JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY VOL. 84, NO. 10, 2024

Comparative Effectiveness of Second-Line Antihyperglycemic Agents for Cardiovascular Outcomes

A Multinational, Federated Analysis of LEGEND-T2DM

Rohan Khera, MD, MS,[a,b,c] Arya Aminorroaya, MD, MPH,[a] Lovedeep Singh Dhingra, MBBS,[a] Phyllis M. Thangaraj, MD, PhD,[a] Aline Pedroso Camargos, PhD,[a] Fan Bu, PhD,[d] Xiyu Ding, MS,[e] Akihiko Nishimura, PhD,[e] Tara V. Anand, BS,[f] Faaizah Arshad, BS,[g] Clair Blacketer, MPH,[h] Yi Chai, PhD,[i] Shounak Chattopadhyay, PhD,[g] Michael Cook, BSc,[e] David A. Dorr, MD, MS,[j] Talita Duarte-Salles, PhD,[k,l] Scott L. DuVall, PhD,[m,n] Thomas Falconer, MS,[f] Tina E. French, RN, CPHQ,[o,p] Elizabeth E. Hanchrow, RN, MSN,[o,p] Guneet Kaur, MS,[q] Wallis C.Y. Lau, BSc, PhD,[r,s,t,u] Jing Li, MS,[v] Kelly Li, BS,[g] Yuntian Liu, MPH,[a,b] Yuan Lu, ScD,[a] Kenneth K.C. Man, BSc, MPH, PhD,[r,s,t,u] Michael E. Matheny, MD, MS, MPH,[o,p] Nestoras Mathioudakis, MD, MHS,[w] Jody-Ann McLeggon, MPH,[f] Michael F. McLemore, RN,[o,p] Evan Minty, MD, MSc,[x] Daniel R. Morales, MD,[q] Paul Nagy, PhD,[w] Anna Ostropolets, MD, PhD,[h] Andrea Pistillo, MSc,[k] Thanh-Phuc Phan, MBA,[y] Nicole Pratt, PhD,[z] Carlen Reyes, MD, PhD,[k] Lauren Richter, MD,[f] Joseph S. Ross, MD, MHS,[b,aa] Elise Ruan, MD,[f] Sarah L. Seager, BS,[bb] Katherine R. Simon, AA,[o,p] Benjamin Viernes, PhD,[m,n] Jianxiao Yang, MS,[cc] Can Yin, MS,[dd] Seng Chan You, MD, PhD,[ee,ff] Jin J. Zhou, PhD,[g,gg] Patrick B. Ryan, PhD,[f] Martijn J. Schuemie, PhD,[hh] Harlan M. Krumholz, MD, SM,[a,b,ii] George Hripcsak, MD, MS,[f] Marc A. Suchard, MD, PhD[g,m,jj,kk]

ABSTRACT

BACKGROUND Sodium-glucose cotransporter 2 inhibitors (SGLT2is) and glucagon-like peptide-1 receptor agonists (GLP-1 RAs) reduce the risk of major adverse cardiovascular events (MACE) in patients with type 2 diabetes mellitus (T2DM). However, their effectiveness relative to each other and other second-line antihyperglycemic agents is unknown, without any major ongoing head-to-head clinical trials.

OBJECTIVES The aim of this study was to compare the cardiovascular effectiveness of SGLT2is, GLP-1 RAs, dipeptidyl peptidase-4 inhibitors (DPP4is), and clinical sulfonylureas (SUs) as second-line antihyperglycemic agents in T2DM.

METHODS Across the LEGEND-T2DM (Large-Scale Evidence Generation and Evaluation Across a Network of Databases for Type 2 Diabetes Mellitus) network, 10 federated international data sources were included, spanning 1992 to 2021. In total, 1,492,855 patients with T2DM and cardiovascular disease (CVD) on metformin monotherapy were identified who initiated 1 of 4 second-line agents (SGLT2is, GLP-1 RAs, DPP4is, or SUs). Large-scale propensity score models were used to conduct an active-comparator target trial emulation for pairwise comparisons. After evaluating empirical equipoise and population generalizability, on-treatment Cox proportional hazards models were fit for 3-point MACE (myocardial infarction, stroke, and death) and 4-point MACE (3-point MACE plus heart failure hospitalization) risk and HR estimates were combined using random-effects meta-analysis.

RESULTS Over 5.2 million patient-years of follow-up and 489 million patient-days of time at risk, patients experienced 25,982 3-point MACE and 41,447 4-point MACE. SGLT2is and GLP-1 RAs were associated with lower 3-point MACE risk than DPP4is (HR: 0.89 [95% CI: 0.79-1.00] and 0.83 [95% CI: 0.70-0.98]) and SUs (HR: 0.76 [95% CI: 0.65-0.89] and 0.72 [95% CI: 0.58-0.88]). DPP4is were associated with lower 3-point MACE risk than SUs (HR: 0.87; 95% CI: 0.79-0.95). The pattern for 3-point MACE was also observed for the 4-point MACE outcome. There were no significant differences between SGLT2is and GLP-1 RAs for 3-point or 4-point MACE (HR: 1.06 [95% CI: 0.96-1.17] and 1.05 [95% CI: 0.97-1.13]).

CONCLUSIONS In patients with T2DM and CVD, comparable cardiovascular risk reduction was found with SGLT2is and GLP-1 RAs, with both agents more effective than DPP4is, which in turn were more effective than SUs. These findings suggest that the use of SGLT2is and GLP-1 RAs should be prioritized as second-line agents in those with established CVD. (*JACC* 2024;84:904–917) © 2024 the American College of Cardiology Foundation. Published by Elsevier. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Listen to this manuscript's audio summary by Editor Emeritus Dr Valentin Fuster on www.jacc.org/journal/jacc.

ISSN 0735-1097/$36.00 https://doi.org/10.1016/j.jacc.2024.05.069

Over the past decade, the therapeutic options for type 2 diabetes mellitus (T2DM) have undergone a significant transformation.[1] Sodium-glucose cotransporter 2 inhibitors (SGLT2is) and glucagon-like peptide-1 receptor agonists (GLP-1 RAs) have expanded the role of antihyperglycemic agents from managing high blood glucose to addressing the elevated cardiovascular risk in patients with T2DM.[2,3] In several large randomized clinical trials (RCTs), SGLT2is and GLP-1 RAs reduced major adverse cardiovascular events (MACE), such as myocardial infarction (MI), hospitalization for heart failure (HF), and cardiovascular mortality.[4-11] Although older antihyperglycemic agents, such as sulfonylureas (SUs), have not undergone similarly comprehensive trials to evaluate their cardiovascular efficacy or safety, some studies suggested their neutral or even harmful effect on cardiovascular outcomes, with increased risk for hypoglycemia.[12-15] Furthermore, direct comparisons of SGLT2is and GLP-1 RAs with dipeptidyl peptidase-4 inhibitors (DPP4is), which are antihyperglycemic agents with neutral effects on MACE, have not been conducted, and no major trials are in progress. Nevertheless, DPP4is and SUs continue to be used in clinical practice and are recommended as second-line T2DM agents in national clinical practice guidelines.[16-19]

Despite the availability of evidence from SGLT2i and GLP-1 RA trials, several gaps in evidence challenge the development of treatment recommendations in T2DM.[16,20,21] Specifically, trials of SGLT2is and GLP-1 RAs were not designed as head-to-head comparisons with older agents but rather as additive

ABBREVIATIONS AND ACRONYMS

CVD = cardiovascular disease

DPP4i = dipeptidyl peptidase-4 inhibitor

EHR = electronic health record

GLP-1 RA = glucagon-like peptide-1 receptor agonist

HF = heart failure

MACE = major adverse cardiovascular event(s)

MI = myocardial infarction

OMOP = Observational Medical Outcomes Partnership

PS = propensity score

RCT = randomized clinical trial

From the [a]Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University, New Haven, Connecticut, USA; [b]Center for Outcomes Research and Evaluation, Yale New Haven Hospital, New Haven, Connecticut, USA; [c]Section of Health Informatics, Department of Biostatistics, Yale School of Public Health, New Haven, Connecticut, USA; [d]Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA; [e]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA; [f]Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, New York, USA; [g]Department of Biostatistics, Fielding School of Public Health, University of California-Los Angeles, Los Angeles, California, USA; [h]Observational Health Data Analytics, Janssen Research and Development, Titusville, New Jersey, USA; [i]Department of Pharmacology and Pharmacy, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China; [j]Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA; [k]Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Barcelona, Spain; [l]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands; [m]Veterans Affairs Informatics and Computing Infrastructure, U.S. Department of Veterans Affairs, Salt Lake City, Utah, USA; [n]Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah, USA; [o]Tennessee Valley Healthcare System, Veterans Affairs Medical Center, Nashville, Tennessee, USA; [p]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA; [q]Division of Population Health and Genomics, School of Medicine, University of Dundee, Dundee, United Kingdom; [r]Research Department of Practice and Policy, School of Pharmacy, University College London, London, United Kingdom; [s]Centre for Medicines Optimisation Research and Education, University College London Hospitals NHS Foundation Trust, London, United Kingdom; [t]Centre for Safe Medication Practice and Research, Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; [u]Laboratory of Data Discovery for Health, Hong Kong Science Park, Hong Kong, China; [v]Data Transformation, Analytics, and Artificial Intelligence, Real World Solutions, IQVIA, Durham, North Carolina, USA; [w]Division of Endocrinology, Diabetes, and Metabolism, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; [x]Faculty of Medicine, O'Brien Institute for Public Health, University of Calgary, Calgary, Alberta, Canada; [y]School of Pharmacy, Taipei Medical University, Taipei, Taiwan; [z]Quality Use of Medicines and Pharmacy Research Centre, UniSA Clinical and Health Sciences, University of South Australia, Adelaide, Australia; [aa]Section of General Medicine and National Clinician Scholars Program, Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut, USA; [bb]Data Transformation, Analytics, and Artificial Intelligence, Real World Solutions, IQVIA, London, United Kingdom; [cc]Department of Computational Medicine, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, California, USA; [dd]Data Transformation, Analytics, and Artificial Intelligence, Real World Solutions, IQVIA, Shanghai, China; [ee]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, South Korea; [ff]Institute for Innovation in Digital Healthcare, Yonsei University College of Medicine, Seoul, South Korea; [gg]Department of Medicine, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, California, USA; [hh]Epidemiology, Office of the Chief Medical Officer, Johnson & Johnson, Titusville, New Jersey, USA; [ii]Section of Cardiovascular Medicine, Department of Health Policy and Management, Yale School of Public Health, New Haven, Connecticut, USA; [jj]Department of Biomathematics, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, California, USA; and the [kk]Department of Human Genetics, David Geffen School of Medicine, University of California-Los Angeles, Los Angeles, California, USA.

Review and acceptance occurred under Dr Valentin Fuster's term as Editor-in-Chief.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

Manuscript received May 15, 2024; accepted May 23, 2024.

SCD = sudden cardiac death
SGLT2i = sodium-glucose cotransporter 2 inhibitor
SMD = standardized mean difference
SU = sulfonylurea
T2DM = type 2 diabetes mellitus

treatments on top of commonly used T2DM agents.[5,6,22] As a result, the relative cardiovascular efficacy of newer vs older agents is not known. Moreover, the comparative cardiovascular efficacy of SGLT2is and GLP-1 RAs has not been evaluated. Thus far, comparative effectiveness assessments have been based on indirect estimates from clinical trials or comparative effectiveness drawn from a limited number of data sources.[23-25] The findings from observational studies using single data sources are challenged by our observations that the uptake of these agents, and therefore the selective pressures on the use of these agents, vary considerably across sites,[18] with the potential to introduce bias in the evaluation of comparative effectiveness. These evidence gaps pose a significant challenge in designing treatment algorithms that rely on the comparative effectiveness and safety of drugs.[21] As a result, there is large variation in clinical practice guidelines and clinical practice with regard to these medications, with many patients initiated on the newer therapies and many others treated with older regimens.[18,26]

To address this, we developed the LEGEND-T2DM (Large-Scale Evidence Generation and Evaluation Across a Network of Databases for Type 2 Diabetes Mellitus) initiative,[27] a large-scale, systematic, federated evaluation of 4.7 million patients with T2DM across multiple international observational data sources, in which we leveraged robust, state-of-the-art methodological and analytical strategies to minimize residual confounding, publication bias, and *P* hacking. Here, we compared the cardiovascular effectiveness of 4 second-line antihyperglycemic agents—SGLT2is, GLP-1 RAs, DPP4is, and SUs—when initiated on the background of metformin therapy in patients with T2DM.

SEE PAGE 918

METHODS

DATA SOURCES. In this study, we included 10 real-world data sources from the LEGEND-T2DM network, including 6 administrative claims and 4 electronic health record (EHR) databases across 4 countries from 1992 to 2021 (**Central Illustration**). These represent data from 6 national-level and 4 health-system data sets from the United States, as well as data sources from Germany, Spain, and the United Kingdom. The vast majority of patient records span from the mid-2000s to the present, covering 2 decades of T2DM treatment as well as the introduction of many second-line antihyperglycemic agents. All LEGEND-T2DM data sources were previously standardized to the Observational Health Data Sciences and Informatics Observational Medical Outcomes Partnership (OMOP) common data model version 5,[28] which mapped international coding systems into standard vocabulary concepts. The use of the OMOP common data model allows a federated analysis of data, without patient-level data sharing, using consistent cohort definitions and study design. All data partners received institutional approval or exemption for their participation. Details of data sources are presented in Supplemental Tables 1 and 2.

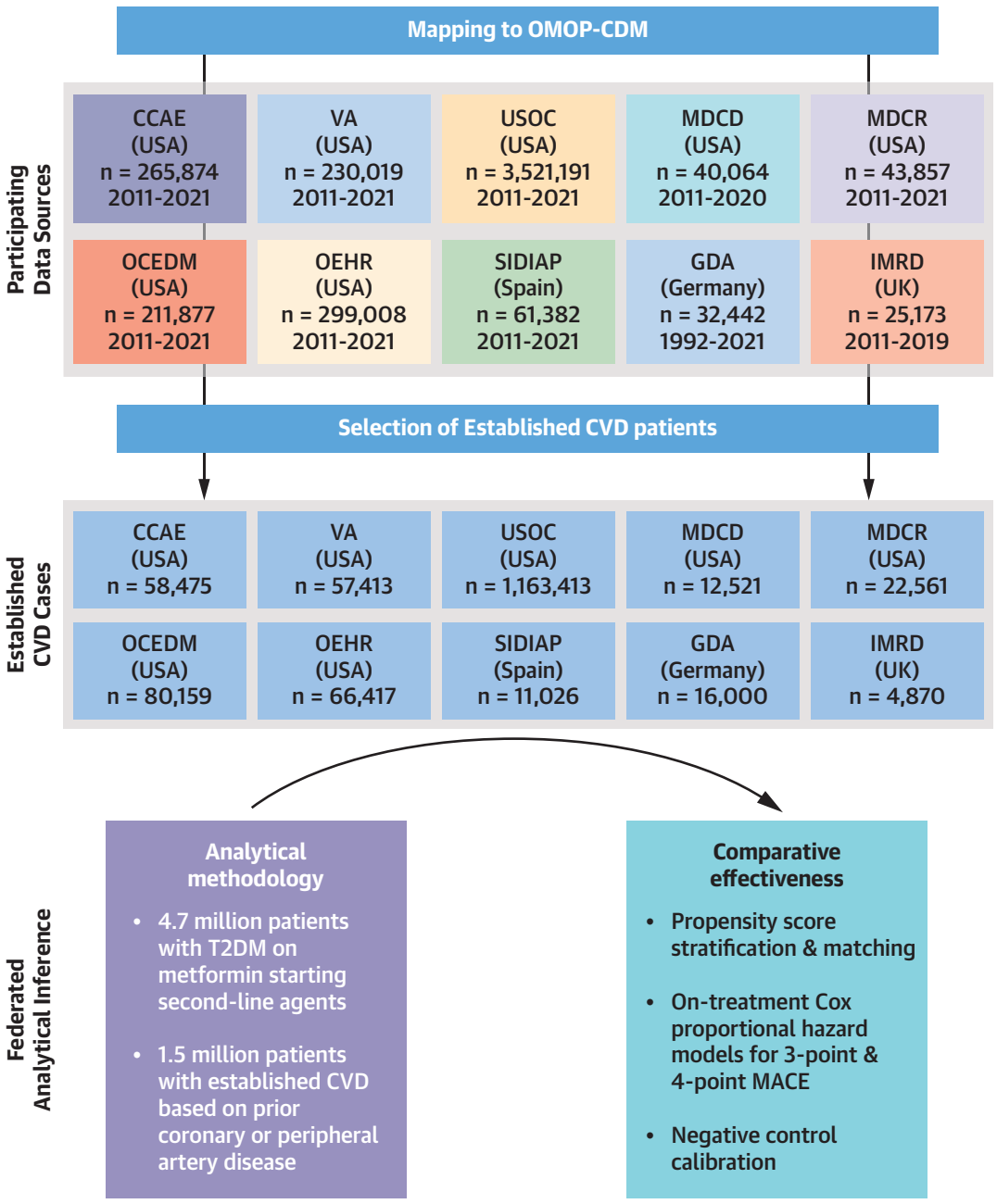
STUDY DESIGN. We identified patients with T2DM and established cardiovascular disease (CVD) on metformin monotherapy who initiated on any of the drug ingredients within one of the SGLT2i, GLP-1 RA, DPP4i, or SU drug classes (Supplemental Methods, Supplemental Table 3).[18] Each exposure cohort thus consisted of new users of each drug class. We require patients with T2DM and CVD to have at least 1 year of prior observation in the database, with at least 3 months of prior metformin use before initiating a second-line agent, and no prior exposure to a comparator second-line or other antihyperglycemic agent or >30 days of insulin exposure (Supplemental Methods). To ensure statistical power, we executed analyses for comparisons and data sources with at least 1,000 patients in each arm. To evaluate the comparative effectiveness of antihyperglycemic agents, we constructed target-comparator-database combinations, in which we compared one antihyperglycemic agent (target) with another agent (comparator) across data sources.

For this, we used a new-user cohort design for target and comparator agents within each data source in order to emulate the hypothetical target trial.[29,30] Methodological principles in the study design were carefully constructed on the basis of the evidence by experts and were leveraged previously to minimize bias and improve reproducibility.[27,31]

STUDY OUTCOMES. Across all data sources and pairwise exposure cohorts, we assessed the relative risks for 2 primary and 4 secondary cardiovascular outcomes. The 2 primary outcomes of interest were: 1) 3-point MACE, including acute MI, stroke, and sudden cardiac death (SCD); and 2) 4-point MACE, which additionally included hospitalization for HF. Secondary outcomes of interest included the 4 individual MACE components. We constructed outcome cohorts on the basis of previously developed phenotypes validated and tested in prior work (Supplemental Methods).[31-36]

For each outcome cohort, we included patients with no events prior to treatment initiation and

CENTRAL ILLUSTRATION **Study Design and Analytical Methodology to Evaluate the Comparative Effectiveness of SGLT2is, GLP-1 RAs, DPP4is, and SUs for Cardiovascular Outcomes**



Khera R, et al. *JACC*. 2024;84(10):904–917.

We leveraged 10 federated data sources from 4 different countries comprising more than 4.7 million individuals with type 2 diabetes mellitus (T2DM) on metformin monotherapy initiating 1 of the 4 second-line agents. We included a subset of 1.5 million patients with established cardiovascular disease (CVD) to assess the comparative cardiovascular effectiveness of these agents. We mapped the data sources to the Observational Medical Outcomes Partnership (OMOP) CDM (common data model) and leveraged robust state-of-the-art methodological and analytical strategies to minimize residual confounding, publication bias, and P hacking. CCAE = IBM MarketScan Commercial Claims and Encounters Data; GDA = Germany Disease Analyser; IMRD = UK-IQVIA Medical Research Data; MACE = major adverse cardiovascular events; MDCD = IBM Health MarketScan Multi-State Medicaid Database; MDCR = IBM Health MarketScan Medicare Supplemental and Coordination of Benefits Database; OCEDM = Optum Clinformatics Extended Data Mart - Date of Death; OEHR = Optum deidentified electronic health record data set; SIDIAP = Information System for Research in Primary Care; USOC = US Open Claims; VA = U.S. Department of Veterans Affairs health care system.

defined continuous drug exposure as consecutive drug prescriptions with <30-day prescription gaps. We considered an on-treatment time-at-risk definition that follows a patient from treatment initiation to treatment discontinuation, which captures direct treatment effects while allowing for escalation with additional T2DM agents.

STATISTICAL ANALYSIS. We used a systematic federated analytical framework to address residual confounding, publication bias, and *P* hacking.[27,37] This framework uses data-driven, large-scale propensity score (PS) adjustment for measured confounding,[38] a large set of negative control outcome experiments to address unmeasured confounding and systematic bias,[39-41] study diagnostics to ensure validity and generalizability,[41,42] and a principled meta-analysis approach to aggregate evidence across data sources. We used standardized vocabularies to construct consistent computable definitions of all study cohorts, covariates, and outcomes. We provide full disclosure of all hypotheses investigated and prespecify and report all analytical procedures in the published protocol of LEGEND-T2DM.[27] To promote open science and avoid publication bias, we have disseminated all results in a publicly available R ShinyApp,[43] and all analytical code is publicly available on GitHub.[44]

Within each data source, we estimated the relative risks for all 6 outcomes between each pair of new-user cohorts, taking one exposure cohort as the target and the other as a comparator. For each pairwise comparison and each data source, we adjusted for measured confounding and improved balance between cohorts by matching and stratifying on PS.[45] We estimated the PS using a data-driven approach that adjusts for a broad range of predefined baseline patient characteristics through regularized regression.[38] These characteristics included demographics, comorbidities, concomitant medication use, and health care use in the period before the initiation of the second-line antihyperglycemic agent. The choice of PS stratification vs matching was based on the approach that achieves a standardized mean difference (SMD) <0.15 across all covariates.[46] When both stratification and matching provided sufficient balance, we preferred stratification over matching and thus reported results on the basis of stratification when available, as the former improves patient inclusion and, therefore, generalizability. We then used Cox proportional hazards models to estimate HRs of each outcome for each comparison, conditional on PS stratification or variable-ratio patient matching.

We also sought to address residual bias, which might persist in observational studies even after PS adjustment that controls for measured confounding.[39,40] For this, we conducted negative control (falsification) outcome experiments for each comparison and outcome, in which the null hypothesis of no differential effect (ie, HR: 1) is believed to be true for each outcome. We selected 100 negative controls through a data-driven algorithm that identifies OMOP condition concept occurrences with similar prevalence to the outcomes of interest that lack evidence of association with exposures in published literature, drug-product labels, and spontaneous reports, which we then confirmed by expert review (Supplemental Methods).[47] We list these negative controls in Supplemental Table 4. From these negative control experiments, we learned an empirical null distribution that informs residual study bias; that is, a deviation from the empirical null across all outcomes represents a quantitative surrogate for the residual bias. We calibrated each original HR estimate to compute a calibrated HR estimate and 95% CI.[39] We declared an HR as significantly different from the null if the calibrated *P* value is <0.05 without considering multiple testing corrections.

We assessed, while blinded to the results, study diagnostics to ensure reliability and generalizability for all comparisons and report only estimates that pass the diagnostics.[41,42] These study diagnostics included: 1) the minimum detectable risk ratio as a metric for statistical power; 2) preference score distributions between the target and comparator cohorts to evaluate empirical equipoise and population generalizability; and 3) cohort balance before and after PS adjustment, defined by the absolute SMDs on extensive patient characteristics. A study passed diagnostics if the minimum detectable risk ratio was <4 and >25% of patients had preference scores between 0.3 and 0.7 in both arms and a maximum SMD <0.15 after PS adjustment. Additional diagnostics for visual examination included: 4) calibration plots on negative control outcomes to examine residual bias; and 5) Kaplan-Meier plots, all publicly available on the R ShinyApp,[43] to visually inspect proportionality assumptions for the Cox models.

We reported all HR estimates, their 95% CIs, and *P* values postcalibration for studies that passed diagnostics. To aggregate evidence across nonoverlapping data sources, we combined all calibrated HR estimates for each comparison using a random-effects meta-analysis approach.[48-50] Furthermore, for each

TABLE 1 Population Size and Follow-Up Time for Initiators of SGLT2is, GLP-1 RAs, DPP4is, and SUs Across Databases

Drug/Data Source	n	On-Treatment Follow-Up Time, d
SGLT2is		
CCAE	10,135	161 (62-537)
GDA	4,337	181 (97-537)
MDCD	1,314	114 (30-383)
MDCR	2,013	112 (34-541)
OCEDM	11,597	118 (55-423)
OEHR	8,996	71 (29-181)
SIDIAP	2,281	322 (103-895)
USOC	196,710	143 (61-525)
VA	7,311	135 (71-443)
Total	244,694	
GLP-1 RAs		
CCAE	7,446	116 (45-441)
MDCR	1,120	94 (29-443)
OCEDM	6,546	98 (36-368)
OEHR	4,525	60 (29-173)
USOC	104,354	112 (52-432)
Total	123,991	
DPP4is		
CCAE	22,801	159 (63-597)
GDA	10,428	221 (97-741)
IMRD	4,024	354 (122-1,174)
MDCD	4,449	142 (56-555)
MDCR	10,541	187 (81-713)
OCEDM	23,918	150 (59-569)
OEHR	18,145	95 (89-312)
SIDIAP	8,126	695 (204-1,916)
USOC	299,835	157 (59-657)
VA	10,969	220 (90-707)
Total	413,236	
SUs		
CCAE	18,093	148 (59-559)
GDA	1,235	216 (119-684)
IMRD	846	350 (106-1,312)
MDCD	6,758	124 (33-503)
MDCR	8,887	159 (59-632)
OCEDM	38,098	177 (85-652)
OEHR	34,751	110 (89-386)
SIDIAP	619	600 (158-1,809)
USOC	562,514	183 (89-752)
VA	39,133	186 (90-689)
Total	710,934	

Values are median (Q1-Q3) unless otherwise indicated.

CCAE = IBM MarketScan Commercial Claims and Encounters Data; DPP4i = dipeptidyl peptidase-4 inhibitor; GDA = Germany Disease Analyzer; GLP-1 RA = glucagon-like peptide-1 receptor agonist; IMRD = UK-IQVIA Medical Research Data; MDCD = IBM Health MarketScan Multi-State Medicaid Database; MDCR = IBM Health MarketScan Medicare Supplemental and Coordination of Benefits Database; OCEDM = Optum Clinformatics Extended Data Mart – Date of Death; OEHR = Optum deidentified electronic health record Dataset; SGLT2i = sodium-glucose cotransporter 2 inhibitor; SIDIAP = Information System for Research in Primary Care; SU = sulfonylurea; USOC = US Open Claims; VA = U.S. Department of Veterans Affairs health care system.

comparison, we performed leave-one-out meta-analyses by calculating the calibrated risk estimates while systematically removing one data source at a time.

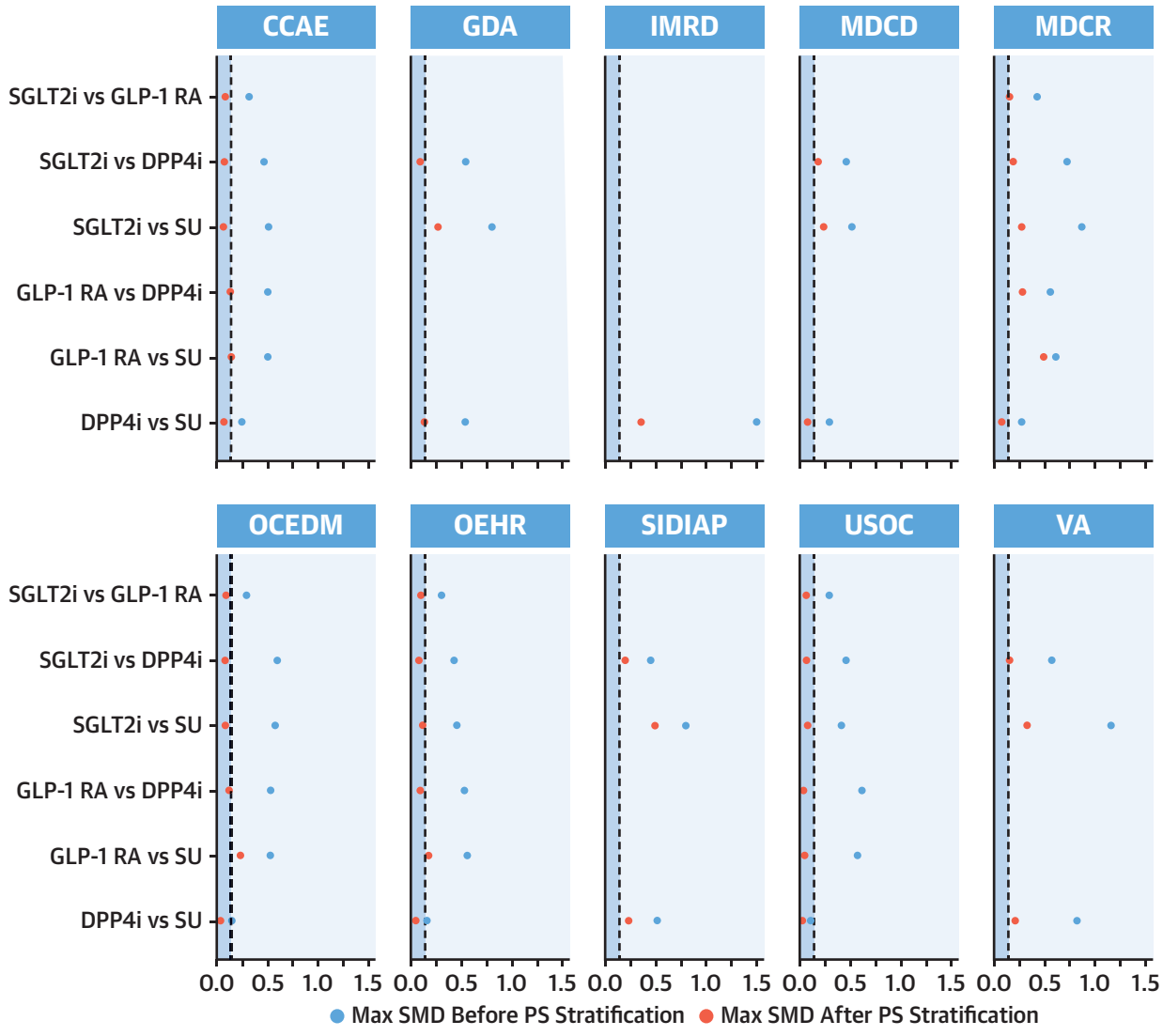
RESULTS

COHORT CHARACTERISTICS. Across 10 federated longitudinal data sources from 4 countries, we identified 4,730,887 patients with T2DM. This included 1,492,855 patients with T2DM and established CVD on metformin monotherapy who initiated 1 of the 4 second-line antihyperglycemic agents and had no prior use of any other antihyperglycemic agents (**Central Illustration**). Among these patients, 244,694 (16.4%) initiated SGLT2is, 123,991 (8.3%) GLP-1 RAs, 413,236 (27.7%) DPP4is, and 710,934 (47.6%) SUs (**Table 1**). U.S. Open Claims with 1,163,413 patients with T2DM and established CVD, followed by Optum Clinformatics Extended Data Mart-Date of Death (n = 80,159) and the Optum deidentified EHR data set (n = 66,417), contributed the most patients to the study population. Median on-treatment time at risk for patients varied by drug class and database between 2.0 and 23.2 months. At least 25% of the patients were exposed to their first drug class for >12 months across a majority of databases (**Table 1**).

ADDRESSING CONFOUNDING ACROSS TARGET-COMPARATOR-DATABASE COMBINATIONS. PS adjustment achieved prespecified covariate balance in patient baseline characteristics for pairwise class comparisons in a majority of databases (**Figure 1**, Supplemental Tables 5 to 47). The maximum SMD across target-comparator-database combinations consistently decreased after PS stratification (**Figure 1**). For example, in the comparison of patients initiating SGLT2is (target) with patients initiating DPP4is (comparator) in the IBM MarketScan Commercial Claims and Encounters database, before PS adjustment, patients initiating SGLT2is were more frequently men and had obesity and HF relative to patients initiating DPP4is. However, after PS adjustment, the SGLT2i and DPP4i populations were well balanced on all demographic and clinical patient characteristics (Supplemental Table 10).

EMPIRICAL EQUIPOISE ACROSS TARGET-COMPARATOR-DATABASE COMBINATIONS. For most data sources, all executed class comparisons were in empirical equipoise (>25% of patients had preference scores between 0.3 and 0.7 in both arms) (Supplemental Figures 1 to 10). The Germany Disease Analyser, Information System for Research in Primary Care, and U.S. Department of Veterans Affairs databases showed less equipoise for comparisons involving SGLT2is and DPP4is. However, in general, PS adjustment achieved sufficient covariate balance in terms of preference score distribution to reduce concerns that measured the estimated effects of

FIGURE 1 **Maximum SMD Before and After Propensity Score Stratification for All Covariates Across Target-Comparator-Database Combinations**

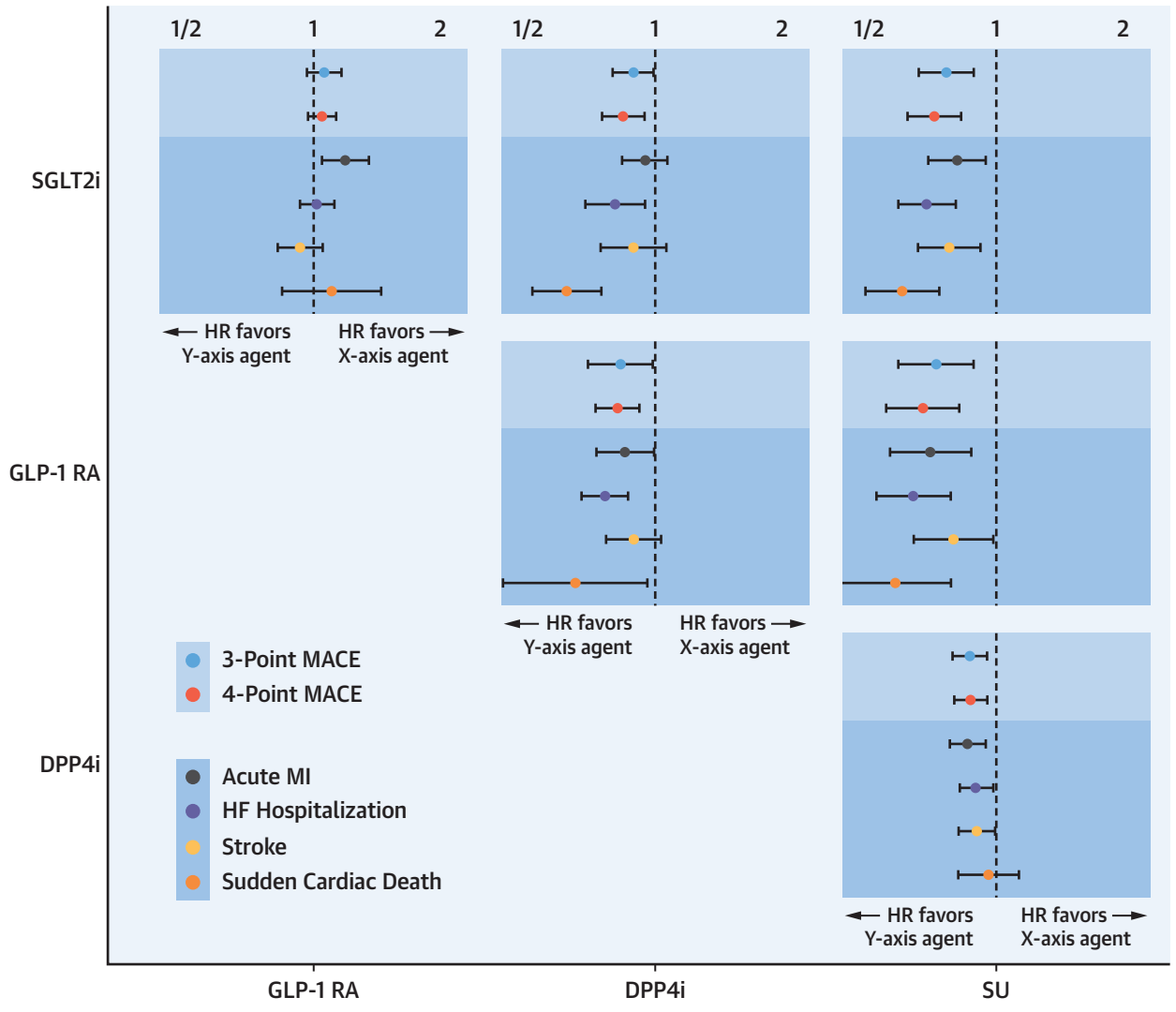


The figure shows to what extent the propensity score (PS) stratification reduced the maximum standardized mean difference (SMD) for all covariates across 6 target-comparator pairs and 10 data sources. CCAE = IBM MarketScan Commercial Claims and Encounters Data; DPP4i = dipeptidyl peptidase 4 inhibitor; GDA = Germany Disease Analyzer; GLP-1 RA = glucagon-like peptide-1 receptor agonist; IMRD = UK-IQVIA Medical Research Data; MDCD = IBM Health MarketScan Multi-State Medicaid Database; MDCR = IBM Health MarketScan Medicare Supplemental and Coordination of Benefits Database; OCEDM = Optum Clinformatics Extended Data Mart - Date of Death; OEHR = Optum deidentified electronic health record data set; SGLT2i = sodium-glucose cotransporter 2 inhibitor; SIDIAP = Information System for Research in Primary Care; SU = sulfonylurea; USOC = US Open Claims; VA = U.S. Department of Veterans Affairs health care system.

baseline confounding biases (Supplemental Figures 11 to 30). Furthermore, the study had limited residual systematic error, where calibration of effect estimates using negative control outcomes resulted in an increase in the proportion of nominal 95% CIs that included 1 for control outcomes across a majority of comparisons and databases (Supplemental Figures 31 to 50). For example, for the comparison of SGLT2is vs DPP4is in the Information System for Research in Primary Care using PS matching, before calibration, the nominal CIs covered 75.0% of control estimates; after calibration, they covered 91.7% (Supplemental Figure 48).

COMPARATIVE EFFECTIVENESS FOR PRIMARY ENDPOINTS. During 1,337,809 patient-years of follow-up, there were 25,982 3-point MACE and 41,447 4-point MACE. SGLT2is and GLP-1 RAs were

FIGURE 2 Meta-Analytical Calibrated HR Estimates for Comparative Effectiveness of SGLT2is, GLP-1 RAs, DPP4is, and SUs for Cardiovascular Outcomes

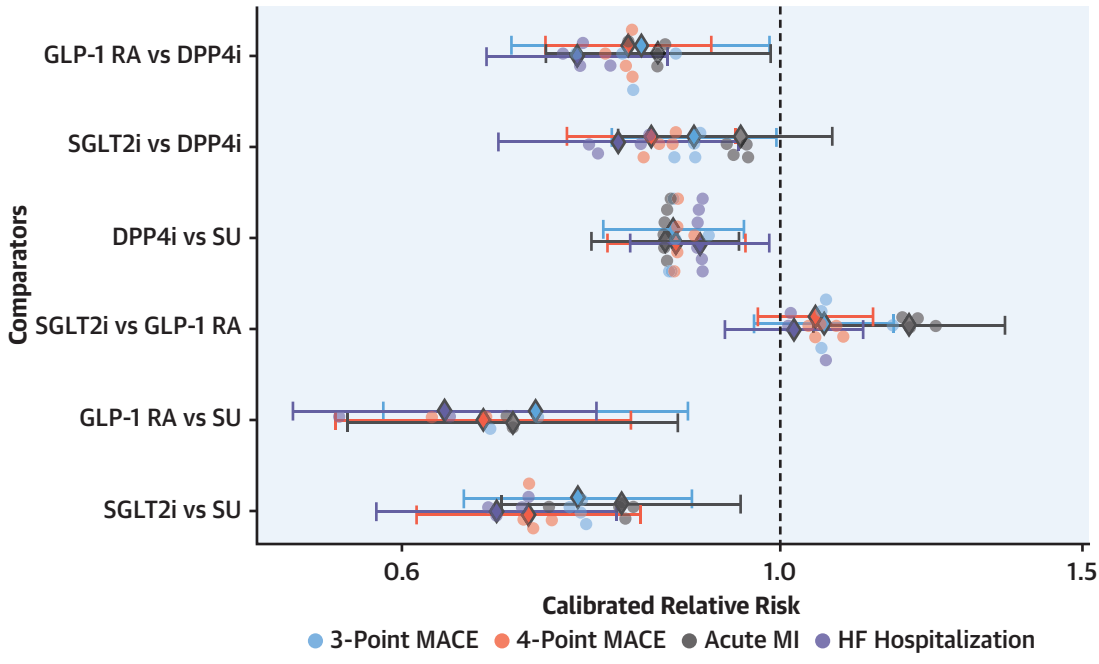


The figure shows the comparative effectiveness of SGLT2is, GLP-1 RAs, DPP4is, and SUs for composite and individual cardiovascular outcomes across 6 target-comparator pairs, with the y-axis representing target and the x-axis showing comparators. Points on the left side of the dashed line represent the HR that favors the target (y-axis), and points on the right side of the dashed line represent the HR that favors the comparator (x-axis). HF = heart failure; MI = myocardial infarction; other abbreviations as in Figure 1.

associated with a lower hazard of 3-point MACE compared with DPP4is (HR: 0.89 [95% CI: 0.79-1.00] and 0.83 [95% CI: 0.70-0.98], respectively) and with even lower HR of 0.76 (95% CI: 0.65-0.89) and 0.72 (95% CI: 0.58-0.88) vs SUs, respectively. DPP4is were associated with a lower risk for 3-point MACE risk vs SUs (HR: 0.87; 95% CI: 0.79-0.95). A consistent pattern was observed for 4-point MACE across the aforementioned comparisons. In direct comparisons of SGLT2is and GLP-1 RAs, there were no differences in either 3-point or 4-point MACE (HR: 1.06 [95% CI: 0.96-1.17] and 1.05 [95% CI: 0.97-1.13], respectively) (Supplemental Tables 48 to 60, Figure 2).

COMPARATIVE EFFECTIVENESS FOR SECONDARY ENDPOINTS. Acute MI. There were 13,536 episodes of acute MI across data sources during follow-up. SGLT2is, GLP-1 RAs, and DPP4is were associated with a lower hazard of acute MI compared with SUs (HR: 0.81 [95% CI: 0.69-0.95], 0.70 [95% CI: 0.56-0.87], and 0.86 [95% CI: 0.77-0.95], respectively).

FIGURE 3 Swarm Plot of Calibrated HR Estimates of Major Outcome Meta-Analysis and Leave-One-Out Meta-Analysis



Circles depict the calibrated relative risk of each leave-one-out study in which 1 original data source is removed from the meta-analysis. Diamonds depict the original meta-analysis effect estimate with all data sources. Points are color-coded by major outcomes, the y-axis represents 6 target-comparator pairs, and the x-axis on log scale measures the calibrated relative risk for each outcome. Abbreviations as in Figures 1 and 2.

Although GLP-1 RAs were associated with a lower risk for acute MI compared with DPP4is (HR: 0.85; 95% CI: 0.73-0.99), the hazard of acute MI was comparable for SGLT2is vs DPP4is (HR: 0.95; 95% CI: 0.83-1.08). Compared with GLP-1 RAs, SGLT2is were associated with a higher hazard of acute MI (HR: 1.19; 95% CI: 1.05-1.35) (Supplemental Tables 48 and 61 to 66, **Figure 2**).

Stroke. There were 13,999 episodes of stroke events across data sources during follow-up. SGLT2is, GLP-1 RAs, and DPP4is were associated with a lower hazard of stroke compared with SUs (HR: 0.77 [95% CI: 0.65-0.92], 0.79 [95% CI: 0.63-0.98], and 0.90 [95% CI: 0.81-0.99], respectively). Compared with DPP4is, SGLT2is and GLP-1 RAs had a comparable hazard of stroke (HR: 0.89 [95% CI: 0.74-1.07] and 0.89 [95% CI: 0.77-1.03], respectively). There was no significant difference between SGLT2is and GLP-1 RAs (HR: 0.93; 95% CI: 0.82-1.05) regarding the hazard of stroke (Supplemental Tables 48 and 67 to 72, **Figure 2**).

SCD. Over follow-up, 3,789 SCDs occurred across data sources. SGLT2is and GLP-1 RAs were associated with a lower hazard of SCD compared with DPP4is (HR: 0.62 [95% CI: 0.51-0.75] and 0.65 [95% CI: 0.44-0.96], respectively) and SUs (HR: 0.60 [95% CI: 0.48-0.73] and 0.57 [95% CI: 0.42-0.78], respectively). DPP4is vs SUs and SGLT2is vs GLP-1 RAs had comparable hazards of SCD (HR: 0.96 [95% CI: 0.81-1.13] and 1.11 [95% CI: 0.84-1.45], respectively) (Supplemental Tables 48 and 73 to 78, **Figure 2**).

HF HOSPITALIZATION. During follow-up, 30,743 HF hospitalizations were recorded across all databases. SGLT2is and GLP-1 RAs were associated with a lower risk for HF hospitalization compared with DPP4is (HR: 0.80 [95% CI: 0.68-0.95] and 0.76 [95% CI: 0.67-0.86], respectively) and SUs (HR: 0.68 [95% CI: 0.58-0.80] and 0.64 [95% CI: 0.52-0.78], respectively). DPP4is were associated with a lower hazard of HF hospitalization compared with SUs (HR: 0.90; 95% CI: 0.82-0.99). SGLT2i and GLP-1 RAs had comparable hazard for HF hospitalization (HR for SGLT2is vs GLP-1 RAs: 1.02; 95% CI: 0.93-1.12) (Supplemental Tables 48 and 79 to 84, **Figure 2**).

SENSITIVITY ANALYSES. Across all outcomes, 98.4% of calibrated relative risk estimates (690 of 701) remained within the CIs of the primary meta-analysis

when systematically removing each data source from the leave-one-out meta-analysis. Within the primary endpoints of 3-point MACE and 4-point MACE and secondary endpoints of acute MI and hospitalization for HF, none of the leave-one-out analyses were outside the CIs of the meta-analysis (**Figure 3**, Supplemental Figures 51 to 56).

DISCUSSION

In this largest multinational, federated study of 1.5 million patients with T2DM and established CVD across 10 cohorts initiating second-line antihyperglycemic agents after metformin monotherapy, SGLT2is and GLP-1 RAs were associated with an 11% and 17% lower risk for cardiovascular events compared with DPP4is, respectively, the class of drugs approved before the U.S. Food and Drug Administration–mandated cardiovascular outcomes trials for antihyperglycemic agents. All 3 agents were associated with a lower risk of cardiovascular events than SUs, with 24% and 28% lower risk with SGLT2is and GLP-1 RAs and 13% lower risk with DPP4is. These patterns were consistent across both ischemic events, such as MI and stroke, as well as for HF. Between SGLT2is and GLP-1 RAs, there was no difference in the composite cardiovascular effectiveness outcomes, 3-point and 4-point MACE, but there was a lower risk for acute MI with GLP-1 RAs compared with SGLT2is. The study accounted for multiple null control outcomes and calibrated all estimates for residual bias on null control outcomes, with multiple sensitivity analyses confirming the findings of the primary analysis.

This builds on the evidence of comparative cardiovascular effectiveness of currently used antihyperglycemic therapies, especially in the sparse landscape of head-to-head RCTs of any of these agents. The CAROLINA (Cardiovascular Outcome Study of Linagliptin Versus Glimepiride in Patients With Type 2 Diabetes) trial is the only RCT that pursued a head-to-head evaluation of the cardiovascular safety of a DDP4i (linagliptin) and an SU (glimepiride) and found no difference between the 2 drugs.[14] In contrast, we found a lower risk for cardiovascular adverse outcomes with DPP4i use compared with SUs. Class-level comparisons and assessments across large and diverse populations likely contribute to these observed differences. Moreover, this does not suggest an absolute benefit from DDP4is but might even reflect a relative harm from SUs, which has been suggested in prior studies.[51-55] There are no ongoing or future trials expected to fill this knowledge gap. Moreover, a few observational studies have used single data sources to evaluate the cardiovascular outcomes associated with these drugs.[56]

Some associations observed in the present study, such as those with DPP4is, contrast with others in which the smaller study population might have necessitated studying a combination of drugs directly against SUs but posed challenges in isolating an effect.[18] These analyses are limited by local patterns of treatment use and challenges with exposure and outcome ascertainment in individual sources that might confound the observed association of these agents with outcomes, further supporting the need for comparative effectiveness evaluations at scale. Our recent work shows that the use of antihyperglycemic agents varies substantially in populations across the world.[18] In the present study, we developed a target trial for a head-to-head comparison of these agents, focusing explicitly on the same disease stage (ie, when escalation to a second-line agent is being done). We conducted the same study across 10 different data sources, ensuring that the effects of the intervention were not limited to the selective pressures or outcome differences in a single population. We ensured empirical equipoise by comparing SMDs before and after PS stratification. Finally, we ensured that unmeasured confounding, potentially manifesting as directional effects across several negative control outcomes, was used to calibrate the PS stratified effect estimates.

The findings have important implications for clinical practice. Our observations support the recent change in clinical practice guidelines recommending first-line use of SGLT2is and GLP-1 RAs among individuals with established CVD.[57-60] However, our recent work shows that a substantial proportion of patients with T2DM are often initiated on DPP4is and SUs despite their limited cardiovascular effectiveness or known superiority on other glycemic or other diabetes-related outcomes.[58-60] There is also evidence that suggests a potentially larger reduction in risk for acute MI with GLP-1 RAs compared with SGLT2is, which was not replicated for stroke. We found no difference among SGLT2is, GLP-1 RAs, and DPP4is in terms of the stroke outcome. However, SGLT2is, GLP-1 RAs, and DPP4is were all associated with a lower risk for stroke compared with SUs. In the absence of head-to-head comparisons of these agents for the stroke outcome, our findings are in line with the collective evidence from cardiovascular outcome trials, in which these agents were compared against placebo.[15,61]

Moreover, we did not observe an exclusive role of decreased HF risk with SGLT2is, as suggested in

current guidelines, as there was no observed difference in HF risk among those receiving SGLT2is or GLP-1 RAs. The original recommendations for HF were based on inference from RCTs of individual agents and emerging evidence. In previous trials of GLP-1 RAs, there was a lack of protocolized assessment of left ventricular ejection fraction, natriuretic peptides, and HF status, which limits inference on the effect on HF outcomes. However, the STEP-HFpEF (Research Study to Investigate How Well Semaglutide Works in People Living With Heart Failure and Obesity) trial, which carefully selected participants with confirmed HF with preserved ejection fraction at baseline and closely monitored HF outcomes, demonstrated improvements in HF outcomes with semaglutide.[62] Moreover, there was no interaction between the recorded ejection fraction and HF outcomes in STEP-HFpEF, suggesting the potential broader role of GLP-1 RAs on HF-related adverse events (NCT04847557, NCT05371496). Our study favors the continued evaluation of GLP-1 RAs as a therapy to reduce HF risk and improve outcomes among those with HF.[63-66]

STUDY LIMITATIONS. First, to capture information across data sources spanning administrative and EHR data sources necessitated the use of administrative definitions of both exposures and outcomes. However, to ensure that the information is accurately captured, we used definitions that have been validated in prior studies.[67,68]

Second, we expect the identification of the exposure and outcomes to vary across data sources. We conducted a random-effects meta-analysis to account for this expected variability. Findings of the leave-one-out sensitivity analyses further ensured the consistency of the effect estimates across data sources. Although we did not explicitly report heterogeneity statistics, the consistency of effect estimates indirectly supports the assumption that our findings are not overly sensitive to any specific data source and further confirms the validity of the observations.

Third, we adopted an alpha criterion of 0.05 without adjusting for multiple testing. Of note, we did not seek to test any key hypotheses, especially with the paucity of any quality evidence in this clinically important domain. Instead, we focused on rigorous methods and thoughtful interpretation that qualitatively assessed the directional consistency of effects across outcomes. Furthermore, we reported 95% CIs of the effect estimates and did not report any *P* values, to draw attention to the effect size rather than significance testing.

Fourth, capturing outcomes for EHR-based data sets, such as for recurrent hospitalizations and deaths, differs from administrative sources that track events across institutions and capture out-of-hospital death events.[16,17] However, we found that sources that capture cross-institution out-of-hospital information did not differ and had similar outcomes as those that focused on institutionally captured death events, further supporting the robustness of the drug association with outcomes.

Fifth, the study did not explicitly address the potential for bias introduced by adherence differences between agents. Nevertheless, pooling of data across multiple data sources with expected variation in adherence likely addresses the issues with differential adherence in a specific group. Moreover, the consistency of effects in analyses that excluded individual data sets also is reassuring.

Sixth, the relatively short time at risk for antihyperglycemic agents might limit inferences on long-term outcomes. The shorter follow-up time is a function of the rigorous on-treatment design of the study, in which patients exit the cohort in case of treatment discontinuation or escalation, which is essential for minimizing the misattribution of treatment effects. Moreover, although a shorter follow-up period with a smaller number of events can bias the effect estimates toward the null, we found differential comparative effectiveness for the included agents.

Seventh, inference on all outcomes can still be subject to residual confounding. However, our use of calibrated estimates explicitly addresses the identification of residual confounding across a series of null control outcomes.

Eighth, the study did not account for the competing risk for death. This was an a priori decision given that the population was not presumably at marked total mortality risk and the computational efficiency of Cox proportional hazards models. The consistency of the composite MACE that included SCD data with the individual outcomes suggests that mortality effects did not meaningfully differ across groups.

Finally, given the evolving indications for these drug classes for the management of T2DM, it may be essential to evaluate their comparative effectiveness as first-line agents without metformin background therapy. Additionally, although the study's scope was limited to class-level analyses, there is an unmet need for further evidence generation on drug-level comparative effectiveness. These evaluations represent the goals of the group's future work.

Furthermore, future comparative effectiveness studies could incorporate outcomes such as hemoglobin A_{1c} and weight to better understand potential mediating effects.

CONCLUSIONS

Our large, rigorous, multinational network study of patients with T2DM and established CVD demonstrated cardiovascular risk reduction with SGLT2is and GLP-1 RAs, with both agents more effective than DPP4is, which in turn were more effective than SUs. These findings suggest that the use of GLP-1 RAs and SGLT2is should be prioritized as second-line agents in those with established CVD.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This study was partially funded through National Institutes of Health grants R01 HL167858, K23 HL153775, R01 LM006910, R01 HG006139, and R01 HL169954 and the U.S. Department of Veterans Affairs under the research priority to "Put VA Data to Work for Veterans" (VA ORD 22-D4V). The funders had no role in the design and conduct of the protocol; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Dr Khera is an associate editor at *JAMA*; and has received support from the National Heart, Lung, and Blood Institute of the National Institutes of Health (awards R01 HL167858 and K23 HL153775) and the Doris Duke Charitable Foundation (award 2022060); has received research support, through Yale University, from Bristol Myers Squibb, Novo Nordisk, and BridgeBio; is a coinventor of U.S. provisional patent applications WO2023230345A1, US20220336048A1, 63/346,610, 63/428,569, 63/484,426, 63/508,315, 63/580,137, 63/606,203, 63/619,241, and 63/562,335, unrelated to the present work; and is a cofounder of Evidence2Health and Ensight-AI, precision health platforms to improve evidence-based cardiovascular care and cardiovascular diagnostics. Dr DuVall has received grants from Alnylam Pharmaceuticals, AstraZeneca Pharmaceuticals, Biodesix, Celgene, Cerner Enviza, GlaxoSmithKline, Janssen Pharmaceuticals, Novartis International, and Parexel International through the University of Utah or the Western Institute for Veteran Research (outside the submitted work). Dr Man has received support from the C.W. Maplethorpe Fellowship, the National Institute of Health Research, European Commission Framework Horizon 2020, the Hong Kong Research Grant Council, and the Innovation and Technology Commission of the Hong Kong Special Administration Region Government outside the submitted work). Dr Morales is supported by a Wellcome Trust Clinical Research Fellowship (214588/Z/18/Z). In the past 3 years, Dr Krumholz has received expenses and/or personal fees from UnitedHealth, Element Science, Aetna, Reality Labs, Tesseract/4Catalyst, F-Prime, the Siegfried and Jensen Law Firm, the Arnold and Porter Law Firm, and the Martin/Baughman Law Firm; is a cofounder of Refactor Health, HugoHealth, and Ensight-AI; and is associated with contracts, through Yale New Haven Hospital, from the Centers for Medicare and Medicaid Services and, through Yale University, from Johnson & Johnson. Dr You is chief technology officer of PHI Digital Healthcare. Dr Ross has received research support through Yale University from Johnson & Johnson to develop methods of clinical trial data sharing, from the Food and Drug Administration for the Yale-Mayo Clinic Center for Excellence in Regulatory Science and Innovation program (grant U01 FD005938), from the Agency for Healthcare Research and Quality (grant R01 HS022882), and from Arnold Ventures; has received research support from the Medical Device Innovation Consortium as part of the National Evaluation System for Health Technology and from the National Heart, Lung, and Blood Institute of the National Institutes of Health (grants R01 HS025164 and R01 HL144644); and was an expert witness at the request of the relator's attorneys, the Greene Law Firm, in a qui tam suit alleging violations of the False Claims Act and Anti-Kickback Statute against Biogen that was settled September 2022. Dr Thangaraj is a coinventor on provisional patent 63/606,203, unrelated to the present work, and is funded by grant 5T32 HL155000-03. Ms Blacketer, Dr Ostropolets, and Dr Ryan are employees of Johnson & Johnson. Dr Schuemie is an employee and shareholder of Johnson & Johnson. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Marc A. Suchard, 6558 Gonda Building, 695 Charles E. Young Drive, South Los Angeles, California 90095-1766, USA. E-mail: msuchard@ucla.edu. X handle: @suchard_group. OR Dr Rohan Khera, 195 Church Street, 6th Floor, New Haven, Connecticut 06510, USA. E-mail: rohan.khera@yale.edu. X handle: @rohan_khera.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: The LEGEND-T2DM initiative is a large-scale, systematic, federated evaluation of individuals with T2DM across multiple international observational data sources, in which we leveraged robust, state-of-the-art methodological and analytical strategies to minimize residual confounding, publication bias, and *P* hacking. In this largest study of 1.5 million patients with T2DM and established CVD across 10 cohorts from 4 countries initiating second-line antihyperglycemic agents—SGLT2is, GLP-1 RAs, DPP4is, or SUs—after metformin monotherapy, SGLT2is and GLP-1 RAs were associated with 11% and 17% lower risk for cardiovascular events compared with DPP4is, respectively. All 3 agents were associated with a lower risk of cardiovascular events than SUs, with 24% and 28% lower risk with SGLT2is and GLP-1 RAs and 12% lower risk with DPP4is. These patterns were consistent across both ischemic events, such as MI and stroke, as well as for HF.

COMPETENCY IN PATIENT CARE: Our study defines strong empirical evidence in support of contemporary clinical practice guidelines suggesting that SGLT2is and GLP-1 RAs are used for those with T2DM and CVD. Our observations also provide support for the exclusive second-line use of SGLT2is and GLP-1 RAs among those with CVD.

REFERENCES

1. Prattichizzo F, La Sala L, Rydén L, et al. Glucose-lowering therapies in patients with type 2 diabetes and cardiovascular diseases. *Eur J Prev Cardiol*. 2019;26:73–80.

2. Hernandez AF, Green JB, Janmohamed S, et al. Albiglutide and cardiovascular outcomes in patients with type 2 diabetes and cardiovascular disease (Harmony Outcomes): a double-blind, randomised placebo-controlled trial. *Lancet*. 2018;392:1519–1529.

3. Solomon SD, McMurray JJV, Claggett B, et al. Dapagliflozin in heart failure with mildly reduced or preserved ejection fraction. *N Engl J Med*. 2022;387:1089–1098.

4. Pfeffer MA, Claggett B, Diaz R, et al. Lixisenatide in patients with type 2 diabetes and acute coronary syndrome. *N Engl J Med*. 2015;373:2247–2257.

5. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2016;375:311–322.

6. Marso SP, Bain SC, Consoli A, et al. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. *N Engl J Med*. 2016;375:1834–1844.

7. Holman RR, Bethel MA, Mentz RJ, et al. Effects of once-weekly exenatide on cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2017;377:1228–1239.

8. Gerstein HC, Colhoun HM, Dagenais GR, et al. Dulaglutide and cardiovascular outcomes in type 2 diabetes (REWIND): a double-blind, randomised placebo-controlled trial. *Lancet*. 2019;394:121–130.

9. Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2019;380:347–357.

10. Zinman B, Bhosekar V, Busch R, et al. Semaglutide once weekly as add-on to SGLT-2 inhibitor therapy in type 2 diabetes (SUSTAIN 9): a randomised, placebo-controlled trial. *Lancet Diabetes Endocrinol*. 2019;7:356–367.

11. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med*. 2017;377:644–657.

12. Azoulay L, Suissa S. Sulfonylureas and the risks of cardiovascular events and death: A methodological meta-regression analysis of the observational studies. *Diabetes Care*. 2017;40:706–714.

13. Vaccaro O, Masulli M, Nicolucci A, et al. Effects on the incidence of cardiovascular events of the addition of pioglitazone versus sulfonylureas in patients with type 2 diabetes inadequately controlled with metformin (TOSCA.IT): a randomised, multicentre trial. *Lancet Diabetes Endocrinol*. 2017;5:887–897.

14. Rosenstock J, Kahn SE, Johansen OE, et al. Effect of linagliptin vs glimepiride on major adverse cardiovascular outcomes in patients with type 2 diabetes: the CAROLINA randomized clinical trial. *JAMA*. 2019;322:1155–1166.

15. Bain S, Druyts E, Balijepalli C, et al. Cardiovascular events and all-cause mortality associated with sulphonylureas compared with other antihyperglycaemic drugs: a Bayesian meta-analysis of survival data. *Diabetes Obes Metab*. 2017;19:329–335.

16. American Diabetes Association Professional Practice Committee. 9. Pharmacologic approaches to glycemic treatment: standards of care in diabetes—2024. *Diabetes Care*. 2024;47:S158–S178.

17. Cosentino F, Grant PJ, Aboyans V, et al. 2019 ESC guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. *Eur Heart J*. 2020;41:255–323.

18. Khera R, Dhingra LS, Aminorroaya A, et al. Multinational patterns of second line antihyperglycaemic drug initiation across cardiovascular risk groups: federated pharmacoepidemiological evaluation in LEGEND-T2DM. *BMJ Med*. 2023;2:e000651.

19. Garber AJ, Handelsman Y, Grunberger G, et al. Consensus statement by the American Association of Clinical Endocrinologists and American College of Endocrinology on the comprehensive type 2 diabetes management algorithm - 2020 executive summary. *Endocr Pract*. 2020;26:107–139.

20. Blonde L, Umpierrez GE, Reddy SS, et al. American Association of Clinical Endocrinology clinical practice guideline: developing a diabetes mellitus comprehensive care plan—2022 update. *Endocr Pract*. 2022;28:923–1049.

21. Tobias DK, Merino J, Ahmad A, et al. Second international consensus report on gaps and opportunities for the clinical translation of precision diabetes medicine. *Nat Med*. 2023;29:2438–2457.

22. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015;373:2117–2128.

23. Fu EL, Clase CM, Janse RJ, et al. Comparative effectiveness of SGLT2i versus GLP1-RA on cardiovascular outcomes in routine clinical practice. *Int J Cardiol*. 2022;352:172–179.

24. Giugliano D, Longo M, Signoriello S, et al. The effect of DPP-4 inhibitors, GLP-1 receptor agonists and SGLT-2 inhibitors on cardiorenal outcomes: a network meta-analysis of 23 CVOTs. *Cardiovasc Diabetol*. 2022;21:42.

25. Gilbert MP, Pratley RE. GLP-1 analogs and DPP-4 inhibitors in type 2 diabetes therapy: review of head-to-head clinical trials. *Front Endocrinol*. 2020;11:178.

26. Nargesi AA, Jeyashanmugaraja GP, Desai N, Lipska K, Krumholz H, Khera R. Contemporary national patterns of eligibility and use of novel cardioprotective antihyperglycemic agents in type 2 diabetes mellitus. *J Am Heart Assoc*. 2021;10:e021084.

27. Khera R, Schuemie MJ, Lu Y, et al. Large-scale evidence generation and evaluation across a network of databases for type 2 diabetes mellitus (LEGEND-T2DM): a protocol for a series of multinational, real-world comparative cardiovascular effectiveness and safety studies. *BMJ Open*. 2022;12:e057977.

28. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc*. 2012;19:54–60.

29. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. 2019;25:1601–1606.

30. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183:758–764.

31. Suchard MA, Schuemie MJ, Krumholz HM, et al. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. *Lancet*. 2019;394:1816–1826.

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

33. You SC, Rho Y, Bikdeli B, et al. Association of ticagrelor vs clopidogrel with net adverse clinical events in patients with acute coronary syndrome undergoing percutaneous coronary intervention. *JAMA*. 2020;324:1640–1650.

34. Wang Y, Desai M, Ryan PB, et al. Incidence of diabetic ketoacidosis among patients with type 2 diabetes mellitus treated with SGLT2 inhibitors and other antihyperglycemic agents. *Diabetes Res Clin Pract*. 2017;128:83–90.

35. Weinstein RB, Ryan PB, Berlin JA, Schuemie MJ, Swerdel J, Fife D. Channeling bias in the analysis of risk of myocardial infarction, stroke, gastrointestinal bleeding, and acute renal failure with the use of paracetamol compared with ibuprofen. *Drug Saf*. 2020;43:927–942.

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

37. Schuemie MJ, Ryan PB, Pratt N, et al. Principles of Large-Scale Evidence Generation and Evaluation Across a Network of Databases (LEGEND). *J Am Med Inform Assoc*. 2020;27:1331–1337.

38. Tian Y, Schuemie MJ, Suchard MA. Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *Int J Epidemiol*. 2018;47:2005–2014.

39. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proc Natl Acad Sci U S A*. 2018;115:2571–2577.

40. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. *Stat Med*. 2016;35:3883–3888.

41. Schuemie MJ, Ryan PB, Hripcsak G, Madigan D, Suchard MA. Improving reproducibility by using high-throughput observational studies with empirical calibration. *Philos Trans A Math Phys Eng Sci*. 2018;376(2128):20170356.

42. Schuemie MJ, Cepeda MS, Suchard MA, et al. How confident are we about observational findings in healthcare: a benchmark study. *Harv Data Sci Rev*. 2020;2(1).

43. LEGEND-T2DM Investigators. Large-Scale Evidence Generation and Evaluation Across a Network of Database for Type 2 Diabetes Mellitus (LEGEND-T2DM). Observational Health Data Sciences and Informatics. Accessed May 3, 2024. https://data.ohdsi.org/LegendT2dmClassEvidenceExplorer/

44. LEGEND-T2DM Investigators. LEGENDT2DM GitHub repository. Accessed May 3, 2024. https://github.com/ohdsi-studies/LegendT2dm

45. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.

46. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28:3083–3107.

47. Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. Accuracy of an automated knowledge base for identifying drug adverse reactions. *J Biomed Inform*. 2017;66:72–81.

48. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7:177–188.

49. Davies MJ, D'Alessio DA, Fradkin J, et al. Management of hyperglycemia in type 2 diabetes, 2018. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care*. 2018;41:2669–2701.

50. Inzucchi SE, Bergenstal RM, Buse JB, et al. Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care*. 2012;35:1364–1379.

51. Richardson TL Jr, Halvorson AE, Hackstadt AJ, et al. Primary occurrence of cardiovascular events after adding sodium-glucose cotransporter-2 inhibitors or glucagon-like peptide-1 receptor agonists compared with dipeptidyl peptidase-4 inhibitors: a cohort study in veterans with diabetes. *Ann Intern Med*. 2023;176:751–760.

52. Kutz A, Kim DH, Wexler DJ, et al. Comparative cardiovascular effectiveness and safety of SGLT-2 inhibitors, GLP-1 receptor agonists, and DPP-4 inhibitors according to frailty in type 2 diabetes. *Diabetes Care*. 2023;46:2004–2014.

53. Thein D, Christiansen MN, Mogensen UM, et al. Add-on therapy in metformin-treated patients with type 2 diabetes at moderate cardiovascular risk: a nationwide study. *Cardiovasc Diabetol*. 2020;19:107.

54. Vashisht R, Patel A, Dahm L, et al. Second-line pharmaceutical treatments for patients with type 2 diabetes. *JAMA Netw Open*. 2023;6:e2336613.

55. Xie Y, Bowe B, Xian H, Loux T, McGill JB, Al-Aly Z. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of major adverse cardiovascular events: emulation of a randomised target trial using electronic health records. *Lancet Diabetes Endocrinol*. 2023;11:644–656.

56. Wang H, Cordiner RLM, Huang Y, et al. Cardiovascular safety in type 2 diabetes with sulfonylureas as second-line drugs: a nationwide population-based comparative safety study. *Diabetes Care*. 2023;46:967–977.

57. Samson SL, Vellanki P, Blonde L, et al. American Association of Clinical Endocrinology consensus statement: comprehensive type 2 diabetes management algorithm—2023 Update. *Endocr Pract*. 2023;29:305–340.

58. Khera R, Dhingra LS, Aminorroaya A, et al. Multinational patterns of second-line anti-hyperglycemic drug initiation across cardiovascular risk groups: a federated pharmacoepidemiologic evaluation in LEGEND-T2DM. Preprint 22283968. *medRxiv*. Posted online December 30, 2022. https://doi.org/10.1101/2022.12.27.22283968

59. McGuire DK, Alexander JH, Johansen OE, et al. Linagliptin effects on heart failure and related outcomes in individuals with type 2 diabetes mellitus at high cardiovascular and renal risk in CARMELINA. *Circulation*. 2019;139:351–361.

60. Rehman MB, Tudrej BV, Soustre J, et al. Efficacy and safety of DPP-4 inhibitors in patients with type 2 diabetes: meta-analysis of placebo-controlled randomized clinical trials. *Diabetes Metab*. 2017;43:48–58.

61. Fei Y, Tsoi M-F, Kumana CR, Cheung TT, Cheung BMY. Network meta-analysis of cardiovascular outcomes in randomized controlled trials of new antidiabetic drugs. *Int J Cardiol*. 2018;254:291–296.

62. Kosiborod MN, Abildstrøm SZ, Borlaug BA, et al. Semaglutide in patients with heart failure with preserved ejection fraction and obesity. *N Engl J Med*. 2023;389:1069–1084.

63. Khan MS, Fonarow GC, McGuire DK, et al. Glucagon-like peptide 1 receptor agonists and heart failure: the need for further evidence generation and practice guidelines optimization. *Circulation*. 2020;142:1205–1218.

64. Rubbo B, Fitzpatrick NK, Denaxas S, et al. Use of electronic health records to ascertain, validate and phenotype acute myocardial infarction: a systematic review and recommendations. *Int J Cardiol*. 2015;187:705–711.

65. Floyd JS, Blondon M, Moore KP, Boyko EJ, Smith NL. Validation of methods for assessing cardiovascular disease using electronic health data in a cohort of veterans with diabetes. *Pharmacoepidemiol Drug Saf*. 2016;25:467–471.

66. Singh S, Fouayzi H, Anzuoni K, et al. Diagnostic algorithms for cardiovascular death in administrative claims databases: a systematic review. *Drug Saf*. 2019;42:515–527.

67. Gliklich RE, Dreyer NA, Leavy MB, eds. *Registries for Evaluating Patient Outcomes: A User's Guide*. Rockville, MD: Agency for Healthcare Research and Quality; 2020.

68. Gliklich RE, Leavy MB, Dreyer NA. In: *Tools and Technologies for Registry Interoperability, Registries for Evaluating Patient Outcomes: A User's Guide*. 3rd Edition, Addendum 2. Rockville, MD: Agency for Healthcare Research and Quality; 2019.

KEY WORDS cardiovascular diseases, comparative effectiveness research, glucagon-like peptide-1 receptor agonists, hypoglycemic agents, sodium-glucose transporter 2 inhibitors, type 2 diabetes mellitus

APPENDIX For supplemental methods, tables, and figures, please see the online version of this paper.