



## Use of real-world evidence data to evaluate the comparative effectiveness of second-line type 2 diabetes medications on chronic kidney disease

Yu Deng<sup>a,\*</sup>, Farhad Ghamsari<sup>b</sup>, Alice Lu<sup>a</sup>, Jingzhi Yu<sup>a</sup>, Lihui Zhao<sup>c</sup>, Abel N. Kho<sup>a</sup>

<sup>a</sup> Center for Health Information Partnerships, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

<sup>b</sup> Dept of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

<sup>c</sup> Dept. of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

### ARTICLE INFO

#### Keywords:

Type 2 diabetes  
Type 2 diabetes treatment  
Chronic kidney disease  
Electronic health records  
Second-line anti-diabetic medication  
Real-world data  
Sodium-glucose cotransporter 2  
Dipeptidyl peptidase 4  
Glucagon-like peptide 1 receptor agonists  
Thiazolidinediones  
Sulfonylureas  
Insulin

### ABSTRACT

Chronic kidney disease (CKD) is a common complication of type 2 diabetes mellitus (T2DM). Approximately one-third of patients with T2DM also have CKD. In clinical trial studies, several anti-diabetic medications (ADM) show evidence of preventing the progression of CKD. Biguanides (e.g., metformin) are widely accepted as the first line medication. However, the comparative effectiveness of second line ADMs on CKD outcomes in T2DM is unclear. In addition, results from clinical trials may not generalize into routine clinical practice. In this study, we aimed to investigate the association of second line ADMs with diagnosed incident CKD, CKD hospitalization, and eGFR < 45 mL/min in T2DM patients using real-world data from electronic health records. Our study found that treatment with sodium-glucose cotransporter 2 (SGLT-2) inhibitors was significantly associated with lower risk of diagnosed CKD incidence in both primary analysis (hazard ratio, 0.43; 95 % CI, [0.22;0.87]; p-value, 0.02) and secondary analysis (hazard ratio, 0.42; 95 % CI, [0.19;0.92]; p-value, 0.03) compared to use of Sulfonylureas (SU) as a second-line ADM. However, significant associations were not observed when using eGFR < 45 mL/min as the endpoint. Treatment with a dipeptidyl peptidase 4 (DPP-4) inhibitor was significantly associated with lower risk of diagnosed incident CKD (hazard ratio, 0.7; 95 % CI, [0.53;0.96]; p-value, 0.03) and lower risk of CKD hospitalization (hazard ratio, 0.6; 95 % CI, [0.37; 0.96]; p-value, 0.04) in the primary analysis. However, both associations were not significant in the sensitivity analysis. We did not observe significant association between use of glucagon-like peptide 1 receptor agonists (GLP-1RA), Thiazolidinediones (TZD), insulin and diagnosed CKD incidence, hospitalization or eGFR < 45 mL/min compared to use of SU as a second-line ADM.

### Introduction

Type 2 diabetes is a major risk factor for chronic kidney disease (CKD) and is the leading cause of end stage renal disease (ESRD) with approximately 20 % – 40 % of patients developing diabetic nephropathy [1]. Sulfonylureas (SU), dipeptidyl peptidase 4 (DPP-4) inhibitor, glucagon-like peptide 1 receptor agonists (GLP-1RA), sodium-glucose cotransporter 2 inhibitor (SGLT-2 inhibitor), Thiazolidinediones (TZD), and Insulin are commonly used as second-line medication in addition to metformin. In the past decade, several randomized clinical trials have shown evidence of the newer anti-diabetic medications (ADM) such as SGLT-2 and GLP-1RA in reducing risk for renal disease outcome compared to placebo [2–6]. However, there are few clinical trials (ongoing or finished) that directly compare the effectiveness of the

newer ADMs to the older ones such as SU and DPP-4 inhibitor. Furthermore, patients in clinical trial studies are more compliant with therapy for a number of reasons (e.g. support from study staff), therefore, it is unclear how well results from clinical trials may apply to the general population in real clinical practice [7]. Electronic health records (EHRs), a major source of real-world evidence data, can facilitate the understanding of treatment effectiveness in clinical practice using patient level data from the routine operation of the healthcare system and complement evidence on the efficacy of medications from randomized controlled trials (RCTs) [8]. Recently, a study using real-world data investigated the comparative effectiveness of SGLT-2 inhibitor, GLP-1RA, DPP-4 inhibitor, and SU in type 2 diabetics for preventing renal disease [9]. This study was conducted within the Department of Veterans Affairs, which serves an older (mean age = 65.46), white (71.95 %),

\* Corresponding author at: Center for health information Partnerships (CHIP), Northwestern University Feinberg School of Medicine, 625 N. Michigan Ave, Suite 1500, Chicago, IL 60611, USA.

E-mail address: [yudeng2015@u.northwestern.edu](mailto:yudeng2015@u.northwestern.edu) (Y. Deng).

<https://doi.org/10.1016/j.jcte.2022.100309>

Received 30 September 2021; Received in revised form 30 May 2022; Accepted 4 October 2022

Available online 10 October 2022

2214-6237/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

male (94.49 %) population. Therefore, additional research is required to compare the effectiveness on renal disease across the multitude of currently available drugs and drug classes. In this study, we compared the effect of commonly used second line anti-diabetic medications including SU, DPP-4 inhibitor, GLP-1RA, SGLT-2 inhibitor, TZD, and Insulin on renal outcomes using EHRs from a large integrated health delivery system.

## Methods

### Study population

The Northwestern Medicine Electronic Data Warehouse (NMEDW) is the primary data repository for all the medical records of patients who receive care within the Northwestern Medicine system [10]. Established in 2007, the NMEDW contains records for over 3.8 million patients, with most EHR data going back to at least 2002, and with some billing claims data going back to 1998 or even earlier [10]. We included patients who met the following criteria: 1) at least one prescription to an ADM 2) at least one diagnosis code for type 2 diabetes (see Supplemental Table 1 for type 2 diabetes diagnosis codes) [11] 3) no excluded diagnoses: pregnancy, type 1 diabetes mellitus 4) at least one year of records in the database before their first ADM exposure and 5) at least three years of continuous ADM prescription records after first ADM exposure. Given that the first SGLT-2 inhibitor was approved by FDA for use in the United States in March 2013, we only included patients who had their first ADM exposure after 2013–03-01. Fig. 1 shows the flow chart of the patient selection process.

We defined an ADM sequence as the chronological series of ADM prescription orders a patient received from the date of their first ADM prescription to one year later (e.g., “Metformin, Glipizide, Glipizide, Glucophage, Sitagliptin, Glipizide”). We then mapped each generic name and brand name to its respective drug class in Anatomical Therapeutic Chemical Classification (e.g., “Biguanides, SU, SU, Biguanides, DPP-4 inhibitor, SU”) [12]. Repeated listings of the same medication in each sequence were understood to be refilled prescription. Therefore, in this example, only the first occurrence of a medication is kept in the final

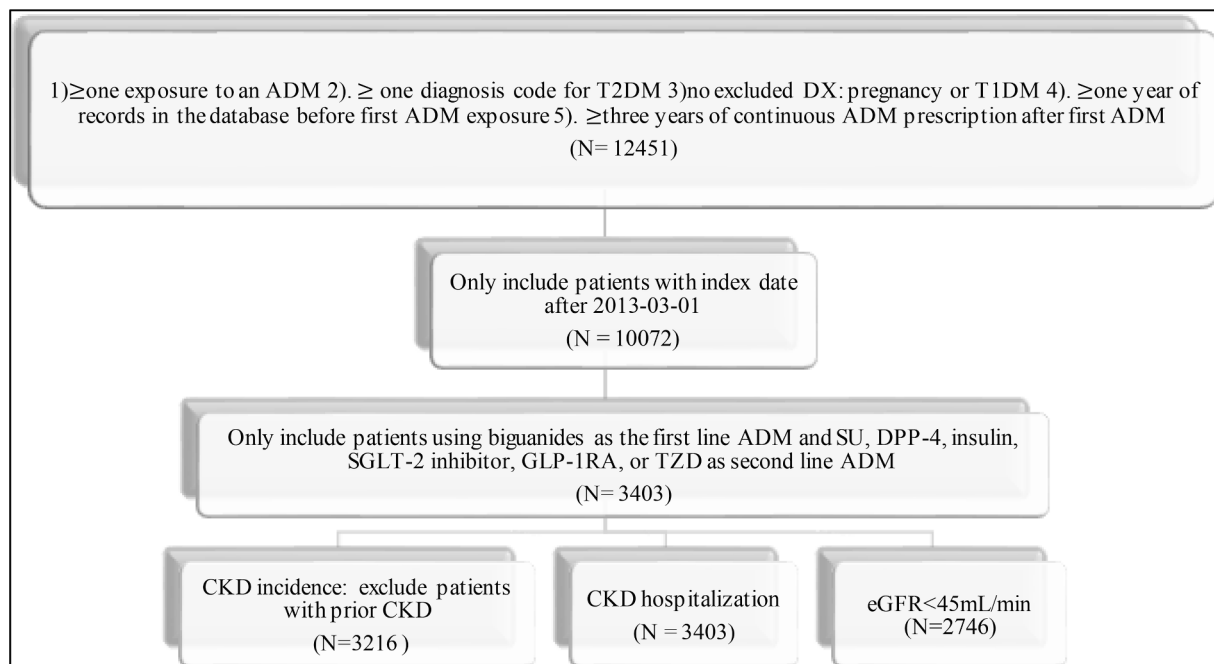
sequence (e.g., “Biguanides, SU, DPP-4 inhibitor”). During the 1-year period, only patients who were treated with biguanides as their first line of medication and had a second medication of either DPP-4 inhibitor, SGLT-2 inhibitor, GLP-1RA, SU, TZD, or Insulin were included in the study.

### Exposure

The exposure in this study was the medication sequence during a patient one-year exposure period as described above. The sequences we were interested in include ‘Biguanides, SU’, ‘Biguanides, DPP-4 inhibitor’, ‘Biguanides, Insulin’, ‘Biguanides, GLP-1RA’, ‘Biguanides, SGLT-2 inhibitor’ and ‘Biguanides, TZD’. A patient might switch to a third medication during the one-year exposure time, we considered them in the same group as the patients who did not. For example, ‘Biguanides, SU, DPP-4’ was considered in the same second-line ADM group as ‘Biguanides, SU’. Patients who were only on Biguanides during the 1-year drug exposure period were excluded from the study. Index date was defined as one year after first the exposure to ADM.

### Renal outcomes

Our primary outcome is diagnosed CKD incidence identified by the first appearance of an associated International Classification of Diseases, 9th/10th revision (ICD9/10) diagnosis code (see Supplemental Table 1) [7]. Patients who had CKD, ESRD or macroalbuminuria before index date were excluded in the analysis for this outcome. In addition to the diagnosed incident CKD event, we evaluated ADMs’ associations with CKD hospitalization. CKD hospitalization were identified based on ICD9/10 codes. Because diagnosis codes may be entered by clinicians in the EHR only for more severe cases of CKD resulting in a significant delay between onset of CKD and an observable diagnosis or admission, we also performed analysis using laboratory value  $eGFR < 45$  mL/min as an additional outcome. We used  $eGFR < 45$  mL/min as an indicator of patients having moderate or more severe CKD and more specific than using a higher threshold of  $< 60$  mL/min. For this outcome, patients who had  $eGFR < 45$  mL/min, CKD diagnosis, or CKD hospitalization



**Fig. 1.** Cohort selection flowchart. Abbreviations: T2DM, type 2 diabetes mellitus; ADM, antidiabetic medication; DX, diagnosis; T1DM, type 1 diabetes; SU, sulfonylureas; DPP-4 inhibitor, dipeptidyl peptidase 4 inhibitors; GLP-1RA, glucagon-like peptide receptor agonists; SGLT-2 inhibitor, sodium-glucose cotransporter 2 inhibitor; TZD, Thiazolidinediones.

prior to index date were excluded from the analysis for this outcome. For any of the outcomes, patients were followed up from their index date until meeting criteria of one of the above outcomes, last hospital visit date, end of 5-year follow up period or end of study date (2019-10-29) whichever comes first.

### Covariates

We included 5 major categories of covariates: demographics, laboratory tests related to CKD, diagnoses from medical history before index date, medications that are known to affect CKD related outcome, insurance status and smoking status. Demographics includes age, gender, race were collected at baseline (index date). Race categories included White, Black, Asian, and other races. We used White race as the reference group. Insurance status categories included Medicaid, self-pay, Medicare, commercial insurance. Commercial insurance was used as the reference group. Laboratory test included hemoglobin A1C (HbA1c), high density lipoprotein cholesterol (HDL), total cholesterol (TOTCHL), serum creatinine, body mass index (BMI). Medications included angiotensin-converting enzyme (ACE) inhibitors, aldosterone receptor antagonists (ARA), angiotensin II receptor blocker (ARB), antiplatelet drugs, beta-blockers, calcium channel blockers, diuretics, statins, and other lipid modifying drugs (loop and thiazide diuretics). Medical history included cardiovascular disease (CVD), congestive heart failure (CHF), hypertension, vascular disease, vascular complication of diabetes (including skin ulcer), diabetic neuropathy, diabetic oculopathy, dyslipidemia, and other diabetic complications (diabetic nephropathy, nephrotic syndrome, nephritis/nephropathy, lower extremity amputations). For CKD hospitalization outcome, prior CKD hospitalization was also included as a covariate.

### Data cleaning

We included the medical history and medication use covariates for any time before the index date. For lab tests, we used the measurement closest to the index date in the prior 2 years. We treated All the lab tests (HbA1c, HDL, TOTCHL, BMI, serum creatinine) as continuous variables. We excluded extreme values that were likely to be erroneous values (e.g. HDL > 100 mg/dL, BMI < 8 kg/m<sup>2</sup>, BMI > 100 kg/m<sup>2</sup>, HbA1c > 20 % (195 mmol/mol), TOTCHL > 1000 mg/dL) from the study. A patient may not have any laboratory test during our 2-year searching time window. In addition, some patients also have missing values in race and insurance status in EHR. To deal with the missing data, we used multiple imputation by chained equation [13]. In our analysis, we used predictive mean matching as the imputation method for continuous variables and polytomous regression imputation for unordered categorical data. We created 10 multiple imputed datasets. For each imputation, we set the number of iterations as 20. We fit Cox regression model on each imputed dataset. We then used Rubin's rule to pool coefficient estimates from 10 Cox regression models [14].

### Statistical analysis

We generated a simple statistical summary and stratified patients by second-line ADM medication. For categorical variables, we conducted Chi-square test for the differences among second-line ADM classes. For continuous variables, we performed *t*-test to test the differences among second-line ADM classes at baseline. To evaluate the association of ADM drug classes and three different CKD-related outcomes, we developed a series of cox proportional hazard regression models. In our baseline model, we included only ADM class variables. In our basic demographic model, we included both ADM class and basic demographic information. In our demographic-medical history model, we included ADM class variables, basic demographics, and medical history. In the full model, we included all the mentioned variables in the Covariates section. Cox regressions were conducted using 'survival' package in R, version 3.6.0.

Multiple data imputation was performed using 'MICE' package in R, version 3.6.0. Estimate pooling from 10 imputed datasets were performed using 'Hmisc' in R, version 3.6.0. Descriptive data statistics were generated using 'TableOne' module in python, version 3.7.3.

### Sensitivity analysis

We conducted a sensitivity analysis to evaluate the robustness of our association findings. In the sensitivity analysis, we excluded patients with missing race, insurance status, HbA1c, total cholesterol, BMI, SBP, eGFR, and serum creatinine instead of imputing the missing data. The same Cox regression models were performed to assess the association of ADMs with the three above mentioned outcomes.

### Results

We identified 3403 patients in our final cohort who started with Biguanides (mainly metformin) as their first ADM and used either a DPP-4 inhibitor, SGLT-2 inhibitor, GLP-1RA, TZD, SU, or an insulin as their second line medication. Table 1 shows the baseline characteristics of patients stratified by second-line ADM classes. For diagnosed CKD incident outcome, we further excluded patients who had CKD diagnosis before the index date, which left us 3216 patients in the final cohort for diagnosed CKD incidence outcome (see Fig. 1). For eGFR outcome, we further excluded patients with prior eGFR < 45 mL/min, which left us 2746 patients in the final cohort for eGFR < 45 mL/min outcome.

Among the 3403 patients, overall, the mean age of the population is 59.8 (S.D = 12.0). Among these patients, 125 (3.67 %) patients have missing data in race. For these who have race information, 2408 (73.5 %) were white, 458 (14.0 %) black, 178 (5.4 %) Asian, 234 (7.1 %) classified as other races; 1853 (54.5 %) patients were male, and 1550 (45.5 %) were female. There are 44 (1.29 %) patients with missing insurance. Among the patients with insurance, 113 (3.5 %) were self-pay, 1706 (50.8 %) had commercial insurance, 128 (3.8 %) were Medicaid, 1407 (41.9 %) were Medicare. For laboratory test, there were 391 (11.49 %) patients with missing HDL-C, 129 (3.79 %) patients with missing HbA1c, 18 (0.5 %) patients with missing BMI, and 462 (13.6 %) patients with missing TOTCHL. For those patients with available laboratory values, the mean of HDL-C is 46.0 mg/dL (S.D. = 12.3), the mean of HbA1c is 7.4 % (57 mmol/mol) (S.D. = 1.5), the mean of TOTCHL is 166.5 mg/dL (S.D. = 40.4), the median of serum creatinine is 0.9 mg/dL ([Q1, Q3] = [0.7–1.0]). Among the 3403 patients, 1192 (35.0 %) patients had DPP-4 inhibitors as second line ADM, 208 (6.1 %) patients had GLP-1RA as second line ADM, 215 (6.3 %) patients had insulin as second ADM, 1348 (39.6 %) used SU as second line ADM, 355 (10.4 %) patients had SGLT-2 inhibitor as second line ADM and 85 (2.5 %) patients had TZD as second line ADM.

Statistical tests showed that age, race, insurance, smoking status, HDL-C, HbA1c, BMI, baseline serum creatinine, eGFR, CKD, CHF, diabetic neuropathy, other complications, vascular disease, VCD, oculopathy, CVD, use of ACE inhibitors, ARA, antiplatelet, diuretics, and statin were significantly different among the six second line ADM groups (see Table 1). In general, patients in newer medication groups (DPP-4 inhibitors, SGLT-2 inhibitors, GLP-1RA) tended to be younger, more likely to be white and have commercial insurance compared to patients in older medication groups (insulin, SU, TZD). In addition, patients in the insulin group tended to have higher HbA1c, lower eGFR and higher prevalence of use of renal related medications and comorbidities compared to other medication groups. In our regression analysis, in the full model, we adjusted for these differences by including demographics, medication history, diagnosis history, and laboratory tests as covariates.

During the 5-year follow up, 232 of 3216 patients (7.2 %) had diagnosed incident CKD (i.e., incident CKD based on ICD codes) with median length of follow-up of 3.23 years. Out of 3403 patients, 109 (3.2 %) patients had CKD hospitalization with median length of follow-up of 3.31 years. Out of 2746 patients, 254 (9.2 %) patients had eGFR < 45

**Table 1**  
Baseline characteristics of study cohort.

	Grouped by second-line medication								P-Value
	Missing	Overall	DPP-4	GLP-1	Insulin	SGLT2	SU	TZD	
n		3403	1192	208	215	355	1348	85	
<b>Basic Demographics</b>									
Age, mean (SD)	0	59.8 (12.0)	59.8 (11.6)	54.9 (10.8)	59.5 (12.3)	56.3 (11.4)	61.2 (12.3)	63.5 (10.9)	<0.001
Gender, n (%)		1853 (54.5)	656 (55.0)	91 (43.8)	111 (51.6)	204 (57.5)	739 (54.8)	52 (61.2)	
<b>Race, n (%)</b>									
Asian	125	178 (5.4)	61 (5.3)	11 (5.7)	8 (4.0)	15 (4.3)	75 (5.7)	8 (9.8)	<0.001
Black		458 (14.0)	136 (11.9)	30 (15.5)	65 (32.2)	21 (6.1)	198 (15.1)	8 (9.8)	
Other races		234 (7.1)	86 (7.5)	9 (4.7)	15 (7.4)	25 (7.2)	97 (7.4)	2 (2.4)	
White		2408 (73.5)	860 (75.2)	143 (74.1)	114 (56.4)	285 (82.4)	942 (71.8)	64 (78.0)	
<b>Insurance, n (%)</b>									
Self-pay	44	118 (3.5)	43 (3.6)	5 (2.5)	13 (6.2)	5 (1.4)	49 (3.7)	3 (3.6)	<0.001
Commercial		1706 (50.8)	622 (52.8)	144 (70.6)	101 (48.3)	241 (68.3)	559 (42.0)	39 (46.4)	
Medicaid		128 (3.8)	34 (2.9)	8 (3.9)	6 (2.9)	11 (3.1)	69 (5.2)		
Medicare		1407 (41.9)	480 (40.7)	47 (23.0)	89 (42.6)	96 (27.2)	653 (49.1)	42 (50.0)	
Smoking status, n (%)		475 (14.0)	167 (14.0)	31 (14.9)	47 (21.9)	36 (10.1)	186 (13.8)	8 (9.4)	0.004
<b>Laboratory test</b>									
HDL-C, mean (SD)	391	46.0 (12.3)	45.9 (12.2)	47.0 (13.0)	46.2 (12.4)	44.6 (12.0)	45.9 (12.2)	49.8 (16.4)	0.033
HbA1c, mean (SD)	129	7.4 (1.5)	7.3 (1.3)	7.2 (1.6)	7.9 (2.0)	7.3 (1.4)	7.5 (1.6)	7.0 (1.1)	<0.001
BMI, mean (SD)	18	34.2 (7.7)	33.4 (7.4)	37.2 (7.4)	34.3 (7.1)	35.9 (8.1)	33.9 (7.9)	33.6 (7.3)	<0.001
TOTCHL, mean (SD)	462	166.5 (40.4)	165.3 (39.7)	168.0 (35.0)	163.4 (48.4)	169.2 (38.2)	167.9 (41.4)	155.3 (31.3)	0.059
Baseline serum creatinine, median [Q1,Q3]	166	0.9 [0.7,1]	0.9 [0.7,1]	0.9 [0.7,1]	0.9 [0.8,1.1]	0.8 [0.7,1]	0.9 [0.8,1]	0.9 [0.8,1.1]	0.005
eGFR, mean (SD)	199	76.4 (28.8)	77.1 (28.7)	62.7 (14.9)	60.3 (22.1)	82.7 (28.8)	79.2 (30.0)	73.1 (28.4)	<0.001
<b>Diagnosis History</b>									
CKD, n (%)		187 (5.5)	62 (5.2)	11 (5.3)	23 (10.7)	10 (2.8)	74 (5.5)	7 (8.2)	0.004
CHF, n (%)		213 (6.3)	62 (5.2)	16 (7.7)	23 (10.7)	16 (4.5)	90 (6.7)	6 (7.1)	0.028
Hypertension, n (%)		2761 (81.1)	983 (82.5)	161 (77.4)	181 (84.2)	274 (77.2)	1089 (80.8)	73 (85.9)	0.089
Dyslipidemia, n (%)		2635 (77.4)	939 (78.8)	152 (73.1)	169 (78.6)	272 (76.6)	1039 (77.1)	64 (75.3)	0.534
Diabetic Neuropathy, n (%)		332 (9.8)	105 (8.8)	19 (9.1)	45 (20.9)	29 (8.2)	127 (9.4)	7 (8.2)	<0.001
Other Complications, n (%)		274 (8.1)	87 (7.3)	17 (8.2)	37 (17.2)	30 (8.5)	93 (6.9)	10 (11.8)	<0.001
Vascular disease, n (%)		330 (9.7)	103 (8.6)	22 (10.6)	40 (18.6)	29 (8.2)	130 (9.6)	6 (7.1)	<0.001
VCD, n (%)		466 (13.7)	147 (12.3)	19 (9.1)	37 (17.2)	58 (16.3)	185 (13.7)	20 (23.5)	0.005
Oculopathy, n (%)		128 (3.8)	28 (2.3)	11 (5.3)	24 (11.2)	13 (3.7)	47 (3.5)	5 (5.9)	<0.001
CVD, n (%)		644 (18.9)	196 (16.4)	43 (20.7)	64 (29.8)	57 (16.1)	268 (19.9)	16 (18.8)	<0.001
hypoglycemia, n (%)		43 (1.3)	6 (0.5)	3 (1.4)	4 (1.9)	6 (1.7)	23 (1.7)	1 (1.2)	0.116
<b>Medication History</b>									
ACE inhibitors, n (%)		2162 (63.5)	765 (64.2)	128 (61.5)	151 (70.2)	204 (57.5)	849 (63.0)	65 (76.5)	0.005
ARA, n (%)		271 (8.0)	82 (6.9)	18 (8.7)	26 (12.1)	26 (7.3)	117 (8.7)	2 (2.4)	0.037
ARB, n (%)		251 (7.4)	97 (8.1)	25 (12.0)	13 (6.0)	25 (7.0)	84 (6.2)	7 (8.2)	0.054
CCBs, n (%)		817 (24.0)	294 (24.7)	47 (22.6)	61 (28.4)	71 (20.0)	321 (23.8)	23 (27.1)	0.274
Antiplatelet, n (%)		1116 (32.8)	376 (31.5)	83 (39.9)	99 (46.0)	99 (27.9)	431 (32.0)	28 (32.9)	<0.001
Beta Blocker, n (%)		1098 (32.3)	374 (31.4)	62 (29.8)	74 (34.4)	99 (27.9)	459 (34.1)	30 (35.3)	0.228
Diuretics, n (%)		1368 (40.2)	466 (39.1)	91 (43.8)	108 (50.2)	124 (34.9)	544 (40.4)	35 (41.2)	0.011
Lipid modifier, n (%)		541 (15.9)	190 (15.9)	38 (18.3)	37 (17.2)	50 (14.1)	204 (15.1)	22 (25.9)	0.111
Statin, n (%)		2284 (67.1)	805 (67.5)	135 (64.9)	154 (71.6)	224 (63.1)	897 (66.5)	69 (81.2)	0.024

This table shows the baseline patient characteristics for CKD hospitalization outcome before missing data imputation. For diagnosed CKD incidence outcome, patients with prior ESRD and prior CKD were excluded from the study cohort. For composite renal outcome, patients with prior ESRD were excluded from the study cohort. Abbreviations: HDL, high density cholesterol, HbA1c, hemoglobin A1c; BMI, body mass index; TOTCHL, total cholesterol; VCD: vascular complications of diabetes; CHF, congestive heart failure; CVD, cardiovascular disease; ACE inhibitor, angiotensin-converting enzyme inhibitors; ARA, aldosterone receptor antagonists; ARB, angiotensin II receptor blocker. DPP-4, dipeptidyl peptidase 4 inhibitors; GLP-1RA, glucagon-like peptide receptor agonists; SGLT-2, sodium-glucose cotransporter 2 inhibitor; TZD, Thiazolidinediones. Measurement units for laboratory tests are as the following: HDL-C, mg/dL; HbA1c, %; BMI, kg/m<sup>2</sup>; TOTCHL, mg/dL; serum creatinine, mg/dL.

\*Other complications include diabetic nephropathy, nephrotic syndrome, nephritis, diabetic retinopathy, cataract, lower extreme amputation.

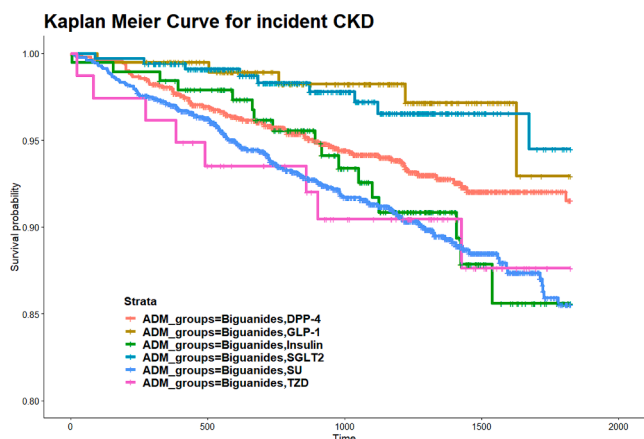
mL/min with median length of follow-up of 3.23 years. Figs. 2–4 showed the unadjusted Kaplan Meier curves for diagnosed incident CKD, CKD hospitalization, and eGFR < 45 min/mL outcomes, respectively. We also examined the change of HbA1c during the 5-year follow up within each second line ADM groups. Patients in the insulin group started with the highest HbA1c among all groups. For this group, HbA1c remained stable during the first 4 years and went up during year 4 to year 5. HbA1c also went up in SGLT-2 group when comparing the fifth year’s value to the baseline value (first two years prior to the first ADM medication). For the

GLP-1RA group, HbA1c in year five slightly decreased from baseline but remained relatively stable overall. For DPP-4, SU, TZD groups, HbA1c increased slightly during the 5-years follow up. (See Supplemental Table 3 for details).

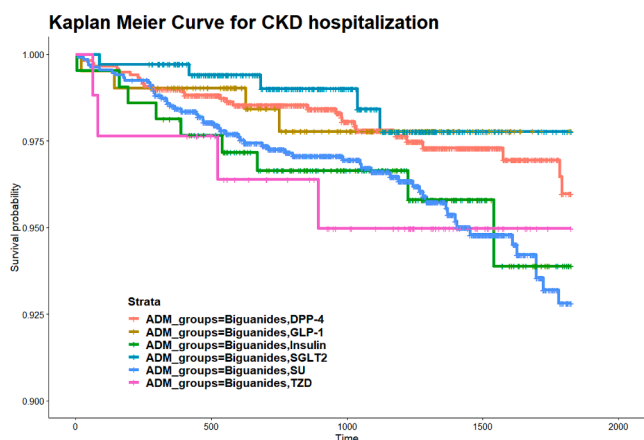
*Risk for diagnosed incident CKD, CKD hospitalization, and eGFR < 45 mL/min outcome in primary analysis*

Table 2 showed the number of events and hazard ratio of all three



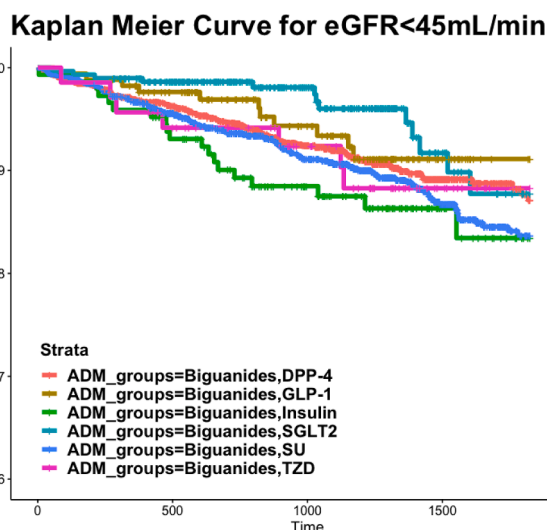


**Fig. 2.** Unadjusted Kaplan Meier curve for incident CKD in different ADM groups. Abbreviations: CKD, chronic kidney disease; ADM, anti-diabetic medication; DPP4, dipeptidyl peptidase 4 inhibitors; GLP-1, glucagon-like peptide receptor agonists; SGLT2 inhibitor, sodium-glucose cotransporter 2 inhibitor; TZD, Thiazolidinediones.



**Fig. 3.** Unadjusted Kaplan Meier curve for CKD hospitalization in different ADM groups. Abbreviations: CKD, chronic kidney disease; ADM, anti-diabetic medication; sDPP4, dipeptidyl peptidase 4 inhibitors; GLP1RA, glucagon-like peptide receptor agonists; SGLT2 inhibitor, sodium-glucose cotransporter 2 inhibitor; TZD, Thiazolidinediones.

renal outcome for each second-line ADM class using SU as reference group in a fully adjusted model in the primary analysis (see [Supplemental Tables 4–6](#) for hazard ratio in baseline model, basic demographics model, basic demographics/medical history model in primary analysis) [7]. Both SGLT-2 inhibitor (HR, 0.43; 95 % CI, [0.22;0.87]; P = 0.02) and DPP-4 inhibitor (HR, 0.71; 95 % CI, [0.53;0.96]; P = 0.03) are significantly associated with lower diagnosed CKD incidence event. GLP-1RA was associated with reduced risk for diagnosed CKD incidence (HR, 0.52; 95 % CI, [0.21;1.29]; P = 0.16), but the P-value is not significant. For CKD hospitalization outcome, DPP-4 is significantly associated with lower risk for CKD hospitalization (HR, 0.60; 95 % CI, [0.37;0.96]; P = 0.03). Use of the other ADMs also did not show significant difference in CKD hospitalization outcome compared to use of SU. For eGFR < 45 mL/min outcome, we observed that SGLT-2 inhibitor was associated with lower risk for eGFR < 45 mL/min (HR, 0.58; 95 % CI, [0.32, 1.07]; P = 0.08). However, this association was not statistically significant. We did not observe significant association between DPP-4 inhibitor and eGFR < 45 mL/min.



**Fig. 4.** Unadjusted Kaplan Meier curve for eGFR < 45 mL/min outcome in different ADM groups. Abbreviations: DPP4, dipeptidyl peptidase 4 inhibitors; GLP1RA, glucagon-like peptide receptor agonists; SGLT2 inhibitor, sodium-glucose cotransporter 2 inhibitor; TZD, Thiazolidinediones.

**Table 2**

Hazard ratio in the fully adjusted Cox regression model in primary analysis.

	Diagnosed CKD incidence (N = 3216, E = 232)		CKD hospitalization (N = 3403, E = 109)		eGFR < 45 mL/min (N = 2746, E = 254)	
	HR (95 % CI)	Pval	HR (95 % CI)	Pval	HR (95 % CI)	Pval
DPP-4	0.71, [0.53;0.96]	0.03	0.6, [0.37;0.96]	0.04	0.95, [0.72;1.26]	0.73
GLP-1RA	0.52, [0.21;1.30]	0.16	1.05, [0.37;3.02]	0.92	0.87, [0.46;1.65]	0.68
Insulin	0.93, [0.54;1.59]	0.80	0.52, [0.24;1.17]	0.11	1.18, [0.71;1.95]	0.52
SGLT-2	0.43, [0.22;0.87]	0.02	0.81, [0.31;2.09]	0.66	0.58, [0.32;1.07]	0.08
TZD	1.03, [0.49;2.13]	0.93	1.25, [0.44;3.70]	0.65	0.96, [0.44;2.08]	0.91

Abbreviations: HR, hazard ratio; DPP-4, dipeptidyl peptidase 4 inhibitors; GLP-1RA, glucagon-like peptide receptor agonists; SGLT-2, sodium-glucose cotransporter 2 inhibitor; TZD, Thiazolidinediones; Pval, p-value; E, number of events.

*Sensitivity analysis*

Table 3 showed hazard ratios for ADMs relative to use of SU in the

**Table 3**

Hazard ratio in the fully adjusted Cox regression model in sensitivity analysis.

	CKD incidence (N = 2364, E = 159)		CKD hospitalization (N = 2499, E = 75)		eGFR < 45 mL/min (N = 2164, E = 197)	
	HR (95 % CI)	Pval	HR (95 % CI)	Pval	HR (95 % CI)	Pval
DPP-4	0.75, [0.52;1.08]	0.12	0.75, [0.42;1.32]	0.32	0.98, [0.71;1.36]	0.90
GLP-1	0.53, [0.19;1.48]	0.23	0.94, [0.26;3.43]	0.92	0.97, [0.48;1.96]	0.93
Insulin	0.72, [0.37;1.43]	0.35	1.00, [0.39;2.57]	1.00	1.49, [0.83;2.65]	0.18
SGLT-2	0.42, [0.19;0.92]	0.03	1.35, [0.50;3.66]	0.56	0.61, [0.31;1.18]	0.14
TZD	1.15, [0.49;2.71]	0.75	0.9, [0.21;3.90]	0.89	0.97, [0.42;2.27]	0.95

Abbreviations: HR, hazard ratio; DPP-4, dipeptidyl peptidase 4 inhibitors; GLP-1RA, glucagon-like peptide receptor agonists; SGLT-2, sodium-glucose cotransporter 2 inhibitor; TZD, Thiazolidinediones; Pval, p-value; E, number of events.

sensitivity analysis in fully adjusted model (see [Supplemental Table 7–9](#) for hazard ratio in baseline model, basic demographics model, basic demographics/medical history model in sensitivity analysis) [7]. Our sensitivity analyses included only those patients with complete data. For the incident CKD event, after excluding patients with missing race, insurance status, HbA1c, insurance status, HDL, total cholesterol, BMI, baseline serum creatinine, and baseline eGFR, 2364 patients remained. Our analysis showed that in the fully adjusted model, SGLT-2 inhibitor use (HR, 0.42; 95 % CI, [0.19, 0.92],  $P = 0.03$ ) was significantly associated with lower incidence of CKD. DPP-4 inhibitor use (HR, 0.75; 95 % CI, [0.52;1.08];  $P = 0.12$ ) was associated with lower diagnosed CKD incidence but the  $P$  value was not statistically significant. For the CKD hospitalization outcome, after excluding patients with missing values, there were 2499 patients left. we did not observe any significant association between a particular second line ADM class relative to SU. For eGFR < 45 mL/min outcome, use of SGLT-2 inhibitor showed lower risk for eGFR < 45 mL/min compared to the reference group, though this association was not statistically significant (HR, 0.61; CI, [0.31,1.18],  $P = 0.14$ ).

## Discussion

Randomized clinical trials and observational studies are two major ways to assess associations between medication and corresponding clinical outcomes [15]. Although randomized clinical trials are rigorously designed, they have several limitations in testing the association between ADMs and renal outcomes among type 2 diabetes patients. First, the major randomized clinical trials compared ADM's effect to a placebo group. To date the ongoing Glycemia Reduction Approaches in Diabetes: A Comparative Effectiveness Study (GRADE) (54) is the only clinical trial (finished or ongoing) that we are aware of that evaluates the comparative effectiveness of newer ADMs vs older ADMs on glucose control [16]. Further, most clinical trials treated second line ADMs as a monotherapy [3,17]. Therefore, it is unclear how the effect changes when the ADMs are used as add-on therapy on top of metformin, which is the most common case in the primary care setting. Secondly, patients in clinical trials are better positioned to adhere to medication and therapy than in real clinical settings. Therefore, the results drawn from clinical trials may not apply to real clinical settings. Observational studies using EHR data like ours could potentially close these gaps by providing evidence from the real-world data of general populations.

In this study, we used EHR data to investigate the association of second-line ADMs with renal outcomes among 3403 type 2 diabetes patients who initiated treatment with metformin. We found that both the use of SGLT-2 inhibitor and DPP-4 inhibitor showed significant association with less incident CKD defined by ICD codes as compared to use of SU as a second-line medication. However, in the sensitivity analysis, only SGLT-2 inhibitor remained significantly associated. For the CKD hospitalization event, DPP-4 inhibitor appeared protective for CKD hospitalization in the primary analysis, but the association was not statistically significant in the sensitivity analysis. For eGFR < 45 mL/min outcome, the association between use of SGLT-2 inhibitors and lower event rate of eGFR < 45 mL/min was marginally significant ( $P = 0.08$ ). Both diagnosed CKD incidence and eGFR < 45 mL/min are ways to identify CKD incidence in EHR. We found that majority (2510/2746) of the patients who had eGFR < 45 mL/min also had diagnosed incident CKD. Out of the 2746 patients who had eGFR < 45 mL/min, 171 patients did not have diagnosed incident CKD. The different results for eGFR < 45 mL/min and ICD based CKD incidence outcome likely reflects the discrepancy of recording time between lab values and diagnosis code.

In each second-line ADM group, around half of the patients switched to a third medication during the one-year exposure period (see [Supplemental Table 10](#) for details). To evaluate the impact of switching to a third or more medication might have on the result, we examined patients who took 2 medications only during the one-year exposure. We observed that the “Biguanide, DPP-4”, “Biguanides, SGLT-2” group had

lower hazard risk for diagnosed CKD incidence compared to the reference group (see [Supplemental Table 10](#)). However, the association was not significant ( $P = 0.19$  for DPP-4 inhibitor group and  $P = 0.13$  for SGLT-2 inhibitor group). This is likely due to small size after restricting to patients taking 2 medications only. In addition, SGLT2 is a relatively new medication, the follow-up time might not be long enough to develop CKD.

The above results from primary and secondary analysis are largely in line with results from placebo-controlled cardiovascular outcomes trials, which showed empagliflozin, canagliflozin, dapagliflozin, and linagliptin all had beneficial effects on indices of CKD [18,19]. For example, the CREDENCE (Canagliflozin and Renal Events in Diabetes with Established Nephropathy Clinical Evaluation) trial showed that type 2 diabetes patients with prior albuminuria assigned canagliflozin (a SGLT-2 inhibitor) had lower risk of composite renal events compared to placebo group [18]. Cooper et al. found linagliptin (a DPP-4 inhibitor) significantly reduced the risk of kidney disease events compared to placebo among type 2 diabetes patients [19]. Groop et al [20] found linagliptin on top of renin-angiotensin-aldosterone system significantly reduced albuminuria among patients with type 2 diabetes and renal dysfunction. We did not find statistically significant difference between GLP-1RA and SU, Insulin and SU, TZD and SU in either primary analysis or sensitivity analysis regarding diagnosed CKD incidence and CKD hospitalization event. For GLP-1RA, previous studies showed inconclusive evidence of its effect in reduction of renal disease [21]. The LEADER [21] study showed GLP-1RA had benefit in reduction of new onset albuminuria while the LIRA-RENAL study showed no improvement in urine ACR, and eGFR change among renal impaired T2DM patients compared to the placebo group [22]. For TZD vs SU, Insulin vs SU, we did not find evidence of one is superior to the other in renal disease outcome in the literature.

Our study has several strengths. First, our clearly defined exposure time starting from first prescription of metformin, with additional prescription of a second line ADM within a year: this ensured comparison of patients who were in a clinically similar diabetes stage. Second, we only included patients who started ADMs after 2013–03-01 which is the time when the first SGLT-2 inhibitor was approved by FDA for use. This ensures that patients using different second-line ADMs have comparable follow-up times and avoids the immortal time bias [23]. Third, we used a rigorously defined patient cohort: we excluded any patients who had type 1 diabetes diagnosis or gestational diabetes. In addition, we only included patients who had 3 continuous years of ADM prescriptions since their first ADM exposure. This makes sure that each patient has sufficient and comparable data depth. Last while there could be confounders that were not included in the study, our list of covariates is relatively comprehensive based on a review of the literature. This list contains basic demographic information, laboratory tests, and a wide range of diagnoses that might not only be associated with renal disease, but also suggest disease severity.

Our study has several limitations. Firstly, for simplicity, we could not distinguish between a switch of medication from an add-on medication. For example, patients who have a DPP-4 inhibitor as the second medication in their sequence could either: 1) have switched entirely from metformin to a DPP-4 inhibitor, or 2) began using metformin and a DPP-4 inhibitor together. Secondly, while we controlled for as many variables/confounders as we could, it is possible that there remain other covariates we did not capture such as diabetes duration, which could affect prescription bias or are still associated with the renal outcomes. Therefore, conclusions drawn from our study should be interpreted with caution. Thirdly, as EHR data is primarily designed for administrative purpose, it is well known that EHR's have missing data issues. Because this is a single site study, patients may have sought care elsewhere. Therefore, it is likely that there were data points that we were unable to capture in this single site study. Additionally, this was a single center study of an academic hospital in a major US city, limiting its generalizability to other populations. Finally, it is known that there is a delay

between the date of a physician attaching ICD codes for CKD and the actual onset date of CKD. To reduce the impact of the above-mentioned delay, we included laboratory data (eGFR < 45 mL/min) which is usually more accurate and evident earlier in EHR as an additional endpoint.

## Conclusion

In conclusion, our study assessed the association of common second-line medications with renal outcomes as compared to SU's from March 2013 through October 2019 using EHRs. Our results showed that SGLT-2 inhibitors were significantly associated with lower diagnosed CKD incidence in both primary and sensitivity analyses and DPP-4 inhibitor was significantly associated with lower diagnosed CKD incidence in primary analysis. However, none of the significant associations were seen in eGFR < 45 mL/min outcome. We did not observe any statistical significance between Insulin and SU, TZD and SU in either primary analysis or sensitivity analysis regarding diagnosed CKD incidence, CKD hospitalization event, or eGFR < 45 mL/min outcome in our dataset. Unlike reported in Xie et al.'s study [9], We also did not observe significant benefit of GLP-1RA compared to SU in CKD hospitalization and diagnosed CKD incidence, possibly due to small sample size. Additional research using multi-site real-world data and larger sample sizes may be needed to confirm the generalizability of our results.

## Contributors

Y.D., L.Z., and A.K. designed the study. Y.D and F.G. performed the data analysis. L.Z. provided expertise data analysis. A.K., L.Z., F.G., and J.Y. contributed to discussion. A.L. contributed to data extraction. Y.D. and F.G. wrote the first draft of the manuscript with all coauthors contributing to revision. A.K. supervised the project. Y.D. and A.K. are 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.

## Patient consent for publication

NA.

## Ethnic approval

The project was reviewed and approved by Northwestern Institutional Review Board.

## Funding

NA.

## CRedit authorship contribution statement

**Yu Deng:** Conceptualization, Methodology, Data curation, Writing – original draft. **Farhad Ghamsari:** Writing – original draft, Writing – review & editing. **Alice Lu:** Data curation. **Jingzhi Yu:** Writing – review & editing. **Lihui Zhao:** Conceptualization, Methodology. **Abel N. Kho:** Conceptualization, Methodology, Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors thank Raymond Kang (Northwestern University) for scientific discussions and technical assistance.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jcte.2022.100309>.

## References

- [1] Alicic RZ, Rooney MT, Tuttle KR. Diabetic kidney disease: challenges, progress, and possibilities. *CJASN* 2017;12(12):2032–45.
- [2] Wilkinson SV, Tomlinson LA, Iwagami M, Stirnadel-Farrant HA, Smeeth L, Douglas I. A systematic review comparing the evidence for kidney function outcomes between oral antidiabetic drugs for type 2 diabetes. *Wellcome Open Res* 2018;3:74.
- [3] Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med* 2017;377(7):644–57.
- [4] Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes | *NEJM* <https://www.nejm.org/doi/full/10.1056/NEJMoa1504720>. Accessed April 27, 2021.
- [5] Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2019;380(4):347–57.
- [6] Heerspink HJL, Stefánsson BV, Correa-Rotter R, Chertow GM, Greene T, Hou F-F, et al. Dapagliflozin in patients with chronic kidney disease. *N Engl J Med* 2020;383(15):1436–46.
- [7] Understanding the Gap Between Efficacy in Randomized Controlled Trials and Effectiveness in Real-World Use of GLP-1 RA and DPP-4 Therapies in Patients With Type 2 Diabetes | *Diabetes Care* <https://care.diabetesjournals.org/content/40/11/1469>. Accessed July 15, 2020.
- [8] Schneeweiss S, Paterno E. Conducting real-world evidence studies on the clinical outcomes of diabetes treatments. *Endocr Rev* 2021;42(5):658–90.
- [9] Xie Y, Bowe B, Gibson AK, McGill JB, Maddukuri G, Yan Y, et al. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of kidney outcomes: emulation of a target trial using health care databases. *Diabetes Care* 2020;43(11):2859–69.
- [10] Northwestern Medicine Enterprise Data Warehouse (NMEDW): Research: Feinberg School of Medicine: Northwestern University <https://www.feinberg.northwestern.edu/research/cores/units/edw.html>. Accessed April 27, 2021.
- [11] Deng Y, Ghamsari F, Lu A, et al.: Comparative Effectiveness of Second Line Anti-Diabetic Medication on Incidence of Chronic Kidney Disease. 10.6084/m9.figshare.13426469.
- [12] WHOCC - ATC/DDD Index [https://www.whooc.no/atc\\_ddd\\_index/](https://www.whooc.no/atc_ddd_index/). Accessed October 9, 2020.
- [13] Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatric Res* 2011;20(1):40–9.
- [14] Barnard J, Rubin DB. Small-sample degrees of freedom with multiple imputation. *Biometrika* 1999;86:948–55.
- [15] Ross JS. Randomized clinical trials and observational studies are more often alike than unlike. *JAMA Intern Med* 2014;174:1557.
- [16] GRADE Study Group: Glycemia Reduction Approaches in Diabetes: A Comparative Effectiveness Study. *clinicaltrials.gov*, 2021. <https://clinicaltrials.gov/ct2/show/NCT01794143>. Accessed April 26, 2021.
- [17] Effects of Linagliptin on Cardiovascular and Kidney Outcomes in People With Normal and Reduced Kidney Function: Secondary Analysis of the CARMELINA Randomized Trial | *Diabetes Care* <https://care.diabetesjournals.org/content/early/2020/05/20/dc20-0279>. Accessed September 9, 2020.
- [18] Cooper ME, Perkovic V, McGill JB, Groop P-H, Wanner C, Rosenstock J, et al. Kidney disease end points in a pooled analysis of individual patient-level data from a large clinical trials program of the dipeptidyl peptidase 4 inhibitor linagliptin in type 2 diabetes. *Am J Kidney Dis* 2015;66(3):441–9.
- [19] Linagliptin Lowers Albuminuria on Top of Recommended Standard Treatment in Patients With Type 2 Diabetes and Renal Dysfunction | *Diabetes Care* [https://care.diabetesjournals.org/content/36/11/3460?ijkey=d7fbf7b1ce210bede909e4a7587050d9cadd67a0&keytype=tf\\_ipsecsha](https://care.diabetesjournals.org/content/36/11/3460?ijkey=d7fbf7b1ce210bede909e4a7587050d9cadd67a0&keytype=tf_ipsecsha). Accessed September 8, 2020.
- [20] Mann JFE, Ørsted DD, Brown-Frandsen K, Marso SP, Poulter NR, Rasmussen S, et al. Liraglutide and renal outcomes in type 2 diabetes. *N Engl J Med* 2017;377(9):839–48.
- [21] Davies MJ, Bain SC, Atkin SL, Rossing P, Scott D, Shamkhalova MS, et al. Efficacy and safety of liraglutide versus placebo as add-on to glucose-lowering therapy in patients with type 2 diabetes and moderate renal impairment (LIRA-RENAL): a randomized clinical trial. *Diabetes Care* 2016;39(2):222–30.
- [22] Suissa S. Lower risk of death with SGLT2 inhibitors in observational studies: real or bias? *Diabetes Care* 2018;41:6–10.
- [23] Marso SP, Daniels GH, Brown-Frandsen K, Kristensen P, Mann JFE, Nauck MA, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes. *N Engl J Med* 2016;375(4):311–22.

## Further reading