WILEY

REVIEW ARTICLE

Observational research on sodium glucose co-transporter-2 inhibitors: A real breakthrough?

Emanuel Raschi MD[1] | Elisabetta Poluzzi PharmD[1] | Gian Paolo Fadini MD[2] |
Giulio Marchesini MD[3] | Fabrizio De Ponti MD[1]

[1]Pharmacology Unit, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

[2]Department of Medicine, University of Padova, Padova, Italy

[3]Unit of Metabolic Diseases & Clinical Dietetics, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

Correspondence
Emanuel Raschi, MD, Alma Mater Studiorum - University of Bologna, Department of Medical and Surgical Sciences, Via Irnerio, 48, 40126 Bologna, Italy.
Email: emanuel.raschi@unibo.it

Funding information
The study was not funded.

Sodium glucose co-transporter-2 inhibitors have attracted the interest of the scientific community following the results from dedicated cardiovascular outcome trials, which demonstrated remarkable reduction in all-cause mortality and other cardiovascular (CV) endpoints with empagliflozin and canagliflozin. These impressive results raised further expectations on real world data from large observational cohort studies. They were designed to address the possible existence of a class effect, and the uncertainty on whether this benefit can be extended from secondary to primary CV prevention of patients with type 2 diabetes. In this review, we collated data from existing observational studies (including the celebrated CVD-REAL cohorts) and critically appraised results and methodological issues with the aim of providing clinical insight, including unsettled aspects, and proposing a research agenda for future investigations.

KEYWORDS
cardiovascular outcome trials, CVD-REAL, real world, safety, sodium glucose co-transporter 2 inhibitors

1 | INTRODUCTION: A TIMELY ACCRUAL OF REAL-WORLD DATA

We are witnessing an unprecedented paradigm shift in diabetology, both in drug development and clinical practice. In order to get marketing authorization, novel antidiabetic drugs have been required to demonstrate safety on major adverse cardiovascular events (MACE) in the so-called cardiovascular outcome trials (CVOTs). The 2008 FDA guidance had a major impact on the cardiovascular (CV) risk assessment of recently approved antidiabetic drugs, with a few classes showing significant reduction in CV mortality, including sodium glucose co-transporter-2 inhibitors (SGLT2-Is).[1] This is going to change the clinical mission of diabetologists from treat to target to treat to benefit and is capturing the interest of general practitioners and cardiologists,[2] with a number of recently conducted or ongoing CVOTs for various antidiabetic medications, including SGLT2-Is.[3]

Findings from the EMPA-REG OUTCOME trial (empagliflozin) and the CANVAS trial programme (canagliflozin) have demonstrated significantly (14%) reduced MACE (a composite of death from CV causes, non-fatal myocardial infarction, or non-fatal stroke), with a remarkable 32% reduction in all-cause mortality in EMPA-REG Outcome.[4,5] Significant renal benefits were also found in both trials, although outcome definitions were dissimilar, thus making direct comparisons not appropriate. Notwithstanding these impressive results, several unresolved issues arose in this medication class. In particular, two issues remain: (1) the possible existence of a class effect, and (2) the uncertainty on whether this benefit can be extended from secondary to primary CV prevention, that is, if it may apply to all patients with type 2 diabetes mellitus (T2DM).

Within this scenario, a number of observational studies on longitudinal healthcare databases were published outside the stringent environment of CVOTs in order to test the actual effectiveness in the so-called real-world evidence (RWE) area. While RWE indicates that the evidence is obtained from the analysis of real-world data (RWD), the term RWD is rather new, and there is considerable variability in its definition: in general, the term RWD is used to define data collected outside randomized controlled trial (RCT) settings, although other definitions diverge from this concept.[6] In the wake of this debate on terminology, we support the use of the term pharmaco-epidemiological research (PER) to describe the clinical setting of observational analysis.

While most opinions converge in considering PER a complementary source of evidence as compared to RCT, it is important to underline that PERs are definitely different from RCTs and have distinct inherent limitations: RCTs do not measure the benefit in clinical practice (low external validity), despite attempting to be pragmatic (e.g. by enrolling patients with T2DM at high CV risk), whereas PERs do not prove causal-effect relationships despite being pragmatically cogent (enrolling a broad and diverse population in clinical practice); thus, their integration offers substantial synergy for risk–benefit assessment and final determination of the place in therapy.[7,8]

Stakeholders are discussing if and how PERs can be a substitute for RCTs and be exploited to support regulatory decision-making: the challenging issue for researchers in increasing validity and reliance on PER (i.e. clinical transferability) is to avoid common flaws in design, balancing measured and unmeasured confounders, conducting sensitivity analyses, and clearly reporting methods and findings.[9]

In this review, we will appraise eight PER studies recently conducted through a cohort design on the CV benefit of SGLT2-Is in order to (1) provide a summary of results with relevant clinical implications, and (2) propose a research agenda for unsettled issues.

2 | OVERVIEW OF RESULTS

2.1 | CVOTs: beyond cardiovascular protection

Impressive findings emerged for CV outcomes from the two CVOTs, including potential renal protective effects. In particular, the EMPA-REG Outcome study showed 38% reduction (hazard ratio [HR] = 0.62; 95% confidence interval [95% CI] = 0.54–0.72) in progression to macro-albuminuria and 39% reduction in incident or worsening nephropathy (0.61; 0.53–0.70), whereas the CANVAS-Renal highlighted 27% reduction in albuminuria progression (0.73; 0.67–0.79) and 40% reduction in the composite renal outcome (0.60; 0.47–0.77) (defined as 40% reduction in the estimated glomerular filtration rate [eGFR], the need for renal replacement therapy, or death from renal causes). However, the marked reduction in CV and all-cause mortality reported in EMPA REG Outcome was not noticed in CANVAS; this difference may be at least partially explained by the characteristics of the population (almost 35% of the CANVAS population was in primary prevention, whereas in EMPA REG Outcome nearly all patients were in secondary prevention).

2.2 | Observational studies: confirming and extending CV benefit in the real world

Observational PER provides an extraordinary opportunity to translate results from CVOTs to clinical routine, especially if timely and accurately performed. The Comparative Effectiveness of Cardiovascular Outcomes in New Users of SGLT-2 Inhibitors (CVD-REAL) study was the first multinational observational investigation analysing 309 056 patients newly initiated on SGLT2-Is versus other glucose-lowering drugs. Notably, these patients were not at increased CV risk (only 13% had established CV disease). In this propensity-matched PER using data from medical claims, primary care/hospital records, and

national registries in six countries (the United States, Germany, Sweden, Norway, Denmark and the UK), the authors assessed the risks of hospitalization for heart failure (HHF), death, and the combined endpoint of heart failure hospitalization or death, which were all lower in patients receiving an SGLT2-I, without significant heterogeneity across countries and with consistent results among multiple sensitivity analyses.[10] In particular, the pooled analysis of all-cause death using patient data from five of the countries (this information was not available for the German registry) showed a 51% reduction favouring patients with SGLT2-Is (0.49; 0.41–0.57; without significant heterogeneity among SGLT2-Is).

Overall, the CVD-REAL study did not examine other hard endpoints, such as MACE (cardiovascular death, myocardial infarction and stoke), which were not available in all cohorts. The CVD-REAL Nordic sub-study examined CV mortality in patients from Sweden, Denmark and Norway (where detailed national registries allowed assessment of the cause of death). Among SGLT2-Is, dapagliflozin accounted for 94% of the total SGLT2-I exposure time. Over a median follow-up of 0.9 years, the study showed that CV mortality was reduced by 47% in patients newly initiated on an SGLT2-I compared with other glucose-lowering medications (0.53; 0.40–0.70).[11] The CVD-REAL Nordic was the only study examining CV mortality, which was also reduced independently of CV disease at baseline. A further sub-analysis of the CVD-REAL Nordic database found that patients newly initiated on dapagliflozin experienced a lower rate of CV events, HHF and all-cause mortality compared with patients newly initiated on DPP4 inhibitors.[12] It appears that the CV benefit was mainly driven by Danish data, whereas the Norwegian cohort was not associated with statistically significant reduction of CV mortality and MACE.

The recent CVD-REAL 2 was designed to cover a large proportion of patients residing outside the United States and Europe and to address uncertain aspects such as the risk of stroke (being more common in Asia and previously found to be non-significantly increased in the EMPA REG Outcome trial). Therefore, this cohort enrolled cases in six countries from Asia Pacific, the Middle East and North America (Australia, Canada, Israel, Japan, Singapore and South Korea). Results confirmed the findings of the former CVD-REAL, with a lower risk of death (0.51; 0.37–0.70), HHF (0.64; 0.50–0.82), and also a reduced incidence of myocardial infarction (0.81; 0.74–0.88) and stroke (0.68; 0.55–0.84) in patients taking SGLT2-Is versus other glucose-lowering medications.[13]

In parallel with CVD-REAL cohorts, additional population-based studies investigated CV outcomes with SGLT2-Is from US (vs. non-gliflozin agents),[14,15] UK Health Improvement Network (vs. other glucose-lowering drugs)[16] and Swedish (DPP4-Is or SGLT2-Is compared with insulin)[17] databases. Although definitions of outcomes as well as patients' characteristics vary across different studies, all consistently highlighted a significantly reduced risk of MACE and other CV endpoints, except for the US cohort on canagliflozin (Table 1). In particular, all-cause mortality was especially reduced (roughly 40% risk reduction) in US, UK and Swedish cohorts, whereas MACE and HHF decreased by a maximum of 33% in the US study. Overall, the magnitude of risk reduction appears comparable to that found in the CVD-REAL studies: this consistency is not completely surprising,

TABLE 1 Summary of characteristics and key CV results of published population-based cohort studies on SGLT2-Is. Unless specified otherwise, outcome results are expressed in terms of hazard ratios with relevant 95% confidence intervals in parentheses

Study	Population	Intervention	Comparator	Main CV outcome(s)	Key results	Main findings on subgroup/sensitivity analyses	Notes
UK cohort[16]	T2DM with 1-yr registration period (electronic medical records from GPs); 20% with history of CVD	Dapagliflozin (4444 patients)	Not exposed to an SGLT2-I (17 680)	ACM, incident CVD (MI and ischaemic heart disease, stroke or TIA, and heart failure or left ventricular dysfunction in low-risk cohort)	ACM: adjIRR = 0.50 (0.33–0.75) Incident CVD: adjIRR = 0.89 (0.61–1.30)	In the low-risk population, adjIRR = 0.44 (0.25–0.78) Sensitivity analyses (DPP4 Is used as a negative control, and high CV risk) consistent in magnitude and direction (except when excluding patients suggestive of T1DM)	Immortal time bias was accounted for Different covariates used as confounders, including BMI Median follow-up 0.9 yr
Swedish cohort[17]	Patients with first-time prescription of either DPP4 or SGLT2-Is, or insulin recorded in national registers; 20% with microvascular disease (matched cohort)	DPP4-Is or SGLT2-Is (21 758 patients)	Insulin (10 979)[a]	ACM, fatal and non-fatal CVM (MI, ischaemic stroke, unstable angina pectoris, heart failure or CVM)	ACM: 0.58 (0.52–0.65) Fatal and non-fatal CVM: 0.84 (0.73–0.97).	In the subgroup of patients with established CV risk at baseline, dapagliflozin reduced CVM (HR = 0.47, 95% CI = 0.24–0.93), but no significance in the larger cohort without CV risk at baseline Data on dapagliflozin slightly more favourable for fatal/ non-fatal CVM	Potential for immortal time bias and time-lag bias[18–21] Median follow-up 1.5 yr (matched cohort)
US cohort[15]	T2DM (US commercial healthcare database); 52% with hypertension and 48% with hyperlipidemia	Canagliflozin (28 149 unique initiators)	Non-gliflozin agents (DPP4-I, GLP1-RA, or a sulphonylurea)	HHF, MACE (admitted to hospital for acute MI, ischaemic stroke, or hemorrhagic stroke), ACM, stroke, unstable angina, MI	HHF (vs. DPP4-Is): 0.70 (0.54–0.92) MACE (vs. DPP4-Is): 0.89 (0.68–1.17) ACM (vs. DPP4-Is): 0.66 (0.25–1.74)	Similar data in the three cohorts (vs. DPP4-Is, GLP1-RAs or sulphonylurea), with higher benefit for the third cohort Consistency in sensitivity analyses (adjusted for baseline HbA1c level, stratified by previous CV events), except for HHF (benefit emerged only in patients with previous HHF)	Propensity score matching and several sensitivity analyses to account for unmeasured confounders Median follow-up 0.6 yr (matched cohort)
EASEL[14]	T2DM with established CV disease with ≥1 yr of observation before the index date (US Military Health System); 31% with chronic complications	New user of SGLT2-Is (12629)[a]	New user of non-SGLT2-Is	ACM, HHF, MACE (ACM, non-fatal MI, and non-fatal stroke), MI, stroke, BKA	ACM: 0.57 (0.49–0.66) HHF: 0.57 (0.45–0.73) MACE: 0.67 (0.60–0.75) BKA: 1.99 (1.12–3.51)	Consistency of CV endpoints in sensitivity analyses (removing individually and collectively patients receiving insulin, sulphonylureas, and thiazolidinediones, having dementia) and subgroup analyses (sex, age, recent insulin/GLP-1 RA use, history of HF, recent HHF, CVD type, renal disease) Some differences emerged only for BKA risk (higher in males, without previous use of GLP-1 RA and no renal disease)	Propensity score matching with >850 variables Immortal time bias cannot be excluded Median follow-up 1.6 yr (intention-to-threat cohort)

(Continues)

TABLE 1 (Continued)

Study	Population	Intervention	Comparator	Main CV outcome(s)	Key results	Main findings on subgroup/sensitivity analyses	Notes
CVD-REAL[10]	T2DM with >1 yr data history in the database before the index date (healthcare records from 6 countries); 13% with history of CVD	New user of SGLT2-Is (as initial or add-on therapy) [max 154 528 patients[a]]	New user of any other oral or injectable glucose-lowering medication	ACM, HHF, MACE[b]	ACM: 0.49 (0.41–0.57) HHF: 0.61 (0.51–0.73) MACE[b]: 0.54 (0.48–0.60)	Consistency in sensitivity analyses (within each country, United States vs. Europe, stepwise removal of specific antidiabetic classes, in- and outpatient hospital visit for HF), including published sub-analysis in patients with/without previous CV disease[22] Non-significant reduced risk in UK cohort Modestly lower risk of MI and stroke in sub-study[23]	Propensity score matching with several variables Potential immortal time bias[18-21] Median follow-up 0.6 yr
CVD-REAL Nordic[11]	T2DM with >1 yr data history in the database before the index date (healthcare records from Denmark, Norway, Sweden); 25% with history of CVD	New user of SGLT2-Is (different definitions among countries) [22830[a]]	New user of any other oral or injectable glucose-lowering medication	CVM, MACE[c], HHF, ACM, atrial fibrillation	CVM: 0.53 (0.40–0.71) MACE[c]: 0.78 (0.69–0.87) HHF: 0.70 (0.61–0.81) ACM: 0.51 (0.45–0.58)	Data confirmed in the sub-study comparing dapagliflozin with DPP4-Is[12] Some variability in subgroup analyses, but reduced risk of MACE only in patients with CV disease at baseline Neutral association for CVM and MACE in aged <65 Non-significant reduced risk in Norwegian cohort	Propensity score matching with several variables Potential immortal time bias[18-21] Median follow-up 0.9 yr
CVD-REAL 2[13]	T2DM with >1 yr data history in the database before the index date (healthcare records from 6 non-EU countries); 27% with history of CVD	New user of SGLT2-Is (as initial or add-on therapy) [235 064 patients[a]]	New user of any other oral or injectable glucose-lowering medication	ACM, HHF, MACE (ACM or HHF, MI, and stroke), MI, stroke	ACM: 0.51 (0.37–0.70) HHF: 0.64 (0.50–0.82) MACE: 0.60 (0.47–0.76)	Neutral association for MI in all countries (except Korea) No differences according to the CV risk at baseline Consistency in sensitivity analyses, with ACM/HHF/MACE attenuated but statistically significant in first new user design No meaningful interactions in subgroup analyses	Propensity score matching with several variables Potential immortal time bias accounted for in sensitivity analyses (first new user design) Median follow-up 1.1 yr
OBSERVE-4D[24]	T2DM (healthcare records from 4 US administrative claims databases); approximately 30% with CVD	New user of canagliflozin (142 800 patients) or other SGLT2-Is (110 897 patients)	New user of non SGLT2-Is (two cohorts), other SGLT2-Is (three cohorts)	HHF, BKA	HHF: 0.82 (0.75–0.89) [intention-to-treat, canagliflozin vs. select non-SGLT2-Is] BKA: 1.01 (0.81–1.25) [intention-to-treat, canagliflozin vs. select non-SGLT2-Is]	No differences when comparing canagliflozin with other SGLT2-Is Incidence of HHF/BKA in patients with established CV disease was approximately 2-fold higher compared to the overall cohorts (and new users of canagliflozin had lower baseline prevalence of renal impairment/urinary tract disease	Full access to source-specific estimates through interactive web-based application Propensity score matching with several variables Immortal time bias cannot be excluded

TABLE 1 (Continued)

Study	Population	Intervention	Comparator	Main CV outcome(s)	Key results	Main findings on subgroup/ sensitivity analyses	Notes
						Consistency in sensitivity analyses (different prior exposure assumptions, time-at-risk windows, event types, propensity score adjustment strategies) and sub-analysis in patients with established CV disease	Median follow-up 1.5 yr (intention-to-treat cohort, highest value among databases)

ACM, all-cause mortality; BKA, atraumatic below-knee lower extremity amputation; CV, cardiovascular; CVD, cardiovascular disease; CVM, cardiovascular mortality; DPP4-I, dipeptidyl peptidase-4 inhibitor; GLP-1RA, glucagon-like peptide-1 receptor agonist; GPs, general practitioners; HF, heart failure; HHF, hospitalization for heart failure; MACE, major adverse cardiovascular events; MI, myocardial infarction; TIA, transient ischemic attack; T1DM, type 1 diabetes mellitus; T2DM, type 2 diabetes mellitus; IRR, incidence rate ratio.
a In propensity matched analyses.
b Composite of HF and death.
c CVM, main diagnosis of MI, main diagnosis of ischaemic or haemorrhagic stroke.

considering that US data in the CVD-REAL study largely contributed to study findings, and that the Nordic sub-study exploited data from Sweden.

The latest OBSERVE-4D directly compared SGLT2-Is for safety and effectiveness: statistically significant differences did not emerge among gliflozins in terms of HHF and atraumatic below-knee lower extremity amputations (BKAs), but meta-analytic estimates showed reduced risk of HHF when comparing canagliflozin or other SGLT2-Is versus non-SGLT2-Is, with no increased risk of BKAs.[24]

The strengths of these real-world cohorts are mainly represented by the catchment area (virtually worldwide with millions of patients) accounting for the high heterogeneity seen in the diabetic population. Moreover, the propensity-score matching adopted in many studies considered hundreds of covariates to minimize the effect of known confounders. Notably, the results were directionally consistent within and across databases, countries, patient subgroups (including those w/wo CV disease), type of SGLT2-I and comparators (insulin, sulphonylureas, thiazolidinediones, incretin-based therapies). Notably, only in the EASEL and CVD-REAL Nordic studies were age and sex tested as subgroups: significant differences emerged only for CV mortality and MACE in patients <65 years (neutral association). In other cohorts, age and sex were only used as matching variables in the propensity score comparison and as covariates to adjust HRs in multivariate survival analyses.

3 | A CRITICAL APPRAISAL OF METHODOLOGICAL ISSUES

3.1 | Administrative claims databases: validity first

Limitations of PER cannot be disregarded and typically stem from the observational nature of the data; outcome data were not collected in a systematic fashion and events of interest were not adjudicated by an independent committee; in routine clinical practice, the primary diagnoses of admission or discharge are based on the opinion of physicians, and may be influenced by reimbursement policies.

Both diabetes and CV diagnoses in claims administrative databases are challenging and the positive predictive values of the adopted algorithm may vary among data sources. Therefore, before pooling data through meta-analysis (MA), validity remains a crucial aspect of study design and conception.[25–28] In particular, validation of the diagnoses is critical in the diabetes area, especially when in multinational epidemiological studies,[29] because the lack of clinical data may compromise optimal algorithms for patients' and outcomes' identification. In addition, administrative claims databases do not allow clinical variables associated with CV risk to be fully captured (or they are only available for a subset of the cohort), such as HbA1c, body mass index, lipid levels, and renal function based on creatinine values (instead of being codified as chronic kidney disease stage).

Although the observational studies on SGLT2-Is exploited validated healthcare databases for PER, relevant peculiarities in data registration should not be disregarded. Notably, mortality rate with SGLT2-Is in the CVD-REAL study is much lower than that in the Swedish study (5.2 vs. 25.6 per 1000 person-years, respectively), and

the mortality rates in the CVD-REAL study are surprisingly highly variable (from the US cohort's rate of 3.1/1000 per year to the Danish cohort's 18.7/1000 per year), thus potentially reflecting some incompleteness in recording death information in the US database.

3.2 | Data analysis: the threat of bias and unmeasured confounders

In the recent past, several methodological challenges have been described in observational studies assessing glucose-lowering medications and CV outcomes, with the majority of publications not adequately addressing appropriate minimization strategies.[30,31] Causal inference from observational research is challenging to be claimed with certainty, and many associations are likely to be the result of residual confounders. In general, for cohort studies, the threshold of potential interest (confidence) for decreased risk has been suggested to start at 0.33 (effect estimates), whereas a risk estimate of 0.5 should be considered in the so-called zone of potential bias.[32] The majority of data from cohort studies on SGLT2-Is lies in this uncertain zone.

Published PERs on SGLT2-Is adopt a propensity score-matched approach, which is increasingly being used in observational studies of cardiovascular interventions as an alternative to conventional covariate adjustment. A propensity score is defined as the probability of a patient being assigned to an intervention, given a set of covariates and all patient characteristics being summarized into a single covariate, thus reducing (although not eliminating) the potential for overfitting and correcting for selection biases and potential confounding. While this technique provides an excellent balance of covariates in most circumstances, it is less precise, and some patients may remain unmatched and hence excluded from the analysis.[33] Moreover, it balances the distribution of characteristics between the compared exposure groups on average, thus there may be pairs that are discordant on special characteristics; finally, there is no clear understanding of the relative performance of matching strategies in subgroup analyses.[34]

The propensity score cannot account for unmeasured confounders such as the healthier behaviour of patients prescribed a new medication, or the profiles of treating clinicians and their early confidence in a new effective drug. In this regard, quasi-experimental methods using instrumental variables (e.g. local variations in prescribing rates) have been proposed to account for selection bias.[35] A recent cohort study, using a US commercial insurance database, achieved balance in clinical variables, thus minimizing the likelihood of residual confounding (provided that an appropriate comparator, new-user design and 1:1 propensity score matching on many proxies of diabetes severity and duration are considered).[36]

Moreover, there are at least three major biases that may affect study results. First, the channelling bias (i.e. the possibility that SGLT2-Is are preferentially prescribed to selected patients) should be considered. This bias may be differently interpreted, also depending on how new drugs are perceived by diabetologists versus general practitioners. For instance, specialized clinicians might be more prone to prescribe new drugs such as SGLT2-Is, although the anticipated risks of infections, hypotension and ketoacidosis may cause a

selection bias towards healthier patients. As a matter of fact, patients receiving an SGLT2-I in the various cohorts prior to propensity matching were younger, had a longer diabetes duration, lower CV burden, but higher microvascular complications. A similar type of channelling emerged in the recent DARWIN-T2D (DApagliflozin Real World evIdeNce in type 2 diabetes) multicentre Italian study,[37] in which dapagliflozin was initially prescribed to difficult-to-treat patients (i.e. younger, with a longer disease duration and suboptimal glycaemic control), although significant benefits were achieved with regard to glucose control, body weight and blood pressure.

The CVD-REAL and Swedish study were also criticized because of the potential existence of the so-called immortal time bias and time-lag bias, which can exaggerate study findings.[18–21] In the CVD-REAL study, the former was due to incorrect attribution of the time between the first antidiabetic prescription and the first SGLT2-I prescription, resulting in potential overestimation of the mortality rate among users of other antidiabetics. The latter emerged when comparing patients receiving second-line drugs with those under third-line therapies; these subjects may be at different stages of diabetes severity with confounding by disease duration. The CVD-REAL 2 study attempted to minimize the likelihood of immortal time bias by performing a first new user design as a conservative sensitivity analysis; notably, the magnitude of the associations for the different CV outcomes was attenuated especially for myocardial infarction (albeit remaining statistically significant) with heterogeneity in point estimates among countries, with Korea mainly contributing to pooled data.

4 | CLINICAL INSIGHT INTO OBSERVATIONAL RESEARCH

Overall, the different population-based studies, and especially CVD-REAL cohorts, confirmed the efficacy emerging from the EMPA-REG Outcome and CANVAS programme in terms of MACE, all-cause mortality, single CV outcomes and HHF reduction, and also extended the CV benefit to myocardial infarction and stroke. Notably, the recently revised US guideline by the American Diabetes Association considers CV risk when choosing the second-line agent after metformin in patients with T2DM: in this scenario, the guideline specifically recommends an agent proven to reduce MACE and CV mortality in patients with established atherosclerotic CV diseases (currently empagliflozin and liraglutide), after considering drug-specific and patient factors (level of the evidence A).[38]

While the results are outstanding from a statistical viewpoint and the relative risk reduction is noteworthy, relevant implications for clinicians may not be obvious and substantiated, especially when communicating and sharing data with patients. In this regard, incidence rates and, most importantly, the number needed to treat to benefit (NNTB), as well the number needed to treat to harm (NNTH), are epidemiological measures of risk that best describe the effect observed in the individual patient.

In Table 2, we report the incidence rates (in published population-based studies) and extracted crude numbers of events to compute NNTB/NNTH. Different considerations can be made: (1) the clinical

TABLE 2 Incidence rates and NNTB/NNTH of the different CV outcomes (exposure to SGLT2-Is), including risk of amputations

Study	MACE		All-cause mortality		HHF		BKAs	
	Incidence rate[a]	NNTB	Incidence rate[a]	NNTB	Incidence rate[a]	NNTB	Incidence rate[a]	NNTH
UK cohort (dapagliflozin vs. non-SGLT2-Is)[16]	1.338[b]	NC	0.527[b]	259	/	/	/	/
Swedish cohort (dapagliflozin only vs. insulin)[17]	1.68	95	0.98	66	/	/	/	/
US cohort (canagliflozin vs. DPP4-Is)[15]	0.99	2523	0.07	5889	0.89	535	/	/
EASEL (SGLT2-Is vs. non-SGLT2-Is)[14]	2.31	55	1.29	62	0.51	158	0.17	743
CVD-REAL (SGLT2-Is vs. non-SGLT2-Is)[10]	0.89	166	0.52	211	0.36	681	/	/
CVD-REAL Nordic (SGLT2-Is vs. non-SGLT2-is)[11]	1.64	206	1.05	76	0.98	220	/	/
CVD-REAL Nordic (dapagliflozin vs. DDP4-Is)[12]	1.86	187	1.03	108	0.99	169	/	/
CVD-REAL 2 (SGLT2-Is vs. non-SGLT2-Is)[13]	1.91	151	0.80	173	1.23	333	/	/
OBSERVE-4D (canagliflozin vs. select non-SGLT2-Is)[24]	/	/	/	/	1.18	58	0.45	NC
OBSERVE-4D (SGLT2-Is vs. select non-SGLT2-Is)[24]	/	/	/	/	0.96	104	0.42	NC

BKAs, atraumatic below-knee lower extremity amputations; DPP4-Is, dipeptidyl peptidase-4 inhibitors; NNTB, number needed to treat to benefit; NNTH, number needed to treat to harm. /: not available. NC: not calculated because no statistically significant difference emerged between exposed and unexposed group.

[a] Data expressed × 100 person-years.

[b] Data for low-risk population (see Table 1 for the definition of MACE, which in this case was defined as incident cardiovascular disease).

relevance of CV benefit in terms of HHF is not as large as expected from the risk reduction estimates when compared to MACE and all-cause mortality; (2) the lowest (outstanding) NNTBs came from the EASEL study; this may be interpreted taking into consideration that the population was a high-risk cohort; (3) the CVD-REAL studies provided similar findings, with the Nordic cohort showing the best estimates, especially the sub-study on dapagliflozin; the reasons are uncertain although the comparison with DPP4-Is may play a role: all published CVOTs on DPP4-Is demonstrate non-inferiority versus placebo, and the relative effect on the risk of HHF remains uncertain, with ongoing controversy over the increased risk for saxagliptin[39-41]; (4) only the UK cohort in the THIN database showed no significant risk reduction for MACE (expressed as incident CVD); (5) the highest NNTBs emerged for canagliflozin in the US cohort compared with DPP4-Is; and (6) overall, it appears that the NNTBs are roughly comparable with those derived from the UKPDS follow-up[42] (in the metformin group comparing intensive vs. conventional therapy, the NNTB for the endpoint death from any cause was 139).

5 | OBSERVATIONAL FINDINGS: WHAT'S MISSING?

5.1 | The knowledge gap on effectiveness

The research agenda still faces some unsettled issues concerning prescribing and optimal management of patients with T2DM, namely: what knowledge gaps are still outstanding after CVD-REAL studies? What type of evidence is required to best meet this clinical need?

In terms of effectiveness, there is interest in verifying whether pharmacological modulation with SGLT2-Is could be part of the armamentarium for treating patients with heart failure, with and without diabetes.[43] Intriguingly, different large RCTs are ongoing (EMPEROR-Preserved, EMPEROR-Reduced, DAPA-HF), which do not include diagnosis of T2DM in the inclusion criteria.[44] Clinicians might also wonder whether or not clinically significant differences exist among

SGLT2-Is through head-to-head comparative research. As discussed below, it appears that the existence of specific effects is a matter of safety rather than an efficacy issue, with the OBSERVE-4D MA of four US databases showing no difference in the estimates of HHF and BKAs.[24] In the modern era, cost-effectiveness analyses are also awaited to actually quantify patient's benefit in terms of Quality Adjusted Life Years, and define the therapeutic gain of an integrated management of patients with T2DM.

A fundamental issue regarding the benefit on renal outcomes observed in CVOTs must be confirmed in clinical practice, where different factors may modify and impact the effect of antidiabetic drugs. None of the CVD-REAL studies and other cohorts have addressed this issue. Data from CVOTs, including a prespecified analysis of renal outcomes in the EMPA-REG Outcome trial, are promising in terms of long-term cardio-renal protection: empagliflozin was associated with a slower progression of kidney disease and lower rates of clinically relevant renal events.[45] However, subjects with chronic kidney disease (eGFR <30 mL/min/1.73 m^2) were excluded from both the CANVAS and EMPA-REG Outcome trials, with relevant contraindication in patients with moderate to severe renal impairment.[46] Conversely, scattered reports suggest that there might be a risk for short-term acute kidney injury, usually in the form of a reversible reduction of GFR, but may occasionally be fatal or require renal replacement therapy,[47] with potential differences among compounds.[48] Warnings of an increased risk of early-onset kidney injury were therefore strengthened by the FDA on the labels of SGLT2-Is, with the latest postmarketing data from the FDA Adverse Event Reporting System confirming a potential signal for acute kidney injury with all approved SGLT2-Is.[49] This effect may be attributed to different mechanisms, including natriuresis and osmotic diuresis-induced hypovolemia, as well as renal haemodynamic effects, mediated by adenosine and increased glucagon, resulting in vasoconstriction of afferent arterioles. Drugs causing vasodilation of the efferent arterioles (namely renin-angiotensin-aldosterone system blockers) might synergistically act with SGLT2-Is in acute eGFR reduction, which may become clinically relevant in subjects with underlying extracellular volume depletion,

that is, those treated with diuretics or receiving non-steroidal anti-inflammatory drugs, giving rise to a new pattern of the clinical entity known as triple whammy.

5.2 | Debated safety issues

There are different uncertainties as regards the risk of BKAs, bone fractures, infections and ketoacidosis, which were not scrutinized in CVD-REAL studies. Genital and urinary tract infections were associated with all members of the class and predictable on the basis of the mechanism of action; although the true rate in clinical practice is still uncertain, the high reporting frequency in the postmarketing phase calls for real-time pharmacovigilance monitoring.[50] The overall evidence suggests that SGLT2-Is increase the risk of genital mycotic infections by four to five times according to the largest MA,[51]

although they are usually mild to moderate, can be adequately treated with standard medical therapy, and drug discontinuation is not required. Conversely, gathered evidence is not consistent for an increased risk of urinary infections, with MAs showing mixed results.[52] However, clinicians should be aware that severe infections have been described, including pyelonephritis and urosepsis, especially in patients with urinary tract outlet obstruction.

The risk of diabetic ketoacidosis (DKA) was publicly disseminated by the FDA, which released a warning in May 2015 based on 20 case reports (both in type 1 diabetes mellitus and T2DM patients), and by the European Medicines Agency in June 2015, which identified a total of 147 cases of DKA in patients treated with SGLT2-Is. A recent literature review analysed RCTs, cohort studies, case reports and pharmacovigilance database studies[53]: DKA incidence was less than 1/1000
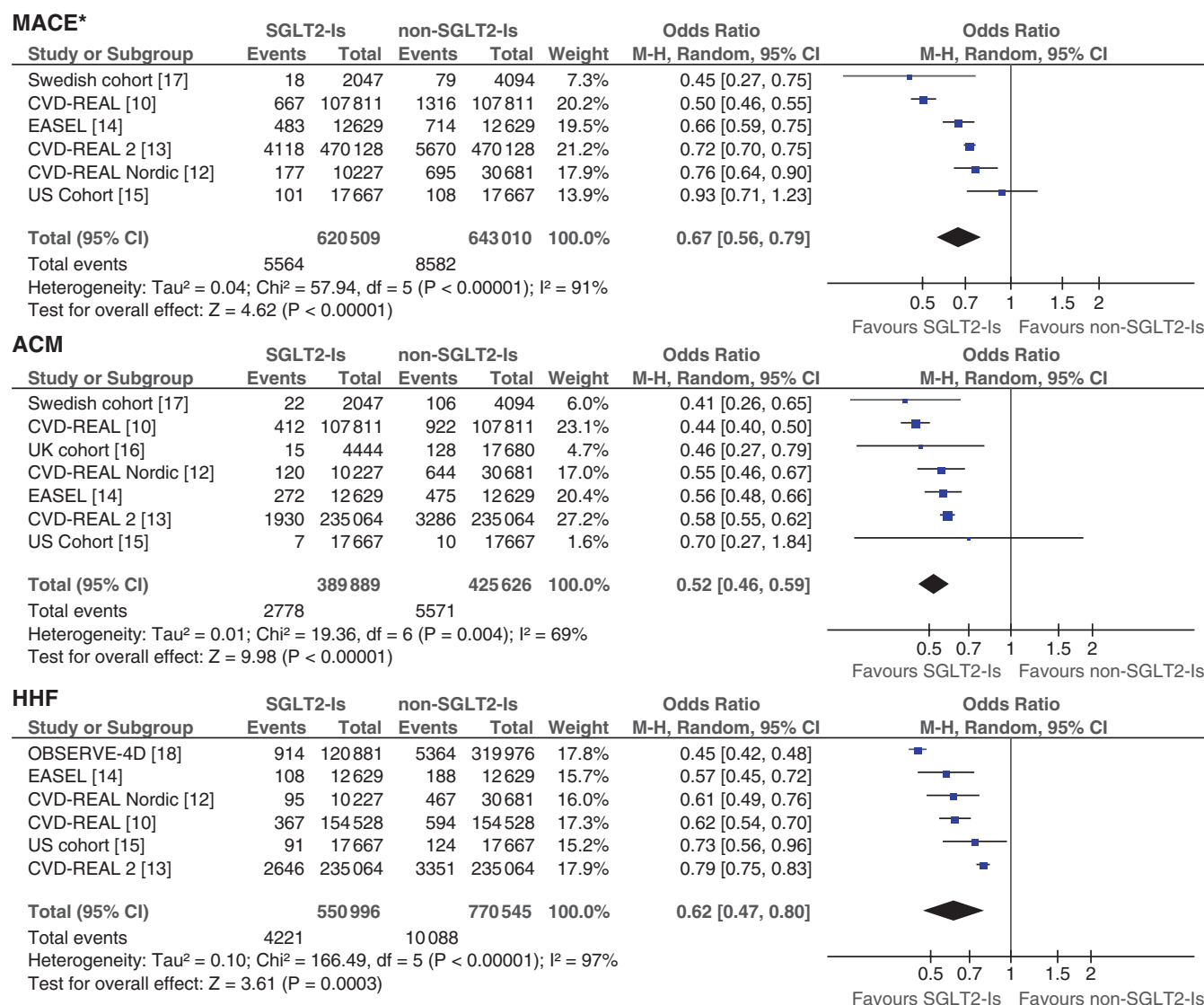
FIGURE 1 Forest plots comparing SGLT2-Is with non-SGLT2-Is for key cardiovascular outcomes. *MACE, major adverse cardiovascular event (this definition may vary among studies); ACM, all-cause mortality; HHF, hospitalization for heart failure. Please refer to individual studies for details. Meta-analysis was performed through RevMan 5.3. Crude numbers were extracted from individual studies and computed using the Mantel–Haenszel method (random effect model); therefore, odds ratios of the individual studies may differ compared with published estimates. Where available, data on intention-to-treat population at low cardiovascular risk were used. Comparator (non-SGLT2-Is) can be represented by all other antidiabetics, insulin or DPP4, depending on studies. The US cohort and OBSERVE-4D study were the only ones on canagliflozin. Data on the CVD-REAL Nordic and Swedish cohort studies are derived from the dapagliflozin dataset

in RCTs (in line with data from an MA[54]) and 1.6/1000 person-years in cohort studies. In case reports and in pharmacovigilance databases, duration of SGLT2-I treatment before DKA onset was extremely variable; overall, DKA is a rare adverse event during SGLT2-I therapy, with fatal episodes representing 1.6% of all reported cases.[55] The latest cohort investigations compared SGLT2-Is with DPP4-Is and found an increased risk with SGLT2-I exposure,[56] especially in patients with diabetic microvascular complications and in those taking diuretics.[57]

Amputations (BKAs) and fractures emerged only for canagliflozin in the CANVAS programme and their biological plausibility is only partially understood. Although a detrimental effect of SGLT2-Is on bone metabolism is biologically possible, studies of bone mineral density are inconclusive so far and the available evidence is not sufficient to determine the existence of a class effect.[58] A recent study of the UK Health Improvement Network database failed to detect an increased risk of treatment-emergent (fragility or any type) fractures in patients initiating dapagliflozin.[59]

The risk of BKAs represents a very relevant current issue. Apart from RCTs, three US observational cohort studies[14,24,60] and two disproportionality analyses of international spontaneous reporting systems, namely, the FDA Adverse Event Reporting System (FAERS) and the WHO Vigibase,[61,62] provided conflicting results as to whether or not this adverse effect concerns all SGLT2-Is.[63] Another critical issue is the mechanistic basis of this side effect,[64] and the multifactorial role of disease and comorbidities. The mechanistic basis is unclear: while gliflozins are presented as selective inhibitors of SGLT, canagliflozin causes a substantial inhibition of other SGLTs (including SGLT1, SGLT3 and SGLT6), but the biological effect is still uncertain.[65] The drivers of amputation in patients with diabetes mellitus are particularly complex and likely multifactorial in many cases, including susceptibility to trauma related to neuropathy, chronic ischaemia because of macrovascular and microvascular dysfunction, infection, and impaired wound healing.

An interesting observation from a US cohort[14] is the possible modification by concomitant use of GLP-1 receptor agonists, potentially mitigating the amputation risk associated with SGLT2-Is. These observations corroborated the hypothesis that specific drug combinations may be associated with a lower rate of adverse events. For instance, pharmacovigilance analyses of the FAERS database concluded that DPP4 inhibitors also reduce the frequency of genital infections associated with SGLT2-Is[66] whereas, conversely, the rate of HF associated with DPP4 inhibitors was moderated by combined treatment with SGLT2-Is.[67] This body of evidence adds further expectation to the results from the DECLARE-TIMI 58 trial, a CVOT similar to CVD-REAL cohorts, since it enrolls patients with and without increased CV risk. The key distinguishing features of this trial are its large size, duration of follow-up (~4 years), a well-characterized subgroup with peripheral artery disease, prospective ascertainment of amputation events and their aetiologies, and nested biosample substudies, thus offering the opportunity to define whether dapagliflozin increases amputation risk, a finding that would support a class effect or, if negative, would suggest an off-target effect specific to canagliflozin.

6 | OBSERVATIONAL RESEARCH: WHAT IS NEXT?

6.1 | Avoiding redundant meta-analyses

From a research perspective, systematic reviews with MA are traditionally performed following the availability of evidence from RCTs and PERs, although the pressure for publication and the general perception of a free-from-bias evidence-based study may generate massive production of unnecessary (redundant) research.[68–70]

The case of SGLT2-Is is not an exception: several MAs of RCTs are accruing,[71] with expected overlapping conclusions on CV benefit (especially when also pooling data from CVOTs),[41] whereas, to the best of our knowledge, no MA has specifically been carried out on observational studies for SGLT2-Is.

MAs are strongly encouraged when individual studies are conflicting, estimates are large and uncertain, or there are some topics not addressed by individual studies. Therefore, we do not believe that an MA of observational studies on CV outcomes *per se* would add clinically important evidence to the existing SGLT2-Is data: cohort studies provided consistent data on CV risk reduction for almost all endpoints (the directions of the estimates are reliable and robust across and

TABLE 3 Summary of findings from the systematic review with Bayesian hierarchical network meta-analysis by Zheng and colleagues[74]

Feature	Details
Population	T2DM
Intervention	SGLT2-Is
Comparator	GLP-1 RAs, DPP4-Is, placebo or no treatment
CV outcome(s)	CVM, ACM, HF events, MI and unstable angina
Study	Network MA of 236 RCTs ≥12 weeks
Key results (efficacy)	CVM = SGLT2-Is vs. DPP4-Is: 0.79 (0.66–0.94); SGLT2-Is vs. GLP-1 RAs: 0.93 (0.78–1.10) ACM = SGLT2-Is vs. DPP4-Is: 0.78 (0.68–0.90), SGLT2-Is vs. GLP-1 RAs: 0.91 (0.79–1.04)
Key results (safety)	No difference between drug classes for any or major hypoglycaemia. SGLT2-Is were associated with reduction in serious adverse events compared to all other controls, whereas GLP-1 RAs were associated with increased risk of adverse events leading to trial withdrawal compared with the control groups
Risk of bias	Risk of attrition bias in 25% of studies. Two out of studies comparing SGLT2Is vs. DPP4-Is were rated with unclear risk of bias for at least two domains
Notes	Low heterogeneity, consistency of results across analyses (frequentist and Bayesian approaches) Breakdown of direct and indirect evidence Data at the individual drug level
Limitations	Only one study compared head-to-head SGLT2-Is vs. GLP-1 RAs, and three or four studies SGLT2-Is vs. DPP4-Is (ACM and CVM, respectively) Statistically significant inconsistency within the network (i.e. ratio between direct and indirect evidence) for some events (e.g. HF), mainly with DPP4-Is

ACM, all-cause mortality; CV, cardiovascular; CVM, cardiovascular mortality; DPP4-Is, dipeptidyl peptidase-4 inhibitors; GLP-1RA, glucagon-like peptide-1 receptor agonist; HF, heart failure; MI, myocardial infarction; RCTs, randomized clinical trials; SGLT2-Is, sodium glucose co-transporter-2 inhibitors.

within cohorts/subpopulations/sensitivity analyses). Thus, as expected, pooling of these data would confirm the robustness of the findings from single studies despite failing to draw firm conclusions because of high heterogeneity (Figure 1). Apart from the limitations of this statistical exercise (no systematic reviews were performed, outcomes overlap only partially, and neither quality nor subgroup analyses were carried out), the statistical and clinical heterogeneity of the studies might hamper the existence of a drug-related clinically significant CV benefit.

6.2 | The added value of informative meta-analysis

A novel informative MA should assess renal outcomes such as progression to albuminuria or serious decline in renal function, unquantified outcomes (infections), and debated issues such as BKAs or fractures.[72] From a methodological viewpoint, individual patient-data MAs and head-to-head RWD from observational studies could represent a viable research option, such as the OBSERVE-4D meta-analytic approach.[24]

Network meta-analyses (NMAs) are increasingly prevalent in medical journals, and capture the interest of clinicians and policy-makers as they provide two types of findings for a specific outcome: the relative treatment effect for all pairwise comparisons (direct and indirect evidence), and a ranking of the treatments. However, the challenges and pitfalls of these statistical approaches must be carefully considered. In the diabetes area, the first NMA by Zaccardi

et al. assessed cardio-metabolic and safety outcomes of 38 RCTs comparing SGLT2-Is with placebo, also providing indirect comparisons among the three drugs; the authors found that all SGLT2-Is increased risk of infections to a similar extent, whereas canagliflozin 300 mg performed better in terms of efficacy than the other inhibitors.[73]

Recently, an NMA of RCTs compared the efficacy of SGLT2-Is, GLP-1 receptor agonists and DPP4 inhibitors for mortality and CV endpoints, and concluded that SGLT2-Is ranked first in terms of CV mortality.[74] This systematic review with an NMA has several strengths and followed current standards for reporting and conducting meta-analytical studies.[75,76] However, clinicians should be advised that focusing on the probability of being ranked first is potentially misleading. In particular, comparisons between SGLT2-Is and GLP-1 receptor agonists or DPP4 inhibitors suffer from both qualitative and quantitative drawbacks, thus making any firm conclusion decidedly premature (Table 3).

7 | SUMMARY AND CONCLUSION

Landmark CVOTs with SGLT2-Is have generated a paradigm shift in the management of patients with T2DM, specifically in those with prior macrovascular disease, with a transition from past algorithms based primarily on glucose and HbA1c control to a more comprehensive strategy focused on CV prevention. This possible benefit was acknowledged by the FDA, which added a new indication to
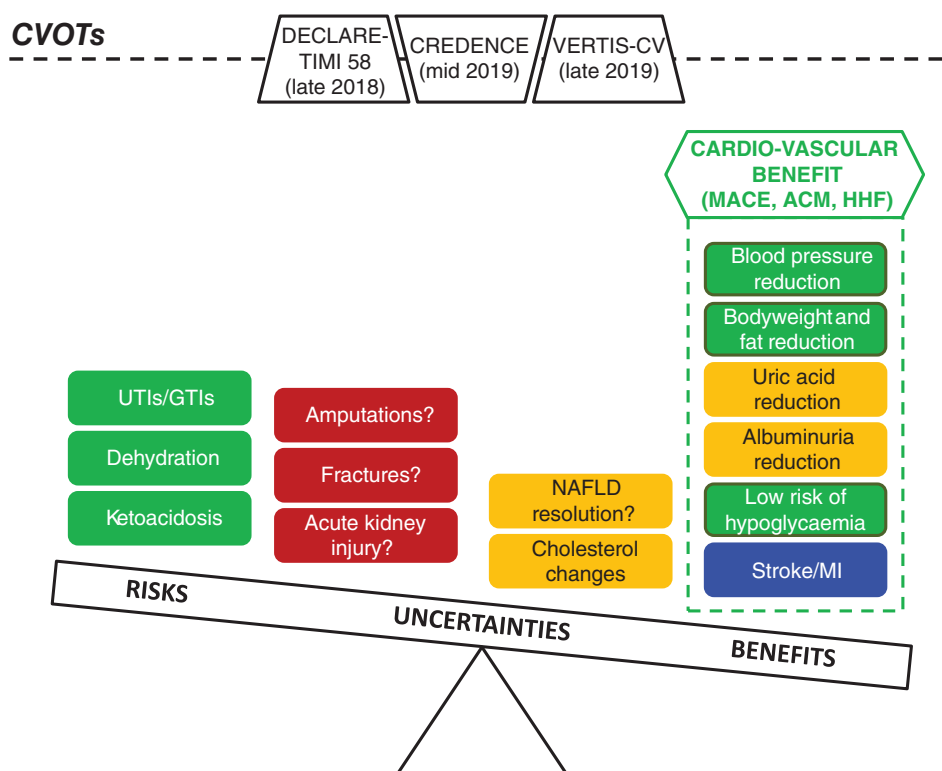
FIGURE 2 Current evidence on the evolving risk–benefit evaluation of SGLT2-Is, with ongoing cardiovascular outcomes trials (CVOTs) potentially affecting final assessment. ACM, all-cause mortality; GTIs, genital tract infections; HHF, hospitalization for heart failure; MACE, major adverse cardiovascular events; MI, myocardial infarction; NAFLD, non-alcoholic fatty liver disease; UTIs, urinary tract infections. Colour coding: green = CONSOLIDATED EVIDENCE (general agreement among types of evidence†); blue = ENCOURAGING EVIDENCE (partial agreement among types of evidence†); orange = PRELIMINARY EVIDENCE (incomplete/pending data, only from a single type of evidence, namely CVOTs); and red = UNCERTAIN EVIDENCE (conflicting data among types of evidence†). †Including systematic reviews with meta-analyses, clinical trials, observational and pharmacovigilance studies

empagliflozin, and is now partially recognized by the 2018 American Diabetes Association guideline, thus increasing the competition to identify the optimal glucose-lowering combination.

Large PER studies such CVD-REALs should be commended for advancing knowledge by pooling such large amounts of prescription data and providing data on heterogeneous cohorts of patients with T2DM. Overall, data from observational cohort studies are not only in agreement with those from RCTs, but also found a larger benefit compared with CVOTs in a population with lower CV risk. Notwithstanding these impressive results and sophisticated statistical techniques, we cannot yet firmly conclude that there is a class effect, and uncertainty remains on safety issues in particular, including BKAs, infections, fractures and renal effects (Figure 2). In essence, confounders were largely minimized through propensity analysis, whereas immortal time bias and time-lag bias were not fully accounted for in CVD-REAL studies.[18,21] The heterogeneity of cohorts stresses the importance of assessing patients for comparability before data pooling. We believe that the strength of CVD-REAL studies lies not in their ability to merge worldwide data, but in capturing the specificities of prescription patterns. In other words, data from a single database should be described as primary analysis (instead of being provided as supplementary material), and discussed in the context of a cross-national comparison.

Considering the consistency of the data and their magnitude, the consensus opinion of the authors is that biases and residual confounders are unlikely to change the direction of results, supporting a clinically significant CV benefit of SGLT2-I use. RTCs remain the best experimental approach to actually inform on efficacy and unsettled issues (renal outcomes), especially with head-to-head comparisons. Instead of performing novel NMAs, the lessons produced by PER studies and pharmacovigilance analyses will instruct the design of next-generation CVOTs focused on lower risk groups or even non-diabetic populations, key emerging outcomes (HF, stroke, peripheral artery disease), unexpected side effects (DKA, fractures, BKAs) and optimal antidiabetic regimens (DPP4 inhibitors/GLP-1 receptor agonists with SGLT2-Is). DECLARE-TIMI 58, CREDENCE and VERTIS-CV will hopefully contribute to defining the evolving risk/benefit profile of SGLT2-Is, clarifying the debate on the use of these medicines in primary CV risk prevention and elucidating the issue of a class effect.

Conflict of interest

E. R., E. P., F. D. P. have no conflicts of interest relevant to the content of the present work. G. M. reports personal fees and other from Sanofi, personal fees and other from NOVO Nordisk, personal fees and other from Eli-Lilly, other from Boehringer-Ingelheim, grants from Merck, other from Glaxo, personal fees and other from Janssen, outside the submitted work. G. P. F. received grant support, lecture or advisory board fees from AstraZeneca, Boehringer-Ingelheim, Eli Lilly, NovoNordisk, Sanofi, Genzyme, Abbott, Novartis, Merck Sharp & Dohme.

Author contributions

E. R. conceived the idea and the overall content of this review, and drafted the manuscript. All the other authors provided substantial contribution to study design. They critically revised the content and approved the final version of the manuscript.

ORCID

Emanuel Raschi http://orcid.org/0000-0003-0487-7996
Gian Paolo Fadini http://orcid.org/0000-0002-6510-2097
Giulio Marchesini http://orcid.org/0000-0003-2407-9860
Fabrizio De Ponti http://orcid.org/0000-0002-0367-9595

REFERENCES

1. Goyat R, Rai P, Chang J, et al. Cardiovascular mortality of oral antidiabetic drugs approved before and after the 2008 US FDA guidance for industry: a systemic review and meta-analysis. *Clin Drug Investig*. 2018;38:491-501.
2. Sattar N, Petrie MC, Zinman B, Januzzi JL Jr. Novel diabetes drugs and the cardiovascular specialist. *J Am Coll Cardiol*. 2017;69:2646-2656.
3. Cefalu WT, Kaul S, Gerstein HC, et al. Cardiovascular outcomes trials in type 2 diabetes: where do we go from here? Reflections from a diabetes care Editors' expert forum. *Diabetes Care*. 2018;41:14-31.
4. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015;373: 2117-2128.
5. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med*. 2017;377: 644-657.
6. Makady A, de Boer A, Hillege H, et al. What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health*. 2017;20:858-865.
7. Najafzadeh M, Gagne JJ, Schneeweiss S. Synergies from integrating randomized controlled trials and real-world data analyses. *Clin Pharmacol Ther*. 2017;102:914-916.
8. Chatterjee S, Davies MJ, Khunti K. What have we learnt from "real world" data, observational studies and meta-analyses. *Diabetes Obes Metab*. 2018;20(suppl 1):47-58.
9. Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? *Clin Pharmacol Ther*. 2017;102:924-933.
10. Kosiborod M, Cavender MA, Fu AZ, et al. Lower risk of heart failure and death in patients initiated on sodium-glucose Cotransporter-2 inhibitors versus other glucose-lowering drugs: the CVD-REAL study (comparative effectiveness of cardiovascular outcomes in new users of sodium-glucose Cotransporter-2 inhibitors). *Circulation*. 2017;136: 249-259.
11. Birkeland KI, Jorgensen ME, Carstensen B, et al. 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.
12. Persson F, Nystrom T, Jorgensen ME, et al. Dapagliflozin is associated with lower risk of cardiovascular events and all-cause mortality in people with type 2 diabetes (CVD-REAL Nordic) when compared with dipeptidyl peptidase-4 inhibitor therapy: a multinational observational study. *Diabetes Obes Metab*. 2018;20:344-351.
13. Kosiborod M, Lam CSP, Kohsaka S, et al. Cardiovascular events associated with SGLT-2 inhibitors versus other glucose-lowering drugs: the CVD-REAL 2 study. *J Am Coll Cardiol*. 2018;71:2628-2639.
14. Udell JA, Yuan Z, Rush T, Sicignano NM, Galitz M, Rosenthal N. Cardiovascular outcomes and risks after initiation of a sodium glucose

Cotransporter 2 inhibitor: results from the EASEL population-based cohort study (evidence for cardiovascular outcomes with sodium glucose Cotransporter 2 inhibitors in the real world). *Circulation*. 2018; 137:1450-1459.

15. Patorno E, Goldfine AB, Schneeweiss S, et al. Cardiovascular outcomes associated with canagliflozin versus other non-gliflozin antidiabetic drugs: population based cohort study. *BMJ*. 2018;360:119.

16. Toulis KA, Willis BH, Marshall T, et al. All-cause mortality in patients with diabetes under treatment with Dapagliflozin: a population-based, open-cohort study in the health improvement network database. *J Clin Endocrinol Metab*. 2017;102:1719-1725.

17. Nystrom T, Bodegard J, Nathanson D, et al. Novel oral glucose-lowering drugs are associated with lower risk of all-cause mortality, cardiovascular events and severe hypoglycaemia compared with insulin in patients with type 2 diabetes. *Diabetes Obes Metab*. 2017;19:831-841.

18. Suissa S. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? *Diabetes Care*. 2018;41:6-10.

19. Nystrom T, Bodegard J, Nathanson D, et al. Comment on Suissa. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? Diabetes Care 2018;41:6-10. *Diabetes Care*. 2018;41: e104-e105.

20. Thuresson M, Cavender MA, Fu AZ, et al. Comment on Suissa. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? Diabetes Care 2018;41:6-10. *Diabetes Care*. 2018;41: e106-e108.

21. Suissa S. Response to comment on Suissa. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? Diabetes Care 2018;41:6-10. *Diabetes Care*. 2018;41:e109-e110.

22. Cavender MA, Norhammar A, Birkeland KI, et al. SGLT-2 inhibitors and cardiovascular risk: an analysis of CVD-REAL. *J Am Coll Cardiol*. 2018;71:2497-2506.

23. Kosiborod M, Birkeland KI, Cavender MA, et al. Rates of myocardial infarction and stroke in patients initiating treatment with SGLT2-inhibitors versus other glucose-lowering agents in real-world clinical practice: results from the CVD-REAL study. *Diabetes Obes Metab*. 2018;20:1983-1987.

24. Ryan PB, Buse JB, Schuemie MJ, et al. Comparative effectiveness of canagliflozin, SGLT2 inhibitors and non-SGLT2 inhibitors on the risk of hospitalization for heart failure and amputation in patients with type 2 diabetes mellitus: a real-world meta-analysis of 4 observational databases (OBSERVE-4D). *Diabetes Obes Metab*. 2018;20:2585-2597.

25. Rawson NSB, D'Arcy C. Healthcare databases for drug safety research: data validity assessment remains crucial. *Drug Saf*. 2018 [Epub ahead of print]. https://doi.org/10.1007/s40264-018-0673-z.

26. Wang SV, Schneeweiss S, Berger ML, et al. Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Value Health*. 2017;20:1009-1022.

27. Bate A. Guidance to reinforce the credibility of health care database studies and ensure their appropriate impact. *Pharmacoepidemiol Saf*. 2017;26:1013-1017.

28. Berger ML, Sox H, Willke RJ, et al. 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.

29. Roberto G, Leal I, Sattar N, et al. Identifying cases of type 2 diabetes in heterogeneous data sources: strategy from the EMIF project. *PLoS One*. 2016;11:e0160648.

30. Patorno E, Patrick AR, Garry EM, et al. Observational studies of the association between glucose-lowering medications and cardiovascular outcomes: addressing methodological limitations. *Diabetologia*. 2014; 57:2237-2250.

31. Patorno E, Garry EM, Patrick AR, et al. Addressing limitations in observational studies of the association between glucose-lowering medications and all-cause mortality: a review. *Drug Saf*. 2015;38:295-310.

32. Grimes DA, Schulz KF. False alarms and pseudo-epidemics: the limitations of observational epidemiology. *Obstet Gynecol*. 2012;120: 920-927.

33. Elze MC, Gregson J, Baber U, et al. Comparison of propensity score methods and covariate adjustment: evaluation in 4 cardiovascular studies. *J Am Coll Cardiol*. 2017;69:345-357.

34. Wang SV, He M, Jin Y, et al. A review of the performance of different methods for propensity score matched subgroup analyses and a summary of their application in peer-reviewed research studies. *Pharmacoepidemiol Drug Saf*. 2017;26:1507-1512.

35. Prentice JC, Conlin PR, Gellad WF, Edelman D, Lee TA, Pizer SD. Capitalizing on prescribing pattern variation to compare medications for type 2 diabetes. *Value Health*. 2014;17:854-862.

36. Patorno E, Gopalakrishnan C, Franklin JM, et al. Claims-based studies of oral glucose-lowering medications can achieve balance in critical clinical variables only observed in electronic health records. *Diabetes Obes Metab*. 2018;20:974-984.

37. Fadini GP, Zatti G, Baldi I, et al. Use and effectiveness of dapagliflozin in routine clinical practice: an Italian multicentre retrospective study. *Diabetes Obes Metab*. 2018;20:1781-1786.

38. American Diabetes Association. 8. Pharmacologic approaches to glycemic treatment: standards of medical care in diabetes-2018. *Diabetes Care*. 2018;41:S73-S85.

39. Li L, Li S, Deng K, et al. Dipeptidyl peptidase-4 inhibitors and risk of heart failure in type 2 diabetes: systematic review and meta-analysis of randomised and observational studies. *BMJ*. 2016;352:i610.

40. Fadini GP, Saragoni S, Russo P, et al. Intraclass differences in the risk of hospitalization for heart failure among patients with type 2 diabetes initiating a dipeptidyl peptidase-4 inhibitor or a sulphonylurea: results from the OsMed health-DB registry. *Diabetes Obes Metab*. 2017;19: 1416-1424.

41. Scheen AJ. Cardiovascular effects of new oral glucose-lowering agents: DPP-4 and SGLT-2 inhibitors. *Circ Res*. 2018;122:1439-1459.

42. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HAW. 10-year follow-up of intensive glucose control in type 2 diabetes. *N Engl J Med*. 2008;359:1577-1589.

43. Lytvyn Y, Bjornstad P, Udell JA, Lovshin JA, Cherney DZI. Sodium glucose Cotransporter-2 inhibition in heart failure: potential mechanisms, clinical applications, and summary of clinical trials. *Circulation*. 2017; 136:1643-1658.

44. Custodio JS Jr, Duraes AR, Abreu M, et al. SGLT2 inhibition and heart failure-current concepts. *Heart Fail Rev*. 2018;23:409-418.

45. Wanner C, Inzucchi SE, Lachin JM, et al. Empagliflozin and progression of kidney disease in type 2 diabetes. *N Engl J Med*. 2016;375: 323-334.

46. Scheen AJ. Pharmacokinetics, pharmacodynamics and clinical use of SGLT2 inhibitors in patients with type 2 diabetes mellitus and chronic kidney disease. *Clin Pharmacokinet*. 2015;54:691-708.

47. Szalat A, Perlman A, Muszkat M, Khamaisi M, Abassi Z, Heyman SN. Can SGLT2 inhibitors cause acute renal failure? Plausible role for altered glomerular hemodynamics and medullary hypoxia. *Drug Saf*. 2018;41:239-252.

48. Perlman A, Heyman SN, Matok I, Stokar J, Muszkat M, Szalat A. Acute renal failure with sodium-glucose-cotransporter-2 inhibitors: analysis of the FDA adverse event report system database. *Nutr Metab Cardiovasc Dis*. 2017;27:1108-1113.

49. Desai M, Yavin Y, Balis D, et al. Renal safety of canagliflozin, a sodium glucose co-transporter 2 inhibitor, in patients with type 2 diabetes mellitus. *Diabetes Obes Metab*. 2017;19:897-900.

50. Raschi E, Parisotto M, Forcesi E, et al. Adverse events with sodium-glucose co-transporter-2 inhibitors: a global analysis of international spontaneous reporting systems. *Nutr Metab Cardiovasc Dis*. 2017;27:1098-1107.

51. Li D, Wang T, Shen S, Fang Z, Dong Y, Tang H. Urinary tract and genital infections in patients with type 2 diabetes treated with sodium-glucose co-transporter 2 inhibitors: a meta-analysis of randomized controlled trials. *Diabetes Obes Metab*. 2017;19:348-355.

52. Filippas-Ntekouan S, Filippatos TD, Elisaf MS. SGLT2 inhibitors: are they safe? *Postgrad Med*. 2018;130:72-82.

53. Bonora BM, Avogaro A, Fadini GP. Sodium-glucose co-transporter-2 inhibitors and diabetic ketoacidosis: an updated review of the literature. *Diabetes Obes Metab*. 2018;20:25-33.

54. Monami M, Nreu B, Zannoni S, Lualdi C, Mannucci E. Effects of SGLT-2 inhibitors on diabetic ketoacidosis: a meta-analysis of randomised controlled trials. *Diabetes Res Clin Pract*. 2017;130:53-60.

55. Fadini GP, Bonora BM, Avogaro A. SGLT2 inhibitors and diabetic ketoacidosis: data from the FDA adverse event reporting system. *Diabetologia*. 2017;60:1385-1389.

56. Fralick M, Schneeweiss S, Patorno E. Risk of diabetic ketoacidosis after initiation of an SGLT2 inhibitor. *N Engl J Med*. 2017;376:2300-2302.

57. Kim YG, Jeon JY, Han SJ, Kim DJ, Lee KW, Kim HJ. 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*. 2018;20:1852-1858.

58. Mannucci E, Monami M. Bone fractures with sodium-glucose co-transporter-2 inhibitors: how real is the risk? *Drug Saf*. 2017;40:115-119.

59. Toulis KA, Bilezikian JP, Thomas GN, et al. Initiation of dapagliflozin and treatment-emergent fractures. *Diabetes Obes Metab*. 2018;20:1070-1074.

60. Yuan Z, DeFalco FJ, Ryan PB, et al. Risk of lower extremity amputations in people with type 2 diabetes mellitus treated with sodium-glucose co-transporter-2 inhibitors in the USA: a retrospective cohort study. *Diabetes Obes Metab*. 2018;20:582-589.

61. Fadini GP, Avogaro A. SGTL2 inhibitors and amputations in the US FDA Adverse Event Reporting System. *Lancet Diabetes Endocrinol*. 2017;5:680-681.

62. Khouri C, Cracowski JL, Roustit M. SGLT-2 inhibitors and the risk of lower-limb amputation: is this a class effect? *Diabetes Obes Metab*. 2018;20:1531-1534.

63. Scheen AJ. Does lower limb amputation concern all SGLT2 inhibitors? *Nat Rev Endocrinol*. 2018;14:326-328.

64. Bonaca MP, Beckman JA. Sodium glucose cotransporter 2 inhibitors and amputation risk: achilles' heel or opportunity for discovery? *Circulation*. 2018;137:1460-1462.

65. Faillie JL. Pharmacological aspects of the safety of gliflozins. *Pharmacol Res*. 2017;118:71-81.

66. Fadini GP, Bonora BM, Mayur S, Rigato M, Avogaro A. Dipeptidyl peptidase-4 inhibitors moderate the risk of genitourinary tract infections associated with sodium-glucose co-transporter-2 inhibitors. *Diabetes Obes Metab*. 2018;20:740-744.

67. Fadini GP, Sarangdhar M, Avogaro A. Pharmacovigilance evaluation of the association between DPP-4 inhibitors and heart failure: stimulated reporting and moderation by drug interactions. *Diabetes Ther*. 2018;9:851-861.

68. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94:485-514.

69. Riaz IB, Khan MS, Riaz H, et al. Disorganized systematic reviews and meta-analyses: time to systematize the conduct and publication of these study overviews? *Am J Med*. 2016;129:339.e11-339.e18.

70. Packer M. Are meta-analyses a form of medical fake news? Thoughts about how they should contribute to medical science and practice. *Circulation*. 2017;136:2097-2099.

71. Raschi E, Poluzzi E, Salvo F, et al. Pharmacovigilance of sodium-glucose co-transporter-2 inhibitors: what a clinician should know on disproportionality analysis of spontaneous reporting systems. *Nutr Metab Cardiovasc Dis*. 2018;28:533-542.

72. Radholm K, Wu JH, Wong MG, et al. Effects of sodium-glucose cotransporter-2 inhibitors on cardiovascular disease, death and safety outcomes in type 2 diabetes – a systematic review. *Diabetes Res Clin Pract*. 2018;140:118-128.

73. Zaccardi F, Webb DR, Htike ZZ, Youssef D, Khunti K, Davies MJ. Efficacy and safety of sodium-glucose co-transporter-2 inhibitors in type 2 diabetes mellitus: systematic review and network meta-analysis. *Diabetes Obes Metab*. 2016;18:783-794.

74. Zheng SL, Roddick AJ, Ghar-Jaffar R, et al. Association between use of sodium-glucose cotransporter 2 inhibitors, glucagon-like peptide 1 agonists, and Dipeptidyl peptidase 4 inhibitors with all-cause mortality in patients with type 2 diabetes: a systematic review and meta-analysis. *JAMA*. 2018;319:1580-1591.

75. Salanti G, Del GC, Chaimani A, et al. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9:e99682.

76. Riley RD, Jackson D, Salanti G, et al. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ*. 2017;358:j3932.