



Using Real-World Data to Predict Findings of an Ongoing Phase IV Cardiovascular Outcome Trial: Cardiovascular Safety of Linagliptin Versus Glimepiride

Elisabetta Patorno,¹
 Sebastian Schneeweiss,¹
 Chandrasekar Gopalakrishnan,¹
 David Martin,² and Jessica M. Franklin¹

Diabetes Care 2019;42:2204–2210 | <https://doi.org/10.2337/dc19-0069>

OBJECTIVE

Using real-world data (RWD) from three U.S. claims data sets, we aim to predict the findings of the CARDiovascular Outcome Trial of LINAgliptin Versus Glimepiride in Type 2 Diabetes (CAROLINA) comparing linagliptin versus glimepiride in patients with type 2 diabetes (T2D) at increased cardiovascular risk by using a novel framework that requires passing prespecified validity checks before analyzing the primary outcome.

RESEARCH DESIGN AND METHODS

Within Medicare and two commercial claims data sets (May 2011–September 2015), we identified a 1:1 propensity score–matched (PSM) cohort of T2D patients 40–85 years old at increased cardiovascular risk who initiated linagliptin or glimepiride by adapting eligibility criteria from CAROLINA. PSM was used to balance >120 confounders. Validity checks included the evaluation of expected power, covariate balance, and two control outcomes for which we expected a positive association and a null finding. We registered the protocol (NCT03648424, ClinicalTrials.gov) before evaluating the composite cardiovascular outcome based on CAROLINA's primary end point. Hazard ratios (HR) and 95% CIs were estimated in each data source and pooled with a fixed-effects meta-analysis.

RESULTS

We identified 24,131 PSM pairs of linagliptin and glimepiride initiators with sufficient power for noninferiority (>98%). Exposure groups achieved excellent covariate balance, including key laboratory results, and expected associations between glimepiride and hypoglycemia (HR 2.38 [95% CI 1.79–3.13]) and between linagliptin and end-stage renal disease (HR 1.08 [0.66–1.79]) were replicated. Linagliptin was associated with a 9% decreased risk in the composite cardiovascular outcome with a CI including the null (HR 0.91 [0.79–1.05]), in line with noninferiority.

CONCLUSIONS

In a nonrandomized RWD study, we found that linagliptin has noninferior risk of a composite cardiovascular outcome compared with glimepiride.

¹Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

²Office of Medical Policy, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD

Corresponding author: Chandrasekar Gopalakrishnan, cgopalakrishnan@bwh.harvard.edu

Received 10 January 2019 and accepted 14 June 2019

This article contains Supplementary Data online at <http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc19-0069/-/DC1>.

© 2019 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <http://www.diabetesjournals.org/content/license>.

See accompanying article, p. 2161.

The 21st Century Cures Act mandates that the U.S. Food and Drug Administration (FDA) establish a program to evaluate the potential use of real-world evidence (RWE) to support a new indication for a drug or postapproval study requirements (1). In the framework for FDA's Real World Evidence Program, real-world data (RWD) includes longitudinal electronic health records, medical claims and billing data, and patient-generated data (2). Nonrandomized database studies are noninterventional clinical study designs in which the study identifies the population and determines the exposure/treatment from data generated before the initiation of the study (2). Health insurance claims data contain structured diagnosis and procedure information, capture patient experiences across the care continuum, and are frequently used by the FDA to evaluate drug safety (3).

Questions remain on whether RWD analyses of claims data can be reliably used for regulatory decision making to evaluate not only the safety of medications but also their effectiveness (4). Frequent criticisms include the lack of prespecified protocols, avoidable design and analytic flaws, data quality issues, and confounding biases that may threaten the validity of findings. Some investigators have designed such studies to match published randomized controlled trials (RCTs) and have successfully replicated (5) or predicted (6–8) trial findings, suggesting that principled methodology can generate accurate information based on RWD in certain cases. Several threats to validity can be assessed before registering the protocol and implementing a nonrandomized database study, so that the decision to move forward with a given study and the analytic plan are well documented. Such a process will instill greater confidence in the ability of nonrandomized database studies to reproduce findings of a comparable RCT.

The CARdiovascular Outcome Trial of LINagliptin Versus Glimepiride in Type 2 Diabetes (CAROLINA) study (9) is an ongoing RCT designed to assess whether linagliptin, a dipeptidyl peptidase 4 (DPP-4) inhibitor, is noninferior and, if so, superior compared with the sulfonylurea glimepiride with respect to cardiovascular events in adults with type 2 diabetes (T2D) at increased risk of cardiovascular

events. Given that medications of both classes are frequently used as second-line therapy after metformin, and because sulfonylureas have been associated with concerns regarding their cardiovascular safety (10), the results of this trial could have a significant impact on clinical practice if practitioners conclude there is a difference in cardiovascular safety between the drugs. Trial recruitment started in 2010, and results are expected in 2019.

Using RWD, we aimed to predict CAROLINA's findings within a framework that requires passing prespecified validity checks before analyzing the primary end point (11). This is part of a series of studies aimed to predict the findings of ongoing trials before their completion.

RESEARCH DESIGN AND METHODS

This study included data from two commercial U.S. health insurance claims data sets (Optum Clinformatics and IBM MarketScan) and fee-for-service Medicare claims data. For each insured individual, the three data sets contain demographic information, health plan enrollment status, longitudinal patient-level information on all reimbursed medical services, inpatient and outpatient diagnoses and procedures, and pharmacy dispensing records, including information on medication start and refill, strength, quantity, and days' supply. Optum and MarketScan are both linked to laboratory test results provided by two national laboratory test provider chains. Through this linkage, results for outpatient laboratory tests are available for a subset of beneficiaries.

Within the three U.S. RWD sources, we identified a cohort of T2D patients 40–85 years old at increased cardiovascular risk who initiated linagliptin or glimepiride from May 2011 (in accordance with the approval of linagliptin in the U.S.) to September 2015, adapting eligibility criteria from CAROLINA (Fig. 1 and Supplementary Table 1). Cohort entry date was the day of the first filled prescription of linagliptin or glimepiride among patients with at least 6 months of continuous enrollment before drug initiation. We used 1:1 propensity score matching (PSM) to control for >120 potential confounders, which were measured during the 6 months before cohort entry and included demographics, calendar time, comorbidities, diabetes-specific complications, use of diabetes

and other medications, and indicators of health care utilization as proxy for overall disease state, care intensity, and surveillance. Laboratory test results, which were available in a 7% subset of the population, were also measured at baseline, although they were not included in the claims-based PS model.

The primary outcome was a composite cardiovascular outcome of hospitalization for myocardial infarction, stroke, or death (Supplementary Table 2), adapted from the CAROLINA study's primary end point of three-point major adverse cardiovascular event (3-P MACE) composite, comprising nonfatal myocardial infarction, nonfatal stroke, or cardiovascular death. The availability of mortality information varied by database. Medicare fee-for-service included complete information on all-cause mortality, MarketScan included information on in-hospital death, and no information was available in Clinformatics. Individual components of the composite cardiovascular outcome were also analyzed as secondary outcomes.

Follow-up started on the day after cohort entry and continued in an "as-treated" approach until treatment discontinuation or switch to a comparator, occurrence of an event of interest, nursing home admission, plan disenrollment, or end of the study period, whichever came first. In case of treatment interruption or discontinuation, we extended the exposure effect window until 30 days after the end of the last prescription's supply. In line with CAROLINA, we allowed study participants who initiated linagliptin or glimepiride to be exposed to nonglimepiride sulfonylureas before cohort entry. Because in the trial patients in both arms could be exposed to nonglimepiride sulfonylureas, we censored patients who added a nonlinagliptin DPP-4 inhibitor but did not censor patients who added a nonglimepiride sulfonylurea.

We purposefully chose an as-treated analysis rather than an intention-to-treat analysis to address the high rate of treatment discontinuation in routine care. The as-treated analysis avoids the substantial exposure misclassification that often occurs when intention-to-treat analyses are applied in RWE studies, which typically bias findings toward the null. Hazard ratios (HR) and 95% CIs were estimated in the PSM

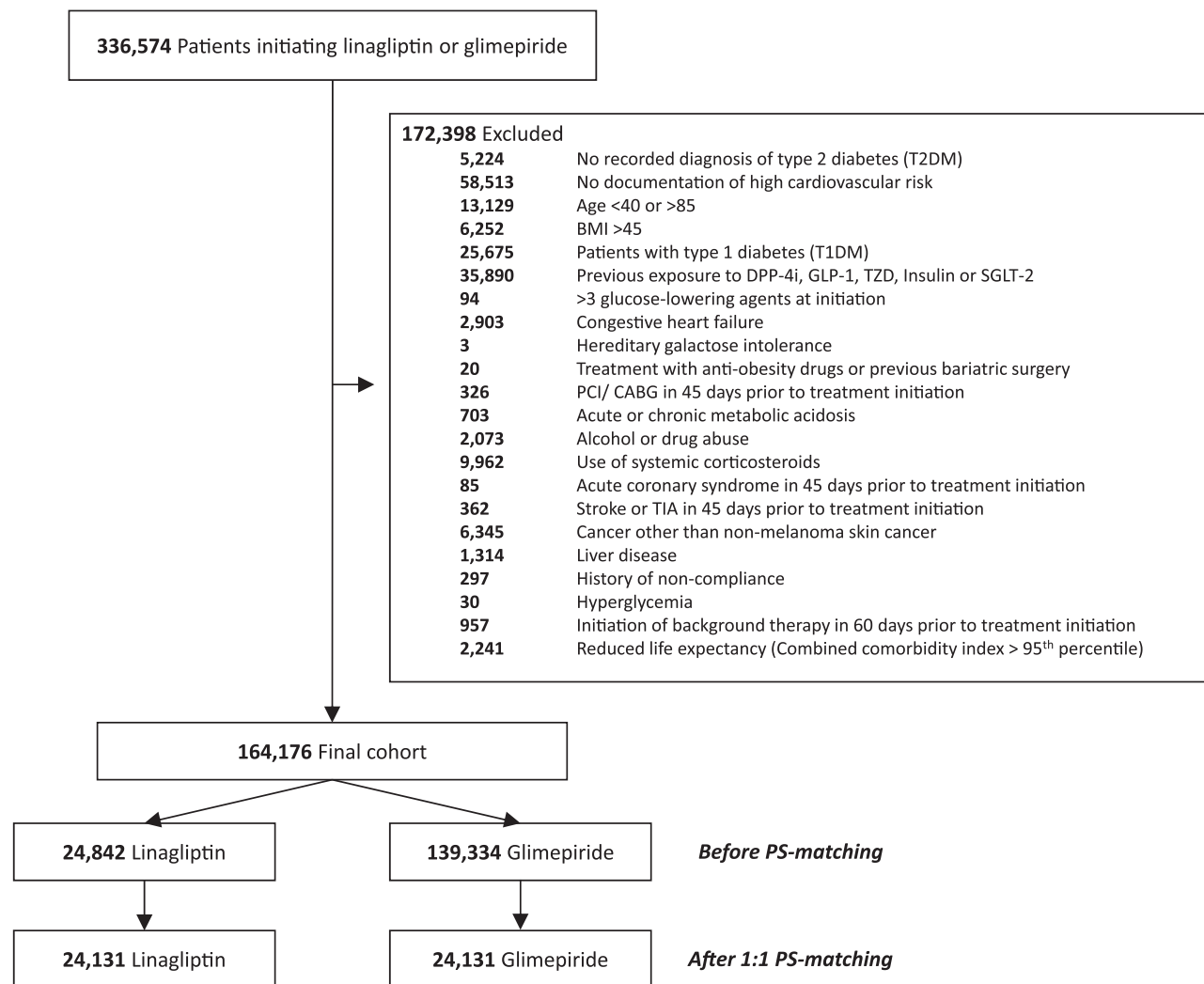


Figure 1—Cohort formation. DPP-4i, DPP-4 inhibitor; GLP-1, glucagon-like peptide 1; PCI/CABG, percutaneous coronary intervention/coronary artery bypass grafting; SGLT-2, sodium-glucose cotransporter 2; TIA, transient ischemic attack; TZD, thiazolidinediones.

cohort using unstratified Cox regression models. Analyses were conducted in each data source separately and then pooled across data sources using a fixed-effects meta-analysis.

Several prospective validity checks were conducted. First, we calculated the expected power for noninferiority of the primary end point at an α level of 0.05, to exclude an upper margin of the 95% CI for the hazard ratio of 1.3, as specified in CAROLINA (9) and mandated by the FDA for cardiovascular outcome trials evaluating new therapies for T2D (12).

Then, we assessed the postmatching balance of potential confounders between exposure groups by calculating standardized differences, with meaningful imbalances set at values >0.1 , and postmatching C-statistic, which is expected to be close to 0.5 when balance

is present (13). The potential for residual confounding by unmeasured factors not included in the PS model was evaluated by inspecting the balance in key baseline laboratory results in the population subset with this information available.

Finally, we evaluated two control outcomes with the aim to replicate known associations. Specifically, we assessed the risk of severe hypoglycemia, defined as an emergency department visit or a hospitalization for hypoglycemia, for which we expected an increased risk associated with the initiation of glimepiride (14), and the risk of incident end-stage renal disease (ESRD), for which we expected a null finding (15) (see Supplementary Table 2 for outcome definitions). ESRD was specifically chosen as a control outcome because of the known preferential prescribing of

linagliptin toward patients with chronic kidney disease (16), which may lead to an apparent elevation in ESRD risk associated with the use of linagliptin.

All analyses were performed using Aetion platform version 3.11 with R version 3.4.2, which has previously been scientifically validated by accurately repeating a range of previously published studies (17) and by replicating (18) or predicting clinical trial findings (6). After all validity checks were met and confidence for accurately predicting results from CAROLINA was achieved, the protocol was registered (NCT03648424, ClinicalTrials.gov), and the primary end point analysis conducted. All individual data were deidentified, the study was approved by the Brigham and Women's Hospital Institutional Review Board, and signed data license agreements were in place for all data sources.

RESULTS

The final eligible study cohort included 164,176 patients with T2D, of whom 24,842 initiated linagliptin and 139,334 initiated glimepiride. Before PSM, linagliptin initiators tended to be younger but had a greater burden of comorbidities such

as hypertension, hyperlipidemia, and chronic kidney disease (Table 1 and Supplementary Table 3). After PSM, we identified 24,131 patient pairs of linagliptin versus glimepiride initiators, and exposure groups achieved excellent covariate balance, with standardized

differences for all covariates of <0.1 (Table 1 and Supplementary Table 4). Compared with the CAROLINA participants, the patients included in our study population were older (mean age 70 vs. 64 years) and more frequently women; however, they had a similar burden of

Table 1—Baseline characteristics of study participants before and after PSM

	Before PSM			After PSM		
	Linagliptin <i>n</i> = 24,842	Glimepiride <i>n</i> = 139,334	St. Diff.	Linagliptin <i>n</i> = 24,131	Glimepiride <i>n</i> = 24,131	St. Diff.
Baseline characteristics						
Age, mean (SD), years	70.32 (7.76)	71.18 (7.72)	−0.11	70.40 (7.76)	70.42 (7.71)	0.00
Male, <i>n</i> (%)	11,872 (47.8)	68,926 (49.5)	−0.03	11,519 (47.7)	11,512 (47.7)	0.00
White race, <i>n</i> (%) ¹	11,554 (71.0)	71,226 (79.8)	−0.21	11,396 (71.8)	11,485 (72.3)	−0.01
Burden of comorbidities						
Combined comorbidity score, mean (SD) ²	1.49 (1.27)	1.21 (1.21)	0.23	1.48 (1.26)	1.47 (1.26)	0.01
Frailty score, mean (SD) ³	0.16 (0.05)	0.16 (0.05)	0.00	0.16 (0.05)	0.16 (0.05)	0.00
Diabetes-related complications, <i>n</i> (%)						
Diabetic nephropathy	1,771 (7.1)	6,350 (4.6)	0.11	1,655 (6.9)	1,646 (6.8)	0.00
Diabetic retinopathy	1,172 (4.7)	5,754 (4.1)	0.03	1,132 (4.7)	1,103 (4.6)	0.00
Diabetes with ophthalmic conditions or procedures	656 (2.6)	3,363 (2.4)	0.01	632 (2.6)	625 (2.6)	0.00
Diabetic neuropathy	2,765 (11.1)	14,093 (10.1)	0.03	2,663 (11.0)	2,656 (11.0)	0.00
Diabetes with peripheral circulatory disorders	1,099 (4.4)	5,309 (3.8)	0.03	1,050 (4.4)	1,039 (4.3)	0.00
Hypoglycemia	708 (2.9)	3,212 (2.3)	0.04	684 (2.8)	695 (2.9)	−0.01
Hyperglycemia	997 (4.0)	3,856 (2.8)	0.07	961 (4.0)	968 (4.0)	0.00
Disorders of fluid electrolyte and acid-base balance	1,114 (4.5)	5,333 (3.8)	0.04	1,083 (4.5)	1,097 (4.5)	0.00
Baseline diabetes therapy						
No. of antidiabetic drugs at cohort entry, mean (SD)	0.81 (0.70)	0.63 (0.59)	0.28	0.80 (0.70)	0.79 (0.69)	0.01
Any use of metformin, <i>n</i> (%)	16,477 (66.3)	88,730 (63.7)	0.05	16,009 (66.3)	16,136 (66.9)	−0.01
Any use of sulfonylureas, <i>n</i> (%)	8,043 (32.4)	21,751 (15.6)	0.40	7,458 (30.9)	7,341 (30.4)	0.01
Any use of meglitinides, <i>n</i> (%)	358 (1.4)	774 (0.6)	0.08	341 (1.4)	293 (1.2)	0.02
Any use of α -glucosidase inhibitors, <i>n</i> (%)	84 (0.3)	240 (0.2)	0.02	74 (0.3)	92 (0.4)	−0.02
Other comorbidities at baseline, <i>n</i> (%)						
Ischemic heart disease	6,194 (24.9)	33,882 (24.3)	0.01	6,016 (24.9)	6,042 (25.0)	0.00
Previous coronary revascularization	1,109 (4.5)	6,428 (4.6)	0.00	1,088 (4.5)	1,096 (4.5)	0.00
Ischemic or hemorrhagic stroke	1,849 (7.4)	9,475 (6.8)	0.02	1,799 (7.5)	1,816 (7.5)	0.00
Heart failure	915 (3.7)	5,144 (3.7)	0.00	896 (3.7)	865 (3.6)	0.01
Peripheral arterial disease or surgery	2,101 (8.5)	9,331 (6.7)	0.07	2,028 (8.4)	2,048 (8.5)	0.00
Hypertension	21,592 (86.9)	114,186 (82.0)	0.14	20,940 (86.8)	20,985 (87.0)	−0.01
Hyperlipidemia	19,154 (77.1)	100,099 (71.8)	0.12	18,581 (77.0)	18,583 (77.0)	0.00
Nondiabetic renal dysfunction	4,668 (18.8)	14,874 (10.7)	0.23	4,385 (18.2)	4,164 (17.3)	0.02
Other medications, <i>n</i> (%)						
ACE inhibitors	10,911 (43.9)	66,820 (48.0)	−0.08	10,670 (44.2)	10,603 (43.9)	0.01
Angiotensin II receptor blockers	8,526 (34.3)	36,352 (26.1)	0.18	8,143 (33.7)	8,284 (34.3)	−0.01
β -Blockers	10,171 (40.9)	55,175 (39.6)	0.03	9,860 (40.9)	10,124 (42.0)	−0.02
Loop diuretics	3,276 (13.2)	16,993 (12.2)	0.03	3,188 (13.2)	3,111 (12.9)	0.01
Statins	18,362 (73.9)	96,296 (69.1)	0.11	17,782 (73.7)	17,772 (73.6)	0.00
Other lipid-lowering drugs	4,110 (16.5)	18,095 (13.0)	0.10	3,959 (16.4)	3,942 (16.3)	0.00
Antiplatelet agents	3,902 (15.7)	18,521 (13.3)	0.07	3,761 (15.6)	3,779 (15.7)	0.00
Oral anticoagulants	1,645 (6.6)	9,700 (7.0)	−0.02	1,615 (6.7)	1,621 (6.7)	0.00
Measures of health care utilization						
Hospitalization within prior 30 days, <i>n</i> (%)	260 (1.0)	1,774 (1.3)	−0.03	255 (1.1)	284 (1.2)	−0.01
Hospitalization during prior 31–180 days, <i>n</i> (%)	1,221 (4.9)	7,346 (5.3)	−0.02	1,174 (4.9)	1,268 (5.3)	−0.02
No. of hospital days, mean (SD)	0.23 (1.64)	0.25 (1.65)	−0.01	0.24 (1.66)	0.25 (1.73)	−0.01
No. of emergency department visits, mean (SD)	0.25 (0.88)	0.25 (0.88)	0.00	0.24 (0.88)	0.25 (0.84)	−0.01
No. of office visits, mean (SD)	4.83 (3.62)	4.09 (3.34)	0.21	4.77 (3.56)	4.81 (3.78)	−0.01
No. of distinct medication prescriptions, mean (SD)	9.06 (3.98)	8.19 (3.68)	0.23	9.01 (3.96)	9.02 (3.89)	0.00

St. Diff., standardized differences (i.e., the difference in means or proportions divided by the pooled SD). ¹Only available in Medicare fee-for-service database. ²See Gagne et al. (26). ³See Kim et al. (27).

comorbidities, such as prior cardiovascular events and renal dysfunction, and patterns of medication use (Table 1). The estimated power exceeded 98% for noninferiority, and the postmatching C-statistic was 0.53 (Table 2). Key baseline laboratory results (HbA_{1c}, lipid levels, estimated glomerular filtration rate, and urinary albumin-to-creatinine ratio [UACR]), which were available in a subset of the population and therefore were not included in the PS adjustment, were equally well balanced (Table 2 and Supplementary Table 4).

The mean (SD) and median (interquartile range) follow-up were 222 (221) and 131 (62, 294) days, respectively. The known association between glimepiride and hypoglycemia (HR 2.38 [95% CI 1.79–3.13]) and the null association between linagliptin and ESRD (HR 1.08 [0.66–1.79]) were correctly estimated, confirming the study's ability to replicate known findings for control outcomes (Table 2 and Supplementary Table 5). Linagliptin was associated with a nonsignificant decrease in the risk of the primary cardiovascular outcome (HR 0.91 [0.79–1.05]) compared

with glimepiride (Table 3), in line with the noninferiority hypothesis of the CAROLINA trial. Within the individual databases, the cardiovascular effect varied from HR 0.96 (95% CI 0.83–1.12) in Medicare (mean age 73 years) to HR 0.76 (95% CI 0.47–1.22) in MarketScan (mean age 66 years) and HR 0.44 (95% CI 0.23–0.87) in Clinformatics (mean age 63 years) (Supplementary Table 6). Analysis of the individual components of the composite cardiovascular outcome produced results consistent with a nonsignificant decreased risk for myocardial infarction (HR 0.87 [95% CI 0.68–1.12]) and

Table 2—Statistical power, patient characteristics, select diagnostics, and control outcomes

Noninferiority power analysis in health care database study

No. patients after 1:1 PSM	48,262
Risk of primary end point per 1,000 patients	20.8
Relative risk assumed by RCT based on equipoise	1
α (two-sided)	0.05
Noninferiority margin	1.3
Statistical power, %	98.6

Selected patient characteristics from CAROLINA trial and from 1:1 PSM health care database study and balance statistics

		Health care database study (N = 48,262)		
	CAROLINA trial (N = 6,041)	Linagliptin (n = 24,131)	Glimepiride (n = 24,131)	aSt. Diff.*
Patient characteristics				
Age, mean (SD), years	64.0 (9.5)	70.40 (7.76)	70.42 (7.71)	0.00
Male sex, n (%)	3,622 (60.0)	11,519 (47.7)	11,512 (47.7)	0.00
White race, n (%) ¹	4,408 (73.0)	11,396 (71.8)	11,485 (72.3)	−0.01
Previous cardiovascular event, n (%)	2,084 (34.5)	8,468 (35.1)	8,537 (35.4)	0.00
Renal dysfunction, n (%)	1,099 (18.2)	4,385 (18.2)	4,164 (17.3)	0.00
Metformin use, n (%)	4,982 (82.5)	16,009 (66.3)	16,136 (66.9)	0.00
Statin use, n (%)	3,866 (64.1)	17,782 (73.7)	17,772 (73.6)	0.00
ACE inhibitors/ARBs, n (%)	4,543 (75.2)	18,381 (76.2)	18,470 (76.5)	0.00
β-Blockers, n (%)	2,296 (38.8)	9,860 (40.9)	10,124 (42.0)	0.00
Baseline glucose-lowering therapy, n (%)				
None	557 (9.2)	8,736 (36.2)	8,836 (36.6)	0.00
1 glucose-lowering agent	3,988 (66.0)	11,574 (48.0)	11,437 (47.4)	0.00
≥2 glucose-lowering agents	1,496 (24.8)	3,821 (15.8)	3,858 (16.0)	0.00
HbA _{1c} , mean (SD), % ²	7.2 (0.6)	8.11 (1.61)	8.15 (1.85)	0.02
HbA _{1c} , mean (SD), mmol/mol ²	55 (6.6)	65 (17.6)	66 (20.2)	0.02
Total cholesterol, mean (SD), mg/dL ²	177 (44)	175.21 (45.85)	176.79 (48.86)	0.03
LDL-cholesterol, mean (SD), mg/dL ²	95 (36)	88.63 (40.20)	87.49 (42.68)	0.03
eGFR, mean (SD), mL/min/1.73 m ²²	77 (20)	81.89 (23.92)	83.91 (23.00)	0.09
UACR >30–300 μg/mg, n (%) ²	1,281 (21.2)	120 (24.9)	123 (27.9)	0.00
UACR >300 μg/mg, n (%) ²	260 (4.3)	48 (10.0)	26 (5.9)	0.03
Post-PSM C-statistic	N.A.	0.53**		
Follow-up, mean (SD) and median (IQR), days	222 (221); 131 (62, 294)			
Control outcomes				
	N events (IR/1,000 PY)	N events (IR/1,000 PY)	HR (95% CI)	
Control end point 1—Severe hypoglycemia (expected positive association with glimepiride)	60 (4.3)	160 (10.4)	2.38 (1.79–3.13) Ref. = linagliptin	
Control end point 2—ESRD (expected null association with linagliptin)	30 (2.1)	32 (2.1)	1.08 (0.66–1.79) Ref. = glimepiride	

ARB, angiotensin receptor blocker; aSt. Diff., absolute standardized difference; eGFR, estimated glomerular filtration rate; IQR, interquartile range; IR, incidence rate; N.A., not applicable; PY, person-years; Ref., reference. ¹Only available in Medicare. ²Only available in commercial data sets for a 7% subset. *Values <0.1 are considered unlikely to cause any meaningful confounding. **A postmatching C-statistic of 0.5 indicates perfect balance across all covariates.

stroke (HR 0.84 [0.64–1.11]) associated with the use of linagliptin and with a null association with all-cause mortality (HR 0.96 [0.79–1.17]), although the availability of mortality information varied by data source (Supplementary Table 5). Results were largely consistent across individual databases and in a sensitivity analysis based on random effects pooling (HR 0.76 [0.51–1.13]).

CONCLUSIONS

In this cohort study using electronic claims data, we found evidence of adequate statistical power, solid confounding control, and the ability to replicate known associations for two control outcomes, which suggests high confidence for accurately predicting results from the CAROLINA trial before the release of its findings. Linagliptin was noninferior to glimepiride and was associated with a nonsignificant 9% decrease in the risk of the primary cardiovascular end point.

The results are consistent with a prior small trial (19) and with noninterventional studies (20,21) that showed a decreased risk of cardiovascular events associated with DPP-4 inhibitors compared with sulfonylureas in younger patients with lower cardiovascular risk, but not in patients with higher risk (20). Because participants in the CAROLINA trial were on average younger compared with our pooled study population (64 vs. 70 years), it is possible that an age-dependent cardiovascular benefit of linagliptin may be slightly larger in the trial. This article was submitted to this journal on 10 January 2019. On 14 February 2019, Boehringer Ingelheim and Eli Lilly announced CAROLINA met its primary end point for 3P-MACE, defined as noninferiority for linagliptin versus glimepiride in adults with T2D with cardiovascular risk (22). Although this press

release points toward an alignment with our findings, full results of CAROLINA, expected to be presented in June at the American Diabetes Association's 79th Scientific Sessions, will reveal whether and to what extent our RWD analysis succeeded in reproducing the CAROLINA findings.

Although many questions on the effectiveness of medications can only be reliably assessed in the setting of an RCT, some questions could potentially be reliably answered using RWE generated from nonrandomized database studies, even in the absence of traditional RCT evidence (23). Developing a process for prospectively identifying such settings and generating RWE that instills high confidence in its validity is essential if there will be a role for nonrandomized studies using RWD in regulatory decision making (2,11).

This study has limitations. First, residual confounding by some unmeasured characteristic(s) cannot be entirely ruled out, although it is likely to be minor. Prespecified validity checks confirmed that the new user–active comparator design combined with the information-rich propensity-score adjustment by >120 variables reduced the potential for confounding by unmeasured covariates substantially. This is shown by the balance in selected laboratory test results, despite them not having been included in the PS model (because they were not recorded in most patients). This multipronged approach to confounding minimization has previously shown success in achieving balance of unmeasured characteristics when studying oral antidiabetic medications, including socioeconomic status (24,25). Prespecified validity checks also confirmed the study's ability to replicate known causal associations for two control outcomes.

Second, information on mortality varied by database, with complete information only in the Medicare fee-for-service database.

Third, because our study reflected the use of linagliptin or glimepiride in routine care, the median follow-up was shorter compared with most cardiovascular outcome trials, which have substantial adherence-improvement measures built-in. Trials generally require long follow-up to accumulate sufficient events for powered analyses. The size of our study population (~50,000 patients) allowed us to achieve powered analyses even with a shorter duration of follow-up. Assuming no time-varying hazards, our shorter-term findings will be generalizable to longer-term trial findings.

Fourth, heterogeneity was observed in the point estimates across the three databases. This is expected, because the databases include different populations with different baseline risks for the primary outcome, similar to the heterogeneity that is observed in the context of RCTs reporting stratified analyses by baseline risk. The small number of events observed in the commercial databases, particularly in Optum, and the resulting imprecise point estimates also contributed to the observed heterogeneity across databases.

Fifth, this is only the first of a series of studies aimed to predict the findings of ongoing trials before their completion, and thus it does not intend to provide conclusive evidence regarding the capacity of RWD analyses to succeed or not in replicating ongoing RCTs.

In conclusion, this large cohort study specifically designed to predict the findings of the ongoing CAROLINA trial before its completion found that linagliptin was noninferior to glimepiride regarding combined cardiovascular events. The

Table 3—Primary and secondary outcomes

	Linagliptin, <i>n</i> = 24,131 <i>N</i> events (IR/1,000 PY)	Glimepiride, <i>n</i> = 24,131 <i>N</i> events (IR/1,000 PY)	HR (95% CI) Ref. = glimepiride
Primary composite cardiovascular outcome	373 (26.7)	458 (29.7)	0.91 (0.79–1.05)
Secondary outcomes:			
Myocardial infarction	111 (7.9)	142 (9.2)	0.87 (0.68–1.12)
Stroke	90 (6.4)	120 (7.8)	0.84 (0.64–1.11)
All-cause mortality*	186 (15.5)	219 (16.6)	0.96 (0.79–1.17)
All-cause mortality in Medicare fee-for-service	180 (19.5)	211 (20.3)	0.97 (0.79–1.18)

IR, incidence rate; PY, patient-years; Ref., reference. *The availability of mortality information varied by database. Medicare fee-for-service included complete information on all-cause mortality, MarketScan only included information on in-hospital death, and no information on mortality was available in Clinformatics.

observed treatment effect, a 9% decrease in the primary cardiovascular end point associated with the initiation of linagliptin, was compatible with a null finding in a population slightly older than CAROLINA. Although a press release from CAROLINA pointed toward an alignment with our study findings, full results from the trial, expected in a few months, will reveal whether and to what extent our RWD analysis succeeded in predicting the CAROLINA findings.

Acknowledgments. The authors thank Dr. Dorothee B. Bartels (BI X GmbH, Ingelheim, Germany) for her involvement in an earlier phase of this study.

Funding. This study was funded by U.S. Food and Drug Administration contract numbers HHSF223201710186C and HHSF223201810146C through the Department of Health and Human Services (HHS) and by the Division of Pharmacoeconomics and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA. E.P. was supported by a career development grant (K08-AG-055670) from the National Institute on Aging.

Duality of Interest. E.P. is co-investigator of investigator-initiated grants to the Brigham and Women's Hospital from GlaxoSmithKline and Boehringer Ingelheim, not directly related to the topic of the submitted work. S.S. is the principal investigator of investigator-initiated grants to the Brigham and Women's Hospital from Bayer, Vertex, and Boehringer Ingelheim unrelated to the topic of this study and is a consultant to WHISCON and to Aetion, a software manufacturer of which he owns equity. His interests were declared, reviewed, and approved by the Brigham and Women's Hospital and Partners HealthCare System in accordance with their institutional compliance policies. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. E.P., S.S., and J.M.F. developed the study protocol. E.P. and C.G. wrote the manuscript and conducted the data analysis. E.P., S.S., C.G., D.M., and J.M.F. contributed to discussion and reviewed and edited the manuscript. E.P. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

- Bonamici S. H.R. 34 - 114th Congress (2015–2016): 21st Century Cures Act [Internet]. 2016. Available from <https://www.congress.gov/bill/114th-congress/house-bill/34>. Accessed November 27 2018
- U.S. Food and Drug Administration. Framework for FDA's Real-World Evidence program [Internet]. 2018. Available from <https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/RealWorldEvidence/UCM627769.pdf>. Accessed December 11 2018
- Platt R, Brown JS, Robb M, et al. The FDA sentinel initiative - an evolving national resource. *N Engl J Med* 2018;379:2091–2093
- Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence—what is it and what can it tell us? *N Engl J Med* 2016;375:2293–2297
- Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008;19:766–779
- Kim SC, Solomon DH, Rogers JR, et al. Cardiovascular safety of tocilizumab versus tumor necrosis factor inhibitors in patients with rheumatoid arthritis: a multi-database cohort study. *Arthritis Rheumatol* 2017;69:1154–1164
- Patrono E, Goldfine AB, Schneeweiss S, et al. Cardiovascular outcomes associated with canagliflozin versus other non-glioflozin antidiabetic drugs: population based cohort study. *BMJ* 2018;360:k119
- Patrono E, Goldfine AB, Schneeweiss S, Glynn R, Liu J, Kim S. Cardiovascular safety of canagliflozin vs. other antidiabetic agents in routine care. *Diabetes* 2017;66(Suppl. 1):1497-P
- Marx N, Rosenstock J, Kahn SE, et al. Design and baseline characteristics of the CARDiovascular Outcome Trial of LINAgliptin Versus Glimpiride in Type 2 Diabetes (CAROLINA®). *Diab Vasc Dis Res* 2015;12:164–174
- 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
- Franklin JM, Glynn RJ, Martin D, Schneeweiss S. Evaluating the use of non-randomized real-world data analyses for regulatory decision making. *Clin Pharmacol Ther* 2019;105:867–877
- Food and Drug Administration (Center for Drug Evaluation and Research). Guidance for Industry: Diabetes Mellitus—Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes [Internet]. 2008. Available from <https://www.fda.gov/downloads/Drugs/Guidances/ucm071627.pdf>. Accessed April 26 2019
- Franklin JM, Rassen JA, Ackermann D, Bartels DB, Schneeweiss S. Metrics for covariate balance in cohort studies of causal effects. *Stat Med* 2014;33:1685–1699
- American Diabetes Association. 8. Pharmacologic approaches to glycemic treatment: *Standards of Medical Care in Diabetes—2018*. *Diabetes Care* 2018;41(Suppl. 1):S73–S85
- Rosenstock J, Perkovic V, Johansen OE, et al.; CARMELINA Investigators. Effect of linagliptin vs placebo on major cardiovascular events in adults with type 2 diabetes and high cardiovascular and renal risk: the CARMELINA randomized clinical trial. *JAMA* 2019;321:69–79
- Patrono E, Gopalakrishnan C, Bartels DB, Brodovitz KG, Liu J, Schneeweiss S. Preferential prescribing and utilization trends of diabetes medications among patients with renal impairment: emerging role of linagliptin and other dipeptidyl peptidase 4 inhibitors. *Endocrinol Diabetes Metab* 2017;1:e00005
- Wang SV, Verpillat P, Rassen JA, Patrick A, Garry EM, Bartels DB. Transparency and reproducibility of observational cohort studies using large healthcare databases. *Clin Pharmacol Ther* 2016;99:325–332
- Fralick M, Kesselheim AS, Avorn J, Schneeweiss S. Use of health care databases to support supplemental indications of approved medications. *JAMA Intern Med* 2018;178:55–63
- Gallwitz B, Rosenstock J, Rauch T, et al. 2-year efficacy and safety of linagliptin compared with glimepiride in patients with type 2 diabetes inadequately controlled on metformin: a randomised, double-blind, non-inferiority trial. *Lancet* 2012;380:475–483
- Fu AZ, Johnston SS, Ghannam A, et al. Association between hospitalization for heart failure and dipeptidyl peptidase 4 inhibitors in patients with type 2 diabetes: an observational study. *Diabetes Care* 2016;39:726–734
- Patrono E, Gopalakrishnan C, Brodovitz KG, et al. Cardiovascular safety of linagliptin compared with other oral glucose-lowering agents in patients with type 2 diabetes: a sequential monitoring programme in routine care. *Diabetes Obes Metab* 2019;21:1824–1836
- Boehringer Ingelheim. Boehringer Ingelheim and Lilly announce the CAROLINA® cardiovascular outcome trial of Tadjenta® met its primary endpoint of non-inferiority compared with glimepiride [Internet]. 2019. Available from <https://www.boehringer-ingelheim.us/press-release/boehringer-ingelheim-and-lilly-announce-carolina-cardiovascular-outcome-trial>. Accessed March 10 2019
- Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? *Clin Pharmacol Ther* 2017;102:924–933
- Patrono E, Gopalakrishnan C, Franklin JM, et al. Claims-based studies of oral glucose-lowering medications can achieve balance in critical clinical variables only observed in electronic health records. *Diabetes Obes Metab* 2018;20:974–984
- Gopalakrishnan C, Gagne JJ, Sarpatwari A, et al. Evaluation of socioeconomic status indicators for confounding adjustment in observational studies of medication use. *Clin Pharmacol Ther* 2019;105:1513–1521
- Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol* 2011;64:749–759
- Kim DH, Schneeweiss S, Glynn RJ. Comparing approaches to measure frailty in Medicare data: deficit-accumulation frailty index versus phenotypic frailty. *J Gerontol A Biol Sci Med Sci* 2018;73:989–990