

Replicating cardiovascular outcome trials of medications used to treat type 2 diabetes using real-world data: A systematic review of observational studies

Wanning Wang BHSc^{1,2}, Wang-Choi Tang BSc^{1,2}, Michael Webster-Clark PhD^{2,3}, Oriana HY Yu MD MSc^{2,3,4}, Kristian B. Filion PhD^{1,2,3}

¹ Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

² Center for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, Quebec, Canada

³ Department of Medicine, McGill University, Montreal, Quebec, Canada

⁴ Division of Endocrinology and Metabolism, Jewish General Hospital/McGill University, Montreal, Quebec, Canada

Correspondence to:

Kristian B. Filion, PhD

Associate Professor and William Dawson Scholar

Departments of Medicine and of Epidemiology, Biostatistics and Occupational Health

McGill University

3755 Cote Ste-Catherine Road, Suite H410.1

Montreal, Quebec, Canada H3T 1E2

Phone: (514) 340-8222 x 28394

Fax: (514) 340-7564

Email: kristian.filion@mcgill.ca

Keywords: Cardiovascular Outcome Trials, Generalizability, Real-world data, Systematic review

Conflict of interest disclosure: None

Sources of Support: Dr. Filion is supported by a William Dawson Scholar award from McGill University. Wanning Wang is supported by a McGill Internal Studentship Award and a CANadian Consortium of Clinical Trial TRAINing platform (CANTRAIN) studentship; CANTRAIN is funded by the Canadian Institutes of Health Research.

Word count: 3295; **Abstract:** 239; **Tables:** 4; **Figures:** 1; **Supplementary tables:** 7; **Supplementary figures:** 5

Twitter: @KB_Filion

ABSTRACT

Background: Cardiovascular outcome trials (CVOTs) are mandated by the U.S. Food and Drug Administration to assess the cardiovascular safety of new antidiabetic medications before entering the market. However, CVOTs often involve highly selective populations and results may not generalize to real-world settings.

Methods: Our study aimed to synthesize observational studies to assess the generalizability of CVOTs to real-world settings. We systematically reviewed observational studies that emulated previous CVOTs for dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide 1 (GLP-1) receptor agonists, and sodium glucose cotransporter-2 (SGLT-2) inhibitors among patients with type 2 diabetes. We searched the MEDLINE, EMBASE and Cochrane databases for observational studies that focused on trial emulation or cross-sectional studies that reported the proportion of real-world patients eligible for completed CVOTs. Two independent reviewers screened articles, extracted data, and assessed study concordance with randomized controlled trial (RCT) results.

Results: Nineteen studies were included in our systematic review, including four cohort studies that emulated previous RCTs and 15 cross-sectional studies that evaluated trial eligibility. Results between RCTs and real-world data (RWD) were concordant for all drug classes in finding non-inferiority. The median eligibility percentage ranged from 13% to 31% for SGLT-2 inhibitor trials and 12% to 43% for GLP-1 receptor agonist trials.

Conclusions: These results suggest that, while RCTs and RWD are concordant in their estimates, the trials lack representativeness. More research is needed on the replication of CVOTs using RWD to understand how different replication methods may impact findings.

INTRODUCTION

Randomized controlled trials (RCTs) are the current gold standard for assessing drug efficacy and for evidence-based regulatory decision making. However, due to their strict inclusion and exclusion criteria, trial populations may differ substantially from the real-world population receiving the medication¹. In addition, there are situations where it may not be suitable to conduct RCTs such as rare diseases or in populations ineligible for RCTs (e.g., children, pregnant women, the elderly). Thus, real world data (RWD) play an important role in the generation of evidence to evaluate the effectiveness and safety of medical interventions. With the adoption of electronic health records, large amounts of RWD on prescription drug use and health outcomes are becoming readily available.

In 2008, the U.S Food and Drug Administration (FDA) issued recommendations for conducting cardiovascular outcome trials (CVOT) to ensure that antidiabetic medications have acceptable cardiovascular risk profiles². More than 13 CVOTs have been conducted on antidiabetic drugs, including dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide 1 (GLP-1) receptor agonists, and sodium glucose cotransporter-2 (SGLT-2) inhibitors. Results from these trials demonstrated cardiovascular safety, with many showing superiority³. However, these trials were highly selective and may not be generalizable to patients in real-world settings. For example, many trials required the presence of high cardiovascular risk factor levels to increase the number of events to achieve greater statistical power⁴⁻⁶.

With regulatory agencies increasingly considering real-world data (RWD) alongside randomized controlled trial (RCT) evidence in their decision-making, there is a growing need to understand how these two types of evidence complement each other. This is particularly true for

antidiabetic medications, which are tested in large CVOTs with strict inclusion criteria. Therefore, we synthesized available information to understand what proportion of patients treated in real-world setting are eligible for CVOTs and compare patient characteristics and estimated treatment effects using RWD and the respective CVOTs via systematic review of observational studies.

METHODS

This systematic review was conducted following guidelines described in the Cochrane Handbook⁷ and reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA)⁸ and Synthesis Without Meta-analysis (SWiM) guidelines⁹ (**Tables S1-S2**). The study protocol was registered in the Open Science Framework platform ([10.17605/OSF.IO/G8ERW](https://doi.org/10.17605/OSF.IO/G8ERW)).

Search strategy

We searched MEDLINE, EMBASE, and the Cochrane databases from inception to June 5th, 2024, to identify studies that used RWD to examine the proportion of patients eligible for CVOTs that emulated these trials. Our predefined search strategy utilized Medical Subject Headings (MeSH) in OVID MEDLINE, Emtree terms in EMBASE, and keywords in all databases. The search strategy was defined in consultation with a medical librarian and is described in detail in **Tables S3-S5**. Briefly, we searched using the following concepts: observational study, RCTs, major adverse cardiovascular outcomes (MACE), T2DM, emulation/replication/comparability. There were no restrictions on language or geographic location. We also conducted a hand search of references included in studies (backward search), studies that have referenced identified key studies (forward search) and previous reviews not captured by our initial database search. We also hand-searched the grey literature using Google Scholar (first 10 pages).

Inclusion and exclusion criteria

We included published, peer-reviewed observational studies (e.g., cohort or nested-case controls studies) that compared the risk of cardiovascular outcomes in real-world patients with T2DM aged 18+ years to that observed in CVOTs. To reduce bias from lack of established temporality, cross-sectional studies were only used to examine eligibility and patient characteristics; they were not included in the assessment of clinical outcomes. The interventions of interest were DPP-4 inhibitors, GLP-1 receptor agonists, and SGLT-2 inhibitors. A list of completed CVOTs and their drug class and molecule can be found in **Table S6**. There were no restrictions on comparators; active comparators such as other antidiabetic medications, standard of care, or lifestyle interventions were eligible for inclusion.

We excluded RCTs, meta-analyses, case reports, case series, letters-to-the-editor, editorials, and commentaries. In addition, we excluded conference abstracts as they often present preliminary results and typically do not undergo rigorous peer review. Full texts published in a language other than English were also excluded.

The primary outcome of interest was MACE, which included cardiovascular death, myocardial infarction (MI), and ischemic stroke. Unstable angina was also included if it was reported as part of the MACE definition. Secondary outcomes were the individual components of MACE, hospitalization for heart failure (HHF), and all-cause mortality. We examined the proportion of the study population eligible for a referenced CVOT as reported by the authors. We also compared the reported patient characteristics of the real-world population and the subpopulation eligible for the CVOTs.

Study selection

Citations generated from the electronic database search were exported to Covidence¹⁰ and duplicates were removed. Two independent reviewers (WW and WT) screened each study title and abstract for potential inclusion. Any study deemed potentially eligible by either reviewer proceeded to full-text review, where the full-text of each study was evaluated independently by both reviewers. Disagreements were resolved by consensus or if needed, with input from a third author (KBF).

Data extraction

Two independent reviewers (WW and WT) extracted data using a piloted Microsoft Excel workbook. Disagreements were resolved via consensus or if needed, with the aid of a third author (KBF).

Assessment of agreement between RCTs and RWD

To assess generalizability of RCT results in RWD studies we used agreement statistics from the DUPLICATE study to compare effect estimates of primary outcomes between the restricted RWD population and the referenced CVOT population^{11,12}. *Full statistical significance agreement* was considered to have occurred when the RWD and RCT estimates and 95% confidence intervals (CIs) were on the same side of the null. *Partial significance agreement* was considered to have occurred when the prespecified noninferiority criteria was met, even if the RWD study indicated superiority¹². Statistical significance agreement was categorized as yes, no, or partial. *Estimate agreement* was considered to have occurred when the effect estimate of the RWD study fell within the 95% CI of the RCT effect estimate. *Standardized difference agreement* was defined by standardized differences $|Z| < 1.96$.

$$Z = \frac{\hat{\theta}_{RWE} - \hat{\theta}_{RCT}}{\sqrt{\hat{\sigma}_{RWE}^2 + \hat{\sigma}_{RCT}^2}} \hat{\theta} \text{ are effect estimates and } \hat{\sigma}^2 \text{ are variances}$$

As per the FDA, all the major CVOTs were designed for non-inferiority with an upper CI limit of 1.3¹³. For our statistical significance agreement analysis, we first assessed whether the trial was able to demonstrate non-inferiority. We then evaluated whether the RWD study was able to replicate the non-inferiority finding within the same margin. If both studies demonstrated non-inferiority, then full statistical significance agreement was established. If the trial achieved non-inferiority but the RWD study achieved superiority, partial statistical significance agreement was established. In addition to non-inferiority, if superiority was established in the trial, we assessed if the RWD study also found superiority to achieve full statistical significance agreement.

Data synthesis

The high degree of clinical heterogeneity among studies and interventions prevented us from conducting a meta-analysis, thus, we reported results following the SWiM reporting guidelines⁹. The main summary measure was the adjusted hazard ratio (HR), as RCTs estimated HRs to assess MACE outcomes. Studies were grouped by drug class and further stratified by RCT emulated. Data were synthesized based on the treatment groups and are presented in tables. Heterogeneity and certainty of evidence were assessed by examining the methods in which the study attempted to replicate the RCT. We captured the exposure definition (intention-to-treat vs on-treatment), active comparators, restriction of the population, follow-up duration, and methods used to reduce confounding.

We conducted several exploratory, post-hoc analyses to examine the potential association of patient characteristics and the proportion of the RWD population eligible for the trial. First, we used scatterplots to identify trends between patient characteristics to the percentage of the population that would have been eligible for their reference CVOT. Second, we calculated

standardized mean differences of patient characteristics between the RWD populations and the trial population for select studies based on data completeness.

RESULTS

Search results

We identified 6,118 potentially relevant publications (**Figure 1**). After removing duplicates, 2,389 articles underwent title and abstract review, and 41 underwent full-text review. A total of 19 studies were included in the systematic review^{14–32}.

Study characteristics

Table 1 summarizes the characteristics of the included studies stratified by drug class. Four cohort studies^{11,21,31,32} examined ten unique populations for MACE. We identified 15 cross-sectional studies that examined the proportion eligible and patient characteristics^{14,16–20,22–30}. The included studies were conducted between 2009 and 2020 in the US (9), Europe (7), and Asia (3). A total of 10 studies assessed GLP-1 receptor agonists^{11,16–18,21,23,25,27,30,31}, 12 assessed SGLT-2 inhibitors^{11,14,18,20–22,24,26–29,32}, and two assessed DPP-2 inhibitors^{11,21}. Within the drug classes, 13 studies evaluated multiple CVOTs (range: 1 to 7 CVOTs; median: 3 per class).

MACE Outcomes in Emulated RCTs

Four studies emulated RCTs using RWD (**Table 2**). Sciannameo et al.²¹ used odds weighting based on RCT subgroup-specific HRs to transport results onto their RWD population; consequently, there was no real-world exposure or comparator for these analyses. As the exposures and outcomes in these analyses were directly from the relevant RCTs, we did not calculate the agreement statistics for this study. For the other three studies, exposure definitions of intention-to-treat (ITT) or on-treatment were used with an active comparator. These studies

also imitated the inclusion criteria and exclusion criteria of the emulated RCTs. They also ensured follow-up was the same duration as the RCT and used propensity score matching to reduce confounding. MACE was defined as a composite endpoint of cardiovascular or all-cause death, stroke, and MI; Franklin et al. emulated TECOS by including angina in their MACE definition¹¹.

For studies that investigated SGLT-2 inhibitors, there were three RCTs examined: EMPA-REG Outcome³³, CANVAS⁴ and DECLARE⁵. Treatment matched the SGLT-2 inhibitors evaluated in the RCTs. However, while the RCTs used a placebo comparator, active comparators (DPP-4 inhibitors) were used in the RWD studies. There was strong agreement between RWD and RCTs, with the three studies achieving full statistical significance agreement and estimate agreement, with HRs from the RWD in the same direction as those from the RCTs and within the RCT's 95% CIs.

There were three analyses of GLP-1 receptor agonists, which examined six RCTs (EXSCEL³⁴, LEADER³⁵, PIONEER-6³⁶, REWIND⁶, SUSTAIN-6³⁷). The RWD studies used ITT and on-treatment exposure definitions. For the emulation of the LEADER trial, liraglutide was compared with sulfonylureas, second-to-third-line antidiabetic drugs, and DPP-4 inhibitors. Across the three different analyses, all showed partial statistical significance agreement. All studies except Abrahami et al. evaluating liraglutide compared to sulfonylureas demonstrated estimate agreement³¹.

There were five analyses of DPP-4 inhibitors, which examined three RCTs (CARMELINA³⁸, SAVOR-TIMI-53³⁹, and TECOS⁴⁰). RCT odds weighting was used in two of the five analyses. In the other three analyses, an on-treatment exposure definition was used, and

second-generation sulfonylureas were used as the comparator. All analyses were able to replicate the non-inferiority findings of the RCTs, and all but one demonstrated estimate agreement.

Secondary Clinical Outcomes in Emulated RCTs

Table 3 summarizes the results for secondary clinical outcomes in the RWD studies and the emulated RCTs. Jang et al, which emulated the EMPA-REG Outcome trial by comparing empagliflozin with sitagliptin, examined select individual components of MACE (MI and stroke), all-cause mortality, and HHF³². All-cause mortality, MI, and HHF had either partial or full statistical significance agreement and estimate agreement. Empagliflozin was found to be non-inferior for stroke in RWD but not the RCT; thus, statistical significance agreement was not met for this outcome. However, there was estimate agreement³². Sciannameo et al. and Franklin et al. examined HHF and/or cardiovascular death when replicating the DECLARE trial^{15,21}. Franklin et al. found both regulatory agreement for superiority and estimate agreement.

Eligibility

Table 4 summarizes the findings when examining the proportion of RWD population eligible (including cross-sectional studies) for CVOTs. In total, there were 59 populations across CVOTs evaluating SGLT-2 inhibitors and GLP-1 receptor agonists. Populations varied in composition, with some from the general population, others only from select populations (inpatient vs outpatient), and others restricted to users of the drug class of interest.

For SGLT-2 inhibitors, RWD populations were compared to a total of four RCTs (CANVAS⁴, DECLARE⁵, EMPA-REG Outcome,³³ and VERTIS-CV⁴¹). The median percentage eligible from the trials ranged between 12.6% to 30.7%. Of note, the Jang et al. study that demonstrated full statistical, estimate, and standardized difference agreement with the EMPA-REG Outcome trial had an 12.6% eligible for the RCT from a population of patients with T2DM

who were new-users of empagliflozin or sitagliptin. For GLP-1 receptor agonists, RWD populations were compared to eight different RCTs (ELIXA⁴², EXSCEL³⁴, FREEDOM-CVO⁴³, HARMONY⁴⁴, LEADER³⁵, PIONEER-6³⁶, REWIND⁶, SUSTAIN-6³⁷). Median percentage eligible from the trials ranged from 11.8% to 42.6%. No included studies evaluated trial eligibility for DPP-4 inhibitors.

We conducted post-hoc analyses for patient characteristics which can be found in **Figures S1-S7**. There were no association between age or sex in RWD populations and percentage of individuals who were eligible for the RCTs. Moreover, there were substantial differences in patient characteristics such as age, sex, use of antidiabetic medication between the RWD populations compared to the populations examined in the respective RCTs.

DISCUSSION

Our study was designed to synthesize observational studies that emulated CVOTs in patients with T2DM and to summarize cross-sectional studies assessing the proportion of real-world populations eligible for previous CVOTs. We found strong agreement between results from RWD and RCTs for SGLT-2 inhibitors, GLP-1 receptor agonists, and DPP-4 inhibitors. Individual components of MACE, HHF, and cardiovascular death also showed consistency between the RWD studies^{15,21,32} and the SGLT-2 inhibitor CVOTs, EMPA-REG Outcome³³ and DECLARE⁵. There was considerable heterogeneity in the percentage of individuals who would have been eligible for the CVOTs of SGLT-2 inhibitors and GLP-1 receptor agonists. Patient characteristics also varied between RWD and RCTs, with no association between patient characteristics and real-world proportion eligible for RCTs.

Our systematic review highlighted the heterogeneity of methods used to emulate RCTs using RWD. As this is an emerging area of research, the studies included in our study had different exposure definitions, active comparators, and methods to reduce confounding. The lack of agreement between Abrahami et al's study and the LEADER trial highlights how differences in comparators from sulfonylureas to broader 2nd or 3rd line antidiabetic drugs influenced the direction of the effect estimate³¹. In addition, some studies utilized an ITT approach where exposure was defined at cohort entry until end of follow-up while others used an on-treatment approach which censored for treatment discontinuation or switching. As these changes may impact the research question and effect estimate, more research is needed to better understand the strengths and limitations of using different methods for replicating RCTs with RWD.

Our study has shown that studies emulating RCTs using RWD to estimate the effects of newer antidiabetic medications on MACE generally demonstrated agreement. However, there was some heterogeneity in our results, which may be due to the types of methods used for replication such as choice of exposure definition (ITT vs on-treatment), choice of active comparator, and methods to adjust for confounding. In addition, there were some differences between RCTs and RWD which may be due to inherent differences between the two designs, such as the source population, exposure (active comparator vs placebo), follow-up duration, and how events are captured and measured. The results may have also differed due to chance. Overall, our systematic review demonstrated that similar conclusions can be obtained when using RWD to emulate RCTs when efforts are made in study design to closely mimic RCTs. While RCTs are the gold-standard for the generation of evidence, regulatory bodies in countries/regions such as Canada, USA and the European Union are increasingly evaluating the use of RWD in decision-making^{2,45,46}. Our

results provide reassurance that with the proper methods, RWD can similarly generate quality evidence.

The proportion of the real-world population that would have been eligible for the RCTs varied across studies but was almost always less than 50%. The lack of generalizability of these RCT populations is driven in part by their design. These trials predominantly included older individuals who possessed one or more cardiovascular disease (CVD) risk factors and generally had longer durations of disease. These factors were necessary for event driven CVOTs to ensure that a sufficient number of events could be accrued in a short time frame³. This requirement explains in part the observed heterogeneity (and large standardized mean differences for patient characteristics) between the RWD and the RCT populations. For example, the RWD populations consistently had lower percentages of established CVD. Given these differences, it is important to utilize studies from routine practice settings to complement findings from highly controlled RCTs to account for differences in patients and practice for clinical decision making³. There are opportunities to use RWD for Phase IV trials and pragmatic trials as highlighted by the FDA².

Previous reviews in the area of RCT emulation using RWD examined the reporting of RWD studies that aimed to replicate RCTs^{47,48}. They found inconsistencies in the reporting of key elements in these studies. For example, in a systematic review of 200 studies aiming to emulate RCTs using observational data, Hansford et al. reported that 43% did not describe all key elements of how the target trial was emulated⁴⁷. In addition, only 37% of studies reported potential unmeasured confounders. A scoping review of 96 studies aiming to emulate a target trial across medical fields also identified unmeasured confounding as the most commonly stated limitation⁴⁸. We found that the cohort studies included in our systematic review all reported the key elements of eligibility criteria, treatment strategy, assignment procedures, outcome, follow-

up, causal contrast, analyses plan and specification of time-zero. However, they did not comment on potential unmeasured confounders. Our study builds upon these previous studies by examining agreement statistics between RWD and RCTs, the proportion of RWD patients who would have been eligible for RCTs, and differences in patient characteristics. The recent development of guidelines for the reporting of target trial emulation may improve reporting in this area⁴⁹.

Our study has many strengths. It included a comprehensive, systematic search that was constructed with the assistance of an experienced librarian. To our knowledge, this is the first systematic review to examine the emulation of CVOTs for newer antidiabetic drugs using RWD. In addition, we assessed the agreement between the RWD studies and RCTs based on regulatory standards. Furthermore, our study is the first to synthesize the generalizability of CVOTs in the real-world by examining the proportion eligible of real-world populations for these CVOTs. We also compared patient characteristics between the CVOTs and RWD to better understand differences in study populations.

Our study has several potential limitations. First, there was a limited number of studies emulating RCTs using RWD among patients with T2DM. These studies also had important heterogeneity in study design, drug class, and analytical approach. Consequently, we conducted a systematic review without formal meta-analysis. Second, as an emerging area of research, there are currently no established guidelines on how to assess the quality of these studies for external validity. While we used the DUPLICATE Team's agreement statistics^{12,15} to assess the estimates achieved from the emulation in comparison to the RCT that was being emulated, these statistics have their own limitations. For example, it is difficult to replicate studies that have shown a null effect, as observational studies often have more precision. This issue is well illustrated by our results for DPP-4 inhibitors where RWD studies achieved superiority while the RCTs reported a

null effect^{15,21}. Moreover, the estimate agreement is contingent on the precision of the RCT. If the RCT has a wider CI, there is a higher probability that the RWD estimate will fall within this interval. Third, to our knowledge, there are no guidelines or standards on reporting or assessing external validity of RCT. Finally, we chose to include cross-sectional studies for our secondary objectives examining the proportion eligible of real-world populations for RCTs and their characteristics, however, these studies had considerable missing data and are at an inherently higher risk of bias.

AUTHORS' CONTRIBUTIONS

WW led the protocol development, conducted the analysis, and drafted the manuscript. WW and WT extracted data. All authors contributed to protocol development, were involved in data interpretation, critically reviewed the manuscript for important intellectual content, and approved the final manuscript. KBF supervised the study and is the guarantor.

FUNDING

WW received funding from a Faculty of Medicine Internal Studentship and a CANadian Consortium of Clinical Trial TRAINing platform (CANTRAIN) studentship; CANTRAIN is funded by the Canadian Institutes of Health Research. KBF holds a William Dawson Scholar award from McGill University.

CONFLICT OF INTEREST

None to declare.

REFERENCES

1. Sherman, R. E. *et al.* Real-World Evidence — What Is It and What Can It Tell Us? *N Engl J Med* **375**, 2293–2297 (2016).
2. Rick Turner, J. FDA’s New Guidance for Industry Addressing Evaluation of the Safety of New Drugs for Improving Glycemic Control: A Case Study in Regulatory Science. *Ther Innov Regul Sci* **55**, 1–5 (2021).
3. Eckel, R. H., Farooki, A., Henry, R. R., Koch, G. G. & Leiter, L. A. Cardiovascular Outcome Trials in Type 2 Diabetes: What Do They Mean for Clinical Practice? *Clin Diabetes* **37**, 316–337 (2019).
4. Neal, B. *et al.* Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *N Engl J Med* **377**, 644–657 (2017).
5. Furtado, R. H. M. *et al.* Dapagliflozin and Cardiovascular Outcomes in Patients With Type 2 Diabetes Mellitus and Previous Myocardial Infarction. *Circulation* **139**, 2516–2527 (2019).
6. Gerstein, H. C. *et al.* Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. *The Lancet* **394**, 121–130 (2019).
7. Cochrane Handbook for Systematic Reviews of Interventions. at <https://training.cochrane.org/handbook/current>
8. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
9. Campbell, M. *et al.* Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ* **368**, 16890 (2020).
10. Covidence - Better systematic review management. *Covidence* at <https://www.covidence.org/>

11. Franklin, J. M. *et al.* Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. *Clinical Pharmacology & Therapeutics* **107**, 817–826 (2020).
12. Wang, S. V., Schneeweiss, S., & RCT-DUPLICATE Initiative Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA* **329**, 1376–1385 (2023).
13. Center for Drug Evaluation and Research Type 2 Diabetes Mellitus: Evaluating the Safety of New Drugs for Improving Glycemic Control Guidance for Industry. (2020).at
<<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/type-2-diabetes-mellitus-evaluating-safety-new-drugs-improving-glycemic-control-guidance-industry>>
14. McGovern, A., Feher, M., Munro, N. & Lusignan, S. de Sodium-Glucose Co-transporter 2 (SGLT2) Inhibitor: Comparing Trial Data and Real-World Use. *Diabetes Ther* **8**, 365–376 (2017).
15. Franklin, J. M. *et al.* Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative. *Circulation* **143**, 1002–1013 (2021).
16. Romera, I. *et al.* A Retrospective Observational Study Examining the Generalizability of Glucagon-Like Peptide 1 Receptor Agonist Cardiovascular Outcome Trials to the Real-World Population with Type 2 Diabetes in Spain: The REPRESENT Study. *Adv Ther* **39**, 3589–3601 (2022).
17. Webb, J. *et al.* Cardiovascular risk profiles: A cross-sectional study evaluating the generalizability of the glucagon-like peptide-1 receptor agonist cardiovascular outcome trials

- REWIND, LEADER and SUSTAIN-6 to the real-world type 2 diabetes population in the United Kingdom. *Diabetes Obes Metab* **24**, 289–295 (2022).
18. Wittbrodt, E., Chamberlain, D., Arnold, S. V., Tang, F. & Kosiborod, M. Eligibility of patients with type 2 diabetes for sodium-glucose co-transporter-2 inhibitor cardiovascular outcomes trials: An assessment using the Diabetes Collaborative Registry. *Diabetes, Obesity and Metabolism* **21**, 1985–1989 (2019).
 19. Eric T. Wittbrodt, P. *et al.* Generalizability of Glucagon-Like Peptide-1 Receptor Agonist Cardiovascular Outcome Trials Enrollment Criteria to the US Type 2 Diabetes Population. **24**, (2018).
 20. Pintat, S. *et al.* Eligibility of patients with type 2 diabetes for sodium–glucose cotransporter 2 inhibitor cardiovascular outcomes trials: a global perspective from the DISCOVER study. *BMJ Open Diab Res Care* **7**, e000627 (2019).
 21. Sciannameo, V. *et al.* Transposition of cardiovascular outcome trial effects to the real-world population of patients with type 2 diabetes. *Cardiovascular Diabetology* **20**, 103 (2021).
 22. Nicolucci, A. *et al.* Generalizability of Cardiovascular Safety Trials on SGLT2 Inhibitors to the Real World: Implications for Clinical Practice. *Adv Ther* **36**, 2895–2909 (2019).
 23. Boye, K. S. *et al.* Generalizability of glucagon-like peptide-1 receptor agonist cardiovascular outcome trials to the overall type 2 diabetes population in the United States. *Diabetes, Obesity and Metabolism* **21**, 1299–1304 (2019).
 24. Birkeland, K. I. *et al.* How representative of a general type 2 diabetes population are patients included in cardiovascular outcome trials with SGLT2 inhibitors? A large European observational study. *Diabetes Obes Metab* **21**, 968–974 (2019).

25. Fan, W., Tong, C. & Wong, N. D. LEADER Trial Eligibility and Preventable Cardiovascular Events in US Adults with Diabetes: the National Health and Nutrition Examination Surveys 2007–2016. *Cardiovasc Drugs Ther* **34**, 737–743 (2020).
26. Zhou, Y.-L. *et al.* Population diversity of cardiovascular outcome trials and real-world patients with diabetes in a Chinese tertiary hospital. *Chinese Medical Journal* **134**, 1317–1323 (2021).
27. Arnold, S. V. *et al.* Real-world use and modeled impact of glucose-lowering therapies evaluated in recent cardiovascular outcomes trials: An NCDR® Research to Practice project. *Eur J Prev Cardiol* **24**, 1637–1645 (2017).
28. Shao, S.-C. *et al.* Sodium glucose co-transporter 2 inhibitors and cardiovascular event protections: how applicable are clinical trials and observational studies to real-world patients? *BMJ Open Diabetes Research and Care* **7**, e000742 (2019).
29. Hinton, W., Feher, M., Munro, N., Joy, M. & Lusignan, S. de Sodium–glucose co-transporter-2 inhibitor cardiovascular outcome trials and generalizability to English primary care. *Diabet Med* **37**, 1499–1508 (2020).
30. Sciannameo, V. *et al.* Enrolment criteria for diabetes cardiovascular outcome trials do not inform on generalizability to clinical practice: The case of glucagon-like peptide-1 receptor agonists. *Diabetes Obes Metab* **22**, 817–827 (2020).
31. Abrahami D. *et al.* Use of Real-World Data to Emulate a Clinical Trial and Support Regulatory Decision Making: Assessing the Impact of Temporality, Comparator Choice, and Method of Adjustment. *Clin. Pharmacol. Ther.* **109**, 452–461 (2021).

32. Jang, H. Y., Kim, I.-W. & Oh, J. M. 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. *Frontiers in Pharmacology* **13**, (2022).
33. Zinman, B. *et al.* Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *N Engl J Med* **373**, 2117–2128 (2015).
34. Drucker, D. J. *et al.* Exenatide once weekly versus twice daily for the treatment of type 2 diabetes: a randomised, open-label, non-inferiority study. *The Lancet* **372**, 1240–1250 (2008).
35. Marso, S. P. *et al.* Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *N Engl J Med* **375**, 311–322 (2016).
36. Husain, M. *et al.* Oral Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *New England Journal of Medicine* **381**, 841–851 (2019).
37. Marso, S. P. *et al.* Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *N Engl J Med* **375**, 1834–1844 (2016).
38. Rosenstock, J. *et al.* Effect of Linagliptin vs Placebo on Major Cardiovascular Events in Adults With Type 2 Diabetes and High Cardiovascular and Renal Risk. *JAMA* **321**, 69–79 (2019).
39. Scirica, B. M. *et al.* Saxagliptin and Cardiovascular Outcomes in Patients with Type 2 Diabetes Mellitus. *New England Journal of Medicine* **369**, 1317–1326 (2013).
40. Green, J. B. *et al.* Effect of Sitagliptin on Cardiovascular Outcomes in Type 2 Diabetes. *New England Journal of Medicine* **373**, 232–242 (2015).
41. Cannon, C. P. *et al.* Cardiovascular Outcomes with Ertugliflozin in Type 2 Diabetes. *New England Journal of Medicine* **383**, 1425–1435 (2020).

42. Pfeffer, M. A. *et al.* Lixisenatide in Patients with Type 2 Diabetes and Acute Coronary Syndrome. *New England Journal of Medicine* **373**, 2247–2257 (2015).
43. Ruff, C. T. *et al.* Subcutaneous infusion of exenatide and cardiovascular outcomes in type 2 diabetes: a non-inferiority randomized controlled trial. *Nat Med* **28**, 89–95 (2022).
44. Hernandez, A. F. *et al.* Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial. *The Lancet* **392**, 1519–1529 (2018).
45. Canada, H. Regulatory review of drugs and devices: Strengthening the use of real world evidence for drugs. (2018).at <<https://www.canada.ca/en/health-canada/corporate/transparency/regulatory-transparency-and-openness/improving-review-drugs-devices/strengthening-use-real-world-evidence-drugs.html>>
46. Use of real-world evidence in regulatory decision making – EMA publishes review of its studies | European Medicines Agency. at <<https://www.ema.europa.eu/en/news/use-real-world-evidence-regulatory-decision-making-ema-publishes-review-its-studies>>
47. Hansford, H. J. *et al.* Reporting of Observational Studies Explicitly Aiming to Emulate Randomized Trials: A Systematic Review. *JAMA Network Open* **6**, e2336023 (2023).
48. Zuo, H., Yu, L., Campbell, S. M., Yamamoto, S. S. & Yuan, Y. The implementation of target trial emulation for causal inference: a scoping review. *Journal of Clinical Epidemiology* **162**, 29–37 (2023).
49. Hansford, H. J. *et al.* Development of the TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guideline. *BMJ Open* **13**, e074626 (2023).

Table 1 Summary of study characteristics for systematic review of observational studies replicating cardiovascular outcome trials among patients with type 2 diabetes

Study	Study design	Data origin	Study Period	Exposure class	RCT referenced	Number of participants in study	Patient population
Abrahami, 2020 ³¹	Cohort Study	UK	2009-2013	GLP-1 Receptor Agonists	LEADER ³⁵	63 297 ^a	Clinical Practice Research Datalink (CPRD)
Franklin, 2020 ¹⁵	Cohort study	USA	2004-2019	GLP-1 Receptor Agonists	LEADER ³⁵	168 690	US Healthcare Claims: OPTUM, IBM, Medicare
Sciannameo, 2021 ²¹	Reanalyses of RCT ^b	Italy	2015-2016	GLP-1 Receptor Agonist	PIONEER-6 ³⁶ , REWIND ⁶ , SUSTAIN-6 ³⁷ , EXSCEL ³⁴ , LEADER ³⁵	139 700	DARWIN- T2DM database
Franklin, 2020 ¹⁵	Cohort study	USA	2004-2019	SGLT-2 inhibitors	DECLARE ⁵ , EMPA-REG ³³ , CANVAS ⁴	305 744 ^c	US Healthcare Claims: OPTUM, IBM, Medicare
Jang, 2022 ³²	Cohort Study	South Korea	2011-2020	SGLT-2 inhibitors	EMPA-REG ³³	23 126	Korean Health Insurance Review and Assessment Service
Sciannameo, 2021 ²¹	Reanalyses of RCT ^b	Italy	2015-2016	SGLT-2 inhibitors	DECLARE ⁵ , EMPA-REG ³³	139 700	DARWIN- T2DM database
Franklin, 2020 ¹⁵	Cohort study	USA	2004-2019	DPP-4 inhibitors	CARMELINA ³⁸ , TECOS ⁴⁰ , SAVOR-TIMI ³⁹	633 432 ^d	US Healthcare Claims: OPTUM, IBM, Medicare
Sciannameo, 2021 ²¹	Reanalyses of RCT ^b	Italy	2015-2016	DPP-4 Inhibitors	SAVOR-TIMI ³⁹ , TECOS ⁴⁰	139 700	DARWIN- T2DM database
Arnold, 2017 ²⁷	Cross-sectional	USA	2017	GLP-1 Receptor Agonists	LEADER ³⁵	87 601	Diabetes Collective Registry (DCR)
Boye, 2018 ²³	Cross-sectional	USA	2012-2017	GLP-1 Receptor Agonists	EXSCEL ³⁴ , LEADER ³⁵ , REWIND ⁶ , SUSTAIN-6 ³⁷	26 110 573	IQIVIA Real World Data Adjudicated Claims linked with EMR and US National Health and Nutrition Examination Surveys (NHANES)
Fan, 2020 ²⁵	Cross-sectional	USA	2007- 2016	GLP-1 Receptor Agonists	LEADER ³⁵	29 629	NHANES
Romera, 2022 ¹⁶	Cross-sectional	Spain	2013- 2019	GLP-1 Receptor Agonists	LEADER ³⁵ , REWIND ⁶ , SUSTAIN-6 ³⁷	24 268	IQIVIA EMR database in Spain
Sciannameo, 2020 ²¹	Cross-sectional	Italy	2015- 2016	GLP-1 Receptor Agonists	LEADER ³⁵ , SUSTAIN-6 ³⁷ , EXSCEL ³⁴ , REWIND ⁶ , PIONEER-6 ³⁶ , HARMONY ⁴⁴	130 380	DARWIN-T2DM database

Webb, 2021¹⁷	Cross-sectional	UK	2018	GLP-1 Receptor Agonists	REWIND ⁶ , LEADER ³⁵ , SUSTAIN-6 ³⁷	33 118	CPRD GOLD
Wittbrodt, 2018¹⁸	Cross-sectional	USA	2009-2012	GLP-1 Receptor Agonist	EXSCEL ³⁴ , REWIND ⁶ , FREEDOM-CVO ⁴³ , LEADER ³⁵ , SUSTAIN-6 ³⁷ , HARMONY ⁴⁴ , ELIXA ⁴²	20 142	NHANES
Arnold, 2017²⁷	Cross-sectional	USA	2017	SGLT-2 inhibitors	EMPA-REG ³³	47 872	Diabetes Collective Registry (DCR)
Birkeland, 2018²⁴	Cross-sectional	Norway, Sweden, Germany, Netherlands	2014-2015	SGLT-2 inhibitors	CANVAS ⁴ , EMPA-REG ³³ , VERTIS-CV ⁴¹	803 836	Germany: Health Insurance funds database Netherlands: Electronic health database Norway and Sweden: National Public Health System
Hinton, 2020²⁹	Cross-sectional	England	2016	SGLT-2 inhibitors	CANVAS ⁴ , DECLARE ⁵ , EMPA-REG ³³ , VERTIS-CV ⁴¹	1 595 445	Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC)
McGovern, 2017¹⁴	Cross-sectional	England	2016	SGLT-2 inhibitors	EMPA-REG ³³	60 327	RCGP RSC
Nicolucci, 2019²²	Cross-sectional	Italy	2016	SLGT-2 inhibitors	EMPA-REG ³³ , CANVAS ⁴ , DECLARE ⁵ , VERTIS-CV ⁴¹	455 622	Italian Association of Diabetologists (AMD) database
Pintat, 2019²⁰	Cross-sectional	38 Countries in Discover study	2017 ^e	SGLT-2 inhibitors	CANVAS ⁴ , DECLARE ⁵ , EMPA-REG ³³ , VERTIS-CV ⁴¹	11 385	DISCOVER prospective observational study
Shao, 2019²⁸	Cross-sectional	Taiwan	2018-2019	SGLT-2 inhibitors	CANVAS ⁴ , DECLARE ⁵ , EMPA-REG ³³ , VERTIS CV ⁴¹	11 650	Chang Gung Research Database
Wittbrodt, 2019¹⁸	Cross-sectional	USA	2013- 2016	SGLT-2 inhibitors	EMPA-REG ³³ , CANVAS ⁴ , DECLARE ⁵ , VERTIS-CV ⁴¹	172 643	DCR
Zhou, 2020²⁶	Cross-sectional	China	2011-2019	SGLT-2 inhibitors	EMPA-REG ³³	214 963	EMR from West China Hospital

T2DM: Type 2 Diabetes Mellitus; EMR: Electronic Medical Records; CPRD: Clinical Practice Research Datalink; DCR: Diabetes Collective Registry; NHANES: National Health and Nutrition Examination Surveys; RCGP: Royal College of General Practitioners; RSC: Research and Surveillance Centre; AMD: Italian Association of Diabetologists; RCT: Randomized Controlled Trial; ^aTotal including liraglutide, sulfonylureas and second-to-third line antidiabetic medications populations; ^bWeighting RCT stratum-specific treatment effects according to proportions of a given characteristics in the target population; ^cTotal from DECLARE, EMPA-REG Outcome and CANVAS RWE populations; ^dTotal from CAREMLINA, TECOS and SAVOR-TIMI RWE population ; ^eTaken from DISCOVER study

Study	RCT	Exposure definition	Exposure	Comparator	Exposure N	Comparator N	HR	95% CI	RCT HR	95% CI	Observed Agreement ^a		
											SA	EA	SD
SGLT-2 inhibitors													
Jang, 2022	EMPA-REG ³³	ITT	Empagliflozin	Sitagliptin	11 563	11 563	0.87	0.79, 0.96	0.86	0.74, 0.99	Y	Y	Y
Franklin, 2020	EMPA-REG ³³	ITT	Empagliflozin	DPP-4i	4 687	2 333	0.83	0.73, 0.94	0.86	0.74, 0.99	Y	Y	Y
Franklin, 2020	CANVAS ⁴	On treatment	Canagliflozin	DPP-4i	5 795	4 347	0.77	0.70, 0.85	0.86	0.75, 0.97	Y	Y	Y
Sciannameo, 2021 ^b	DECLARE ⁵	N/A	N/A	N/A	N/A	N/A	0.94	0.84, 1.04	0.93	0.84, 1.03	-	-	-
Sciannameo, 2021 ^b	EMPA-REG ³³	N/A	N/A	N/A	N/A	N/A	0.88	0.74, 1.03	0.86	0.74, 0.99	-	-	-
GLP-1 Receptor Agonists													
Abrahami, 2020	LEADER ³⁵	ITT	Liraglutide	SU	1 868	25 895	1.03	0.82, 1.30	0.87	0.78, 0.97	N	N	Y ^d
Abrahami, 2020	LEADER ³⁵	ITT	Liraglutide	2 nd – 3 rd line ^c	1 864	32 899	0.97	0.78, 1.22	0.87	0.78, 0.97	N	Y	Y ^d
Franklin, 2020	LEADER ³⁵	On treatment	Liraglutide	DPP-4i	4 668	4 672	0.82	0.76, 0.87	0.87	0.78, 0.97	Y	Y	Y
Sciannameo, 2021 ^b	EXSCEL ³⁴	N/A	N/A	N/A	N/A	N/A	0.92	0.82, 1.02	0.91	0.83, 1.00	-	-	-
Sciannameo, 2021 ^b	LEADER ³⁵	N/A	N/A	N/A	N/A	N/A	0.88	0.77, 0.99	0.87	0.78, 0.97	-	-	-
Sciannameo, 2021 ^b	PIONEER-6 ³⁶	N/A	N/A	N/A	N/A	N/A	0.76	0.41, 1.10	0.79	0.57, 1.11	-	-	-
Sciannameo, 2021 ^b	REWIND ⁶	N/A	N/A	N/A	N/A	N/A	0.87	0.76, 0.98	0.88	0.79, 0.99	-	-	-
Sciannameo, 2021 ^b	SUSTAIN-6 ³⁷	N/A	N/A	N/A	N/A	N/A	0.73	0.47, 0.99	0.74	0.58, 0.95	-	-	-
DPP-4 Inhibitors													
Franklin, 2020	CARMELIN ³⁸	On treatment	Linagliptin	2 nd – gen SU	3 494	3 485	0.90	0.84, 0.96	1.02	0.89, 1.17	P	Y	Y
Franklin, 2020	SAVOR-TIMI ³⁹	On treatment	Saxagliptin	2 nd – gen SU	8 280	8 212	0.81	0.76, 0.86	1.00	0.89, 1.12	P	N	Y
Franklin, 2020 ^e	TECOS ⁴⁰	On treatment	Sitagliptin	2 nd – gen SU	7 257	7 266	0.89	0.86, 0.91	0.98	0.88, 1.09	P	Y	Y
Sciannameo, 2021 ^b	SAVOR-TIMI ³⁹	N/A	N/A	N/A	N/A	N/A	0.99	0.87, 1.10	1.00	0.89, 1.12	-	-	-
Sciannameo, 2021 ^b	TECOS ⁴⁰	N/A	N/A	N/A	N/A	N/A	0.97	0.87, 1.06	0.98	0.88, 1.09	-	-	-

Table 2 Summary of MACE outcomes for RWD and RCT studies grouped by exposure class

Abbreviations: HR: Hazard Ratio; CI: Confidence Interval; ITT: intention-to-treat, DPP-4i: Dipeptidyl peptidase 4 inhibitors; SU: sulfonylureas; 2nd – gen SU: Second generation sulfonylureas; RCT: Randomized Controlled Trial;

^aSA: Full statistical significance agreement (Y/N/P) – determines if the RWD and RCT have estimates and CIs on the same side of the null; P: partial significance agreement- met the prespecified noninferiority criteria even though the database study may have indicated superiority, EA: Estimate agreement (Y/N) – determines if the effect estimate of the RWD falls within the 95% confidence interval of the RCT effect estimate. SD: Standardized difference agreement is defined by standardized differences $|Z| < 1.96$ (Y = yes, N = no).

^b Weighting RCT stratum-specific treatment effects according to proportions of a given characteristics in the target population; ^csecond to third line antidiabetic drugs; ^d Calculated by hand; ^eMACE including angina;

Study	RCT	Exposure definition	Exposure	Comparator	Exposure N	Comparator N	Outcome	HR	95% CI	RCT HR	RCT 95% CI	Observed Agreement ^a		
												SA	EA	SD
Jang, 2022	EMPA-REG ³³	ITT	Empagliflozin	Sitagliptin	11 563	11 563	All cause death	0.78	0.67, 0.91	0.68	0.57, 0.82	Y	Y	Y
Jang, 2022	EMPA-REG ³³	ITT	Empagliflozin	Sitagliptin	11 563	11 563	MI	0.91	0.76, 1.08	0.87	0.70, 1.09	Y	Y	Y
Jang, 2022	EMPA-REG ³³	ITT	Empagliflozin	Sitagliptin	11 563	11 563	Stroke	0.89	0.75, 1.05	1.18	0.89, 1.56	N	Y	Y
Jang, 2022	EMPA-REG ³³	ITT	Empagliflozin	Sitagliptin	11 563	11 563	HHF	0.85	0.75, 0.95	0.65	0.50, 0.85	Y	Y	Y
Franklin, 2020	DECLARE ⁵	ITT	Dapagliflozin	DPP-4i	8 582	8 578	HHF and CV death	0.70	0.59, 0.82	0.83	0.73, 0.95	Y	Y	Y ^b
Sciannameo, 2021^c	DECLARE ⁵	N/A	N/A	N/A	N/A	N/A	HHF and/or CV death	0.86	0.73, 0.99	0.83	0.73, 0.95	-	-	-

Table 3 Summary of secondary outcomes of all-cause mortality, MI, Stroke, HHF and CV death for RWD and RCT studies for SGLT-2 Inhibitors

Abbreviations: HR: Hazard Ratio; CI: Confidence Interval; ITT: intention-to-treat; HHF: hospitalization due to heart failure; CV: cardiovascular; RCT: Randomized Controlled Trial; DPP-4i: Dipeptidyl peptidase 4 inhibitors

^a SA: Full statistical significance agreement (Y/N/P) – determines if the RWD and RCT have estimates and CIs on the same side of the null; P: partial significance agreement- met the prespecified noninferiority criteria even though the database study may have indicated superiority, EA: Estimate agreement (Y/N) – determines if the effect estimate of the RWD falls within the 95% confidence interval of the RCT effect estimate. SD: Standardized difference agreement is defined by standardized differences $|Z| < 1.96$ (Y = yes, N = no).

^b Calculated by hand; ^cWeighting RCT stratum-specific treatment effects according to proportions of a given characteristics in the target population

Table 4 Summary of proportion of real-world patients eligible for cardiovascular outcome trials examining newer antidiabetic drugs.

Study	RCT	Population Name	Population Size	Eligible for RCT (%)	Median (25 th , 75 th percentile)
<u>SGLT-2 Inhibitors</u>					
Birkeland, 2018	CANVAS ⁴	Total population with T2DM	803,836	34.0	
Hinton, 2020	CANVAS ⁴	Total population with T2DM	84,394	17.0	
Nicolucci, 2019	CANVAS ⁴	Total population with T2DM	45,566	29.4	
Pintat, 2019	CANVAS ⁴	Total population with T2DM	11,385	19.9	
Shao, 2019	CANVAS ⁴	Canagliflozin new users	1,091	57.3	
Wittbrodt, 2019	CANVAS ⁴	Total population with T2DM	172,643	32.0	
Summary	CANVAS ⁴				30.7 (22.3, 33.5)
Birkeland, 2018	DECLARE ⁵	Total population with T2DM	803,836	59.0	
Hinton, 2020	DECLARE ⁵	Total population with T2DM	84,394	27.0	
Nicolucci, 2019	DECLARE ⁵	Total population with T2DM	45,566	55.9	
Pintat, 2019	DECLARE ⁵	Total population with T2DM	11,385	40.5	
Shao, 2019	DECLARE ⁵	Dapagliflozin new users	4,748	50.4	
Wittbrodt, 2019	DECLARE ⁵	Total population with T2DM	172,643	44.0	
Summary	DECLARE ⁵				47.2 (41.4, 54.5)
Arnold, 2017	EMPA-REG ³³	SGLT-2 inhibitor users	2,497	5.2	
Arnold, 2017	EMPA-REG ³³	Total population with T2DM	182,525	26.2	
Birkeland, 2018	EMPA-REG ³³	Total population with T2DM	803,836	21.0	
Hinton, 2020	EMPA-REG ³³	Total population with T2DM	84,394	7.0	
Jang, 2022	EMPA-REG ³³	Empagliflozin, sitagliptin new users	384,579	12.6	
McGovern, 2017	EMPA-REG ³³	RCGP RSC Total T2DM	60,327	15.7	
McGovern, 2017	EMPA-REG ³³	RCGP RSC SGLT2-users	1,642	11.1	
Nicolucci, 2019	EMPA-REG ³³	Total population with T2DM	45,566	11.7	
Pintat, 2019	EMPA-REG ³³	Total population with T2DM	11,385	7.1	
Shao, 2019	EMPA-REG ³³	Empagliflozin new users	11,650	18.7	
Wittbrodt, 2019	EMPA-REG ³³	Total population with T2DM	172,643	26.0	
Zhou, 2020	EMPA-REG ³³	Inpatients	48,257	17.4	
Zhou, 2020	EMPA-REG ³³	Outpatients	36,857	7.2	
Summary	EMPA-REG ³³				12.6 (7.2, 18.7)
Birkeland, 2018	VERTIS-CV ⁴¹	Total population with T2DM	803,836	17.0	
Hinton, 2020	VERTIS-CV ⁴¹	Total population with T2DM	84,394	7.0	
Nicolucci, 2019	VERTIS-CV ⁴¹	Total population with T2DM	45,566	12.8	
Pintat, 2019	VERTIS CV ⁴¹	Total population with T2DM	11,385	7.2	
Shao, 2019	VERTIS-CV ⁴¹	Total population with T2DM	33,118	19.2	
Wittbrodt, 2019	VERTIS CV ⁴¹	Total population with T2DM	172,643	27.0	
Summary	VERTIS CV ⁴¹				14.9 (8.6, 18.7)
<u>GLP-1 Receptor Agonists</u>					
Wittbrodt, 2018	ELIXA ⁴²	Adults likely to have T2DM	23,941,512	6.4	
Boye, 2018	EXSCEL ³⁴	Total population with T2DM	26,110,573	15.9	
Sciannameo, 2020	EXSCEL ³⁴	Total population with T2DM	16,544	13.4	
Wittbrodt, 2018	EXSCEL ³⁴	Adults likely to have T2DM	24,062,453	47.2	
Summary	EXSCEL ³⁴				15.9 (14.7, 31.6)
Wittbrodt, 2018	FREEDOM-CVO ⁴³	Adults likely to have T2DM	23,941,512	15.5	

Sciannameo, 2020	HARMONY ⁴⁴	Total population with T2DM	10,208	9.5	
Wittbrodt, 2018	HARMONY ⁴⁴	Adults likely to have T2DM	24,062,453	8.0	
Abrahami, 2020	LEADER ³⁵	New users of liraglutide	159,660	1.2	
Arnold, 2017	LEADER ³⁵	GLP-1 RA users	5,249	6.0	
Arnold, 2017	LEADER ³⁵	Total population with T2DM	182,525	48.0	
Boye, 2018	LEADER ³⁵	Total population with T2DM	26,110,573	12.9	
Fan, 2020	LEADER ³⁵	Total population with T2DM	800	15.4	
Romera, 2022	LEADER ³⁵	Total population with T2DM	24,268	10.1	
Sciannameo, 2020	LEADER ³⁵	Total population with T2DM	10,061	9.4	
Webb, 2021	LEADER ³⁵	Total population with T2DM	33,118	13.3	
Wittbrodt, 2018	LEADER ³⁵	Adults likely to have T2DM	23,941,512	12.8	
Summary	LEADER ³⁵				12.8 (9.4, 13.3)
Sciannameo, 2020	PIONEER-6 ³⁶	Total population with T2DM	39,726	34.1	
Boye, 2018	REWIND ⁶	Total population with T2DM	26,110,573	42.6	
Romera, 2022	REWIND ⁶	Total population with T2DM	24,268	53.6	
Sciannameo, 2020	REWIND ⁶	Total population with T2DM	37,574	35.8	
Webb, 2021	REWIND ⁶	Total population with T2DM	33,118	44.4	
Wittbrodt, 2018	REWIND ⁶	Adults likely to have T2DM	23,941,512	22.4	
Summary	REWIND ⁶				42.6 (35.8, 44.4)
Boye, 2018	SUSTAIN-6 ³⁷	Total population with T2DM	26,110,573	13.0	
Romera, 2022	SUSTAIN-6 ³⁷	Total population with T2DM	24,268	10.4	
Sciannameo, 2020	SUSTAIN-6 ³⁷	Total population with T2DM	9,942	10.1	
Webb, 2021	SUSTAIN-6 ³⁷	Total population with T2DM	33,118	13.5	
Wittbrodt, 2018	SUSTAIN-6 ³⁷	Adults likely to have T2DM	23,941,512	11.8	
Summary	SUSTAIN-6 ³⁷				11.8 (10.4, 13.0)

Abbreviations: RCT: Randomized Controlled trial T2DM: Type 2 Diabetes Mellitus; SGLT-2: Sodium-Glucose Transport Protein 2;

GLP-1 RA: Glucagon-Like Peptide-1 Receptor Agonist

Figure Legend

Figure 1 PRISMA flow diagram of study inclusion for systematic review of replication of CVOTs using RWD in patients with T2DM

Supplementary Materials Titles

Table S1 PRISMA Checklist for systematic review of observational studies assessing generalizability of CVOTS

Table S2 SWiM Reporting Checklist for systematic review of observational studies assessing generalizability of CVOTS

Table S3 Search strategy for systematic review of observational studies assessing generalizability of CVOTS using OVID MEDLINE database

Table S4 Search strategy for systematic review of observational studies assessing generalizability of CVOTS using OVID EMBASE + EMBASE CLASSIC database

Table S5 Search strategy for systematic review of observational studies assessing generalizability of CVOTS using Cochrane CENTRAL database

Table S6 A list of completed cardiovascular outcome trials and their drug class and molecule

Table S7 Standardized differences computed for observational study compared to RCT hazard ratios

Figure S1 Distribution of percent eligible for inclusion in CVOT of SGLT-2 inhibitors in real-world populations by mean age (i) and by sex (ii)

Figure S2 Distribution of percent eligible for inclusion in CVOT of GLP-1 Receptor Agonists in real-world populations by mean age (i) and by sex (ii)

Figure S3 Standardized Mean Differences of patient characteristics between GLP-1 Receptor Agonist CVOTs and the study population from the RWD study from Sciannameo, 2020

Figure S4 Select standardized mean difference of patient characteristics for CVOT and RWD for the DECLARE trial

Figure S5 Select standardized mean difference of patient characteristics for CVOT and RWD for CANVAS trial

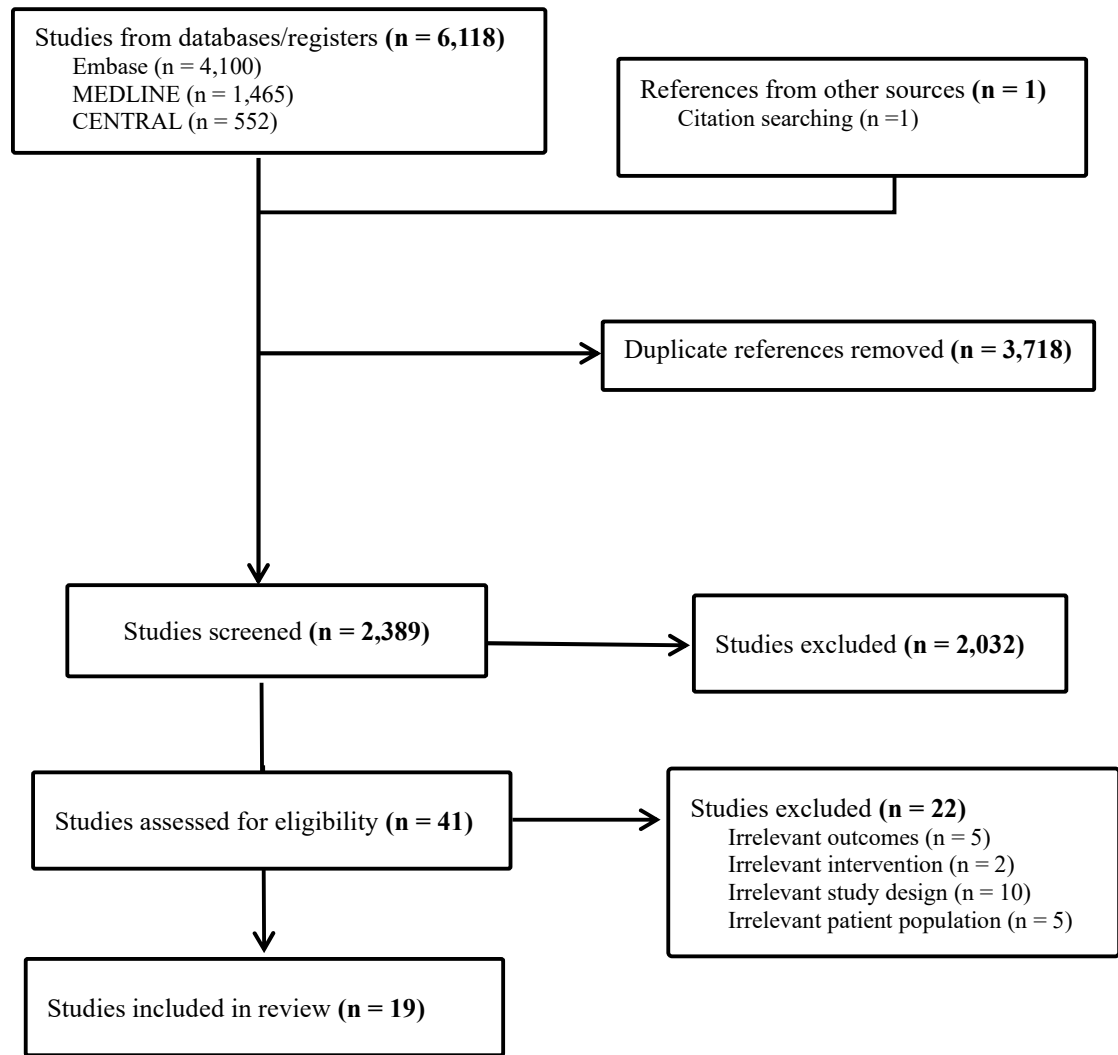


Figure 1