Second-Line Pharmaceutical Treatments for Patients with Type 2 Diabetes

Rohit Vashisht, PhD; Ayan Patel, MS; Lisa Dahm, PhD; Cora Han, JD; Kathryn E. Medders, PhD; Robert Mowers, Pharm D; Carrie L. Byington, MD; Suneil K. Koliwad, MD, PhD; Atul J. Butte, MD, PhD

Abstract

IMPORTANCE Assessing the relative effectiveness and safety of additional treatments when metformin monotherapy is insufficient remains a limiting factor in improving treatment choices in type 2 diabetes.

OBJECTIVE To determine whether data from electronic health records across the University of California Health system could be used to assess the comparative effectiveness and safety associated with 4 treatments in diabetes when added to metformin monotherapy.

DESIGN, SETTING, AND PARTICIPANTS This multicenter, new user, multidimensional propensity score–matched retrospective cohort study with leave-one-medical-center-out (LOMCO) sensitivity analysis used principles of emulating target trial. Participants included patients with diabetes receiving metformin who were then additionally prescribed either a sulfonylurea, dipeptidyl peptidase-4 inhibitor (DPP4I), sodium-glucose cotransporter-2 inhibitor (SGLT2I), or glucagon-like peptide-1 receptor agonist (GLP1RA) for the first time and followed-up over a 5-year monitoring period. Data were analyzed between January 2022 and April 2023.

EXPOSURE Treatment with sulfonylurea, DPP4I, SGLT2I, or GLP1RA added to metformin monotherapy.

MAIN OUTCOMES AND MEASURES The main effectiveness outcome was the ability of patients to maintain glycemic control, represented as time to metabolic failure (hemoglobin A1c [HbA1c] ≥7.0%). A secondary effectiveness outcome was assessed by monitoring time to new incidence of any of 28 adverse outcomes, including diabetes-related complications while treated with the assigned drug. Sensitivity analysis included LOMCO.

RESULTS This cohort study included 31,852 patients (16,635 [52.2%] male; mean [SD] age, 61.4 [12.6] years) who were new users of diabetes treatments added on to metformin monotherapy. Compared with sulfonylurea in random-effect meta-analysis, treatment with SGLT2I (summary hazard ratio [sHR], 0.75 [95% CI, 0.69-0.83]; I² = 37.5%), DPP4I (sHR, 0.79 [95% CI, 0.75-0.84]; I² = 0%), GLP1RA (sHR, 0.62 [95% CI, 0.57-0.68]; I² = 23.6%) were effective in glycemic control; findings from LOMCO sensitivity analysis were similar. Treatment with SGLT2I showed no significant difference in effectiveness compared with GLP1RA (sHR, 1.26 [95% CI, 1.12-1.42]; I² = 47.3%; no LOMCO) or DPP4I (sHR, 0.97 [95% CI, 0.90-1.04]; I² = 0%). Patients treated with DPP4I and SGLT2I had fewer cardiovascular events compared with those treated with sulfonylurea (DPP4I: sHR, 0.84 [95% CI, 0.74-0.96]; I² = 0%; SGLT2I: sHR, 0.78 [95% CI, 0.62-0.98]; I² = 0%). Patients treated with a GLP1RA or SGLT2I were less likely to develop chronic kidney disease (GLP1RA: sHR, 0.75 [95% CI 0.6-0.94]; I² = 0%; SGLT2I: sHR, 0.77 [95% CI, 0.61-0.97]; I² = 0%), kidney failure (GLP1RA: sHR, 0.69 [95% CI, 0.56-0.86]; I² = 9.1%; SGLT2I: sHR, 0.72 [95% CI, 0.59-0.88]; I² = 0%), or...
Abstract (continued)

hypertension (GLP1RA: sHR, 0.82 [95% CI, 0.68-0.97]; $I^2 = 0$%; SGLT2I: sHR, 0.73 [95% CI, 0.58-0.92]; $I^2 = 38.5$%) compared with those treated with a sulfonylurea. Patients treated with an SGLT2I, vs a DPP4I, GLP1RA, or sulfonylurea, were less likely to develop indicators of chronic hepatic dysfunction (sHR vs DPP4I, 0.68 [95% CI, 0.49-0.95]; $I^2 = 0$%; sHR vs GLP1RA, 0.66 [95% CI, 0.48-0.91]; $I^2 = 0$%; sHR vs sulfonylurea, 0.60 [95% CI, 0.44-0.81]; $I^2 = 0$%), and those treated with a DPP4I were less likely to develop new incidence of hypoglycemia (sHR, 0.48 [95% CI, 0.36-0.65]; $I^2 = 22.7$%) compared with those treated with a sulfonylurea.

CONCLUSIONS AND RELEVANCE

These findings highlight familiar medication patterns, including those mirroring randomized clinical trials, as well as providing new insights underscoring the value of robust clinical data analytics in swiftly generating evidence to help guide treatment choices in diabetes.


Introduction

The annual prevalence of type 2 diabetes remains high, impacting an estimated 11.3% of the US population and 9.3% of people worldwide. Current clinical treatment guidelines recommend targeting glycated hemoglobin A1C (HbA1c) less than 7% for most adults (to convert to proportion of total hemoglobin, multiply by 0.01), although there is heterogeneity across recommendations, ranging between 6.5% and 8%, depending on the target population and risk of various microvascular and macrovascular complications. Moderate glycemic control, defined by HbA1c within 7% to 8%, has been shown to improve macrovascular and microvascular outcomes in individuals with diabetes. Metformin is a preferred glucose-lowering drug and is often recommended at treatment initiation for patients with diabetes due to its tolerability, efficacy, and low cost. However, intensification of treatment with an additional agent is often needed to maintain glycemic levels in the recommended range. Indeed, an estimated one-third of the US population with HbA1c greater than 7% is prescribed an additional drug to maintain glycemic control. However, the choice of such an add-on treatment is ambiguous, owing to the lack of direct comparisons of currently available agents. Additionally, evidence emerging from the characterization of clinical data suggests that the actual choices clinicians make in adding another agent onto metformin is highly variable across health systems, both in the US and worldwide. Therefore, to help inform clinical decision-making, it is imperative to fill this knowledge gap by comparing the effectiveness and safety of agents added on to metformin in a head-to-head fashion. However, such rigorous comparison across multiple therapeutic options is simply not feasible in the context of classical randomized clinical trials (RCTs) for myriad reasons, including but not limited to cost, time, patient follow-ups, and the combinatorial nature of diabetes treatment options.

We systematically evaluated the comparative effectiveness and safety associated with 4 categories of diabetes drugs: sulfonylurea, dipeptidyl peptidase-4 inhibitor (DPP4I), glucagon-like peptide-1 receptor agonist (GLP1RA), and sodium-glucose cotransporter-2 inhibitor (SGLT2I) when added onto existing metformin therapy using a clinical database covering 8 million patients across 5 University of California (UC) academic health centers that have been linked to form UC Health (eMethods in Supplement 1). Our objective did not involve the emulation of any specific trial; rather, we endeavored to conduct an analysis following the conventions observed within a customary clinical trial framework under the purview of target trial emulation from observational data. Our study involved identifying new users of a given diabetes treatment, creating matched cohorts of patients prescribed a specific drug and a comparator agent, evaluating their effectiveness and safety in maintaining glycemic control, and assessing adverse outcomes over 5 years at each site. We
performed meta-analysis of evidence gathered from each site, followed by leave-1-medical-center-out (LOMCO) influence analysis to assess the replicability and robustness (eMethods in Supplement 1).

Methods
For this cohort study, the institutional review boards across the UC Health system have determined the research use of the deidentified clinical database was not human participants research; therefore, this study was exempt from further approval and informed consent. The study is reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. The analysis was conducted between January 2022 to April 2023 using data from electronic health records (EHR) starting January 2012 to April 2023. Patient demographic data were extracted from EHR. Race and ethnicity were not included in analysis because of incompleteness in the EHR. The patients were followed up for a 5-year monitoring period after treatment initiation. Detailed methods are provided in the eMethods in Supplement 1.

Statistical Analysis
The clinical data from EHRs of more than 8 million individuals across UC Health were extracted, transformed, and loaded into a standardized Observational Medical Outcome Partnership Common Data Model (OMOP-CDM) with an overall data quality score of 95% based on plausibility, conformance, and completeness.17,18 New users of diabetes treatments were identified using predefined inclusion and exclusion criteria. Patients from each treatment cohort were 1:1 matched with patients in a comparator cohort based on high-dimensional propensity scores estimated using adaptive least absolute shrinkage and selection operator,19 leveraging their extensive pretreatment clinical history at each site individually. The Cox proportional hazards model was used to calculate hazard ratios (HRs) assessing the comparative effectiveness and safety of treatments in each of the matched comparator-treatment pair at each site independently. The evidence from each site were summarized using random-effects meta-analysis and further assessed for stability and robustness by implementing LOMCO influence analysis approach. P values were corrected for multiple hypotheses using false discovery rate correction. The summary HR (sHR) was considered reliable, stable, and significant if its 95% CI did not span 1, had $P$ less than 60%, had corrected $P < .10$, and remained stable based on LOMCO analysis. All the calculations were performed using R statistical software version 3.6.3 (R Project for Statistical Computing). A detailed description of the methods is provided in the eMethods in Supplement 1.

Results
Population Characteristics
A total of 31,852 patients (16,635 [52.2%] male; mean [SD] age, 61.4 [12.6] years) who were new users of sulfonylurea, DPP4I, SGLT2I, and GLP1RA ages 18 years and older and who satisfied the study inclusion and exclusion criteria were identified across all of UC Health (Table; eMethods and eAppendix 1 in Supplement 1). Among these, the sulfonylurea group was the largest, with 14,029 new users, followed by DPP4I (7,434 new users), GLP1RA (5,472 new users), and SGLT2I (4,917 new users) distributed across the sites. There was no statistical difference in age of females (mean [SD] age, 61.6 [3.6] years) and males (mean [SD] age, 61.4 [2.5] years) at the time of add-on treatment initiation across the sites, which is also in accordance with the known predominance of relatively older individuals receiving care across UC Health. Mean HbA1c at the time of treatment initiation varied from 7.4% (GLP1RA) to 8.3% (sulfonylurea) (Table). HbA1c values were highly variable in our patient population viewed across UC Health. This included cohorts receiving each of the add-on medications, except for GLP1RA, which was not as
### Table. Characteristics of the New Users of Diabetes Treatments in Addition to Metformin Considered for Comparative Effectiveness and Safety Analysis

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patients, No. (%)</th>
<th>UC-1</th>
<th>UC-2</th>
<th>UC-3</th>
<th>UC-4</th>
<th>UC-5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sulfonylureas</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New users, No.</td>
<td></td>
<td>2409</td>
<td>1991</td>
<td>3140</td>
<td>1754</td>
<td>4735</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>1223 (50.8)</td>
<td>886 (44.5)</td>
<td>1449 (46.2)</td>
<td>766 (43.7)</td>
<td>2069 (43.7)</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td>1186 (49.2)</td>
<td>1105 (55.5)</td>
<td>1691 (53.9)</td>
<td>988 (56.3)</td>
<td>2666 (56.3)</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td></td>
<td>56.8 (13.6)</td>
<td>60.8 (14.3)</td>
<td>61.8 (12.4)</td>
<td>62.6 (13.4)</td>
<td>63.9 (12.8)</td>
</tr>
<tr>
<td>HbA1c, mean (SD), %a</td>
<td></td>
<td>8.74 (2.1)</td>
<td>7.57 (1.8)</td>
<td>8.60 (2.4)</td>
<td>8.40 (2.0)</td>
<td>8.24 (2.2)</td>
</tr>
<tr>
<td>Baseline comorbidityb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td></td>
<td>191 (7.9)</td>
<td>256 (12.9)</td>
<td>283 (9)</td>
<td>164 (9.4)</td>
<td>351 (7.4)</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td>69 (2.9)</td>
<td>68 (3.4)</td>
<td>69 (2.2)</td>
<td>43 (2.5)</td>
<td>115 (2.4)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td></td>
<td>54 (2.2)</td>
<td>70 (3.5)</td>
<td>69 (2.2)</td>
<td>31 (1.8)</td>
<td>126 (2.7)</td>
</tr>
<tr>
<td>Heart Failure</td>
<td></td>
<td>95 (3.9)</td>
<td>148 (7.4)</td>
<td>181 (5.8)</td>
<td>105 (6.0)</td>
<td>171 (3.6)</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td></td>
<td>105 (4.4)</td>
<td>157 (7.9)</td>
<td>186 (5.9)</td>
<td>164 (9.4)</td>
<td>433 (9.1)</td>
</tr>
<tr>
<td>Baseline health care visits, mean (SD), No.c</td>
<td></td>
<td>27.86 (41.39)</td>
<td>29.99 (55.51)</td>
<td>38.44 (56.69)</td>
<td>29.16 (51.47)</td>
<td>23.46 (41.55)</td>
</tr>
<tr>
<td><strong>Dipeptidyl peptidase-4 inhibitor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New users, No.</td>
<td></td>
<td>1033</td>
<td>1226</td>
<td>1077</td>
<td>696</td>
<td>3402</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>518 (50.2)</td>
<td>621 (50.7)</td>
<td>531 (49.3)</td>
<td>352 (50.6)</td>
<td>1679 (49.4)</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td>515 (49.9)</td>
<td>605 (49.4)</td>
<td>546 (50.7)</td>
<td>344 (49.4)</td>
<td>1723 (50.7)</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td></td>
<td>64.7 (14.0)</td>
<td>65.5 (12.6)</td>
<td>62.9 (13.1)</td>
<td>65.8 (14.6)</td>
<td>65.9 (12.8)</td>
</tr>
<tr>
<td>HbA1c, mean (SD), %a</td>
<td></td>
<td>7.9 (2.1)</td>
<td>7.0 (1.4)</td>
<td>7.9 (1.6)</td>
<td>7.6 (1.8)</td>
<td>7.5 (1.9)</td>
</tr>
<tr>
<td>Baseline comorbidityb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td></td>
<td>122 (11.8)</td>
<td>161 (13.1)</td>
<td>110 (10.2)</td>
<td>84 (12.1)</td>
<td>262 (7.7)</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td>56 (5.4)</td>
<td>42 (3.4)</td>
<td>26 (2.4)</td>
<td>28 (4.0)</td>
<td>86 (2.5)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td></td>
<td>30 (2.9)</td>
<td>37 (3.0)</td>
<td>36 (3.3)</td>
<td>17 (2.4)</td>
<td>63 (1.9)</td>
</tr>
<tr>
<td>Heart Failure</td>
<td></td>
<td>57 (5.5)</td>
<td>109 (8.9)</td>
<td>73 (6.8)</td>
<td>46 (6.6)</td>
<td>140 (4.1)</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td></td>
<td>98 (9.5)</td>
<td>143 (11.7)</td>
<td>77 (7.2)</td>
<td>76 (10.9)</td>
<td>270 (7.9)</td>
</tr>
<tr>
<td>Baseline health care visits, mean (SD), No.c</td>
<td></td>
<td>53.61 (87.68)</td>
<td>42.49 (69.36)</td>
<td>53.97 (81.84)</td>
<td>35.19 (59.66)</td>
<td>34.13 (73.05)</td>
</tr>
<tr>
<td><strong>Sodium-glucose cotransporter-2 inhibitor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New users, No.</td>
<td></td>
<td>754</td>
<td>923</td>
<td>531</td>
<td>486</td>
<td>2223</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>337 (44.7)</td>
<td>381 (41.3)</td>
<td>208 (39.2)</td>
<td>190 (39.1)</td>
<td>865 (38.9)</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td>417 (55.3)</td>
<td>542 (58.7)</td>
<td>323 (60.8)</td>
<td>296 (60.9)</td>
<td>1358 (61.1)</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td></td>
<td>59.9 (14.1)</td>
<td>63.3 (13.8)</td>
<td>61.9 (14.4)</td>
<td>64.2 (12.2)</td>
<td>60.6 (12.9)</td>
</tr>
<tr>
<td>HbA1c, mean (SD), %a</td>
<td></td>
<td>8.3 (2.0)</td>
<td>7.3 (1.8)</td>
<td>8.1 (2.1)</td>
<td>7.9 (2.7)</td>
<td>7.5 (1.7)</td>
</tr>
<tr>
<td>Baseline comorbidityb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td></td>
<td>120 (15.9)</td>
<td>224 (24.3)</td>
<td>135 (25.4)</td>
<td>104 (21.4)</td>
<td>261 (11.7)</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td>36 (4.8)</td>
<td>38 (4.1)</td>
<td>26 (4.9)</td>
<td>13 (2.7)</td>
<td>28 (1.3)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td></td>
<td>40 (5.3)</td>
<td>60 (6.5)</td>
<td>45 (8.5)</td>
<td>34 (7.0)</td>
<td>92 (4.1)</td>
</tr>
<tr>
<td>Heart Failure</td>
<td></td>
<td>80 (10.6)</td>
<td>178 (19.3)</td>
<td>112 (21.1)</td>
<td>82 (16.9)</td>
<td>192 (8.6)</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td></td>
<td>67 (8.9)</td>
<td>148 (16.0)</td>
<td>68 (12.8)</td>
<td>55 (11.3)</td>
<td>171 (7.7)</td>
</tr>
<tr>
<td>Baseline health care visits, mean (SD), No.c</td>
<td></td>
<td>58.86 (92.29)</td>
<td>75.78 (114.83)</td>
<td>95.86 (121.38)</td>
<td>70.91 (95.83)</td>
<td>51.56 (79.74)</td>
</tr>
</tbody>
</table>

(continued)
variable (Table). We also saw a significant difference in the HbA1c levels of females vs males at the
time of add-on treatment initiation irrespective of which agent was being prescribed.

**Matched Cohorts**

We analyzed 30 pairs of comparator and treatment cohorts in which each pair represented a combination of 6 drug pairs from 4 categories of individual drugs prescribed as an add-on to metformin across the 5 sites. Patients from each treatment cohort were matched 1:1 with patients in a comparator cohort based on their estimated propensity scores at each site independently. Propensity score was estimated by accounting for a mean (SD) 4688.2 (643.2) pretreatment clinical covariates, along with age and sex of patients in each comparator-treatment pair (eAppendix 2 in Supplement 1). For instance, Figure 1 illustrates an example of propensity score matching for 1033 patients treated with a DPP4I (treatment cohort) and 2409 patients treated with a sulfonylurea (comparator cohort) at UC-1. In this example, the propensity scores were estimated by adjusting for 4429 clinical covariates, including medical procedures (556 covariates), laboratory measurements (832 covariates), unique orders of various drugs (1561 covariates), and medical conditions (1480 covariates) observed on or at least 365 days prior to the treatment initiation, in addition to age and sex. After matching, all clinical features (100%) of the 4429 covariates were balanced (absolute SMD < 0.10), resulting in a final cohort of 1756 matched patients (Figure 1C). Detailed illustrations of the estimated propensity score distribution before and after matching along with the covariate balance diagnostic of each comparator-treatment pair for each site is provided in eAppendix 2 in Supplement 1.

**Comparative Effectiveness**

We analyzed the effectiveness in each matched comparator-treatment pair independently at each UC Health site in terms of a given patient's ability to maintain glycemic control and then summarized the results using random-effect meta-analysis across all UCs with LOMCO influence analysis. The patients who received a DPP4I (summary hazard ratio [sHR], 0.79 [95% CI, 0.75-0.84]; $I^2 = 0\%$), GLP1RA (sHR, 0.62 [95% CI, 0.57-0.68]; $I^2 = 23.6\%$), or SGLT2I (sHR, 0.75 [95% CI, 0.69-0.83]; $I^2 = 37.5\%$) were significantly less likely to experience suboptimal glycemic control compared with those who were treated with a sulfonylurea with metformin during the 5-year posttreatment period.
follow-up (Figure 2). The summary estimates of DPP4I, GLP1RA, and SGLT2I compared with sulfonylurea were found to be reliable based on the LOMCO influence analysis, indicating their robustness. There was no significant difference in the effectiveness associated with SGLT2I compared with DPP4I (sHR, 0.97 [95% CI, 0.90-1.04]; $I^2 = 0\%$). Across UCs, we observed better effectiveness of GLP1RA compared with DPP4I (sHR, 0.81 [95% CI, 0.74-0.88]; $I^2 = 0\%$).

![Figure 1. Propensity Score (PS) Matching and Covariate Balance Diagnostics](image)

![Figure 2. Comparative Effectiveness of Diabetes Treatments for the Ability of Patients to Maintain Glycemic Control Across the University of California (UC) Health System](image)

<table>
<thead>
<tr>
<th>Comparator</th>
<th>Treatment</th>
<th>No.</th>
<th>sHR (95% CI)</th>
<th>$I^2$, %</th>
<th>Stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPP4I</td>
<td>GLP1RA</td>
<td>5968</td>
<td>0.81 (0.74-0.88)</td>
<td>0</td>
<td>Yes</td>
</tr>
<tr>
<td>DPP4I</td>
<td>SGLT2I</td>
<td>6482</td>
<td>0.97 (0.9-1.04)</td>
<td>0</td>
<td>Yes</td>
</tr>
<tr>
<td>GLP1RA</td>
<td>SGLT2I</td>
<td>6678</td>
<td>1.26 (1.12-1.42)</td>
<td>47.3</td>
<td>No</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>DPP4I</td>
<td>13070</td>
<td>0.79 (0.75-0.84)</td>
<td>0</td>
<td>Yes</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>GLP1RA</td>
<td>6318</td>
<td>0.62 (0.57-0.68)</td>
<td>23.6</td>
<td>Yes</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>SGLT2I</td>
<td>7130</td>
<td>0.75 (0.69-0.83)</td>
<td>37.5</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Glycemic control was defined as time to HbA1c of 7% or greater (to convert to proportion of total hemoglobin, multiply by 0.01). Summary hazard ratios (sHRs) and 95% CIs were obtained from random-effect meta-analysis across the UC Health system, the heterogeneity ($I^2$) of the sHRs and the stability of the sHRs are based on leave-1-medical-center-out influence analysis. The sub-forest plot shows the HRs of the comparative effectiveness of patients treated with DPP4I compared with sulfonylureas at each UC Health site along with the sHR obtained from random effect meta-analysis across UC Health as an example. Size of square indicates sample size. Detailed forest plots of each drug pair comparison at each UC Health site along with summary estimates across UC Health, including leave-1-medical-center-out influence analysis, are provided in eAppendix 3 in Supplement 1.
Patients treated with a SGLT2I were more likely to experience metabolic failure than those treated with a GLP1RA (sHR, 1.26 [95% CI, 1.12-1.42]; \( I^2 = 47.3\% \)). However, this estimated summary difference between SGLT2I and GLP1RA was considered inconclusive because it was not stable based on the LOMCO analysis (Figure 2). A detailed illustration of the hazard ratios of each matched comparator-treatment pair at each site along with their summary estimates across all 5 UC Health sites, including LOMCO metric, is provided in eAppendix 3 in Supplement 1.

**Comparative Analysis of Diabetes Complications and Adverse Outcomes**

We also analyzed each matched comparator and treatment pair independently at each UC Health site to evaluate the time it took for a new adverse outcome to occur within the 5-year follow-up period after treatment initiation, and then summarized the results using random-effect meta-analysis across all UCs with LOMCO influence analysis (eAppendix 4 and the eFigure in Supplement 1). Compared with patients treated with a sulfonylurea, patients treated with DPP4I were less likely to develop all-cause cardiovascular disease (sHR, 0.84 [95% CI, 0.74-0.96]; \( I^2 = 0\% \)) and specifically myocardial infarction (sHR, 0.75 [95% CI, 0.61-0.92]; \( I^2 = 0\% \)) (eFigure in Supplement 1). Compared with patients treated with a sulfonylurea, patients treated with SGLT2I also had lower hazards for all-cause cardiovascular disease (sHR, 0.78 [95% CI, 0.62-0.98]; \( I^2 = 0\% \)) and specifically stroke (hemorrhagic and ischemic combined: sHR, 0.55 [95% CI, 0.35-0.86]; \( I^2 = 0\% \)) (eFigure in Supplement 1). Patients treated with GLP1RA or SGLT2I, compared with those treated with a sulfonylurea, had a lower hazard for kidney failure (GLP1RA: sHR, 0.69 [95% CI, 0.56-0.86]; \( I^2 = 9.1\% \); SGLT2I: sHR, 0.72 [95% CI, 0.59-0.88]; \( I^2 = 0\% \)) as well as chronic kidney disease (GLP1RA: sHR, 0.75 [95% CI 0.6-0.94]; \( I^2 = 0\% \); SGLT2I: sHR, 0.77 [95% CI, 0.61-0.97]; \( I^2 = 0\% \)) (eFigure in Supplement 1). There were no differences in the renal outcomes of among patients treated with a GLP1RA vs a DPP4I, nor when comparing individuals treated with an SGLT2I vs a GLP1RA. However, patients treated with an SGLT2I had lower hazards for chronic kidney disease and kidney failure compared with those treated with a DPP4I. Patients treated with an SGLT2I had lower hazards of chronic liver disease after treatment compared with those treated with a DPP4I, GLP1RA, or sulfonylurea (eFigure in Supplement 1).

Hypoglycemia is an important concern whenever diabetes treatments are intensified to maintain strict glycemic control.\(^{20,21}\) We found that patients treated with a DPP4I had a lower hazard for hypoglycemia compared with those treated with a sulfonylurea (sHR, 0.48 [95% CI, 0.36-0.65]; \( I^2 = 22.7\% \)). Separately, we found that patients treated with a GLP1RA or SGLT2I had lower hazards for hypertension compared with those receiving a sulfonylurea (GLP1RA: sHR, 0.82 [95% CI, 0.68-0.97]; \( I^2 = 0\% \); SGLT2I: sHR, 0.73 [95% CI, 0.58-0.92]; \( I^2 = 38.5\% \)) (eFigure in Supplement 1). There was no conclusive evidence regarding other acute adverse outcomes, such as abdominal pain, among all the comparator-treatment pairs, except for the GLP1RA group compared with the sulfonylurea group (sHR, 1.24 [95% CI, 1.04-1.49]; \( I^2 = 0\% \)) (eFigure in Supplement 1). We found that both treatment with an SGLT2I (compared with GLP1RA), and treatment with a GLP1RA (compared with sulfonylurea) had lower hazards for development of acute adverse outcomes, such as vomiting and nausea (eFigure in Supplement 1). With respect to bone fracture, we noted a lower hazard among patients treated with a SGLT2I compared with a DPP4I (sHR, 0.64 [95% CI, 0.46-0.88]; \( I^2 = 0\% \)) while no conclusive differences were found for other treatment comparisons (eFigure in Supplement 1). A detailed illustration of the hazard ratios of each matched comparator-treatment pair at each UC Health site along with their summary estimates across UCs, including LOMCO metrics, is provided in eAppendix 3 in Supplement 1.

**Discussion**

The findings of this cohort study are meaningful with respect to guiding clinical care, especially when compared with RCT data and traditional observational studies. First, our analysis found that compared with sulfonylurea, treatment with either a GLP1RA, SGLT2I, or DPP4I was associated with effectively maintaining glycemic control in patients with diabetes when added to metformin.
monotherapy over a 5-year monitoring period (Figure 3). This finding is important because the drugs in question are not similarly priced. Moreover, whereas the concept of secondary failure, including in a 5-year period, is established for sulfonylurea, clinicians are grappling with whether there is differential effectiveness between the other classes as adjunctive agents to metformin. Our data-driven findings suggest more equivalency across classes with respect to HbA1c control than may have been found from separate RCTs without a head-to-head element across classes.

We also found that treatment with a GLP1RA was associated with better glycemic control than a DPP4I. This mirrors what is expected based on the complementary mechanisms of action for each of these classes and validates findings from prior RCTs. By contrast, the comparison between SGLT2I and DPP4I yielded inconclusive results. We also observed a higher risk of metabolic failure among patients treated with a SGLT2I compared with those treated with a GLP1RA, although the stability and robustness of this evidence were weakened once our LOMCO influence analysis was conducted.

Our study found that adding DPP4I or SGLT2I to metformin monotherapy was associated with a significantly lower risk of new cardiovascular disease compared with sulfonylurea. At least for SGLT2I, this finding again mirrors what might be found in RCTs, which are increasingly showing that SGLT2Is are protective against congestive heart failure. Additionally, patients treated with a DPP4I had lower risk of myocardial infarction, and patients treated with an SGLT2I had lower risk of stroke. This too, aligns with findings from RCTs, which consistently show that GLP1RA and SGLT2I have potential for cardiovascular (GLP1RA) and cardiac (SGLT2I) benefit. The comparative evaluation of a specific DPP4I, linagliptin, vs a specific sulfonylurea, glimepiride, when added to metformin using a noninferiority design in the CAROLINA trial did not reveal a cardiovascular advantage to DPP4I. The trial’s findings suggested that when supplementary glucose-lowering therapy is needed, the DPP4I’s favorable attributes, including its lower risk of hypoglycemia and weight gain, make it a potentially suitable choice. In contrast, our analysis indicated not only that addition of a DPP4I was associated with cardiovascular benefits compared with addition of an sulfonylurea following metformin, but also that DPP4I treatment vs sulfonylurea was associated with superior glycemic control and lower likelihood of hypoglycemia, observations that align with CAROLINA, emphasizing DPP4I as a viable second-line option.

Figure 3. Visual Summary of the Findings From the Comparative Effectiveness and Safety Study Across the University of California Health System
Adding a GLP1RA or SGLT2I to metformin also showed lower risks of new chronic kidney disease and kidney failure compared with a sulfonylurea.25-29 Once more, this comports with accumulated RCT findings, at least for SGLT2I. SGLT2I were associated with a lower risk of new chronic liver disease compared with DPP4I. Moreover, when added to metformin, SGLT2I, DPP4I, and GLP1RA were associated with lower risks of new nausea, abdominal pain, hypertension, and hypoglycemia compared with sulfonylurea. These findings contrast with those from RCTs evaluating GLP1RAs against placebo, in which nausea is highlighted as a class-wide adverse effect. Overall, our analysis highlights the effectiveness of diabetes medications, including SGLT2I, DPP4I, and GLP1RA, in glycemic control and accurately assesses impacts with respect to diabetes complications and potential adverse effects for patients with diabetes. Whereas some of our findings with respect to effectiveness and ancillary benefits mirror what has been found in much more costly RCTs, they extend across the entire class (whereas RCTs usually focus on only 1 drug) and highlight associations that either have not yet been probed by RCTs or that were not included in prior primary or secondary prespecified RCT outcomes.

The understanding that the cardioprotective effects of newer classes of antidiabetic drugs are primarily attributed to their pleiotropic properties rather than only to their glucose-lowering abilities has played a significant role in the broader management strategies.26,30 Sulfonylurea is a preferred add-on treatment to metformin monotherapy, potentially due to its low cost and long track record.5 This choice is widely practiced, including in health care systems like UC Health.11,12 However, sulfonylurea can cause relatively higher incidence of hypoglycemia compared with newer diabetes medications, which may limit their use.7 Our analysis showed that GLP1RA, SGLT2I, and DPP4I, the newer diabetes medications, were significantly associated with more effective glycemic control maintenance when added to metformin monotherapy compared with a sulfonylurea. Additionally, treatment with a DPP4I, compared with a sulfonylurea, was not only associated with effectively maintaining glycemic control but also a lower risk of hypoglycemia and adverse cardiac outcomes.

The GRADE trial comparing glargine (insulin), glimepiride (sulfonylurea), liraglutide (GLP1RA), and sitagliptin (DPP4I) pairwise indicated a similar efficacy of these 4 drugs in reducing HbA1c, with better efficacy of glargine and liraglutide when added to metformin monotherapy and no significant difference in the incidence of macrovascular or microvascular outcomes.32,33 Like the evidence from the GRADE trial, our findings indicated better effectiveness associated with a GLP1RA compared with a sulfonylurea when added to metformin monotherapy in maintaining glycemic control, in addition to their protective associations against cardiac, renal, and liver disorders as well as hypertension. Additionally, our analysis showed better effectiveness associated with use of a SGLT2I, which was not considered in the GRADE trial, in maintaining glycemic control along with their protective associations against cardiac, renal, and liver disorders, as well as hypertension. Thus, our study underscores the significance of evidence in not only reaffirming anticipated results but also generating evidence for drugs that were not included in the GRADE RCT.32,33

Previous investigations using clinical data have improved our understanding of the effectiveness and safety of treatments for diabetes.27,34,35 Recent evidence from OptumLabs, a nationwide claims database in the US, has provided similar findings to the GRADE trial, highlighting the importance of generating timely evidence.36 Initiatives such as RCT-Duplicate underscore the importance of cautiously evaluating clinical data, such as from EHR, owing to various types of underlying bias,38-40 as well as the need of further research to build frameworks to better emulate target trials. Our investigation presents a comprehensive framework for generating meaningful evidence using clinical data, incorporating the principles of emulating target trials in alignment with efforts, such as those represented by RCT-Duplicate.57 The reliability of evidence is bolstered through independent analyses of each UC Health campus, accounting for unique medical practices as well as correcting for multiple hypothesis testing. To assess evidence credibility, we performed LOMCO influence analysis, discerning the comparative strength of data from different health systems. The framework introduced in this study holds promise for informing future studies, facilitating the generation of meaningful evidence to drive effective medical decision-making.
Limitations
This study has some limitations. Our study did not consider race and ethnicity of patients due to incompleteness of capturing such data elements in EHR, especially in outpatient and ambulatory encounters, as well as other factors, such as their socioeconomic backgrounds or the cost of treatment. We binarized (ie, present or absent) medical procedures, diagnoses, medications, and laboratory measurements to use as surrogate markers of the clinical state of a patient to adjust for potentially observed confounders. This approach has the potential to obscure the underlying rationale for treatment decisions established at the baseline.

The study population represents patients receiving care across California in any of the UC Health sites. Patients from outside of UC Health may occasionally receive care at a facility within UC Health, but we estimate that the proportion of such patients is small. However, this concept of fragmented care delivery with components of care delivered across multiple unconnected systems is both a potential limitation of our study and a potential limitation to the interpretation of our results by health systems located outside of California. However, UC Health represents a large number of patients, and California is a diverse state, suggesting that our results may be broadly applicable. Moreover, the issue of fragmented care delivery is a weakness of the US health care system as a whole and highlights the importance of building electronic health record tools with increasing capacity to capture care delivered within and outside of any given health system through interoperability. Given all of this, we acknowledge that determining the true status of new users of any given medication added on to metformin is challenging, given current tools and the limitations. We discuss limitations further in eAppendix 5 in Supplement 1.

Conclusions
This cohort study used clinical data from 5 academic health centers within UC Health to assess the comparative effectiveness and safety associated with 4 pharmaceutical treatments added to metformin monotherapy for diabetes using the principles of emulating target trial. Findings from this study indicated that GLP1RAs, SGLT2Is, or DPP4Is, compared with sulfonylureas, were each effective in maintaining glycemic control in individuals with diabetes when added to metformin. Moreover, the study found that clinical data could be used to define drug-specific patterns with respect to the prevention of diabetes-associated complications, as well as potential drug-related adverse effects. Some of these patterns mirrored those found in dedicated RCTs, whereas others were new and likely would not have been possible to identify without access to data from so many patients. As such, statistically rigorous clinical data analytics could be used to rapidly generate evidence for both effectiveness and safety associated with pharmaceutical interventions to aid in enforcing beneficial shifts in clinical decision-making.
Author Contributions: Dr Vashisht had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Vashisht, Medders, Koliwad, Butte.

Acquisition, analysis, or interpretation of data: Vashisht, Patel, Dahm, Han, Medders, Mowers, Byington.

Drafting of the manuscript: Vashisht, Koliwad, Butte.

Critical review of the manuscript for important intellectual content: Vashisht, Patel, Dahm, Han, Medders, Mowers, Byington, Koliwad.

Statistical analysis: Vashisht.

Obtained funding: Vashisht.

Administrative, technical, or material support: Vashisht, Patel, Dahm, Han, Mowers, Byington.

Supervision: Vashisht, Dahm, Koliwad.

Conflict of Interest Disclosures: Dr Byington reported personal fees from Becton Dickinson outside the submitted work. Dr Butte reported serving as a cofounder and consultant to Personalis and NuMedi; serving as a consultant to Mango Tree Corporation, Samsung, 10x Genomics, Helix, Pathway Genomics, and Verinata (Illumina); serving on paid advisory panels or boards for Geisinger Health, Regenstrief Institute, Gerson Lehman Group, AlphaSights, Covance, Novartis, Genentech, and Merck, and Roche; owning stock in Personalis, NuMedi, Apple, Meta, Alphabet, Microsoft, Amazon, Snap, 10x Genomics, Illumina, Regeneron, Sanofi, Pfizer, Royalty Pharma, Moderna, Sutro, Doximity, BioNTech, Invitae, Pacific Biosciences, Editas Medicine, Nuna Health, Assay Depot, and Ver24seven; receiving personal fees from Johnson and Johnson, Roche, Genentech, Pfizer, Merck, Lilly, Takeda, Varian, Mars, Siemens, Optum, Abbott, Celgene, AstraZeneca, AbbVie, Westat, Stanford University, NuMedi, Personalis, Dartmouth University, Boston Children's Hospital, Johns Hopkins University, Endocrine Society, Alliance for Academic Internal Medicine, Children's Hospital of Philadelphia, University of Pittsburgh Medical Center, Cleveland Clinic, University of Utah, Society of Toxicology, Mayo Clinic, Washington University in Saint Louis, and University of Michigan; receiving grants from the National Institutes of Health, Peraton, Genentech, Johnson and Johnson, US Food and Drug Administration (FDA), Robert Wood Johnson Foundation, Leon Lowenstein Foundation, Intervenn Foundation, Priscilla Chan and Mark Zuckerberg, the Barbara and Gerson Bakar Foundation, March of Dimes, Juvenile Diabetes Research Foundation, California Governor's Office of Planning and Research, California Institute for Regenerative Medicine, L'Oreal, and Progenity outside the submitted work. No other disclosures were reported.

Funding/Support: This work was supported by the University of California, San Francisco (UCSF), Clinical and Translational Science Institute (award No. UL1TR001872). Partial grant support was provided through the FDA (grant No. U01FD005978) to the UCSF–Stanford Center of Excellence in Regulatory Sciences and Innovation. Partial support was provided by the UCSF Bakar Computational Health Sciences Institute and the University of California Health Center for Data-driven Insights and Innovation.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The content reflects the views of the authors and should not be construed to represent the views or policies of the National Institutes of Health or FDA.

Data Sharing Statement: See Supplement 2.

Additional Contributions: David Gonzalez, BIT, and Robert William Follett, BIT (University of California Health) provided support with the University of California Health Data Warehouse. They were not compensated for this work outside of their usual salaries.

REFERENCES


SUPPLEMENT 1.
eMethods.
eAppendix 1. Cohort Selection
eAppendix 2. Patient Matching
eAppendix 3. Comparative Effectiveness and Safety
eFigure. Forest Plot of Summary Estimates of Secondary Outcomes
eAppendix 4. Secondary Outcomes
eAppendix 5. Discussion and Limitations

SUPPLEMENT 2.
Data Sharing Statement