Diabetes mellitus is a risk factor for coronary heart disease, ischemic stroke, and peripheral arterial occlusion (macrovascular disease), as well as chronic kidney disease, neuropathy, and retinopathy (microvascular disease). Antihyperglycemic treatments reduce the risk of microvascular conditions, but their effect on macrovascular events is uncertain (1, 2). Years ago, the University Group Diabetes Program suggested that sulfonylureas and phenformin might increase the risk of major adverse cardiovascular events (3). After similar concerns were raised about two PPAR-γ agonists (4, 5), the FDA convened an Advisory Committee in 2008 to discuss how to evaluate the cardiovascular safety of diabetes drugs (6–8).

Recommendations of the Advisory Committee The Committee recommended a two-stage evaluation: (a) a preapproval analysis to exclude an unacceptable (80%) increase in cardiovascular risk, and (b) a postapproval trial to exclude a 30% increase in risk (9). The Committee focused on comparisons of the study drug to treatments comprising the standard of care, rather than on whether differences in glycemic control would lead to different outcomes. Patients in both treatment arms were to be controlled to comparable levels of HbA1c. The Committee acknowledged the possibility that HbA1c-lowering might decrease cardiovascular risk “despite the failure of trials to show that benefit.” Further, the Committee recommended that patients in both treatment arms should be managed to achieve similar profiles for known cardiovascular [...]

Find the latest version:
http://jci.me/99820/pdf
Diabetes mellitus is a risk factor for coronary heart disease, ischemic stroke, and peripheral arterial occlusion (macrovascular disease), as well as chronic kidney disease, neuropathy, and retinopathy (microvascular disease). Antihyperglycemic treatments reduce the risk of microvascular conditions, but their effect on macrovascular events is uncertain (1, 2). Years ago, the University Group Diabetes Program suggested that sulfonylureas and phenformin might increase the risk of major adverse cardiovascular events (3). After similar concerns were raised about two PPARγ agonists (4, 5), the FDA convened an Advisory Committee in 2008 to discuss how to evaluate the cardiovascular safety of diabetes drugs (6–8).

Recommendations of the Advisory Committee
The Committee recommended a two-stage evaluation: (a) a preapproval analysis to exclude an unacceptable (80%) increase in cardiovascular risk, and (b) a postapproval trial to exclude a 30% increase in risk (9). The Committee focused on comparisons of the study drug to treatments comprising the standard of care, rather than on whether differences in glycemic control would lead to different outcomes. Patients in both treatment arms were to be controlled to comparable levels of HbA1c. The Committee acknowledged the possibility that HbA1c-lowering might decrease cardiovascular risk “despite the failure of trials to show that benefit.” Further, the Committee recommended that patients in both treatment arms should be managed to achieve similar profiles for known cardiovascular risk factors including blood pressure and LDL-cholesterol.

The Committee also considered ethical issues. Saul Genuth (10) recommended “not to allow any participant in the trial to be...continuously above some ethically acceptable level that exposes them to microvascular risk...We can’t let people go for four or five years with hemoglobin A1cs above...8% for sure” (7, 8). Thirty years earlier, the Belmont Report concluded that risks to human subjects in biomedical research should be reduced to those necessary to achieve the research objective (11). Thus, in conducting a cardiovascular outcome study, the Committee recommended that patients should not be exposed to increased risk of microvascular complications due to suboptimal glycemic control.

Results from recent outcome studies
Table 1 summarizes data from 12 recent cardiovascular outcome studies conducted with five classes of drugs (thiazolidinediones, dipeptidylpeptidase-4 inhibitors, SGLT2 inhibitors, GLP1 receptor agonists, and a dopamine D2 receptor agonist) (12–25). Six studies met statistical criteria for superiority with respect to a composite endpoint for major adverse cardiovascular events (cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke). Unfortunately, these trials have not always conformed to the principles advocated by the FDA Advisory Committee.

Study design: selection of therapy for the comparator arm. The RECORD study compared rosiglitazone to a combination of metformin and a sulfonylurea. The Committee warned against using comparator drugs such as sulfonylureas for which prior studies suggested the possibility of increased cardiovascular risk. In trials other than RECORD, the study drug was compared to placebo (Table 1). For the most part, management of diabetes and cardiovascular disease was left to the discretion of the “responsible physician.” While some studies incorporated “rescue” criteria and recommended preferred treatments for excessive glycemia or blood pressure, the study design did not ensure that patients in the placebo arm achieved comparable control of HbA1c or blood pressure.

Ethical implications: glycemic control. While the sitagliptin study (TECOS) enrolled patients with relatively good glycemic control (mean baseline HbA1c of 7.2%), other studies enrolled some patients with suboptimal glycemic control (mean baseline HbA1c levels ranging from 7.7% to 8.7%). As expected, patients treated with study drugs experienced prompt reductions in mean HbA1c levels, but placebo-treated patients experienced little or no improvement in mean HbA1c — despite the glycemic rescue provisions noted above. Thus, patients on study drugs had substantially lower mean HbA1c levels in all the studies — with the exception of the TECOS study where both groups of patients were relatively well controlled even at baseline. For most of the study duration, mean HbA1c levels for placebo-treated patients hovered around 8% in studies of empagliflozin, canagliflozin, liraglutide, exenatide, alogliptin, and saxagliptin (12, 13, 15, 17, 19, 22). In the SUSTAIN-6 study with semaglutide, patients in the comparator arm exhibited even higher mean HbA1c levels (i.e., ~8.5%) (21). Inasmuch as a substantial percentage of patients had HbA1c levels above the mean, many patients were permitted...
to have HbA1c levels exceeding what Dr. Genuith defined as “ethically acceptable levels” for as long as five years (6–8), and were therefore potentially exposed to increased risk of microvascular disease.

**How should these results be interpreted?**

These ethical shortcomings also introduce challenges in data interpretation. For example, based on secondary analyses of secondary outcomes, empagliflozin, canagliflozin, and liraglutide have been reported to slow progression of diabetic kidney disease (16, 17, 20). Nevertheless, because of the substantially lower mean HbA1c levels in patients treated with those study drugs, the question remains to what extent the beneficial effect on kidney disease is attributable to improved glycemic control in the treatment arm, rather than a putative benefit related to the drug’s specific mechanism of action.

Treatment guidelines are beginning to cite recent cardiovascular outcome studies as evidence in favor of selecting specific drugs (e.g., empagliflozin, liraglutide, semaglutide, and canagliflozin) as first choices to be added to metformin therapy in patients with established atherosclerotic cardiovascular disease (26). This conclusion seems premature. Strictly speaking, these cardiovascular outcome studies demonstrated improved cardiovascular outcomes when compared with placebo — not when compared with other drugs. In this context, it is important also to account for the blood pressure-lowering effect of empagliflozin, canagliflozin, liraglutide, and semaglutide (15, 17, 19, 21). Before advocating that physicians should select one of these drugs, it would be informative to conduct a study in which both study arms have comparable mean levels of both HbA1c and blood pressure. For example, an SGLT2 inhibitor or a GLP1 receptor agonist could be compared to a DPP4 inhibitor in combination with generic antihypertensive medications. Such a study would reflect more accurately the real-world options available to physicians and patients. The FDA recognized the potentially important confounding impact of HbA1c and blood pressure at the FDA Advisory Committee on the EMPA-REG trial. Specifically, the FDA “backgrounder” document (27) inquired whether “differences in glycemic control could have contributed directly or indirectly to differences in observed outcomes” and also “whether or how differences [in blood pressure] contributed to the observed outcomes.” In this context, the pioneering studies that established the unique benefit of angiotensin II receptor blockers in diabetic kidney disease compared these drugs to medications with equivalent effects on blood pressure (28, 29). Furthermore, before comparing cardio-

---

**Table 1. Summary of cardiovascular outcome trials of diabetes drugs (2005–2017)**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Study name</th>
<th>Inclusion criteria</th>
<th>N</th>
<th>Mean duration</th>
<th>Comparator</th>
<th>Baseline mean HbA1c</th>
<th>HR – MACE (95% CI)</th>
<th>P value* (superiority)</th>
<th>Literature references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pioglitazone</td>
<td>PROactive</td>
<td>Macrovascular disease</td>
<td>5,238</td>
<td>2.9 yrs</td>
<td>Placebo</td>
<td>7.8%/7.9%</td>
<td>0.84 (0.72–0.98)</td>
<td>0.027</td>
<td>23</td>
</tr>
<tr>
<td>Rosiglitazone</td>
<td>RECORD</td>
<td>Monotherapy with metformin or SU</td>
<td>4,447</td>
<td>5.5 yrs</td>
<td>Sulfonylurea + metformin</td>
<td>7.9%/7.9%</td>
<td>0.95 (0.78–1.17)</td>
<td>–</td>
<td>24</td>
</tr>
<tr>
<td>Sitagliptin</td>
<td>TECOS</td>
<td>Established CV disease</td>
<td>14,671</td>
<td>3.0 yrs</td>
<td>Placebo</td>
<td>7.2%</td>
<td>0.98 (0.89–1.08)</td>
<td>0.65</td>
<td>14</td>
</tr>
<tr>
<td>Saxagliptin</td>
<td>SAVOR-TIMI 53</td>
<td>History of CV disease or high CV risk</td>
<td>16,492</td>
<td>2.1 yrs</td>
<td>Placebo</td>
<td>8.0%/8.0%</td>
<td>1.00 (0.89–1.12)</td>
<td>0.99</td>
<td>12</td>
</tr>
<tr>
<td>Alogliptin</td>
<td>EXAMINE</td>
<td>Acute coronary syndrome</td>
<td>5,380</td>
<td>1.5 yrs</td>
<td>Placebo</td>
<td>8.0%/8.0%</td>
<td>0.96 (&lt;1.16)</td>
<td>0.32</td>
<td>13</td>
</tr>
<tr>
<td>Empagliflozin</td>
<td>EMPA-REG</td>
<td>Established CV disease</td>
<td>7,028</td>
<td>2.6 yrs</td>
<td>Placebo</td>
<td>8.07%/8.08%</td>
<td>0.86 (0.74–0.99)</td>
<td>0.04</td>
<td>15, 16</td>
</tr>
<tr>
<td>Canagliflozin</td>
<td>CANVAS</td>
<td>ASCVD or ≥2 CV risk factors</td>
<td>10,142</td>
<td>3.6 yrs</td>
<td>Placebo</td>
<td>8.2%/8.2%</td>
<td>0.86 (0.75–0.97)</td>
<td>0.02</td>
<td>17</td>
</tr>
<tr>
<td>Lixisenatide</td>
<td>ELIXA</td>
<td>Acute coronary syndrome</td>
<td>6,068</td>
<td>1.9 yrs</td>
<td>Placebo</td>
<td>7.7%/7.6%</td>
<td>1.02 (0.89–1.17)</td>
<td>0.81</td>
<td>18</td>
</tr>
<tr>
<td>Liraglutide</td>
<td>LEADER</td>
<td>High CV risk</td>
<td>9,340</td>
<td>3.8 yrs</td>
<td>Placebo</td>
<td>8.7%/8.7%</td>
<td>0.87 (0.78–0.97)</td>
<td>0.01</td>
<td>19, 20</td>
</tr>
<tr>
<td>Semaglutide</td>
<td>SUSTAIN-6</td>
<td>Established CVD, CV, or CHF</td>
<td>3,297</td>
<td>2 yrs</td>
<td>Placebo</td>
<td>8.7%/8.7%</td>
<td>0.74 (0.58–0.95)</td>
<td>0.02</td>
<td>21</td>
</tr>
<tr>
<td>Exenatide</td>
<td>EXSCEL</td>
<td>Prior CVD (73%)</td>
<td>14,752</td>
<td>3.2 yrs</td>
<td>Placebo</td>
<td>8.1%/8.1%</td>
<td>0.91 (0.83–1.00)</td>
<td>0.06</td>
<td>22</td>
</tr>
<tr>
<td>Bromocriptine</td>
<td>Cycloset Safety Trial</td>
<td>T2D. Stable Rx for &gt; 30 days.</td>
<td>3,070</td>
<td>0.5–1 yr</td>
<td>Placebo</td>
<td>7.0%/7.0%</td>
<td>0.48 (0.23–1.00)</td>
<td>0.05</td>
<td>25</td>
</tr>
</tbody>
</table>

The table summarizes information on cardiovascular outcome trials evaluating the impact of individual glucose-lowering drugs on the risk of major adverse cardiovascular events. Because several drugs were approved prior to 2008, their cardiovascular outcome trials were initiated prior to the time that the FDA began to require formal cardiovascular outcome trials for all drugs to treat type 2 diabetes: pioglitazone, rosiglitazone, and sitagliptin. For completeness, we have included the Cycloset Safety Trial of bromocriptine in the table, but have not discussed it in the text because of challenges to interpreting the data. Specifically, a relatively low percentage of patients actually completed the study (53% of bromocriptine-treated patients vs. 68% in placebo-treated patients). Moreover, a substantial number of patients were lost to follow-up: 5.6% (bromocriptine) and 5.6% (placebo). The table presents hazard ratios based on data for the three-component composite for major adverse cardiovascular events (MACE-3): cardiovascular death, nonfatal myocardial infarction, and nonfatal stroke. In some cases, MACE-3 was not prespecified as the study’s original primary outcome (PROactive and RECORD); in those cases, we have presented nominal P values uncorrected for multiple comparisons. In the case of ELIXA, the table presents data on four-component MACE-4 (components of MACE-3 plus hospitalization for unstable angina). However, because hospitalization for unstable angina represented less than 2.5% of the major adverse cardiovascular events in ELIXA, it is likely that the hazard ratio for MACE-3 would have been very similar. *Values in bold signify P <0.05 (i.e., nominal statistical significance). ASCVD, atherosclerotic cardiovascular disease; CHF, congestive heart failure; CKD, chronic kidney disease; CV, cardiovascular; CVD, cardiovascular disease; HR, hazard ratio; Rx, therapy; SU, sulfonylurea; T2D, type 2 diabetes.*
vascular outcomes in two separate studies, it is important to inquire whether the study populations were indeed comparable. For example, the TECOS study of sitagliptin (14) enrolled patients with a mean baseline HbA1c of 7.2% (i.e., markedly lower than the levels in the studies of SGLT2 inhibitors or GLP1 receptor agonists). In other words, it may not be appropriate to compare data derived from patients with relatively good glycemic control (i.e., mean HbA1c of 7.2%) to less-well-controlled patients (i.e., mean baseline HbA1c levels of 8.2% to 8.7% for most of the other cardiovascular outcome trials).

Conclusions
In the future, one would hope that studies would be designed to assure better protection of research subjects. This might be accomplished if treatment algorithms were provided to make sure that the responsible physicians actually prescribe diabetes drugs and antihypertensive agents to achieve comparable mean levels of HbA1c and blood pressure in all treatment arms. Data and Safety Monitoring Boards should review unblinded data on HbA1c and cardiovascular risk factors to make sure that appropriate actions are taken. Smith et al. (9) have questioned whether it is justified to require large outcome studies for all new antihyperglycemic medications. Their argument was based on their assessment that there is no longer a reason to believe a priori that diabetes drugs are likely to increase cardiovascular risk. While we share their assessment about drug-induced increases in cardiovascular risk, we are nevertheless convinced that recent outcome studies have provided important clinical data that go well beyond the original objective of ruling out increased risk of major adverse cardiovascular events. For example, these studies have detected previously unexpected safety issues — especially informative for uncommon side effects that require time to develop (e.g., increased risks of bone fracture, ref. 30; amputation, ref. 17; and possibly bladder cancer, ref. 23; or hospitalization for heart failure, ref. 12). Furthermore, despite our critique of the design of the recent cohort of cardiovascular outcome studies (12–25), it is likely that improved study design would enable future trials to provide high-quality data to guide important therapeutic decisions. While large outcome studies are expensive, the expense must be judged in comparison to the very substantial sales of successful diabetes drugs. Most importantly, there is great value in providing high-quality information to patients and physicians so that they can make informed decisions. With the wisdom of hindsight, it is not entirely surprising that recent cardiovascular outcome studies have fallen short of perfection. With an eye to the future, we believe that these historical studies have provided valuable lessons that can be applied to improve study design — thereby increasing the value of future studies.

Address correspondence to: Simeon I. Taylor, University of Maryland School of Medicine, MSTF 357-B, 655 West Baltimore Street, Baltimore, Maryland 21201, USA. Phone: 301.980.5272; Email: staylor2@som.umaryland.edu.

