**Name of Journal:** *World Journal of Diabetes*

**Manuscript No:** 41697

**Manuscript Type:** ORIGINAL ARTICLE

***Observational Study***

**Using real world data to assess cardiovascular outcomes of two antidiabetic treatment classes**

Stapff MP. Real world data and cardiovascular outcomes

Manfred Paul Stapff

**Manfred Paul Stapff,** CMO, TriNetX Inc., Cambridge, MA 02140, United States

**ORCID number:** Manfred Paul Stapff (0000-0002-6127-1109).

**Author contributions:** Stapff MP developed the scientific concept, literature search, study design, applied the data querying, result interpretation, scientific discussion, and prepared the manuscript.

**Institutional review board statement:** As a federated network TriNetX received a waiver from Western IRB since only aggregated counts, statistical summaries of de-identified information, but no protected health information is received, and no study specific activities are performed in retrospective analyses.

**Informed consent statement:** This was an observational study based on analyses of anonymized electronic medical records describing real world treatment. No intervention or any study specific activity was done. Therefore, no informed consent was necessary and would even have been not feasible considering the anonymized and retrospective character of the analysis.

**Conflict-of-interest statement:** The author is employee of TriNetX Inc., the data network and analytics platform used for this publication. TriNetX as a company was not involved in the design of the study; the collection, analysis, and interpretation of data; writing the report; or the decision to submit the report for publication. The author does not declare conflicting interests (including but not limited to commercial, personal, political, intellectual, or religious interests).

**STROBE Statement:** The author has read the STROBE Statement-checklist of items, and the manuscript was prepared and revised according to the STROBE Statement-checklist of items.

**Open-Access:** This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

**Manuscript source:** Unsolicited manuscript

**Correspondence to: Manfred Paul Stapff, MD, PhD, Chief Medical Officer,** CMO, TriNetX Inc., 125 Cambridgepark Drive, Ste 500, Cambridge, MA 02140, United States. [Manfred.stapff@trinetx.com](mailto:Manfred.stapff@trinetx.com)

**Telephone:** +1-857-2856043

**Received:** August 22, 2018

**Peer-review started:** August 22, 2018

**First decision:** October 4, 2018

**Revised:** October 10, 2018

**Accepted:** November 15, 2018

**Article in press:**

**Published online:**

**Abstract**

***AIM***

To evaluate the effect on cardiovascular outcomes of sodium-glucose co-transporter-2 (SGLT2) inhibitors in a real world setting by analyzing electronic medical records.

***METHODS***

We used TriNetX, a global federated research network providing statistics on electronic health records (EHR) from approximately 38 Million patients in 35 Health Care Organizations predominately in the USA. The records of 46,909 patients who had taken SGLT2 inhibitors were compared to 189,120 patients with dipeptidyl peptidase (DPP) 4 inhibitors. We identified five potential confounding factors and built respective strata: elderly, hypertension, chronic kidney disease (CKD), co-medication with insulin or with metformin. Cardiovascular events were counted as stroke (ICD10 code: I63) or myocardial infarction (ICD10: I21) occurring within three years after the first instance of the respective medication in the patients’ records.

***RESULTS***

Of the 46909 patients with SGLT2 inhibitors in their EHR, 1667 patients (3.6%) had an ICD code for stroke or for myocardial infarction within the first three years after the first instance of the medication. In the control group there were 10680 events of 189120 patients (5.6%), which represents a risk ratio of 0.63 (95%CI: 0.60–0.66). The overall incidence of stroke or myocardial infarction in the strata with a potential confounding risk factor reached from 4.9% in patients taking metformin to 12.5% in the stratum with the highest risk (concomitant CKD). In all strata the difference in risk of experiencing a cardiovascular event was similarly in favor of SGLT2 versus control, with Risk Ratio ranging from 0.62 to 0.81.

***CONCLUSION***

Real world data replicated the results from randomized clinical trials, confirmed the cardiovascular advantages of SGLT2 inhibitors, and showed the applicability to the US American population.

**Key words:** Cardiovascular events; Clinical trials; Diabetes; Dipeptidyl peptidase 4 inhibitors; Electronic medical records; Real world evidence; Sodium-glucose co-transporter-2 inhibitors

**© The Author(s) 2018.** Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** Cardiovascular advantages of sodium-glucose co-transporter-2 (SGLT2) inhibitors were shown in complex clinical trials or countries with large registries, but it was unclear whether these findings can be applied to routine medical practice in the US. This real world analysis from 46909 patients with SGLT2 inhibitors revealed a 0.63 (95%CI: 0.60–0.66) risk ratio of SGLT2 inhibitors compared to 189120 patients with dipeptidyl peptidase 4 inhibitors. This analysis of electronic health records could replicate the results of randomized clinical trials, which supports the usefulness of such real world studies, *e.g.,* for long-term outcome or for safety observations.

Stapff MP. Using real world data to assess cardiovascular outcomes of two antidiabetic treatment classes. *World J Diabetes* 2018; In press

**INTRODUCTION**

An estimated 30.3 million people of all ages - or 9.4% of the United States population - had diabetes in 2015[[[1]](#endnote-1)]. It is expected that the world prevalence of diabetes among adults will increase to 7.7% and 439 million adults by 2030. Between 2010 and 2030, there will be a 69% increase in numbers of adults with diabetes in developing countries and a 20% increase in developed countries[[[2]](#endnote-2)].

While short-term treatment targets focus on normalization of values for glucose and hemoglobin A1c, the long-term objective is to avoid late stage complications of diabetes and end-organ damage. Up to 70% of patients with diabetes type II (T2DM) also have arterial hypertension[1] and are thus exposed to an increased risk of experiencing a stroke or heart attack. It is therefore important that treatment paradigms for T2DM consider the long-term cardiovascular risk.

In 2015, the EMPA-REG OUTCOME trial found a significant mortality benefit of sodium-glucose co-transporter-2 (SGLT2) inhibitors versus placebo[[[3]](#endnote-3)]. Because the findings were unexpected and unprecedented and not linked to obvious mechanistic pathway, it was suggested that the results need to be replicated in future investigations[[[4]](#endnote-4)]. Recently, CVD-REAL Nordic, a multinational observational study, analyzed the cardiovascular mortality and morbidity in patients with T2DM following initiation SGLT2 inhibitors[[[5]](#endnote-5)]. CVD-REAL Nordic was an observational analysis of individual patient-level data from national registries in three Scandinavian countries showing that SGLT2 inhibitor use was associated with reduced cardiovascular disease and cardiovascular mortality.

The objective of the following analysis was to support or to contradict the results of EMPA-REG OUTCOME and CVD-REAL Nordic by using electronic medical records (EMR) from a predominately United States-based research network, thus evaluating the representativity of these results outside the experimental setting of a randomized clinical trial and beyond an European population, respectively.

**MATERIALS AND METHODS**

We used TriNetX, a global federated research network providing access to statistics on EMR (diagnoses, procedures, medications, laboratory values, genomic information) from approximately 38 million patients in 35 large Health Care Organizations predominately in the United States. As a federated network TriNetX received a waiver from Western IRB since only aggregated counts, statistical summaries of de-identified information, but no protected health information is received, and no study specific activities are performed in retrospective analyses. Details of the network have been described elsewhere[[[6]](#endnote-6),[[7]](#endnote-7),[[8]](#endnote-8)]. All analyses were done in the TriNetX “Analytics” network using the browser based real time analytics features. At the time of the analysis, in June 2018, we analyzed the EMR of 46909 patients in the network who had an instance of any SGLT2 inhibitor (empagliflozin or dapagliflozin or canagliflozin) any time within the past ten years in their electronic medical record. As a comparator group we chose patients who had taken dipeptidyl peptidase (DPP) 4 inhibitors (linagliptin or alogliptin or sitagliptin or saxagliptin) during the same time and found *n* = 189120 patients. Using a Bayesian statistical approach[9] on demographics and pre-existing (baseline) comorbidities of the two groups, we identified five potential confounding factors and built strata with the following criteria: age ≥ 60 years, presence of hypertension [International Classification of Diseases (ICD)10 code I10] presence of CKD (ICD10 code N18), co-medication with insulin, and co-medication with metformin. Analyzing strata separately allowed addressing potential bias in the federated data model without direct access to the individual data sets on the patient level.

Cardiovascular events have been counted by selecting any stroke (ICD10 code I63) or myocardial infarction (ICD10 code I21) occurring during a three-year observation period after the first instance of above mentioned medication in the patients’ records.

The risks of experiencing an event in each stratum have been calculated by dividing the number of patients with an event (numerator) by the total number of patients with the respective medication in each stratum (denominator). The risk ratios for SGLT2 inhibitors versus the comparator group were calculated by dividing the risk for each SGLT2 stratum by the risk in each corresponding DPP4 stratum.

**RESULTS**

Of the 46909 patients taking SGLT2 inhibitors 1667 patients (3.6%) had an ICD code for stroke or myocardial infarction during their three-year observation period, compared to 10680 of 189120 (5.6%) in the control group (Table 1). This translates into a risk ratio of 0.63 without any correction for potential bias (*P* < 0.001; 95%CI: 0.60-0.66).

SGLT2 inhibitors carry a contra-indication for renal insufficiency[9]. Indeed, the percentage of patients with CKD was only 4% in the SGLT2 group, compared to 8% in the control group. While the groups were similar in gender distribution (53% and 52% male, respectively) and low density lipoprotein as well as high density lipoprotein levels, the SGLT2 group was younger than the control group (mean age 59 *vs* 66) and had more patients with concomitant hypertension (45% *vs* 41%). There were also differences in the use of insulin (32% *vs* 19%) and metformin (52% *vs* 33%). To balance for these potential confounding factors strata have been built for age ≥ 60 years, CKD, hypertension, and antidiabetic co-medication (insulin and metformin). The overall incidence of stroke or myocardial infarction in each stratum reached from 4.9% in to 12.5%. In all strata the difference in risk of experiencing a cardiovascular event in the SGLT2 group *vs* control was similarly in favor of SGLT2, with risk ratios ranging from 0.62 (co-medication insulin) to 0.81 (patients with CKD) (Table 2).

**DISCUSSION**

Drug therapy of type II diabetes mellitus should both, bring glucose and hemoglobin A1c values into an acceptable and stable range, and reduce the likelihood of end organ damage or cardiovascular events.

Several studies and meta-analyses have suggested a positive effect on cardiovascular outcomes by the SGLT2 inhibitor class[9,10]. EMPA-REG OUTCOME and CANVAS were randomized placebo controlled prospective trials using empagliflozin[4] and canagliflozin[[[9]](#endnote-9)], respectively.

A recent observational cohort study observed protective effects of SGLT2 inhibitors compared to sulfonylureas by a database analysis[10]. Another study, CVD-REAL Nordic was the first large observational analysis in real world settings in three Scandinavian countries evaluating the cardiovascular benefits of this class, and it too showed that SGLT2 inhibitor use was associated with reduced cardiovascular disease and cardiovascular mortality compared with use of other glucose-lowering drugs[4]. Such real-world studies are less complicated and significantly less costly than traditional prospective randomized clinical outcomes trials. In addition, the reduced number of eligibility criteria ensures that the study results are representative and applicable to a much wider population. Recently, another study confirmed that real-world data analyses of patients receiving routine care provide findings similar to those found in a randomized clinical trial and even may support (supplemental) regulatory applications[[[10]](#endnote-10)]. Real world evidence can sometimes complement or even replace randomized controlled trials, but prejudices and reservations have so far limited their acceptance[11].

Therefore, the underlying data sources must be reliable, and the methods used have to be defined in advance to avoid “data dredging” based on the findings[11]. Furthermore, the data usually come from non-consented patients and therefore highest standards on data privacy must be ensured.

The present study has been undertaken to evaluate whether the results of the EMPA-REG OUTCOME and CVD-REAL Nordic studies can be replicated in a federated network of EMR and if they can be applied to a predominantly United States American population. As control we chose DPP4 inhibitors because they represent another homogeneous and relatively new non-metformin class. We found a significantly lower incidence of strokes or myocardial infarctions in the SGLT2 group within the three-year observation period than in the control group.

In a federated data network, individual data sets never leave the source, *i.e.*, the data warehouse of a healthcare organization. Instead, the analyses are done based on aggregated statistical counts. At the time of this analysis our platform limited the methods which could be applied to correct for potential confounding factors, for example pair matching or propensity score matching (PSM). PSM is a popular method of preprocessing data for causal inference, but it has been discussed controversially as it may accomplish the opposite of its intended goal, i.e. increasing imbalance or bias[11]. In addition, the censoring by PSM, excluding certain patients from the analysis, reduces the sample size and the representativity of a diverse patient population, thus re-introducing the criticism often applied to randomized clinical trials with their very restrictive eligibility criteria.

We chose therefore to build subgroups of the study population according to the presence of potentially confounding factors and to test these strata individually. SGLT2 inhibitors have a contraindication for renal insufficiency and are a relatively new class of antidiabetics with less long-term experience than comparator classes, for example metformin or the DPP4 inhibitors. One can therefore assume that the treatment decision by prescribing physicians may be driven by a patient’s renal function, the patient’s age and other potential risk factors. Indeed, we found in the SGLT2 group a lower mean age, similarly as in CVD-REAL Nordic before matching. Furthermore, SGLT2 group had fewer patients with CKD than the comparator group. While in prospective randomized clinical trials such factors usually get balanced by randomization, they must be corrected for when a retrospective analysis is done. We therefore created five strata based on age ≥ 60 years, hypertension, CKD, insulin therapy or metformin therapy and tested the event rates individually in each of these subgroups. The fact that the overall highest event rate was found in the higher risk stratum (patients with CKD) provides internal validation for the selection of the strata.

All strata showed very similar hazard ratios for cardiovascular events (according to our definition using ICD10 codes for myocardial infarction or stroke) which were consistently in favor of the SGLT2 inhibitor group, *i.e.*, between 0.62 and 0.81. This generally confirms the findings of the CVD-REAL Nordic study where the risk ratio for cardiovascular mortality and for major cardiovascular events was in a similar range, 0.53 and 0.78, respectively.

***Limitations***

Due to the nature of the design (retrospective, non-randomized) and data analysis (federated, aggregated strata) this study could be done very quickly, simplistic, and with minimal cost, but may have several limitations. Non-randomized comparisons bear the risk that patients’ disease state influence the treatment decision and thus introduce imbalances. We limited balancing for confounders to five major factors and did not further correct for residual potentially confounding factors like other co-morbidities, duration of diabetes, glucose or HBa1c values, concomitant medications or length of exposure to concomitant treatment. Our outcome criteria were simply the ICD10 codes for myocardial infarction or stroke, relying on correct coding at the source, without differentiation between morbidity and mortality. Despite one specific compound was numerically dominating in each group (SGLT2: canagliflozin 78%, DPP4: sitagliptin 69%) we consider the results as representative for a class but not robust enough for a comparison of two individual compounds.

Real world studies depend on the prescribing and documentation behavior of the data providing institutions. We used EMR in structured form rather than claims data. This has the advantage of quite complete medical information coming from the respective Health Care Organization, but data may lack if a patient visits another institution. This applies especially to medication and prescription refills. We defined an observation period of three years, but we could not validate whether the patients actually stayed with their medication for the whole period as we defined the treatment group just on the fact of one documentation of SGLT2 or DPP4 in their records. Insofar a difference in compliance or persistence between the groups could introduce a potential imbalance, but the approach would be similar to the intent-to treat principle which is applied in randomized clinical trials.

Furthermore, under documentation or differences in completeness of the medical records between comparator groups need to be taken into consideration as well. Looking for a potential documentation bias we found a similar data density in the SGLT2 cohort compared to control (Table 3).

Theoretically, one could assume that more events had been found in the control group just because this patient cohort was better documented. In real world studies consideration of different therapeutic settings and documentation completeness is important, e.g. when comparing oral vs injectable medication or inpatient vs outpatient procedures. However, SGLT2 inhibitors and DPP4 inhibitors are both taken orally and prescribed in similar settings. In addition, our data found overall about 20% more events in the DPP4 group, but the density of facts per patient in the documentation of this group was only 6% higher. Therefore, a documentation bias as an explanation for the difference in CV events in this study is very unlikely.

In conclusion, this study was conducted by analyzing EMR of approximately 38 million patients from 35 healthcare organizations, mainly from the United States. This real world clinical setting allows the analysis of data from patients with a much broader cardiovascular risk profile than the much selected population in randomized clinical trials. The federated structure of this network ensures the highest level of data privacy standards but poses some restrictions on the analytics possible, for example matching by propensity scores. Despite these limitations: (1) this analysis could replicate the results from much more complex and costly studies on the same topic, which validates our methods and the quality of data in the network; (2) our analysis shows that the cardiovascular advantages of SGLT2 inhibitors found in the Scandinavian CVD-REAL Nordic study can be applied to the United States American population.

**Article Highlights**

***Research background***

Therapy of diabetes mellitus intends to control blood glucose values, to prevent or delay diabetic complications such as chronic kidney disease or retinopathy, and to reduce the likelihood of cardiovascular events like myocardial infarction or stroke. Several randomized clinical trials and sophisticated European registries have suggested that sodium-glucose co-transporter-2 (SGLT2) inhibitors may have an advantage in preventing cardiovascular events.

***Research motivation***

Randomized clinical trials are conducted on highly selected patient populations and follow very artificial treatment protocols. This makes it sometimes questionable whether the results are representative and can be applied to routine medical practice.

***Research objectives***

To evaluate if the positive results of randomized clinical trials with SGLT2 inhibitors can be confirmed by real world data coming from actual medical routine practice in the United States.

***Research methods***

A federated research network was used allowing analyses of electronic medical records (EMR) from 38 Million patients in 35 large Health Care Organizations predominately in the United States. Cardiovascular events have been counted occurring during a three-year observation period after start of a therapy with an SGLT2 inhibitor and compared to a control group starting dipeptidyl peptidase 4 inhibitors. Comorbidity strata have been created to address potential confounders.

***Research results***

In the overall cohort and in all comorbidity strata the risk of experiencing a cardiovascular event was similarly in favor of SGLT2, with risk ratios ranging from 0.62 to 0.81.

***Research conclusions***

The analysis of data from patients with a much broader cardiovascular risk profile than the much selected population in randomized clinical trials could replicate the results of such trials. This validates the methods, the quality of data in the network, and allows extrapolation of the trial results to the general patient population.

***Research perspectives***

Sophisticated analyses of high quality EMR can complement costly, complex and lengthy randomized clinical trials, can assess their representativity for actual medical practice in real world, and may, in certain instances, even be able to replace them.

**REFERENCES**

1 **Centers for Disease Control and Prevention.** National Diabetes Statistics Report. 2017. Atlanta, GA: Centers for Disease Control and Prevention, U.S. Dept of Health and Human Services; 2017 Available from: URL: https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf

2 **Cinek O**, Kramna L, Mazankova K, Odeh R, Alassaf A, Ibekwe MU, Ahmadov G, Elmahi BME, Mekki H, Lebl J, Abdullah MA. The bacteriome at the onset of type 1 diabetes: A study from four geographically distant African and Asian countries. *Diabetes Res Clin Pract* 2018; **144**: 51-62 [PMID: 30121305 DOI: 10.1016/j.diabres.2009.10.007]

3 **Bhatt AS**. Digesting New Developments in Biosensors. *N Engl J Med* 2018; **379**: 686-688 [PMID: 30110595 DOI: 10.1056/NEJMoa1504720]

4 **Kaul S**. Is the Mortality Benefit With Empagliflozin in Type 2 Diabetes Mellitus Too Good To Be True? *Circulation* 2016; **134**: 94-96 [PMID: 27400894 DOI: 10.1161/CIRCULATIONAHA.116.022537]

5 **Birkeland KI**, Jørgensen ME, Carstensen B, Persson F, Gulseth HL, Thuresson M, Fenici P, Nathanson D, Nyström T, Eriksson JW, Bodegård J, Norhammar A. Cardiovascular mortality and morbidity in patients with type 2 diabetes following initiation of sodium-glucose co-transporter-2 inhibitors versus other glucose-lowering drugs (CVD-REAL Nordic): a multinational observational analysis. *Lancet Diabetes Endocrinol* 2017; **5**: 709-717 [PMID: 28781064 DOI: 10.1016/S2213-8587(17)30258-9]

6 **Stacey J,** Mehta M. Using EHR Data Extraction to Streamline the Clinical Trial Process. *Clinical Researcher* 2017; 4: 2-7 [DOI: 10.14524/CR-17-0004]

7 **Stapff M.** Use of Electronic Health Data in Clinical Development. Pharm. Ind. **79,** Nr. 2, 204–210. ECV Editio Cantor Verlag, Aulendorf, Germany (2017) Available from: https://www.trinetx.com/wp-content/uploads/2018/05/Use-of-Electronic-Health-Data-in-Clinical-Development.pdf

8 **Stapff M.** Use of Electronic Health Records for Development and Feasibility Testing of Clinical Trial Protocols. DIA 28th EuroMeeting, April 6-8, 2016, Hamburg, Germany Available from:

<https://www.diaglobal.org/productfiles/4124219/16101_pgm.pdf>

9 **Oliphant TE.** A Bayesian perspective on estimating mean, variance, and standard-deviation from data. *Faculty Publications* 2006; 278 Available from: URL: http://hdl.lib.byu.edu/1877/438

10 Prescribing information canagliflozin 07/2017, Available from: URL: https://www.invokana.com/prescribing-information.pdf

11 **Sonesson C**, Johansson PA, Johnsson E, Gause-Nilsson I. Cardiovascular effects of dapagliflozin in patients with type 2 diabetes and different risk categories: a meta-analysis. *Cardiovasc Diabetol* 2016; **15**: 37 [PMID: 26895767 DOI: 10.1186/s12933-016-0356-y]

12 **Wu JH**, Foote C, Blomster J, Toyama T, Perkovic V, Sundström J, Neal B. Effects of sodium-glucose cotransporter-2 inhibitors on cardiovascular events, death, and major safety outcomes in adults with type 2 diabetes: a systematic review and meta-analysis. *Lancet Diabetes Endocrinol* 2016; **4**: 411-419 [PMID: 27009625 DOI: 10.1016/S2213-8587(16)00052-8]

13 **Neal B**, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, Shaw W, Law G, Desai M, Matthews DR; CANVAS Program Collaborative Group. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *N Engl J Med* 2017; **377**: 644-657 [PMID: 28605608 DOI: 10.1056/NEJMoa1611925]

14 **Matsuda H**, Mullapudi ST, Yang YHC, Masaki H, Hesselson D, Stainier DYR. Whole-Organism Chemical Screening Identifies Modulators of Pancreatic β-Cell Function. *Diabetes* 2018; **67**: 2268-2279 [PMID: 30115653 DOI: 10.2337/db18-1493-P]

15 **Fralick M**, Kesselheim AS, Avorn J, Schneeweiss S. Use of Health Care Databases to Support Supplemental Indications of Approved Medications. *JAMA Intern Med* 2018; **178**: 55-63 [PMID: 29159410 DOI: 10.1001/jamainternmed.2017.3919]

16 **Franklin JM**, Schneeweiss S. When and How Can Real World Data Analyses Substitute for Randomized Controlled Trials? *Clin Pharmacol Ther* 2017; **102**: 924-933 [PMID: 28836267 DOI: 10.1002/cpt.857]

17 **Berger ML**, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, Madigan D, Makady A, Schneeweiss S, Tarricone R, Wang SV, Watkins J, Daniel Mullins C. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf* 2017; **26**: 1033-1039 [PMID: 28913966 DOI: 10.1002/pds.4297]

18 **King G,** Nielsen R. Why propensity scores should not be used for matching. 2016 Available from: http://gking.harvard.edu/files/gking/files/psnot.pdf. Accessed June 28, 2018

**P-Reviewer:** Senol MG, Avtanski D **S-Editor:** Dou Y **L-Editor: E-Editor:**

**Specialty type:** Endocrinology and metabolism

**Country of origin:** United States

**Peer-review report classification**

Grade A (Excellent): A

Grade B (Very good): 0

Grade C (Good): C

Grade D (Fair): 0

Grade E (Poor): 0

**Table 1 Patient characteristics and results before correcting for potential confounding factors**

|  |  |  |
| --- | --- | --- |
|  | **SGLT2** | **Control** |
| *n* | 46909 | 189120 |
| mean age | 59 | 66 |
| SD age | 11 | 13 |
| percent male | 53% | 52% |
|  |  |  |
| comorbidities: |  |  |
| hypertension (I10) | 45% | 41% |
| CKD (N18) | 4% | 8% |
|  |  |  |
| co-medication: |  |  |
| on insulin | 32% | 19% |
| on metformin | 52% | 33% |
|  |  |  |
| LDL cholesterol (mg/dL) | 91.6 | 93.1 |
| HDL cholesterol (mg/dL) | 43.6 | 43.2 |
|  |  |  |
| after index event: |  |  |
| total stroke (I63) or MI (I21) | 12347 (5.2 %) | |
| *n* in group | 1667 | 10680 |
| percent in group | 3.6% | 5.6% |
| RR SGLT2 *vs* control | 0.63 | |

SGLT2: Sodium-glucose co-transporter-2; RR: Risk ratio; SD: Standard deviation; CKD: Chronic kidney disease; LDL: Low density lipoprotein; HDL: High density lipoprotein.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Stratum 1** | | **Stratum 2** | | **Stratum 3** | | **Stratum 4** | | **Stratum 5** | |
|  | **< 60 yr** | | **hypertension** | | **CKD** | | **insulin** | | **metformin** | |
|  |  | |  | |  |  |  | |  | |
|  | SGLT2 | control | SGLT2 | control | SGLT2 | control | SGLT2 | control | SGLT2 | control |
| *n* | 23594 | 131219 | 27499 | 115703 | 3786 | 34388 | 24395 | 90978 | 37762 | 136569 |
| patients with stroke or MI | 9784 (6.3 %) | | 10827 (7.6 %) | | 4755 (12.5 %) | | 8976 (7.8 %) | | 8629 (4.9 %) | |
| *n* in group | 1077 | 8707 | 1452 | 9375 | 391 | 4364 | 1275 | 7701 | 1394 | 7235 |
| percent in group | 4.6% | 6.6% | 5.3% | 8.1% | 10.3% | 12.7% | 5.2% | 8.5% | 3.7% | 5.3% |
| RR SGLT2 *vs* control | 0.69 | | 0.65 | | 0.81 | | 0.62 | | 0.70 | |

**Table 2 Results in the subgroups (strata) in patients with potential confounding factors**

SGLT2: Sodium-glucose co-transporter-2; RR: Risk ratio; CKD: Chronic kidney disease.

**Table 3 Data density in the two comparator cohorts**

|  |  |  |
| --- | --- | --- |
|  | **SGLT2** | **control** |
| Total facts | 54852092 | 261813664 |
| Avg facts per patient | 1143 | 1325 |
| Avg diagnosis facts per patient | 231 | 262 |
| Eastern United States, patients (%) | 68 | 69 |
| Western United States, patients (%) | 32 | 31 |

SGLT2: Sodium-glucose co-transporter-2.

1. [↑](#endnote-ref-1)
2. [↑](#endnote-ref-2)
3. [↑](#endnote-ref-3)
4. [↑](#endnote-ref-4)
5. [↑](#endnote-ref-5)
6. [↑](#endnote-ref-6)
7. [↑](#endnote-ref-7)
8. [↑](#endnote-ref-8)
9. [↑](#endnote-ref-9)
10. [↑](#endnote-ref-10)