



Lower Risk of Death With SGLT2 Inhibitors in Observational Studies: Real or Bias?

Samy Suissa

Diabetes Care 2018;41:6–10 | <https://doi.org/10.2337/dc17-1223>

Two recent observational studies reported a remarkably lower rate of all-cause death associated with sodium–glucose cotransporter 2 inhibitor (SGLT2i) use in all patients with type 2 diabetes and not only those at increased cardiovascular risk. The >50% lower mortality rates reported in these studies are much greater than those found in the BI 10773 (Empagliflozin) Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (EMPA-REG OUTCOME) and CANagliflozin cardiovascular Assessment Study (CANVAS) randomized trials. We show that these observational studies are affected by time-related biases, including immortal time bias and time-lag bias, which tend to exaggerate the benefits observed with a drug. The Comparative Effectiveness of Cardiovascular Outcomes in New Users of SGLT-2 Inhibitors (CVD-REAL) study, based on 166,033 users of SGLT2i and 1,226,221 users of other glucose-lowering drugs (oGLD) identified from health care databases of six countries, was affected by immortal time bias. Indeed, the immortal time between the first oGLD prescription and the first SGLT2i prescription was omitted from the analysis, which resulted in increasing the rate of death in the oGLD group and thus producing the appearance of a lower risk of death with SGLT2i use. The Swedish study compared 10,879 SGLT2i/dipeptidyl peptidase 4 inhibitor (DPP-4i) users with 10,879 matched insulin users. Such comparisons involving second-line therapies with a third-line therapy can introduce time-lag bias, as the patients may not be at the same stage of diabetes. This bias is compounded by the fact that the users of insulin had already started their insulin before cohort entry, unlike the new users of SGLT2i. Finally, the study also introduces immortal time bias with respect to the effects of SGLT2i relative to DPP-4i. In conclusion, the >50% lower rate of death with SGLT2i in type 2 diabetes reported by two recent observational studies is likely exaggerated by immortal time and time-lag biases. It thus remains uncertain whether the benefit seen with empagliflozin in the EMPA-REG OUTCOME trial applies to all SGLT2i and to all patients with type 2 diabetes, not only those at increased cardiovascular risk. While observational studies can provide crucial real-world evidence for the effects of medications, they need to be carefully conducted to avoid such major time-related biases.

The findings from the BI 10773 (Empagliflozin) Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (EMPA-REG OUTCOME) randomized trial of substantial reductions in mortality with the sodium–glucose cotransporter 2 inhibitor (SGLT2i) empagliflozin in patients with type 2 diabetes at increased cardiovascular risk triggered several questions on this class of medications (1). Indeed, it remained uncertain whether this benefit applied to the entire class of SGLT2i and whether it applied to all patients with type 2 diabetes and not only those at increased cardiovascular risk.

Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, Canada
Department of Epidemiology, Biostatistics, and Occupational Health and Department of Medicine, McGill University, Montreal, Canada

Corresponding author: Samy Suissa, samy.suissa@mcgill.ca.

Received 19 June 2017 and accepted 2 October 2017.

© 2017 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <http://www.diabetesjournals.org/content/license>.

Observational studies conducted to address these questions have recently been published (2,3). These studies reported that SGLT2i use in broad populations with type 2 diabetes, not only those at increased cardiovascular risk, was associated with >50% lower rates of all-cause mortality, in addition to importantly lower incidence of major cardiovascular events. It is certainly tempting to accept these results in view of their apparent consistency with the 32% reduction in all-cause mortality found in the EMPAREG OUTCOME randomized trial. However, some sources of bias characteristic of such database observational studies could have affected and exaggerated the reported results.

In this perspective, I describe how the two recent observational studies investigating the association between SGLT2i use and mortality are affected by time-lag and immortal time biases.

THE CVD-REAL OBSERVATIONAL STUDY

Comparative Effectiveness of Cardiovascular Outcomes in New Users of SGLT-2 Inhibitors (CVD-REAL) is an observational study using health care data, including medical claims, primary care and hospital records, and national registries, from six countries, namely Norway, Denmark, Sweden, Germany, the U.K., and the U.S. (2). Patients with type 2 diabetes, newly initiated on an SGLT2i between November 2012 and November 2016, were identified and propensity score matched to patients initiating an “other glucose-lowering drug” (oGLD) during that period. Treatment initiation was defined as no prescriptions of that class given during the preceding year. Patients were followed from treatment initiation until the end of this treatment, death, or the end of data availability. Hazard ratios (HRs) of death and hospitalization for heart failure were estimated by country and pooled across countries.

A total of 166,033 new SGLT2i and 1,226,221 oGLD users were identified. After propensity matching, there were 154,528 patients in each treatment group. The SGLT2i exposure time was divided as canagliflozin (53%), dapagliflozin (42%), and empagliflozin (5%). There were 412 deaths during 79,888 person-years follow-up in the SGLT2i group (incidence rate 5.2 per 1,000 person-years) compared with 922 deaths during 74,102

person-years follow-up in the oGLD group (incidence rate 12.4 per 1,000 person-years). Use of SGLT2i was associated with a 51% lower rate of death than oGLD (HR 0.49 [95% CI 0.41–0.57]; $P < 0.001$).

IMMORTAL TIME BIAS IN THE CVD-REAL STUDY

Immortal time bias was introduced in the CVD-REAL study by the approach used to form the cohort (4). Generally, a cohort study that compares the initiation of a study drug (SGLT2i) to the initiation of a comparator drug (oGLD) must account for all follow-up from the time of the first occurrence of either drug in the cohort, which was not done in the CVD-REAL study. Excluding or misclassifying some of the follow-up time, in particular the time between initiation of the comparator oGLD and initiation of the study drug SGLT2i, will result in immortal time bias (4).

To illustrate this issue, we use the U.S. Truven MarkenScan cohort (the largest of the six databases), where there were 123,648 new SGLT2i users and 712,426 new oGLD users before matching.

Supplementary Table 5 of the article by Kosiborod et al. (2) shows that compared with the initiators of oGLD, the initiators of SGLT2i were more frequent users in the year before initiation of metformin (80.2% vs. 41.5%), sulfonylureas (41.2% vs. 22.9%), dipeptidyl peptidase 4 inhibitors (DPP-4i) (36.9% vs. 13.1%), thiazolidinediones (11.1% vs. 4.9%), glucagon-like peptide 1 receptor agonists (22.4% vs. 4.6%), and insulin (30.8% vs. 17.2%). Thus, we can presume that most, if not all, of the initiators of SGLT2i had been prior initiators of oGLD (the comparator drug) sometime between the study entry date (November 2012) and the day they initiated SGLT2i. The time between the first oGLD prescription (comparator drug) and the first SGLT2i prescription (study drug) is called “immortal” (4).

Immortal time bias is introduced by omitting this time between the first oGLD prescription and the first SGLT2i prescription from the design and the analysis. This time is called “immortal” (thick red line in Fig. 1) because the patient must be alive to have received their first SGLT2i prescription (4). This immortal person-time should not be omitted but

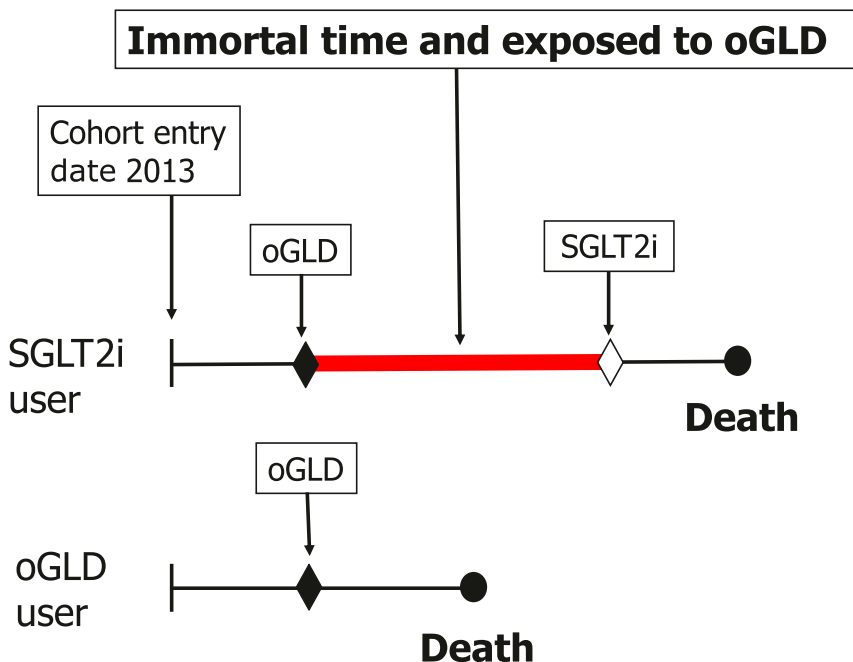


Figure 1—Depiction of immortal time bias: description of SGLT2i-exposed and oGLD-exposed patients who die of any cause according to the definition used in the CVD-REAL observational study (2). The top patient initiated treatment with an oGLD and subsequently switched to or added an SGLT2i, but the patient was classified as an SGLT2i user. The time between the first oGLD prescription and the first SGLT2i prescription is thus immortal (thick red line), since the subject must survive to receive this first SGLT2i prescription, but is not included as exposed to oGLD, leading to immortal time bias.

rather classified as “oGLD-exposed” until the start of SGLT2i, at which point the remaining person-time can be classified as “SGLT2i-exposed.” For example, consider two similar patients who received a first oGLD prescription at the same time point during the study period (Fig. 1). According to the approach of the CVD-REAL study, if the first patient subsequently received an SGLT2i, they would be classified as SGLT2i-exposed while the second patient who died after their first oGLD prescription would be classified as oGLD-exposed (Fig. 1). However, while the first patient had to survive to receive an SGLT2i, the time they were on the oGLD and survived (immortal) was not counted in the oGLD risk calculation according to the approach of the CVD-REAL study. Since the denominator to compute the rate of death under oGLD includes all person-time during oGLD exposure, excluding this immortal time will result in an overestimate of the mortality rate among the remaining oGLD users, as the denominator of the rate will be underestimated by the excluded immortal person-time.

Thus, the omission of this immortal person-time in the design and in the analysis of this study from the at-risk period of the oGLD-exposed group will lead to immortal time bias (5). To avoid this bias, one can use Poisson-type regression techniques that classify all person-time in the cohort according to oGLD or SGLT2i exposure. One can also use Cox-type models with time-dependent exposure that allow classifying this immortal person-time as oGLD-exposed until the start of SGLT2i, at which point the remaining person-time can be classified as SGLT2i-exposed. Finally, if the study requires matching on or adjustment by propensity scores, a prevalent new-user design with time-conditional propensity scores can be used to avoid this bias (6).

THE SWEDISH OBSERVATIONAL STUDY

The observational study by Nyström et al. (3) used health care data from Sweden to form a cohort of patients with type 2 diabetes, with cohort entry at the first prescription for a DPP-4i or SGLT2i, or of insulin, between 1 July 2013 and 31 December 2014. The patients receiving a DPP-4i or SGLT2i were propensity score matched to patients receiving insulin during that period. Patients were followed

from the first prescription in that period until the end of this treatment, death, or the end of the study period. HRs of death comparing SGLT2i/DPP-4i use with insulin use were estimated by the Cox proportional hazards model.

A total of 12,544 SGLT2i/DPP-4i users and 25,059 insulin users were identified. After propensity matching, there were 10,879 patients in each treatment group. There were 330 deaths during 16,304 person-years follow-up in the SGLT2i/DPP-4i group (incidence rate 25.6 per 1,000 person-years) compared with 554 deaths during 16,306 person-years follow-up in the insulin group (incidence rate 45.7 per 1,000 person-years). Use of SGLT2i/DPP-4i was associated with a 44% lower rate of death than insulin (HR 0.56 [95% CI 0.49–0.64]; $P < 0.001$).

TIME-LAG AND IMMORTAL TIME BIASES IN THE SWEDISH STUDY

The Swedish cohort study compared exposure time under SGLT2i/DPP-4i with

exposure time under insulin. Such comparisons involving second-line therapies with a third-line therapy can introduce time-lag bias. Indeed, patients initiating a second-line therapy may not be at the same stage of diabetes as those initiating a third-line therapy; their comparison can induce confounding by disease duration, as longer duration of diabetes may be associated with higher mortality, independently of age. This is illustrated in Fig. 2A, where the patient receiving the comparator insulin is far along their disease course after having previously used several oGLD, while the patient receiving the study drug SGLT2i/DPP-4i is earlier in their disease, which confounds their comparison. Indeed, the patient on insulin is expected to have a shorter survival (thick blue line in Fig. 2A), while the patient on SGLT2i/DPP-4i is expected to have a longer survival simply on the basis of disease duration. An unconfounded comparison of survival requires matching on disease duration, and

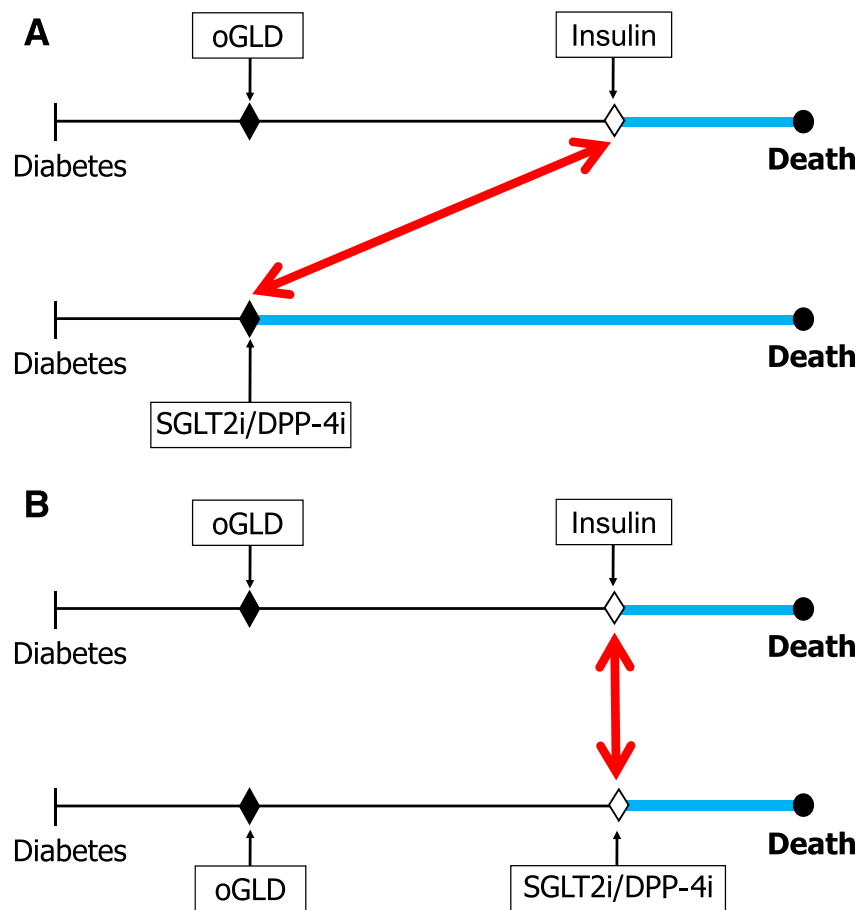


Figure 2—Depiction of A) time-lag bias in comparing a second-line drug (SGLT2i/DPP-4i) used at an earlier stage of diabetes with third-line insulin and B) cohort design that controls for time-lag bias by comparing two patients at the same stage of diabetes and with similar previous medication use (oGLD).

possibly also on prior medication use, with the cohort formed as in Fig. 2B rather than the time-lag comparison suggested by Fig. 2A. This can be done using straightforward matching of subjects on disease duration and prior medication use or, if the study requires matching on propensity scores, using a prevalent new-user design with time-conditional propensity scores (6).

A second issue with the study design is that while the first SGLT2i prescription was clearly the first ever (SGLT2i entered the Swedish market in July 2013), making these new users of this treatment, this was not the case for the older DPP-4i and insulin. In particular, patients in the reference insulin group were likely not new users of insulin and could have used it for long before the study entry date (Fig. 3). This comparison of incident users of SGLT2i with prevalent users of insulin compounds the potential for confounding bias by disease severity. Moreover, 6% of the insulin group had previously used a DPP-4i, making the comparison with SGLT2i/DPP-4i unclear.

Finally, the study also introduces immortal time bias with respect to the effects of SGLT2i relative to DPP-4i. Indeed, patients with both drug classes were “primarily included in the SGLT2 inhibitor group and secondly in the DPP-4 inhibitor group. For example, a patient filling a DPP-4 inhibitor prescription prior to an SGLT2 inhibitor prescription was placed

in the SGLT2 inhibitor group” (3). As a result of this definition, the time between the first DPP-4i prescription and the first SGLT2i prescription will be immortal (similar to Fig. 1). It should not be omitted but rather classified as DPP-4i-exposed until the start of SGLT2i, at which point the remaining person-time can be classified as SGLT2i-exposed. Here again, this immortal time bias would result in an underestimate of the DPP-4i effect on mortality.

DISCUSSION

The two recent observational studies conducted to evaluate the effectiveness of SGLT2i use in a real-world setting, including all patients with type 2 diabetes and not only those at increased cardiovascular risk, reported remarkably lower rates of all-cause mortality, in addition to an importantly lower incidence of major cardiovascular events. We showed that these studies, by their design, are affected by immortal time and time-lag biases, which have been described and identified as occurring frequently in the context of metformin and cancer incidence and mortality (7). These biases tend to overestimate the “effectiveness” of a drug, particularly exaggerating the reduction in mortality. Thus, the >50% lower rates of all-cause mortality associated with SGLT2i use reported in these studies are likely an amplification of the real effect.

We focused particularly on the time-related biases in these two recent observational studies, but one cannot overlook other potential sources of bias. First, the mortality rate on the SGLT2i study drug from the CVD-REAL study is much lower than that in the Swedish study (5.2 vs. 25.6 per 1,000 person-years, respectively). Moreover, the mortality rates in the CVD-REAL study are surprisingly highly variable, from the U.S. Truven MarketScan cohort’s rate of 3.1 per 1,000 per year to the Danish cohort’s 18.7 per 1,000 per year. The authors did not discuss reasons for this sixfold span, which may reflect some incompleteness of death information in the U.S. database, which accounts for >50% of the overall data in the CVD-REAL study. These differences can introduce bias if the various sources of data that contribute to the U.S. Truven MarketScan database have differential completeness on mortality data and vary on prescribing access to newer drugs such as SGLT2i. Second, propensity scores are useful for balanced comparisons but cannot account for unmeasured and unknown confounders such as the healthier behavior of patients prescribed a new drug such as an SGLT2i or the profile of treating physicians who are early adopters of a new effective drug class.

The remarkable mortality findings of the two observational studies can be inaccurately perceived as compatible with the reported results of the EMPA-REG OUTCOME randomized trial, of a 32% reduction in all-cause mortality with empagliflozin (1). However, the EMPA-REG OUTCOME trial included patients with type 2 diabetes at increased cardiovascular risk, with empagliflozin appearing to have specific cardiovascular effects (1,8,9). The observational studies, which included all patients with type 2 diabetes, did not report the effects stratified by cardiovascular risk. However, if the impact on mortality is more important in those patients at increased cardiovascular risk, we can hypothesize that the 32% reduction in mortality would be somewhat lower in a broader population with diabetes. This premise seems to be supported partly by the recent CANagliflozin Cardiovascular Assessment Study (CANVAS) randomized trial of the SGLT2i canagliflozin, based on over 10,000 patients with type 2 diabetes, which reported a lower 13% reduction in the risk of all-cause death (HR 0.87 [95% CI 0.74–1.01]) for this SGLT2i

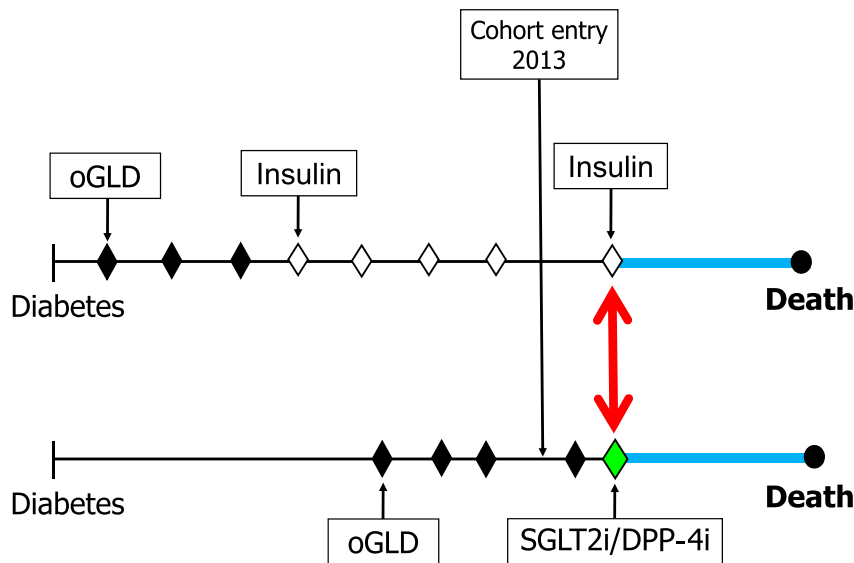


Figure 3—Depiction of prevalent/incident bias from comparing patients at their first-ever SGLT2i prescription (incident users) to patients at their first insulin prescription after 2013, who could also have used insulin previously (prevalent users), leading to potential confounding bias by disease severity.

compared with placebo (10). In the CANVAS trial population, 66% had a history of cardiovascular disease but, regrettably, the HR of all-cause mortality was not reported stratified by history of cardiovascular disease. However, the HR for the composite outcome (death from cardiovascular causes, nonfatal myocardial infarction, or nonfatal stroke), which was 0.86 (95% CI 0.75–0.97) overall, was 0.82 (95% CI 0.72–0.95) for the patients with a history of cardiovascular disease and 0.98 (95% CI 0.74–1.30) among those with no such history (10).

Conclusion

The >50% lower risk of all-cause death reported by the recent observational studies of SGLT2i effectiveness is inconsistent with the reductions found in the recent large randomized trials of this class of drugs. It seems likely that immortal time and time-lag biases exaggerated the mortality effects reported in the observational studies. Consequently, it still remains uncertain whether the clear and significant

mortality reduction shown with empagliflozin also applies to all SGLT2i drugs and whether it also applies to all patients with type 2 diabetes and not only those at increased cardiovascular risk. The answer will come from upcoming large randomized trials and well-conducted observational studies free of time-related biases such as immortal time and time-lag biases.

Funding and Duality of Interest. S.S. is the recipient of the James McGill Chair. S.S. has received research grants and has participated in advisory board meetings or as a speaker at conferences for AstraZeneca, Bayer Pharmaceuticals, Boehringer Ingelheim, Bristol-Myers Squibb, Merck, and Novartis. No other potential conflicts of interest relevant to this article were reported.

References

1. Zinman B, Wanner C, Lachin JM, et al.; EMPA-REG OUTCOME Investigators. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med* 2015;373:2117–2128
2. Kosiborod M, Cavender MA, Fu AZ, et al. Lower risk of heart failure and death in patients initiated on SGLT-2 inhibitors versus other glucose-lowering

drugs: the CVD-REAL study. *Circulation* 2017;136:249–259

3. Nyström T, Bodegard J, Nathanson D, Thuresson M, Norhammar A, Eriksson JW. 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
4. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol* 2008;167:492–499
5. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf* 2007;16:241–249
6. Suissa S, Moodie EE, Dell’Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiol Drug Saf* 2017;26:459–468
7. Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care* 2012;35:2665–2673
8. McMurray J. EMPA-REG - the “diuretic hypothesis.” *J Diabetes Complications* 2016;30:3–4
9. Ferrannini E, Mark M, Mayoux E. CV protection in the EMPA-REG OUTCOME trial: a “thrifty substrate” hypothesis. *Diabetes Care* 2016;39:1108–1114
10. Neal B, Perkovic V, Mahaffey KW, et al.; CANVAS Program Collaborative Group. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med* 2017;377:644–657