



## ORIGINAL RESEARCH ARTICLE

# The EMPagliflozin compaRative effectiveness and SafEty (EMPRISE) study programme: Design and exposure accrual for an evaluation of empagliflozin in routine clinical care

Elisabetta Patorno<sup>1</sup>  | Mehdi Najafzadeh<sup>1</sup> | Ajinkya Pawar<sup>1</sup> | Jessica M. Franklin<sup>1</sup> | Anouk Déruaz-Luyet<sup>2</sup>  | Kimberly G. Brodovicz<sup>3</sup> | Adrian J. Santiago Ortiz<sup>1</sup> | Lily G. Bessette<sup>1</sup> | Martin Kulldorff<sup>1</sup> | Sebastian Schneeweiss<sup>1</sup>

<sup>1</sup>Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

<sup>2</sup>Boehringer Ingelheim GmbH, Ingelheim, Germany

<sup>3</sup>Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA

## Correspondence

Elisabetta Patorno, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont Street, Suite 3030, Boston, MA 02120, USA. Email: epatorno@bwh.harvard.edu

## Funding information

Boehringer-Ingelheim, Grant/Award Number: Research grant to Brigham and Women's Hospital; National Institute on Aging, Grant/Award Number: RO1GM108999; National Institute of General Medical Sciences, Grant/Award Number: RO1GM108999; Bayer; Boehringer Ingelheim

## Abstract

**Background:** The EMPA-REG OUTCOME trial showed that empagliflozin reduced the risk of cardiovascular death and hospitalization for heart failure (HHF) in diabetic patients with cardiovascular disease. EMPRISE is a study programme on the effectiveness, safety and healthcare utilization of empagliflozin in routine care, leveraging real-world data from two commercial and one federal US data sources from 2014 to 2019.

**Objectives:** To describe rationale and design of EMPRISE, assess ability to minimize confounding and evaluate the time to reach sufficient statistical power for a key study outcome, HHF, using baseline information from the first year of EMPRISE.

**Methods:** In 3 claims data sets, we identified a 1:1 propensity score (PS)-matched cohort of diabetic patients  $\geq 18$  years initiating empagliflozin or a dipeptidyl peptidase-4 inhibitor (DPP4i), resulting in 6643 total pairs. The PS model included  $>140$  baseline covariates. We measured covariate balance via standardized differences (SD) and postmatching c-statistic. We computed the incidence rate (IR) of HHF, predicted exposure accrual over time and calculated expected power.

**Results:** After PS matching, patient characteristics were balanced with SD  $< 0.1$  and c-statistic between 0.54 and 0.59. The population IR of HHF was 4.4 per 1000 person-years using a specific HHF definition and 14.8 using a broader HHF definition. In our projection, 80%-powered analyses would require a minimum of 169 HHF events, expected to accumulate by year 3 (specific definition) or year 2 (broader definition).

**Conclusion:** Baseline information from EMPRISE provided evidence of solid confounding control and adequate exposure accrual with expected powered analyses for the primary outcomes.

## KEYWORDS

comparative effectiveness, confounding (epidemiology), empagliflozin, heart failure, real-world data, study validity, type 2 diabetes

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors. *Endocrinology, Diabetes & Metabolism* published by John Wiley & Sons Ltd.

## 1 | BACKGROUND

The cardiovascular outcome trial EMPA-REG OUTCOME<sup>1</sup> showed that empagliflozin, a sodium-glucose cotransporter-2 (SGLT2) inhibitor, reduces the relative risk of cardiovascular death by 38% (HR 0.62; 95% CI: 0.49-0.77), all-cause mortality by 32% (HR 0.68; 95% CI: 0.57-0.82) and hospitalization for heart failure by 35% (HR 0.65; 95% CI: 0.50-0.85) when added onto standard of care in patients with type 2 diabetes (T2D) and established cardiovascular disease.

However, the beneficial effects seen in the EMPA-REG OUTCOME trial are yet to be evaluated in routine clinical care, which includes patients across a broader spectrum of cardiovascular risk. Moreover, the information on unintended harms (eg bone fractures, ketoacidosis, lower limb amputations) potentially associated with some SGLT2 inhibitors<sup>2-5</sup> has been rapidly accumulating. The impact on healthcare resource utilization and costs has also not been fully evaluated in routine clinical care. Real-world data routinely generated in the course of healthcare delivery for millions of patients can fill these evidence gaps and inform regulatory and coverage decision-making,<sup>6-8</sup> as recently recommended by the 21st Century Cures Act and the Prescription Drug User Fee Act.<sup>9,10</sup>

The EMPagliflozin compaRative effectiveness and SafEty (EMPRISE) programme of studies aims to assess the comparative effectiveness, safety and impact on healthcare utilization of empagliflozin, based on real-world data from two commercial and Medicare databases in the United States. The study will collect accumulating data on empagliflozin for a period of five years following the date of approval in the United States, 1 August 2014 through 30 September 2019. In the context of noninterventional studies of a newly available medication with prospectively accumulating real-world data, baseline information from the early stages can provide valuable insights regarding study validity and inform projections of exposure accrual over time and the resulting statistical power.<sup>11,12</sup> These elements can be crucial to determine the level of confidence in future findings that may inform prompt decision-making with regard to diabetes treatment.

We sought to describe the rationale and study design of EMPRISE and, using the first year of data, to (a) assess the ability to overcome confounding and achieve high study validity by measuring baseline comparability of treatment groups in the study population; and (b) assess when adequate statistical power will be achieved using projected drug exposure accrual.

## 2 | METHODS

### 2.1 | Data sources

This study includes data from two commercial US health insurance data sets (Optum Clinformatics and IBM MarketScan) with nationwide commercial coverage including some Medicare Advantage plans. As a third data source, we included fee-for-service Medicare, a US federal health insurance programme which provides health care

to Americans aged 65 years or older and patients with disabilities. The three data sources together cover about 200 million lives in the United States. For each insured individual, the three data sets contain demographic information, health plan enrolment status, longitudinal patient-level information on all reimbursed medical services, both inpatient and outpatient diagnoses and procedures along with pharmacy dispensing records, including information on medication start and refill, strength, quantity and days' supply. Both Optum and MarketScan are linked to laboratory test results provided by two national laboratory test provider chains. Through this linkage, results for outpatient laboratory tests are available for a subset of beneficiaries. All three data sources have been extensively used in pharmacoepidemiologic research.<sup>13</sup> Information on mortality is available in Optum and MarketScan through linkage with the Social Security Administration Death Master File and in-hospital deaths in MarketScan. The Death Master File was limited in its completeness by a policy change in 2011 concerning the extent of the Social Security Administration disclosure of death records received from states.<sup>14</sup> EMPRISE has complete information on date and cause of death in Medicare fee-for-service patients through linkage with the National Death Index.

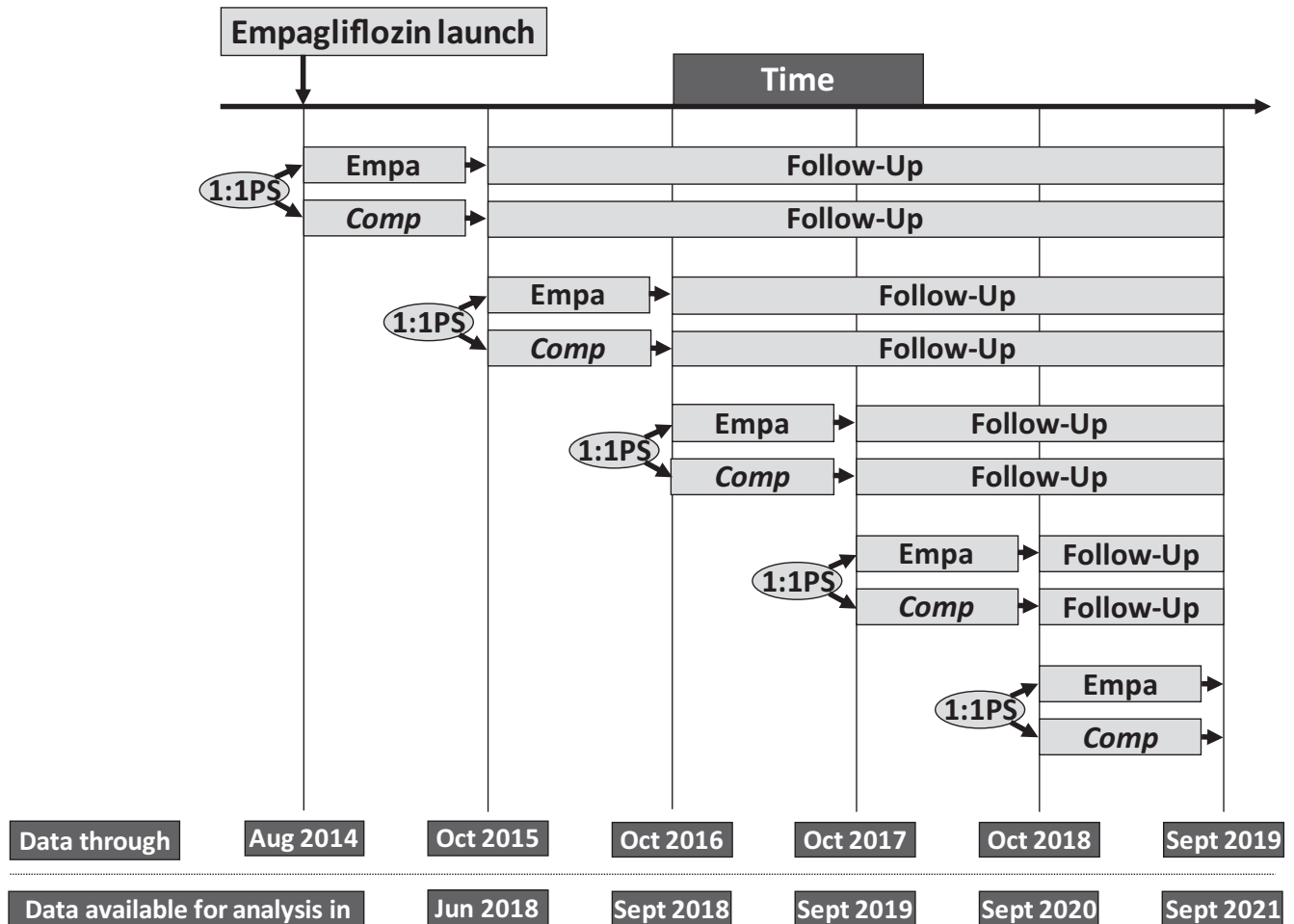
All individual data were de-identified, the study was approved by the Brigham and Women's Hospital institutional review board, and signed data licence agreements were in place for all data sources. The study was registered at EnCEPP (EUPAS20677) and on ClinicalTrials.gov (NCT03363464).

### 2.2 | Study design

EMPRISE is a sequentially built new-user active-comparator cohort study that includes four planned interim analyses and one final analysis, using data from August 2014 to September 2019 (Figure 1). For this manuscript, we focus on the first year of data.

New users of empagliflozin are 1:1 propensity score matched with initiators of dipeptidyl peptidase-4 (DPP-4) inhibitor. Cohort entry was the day of the first filled prescription of empagliflozin or a DPP-4 inhibitor, defined as no use of either SGLT2 inhibitors or DPP-4 inhibitors in the previous year among patients with at least a year of continuous enrolment prior to cohort entry. Study participants were patients aged 18 years or older with T2D, that is, with an inpatient or outpatient T2D diagnosis recorded during the year prior to drug initiation. Patients were excluded from the study if they had a diagnosis of type 1 diabetes, history of secondary or gestational diabetes, malignancy, end-stage renal disease, human immunodeficiency virus, organ transplant or a nursing home admission during the year prior to cohort entry (Table S1 and Figure S1).

This study design, based on sequentially built cohorts of new users, reduces confounding arising from differences between patients prescribed with the two treatments under investigation.<sup>7,12,15</sup> Furthermore, the study inclusion of new users of either a empagliflozin or a DPP-4 inhibitor, who had no use of either SGLT2 inhibitors or DPP-4 inhibitors during the year prior to cohort entry, reduces



**FIGURE 1** Basic schematic for sequential new-user cohort creation and timeline for data accumulation and expected results availability

chances of time lag and immortal time biases that have plagued other real-world data studies of oral antidiabetics.<sup>16,17</sup> DPP-4 inhibitors were chosen as the primary comparator group because they represent a comparable therapeutic alternative in a similar position in T2D treatment pathway, have similar glycemic efficacy and hypoglycaemia risk, and have demonstrated to be neutral on cardiovascular outcomes;<sup>18-22</sup> this enhances clinical equipoise across treatment groups and reduces confounding.<sup>23,24</sup> Other comparators are considered for future analyses. As randomized controlled studies (RCTs) on SGLT2 inhibitors<sup>1,2</sup> and other antidiabetic agents<sup>25,26</sup> are contributing new evidence that is prompting labelling changes<sup>27,28</sup> and shaping clinical guidelines,<sup>29</sup> the EMPRISE protocol specifies for the use of additional comparison groups that over time may become more clinically meaningful to serve as contrasts for empagliflozin in particular population subgroups, such as patients with cardiovascular or renal disease, that is, glucagon-like peptide-1 (GLP-1) receptor agonists. In anticipation of time-related changes in prescribing patterns, EMPRISE is designed to include tight 1:1 matching on calendar time.

Follow-up for study outcomes started on the day after the initiation of treatment and continued in an "as-treated" approach until discontinuation or switch to a drug in the comparator class, the occurrence of an outcome of interest, a nursing home admission,

death, healthcare plan disenrolment or end of the study period (30 September 2015 for this first interim analysis), whichever came first. We extended the exposure effect window until 30 days after the end of the last prescription's supply. In order to address more fully potential exposure misclassification, informative censoring and nature of the specific outcome assessed,<sup>24</sup> planned sensitivity analyses include the redefinition of the exposure risk window to 14 or 90 days after the end of the last prescription's days' supply interval, and the application of an intention-to-treat type approach, which will carry forward the initial exposure until the occurrence of a study event, healthcare plan disenrolment, admission to a nursing home or the end of the study period.

### 2.3 | Study outcomes

EMPRISE will assess several effectiveness, safety and healthcare utilization outcomes (Figure 2). Primary outcomes of interest include the following: a 3-point major adverse cardiovascular events (MACE) outcome, including nonfatal myocardial infarction, nonfatal stroke or cardiovascular mortality, hospitalization for heart failure (HHF) and all-cause mortality. In prior investigations, the positive predictive values of the validated ICD-9-code-based algorithms used to define

<b>Effectiveness</b>	- 3P-MACE - CV mortality - Myocardial infarction	- Stroke - HHF - All-cause death	- ESRD - Coronary revascularization - Treatment for retinopathy
<b>Safety</b>	- Severe hypoglycaemia - Lower limb amputation	- Bone fractures - DKA	- AKI requiring dialysis - Urinary tract cancers
<b>Resources and costs</b>	- Hospitalizations - Length of stay - Outpatient visits	- ED visits - No. filled drugs - Total care costs	- Inpatient costs - Outpatient costs - Pharmacy costs

**FIGURE 2** EMPRISE study outcomes of interest. 3P-MACE: three-point major cardiovascular event; AKI, acute kidney injury; CV, cardiovascular; DKA, diabetic ketoacidosis; ESRD, end-stage renal disease; ED, emergency department; No., number

these cardiovascular primary outcomes were 84% or higher.<sup>30-33</sup> Detailed definitions for all study outcomes are available in Table S2.

For the scope of this manuscript, we focused on HHF, defined as a hospitalization with a discharge diagnosis of heart failure in the primary position (positive predictive value [PPV] = 84%-100%),<sup>33</sup> the primary effectiveness outcome with the lowest expected incidence rate. This was to ensure that sufficient power was achieved for other primary outcomes. We also explored a broader definition of HHF, defined as a hospitalization with a discharge diagnosis of heart failure in any position (PPV = 79%-96%).<sup>33</sup>

## 2.4 | Patient characteristics

Patient baseline characteristics were measured on the basis of enrolment information and claims during the 12 months before and including the date of cohort entry.

Covariates of interest included demographics, calendar time (in quarters and days), comorbidities, diabetes-specific complications, use of diabetes drugs, use of other medications, indicators of health-care utilization as proxy for overall disease state, care intensity and surveillance, and laboratory test results, where available (Table 1; Appendix S1, Table S3). Patient characteristics were defined using ICD-9 diagnosis codes, CPT-4 procedure codes, NDC pharmacy codes and LOINC codes. Particular emphasis was placed on the identification of claims-measured indicators of diabetes severity, including number of glucose-lowering medications at index date and specific past or concurrent diabetes therapy, diabetic nephropathy, neuropathy, retinopathy, diabetic foot and lower limb amputations, number of HbA1c or glucose tests ordered. To address the potential evolving prescribing preferences towards empagliflozin over time for patients with history of cardiovascular disease, we captured and matched upon cardiovascular conditions and medications with high granularity (Table 1; Appendix S1, Table S3).

For a subset of about 45%-50% of patients in Optum and 5%-10% in MarketScan, laboratory test results during the 12 months before cohort entry were available through linkage with two national laboratory test provider chains. In this subpopulation, laboratory test results were used to assess potential residual confounding by unmeasured factors not considered in the 1:1 propensity score matching on claims information only, a method that has demonstrated the success in confounding control in studies of patients with T2D.<sup>23</sup> The value closest to the cohort entry date (within the 1-year baseline period) was considered.

## 2.5 | Analysis

The statistical analysis plan for EMPRISE is reported in the Appendix S1, Summary of EMPRISE statistical analysis plan. Here, we focus on the analyses required to address the study validity.

To assess the ability of EMPRISE to overcome confounding, we cross-tabulated baseline patient characteristics by initiation of empagliflozin or DPP-4 inhibitors between August 2014 and September 2015 for each of the three cohorts. We evaluated frequencies and percentages for binary variables; and means (standard deviations) and medians (25th and 75th interquartile range) for continuous variables. Within each cohort, we estimated an exposure propensity score (PS) using a multivariable logistic regression predicting the initiation of empagliflozin vs. a DPP-4 inhibitor, conditional upon over 140 predefined baseline characteristics (Appendix S1: Table S3).<sup>34</sup> We 1:1 PS-matched patients using the nearest neighbour methodology with a maximum caliper of 0.01 of the PS.<sup>35,36</sup> Postmatching covariate balance between empagliflozin and DPP-4 inhibitors was assessed by the calculation of two metrics for covariate balance: standardized differences, with meaningful imbalances set at values >0.1,<sup>37</sup> and postmatching discrimination using the c-statistic from the PS model refit in the matched sample, which is expected to be

**TABLE 1** Selected baseline characteristics of empagliflozin and DPP-4 inhibitors before and after propensity score matching

Baseline characteristics	Unmatched			Propensity score matched		
	Overall study population			Overall study population		
	DPP-4 inhibitors (N = 170 841)	Empagliflozin (N = 7089)	St. Diff.	DPP-4 inhibitors (N = 6643)	Empagliflozin (N = 6643)	St. Diff.
<b>Demographics</b>						
Age; mean (SD)	67.25 (9.72)	57.12 (8.85)	1.09	57.36 (8.97)	57.40 (8.89)	0.00
Male; n (%)	81 890 (47.9%)	3761 (53.1%)	-0.01	3555 (53.5%)	3529 (53.1%)	0.00
White race <sup>a</sup> ; n (%)	70 390 (75.5%)	70 390 (75.5%)	0.24	70 390 (75.5%)	70 390 (75.5%)	0.02
<b>Burden of comorbidities</b>						
Combined comorbidity score <sup>b</sup> ; mean (SD)	0.44 (1.65)	0.10 (1.22)	0.23	0.61 (1.40)	0.61 (1.39)	0.00
Frailty score <sup>c</sup> , mean (SD)	0.15 (0.05)	0.14 (0.04)	0.22	0.14 (0.04)	0.14 (0.04)	0.00
<b>Diabetes-related complications</b>						
Diabetic nephropathy; n (%)	15 991 (9.4%)	462 (6.5%)	0.01	408 (6.1%)	427 (6.4%)	0.00
Diabetic retinopathy; n (%)	13 854 (8.1%)	496 (7.0%)	0.00	474 (7.1%)	456 (6.9%)	0.00
Diabetic neuropathy; n (%)	29 840 (17.5%)	1102 (15.5%)	0.00	1058 (15.9%)	1010 (15.2%)	0.00
Diabetes with peripheral circulatory disorders; n (%)	11 325 (6.6%)	265 (3.7%)	0.01	232 (3.5%)	244 (3.7%)	0.00
Diabetic Foot; n (%)	4443 (2.6%)	142 (2.0%)	0.00	136 (2.0%)	130 (2.0%)	0.00
Hypoglycaemia; n (%)	10 511 (6.2%)	379 (5.3%)	0.00	335 (5.0%)	354 (5.3%)	0.00
<b>Features of diabetes medication initiation and baseline diabetes therapy</b>						
No. antidiabetic drugs at cohort entry; mean (SD)	2.14 (0.78)	2.31 (0.97)	-0.19	2.24 (0.84)	2.26 (0.94)	-0.02
Naive new user; n (%)	96 148 (56.3%)	1683 (23.7%)	0.05	1638 (24.6%)	1661 (25.0%)	0.00
Monotherapy; n (%)	110 469 (64.7%)	2291 (32.3%)	0.05	2301 (34.6%)	2265 (34.1%)	0.00
Dual therapy with metformin; n (%)	38 130 (22.3%)	787 (11.1%)	0.03	742 (11.2%)	768 (11.6%)	0.00
Concomitant initiation or current use of metformin; n (%)	106 505 (62.3%)	4209 (59.4%)	0.00	3922 (59.0%)	3954 (59.5%)	0.00
Concomitant initiation or current use of sulfonylureas; n (%)	56 152 (32.9%)	1821 (25.7%)	0.01	1675 (25.2%)	1740 (26.2%)	0.00
Concomitant initiation or current use of insulin; n (%)	18 070 (10.6%)	1600 (22.6%)	-0.03	1406 (21.2%)	1402 (21.1%)	0.00
Past use of metformin; n (%)	24 742 (14.5%)	1305 (18.4%)	-0.01	1236 (18.6%)	1230 (18.5%)	0.00
Past use of sulfonylureas; n (%)	16 522 (9.7%)	652 (9.2%)	0.00	593 (8.9%)	587 (8.8%)	0.00
Past use of insulin; n (%)	6638 (3.9%)	658 (9.3%)	-0.02	580 (8.7%)	554 (8.3%)	0.00
<b>Life style factors</b>						
Obesity; n (%)	38 579 (22.6%)	2189 (30.9%)	-0.02	1994 (30.0%)	2003 (30.1%)	0.00
Overweight; n (%)	7966 (4.7%)	267 (3.8%)	0.00	254 (3.8%)	250 (3.8%)	0.00
Smoking; n (%)	23 880 (14.0%)	675 (9.5%)	0.01	625 (9.4%)	644 (9.7%)	0.00
<b>Other comorbidities at baseline</b>						
Ischaemic heart disease; n (%)	44 534 (26.1%)	1113 (15.7%)	0.02	1026 (15.4%)	1049 (15.8%)	0.00
Previous coronary revascularization; n (%)	13 246 (7.8%)	253 (3.6%)	0.02	219 (3.3%)	243 (3.7%)	0.00
Ischaemic or haemorrhagic stroke; n (%)	17 419 (10.2%)	329 (4.6%)	0.02	315 (4.7%)	313 (4.7%)	0.00
Heart failure; n (%)	18 653 (10.9%)	302 (4.3%)	0.02	280 (4.2%)	292 (4.4%)	0.00
Peripheral arterial disease or surgery; n (%)	18 344 (10.7%)	328 (4.6%)	0.02	312 (4.7%)	314 (4.7%)	0.00
Hypertension; n (%)	141 117 (82.6%)	5291 (74.6%)	0.01	4947 (74.4%)	4941 (74.3%)	0.00

(Continues)

TABLE 1 (Continued)

Baseline characteristics	Unmatched			Propensity score matched		
	Overall study population			Overall study population		
	DPP-4 inhibitors (N = 170 841)	Empagliflozin (N = 7089)	St. Diff.	DPP-4 inhibitors (N = 6643)	Empagliflozin (N = 6643)	St. Diff.
<b>Demographics</b>						
Hyperlipidemia; n (%)	139 697 (81.8%)	5622 (79.3%)	0.00	5249 (79.0%)	5246 (78.9%)	0.00
Oedema; n (%)	18 375 (10.8%)	449 (6.3%)	0.02	450 (6.8%)	431 (6.5%)	0.00
Nondiabetic renal dysfunction; n (%)	34 829 (20.4%)	612 (8.6%)	0.03	624 (9.4%)	589 (8.9%)	0.00
Chronic kidney disease; n (%)	29 889 (17.5%)	446 (6.3%)	0.03	438 (6.6%)	432 (6.5%)	0.00
COPD; n (%)	18 581 (10.9%)	383 (5.4%)	0.02	370 (5.6%)	367 (5.5%)	0.00
Obstructive sleep apnoea; n (%)	17 053 (10.0%)	1068 (15.1%)	-0.01	961 (14.5%)	977 (14.7%)	0.00
Pneumonia; n (%)	7287 (4.3%)	174 (2.5%)	0.01	151 (2.3%)	163 (2.5%)	0.00
Liver disease; n (%)	11 734 (6.9%)	486 (6.9%)	0.00	435 (6.5%)	455 (6.8%)	0.00
Osteoarthritis; n (%)	40 824 (23.9%)	1089 (15.4%)	0.02	963 (14.5%)	1033 (15.5%)	0.00
<b>Other Medications</b>						
Angiotensin converting enzyme inhibitors; n (%)	79 712 (46.7%)	3293 (46.5%)	0.00	3085 (46.4%)	3082 (46.4%)	0.00
Angiotensin II receptor blockers; n (%)	52 655 (30.8%)	2126 (30.0%)	0.00	1925 (29.0%)	1974 (29.7%)	0.00
Beta-blockers; n (%)	71 567 (41.9%)	2220 (31.3%)	0.02	2057 (31.0%)	2078 (31.3%)	0.00
Calcium-channel blockers; n (%)	53 432 (31.3%)	1562 (22.0%)	0.02	1590 (23.9%)	1467 (22.1%)	0.00
Thiazide diuretics; n (%)	24 190 (14.2%)	808 (11.4%)	0.01	805 (12.1%)	759 (11.4%)	0.00
Loop diuretics; n (%)	28 161 (16.5%)	668 (9.4%)	0.02	586 (8.8%)	622 (9.4%)	0.00
Nitrates; n (%)	13 163 (7.7%)	310 (4.4%)	0.01	293 (4.4%)	291 (4.4%)	0.00
Other hypertension drugs; n (%)	13 648 (8.0%)	320 (4.5%)	0.01	321 (4.8%)	306 (4.6%)	0.00
Digoxin; n (%)	4568 (2.7%)	095 (1.3%)	0.01	079 (1.2%)	090 (1.4%)	0.00
Statins; n (%)	118 401 (69.3%)	4631 (65.3%)	0.00	4276 (64.3%)	4301 (64.7%)	0.00
Antiplatelet agents; n (%)	24 496 (14.3%)	668 (9.4%)	0.01	605 (9.1%)	634 (9.5%)	0.00
Oral anticoagulants; n (%)	13 447 (7.9%)	295 (4.2%)	0.02	283 (4.3%)	284 (4.3%)	0.00
Opioids; n (%)	56 179 (32.9%)	2331 (32.9%)	0.00	2177 (32.8%)	2188 (32.9%)	0.00
<b>Measures of healthcare utilization</b>						
Previous hospitalization; n (%)	24 336 (14.2%)	549 (7.7%)	0.02	488 (7.3%)	519 (7.8%)	0.00
Hospitalization within prior 30 d; n (%)	7601 (4.4%)	063 (0.9%)	0.02	055 (0.8%)	063 (0.9%)	0.00
No. emergency department visits; mean (sd)	0.73 (2.02)	0.31 (1.40)	0.24	0.31 (1.44)	0.30 (1.17)	0.01
No. office visits; mean (SD)	9.49 (7.26)	8.76 (6.36)	0.11	8.63 (6.62)	8.68 (6.33)	-0.01
Endocrinologist visit within prior 30 d; n (%)	13 066 (7.6%)	1401 (19.8%)	-0.03	1117 (16.8%)	1118 (16.8%)	0.00
Internal medicine visit within prior 30 d; n (%)	111 623 (65.3%)	4171 (58.8%)	0.01	4074 (61.3%)	4003 (60.2%)	0.00
Cardiologist visit within prior 30 d; n (%)	18 618 (10.9%)	394 (5.6%)	0.02	366 (5.5%)	378 (5.7%)	0.00
No. distinct medication prescriptions; mean (SD)	12.16 (6.06)	12.42 (6.08)	-0.04	12.19 (6.23)	12.25 (6.05)	-0.01
<b>Laboratory test results<sup>d</sup></b>						
HbA1c (%); mean (SD)	8.41 (1.82)	8.49 (1.76)	-0.04	8.62 (1.84)	8.48 (1.76)	0.08

(Continues)

TABLE 1 (Continued)

Baseline characteristics	Unmatched			Propensity score matched		
	Overall study population			Overall study population		
	DPP-4 inhibitors (N = 170 841)	Empagliflozin (N = 7089)	St. Diff.	DPP-4 inhibitors (N = 6643)	Empagliflozin (N = 6643)	St. Diff.
Demographics						
Patients with HbA1c results available; n (%)	15 768 (20.3%)	943 (16.2%)	0.01	1002 (18.6%)	864 (16.0%)	0.01
Creatinine (mg/dL); mean (SD)	0.96 (0.34)	0.89 (0.23)	0.24	0.91 (0.27)	0.89 (0.23)	0.08
Patients with creatinine results available; n (%)	16 248 (20.9%)	1044 (18.0%)	0.01	1046 (19.4%)	952 (17.7%)	0.00
Total cholesterol (mg/dL); mean (SD)	179.38 (47.09)	177.44 (47.46)	0.04	177.75 (45.04)	177.92 (48.05)	0.00
Patients with total cholesterol results available; n (%)	14 494 (18.7%)	944 (16.2%)	0.01	955 (17.7%)	863 (16.0%)	0.00
LDL level (mg/dL); mean (sd)	90.82 (40.13)	87.09 (38.55)	0.09	88.92 (39.42)	87.47 (38.91)	0.04
Patients with LDL results available; n (%)	14 984 (19.3%)	943 (16.2%)	0.01	976 (18.1%)	864 (16.0%)	0.01
HDL level (mg/dL); mean (SD)	45.74 (21.76)	44.44 (13.04)	0.07	43.96 (12.59)	44.39 (13.04)	-0.03
Patients with HDL results available; n (%)	14 325 (18.5%)	926 (15.9%)	0.01	942 (17.5%)	846 (15.7%)	0.00

Abbreviations: DPP-4: dipeptidyl peptidase-4; St. Diff: standardized differences, that is, the difference in means or proportions divided by the pooled standard deviation [Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine* 2009;28:3083-107]; SD: standard deviation; Q: quarter; GLP-1 RA: glucagon-like peptide-1 receptor agonists; COPD: chronic obstructive pulmonary disease; BUN: blood urea nitrogen; HbA1c: haemoglobin A1c; LDL: low-density lipoprotein; HDL: high-density lipoproteins.

<sup>a</sup>Only available in Medicare fee-for-service.

<sup>b</sup>Gagne et al.<sup>42</sup>

<sup>c</sup>Kim et al.<sup>56</sup>

<sup>d</sup>Only available in Optum Clinformatics and Truven MarketScan.

close to 0.5 if multivariate balance has been achieved.<sup>38</sup> We also assessed the postmatching balance achieved for selected laboratory test results, which have not been included in the PS model, including HbA1c, creatinine and lipid levels.

We developed a model to predict the time when a statistical power of 80% could be achieved for the two HHF outcome definitions across the three EMPIRE data sources. For this purpose, we first used the power calculation formula for a two-group survival analysis to calculate the minimal number of events required to achieve sufficient power to reject the null hypothesis for each of the two HHF outcome definitions. We then simulated cohorts of new empagliflozin and DPP-4 users based on the expected number of users at different years of EMPIRE. Specifically, the annual utilization rates of empagliflozin based on market data (August 2014 until November 2017) provided by the manufacturer and the number of empagliflozin new users captured in the first year of EMPIRE (August 2014 to September 2015) was used to estimate the expected number of empagliflozin new users in the subsequent years. Incidence of HHF outcomes or censoring events was modelled based on the rates observed in our first interim analysis. The model identified the earliest time at which the total number of HHF events exceeds the required minimal threshold determined by the power calculation. In sensitivity analyses, we tested the robustness of the

main findings under different scenarios for censoring rates and incidence rates of HHF outcomes (Appendix S1, Additional details of the model used to determine the timing of achieving statistical power for HHF outcomes).

Analyses were conducted using the validated Aetion platform<sup>39</sup> and R version 3.1.2, SAS 9.4 Statistical Software (SAS Institute Inc), and Arena, version 15.00 (Rockwell Automation) for simulation.

### 3 | RESULTS

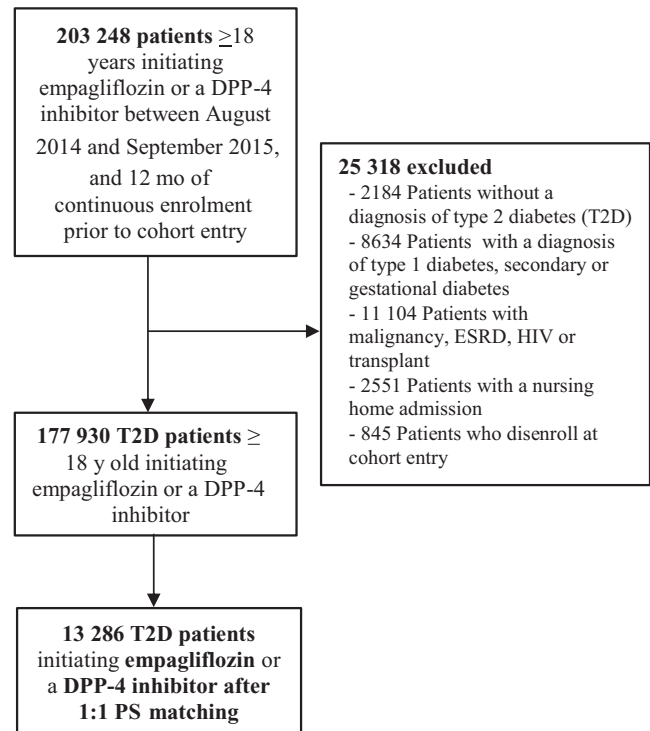
After the implementation of inclusion and exclusion criteria for the first interim analysis, we identified 7096 patients who initiated empagliflozin (1446 initiators in Optum, 4368 in MarketScan and 1282 in Medicare) and 171 358 patients who initiated DPP-4 inhibitors (23 693 initiators in Optum, 53 918 in MarketScan and 93 747 in Medicare) between 1 August 2014 and 30 September 2015 (Figure 3; Figure S1). After 1:1 PS matching, most empagliflozin initiators (6643; 93.6%) were successfully matched to an initiator of DPP-4 inhibitor, with 1317 matched pairs in Optum, 4075 pairs in MarketScan and 1251 pairs in Medicare. Across all data sources, compared with initiators of DPP-4 inhibitors, patients initiating empagliflozin before PS matching were generally younger, more frequently male, less frail as

measured by the Claims-Based Frailty Index (CFI),<sup>40</sup> and had a lower general burden of comorbidities as measured by the combined comorbidity score<sup>41</sup> and by the prevalence of individual comorbidities at baseline, including ischaemic heart disease, stroke, heart failure, peripheral vascular disease, chronic kidney disease and diabetic nephropathy (Table 1; Table S3); they also experienced fewer hospitalizations or emergency department visits during the year prior to treatment initiation. Conversely, empagliflozin initiators had higher prevalence of diabetic neuropathy or obesity, higher baseline use of insulin or glucagon-like peptide (GLP)-1 receptor agonists, higher number of antidiabetic medications at cohort entry, were less likely to be naïve new users, defined as not having used any diabetes treatment during the one year prior to cohort entry and were more likely to have had a visit with an endocrinologist prior to treatment initiation. After PS matching, all baseline patient characteristics were well balanced between initiators of empagliflozin and DPP-4 inhibitors with standardized differences smaller than 0.1 (Table 1; Table S4). For the subset of the population with laboratory values, available test results for HbA1c and creatinine were well balanced after PS matching despite not having been components of the PS model. Depending on the individual data source, the PS model c-statistic moved from 0.82-0.87 before matching to 0.54-0.59 after matching (Table S5).

### 3.1 | Population incidence of HHF, projected utilization trends and power considerations

After PS matching, the population mean (standard deviation) follow-up time was 109 (78) days in Optum, 133 (96) days in MarketScan and 114 (86) days in Medicare, for an overall mean follow-up time of 125 (91) days (Table 2). Most patients were censored due to the end of the 1-year study period, that is, 30 September 2015, (59.8% in Optum, 49.7% in MarketScan and 58.7% in Medicare) or to the discontinuation of the original cohort-defining treatment (32.5% in Optum, 35.5% in MarketScan and 35.6% in Medicare) (Table S6). As subsequent years of data accumulate over the EMPRISE study period, follow-up time will increase. The population incidence rate of the more specific HHF definition was 2.53 per 1000 person-years in Optum, 3.03 in MarketScan and 11.51 in Medicare, and 4.40 across the three data sources (Table 2). For the broader HHF definition, the overall incidence rate was 14.78 per 1000 person-years (Table 2).

Based on our power calculations, to detect a clinically meaningful effect on both HHF outcomes in line with the hazard ratio observed in the EMPA-REG OUTCOME trial (HR = 0.65), a minimum of 169 overall events among initiators of empagliflozin or DPP-4 inhibitors are needed to achieve 80% power for rejecting the null hypothesis for each of the two HHF outcomes at 5% significance level. Our model was based on the expected number of empagliflozin initiators at different years of EMPRISE (Table S7) and the observed rates of HHF outcomes or censoring events. It suggested that the number of accumulated events will exceed the threshold of 169 events by the end of year 3 for the more specific HHF definition, assuming



**FIGURE 3** Flowchart of overall study population. DPP-4, dipeptidyl peptidase-4; ESRD, end-stage renal disease; HIV, human immunodeficiency virus; PS, propensity score

~37 000 accrued empagliflozin initiators, and by the end of year 2 for the broader definition, assuming ~17 000 accrued empagliflozin initiators. (Figure 4; Tables S8, S9). Because the periodic analyses in a prospective database study are not used to potentially end the study, to trigger regulatory action, to initiate follow-up studies or to prevent the publication or dissemination of the results from subsequent analyses, unlike a clinical trial, primary analyses will not be corrected for multiple looks over time via group sequential analysis. Fully adjusted hazard ratios and 95% CIs will be estimated at each of the four interim analyses and the final analysis for all EMPRISE outcomes (Appendix S1, Summary of EMPRISE statistical analysis plan).

## 4 | DISCUSSION

In a first interim analysis from the first year of EMPRISE, a new-user active-comparator cohort study, we demonstrated solid confounding control, as measured by the superior balance across treatment groups in a wide range of potential confounding factors and their proxies after propensity score matching and confirmed that we will reach adequate patient accrual rates for the achievement of powered interim analyses for all primary outcomes.

In the EMPA-REG OUTCOME trial,<sup>1</sup> empagliflozin reduced the relative risk of cardiovascular death, all-cause mortality and hospitalization for heart failure, in patients with T2D and established cardiovascular disease. These findings prompted the US Food and Drug Administration and other regulatory agencies to expand empagliflozin's



**TABLE 2** Population incidence rates for HHF outcomes within individual data sets and overall

	OPTUM Commercial + Medicare Advantage	MARKETSCAN Commercial + Medicare Advantage	MEDICARE Fee-for-service	Overall Study Population
Total 1:1 PS-matched pairs	2634	8150	2502	13 286
Outcomes				
Specific HHF outcome <sup>a</sup>				
Events	2	9	<11 <sup>b</sup>	20
Follow-up, mean (SD)	109.4 (71.7)	133.1 (96.3)	113.8 (86.0)	124.7 (90.6)
Incidence rate per 1000 person-years (95% CI)	2.53 (0.98-6.05)	3.03 (1.05-5.01)	11.51 (3.99-19.04)	4.40 (2.47-6.33)
Broad HHF outcome <sup>c</sup>				
Events	9	9	26	44
Follow-up, mean (SD)	109.2 (71.7)	132.9 (96.2)	113.3 (85.9)	124.5 (90.6)
Incidence rate per 1000 person-years (95% CI)	11.42 (3.96-18.89)	8.76 (5.39-12.13)	41.11 (26.87-55.35)	14.78 (11.24-18.32)

Abbreviations: CI, confidence interval; HHF, Heart failure hospitalization; PS, propensity score; SD, standard deviation.

<sup>a</sup>Defined as a hospitalization with a discharge diagnosis of heart failure in the primary position.

<sup>b</sup>In accordance with the data use agreement, we do not report information for frequency cells with <11 cases. These are noted as <11.

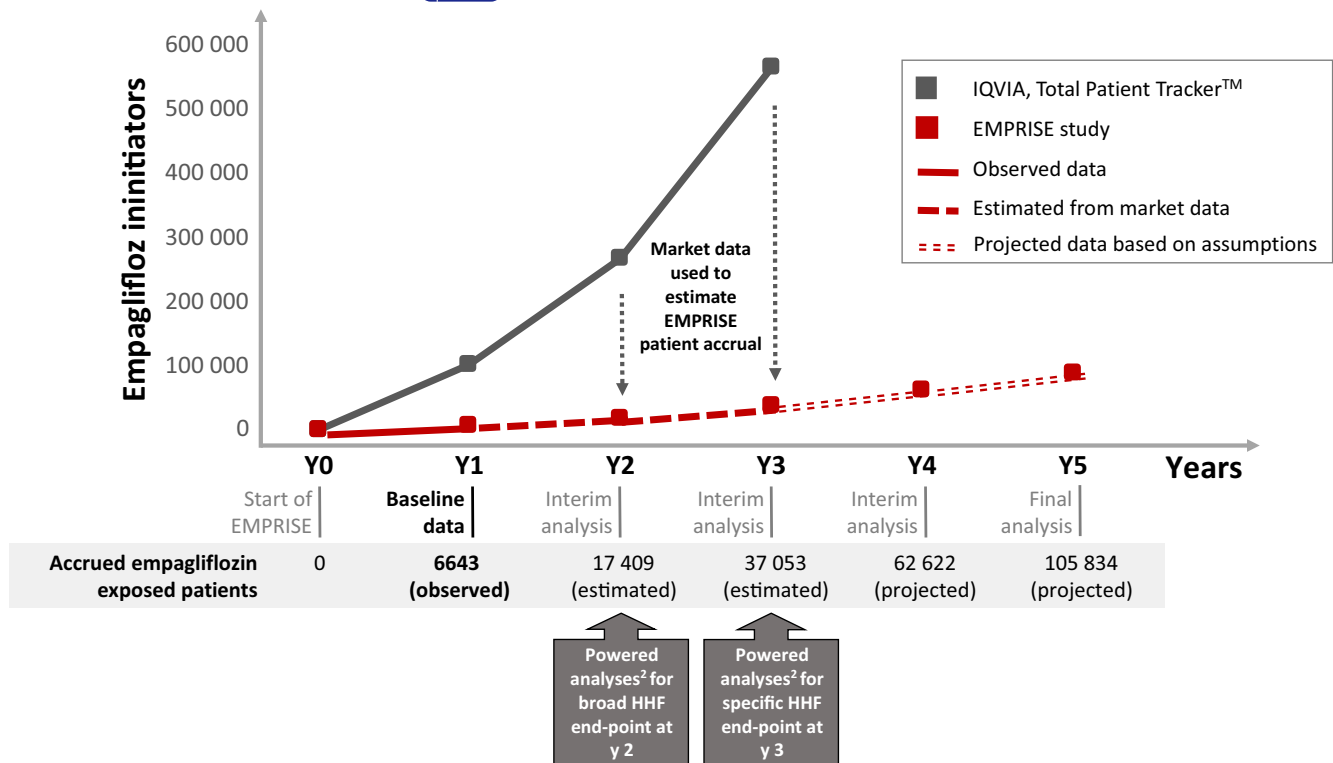
<sup>c</sup>Defined as a hospitalization with a discharge diagnosis of heart failure in any position.

indication to reduce the risk of cardiovascular death in adult patients with T2D and cardiovascular disease<sup>42</sup> and led to changes in major clinical guidelines with regard to diabetes treatment among patients with cardiovascular disease,<sup>29,43</sup> which affects approximately one-third of all patients with T2D.<sup>44</sup> EMPRISE is an ongoing study programme on the comparative effectiveness, safety and healthcare utilization of empagliflozin based on real-world data among patients with T2D as treated in routine clinical care, which will collect accumulating data on empagliflozin for a period of five years following the date of approval in the United States, 1 August 2014 and will complement the findings from the EMPA-REG OUTCOME trial.

In the context of diabetes treatment, long-term outcome trials required by regulators<sup>45,46</sup> to assess the cardiovascular safety of antidiabetic medications have yielded important information on both the benefits and risks associated with glucose-lowering agents. Nonetheless, RCTs may not be able to answer all questions<sup>47</sup> as (a) they are less reflective of the patients and treatment patterns encountered in routine care; (b) are not powered to detect rare noncardiovascular risks that may become evident in larger and more broadly defined populations<sup>48</sup>; and (c) are often placebo-controlled, thus not directly addressing the question faced by prescribers as several alternative medications are available. Studies based on real-world data sources offer a tremendous opportunity to fill these gaps and guide decision-making.<sup>49</sup> Over the past 20 years, tremendous advances have been made

and new design and analytic standards are emerging to limit avoidable design flaws, for example, time-related biases,<sup>16,17</sup> in which this monitoring programme incorporates to achieve high validity findings.<sup>50-52</sup>

EMPRISE was designed to enhance clinical equipoise across treatment groups and minimize chances of confounding and time-related biases.<sup>16,17,24,53</sup> Specifically, EMPRISE (a) does not use a hierarchical definition of the exposure allowing switchers from a DPP-4 inhibitor to empagliflozin to enter the cohort as empagliflozin initiators resulting in possible immortal time bias,<sup>16,17</sup> but rather includes new users of empagliflozin or a DPP-4 inhibitor, without any use of either SGLT2 or DPP-4 inhibitors during the year prior to cohort entry<sup>7</sup>; (b) does not contrast empagliflozin to diabetes agents used at the extremes of the treatment algorithm for T2D, for example, metformin or insulin, but includes comparators (ie DPP-4 inhibitors) recognized as comparable therapeutic alternatives for patients with T2D,<sup>54</sup> thus improving clinical equipoise for diabetes severity and duration between exposure groups and reducing chances of time-lag bias<sup>55</sup>; and (c) employs an extensive propensity score adjustment for many proxies of diabetes severity and duration, for example, baseline use of insulin and other diabetes agents, diabetes-related complications and measures of healthcare utilization, which have demonstrated success in balancing measured and unmeasured characteristics in studies of patients with T2D<sup>23</sup> and which can also limit time-lag



**FIGURE 4** Timing for accrual of PS-matched empagliflozin exposed patients and achievement of powered analyses for HHF outcomes in EMPRISE<sup>1</sup>. HHF, heart failure hospitalization; PS, propensity score. 1 For details on the underlying model, see Appendix S1<sup>1</sup> "Additional details of the model used to determine the timing of achieving statistical power for HHF outcomes" 2 Powered analyses to reach a minimum of 169 HHF events

bias.<sup>55</sup> Furthermore, the inclusion of head-to-head comparisons of specific alternative diabetes treatment options answers the more clinically relevant question of which medication to choose for optimal diabetes care, and the broadly defined population comprised of patients as treated in routine care without restrictions enables assessment of the effects of empagliflozin across a broad range of patients with diabetes.

However, residual confounding by some unmeasured characteristic(s) cannot be entirely ruled out although it is likely to be minor. Selected laboratory test results, including HbA1c and creatinine, were well balanced after propensity score adjustment, despite not having been included in the propensity score model as only available for a subset of the population; previous work also demonstrated that a new-user active-comparator cohort study paired with propensity score matching on many proxies of diabetes severity and duration improves balance in covariates typically unmeasured in administrative claims data sets, for example, duration of diabetes.<sup>23</sup> Even though heart failure outcomes were defined using previously validated claims-based algorithms with high positive predictive value,<sup>33</sup> some extent of outcome misclassification remains a possibility. At this stage of EMPRISE, the short duration of follow-up, mainly driven by the availability for analysis of only 1 year of empagliflozin use, limits the assessment of the long-term effects of empagliflozin. However, the decreased risk of HHF observed in RCTs appeared equally early. Thus, the short follow-up observed in the current study is not expected to affect the assessment of HHF.

## 5 | CONCLUSION

EMPRISE is a study programme on the comparative effectiveness, safety and healthcare utilization of empagliflozin in patients with T2D as treated in routine care. Baseline information from the first year of empagliflozin use provides evidence of solid confounding control and adequate exposure accrual with expected powered analyses for all primary outcomes by the end of year 3 of the study programme. These elements are crucial to inform the level of confidence in future findings.

## ACKNOWLEDGMENTS

This study was supported by a research grant to the Brigham and Women's Hospital from Boehringer Ingelheim. The study was conducted by the authors independent of the sponsor. The authors retained the right of publication and determined the final wording of the manuscript. EP was supported by a career development grant K08AG055670 from the National Institute on Aging. MK was supported by the National Institute of General Medical Sciences, grant RO1GM108999. EP is investigator of investigator-initiated grants to the Brigham and Women's Hospital from GSK, not related to the topic of the submitted work. SS is the principal investigator of investigator-initiated grants to the Brigham and Women's Hospital from Bayer, Vertex and Boehringer Ingelheim unrelated to the topic of this study. He is a consultant to WHISCON and to Aetion, a software manufacturer of which he owns equity. His interests were declared,

reviewed and approved by the Brigham and Women's Hospital and Partners HealthCare System in accordance with their institutional compliance policies.

#### AUTHOR CONTRIBUTIONS

EP, SS, MN, JMF, ADL, KGB and MK designed the study. EP, SS, AP, AJSO and LGB involved in conduct/data collection. MN, AP and JMF performed the analysis. EP, SS, MN, AP, JMF, ADL, KGB, MK, AJSO and LGB wrote the manuscript.

#### ETHICS AND DATA ACCESSIBILITY

The study was approved by the Brigham and Women's Hospital Institutional Review Board. Signed data licence agreements were in place for all data sources.

#### ORCID

Elisabetta Patorno  <https://orcid.org/0000-0002-8809-9898>

Anouk Déruaz-Luyet  <https://orcid.org/0000-0001-7396-7392>

#### REFERENCES

- Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in Type 2 diabetes. *N Engl J Med*. 2015;373(22):2117-2128.
- Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and cardiovascular and renal events in Type 2 diabetes. *N Engl J Med*. 2017;377(7):644-657.
- Administration. (USFaD). FDA Drug Safety Communication: FDA revises label of diabetes drug canagliflozin (Invokana, Invokamet) to include updates on bone fracture risk and new information on decreased bone mineral density. Accessed September 10, 2015.
- Administration. (USFaD). Drug Safety Communication: FDA confirms increased risk of leg and foot amputations with the diabetes medicine canagliflozin (Invokana, Invokamet, Invokamet XR). Accessed May 16, 2017.
- Administration. (FaD). FDA Drug Safety Communication: FDA revises labels of SGLT2 inhibitors for diabetes to include warnings about too much acid in the blood and serious urinary tract infections. Accessed December 4, 2015.
- Schneeweiss S. Improving therapeutic effectiveness and safety through big healthcare data. *Clin Pharmacol Ther*. 2016;99(3):262-265.
- Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. *Pharmacoepidemiol Drug Saf*. 2010;19(8):858-868.
- Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? *Clin Pharmacol Ther*. 2017;102(6):924-933.
- S. B. H.R.34—21st Century Cures Act. 114th Congress (2015–2016); December 13, 2016; became Public Law No: 114-255.
- Administration (USFaD). Prescription Drug User Fee Act (PDUFA) – PDUFA VI: Fiscal Years 2018–2022. <https://www.fda.gov/forindustry/userfees/prescriptiondruguserfee/ucm446608.htm>. Accessed July 27, 2018.
- Franklin JM, Rassen JA, Bartels DB, Schneeweiss S. Prospective cohort studies of newly marketed medications: using covariate data to inform the design of large-scale studies. *Epidemiology*. 2014;25(1):126-133.
- Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. *Clin Pharmacol Ther*. 2011;90(6):777-790.
- Seeger JDDG. Commercial insurance databases. In: Strom BLKSE, Hennessy S, eds. *Pharmacoepidemiology*. Philadelphia, PA: Wiley&Sons; 2012.
- Important notice: change in public death master file records. 2011. <https://classic.ntis.gov/assets/pdf/import-change-dmf.pdf>
- Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158(9):915-920.
- Suissa S. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? *Diabetes Care*. 2018;41(1):6-10.
- Suissa S. Reduced mortality with sodium-glucose cotransporter-2 inhibitors in observational studies: avoiding immortal time bias. *Circulation*. 2018;137(14):1432-1434.
- Inzucchi SE, Bergenstal RM, Buse JB, et al. Management of hyperglycemia in type 2 diabetes, 2015: a patient-centered approach: update to a position statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2015;38(1):140-149.
- Green JB, Bethel MA, Armstrong PW, et al. Effect of sitagliptin on cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2015;373(3):232-242.
- Rosenstock J, Perkovic V, Johansen OE, et al. Effect of linagliptin vs placebo on major cardiovascular events in adults with type 2 diabetes and high cardiovascular and renal risk: the CARMELINA randomized clinical trial. *JAMA*. 2019;321(1):69-79.
- Scirica BM, Bhatt DL, Braunwald E, et al. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *N Engl J Med*. 2013;369(14):1317-1326.
- White WB, Cannon CP, Heller SR, et al. Alogliptin after acute coronary syndrome in patients with type 2 diabetes. *N Engl J Med*. 2013;369(14):1327-1335.
- Patorno E, Gopalakrishnan C, Franklin JM, et al. Claims-based studies of oral glucose-lowering medications can achieve balance in critical clinical variables only observed in electronic health records. *Diabetes Obes Metab*. 2018;20(4):974-984.
- Patorno E, Patrick AR, Garry EM, et al. Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations. *Diabetologia*. 2014;57(11):2237-2250.
- Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2016;375(4):311-322.
- Marso SP, Bain SC, Consoli A, et al. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. *N Engl J Med*. 2016;375(19):1834-1844.
- Administration. (USFaD). FDA News Release. FDA approves Jardiance to reduce cardiovascular death in adults with type 2 diabetes; Accessed December 2, 2016.
- ME. T. FDA Grants Liraglutide Cardiovascular Events Indication - Medscape - Aug 25, 2017.
- Davies MJ, D'Alessio DA, Fradkin J, et al. Management of Hyperglycemia in Type 2 Diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care*. 2018;41(12):2669-2701.
- Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, Solomon DH. 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*. 2004;148(1):99-104.
- Wahl PM, Rodgers K, Schneeweiss S, et al. Validation of claims-based diagnostic and procedure codes for cardiovascular and gastrointestinal serious adverse events in a commercially-insured population. *Pharmacoepidemiol Drug Saf*. 2010;19(6):596-603.

32. Tirschwell DL, Longstreth WT Jr. Validating administrative data in stroke research. *Stroke*. 2002;33(10):2465-2470.
33. Saczynski JS, Andrade SE, Harrold LR, et al. A systematic review of validated methods for identifying heart failure using administrative data. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 1):129-140.
34. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127(8 Pt 2):757-763.
35. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. One-to-many propensity score matching in cohort studies. *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 2):69-80.
36. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat*. 2011;10(2):150-161.
37. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083-3107.
38. Franklin JM, Rassen JA, Ackermann D, Bartels DB, Schneeweiss S. Metrics for covariate balance in cohort studies of causal effects. *Stat Med*. 2014;33(10):1685-1699.
39. Wang SV, Verpillat P, Rassen JA, Patrick A, Garry EM, Bartels DB. Transparency and reproducibility of observational cohort studies using large healthcare databases. *Clin Pharmacol Ther*. 2016;99(3):325-332.
40. Kim DH, Glynn RJ, Avorn J, et al. Validation of a claims-based frailty index against physical performance and adverse health outcomes in the health and retirement study. *J Gerontol A Biol Sci Med Sci*. 2019;74(8):1271-1276.
41. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol*. 2011;64(7):749-759.
42. Administration. USFaD. FDA News Release. FDA approves Jardiance to reduce cardiovascular death in adults with type 2 diabetes; Accessed December 2, 2016.
43. American Diabetes A. 8. Pharmacologic approaches to glycemic treatment: standards of medical care in diabetes-2018. *Diabetes Care*. 2018;41(Suppl 1):S73-S85.
44. Einarson TR, Acs A, Ludwig C, Panton UH. 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(1):83.
45. Administration. USFaD. Guidance for Industry. Diabetes Mellitus – Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes; Accessed December 2008.
46. European Medicines Agency. Guideline on clinical investigation of medicinal products in the treatment or prevention of diabetes mellitus;2012.
47. Cefalu WT, Kaul S, Gerstein HC, et al. Cardiovascular outcomes trials in type 2 diabetes: Where do we go from here? Reflections from a diabetes care editors' expert forum. *Diabetes Care*. 2018;41(1):14-31.
48. Fralick M, Schneeweiss S, Patorno E. Risk of diabetic ketoacidosis after initiation of an SGLT2 inhibitor. *N Engl J Med*. 2017;376(23):2300-2302.
49. Jarow JP, LaVange L, Woodcock J. Multidimensional evidence generation and FDA regulatory decision making: defining and using "Real-World" data. *JAMA*. 2017;318(8):703-704.
50. Toh S, Hampp C, Reichman ME, et al. Risk for hospitalized heart failure among new users of saxagliptin, sitagliptin, and other anti-hyperglycemic drugs: a retrospective cohort study. *Ann Intern Med*. 2016;164(11):705-714.
51. Gagne JJ, Glynn RJ, Rassen JA, et al. Active safety monitoring of newly marketed medications in a distributed data network: application of a semi-automated monitoring system. *Clin Pharmacol Ther*. 2012;92(1):80-86.
52. Mayer F, Kirchmayer U, Coletta P, et al. Safety and effectiveness of direct oral anticoagulants versus vitamin K antagonists: pilot implementation of a near-real-time monitoring program in Italy. *J Am Heart Assoc*. 2018;7(6).Article Number: e008034.
53. Patorno E, Garry EM, Patrick AR, et al. Addressing limitations in observational studies of the association between glucose-lowering medications and all-cause mortality: a review. *Drug Saf*. 2015;38(3):295-310.
54. Inzucchi SE, Matthews DR. Management of Hyperglycemia in Type 2 Diabetes American Diabetes A, European Association for the Study of Diabetes Position Statement Writing G. Response to Comments on Inzucchi et al. Management of Hyperglycemia in Type 2 Diabetes, 2015: A Patient-Centered Approach. Update to a Position Statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2015;38:140-149. *Diabetes Care*. Aug 2015;38(8):e128-129.
55. Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care*. 2012;35(12):2665-2673.
56. Kim DH, Schneeweiss S, Glynn RJ. Comparing approaches to measure frailty in medicare data: deficit-accumulation frailty index versus phenotypic frailty. *J Gerontol A Biol Sci Med Sci*. 2018;73(7):989-990.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Patorno E, Najafzadeh M, Pawar A, et al. The EMPagliflozin compaRative effectIveness and SafETy (EMPRISE) study programme: Design and exposure accrual for an evaluation of empagliflozin in routine clinical care. *Endocrinol Diab Metab*. 2020;3:e00103. <https://doi.org/10.1002/edm2.103>