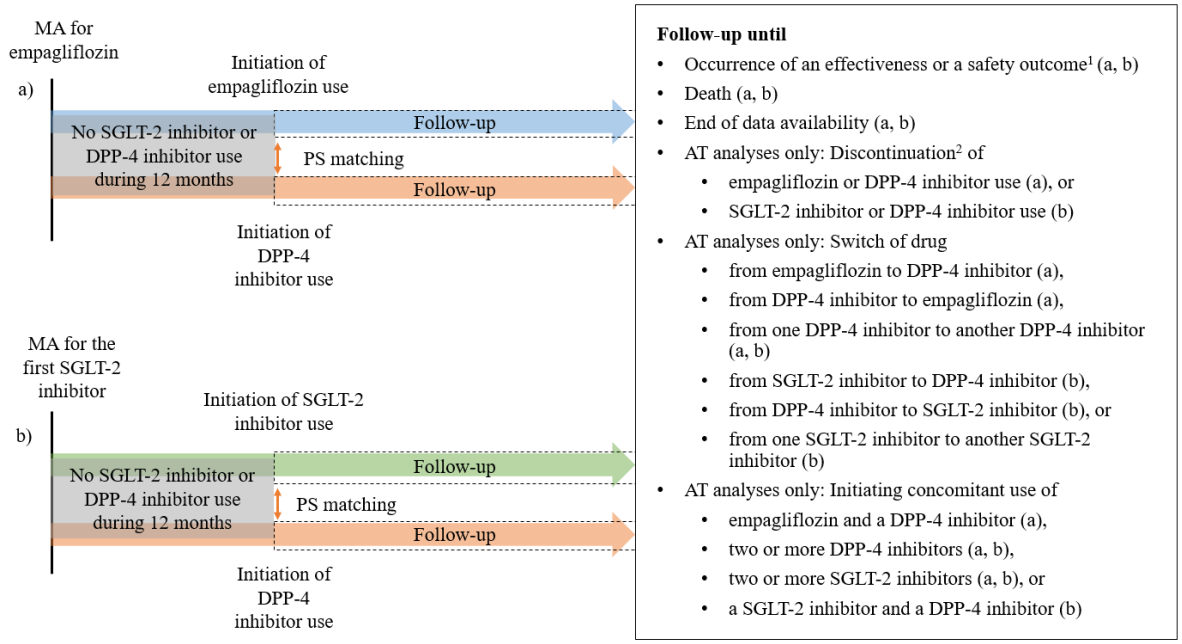
Secondary objective 15: Compare the cost of care outcomes during the follow-up period.

9. RESEARCH METHODS

9.1 STUDY DESIGN

This will be a non-interventional, multi-country cohort study using existing data and including adults (≥18 years) with a diagnosis of T2DM. All data will be obtained from electronically recorded longitudinal secondary data sources, separately in each included country. In each country, patients initiating treatment with empagliflozin or any other SGLT-2 inhibitor will be compared with PS-matched patients initiating treatment with any DPP-4 inhibitor (Figure 1; specific comparisons presented in section 8).

Figure 1 Overview of the study periods and comparison pairs in the analyses: a) Analyses comparing empagliflozin use with DPP-4 inhibitor use; b) Analyses comparing SGLT-2 inhibitor use with DPP-4 inhibitor use.



AT=as treated; DPP-4=dipeptidyl peptidase-4; MA=marketing authorization; PS=propensity score; SGLT-2=sodium-glucose cotransporter-2.
1 In analyses investigating effectiveness or safety outcomes, the occurrence of the outcome in question will be observed until the end of the follow-up (e.g. while investigating hospitalization for heart failure, the follow-up will not end at the occurrence of a stroke).
2 In analyses investigating safety outcomes, an exposure risk window will be accounted for after discontinuation.

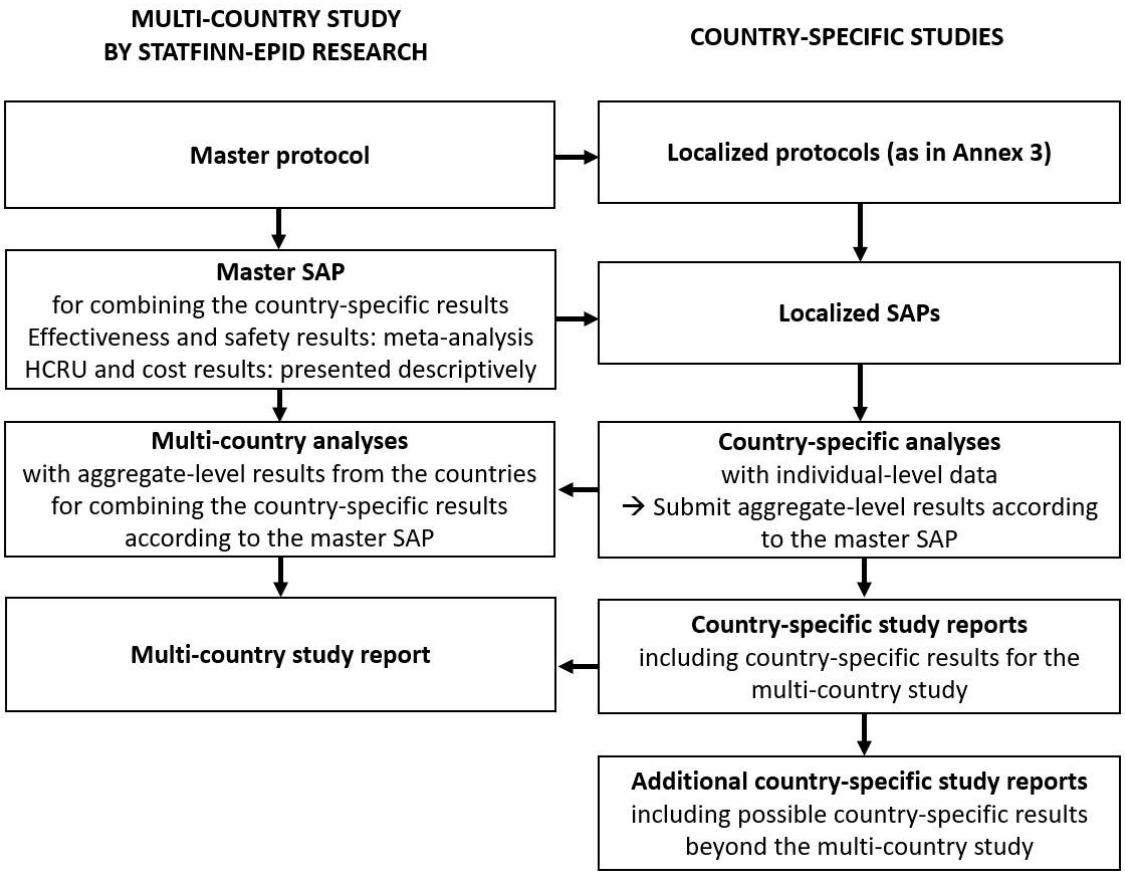
By comparing new users of empagliflozin or new users of any SGLT-2 inhibitor with an active comparator, i.e. new users of any DPP-4 inhibitor, confounding by indication can be decreased as the compared sub-cohorts have similar indications [R18-3506]. Further, the

inclusion of eligible patients from multiple countries increases the power of the analyses and generalizability of the results.

9.1.1 Country-specific studies as components of this multi-country study

This master protocol is an overarching protocol for the multi-country study (Figure 2). Country-specific adaptations of this master protocol will be described in localized protocols (Figure 2), written separately in each included country. As illustrated in Figure 2, individual-level data will be analyzed as part of the country-specific studies. To meet the objectives of this multi-country study, aggregate-level results obtained from the countries will be combined in the multi-country study: effectiveness and safety results will be pooled in a meta-analysis, while HCRU and cost results will be presented descriptively.

Figure 2 Overview of the country-specific studies as components of the multi-country study.



HCRU=health care resource utilization; SAP=statistical analysis plan.

9.2 SETTING

9.2.1 Overview

The study will be performed at least in 11 countries: Denmark, Finland, Germany, Israel, Japan, Norway, South Korea, Spain, Sweden, Taiwan, and the United Kingdom. Additional countries may be included during the course of this study. The dates of marketing authorizations (MA) of empagliflozin, the first available SGLT-2 inhibitor, and the first available DPP-4 inhibitor in the participating countries are presented in **Table 1**. If additional countries are included in the study, the corresponding MA dates will be specified in the localized protocol.

Table 1 Marketing authorizations of empagliflozin, first available SGLT-2 inhibitor, and first available DPP-4 inhibitor in each study country.

Country	Date of empagliflozin MA	Date of first SGLT-2 inhibitor MA	Date of first DPP-4 inhibitor MA
Denmark	May 2014	November 2012	March 2007
Finland	May 2014	November 2012	March 2007
Germany	May 2014	November 2012	March 2007
Israel	August 2015	August 2014	June 2008
Japan	December 2014	January 2014	October 2009
Norway	May 2014	November 2012	March 2007
South Korea	August 2014	November 2013	September 2007
Spain	May 2014	November 2012	March 2007
Sweden	May 2014	November 2012	March 2007
Taiwan	May 2016	May 2016	March 2009
The United Kingdom	August 2014	November 2012	April 2007

DPP-4=dipeptidyl peptidase-4; MA=marketing authorization; SGLT-2=sodium-glucose cotransporter-2.

New users of empagliflozin will be identified between the MA date of empagliflozin (May 2014 onwards) and the end of data availability (not later than December 2018) in each country (Table 1, Figure 1). A similar approach will be utilized in defining new users of any SGLT-2 inhibitors, and new users will be identified between the MA date of first available SGLT-2 inhibitor (November 2012 onwards) and end of data availability. New users of any DPP-4 inhibitor will be identified during the corresponding study periods of empagliflozin and any SGLT-2 inhibitor (Table 2, Figure 1).

Table 2 Definitions of investigated sub-cohorts of patients with type 2 diabetes mellitus included in this study.

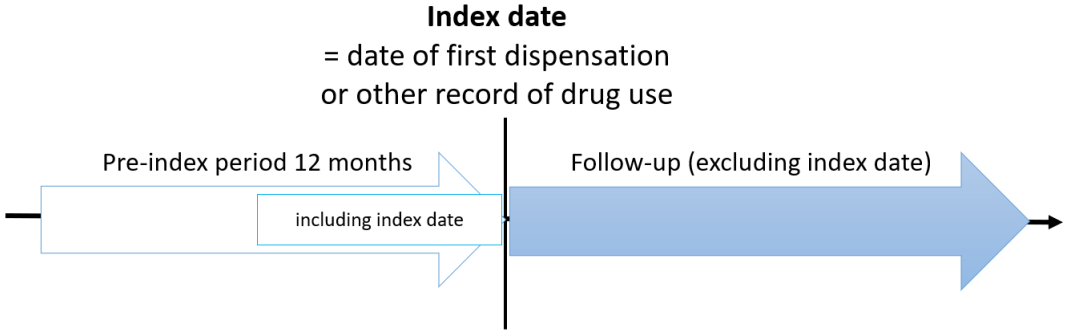
Sub-cohort	Definition	The index date
Empagliflozin sub-cohort	Patients who initiate[1] empagliflozin use during the study period and meet criteria in Figure 1 and Figure 4 in section 9.2.2.	The first dispensation date or the first date of any other record of the empagliflozin use after the MA date[2] of empagliflozin
Any SGLT-2 inhibitor sub-cohort	Patients who initiate[1] any SGLT-2 inhibitor use during the study period and meet criteria in Figure 1 and Figure 4 in section 9.2.2.	The first dispensation date or the first date of any other record of any SGLT-2 inhibitor use after the MA date[3] of first SGLT-2 inhibitor
Any DPP-4 inhibitor sub-cohort	Patients who initiate[1] any DPP-4 inhibitor use during the study period and meet criteria in Figure 1 and Figure 4 in section 9.2.2.	The first dispensation date or the first date of any other record of any DPP-4 inhibitor use during the study period[4].

[1] New use will be defined as having no dispensations or any other record of drug use for any SGLT-2 inhibitor or any DPP-4 inhibitor during the previous 12 months; [2] Range from May 2014 to May 2016; [3] Range from November 2012 to May 2016; [4] The study period starts at empagliflozin MA date in analyses with empagliflozin and at first SGLT-2 inhibitor MA date in analyses with any SGLT-2 inhibitors.

T2DM=type 2 diabetes mellitus; DPP-4=dipeptidyl peptidase-4; MA=marketing authorization; SGLT-2=sodium glucose cotransporter 2.

New use of empagliflozin, any SGLT-2 inhibitor, and any DPP-4 inhibitor will be defined as not having any dispensation or any other record of the drug use of these drugs during the preceding 12 months. The index date will be defined as the date of the first dispensation or any other record of the use of the sub-cohort-defining drug, forming three sub-cohorts: empagliflozin sub-cohort, any SGLT-2 inhibitor sub-cohort, and any DPP-4 inhibitor sub-cohort (Table 2). Data on comorbidities, other drug use, and clinical variables will be obtained from the pre-index period, i.e. 12 months before (and including the date of) the index date (Figure 3).

Figure 3 Definitions of the pre-index period and follow-up in relation to the index date.



9.2.2 Study population

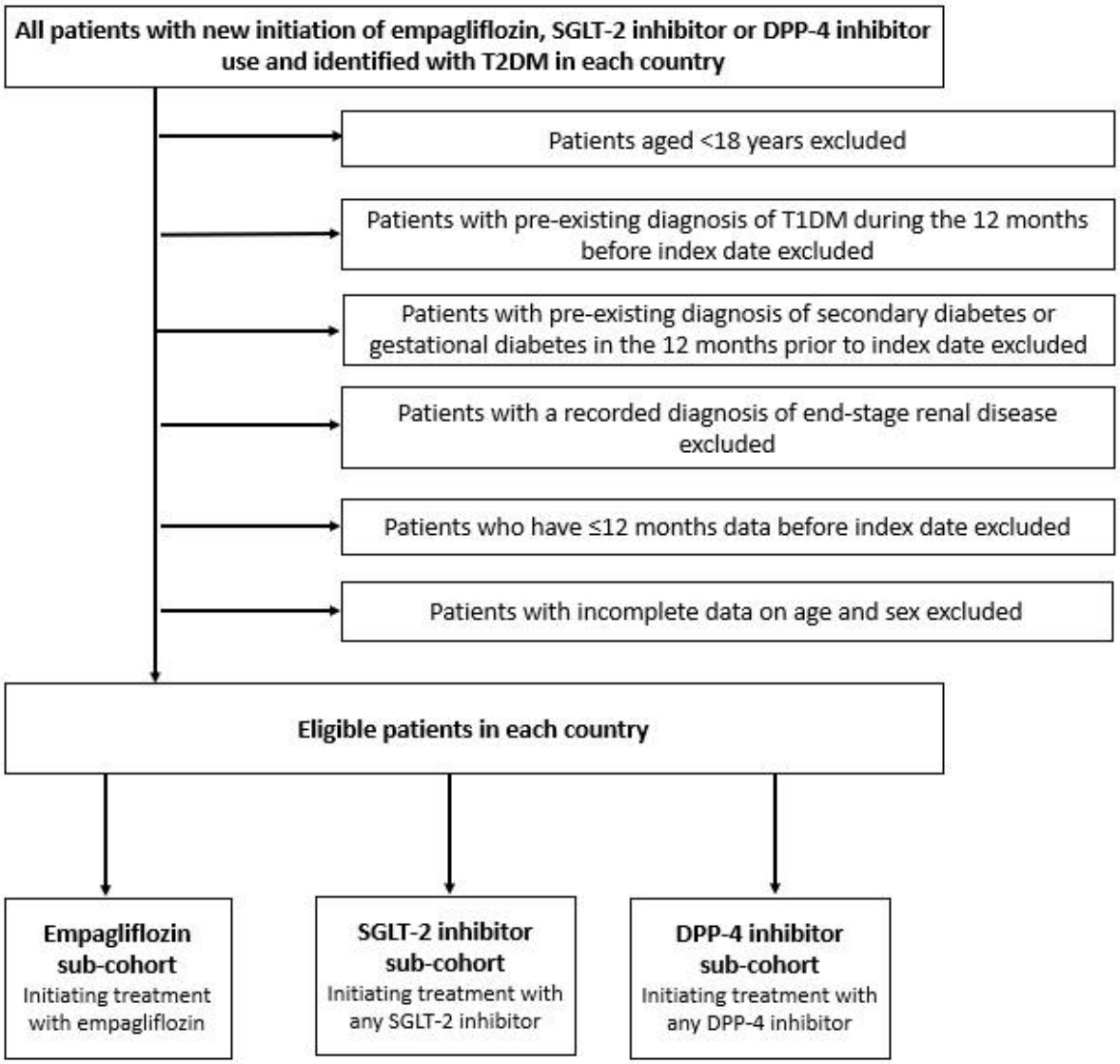
To study real-world treatment patterns, the study population will include a broad spectrum of users of empagliflozin, any SGLT-2 inhibitor, and any DPP-4 inhibitor. The following inclusion and exclusion criteria will be used (definitions detailed in ANNEX 5):
Inclusion criteria (Figure 4):
Dispensation or any other record of empagliflozin, any SGLT-2 inhibitor, or any DPP-4 inhibitor use during the study period (Figure 1)
No dispensation or any other record of any other SGLT-2 inhibitor or DPP-4 inhibitor[1] use during the 12 months preceding the index date, and
Having a diagnosis of T2DM before the index date, based on ICD-10 codes or other available data[2].
Exclusion criteria (Figure 4):
Aged <18 years on the first dispensation date or date of the first record of empagliflozin, any SGLT-2 inhibitor or any DPP-4 inhibitor use,
Pre-existing diagnosis of T1DM during the 12 months before the index date[2],

- → Pre-existing diagnosis of secondary diabetes or gestational diabetes in the 12 months prior to the index date[2],
- → Having a diagnosis of ESRD during the 12 months before the index date,
- → <12 months of available data before the index date, and/or no complete history of drug dispensations/other records of drug use during this period, and
- → Missing or ambiguous data on age or sex.

[1] Including possible SGLT-2 inhibitors or DPP-4 inhibitors that are marketed locally and are not study drugs of this multi-country study (Table 3). Patients are also excluded if they are at the index date dispensed both a SGLT-2 inhibitor and a DPP-4 inhibitor, or if records of use are available for both, in free or fixed-dose combination.

[2] The definition of T2DM, T1DM, and other types of diabetes, including the used diagnosis, drug or other data and the most appropriate time-window prior to the index date, will be detailed in the localized protocols according to validated or standard definitions used in the country and data source.

Figure 4 Flow chart



DPP-4=dipeptidyl peptidase-4; SGLT-2=sodium-glucose cotransporter-2; T1DM=type 1 diabetes mellitus; T2DM=type 2 diabetes mellitus.

These criteria ensure that only patients with T2DM, and not any other type of diabetes, are included in the study. In addition, the exclusion of patients with unknown drug use data due to, for example, nursing home stay, ensures that the included patients have sufficient exposure data available.

9.2.3 Study period and follow-up

The duration of the study period depends on the dates of the marketing authorizations of the study drugs (Table 1) and data availability in each country's data sources and shall be detailed in each localized protocol. For all PS-matched patients, the follow-up will begin on the index date and potentially continue for all as long as data are available in the used data sources, up to December 2018. However, if the follow-up ends (see below) at the same (index) date, the patient will not contribute to the outcome analyses, i.e., that patient will account for 0 person-years and 0 events, even if follow-up ended due to an outcome happening at the index date (e.g. death). The minimum follow-up time will be 1 day for each patient who contributes to the outcome analyses, and the follow-up time contributing to the analyses starts one day after the index date. Specifically, follow-up time will be 1 day for any patient whose follow-up ends at the day following the index date. The duration of follow-up will be computed as follows: (date of ending the follow-up – index date). The rationale for this decision to handle these expectedly rare events is that in several countries the timing of event occurrence can only be determined with a 1-day precision. Each patient will be included only once in each sub-cohort.

In the main analyses, the follow-up will end on the date of the first occurrence of any of the following events: occurrence of an effectiveness or a safety outcome (when studying outcomes related to effectiveness or safety), death, discontinuation of the initial drug, switch to any other study drug (empagliflozin, any SGLT-2 inhibitor, any DPP-4 inhibitor)[3], initiation of concomitant use of empagliflozin/SGLT-2 inhibitor and a DPP-4 inhibitor (either as free or fixed-dose combinations)[4] or two drugs within the same class (i.e. two SGLT-2 inhibitors or two DPP-4 inhibitors), or end of data availability (Figure 1). Depending on the data source, the end of data availability could result from, for example, moving to another region or admission to institutional care. When studying the HCRU and cost of care outcomes, the follow-up is not censored at the occurrence of an effectiveness or a safety outcome.

9.3 VARIABLES

The variable definitions presented in this master protocol are master definitions, which shall be adapted in each country and detailed in the localized protocols (ANNEX 3). The master definitions concerning diagnoses and procedures are in this master protocol provided using the International Classification of Diseases, both 9th (ICD-9) and 10th revision (ICD-10) codes. The master definitions concerning drugs are in this master protocol provided using the Anatomical Therapeutic Chemical (ATC) codes. Both ICD and ATC codes shall be adapted

[3] The follow-up time will be censored also at switch to SGLT-2 inhibitors or DPP-4 inhibitors that are marketed locally and are not study drugs in this multi-country study (Table 3).

[4] The follow-up time will be censored also at initiation of concomitant use of SGLT-2 inhibitors or DPP-4 inhibitors that are marketed locally and are not study drugs in this multi-country study (Table 3).

to locally used coding systems (or other definitions) and detailed in the localized protocols. The diagnoses in the localized protocols should be identified from recorded diagnoses, claims or other applicable secondary data. Furthermore, the used data sources shall be detailed in section 9.4, providing information on the type of data and healthcare setting where the data have been retrieved, including whether the drug use data are based on dispensations or other records (e.g. prescriptions) and whether primary care is included.

9.3.1 Exposures: study drugs

Main exposures: empagliflozin and any SGLT-2 inhibitor
The main exposures in this study are incident use of empagliflozin (ATC codes: A10BK03, A10BD20) and incident use of any SGLT-2 inhibitor. The globally available SGLT-2 inhibitors to be included in any SGLT-2 inhibitor sub-cohort are canagliflozin, dapagliflozin, empagliflozin, ertugliflozin, and ipragliflozin (Table 3). Fixed-dose combination drugs with a SGLT-2 inhibitor and another drug other than a DPP-4 inhibitor (comparator) will be included. Other available SGLT-2 inhibitors available in specific countries may be included as study drugs in the country-specific studies (ANNEX 3). The results of such country-specific analyses will not be included in the combined results of the multi-country study.

For the secondary objectives on safety, HCRU and costs (see sections 8.4 and 8.5), the main exposure of interest is exclusively empagliflozin, and thus any SGLT-2 inhibitor users will not be analyzed.

Comparator drugs: any DPP-4 inhibitor
Incident use of any DPP-4 inhibitor is the main comparator in the analyses. The globally available DPP-4 inhibitors to be included in any DPP-4 inhibitor sub-cohort are presented inTable 3. Fixed-dose combination drugs with a DPP-4 inhibitor and another drug other than a SGLT-2 inhibitor (main exposure) will be included. The DPP-4 inhibitors will be used as the active comparator sub-cohort because they are used in the treatment of the same disease stage of T2DM as SGLT-2 inhibitors [P15-09840]. Further, most DPP-4 inhibitors have not been associated with a risk of CV outcomes [P18-04466]. [redacted]
In localized protocols, alternative comparator drugs may be included. For example, DPP-4 inhibitors are used as a first-line treatment of T2DM in Japan and, consequently, alternative comparator drugs reflecting local treatment practices will have to be included in the country-specific analyses. The results of such country-specific analyses with alternative comparator drugs will not be included in the combined results of the multi-country study (ANNEX 3).

Table 3 Study drugs forming the sub-cohorts.

Sub-cohorts and study drugs	**ATC code**
Empagliflozin (including fixed-dose combinations with any other drug than DPP-4 inhibitor1)	
Empagliflozin	A10BK03 (A10BX12[2])
Empagliflozin and metformin	A10BD20
Any SGLT-2 inhibitor (including fixed-dose combinations with a SGLT-2 inhibitor and any other drug than DPP-4 inhibitor1)	
Dapagliflozin	A10BK01 (A10BX09[2])
Dapagliflozin and metformin	A10BD15
Empagliflozin	A10BK03 (A10BX12[2])
Empagliflozin and metformin	A10BD20
Canagliflozin	A10BK02 (A10BX11[2])
Canagliflozin and metformin	A10BD16
Ertugliflozin	A10BK04
Ertugliflozin and metformin	A10BD23
Ipragliflozin	A10BK05
Any DPP-4 inhibitor (including fixed-dose combinations with a DPP-4 inhibitor and any other drug than SGLT-2 inhibitor[1])	
Sitagliptin	A10BH01
Sitagliptin and metformin	A10BD07
Sitagliptin and pioglitazone	A10BD12
Sitagliptin and simvastatin	A10BH51
Vildagliptin	A10BH02
Vildagliptin and metformin	A10BD08
Saxagliptin	A10BH03
Saxagliptin and metformin	A10BD10
Alogliptin	A10BH04
Alogliptin and pioglitazone	A10BD09

Table 3 (cont'd) Study drugs forming the sub-cohorts.

Sub-cohorts and study drugs	**ATC code**
Any DPP-4 inhibitor (including fixed-dose combinations with a DPP-4 inhibitor and any other drug than SGLT-2 inhibitor1)	
Alogliptin and metformin	A10BD13
Linagliptin	A10BH05
Linagliptin and metformin	A10BD11
Gemigliptin	A10BH06
Gemigliptin and rosuvastatin	A10BH52
Gemigliptin and metformin	A10BD18
Evogliptin	A10BH07
Evogliptin and metformin	A10BD22

ATC=Anatomical Therapeutic Chemical; DPP-4=dipeptidyl peptidase-4; SGLT-2=sodium-glucose cotransporter-2
[1] Includes combinations with other classes of antidiabetic drugs.
[2] A previous version of the ATC code

Concomitant use of a SGLT-2 inhibitor and a DPP-4 inhibitor excluded
Concomitant use of a SGLT-2 inhibitor and a DPP-4 inhibitor will not be accounted for in this master protocol or in the combined analyses of the multi-country study. As described in section 9.2.2, the inclusion and exclusion criteria require that patients have no previous use of any SGLT-2 inhibitor or any DPP-4 inhibitor (also those marketed locally but not listed as study drugs in Table 3), in free or fixed-dose combination.

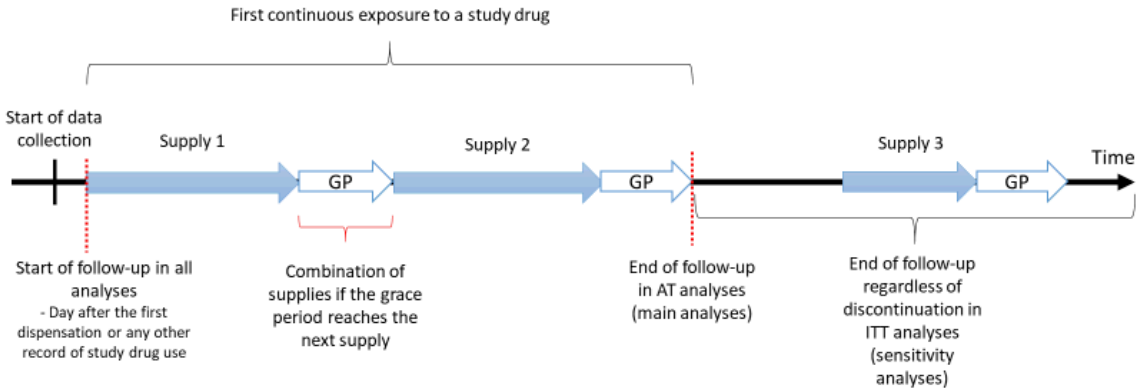
During the follow-up, patients are censored if they switch to any SGLT-2 inhibitor or DPP-4 inhibitor, or initiate a concomitant use of study drugs (free or fixed-dose combinations), as described in section 9.2.3 and in Figure 1.

However, concomitant use of a SGLT-2 inhibitor and a DPP-4 inhibitor may be detailed in the localized protocols to be included in the country-specific analyses; these results will not be included in the combined results of the multi-country study (ANNEX 3).

Definition of exposure periods
Exposure periods will be defined based on available data on drug dispensations or other records of the drug use in each country. If data on drug dispensations or prescriptions is not available in the data source, alternative definitions for exposure periods will be used, provided in ANNEX 4.

Drug use will be assumed to begin on the date of a dispensation or the date of any other record of the drug use. A supply will indicate the duration of exposure after a dispensation or any other record of the drug use (Figure 5). Supply will be defined for each dispensation or any other record of drug use based on available data on the daily dose, as detailed in the localized protocols. For example, the duration of the supply can be derived from the dispensed amount (total milligrams dispensed) and the daily dose (milligrams). If a daily dose is not readily available, the daily dose shall be estimated as appropriate in the local setting. If a subsequent supply occurs before the previous supply has finished, the start of the subsequent supply will be shifted. For example, if a patient receives the first supply for 60 days and refills at day 50 for another 60 days, the duration of the two supplies combined will be 50+10+60=120 days, after shifting the start of the second supply with 10 days for which the patient had supply in storage from the first supply. To avoid artificially long exposure periods, however, a subsequent supply can be shifted with a maximum of 14 days, which is considered a reasonable time for patients to refill prior to the end of their ongoing supply. Further, a subsequent supply cannot be shifted over the grace period (as defined below).

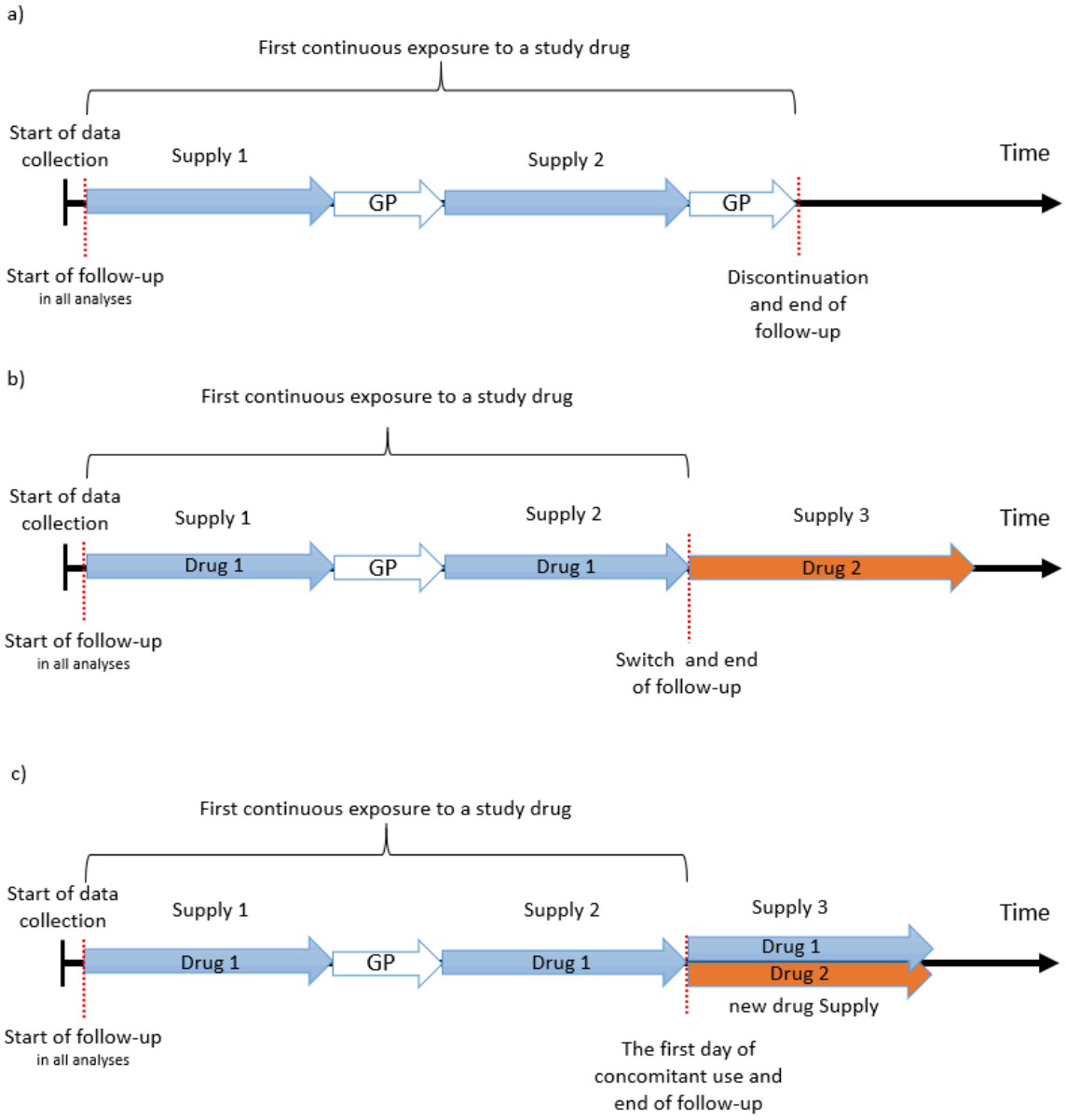
Figure 5 Illustration of combining exposure periods with grace periods and the differences between the AT and ITT approaches.



AT=as-treated; GP=grace period; ITT=intention to treat

A grace period of 100% of the duration of the most recent supply will be included in the exposure period to account for uncertainty related to actual drug use patterns. The grace period will be defined from the most recent supply. In the above example on two overlapping supplies with 60 days, totaling in 120 days, the grace period is 60 days (100% of the most recent supply). Overlapping supplies and grace periods will be combined as exposure periods (Figure 5).
The discontinuation will be defined as the date of ending the grace period (Figure 6). The switch of a drug is defined as replacing one study drug by another, and no grace period will be used (Figure 6). The concomitant use addresses to the simultaneous use of study drugs, and the first date of concomitant use will end the follow-up, i.e., no grace period will be used (Figure 6).

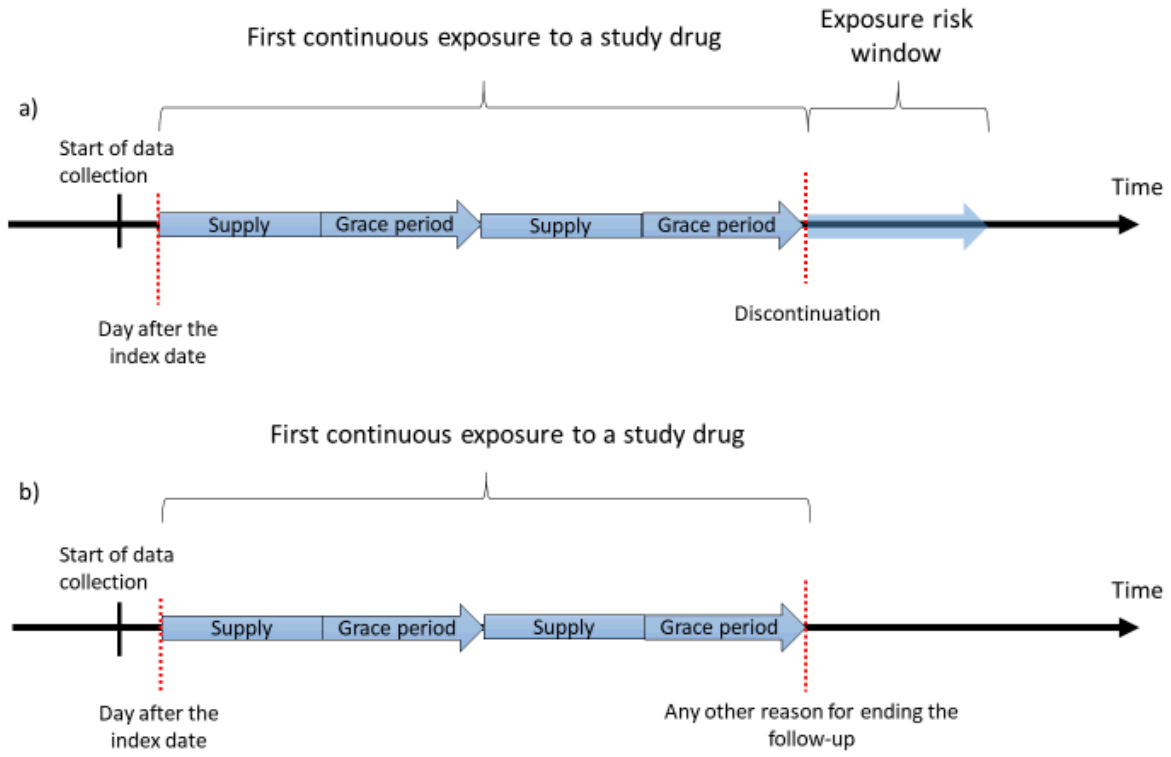
Figure 6 Illustration of a study drug a) discontinuation; b) switch and c) concomitant use



An ‘as-treated ‘(AT) approach will be utilized in the main analyses. This means that the follow-up is censored at discontinuation, switch to other study drug, or concomitant use, as detailed in Figure 1. The outcomes related to effectiveness, safety, HCRU, and the cost of care will be observed until the end of the first continuous exposure to a study drug (Figure 5).

When investigating the safety outcomes, a modified AT approach will be utilized in the main analyses with an additional exposure risk window of 30 days after discontinuation of drug use (Figure 7), to consider potential lag in the occurrence of safety outcomes. The exposure risk window will be accounted for only in case of discontinuation of the drug use, and not when the follow-up ends at the switch of drug and starting concomitant use (Figure 1). Consequently, the risk period in these analyses will consist of the supply, grace period, and exposure risk window. The exposure risk window will not be used after switch or starting concomitant use because the safety associated with the initial drug is the main interest in these analyses.

Figure 7 Illustration of the utilization of 30-day exposure risk window in the analyses with safety outcomes. Exposure risk window will be accounted for if the follow-up ends at discontinuation of drug use (a) but not at discontinuation for any other reason (b).



9.3.2 Outcomes

The outcome definitions shall be detailed in the localized protocols, using the master definitions presented in this master protocol (Table 4; more detailed definitions in ANNEX 6).

Primary and secondary effectiveness outcomes

The primary and secondary effectiveness outcomes (Table 4) will be observed during the study period, and the occurrence of the outcome in question, excluding secondary outcome 6b, will be observed until the end of the follow-up.

Secondary renal effectiveness outcomes

The baseline value for the secondary renal effectiveness outcomes 6-8 (Table 4) will be considered as the creatinine and/or albuminuria value that is closest to the index date during the pre-index period (Figure 3). The date of outcome for outcomes 6a, 7, and 8 is the date when the decline to abnormal status is observed. However, the first 4 weeks after the index

date will be disregarded in all analysis for the secondary outcomes 6-8, as empagliflozin seems to have some initial effects on eGFR decline [P18-09763].

If countries can define inpatient and outpatient visits, they could be preferable to use outpatient measurements.

Definitions of different levels of eGFR and albuminuria are presented in ANNEX 6. The eGFR decline will be measured by utilizing serum creatinine and the change from normal kidney function (≥60ml/min/1,73m^2) to abnormal kidney function (<60ml/min/1,73m^2) will be analyzed. The eGFR will be based on serum creatinine (SCr) measurements and the 2009 Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) creatinine equation [R13-4387]: $141*\min(SCr/\kappa, 1)^{\alpha} *\max(SCr/\kappa, 1)^{-1.209}*0.993^{Age}$ [*1.018 if female] [*1.159 if black], where κ is 0.7 for female and 0.9 for male whereas α is -0.329 for female and -0.411 for male. If 'black' is missing, the factor of 1 will be used. The min is the minimum of SCr/κ or 1, and max is the maximum of SCr/κ or 1. In addition, the progression from normoalbuminuria (<30mg/g) to micro- or macroalbuminuria (≥30mg/g) will be based on the urine albumin-to-creatinine ratio.

Secondary safety outcomes
The safety outcomes (Table 4) will be observed during the study period, but if the follow-up ends at discontinuation, an exposure risk window will be accounted for (Figure 7). The occurrence of the outcome in question will be observed until the end of the follow-up. These analyses will focus on empagliflozin because the safety profile of the SGLT-2 inhibitors as a group might not be comparable [P18-02804, P19-04731, P18-01920].

Secondary outcomes on health care resource utilization and cost of care
Availability of cost data may vary across countries. If costs are recorded in the registers directly, they will be used as such. If not, costs of outpatient visits, inpatient stays, and the drug use will be evaluated indirectly by country-specific standards and associated with the observed number of different encounter types (Table 4). For example, the diagnosis-related group information about the cost of the outpatient visit, inpatient stay of 24-hours of a specific type, or general information about country-specific drug costs can be used and assigned to the observed number of resource use of this type. Quantities of HCRU for patients on each treatment and the cost (average cost or unit cost) for each type of HCRU will be reported separately, in 2018 values (or nearest value year). Due to differences in availability of data on costs for different types of HCRU in each country, such as administrative cost data or list prices, the methods used for calculating and (if relevant) inflating costs associated with each HRCU will be reported.

HCRU and costs (Table 4) will be measured over the follow-up periods, as described at 9.2.3 section, however, the occurrence of any other effectiveness outcome or a safety outcome will not end the follow-up. The HCRU and cost of care outcomes will not be included in the meta-analysis; instead, the HCRU and cost results will be reported descriptively by country.

Country-specific outcomes beyond this multi-country study
Country-specific primary and secondary outcomes can be included in the analyses, as specified in the localized protocols. The obtained results of such country-specific analyses on further outcomes will not be included in the combined results of the multi-country study (ANNEX 3).

Table 4 Master definitions for primary and secondary outcomes included in this study.

Outcomes corresponding to each objective	**Definition[1]**
Primary effectiveness outcomes	
Primary outcome 1: Hospitalization for heart failure[2]	Definition 1a: Primary diagnosis[3] associated with hospital admission Definition 1b: Any diagnosis[2] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters, or dispensation or any other record of the high-ceiling diuretics (loop diuretics).
Primary outcome 2: All-cause mortality	As available in the data sources
Primary outcome 3: Composite outcome including Hospitalization for heart failure[2] All-cause mortality	As for primary outcome 1a and primary outcome 2
Primary outcome 4: Composite outcome including MI[2] Stroke All-cause mortality	For MI and stroke, any diagnosis[2] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters For all-cause mortality, as for primary outcome 2
Secondary cardiovascular effectiveness outcomes	
Secondary outcome 1: CV mortality	As available in the data sources
Secondary outcome 2: composite outcome including Hospitalization for heart failure2 CV mortality	As for primary outcome 1a and secondary outcome 1.

Table 4 (cont'd) Master definitions for primary and secondary outcomes included in this study.

Outcomes corresponding to each objective	**Definition[1]**
Secondary cardiovascular effectiveness outcomes	
Secondary outcome 3: 3-point MACE, a composite outcome including MI2 Stroke2 CV mortality	As for primary outcome 4 and secondary outcome 1.
Secondary outcome 4: Coronary revascularization procedure	Any diagnosis/procedure[3] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters
Secondary renal effectiveness outcomes	
Secondary outcome 5: ESRD	Any diagnosis/procedure/laboratory measurement[3] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters
Secondary outcome 6: eGFR decline[4]	Definition 6a: Decline from normal kidney function (≥60ml/min/1,73m^2) to abnormal kidney function (<60ml/min/1,73m^2). Definition 6b: Change in eGFR over time.
Secondary outcome 7: Progression from normoalbuminuria to micro- or macroalbuminuria[4]	Progression from normoalbuminuria (<30mg/g) to micro- or macroalbuminuria (≥30mg/g).
Secondary outcome 8: Composite outcome including eGFR decline and progression to micro- or macroalbuminuria[4]	As for secondary outcome 6a and secondary outcome 7.
Secondary safety outcomes	
Secondary outcome 9: Bone fracture	Any diagnosis/procedure[3] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters

Table 4 (cont'd) Master definitions for primary and secondary outcomes included in this study.

Outcomes corresponding to each objective	**Definition**[1]
Secondary safety outcomes	
Secondary outcome 10: Diabetic ketoacidosis	Any diagnosis[3] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters
Secondary outcome 11: Severe hypoglycemia	Any diagnosis[3] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters
Secondary outcome 12: Lower-limb amputation	Any diagnosis/procedure[3] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters
Secondary outcome 13 Acute kidney injury requiring dialysis	Any diagnosis/procedure[3] associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters
Secondary outcomes on HCRU and cost of care	
Secondary outcome 14: HCRU	Outpatient healthcare visits Primary care visits, specialist outpatient care visits or office visits, excluding emergency care, according to data availability Emergency department visits Inpatient healthcare visits Hospital admissions Inpatient days Length of stay Drug use in outpatient and inpatient care Dispensations/other records of the drug use
Secondary outcome 15: Cost of care	Costs of HCRU (as in secondary outcome 14)

1 The outcome definitions shall be detailed in the localized protocols, using the master definitions presented in this master protocol; 2 As heart failure, MI and stroke are defined from diagnosis without mortality data, the outcomes include both fatal and non-fatal events; 3 Detailed variable definitions for the outcomes are available in ANNEX 6; 4 Based on the Kidney Disease Improving Global Outcomes (KDIGO) 2012 recommendations (ANNEX 6).

CV=cardiovascular; eGFR=estimated glomerular filtration rate; ESRD=end-stage renal disease; HCRU=health care resource utilization; MACE=major adverse cardiovascular event; MI=myocardial infarction.

9.3.3 Covariates at baseline (baseline characteristics)

The types of covariates at baseline (baseline characteristics) utilized in this study and their measurement periods are listed in. Table 5. The complete list of covariates, representing each covariate type, and their operational definitions are available in ANNEX 7. The master definitions for the covariates presented in this master protocol shall be adapted to local settings and described in the localized protocols, for country-level analyses (ANNEX 3). Further, the availability of the covariates will vary between the data sources and will be specified in the localized protocols.

Additional covariates at baseline, or also during the follow-up, can be included in the additional country-specific analyses (ANNEX 3). The obtained results of such country-specific analyses with additional covariates will not be included in the combined results of this multi-country study.

Table 5 Types of covariates at baseline.

Covariate[1,2] at baseline	**Measurement period**
Sociodemographic characteristics	The index date (closest to)
Covariates related to lifestyle	The index date (closest to)
Diabetes complications	12 months preceding the index date (inclusive)
Other comorbidities	12 months preceding the index date (inclusive)
Laboratory values	12 months preceding the index date (inclusive)
Prior/concomitant use of other antidiabetic drugs	12 months preceding the index date and at the index date
Prior use of other drugs	12 months preceding the index date (inclusive)
Healthcare resource utilization covariates	12 months preceding the index date (inclusive)
Cost covariates	12 months preceding the index date (inclusive)

The complete list of covariates, representing each covariate type, and their operational definitions are available in ANNEX 7. These definitions shall be adapted to local settings and described in the localized protocols. The availability of the variables shall also be specified in the localized protocols; [2] In the localized protocols, additional variables may be added.

9.4 DATA SOURCES

This study will be based on several nationwide and regional data sources of observational data in 11 countries across Europe and Asia, namely Denmark, Finland, Germany, Israel, Japan, Norway, South Korea, Spain, Sweden, Taiwan, and the United Kingdom. This will ensure a sufficient number of drug users. Additional countries may be included during the course of the study.
Data sources eligible for this study include nationwide healthcare registers, regional quality registers, regional high-quality medical health records, and other health claims data. The utilized data sources, to be specified in the localized study protocols (ANNEX 3), will have a sufficient level of key data elements to successfully conduct the study. The key data elements in the utilized data sources will be the identification of patients using empagliflozin, any SGLT-2 inhibitor, or any DPP-4 inhibitor and availability of patient-level data on exposures, outcomes, and covariates, including sociodemographic characteristics, other drugs, other diagnoses, and HCRU. Moreover, the utilized data sources shall be capable of delivering insights for the study in a reasonable timeframe, including consideration of data lag time and lag time due to administrative issues.

9.5 STUDY SIZE

Table 6 shows the sample size required to detect a 20% or 30% decrease in the outcome rate (HR 0.8 or 0.7, respectively) with 80% power and 5% alpha level. Varying expected outcome event rates (0.5-3% per patient-year) are presented assuming that each patient will contribute either 2 years, 1 year or 0.5 years of follow-up time. For example, in the EMPA-REG OUTCOME® study [P18-01920] study the heart failure rate was approximately 1.5% per year in the non-empagliflozin arm, with HR = 0.65 for the empagliflozin arm.

Table 6 The sample size required to detect 20% or 30% decrease (HR 0.80 or 0.70) in the risk of outcome with 80% power as a function of yearly outcome event rate (ranging from 0.5% to 3.0%), and assuming either 2.0, 1.0 or 0.5 years of average follow-up time.

Event rate	**0.5%**	**1.0%**	**1.5%**	**2.0%**	**2.5%**	**3.0%**
HR	**0.80**	**0.80**	**0.80**	**0.80**	**0.80**	**0.80**
Average follow-up time: 2.0 years						
Sample required, per sub-cohort	35,000	18,000	12,000	9,000	7,000	6,000
Average follow-up time: 1.0 year						
Sample required, per sub-cohort	71,000	35,000	24,000	18,000	14,000	12,000
Average follow-up time: 0.5 years						
Sample required, per sub-cohort	142,000	71,000	47,000	35,000	28,000	24,000
Event rate	**0.5%**	**1.0%**	**1.5%**	**2.0%**	**2.5%**	**3.0%**
HR	**0.70**	**0.70**	**0.70**	**0.70**	**0.70**	**0.70**
Average follow-up time: 2.0 years						
Sample required, per sub-cohort	15,000	7,000	5,000	4,000	3,000	2,000
Average follow-up time: 1.0 year						
Sample required, per sub-cohort	30,000	15,000	10,000	7,000	6,000	5,000
Average follow-up time: 0.5 years						
Sample required, per sub-cohort	60,000	30,000	20,000	15,000	12,000	10,000

HR=hazard ratio.

For example, with a baseline outcome rate of 1.5% per year, the analyses are sufficiently powered (>80%) to detect 30% decrease in the baseline rate with 20,000 patients per sub-cohort, i.e., 40,000 patients in total, and 0.5 years of follow-up per patient.

The estimated number of users of empagliflozin based on selected real-world datasets is presented in Table 7. The expected total number of patients in the study is at least over 150,000, which is sufficient for most of the comparisons presented in. Table 6. However, the final power in the meta-analyses will also depend on the heterogeneity in the results across the study countries.

Table 7 The estimated number of empagliflozin users in selected databases by study country.

Country	**N**
Denmark	15,000
Finland	25,000
Germany	2,100
Israel	6,000
Japan	5,600
Norway	11,000
South Korea	39,000
Spain	12,000
Sweden	17,600
Taiwan	14,000
The United Kingdom	3,000-5,000

9.6 DATA MANAGEMENT

9.6.1 Country-specific studies

Detailed data management procedures for the country-specific studies, including the used statistical software in each country, will be detailed in localized protocols (ANNEX 3**)**. Each country will collect the variables from their data sources and comply with procedures that include checking electronic files, maintaining security and data confidentiality, following analysis plans, and performing quality checks for all programs. Each country will maintain any patient-identifying information securely on site according to up-to-date standard operating procedures. They will also maintain appropriate data storage, and archiving procedures will be followed with a periodic backup of files.

Each country will perform analyses with individual-level data separately as described in sections 9.7.1-9.7.6 (see also Figure 2). After performing the analyses with individual-level data in each country, aggregate-level results will be collected for combining the results from each country, in a pre-specified (similar) format to be detailed in the master statistical analysis plan (SAP). The master SAP will include template tables for collecting result data from each country.

9.6.2 A multi-country study combining country-specific results

After receiving the aggregate-level results from each country, the results will be combined by [redacted] (see also Figure 2). The meta-analysis will be performed using the R language [R18-3507]. The aggregate-level results will be stored securely according to up-to-date standard operating procedures. Maintaining appropriate data storage includes periodic backup of files and archiving procedures. [redacted] will comply with procedures that include checking electronic files, maintaining security and data confidentiality, following analyses plans, and performing quality checks for all programs. The source code of data management and data analyses is kept for inspection for five years after the publication of results. Access to the study data cannot be given to any third parties, neither the study data can be used for other purposes than described in this master protocol.

9.7 DATA ANALYSIS

All analyses described in sections 9.7.1–9.7.6 will be performed separately by each country. Analyses for combining the country-specific results described in section 9.7.7 and 9.7.8 will be performed by [redacted]

Additional analyses, other than those described in this master protocol for the multi-country study, may be detailed in localized study protocols (ANNEX 3). The obtained results of any additional analyses will not be included in the combined results of the multi-country study.

9.7.1 Propensity score methodology

PS matching will be used to reduce confounding in the comparative analyses across study sub-cohorts. Pairwise PS models between i) empagliflozin sub-cohort and DPP-4 sub-cohort and ii) SGLT-2 sub-cohort and DPP-4 sub-cohort will be estimated using logistic regression including all available covariates as indicated in ANNEX 7 (types of covariates are also outlined in Table 5). Propensity scores, obtained from the logistic regression models, indicate the predicted treatment probability of being in the main exposure sub-cohort, i.e. i) empagliflozin user or ii) any SGLT-2 inhibitor user. [redacted]
PS matching will be done in a 1:1 ratio using the nearest-neighbor algorithm, using calipers of width equal to 0.2 of the standard deviation of the logit of the PS (matches of the main exposure sub-cohort and the comparator sub-cohort are formed if the caliper width is not exceeded for difference of the logits of PS). One-to-one matching will be used as one-to-many matching introduces several complications while providing only a minimal gain in estimation precision [R17-1267].
In the matching procedure, matched comparator sub-cohort patients (any DPP-4 inhibitor sub-cohort) will be removed from the pool of eligible patients among the comparator sub-cohort so that a matched patient is no longer available for consideration as a potential match for another exposed patient. If there are multiple potential matches (controls) available for a patient in the main exposure sub-cohort (empagliflozin / any SGLT-2 inhibitor sub-cohort), then the controls will be sorted by (ascending) absolute difference in the logit of the PS, and the control on the top of this list will be selected. In case of a tie, the choice is made randomly.
Evaluation of the success of the matching procedure will be based on a standardized difference (detailed below) of a covariate between treatment groups. A standardized difference of a covariate of less than 0.1 between treatment groups in this study indicates a negligible difference in the mean or prevalence of the covariate. If larger standardized differences than 0.1 for some covariates exist after matching, they will be used in post-matching adjustments (i.e. PS-variables with standardized difference >0.1 will be adjusted for in the analyses). Further adjustments can also be made as appropriate in additional models.

The achieved PS balance will be investigated by plotting the PS-distribution and tabulating for each PS-variable the standardized difference among the main exposure sub-cohorts and the comparator sub-cohort. Standardized differences for the covariates included in the calculation of the PS will be calculated according to the unweighted formula provided by Austin [R18-3545]. The formula for standardized difference for a dichotomous covariate is:

$$d = \frac{|p_1 - p_2|}{\sqrt{\frac{p_1(1-p_1) + p_2(1-p_2)}{2}}}$$

where p_1 and p_2 denote the prevalence in exposed and unexposed patients, respectively. For a continuous variable the formula is:

$$d = \frac{|\overline{x_1} - \overline{x_2}|}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

where $\overline{x_1}$ and $\overline{x_2}$ denote the sample means; s_1^2 and s_2^2 sample variances in exposed and unexposed patients, respectively. For a categorical variable with multiple levels, the method proposed by Dalton [R18-3515] will be used.

9.7.2 Baseline characteristics / descriptive statistics

For all included patients, descriptive statistics will be generated separately in each sub-cohort. Continuous covariates will be described by the mean, standard deviation (SD), median, 25^{th,} and 75th percentiles, minimum and maximum. Categorical covariates and continuous covariates that were also categorized will be described by proportion and frequency in each category.

Patient characteristics at cohort entry for the main exposure sub-cohort and the comparator sub-cohort patients and the standardized differences before and after matching will be presented as in Table 8 below.

Table 8 Illustration of baseline characteristics table. Among the main exposure sub-cohorts (empagliflozin/ SGLT-2 inhibitor sub-cohort) and the comparator sub-cohort (DPP-4 sub-cohort), baseline characteristics variables will be presented before and after matching and their standardized differences. A separate table will be provided for the empagliflozin – any DPP-4 inhibitor sub-cohorts and the SGLT-2 inhibitor – any DPP-4 inhibitor sub-cohorts.

	Before matching			**After matching**		
	Main exposure sub-cohort	**Comparator sub-cohort**	**Std.**	**Main exposure sub-cohort**	**Comparator sub-cohort**	**Std.**
Continuous variable 1 (e.g. age) that has been also categorized						
Category 1 (e.g. <39 years) N (%)	n (p.pp)	n (p.pp)	-	n (p.pp)	n (p.pp)	-
Category 2 (e.g. 40-60 years) N (%)	n (p.pp)	n (p.pp)	-	n (p.pp)	n (p.pp)	-
Category 3 (e.g. >60 years) N (%)	n (p.pp)	n (p.pp)	-	n (p.pp)	n (p.pp)	-
Mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xx (based on mean)	x.xx (x.xx)	x.xx (x.xx)	x.xx (based on mean)
Median (Q1, Q3)	x.xx (x.xx, x.xx)	x.xx (x.xx, x.xx)		x.xx (x.xx, x.xx)	x.xx (x.xx, x.xx)	
Min, Max	x.xx, x.xx	x.xx, x.xx		x.xx, x.xx	x.xx, x.xx	

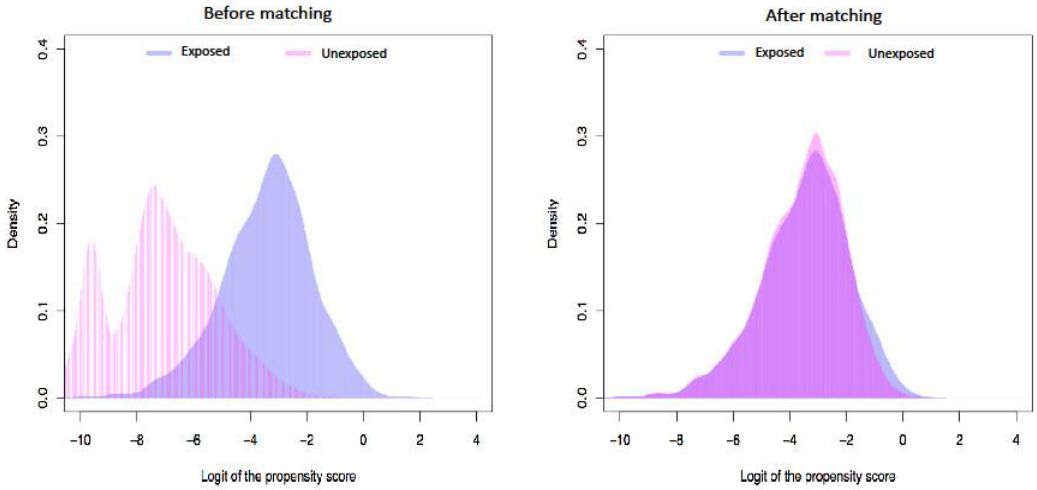
Table 8 (cont'd) Illustration of baseline characteristics table. Among the main exposure sub-cohorts (empagliflozin/ SGLT-2 inhibitor sub-cohort) and the comparator sub-cohort (DPP-4 sub-cohort), baseline characteristics variables will be presented before and after matching and their standardized differences. A separate table will be provided for the empagliflozin – any DPP-4 inhibitor sub-cohorts and the SGLT-2 inhibitor – any DPP-4 inhibitor sub-cohorts.

	Before matching			**After matching**		
	Main exposure sub-cohort	**Comparator sub-cohort**	**Std.**	**Main exposure sub-cohort**	**Comparator sub-cohort**	**Std.**
Categorical variable 1 (e.g. sex)						
Category 1 (e.g. male) N (%)	n (p.pp)	n (p.pp)	x.xx	n (p.pp)	n (p.pp)	x.xx
Category 2 (e.g. female) N (%)	n (p.pp)	n (p.pp)	-	n (p.pp)	n (p.pp)	-
Continuous variable 2 that has not been categorized						
Mean (SD)	x.xx (x.xx)	x.xx (x.xx)	x.xx (based on mean)	x.xx (x.xx)	x.xx (x.xx)	x.xx (based on mean)
Median (Q1, Q3)	x.xx (x.xx, x.xx)	x.xx (x.xx, x.xx)		x.xx (x.xx, x.xx)	x.xx (x.xx, x.xx)	
Min, Max	x.xx, x.xx	x.xx, x.xx		x.xx, x.xx	x.xx, x.xx	

Std.=standardized difference; SD=standard deviations; Q1=first quartile; Q3=third quartile.

The propensity score distributions (in logit scale) will be presented before and after matching as illustrated in Figure 8.

Figure 8 Illustration of propensity score distribution among the main exposure sub-cohorts (empagliflozin/ any SGLT-2 inhibitor sub-cohorts; "exposed" in the figure) and among the comparator sub-cohort (any DPP-4 inhibitor sub-cohort; "unexposed" in the figure) before and after matching. Separate figures will be provided for the empagliflozin – any DPP-4 inhibitor sub-cohorts and any SGLT-2 inhibitor – any DPP-4 inhibitor sub-cohorts.



DPP-4=dipeptidyl peptidase-4; SGLT-2=sodium-glucose cotransporter-2.

9.7.3 Primary effectiveness outcomes

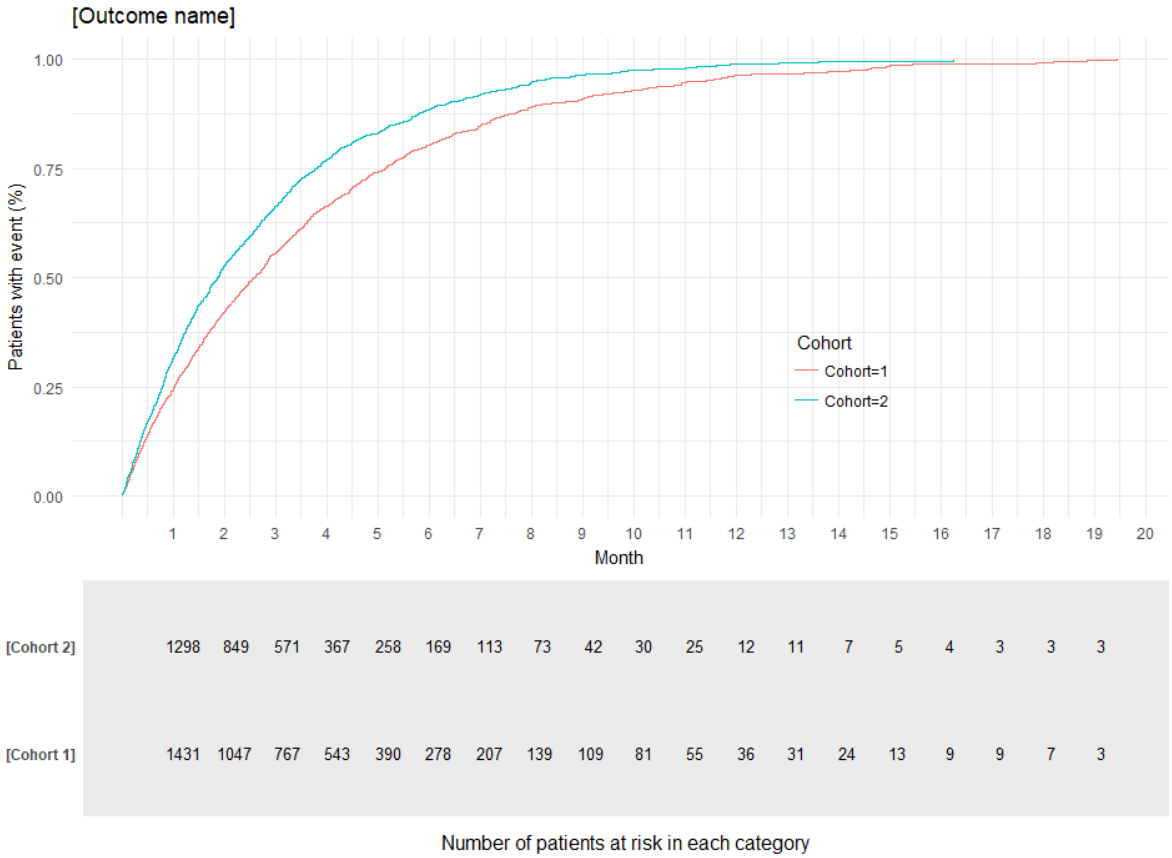
Primary analyses
For the primary effectiveness outcomes, incidence rates (events per person-years) with corresponding 95% confidence intervals (CI) will be calculated, separately for each sub-cohort. An example illustrating the result output is provided in Table 9. In addition, the competing risk-adjusted (i.e. treating death as competing event) cumulative incidence plots will be provided, and adjusted HRs with 95% CIs estimated using a Cox proportional hazard model. Components of all-cause mortality that are not part of the particular outcome will be considered as competing risk in the analyses of cumulative incidence. The cumulative incidence plots will be accompanied by a number of patients at risk at different time points during follow-up, as illustrated in Figure 9.
As described in section 9.7.1, the main Cox model will be adjusted for unbalanced PS-variables at baseline. If there is a concern of residual confounding, analyses with further adjustments can be included in additional models to be detailed in localized study protocols. The possible additional models may include more adjusting covariates at baseline (also other than those included in ANNEX 7) and may include time-dependent exposures to drugs listed in ANNEX 7. The obtained results of these additional models will not be included in the combined results of the multi-country study.

Table 9 The number of outcome events, patients at risk, incidence rates and hazard ratio with 95% confidence intervals among empagliflozin/any SGLT-2 inhibitor sub-cohort and among any DPP-4 inhibitor sub-cohort.

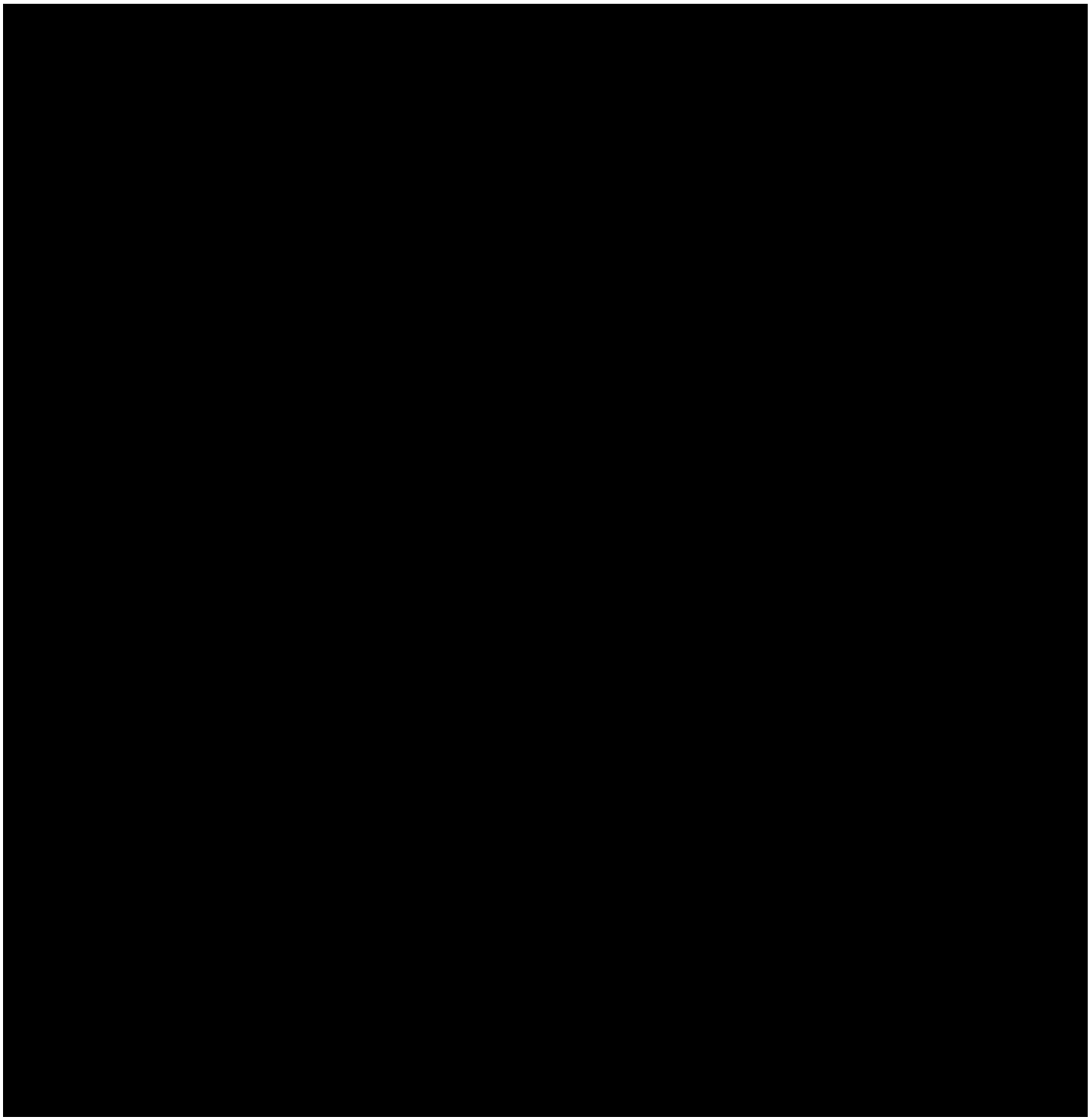
Outcome: (e.g. heart failure)	**Any DPP-4 inhibitor sub-cohort**	**Empagliflozin/Any SGLT-2 sub-cohort**
Events (person-years)	x.xx (y.yy)	x.xx (y.yy)
N at risk	N	N
Incidence rate (95% CI)	x.xx (x.xx)	x.xx (x.xx)
Adjusted HR (95% CI)	1 (reference)	x.xx (x.xx)

CI=confidence interval; DPP-4=dipeptidyl peptidase-4; HR=hazard ratio; SGLT-2=sodium-glucose cotransporter-2.

Figure 9 Illustration of cumulative incidence plot. Death is treated as a competing event.



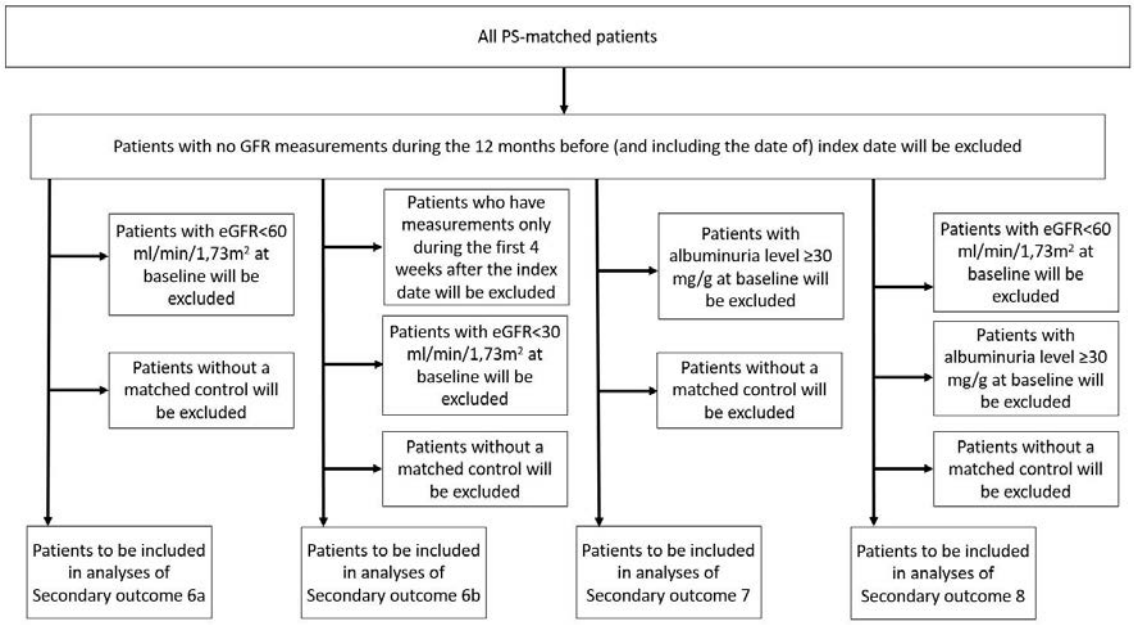
Alternative analyses with additional exclusion criteria
The primary analyses, as described above, will be re-performed using the following additional exclusion criteria: History of diagnosis of malignant neoplasm during the 5 years before the index date (ANNEX 8). Since applying this criterion will change the study population, the PS matching and balance analyses will also be re-performed.



9.7.5 Secondary outcomes: effectiveness, safety, HCRU and costs of care

Effectiveness outcomes
The analysis of secondary outcomes 1-6a, 7, and 8 (Table 4) is done in a similar way as the analysis of primary effectiveness outcomes (section 9.7.3) and will be investigated for both empagliflozin and SGLT-2 inhibitor exposure.
The analyses for secondary outcomes 6a, 7, and 8 (Table 4) will be based on the PS-matched patients as described in section 9.7.1; however, additional exclusion criteria will be used (Figure 10). Persons with abnormal status at baseline will be excluded because no decline can be observed for those persons. Therefore, persons with eGFR<60 ml/min/1,73m^2 will be excluded from analyses for secondary outcome 6a and 8. Similarly, persons with albuminuria level ≥30 mg/g will be excluded from analyses for secondary outcome 7 and 8. Persons with eGFR<30 ml/min/1,73m^2 will be excluded from analyses for secondary outcome 6b. In addition, persons with no GFR measurements during the 12 months before (and including the date of) index date will be excluded from analyses regarding secondary outcomes 6-8. In the analyses regarding secondary outcome 6b, persons who have measurements only during the first 4 weeks after the index date will be excluded (Figure 10). In all analyses for secondary outcomes 6-8, any measurements during the first 4 weeks after the index date will be disregarded (as described in Section 9.3.2).

Figure 10 Flow chart for secondary renal effectiveness outcomes

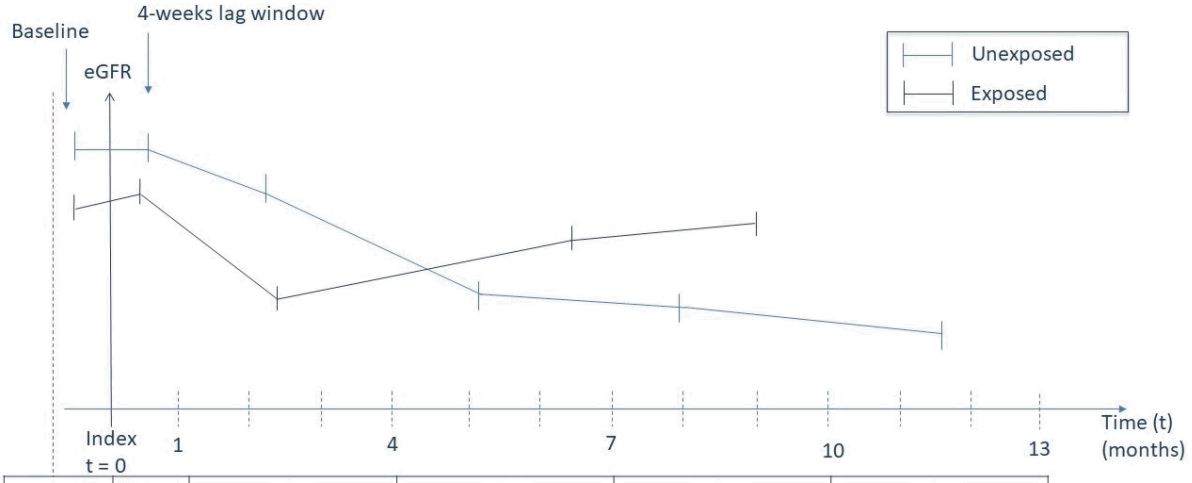


PS=propensity score; GFR=glomerular filtration rate; eGFR=estimated glomerular filtration rate

Regarding secondary outcome 6a, 7 and 8, the number and proportion of patients who stay or move to the lower/higher eGFR/albuminuria state between the baseline and the last available measurement will be summarized in a N x N table (where N is the number of possible states (Figure 10). Each study cohort will be described separately. In this table all patients in the matched cohort will be included, regardless of missing values or abnormal states at baseline or at last available measurement.

Regarding secondary outcome 6b (Table 4), the population trend for the change of eGFR will be characterized by sub-cohorts using moving averages over time. Specifically, the mean values overall patients with existing measurements will be calculated in three-month windows starting with four weeks after the index date, securing the lag period. Each patient contributes equally within the time-window, meaning that patients with more than one measurement within the time-window will contribute with the mean value of their measurements. The percentage of patients having at least one measure will be calculated for each time-window. The trends will be visualized by sub-cohorts as illustrated in Figure 11.

Figure 11 Illustration of the eGFR change by sub-cohort.



Patients with no measurements	Unexposed	34/200	29/200	52/200	75/200
	Exposed	43/200	35/200	92/200	200/200

eGFR=estimated glomerular filtration rate

Further, the slope of eGFR change (average change per unit time: (observed value in time-window – baseline value)/time(years)) between the baseline eGFR and the last measured value within pre-fixed time-window will be compared between sub-cohorts using linear regression model adjusted for baseline eGFR value and variables not balanced after PS matching. The following time-windows will be used:

- from the start of week 5 until the end of follow-up (AT approach),
- from the start of week 5 until the end of 12 months or the end of follow-up (whichever comes first) and will include 3 month-interval analysis if sufficient sample sizes exist
- from the start of 13th month until the end of 24 months or the end of follow-up (whichever comes first)
- moving along yearly until the end of data availability

The follow-up is not censored in the occurrence of an effectiveness or safety outcome. All patients with baseline eGFR >30 ml/min/1,73m2 and at least 1 measurement in the corresponding time-window will be included in the analyses. The regression coefficients will be presented together with 95% CIs. In country-specific analyses, alternative methods (e.g. mixed effect models) can be used in case of sufficient data availability. These country-specific results will not be included in the meta-analysis.

When excluding patients from the analyses of outcomes 6a, 6b, 7 and 8, the same matched cohort that will be used for primary and other secondary outcomes, will be used, i.e. no re-matching will be done. Instead, to ensure balance across cohorts, when excluding patients from the exposed sub-cohort, the corresponding matched control will also be excluded. Similarly, when excluding patients from the unexposed, the corresponding matched exposed patient will also be excluded. After this, the balance will be checked, and the unbalanced variables will be adjusted for in the analyses.

Safety outcomes
The safety outcomes will be investigated for empagliflozin exposure only. In cases where the follow-up is censored at discontinuation of drug use, the exposure risk window (Figure 7) will be accounted for.

The analysis of safety outcomes will include the description of the number of outcome events, patients at risk, incidence rates with the 95% CIs, and cumulative incidence function plots for the safety outcomes. All-cause mortality will be considered as a competing risk in the analyses of cumulative incidence.

The HRs together with the 95% CIs will be estimated by Cox regression for all time-to-event outcomes. The HRs will be estimated using models adjusted for the unbalanced PS-variables at baseline, as described in section 9.7.1. All safety outcomes will be analyzed with the time-to-event approach.

In the multi-country study, the main approach in the analyses of safety outcomes is to study the time to the first event. In country-specific analyses, recurrent safety outcomes can be considered, as described in local SAPs. These country-specific analyses will not be included in the meta-analyses.

Table 10 Description of transitions between eGFR and albuminuria states from baseline to the last available measurement.

At baseline			At last measurement				Missing data				
Sub-cohort*:			eGFR 1 albuminuria 1	eGFR 1 albuminuria 0	eGFR 0 albuminuria 1	eGFR 0 albuminuria 0	eGFR 0 albuminuria 2	eGFR 2 albuminuria 0	eGFR 1 albuminuria 2	eGFR 2 albuminuria 1	eGFR 2 albuminuria2
eGFR	albu.	N									
1	1	N_{11}	n_{1111} (%**)	n_{1110} (%)	n_{1101} (%)	n_{1100} (%)	n_{1102}	n_{1120}	n_{1112}	n_{1121}	n_{1122}
1	0	N_{10}	n_{1011} (%)	n_{1010} (%)	n_{1001} (%)	n_{1000} (%)	n_{1002}	n_{1020}	n_{1012}	n_{1021}	n_{1022}
0	1	N_{01}	n_{0111} (%)	n_{0110} (%)	n_{0101} (%)	n_{0100} (%)	n_{0102}	n_{0120}	n_{0112}	n_{0121}	n_{0122}
0	0	N_{00}	n_{0011} (%)	n_{0010} (%)	n_{0001} (%)	n_{0000} (%)	n_{0002}	n_{0020}	n_{0012}	n_{0021}	n_{0022}
0	2	N_{02}	n_{0211}	n_{0210}	n_{0201}	n_{0200}	n_{0202}	n_{0220}	n_{0212}	n_{0221}	n_{0222}
2	0	N_{20}	n_{2011}	n_{2010}	n_{2001}	n_{2000}	n_{2002}	n_{2020}	n_{2012}	n_{2021}	n_{2022}
1	2	N_{12}	n_{1211}	n_{1210}	n_{1201}	n_{1200}	n_{1202}	n_{1220}	n_{1212}	n_{1221}	n_{1222}

Table 10 (cont'd) Description of transitions between eGFR and albuminuria states from baseline to the last available measurement.

At baseline			At last measurement				Missing data				
Sub-cohort*:			eGFR 1 albuminuria 1	eGFR 1 albuminuria 0	eGFR 0 albuminuria 1	eGFR 0 albuminuria 0	eGFR 0 albuminuria 2	eGFR 2 albuminuria 0	eGFR 1 albuminuria 2	eGFR 2 albuminuria 1	eGFR 2 albuminuria2
eGFR	albu.	N									
2	1	N_{21}	n_{2111}	n_{2110}	n_{2101}	n_{2100}	n_{2102}	n_{2120}	n_{2112}	n_{2121}	n_{2122}
2	2	N_{22}	n_{2211}	n_{2210}	n_{2201}	n_{2200}	n_{2202}	n_{2220}	n_{2212}	n_{2221}	n_{2222}

* This table will be presented for all sub-cohorts (Empagliflozin/SGLT-2/DPP-4);

**The denominators for the percentages calculations will be the numbers from the third column (N_{11},N_{10},etc)

eGFR=estimated glomerular filtration rate; albu.= albuminuria;

0=abnormal; 1=normal; 2=missing.

HCRU outcomes
The HCRU outcomes (Table 4) will be investigated for empagliflozin exposure only. HCRU outcomes will be analyzed separately in each sub-cohorts, i.e. empagliflozin and DPP-4 inhibitors, by using the AT approach, applying the specific reasons for ending the follow-up (see section 9.3.2). Only time observed during follow-up will be considered in the analyses, although inpatient stays may extend outside the study period. For example, when reporting the length of inpatient stay in the case a patient is admitted before the end of a study period, but not discharged until its end, the days going past the end of the study period will not be considered (right truncation). The same procedure will be applied when calculating outcomes per member per month and a hospital stay extends over different monthly windows.

The total number of inpatient days will be counted and reported for each patient during follow-up. The mean, standard deviation, median, 25^{th,} and 75th quartiles, minimum, and maximum will be reported for the crude number of inpatient days per patient, and as the proportion of inpatient days per each patient´s follow-up time. The number of inpatient days will also be calculated and reported per member per month.

For outpatient healthcare visits, inpatient healthcare visits and drug use (Table 4), the number of events (total), person-years, and incidence rates (number of events per person-year) will be calculated and reported. Each of these three resource utilization types will be analyzed separately. A total number of inpatient visits, outpatient visits, and dispensations or any other records of the drug use (each separately) will be compared across treatment groups using a Poisson regression model. The logarithm of follow-up time will be used as an offset in the model. However, if the Poisson assumptions are violated, another modeling option can be used (e.g. zero-inflated model or negative-binomial). When counting the number of events and person-years for healthcare visits, inpatient days will be censored (i.e. not counted as risk time). The number of visits will also be calculated and reported per member per month.

The first inpatient visit (hospitalization) will be analyzed separately by reporting the number of events, person-time, and incidence rate. Further, HRs with 95% confidence intervals will be estimated using the Cox proportional hazard models, comparing treatment groups and adjusting for unbalanced PS-variables. In this analysis, time after the first hospitalization will be censored.

As described in section 9.7.1, the main Cox and Poisson models will be adjusted for unbalanced PS-variables at baseline. In addition, as the event rate can increase as a function of time since the start of follow-up, year of follow-up will be additionally adjusted for.

In country-specific analyses, HCRU outcomes can be additionally analyzed separately by type: e.g. CV vs. non-CV, diabetes-related vs. non-diabetes related resources.

Cost of care outcomes
The cost of care outcomes (Table 4) will be investigated for empagliflozin exposure only. Cost of care outcomes will be analyzed separately in each sub-cohorts, i.e. empagliflozin and DPP-4 inhibitors, by using the AT approach, applying the specific reasons for ending the follow-up (see section 9.3.2).

Costs will be reported overall and separately based on the type of service: inpatient healthcare visits, outpatient healthcare visits, and drug use (Table 4). For the primary analysis of cost outcomes, the Lin's method will be applied [R17-2088]. Specifically, the average cost accrued per member in a given month (i.e. cost per member per month) will be estimated. This involves calculating the average cost among members who are alive at the beginning of a month. Cost per member per month then will be weighted by Kaplan-Meier probability of survival for each month and then summed up over available months of the study period. Poisson regression will also be used to estimate the relationship between total costs and treatment (empagliflozin vs. any DPP-4 inhibitor), controlling for patient covariates. For the regression analyses, patients' observed costs within each month will be weighted by inverse probability of not being censored, due to potential informative censoring [R18-3512]. After weighting, the formal comparison of secondary cost outcomes between study sub-cohorts will be done by estimating rate ratios using the Poisson regression with 95% CI, adjusting for potential confounders. The logarithm of follow-up time will be used as an offset in the model. However, if the Poisson assumptions are violated, another modeling option can be used (e.g. zero-inflated model, negative-binomial, other link function with or without data transformation).

As described in section 9.7.1, the main Poisson models will be adjusted for unbalanced PS-variables at baseline. In addition, as the event rate can increase as a function of time since the start of follow-up, year of follow-up will be additionally adjusted for.
In country-specific analyses, cost outcomes can be additionally analyzed separately by type: e.g. CV vs. non-CV, diabetes-related vs. non-diabetes related costs. The obtained results of these additional analyses will not be included in the combined results of the multi-country study.

[redacted]

[redacted]

9.7.7 Combining effectiveness results from countries: meta-analyses

Meta-analyses will be performed for the primary and secondary effectiveness outcomes and the safety outcomes, to combine individual country-level results. In the main meta-analyses, all countries will be included. In additional analyses, several subgroups of countries (e.g. Nordic, European, and Asian countries separately) will be included.

Prior to conducting meta-analyses, heterogeneity across the countries in terms of study design and conduct will be assessed by considering deviations from this master protocol for each country. Specifically, the variability in population characteristics (including data availability), study design and statistical methods used will be investigated. Statistical heterogeneity will be subsequently assessed using:

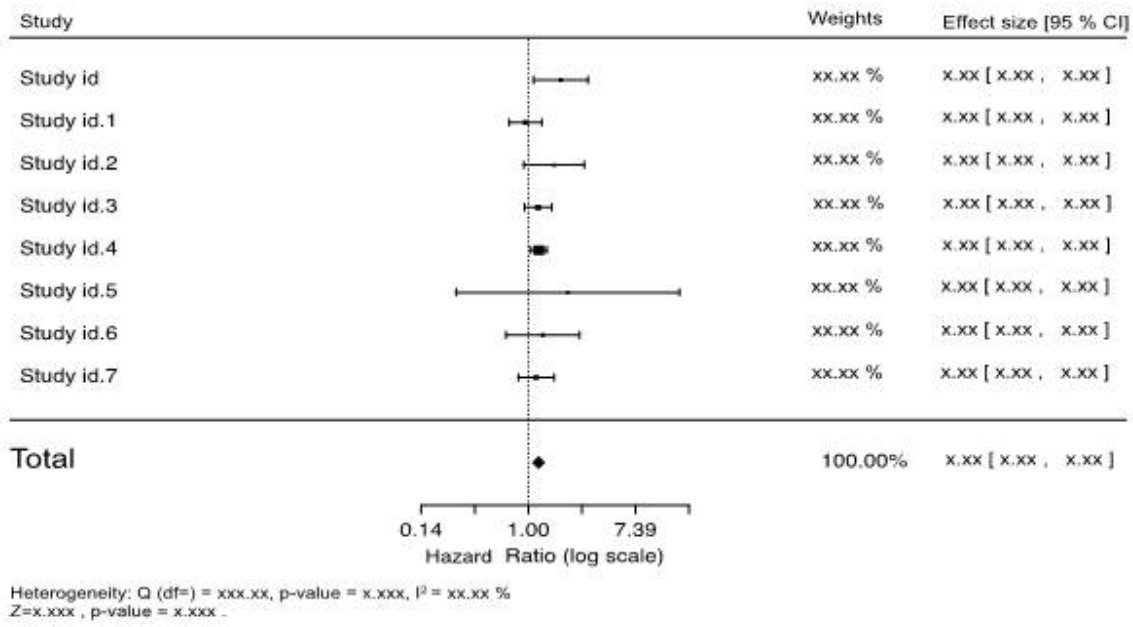
- The estimated total heterogeneity.
- The Chi-squared test (significance level: 0.1).
- I^2 statistic (0% to 40%: might not be important; 30% to 60%: may represent moderate heterogeneity; 50% to 90%: may represent substantial heterogeneity; 75% to 100%: considerable heterogeneity).

If high levels of heterogeneity exist (I^2 >=50% or P<0.1), the study design and characteristics in the included country-specific studies will be reviewed and the possible source of heterogeneity discussed. Additional meta-analyses may also be performed excluding countries leading to high heterogeneity.
Results of the meta-analyses will be derived using a random-effects model, and will be presented in a forest plot including the following information:

- Study identifiers.
- Effect size and 95% CI for each study included in the analysis (on a log scale).
- The weights allocated to each study.
- The combined estimated HR with 95% CI (on a log scale).
- The estimated amount of heterogeneity.
- Value of the Chi^2 test statistic, with the number of degrees of freedom and p-value.
- The proportion of variability explained by heterogeneity (I^2).
- The value of the Z score for overall effect with the p-value.

An example figure is given below in Figure 12

Figure 12 Illustration of meta-analysis results.



CI=confidence interval.

9.7.8 Presenting the HCRU and cost of care results from countries descriptively

HCRU and cost of care outcomes (Table 4) will be presented descriptively by country, but not combined by meta-analysis methods.

9.7.9 Adjustment for multiple comparisons

Will not be applied.

9.8 QUALITY CONTROL

9.8.1 Country-specific studies: localized protocols, localized SAPs, country-specific analyses, and country-specific study reports

Each study country will localize this master protocol (Figure 2). For the localized protocols, localized SAPs, country-specific analyses, and country-specific study reports, quality control procedures will be described in each localized protocol.

9.8.2 Multi-country study: master protocol, master SAP, analyses with aggregate-level results from the countries, and final report

Master protocol, master SAP, and multi-country report
All key study documents of the multi-country study, the master protocol, the master SAP, and the multi-country report will undergo quality-control review, senior scientific review, and editorial review.
The study will be conducted as specified in this master protocol. The principal investigator, the co-investigators and the sponsors of the study must approve all revisions to the master protocol. All changes are documented as master protocol amendments.

This master protocol has been written following the Code of Conduct by the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) [R18-3511]. The protocol also follows the key elements of the Guideline for Good Pharmacoepidemiology Practices by International Society for Pharmacoepidemiology. [redacted] the principal investigator, co-investigators, the Sponsors and individuals acting on their behalf commit to adhere to the rules of the ENCePP Code of Conduct in their entirety.

Multi-country analyses: combining the country-specific results

For the multi-country analyses, [redacted] will receive aggregate-level data from each country, to be combined. Standard operating procedures of [redacted] will be used in the conduct of the combined multi-country analyses, including the meta-analysis and presenting results descriptively. These procedures include internal quality audits, rules for secure and confidential data storage, methods to maintain and archive project documents, and quality control procedures for programming. All work for the multi-country analyses will be subject to quality control and documentation procedures to make certain the multi-country report is accurate and thorough, and the multi-country analyses can be reproduced. If the data do not permit the meta-analysis as planned or if clarifying analyses are required, the missing or the additional information and results will be included in the multi-country report and the corresponding explanation given.

Quality control will also be performed on the retrieved aggregate data from each country, including controlling the inclusion of necessary results from each country. If data are missing or incorrect, the dataset is sent back to the local contract points for correction.

All programs for data management and data analyses for the combined multi-country analyses will be written by the study statistician(s). Quality control check of these programs will be carried out by a statistician other than the one who writes the program. All processes from data management leading to dissemination of study results will undergo quality control checks for programs, result tables, and written text. A detailed audit trail of all documents (programs, result tables, reports) along with quality control processes will be maintained.

9.9 LIMITATIONS OF THE RESEARCH METHODS

9.9.1 Data sources and variables

As this study will incorporate data sources from several countries, the availability, coverage, and validity of information will vary. For example, laboratory values will not be available in all countries, limiting the investigation of renal effectiveness outcomes. Further, the available laboratory values might be affected by information bias, if measurements are disproportionally frequently available for patients with decreased renal function. There will also be limitations specific to each data source. Some data sources might provide samples that have been selected on the basis of, for example, income status and capability to use healthcare services.

9.9.2 Identifying patients with T2DM

The validity of T2DM diagnoses will vary between countries due to varying recording practices. However, the participating countries shall define the most appropriate strategy of identifying patients with T2DM from the utilized data sources. The beginning of T2DM symptoms cannot be determined from the utilized data sources, whereas the date of diagnosis usually can be determined. Therefore, the duration of the disease cannot be determined.

9.9.3 Drug exposure

Secondary data sources do not include information on actual drug use patterns and, therefore, the drug exposure in this study may be misclassified. Further, the data sources might not include complete information on drug dispensations or other records of drug use. Prescription data is one further potential source for exposure misclassification as the dispensation date cannot be determined.

The drugs of interest are relatively new, and the available follow-up data will be limited to up to approximately six years for SGLT-2 inhibitors and up to approximately four years for empagliflozin (as clarified in section 9.2.1).

9.9.4 Bias and confounding

Confounding by indication cannot be completely excluded from this study although several approaches will be used to minimize confounding (section 9.12). Further, there will be residual confounding related to, for example, incomplete recording of diagnoses and potential lack of variables in the data sources to be utilized.

9.9.5 Analyses

This will be an observational study and, therefore the results will indicate a correlation between drug use and selected outcomes. Causality cannot be determined in observational studies. Potential heterogeneity in the results provided by the countries will be addressed with the random-effects model.

9.10 OTHER ASPECTS

None specified.

9.11 SUBJECTS

Not applicable as this is a cohort study.

9.11.1 Cases

Not applicable as this is a cohort study.

9.11.2 Controls

Not applicable as this is a cohort study.

9.12 BIAS

This study will include new users of empagliflozin, any SGLT-2 inhibitor, and any DPP-4 inhibitor. Immortal time bias [P07-06047] will be minimal as the follow-up will begin at the initiation of drug use. Adherence to drug use will be unknown. Potential bias from exposure misclassification will be addressed with grace periods and the ITT analyses.

Comparing the outcomes between users of empagliflozin/any SGLT-2 inhibitor and any DPP-4 inhibitor will decrease confounding by indication, as the comparison group will have a similar indication to drug use than the patients in the empagliflozin/SGLT-2 sub-cohorts [R18-3506]. [redacted]
The compared sub-cohorts will be matched on the basis of PS. The PS will be computed on the basis of a wide range of covariates related to sociodemographic characteristics, lifestyle, diabetes complications, other comorbidities, laboratory values, prior/concomitant use of other antidiabetic drugs, prior/concomitant use of other drugs, healthcare resource utilization covariates, and cost covariates. The PS matching will decrease the systematic differences between the sub-cohorts and, thus, decrease confounding.

10. PROTECTION OF HUMAN SUBJECTS

Per design, this non-interventional study utilizes secondary data. Thus, the study does not affect the treatment or health outcomes of the study individuals. The study individuals will not be contacted in any phase of the study.

10.1 COUNTRY-SPECIFIC STUDIES

For the country-specific studies, each localized protocol will detail the protection of human subjects, including applicable applications for ethical approval and data protection according to national legislation.

10.2 MULTI-COUNTRY STUDY

For the multi-country study [redacted] will receive exclusively aggregate-level results from each country, in which study individuals in the countries cannot be identified. [redacted] will not have access to the individual-level data at any time of the study. Thus, no ethical approval is required for the multi-country study.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Not applicable. Data is anonymized, extracted, analyzed, validated and reported in aggregate. There is no potential that any employee of BI or agent working on behalf of BI will access individual patient data in which the patient may be identified during data compilation, data reporting or data analysis.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Manuscripts describing this work will be submitted for publication in peer-review journals. Results may also be submitted for presentation at scientific conferences.

13. REFERENCES

13.1 PUBLISHED REFERENCES

P07-06047 Suissa S. Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf 2007;16:241–9. doi:10.1002/pds.1357

P15-09840 Association AD. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes—2018. Diabetes Care 2018;41:S73–85. doi:10.2337/dc18-S008

P17-01701 Hsia DS, Grove O, Cefalu WT. An update on sodium-glucose co-transporter-2 inhibitors for the treatment of diabetes mellitus. Curr Opin Endocrinol Diabetes Obes 2017;24:73–9. doi:10.1097/MED.0000000000000311

P17-05895 Kosiborod M, Cavender MA, Fu AZ, et al. Lower Risk of Heart Failure and Death in Patients Initiated on Sodium-Glucose Cotransporter-2 Inhibitors Versus Other Glucose-Lowering Drugs: The CVD-REAL Study (Comparative Effectiveness of Cardiovascular Outcomes in New Users of Sodium-Glucose Cotransporter-2 Inhibitors). Circulation 2017;136:249–59. doi:10.1161/CIRCULATIONAHA.117.029190

P17-08964 Birkeland KI, Jørgensen ME, Carstensen B, et al. Cardiovascular mortality and morbidity in patients with type 2 diabetes following initiation of sodium-glucose co-transporter-2 inhibitors versus other glucose-lowering drugs (CVD-REAL Nordic): a multinational observational analysis. Lancet Diabetes Endocrinol 2017;5:709–17. doi:10.1016/S2213-8587(17)30258-9

P18-01920 Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. New England Journal of Medicine 2015;373:2117–28. doi:10.1056/NEJMoa1504720

P18-02804 Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. New England Journal of Medicine 2017;377:644–57. doi:10.1056/NEJMoa1611925

P18-04466 Scheen AJ. Cardiovascular Effects of New Oral Glucose-Lowering Agents: DPP-4 and SGLT-2 Inhibitors. Circ Res 2018;122:1439–59. doi:10.1161/CIRCRESAHA.117.311588

P18-06283 Kosiborod M, Lam CSP, Kohsaka S, et al. Cardiovascular Events Associated With SGLT-2 Inhibitors Versus Other Glucose-Lowering Drugs: The CVD-REAL 2 Study. J Am Coll Cardiol 2018;71:2628–39. doi:10.1016/j.jacc.2018.03.009

P18-09763 Wanner C, Heerspink H, Zinman B, et al. Empagliflozin and Kidney Function Decline in Patients with Type 2 Diabetes: A Slope Analysis from the EMPA-REG OUTCOME Trial. J Am Soc Nephrol 2018;29:2755–69. doi:10.1681/ASN.2018010103

P18-10828 International Diabetes Federation. IDF Clinical Practice Recommendations for managing Type 2 Diabetes in Primary Care. 2017.https://www.idf.org/component/attachments/attachments.html?id=1270&task=download (accessed 25 Sep 2018).

P18-10986 Ueda P, Svanström H, Melbye M, et al. Sodium glucose cotransporter 2 inhibitors and risk of serious adverse events: nationwide register based cohort study. BMJ 2018;363. doi:10.1136/bmj.k4365

P19-04731 Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. New England Journal of Medicine 2019;380:347–57. doi:10.1056/NEJMoa1812389

R13-4387 Kidney International Supplements | KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease | ScienceDirect.com. https://www.sciencedirect.com/journal/kidney-international-supplements/vol/3/issue/1 (accessed 13 Mar 2019)

R15-1988 Chen L, Magliano DJ, Zimmet PZ. The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives. Nature Reviews Endocrinology 2012;8:228–36. doi:10.1038/nrendo.2011.183

R17-1267 Wang SV, Schneeweiss S, Rassen JA. Optimal Matching Ratios in Drug Safety Surveillance: Epidemiology 2014;25:772–3. doi:10.1097/EDE.0000000000000148

R17-2088 Lin DY, Feuer EJ, Etzioni R, et al. Estimating medical costs from incomplete follow-up data. Biometrics 1997;53:419–34.

R18-3505 Persson F, Nyström T, Jørgensen ME, et al. Dapagliflozin is associated with lower risk of cardiovascular events and all-cause mortality in people with type 2 diabetes (CVD-REAL Nordic) when compared with dipeptidyl peptidase-4 inhibitor therapy: A multinational observational study. Diabetes Obes Metab 2018;20:344–51. doi:10.1111/dom.13077

R18-3506 Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep 2015;2:221–8. doi:10.1007/s40471-015-0053-5

R18-3507 R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

K18-3508 Лфтп НЬб Лшь Н-Об Зфкл О-Нб уе фдю Mortality and causes of death in a national sample of type 2 diabetic patients in Korea from 2002 to 2013. Cardiovascular Diabetology 2016;15:131. doi:10.1186/s12933-016-0451-0

R18-3509 Einarson TR, Acs A, Ludwig C, et al. Economic Burden of Cardiovascular Disease in Type 2 Diabetes: A Systematic Review. Value Health;21:881–90. doi:10.1016/j.jval.2017.12.019

R18-3510 Einarson TR, Acs A, Ludwig C, et al. Prevalence of cardiovascular disease in type 2 diabetes: a systematic literature review of scientific evidence from across the world in 2007-2017. Cardiovasc Diabetol 2018;17:83. doi:10.1186/s12933-018-0728-6

R18-3511 ENCePP. Code of Conduct. www.encepp.eu/code_of_conduct/ (accessed 13 May 2015).

R18-3512 Huang Y. Cost analysis with censored data. Med Care 2009;47:S115-119. doi:10.1097/MLR.0b013e31819bc08a

R18-3515 Dongsheng Yang, Jarrod E. Dalton. A unified approach to measuring the effect size between two groups using SAS. http://support.sas.com/resources/papers/proceedings12/335-2012.pdf

R18-3545 Austin PC. Assessing balance in measured baseline covariates when using many-to-one matching on the propensity-score. Pharmacoepidemiology and Drug Safety 2008;17:1218–25. doi:10.1002/pds.1674

13.2 UNPUBLISHED REFERENCES

Not applicable

ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

c26457869-01 Multi-country non-interventional study on the effectiveness of Empagliflozin in adult patients with type 2 diabetes in Europe and Asia; version 1.0; 29 November 2018

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

Study title: Multi-country non-interventional study on the effectiveness and safety of Empagliflozin in adult patients with type 2 diabetes in Europe and Asia

EU PAS Register® number: EUPAS27606
Study reference number (if applicable): 1245.195 (sponsor study number)

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[5]	☒	☐	☐	6
1.1.2 End of data collection[6]	☒	☐	☐	6
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☒	☐	☐	6
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	6
1.1.6 Final report of study results.	☒	☐	☐	6.0

Comments:

Progress reports are not conducted in this study, only interim reports.

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	7
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	9.2.2
2.1.4 Which hypothesis(-es) is (are) to be tested?	☒	☐	☐	8, 9.7
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.1

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

[6] Date from which the analytical dataset is completely available.

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	8
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	9.7
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☒	☐	☐	11

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	9.2.1
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	9.2.1
	4.2.2 Age and sex	☒	☐	☐	9.2.2
	4.2.3 Country of origin	☒	☐	☐	9.2.1
	4.2.4 Disease/indication	☒	☐	☐	9.2.2
	4.2.5 Duration of follow-up	☒	☐	☐	9.2.3
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	9.2.2

Comments:

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
5.1	Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.2.3 9.3.1
5.2	Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☒	☐	
5.3	Is exposure categorised according to time windows?	☒	☐	☐	9.3.1
5.4	Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	9.3.1
5.5	Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☒	☐	
5.6	Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	9.3.1

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.3.2
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☐	☒	☐	
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYs, health care services utilisation, burden of disease or treatment, compliance, disease management)	☒	☐	☐	9.3.2 9.7.5

Comments:

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	9.7, 9.9.4 9.12
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☐	☒	☐	
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.7, 9.9.3 9.12

Comments:

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☒	☐	☐	9.3.3 9.3.4, 9.7

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.4

Section 9: Data sources		**Yes**	**No**	**N/A**	**Section Number**
	9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.4
	9.1.3 Covariates and other characteristics?	☒	☐	☐	9.4
9.2	Does the protocol describe the information available from the data source(s) on:				
	9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage,prescriber)	☐	☐	☒	
	9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☐	☐	☒	
	9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☐	☐	☒	
9.3	Is a coding system described for:				
	9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9.3.1
	9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	9.3.2
	9.3.3 Covariates and other characteristics?	☒	☐	☐	9.3.3 Annex 7
9.4	Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☐	☐	☒	

Comments:

Questions 9.2 and 9.4: The utilized data sources and linkage methods will be further specified and described in the localized study protocols.

Section 10: Analysis plan		**Yes**	**No**	**N/A**	**Section Number**
10.1	Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.3.1 9.7
10.2	Is study size and/or statistical precision estimated?	☒	☐	☐	9.5
10.3	Are descriptive analyses included?	☒	☐	☐	9.7.2
10.4	Are stratified analyses included?	☒	☐	☐	9.7.3-9.7.6
10.5	Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7
10.6	Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7	Does the plan describe methods for handling missing data?	☒	☐	☐	9.8.2
10.8	Are relevant sensitivity analyses described?	☒	☐	☐	9.7.4 9.7.6

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.6
11.2 Are methods of quality assurance described?	☒	☐	☐	9.8
11.3 Is there a system in place for independent review of study results?	☐	☐	☒	

Comments:

Questions 11.1 and 11.2: The country-specific data management and quality control measures will be further specified and described in the localized study protocols.

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☐	☒	☐	
12.1.2 Information bias?	☒	☐	☐	9.9.4
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	9.9.4, 9.7.7
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	9.5.

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	10.1, 10.2
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	
13.3 Have data protection requirements been described?	☒	☐	☐	10.1

Comments:

This protocol has not yet undergone any ethical review procedure.

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	5.0

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	12.0

15.2	Are plans described for disseminating study results externally, including publication?	☒	☐	☐	12.0

Comments:

This is a Master protocol

ANNEX 3. REQUIREMENTS FOR THE LOCALIZED PROTOCOLS

MULTI-COUNTRY STUDY: Section of the master protocol	**COUNTRY-SPECIFIC STUDIES: To be specified in localized protocols**	
	Required specifications	**Optional additional specifications; the obtained results of these country-specific additions will not be included in the combined results of the multi-country study**
6 MILESTONES	Timelines for the country-specific studies	-
8 RESEARCH QUESTION AND OBJECTIVES	Specification of which objectives can be addressed	Additional objectives
9.2 SETTING	Marketing authorization dates for study drugs (if new country for which the date is not included in the master protocol) Study period and follow-up, depending on the dates of the marketing authorizations of the study drugs and data availability in each country's data sources	-
9.2.2 Study population	Used definition for T2DM	-
9.3 VARIABLES	Adaptation of all master variables i.e. locally used definitions for all variables, including coding systems or other definitions (for all variables under section 9.3) Availability of the variables listed in the master protocol (for all variables under section 9.3)	Definitions for additional variables, beyond the variables listed in the master protocol
9.3.1 Exposures: study drugs	Definition of a daily dose for the study drugs Exposure period	Additional SGLT-2 inhibitors, DPP-4 inhibitors or other drugs included in country-specific studies. Alternative comparator drugs Additional definitions for considering concomitant use of a SGLT-2 inhibitor and a DPP-4 inhibitor, of dose effects
9.3.2 Outcomes	-	Further outcomes
9.3.3 Covariates at baseline (baseline characteristics)	-	Additional covariates, to be used in additional analyses
9.3.4 Other variables required for analyses	-	Additional variables, to be used in additional analyses (e.g. for stratification)

MULTI-COUNTRY STUDY: Section of the master protocol	**COUNTRY-SPECIFIC STUDIES: To be specified in localized protocols**	
	Required specifications	**Optional additional specifications; the obtained results of these country-specific additions will not be included in the combined results of the multi-country study**
9.4 DATA SOURCES	Utilized data sources, including whether dispensation data or other records of the drug use are utilized and whether primary care is included	-
9.6 DATA MANAGEMENT	Detailed data management procedures including the used statistical software	-
9.7 DATA ANALYSIS	-	Any additional analyses, including e.g. additional models with additional adjusting variables, models considering time-dependent confounding or analyses considering dose effects
9.8 QUALITY CONTROL	Description of quality control procedures for the localized protocols, localized SAPs, country-specific analyses, and country-specific study reports	-
9.9 LIMITATIONS OF THE RESEARCH METHODS 9.12 BIAS	Any country-specific limitations or sources of bias	
10 PROTECTION OF HUMAN SUBJECTS	Details on the protection of human subjects, including applicable applications for ethical approval and data protection according to national legislation	

DPP-4=dipeptidyl peptidase-4; ICD-10=International Classification of Diseases, 10th revision; SAP=statistical analysis plan; SGLT-2=sodium-glucose cotransporter-2; T1DM=type 1 diabetes mellitus; T2DM=type 2 diabetes mellitus.

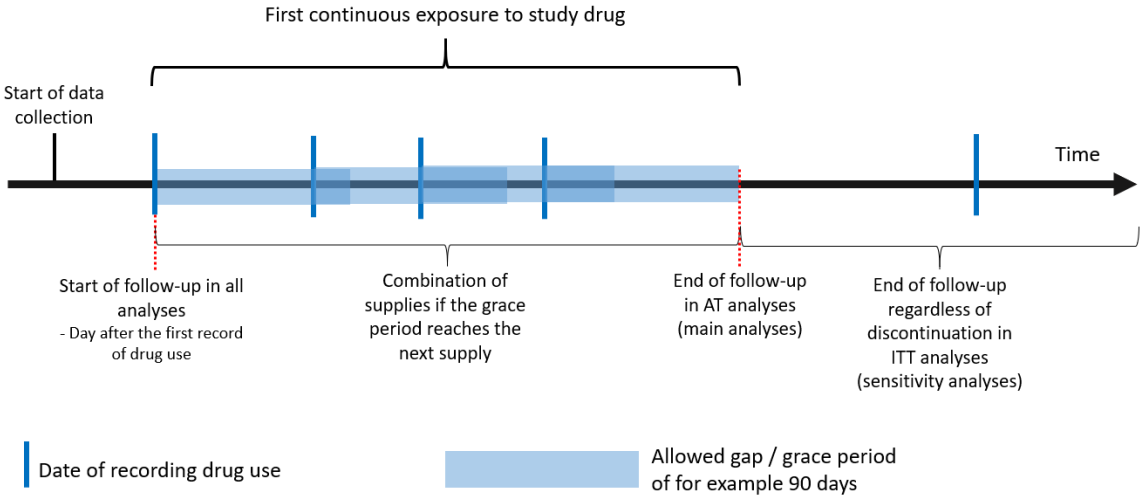
ANNEX 4. ALTERNATIVE DEFINITION OF EXPOSURE PERIODS (DRUG SUPPLY UNAVAILABLE)

If there is no sufficient data available to define drug supplies (such as data on prescriptions or dispensations) in the local data source, exposure periods of study drugs can be defined based on other available records of drug use, such as manual records. Drug exposure will be assumed to begin on the date of a record of drug use, or at on the first treatment day if specified.

If defining drug supplies is not feasible from the available data sources, only the grace period can be used in the definition of drug exposure. This definition of grace period should reflect local treatment practices, i.e. which gap is allowed between two consecutive records of drug use in continuous exposure to the drug. In ANNEX 4 Figure 1 below, a grace period of 90 days is used as an example. If a new record is not observed during the grace period, the exposure will end at the end of that grace period.

If a subsequent supply occurs before the previous supply has finished, the start of the subsequent supply might be shifted considering local treatment practices.

ANNEX 4 Figure. 1: Illustration of combining exposure periods with grace periods and the differences between the AT and ITT approaches.



AT=as-treated; ITT=intention to treat

ANNEX 5. DEFINITIONS FOR INCLUSION AND EXCLUSION CRITERIA

	ICD-9	ICD-10	ATC	Comments
Inclusion criteria				
Dispensation or any other record of empagliflozin, any SGLT-2 inhibitor, or any DPP-4 inhibitor use during the study period			Empagliflozin (including fixed-dose combinations with other drugs than DPP-4 inhibitors): A10BK03 A10BD20 Any SGLT-2 inhibitor (including fixed-dose combinations with a SGLT-2 inhibitor and another drug than DPP-4 inhibitor): A10BK01 A10BD15 A10BK03 A10BD20 A10BK02 A10BD16 A10BK04 A10BD23 Any DPP-4 inhibitor (including fixed-dose combinations with a DPP-4 inhibitor and another drug than SGLT-2 inhibitor): A10BH01 A10BD07 A10BD12 A10BH51	Includes both free and fixed-dose combination, also at index date
No dispensation or any other record of any other SGLT-2 inhibitor or DPP-4 inhibitor use during the preceding 12 months including at index date	-	-		

	ICD-9	ICD-10	ATC	Comments
			A10BH02 A10BD08 A10BH03 A10BD10 A10BH04 A10BD09 A10BD13 A10BH05 A10BD11	
Having a diagnosis of T2DM before the index date, based on ICD-10 codes or other available data	250.x0, 250.x2	E11	≥1 filled dispensation/any other record of metformin use at any time in the past: A10BA02 metformin A10BD15 Dapagliflozin and metformin A10BD20 Empagliflozin and metformin A10BD16 Canagliflozin and metformin A10BD23 Ertugliflozin and metformin A10BD07 Sitagliptin and metformin A10BD08 Vildagliptin and metformin A10BD10 Saxagliptin and metformin A10BD13 metformin and alogliptin2 A10BD11 Linagliptin and	Most appropriate time-window prior to the index date will be detailed in the localized protocols according to validated or standard definitions used in the country and data source.

	ICD-9	ICD-10	ATC	Comments
			metformin A10BD05 metformin and pioglitazone A10BD03 metformin and rosiglitazone	
Exclusion criteria				
Aged <18 years on the first dispensation date or date of the first record of empagliflozin, any SGLT-2 inhibitor or any DPP-4 inhibitor use	-	-	-	
Type 1 diabetes mellitus	250.x1 or 250.x3	E10	-	
Secondary diabetes	249.xx	E08, E09, E13	-	Diabetes mellitus due to underlying condition, drug or chemical induced diabetes mellitus, or other specified diabetes mellitus
Gestational diabetes	648.8x	O24.4, O24.1	-	Gestational diabetes mellitus or pre-existing type 2 diabetes mellitus, in pregnancy, childbirth, and the puerperium
Having a diagnosis of ESRD during the 12 months before the index date	At least 2 measurements separated by at least 30 days (but no more than 12 months) ≥2 of the following diagnosis or procedure codes (either inpatient or	At least 2 measurements separated by at least 30 days (but no more than 12 months) Defined as 2 codes (either inpatient or	Erythropoietin B03XA01 Calcium acetate V03AE07 Calcium acetate+Magnesium carbonate V03AE04 Lanthanum carbonate V03AE03 Sevelamer V03AE02 Sucroferric oxyhydroxide	Having any of the following: eGFR <15, diagnoses or procedure related to ESRD, or diagnoses or procedures related to kidney transplant

	ICD-9	ICD-10	ATC	Comments
	outpatient), separated by at least 30 days **Diagnoses:** 585.5x, 585.6x, V56.0x, V56.8x, V45.1x **Procedures:** 39.95, 54.98 Kidney transplant, defined as ≥1 of the following diagnosis or procedure codes (inpatient or outpatient): **Diagnoses:** V42.0x, 996.81 **Procedures:** 55.6x	outpatient), separated by at least 30 days **Codes include:** **Renal dialysis:** Z99.2 **Procedures related to renal dialysis:** 5A1D00Z, 5A1D60Z, 3E1M39Z **Chronic kidney disease, Stage V (for ESRD with no mention of dialysis):** N18.5 **End-stage renal disease (for ESRD with dialysis):** N18.6 **Encounter for dialysis:** Z49.31, Z49.32 **Defined as either 1 inpatient or 1 outpatient code** Codes include: **Kidney transplant status:**	V03AE05 Ferric citrate V03AE08 Except for patients with a diagnosis of primary hyperparathyroidism (ICD-10: E21.0)	

	ICD-9	ICD-10	ATC	Comments
		Z4822, Z940, OTS00ZZ, OTS10ZZ, OTY00Z x, OTY10Zx **Complications of transplanted kidney:** Z4822, 0TY00Z0, 0TY00Z1, 0TY00Z2, 0TY10Z0, 0TY10Z1, 0TY10Z2 **Transplant of the kidney:** 0TY00Zx, 0TY10Zx, Z940 (Exclude 0TS00ZZ, OTS10ZZ)		
<12 months of available data before the index date, and/or no complete history of drug dispensations/other records of the drug use during this period	-	-	-	For example, known gaps in the drug data, or known inadequate validity of the drug data
Missing or ambiguous data on age or sex	-	-	-	As available in the data source

DPP-4=dipeptidyl peptidase-4; eGFR=estimated glomerular filtration rate; ESRD=end-stage renal disease; ICD-9=International Classification of Diseases, 9th revision; ICD-10=International Classification of Diseases, 10th revision; SGLT-2=sodium-glucose cotransporter-2; T2DM=type 2 diabetes mellitus.

ANNEX 6. VARIABLE DEFINITIONS FOR OUTCOMES

Outcome	ICD-9	ICD-10	Comments
Primary effectiveness outcomes			
Hospitalization for heart failure*	428.x, 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.91, 404.03, 404.13, 404.93	I09.81, I11.0, I13.0, I13.2, I50.1x, I50.2x, I50.3x, I50.4x, 150.9	
Any diagnosis of heart failure associated with healthcare encounters, including hospitalizations and specialist outpatient and primary care encounters, or dispensation or any other record of the high-ceiling diuretics (loop diuretics)	428.x, 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.91, 404.03, 404.13, 404.93	I09.81, I11.0, I13.0, I13.2, I50.1x, I50.2x, I50.3x, I50.4x, 150.9	**ATC for high -ceiling diuretics (loop diuretics):** C03C
Myocardial infarction	410.X, excluding 410.x2	I21.0x, I21.1x, I21.2x, I21.3x, I21.4x, I22.0x, I22.0x, I22.1x, I22.2x, I22.8x, I22.9x, I25.2x	
Stroke	430.xx, 431.xx, 433.x1, 434.xx (excluding 434.x0), 436.x	**Subarachnoid hemorrhage (SAH):** I60xx **Intracerebral hemorrhage (ICH):** I610, I611, I612, I613, I614, I615, I616, I618, I619 **Occlusion and stenosis of precerebral arteries with cerebral infarction:** I6300, I63011, I63012, I63013, I63019, I6302, I63031, I63032, I63033, I63039, I6309, I6310, I63111, I63112, I63113, I63119, I6312, I63131, I63132, I63133, I63139, I6319, I6320, I63211,	

Outcome	ICD-9	ICD-10	Comments
		I63212, I63213, I63219, I6322, I63231, I63232, I63233, I63239, I6329, I6330, I6331x, I6332x, I6333x, I6334x, I6339, I6340, I6341x, I6342x, I6343x, I6344x, I6349, I6350, I6351x, I6352x, I6353x, I6354x, I6359, I636, I638, I639, I6789 **Acute, but ill-defined cerebrovascular events**: I6789	
All-cause mortality	Any code	Any code	
Secondary cardiovascular effectiveness outcomes			
CV mortality	390-459	I00 to I99	
Coronary revascularization procedure	a) PTCA, inpatient procedure: 00.66, 36.01, 36.02, 36.03, 36.05, 36.09 b) (no ICD-9) c) Stenting, inpatient procedure: 36.06, 36.07 d) CABG, inpatient procedure: 36.1x, 36.2x e) (no ICD-9)	**a) PTCA** (no ICD-10 available) **b) coronary bypass:** 02.10083, 02.1008x, 02.1009x, 02.100Ax, 02.100Jx, 02.100Kx, 02.100Zx, 02.10483, 02.10488, 02.10489, 02.1048x, 02.1049x, 02.1049x, 02.104Ax, 02.104Jx, 02.104Kx, 02.104Zx, 02.1108x, 02.1109x, 02.110Ax, 02.110Jx, 02.110Kx, 02.110Zx, 02.1148x, 02.1149x, 02.114Ax, 02.114Jx, 02.114Kx, 02.114Zx, 02.1208x, 02.1209x, 02.120Ax, 02.120Jx, 02.120Kx, 02.120Zx, 02.1248x, 02.1249x, 02.124Ax, 02.124Jx, 02.124Kx, 02.124Zx, 02.1308x, 02.1309x, 02.130Ax, 02.130Jx, 02.130Kx, 02.130Zx, 02.1348x, 02.1349x, 02.134Ax, 02.134x, 02.134Kx, 02.134Zx, 02.1K0Zx, 02.1K4Zx, 02.1L08X, 02.1L09x, 02.IL0Ax, 02.1L0Jx, 02.IL0Kx, 02.1l0Zx, 02.1L0Zx, 02.1L48X, 02.1L49x,	Having diagnoses or procedures related to a) PTCA, b) coronary bypass, c) stenting, d) CABG, or e) transmyocardial revascularization

Outcome	ICD-9	ICD-10	Comments
		02.1L4xx, 02.7004x, 02.7006x, 02.7007x, 02.700Dx, 02.700Ex, 02.700Fx, 02.700Gx, 02.700Tx, 02.700Zx, 02.703xx, 02.704xx, 02.7104x, 02.7105x, 02.7106x, 02.7107x, 02.710Dx, 02.710Ex, 02.710Fx, 02.710Gx, 02.710Tx, 02.710Zx, 02.7134x, 02.7135x, 02.7136x, 02.7137x, 02.713Dx, 02.713Ex, 02.713Fx, 02.713Gx, 02.713Tx, 02.713Zx, 02.7144x, 02.7145x, 02.7146x, 02.7147x, 02.714Dx, 02.714Ex, 02.714Fx, 02.714Gx, 02.714Tx, 02.714Zx, 02.7204x, 02.72056, 02.72066, 02.7206Z, 02.72076, 02.7207Z, 02.720Dx, 02.720Ex, 02.720Fx, 02.720Tx, 02.720Zx, 02.7234x, 02.7235x, 02.7236x, 02.7237x, 02.7237x, 02.723Dx, 02.723Ex, 02.723Fx, 02.723Gx, 02.723Tx, 02.723Tx, 02.723Zx, 02.723Zx, 02.7244x, 02.7245x, 02.7246x, 02.7247x, 02.724Dx, 02.724Ex, 02.724Fx, 02.724Gx, 02.724Tx, 02.7304x, 02.7304x, 02.7305x, 02.7306x, 02.7307x, 02.730Dx, 02.730Ex, 02.730Fx, 02.730Gx, 02.730Tx, 02.730Zx, 02.7334x, 02.7335x, 02.7336x, 02.7337x, 02.733Dx, 02.733Ex, 02.733Fx, 02.733Gx, 02.733Tx, 02.733Zx, 02.7344x, 02.7345x, 02.7346x, 02.7347x, 02.734Dx, 02.734Ex, 02.734Fx, 02.734Gx, 02.734Tx, 02.734Zx, 02.C00Zx, 02.C03Zx, 02.C04Zx, 02.C10Zx, 02.C13Zx, 02.C14Zx, 02.C20Zx, 02.C23Zx, 02.C24Zx, 02.C30Zx, 02.C33Zx, 02.C34Zx	

Outcome	ICD-9	ICD-10	Comments
		c) Stenting: 02ю7004чб 02ю7005чб 02ю7006чб 02ю7007чб 02ю700Вчб 02ю700Учб 02ю700Ачб 02ю700Пчб 02ю700Ечб 02ю700Ячб 02ю7034чб 02ю7035чб 02ю7036чб 02ю7037чб 02ю703Вчб 02ю703Учб 02ю703Ачб 02ю703Пчб 02ю703Ечб 02ю703Ячб 02ю7044чб 02ю7045чб 02ю7046чб 02ю7047чб 02ю704Вчб 02ю704Учб 02ю704Ачб 02ю704Пчб 02ю704Ечб 02ю704Ячб 02ю7104чб 02ю7105чб 02ю7106чб 02ю7107чб 02ю710Вчб 02ю710Учб 02ю710Ачб 02ю710Пчб 02ю710Ечб 02ю710Ячб 02ю7134бб 02ю7134Яб 02ю7135бб 02ю7135Яб 02ю7136бб 02ю7136Яб 02ю7137бб 02ю7137Яб 02ю713Вбб 02ю713ВЯб 02ю713Убб 02ю713УЯб 02ю713Абб 02ю713АЯб 02ю713Пбб 02ю713ПЯб 02ю713Ебб 02ю713ЕЯб 02ю713Ябб 02ю713ЯЯб 02ю7144бб 02ю7144Яб 02ю7145бб 02ю7145Яб 02ю7146бб 02ю7146Яб 02ю7147бб 02ю7147Яб 02ю714Вбб 02ю714ВЯб 02ю714Убб 02ю714УЯб 02ю714Абб 02ю714АЯб 02ю714Пбб 02ю714ПЯб 02ю714Ебб 02ю714ЕЯб 02ю714Ябб 02ю714ЯЯб 02ю7204бб 02ю7204Яб 02ю7205бб 02ю7205Яб 02ю7206бб 02ю7206Яб 02ю7207бб 02ю7207Яб 02ю720Вбб 02ю720ВЯб 02ю720Убб 02ю720УЯб 02ю720Абб 02ю720АЯб 02ю720Пбб 02ю720ПЯб 02ю720Ебб 02ю720ЕЯб 02ю720Ябб 02ю720ЯЯб 02ю7234бб 02ю7234Яб 02ю7235бб 02ю7235Яб 02ю7236бб 02ю7236Яб	

Outcome	ICD-9	ICD-10	Comments
		02ю72376б 02ю7237Яб 02ю723Вбб 02ю723ВЯб 02ю723Убб 02ю723УЯб 02ю723Абб 02ю723АЯб 02ю723Пбб 02ю723ПЯб 02ю723Ебб 02ю723ЕЯб 02ю723Ябб 02ю723ЯЯб 02ю72446б 02ю7244Яб 02ю7245бб 02ю7245Яб 02ю72466б 02ю7246Яб 02ю7247бб 02ю7247Яб 02ю724Вбб 02ю724ВЯб 02ю724Убб 02ю724УЯб 02ю724Абб 02ю724АЯб 02ю724Пбб 02ю724Ебб 02ю724ЕЯб 02ю724Ябб 02ю724ЯЯб 02ю73046б 02ю7304Яб 02ю7305бб 02ю7305Яб 02ю73066б 02ю7306Яб 02ю7307бб 02ю7307Яб 02ю730Вбб 02ю730ВЯб 02ю730Убб 02ю730УЯб 02ю730Абб 02ю730АЯб 02ю730Пбб 02ю730ПЯб 02ю730Ебб 02ю730ЕЯб 02ю730Ябб 02ю730ЯЯб 02ю73346б 02ю7334Яб 02ю7335бб 02ю7335Яб 02ю73366б 02ю7336Яб 02ю7337бб 02ю7337Яб 02ю733Вбб 02ю733ВЯб 02ю733Убб 02ю733УЯб 02ю733Абб 02ю733АЯб 02ю733Пбб 02ю733ПЯб 02ю733Ебб 02ю733ЕЯб 02ю733Ябб 02ю733ЯЯб 02ю73446б 02ю7344Яб 02ю7345бб 02ю7345Яб 02ю73466б 02ю7347бб 02ю7347Яб 02ю34Вбб 02ю734ВЯб 02ю734Убб 02ю734УЯб 02ю734Абб 02ю734АЯб 02ю734Пбб 02ю734ПЯб 02ю734Ебб 02ю734ЕЯб 02ю734Ябб 02ю734ЯЯб 02юС00Ябб 02юС03Ябб 02юС03ЯЯб 02юС04Ябб 02юС10Ябб	

Outcome	ICD-9	ICD-10	Comments
		02юC13Я6б 02юC13ЯЯб 02юC14Я6б 02юC20Я6б 02юC23Я6б 02юC23ЯЯб 02юC24Я6б 02юC30Я6б 02юC33ЯЯб 02юC34Я6б **d) CABG** (No code available) **e) Transmyocardial revascularization:** 021K0Z5, 021K4Z5, 021L0Z5, 021L4Z5, 02QA3ZZ, 02QA4ZZ, 02QB3ZZ, 02QB4ZZ, 02QC3ZZ, 02QC4ZZ	
Secondary renal effectiveness outcomes			
End-stage renal disease	a) ESRD, defined as having ≥2 of the following diagnosis or procedure codes (either inpatient or outpatient), separated by at least 30 days Diagnoses: 585.5x, 585.6x, V56.0x, V56.8x, V45.1x Procedures: 39.95, 54.98 b) Kidney transplant, defined as ≥1 of the following diagnosis or procedure codes	a) **ESRD,** defined as 2 codes (either inpatient or outpatient), separated by at least 30 days Codes include: **Renal dialysis:** Z99.2 **Procedures related to renal dialysis:** 5A1D00Z, 5A1D60Z, 3E1M39Z **Chronic kidney disease, Stage V (for ESRD with no mention of dialysis):** N18.5 **End-stage renal disease (for ESRD with dialysis):** N18.6 **Encounter for dialysis:** Z49.31, Z49.32	Having any of the following a) diagnoses or procedure related to ESRD, or b) diagnoses or procedures related to kidney transplant

Outcome	ICD-9	ICD-10	Comments
	(inpatient or outpatient): Diagnoses: V42.0x, 996.81 Procedures: 55.6x.	**Kidney transplant, defined as either 1 inpatient or 1 outpatient code** Codes include: **Kidney transplant status:** Z4822, Z940, OTS00ZZ, OTS10ZZ, OTY00Z x, OTY10Zx **Complications of transplanted kidney:** Z4822, 0TY00Z0, 0TY00Z1, 0TY00Z2, 0TY10Z0, 0TY10Z1, 0TY10Z2 **Transplant of the kidney:** 0TY00Zx, 0TY10Zx, Z940 (Exclude 0TS00ZZ, OTS10ZZ)	
Decline from normal kidney function (≥60ml/min/1,73m2) to abnormal kidney function (<60ml/min/1,73m2).			Categorized according to the following serum creatinine values (ml/min/1,73m2) **Normal kidney function**: ≥60 **Abnormal kidney function:** <60<15
Progression from normoalbuminuria to micro- or macroalbuminuria			Categorized according to the following albuminuria levels (mg/g) **Normoalbuminuria:** <30 **Micro- or macroalbuminuria:** ≥30>30
Secondary safety outcomes			
Bone fracture	**Hip fracture diagnosis:**	**Osteoporosis with pathological fracture:**	

Outcome	ICD-9	ICD-10	Comments
	ICD-9 diagnosis (820.xx, 733.14, 733.96) during hospitalization AND procedure code (ICD-9: 78.55, 79.05, 79.15, 79.25, 79.35, 79.65) during hospitalization **Pelvis fracture diagnosis:** 808.xx, 733.98 **Radius/ulna fracture diagnosis:** (ICD-9: 813.xx, 733.12) AND procedure (ICD-9: 78.53, 79.02, 79.12, 79.22, 79.32, 79.62) within 30 days of fracture date **Humerus fracture diagnosis:** (ICD-9: 812.xx, 733.11) AND procedure (ICD-9: 78.52, 79.01, 79.11, 79.21, 79.31, 79.61) within 30 days of fracture date	M80.021, M80.022, M80.029, M80.031, M80.032, M80.039, M80.051, M80.052, M80.059, M80.821, M80.822, M80.829, M80.831, M80.832, M80.839, M80.851, M80.852, M80.859, **Disorders of continuity of bone:** M84.350, M84.359, M84.421, M84.422, M84.429, M84.431, M84.432, M84.433, M84.52x, M84.53x, M84.55x, M84.62x, M84.63x, M84.65x, M84.75x, **Fracture of lumbar spine and pelvis:** S32.30x, S32.31x, S32.39x, S32.40x, S32.41x, S32.42x, S32.43x, S32.44x, S32.45x, S32.46x, S32.47x, S32.48x, S32.49x, S32.50x, S32.51x, S32.59x, S32.60x, S32.61x, S32.69x, S32.81x, S32.82x, S32.89x, S32.9x **Femur fracture:** S72.01x, S72.02x, S72.03x, S72.04x, S72.05x, S72.06x, S72.09x, S72.11x, S72.12x, S72.13x, S72.14x, S72.21x, S72.22x, S72.23x, S72.24x, S72.25x, S72.26x, S79.00x, S79.01x, S79.09x **Radius/ulna fracture diagnosis:** S52.01x, S52.02x, S52.03x, S52.04x, S52.09x, S52.10x, S52.11x, S52.12x, S52.13x, S52.18x, S52.2x, S52.3x, S52.5x, S52.61x, S52.62x, S52.69x, S52.90x, S52.91x, S52.92x, S59.00x, S59.01x, S59.02x, S59.03x, S59.04x, S59.09x, S59.10x S59.11x, S59.12x, S59.13x, S59.14x, S59.19x, S59.20x, S59.21x, S59.22x, S59.23x, S59.24x, S59.29x	

Outcome	ICD-9	ICD-10	Comments
		Humerus fracture diagnosis: S42.20x, S42.21x, S42.22x, S42.23x, S42.24x, S42.25x, S42.26x, S42.27x, S42.29x, S42.3x, S42.4x, S49.01x, S49.02x, S49.03x, S49.04x, S49.09x, S49.11x, S49.12x, S49.13x, S49.14x, S49.19x **Femur fracture:** S72.01x, S72.02x, S72.03x, S72.04x, S72.05x, S72.06x, S72.09x, S72.11x, S72.12x, S72.13x, S72.14x, S72.21x, S72.22x, S72.23x, S72.24x, S72.25x, S72.26x, S79.00x, S79.01x, S79.09x	
Diabetic ketoacidosis	Inpatient diagnosis: 250.1x	E08.1, E09.1, E10.1, E13.1	
Severe hypoglycemia	**Any-position ED or primary inpatient ICD-9 diagnosis:** 251.0, 251.1x, 251.2x, or 250.8x. **Outcomes identified by 250.8x are not included if they co-occur with one of the following diagnoses:** 259.8, 272.7, 681.xx, 682.xx, 686.9, 707.1x, 707.2x, 707.8, 707.9, 709.3, 730.0x, 730.1x, 730.2x, 731.8.	E160, E161, E162	

Outcome	ICD-9	ICD-10	Comments
Lower-limb amputation	84.10-84.18	(not available)	
Acute kidney injury that requires dialysis	**Inpatient ICD-9 diagnosis:** 584.5x ARF with a lesion of tubular necrosis 584.6x ARF with a lesion of renal cortical necrosis 584.7x ARF with lesion of renal medullary [papillary] necrosis 584.8x ARF with another specified pathological lesion in the kidney 584.9x ARF, unspecified **AND any of the following inpatient codes (within the same claim):** 39.95 hemodialysis V45.1 renal dialysis status V56.0 extracorporeal dialysis V56.1 fitting and adjustment of a dialysis catheter	N17	
Secondary outcomes on healthcare resource utilization and cost of care			

Outcome	ICD-9	ICD-10	Comments
Healthcare resource utilization			Outpatient healthcare visits • Primary care visits, specialist outpatient care visits or office visits, excluding emergency care, according to data availability • Emergency department visits Inpatient healthcare visits • Hospital admissions • Inpatient days • Length of stay Drug use in outpatient and inpatient care • Dispensations/other records of the drug use
Cost of care			Costs associated with healthcare resource utilization outcomes

ARF=acute renal failure; CABG=coronary artery bypass grafting; CV=cardiovascular; eGFR=estimated glomerular filtration rate; ED=emergency department; ESRD=end-stage renal disease; ICD-9=International Classification of Diseases, 9th revision; ICD-10=International Classification of Diseases, 10th revision; PTCA=percutaneous transluminal coronary angioplasty.

ANNEX 7. DEFINITIONS FOR COVARIATES AT BASELINE

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Sociodemographic characteristics (closest to index date)				
Age	By year	-	-	-
	By category (18-54, 55-64, 65-74, 75+)	-	-	-
Sex	female, male	-	-	-
Race	Categorization as applicable in the country	-	-	-
Socioeconomic status	By category (low, intermediate, high). Can include e.g. educational level or income level, as deemed appropriate by the countries	-	-	-
Calendar time by 3-month block	By 3-month block	-	-	-
	By day within the 3-month block	-	-	-
Covariates related to lifestyle (closest to index date)				
Obesity	**Having:** a) recorded diagnoses/procedures related to obesity, or b) filled dispensations/other records of the drug use related to obesity care		Diagnoses: E66, excluding E66.3	A08A anti-obesity preparations, excl. diet products
Overweight	-	278.02, V85.2x	E66.3	-
Smoking	-	V15.82, 305.1x, 649.0x, 989.84	F17, Z72.0, T65.2	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Alcohol abuse or dependence	-	291.xx, 303.xx, 305.0x, 571.0x, 571.1x, 571.2x, 571.3x, 357.5x, 425.5x, E860.0x, V11.3x	F10, K70, G62.1, I42.6, Z71.4	-
Drug abuse or dependence	-	292.xx, 304.xx, 305.2x-305.9x, 648.3x	F11; F12, F13, F14, F15, F16, F18, F19, Z71.5	-
Diabetes complications (12 months preceding index date) *				
Diabetic retinopathy	-	362.0x	E103.1x, E103.2x, E103.3x, E103.4x, E103.5x, E113.1x, E113.2x, E113.3x, E113.4x, E113.5x	-
Diabetes with other ophthalmic manifestations	Diagnoses related to Diabetes with ophthalmic manifestation (without mention of other retinal disorders), Diabetic cataract, Diabetic glaucoma	250.5x (without 362.01-362.07), 366.41 (diabetic cataract), 365.44 (diabetic glaucoma)	H28, H42, E08.36, E08.39, E09.36, E09.39, E10.36, E10.39, E11.36, E11.39, E13.36, E13.39	-
Retinal detachment, vitreous hemorrhage, vitrectomy	Related diagnoses or procedures	Diagnoses: 361.9x, 379.23 Procedure: 14.7x	E103.5x, E113.1x, E113.2x, E113.3x, E113.4x, E113.5x	-
Retinal laser coagulation therapy	-	14.24, 14.34, 14.54	08.5x, 08.Qx	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Diabetic neuropathy	-	250.6x, 357.2x, 337.1	G99.0, G59.0, G63.2, E08.4, E09.4, E10.4, E11.4, E13.4	-
Diabetic nephropathy	-	250.4x, 583.81	E092.1x, E092.2x, E092.9x, E102.2x, E102.9x, E106.5x, E112.1x, E112.2x, E112.9x, E132.2x, E132.9x,	-
Hypoglycemia	Diagnoses related to Hypoglycemia, or Hypoglycemic events	Diagnoses: 251.1x, 251.20, 962.30 Also, 250.8x as long as none of the following codes are co-occurring diagnoses (i.e. same day): 259.8, 272.7, 681.xx, 682.xx, 686.9x, 707.xx, 709.3, 730.0-730.2, 731.8	E13649, E1369, E160, E161, E162	-
Hyperglycemia	-	790.29	R73.0x, R73.9	-
Disorders of fluid electrolyte and acid-base balance	-	276.xx	E870, E871, E872, E873, E874, E869, E860, E877.1, E877.0, E877.9, E875, E876, E878	-
Diabetic ketoacidosis		250.1x	E1010, E1310	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Hyperosmolar hyperglycemic nonketotic syndrome	-	250.2x	E110.0, E110.1, E130.0, E130.1	-
Diabetes with peripheral circulatory disorders	-	250.7x, 443.81	E105.x, E.115.x, E.116.x, E.135.x	-
Diabetic foot	-	707.1x	L97.90x, L97.91x, L97.92x, I70.23x, I70.24x, I70.33x, I70.34x, I70.43x, I70.44x, I70.53x, I70.54x, I70.63x, I70.64x, I70.73x, I70.74x, L97.10x, L97.11x, L97.12x, L97.20x, L97.21x, L97.22x, L97.30x, L97.31x, L97.32x, L97.40x, L97.41x, L97.42x, L97.50x, L97.51x, L97.52x, L97.80x, L97.81x, L97.82x, E116.21, E136.21, E106.21	-
Gangrene	-	785.4x	I70.2x, I70.3x, I70.4x, I70.5x, I70.6x, I70.7x, E105.2x, E115.2x, E135.2x	-
Lower extremity amputation	Diagnoses or procedures related to lower limb amputation	Diagnoses: V49.7x (excluding V49.76 or V49.77) Procedures: 84.10-84.17	Z89	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Osteomyelitis	-	730.xx	M860xx, M860x,	-
Skin infections	-	680.xx-686.xx	L02.0x, L02.1x, L02.2x, L02.4x, L02.5x, L02.6x, L02.7x, L02.8x, L02.9x, L03.0x, K12.2, L03.1x, L03.2x, L03.3x, L03.8, L03.9, L98.3, L04.0, L04.1, L04.2, L04.3, L04.8, L04.9, L01.0x, L01.1, L05.0x, L05.9x, L08.0, L88, L08.8x, L92.8, L98, B78.1, E832, L08.82, L08.89, L08.9	-
Erectile dysfunction	-	607.84	N52.0x, N52.1x, N52.3x, N52.8x, N52.9x	-
Diabetes with unspecified complication	-	250.9x	E118, E138, E108	-
Diabetes mellitus without mention of complications	-	250.0x	E109, E119, E139	-
Other comorbidities (12 months preceding index date)				
Hypertension	1 inpatient or 2 outpatient encounters with any diagnosis codes for hypertension	401.x – 405.x	I10, I16.0, I16.1, I11.9, I11.0, I12.9, I12.0, I13.10, I13.0, I13.11, I13.2, I15.0, I15.8, I15.1, N26.2, I15.2, I15.9	-
Hyperlipidemia	-	272.0x-272.4x	E780.x, E782x E784, E785	-
Ischemic heart disease	-	410.xx-414.xx	I20.0, I20.1x, I20.8x, I20.9x, I21.01, I21.02, I21.09, I21.19, I21.2x, I21.3x, I21.4x, I22.0,	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
			I22.1x, I22.2x, I22.8x, I22.9x, I24.0x, I24.1, I24.8x, I24.9x, I25.1x, I25.2x, I25.3x, I25.6x, I25.7x, I25.8x, I25.9x	
Acute MI	-	410.xx	I21.9x, I21.1x, I21.2x, I21.4x, I21.3x, I22.8x,	-
Acute coronary syndrome or unstable angina	-	411.xx	I20.0x, I24.0x, I24.1x, I24.8x, I24.9x, I25.1x, I25.7x	-
Old MI	-	412.xx	I25.2	-
Stable angina	-	413.xx	I20.1, I20.8, I20.9, I2.1x, I25.7x	-
Coronary atherosclerosis and other forms of chronic ischemic heart disease	-	414.xx	I25.1x, I25.7x, I25.8x, I25.3x, I25.5x, I25.6x, I25.9x	-
Other atherosclerosis	Diagnoses related to Arteriosclerotic cardiovascular disease, or Generalized and unspecified atherosclerosis	429.2x 440.9x	I25.1x, I70.90, I70.91	-
Previous cardiac procedure	Diagnoses or procedures related to CABG, PTCA, Stent, or Transmyocardial revascularization	00.66, 36.01, 36.02, 36.03, 36.04, 36.05, 36.06, 36.07, 36.09, 36.1x, 36.2x	(No ICD-10 available) (No ICD-10 available) (No ICD-10 available) Z95.5, Z98.61, Z95.1	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
History of CABG or PTCA		V45.81, V45.82		
Any stroke	-	430.xx, 431.xx, 433.xx, 434.xx, 436.xx	I63.5x, I63.6x, I63.8x, I63.9x, I65.0x, I65.1x, I65.2x, I66.0x, I66.1x, I66.2x, I66.3x, I66.8x, I66.9x	-
Ischemic stroke (with and without mention of cerebral infarction)	-	433.xx, 434.xx, 436.xx	I63.xx I65.xx I66.xx	-
Hemorrhagic stroke	-	430.xx, 431.xx	I60.xx, I61.xx,	-
Transient cerebral ischemia and related syndromes	-	435.xx	G45.x	-
Other cerebrovascular disease	-	432.xx, 437.xx	I00.xx, I0x.x, I3x.xx, I40.xx, I41.xx, I42.xx, I43.xx, I71.xx, I72.xx, I79.xx, I73.1x, I74.xx, I75.xx, I77.xx, I62.0x, I62.1x, I62.9x, I67.0, I67.2, I67.4, I67.5, I67.6, I67.7, I67.8, I67.9, I68.0, I68.2x, M30.xx, M31.xx, M32.1x, G45.4x, G46.0x, G46.1x, G46.2x, G46.3x, G46.4x, G46.5x, G46.6x, G46.7x, G46.8x,	-
Late effects of	-	438.xx	I69.0x, I69.1x, I69.2x, I69.3x,	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
cerebrovascular disease			I69.8x, I69.9x,	
Cerebrovascular procedure	Procedures related to Carotid bypass, or Cerebrovascular revascularization	39.28 00.61 - 00.65, 38.11, 38.12	(no codes available)	-
CHF	-	428.x, 398.91, 402.01, 402.11, 402.91, 404.01, 404.11, 404.91, 404.03, 404.13, 404.93	I09.81, I11.0, I13.0, I13.2, I50.1, I50.20, I50.21, I50.22, I50.23, I50.31, I50.32, I50.33, I50.40, I50.41, I50.42, I50.43, I50.9	-
Peripheral vascular disease or surgery	Recorded diagnoses related to *a)* peripheral vascular disease or procedures related to *b)* lower-extremity endarterectomy, stenting, angioplasty, or atherectomy, c) Lower-extremity bypass, OR d) Other peripheral vascular surgery	a) 440.20 – 440.24, 440.29 – 440.32, 440.3, 440.4, 443.9 b) 38.18, 38.19 c) 39.25, 39.29 d) 38.08, 38.09, 38.38, 38.39,38.48, 38.49, 39.5x, 39.9x	**Peripheral vascular disease:** I70.2x, I70.3x, I70.4x, I70.5x, I70.6x, I70.7x, I70.9x **Lower extremity endarterectomy, stenting, angioplasty, or atherectomy:** (no ICD-10) **Lower-extremity bypass:** (Aorta-iliac femoral bypass) 04.1x, 04.7x, **Other peripheral vascular surgery:** (no ICD-10)	-
Atrial fibrillation	-	427.3x	I48.0x, I48.1x, I48.2x, I48.3x, I48.4x, I48.9x	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Other cardiac dysrhythmia	-	427.xx, exclude 427.5x (cardiac arrest) and 427.3x	I48.0x, I48.1x, I48.2x, I48.3x, I48.4x, I48.9x, I49.0x, I49.4x, I49.1x, I49.3x, I49.5x, I49.8x, I49.9, R00.1, Exclude I46.2x, I46.8x, I46.9 (cardia arrest)	-
Cardiac conduction disorders	-	426.xx	I44.0x, I44.1x, I44.2x, I44.3x, I44.4x, I44.5x, I44.6x, I44.7x, I45.0x, I45.1x, I45.4x, I45.2x, I45.3x, I45.5x, I45.6x, I45.8x, I45.9x	-
Other CV disease	Including all registered CV events at any time prior to index date	390.xx -398.xx, 420.xx -425.xx, 441.xx -447.xx (except 442.1x, 443.81, 443.9x)	I00-I09, I30-I43, I71-I75, I77, I79, M30, M31 (Except I72.2, I79.8, I73.9)	-
Edema	-	782.3x	R60.0x, R60.1x, R60.9x	-
COPD	-	491.xx, 492.xx, or 496.xx	J44	-
Asthma	-	493.xx	J45.2x, J45.3x, J45.4x, J45.9xx	-
Obstructive sleep apnea	-	327.23	G47.33	-
Pneumonia	-	480.xx – 486.xx, 487.0x, 507.xx	J12.0x, J12.1x, J12.2x, J12.8x, J12.3x, J12.9x, J13, J18.1x, J15, J15.1x, J14, J15.4x, J15.3, J15.2x, J15.8x, J15.5x, J15.6x, A48.1x, J15.9x, J15.7x, J16.0x, J16.8x, B25,	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
			A37.0x, A37.8x, A7.9x, A22.1x, B44.0x, J17, B77.8x, J17, J18.0x, J18.9x, J10.0x, J11.0X, J12.9x, J69.0x, J69.1, J69.8x	
Renal dysfunction (non-diabetic)	Diagnoses related to Acute renal disease, or Chronic kidney disease, or Hypertensive nephropathy, or Miscellaneous Renal Insufficiency, or Dialysis	Acute renal disease 572.4x, 580.xx, 584.xx, 791.2x, 791.3x **Chronic kidney disease**: 585.xx **Hypertensive nephropathy**: 403.xx, 404.xx **Miscellaneous Renal Insufficiency**: 274.10, 440.1x, 442.1x, 453.3x, 581.xx, 593.xx, 753.0x, 753.3x, 866.00, 866.01, 866.1x **Dialysis Procedure:** 39.95, 54.98 **Diagnosis:** V56.0x, V56.8x,	Acute renal disease: K76.7, N00.x, N01.x, N08, N17.x, R82.x, **Chronic kidney disease:** N18.5, N18.6 **Hypertensive nephropathy**: I12.9, I12.0, I13.0, I13.10, I13.11, I13.2 **Miscellaneous Renal Insufficiency:** E082.1, E082.2, E082.9, N08, N02.8, N04.7, N04.8, N04.9, N28.83, N28.81, N28.1, N11.1, N13.0, N13.1, N13.5, N11.1, N13.8, N13.4, R80.2, N13.7, N13.71, N13.72x, N13.9, N28.9, N28.89, N28.82, N28.9, N29, Q60.0, Q60.1, Q60.2, Q60.3, Q60.4, Q60.5, Q60.6x, Q63.0, Q63.1, Q63.2, Q63.3, Q63.8, Q63.9, S37.001A, S37.002A, S37.009A, S37.011A, S37.012A, S37.019A, S37.021A, S37.022A, S37.029A **Dialysis**: Z99.2	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
		V45.1x		
Acute renal disease	-	572.4x, 580.xx, 584.xx, 791.2x, 791.3x	K76.7, N00.x, N01.x, N08, N17.x, R82.x,	-
Chronic renal insufficiency	-	582.xx, 583.xx, 585.xx, 586.xx, 587.xx	N03.0, N03.1, N03.2, N03.3, N03.4, N03.5, N03.6, N03.7, N03.8, N03.9, N05.0, N05.1, N05.2, N05.3, N05.6, N05.7, N05.8, N05.9, N06.0, N06.1, N06.2, N06.3, N06.4, N06.5, N06.6, N06.9, N07.3, N07.4, N07.5, N07.7, N07.8, N07.9, N08, N14.0, N14.1, N14.2, N14.3, N14.4, N15, N15.8, N15.9, N16, N17.1, N17.2, N18.1, N18.2, N18.3, N18.4, N18.5, N18.6, N18.9, N19, N26.1, N26.9, E092.1, E092.2, M32.14, M32.15, M35.04,	-
Chronic renal insufficiency without chronic kidney disease		582.xx, 583.xx, 585.9x, 586.xx, 587.xx	(no ICD-10)	
Chronic kidney disease Stage 1-2		(no ICD-9)	(no ICD-10)	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Chronic kidney disease stage 3-6		(no ICD-9)	(no ICD-10)	
Chronic kidney disease	-	585.xx	N18.5, N18.6	-
Chronic kidney disease stage 3-4	-	585.3x-585.4x	N18.3, N18.4	-
Hypertensive nephropathy	-	403.xx, 404.xx		-
Miscellaneous Renal insufficiency		274.10, 440.1x, 442.1x, 453.3x, 581.xx, 593.xx, 753.0x, 753.3x, 866.00, 866.01, 866.1x	E082.1, E082.2, E082.9, N08, N02.8, N04.7, N04.8, N04.9, N28.83, N28.81, N28.1, N11.1, N13.0, N13.1, N13.5, N11.1, N13.8, N13.4, R80.2, N13.7, N13.71, N13.72x, N13.9, N28.9, N28.89, N28.82, N28.9, N29, Q60.0, Q60.1, Q60.2, Q60.3, Q60.4, Q60.5, Q60.6x, Q63.0, Q63.1, Q63.2, Q63.3, Q63.8, Q63.9, S37.001A, S37.002A, S37.009A, S37.011A, S37.012A, S37.019A, S37.021A, S37.022A, S37.029A	
Dialysis		**Procedure:** 39.95, 54.98 **Diagnosis:** V56.0x, V56.8x, V45.1x	Z99.2	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Liver disease	-	**Diagnoses:** 070.xx, 570.xx-573.xx 456.0x-456.2x, 576.8x, 782.4x, 789.5x ICD-9 **Procedures:** 39.1x, 42.91	B18, I85.0, I85.9, I86.4, I98.2, K70.0-K70.3, K70.4, K70.9, K71.1, K71.3-K71.5, K71.7, K72.1, K72.9, K73, K74, K76.0, K76.2-K76.4, K76.5, K76.6, K76.7K76.8, K76.9, Z94.4	-
Osteoarthritis	Diagnoses related to polyarthrosis, coxarthrosis, gonarthrosis, arthrosis of first carpometacarpal joint, and another arthrosis	715.xx	M15.0, M15.1, M15.2, M15.3, M15.4, M15.8, M15.9, M16.0, M16.1, M16.2, M16.3, M16.4, M16.5, M16.6, M16.7, M16.9, M17.0, M17.1, M17.2, M17.3, M17.4, M17.5, M17.9, M18.0, M18.1, M18.2, M18.3, M18.4, M18.5, M18.9, M19.01, M19.02, M19.03, M19.04, M19.07, M19.11, M19.12, M19.13, M19.14, M19.17, M19.21, M19.22, M19.23, M19.24, M19.27, M19.90, M19.91, M19.92, M19.93, M19.9	-
Other arthritis, arthropathies and musculoskeletal pain	-	710.xx−714.xx, 716.xx−719.xx, 725.xx−729.xx (excluding 729.2x)]	M00-M08 (except M04.1), M11-M14, M22-M25, M32-M36, M43.3, M43.4, M43.5, M60-M72 (except M62.830), M75-M79, (except M79.2), R26.2, R29.898,	-
Dorsopathies	-	720.xx-724.xx	M43.2, M43.6, M45-M54, M62.830	-
Bone fractures	-	733.1x, 800.xx-	M80.0x, M80.8x, M84.4x, M84.5x,	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
		829.xx	M84.6x, M84.7x, M48.5x, S02.0X, S02.1x, S02. 11x, S02.2x, S02.3x, S02.9x, S02.4x, S06.5x, S02.6x, S02.9x, S06.33, S06.36, S06.4X, S06.5X, S06.6X, S06.8x, S06.9x, S12.9X, S12.0x,S12.1x, S12.2x, S12.3x, S12.4x, S12.5x, S12.6x, S12.8x, S12.9, S22.0x, S32.0x, S12.0x, S12.1x, S12.2x, S12.3x, S14.1, S12.0x, S12.1x, S12.2x, S12.3x, S14.1x, S22.0x, S22.2x, S22.3x, S22.4x, S22.5x, S24.1x, S32.0x, S32.3x, S32.6x, S32.8x, S32.9x, S34.1x, S32.0x, S32.4x, S32.5x, S34.1x, S42.0x, S42.1x, S42.2x, S42.3x, S42.4x, S42.9x, S49.1x, S52.0x, S52.1x, S52.2x, S52.3x, S52.5x, S52.6x, S52.9x, S59.0x, S59.2x, S62.0x, S62.1x, S62.2x, S62.3x, S62.5x, S62.6x, S62.9x, S72.0x, S72.1x, S72.2x, S72.3x, S72.4x, S72.8x, S72.9x, S79.0x, S79.1x, S82.0x, S82.1x, S82.2x, S82.3x, S82.4x, S82.5x, S82.6x, S82.8x, S82.9x, S89.0x, S89.1x, S89.2x, S89.3x, S92.0x, S92.1x, S92.2x, S92.3x, S92.4x, S92.5x, S92.8x, S92.9x, S99.0x, S99.1x, S99.2x, T14.8	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Falls	Diagnoses related to accidental falls or history of falls or dispensations/other records of the drug use of osteoporosis drugs	a) E880.0x-E888.9x b) V15.88	a) W00-W19 b) Z91.81	M05BA Bisphosphonates M05BB Bisphosphonates, combinations1 G03XC01 Raloxifene H05AA02 Teriparatide H05BA Calcitonin preparations
Osteoporosis	-	733.0x	M80, M81	-
Hyperthyroidism	-	242.1x, 242.3x, 242.9x	E05.1, E05.2, E05.9	-
Hypothyroidism	-	243.xx, 244.xx	E00, E01.8, E02, E03, E89.0	-
Other disorders of the thyroid gland	Excluding hyperthyroidism and hypothyroidism	240.xx – 246.xx,	E01-E07	-
Depression	-	293.83, 296.2x. 296.3x, 298.0x, 300.4x, 309.0x, 309.1x, 309.28, 311.xx	F03.6, F32 (except F32.8), F33 (except F33.8), F34.1, F43.21, F43.23	-
Anxiety	-	293.84, 300.0x, 300.2x, 300.3x, 309.24, 308.0x, 309.81	F06.4, F40-F42 (except F42.4), F43.0, F43.1, F43.22	-
Sleep disorder	-	307.4x, 327.0x, 327.2x 780.5x, 347.xx	F51 (except F51.13), G47 (except G47.4-G47.6)	-

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Dementia	-	290.xx, 294.xx, 330.xx, 331.xx	F01-F04, F05.1, F06.0, F06.1, F06.8, E75.0, E75.1, E75.23, E75.25, E75.29, E75.4, F84.2, G30, G31, G93.7, G93.9,	-
Delirium	-	290.11, 290.3x, 290.41, 291.0x, 292.81, 293.xx, 348.3x, 349.82	F01.51, F03.90, F05, F1x.121, F1x.221 (except F17.221), F1x.231, F1x.921, G93.4, I67.83, G92	-
Psychosis	-	290.8x, 290.9x, 295.xx, 297.xx, 298.xx, 299.xx, 780.1x	F03.90, F20, F22-F29, F32.2, F32.3, F33.3, F44.89, F84 (except F84.2), R44.0, R44.2, R44.3	-
Frailty	The sum of weights related to pre-specified conditions (as in Kim et al. J Gerontol A Biol Sci Med Sci. 2018 Jun 14;73(7):980-987)			-
Foot ulcer		707.1x, 440.23	I70.2x, I70.3x, I70.4x, I70.5x, I70.6x, I70.7x, L97.1x, L97.2x, L97.3x, L97.4x, L97.5x, L97.8x, L97.9x,	
Cellulitis or abscess of toe		680.7x, 682.7x	L02.61, L02.62, L02.63, L03.11, L03.12	
Hypertension 1 code		401.xx-405.xx	I10.x, I11.x, I13.x, I15.x, I16.x, N26.2	
Imaging		88.48, 88.77	B3.4x, B4.0x, B4.1x, B4.4x,	
Glaucoma or cataracts		365, 366	(not defined)	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
LABORATORY VALUES				
HbA1c (%)	As available in the data source			
Low-density lipoprotein (LDL) level (mg/dl)	As available in the data source			
High-density lipoprotein (HDL) level (mg/dl)	As available in the data source			
Total cholesterol (mg/dl)	As available in the data source			
Triglyceride level (mg/dl)	As available in the data source			
Creatinine (mg/dl)	As available in the data source			
Glomerular Filtration Rate (mL/min/1,73m2)	As available in the data source			
BUN (mg/dl)	As available in the data source			
BNP	As available in the data source			
NT-proBNP	As available in the data source			
Prior/concomitant use of other antidiabetic drugs (12 months preceding index date and at index date)				
N antidiabetic substances at index date	N antidiabetic substances (including study drugs listed in **Table 3** and other antidiabetic substances) on the day of initiation of the study drug			Unique substances in the ATC class A10 drugs used in diabetes. Fixed-dose combination products with multiple

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
				substances count as multiple drugs, e.g. empagliflozin+metformin results in 2 substances
Naive new use of antidiabetic drugs	No use of any antidiabetic drug in the prior 12 months			No use of drugs in the ATC class A10 drugs used in diabetes, apart from the index drug
Initiation of the study drug (empagliflozin/any SGLT-2 inhibitor/any DPP-4 inhibitor) as monotherapy	Study drug (empagliflozin/any SGLT-2 inhibitor/ any DPP-4 inhibitor) initiated as monotherapy, i.e., no filled dispensation/ no other records of the drug use of any anti-diabetic drugs in the 12 months prior to drug initiation AND no concomitant initiation of any anti-diabetic drugs at index date			For study drugs: see **Table 3**, excluding the listed fixed-dose combinations For any anti-diabetic drugs: A10 drugs used in diabetes
Dual therapy with metformin (without the use of other antidiabetic drugs)	≥2 filled dispensations/other records of the drug use of metformin in the 183 days prior to index date and current use (i.e. days' supply overlap) at index date AND no filled dispensations or other records of the drug use of other drugs used in diabetes in the 12 months prior to the index date			A10BA02 metformin For other antidiabetic drugs: A10 drugs used in diabetes

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Concomitant initiation or current use of other antidiabetic drugs	≥1 filled dispensation/any other record of the drug use of other antidiabetic drugs at index date of the drug of interest OR ≥1 filled dispensation/any other record of the drug use of other antidiabetic drugs 12 months prior to index date with current use at index date (i.e. days of supply overlap with index date)			A10 drugs used in diabetes and The drugs/drug classes listed below* separately
Past use of other antidiabetic drugs	≥1 filled dispensation/any other record of the drug use of other antidiabetic drugs 12 months prior to index date without current use at index date (i.e. NO overlapping days of supply with index date)			A10 drugs used in diabetes and The drugs/drug classes listed below* separately
Metformin*	(excluding fixed-dose combinations with metformin and the study drugs listed in Table 3; including fixed-dose combinations with metformin and other drugs)			A10BA02 metformin A10BD05 metformin and pioglitazone A10BD03 metformin and rosiglitazone3
Sulfonylureas 2nd generation*	Glimepiride, glipizide, glibornuride (also known as glyburide)			A10BB12 glimepiride A10BB07 glipizide
Glucagon-like peptide-1 receptor agonists*	-			A10BJ Glucagon-like peptide-1 analog
Thiazolidinediones*	-			A10BG Thiazolidinediones A10BD03 metformin and

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
				rosiglitazone A10BD04 glimepiride and rosiglitazone A10BD05 metformin and pioglitazone A10BD06 glimepiride and pioglitazone
Meglitinides*	-			A10BX02 repaglinide A10BX03 nateglinide
Insulin*	-			A10A insulins and analogs
Alpha-glucosidase inhibitors*	-			A10BF Alpha-glucosidase inhibitors
Any use of pramlintide	≥1 filled dispensation/any other record of the drug use of pramlintide 12 months prior to the index date			
Any use of 1st generation sulfonylureas	≥1 filled dispensation/any other record of the drug use of 1st generation sulfonylureas 12 months prior to the index date (acetohexamide, chlorpropamide, tolazamide, tolbutamide) in the 12-month period prior to the day of initiation of the drug of interest			A10BB03 tolbutamide
Prior use of other drugs (12 months preceding index date				
Angiotensin-converting-enzyme inhibitor	≥1 filled dispensation/any other record of the drug use			C09A Angiotensin-converting-enzyme inhibitors, plain

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
				C09B Angiotensin-converting-enzyme inhibitors, combinations
Angiotensin II receptor blocker	≥1 filled dispensation/any other record of the drug use			C09C angiotensin ii antagonists, plain C09D angiotensin ii antagonists, combinations
Beta-blocker	≥1 filled dispensation/any other record of the drug use			C07 beta-blocking agents
Calcium channel blocker	≥1 filled dispensation/any other record of the drug use			C08 calcium channel blockers C07FB Beta-blocking agents and calcium channel blockers C09BB Angiotensin-converting-enzyme inhibitors and calcium channel blockers C09DB Angiotensin II antagonists and calcium channel blockers
Thiazides	≥1 filled dispensation/any other record of the drug use			C03A low-ceiling diuretics, thiazides
Loop diuretics	≥1 filled dispensation/any other record of the drug use			C03C high-ceiling diuretics
Other diuretics	≥1 filled dispensation/any other record of the drug use			C03D potassium-sparing agents

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Nitrates	≥1 filled dispensation/any other record of the drug use Including glyceryl trinitrate (also known as nitroglycerin), isosorbide dinitrate, isosorbide mononitrate, ranolazine			C01DA02 glyceryl trinitrate C01DA08 isosorbide dinitrate C01DA14 isosorbide mononitrate C01EB18 ranolazine
Other hypertension drugs	≥1 filled dispensation/any other record of the drug use Including doxazosin, eplerenone, prazosin, terazosin, clonidine, guanabenz (no ATC code), guanadrel (no ATC code), guanethidine, guanfacine, hydralazine, methyldopa, metirosine (also known as metyrosine), reserpine, minoxidil, aliskiren			C02CA04 doxazosin C03DA04 eplerenone C02CA01 prazosin G04CA03 terazosin C02AC01 clonidine C02AC02 guanfacine C02DB02 hydralazine C09XA02 aliskiren C09XA52 aliskiren and hydrochlorothiazide
Digoxin	≥1 filled dispensation/any other record of the drug use			C01AA05
Valsartan and sacubitril	≥1 filled dispensation/any other record of the drug use			C09DX04
Antiarrhythmic drugs	≥1 filled dispensation/any other record of the drug use			C01B antiarrhythmics, class I and III
COPD or asthma medications	≥1 filled dispensation/any other record of the drug use Including: Fluticasone/salmeterol (fixed combo), budesonide/formoterol (fixed combo), mometasone/formoterol (fixed			R03AK06 salmeterol and fluticasone R03AK07 formoterol and budesonide R03AC13 formoterol

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
	combo), aformoterol, formoterol, salmeterol, salbutamol (also known as albuterol/levalbuterol),orciprenaline (also known as metaproterenol), pirbuterol, terbutaline, ipratropium, tiotropium, theophylline, montelukast, zafirlukast, zileuton (no ATC code), aclidinium, indacaterol, olodaterol, umeclidinium			R03AC12 salmeterol R03AC02 salbutamol R03CC03 terbutaline1 R01AX03 ipratropium bromide2 R03BB01 ipratropium bromide R03BB04 tiotropium bromide R03DA04 theophylline R03DC03 montelukast R03DC01 zafirlukast2,3,4 R03BB05 aclidinium bromide R03AL05 formoterol and aclidinium bromide R03AC18 indacaterol R03AC19 olodaterol R03BB07 umeclidinium bromide R03AL01 fenoterol and ipratropium bromide1,2,3 R03AL02 salbutamol and ipratropium bromide2 R03AL03 vilanterol and umeclidinium bromide R03AL04 indacaterol and glycopyrronium bromide

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
				R03AL05 formoterol and aclidinium bromide R03AL06 olodaterol and tiotropium bromide R03AL08 vilanterol, umeclidinium bromide, and fluticasone furoate R03AL09 formoterol, glycopyrronium bromide, and beclometasone
Statin	≥1 filled dispensation/any other record of the drug use			C10AA HMG CoA reductase inhibitors C10BA HMG CoA reductase inhibitors in combination with other lipid modifying agents C10BX HMG CoA reductase inhibitors, other combinations
PCSK-9 inhibitors	≥1 filled dispensation/any other record of the drug use			C10AX13 Evolocumab Bococizumab C10AX14 Alirocumab
Other lipid-lowering drugs, excluding statins	≥1 filled dispensation/any other record of the drug use			C10 lipid modifying agents, excluding C10AA HMG CoA reductase inhibitors C10BA HMG CoA reductase inhibitors in combination with other

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
				lipid modifying agents C10BX HMG CoA reductase inhibitors, other combinations
Antiplatelet	≥1 filled dispensation/any other record of the drug use Including Aspirin alone, clopidogrel, prasugrel, ticlopidine, aspirin-dipyridamole, dipyridamole alone, cilostazol, ticagrelor			B01AC06 acetylsalicylic acid B01AC04 clopidogrel B01AC22 prasugrel B01AC05 ticlopidine B01AC07 dipyridamole B01AC23 cilostazol B01AC24 ticagrelor
Anticoagulants	≥1 filled dispensation/any other record of the drug use Including warfarin, dabigatran, rivaroxaban, apixaban			B01AA03 warfarin B01AE07 dabigatran etexilate B01AF01 rivaroxaban B01AF02 apixaban
Heparin and other low-molecular-weight heparins	≥1 filled dispensation/any other record of the drug use Including heparin (including tinzaparin) dalteparin, enoxaparin			B01AB Heparin group B01AE07 dabigatran etexilate
Nonsteroidal anti-inflammatory drugs	≥1 filled dispensation/any other record of the drug use			M01A anti-inflammatory and antirheumatic products, non-steroids
Oral corticosteroids	≥1 filled dispensation/any other record of the drug use of oral formulations Including cortisone, hydrocortisone, prednisone, prednisolone,			H02AB10 cortisone H02AB09 hydrocortisone H02AB07 prednisone2

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
	methylprednisolone, triamcinolone, dexamethasone, betamethasone			H02AB06 prednisolone H02AB04 methylprednisolone H02AB08 triamcinolone H02AB02 dexamethasone H02AB01 betamethasone
Bisphosphonates	≥1 filled dispensation/any other record of the drug use			M05BA Bisphosphonates M05BB Bisphosphonates, combinations
Opioids	≥1 filled dispensation/any other record of the drug use			N02A
Antidepressants	≥1 filled dispensation/any other record of the drug use			N06A
Antipsychotics	≥1 filled dispensation/any other record of the drug use			N05A
Anticonvulsants	≥1 filled dispensation/any other record of the drug use			N03A
Lithium	≥1 filled dispensation/any other record of the drug use			N05AN01
Benzodiazepines	≥1 filled dispensation/any other record of the drug use			N05BA Benzodiazepine derivatives N05CD Benzodiazepine derivatives
Other anxiolytics/hypnotics	≥1 filled dispensation/any other record of the drug use			N05CF02 zolpidem N05CF01 zopiclone

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
	Including eszopiclone, zaleplon, zolpidem, chloral hydrate, diphenhydramine, doxylamine, ethchlorvynol, glutethimide, methaqualone, buspirone			R06AA02 diphenhydramine1,2 R06AA09 doxylamine3 N05BE01 buspirone
Agents for dementia	≥1 filled dispensation/any other record of the drug use			N06D anti-dementia drugs
Anti-Parkinson agents	≥1 filled dispensation/any other record of the drug use			N04 anti-Parkinson drugs
HEALTHCARE RESOURCE UTILIZATION COVARIATE (12 months preceding index date)				
Combined comorbidity score	**The sum of weights related to pre-specified comorbidities** (as in Gagne et al. J Clin Epidemiol. 2011 Jul;64(7):749-59) **Conditions with weight of 5:** Metastatic cancer **Conditions with weight of 2:** Congestive heart failure; dementia; renal failure; weight loss **Conditions with weight of 1:** Hemiplegia; alcohol abuse; any tumor; cardiac arrhythmias; chronic pulmonary disease;	a) 196.x–199.x b) 402.01, 402.11, 402.91, 425.x, 428.x, 429.3 c) 290.x, 331.0, 331.1, 331.2 d) 403.11, 403.91, 404.12, 404.92, 585.x, 586.x,	a) C45.9, C77.x-C80.x b) A18.84, I09.9, I11.0, I13.0, I13.2, I25.5, I42.x, I43.x, I50.x, I51.7, P29.0 c) F01.x–F03.x, F05, G30.x, G31.01, G31.09, G31.1 d) I12.0, I13.x, N03.2–N03.7, N05.2– N05.7, N18.x, N19.x, N25.0, Z39.32, Z48.22, Z49.0x, Z49.31, Z91.15, Z94.0, Z99.2 e) E40.x-E46.x, E64.0, R63.4, R64 f) G04.1, G11.4, G80.1, G80.2, G81.x, G82.x, G83.x g) F10.x, E52, G62.1, I42.6, K29.2x, K70.0, K70.3x,	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
	coagulopathy; complicated diabetes; deficiency anemias; fluid and electrolyte disorders; liver disease; peripheral vascular disorder; psychosis; pulmonary circulation disorders; Conditions with weight of -1 HIV/AIDS; hypertension	V42.0, V45.1, V56.0, V56.8 e) 260.x–263.x f) 342.x, 344.x g) 291.1, 291.2, 291.5, 291.8, 291.9, 303.9-303.93, 305.0-305.03, V11.3 h) 140.x–171.x, 174.x–195.x, 200.x–208.x, 273.0, 273.3, V104.6 i) 426.10,	K70.9, T51.x, Z71.4x, Z65.8 h) C00.x-C26.x, C30.x–C34.x, C37.x-C41.x, C43.x, C45.x–C58.x, C60.x– C75.x, C76.x, C81.x-94.3x, C94.8x, C95.x, C96.0-C96.4, C96.9, C96.A, C96.Z, D45, D89, Z85.46 i) I44.0, I44.1, I44.3x-145.2, I45.4-I45.8x, I45.9, I47.x–I49.x, R00.0, R00.1, R00.8, T82.1x, Z45.0x, Z95.0, Z95.810, Z95.818, Z95.9 j) I26.0x, I27.2-I27.9, J40.x-J47.x, J60.x–J67.x, J68.4, J70.1, J70.3 k) D65.x-D68.x, D69.1, D69.3–D69.6 l) E10.2x–E10.8, E11.2x–E11.8, E12.2x– E12.8, E13.2x–E13.8x m) D50.1-D50.9, D51.x–D53.x, D64.9 n) E22.2, E86.x, E87.x o) B18.x, I85.x, I86.4, K70.x, K71.1, K71.3–K71.5, K71.7, K72.1x, K72.9.x, K73.x-	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
		426.11, 426.13, 426.2-426.4, 426.50-426.53, 426.6-426.8, 427.0, 427.2, 427.31, 427.6, 427.9, 785.0, V45.0, V53.3 j) 415.0, 416.8, 416.9, 491.x–494.x, 496.x k) 286.0-286.9, 287.1, 287.3–287.5 l) 250.4–250.73,	K74.x, K75.4, K75.81, K76.0, , K76.2–K76.9, Z48.23, Z94.4 p) E08.51, E08.52, E09.51, E09.52, E10.51, E10.52, E11.51, E13.51, E13.52, I67.0, I70.x, I71.x, I73.1, I73.8x, I73.9, I77.1, I77.71-I77.74, I177.79, I79.x, K55.1, K55.8, K55.9, Z95.82x, Z95.9 q) F20.x, F22-25.x, F28.x, F29.x, F30.x-F33.x, F34.8, F34.9, F39.x, F44.89, F84.3 r) I26.x, I27.x, I28.0, I28.8, I28.9 s) B20.x t) I10.x, I11.x–I13.x, I15.x, N26.2	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
		250.90-250.93 m) 280.1–281.9, 285.9 n) 276.x o) 070.32, 070.33, 070.54, 456.0, 456.1, 456.20, 456.21, 571.0, 571.2, 571.3, 571.40-571.49, 571.5, 571.6, 571.8, 571.9, 572.3, 572.8, V42.7 p) 440.x, 441.2, 441.4,		

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
		441.7, 441.9, 443.1– 443.9, 447.1, 557.1, 557.9, V43.4 q) 295.x- 298.99, 299.1, 299.11 r) 416.x, 417.9 s) 042.x– 044.x t) 401.1, 401.9, 402.10, 402.90, 404.10, 404.90, 405.11, 405.19, 405.91, 405.99		

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Total N distinct diagnosis codes	3rd digit level ICD diagnoses			
Number of different medications	As available in the data source, for example number of unique 7-digit ATC codes			
Any hospitalization	As available in the data source			
Any hospitalization within prior 30 days	As available in the data source			
Any hospitalization during prior 31-365 days	As available in the data source			
Number of hospitalizations	As available in the data source			
Number of hospital days	As available in the data source			
Number of emergency department visits	As available in the data source			
Number of office visits	As available in the data source			
Endocrinologist visit	As available in the data source			
Endocrinologist visit in 30 days prior	As available in the data source			
Endocrinologist visit in 31 to 365 days prior	As available in the data source			
Number of	As available in the data source			

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
endocrinologist visits				
Internal medicine (IM) / family medicine (FM) visits	As available in the data source			
Number of IM/FM visits	As available in the data source			
IM/FM visit (30 days prior)	As available in the data source			
IM/FM visit (31 to 365 days prior)	As available in the data source			
Cardiologist visit	As available in the data source			
Number of cardiologist visits	As available in the data source			
Cardiologist visit (30 days prior)				
Cardiologist visit (31 to 365 days prior)				
Electrocardiogram	-	89.51, 89.52	4A.020FZ, 4A.02XFZ, 4A.0204Z, 4A.02X4A, 4A.02X4Z	
Number electrocardiograms received	-	89.51, 89.52		
Use of glucose test strips	As available in the data source			

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Number of HbA1c tests ordered	As available in the data source			
Number of glucose tests ordered	As available in the data source			
Number of lipid tests ordered	As available in the data source			
Number of creatinine tests ordered	As available in the data source			
Number of BUN tests ordered	As available in the data source			
Number of tests for microalbuminuria	As available in the data source			
Cost covariates (12 months preceding index date)				
The total cost of care	Total charges for inpatient, outpatient, and pharmacy services			
Inpatient cost	Total charges for inpatient services			
Total CV-related inpatient cost	Total charges for inpatient services for hospitalization with a primary discharge diagnosis of a cardiovascular condition	Cardiovascular condition (ICD-9 390-459)	Cardiovascular condition (ICD-10 I00 to I99)	

Covariate	Definition (when applicable)	ICD-9	ICD-10	ATC
Total non-CV-related inpatient cost	Total charges for inpatient services for hospitalization with a primary discharge diagnosis of a non-cardiovascular condition	Non-cardiovascular condition (any ICD-9 expect for 390-459)	Non-cardiovascular condition (any ICD-10 expect for I00 to I99)	
Total outpatient cost	Total charges for outpatient services			
Total CV-related outpatient cost	Total charges for outpatient services for physician visits associated with a cardiovascular diagnosis	Cardiovascular diagnosis (ICD-9 390-459)	Cardiovascular condition (ICD-10 I00 to I99)	
Total non-CV-related outpatient cost	Total charges for outpatient services for physician visits associated with a non-CV diagnosis	Non-CV diagnosis (any ICD-9 except for 390-459)	Non-cardiovascular condition (any ICD-10 expect for I00 to I99)	
Total pharmacy cost	Total charges for pharmacy services			
Total pharmacy cost for antidiabetic medications	Total charges associated with the use of antidiabetic medications			
Total pharmacy cost for non-antidiabetic medications	Total charges associated with the use of medications other than antidiabetics			

AIDS=acquired immunodeficiency syndrome; ATC=Anatomical Therapeutic Chemical; BNP=B-type natriuretic peptide; BUN=blood urea nitrogen; CABG=coronary artery bypass grafting; CHF=congestive heart failure; COPD=chronic obstructive pulmonary disease; CV=cardiovascular; GLP=glucagon-like peptide; HbA1c=glycated hemoglobin A1c; HIV=human immunodeficiency virus; ICD-9=International Classification of Diseases, 9th revision; ICD-10=International Classification of Diseases, 10th revision; IM=internal medicine; MI=myocardial infarction; FM=family medicine; NT=N-terminal; PTCA=percutaneous transluminal coronary angioplasty.

ANNEX 8. DIAGNOSIS CODES FOR OTHER VARIABLES REQUIRED FOR ANALYSES

Diagnosis	ICD-9	ICD-10
Malignant neoplasm	140.xx-208.xx (except 173.xx, non-melanoma skin cancer)	C00.x-C96.x, D03.0x, Z85 (except C44, C4A, C7A, C7B)

ICD-9=International Classification of Diseases, 9th revision; ICD-10=International Classification of Diseases, 10th revision.

APPROVAL / SIGNATURE PAGE

Document Number: c26457869 **Technical Version Number:2.0**

Document Name: 1245-0195-protocol-v1-0-final 2018-10-31-clean-2018-11-02

Title: Multi-country non-interventional study on the effectiveness and safety of Empagliflozin in adult patients with type 2 diabetes in Europe and Asia

Signatures (obtained electronically)

Meaning of Signature	Signed by	Date Signed
Approval-Team Member Medicine	[redacted]	15 Oct 2019 09:46 CEST
Approval [redacted] Safety Evaluation Therapeutic Area		15 Oct 2019 11:02 CEST
Author-On behalf of Coauthor		15 Oct 2019 16:06 CEST
Approval-On behalf of [redacted] or [redacted] or [redacted]		15 Oct 2019 19:06 CEST

(Continued) Signatures (obtained electronically)

Meaning of Signature	Signed by	Date Signed