



Mitigating Cardiovascular Risk in Type 2 Diabetes With Antidiabetes Drugs: A Review of Principal Cardiovascular Outcome Results of EMPA-REG OUTCOME, LEADER, and SUSTAIN-6 Trials

Sanjay Kaul

Diabetes Care 2017;40:821–831 | <https://doi.org/10.2337/dc17-0291>

The U.S. Food and Drug Administration (FDA) issued a diabetes guidance in 2008 mandating that all new antidiabetes drugs rule out excess cardiovascular (CV) risk, defined as an upper bound of the two-sided 95% CI for major adverse CV events (MACE) of less than 1.80 preapproval and 1.30 postapproval. Over 25 large, prospective, randomized, controlled clinical trials involving nearly 195,000 subjects thus far have been completed or are ongoing in accordance with this guidance. The results of seven trials have been presented so far—three with dipeptidyl peptidase 4 inhibitors, one with a sodium–glucose cotransporter 2 (SGLT2) inhibitor, and three with glucagon-like peptide 1 receptor agonists (GLP-1 RA). While all seven trials showed noninferiority in the rate of MACE with the use of these agents compared with placebo, three of them revealed CV benefits. Treatment with empagliflozin (an SGLT2 inhibitor) and treatment with liraglutide (a GLP-1 RA) both significantly reduced the risk of MACE, mortality from CV causes, and mortality from any cause when compared with placebo. Treatment with semaglutide, another GLP-1 RA, showed a significantly lower rate of MACE but not mortality from CV or any cause compared with placebo. In all of the trials, the effects of treatment on outcomes were out of proportion to the small differences in glycemic control levels, suggesting that the effects observed were likely unrelated to differences in the glucose-lowering efficacy. Overall, the results of these trials yield a favorable benefit-risk balance for these therapies in mitigating CV risk in patients with type 2 diabetes. More research is needed to elucidate the underlying mechanisms and confirm whether the CV benefits are a class effect or whether the benefits persist in patients without established CV disease or are evident even in patients without diabetes.

Be skeptical, ask questions, demand proof. Demand evidence. Don't take anything for granted. But here's the thing: when you get proof, you need to accept the proof. And we're not that good at doing that.

—Michael Specter (1)

Diabetes confers a high risk for CV disease (CVD). Therapies to lower lipids and blood pressure are proven interventions to reduce the risk of morbidity and mortality from CVD in patients with diabetes. However, it is not clear whether CVD is also prevented by

Cedars-Sinai Medical Center, Los Angeles, CA
Corresponding author: Sanjay Kaul, sanjay.kaul@cshs.org.

Received 7 February 2017 and accepted 12 April 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>.

See accompanying article, p. 813.

therapies targeting glycemic control. Historically, new therapies for diabetes were marketed on the basis of well-tolerated improvements in glycemic control. Primarily in response to the findings of possible increased CV risk with rosiglitazone (2), the U.S. Food and Drug Administration (FDA) issued a diabetes guidance in 2008 mandating that all new antidiabetes drugs rule out excess CV risk, defined as an upper bound of the two-sided 95% CI for major adverse CV events (MACE) of less than 1.80 preapproval and 1.30 postapproval (3). Over 25 large, prospective, randomized, controlled clinical trials involving nearly 195,000 subjects thus far have been completed or are ongoing in accordance with this guidance. The results of the first seven trials—three with dipeptidyl peptidase 4 inhibitors (4–6), three with glucagon-like peptide 1 receptor agonists (GLP-1 RA) (7–9), and one with a sodium–glucose cotransporter 2 (SGLT2) inhibitor (10) were reported between 2013 and 2016. All seven trials met the primary objective to exclude an unacceptable level of CV risk as defined in the FDA guidance. In September 2015, the BI 10773 (Empagliflozin) Cardiovascular Outcome Event Trial in Type 2 Diabetes Mellitus Patients (EMPA-REG OUTCOME) was the first large, prospective, randomized, placebo-controlled trial to report a CV benefit of an antidiabetes drug (10). Since then, two additional placebo-controlled trials have reported favorable outcomes: Liraglutide Effect and Action in Diabetes: Evaluation of Cardiovascular Outcome Results—A Long Term Evaluation (LEADER), which demonstrated cardiovascular benefit with once-daily treatment with the GLP-1 RA liraglutide (8), and Trial to Evaluate Cardiovascular and Other Long-term Outcomes With Semaglutide in Subjects With Type 2 Diabetes (SUSTAIN-6), which unexpectedly reported favorable results with the long-acting GLP-1 RA semaglutide (9). Both empagliflozin and liraglutide, but not semaglutide, are approved in the U.S. as “an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus.” On 2 December 2016, the FDA announced approval for expanded indication of empagliflozin “to reduce the risk of CV death in adult patients with type 2 diabetes mellitus and cardiovascular disease” (11), thereby becoming the first antidiabetes drug to receive such a claim. A regulatory

decision regarding CV risk reduction claim for liraglutide has not been made at the time of this review.

The purpose of this report is to review the principal CV outcome results of the three trials that reported favorable outcomes. As none of the trials were prospectively designed with the expressed intent of demonstrating a CV benefit of the new antidiabetes therapy, the focus of this review is to describe whether the quality and the quantity of evidence is sufficient to support a valid inference of “superiority” of CV risk reduction. In addition, the key deliberations from the FDA’s 28 June 2016 Meeting of the Endocrinologic and Metabolic Drugs Advisory Committee (EMDAC) convened to discuss whether the findings from EMPA-REG OUTCOME establish that empagliflozin is effective in reducing CV risk will be summarized.

TRIAL DESIGN AND DEMOGRAPHICS

All three trials were designed as safety studies where the primary objective was to rule out unacceptable CV risk as mandated by the 2008 guidance. These trials generally specify that noninferiority and superiority hypotheses will be tested in a sequential manner in their analysis plans regardless of what the trial is initially powered to show. EMPA-REG OUTCOME was initiated in September 2010 prior to the approval of empagliflozin in August 2014. The approval was based on meeting the preapproval exclusion of 1.8 risk margin derived from a pooled analysis of phase 2/3 trials ($n = 54$ MACE) plus interim analysis of EMPA-REG OUTCOME ($n = 142$ MACE, 85/3,046 empagliflozin vs. 57/1,513 placebo, hazard ratio [HR] 0.74 [99.98% CI 0.39, 1.39]). The postapproval exclusion of 1.3 risk margin was not met in the interim analysis (12). LEADER was initiated in September 2010 following the approval of liraglutide in January 2010. The clinical development program was completed before the FDA guidance was issued in December 2008, but retrospective analyses of CV events from the combined phase 2/3 trials versus active comparators and placebo ($n = 39$ MACE) showed that liraglutide met the preapproval, but not the postapproval, standard for ruling out unacceptable increase in CV risk (HR 0.73 [95% CI 0.38, 1.41]) (13). The FDA therefore required a separate postapproval study

of CV safety. SUSTAIN-6 was initiated in February 2013 as a preapproval trial aimed at enhancing the probability that regulatory guidance was met in the development program.

CV safety trials conducted to meet the FDA guidance generally use an efficient trial design that enrolls patients with more advanced atherosclerotic CV risk or established CVD (“enriched” population) to accrue sufficient events in a timely manner. However, a major limitation of such an approach is that the safety population is not representative of patients in ambulatory diabetes care, thereby raising questions about generalizability. EMPA-REG OUTCOME and SUSTAIN-6 were event-driven trials requiring at least 691 and 122 primary end point events to rule out postapproval HR of 1.3 and preapproval HR of 1.8, respectively. The primary end point was time from randomization to the first occurrence of an adjudication committee–confirmed 3-point MACE—a composite of CV death, nonfatal myocardial infarction (MI), or nonfatal stroke (9,10). The duration of the LEADER trial was driven by the number of MACE (at least 611 events) and by time (a minimum period of 3.5 years) (8). The primary hypothesis of noninferiority was analyzed with pooled doses of 10 mg and 25 mg empagliflozin (10), 0.6 to 1.8 mg liraglutide (85% of the total exposure was to 1.8 mg dose) (8) and pooled doses of 0.5 mg and 1 mg semaglutide (9) versus placebo. A superiority hypothesis was tested in a prespecified hierarchy after noninferiority had been initially established in EMPA-REG OUTCOME (for both primary outcome and key secondary outcome) and LEADER (for only primary outcome), but it was not prespecified in SUSTAIN-6. The primary results were analyzed following the intent-to-treat (ITT) principle with on-treatment (OT) or per-protocol (PP) analyses reported as sensitivity analyses.

Patients had a long-standing history of diabetes (mean duration of 12.7 to 13.9 years) with baseline mean HbA_{1c} ranging from 8.1% (65 mmol/mol) to 8.7% (72 mmol/mol). The prevalence of high CV risk, including history of established CVD and heart failure, is summarized in Table 1. Most study patients were receiving optimal CV risk management at baseline, as shown by a high proportion of patients receiving antihypertensive, lipid-lowering, and antiplatelet medications

Table 1—Trial design and demographics

| Variable | EMPA-REG OUTCOME (N = 7,020) | LEADER (N = 9,340) | SUSTAIN-6 (N = 3,297) |
|--|--|--|---|
| Treatment intervention | Empagliflozin 10 and 25 mg vs. placebo | Liraglutide 0.6–1.8 mg vs. placebo | Semaglutide 0.5–1.0 mg vs. placebo |
| Main inclusion criteria | Preexisting CVD | ≥50 years + preexisting CVD, CKD, HF; ≥60 years + CVD risk factors | ≥50 years + preexisting CVD; ≥60 years + CVD risk factors |
| HbA _{1c} inclusion criteria | 7.0–10.0 | >7.0 | >7.0 |
| Mean age (years) | 63.1 | 64.3 | 64.6 |
| Female | 28 | 36 | 39 |
| White | 72 | 77.5 | 83 |
| Black | 5 | 8.3 | 6.7 |
| Asian | 22 | 10 | 8.3 |
| North America (U.S.) | 20 (17) | 30 (27) | (34) |
| Mean BMI (kg/m ²) | 30.6 | 32.5 | 32.8 |
| Mean baseline HbA _{1c} % (mmol/mol) | 8.1 (65) | 8.7 (72) | 8.7 (72) |
| Diabetes duration >10 years | 57 | Mean 12.8 years | Mean 13.9 years |
| Current cigarette smoker | 13 | 12 | 55 (smoking history) |
| History of hypertension | 94 | 90 | 93 |
| History of CVD | 99 | 81 | 83 |
| Prior MI/stroke or TIA | 47/23 | 31/16 | 33/15 |
| Statin use | 77 | 72 | 73 |
| History of cardiac failure | 10 | 18 | 24 |
| eGFR <60 mL/min/1.73 m ² | 26 | 25 | 28 |
| Completed study | 97 | 96.8 | 98 |
| Vital status known | 99.2 | 99.7 | 99.6 |

Data are % unless otherwise indicated. eGFR was measured according to CKD-EPI criteria. CKD, chronic kidney disease; HF, heart failure; TIA, transient ischemic attack.

(8–10). Because the trials were designed to establish “glycemic equipoise” between the treatment groups (to minimize the confounding impact of differential glycemic effects on CV safety), treatment intensification with other oral hypoglycemic agents or insulin was more prevalent in the placebo group. Over 97% completed the study, and vital status was available in >99% of patients, indicating excellent trial conduct and patient retention. Treatment discontinuation

rates due to adverse events favored empagliflozin in EMPA-REG OUTCOME (17% vs. 19% placebo, *P* < 0.01) and placebo in LEADER (9.5% vs. 7.3% placebo, *P* < 0.001) and were similar across treatment groups in SUSTAIN-6 (20% overall). However, there were a limited number of women (28% to 39%) and nonwhite subjects (15% to 27%) enrolled, raising questions about generalizability. Differences in baseline characteristics of the patient population recruited as well as in trial

design and protocol make it difficult to compare results from these trials and inappropriate to reliably assess relative benefits of therapies.

CARDIOMETABOLIC OUTCOMES

The impact on cardiometabolic factors is shown in Table 2. Small but statistically significant reductions in HbA_{1c}, systolic blood pressure (SBP), diastolic blood pressure (DBP), and weight and small increases in LDL and HDL cholesterol were

Table 2—Impact on cardiometabolic factors

| End point | EMPA-REG OUTCOME | | | LEADER | | | SUSTAIN-6 | | | | |
|--------------------------------|------------------|------|----------------|----------|------|----------------|-----------|-------------|-------------|----------------|---------|
| | Baseline | AD | <i>P</i> value | Baseline | AD | <i>P</i> value | Baseline | AD | | <i>P</i> value | |
| | | | | | | | | 0.5 mg | 1.0 mg | 0.5 mg | 1.0 mg |
| HbA _{1c} % (mmol/mol) | 8.1 (65) | −0.3 | NR | 8.7 (72) | −0.4 | <0.001 | 8.7 (72) | −0.66 (−12) | −1.05 (−15) | <0.0001 | <0.0001 |
| SBP, mmHg | 135 | −4.0 | NR | 136 | −1.2 | <0.001 | 135.6 | −1.3 | −2.6 | NS | <0.001 |
| DBP, mmHg | 77 | −1.0 | NR | 77 | +0.6 | 0.004 | 77 | −0.04 | +0.14 | NS | NS |
| Weight, kg | 86 | −2.0 | NR | 92 | −2.3 | <0.001 | 92.1 | −2.9 | −4.4 | <0.0001 | <0.0001 |
| LDL cholesterol, mg/dL | 86 | +5.3 | NR | 89.5 | −1.6 | 0.02 | 82.3 | −3.3 | −0.8 | <0.05 | NS |
| HDL cholesterol, mg/dL | 44.5 | +2.0 | NR | 45.5 | +0.3 | 0.07 | 43.7 | 0 | +1.7 | NS | <0.001 |
| Heart rate, bpm | 71 | 0.0 | NR | 72 | +3.0 | <0.001 | 72 | +2.0 | +2.5 | <0.0001 | <0.0001 |

AD, absolute difference relative to baseline (− implying lower and + implying higher values with active treatment); NR, not reported; NS, not significant.

observed in EMPA-REG OUTCOME (10). There were no changes observed in heart rate. In contrast, treatment with liraglutide was associated with small but statistically significant increases in DBP and heart rate while lowering HbA_{1c}, SBP, and weight significantly (8). Reductions in HbA_{1c} and weight were greater with semaglutide compared with the other two agents, and semaglutide also lowered SBP and increased heart rate significantly but had small effect on DBP and LDL and HDL cholesterol (9). Overall, the magnitude of treatment effect on cardiometabolic factors was modest and of unclear clinical significance.

CV OUTCOMES

The number of primary end point events were higher than those assumed in the power calculation in all three trials with LEADER yielding 1,306 (vs. 611 assumed), EMPA-REG OUTCOME 772 (vs. 691 assumed), and SUSTAIN-6 254 (vs. 122 assumed) (Table 3). In the placebo arm, the incidence rates of the 3-point MACE in SUSTAIN-6 and EMPA-REG OUTCOME were slightly higher compared with LEADER (44 and 43.9 vs. 39/1,000 patient-years); CV death and all-cause deaths incidence rates were higher in EMPA-REG OUTCOME compared with the other two trials, indicating a higher risk profile. However, the total number of 3-point MACE was highest in LEADER because of the greater number of patients enrolled and longer follow-up (3.5 to 5 years, median 3.8 years).

The primary end point results are summarized in Table 3. The exclusion of postapproval 1.3 risk margin was met in EMPA-REG OUTCOME and LEADER, thereby establishing CV safety for empagliflozin and liraglutide, respectively. In addition, the prespecified criterion for superiority (excluding 1.0 risk margin) was also met in both trials. In SUSTAIN-6, the exclusion of preapproval 1.8 risk margin was met. However, despite the unexpected finding of a significant reduction in risk, it is questionable whether superiority can be reliably inferred because it was not prespecified in the testing hierarchy. It is also debatable whether an exclusion of 1.3 risk margin can be inferred, as the total number of events ($n = 254$) is much lower than required per the FDA guidance ($n = 611$ events) (3). In EMPA-REG OUTCOME, the 4-point MACE, a key secondary outcome that was prespecified

as part of the testing hierarchy, was not significantly reduced. In contrast, the expanded MACE, one of the secondary outcomes, was significantly reduced in LEADER and SUSTAIN-6. However, the expanded MACE was not prespecified in these two trials as part of the testing hierarchy that controlled the type 1 error, so the findings would be better viewed as exploratory or hypothesis generating rather than confirmatory.

In EMPA-REG OUTCOME, there is heterogeneity of treatment effect on the individual components of the primary end point with significant reduction observed in CV death and statistically nonsignificant reduction in nonfatal MI and increase in nonfatal stroke. In contrast, all three components of the primary end point contributed to the reduced risk with liraglutide, and the HR for CV death was statistically significant. In SUSTAIN-6, the reduction in 3-point MACE was driven by a significant reduction in nonfatal stroke with statistically nonsignificant reduction in nonfatal MI and no effect on CV death. A notable difference from the EMPA-REG OUTCOME is that silent MIs were adjudicated in LEADER and SUSTAIN-6, yielding a statistically nonsignificant advantage for liraglutide (absolute risk difference [ARD] -0.3% , HR 0.86) and semaglutide (ARD -0.2% , HR 0.57). Secondary end point of all-cause death was significantly reduced in EMPA-REG OUTCOME and LEADER but not in SUSTAIN-6; hospitalization for heart failure was reduced only in EMPA-REG OUTCOME; and hospitalization for unstable angina was not reduced in any trial. Because of the testing hierarchy used and the lack of control of type 1 error, reduction in heart failure outcomes in EMPA-REG OUTCOME should be viewed as hypothesis generating that requires confirmation in future studies. In contrast, CV death reduction was based on a large number of events (309 in EMPA-REG OUTCOME and 497 in LEADER), was clinically important, and was statistically robust, yielding overwhelming evidence of benefit that does not require confirmation (14).

Sensitivity analyses performed to investigate the robustness of the results are summarized in Table 4. Superiority of CV death but not 3-point MACE was established across all sensitivity analyses in EMPA-REG OUTCOME; four out of six sensitivity analyses overturned

superiority for 3-point MACE, indicating fragility of evidence (12). In contrast, all predefined sensitivity analyses supported the robustness of the primary analysis in LEADER and SUSTAIN-6 (8,9).

Consistent treatment effects were seen across relevant subgroups in EMPA-REG OUTCOME for CV death, although there was some heterogeneity with respect to the primary end point for age and baseline HbA_{1c} (unadjusted $P = 0.01$) (10). A renal function-treatment interaction (unadjusted $P = 0.01$) was observed in LEADER but not EMPA-REG OUTCOME or SUSTAIN-6, i.e., treatment benefit with regards to primary end point was only evident in those with moderate or severe renal impairment (estimated glomerular filtration rate [eGFR] <60 mL/min) (8). A significant interaction (unadjusted $P = 0.04$) was also seen in LEADER for baseline risk of CVD (HR of 1.20 for those aged ≥ 60 years plus risk factors for CVD vs. HR of 0.83 for those aged ≥ 50 years plus established CVD/chronic kidney disease) (8). Such qualitative interactions (i.e., point estimates going in opposite directions) are unreliable and seldom replicable, as evidenced by subgroup findings in SUSTAIN-6 (HR of 1.0 for those with risk factors for CVD vs. HR of 0.72 for those with established CVD, interaction $P = 0.49$). It is important to note that the significant interactions reported in the trials were not adjusted for multiple comparisons, thereby inflating the likelihood of false positive results. Conversely, subgroup analyses lack the statistical power to capture true positive interactions and are thus also prone to false negative results.

The delayed separation of the Kaplan-Meier curves in LEADER (>12 months for CV death and >18 months for all-cause deaths and hospitalization for heart failure) (8) contrasts with the early separation of curves in EMPA-REG OUTCOME (<3 months) (10). Another notable finding is that the favorable CV outcome benefit observed in LEADER and SUSTAIN-6 contrasts with the null results seen with another GLP-1 RA, lixisenatide, in the Evaluation of Lixisenatide in Acute Coronary Syndrome (ELIXA) trial, which enrolled 6,000 patients within 180 days of acute coronary syndrome (7). Although the exact reasons are not clear, this discrepancy might be related to differences in pharmacokinetic and pharmacodynamic properties—lixisenatide is a

Table 3—CV outcomes

| Variable | EMPA-REG OUTCOME | | LEADER | | SUSTAIN-6 | |
|--|------------------------------|------------------------------|------------------------------|------------------------------|------------------------------|-----------------------------|
| | Empagliflozin | Placebo | Liraglutide | Placebo | Semaglutide | Placebo |
| <i>n</i> | 4,687 | 2,333 | 4,668 | 4,672 | 1,648 | 1,649 |
| Follow-up (median, years) | 3.2 | 3.1 | 3.8 | 3.8 | 2.1 | 2.1 |
| Primary outcome (3-point MACE) | 490 (10.5%) 37.4/1,000 PY | 282 (12.1%) 43.9/1,000 PY | 608 (13.0%) 34.1/1,000 PY | 694 (14.9%) 39.1/1,000 PY | 108 (6.6%) 32.4/1,000 PY | 146 (8.9%) 44.4/1,000 PY |
| | HR 0.86 (95% CI 0.74, 0.99) | | HR 0.87 (95% CI 0.78, 0.97) | | HR 0.74 (95% CI 0.58, 0.95) | |
| | <i>P</i> = 0.04 | | <i>P</i> = 0.01 | | <i>P</i> = 0.02 | |
| CV death | 172 (3.7%) 12.4/1,000 PY | 137 (5.9%) 20.2/1,000 PY | 219 (4.7%) 12.1/1,000 PY | 278 (6.0%) 16.1/1,000 PY | 44 (2.7%) 12.9/1,000 PY | 46 (2.8%) 13.5/1,000 PY |
| | HR 0.62 (95% CI 0.49, 0.77) | | HR 0.78 (95% CI 0.66, 0.93) | | HR 0.98 (95% CI 0.65, 1.48) | |
| Nonfatal MI* | 213 (4.5%) 16/1,000 PY | 121 (5.2%) 18.5/1,000 PY | 281 (6.0%) 16/1,000 PY | 317 (6.8%) 18/1,000 PY | 47 (2.9%) 14.0/1,000 PY | 64 (3.9%) 19.2/1,000 PY |
| | HR 0.87 (95% CI 0.70, 1.09) | | HR 0.86 (95% CI 0.73, 1.00) | | HR 0.74 (95% CI 0.51, 1.08) | |
| Nonfatal stroke | 150 (3.2%) 11.2/1,000 PY | 60 (2.6%) 9.1/1,000 PY | 159 (3.4%) 9/1,000 PY | 177 (3.8%) 10/1,000 PY | 27 (1.6%) 8/1,000 PY | 44 (2.7%) 13.1/1,000 PY |
| | HR 1.18 (95% CI 0.89, 1.56) | | HR 0.86 (95% CI 0.71, 1.06) | | HR 0.61 (95% CI 0.38, 0.99) | |
| Key secondary outcome (4-point or expanded MACE) | 599 (12.8%) 46.4/1,000 PY | 333 (14.3%) 52.5/1,000 PY | 948 (20.3%) 53/1,000 PY | 1,062 (22.7%) 60/1,000 PY | 199 (12.1%) 61.7/1,000 PY | 264 (16%) 83.6/1,000 PY |
| | HR 0.89 (95% CI 0.78, 1.01) | | HR 0.88 (95% CI 0.81, 0.96) | | HR 0.74 (95% CI 0.62, 0.89) | |
| Hospitalization for heart failure | 126 (2.7%) 9.4/1,000 PY | 95 (4.1%) 14.5/1,000 PY | 218 (4.7%) 12/1,000 PY | 248 (5.3%) 14/1,000 PY | 59 (3.6%) 17.6/1,000 PY | 54 (3.3%) 16.1/1,000 PY |
| | HR 0.65 (95% CI 0.50, 0.85) | | HR 0.87 (95% CI 0.73, 1.05) | | HR 1.11 (95% CI 0.77, 1.61) | |
| All-cause death | 269 (5.7%) 19.4/1,000 PY | 194 (8.3%) 28.6/1,000 PY | 381 (8.2%) 21/1,000 PY | 447 (9.6%) 25/1,000 PY | 62 (3.8%) 18.2/1,000 PY | 60 (3.6%) 17.6/1,000 PY |
| | HR 0.68 (95% CI 0.57, 0.82) | | HR 0.85 (95% CI 0.74, 0.97) | | HR 1.05 (95% CI 0.74, 1.50) | |
| Hospitalization for unstable angina | 133 (2.8%) 10/1,000 PY | 66 (2.8%) 10/1,000 PY | 122 (2.6%) 7/1,000 PY | 124 (2.7%) 7/1,000 PY | 22 (1.3%) 6.5/1,000 PY | 27 (1.6%) 8/1,000 PY |
| | HR 0.99 (95% CI 0.74, 1.34) | | HR 0.98 (95% CI 0.76, 1.26) | | HR 0.82 (95% CI 0.47, 1.44) | |

Data for incidence (%) and incidence rates per 1,000 patient-years (PY) are shown. 3-point MACE: CV death, nonfatal MI, or nonfatal stroke; 4-point MACE in EMPA-REG OUTCOME (3-point MACE or hospitalization for unstable angina); expanded MACE in LEADER (3-point MACE, coronary revascularization, or hospitalization for unstable angina or heart failure); expanded MACE in SUSTAIN-6 (3-point MACE, coronary or peripheral revascularization, or hospitalization for unstable angina or heart failure). * Silent MIs were screened and adjudicated in LEADER and SUSTAIN-6 but not in EMPA-REG OUTCOME.

Table 4—Sensitivity analyses

| End point | EMPA-REG OUTCOME | | | LEADER | | |
|---|------------------|---------|----------------|--------|---------|----------------|
| | HR | ARD (%) | <i>P</i> value | HR | ARD (%) | <i>P</i> value |
| 3-point MACE (FAS) | 0.86 | −1.6 | 0.04 | 0.87 | −1.9 | 0.01 |
| 3-point MACE (PP) | 0.86 | −1.5 | 0.052 | 0.86 | −1.7 | 0.01 |
| 3-point MACE (OT) | 0.87 | −1.0 | 0.090 | 0.83 | −1.6 | 0.01 |
| 3-point MACE (FAS + silent MIs*) | 0.91 | NR | NS | | | |
| 3-point MACE (FAS – nonassessable deaths) | 0.90 | NR | NS | | | |
| 3-point MACE (imputation for missing data) | 0.86 | NR | <0.05 | | | |
| 3-point MACE (+ all-cause deaths**) | 0.85 | NR | <0.05 | | | |
| CV death (FAS) | 0.62 | −2.2 | <0.001 | 0.78 | −1.3 | 0.007 |
| CV death (OT) | 0.59 | −1.5 | 0.0002 | | | |
| CV death (FAS – nonassessable deaths) | 0.59 | −1.4 | 0.0004 | | | |
| CV death (worst-case missing data analysis) | 0.75 | −1.5 | 0.008 | 0.83 | −1.0 | 0.03 |

For ARD, – implies treatment benefit. OT includes events in FAS \leq 30 days after last intake of trial medication. 3-point MACE includes CV death, nonfatal MI, or nonfatal stroke. *P* values for superiority are shown. For sensitivity analyses of EMPA-REG OUTCOME results: 124/309 (40%) of CV deaths were adjudicated as “nonassessable” but presumed to be CV. Multiple imputation method for missing follow-up was used to assess the impact of missing data on 3-point MACE. Worst-case missing data analysis for CV death assumes all missing subjects on treatment dead ($n = 36$) and on placebo alive ($n = 17$). Worst-case missing data analysis for CV death were estimated for LEADER based on unknown vital status in 12 patients on liraglutide and 17 on placebo. FAS, full analysis set (corresponds to ITT analysis); NR, not reported; NS, not significant ($P > 0.05$). *Silent MIs were screened and adjudicated toward 3-point MACE in LEADER but not in EMPA-REG OUTCOME. Information regarding silent MIs was collected in EMPA-REG OUTCOME based only on ECG criterion (but not reviewed or adjudicated to verify whether it was a silent MI) in 51% patients (3,589 of 7,020) who did not have baseline ECG abnormalities, absence of postbaseline ECG evaluation, or without intervening ECG changes unrelated to event. **There were 135 additional all-cause deaths following time to MACE (51 in placebo and 84 in empagliflozin treatment arms).

once-daily, short-acting prandial GLP-1 RA that acts primarily on postprandial glucose compared with longer-acting liraglutide and semaglutide that act on both fasting and postprandial glucose. Another explanation for the contrasting results might be the trial differences—LEADER enrolled lower-risk patients (placebo arm 3-point MACE incidence rate of 39/1,000 patient-years vs. 64/1,000 patient-years in ELIXA) and had longer follow-up (3.8 vs. 2.1 years). Finally, the GLP-1 RA–induced increase in heart rate does not present an increased CV risk. Previous epidemiological studies suggest that elevated heart rate is independently associated with increased CV morbidity and mortality, however, this relationship might be confounded (15). The exact mechanism for increased heart rate remains unclear (either a direct effect on the sinoatrial node where GLP-1 receptor is known to be expressed or indirectly via modulation of the autonomic nervous system). It remains to be seen whether a pronounced increase in heart rate may be associated with adverse outcomes in vulnerable patients such as those with advanced heart failure.

The strength of evidence as assessed by *P* value, minimum Bayes factor (14,16),

and number needed to treat (NNT) is summarized for EMPA-REG OUTCOME and LEADER in Table 5. The Bayes factor overcomes a key limitation of the *P* value that overestimates the evidence against the null (14,16). For example, the *P* value of 0.038 for 3-point MACE in EMPA-REG OUTCOME translates into a minimum Bayes factor of 0.131, which means the evidence supports the null hypothesis approximately one-eighth as strongly as it does the alternative. This reduces the null probability from 50% pretrial to 10% posttrial. This does not represent strong evidence against the null and thus requires independent confirmation in a subsequent trial (14). In contrast, the null probability for the 3-point MACE in LEADER is reduced from 50% pretrial to 4% posttrial, indicating moderate to strong evidence against the null. For all-cause and CV mortality in EMPA-REG OUTCOME, the nominal *P* value of 0.0001 translates into a Bayes factor of 0.0006 (1/1,815) and 0.0004 (1/2,358), respectively, which reduces the extremely skeptical prior null probability of 95% to <0.5% posttrial, indicating very strong evidence against the null. For all end points except 3-point MACE, the

evidence is relatively stronger in support of empagliflozin compared with liraglutide as assessed by the Bayes factor. This is also consistent with the lower NNTs in favor of empagliflozin.

MICROVASCULAR OUTCOMES

A secondary outcome of the EMPA-REG OUTCOME and LEADER trials was a composite renal and retinal microvascular outcome (Table 6) (8,10). In SUSTAIN-6, retinopathy and nephropathy outcomes were assessed separately (9). Treatment with empagliflozin significantly reduced the composite microvascular outcome and the renal outcome (17). Of note, the latter was neither prespecified nor adjudicated in EMPA-REG OUTCOME (12). Composite retinal outcomes have not been reported so far for this trial, although the individual component results are published (17). In the LEADER trial, the incidence of composite microvascular outcome (both prespecified and adjudicated) was lower with liraglutide (8), mainly due to a significantly lower rate of nephropathy events. As with the other two treatment interventions, the renal outcome was also favorably impacted by semaglutide, primarily driven by a “softer” component such as persistent macroalbuminuria that is of unclear clinical relevance. The incidence of retinal outcome was nonsignificantly higher with liraglutide treatment. A significant increase in the incidence of retinal outcome was also observed with semaglutide in SUSTAIN-6, with >80% of events occurring in subjects with evidence of preexisting retinopathy at baseline. This finding raises the question of retinal monitoring with the use of GLP-1 RA. It is not clear whether these events can be attributed to rapid glucose lowering on progression of diabetic retinopathy as has been previously described (18). This merits careful evaluation in future clinical trials and postmarketing registries.

MECHANISM OF CV RISK REDUCTION

The exact mechanism underlying CV benefit in any one of these trials is not clear. To what extent the favorable effects on cardiometabolic factors such as blood pressure, body weight, or glycemic control (a Steno-2–like effect) contribute to the overall benefit in EMPA-REG OUTCOME remains unclear. Given the rapid onset of treatment effect (curves

Table 5—Evaluating strength of evidence of CV outcomes using Bayes factor and NNT

| End point | Trial | Effect size | | P value (z score) | Minimum Bayes factor | Decrease in probability of null hypothesis, % | | Strength of evidence |
|-----------------------------------|------------------|-------------|--------------|----------------------|----------------------|--|--------------------|----------------------|
| | | HR | NNT | | | From | To no less than | |
| 3-point MACE | EMPA-REG OUTCOME | 0.86 | 63 | 0.038 (2.02) | 0.131 | 95 | 54 | Moderate |
| | | | | | | 75 | 28 | |
| | | | | | | 50 | 12 | |
| LEADER | 0.87 | 66 | 0.01 (2.55) | 0.038 | 95 | 42 | Moderate to strong | |
| | | | | | 75 | 10 | | |
| | | | | | 50 | 4 | | |
| All-cause deaths | EMPA-REG OUTCOME | 0.68 | 39 | 0.0001 (3.94) | 0.0006 | 95 | 0.49 | Very strong |
| | | | | | | 75 | 0.16 | |
| | | | | | | 50 | 0.06 | |
| LEADER | 0.85 | 98 | 0.017 (2.39) | 0.057 | 95 | 52 | Moderate to strong | |
| | | | | | 75 | 15 | | |
| | | | | | 50 | 5 | | |
| CV deaths | EMPA-REG OUTCOME | 0.62 | 45 | 0.0001 (3.87) | 0.0004 | 95 | 0.38 | Very strong |
| | | | | | | 75 | 0.13 | |
| | | | | | | 50 | 0.04 | |
| LEADER | 0.78 | 104 | 0.007 (2.71) | 0.024 | 95 | 31 | Strong | |
| | | | | | 75 | 7 | | |
| | | | | | 50 | 2 | | |
| Hospitalization for heart failure | EMPA-REG OUTCOME | 0.65 | 71 | 0.0017 (2.93) | 0.0137 | 95 | 11 | Strong |
| | | | | | | 75 | 4 | |
| | | | | | | 50 | 1 | |
| LEADER | 0.87 | NE | 0.15 (1.42) | 0.357 | 95 | 87 | Weak | |
| | | | | | 75 | 52 | | |
| | | | | | 50 | 26 | | |

Bayes theorem: posterior odds = prior odds × evidence (Bayes factor). Bayes factor = probability (data/ H_0)/probability (data/ H_1) (likelihood ratio); H_0 = null hypothesis; H_1 = alternative hypothesis. Minimum Bayes factor = $\exp(-0.5z^2)$. Odds = probability/(1 - probability). Probability = odds/(1 + odds). NNT to prevent one event over 3 years, calculated as inverse of absolute risk difference based on Kaplan-Meier curve estimates, is only reported for statistically significant differences. NE, not estimated because of lack of statistically significant difference.

separate as early as 2–3 months) and modest effects on these factors, the impact if any might be small. And it is unlikely to be mediated by an antiatherothrombotic effect, given the lack of effect on MI and stroke. In the Steno-2 trial of multifactorial intervention, the 53% reduction in death and MI was observed at a median follow-up of 7.8 years with delayed separation of the Kaplan-Meier curves and associated with larger changes in HbA_{1c} (−1.1% vs. −0.3%), SBP (−15 vs. −4 mmHg), DBP (−5 vs. −1 mmHg), LDL cholesterol (−45 vs. +5.3 mg/dL), and weight (+1.1 vs. −2.0 kg) (19). The observation that hospitalization for heart failure was reduced by 35% and that half of the CV mortality advantage was driven by reduction in worsening heart failure and sudden cardiac death support a possible hemodynamic or antiarrhythmic effect, possibly mediated via modulation of the renin-angiotensin-aldosterone system pathway. Inhibition of sodium absorption in the kidney by an SGLT2 inhibitor

results in natriuresis leading to osmotic diuresis with attendant reduction in plasma volume and SBP and amelioration of renal hyperfiltration, thereby improving renal function (17). This might suggest a potential role of empagliflozin influencing the cardiorenal axis in mitigating CV risk. Of note, patients with history of heart failure (10% prevalence) experienced similar CV outcome benefits as those without heart failure in EMPA-REG OUTCOME (20). Accordingly, the sponsor has announced plans to conduct two trials to evaluate empagliflozin in patients with chronic heart failure and preserved ejection fraction (HFpEF) and heart failure and reduced ejection fraction (HFrEF) both with or without type 2 diabetes (21,22). Others have suggested a “thrifty substrate” hypothesis wherein under conditions of mild hyperketonemia (as seen with SGLT2 inhibitor), β -hydroxybutyrate (“superfuel”) is preferentially utilized by the heart in preference to fatty acids, resulting in improved

transduction of VO₂ into myocardial efficiency (23). This, together with increased hematocrit, improves oxygen delivery, further enhancing myocardial performance (23). Future studies aimed at these targets should help clarify the mechanistic pathways.

The consistent treatment effect on MI and stroke and delayed separation of Kaplan-Meier curves seen in the trials of liraglutide and semaglutide is more suggestive of a potential antiatherothrombotic effect mediated via favorable impact on cardiometabolic factors. In addition to glycemic control and weight loss, GLP-1 RA improve other possible risk factors including blood pressure, inflammatory markers, insulin sensitivity, and lipid profile and delay the progression of atherosclerotic disease (24). Animal studies have indicated that GLP-1 receptor activation in heart tissue may have benefits including improved left ventricular (LV) function and protection from ischemic reperfusion injury (24). However, three

Table 6—Microvascular outcomes

| End point | EMPA-REG OUTCOME | | LEADER | | SUSTAIN-6 | |
|---|-----------------------------|-------------------|-----------------------------|------------------|-----------------------------|------------------|
| | Empagliflozin | Placebo | Liraglutide | Placebo | Semaglutide | Placebo |
| Composite microvascular outcome (renal plus retinal events) | 577/4,132 (14%) | 424/2,068 (20.5%) | 355/4,668 (7.6%) | 416/4,672 (8.9%) | NA | NA |
| | HR 0.62 (95% CI 0.54, 0.70) | | HR 0.84 (95% CI 0.73, 0.97) | | NA | |
| | <i>P</i> < 0.001 | | <i>P</i> = 0.02 | | | |
| New or worsening nephropathy (composite renal outcome) | 525/4,124 (12.7%) | 388/2,061 (18.8%) | 268/4,668 (5.7%) | 337/4,672 (7.2%) | 62/1,648 (3.8%) | 100/1,649 (6.1%) |
| | HR 0.61 (95% CI 0.53, 0.70) | | HR 0.78 (95% CI 0.67, 0.92) | | HR 0.64 (95% CI 0.46, 0.88) | |
| | <i>P</i> < 0.0001 | | <i>P</i> = 0.003 | | <i>P</i> = 0.005 | |
| New-onset persistent macroalbuminuria | 459/4,091 (11.2%) | 330/2,033 (16.2%) | 161/466 (3.4%) | 215/4,672 (4.6%) | 44/1,648 (2.7%) | 81/1,649 (4.9%) |
| | HR 0.62 (95% CI 0.54, 0.72) | | HR 0.74 (95% CI 0.60, 0.91) | | HR 0.54 (95% CI 0.37, 0.77) | |
| | <i>P</i> < 0.0001 | | | | <i>P</i> = 0.001 | |
| Doubling of serum creatinine | 70/4,645 (1.5%) | 60/2,323 (2.6%) | 87/4,668 (1.9%) | 97/4,672 (2.1%) | 18/1,648 (1.1%) | 14/1,649 (0.8%) |
| | HR 0.56 (95% CI 0.39, 0.79) | | HR 0.88 (95% CI 0.66, 1.18) | | HR 1.28 (95% CI 0.64, 2.58) | |
| | <i>P</i> = 0.0009 | | | | <i>P</i> = 0.48 | |
| Renal replacement therapy | 13/4,687 (0.3%) | 15/2,333 (0.6%) | 56/4,668 (1.2%) | 64/4,672 (1.4%) | 11/1,648 (0.7%) | 12/1,649 (0.7%) |
| | HR 0.45 (95% CI 0.21, 0.97) | | HR 0.87 (95% CI 0.61, 1.24) | | HR 0.91 (95% CI 0.40, 2.07) | |
| | <i>P</i> = 0.041 | | | | <i>P</i> = 0.83 | |
| Renal death | 3/4,687 (0.1%) | 0/2,333 (0.0%) | 8/4,668 (0.2%) | 5/4,672 (0.1%) | NA | NA |
| | NA | | HR 1.59 (95% CI 0.52, 4.87) | | NA | |
| Retinopathy (composite outcome) | NR | NR | 106/4,668 (2.3%) | 29/4,672 (2.0%) | 50/1,648 (3.0%) | 29/1,649 (1.8%) |
| | NR | | HR 1.15 (95% CI 0.87, 1.52) | | HR 1.76 (95% CI 1.11, 2.78) | |
| | | | <i>P</i> = 0.33 | | <i>P</i> = 0.02 | |

Nephropathy is defined as the new onset of macroalbuminuria (urine albumin-to-creatinine ratio >300 mg/g) or a doubling of the serum creatinine level and an eGFR of ≤ 45 mL per minute per 1.73 m^2 , the need for continuous renal-replacement therapy, or death from renal disease. Retinopathy is defined as the need for retinal photocoagulation or treatment with intravitreal agents, vitreous hemorrhage, or the onset of diabetes-related blindness. Nephropathy was a prespecified exploratory adjudicated outcome in LEADER and SUSTAIN-6 but not in EMPA-REG OUTCOME. NA, not available; NR, not reported.

smaller, placebo-controlled trials have failed to demonstrate favorable effects of liraglutide on LV systolic function and exercise capacity in patients with diabetes and coronary artery disease (25), on myocardial energetics and posthospitalization clinical stability in patients (with or without diabetes) with advanced heart failure and reduced LV ejection fraction (26), or on LV ejection fraction in patients with chronic heart failure (with or without diabetes) and reduced LV ejection fraction (27). Patients with heart failure and reduced LV ejection fraction treated with liraglutide suffered numerically more serious adverse cardiac events than those on placebo, including death or hospitalization for heart failure and arrhythmia (26,27). It is not clear whether the adverse cardiac events are related to the increase in heart rate with liraglutide.

The imbalance in the new use of insulin with two- to threefold greater use in the placebo arm of the three trials is

unlikely to contribute to the treatment effect. While observational studies associate insulin use with increased CV risk, a large-scale, randomized, controlled trial showed that insulin glargine, as compared with placebo, was not associated with an increased CV risk (28).

KEY DISCUSSION POINTS AT THE EMDAC PANEL FOR sNDA FOR EMPAGLIFLOZIN

On 28 June 2016, the FDA convened an EMDAC panel to discuss the supplemental New Drug Application (sNDA) seeking expanded indication for empagliflozin to reduce the incidence of CV death in adult patients with type 2 diabetes and established CVD. The panel narrowly voted 12–11 in favor of granting the claim. Although the panel agreed unanimously that the trial had successfully established CV safety of empagliflozin, there were differing opinions regarding establishment of cardiovascular benefit. At the crux of

the debate was whether a single trial designed to demonstrate CV “safety” for a 3-point MACE primary end point could support a label claim of benefit for the secondary end point of CV death reduction based on “substantial evidence of effectiveness,” especially when there is no regulatory precedence for such an action in this space.

The FDA reviewers and some panel members raised issues with the high number of CV deaths characterized as “nonassessable” but presumed to be CV deaths. About 40% of CV deaths (124/309) and 26.8% of all-cause deaths (124/463) were categorized as nonassessable (12). Typically, nonassessable deaths make up no more than 10–20% of all deaths in CV trials (12). Nearly one-third of the absolute reduction in CV death (0.8% out of 2.2%) was driven by reduction in these nonassessable deaths (10,12). Excluding nonassessable deaths from the primary analysis overturned

superiority for 3-point MACE from HR 0.86 (95% CI 0.74, 0.99) to 0.90 (0.77, 1.06) (12) (Table 4) but had no substantial impact on CV death benefit, which remained robust: 0.59 (0.44, 0.79) vs. 0.62 (0.49, 0.77) (12) (Table 4). In a post hoc analysis, addition of all-cause deaths to the 3-point MACE (51 additional events with placebo and 84 additional events with empagliflozin) preserved the CV benefit of empagliflozin, HR 0.85 (95% CI 0.74, 0.97) (12). Some committee members argued that even though both CV and all-cause deaths were prespecified as secondary end points, there was no prespecified α -adjustment for these end points and they were not formally included in the statistical hierarchical testing strategy, which included a stepwise evaluation of noninferiority followed by superiority of 3-point and 4-point MACE. Because superiority of the 4-point MACE was clearly not met ($P = 0.079$) and the reliability of superiority of the 3-point MACE was questionable as it was not established in 4/6 sensitivity analyses (12) (Table 4), all subsequent analyses, including deaths, are deemed exploratory (“hypothesis generating”), requiring confirmation in subsequent trials. Others felt that CV and all-cause death trumped all end points and that CV and all-cause death reduction was based on a large number of events (309 CV and 463 all-cause deaths), it was clinically important, statistically robust yielding overwhelming evidence of benefit (“proof beyond reasonable doubt”) that remained significant even after adjustments for missing data assuming the worst-case scenario (Table 4) and multiple comparisons. Furthermore, the reduction in deaths was consistently seen with both 10 and 25 mg doses, which is akin to two separate trials embedded within one trial. Thus, the quality and the quantity of evidence was sufficient to support the FDA’s “substantial” evidence criterion of effectiveness based on a single trial, i.e., a single multicenter study of excellent design providing highly reliable and statistically strong evidence of an important clinical benefit (14). A similar conclusion was also reached independently by the reviewers within the FDA’s Division of Cardiovascular and Renal Products (12).

Another concern raised was that “silent MIs” were not adjudicated toward the primary end point (12). Silent MIs are common in diabetes, accounting for up to

one-third of all MIs, and they confer increased mortality risk (29,30). Thus, there are legitimate arguments for including silent MIs in the overall adjudication of MIs in CV outcome trials. However, because of challenges regarding accurate ascertainment, especially the inability to capture the precise timing and therefore reliably estimate time to event, this end point has not been consistently included in CV outcome trials. For example, some trials included silent MIs toward the overall adjudication of MIs (8,9,29,31,32), but other contemporary trials did not (4–7), thereby raising questions regarding standardization of end points in CV outcome trials. In contrast to LEADER, in EMPA-REG OUTCOME, silent MIs were assessed (but not adjudicated toward either primary outcome or MI) based only on electrocardiogram (ECG) criteria in about half of the patients who did not have baseline ECG abnormalities, or who had baseline or postbaseline ECGs available for evaluation, or where intervening ECG changes were unrelated to event (12). As such, these results are subject to missing data. Nonetheless, these events occurred in 15/1,211 (1.2%) in the placebo group and 38/2,378 (1.6%) in the empagliflozin group, yielding an HR of 1.28 (95% CI 0.70, 2.33) (12). Counting these “silent MI” events in an exploratory analysis would overturn statistical significance for the 3-point MACE: HR 0.91 (95% CI 0.73, 1.13) (12) (Table 4).

Several panelists argued that lack of a clear mechanistic explanation underlying CV benefit in EMPA-REG OUTCOME is a major limitation, calling into question the validity of the findings. This is a rather uncharitable criticism, as outcome trials are not designed to yield mechanistic insights (14). These should be explored in future investigations.

The increase in the hazard for stroke with empagliflozin, although not statistically significant, was also a subject of deliberation. In the ITT analysis, the HR for total stroke is 1.18 (95% CI 0.89, 1.56). In some subgroups, such as those enrolled from Europe (representing 41% of the overall trial cohort), on loop diuretics, with history of atrial fibrillation, or those with baseline HbA_{1c} >8.5%, the HR exceeds 2.0 (12). Disability associated with stroke was not formally assessed. However, fatal strokes, a proxy for large disabling strokes, occurred infrequently—27 out of 233 total strokes (11.6%)—and the

HR for fatal stroke of 0.72 (95% CI 0.33, 1.55), $P = 0.41$ (12), is reassuring. On-treatment analysis that limits assessment to events occurring within 30 days of last study drug yields an HR of 1.08 (95% CI 0.81, 1.45) (12). No significant associations were observed between stroke and changes in hematocrit and blood pressure or volume depletion. Thus, the clinical relevance of the numerical imbalance in stroke is unclear. While the finding might represent a play of chance, it remains a potential concern for this and other products within the drug class (12) and therefore merits careful assessment in future studies.

Panel members largely agreed that the reductions in heart failure and adverse renal outcomes shown in EMPA-REG OUTCOME were not sufficient (or controlled for type 1 error) to establish conclusive benefits and therefore they were deemed not actionable for regulatory decision making.

On 2 December 2016, the FDA announced approval for expanded indication of empagliflozin “to reduce the risk of CV death in adult patients with type 2 diabetes mellitus and cardiovascular disease” (11), making empagliflozin the first antidiabetes drug to receive a CV risk reduction claim. On the basis of the results of the EMPA-REG OUTCOME and LEADER trials, the American Diabetes Association’s *Standards of Medical Care in Diabetes—2017* recommends adding empagliflozin (or liraglutide) for patients with established CVD to reduce the risk of mortality (33). The Canadian Diabetes Association Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada had previously endorsed addition of empagliflozin to antihyperglycemic therapy to reduce the risk of CV and all-cause mortality in people with clinical CVD in whom glycemic targets are not met (34). More research is needed to elucidate the underlying mechanisms and confirm whether CV benefits are a class effect or whether the benefits persist in patients without established CVD or are evident even in patients without diabetes.

CONCLUSIONS

Overall, the results of the EMPA-REG OUTCOME and LEADER trials represent a clinical breakthrough representing the first two antidiabetes interventions to unequivocally show CV risk

reduction in patients with type 2 diabetes. Although SUSTAIN-6 unexpectedly yielded a significant reduction in MACE in the preapproval trial, it is questionable whether an inference of superiority is reliable or credible. The CV benefit is likely unrelated to glucose-lowering efficacy of these drugs. The mean on-treatment HbA_{1c} level in both these trials ranged from 7.6% (60 mmol/mol) and 8.3% (67 mmol/mol), respectively. In contrast, previous trials demonstrated that lowering HbA_{1c} levels to less than 7% was not associated with CV benefits compared with less intensive glycemic control (32,35). Questions have also been raised regarding whether intensive glucose control in type 2 diabetes yields unequivocal evidence of benefit on hard microvascular complications, such as vision loss or renal failure (35). These findings raise questions regarding whether targeting glycemic control should remain the principal regulatory criterion for marketing authorization for antidiabetes drugs or continue to be the focus of guideline recommendations, which currently promote individualized glycemic targets for patients based on their comorbidities, propensity for hypoglycemia, and capacity to carry out the treatment plan (33).

The lack of CV safety signals in all the seven trials that have published results so far calls into question the wisdom of a default approach that assumes all antidiabetes drugs are suspected of CV harm unless proven otherwise (36). Perhaps a more selective and targeted strategy that is informed by adverse signals observed during the preclinical phase of drug development, plausible mechanisms of risk, or a known class effect would offer a more enlightened and resource-sensitive approach to assessment of CV safety of antidiabetes drugs (37). While one might argue that the results of these trials vindicate the 2008 guidance, it is perhaps time to move beyond the restricted focus of ruling out unacceptable CV harm in high-risk patients over short-term follow-up (a somewhat artificial scenario mandated by the guidance with limited generalizability) to designing pragmatic trials aimed at yielding tangible long-term benefit in microvascular and macrovascular outcomes in lower-risk patients who are more representative of the patients encountered in daily clinical practice. This will no doubt be a time-consuming and resource-heavy endeavor,

but it would be worth the investment to fully understand the impact of treatment options on all relevant outcomes. Just because a therapy lowers glucose does not necessarily mean that it has other clinically beneficial effects.

Duality of Interest. S.K. has a consultant or advisory relationship with Boehringer Ingelheim, sponsor of empagliflozin; Eli Lilly, collaborator with Boehringer Ingelheim for empagliflozin; and Novo Nordisk, sponsor of liraglutide and semaglutide. No other potential conflicts of interest relevant to this article were reported.

References

1. Michael Specter. The danger of science denial. Available from https://www.ted.com/talks/michael_specter_the_danger_of_science_denial. Accessed 13 May 2017
2. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes [published correction appears in *N Engl J Med* 2007;357:100]. *N Engl J Med* 2007;356:2457–2471
3. U.S. Food and Drug Administration. Guidance for industry: diabetes mellitus—evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes [Internet]. Silver Spring, MD, 2008. Available from <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf>. Accessed 11 April 2017
4. White WB, Cannon CP, Heller SR, et al.; EXAMINE Investigators. Alogliptin after acute coronary syndrome in patients with type 2 diabetes. *N Engl J Med* 2013;369:1327–1335
5. Scirica BM, Bhatt DL, Braunwald E, et al.; SAVOR-TIMI 53 Steering Committee and Investigators. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *N Engl J Med* 2013;369:1317–1326
6. Green JB, Bethel MA, Armstrong PW, et al.; TECOS Study Group. Effect of sitagliptin on cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2015;373:232–242
7. Pfeffer MA, Claggett B, Diaz R, et al.; ELIXA Investigators. Lixisenatide in patients with type 2 diabetes and acute coronary syndrome. *N Engl J Med* 2015;373:2247–2257
8. Marso SP, Daniels GH, Brown-Frandsen K, et al.; LEADER Steering Committee; LEADER Trial Investigators. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2016;375:311–322
9. Marso SP, Bain SC, Consoi A, et al.; SUSTAIN-6 Investigators. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. *N Engl J Med* 2016;375:1834–1844
10. 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
11. U.S. Food and Drug Administration. JARDIANCE (empagliflozin) tablets, for oral use: highlights of prescribing information. Available from http://www.accessdata.fda.gov/drugsatfda_docs/label/2016/204629s008lbl.pdf. Accessed 30 January 2017
12. U.S. Food and Drug Administration. FDA briefing document: Endocrine and Metabolic Drug Advisory Committee Meeting, June 28, 2016. Available from <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM508422.pdf>. Accessed 30 January 2017
13. Marso SP, Lindsey JB, Stolker JM, et al. Cardiovascular safety of liraglutide assessed in a patient-level pooled analysis of phase 2: 3 liraglutide clinical development studies. *Diab Vasc Dis Res* 2011;8:237–240
14. Kaul S. Is the mortality benefit with empagliflozin in type 2 diabetes mellitus too good to be true? *Circulation* 2016;134:94–96
15. Perret-Guillaume C, Joly L, Benetos A. Heart rate as a risk factor for cardiovascular disease. *Prog Cardiovasc Dis* 2009;52:6–10
16. Goodman SN. Toward evidence-based medical statistics. 2: the Bayes factor. *Ann Intern Med* 1999;130:1005–1013
17. Wanner C, Inzucchi SE, Lachin JM, et al.; EMPA-REG OUTCOME Investigators. Empagliflozin and progression of kidney disease in type 2 diabetes. *N Engl J Med* 2016;375:323–334
18. Shurter A, Genter P, Ouyang D, Ipp E. Euglycemic progression: worsening of diabetic retinopathy in poorly controlled type 2 diabetes in minorities. *Diabetes Res Clin Pract* 2013;100:362–367
19. Gaede P, Vedel P, Larsen N, Jensen GVH, Parving H-H, Pedersen O. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. *N Engl J Med* 2003;348:383–393
20. Fitchett D, Zinman B, Wanner C, et al.; EMPA-REG OUTCOME® trial investigators. Heart failure outcomes with empagliflozin in patients with type 2 diabetes at high cardiovascular risk: results of the EMPA-REG OUTCOME® trial. *Eur Heart J* 2016;37:1526–1534
21. EMPagliflozin outcome tRial in Patients With chrOnic heaRt Failure With Reduced Ejection Fraction (EMPEROR-Reduced). *ClinicalTrials.gov* identifier NCT03057977. Available from <https://clinicaltrials.gov/ct2/show/NCT03057977>. Accessed 9 April 2017
22. EMPagliflozin outcome tRial in Patients With chrOnic heaRt Failure With Preserved Ejection Fraction (EMPEROR-Preserved). *ClinicalTrials.gov* identifier NCT03057951. Available from <https://clinicaltrials.gov/ct2/show/NCT03057951>. Accessed 9 April 2017
23. Ferrannini E, Mark M, Mayoux E. CV protection in the EMPA-REG OUTCOME trial: a “thrifty substrate” hypothesis. *Diabetes Care* 2016;39:1108–1114
24. Ussher JR, Drucker DJ. Cardiovascular actions of incretin-based therapies. *Circ Res* 2014;114:1788–1803
25. Kumarathurai P, Anholm C, Nielsen OW, et al. Effects of the glucagon-like peptide-1 receptor agonist liraglutide on systolic function in patients with coronary artery disease and type 2 diabetes: a randomized double-blind placebo-controlled crossover study. *Cardiovasc Diabetol* 2016;15:105
26. Margulies KB, Hernandez AF, Redfield MM, et al.; NHLBI Heart Failure Clinical Research Network. Effects of liraglutide on clinical stability among patients with advanced heart failure and

- reduced ejection fraction: a randomized clinical trial. *JAMA* 2016;316:500–508
27. Jorsal A, Kistorp C, Holmager P, et al. Effect of liraglutide, a glucagon-like peptide-1 analogue, on left ventricular function in stable chronic heart failure patients with and without diabetes (LIVE)-a multicentre, double-blind, randomised, placebo-controlled trial. *Eur J Heart Fail* 2017;19:69–77
28. Gerstein HC, Bosch J, Dagenais GR, et al.; ORIGIN Trial Investigators. Basal insulin and cardiovascular and other outcomes in dysglycemia. *N Engl J Med* 2012;367:319–328
29. Burgess DC, Hunt D, Li L, et al. Incidence and predictors of silent myocardial infarction in type 2 diabetes and the effect of fenofibrate: an analysis from the Fenofibrate Intervention and Event Lowering in Diabetes (FIELD) study. *Eur Heart J* 2010;31:92–99
30. Zhang ZM, Rautaharju PM, Prineas RJ, et al. Race and sex differences in the incidence and prognostic significance of silent myocardial infarction in the Atherosclerosis Risk in Communities (ARIC) Study. *Circulation* 2016;133:2141–2148
31. Dormandy JA, Charbonnel B, Eckland DJ, et al.; PROactive Investigators. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. *Lancet* 2005;366:1279–1289
32. Gerstein HC, Miller ME, Genuth S, et al.; ACCORD Study Group. Long-term effects of intensive glucose lowering on cardiovascular outcomes. *N Engl J Med* 2011;364:818–828
33. American Diabetes Association. *Standards of Medical Care in Diabetes—2017*. *Diabetes Care* 2017;40(Suppl. 1):S1–S135
34. Diabetes Canada. Pharmacologic management of type 2 diabetes: November 2016 interim update. Available from http://guidelines.diabetes.ca/browse/chapter13_2016. Accessed 30 January 2017
35. Rodríguez-Gutiérrez R, Montori VM. Glycemic control for patients with type 2 diabetes mellitus: our evolving faith in the face of evidence. *Circ Cardiovasc Qual Outcomes* 2016;9:504–512
36. Smith RJ, Goldfine AB, Hiatt WR. Evaluating the cardiovascular safety of new medications for type 2 diabetes: time to reassess? *Diabetes Care* 2016;39:738–742
37. Committee for Medicinal Products for Human Use. *Guideline on Clinical Investigation of Medicinal Products in the Treatment or Prevention of Diabetes Mellitus*. London, European Medicines Society, 2012