Diabetologia
https://doi.org/10.1007/s00125-024-06126-3

ARTICLE

Effectiveness and safety of empagliflozin: final results from the EMPRISE study

Phyo T. Htoo[1] · Helen Tesfaye[1] · Sebastian Schneeweiss[1] · Deborah J. Wexler[2] · Brendan M. Everett[3] · Robert J. Glynn[1] · Niklas Schmedt[4] · Lisette Koeneman[5] · Anouk Déruaz-Luyet[4] · Julie M. Paik[1,6] · Elisabetta Patorno[1]

Received: 18 October 2023 / Accepted: 5 February 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Aims/hypothesis Limited evidence exists on the comparative safety and effectiveness of empagliflozin against alternative glucose-lowering medications in individuals with type 2 diabetes with the broad spectrum of cardiovascular risk. The EMPagliflozin compaRative effectIveness and SafEty (EMPRISE) cohort study was designed to monitor the safety and effectiveness of empagliflozin periodically for a period of 5 years with data collection from electronic healthcare databases.

Methods We identified individuals ≥18 years old with type 2 diabetes who initiated empagliflozin or dipeptidyl peptidase-4 inhibitors (DPP-4i) from 2014 to 2019 using US Medicare and commercial claims databases. After 1:1 propensity score matching using 143 baseline characteristics, we identified four a priori-defined effectiveness outcomes: (1) myocardial infarction (MI) or stroke; (2) hospitalisation for heart failure (HHF); (3) major adverse cardiovascular events (MACE); and (4) cardiovascular mortality or HHF. Safety outcomes included lower-limb amputations, non-vertebral fractures, diabetic ketoacidosis (DKA), acute kidney injury (AKI), severe hypoglycaemia, retinopathy progression, and short-term kidney and bladder cancers. We estimated HRs and rate differences (RDs) per 1000 person-years, overall and stratified by age, sex, baseline atherosclerotic cardiovascular disease (ASCVD) and heart failure.

Results We identified 115,116 matched pairs. Compared with DPP-4i, empagliflozin was associated with lower risks of MI/stroke (HR 0.88 [95% CI 0.81, 0.96]; RD −2.08 [95% CI (−3.26, −0.90]), HHF (HR 0.50 [0.44, 0.56]; RD −5.35 [−6.22, −4.49]), MACE (HR 0.73 [0.62, 0.86]; RD −6.37 [−8.98, −3.77]) and cardiovascular mortality/HHF (HR 0.57 [0.47, 0.69]; RD −10.36 [−12.63, −8.12]). Absolute benefits were larger in older individuals and in those with ASCVD/heart failure. Empagliflozin was associated with an increased risk of DKA (HR 1.78 [1.44, 2.19]; RD 1.59 [1.08, 2.09]); decreased risks of AKI (HR 0.62 [0.54, 0.72]; RD −2.39 [−3.08, −1.71]), hypoglycaemia (HR 0.75 [0.67, 0.84]; RD −2.46 [−3.32, −1.60]) and retinopathy progression (HR 0.78 [0.63, 0.96)]; RD −9.49 [−16.97, −2.10]); and similar risks of other safety events.

Conclusions/interpretation Empagliflozin relative to DPP-4i was associated with risk reductions of MI or stroke, HHF, MACE and the composite of cardiovascular mortality or HHF. Absolute risk reductions were larger in older individuals and in those who had history of ASCVD or heart failure. Regarding the safety outcomes, empagliflozin was associated with an increased risk of DKA and lower risks of AKI, hypoglycaemia and progression to proliferative retinopathy, with no difference in the short-term risks of lower-extremity amputation, non-vertebral fractures, kidney and renal pelvis cancer, and bladder cancer.

Keywords Cardiovascular diseases · DPP-4i · Empagliflozin · Heart failure · SGLT2i · Type 2 diabetes mellitus

Abbreviations

AKI	Acute kidney injury
ASCVD	Atherosclerotic cardiovascular disease
CIF	Cumulative incidence function
CKD	Chronic kidney disease
DKA	Diabetic ketoacidosis
DPP-4i	Dipeptidyl peptidase-4 inhibitors
EMPRISE	EMPagliflozin compaRative effectIveness and SafEty
ESKD	End-stage kidney disease
GLP-1RA	Glucagon-like peptide-1 receptor agonists
HF	Heart failure

Extended author information available on the last page of the article

Published online: 21 March 2024

Research in context

What is already known about this subject?

- Empagliflozin has demonstrated cardiorenal benefits in individuals with a history of atherosclerotic cardiovascular disease (ASCVD) or heart failure (HF)
- Questions remain on its safety and effectiveness in broader populations under-represented in trials since prior studies included a few empagliflozin users, focused on narrow populations or presented limited safety outcomes

What is the key question?

- What is the comparative safety and effectiveness of empagliflozin relative to dipeptidyl peptidase-4 inhibitors (DPP-4i) in individuals with type 2 diabetes, overall and across broad patient subgroups in clinical practice?

What are the new findings?

- Compared with DPP-4i, empagliflozin was associated with a small risk reduction of myocardial infarction or stroke and large risk reductions of hospitalisation for HF, major adverse cardiovascular events, progression to end-stage kidney disease, and cardiovascular-specific and all-cause mortality, with larger rate differences in older individuals and in those with a history of ASCVD or HF, and no differences between male and female individuals
- For safety outcomes, we observed an increased risk of diabetic ketoacidosis and decreased risks of acute kidney injury, severe hypoglycaemia and retinopathy progression, with similar short-term risks of lower-extremity amputation, non-vertebral fractures, kidney and renal pelvis cancer, and bladder cancer, in patients initiating empagliflozin vs DPP-4i

How might this impact on clinical practice in the foreseeable future?

- This study adds to the accumulating evidence on the safety profile of empagliflozin in clinical practice. Its cardiovascular effectiveness may differ by age and ASCVD/HF history

HHF	Hospitalisation for heart failure
MACE	Major adverse cardiovascular events
MI	Myocardial infarction
NNH	Number needed to harm
NNT	Number needed to treat
PPV	Positive predictive value
PS	Propensity score
PY	Person-years
RD	Rate difference
SGLT2i	Sodium–glucose cotransporter 2 inhibitor

Introduction

Atherosclerotic cardiovascular disease (ASCVD) and heart failure (HF) are the leading causes of morbidity and mortality in individuals with type 2 diabetes [1]. Placebo-controlled trials have demonstrated the cardiovascular benefits of empagliflozin, a sodium–glucose cotransporter 2 inhibitor (SGLT2i), in individuals with established ASCVD or HF. However, questions remain on how its benefits and safety compare against alternative glucose-lowering medications in individuals with type 2 diabetes with more diverse patient characteristics than those individuals enrolled in RCTs [2–4].

To date, no RCTs have directly compared the effectiveness and safety of empagliflozin with alternative glucose-lowering medications [5, 6]. Previous studies reporting the benefits of SGLT2i in routine care included only a small number of empagliflozin users, reported limited data on safety events or were restricted to individuals with specific conditions [7–14]. Comparative evidence on the safety and effectiveness of empagliflozin vs alternative medications could help balance their benefits against the potential adverse effects.

EMPagliflozin compaRative effectIveness and SafEty (EMPRISE) is a sequential cohort study designed to monitor the safety and effectiveness of empagliflozin periodically using three US electronic healthcare databases for 5 years (1 August 2014 to 30 September 2019) [9–12, 15]. In this final-year report, we provided a final, comprehensive assessment of all cardiorenal effectiveness and safety outcomes specified in the EMPRISE monitoring programme (EnCEPP registration no. EUPAS20677 and ClinicalTrials.gov registration no. NCT03363464) [15], including the outcomes that have never been reported in prior interim analyses: cardiovascular mortality, progression to end-stage kidney disease (ESKD) and several safety outcomes, with the analyses further stratified by age, sex, history of ASCVD and HF.

Methods

We emulated a comparative safety and effectiveness trial of alternative glucose-lowering medications (electronic supplementary material [ESM] Table 1) using nationwide Medicare data, Optum's de-identified Clinformatics Data Mart Database and IBM Marketscan. These electronic databases contain longitudinal patient-level data on demographics, insurance enrolment, diagnoses and procedures for reimbursed medical services, and pharmacy drug dispensing records for eligible beneficiaries. The Mass General Brigham Institutional Review Board has approved the study protocol. Data use agreements were in place.

Study population The study population included individuals with type 2 diabetes (aged over 65 years for Medicare; over 18 years for other databases) who initiated empagliflozin or dipeptidyl peptidase-4 inhibitors (DPP-4i) (sitagliptin, saxagliptin, linagliptin, alogliptin) between 1 August 2014 (date of the first approval of empagliflozin in the USA) and 30 September 2019. Since the Medicare fee-for-service database covers all older adults 65 years or above in the USA, and Clinformatics and Marketscan databases cover >70 million geographically diverse commercial insurance beneficiaries in the USA, our study population is expected to be representative of the source population of individuals with type 2 diabetes diagnosis, including all older adults over 65 years and adults over 18 years with commercial insurance plans. Participants entered the cohort on the date of the first filled prescription of either study drug after a 12 month baseline period with no prescription fills for any SGLT2i or DPP-4i. We required eligible participants to have continuous coverage in insurance plans during this 12 month baseline period: Part A (inpatient services), B (outpatient and physician services) and D (prescription medications) plans for Medicare, and respective commercial plans for other databases. Individuals with both SGLT2i and DPP-4i prescriptions and individuals with >1 DPP-4i agents on the cohort entry date were excluded. We restricted the cohort to individuals with a recorded diagnosis of type 2 diabetes prior to or upon cohort entry and excluded individuals with recorded diagnoses of type 1 or secondary diabetes, malignancy, ESKD or kidney replacement therapy, human immunodeficiency virus, solid organ transplant or a nursing home admission at baseline (ESM Table 2, ESM Fig. 1).

Study follow-up began from 1 day after cohort entry until the earliest occurrence of: discontinuation of the index drug, switching to the comparator drug, switching from the initial drug to another agent within the same class, disenrollment, death, end of the study (30 September 2019) or a study outcome. Participants were considered exposed to the index drug until 60 days after the end of the days-supply of the last prescription.

Outcomes Primary effectiveness outcomes included: (1) a composite of myocardial infarction (MI) or stroke; (2) hospitalisation for heart failure (HHF) (defined as an HF diagnosis in the primary discharge position); (3) major adverse cardiovascular events (MACE), defined as a composite of MI, stroke or cardiovascular mortality; and (4) a composite outcome of cardiovascular death or HHF. Secondary outcomes included HHF more broadly defined as an HF diagnosis in all discharge positions, the individual components of the MACE outcome, all-cause mortality, unstable angina hospitalisation, coronary artery revascularisation and progression to ESKD (among individuals with chronic kidney disease [CKD] stages 3–4). Safety outcomes were a lower-limb amputation procedure, acute kidney injury (AKI) hospitalisation, diabetic ketoacidosis (DKA) hospitalisation, a non-vertebral fracture, severe hypoglycaemia requiring inpatient or emergency care, progression to proliferative retinopathy (including its complications or treatments), kidney and renal pelvis cancer and bladder cancer [16–21]. To assess progression to the proliferative retinopathy outcome, we limited analyses to individuals with history of non-proliferative retinopathy at baseline to allow sufficient time to develop the outcomes.

We report detailed outcome definitions in ESM Table 3. Primary outcomes were defined according to validated claims-based definitions, with high specificity (93–98%) and positive predictive value (PPV) (>98%) [22–24]. Date of death was ascertained from the Vital Status and the Social Security Administration (SSA) data, which have been validated against death certificate data and capture >95% of deaths in older adults aged >65 years in the USA [25, 26]. Cause of death was ascertained from the National Death Index considering only diagnoses in the primary position [27], and was reported only by Medicare data.

Potential confounders and baseline subgroups We identified 143 covariates a priori based on literature review and clinical knowledge: participant demographics (age, sex, race, census region), calendar time of cohort entry, modified Charlson/Elixhauser combined comorbidity score [28], indicators of frailty and validated claims-based frailty index [29], diabetes complications, glucose-lowering medication use on cohort entry and during baseline, CVDs, other comorbidities, medications for chronic diseases and measures of healthcare utilisation in various healthcare settings as a proxy for the intensity of care and surveillance. These potential confounders were identified using administrative enrolment data (sex and race), diagnosis or procedure codes, and National Drug Codes during the baseline period.

Laboratory results data were available in a subset (~20%) of the population (ESM Table 4).

We stratified analyses by the following subgroups: age (≥65 vs <65 years), sex (male vs female), history of ASCVD at baseline (defined as a diagnosis for any of the conditions: MI, angina, coronary atherosclerosis or other forms of chronic ischaemic heart disease, coronary procedure, ischaemic stroke, peripheral arterial disease or surgery, or lower-extremity amputation) and history of HF at baseline.

Statistical analyses Within each database, we matched initiators of empagliflozin and DPP-4i 1:1 using the propensity score (PS), which is the predicted probability of initiating empagliflozin relative to DPP-4i based on the measured covariates using multivariable logistic regression [30]. Laboratory results, available in a subset of the data, were not used in PS matching but were used to assess post-matching covariate balance. We performed PS matching separately within each database using the nearest neighbour matching algorithm without replacement [31], with the maximum allowed difference (calliper) of 0.01 in PS between treatments [31].

To allow more extensive control of baseline cardiovascular risk factors and evolving treatment indications over time, within each database, we estimated and matched PS separately in individuals with and without baseline CVD conditions (ASCVD and/or HF) and within each calendar time block (before and after October 2017, when the change in treatment guidelines occurred) [32]. Post-matching covariate balance was assessed using absolute standardised mean differences (ASD) [33] and the post-matching C statistic of the model predicting the exposure based on baseline covariates (0.5 indicating satisfactory balance) [34].

To conduct pooled analyses, we concatenated the three databases after PS matching and estimated HRs and rate differences (RDs) in the final stacked database using the stratified likelihood. We did not use the random effects meta-analysis since it could produce biased results due to the small number of databases we were pooling [35]. We estimated HRs using the Fine and Gray sub-distribution hazards regression and RDs using the Mantel–Haenszel method [36]. The absolute and relative heterogeneity of treatment effect across subgroups was detected by the Wald test for homogeneity [36]. We presented the Aalen Johansen cumulative risks of outcome over the follow-up using cumulative incidence function (CIF) plots [36]. The 1 year numbers needed to treat (NNT) or harm (NNH) were estimated from these plots.

We presented time-updated plots of HRs for the primary outcomes over the duration of the EMPRISE study.

Sensitivity analyses To reduce unmeasured confounding and informative censoring, we conducted the following sensitivity analyses. First, we rematched empagliflozin and DPP-4i initiators using laboratory results data: HbA_{1c} and eGFR in addition to claims-based variables among individuals for whom laboratory results data were available (~20% of the cohort). Second, to reduce unmeasured confounding by kidney function, we restricted the study cohort to those having a baseline metformin prescription, which is the recommended first line therapy for those without severely compromised kidney function [37]. Third, we excluded individuals with baseline insulin prescriptions to reduce unmeasured confounding due to uncontrolled blood glucose level and diabetes severity. Fourth, we performed 1:1 high-dimensional PS matching, which enriched the original PS with 200 additional empirically identified covariates, based on thousands of candidate covariates in different care settings [38]. The algorithm automatically selects covariates based on their confounding potential and has been shown to improve adjustment for unmeasured confounding [38]. Fifth, to quantify the impact of unmeasured confounding on the estimates, we conducted bias analyses evaluating the estimates adjusted for HbA_{1c} or eGFR, assuming strong residual associations between change in HbA_{1c} or eGFR and the risk of cardiovascular outcomes [39]. Sixth, to address potential exposure misclassification, we varied the exposure assessment window from 60 to 30 days before censoring for treatment discontinuation or switching. Seventh, to account for potential informative censoring, we conducted intent-to-treat (ITT) analyses without censoring for treatment discontinuation or switching, following individuals until 2 years after cohort entry. Eighth, to adjust for informative censoring, we performed censoring-weighted analyses which created pseudo-populations in which censoring due to treatment discontinuation/switching was independent of baseline covariates, allowing up to 1 year of follow-up. Ninth, we repeated the analyses using sitagliptin (a frequently used DPP-4i) as the comparator since it has been consistently demonstrated to be cardiovascular neutral across trials [40]. Tenth, to allow longer follow-up, we repeated the analyses among individuals who had at least 1 and 2 years of follow-up. In these analyses, follow-up started from 1 and 2 years post index until the end of available follow-up.

All analyses were performed using the Aetion Evidence Platform (2023), Aetion Substantiate software for real-world data analysis validated for a range of studies (Aetion, USA, https://www.aetion.com) [41], with R version 4.2 (R Foundation for Statistical Analysis, Vienna, Austria) and SAS 9.4 Statistical Software (SAS Institute, Cary, NC, USA).

Results

Cohort characteristics The overall study population included 136,937 empagliflozin and 599,537 DPP-4i initiators who fulfilled the eligibility criteria. After 1:1 PS matching, the population included 115,116 individuals in each treatment group (ESM Fig. 2).

Before PS matching, participants initiating empagliflozin were younger (62 vs 67 years) and were less likely to be female (44% vs 52%) compared with those initiating DPP-4i. Although the proportion of individuals on metformin at baseline was similar between empagliflozin and DPP-4i initiators (66% vs 63%), empagliflozin initiators were more likely to have used insulin (16% vs 10%) and glucagon-like peptide-1 receptor agonists (GLP-1RA) (16% vs 2%) on cohort entry. The prevalence of CKD was lower in participants initiating empagliflozin (9% vs 18%), while the prevalence of baseline CVD was approximately similar between empagliflozin and DPP-4i (35% vs 38% ASCVD/HF). After PS matching, all these differences were balanced. Approximately 33% of the matched individuals had history of ASCVD or HF at baseline. Older adults over the age of 65 years constitute approximately 52% of the matched population. Laboratory results were also balanced (Table 1, ESM Table 4).

After PS matching, the median follow-up time was approximately 5 months (interquartile range: 3–10 months) in both empagliflozin and DPP-4i initiators. The most common reason for censoring was treatment discontinuation or end of study across different outcomes (ESM Table 5). Approximately 20% of the cohort (24,772 empagliflozin and 23,331 DPP-4i initiators) had follow-up time greater than 1 year.

Effectiveness and safety outcomes After matching, we identified 13.2 and 15.3 events per 1000 person-years (PY) for the composite of MI or stroke among empagliflozin and DPP-4i initiators, with a corresponding HR (95% CI) of 0.88 (0.81, 0.96) and RD per 1000 PY (95% CI) of −2.08 (−3.26, −0.90). For the HHF outcome, we identified 5.0 and 10.3 events per 1000 PY with an HR of 0.50 (0.44, 0.56) and an RD of −5.35 (−6.22, −4.49). The rates of MACE outcome in Medicare were lower in empagliflozin vs DPP-4i initiators (22.4 vs 28.7 events/1000 PY; HR 0.73 [0.62, 0.86]; RD −6.37 [−8.98, −3.77]). The risk of a composite of cardiovascular death or HHF in Medicare was also lower in empagliflozin vs DPP-4i initiators (14.1 vs 24.4 events/1000 PY; HR 0.57 [0.47, 0.69]; RD −10.36 [−12.63, −8.12]). Secondary outcomes showed similar patterns (Table 2). Empagliflozin lowered rates of cardiovascular mortality (HR 0.61 [0.45, 0.83]; RD −3.10 [−4.40, −1.82]), all-cause mortality (HR 0.62 [0.56, 0.70]; RD −3.90 [−4.78, −3.01]) and ESKD (HR 0.45 [0.35, 0.58)]; RD −20.82 [−27.39, −14.42)]). The estimated NNTs at 1 year ranged from 102 for the composite of cardiovascular death or HHF outcome to 510 for the composite of MI or stroke.

Regarding the safety outcomes, the risks of lower-limb amputation (HR 1.07 [0.89, 1.28]; RD 0.17 [−0.37,0.72]) and non-vertebral fractures (HR 1.08 [0.92, 1.26]; RD 0.25 [−0.37, 0.88]) were similar between empagliflozin and DPP-4i initiators. Empagliflozin was associated with a higher risk of DKA (HR 1.78 [1.44, 2.19]; RD 1.59 [1.08, 2.09]) and lower risks of AKI (HR 0.62 [0.54, 0.72]; RD −2.39 [−3.08, −1.71]) and severe hypoglycaemia (HR 0.75 [0.67, 0.84]; RD −2.46 [−3.32, −1.60]) than DPP-4i. There was no difference in the risk of kidney and bladder cancers between treatments. In participants with baseline non-proliferative retinopathy, the risk of progression to proliferative retinopathy was lower in empagliflozin vs DPP-4i initiators (HR 0.78 [0.63, 0.96]; RD −9.49 [−16.97, −2.10]). For DKA, the estimated NNH at 1 year was 693, while for the AKI, the 1 year NNT was 421 (Table 2). Database-specific estimates were overall consistent across databases (ESM Table 6).

Time-updated plots of HRs showed that the estimates were relatively consistent throughout 5 years after marketing, with fewer events and less precise CIs in the earlier years of the EMPRISE study. The estimates for rare outcomes like fractures and DKA fluctuated throughout the early years of the study due to the small number of events (Fig. 1).

Consistent with HR and RD estimates, CIF curves showed lower risks of the composite outcome of MI or stroke and MACE among individuals initiating empagliflozin relative to DPP-4i. The risks of HHF, the composite outcome of cardiovascular death or HHF, and ESKD were lower in individuals initiating empagliflozin relative to those initiating DPP-4i (Fig. 2, ESM Fig. 3).

Subgroup analyses Overall, across all outcomes, the absolute risk reductions were larger in participants with baseline history of ASCVD or HF than in those without these conditions, with p values for homogeneity varying from <0.0001 to 0.08. Relative and absolute risk reductions in HHF and the composite of HHF and cardiovascular death were consistently seen independently of baseline ASCVD and HF (Fig. 3).

Stratified analyses by age showed that the relative risk reduction of the composite outcome of MI or stroke was slightly larger in older than younger individuals (p value for homogeneity 0.54). Across all outcomes, absolute RDs were larger in individuals 65 years and older. Relative and absolute risk reductions were similar between male and female participants, with p values for homogeneity ranging from 0.45 to 0.96 (ESM Fig. 4). Secondary outcomes

Table 1 Participant characteristics for 1:1 PS-matched initiators of empagliflozin vs DPP-4i pooled across three databases

Characteristic	Before matching			After matching		
	Empagliflozin *N*=136,937	DPP-4i *N*=599,537	SMD	Empagliflozin *N*=115,116	DPP-4i *N*=115,116	SMD
Demographics						
Age	62.0±8.7	67.5±9.1	0.62	62.5±8.6	62.5±8.7	<−0.001
Sex, female	60,000 (43.8)	312,601 (52.1)	0.17	51,729 (44.9)	51,661 (44.9)	<0.0001
Race categories[a]						
White	65,083 (72.6)	306,365 (69.8)	−0.06	54,476 (71.7)	54,652 (72.0)	0.01
Black	8863 (9.9)	51,736 (11.8)	0.06	7763 (10.2)	7669 (10.1)	<−0.001
Asian	3245 (3.6)	23,061 (5.3)	0.08	2966 (3.9)	2983 (3.9)	<0.0001
Hispanic	8168 (9.1)	37,214 (8.5)	−0.02	7090 (9.3)	6962 (9.2)	−0.001
Other or unknown	4268 (4.8)	20,845 (4.7)	−0.00	3657 (4.8)	3686 (4.9)	<0.01
Burden of comorbidities						
Combined comorbidity score[b]	1.2±1.7	1.4±1.9	0.10	1.2±1.7	1.2±1.6	0.01
Frailty score	0.2±0.0	0.2±0.1	0.22	0.2±0.0	0.2±0.0	<0.0001
Lifestyle-related factors						
Overweight	12,998 (9.5)	49,660 (8.3)	−0.04	11,112 (9.7)	11,167 (9.7)	<0.0001
Obesity	58,131 (42.5)	174,670 (29.1)	−0.28	45,585 (39.6)	45,465 (39.5)	<−0.001
Smoking	25,956 (19.0)	108,641 (18.1)	−0.02	21,572 (18.7)	21,611 (18.8)	<0.01
Diabetes-related conditions						
Diabetic nephropathy	18,184 (13.3)	90,387 (15.1)	0.05	14,776 (12.8)	14,996 (13.0)	0.01
Diabetic retinopathy	14,752 (10.8)	60,462 (10.1)	−0.02	11,579 (10.1)	11,599 (10.1)	<0.0001
Diabetic neuropathy	30,559 (22.3)	124,768 (20.8)	−0.04	24,172 (21.0)	24,260 (21.1)	<0.01
Diabetic foot	3304 (2.4)	15,752 (2.6)	0.01	2626 (2.3)	2632 (2.3)	<0.0001
Hypoglycaemia	15,755 (11.5)	55,562 (9.3)	−0.07	12,448 (10.8)	12,426 (10.8)	<0.0001
Hyperglycaemia	71,161 (52.0)	213,042 (35.5)	−0.34	56,140 (48.8)	56,361 (49.0)	<0.01
DKA	464 (0.3)	2537 (0.4)	0.02	393 (0.3)	418 (0.4)	0.02
Hyperosmolar hyperglycaemic nonketosis	1256 (0.9)	4917 (0.8)	−0.01	1017 (0.9)	1022 (0.9)	<0.0001
Diabetes treatment						
Number of diabetes medications on cohort entry[c]	1.4±0.9	1.2±0.8	−0.26	1.2±0.9	1.2±0.8	<0.0001
Initiation of empagliflozin or comparator monotherapy	6257 (4.6)	43,797 (7.3)	0.11	6112 (5.3)	6083 (5.3)	<0.0001
Metformin (concurrent use)[d]	89,861 (65.6)	378,717 (63.2)	−0.05	75,744 (65.8)	75,787 (65.8)	<0.0001
Sulfonylureas, second generation (concurrent use)[d]	36,516 (26.7)	194,303 (32.4)	0.13	32,062 (27.9)	32,141 (27.9)	<0.0001
Thiazolidinediones (concurrent use)[d]	8381 (6.1)	30,639 (5.1)	−0.04	6703 (5.8)	6817 (5.9)	<0.01
GLP-1RA (concurrent use)[d]	22,060 (16.1)	11,208 (1.9)	−0.51	8468 (7.4)	8380 (7.3)	<−0.001
Insulins (concurrent use)[d]	21,281 (15.5)	56,808 (9.5)	−0.18	14,682 (12.8)	14,637 (12.7)	<−0.001
Other comorbidities						
Acute MI	3482 (2.5)	13,461 (2.2)	−0.02	2538 (2.2)	2520 (2.2)	<0.0001
MI sequelae/old MI	6429 (4.7)	25,593 (4.3)	−0.02	4978 (4.3)	4941 (4.3)	<0.0001
Unstable angina	4261 (3.1)	15,444 (2.6)	−0.03	3159 (2.7)	3147 (2.7)	<0.0001
Coronary atherosclerosis	33,272 (24.3)	144,072 (24.0)	−0.01	26,450 (23.0)	26,320 (22.9)	<−0.001
Coronary procedure	3969 (2.9)	11,537 (1.9)	−0.07	2749 (2.4)	2733 (2.4)	<0.0001
HF	11,913 (8.7)	67,900 (11.3)	0.09	9706 (8.4)	9727 (8.4)	<0.0001
Cardiomyopathy	4730 (3.5)	22,809 (3.8)	0.02	3803 (3.3)	3803 (3.3)	<0.0001
Atrial fibrillation	11,029 (8.1)	61,062 (10.2)	0.07	9247 (8.0)	9109 (7.9)	<−0.001
Ischaemic stroke	10,403 (7.6)	58,687 (9.8)	0.08	8784 (7.6)	8764 (7.6)	<0.0001
Peripheral arterial disease	11,111 (8.1)	67,335 (11.2)	0.11	9405 (8.2)	9336 (8.1)	<−0.001
CKD stage 3–4	9124 (6.7)	84,348 (14.1)	0.24	8225 (7.1)	8373 (7.3)	0.01
Proteinuria	7183 (5.2)	31,510 (5.3)	<0.01	5609 (4.9)	5574 (4.8)	<−0.001
COPD	11,016 (8.0)	67,181 (11.2)	0.11	9546 (8.3)	9822 (8.5)	0.01
Obstructive sleep apnoea	27,328 (20.0)	81,427 (13.6)	−0.17	21,122 (18.3)	21,142 (18.4)	<0.01

Table 1 (continued)

Characteristic	Before matching			After matching		
	Empagliflozin *N*=136,937	DPP-4i *N*=599,537	SMD	Empagliflozin *N*=115,116	DPP-4i *N*=115,116	SMD
Non-alcoholic steatohepatitis/fatty liver	9053 (6.6)	28,455 (4.7)	−0.08	7192 (6.2)	7225 (6.3)	<0.01
Dementia	2822 (2.1)	32,050 (5.3)	0.17	2584 (2.2)	2565 (2.2)	<0.0001
Other medications						
ACEI and ARBs	101,778 (74.3)	441,932 (73.7)	−0.01	84,579 (73.5)	84,728 (73.6)	<0.01
β-blockers	52,067 (38.0)	243,203 (40.6)	0.05	42,828 (37.2)	42,778 (37.2)	<0.0001
Calcium channel blockers	37,182 (27.2)	191,394 (31.9)	0.10	31,550 (27.4)	31,429 (27.3)	<−0.001
Nitrates and other antianginal agents	9928 (7.3)	44,991 (7.5)	0.01	7814 (6.8)	7891 (6.9)	<0.01
Thiazides	18,768 (13.7)	87,117 (14.5)	0.02	15,658 (13.6)	15,564 (13.5)	<−0.001
Loop diuretics	16,818 (12.3)	97,642 (16.3)	0.11	13,796 (12.0)	13,840 (12.0)	<0.0001
Mineralocorticoid receptor antagonists	6019 (4.4)	25,168 (4.2)	−0.01	4688 (4.1)	4670 (4.1)	<0.0001
Digoxin	1841 (1.3)	12,927 (2.2)	0.07	1553 (1.3)	1530 (1.3)	<0.0001
Antiarrhythmics	2335 (1.7)	12,650 (2.1)	0.03	1933 (1.7)	1907 (1.7)	<0.0001
Anticoagulants	9759 (7.1)	51,161 (8.5)	0.05	8157 (7.1)	8079 (7.0)	<−0.001
Antiplatelets	16,738 (12.2)	70,380 (11.7)	−0.02	13,165 (11.4)	13,214 (11.5)	<0.01
Statins	102,030 (74.5)	429,097 (71.6)	−0.07	84,323 (73.3)	84,310 (73.2)	<−0.001
PCSK9 inhibitors and other lipid-lowering agents	21,500 (15.7)	88,928 (14.8)	−0.03	17,206 (14.9)	17,190 (14.9)	<0.0001
Corticosteroids (oral)	24,323 (17.8)	103,588 (17.3)	−0.01	20,509 (17.8)	20,580 (17.9)	<0.01
Opioids	41,543 (30.3)	194,860 (32.5)	0.05	34,766 (30.2)	34,754 (30.2)	<0.0001
Measures of healthcare utilisation						
Visit to an internist (−30 days to cohort entry)	86,673 (63.3)	407,677 (68.0)	0.10	75,659 (65.7)	76,013 (66.0)	<0.01
Visit to a cardiologist (−30 days to cohort entry)	16,093 (11.8)	71,271 (11.9)	<0.01	12,799 (11.1)	12,709 (11.0)	<0.01
HbA_{1c} test order (number of tests)	2.6±1.4	2.4±1.4	−0.16	2.5±1.4	2.5±1.4	<0.0001
Glucose test and monitoring (number of tests)	0.9±2.3	0.8±3.6	−0.02	0.8±2.4	0.8±1.9	<−0.001
Microalbuminuria/proteinuria test order (number of tests)	1.0±1.0	0.9±1.0	−0.10	1.0±1.0	1.0±1.0	<0.0001
Hospitalisations (number)	0.1±0.5	0.2±0.6	0.13	0.1±0.4	0.1±0.4	<0.0001
Length of stay (−30 days to cohort entry)	0.1±0.9	0.4±2.1	0.16	0.1±0.9	0.1±0.8	0.01
Emergency visits (number)	0.5±1.3	0.6±1.3	0.08	0.5±1.3	0.5±1.2	0.00
Distinct brand-name medications (number)	2.8±1.8	2.6±1.7	−0.16	2.6±1.7	2.6±1.7	0.00
Laboratory results						
HbA_{1c}, mmol/mol	75.0±2.0	73.0±2.0	0.09	75.0±2.0	74.0±2.0	0.04
HbA_{1c}, %[e]	9.0±2.3	8.8±2.3	0.09	9.0±2.3	8.9±2.3	0.04
eGFR, ml/min per 1.73 m^2 [e]	85.2±22.0	78.5±24.7	0.29	85.1±22.0	83.7±23.0	0.06

Data are *n* (%) or mean±SD

Baseline characteristics were measured during 12 months prior to and including the index date (cohort entry date) unless otherwise stated. For ASD, <0.1 was suggested as a measure of satisfactory balance as in Austin, 2009 [33]

[a]Race information is available only in Medicare and Clinformatics administrative enrolment data, i.e. beneficiary summary data files, and not reported in Marketscan

[b]Calculated using the weights in Gagne et al, 2011 [28]

[c]Number of diabetes medications calculated here did not include the index medications

[d]Concurrent use on the index date was defined as the overlap of days-supply of baseline medication with the cohort entry date

[e]Available for a subset (~20%) of participants, thus not included in in the PS model

ACEI, angiotensin-converting enzyme inhibitors; ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; PCSK9, proprotein convertase subtilisin/kexin type 9 serine protease; SMD, standardised mean differences

stratified by subgroups showed similar patterns (ESM Table 7).

Sensitivity analyses Sensitivity analyses and quantitative bias analyses produced consistent results with the primary findings (ESM Tables 8–10, ESM Figs 5–7). Rematching

Table 2 Incidence rates and treatment effect estimates for PS-matched initiators of empagliflozin vs DPP-4i

Primary and secondary outcomes	Empagliflozin *N* events (IR/1000 PY)	DPP-4i *N* events (IR/1000 PY)	Empagliflozin vs DPP-4i		NNT or NNH at 1 year
			HR (95% CI)	RD/1000 PY (95% CI)	
Primary outcomes					
Composite of MI or stroke	1051 (13.2)	1188 (15.3)	0.88 (0.81, 0.96)	−2.08 (−3.26, −0.90)	510
Hospitalisation for HF	397 (5.0)	804 (10.3)	0.50 (0.44, 0.56)	−5.35 (−6.22, −4.49)	199
MACE outcome[a]	638 (22.4)	842 (28.7)	0.73 (0.62, 0.86)	−6.37 (−8.98, −3.77)	162
Composite of cardiovascular death or hospitalisation for HF[a]	403 (14.1)	718 (24.4)	0.57 (0.47, 0.69)	−10.36 (−12.63, −8.12)	102
Secondary outcomes					
Hospitalisation for HF (broad)	1871 (23.6)	2651 (34.5)	0.71 (0.67, 0.75)	−10.83 (−12.52, −9.13)	99
MI	657 (8.3)	755 (9.7)	0.86 (0.78, 0.96)	−1.45 (−2.39, −0.52)	715
Stroke	400 (5.0)	435 (5.6)	0.92 (0.80, 1.05)	−0.56 (−1.28, 0.15)	1964
Cardiovascular mortality[a]	138 (4.8)	234 (7.9)	0.61 (0.45, 0.83)	−3.10 (−4.40, −1.82)	341
All-cause mortality	485 (6.1)	779 (10.0)	0.62 (0.56, 0.70)	−3.90 (−4.78, −3.01)	256
Unstable angina	207 (2.6)	236 (3.0)	0.88 (0.73, 1.06)	−0.43 (−0.95, 0.09)	2315
Coronary revascularisation	853 (10.7)	830 (10.7)	1.02 (0.93, 1.13)	0.05 (−0.97, 1.07)	9553
ESKD[b]	81 (16.8)	192 (37.6)	0.45 (0.35, 0.58)	−20.82 (−27.39, −14.42)	53
Safety outcomes					
Lower-limb amputations	252 (3.2)	233 (3.0)	1.07 (0.89, 1.28)	0.17 (−0.37, 0.72)	5595
Non-vertebral fractures	330 (4.1)	303 (3.9)	1.08 (0.92, 1.26)	0.25 (−0.37, 0.88)	3866
DKA[c]	273 (3.4)	143 (1.8)	1.78 (1.44, 2.19)	1.59 (1.08, 2.09)	693
AKI[d]	284 (3.6)	464 (6.0)	0.62 (0.54, 0.72)	−2.39 (−3.08, −1.71)	421
Hypoglycaemia	500 (6.3)	680 (8.7)	0.75 (0.67, 0.84)	−2.46 (−3.32, −1.60)	432
Kidney and renal pelvis cancer	69 (0.9)	69 (0.9)	1.00 (0.70, 1.43)	−0.02 (−0.31, 0.27)	61,133
Bladder cancer	75 (0.9)	67 (0.9)	1.03 (0.72, 1.49)	0.08 (−0.21, 0.38)	12,525
Retinopathy progression[e]	154 (30.6)	195 (40.1)	0.78 (0.63, 0.96)	−9.49 (−16.97, −2.10)	122

Empagliflozin vs DPP-4i, *N* matched pairs=115,116 (43,244 in Medicare data)

[a]MACE outcome includes hospitalisation for MI, or ischaemic or haemorrhagic stroke, and cardiovascular-specific mortality; cardiovascular-specific mortality data were available only in the Medicare database, the only database where linkage with the National Death Index was possible

[b]Restricted to individuals with CKD stage 3 and above. No. of matched pairs 8072

[c]Defined using diagnosis codes in inpatient setting, with any diagnosis fields on hospital discharge, following a validated algorithm shown to have high specificity and PPV

[d]Defined using diagnosis codes in inpatient setting, with primary diagnosis field on hospital discharge, following a validated algorithm shown to have high specificity and PPV

[e]The outcome was defined as a composite of proliferative diabetic retinopathy, onset of vitreous haemorrhage and initiation of intravitreal anti-vascular endothelial growth factor injection or panretinal photocoagulation. Analyses were restricted to individuals with diagnoses of non-proliferative retinopathy and without the baseline history of the conditions included in the outcome definition. No. of matched individuals eligible for this analysis was 7839

IR, incidence rate

treatments using laboratory results also provided similar findings (ESM Fig. 7).

Discussion

In this comparative effectiveness and safety study of 230,232 individuals with type 2 diabetes, empagliflozin relative to DPP-4i was associated with large risk reductions of HHF, MACE and the composite outcome of cardiovascular death or HHF, and a small risk reduction of the composite of MI or stroke, with absolute NNTs at 1 year ranging from 102 to 510. Absolute risk reductions were larger in older individuals and in those who had history of ASCVD or HF. Regarding the safety outcomes, empagliflozin was associated with an increased risk of DKA (with 1 year NNH of 693) and lower risks of AKI, severe hypoglycaemia and progression to proliferative

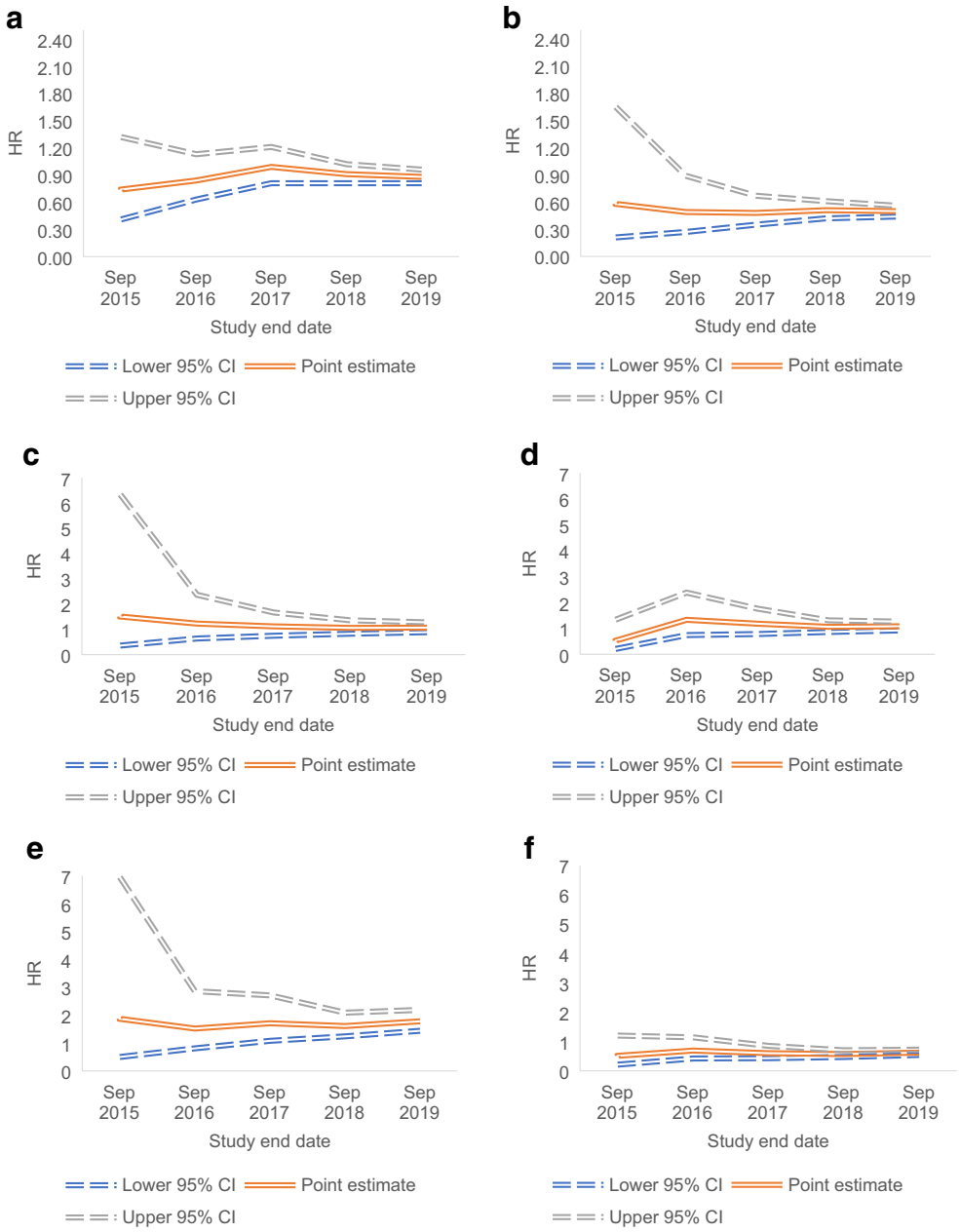
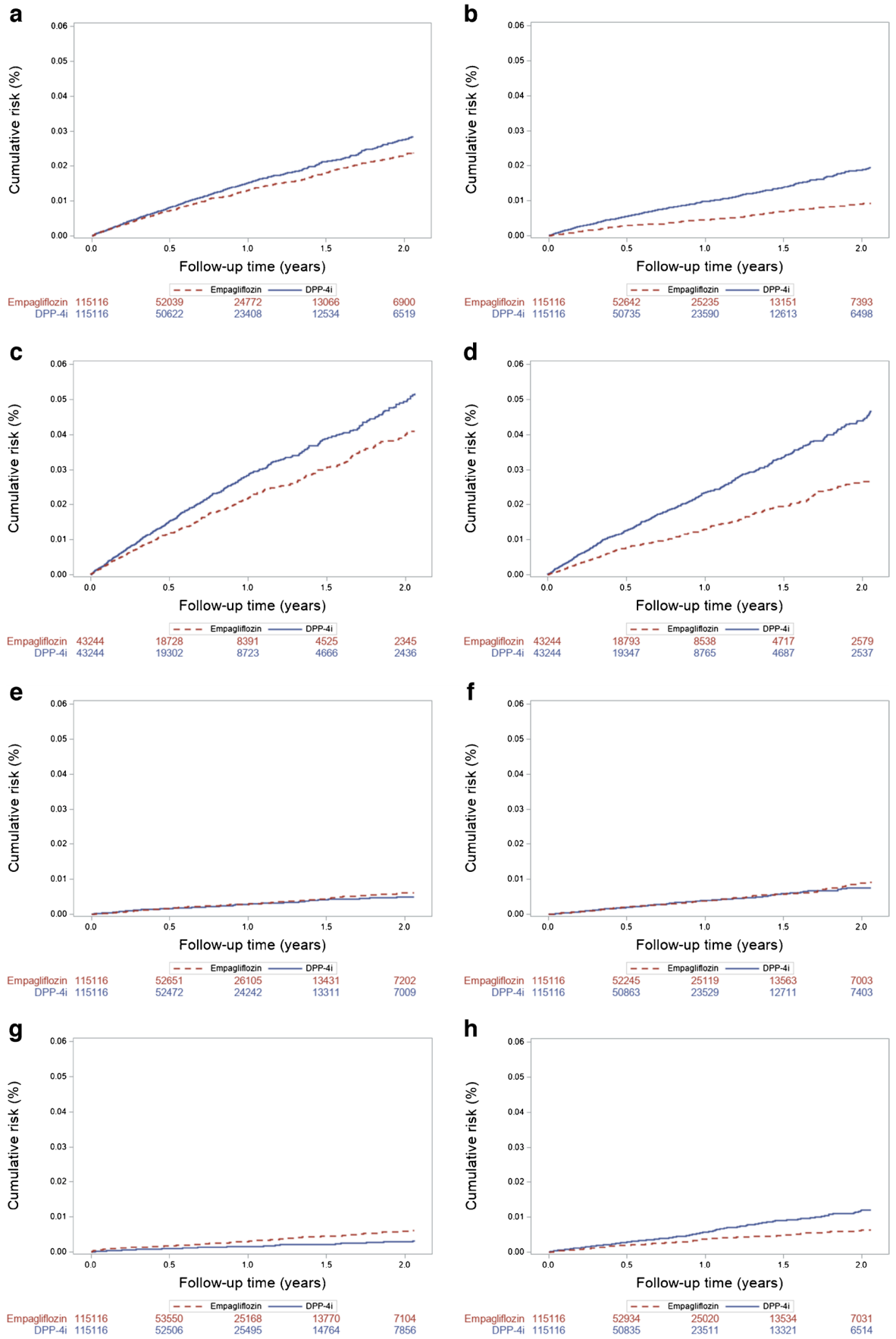


Fig. 1 Time-updated plots of HRs from the EMPRISE study (2014–2019). (**a**) Composite MI or stroke. (**b**) Hospitalisation for HF. (**c**) Lower-limb amputations. (**d**) Non-vertebral fractures. (**e**) DKA. (**f**) AKI

retinopathy (with 1 year NNTs of 421, 432 and 122, respectively) when compared with DPP-4i. There was no difference in the short-term risks of lower-extremity amputation, non-vertebral fractures, kidney and renal pelvis cancer and bladder cancer.

Our study adds to the accumulating evidence on the cardiovascular effectiveness and safety of empagliflozin, complementing the evidence from RCTs [2–4]. Our findings were consistent with previous studies [7–13], and with estimates from the interim reports of the EMPRISE study, with greater precision for both cardiovascular and safety outcomes (Fig. 1) [9–13].

The effects of SGLT2i on MI and stroke outcomes are not well established across RCTs. The EMPA-REG OUTCOME trial reported a numerical 13% risk reduction of MI, while the estimates for other SGLT2i agents varied across trials depending on the population and examined subgroups [2, 6]. A meta-analysis of placebo-controlled RCTs reported an 11% risk reduction of MI in participants randomised to SGLT2i vs those randomised to placebo (HR 0.89 [0.80,


a
Cumulative risk (%)
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0.0
0.5
1.0
1.5
2.0
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 115116 52039 24772 13066 6900
DPP-4i 115116 50622 23408 12534 6519
b
Cumulative risk (%)
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 115116 52642 25235 13151 7393
DPP-4i 115116 50735 23590 12613 6498
c
Cumulative risk (%)
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 43244 18728 8391 4525 2345
DPP-4i 43244 19302 8723 4666 2436
d
Cumulative risk (%)
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 43244 18793 8538 4717 2579
DPP-4i 43244 19347 8765 4687 2537
e
Cumulative risk (%)
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 115116 52651 26105 13431 7202
DPP-4i 115116 52472 24242 13311 7009
f
Cumulative risk (%)
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 115116 52245 25119 13563 7003
DPP-4i 115116 50863 23529 12711 7403
g
Cumulative risk (%)
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 115116 53550 25168 13770 7104
DPP-4i 115116 52506 25495 14764 7856
h
Cumulative risk (%)
Follow-up time (years)
Empagliflozin
DPP-4i
Empagliflozin 115116 52934 25020 13534 7031
DPP-4i 115116 50835 23511 13321 6514

◂**Fig. 2** Cumulative risk of primary cardiovascular and safety outcomes. (**a**) Composite MI or stroke. (**b**) Hospitalisation for HF. (**c**) MACE. (**d**) Composite of cardiovascular death or hospitalisation for HF. (**e**) Lower-limb amputations. (**f**) Non-vertebral fractures. (**g**) DKA. (**h**) AKI. Cardiovascular death data were only available in the Medicare database. MACE includes hospitalisation for MI, or ischaemic or haemorrhagic stroke, or cardiovascular-specific mortality

0.98]), which is close to the 12% risk reduction reported in this study [6]. The same meta-analysis reported a 15% risk reduction (HR 0.85 [0.76, 0.95]) in a subgroup of participants with a history of ASCVD, which is also consistent with our findings [6]. In our study, the absolute risk reduction of the composite of MI or stroke was larger in participants with history of ASCVD or HF compared with participants without these conditions, which supports the current Standards of Care recommending SGLT2i in individuals with history of ASCVD or HF [37].

The evidence on the effect of SGLT2i on cardiovascular mortality is also conflicting across different agents. Data from cardiovascular outcome trials suggest that empagliflozin may offer the largest risk reduction for cardiovascular mortality within the SGLT2i class, with a risk reduction of 38% reported in the EMPA-REG OUTCOME trial [2]. Consistent with these previous findings, we observed relative risk reductions in cardiovascular mortality of 39% in the overall population and 40% in participants with history of ASCVD (ESM Table 7).

Regarding HHF outcomes, our study adds to the current evidence base that empagliflozin offers consistent risk reductions across broad subgroups of individuals, including those without history of HF [9–12, 42]. For the composite outcome of HHF or cardiovascular death, we observed a relative risk reduction of 43% in the overall population, and 41% in participants with history of ASCVD in older Medicare individuals, which was consistent with the 44% risk reduction reported in the EMPA-REG OUTCOME trial. In terms of absolute benefit, such reduction in risk corresponded to the NNT of 102, which is the number of individuals needed to be on empagliflozin for 1 year to reduce one additional case of HHF or cardiovascular death.

Empagliflozin was also found to reduce the risk of progression to ESKD in individuals with CKD stages 3–4, in line with evidence from RCTs [6]. Most individuals who developed ESKD outcomes were older than 65 years, and had follow-up longer than 90 days.

Although empagliflozin demonstrated an overall favourable safety profile across cardiovascular outcome trials, these trials were not powered to evaluate safety events. In this study, due to a large number of individuals using empagliflozin in routine clinical practice, we were able to estimate the relative risk of most safety events with reasonable precision.

Since SGLT2i affects multiple bone and mineral metabolism hormones, including fibroblast growth factor-23 and vitamin D, there has been concern that the drug class could increase the risk of fractures [2, 43]. There have also been concerns about the potential increased risk of lower-extremity amputations with SGLT2i; however, in this study, we did not find evidence of an increased risk of fractures or lower-extremity amputations in individuals initiating empagliflozin, which is in line with the current evidence base [2, 44, 45]. We found an increased risk of hospitalisation for DKA in individuals initiating empagliflozin compared with those initiating DPP-4i, in line with an established class effect for SGLT2i as shown in both randomised and non-randomised studies [2, 46]. However, the estimated number of individuals required to receive empagliflozin for 1 year to increase one additional case of DKA was large (n=693). Regarding AKI, there was initial concern that SGLT2i could increase risk of AKI, since SGLT2i can cause acute changes in eGFR after treatment initiation [47]. However, evidence from later RCTs suggested that empagliflozin reduced AKI risk [2]. We found a risk reduction of AKI with empagliflozin, which is in line with evidence from the RCTs and other population-based observational studies of SGLT2i [2, 48]. Finally, our findings with respect to severe hypoglycaemic events, short-term cancer outcomes and retinopathy progression were consistent with prior literature [2, 49]. A post hoc analysis of the EMPA-REG OUTCOME trial reported that empagliflozin reduced the risk of diabetic retinopathy relative to placebo, though CIs were wide (HR 0.78 [0.54, 1.12]) [49]. To assess retinopathy progression, we limited analyses to participants with history of retinopathy to allow assessment of the outcome within a short follow-up time and to limit confounding by indication.

Our study has limitations. We cannot entirely exclude residual confounding. A previous study reported that a new-user, active-comparator design paired with the adjustment for a large number of claims-based confounder proxies using PS matching could balance the variables not available in the claims data [50]. Indeed, we achieved balance in the laboratory results and biomarkers, which are available in a subset of the populations, and rematching the treatment groups using laboratory results did not meaningfully change the primary analytical findings in these databases. Additionally, our bias analyses suggest that, after adjusting for 143 baseline characteristics through 1:1 PS matching [30], and further incorporating 200 additional prognostic covariates through high-dimensional PS matching [38], residual confounding due to HbA_{1c} or eGFR appears unlikely to materially change our estimates. We additionally addressed confounding due to evolving treatment indications over time. For example, during the time frame of this study, individuals with a history of CVD could have become more likely to be prescribed with empagliflozin as opposed to DPP-4i due to the change in guideline recommendations over time, and this could have led to biased findings. We therefore matched empagliflozin

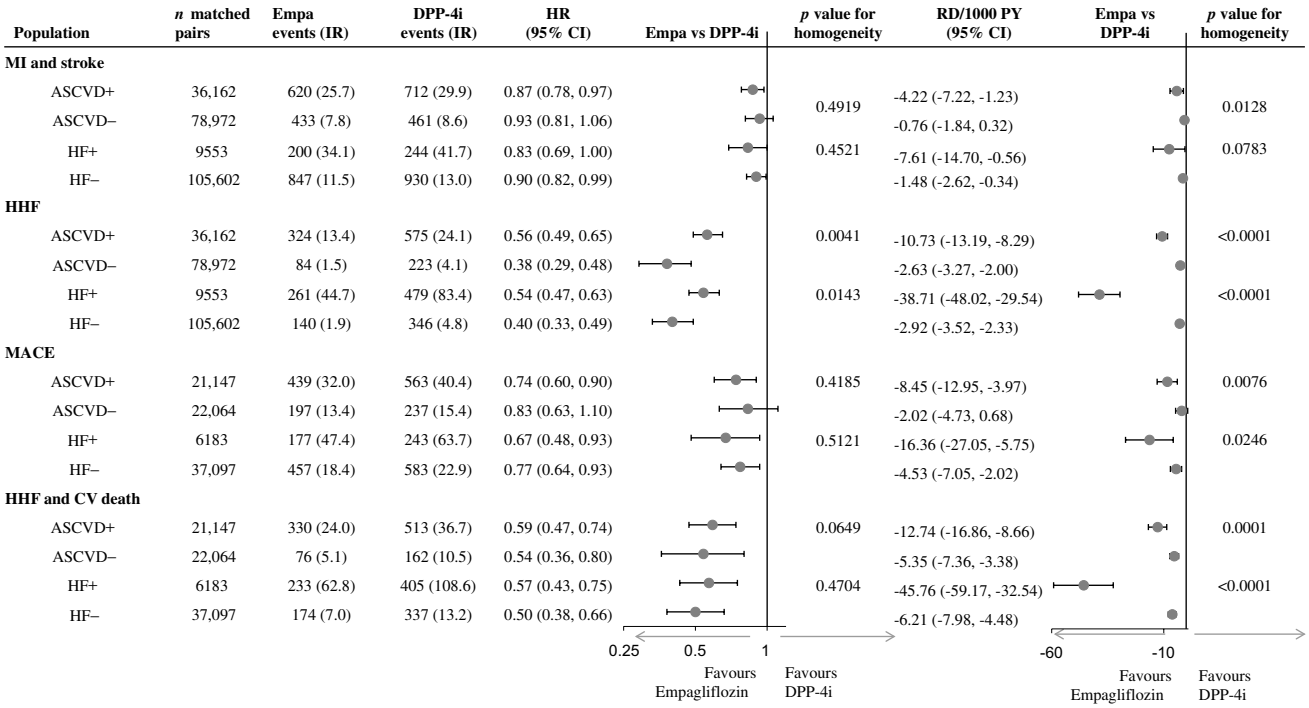


Fig. 3 Subgroup analyses by history of ASCVD or HF. History of ASCVD was defined as history of MI, angina, coronary atherosclerosis and other forms of chronic ischaemic heart disease, coronary procedure, HF, ischaemic stroke, peripheral arterial disease or surgery, or lower-extremity amputation. MACE refers to hospitalisation for MI, stroke, or cardiovascular-specific mortality. +, history of; −, no history of; CV, cardiovascular; Empa, empagliflozin; IR, incidence rates per 1000 person-years

and DPP-4i initiators separately during each calendar time block and within each baseline CVD subgroup. Time-updated plots of HRs showed the relatively stable estimates for the outcomes over time, which supports the strength of our confounding control approach. Our outcome definitions were based on claims-based algorithms validated to have high specificity but low sensitivity. Highly specific outcome definitions have been shown to have reduced bias in the relative risk estimates due to outcome misclassification [36]. Although our cardiovascular death outcome definition is highly accurate [25, 27], the cause of death data were available only in Medicare enrolees aged >65 years.

The follow-up duration in our study reflected the persistence on treatment typically observed in clinical practice and may be too short to observe risk reductions for the MI or stroke outcomes, which may require individuals to be on medications for longer periods of time. However, our analyses showed that a substantial number of individuals (24,772 empagliflozin and 23,331 DPP-4i initiators) remained on medications at 1 year of follow-up (ESM Table 10).

Conclusions In this comparative safety and effectiveness study, after extensive confounding control, participants initiating empagliflozin were associated with reductions in the risk of HHF, cardiovascular mortality and MACE outcomes (NNTs ranging from 102 to 199), when compared with participants initiating DPP-4i. Absolute risk reductions were larger in individuals with ASCVD or HF history and in older individuals, while they were similar between male and female individuals. Safety findings showed an increased risk for DKA relative to DPP-4i, corresponding to a 1 year NNH of 693, and reduced risks of AKI, severe hypoglycaemia and retinopathy progression, with similar risks of other safety events.

Supplementary Information The online version of this article (https://doi.org/10.1007/s00125-024-06126-3) contains peer-reviewed but unedited supplementary material.

Acknowledgements Some of the data were presented as an abstract at the 38th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE) in 2022.

Data availability Raw data used in the manuscript are not available to share due to the data user agreements in place but are available for purchase through the data vendors.

Funding This study was supported by a research grant to the Brigham and Women's Hospital from Boehringer Ingelheim. The authors had full control of the design and conduct of the study and the interpretation of the study's findings. The authors retained the right of publication and determined the final wording of the manuscript. PTH was supported by a training grant (5T32DK007527) from the National Institute of Diabetes and Digestive and Kidney Diseases and is currently supported by a post-doctoral grant (4-22-PDFPM-15) from the

American Diabetes Association. EP was supported by a career development grant (K08AG055670) from the National Institute on Aging and research grants from the Patient-Centered Outcomes Research Institute (DB-2020C2-20326) and the Food and Drug Administration (5U01FD007213).

Authors' relationships and activities PTH previously worked at Johnson & Johnson on unrelated work. SS reports investigator-initiated grants to the Brigham and Women's Hospital from Boehringer Ingelheim International Gmbh, and owns equity in a software manufacturer, Aetion, Inc. DJW reports serving on data monitoring committees for Novo Nordisk, consulting for Elsevier and UpToDate, and grants from PCORI. BME reports consulting for Janssen, Eli Lilly and Company, Provention Bio, Ipsen Pharmaceuticals, the NIDDK, the American Heart Association and UptoDate; and grants from PCORI and Novo Nordisk outside the submitted work. RJG reports grants from Amgen, AstraZeneca, Kowa, Novartis and Pfizer outside the submitted work. LK is an employee of Eli Lilly and Company, and owns stock in Eli Lilly and Company. NS and AD-L are employees of Boehringer Ingelheim International GmbH. The authors declare that there are no other relationships or activities that might bias, or be perceived to bias, their work.

Contribution statement All authors approved the final version to be published. PTH contributed to the study design, data analysis and initial development of the manuscript, and is the guarantor of this work. HT contributed to the data analysis and critical review of the manuscript. EP contributed to the study conception and design and critical review of the manuscript. DJW, JMP and BME contributed to the study conception and critical review of the manuscript. RJG and SS contributed to the study design, statistical input and critical review of the manuscript. NS, LK and AD-L participated in the study design, study conception and critical review of the manuscript. PTH and EP are responsible for the integrity of the work as a whole and are guarantors of the study.

References

1. Emerging Risk Factors Collaboration, Di Angelantonio E, Kaptoge S et al (2015) Association of cardiometabolic multimorbidity with mortality. JAMA 314(1):52–60. https://doi.org/10.1001/jama.2015.7008
2. Zinman B, Wanner C, Lachin JM et al (2015) Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med 373(22):2117–2128. https://doi.org/10.1056/NEJMoa1504720
3. Anker SD, Butler J, Filippatos G et al (2021) Empagliflozin in heart failure with a preserved ejection fraction. N Engl J Med 385(16):1451–1461. https://doi.org/10.1056/nejmoa2107038
4. Packer M, Anker SD, Butler J et al (2020) Cardiovascular and renal outcomes with empagliflozin in heart failure. N Engl J Med 383(15):1413–1424. https://doi.org/10.1056/nejmoa2022190
5. D'Andrea E, Kesselheim AS, Franklin JM, Jung EH, Hey SP, Patorno E (2020) Heterogeneity of antidiabetic treatment effect on the risk of major adverse cardiovascular events in type 2 diabetes: a systematic review and meta-analysis. Cardiovasc Diabetol 19(1):154. https://doi.org/10.1186/s12933-020-01133-1
6. Zelniker TA, Wiviott SD, Raz I et al (2019) SGLT2 inhibitors for primary and secondary prevention of cardiovascular and renal outcomes in type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. Lancet 393(10166):31–39. https://doi.org/10.1016/s0140-6736(18)32590-x
7. Nyström T, Toresson Grip E, Gunnarsson J et al (2023) Empagliflozin reduces cardiorenal events, healthcare resource use and mortality in Sweden compared to dipeptidyl peptidase-4 inhibitors: real world evidence from the Nordic EMPRISE study. Diabetes Obes Metab 25(1):261–271. https://doi.org/10.1111/dom.14870
8. Karasik A, Lanzinger S, Chia-Hui Tan E et al (2023) Empagliflozin cardiovascular and renal effectiveness and safety compared to dipeptidyl peptidase-4 inhibitors across 11 countries in Europe and Asia: results from the EMPagliflozin compaRative effectIveness and SafEty (EMPRISE) study. Diabetes Metab 49(2):101418. https://doi.org/10.1016/j.diabet.2022.101418
9. Htoo PT, Tesfaye H, Schneeweiss S et al (2022) Comparative effectiveness of empagliflozin vs liraglutide or sitagliptin in older adults with diverse patient characteristics. JAMA Netw Open 5(10):e2237606. https://doi.org/10.1001/jamanetworkopen.2022.37606
10. Desai RJ, Glynn RJ, Everett BM et al (2022) Comparative effectiveness of Empagliflozin in reducing the burden of recurrent cardiovascular hospitalizations among older adults with diabetes in routine clinical care. Am Heart J 254:203–215. https://doi.org/10.1016/j.ahj.2022.09.008
11. Patorno E, Pawar A, Franklin JM et al (2019) Empagliflozin and the risk of heart failure hospitalization in routine clinical care. Circulation 139(25):2822–2830. https://doi.org/10.1161/circulationaha.118.039177
12. Patorno E, Pawar A, Wexler DJ et al (2022) Effectiveness and safety of empagliflozin in routine care patients: results from the EMPagliflozin compaRative effectIveness and SafEty (EMPRISE) study. Diabetes Obes Metab 24(3):442–454. https://doi.org/10.1111/dom.14593
13. D'Andrea E, Wexler DJ, Kim SC, Paik JM, Alt E, Patorno E (2023) Comparing effectiveness and safety of SGLT2 inhibitors vs DPP-4 inhibitors in patients with type 2 diabetes and varying baseline HbA1c levels. JAMA Intern Med 183(3):242–254. https://doi.org/10.1001/jamainternmed.2022.6664
14. Fu EL, Patorno E, Everett BM et al (2023) Sodium-glucose cotransporter 2 inhibitors vs. sitagliptin in heart failure and type 2 diabetes: an observational cohort study. Eur Heart J 44(24):2216–2230. https://doi.org/10.1093/eurheartj/ehad273
15. Patorno E, Najafzadeh M, Pawar A et al (2020) The EMPagliflozin compaRative effectIveness and SafEty (EMPRISE) study programme: design and exposure accrual for an evaluation of empagliflozin in routine clinical care. Endocrinol Diabetes Metab 3(1):e00103. https://doi.org/10.1002/edm2.103
16. Karter AJ, Warton EM, Lipska KJ et al (2017) Development and validation of a tool to identify patients with type 2 diabetes at high risk of hypoglycemia-related emergency department or hospital use. JAMA Int Med 177(10):1461. https://doi.org/10.1001/jamainternmed.2017.3844
17. Bobo WV, Cooper WO, Epstein RA Jr, Arbogast PG, Mounsey J, Ray WA (2011) Positive predictive value of automated database records for diabetic ketoacidosis (DKA) in children and youth exposed to antipsychotic drugs or control medications: a Tennessee Medicaid study. BMC Med Res Methodol 11:157. https://doi.org/10.1186/1471-2288-11-157
18. Newton KM, Wagner EH, Ramsey SD et al (1999) The use of automated data to identify complications and comorbidities of diabetes: a validation study. J Clin Epidemiol 52(3):199–207. https://doi.org/10.1016/s0895-4356(98)00161-9
19. Waikar SS, Wald R, Chertow GM et al (2006) Validity of international classification of diseases, ninth revision, clinical modification codes for acute renal failure. J Am Soc Nephrol 17(6):1688–1694. https://doi.org/10.1681/asn.2006010073
20. Wright NC, Daigle SG, Melton ME, Delzell ES, Balasubramanian A, Curtis JR (2019) The design and validation of a new algorithm to identify incident fractures in administrative claims data. J Bone Miner Res 34(10):1798–1807. https://doi.org/10.1002/jbmr.3807
21. Setoguchi S, Solomon DH, Glynn RJ, Cook EF, Levin R, Schneeweiss S (2007) Agreement of diagnosis and its date for hematologic

malignancies and solid tumors between medicare claims and cancer registry data. Cancer Causes Control 18(5):561–569. https://doi.org/10.1007/s10552-007-0131-1
22. Jones SA, Gottesman RF, Shahar E, Wruck L, Rosamond WD (2014) Validity of hospital discharge diagnosis codes for stroke: the Atherosclerosis Risk in Communities Study. Stroke 45(11):3219–3225. https://doi.org/10.1161/strokeaha.114.006316
23. Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, Solomon DH (2004) Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. Am Heart J 148(1):99–104. https://doi.org/10.1016/j.ahj.2004.02.013
24. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF (2005) Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Med Care 43(5):480–485. https://doi.org/10.1097/01.mlr.0000160417.39497.a9
25. Hill ME, Rosenwaike I (2001) The Social Security Administration's Death Master File: the completeness of death reporting at older ages. Soc Secur Bull 64(1):45–51
26. Research Data Assistance Center (ResDAC) (2020) Death information in the Research Identifiable Medicare Data. Available from https://resdac.org/articles/death-information-research-identifiable-medicare-data2021. Accessed 20 Mar 2023
27. Olubowale OT, Safford MM, Brown TM et al (2017) Comparison of expert adjudicated coronary heart disease and cardiovascular disease mortality with the national death index: results from the REasons for Geographic And Racial Differences in Stroke (REGARDS) Study. J Am Heart Assoc 6(5):e004966. https://doi.org/10.1161/jaha.116.004966
28. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S (2011) A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol 64(7):749–759. https://doi.org/10.1016/j.jclinepi.2010.10.004
29. Kim DH, Patorno E, Pawar A, Lee H, Schneeweiss S, Glynn RJ (2020) Measuring frailty in administrative claims data: comparative performance of four claims-based frailty measures in the U.S. Medicare data. J Gerontol A Biol Sci Med Sci 75(6):1120–1125. https://doi.org/10.1093/gerona/glz224
30. Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70(1):41–55. https://doi.org/10.1093/biomet/70.1.41
31. Ripollone JE, Huybrechts KF, Rothman KJ, Ferguson RE, Franklin JM (2018) Implications of the propensity score matching paradox in pharmacoepidemiology. Am J Epidemiol 187(9):1951–1961. https://doi.org/10.1093/aje/kwy078
32. Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA (2011) Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. Clin Pharmacol Ther 90(6):777–790. https://doi.org/10.1038/clpt.2011.235
33. Austin PC (2009) Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med 28(25):3083–3107. https://doi.org/10.1002/sim.3697
34. Franklin JM, Rassen JA, Ackermann D, Bartels DB, Schneeweiss S (2014) Metrics for covariate balance in cohort studies of causal effects. Stat Med 33(10):1685–1699. https://doi.org/10.1002/sim.6058
35. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2011) Introduction to meta-analysis. Wiley, Hoboken
36. Lash TL, Rothman KJ, VanderWeele TJ, Haneuse S (2020) Modern epidemiology. Wolters Kluwer, Philadelphia
37. ElSayed NA, Aleppo G, Aroda VR et al (2022) 9. Pharmacologic approaches to glycemic treatment: standards of care in diabetes—2023. Diabetes Care 46(Suppl 1):S140–S157. https://doi.org/10.2337/dc23-S009
38. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA (2009) High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology 20(4):512–522. https://doi.org/10.1097/EDE.0b013e3181a663cc
39. Schneeweiss S (2006) Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf 15(5):291–303. https://doi.org/10.1002/pds.1200
40. Monami M, Ahrén B, Dicembrini I, Mannucci E (2013) Dipeptidyl peptidase-4 inhibitors and cardiovascular risk: a meta-analysis of randomized clinical trials. Diabetes Obes Metab 15(2):112–120. https://doi.org/10.1111/dom.12000
41. Wang SV, Verpillat P, Rassen JA, Patrick A, Garry EM, Bartels DB (2016) Transparency and reproducibility of observational cohort studies using large healthcare databases. Clin Pharmacol Ther 99(3):325–332. https://doi.org/10.1002/cpt.329
42. Thomsen RW, Knudsen JS, Kahlert J et al (2021) Cardiovascular events, acute hospitalizations, and mortality in patients with type 2 diabetes mellitus who initiate empagliflozin versus liraglutide: a comparative effectiveness study. J Am Heart Assoc 10(11):e019356. https://doi.org/10.1161/jaha.120.019356
43. Rau M, Thiele K, Hartmann NK et al (2022) Effects of empagliflozin on markers of calcium and phosphate homeostasis in patients with type 2 diabetes - data from a randomized, placebo-controlled study. Bone Rep 16:101175. https://doi.org/10.1016/j.bonr.2022.101175
44. Zhuo M, Hawley CE, Paik JM et al (2021) Association of sodium-glucose cotransporter–2 inhibitors with fracture risk in older adults with type 2 diabetes. JAMA Netw Open 4(10):e2130762. https://doi.org/10.1001/jamanetworkopen.2021.30762
45. Inzucchi SE, Iliev H, Pfarr E, Zinman B (2018) Empagliflozin and assessment of lower-limb amputations in the EMPA-REG OUTCOME trial. Diabetes Care 41(1):e4–e5. https://doi.org/10.2337/dc17-1551
46. Fralick M, Schneeweiss S, Redelmeier DA, Razak F, Gomes T, Patorno E (2021) Comparative effectiveness and safety of sodium-glucose cotransporter-2 inhibitors versus metformin in patients with type 2 diabetes: an observational study using data from routine care. Diabetes Obes Metab 23(10):2320–2328. https://doi.org/10.1111/dom.14474
47. Lytvyn Y, Bjornstad P, van Raalte DH, Heerspink HL, Cherney DZI (2020) The new biology of diabetic kidney disease—mechanisms and therapeutic implications. Endocr Rev 41(2):202–231. https://doi.org/10.1210/endrev/bnz010
48. Nadkarni GN, Ferrandino R, Chang A et al (2017) Acute kidney injury in patients on SGLT2 inhibitors: a propensity-matched analysis. Diabetes Care 40(11):1479–1485. https://doi.org/10.2337/dc17-1011
49. Inzucchi SE, Wanner C, Hehnke U et al (2019) Retinopathy outcomes with empagliflozin versus placebo in the EMPA-REG OUTCOME trial. Diabetes Care 42(4):e53–e55. https://doi.org/10.2337/dc18-1355
50. Patorno E, Gopalakrishnan C, Franklin JM et al (2018) Claims-based studies of oral glucose-lowering medications can achieve balance in critical clinical variables only observed in electronic health records. Diabetes Obes Metab 20(4):974–984. https://doi.org/10.1111/dom.13184

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Phyo T. Htoo[1] · Helen Tesfaye[1] · Sebastian Schneeweiss[1] · Deborah J. Wexler[2] · Brendan M. Everett[3] · Robert J. Glynn[1] · Niklas Schmedt[4] · Lisette Koeneman[5] · Anouk Déruaz-Luyet[4] · Julie M. Paik[1,6] · Elisabetta Patorno[1]

✉ Elisabetta Patorno
epatorno@bwh.harvard.edu

[1] Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[2] Massachusetts General Hospital Diabetes Center, Harvard Medical School, Boston, MA, USA

[3] Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

[4] Global Epidemiology, Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany

[5] Global Medical Affairs, Lilly Deutschland GmbH, Bad Homburg vor der Höhe, Germany

[6] Division of Renal (Kidney) Medicine, Brigham and Women's Hospital, Boston, MA, USA