

Post-Authorization Safety Study of Hospitalization for Acute Kidney Injury in Patients With Type 2 Diabetes Exposed to Dapagliflozin in a Real-World Setting

Journal: *Drug Safety*

Catherine B. Johannes, PhD¹ ([0000-0002-0586-9886](#)); Daniel C. Beachler, PhD² ([0000-0003-2788-3061](#)); J. Bradley Layton, PhD³ ([0000-0003-0994-5820](#)); Heather E. Danysh, PhD¹ ([0000-0003-3014-7508](#)); Ryan Ziemiecki, MS⁴ ([0000-0002-1774-7034](#)); Alejandro Arana, MD⁵ ([0000-0002-1593-3124](#)); Jade Dinh, MPH^{2,6} ([0000-0003-0386-2776](#)); Ling Li, MSPH²; Brian Calingaert, MS³ ([0000-0001-8177-6326](#)); Manel Pladevall-Vila, MD^{5,7} ([0000-0002-9359-6055](#)); Phillip R. Hunt, ScD⁸ ([0000-0002-9559-175X](#)); Hungta Chen, PhD⁸; Cecilia Karlsson, MD⁹ ([0000-0002-4299-8775](#)); Kristina Johnsson, MD⁹; Alicia Gilsenan, PhD³ ([0000-0002-9266-1417](#))

¹ Department of Pharmacoepidemiology and Risk Management, RTI Health Solutions, Waltham, MA, USA

² Department of Safety and Epidemiology, HealthCore, Inc., Wilmington, DE, USA

³ Department of Pharmacoepidemiology and Risk Management, RTI Health Solutions, Research Triangle Park, NC, USA

⁴ Department of Biostatistics, RTI Health Solutions, Research Triangle Park, NC, USA

⁵ Department of Pharmacoepidemiology and Risk Management, RTI Health Solutions, Barcelona, Spain

⁶ Department of Research Operations, HealthCore, Inc., Wilmington, DE, USA

⁷ The Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, USA

⁸ BioPharmaceuticals Business Unit, AstraZeneca, Gaithersburg, MD, USA

⁹ BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

Corresponding Author:

Catherine Johannes

Department of Pharmacoepidemiology and Risk Management

RTI Health Solutions

307 Waverley Oaks Road, Suite 101

Waltham, MA 02452 USA

Phone: 781.434.1784

Email: cjohannes@rti.org

Online Resource 1

Post-Authorization Safety Study of Hospitalization for Acute Kidney Injury in Patients With Type 2 Diabetes Exposed to Dapagliflozin in a Real-World Setting

TABLE OF CONTENTS

LIST OF TABLES	3
LIST OF FIGURES.....	3
1 METHODS AND RESULTS.....	19
1.1 Propensity Score Modeling Approach.....	19
1.2 Estimation of Pooled Adjusted Incidence Rate Ratio.....	19
1.3 Assessment of the Potential Effect of Unmeasured Confounders (Quantitative Bias Analysis Methods, Results, and Interpretation).....	19
1.4 Definitions for Index Therapy Type Categories.....	21
2 REFERENCES.....	22

LIST OF TABLES

Table S1. Covariate Medical Conditions at Baseline	4
Table S2. Covariate Medications at Baseline	6
Table S3. Lifestyle, Demographic, and Health Care Utilization Variables at Baseline	8
Table S4. Results of the Cohort Selection Process; Counts of Treatment Episodes After Exclusions and Final Matched Cohorts for the Combined Dapagliflozin and Comparator GLD Exposure Groups	11
Table S5. Selected Baseline Characteristics, Analysis Sample After Propensity Score Trimming	13

LIST OF FIGURES

Figure S1. Study Design, Cohort Eligibility, and Inclusion Criteria.....	17
Figure S2. Sensitivity Analysis, Adjusted Incidence Rate Ratios Under Varying Assumptions of Unmeasured Confounding Compared With the Observed Incidence Rate Ratio Estimate From the Pooled Analysis	20

PART A. Supplemental Tables and Figure

Table S1. Covariate Medical Conditions at Baseline

Variable	Evaluation period	Coding
Hospitalization for a serious condition that requires intensive care ^a	During the 180 days before the index date	0 = Absent 1 = Present
Number of hospital stays		0, 1, 2 or more
Total number of hospital days		Mean (SD) Median (IQR) Min, max
Medical comorbidities		
Indicators of diabetes severity		
Diabetic nephropathy	During all available time up to and including the index date	0 = Absent 1 = Present
Peripheral neuropathy	During all available time up to and including the index date	0 = Absent 1 = Present
Peripheral vascular disease ^b	During all available time up to and including the index date	0 = Absent 1 = Present
Retinopathy	During all available time up to and including the index date	0 = Absent 1 = Present
Coronary heart disease	During all available time up to and including the index date	0 = Absent 1 = Present
Cerebrovascular disease	During all available time up to and including the index date	0 = Absent 1 = Present
Amputation	During all available time up to and including the index date	0 = Absent 1 = Present
Time since first diagnosis of T2DM ^c	During all available time up to and including the index date	Mean (SD) Median (IQR) Min, max
Other medical conditions		
Kidney and genitourinary stones	During all available time up to and including the index date	0 = Absent 1 = Present
Hypertension	During all available time up to and including the index date	0 = Absent 1 = Present
Heart failure	During all available time up to and including the index date	0 = Absent 1 = Present
Liver disease	During all available time up to and including the index date	0 = Absent 1 = Present
Cardiovascular disease other than ischemic heart disease, heart failure, or peripheral vascular disease	During all available time up to and including the index date	0 = Absent 1 = Present
Chronic obstructive pulmonary disease, emphysema, respiratory insufficiency	During all available time up to and including the index date	0 = Absent 1 = Present
Systemic connective tissue disorders	During all available time up to and including the index date	0 = Absent 1 = Present

Variable	Evaluation period	Coding
Rheumatoid arthritis	During all available time up to and including the index date	0 = Absent 1 = Present
Other autoimmune disorders	During all available time up to and including the index date	0 = Absent 1 = Present
Osteoarthritis	During all available time up to and including the index date	0 = Absent 1 = Present
Polymyalgia rheumatica	During all available time up to and including the index date	0 = Absent 1 = Present
Urinary infections (chronic or recurring)	During all available time up to and including the index date	0 = Absent 1 = Present
Colon polyps	During all available time up to and including the index date	0 = Absent 1 = Present
Crohn's disease	During all available time up to and including the index date	0 = Absent 1 = Present
Ulcerative colitis	During all available time up to and including the index date	0 = Absent 1 = Present
Pancreatitis	During all available time up to and including the index date	0 = Absent 1 = Present
Immunosuppressive diseases, such as HIV/AIDS	During all available time up to and including the index date	0 = Absent 1 = Present
Peptic ulcer disease	During all available time up to and including the index date	0 = Absent 1 = Present
Dementia	During all available time up to and including the index date	0 = Absent 1 = Present
Asthma	During all available time up to and including the index date	0 = Absent 1 = Present
Hyperlipidemia	During all available time up to and including the index date	0 = Absent 1 = Present
All malignancies other than non-melanoma skin cancer	During all available time up to but not including the index date	0 = Absent 1 = Present
Trauma ^d	During the 180 days before the index date	0 = Absent 1 = Present

CPRD = Clinical Practice Research Datalink; HIRD = HealthCore Integrated Research Database; HIV = human immunodeficiency virus; IQR = interquartile range; SD = standard deviation; T2DM = type 2 diabetes mellitus.

^a Variable included in the HIRD and Medicare data only.

^b Includes peripheral artery disease.

^c Variable included only in CPRD data.

^d Assessed in the 180 days before the index date.

Table S2. Covariate Medications at Baseline

Medication category	Evaluation period	Coding
Insulin	Current use on the index date	0 = Absent 1 = Present
Antihypertensives/diuretics	During the 180 days up to and including the index date	0 = Absent 1 = Present
Antiarrhythmics	During the 180 days up to and including the index date	0 = Absent 1 = Present
Digoxin	During the 180 days up to and including the index date	0 = Absent 1 = Present
Nitrates	During the 180 days up to and including the index date	0 = Absent 1 = Present
Lipid-modifying agents	During the 180 days up to and including the index date	0 = Absent 1 = Present
Non-steroidal anti-inflammatory drugs (NSAIDs)	During the 180 days up to and including the index date	0 = Absent 1 = Present
Systemic corticosteroids ^a	During the 180 days up to and including the index date	0 = Absent 1 = Present
Inhaled systemic corticosteroids	During the 180 days up to and including the index date	0 = Absent 1 = Present
Zoledronic acid	During the 180 days up to and including the index date	0 = Absent 1 = Present
Acetaminophen	During the 180 days up to and including the index date	0 = Absent 1 = Present
Antibiotics (all types)	During the 180 days up to and including the index date	0 = Absent 1 = Present
Anticonvulsants	During the 180 days up to and including the index date	0 = Absent 1 = Present
Antifungal agents	During the 180 days up to and including the index date	0 = Absent 1 = Present
Antituberculars	During the 180 days up to and including the index date	0 = Absent 1 = Present
Methotrexate	During the 180 days up to and including the index date	0 = Absent 1 = Present
Antineoplastic agents other than methotrexate	During the 180 days up to and including the index date	0 = Absent 1 = Present
Systemic antivirals	During the 180 days up to and including the index date	0 = Absent 1 = Present
Aspirin and antiplatelets other than clopidogrel	During the 180 days up to and including the index date	0 = Absent 1 = Present
Anticoagulants	During the 180 days up to and including the index date	0 = Absent 1 = Present
Number of GLD classes ^b used within the 12 months before the index date	During the 365 days before but not including the index date	0, 1-2, 3-4, 5-8

Medication category	Evaluation period	Coding
Number of GLD classes ^b used from 12 to 24 months before the index date	During the 12-24 months before but not including the index date	0, 1-2, 3-4, 5-8, not applicable ^c
Number of GLD classes ^b used more than 24 months before the index date	More than 24 months before but not including the index date	0, 1-2, 3-4, 5-8, not applicable ^c
Type of index therapy	During these time intervals: • <i>Interval 1</i> = the 90 days before (and not including) the index date • <i>Interval 2</i> = study drug on the index date	• Index monotherapy • Combined index therapy with no prior treatment • Add-on index therapy • Switched-to index therapy • Add-on and switched-to index therapy • Non-evaluable^d

DPP-4 = dipeptidyl peptidase-4; GLD = glucose-lowering drug; GLP-1 = glucagon-like peptide-1; NSAID = nonsteroidal anti-inflammatory drug.

^a Does not include inhaled corticosteroids.

^b GLD classes that were considered: insulin, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, GLP-1 receptor agonists, biguanides (metformin), alpha glucosidase inhibitors, and meglitinides.

^c For patients with insufficient lookback time.

^d For patients with insufficient follow-up time in the study to assess the 90-day add-on/switch requirement.

Table S3. Lifestyle, Demographic, and Health Care Utilization Variables at Baseline

Lifestyle and other variables	Evaluation period	Coding
Age groups (years) ^{a, b}	At the index date	18-34 35-44 45-54 55-64 65-74 75+
Age (years) ^{a, b}	At the index date	Mean (SD) Median (IQR) Min, max
Sex ^b	At the index date	1 = Male 2 = Female
Calendar year of the index date ^b	At the index date	2013 2014 2015 2016 2017 2018
Geographic region (UK) ^c	At the index date	England Northern Ireland Scotland Wales
Geographic region (US) ^d	At the index date	Midwest Northeast South West
Time since approval of dapagliflozin, years	At the index date	Mean (SD) Median (IQR) Min, max
Duration of time before the index date ^e	Number of days from cohort entry date up to and including the index date	Mean (SD) Median (IQR) Min, max
Body mass index (BMI) (using recorded BMI variable or data on height and weight) ^c	During all available history (including the index date); used the most recent value recorded	1 = < 20 (underweight) 2 = 20 to < 25 (normal) 3 = 25 to < 30 (overweight) 4 = 30 to < 40 (obese) 5 = ≥ 40 (severely obese) 6 = Unknown/missing
Smoking history ^c	During all available history (including the index date); used the most recent value recorded	1 = Current 2 = Former 3 = Non-smoker 4 = Unknown/missing

Lifestyle and other variables	Evaluation period	Coding
Alcohol use ^c	During all available history (including the index date); used the most recent value recorded	1 = Non-drinker 2 = Low to moderate intake (1-6 units/week) 3 = Heavy or very heavy intake (7+ units/week) 4 = Drinker, unknown quantity 5 = Unknown/missing
History of alcohol abuse ^c	During all available history (including the index date)	0 = No 1 = Yes
Socioeconomic status		
Index of multiple deprivation ^c	During all available history (including the index date); used the most recent value recorded	Quintiles: Q1 (least deprived) through Q5 (most deprived)
Low-income subsidy status ^f	At the index date	0 = No 1 = Yes
Race/ethnicity ^f	At the index date	1 = White 2 = Black 3 = Asian 4 = Hispanic 5 = Other 6 = Unknown
HbA1c, % ^c	The most recent recorded laboratory value in the 180 days before the index date	< 7.0 ($\leq$ 53 mmol/mol) 7.0-9.9 (53-85 mmol/mol) $\geq$ 10.0 ($\geq$ 86 mmol/mol) Unknown
HbA1c test ^d	Number of HbA1c measurements in the 180 days before the index date	0, 1, 2, 3+
Health care utilisation		
Number of outpatient visits ^g	In the 180 days before but not including the index date	0, 1, 2-3, 4+
Number of hospitalisations	In the 180 days before but not including the index date	0, 1+
Number of emergency department visits	In the 180 days before but not including the index date	0, 1, 2+
Number of speciality care visits	In the 180 days before but not including the index date	0, 1, 2-3, 4+

BMI = body mass index; CPRD = Clinical Practice Research Datalink; GP = general practitioner; HbA1c = glycated hemoglobin; HIRD = HealthCore Integrated Research Database; IQR = interquartile range; Q1, Q5 = first and fifth quintiles; SD = standard deviation; UK = United Kingdom; US = United States.

^a CPRD included age groups 18 years or older all age groups; the HIRD included age groups 18-64 years; Medicare included age groups 65-74 and 75+ years.

^b Matching variable.

^c CPRD only.

^d The HIRD and Medicare only.

^e CPRD: Duration of continuous registration in an up-to-standard practice before the index date; the HIRD: Duration of continuous enrolment in medical and pharmacy coverage before the index date; Medicare: Duration of continuous enrolment (Parts A, B, and D) before the index date.

^f Medicare only.

^g CPRD: visits to GP or outpatient hospital clinic; the HIRD and Medicare: all non-inpatient visits.

Table S4. Results of the Cohort Selection Process: Counts of Treatment Episodes After Exclusions and Final Matched Cohorts for the Combined Dapagliflozin and Comparator GLD Exposure Groups

Type of exclusion	CPRD		HIRD		Medicare	
	Dapagliflozin	Comparator GLD	Dapagliflozin	Comparator GLD	Dapagliflozin	Comparator GLD
<i>Total number of potential treatment episodes during the study period</i>	14,881	114,639	46,581	1,014,001	50,262	2,424,555
CPRD-only exclusions:						
Prior to practice up-to-standard date	0	0	NA	NA	NA	NA
Aged < 18 years	0	25	NA	NA	NA	NA
< 180 days of lookback	893	14,343	NA	NA	NA	NA
HIRD-only exclusions:						
Aged < 18 or ≥ 65 years	NA	NA	4,663	215,794	NA	NA
< 180 days of lookback	NA	NA	7,963	269,275	NA	NA
Medicare-only exclusions:						
Aged < 65 years	NA	NA	NA	NA	1,178	47,798
Enrolled because of disability or end-stage renal disease	NA	NA	NA	NA	≈ 20 ^a	9,661
Not a resident of a US state or District of Columbia at the index date	NA	NA	NA	NA	41	4,992
< 6 months of enrollment	NA	NA	NA	NA	15,554	828,088
Index date coincided with the last date meeting Medicare enrollment criteria (i.e., patient had no follow-up time)	NA	NA	NA	NA	< 11 ^b	394
<i>Potential treatment episodes meeting the inclusion criteria after applying the data source-specific exclusions</i>	13,988	100,271	33,955	528,932	33,470	1,533,622
Diagnosis of type 1 diabetes mellitus on or before the index date	677	3,411	3,347	40,925	2,941	141,937

Type of exclusion	CPRD		HIRD		Medicare	
	Dapagliflozin	Comparator GLD	Dapagliflozin	Comparator GLD	Dapagliflozin	Comparator GLD
Dapagliflozin use before the study period	0	0	0	0	0	0
Other SGLT2 inhibitor medication use on or before the index date	174	12,039	6,952	142,609	5,132	215,909
Comparator GLD use starts during dapagliflozin exposure period	0	11	0	7,207	0	5,713
<i>Potential treatment episodes meeting the inclusion criteria after applying the diabetes-related exclusions</i>	13,137	84,810	23,656	337,226	25,397	1,170,063
Diagnosis of acute kidney injury in the 180 days before and including the index date	34	1,213	130	5,214	517	75,300
Chronic kidney disease before and including the index date	1,041	17,327	2,351	37,664	6,801	394,198
<i>Final number of treatment episodes eligible for cohort selection</i>	12,062	66,270	21,175	294,348	18,079	700,565
<i>Final number of treatment episodes selected into the cohort^c</i>	12,051	43,296	21,173	239,421	18,079	271,180

CMS = Center for Medicare and Medicaid Services; GLD = glucose-lowering drug; NA = not applicable; SGLT2 = sodium-glucose cotransporter 2; US = United States.

^a Approximate value reported to prevent the derivation of reportable values in other cells.

^b According to CMS policy, any cell with a value of 1 to 10 cannot be reported.

^c Final number of treatment episodes selected after matching and before propensity score trimming.

Table S5. Selected Baseline Characteristics, Analysis Sample After Propensity Score Trimming

	CPRD		HIRD		Medicare	
	Dapagliflozin (n = 10,341)	Comparator GLD (n = 36,276)	Dapagliflozin (n = 18,777)	Comparator GLD (n = 215,908)	Dapagliflozin (n = 15,368)	Comparator GLD (n = 236,576)
Age (years) ^a						
Mean (SD)	57.1 (10.4)	57.8 (10.9)	51.6 (8.6)	51.5 (8.8)	70.6 (5.0)	70.6 (5.0)
Sex, n (%)						
Male	6,029 (58.3)	21,444 (59.1)	10,107 (53.8)	114,973 (53.3)	7,601 (49.5)	117,588 (49.7)
Female	4,312 (41.7)	14,832 (40.9)	8,670 (46.2)	100,935 (46.7)	7,767 (50.5)	118,988 (50.3)
Race/ethnicity, ^b n (%)						
Asian	NA	NA	NA	NA	717 (4.7)	10,106 (4.3)
Black	NA	NA	NA	NA	989 (6.4)	19,278 (8.1)
Hispanic	NA	NA	NA	NA	591 (3.8)	9,359 (4.0)
White	NA	NA	NA	NA	12,302 (80.0)	184,443 (78.0)
Other ^c	NA	NA	NA	NA	372 (2.4)	7,479 (3.2)
Unknown	NA	NA	NA	NA	397 (2.6)	5,911 (2.5)
Body mass index (kg/m ²), ^d n (%)						
< 20 (underweight)	19 (0.2)	177 (0.5)	NA	NA	NA	NA
20 to < 25 (normal)	364 (3.5)	2,777 (7.7)	NA	NA	NA	NA
25 to < 30 (overweight)	2,070 (20.0)	9,542 (26.3)	NA	NA	NA	NA
30 to < 40 (obese)	5,654 (54.7)	17,429 (48.0)	NA	NA	NA	NA
40+ (severely obese)	2,059 (19.9)	5,306 (14.6)	NA	NA	NA	NA
Unknown	175 (1.7)	1,045 (2.9)	NA	NA	NA	NA
Comedications in the 180 days before or on the index date, n (%)						
Antihypertensives / diuretics	7,209 (69.7)	24,540 (67.6)	12,665 (67.4)	140,497 (65.1)	12,344 (80.3)	189,164 (80.0)
Lipid-modifying agents	7,985 (77.2)	27,123 (74.8)	11,421 (60.8)	122,803 (56.9)	11,744 (76.4)	175,555 (74.2)

	CPRD		HIRD		Medicare	
	Dapagliflozin (n = 10,341)	Comparator GLD (n = 36,276)	Dapagliflozin (n = 18,777)	Comparator GLD (n = 215,908)	Dapagliflozin (n = 15,368)	Comparator GLD (n = 236,576)
Insulin use at the index date, n (%)	1,053 (10.2)	1,716 (4.7)	2,456 (13.1)	22,507 (10.4)	2,303 (15.0)	28,376 (12.0)
Type of index therapy, ^e n (%)						
Index monotherapy with no prior treatment	209 (2.0)	655 (1.8)	1,564 (8.3)	20,202 (9.4)	1,676 (10.9)	18,345 (7.8)
Combined index therapy with no prior treatment	143 (1.4)	945 (2.6)	1,428 (7.6)	21,654 (10.0)	829 (5.4)	25,523 (10.8)
Add-on index therapy	6,432 (62.2)	25,806 (71.1)	12,632 (67.3)	135,701 (62.9)	8,668 (56.4)	127,632 (53.9)
Switched-to index therapy	428 (4.1)	2,479 (6.8)	428 (2.3)	4,341 (2.0)	1,157 (7.5)	21,229 (9.0)
Add-on and switched-to index therapy	2,770 (26.8)*	5,142 (14.2)*	1,463 (7.8)	11,453 (5.3)	2,277 (14.8)	31,738 (13.4)
Nonevaluable ^f	359 (3.5)	1,249 (3.4)	1,262 (6.7)	22,557 (10.4)	761 (5.0)	12,109 (5.1)
Indicators of diabetes severity, n (%)						
Retinopathy	2,759 (26.7)	8,896 (24.5)	4,376 (23.3)	45,394 (21.0)	5,318 (34.6)	71,901 (30.4)
Peripheral neuropathy	311 (3.0)	987 (2.7)	344 (1.8)	3,776 (1.7)	844 (5.5)	12,275 (5.2)
Peripheral vascular disease ^g	307 (3.0)	1,215 (3.3)	4,197 (22.4)	44,186 (20.5)	5,318 (34.6)	74,543 (31.5)
Coronary heart disease	1,269 (12.3)	4,406 (12.1)	1,733 (9.2)	19,736 (9.1)	4,538 (29.5)	67,315 (28.5)
Cerebrovascular disease	443 (4.3)	1,898 (5.2)	311 (1.7)	4,022 (1.9)	1,860 (12.1)	27,501 (11.6)
Amputation	75 (0.7)	315 (0.9)	47 (0.3)	612 (0.3)	64 (0.4)	1,471 (0.6)
Number of GLD classes ^h used within 12 months ⁱ before the index date						
0	88 (0.9)	525 (1.4)	2,693 (14.3)	38,895 (18.0)	1,101 (6.1)	23,793 (8.8)
1-2	7,094 (68.6)*	31,612 (87.1)*	13,684 (72.9)	162,426 (75.2)	10,550 (58.4)*	202,744 (74.8)*
3-4	3,127 (30.2)*	4,099 (11.3)*	2,398 (12.8)*	14,574 (6.8)*	6,186 (34.2)*	43,686 (16.1)*
5-8	32 (0.3)	40 (0.1)	NR	13 (< 0.1)	242 (1.3)	957 (0.4)

	CPRD		HIRD		Medicare	
	Dapagliflozin (n = 10,341)	Comparator GLD (n = 36,276)	Dapagliflozin (n = 18,777)	Comparator GLD (n = 215,908)	Dapagliflozin (n = 15,368)	Comparator GLD (n = 236,576)
Healthcare utilization in the 180 days before the index date						
Number of outpatient visits, ^j n (%)						
0	431 (4.2)	1,399 (3.9)	270 (1.4)	4,593 (2.1)	763 (5.0)	14,671 (6.2)
1	864 (8.4)	3,043 (8.4)	716 (3.8)	9,507 (4.4)	972 (6.3)	16,972 (7.2)
2 or more	9,046 (87.5)	31,834 (87.8)	17,791 (94.7)	201,808 (93.5)	13,633 (88.7)	204,933 (86.6)
Number of hospitalizations, n (%)						
0	9,400 (90.9)	32,749 (90.3)	18,222 (97.0)	207,781 (96.2)	14,704 (95.7)	224,488 (94.9)
1	685 (6.6)	2,471 (6.8)	497 (2.6)	7,247 (3.4)	498 (3.2)	8,880 (3.8)
2 or more	256 (2.5)	1,056 (2.9)	58 (0.3)	880 (0.4)	166 (1.1)	3,208 (1.4)

CPRD = Clinical Practice Research Datalink; GLD = glucose-lowering drug; HIRD = HealthCore Integrated Research Database; NA = not applicable; NR = not reportable due to small cell number; SD = standard deviation.

* Absolute standardized difference (StDiff) > 0.20.

Note: For HIRD data, any cell with a count of 1-10 or any cell that allows a count of 1-10 to be derived from other reported cells or information cannot be reported.

^a Eligible patients were aged 18 years or older in CPRD, 18-64 years in the HIRD, and 65 years or older in Medicare.

^b Data on race/ethnicity were available only in Medicare.

^c Includes patients categorized as Other or as North American Native in Medicare.

^d Data on body mass index were available only in CPRD.

^e Detailed definitions for the index therapy type categories are provided in Section 1.4.

^f Patients who did not have sufficient follow-up time to assess the 90-day add-on/switch requirement.

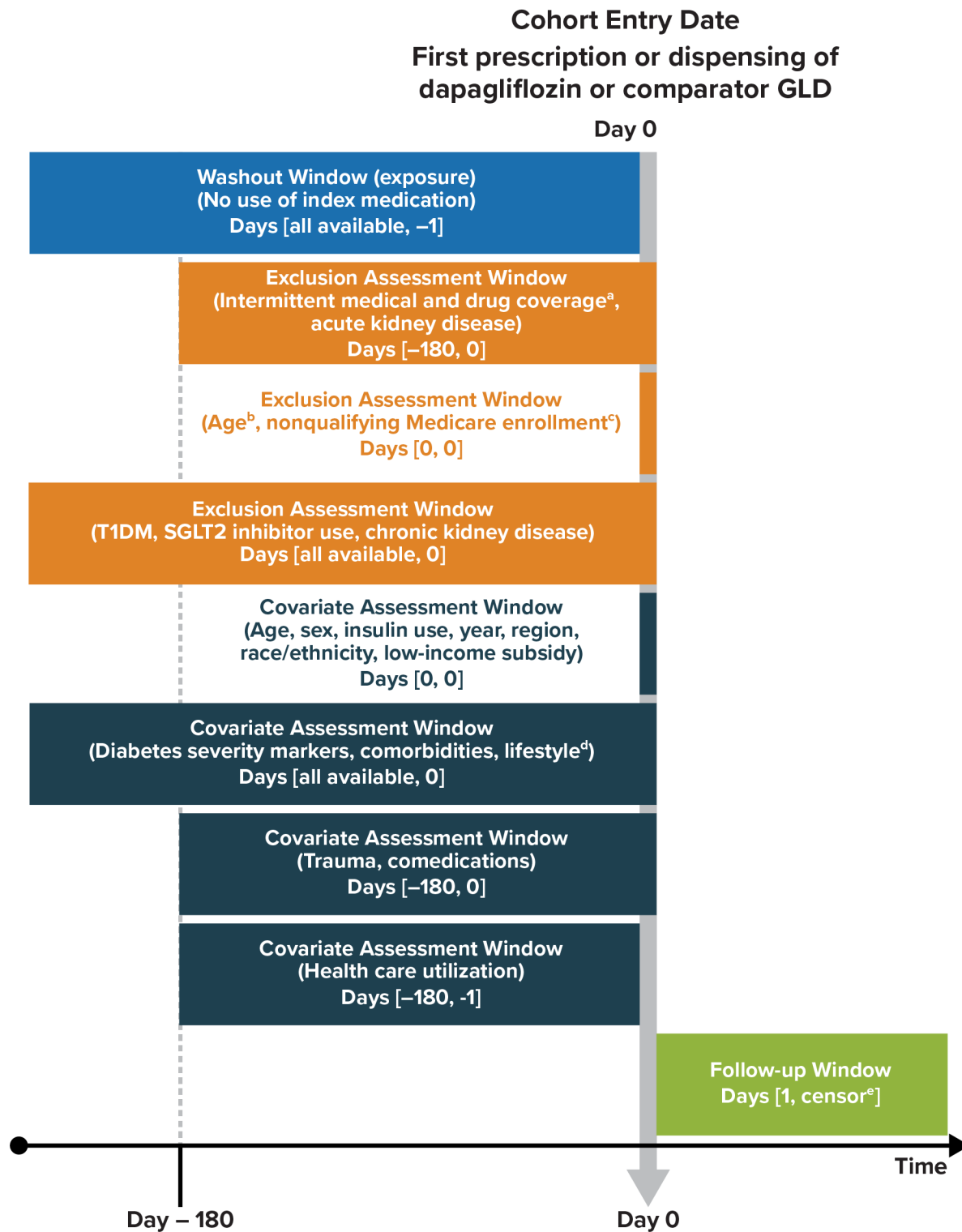
^g Includes peripheral artery disease.

^h Glucose-lowering drug classes that were considered were insulin, sulfonylureas, thiazolidinediones, dipeptidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide-1 (GLP-1) receptor agonists, biguanides (metformin), alpha glucosidase inhibitors, and meglitinides.

ⁱ Those with at least 180 days of available lookback data before the index date were eligible for inclusion in the study, and therefore some patients had less than 12 months of available lookback data.

^j Outpatient visits included general practitioner and outpatient hospital visits.

Figure S1. Study Design, Cohort Eligibility, and Inclusion Criteria



CPRD = Clinical Practice Research Datalink; GLD = glucose-lowering drug; hAKI = hospitalization for acute kidney injury;
HIRD = HealthCore Integrated Research Database; SGLT2 = sodium-glucose cotransporter 2; T1DM = type 1 diabetes mellitus.

Note: Schematic based on the framework of graphical representation for visualizing longitudinal study designs proposed by Schneeweiss et al¹. The lighter blue box represents the assessment window for use of the index study medication, the orange boxes represent the assessment windows for the exclusion criteria, the darker blue boxes represent the assessment windows for covariate variables, and the green box represents the follow-up period.

^a CPRD, registered in an up-to-standard participating general medical practice. HIRD, complete pharmacy and medical coverage in a health insurance plan with no enrollment gaps greater than 30 days; Medicare, enrolled in fee-for-service insurance in Parts A (hospital insurance), B (medical insurance), and D (prescription drug coverage).

^b CPRD, ages < 18 years; HIRD, ages < 18 or > 64 years; Medicare, ages < 65 years.

^c Medicare, enrolled because of disability or end-stage renal disease; nonresident of a US state or the District of Columbia; enrolled in managed care coverage.

^d Body mass index, smoking history, and alcohol use were available in CPRD GOLD data but were not available in HIRD or Medicare.

^e Discontinuation of the study medication (30 days after the end of the days' supply of the last consecutive prescription or dispensing), hAKI event, death, end of patient-specific data or eligibility in each data source or end of the study period, initiation of an SGLT2 inhibitor, diagnosis of T1DM.

PART B. Supplemental Methods and Results

1 Methods and Results

1.1 Propensity Score Modeling Approach

In each data source, propensity score models were built to estimate the probability of an individual receiving dapagliflozin versus a comparator glucose-lowering drug (GLD). To select covariates for the propensity score models from a list of preidentified potential confounders (potential covariates are listed in Tables S1-S3), separate Cox proportional hazards regression models were fit for each potential covariate to evaluate the covariate's influence on the association of dapagliflozin exposure with the outcome hospitalization for acute kidney injury. Across all data sources, the base set of covariates for these Cox models was age at the index date, sex, duration of lookback time, primary care practice or geographic region, an indicator variable for whether the index medication was "added on" or "switched to" (detailed definitions for the index therapy type are described in Section 1.4), an indicator variable for insulin use at the index date, and calendar year of the index date. All remaining potential candidate variables (medical conditions, medications, and lifestyle and healthcare resource utilization variables) were included in separate base models. Variables from the Cox models for which the resulting treatment-related hazard ratio (HR) met one of two conditions—(1) change in the absolute value of the HR estimate of more than 0.005 or (2) change in the value of the HR estimate of more than 0.05%—were selected for inclusion in the final propensity score model. Variables could also be selected for the propensity score model based on clinical relevance. Propensity scores were then calculated for each treatment episode by fitting a multivariate logistic regression model with exposure as the dependent variable (0 = comparator GLD initiator, 1 = dapagliflozin initiator), the base set of covariates, and all other data source-specific covariates that were identified in the variable selection process as independent variables. The variables included in each data source-specific propensity score model are displayed in each of the standardized differences plots presented in the main manuscript (Fig 1, 2, and 3).

1.2 Estimation of Pooled Adjusted Incidence Rate Ratio

After estimating the adjusted incidence rate ratio (IRR) of hospitalization for acute kidney injury (hAKI) within each data source, the adjusted IRR estimates were pooled across the data sources to generate an overall (i.e., pooled) adjusted IRR estimate. First, the data source-specific IRR estimates were evaluated for homogeneity by examining the value and direction of the IRRs and their corresponding 95% confidence intervals (CIs). Next, within each data source, event counts and person-time were aggregated by exposure category and data source-specific propensity score stratum. Using Mantel-Haenszel methods,² the stratum-specific estimates were pooled across all data sources.

The I^2 index³ was used to assess statistical heterogeneity between the data source-specific estimates, i.e., the amount of between-study variation in the IRRs, and thus the appropriateness of pooling estimates from the data sources. For this study, if the calculated I^2 index was below 50%, the pooled Mantel-Haenszel-adjusted IRR was reported.

1.3 Assessment of the Potential Effect of Unmeasured Confounders (Quantitative Bias Analysis Methods, Results, and Interpretation)

A quantitative bias analysis was performed to evaluate the effect of a potential unmeasured confounding variable on the observed IRR estimates using the method described by Lash et al.⁴ Three bias parameters were defined for this analysis: (1) the expected association between the unmeasured confounder and the IRR, (2) the prevalence of the unmeasured confounder among those exposed to dapagliflozin, and (3) the prevalence of the unmeasured confounder among those without dapagliflozin exposure. Using the three bias parameters, the following formula was used to calculate the overall, pooled association between dapagliflozin and the outcome of interest expected after adjustment for the unmeasured confounder:

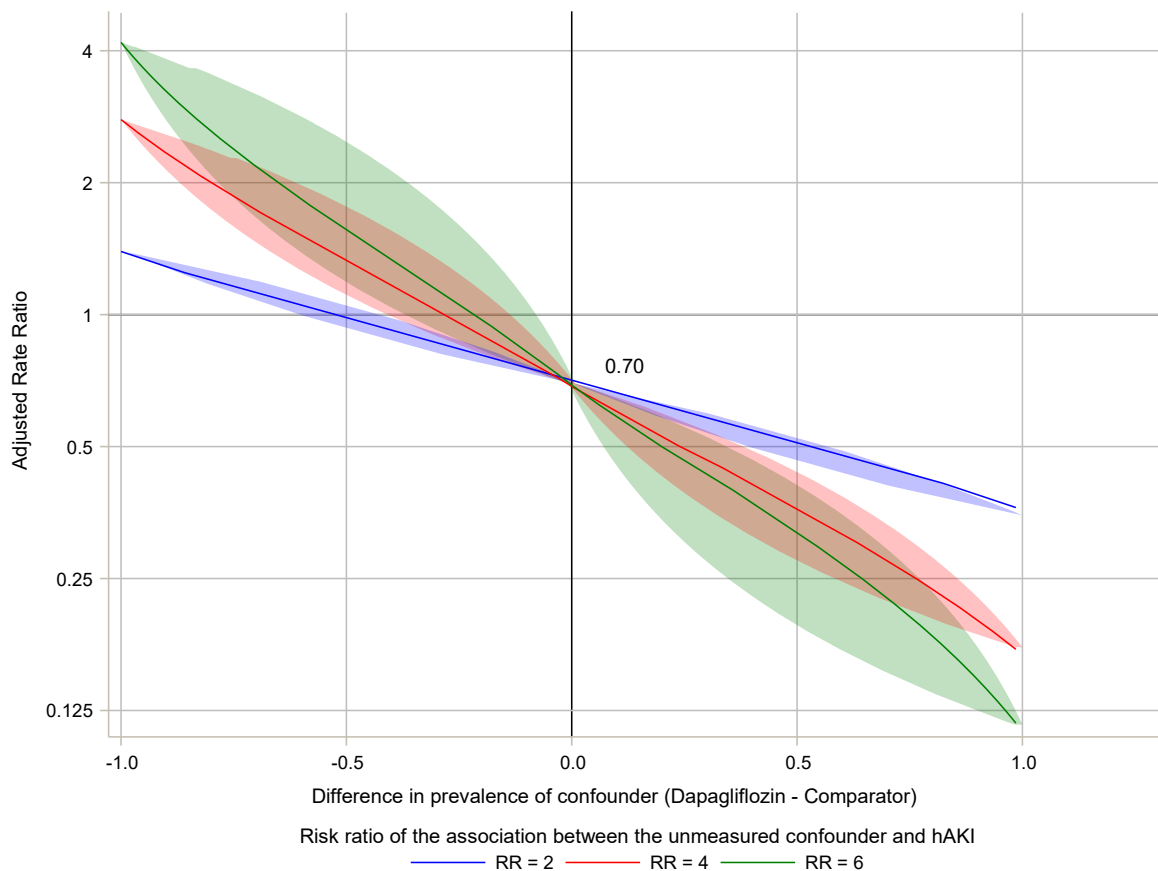
$$IRR_{adj} = IRR_{obs} (IRR_{CD\rho_0} + [1 - \rho_0] / IRR_{CD\rho_1} + [1 - \rho_1]),$$

where IRR_{adj} is the IRR associating dapagliflozin exposure with the outcome adjusted for the unmeasured confounder, IRR_{obs} is the observed IRR without adjustment for the unmeasured confounder, IRR_{CD} is the IRR associating the unmeasured confounder with the disease, and ρ_1 and ρ_0 are the estimated proportions of patients with the unmeasured confounder in the exposed and unexposed groups, respectively.⁴

The IRRs were estimated under multiple variations of two different scenarios: (1) ranging the association of the unmeasured confounder with the outcome from a risk ratio (RR) = 1.5 to RR = 4.5 and (2) varying the difference in prevalence of the confounder in each treatment group from -100% (i.e., present in every patient in the comparator GLD group but not present in any patient in the dapagliflozin group) to 100% (i.e., present in every patient in the dapagliflozin group but not present in any patient in the comparator GLD group).

The resulting IRRs for the pooled, overall analysis were plotted under these varying assumptions of unmeasured confounding compared with the observed IRR estimate from the pooled analysis (Figure S2).

Figure S2. Sensitivity Analysis, Adjusted Incidence Rate Ratios Under Varying Assumptions of Unmeasured Confounding Compared With the Observed Incidence Rate Ratio Estimate From the Pooled Analysis



hAKI = hospitalization for acute kidney injury (the study outcome); RR = risk ratio.

As no specific unmeasured confounder was being evaluated, three scenarios of the association of hypothetical unmeasured confounders with hAKI were evaluated and displayed in the figure—risk ratio (RR) = 1.5 (moderate association, in blue), RR = 3.0 (strong association, in red), and RR = 4.5 (very strong association, in green). An underlying assumption is that these confounder relationships were independent and unassociated with other, measured confounders. For a factor to have been a confounder of the relationship between dapagliflozin and hAKI, it must have been imbalanced between the two index treatment groups. For each hypothetical confounder scenario,

we calculated IRRs corrected for the unmeasured confounder with imbalances ranging from –100% (present in every comparator GLD patient and not present in any dapagliflozin patient) to 100% (present in every dapagliflozin patient and not present in any comparator GLD patient). In Figure S2, the IRR values at 0% imbalance in each scenario are the observed IRRs estimated in the overall, pooled analysis, as it assumed no imbalance of the unmeasured confounder, and thus no confounding. For each level of covariate imbalance, an IRR corrected for the hypothetical unmeasured confounder was calculated based on assumptions regarding the strength of the confounder in that scenario and the confounder prevalences in the two treatment groups.

A given covariate imbalance may be achieved through numerous combinations of confounder levels (e.g., an imbalance of 10% can result from confounder prevalences of 10% in dapagliflozin patients and 0% in comparator GLD patients or prevalences of 90% in dapagliflozin patients and 80% in comparator GLD patients). In Figure S2, the colored bands for each confounding scenario represent the minimum and maximum possible corrected IRRs at each level of imbalance, and the solid line represents the mean corrected IRR at each imbalance level.

In the worst-case scenario of having a hypothetical moderate confounder ($RR = 1.5$) in which groups would be completely imbalanced (0% prevalence in the dapagliflozin group and 100% prevalence in the comparator GLD group), the maximum true IRR would be 1.05; any imbalance less extreme would result in IRRs lower than 1.05. A hypothetical moderate confounder ($RR = 1.5$) would require an imbalance of at least approximately –90% (i.e., higher prevalence in the comparator GLD patients) to mask a true IRR greater than 1.00. If the hypothetical unmeasured confounder had a stronger independent relationship with the outcome, $RR = 3.0$ or 4.5 , then a smaller imbalance would be required.

For context, in the CPRD cohort, current smoker status had an imbalance of only 2.1% (16.2% in dapagliflozin users, 18.3% in comparator GLD users) in the full, untrimmed sample. Smoking is a risk factor for kidney diseases; current smoking has a meta-analyzed association with diabetic nephropathy of odds ratio (OR) of 1.62 or 1.70, depending on the analysis⁵; incident chronic kidney disease, OR of 1.34; and incident end-stage kidney disease, OR of 1.91.⁶ Similarly, male sex has been reported to be associated with increased risk of acute kidney injury in the general population, with IRR estimates of approximately 1.3 to 1.7⁷; in all three data sources in the present study, the imbalance of sex was less than 2%.

Strong risk factors for acute kidney injury include more acute conditions such as critical illnesses, sepsis, hypovolemia, receipt of intravenous contrast agents, trauma, surgical procedures, and preexisting kidney disease.⁸⁻¹⁰ However, patients with chronic kidney disease were excluded at baseline. Additionally, in all data sources, trauma was well balanced in the full cohort, and the proportion of patients with prior hospitalization was low, suggesting that the prevalence of a hypothetical potentially confounding hospitalized procedure or condition would not be high enough to meaningfully impact the results if it was not included in the analysis.

1.4 Definitions for Index Therapy Type Categories

Table 3 and Table S5 present the frequency distributions of the index therapy type in the full samples (i.e., before propensity score trimming) and in the propensity score-trimmed samples, respectively. The index medication could be initiated as monotherapy, added to another GLD, switched from another GLD to the index medication, or initiated as index combined therapy (more than one drug was initiated on the index date). For creation of the index therapy type categories, three intervals of time were considered: *interval 1* = the 90 days before (and not including) the index date; *interval 2* = study drug index date; *interval 3* = the 90 days after (and not including) the index date.

The following categories of index medication exposure were created based on individual drugs (i.e., the drug substance), not by drug class:

Index monotherapy with no prior treatment. Only a single index drug substance was prescribed or dispensed on the index date (interval 2), and there was no prescription or dispensing for a GLD or insulin in interval 1. Note that interval 3 could be less than 90 days and the definition of index monotherapy would still apply.

Index combined therapy with no prior treatment. Multiple drug substances were prescribed or dispensed at interval 2, and there was no prescription or dispensing for any GLD or insulin in interval 1. Note that interval 3 could be less than 90 days and the definition of index combined therapy would still apply.

Add-on index therapy. A GLD or insulin other than the index medication was prescribed or dispensed during interval 1, and then a subsequent prescription or dispensing for the same drug substance was identified during interval 2 or 3. If multiple drug substances were identified during interval 1, then *all* these drug substances would need a new prescription or dispensing during interval 2 or 3 to fit this category.

Switched-to index therapy. A drug substance(s) other than the index GLD was prescribed or dispensed during interval 1 and had no subsequent prescriptions or dispensings during intervals 2 and 3. If multiple drug substances were identified in interval 1, no additional prescriptions or dispensings for any of these substances could occur in intervals 2 and 3.

Add-on and switched-to index therapy. A patient had multiple drug substances with a reported prescription or dispensing in interval 1 and the following two criteria were met:

- At least one drug substance had a prescription or dispensing during interval 2 or interval 3.
- At least one drug substance had no prescription or dispensing during interval 2 and interval 3.

Non-evaluable index treatment. A drug substance(s) was prescribed or dispensed during interval 1 and the patient had less than 90 days of follow-up; therefore, it could not be determined if there was an add-on or a switch or both. Note that an outcome of interest was not allowed to truncate the patient's record for the assignment of medication type.

2 References

1. Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, Happe L, Arlett P, et al.: Graphical Depiction of Longitudinal Study Designs in Health Care Databases. *Ann Intern Med* 170: 398-406, 2019
2. Rothman KJ, Greenland S, Lash TL, editors: *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008.
3. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al.: Cochrane handbook for systematic reviews of interventions, version 6.0. Cochrane. July 2019. Accessed 12 August 2020. <https://training.cochrane.org/handbook/>
4. Lash TL, Fox MP, Fink AK: Unmeasured and unknown confounders. In: *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York, NY: Springer 2009. p. 59-78.
5. Su S, Wang W, Sun T, Ma F, Wang Y, Li J, et al.: Smoking as a risk factor for diabetic nephropathy: a meta-analysis. *Int Urol Nephrol* 49: 1801-1807, 2017
6. Xia J, Wang L, Ma Z, Zhong L, Wang Y, Gao Y, et al.: Cigarette smoking and chronic kidney disease in the general population: a systematic review and meta-analysis of prospective cohort studies. *Nephrol Dial Transplant* 32: 475-487, 2017
7. Grams ME, Sang Y, Ballew SH, Gansevoort RT, Kimm H, Kovesdy CP, et al.: A meta-analysis of the association of estimated GFR, albuminuria, age, race, and sex with acute kidney injury. *Am J Kidney Dis* 66: 591-601, 2015
8. Finlay S, Bray B, Lewington AJ, Hunter-Rowe CT, Banerjee A, Atkinson JM, et al.: Identification of risk factors associated with acute kidney injury in patients admitted to acute medical units. *Clin Med (Lond)* 13: 233-238, 2013

9. Hu J, Chen R, Liu S, Yu X, Zou J, Ding X: Global incidence and outcomes of adult patients with acute kidney injury after cardiac surgery: a systematic review and meta-analysis. *J Cardiothorac Vasc Anesth* 30: 82-89, 2016
10. Skinner DL, Hardcastle TC, Rodseth RN, Muckart DJ: The incidence and outcomes of acute kidney injury amongst patients admitted to a level I trauma unit. *Injury* 45: 259-264, 2014