Using real-world data for supporting regulatory decision making: Comparison of cardiovascular and safety outcomes of an empagliflozin randomized clinical trial versus real-world data

Ha Young Jang, In-Wha Kim and Jung Mi Oh*

College of Pharmacy and Research Institute of Pharmaceutical Sciences, Seoul National University, Seoul, South Korea

Aims: In countries where a randomized clinical trial (RCT) is difficult to perform, a real-world evidence (RWE) study with a design similar to an RCT may be an option for drug regulatory decision-making. In this study, the objective was to find out to what extent the safety of empagliflozin from the RWE study in Korea is different from the one in RCT by emulating the design of foreign RCT. The outcome covers various safety outcomes including cardiovascular safety.

Methods: The EMPA-REG OUTCOME trial (NCT01131676) was selected for comparison. The inclusion/exclusion criteria and follow-up method for the RWE were matched to the comparison RCT. Major adverse cardiovascular events (MACEs) were used as a primary outcome and 15 other outcomes were also included for analysis.

Result: We followed 23,126 matched patients with type 2 diabetes mellitus (11,563 empagliflozin users and 11,563 sitagliptin users) for 2.7 years (median). Empagliflozin use was associated with a significantly decreased risk of MACEs [EMPA-REG DUPLICATE RWE: adjusted HR 0.87, 95% confidence interval (CI) 0.79–0.96]. The predefined estimate agreement, regulatory agreement, and standardized difference for RCT duplication were achieved [EMPA-REG OUTCOME RCT: adjusted HR 0.86, 95% CI (0.74–0.99)]. According to the predefined criteria for 15 outcomes, 10 outcomes were evaluated as good, and three as moderate.

Conclusion: Our study results suggest that RWE in one country in comparison with an RCT has the potential for providing evidence for future regulatory decision-making in an environment where RCT could not be performed.

KEYWORDS
real-world evidence, randomized controlled trial, emulation analysis, diabetes mellitus, sitagliptin, empagliflozin
Introduction

Randomized controlled trials (RCTs) are generally regarded as the gold standard for regulatory decision making. Given the growing trend of globalization and the need to make new or extended-use medicines rapidly available to patients worldwide, the RCTs are usually conducted in multi-regional clinical settings (Quan et al., 2017). However, since most clinical trials are conducted in the US and Europe, the proportion of Asians is relatively low. It has been reported that the proportion of clinical trials in Korea among the total clinical trial is about 3% (ClinicalTrials.gov, 2022). Data from multi-regional clinical trials (MRCTs) are submitted to regulatory agencies, which currently find it difficult to evaluate such data for drug approval (Sohn et al., 2019). The main reason is that clinical trial subjects are of different races. Furthermore, it is difficult to conduct additional clinical trials for regulatory decisions like expanding drug indications or adding side effects information, due to time and cost (Revicki and Frank, 1999; Garrison et al., 2007).

Real-world evidence (RWE) is clinical evidence concerning the potential benefits or risks of a medication derived from analysis of real-world data (RWD). RWE has a relative advantage over RCTs because it enables a long-term follow-up study or research on rare populations. In the United States, the 21st Century Cures Act, passed in 2016, placed additional focus on the use of RWE to support regulatory decision making, including adding/modifying an indication, use in a new population, and adding comparative effectiveness or safety information (Food-and-Drug-Administration-FDA, 2018a; Food-and-Drug-Administration-FDA, 2018b; Food-and-Drug-Administration-FDA, 2019a; Food-and-Drug-Administration-FDA, 2019b). With a rise in observational COVID-19 study dissemination, this trend is being accelerated (Pundi et al., 2020). Rather than performing additional RCT in every country to verify new indications or side effects, performing an RWE study in other races and medical-practice conditions could be an alternative way. If the design and analysis method of the RWE study are implemented as closely as possible with the RCT, it will be easier to make regulatory decisions based on comparisons of results of RWEs and RCTs.

Empagliflozin is a sodium-glucose cotransporter 2 (SGLT2) inhibitor drug approved by US Food and Drug Administration (FDA) in 2014 for the treatment of type 2 diabetes (T2DM). After its approval for T2DM, several RCTs have been performed to demonstrate the safety of empagliflozin for other outcomes (Zinman et al., 2015; Packer et al., 2020). New indications such as reducing the risk of cardiovascular death in adults with T2DM and established cardiovascular disease and hospitalization for heart failure in adults with heart failure were added under FDA approval (Dailymed-Prescribing-information, 2022). However, the Ministry of Food and Drug Safety (MFDS) in Korea has not yet recognized the safety of empagliflozin for cardiovascular disease. This is because sufficient evidence has not been provided for whether the indication, “reducing the risk of cardiovascular death” could be demonstrated for Koreans as well. For this reason, empagliflozin has not yet been approved for reducing the risk of cardiovascular disease (MFDS-Prescribing-information, 2021).

In this study, we aimed to investigate to what extent the safety of empagliflozin from the RWE study in Korea is different from the one in RCT by emulating the design of foreign RCT. The outcome covers various safety outcomes including cardiovascular safety. We applied a RCT emulation analysis process that would be acceptable for regulation (Franklin and Schneeweiss, 2017; Franklin et al., 2020). If there were any discrepancies between the RCT and RWE, we investigated the circumstances under which this inconsistency occurs.

Methods

Study design and data sources

The study drug was selected through a pre-determined process (Supplementary Figure S1). Firstly, drugs that need to be re-evaluated under MFDS (date of announcement: 2021-01-24) were assessed (number of drugs: 498) (Supplementary Table S1). Secondly, according to the selection criteria set by the research team, 91 drugs were considered having high demand for safety evaluation. Of those, the anti-diabetic medications consisting largest number of drugs (number of drugs: 7) were selected (Supplementary Table S2). An additional selection process was carried out with considering each drugs’ adverse reaction profiles. Finally, empagliflozin and its pivotal study (EMPA-REG Outcome) were selected as a target drug and a target trial, respectively. This 1:1 matched cohort study included patients with type 2 diabetes mellitus and high cardiovascular risk, using the same inclusion/exclusion criteria, follow-up method and outcome definitions of a target RCT. The study assessed the effect of empagliflozin versus sitagliptin on cardiovascular and several safety outcomes of empagliflozin. The EMPA-REG OUTCOME trial (NCT01131676) (Zinman et al., 2015) was selected to target emulation (Franklin et al., 2020; Franklin et al., 2021). The EMPA-REG OUTCOME study provided strong evidence that the SGLT2 inhibitor empagliflozin protects against major adverse cardiovascular events (MACEs) and other secondary outcomes (Zinman et al., 2015).

The analyzed health insurance data was officially provided by the Korean Health Insurance Review & Assessment Service (HIRA) (Kim et al., 2017). The insurance data included demographic, diagnosis, procedure, and prescription data of...
patients. The requirement for written informed consent from participants was waived because all participants were anonymized using a randomized identification number. This study was approved by the institutional review board of Seoul National University (IRB No. E2101/001-003). This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline (von Elm et al., 2014).

Study patients

The target population is patients with T2DM and established cardiovascular disease. Patients who had been diagnosed with T2DM were included from 2011 to 2020, with a 3 years of study index period between May 2016 and May 2018. The period between January 2011 to May 2016 was used as a screening period for applying inclusion/exclusion criteria. Patients were selected according to the same inclusion/exclusion criteria as a RCT (Supplementary Table S3). All patients (≥18 years) had established cardiovascular disease and received empagliflozin or sitagliptin for the first time. Note that according to 2013 American College of Cardiology and American Heart Association guideline, patients who have been diagnosed with an established cardiovascular disease are classified as a high-risk group (Karmali et al., 2014). Therefore, included patients were considered as having high cardiovascular disease risks. We selected an active comparator (sitagliptin) as a proxy for the placebo, because it is well known for observational studies, that a non-user comparator group can differ substantially from actively treated patients, unlike RCTs (Food-and-Drug-Administration-FDA, 2013). Many other studies have also selected Dipeptidyl peptidase-4 inhibitors as comparators for assessment of SGLT-2 safety (Kim et al., 2018; Douros et al., 2020; Lee et al., 2020; Seong et al., 2020; Han et al., 2021). The index date was defined as the very first date each drug was prescribed.

Key variables

Individuals were followed-up until May 2020, and outcomes were recorded between each individual’s index date and May 2020. MACEs outcome from the EMPA-REG OUTCOME trial was used as a primary outcome. Since HIRA does not provide cause of death information, modified MACEs (all-cause death, myocardial infarction (MI), and stroke) was applied (Yeom et al., 2015). A total of seven cardiovascular outcomes were analyzed: all-cause death, MI, hospitalization for unstable angina, coronary revascularization procedure, stroke, transient ischemic attack, and hospitalization for heart failure.

Eight safety outcomes were also analyzed: hypoglycemic events, urinary tract infections (UTIs), genital infections, volume depletion, acute kidney injury (AKI), diabetic ketoacidosis (DKA), thromboembolic events, and bone fracture. The operational definitions of outcomes were defined using the Korean Standard Classification of Diseases-9 codes or procedure codes and were directly matched to each Regulatory Activities Preferred Term (MedDRA PT) in the RCT (Supplementary Table S4). To minimize confounding variables (e.g., selection bias) as much as possible, 72 covariates were included viz. Demographics, comorbidities, and disease/outcome specific variables. Of those, the main variables included are as follows: Seven types of glucose-lowering therapies (Metformin, Insulins, Sulfonylureas, Glitazones, Glucagon-like peptide-1 agonists, Alpha-glucosidase inhibitors, and Meglitinides) [Diabetes treatment strategies], time since type 2 diabetes mellitus [Duration of continuous enrolment], number of inpatient/outpatient visit [Indicators of health care utilization of the patients], five types of cardiovascular risk factors (Coronary artery disease (CAD), Multi vessel CAD, MI, Coronary Artery Bypass Graft, and Stroke with proper cardiovascular procedures) [history of cardiovascular procedures]. All covariates within the preceding 1 year of index date were evaluated.

Statistical analysis

Statistical analyses were performed for the intention-to-treat population. Each time an outcome was analyzed, a new cohort was constructed after excluding patients with a history of the corresponding outcome. Patients were followed up until the earliest of events, the date of last follow-up, the date of switching diabetic medication to the other comparison group, or the end of the study period. The maximum follow-up period was set at 48 months (same as in the RCT). Empagliflozin users were matched 1:1 to sitagliptin users and the distribution of the propensity score was inspected (Parsons, 2001). A standardized difference >0.1 was regarded as a sign of imbalance (NCSS-statistical-Software, 2017). As same with RCT, the age and sex-adjusted multivariate Cox proportional hazard regression was used to estimate the hazard ratio (HR) of empagliflozin for the cardiovascular outcome, with a 95% confidence interval (CI). For a safety outcome model, logistic regression was used to the odds ratio (OR) of empagliflozin.

Sensitivity analyses were performed the same as with the RCTs in two ways. First, patients who received at least one dose of the study drug were observed until ≤30 days after a patient’s last intake of medication. Additionally, we followed up patients who received the study drug for ≥30 days (cumulative) including events that only occurred ≤30 days after a patient’s last intake of medication (“as-treated” analysis). Analyses were performed with SAS Enterprise Guide version 7.1 (SAS Institute Inc., Cary, NC, United States).
RCT-RWE agreement assessment

We defined three metrics below to make a binary decision on whether an RCT was successfully emulated, considering statistical significance, directionality, and CIs associated with the corresponding RWE study. The agreement criteria suggested by Franklin et al. were used for determining each agreement (Franklin et al., 2020). First, we defined regulatory agreement (RA) as the ability of the RWE study to emulate the direction and statistical significance of the randomized trial finding. A secondary agreement metric was the estimate agreement (EA), defined as an RWE estimate that lies within the RCT 95% CI. We also conducted hypothesis tests to evaluate whether there was a difference in findings by calculating the standardized difference (SD) between the RCT and RWE effect estimates. We considered a p-value < 0.05 (where SD is greater than 1.96) statistically significant for the SD agreement. For comparison of results, HRs for cardiovascular outcomes and ORs for safety outcomes were calculated and compared (HRs were not provided for safety outcomes in an RCT). We defined the emulation result as “good” or “moderate” if all three agreements or two of three agreements were achieved, respectively. If the emulation result achieved ≤ one of the agreements, we defined the result as ‘fail’.

Results

A total of 932,465 patients (age ≥18 years) diagnosed with diabetes who received empagliflozin or sitagliptin were identified. New empagliflozin or sitagliptin users (n = 384,579) were selected (Figure 1). Among 98,733 patients who have high cardiovascular disease risks, an eligible study cohort with 48,545 patients remained after excluding patients who do not meet predefined inclusion criteria. Sitagliptin users were older and visited clinics more frequently (inpatient/outpatient) than empagliflozin users (Table 1). A later index date of empagliflozin users was observed compared to sitagliptin users. Compared to sitagliptin users, empagliflozin users were more often diagnosed with coronary artery disease (including coronary revascularization) and had fewer strokes.

After 11,563 empagliflozin users were matched to sitagliptin users, the above differences (age, number of clinic visits, index date, cardiovascular risk factors, comediations, and comorbidities) were reduced, and both groups were well balanced. Standardized differences were well below 0.1 for all 72 covariates. Median length of follow-up (2.7 years; median duration of anti-diabetic medications prescription during follow-up [1.7 (interquartile range 0.5–2.4) years]; and mean age of patients [55.6 years; men: 58.9% (n = 13,628)] were shown. In the other nine study cohorts for evaluating safety outcomes, the two drug user groups were also well balanced after 1:1 matching (Supplementary Tables S5–S13).
TABLE 1 Baseline characteristics.

| Variables                      | Pre-Match | STD | Post-Match | STD |
|--------------------------------|-----------|-----|------------|-----|
| Sex                            | 20,289 (55) | -0.04 | 6,799 (58.8) | 0.004 |
| Age                            | 60.4 ± 11.4 | -0.04 | 55.5 ± 11.3 | 0.008 |
| Insurance type                 |           |     |            |     |
| Normal                         | 34,434 (93.4) | 0.06 | 10,927 (94.5) | 0.01 |
| Medicaid                       | 11,065 (94.7) |       | 36 (4.3) | 0.036 |
| No change                      | 36 (3.8) |       | 36 (3.8) | 0.036 |
| Number of Inpatient visit      | 0.8 ± 2.2 | -0.2 | 0.5 ± 1.1 | 0.01 |
| Number of outpatient visit     | 28.3 ± 27.5 | -0.1 | 25.3 ± 23.1 | 0.01 |
| Time since type 2 diabetes mellitus ≤ 1 year | 18,197 (49.4) | 0.1 | 5,043 (43.6) | 0.007 |
| >1–5 years                     | 16,927 (45.9) | 0.06 | 5,941 (50.9) | 0.04 |
| >5 years                       | 1737 (4.7) |       | 693 (6) | 0.01 |
| Charlson comorbidity index     |           |     |            |     |
| 0                              | 1,521 (4.1) | 0.03 | 469 (4.1) | 0.004 |
| 1                              | 3,361 (9.1) |       | 1,150 (10) | 0.003 |
| 2                              | 4,957 (13.4) | 0.08 | 1,634 (14.1) | 0.008 |
| 3                              | 17,337 (47.4) | -0.03 | 8,423 (72.1) | -0.01 |
| CV risk factor                 |           |     |            |     |
| CAD                            | 32,597 (88.4) | 0.1 | 10,657 (92.2) | 0.01 |
| Multi vessel CAD               | 16,230 (44.0) | 0.2 | 5,944 (72.1) | 0.003 |
| MI                             | 1920 (5.2) |       | 877 (9.1) | 0.003 |
| CABG                           | 7,065 (20.0) | 0.2 | 2,966 (28.2) | 0.01 |
| Stroke                         | 5,209 (14.4) | -0.2 | 1,030 (14.1) | 0.003 |
| PAD                            | 2,716 (7.3) | -0.03 | 637 (9.9) | 0.003 |
| DM circulation                 | 4,921 (13.4) | 0.06 | 1,701 (15.2) | 0.003 |
| DM foot                        | 3 (0.0) |       | 3 (0.0) | 0.003 |
| DM nephropathy                 | 2,165 (6.4) | 0.08 | 965 (8.4) | 0.06 |
| DM neuropathy                  | 5,274 (14.3) | -0.02 | 1,620 (14) | 0.01 |
| DM other Complications         | 27,266 (74.1) | -0.06 | 8,288 (71.7) | 0.003 |
| Hyperglycemia                  | 694 (1.9) | -0.04 | 160 (1.4) | 0.003 |
| Comorbidities                  |           |     |            |     |
| Hypertension                   | 28,072 (78.3) | 0.01 | 9,095 (78.7) | 0.003 |
| Edema                          | 3,490 (9.5) | -0.01 | 1,065 (9.2) | -0.003 |
| Kidney stone                   | 580 (1.6) |       | 189 (1.6) | 0.003 |
| Osteoarthritis                 | 13,169 (35.7) | -0.1 | 3,561 (30.8) | 0.001 |
| Other arthritis                | 9,104 (24.7) | -0.06 | 2,484 (21.3) | 0.02 |
| PUD                            | 9,380 (25.5) | -0.03 | 2,735 (25.6) | 0.01 |
| Pancreatitis                   | 342 (0.9) |       | 105 (0.9) | 0.003 |
| UC                             | 59 (0.2) |       | 12 (0.1) | 0.003 |
| CVA                            | 15 (0.0) |       | 5 (0.0) | 0.003 |
| Asthma                         | 5,421 (14.7) | -0.02 | 1,596 (13.8) | 0.002 |
| COPD                           | 1,349 (3.7) | -0.07 | 300 (2.6) | 0.003 |
| Bladder stone                  | 29 (0.1) |       | 5 (0.0) | 0.003 |
| Dementia                       | 5,993 (16.3) | -0.2 | 1,141 (9.9) | 0.003 |
| Electrolyte Imbalance          | 2,153 (6.4) | -0.05 | 585 (5.1) | 0.006 |
| Glaucoma/Cataract              | 10,509 (28.5) | -0.03 | 3,127 (27) | 0.005 |

(Continued on following page)
Comparison of baseline characteristics between RCT and RWE

A lower proportion of men and a lower mean age were observed in our RWE cohort than in the corresponding RCT. (Table 2). Compared to the RCT, the RWE cohort was more often diagnosed with coronary artery disease (including coronary revascularization) and had fewer MIs, strokes, and peripheral artery disease. Rates of patients receiving glucose-lowering therapies were generally similar between the RCT and the RWE, except for the use of insulin. However, the proportions of patients who have been more than 5 years since their diagnosis of T2DM were 82.0 and 6.0% in RCT and RWE, respectively (p-value < 0.001).
RCT-RWE agreement on cardiovascular outcomes

From the results of RWE, empagliflozin was associated with a significantly decreased risk of MACEs (HR 0.87, 95% CI 0.79–0.96), all-cause mortality (HR 0.78, 95% CI 0.67–0.91), and heart failure (HR 0.85, 95% CI 0.75–0.95) comparing to sitagliptin (Table 3). MI, stroke, hospitalization for unstable angina, coronary revascularization, and transient ischemic attack were not significantly associated with empagliflozin use. As mentioned above, empagliflozin was related to a significantly decreased risk of MACEs [EMPA-REG DUPLICATE RWE: adjusted HR 0.87, 95% confidence interval (CI) 0.79–0.96]. The predefined estimate agreement, regulatory agreement, and standardized difference for RCT duplication were achieved (Figure 2) [EMPA-REG OUTCOME RCT: adjusted HR 0.86, 95% (CI) 0.74–0.99]. All of the eight cardiovascular outcomes except stroke achieved three agreements (RA/EA/SD) (point estimate HR in RCT and RWE = 0.86:0.87 [MACEs], 0.68:0.78 [all-cause death], 0.87:0.91 [MI], 0.99:0.94 [hospitalization for unstable angina], 0.86:0.94 [coronary revascularization], 0.85:0.88 [transient ischemic attack], and 0.65:0.85 [hospitalization for heart failure]). For stroke, the HR estimate of RWE 0.89 was in the opposite direction to that of RCT (disagreement of RA [point estimate HR of RCT: 1.18]), and two of three agreements (EA/SD) were achieved.

RCT-RWE agreement on safety outcomes

For safety outcomes from RWE, empagliflozin was associated with lowered risk of hypoglycemia (OR 0.70, 95% CI 0.59–0.84), UTI (OR 0.87, 95% CI 0.81–0.94), AKI (OR 0.70, 95% CI 0.59–0.82), and volume depletion (OR 0.84, 95% CI 0.76–0.94) comparing to sitagliptin. Alternatively, the risk of genital infections significantly increased (OR 1.49, 95% CI 1.35–1.65) compared to sitagliptin. No significant associations were identified in DKA, thromboembolic event, and fracture (Table 4).
In regulatory agreement, empagliflozin showed significantly lowered risk in RWE, whereas the RCT reported a non-significant effect on the hypoglycemic adverse event, UTI, and volume depletion. An estimate agreement was achieved for 6 of the 8 emulations, with the exception of a hypoglycemic adverse event (OR: 0.70) and genital infections (OR: 1.49) where the emulation estimates were below the lower 95% CI bound from the RCT (OR: 1.00; 95% CI: 0.89–1.11 and OR: 3.74; 95% CI: 2.70–5.19 for hypoglycemic adverse event and genital infection, respectively). Statistically significant disagreements in SDs were shown (SD: −3.3 and −5.3 for hypoglycemic adverse event and genital infections, respectively).

### Sensitivity analyses

After follow-up of patients who received at least one dose of study drugs until ≤30 days after the last intake of medication, similar results (HR for MACEs: 0.88; 95% CI: 0.77–0.99) were obtained (Supplementary Table S14). Additional sensitivity analysis (including patients who received study drugs for ≥30 days including only events that occurred ≤30 days after a patient’s last intake of medications) did not produce meaningful changes in the study findings (HR for MACEs: 0.87; 95% CI: 0.79–0.96) (Supplementary Table S15). All three agreements remained ‘Y’ for MACEs in both sensitivity analyses. In the same sensitivity analyses for eight safety outcomes, at least two of the three agreements were achieved in six safety outcomes (UTI, AKI, volume depletion, DKA, thromboembolic event, and fracture) (Supplementary Tables S16, S17). The hypoglycemic adverse event and genital infections still failed to show sufficient agreements, as in the main analysis.

### Discussion

Our study analyzed patients with high cardiovascular disease risks that were prescribed empagliflozin or sitagliptin for...
emulation of a pre-existing RCT. The primary objective of the study was to evaluate to what extent the safety of empagliflozin from the RWE study in Korea is different from the one in RCT by emulating the design of foreign RCT. This study emulated the cardiovascular outcomes including other safety outcomes of the EMPA-REG OUTCOME RCT in Korea. According to pre-specified agreement standards, successful agreements were achieved in cardiovascular disease including MACEs. For all outcomes, 14 of the 16 RCT outcomes including safety outcomes were successfully reproduced (graded as “good” or “moderate”). Our study results suggested that RWE can emulate RCT results satisfactorily and have the potential for providing evidence for future regulatory decision-making when RCT evidence is not available in Korea.

As shown in other studies, one must always keep in mind that some discrepancies may occur due to differences in study samples, study designs, or statistical methods. To date, various RWE studies have reported on the safety of SGLT-2 inhibitors including empagliflozin. There were discrepancies between findings, for example, the beneficial effect of SGLT-inhibitors on MACEs has been reported (Persson et al., 2018; Filion et al., 2020; Dave et al., 2021). However, two other studies have reported non-significant results in MACEs (Norhammar et al., 2019; Jeon et al., 2021). In other safety outcomes, Lega et al. reported a decreased risk of UTIs (Lega et al., 2019), while another study reported an association with an increased risk of UTI (Han et al., 2021). SGLT2 inhibitor use was associated with an elevated DKA risk (Wang et al., 2019); however, this study was not in Korea (Kim et al., 2018). We found both adverse (Ueda et al., 2018) and beneficial (Toulis et al., 2018) effects on fracture, although most results were non-significant. Most studies have reported decreased risks of SGLT2 inhibitors on AKI or impairment in renal function (Nadkarni et al., 2017; Cahn et al., 2019; Heerspink et al., 2020; Koh et al., 2021). Therefore, our study focused on emulating an existing RCT design and thereby confirming that the same results can be obtained from RWE. We have demonstrated SGLT-2 inhibitors’ associations with decreased cardiovascular outcomes including reducing MACEs and heart failure. Our results were consistent with the results of the target trial, and other studies including RCTs [MACEs (Mascolo et al., 2021) and heart failure (Kramer et al., 2010; Mascolo et al., 2021; Requena-Ibanez et al., 2021; Santos-Gallego et al., 2021; Ferreira et al., 2022; Neuen et al., 2022; Requena-Ibanez et al., 2022; Sauer, 2022)] which show that SGLT could induce reverse cardiac remodeling and improving quality of life, and also reduce myocardial fibrosis.

However, despite the substantial effort, there were disagreements between the RCT and RWE in several outcomes. Stroke is a well-known disease that can be captured with a high accuracy because of its seriousness. The incidence rates were similar between RCT and RWE results. However, our study result suggested that empagliflozin was associated with a decreased risk of stroke (although not significant) unlike its non-significant increase in the RCT. Several meta-analyses including all trials do show reductions

![FIGURE 2](https://example.com/figure2.png)

**FIGURE 2**  
RCT-RWE agreements plots. EA, estimate agreement; MACEs, major adverse cardiovascular events; RA, regulatory agreement; RCT, randomized clinical trial; RWE, real-world evidence; SD, standardized difference.
in hemorrhagic stroke (Tsai et al., 2021) and in total stroke (Mascolo et al., 2021), which supports our results. Also, SGLT-2 inhibitors seem to reduce atrial fibrillation (Pandey et al., 2021), which can also explain the stroke protection. It seems reason for the discrepancy is not clear. Ethnic factors may have been involved because over 70% of patients were Caucasian, and only 20% were Asian in the RCT (Zinman et al., 2015). Asians are reported to have a lower risk of cardiovascular disease than other races (Jung et al., 2015). As this study was conducted on Koreans, the proportion of patients with a history of severe diseases such as MI or stroke was small at baseline, and the age and severity of diabetes (time since onset of T2DM) were also lower than those of the RCT. In the subgroup analysis reported by the RCT, empagliflozin was reported to have a HR of 0.88 and 1.48 for Caucasians and Blacks for MACEs respectively and 0.68 for Asians (Zinman et al., 2015). Another study showed the protective effect of the SGLT-2 inhibitors against stroke in Koreans (Han et al., 2021); therefore, racial factors may have influenced our findings.

Another hypothesis includes a possibility of physicians’ reluctance to prescribe empagliflozin because of its known side effects. It has been reported that cardiologists may be reluctant to prescribe SGLT2 inhibitors due to concerns of adverse effects (Vardeny and Vaduganathan, 2019). Owing to incomplete knowledge of its benefits and/or risks (Das et al., 2018), concerns with SGLT2 inhibitors have led to decreased use in clinical practice (Vaduganathan et al., 2018). The drug approval date of empagliflozin was May 2016 in Korea, and physicians may have paid attention to prescribing in the early stage of approving the index period (2016–2018) of this study. Typically, patients tend not to use drugs when they are not in good health (Glynn et al., 2001) and this phenomenon can be observed in a study that reported excessively large protective effects on cardiovascular disease by using statins (Glynn et al., 2006). In the case of a new drug, this point should be taken into account because physicians often intend to prescribe the medication to a person who is expected to be relatively healthy and has a good prognosis. This trend is expected to be more prominent in outcomes such as stroke and genital infection in which the point estimate was reported as one or higher in RCTs. The HR point estimate of such an outcome in RWE is either reversed or

### TABLE 4 RCT-RWE agreement for each safety outcome.

| Outcomes                        | EMPA-REG Outcome * (RCT) | EMPA-REG Duplicate (RWE) | STD | Agreement |
|---------------------------------|---------------------------|---------------------------|-----|-----------|
|                                 | Rate (%)                  | OR (95%CI)                | Rate (%) | OR (95%CI) | RA | EA | SD |
| Hypoglycemic adverse event      |                           |                           |     |           |
| Sitagliptin                     | 27.9                      | 1.00 (0.89–1.11)          | 2.6 | 0.70 (0.59–0.84) | -3.3 | N | N | N |
| Empagliflozin                   | 27.8                      | 1.9                       |     |           |
| Urinary tract infection         |                           |                           |     |           |
| Sitagliptin                     | 18.1                      | 0.99 (0.87–1.13)          | 23.3 | 0.87 (0.81–0.94) | -1.7 | N | Y | Y |
| Empagliflozin                   | 18.0                      | 20.9                      |     |           |
| Genital infection               |                           |                           |     |           |
| Sitagliptin                     | 1.8                       | 3.74 (2.70–5.19)          | 7.9 | 1.49 (1.35–1.65) | -5.3 | Y | N | N |
| Empagliflozin                   | 6.4                       | 11.4                      |     |           |
| Acute kidney injury             |                           |                           |     |           |
| Sitagliptin                     | 6.6                       | 0.78 (0.63–0.96)          | 3.3 | 0.70 (0.59–0.82) | -0.8 | Y | Y | Y |
| Empagliflozin                   | 5.2                       | 2.3                       |     |           |
| Volume depletion                |                           |                           |     |           |
| Sitagliptin                     | 4.9                       | 1.04 (0.82–1.30)          | 7.1 | 0.84 (0.76–0.94) | -1.7 | N | Y | Y |
| Empagliflozin                   | 5.1                       | 6.1                       |     |           |
| Diabetic ketoacidosis           |                           |                           |     |           |
| Sitagliptin                     | 0.04                      | 1.99 (0.2–17.8)           | 0.38 | 1.09 (0.72–1.64) | -0.5 | Y | Y | Y |
| Empagliflozin                   | 0.1                       | 0.42                      |     |           |
| Thromboembolic event            |                           |                           |     |           |
| Sitagliptin                     | 0.9                       | 0.75 (0.42–1.31)          | 4.3 | 0.92 (0.80–1.05) | 0.7 | Y | Y | Y |
| Empagliflozin                   | 0.6                       | 3.9                       |     |           |
| Fracture                        |                           |                           |     |           |
| Sitagliptin                     | 3.9                       | 0.98 (0.76–1.27)          | 13.8 | 0.94 (0.87–1.03) | -0.3 | Y | Y | Y |
| Empagliflozin                   | 3.8                       | 13.1                      |     |           |

EA, estimate agreement; HR, hazard ratio; CI, confidence interval; MACEs, major adverse cardiovascular events; RA, regulatory agreement; RCT, randomized clinical trial; RWE, real-world evidence; STD, standardized difference; SD, standardized difference; Y, yes; N, no.
much lower than the value reported in the RCT. Stroke and genital infection showed HRs and ORs of 1.18 and 3.74 in the RCT, and 0.89 and 1.49 in our RWE study, respectively. Therefore, there is a possibility that undetected selection bias exists in our study.

In the hypoglycemic event, there was a >10-fold difference between the incidence in a RCT and that in RWE. The hypoglycemic event was less likely to be captured in real-world claim data, as shown in the event rates. Kim et al. reported that there is a possibility of underestimating the frequency of the hypoglycemic events when using HIRA data (Kim et al., 2016). Other studies share similar problems, showing the accuracy of diagnosis could be low owing to the nature of claims data because hypoglycemic events that can be self-treated do not need any medical management (Task Force Team for Basic Statistical Study of Korean Diabetes Mellitus of Korean Diabetes Association et al., 2013; Park et al., 2018). It appears that physicians in Korea consider hypoglycemic events to be temporary and do not often record a diagnostic code. Similarly, two observational studies in Korea showed low event rates of hypoglycemia (6.3%, self-reported outcome) (Hong et al., 2019), and 2.4 per 100 person-year (insurance claim data) (Han et al., 2021). The discrepancy in event rates could have led to the disagreement in treatment effect estimates. The event rate appears to be an important factor when conducting the RCT emulation study.

The intention-to-treat approach was applied in our study, and the median duration of observation time was 2.7 and 3.1 years in RWE and RCT studies, respectively. Adherence to medications in the RWE is often poor compared with the RCT (Freemantle et al., 2013), and the median duration of treatment was 1.7 (RWE) and 2.6 (RCT) years in this study. In sensitivity analysis, as-treatment analyses were performed to test whether our main outcome was affected by adherence. Similar results were obtained, and shorter duration of use for empagliflozin provided a benefit on several outcomes.

There are several limitations in our study. We tried to emulate as much of an RCT as possible, including inclusion and exclusion criteria, exposures, and results; however, because of the limitations of the healthcare database, accurate emulation was not possible. Our study is a retrospective cohort design and not all information is included in the HIRA data (e.g., lab results for blood glucose test, urine culture test, or body weight). Therefore, although we adjusted for all possible confounders, there still may be residual confounding factors present. There were regulatory disagreements in UTIs and volume depletion outcomes, indicating potential for residual confounding factors related to these outcomes. Additionally, note that unlike RCT, RWE cannot provide the exact cause and effect, and it could only show a significant association. The ultimate goal of our study was to utilize relevant RWE for regulatory decisions when no RCT evidence is available. The results of RCT and RWE are not always consistent. As mentioned above, event rates for testing specificity of outcome definition should be addressed. In addition, consideration of characteristics such as study participants, real-world clinical settings, and data availability might be important for enhancing the validity of study.

Our study results suggest that RWE emulating foreign RCT has the potential for providing evidence for future regulatory decision-making in an environment where RCT could not be performed. Further research is needed to determine whether RWE findings can be reliable evidence in various clinical settings or specific patient groups.

Data availability statement

The datasets presented in this article are not readily available because Viewing of data that shows all the records of a patient are difficult to share owing to the policy of the Korean National Health Insurance Service. It can only be viewed in anonymized form when analyzed. Therefore, if there is a request for original data, the statistical data obtained after the desired statistical processing on the server will be shared. Requests to access the datasets should be directed to National Health Insurance Service, nhiss. nhis.or.kr.

Ethics statement

The studies involving human participants were reviewed and approved by the Seoul National University (IRB No. E2101/001-003). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

HJ contributed to the conception and design of the study, data acquisition, analysis and interpretation of results, drafted, and revised the manuscript. JO and I-WK contributed to the conception and design of the study, analysis and interpretation of results, and revised the manuscript.

Funding

This research was supported by a grant (20173MFDS171) from the Ministry of Food and Drug Safety in 2020.

Acknowledgments

This study used HIRA research data (M20200708641) from the HIRA.
Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.928121/full#supplementary-material

References

Cahn, A., Melzer-Cohen, C., Pollack, R., Chodick, G., and Shalev, V. (2019). Acute renal outcomes with sodium-glucose co-transporter-2 inhibitors: Real-world data analysis. Diabetes Obes. Metab. 21 (2), 340–348. doi:10.1111/dob.13532

ClinicalTrials.gov (2022). Number of clinical trials conducted by region according to ClinicalTrials.gov [Online]. Available: https://clinicaltrials.gov/ct2/results?map=map-ES (Accessed 04 2022, 06).

Dailymed-Prescribing-information (2022). Prescribing information: Empagliflozin (JARDIANE) [Online]. Available: https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=19c8-39e2-02be-2322cb6357f8&type=display (Accessed 05 2021, 01).

Das, S. R., Eventt, B. M., Bırcher, K. K., Brown, J. M., Cefalu, W. T., Jaruzi, J. L., Jr., et al. (2018). 2018 ACC expert consensus decision pathway on novel therapies for cardiovascular risk reduction in patients with type 2 diabetes and atherosclerotic cardiovascular disease: A report of the American College of Cardiology task Force on expert consensus decision pathways. J. Am. Coll. Cardiol. 72 (24), 3200–3223. doi:10.1016/j.jacc.2018.09.020

Dave, C. V., Kim, S. C., Goldfine, A. B., Glynn, R. J., Tong, A., and Patorno, E. (2021). Risk of cardiovascular outcomes in patients with type 2 diabetes after addition of SGLT2 inhibitors versus sulfonylureas to baseline GLP-1RA therapy. Circulation 143 (8), 779–779. doi:10.1161/CIRCULATIONAHA.120.047965

Douroso, A., Lin, I. M., Frailik, M., Dell’Aniello, S., Shah, B. R., Ronkole, P. E., et al. (2020). Sodium-glucose co-transporter-2 inhibitors and the risk for diabetic ketoacidosis: A multicenter cohort study. Ann. Intern. Med. 173 (6), 417–425. doi:10.7326/M20-0289

Ferreira, J. P., Butler, J., Zannad, F., Filipappolou, G., Schueier, E., Steul, D., et al. (2022). Mineralocorticoid receptor antagonists and empagliflozin in patients with heart failure and preserved ejection fraction. J. Am. Coll. Cardiol. 79 (12), 1129–1137. doi:10.1016/j.jacc.2022.01.029

Filion, K. B., Lin, I. M., Yu, O. H., Dell’Aniello, S., Douros, A., Shah, B. R., et al. (2020). Sodium-glucose co-transporter 2 inhibitors and risk of major adverse cardiovascular events: Multi-database retrospective cohort study. BMJ 370, m3342. doi:10.1136/bmj.m3342

Food-and-Drug-Administration (FDA) (2013). Best practices for conducting and reporting pharmacoeconomicologic safety studies using electronic healthcare data. Silver Spring, Maryland: U.S. Food and Drug Administration.

Food-and-Drug-Administration (FDA) (2018a). Framework for FDA’s real-world evidence program. Silver Spring, Maryland: U.S. Food and Drug Administration.

Food-and-Drug-Administration (FDA) (2018b). Submitting documents utilizing real-world data and real-world evidence to FDA for drugs and biologies [Online]. Available: https://www.fda.gov/media/99447/download (Accessed 09 2021, 24).

Food-and-Drug-Administration (FDA) (2018b). Use of electronic health record data in clinical investigations guidance for industry [Online]. Available: https://www.fda.gov/media/99447/download (Accessed 09 2021, 24).

Franklin, J. M., and Schneeweiss, S. (2017). When and how can real world data analyses substitute for randomized controlled trials? Clin. Pharmacol. Ther. 102 (6), 924–933. doi:10.1002/cpt.857

Freemantele, N., Marston, L., Walters, K., Wood, J., Reynolds, M. R., and Petersen, J. (2013). Making inferences on treatment effects from real world data: Propensity scores, confounding by indication, and other perils for the unwary in observational research. BMJ 347, f6409. doi:10.1136/bmj.f6409

Garrison, L. P., Jr., Neumann, P. J., Erickson, P., Marshall, D., and Mullins, C. D. (2007). Using real-world data for coverage and payment decisions: The ISPOR real-world data task Force report. Value Health 10 (5), 326–335. doi:10.1111/j.1524-4733.2007.00386.x

Glynn, R. J., Knight, E. L., Levin, R., and Avorn, J. (2001). Paradoxical relations of drug treatment with mortality in older persons. Epidemiology 12 (6), 682–689. doi:10.1097/00001648-200111000-00017

Glynn, R. J., Schneeweiss, S., Wang, P. S., Levin, R., and Avorn, J. (2006). Selective prescribing led to overestimation of the benefits of lipid-lowering drugs. J. Clin. Epidemiol. 59 (8), 819–828. doi:10.1016/j.jclinepi.2005.12.012

Han, S. J., Ha, K. H., Lee, N., and Kim, D. J. (2021). Effectiveness and safety of sodium-glucose co-transporter-2 inhibitors compared with dipeptidyl peptidase-4 inhibitors in older adults with type 2 diabetes: A nationwide population-based study. Diabetes Obes. Metab. 23 (3), 682–691. doi:10.1111/dob.14261

Herspink, H. J. L., Karasik, A., Thurosson, M., Melner-Cohen, C., Chodick, G., Khant, K., et al. (2020). Kidney outcomes associated with use of SGLT2 inhibitors in real-world clinical practice (CVD-REAL 3): A multinational observational cohort study. Lancet. Diabetes Endocrinol. 8 (1), 27–35. doi:10.1016/S2213-8571(19)30084-5

Hong, A. R., Koo, B. K., Kim, S. W., Yi, K. H., and Moon, M. K. (2019). Efficacy and safety of sodium-glucose cotransporter-2 inhibitors in Korean patients with type 2 diabetes mellitus in real-world clinical practice. Diabetes Metab. J. 43 (5), 590–606. doi:10.4093/dmj.2018.0134

Jeon, J. Y., Ha, K. H., and Kim, D. J. (2021). Cardiovascular safety of sodium glucose cotransporter 2 inhibitors as add-on to metformin monotherapy in patients with type 2 diabetes mellitus. Diabetes Metab. J. 45 (4), 505–514. doi:10.4093/dmj.2020.0057

Jung, K. J., Jang, Y., Oh, D. J., Oh, B. H., Lee, S. H., Park, S. W., et al. (2015). The ACC/AHA 2013 pooled cohort equations compared to a Korean Risk Prediction Model for atherosclerotic cardiovascular disease.therosclerotic cardiovascular disease. Atherosclerosis 242 (3), 367–375. doi:10.1016/j.atherosclerosis.2015.07.033

Karmali, K. N., Goff, D. C., Jr., Ning, H., and Lloyd-Jones, D. M. (2014). A systematic examination of the 2013 ACC/AHA pooled cohort risk assessment tool for atherosclerotic cardiovascular disease. J. Am. Coll. Cardiol. 64 (10), 959–968. doi:10.1016/j.jacc.2014.06.1186

Kim, H. M., Seong, J. M., and Kim, J. (2016). Risk of hospitalization for hypoglycemia among older Korean people with diabetes mellitus: Interactions between treatment modalities and comorbidities. Med. Baltim. 95 (42), e5016. doi:10.1016/MED.0000000000000016

Kim, J. A., Yoon, S. Kim, L. Y., and Kim, D. S. (2017). Towards actualizing the value potential of Korea health insurance review and assessment (HIRA) data as a resource for health research: Strengths, limitations, applications, and strategies for optimal use of HIRA data. J. Korean Med. Sci. 32 (5), 718–728. doi:10.3346/jkms.2017.32.5.718

Kim, Y. G., Jeon, J. Y., Han, S. J., Kim, D. J., Lee, K. W., and Kim, H. J. (2018). Sodium-glucose co-transporter-2 inhibitors and the risk of ketoacidosis in patients with type 2 diabetes mellitus: A nationwide population-based cohort study. Diabetes Obes. Metab. 20 (8), 1852–1858. doi:10.1111/dob.13297

Jang et al. 10.3389/fphar.2022.928121
