

# Identifying Risk Factors for Diabetic Ketoacidosis Associated with SGLT2 Inhibitors: a Nationwide Cohort Study in the USA



Michael Fralick, MD, PhD, SM<sup>1,2,3</sup> , Donald A Redelmeier, MD, MSc<sup>3</sup>,  
Elisabetta Patorno, MD, DrPh<sup>1</sup>, Jessica M Franklin, PhD<sup>1</sup>, Fahad Razak, MD, MSc<sup>4,5</sup>,  
Tara Gomes, PhD<sup>3,5</sup>, and Sebastian Schneeweiss, MD, ScD<sup>1</sup>

<sup>1</sup>Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Suite 3030, Boston, MA, USA; <sup>2</sup>Sinai Health System, Department of Medicine, University of Toronto, Toronto, ON, Canada; <sup>3</sup>ICES, Toronto, ON, Canada; <sup>4</sup>St Michael's Hospital, Department of Medicine, University of Toronto, Toronto, ON, Canada; <sup>5</sup>Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, ON, Canada.

## ABSTRACT:

**INTRODUCTION:** Sodium glucose co-transporter-2 inhibitors (SGLT2) are commonly prescribed to patients with type 2 diabetes mellitus, but can increase the risk of diabetic ketoacidosis. Identifying patients prone to diabetic ketoacidosis may help mitigate this risk.

**METHODS:** We conducted a population-based cohort study of adults initiating SGLT2 inhibitor use from 2013 through 2017. The primary objective was to identify potential predictors of diabetic ketoacidosis. Two machine-learning methods were applied to model high-dimensional pre-exposure data: gradient boosted trees and least absolute shrinkage and selection operator (LASSO) regression. We rank ordered the variables produced from LASSO by the size of their estimated coefficient (largest to smallest). With gradient boosted trees, a relative importance measure for each variable is provided rather than a coefficient. The "top variables" were identified after reviewing the distributions of the effect estimates from LASSO and gradient boosted trees to identify where there was a substantial decrease in variable importance. The identified predictors were then assessed in a logistic regression model and reported as odds ratios (ORs) with 95% confidence intervals (CIs).

**RESULTS:** We identified 111,442 adults who started SGLT2 inhibitor use. The mean age was 57 years, 44% were female, the mean hemoglobin A1C was 8.7%, and the mean creatinine was 0.89 mg/dL. During a mean follow-up of 180 days, 192 patients (0.2%, i.e., 2 per 1000) were diagnosed and hospitalized with diabetic ketoacidosis (DKA) and 475 (0.4%, i.e., 4 per 1000) were diagnosed in either an inpatient or outpatient setting. Using gradient boosted trees, the strongest predictors were prior DKA, baseline hemoglobin A1C level, baseline creatinine level, use of medications for dementia, and baseline bicarbonate level. Using LASSO regression not including laboratory test results due to missing data, the strongest predictors were prior DKA, digoxin use, use of medications for dementia, and recent hypoglycemia. The logistic regression model incorporating the variables

identified from gradient boosted trees and LASSO regression suggested the following pre-exposure characteristics had the strongest association with a hospitalization for DKA: use of dementia medications (OR = 7.76, 95% CI 2.60, 23.1), prior intracranial hemorrhage (OR = 11.5, 95% CI 1.46, 91.1), a prior diagnosis of hypoglycemia (OR = 5.41, 95% CI 1.92, 15.3), prior DKA (OR = 2.45, 95% CI 0.33, 18.0), digoxin use (OR = 4.00, 95% CI 1.21, 13.2), a baseline hemoglobin A1C above 10% (OR = 3.14, 95% CI 1.95, 5.06), and baseline bicarbonate below 18 mmol/L (OR 5.09, 95% CI 1.58, 16.4).

**CONCLUSION:** Diabetic ketoacidosis affected approximately 2 per 1000 patients starting to use an SGLT2 inhibitor. We identified both anticipated, e.g., low baseline serum bicarbonate, and unanticipated, e.g., digoxin, dementia medications, risk factors for SGLT2 inhibitor-induced DKA.

J Gen Intern Med 36(9):2601–7  
DOI: 10.1007/s11606-020-06561-z  
© Society of General Internal Medicine 2021

## INTRODUCTION

Sodium glucose co-transporter 2 (SGLT2) inhibitors are commonly used for the treatment of type 2 diabetes mellitus.<sup>1,2</sup> Their primary mechanism of action lowers plasma glucose by inhibiting reabsorption at the nephron.<sup>1,2</sup> Two of the SGLT2 inhibitors, empagliflozin and canagliflozin, also reduce the risk of myocardial infarction, stroke, and cardiovascular mortality.<sup>3,4</sup> Because of the reduction in cardiovascular events, clinical trials are underway to test these medications in patients with cardiovascular disease who do not have diabetes.<sup>5</sup> One trial, DAPA-HF, recently demonstrated a lower risk of cardiovascular death or heart failure with dapagliflozin compared to placebo, regardless of whether the patient had diabetes.<sup>6</sup> As a result, SGLT2 inhibitors may soon be indicated for a much wider patient population which highlights the importance of rare adverse events for this class of medications.

Received July 22, 2020

Accepted December 22, 2020

Published online February 9, 2021

Important adverse events of SGLT2 inhibitors identified in clinical trials included mycotic genital infection and excessive volume depletion.<sup>1,3,4,7</sup> One additional adverse event not initially detected in clinical trials, but subsequently identified after widespread use, was euglycemic diabetic ketoacidosis.<sup>8–10</sup> Diabetic ketoacidosis is typically a complication of type 1 diabetes mellitus rather than type 2 diabetes and can be fatal.<sup>11,12</sup> Because it is uncommon among adults with type 2 diabetes, initial cases of SGLT2 inhibitor-related ketoacidosis were unexpected and a diagnostic challenge for clinicians.<sup>8</sup> In May 2015,<sup>13</sup> the Food and Drug Administration (FDA) released a warning and subsequently updated the drug label to include diabetic ketoacidosis as a side effect of SGLT2 inhibitors.<sup>14</sup>

Four observational studies<sup>8,10,15,16</sup> and two randomized trials<sup>17,18</sup> have subsequently demonstrated that SGLT2 inhibitors are associated with diabetic ketoacidosis. However, these studies have not identified the risk factors for this adverse event. Risk factors for diabetic ketoacidosis generally fall into three categories: non-adherence to insulin, intercurrent illness (e.g., severe infection), and cardiovascular events (e.g., myocardial infarction, stroke). Prior to SGLT2 inhibitors, no specific medications were consistently shown to be associated with diabetic ketoacidosis. Because it is well established that diabetic ketoacidosis with SGLT2 inhibitors is rare, being able to identify risk factors for this could help mitigate this potential patient harm among patients prescribed an SGLT2 inhibitor. This is particularly important due to expanding indications for these medications.<sup>19</sup> The objective of this study was to identify patient-level characteristics associated with an increased risk of diabetic ketoacidosis for patients receiving an SGLT2 inhibitor.

## METHODS

### Study Population

We conducted a population-based, new-user, cohort study using the nationwide US commercial insurance claims database Optum® Clinformatics® Data Mart.<sup>20</sup> This database provides individual-level de-identified data on demographics, healthcare utilization, diagnoses, diagnostic tests and procedures, outpatient laboratory results, and pharmacy dispensing of drugs to over 13 million people in the USA. It has been widely used to understand the safety and effectiveness of medications used in routine care.<sup>20–22</sup>

We included adults with type 2 diabetes mellitus over age 18 who were newly prescribed an SGLT2 inhibitor (empagliflozin, canagliflozin, dapagliflozin) between March 29, 2013 (date of approval of the first SGLT2 inhibitor) and September 30, 2017 (last available database update). Patients with diabetes mellitus type 2 were identified using the International Classification of Diseases, Ninth Revision (ICD-9) and ICD-10 codes. The cohort entry date was the date of the first prescription for an SGLT2 inhibitor. A new user of an SGLT2 inhibitor was defined as an adult without a prior prescription for an SGLT2 inhibitor in the preceding 180 days.

Patients with insufficient baseline data (i.e., less than 180 days of available data) or a diagnosis of type 1 diabetes were excluded. For our primary analysis, we included patients with a prior history of DKA (in a sensitivity analysis, these patients were excluded).

The Brigham and Women's Hospital Institutional Review Board provided ethics approval and a valid data use agreement for the database was in place.

### Cohort Follow-Up

Follow-up began on the day after the first SGLT2 inhibitor prescription was filled and continued until the end of the study period (i.e., September 30, 2017), end of continuous health plan enrollment, first study outcome, discontinuation of SGLT2 inhibitor, 365 days, or death (whichever came first). An SGLT2 inhibitor was considered discontinued if 30 days elapsed after the expiration of the last prescription's supply without being refilled (data censored on day 31 onwards).

### Study Outcomes

The primary objective was to identify predictors of DKA among patients prescribed an SGLT2 inhibitor. A diagnosis of DKA was defined as one of (i) hospitalization with DKA as the primary diagnosis, (ii) hospitalization with DKA as a secondary diagnosis, and (iii) outpatient diagnosis of DKA. The primary analysis was restricted to hospitalizations with DKA. Secondary analyses were restricted to (i) alone expanded and to (i), (ii), or (iii). In all cases, DKA was identified using ICD-9 or ICD-10 codes ([Appendix](#)). Prior studies have utilized claims data to estimate the risk of DKA in various clinical scenarios.<sup>10,23,24</sup> In addition, two prior studies suggest these codes have a specificity of about 92%, a sensitivity of 68%, and a positive predictive value above 90%.<sup>25,26</sup>

### Baseline Covariates

For the primary analysis, covariates were assessed during the 180 days before cohort entry. In a secondary analysis, covariates were assessed in the 60 days before cohort entry since diagnoses immediately preceding the prescription may be especially relevant. Data included chronic medical conditions (e.g., cardiovascular disease), proxies for diabetes severity (e.g., hemoglobin A1C), established risk factors for diabetic ketoacidosis (e.g., prior event, current insulin use, recent infection), healthcare utilization (e.g., recent hospitalization, emergency department visit, or surgery), prescriber information (e.g., endocrinologist, general practitioner), other diabetes medications (e.g., metformin, sulfonylureas) and non-diabetes-related medications (e.g., diuretics, steroids). These covariates were a priori selected based on prior literature, clinical experience, and expert opinion (Table 1).<sup>27,28</sup> Importantly, some of the laboratory measurements (e.g., serum bicarbonate) may appear to be causal intermediates; however, we only included laboratory values that preceded the

Table 1 Baseline Characteristics

	SGLT2 inhibitor use
	(N = 111,442)
Age, mean (sd)	57.3 (11.2)
Sex, female	49,397 (44.3)
Diabetes-related	
Hemoglobin A1C, mean (sd) %*	8.72 (1.80)
Creatinine, mean (sd) mg/dL*	0.89 (0.25)
Hypoglycemia	1088 (1.0)
Prior diabetic ketoacidosis	530 (0.5)
Diabetic nephropathy	9621 (8.6)
Diabetic neuropathy	17,975 (16.1)
Diabetic retinopathy	5071 (4.6)
Chronic kidney disease	3827 (3.4)
Ischemic heart disease	15,007 (13.5)
Ischemic stroke or TIA	1838 (1.6)
Peripheral vascular disease	5324 (4.8)
Other comorbid conditions	
Heart failure	4158 (3.7)
Hypertension	82,960 (74.4)
Dyslipidemia	82,061 (73.6)
Obese or overweight	35,837 (32.2)
Cancer	6331 (5.7)
Delirium	473 (0.4)
Pneumonia	1604 (1.4)
Urinary tract infection	6438 (5.8)
Liver disease	741 (0.7)
Underweight	79 (0.1)
Smoking	10,168 (9.1)
Bicarbonate, mean (sd) mmol/L*	24.1 (3.1)
Diabetes medications	
Metformin	68,715 (61.7)
Insulin	26,925 (24.2)
Glucagon-like peptide-1 analogue	15,283 (13.7)
Dipeptidyl-peptidase 4 inhibitor	20,440 (18.3)
Sulfonylurea	39,616 (35.5)
Other medications	
Angiotensin converting enzyme inhibitor	46,527 (41.7)
Angiotensin receptor blocker	29,625 (26.6)
Beta blocker	26,233 (23.5)
Digoxin	1053 (0.9)
Antiplatelet	7790 (7.0)
Antipsychotics	2525 (2.3)
Dementia medications	689 (0.6)
Oral steroid	10,981 (9.9)
Statin	71,905 (64.5)
Proton pump inhibitors	22,028 (19.8)
Antibiotics	35,409 (31.8)
Healthcare utilization	
Number of outpatient visits, mean (sd)	4.7 (3.7)
Inpatient hospitalization	434 (0.4)
Emergency room visit	11,213 (10.1)
Primary care doctor visit	69,612 (62.5)
Endocrinologist visit	22,776 (20.4)
A1C ordered	92,714 (83.2)
Flu shot	13,238 (11.9)
Pneumococcal vaccine	4865 (4.4)
Surgical procedure	1625 (1.5)

Legend - numbers represent row totals (%) unless otherwise specified; SGLT2 = sodium glucose co-transporter 2; sd = standard deviation; surgical procedure included knee replacement, hip replacement, cholecystectomy, or appendectomy; \* available for approximately 34% of patients

prescription for the SGLT2 inhibitor. Thus, the laboratory values by definition are not on the causal pathway between SGLT2 inhibitor use and diabetic ketoacidosis.

## Statistical Analysis

Each patient had 83 covariates identified during the baseline period before being prescribed an SGLT2 inhibitor. Because

DKA was rare, including all covariates can cause over-fitting in a logistic regression model. Instead, we applied two machine learning techniques for identifying variables that might be associated with SGLT2 inhibitor-related DKA: least absolute shrinkage and selection operator (LASSO) regression and gradient boosted trees. These two approaches were selected because they are two of the most commonly applied supervised machine learning techniques.

LASSO regression can handle high-dimensional data (i.e., a large number of predictors relative to outcomes), even with substantial collinearity among covariates.<sup>29</sup> Because LASSO cannot handle missing data, we included only baseline covariates with complete data (i.e., excluding serum creatinine, serum bicarbonate, serum hemoglobin A1C). We performed LASSO using the glmnet package available in R and standardized the predictors by their individual standard deviation (sd) so that the odds ratios (OR) produced by LASSO were on a consistent scale.

Gradient boosted trees can accommodate missing data, and it has been shown to have good predictive performance across a wide range of problems.<sup>29,30</sup> Since gradient boosting effectively handles missing data, we retained variables with missing data. The model tuning parameters were selected using a grid search of varying number of trees (0 to 12,000), interaction depth (1 or 3), shrinkage factor (0.001, 0.01, or 0.1), and bag fraction (0.4, 0.5) that optimized the standard loss function for classification (i.e., Bernoulli deviance) (Appendix).<sup>29,30</sup>

To quantify the association between potential predictors and the risk of DKA, we included variables identified using LASSO regression and gradient boosted trees in a logistic regression model. We rank ordered the variables produced from LASSO by the size of their estimated coefficient (largest to smallest). With gradient boosted trees, a relative importance (RI) measure for each variable is provided rather than a coefficient.<sup>29,30</sup> The relative importance measure was then sorted from largest to smallest to identify the variables with the largest association with diabetic ketoacidosis.<sup>29,30</sup> We selected the top variables from LASSO (i.e., largest coefficient) and gradient boosted trees (i.e., largest relative importance) as candidates for subsequent logistic regression. The “top variables” were identified after visually reviewing the distributions of the effect estimates from LASSO and gradient boosted trees to identify where there was a substantial decrease in variable importance. There was no specific cut-off or analytic metric used because both approaches would require arbitrary cut-offs and there is no specific literature to indicate this approach is robust. Partial dependence plots were reviewed for the top variables that were continuous to aid in model interpretability.<sup>31</sup> A partial dependence plot displays the marginal effect of the continuous variable on the outcome.<sup>31</sup> An alternative approach is to arbitrarily dichotomize variables (e.g., hemoglobin A1C less than 7%), but doing so decreases statistical power and limits the ability to identify potentially clinically relevant cut-points.

Three predefined sensitivity analyses were performed. First, the original cohort was re-analyzed using only baseline characteristics in the 60 days rather than 180 days before being



prescribed an SGLT2 inhibitor. A shorter baseline period was used under the assumption that perhaps variables identified closer to the index date (i.e., up to 60 days before) might be more relevant than ones identified further from the index date (e.g., up to 180 days before). Second, a new cohort of patients was constructed that excluded any patients with a prior diagnosis of diabetic ketoacidosis. Excluding patients with a prior diagnosis of diabetic ketoacidosis was considered in the event that prior diabetic ketoacidosis is the strongest predictor and thus overshadows other potentially relevant characteristics. Third, an additional cohort of patients was constructed to include those with type 1 diabetes mellitus, since some patients with type 1 diabetes mellitus have been prescribed an SGLT2 inhibitor off-label.

## RESULTS

### Study Population

A total of 111,442 patients satisfied study inclusion and exclusion criteria (Fig. 1). The mean age was 57 years and 44% were female. The mean hemoglobin A1C was 8.7% (sd = 1.8) and the mean creatinine was 0.89 (sd = 0.25). Overall, 62% were prescribed metformin and 24% were prescribed insulin (Table 1). Risk factors for diabetic ketoacidosis were infrequent (e.g., prior diabetic ketoacidosis). Over a mean follow-up of approximately 180 days, 475 patients were diagnosed with diabetic ketoacidosis (inpatient or outpatient).

### Predictors Identified by LASSO Regression

In the LASSO model of hospitalization with a diagnosis of DKA ( $N=192$ ), 54 of the 80 predictors had an odds ratio (OR) of 1, and 4 had an OR close to 1 (i.e., 1.01 to 1.09). Of the remaining 22 variables, the largest predictors were a prior diagnosis of DKA (OR = 4.8), use of dementia medications (OR = 4.2), hypoglycemia (OR = 2.4), digoxin use (OR = 2.0), insulin use (OR = 1.8), diabetic retinopathy (OR = 1.8), heparin use (OR = 1.7), or gastrointestinal bleed (OR = 1.62). Similar findings were observed in analyses restricted to DKA as the primary diagnosis ( $N=125$ ), and analyses expanded to include outpatient diagnoses ( $N=475$ ). Similarly, limiting the baseline time period to 60 days yielded comparable results (Appendix Table 1).

When we restricted the cohort to only patients without a past episode of DKA, the strongest predictors for a hospitalization with DKA were use of dementia medications (OR = 3.7) and use of digoxin (OR = 2.1). When we broadened the cohort to include patients with a diagnosis of type 1 diabetes, the strongest predictors were intracranial hemorrhage (OR = 4.7), use of dementia medications (OR = 4.6), prior DKA (OR = 4.1), and a diagnosis of type 1 diabetes mellitus (OR = 3.3).

### Predictors Identified by Gradient Boosted Trees

The selected model with the lowest Bernoulli deviance had 1257 trees, a lambda of 0.1, interaction depth of 1, and a bag

fraction of 0.4 though differences in the Bernoulli deviance across the tuned model hyper-parameters were small (i.e., generally less than 0.01). In the fitted gradient boosted tree model that included a hospitalization with a diagnosis of DKA ( $N=192$ ), 63 of the 83 predictors had a relative importance of 0 and 9 predictors had a relative importance near 0 (i.e., between 0 and 0.1). Of the remaining 11 variables, the largest predictors were baseline hemoglobin A1C (RI = 55.9), baseline creatinine (RI = 40.0), use of dementia medications (RI = 1.1), prior diagnosis of DKA (RI = 1.0), and serum bicarbonate (RI = 0.6). Similar findings were observed in analyses restricted to hospitalizations with DKA as the primary diagnosis ( $N=125$ ), or in analyses expanded to include outpatient diagnoses ( $N=475$ ). In addition, analyses limited to a baseline time period to 60 days rather than 180 days yielded comparable findings (Appendix Table 2). When we restricted the cohort to only patients without a past episode of DKA, the strongest predictors were serum creatinine (RI = 51.4), hemoglobin A1C (RI = 46.4), and serum bicarbonate (RI = 0.92). When we broadened the cohort to include patients with a diagnosis of type 1 diabetes, the strongest predictors were baseline hemoglobin A1C (RI = 54.5) and baseline creatinine (RI = 28.9).

### Predictors Analyzed Using Logistic Regression

A logistic regression model included the variables that were consistently identified using either LASSO or gradient boosted trees: prior diabetic ketoacidosis, hypoglycemia, digoxin, dementia medications, delirium, intracranial hemorrhage, hemoglobin A1C, creatinine, and bicarbonate (Table 2). The cut-offs for hemoglobin A1C, serum bicarbonate, and creatinine were identified using partial dependency plots that indicated clear transition points in the predicted probability of DKA. Results are also provided for the logistic regression model only including variables in the preceding 60 days (Table 3).

## DISCUSSION

In this study of over 100,000 adults who started on an SGLT2 inhibitor, overall 475 (4 per 1000) were subsequently diagnosed with diabetic ketoacidosis in the inpatient or outpatient setting, with 192 in the inpatient setting (2 per 1000) over a mean follow-up of approximately 180 days. Using machine learning techniques, both anticipated (i.e., prior DKA, low serum bicarbonate, and a hemoglobin A1C above 10%) and unanticipated (i.e., digoxin, dementia medications) risk factors for DKA were identified. These findings were robust across various sensitivity analyses and highlight a role of machine learning for identifying potential risk factors for rare adverse events.

Preventing SGLT2 inhibitor-related DKA is important because it can be life threatening and an easily over-looked diagnosis for several reasons.<sup>8,12</sup> First, DKA is typically associated with type 1 diabetes mellitus, rather than type 2 diabetes.<sup>12,32</sup> Since SGLT2 inhibitors are prescribed to patients with

### REASONS FOR EXCLUSION

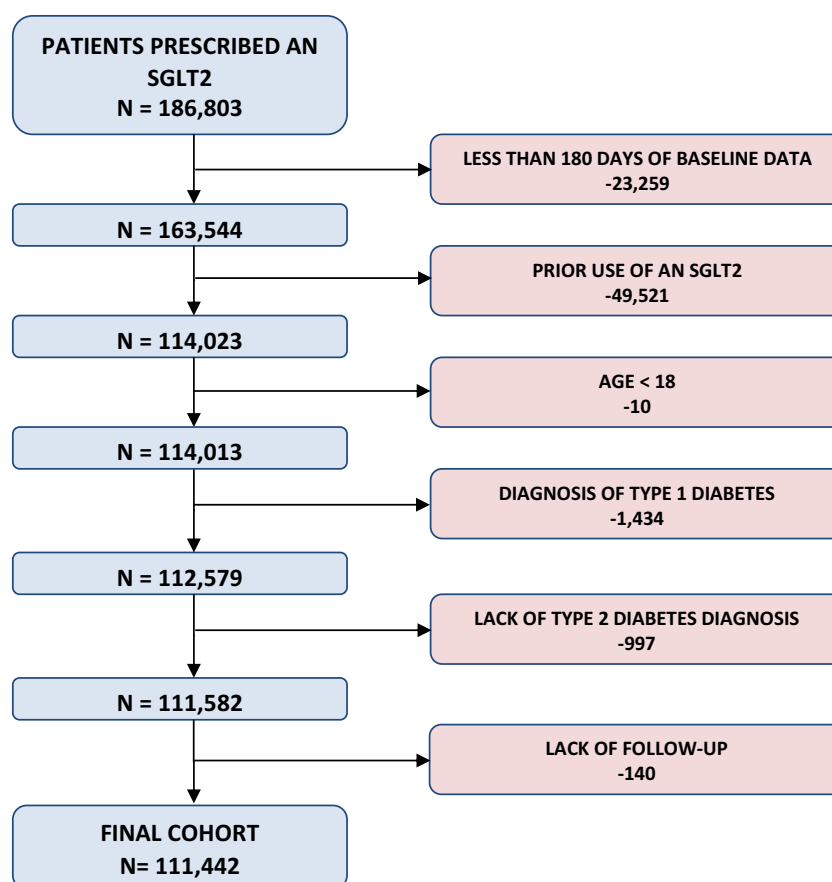


Figure 1 Cohort entry criteria. Legend: SGLT2 = sodium glucose co-transporter 2 inhibitor; all criteria assessed in the preceding 180 days with the exception of a diagnosis of type 2 diabetes which was assessed up to 1000 days prior.

type 2 diabetes mellitus, DKA is not always considered.<sup>5</sup> Second, DKA was traditionally considered to be caused by

profound insulin deficiency rather than an adverse event from medications.<sup>12</sup> Third, patients with SGLT2 inhibitor-related

Table 2 Risk Factors for Diabetic Ketoacidosis Using Variables Identified Up to 180 Days Prior to SGLT2 Inhibitors

	Inpatient diagnosis of DKA			Inpatient or outpatient diagnosis of DKA		
	OR	95% CI		OR	95% CI	
Prior DKA	2.43	0.33	17.82	24.07	14.28	40.59
Hypoglycemia	5.43	1.92	15.32	2.46	0.98	6.17
Hemoglobin	3.16	1.97	5.08	1.89	1.38	2.59
A1C > 10%						
Bicarbonate < 18 mmol/L	5.10	1.58	16.42	2.93	1.17	7.32
Creatinine ≤ 0.5 mg/dL	2.55	0.79	8.17	1.03	0.33	3.26
Creatinine > 1.5 mg/dL	0.50	0.06	3.86	1.54	0.66	3.60
Delirium	3.22	0.68	15.14	2.26	0.75	6.89
Dementia	7.65	2.57	22.76	3.47	1.34	8.96
medication						
Digoxin	3.97	1.20	13.09	5.45	2.71	10.93
Intracranial hemorrhage	11.46	1.45	91.38	5.30	0.71	39.52

SGLT2 = sodium glucose co-transporter 2; OR = odds ratio; CI = confidence interval; DKA = diabetic ketoacidosis; dementia medication = donepezil, memantine, or rivastigmine. For creatinine the reference group was between 0.5 mg/dL and 1.5 mg/dL

Table 3 Risk Factors for Diabetic Ketoacidosis Using Variables Identified Up to 60 Days Prior to SGLT2 Inhibitors

	Inpatient diagnosis of DKA			Inpatient or outpatient diagnosis of DKA		
	OR	95% CI		OR	95% CI	
Prior DKA	5.14	0.63	42.08	56.17	29.54	106.83
Hemoglobin	2.55	1.50	4.32	1.92	1.34	2.73
A1C > 10%						
Bicarbonate < 18 mmol/L	3.91	0.94	16.34	3.00	1.07	8.42
Creatinine ≤ 0.5 mg/dL	3.09	0.96	9.96	1.25	0.39	4.02
Creatinine > 1.5 mg/dL	0.73	0.10	5.54	1.40	0.50	3.93
Acute kidney injury diagnosis	1.65	0.18	15.17	0.45	0.08	2.41
Delirium	4.24	0.43	41.58	5.65	1.10	29.06
Dementia	13.28	4.44	39.7	5.97	2.25	15.85
medication						
Digoxin	4.20	0.97	18.21	4.59	1.78	11.81

Legend: SGLT2 = sodium glucose co-transporter 2; OR = odds ratio; CI = confidence interval; DKA = diabetic ketoacidosis; dementia medication = donepezil, memantine, or rivastigmine. For creatinine the reference group was between 0.5 and 1.5

DKA can have normal or mildly elevated blood glucose levels.<sup>8,9,32</sup> This is atypical for DKA, because patients typically present with blood glucose levels that are five- to tenfold higher than normal. Indeed, many physicians do not initially recognize DKA due to the near-normal glucose levels.<sup>8</sup> Educating clinicians about SGLT2 inhibitor-related DKA may help raise awareness and identify strategies to prevent it from occurring.

We identified prior DKA, digoxin, dementia medications, serum bicarbonate less than 18 mmol/L, and a hemoglobin A1C greater than 10% as some of the important risk factors. Prior DKA, a low serum bicarbonate, and an elevated hemoglobin A1C may seem intuitive since the former suggests a metabolic acidosis may already be present and the latter suggests poorly controlled diabetes. However, many associations in medicine can appear intuitive, but empirical data help to support informed care. Other predictors were surprising (i.e., digoxin, dementia medications, intracranial hemorrhage), and thus, it is unknown whether these findings are spurious, surrogate markers of underlying risk, or potentially directly related to an increased risk of DKA with SGLT2 inhibitors. Of note, both digoxin and SGLT2 inhibitors are a substrate for p-glycoprotein.<sup>33,34</sup> A recent pharmacokinetic study of healthy volunteers identified that digoxin modestly increased serum concentration of empagliflozin which may help explain the increased odds of DKA we observed compared to adults not prescribed digoxin.<sup>33,34</sup> However, digoxin may instead be a surrogate for cardiovascular disease severity, rather than being causally linked to SGLT2 inhibitor-associated diabetic ketoacidosis.

It is unclear why dementia medications might be associated with a higher risk of DKA. SGLT2 inhibitors are generally not metabolized by cytochrome P450 enzymes and instead are eliminated by glucuronidation via UGT1A9 and UGT2B4. Neither donepezil nor memantine should affect glucuronidation, but memantine is eliminated by tubular secretion which could alter plasma levels of SGLT2 inhibitors which are renally cleared and act at the proximal convoluted tubule.<sup>35</sup> An alternative explanation is that the increased risk of DKA might be related to dementia severity, and dementia medication use is a proxy for severity, or that dementia medications are a sign of polypharmacy. It is also unclear why prior intracranial hemorrhage was a seemingly important predictor. It may represent a surrogate of recent hospitalization, illness severity, or perhaps a spurious finding.

Despite observational studies, some clinicians remain skeptical that DKA can be caused by an SGLT2 inhibitor.<sup>36,37</sup> For example, two prior meta-analyses of clinical trials found no association between SGLT2 inhibitors and DKA.<sup>36,37</sup> However, the average number of patients who received an SGLT2 inhibitor was about 500 in each of the trials. Since the overall rate is approximately 3–8 per 1000 person-years, the majority of those trials were underpowered to detect diabetic ketoacidosis. Moreover, our study identified that a hemoglobin A1C above 10% is a strong risk factor yet the majority of trials included excluded patients who had a hemoglobin A1C above 10%.<sup>36,37</sup>

The cardiovascular outcome trial for dapagliflozin ( $N = 17,160$ ) included patients who had a hemoglobin A1C up to 12%, and found a twofold higher rate of DKA (hazard ratio [HR] 2.18, 95% confidence interval [CI] 1.10–4.30).<sup>17</sup> Furthermore, the renal outcome trial for canagliflozin ( $N = 4410$ ) also included patients who had a hemoglobin A1C of up to 12% and also found an increased rate of DKA in patients randomized to canagliflozin (HR = 10.80, 95% CI 1.39–83.65).

Unlike the recent clinical trials identifying an increased risk of DKA with SGLT2 inhibitors, our study lacked diagnostic certainty in identifying DKA. While ICD codes are popular, they are imperfect and can result in misclassification. For example, we observed a higher odds ratio for prior DKA when our outcome definition included an outpatient diagnosis of DKA as opposed to only an inpatient diagnosis of DKA. The higher odds ratio with prior DKA may represent re-recording of prior events rather than a truly new DKA event. For this reason, the results from our models that defined outcomes based on inpatient diagnostic codes might be more accurate. We also lacked complete laboratory data (i.e., only one-third had a baseline hemoglobin A1C) in addition to not having data related to other potential risk factors for DKA including body mass index, dietary intake, alcohol use, and genetic markers.<sup>12</sup> Furthermore, diagnostic codes for variables such as hypoglycemia, delirium, and smoking are imperfect measures and likely underestimate the prevalence of these conditions. Similarly, there were considerable amounts of missing data for laboratory measures, and thus, the mere fact that they were performed may be an indicator of underlying illness severity or concern by the attending physician. These gaps are an important area for future research.

For patients with multiple risk factors, further laboratory monitoring might help to risk stratify these patients. For example, if laboratory testing identifies a low serum bicarbonate level, then these should be worked up accordingly and an SGLT2 inhibitor should be initially avoided. Of course, this recommendation is pragmatic and not formally tested in our study.

## CONCLUSION

SGLT2 inhibitors are an effective class of medications for adults with type 2 diabetes mellitus, but DKA remains an important risk. We applied machine learning methods and identified both anticipated (i.e., prior DKA, low serum bicarbonate, elevated hemoglobin A1C) and unanticipated (i.e., digoxin) risk factors to advance our understanding of SGLT2 inhibitor-related DKA. Additional studies are required to confirm our findings, but patients with multiple risk factors for SGLT2 inhibitor-related diabetic ketoacidosis may benefit from laboratory testing prior to initiation of an SGLT2 inhibitor, closer monitoring for diabetic ketoacidosis, or alternative medications to manage their diabetes.

---

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11606-020-06561-z>.

**Acknowledgments:** We appreciate the feedback from David Dai on an earlier version of this manuscript. He was not compensated for his work.

**Corresponding Author:** Michael Fralick, MD, PhD, SM; ICES, Toronto, ON, Canada (e-mail: mike.fralick@mail.utoronto.ca).

**Funding** The study was funded by the Division of Pharmacoepidemiology and Pharmacoeconomics.

# Compliance with Ethical Standards:

**Conflict of Interest:** EP is an investigator of investigator-initiated grants to the Brigham and Women's Hospital from GSK, not related to the topic of the submitted work. SS is participating in investigator-initiated grants to the Brigham and Women's Hospital from Boehringer Ingelheim unrelated to the topic of this study. He is a consultant to Aetion Inc., 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.

## REFERENCES

1. Wu JHY, Foote C, Blomster J, et al. Effects of sodium-glucose cotransporter-2 inhibitors on cardiovascular events, death, and major safety outcomes in adults with type 2 diabetes: A systematic review and meta-analysis. *Lancet Diabetes Endocrinol*. 2016;4(5):411-419.
2. Tahrani A, Barnett A, Bailey C. SGLT inhibitors in management of diabetes. *Lancet Diabetes Endocrinol*. 2013;1:140-151.
3. Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes. *N Engl J Med*. 2017;377(7):644-657.
4. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015;373(22):2117-2128.
5. Zelniker TA, Braunwald E. Cardiac and renal effects of sodium-glucose co-transporter 2 inhibitors in diabetes. *J Am Coll Cardiol*. 2018;1-11.
6. McMurray JJV, Solomon SD, Inzucchi SE, et al. Dapagliflozin in patients with heart failure and reduced ejection fraction. *N Engl J Med*. 2019;1-13.
7. Fralick M, MacFadden D. A hypothesis for why sodium glucose co-transporter 2 inhibitors have been found to cause genital infection, but not urinary tract infection. *Diabetes Obes Metab*. 2020.
8. Meyer EJ, Gabb G, Jesudason D. SGLT2 inhibitor-associated euglycemic diabetic ketoacidosis: a South Australian clinical case series and Australian spontaneous adverse event notifications. *Diabetes Care*. 2018;(August 2017):dc171721.
9. Ogawa W, Sakaguchi K. Euglycemic diabetic ketoacidosis induced by SGLT2 inhibitors: possible mechanism and contributing factors. *J Diabetes Investig*. 2016;7(2):135-138.
10. Fralick M, Schneeweiss S, Patorno E. Risk of diabetic ketoacidosis after initiation of an SGLT2 inhibitor. *N Engl J Med*. 2017;376(23):2300-2302.
11. Basu A, Close CF, Jenkins D, Krentz AJ, Natrass M, Wright AD. Persisting mortality in diabetic ketoacidosis. *Diabet Med*. 1993;10(3):282-284.
12. Umpierrez G, Korytkowski M. Diabetic emergencies — ketoacidosis, hyperglycaemic hyperosmolar state and hypoglycaemia. *Nat Rev Endocrinol*. 2016;12(4):222-232.
13. Fadini GP, Avogaro A. SGLT2 inhibitors and amputations in the US FDA Adverse Event Reporting System. *Lancet Diabetes Endocrinol*. 2017;5(9):680-681.
14. Center for Drug Evaluation and Research. Drug safety and availability - FDA drug safety communication. <https://www.fda.gov/Drugs/DrugSafety/ucm475463.htm>. Published 2015. Accessed March 2, 2018.
15. Fadini GP, Bonora BM, Avogaro A. SGLT2 inhibitors and diabetic ketoacidosis: data from the FDA Adverse Event Reporting System. *Diabetologia*. 2017;60(8):1385-1389.
16. Blau JE, Tella SH, Taylor SI, Rother KI. Ketoacidosis associated with SGLT2 inhibitor treatment: analysis of FAERS data. *Diabetes Metab Res Rev*. 2017;33(8).
17. Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2018;375(4):311-322.
18. Perkovic V, Jardine MJ, Neal B, et al. Canagliflozin and renal outcomes in type 2 diabetes and nephropathy. *N Engl J Med*. 2019;380(24):2295-2306.
19. Zelniker TA, Braunwald E. Cardiac and renal effects of sodium-glucose co-transporter 2 inhibitors in diabetes: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2018;72(15):1845-1855.
20. Fralick M, Kim SC, Schneeweiss S, Kim D, Redelmeier DA, Patorno E. Fracture risk after initiation of use of canagliflozin: a cohort study. *Ann Intern Med*. 2019;170(3):155-163.
21. Fralick M, Sacks CA, Kesselheim AS. Assessment of use of combined dextromethorphan and quinidine in patients with dementia or Parkinson disease after US Food and Drug Administration approval for pseudobulbar affect. *JAMA Intern Med*. 2019;179(2):224-230.
22. Patorno E, Pawar A, Franklin JM, et al. Empagliflozin and the risk of heart failure hospitalization in routine clinical care. *Circulation*. 2019;139(25):2822-2830.
23. Jensen ML, Persson F, Andersen GS, et al. Incidence of ketoacidosis in the Danish type 2 diabetes population before and after introduction of sodium-glucose cotransporter 2 inhibitors—a nationwide, retrospective cohort study, 1995–2014. *Diabetes Care*. 2017;40(5):e57-e58.
24. Ueda P, Svanström H, Melbye M, et al. Sodium glucose cotransporter 2 inhibitors and risk of serious adverse events: nationwide register based cohort study. *BMJ*. 2018;363:k4365.
25. Curtis JR, To T, Muirhead S, Cummings E, Daneman D. Recent trends in hospitalization for diabetic ketoacidosis in ontario children. *Diabetes Care*. 2002;25(9):1591-1596.
26. Ray WA, Bobo W V, Cooper WO, Epstein RA, Arbogast PG, Mounsey J. Positive predictive value of automated database records for diabetic ketoacidosis (DKA) in children and youth exposed to antipsychotic drugs or control medications: a tennessee medicaid study. *BMC Med Res Methodol*. 2011;11:157.
27. Rathmann W, Kostev K. Fracture risk in patients with newly diagnosed type 2 diabetes: a retrospective database analysis in primary care. *J Diabetes Complications*. 2015;29(6):766-770.
28. Leslie W, Rubin M, Schwartz A, Kanis J. Type 2 diabetes and bone. *J Bone Min Res*. 2012;27:2231-2237.
29. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. Springer Texts Stat.:618.
30. Hoon KS, Yeo KC, Azam S, Shunmugam B, De Boer F. Critical review of machine learning approaches to apply big data analytics in DDoS forensics. In: 2018 International Conference on Computer Communication and Informatics (ICCCI). IEEE; 2018:1-5.
31. Friedman J. Greedy Function Approximation: a Gradient Boosting Machine. Author ( s ): Jerome H . Friedman Source : The Annals of Statistics , Vol . 29 , No . 5 ( Oct . , 2001 ) , pp . 1189-1232 Published by : Institute of Mathematical Statistics Stable URL : <http://www.AnnStat>. 2001;29(5):1189-1232.
32. Rosenstock J, Ferrannini E. Euglycemic diabetic ketoacidosis: a predictable, detectable, and preventable safety concern with SGLT2 inhibitors. *Diabetes Care*. 2015;38(9):1638-1642.
33. Macha S, Sennewald R, Rose P, et al. Lack of clinically relevant drug-drug interaction between empagliflozin, a sodium glucose cotransporter 2 inhibitor, and verapamil, ramipril, or digoxin in healthy volunteers. *Clin Ther*. 2013;35(3):226-235.
34. Food and Drug Administration. Canagliflozin drug label (2017). Available online: <https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&AppNo=204042> (Accessed 6 Sep 2017). [http://www.accessdata.fda.gov/drugsatfda\\_docs/label/2015/204042s0051bl.pdfSkeith](http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/204042s0051bl.pdfSkeith). Accessed November 19, 2016.
35. U.S. Food and Drug Administration (FDA). Memantine product monograph. 2018.
36. Storgaard H, Gluud LL, Bennett C, et al. Benefits and harms of sodium-glucose co-transporter 2 inhibitors in patients with type 2 diabetes: a systematic review and meta-analysis. *PLoS One*. 2016;11(11):1-23.
37. Monami M, Nreu B, Zannoni S, Lualdi C, Mannucci E. Effects of SGLT-2 inhibitors on diabetic ketoacidosis: a meta-analysis of randomised controlled trials. *Diabetes Res Clin Pract*. 2017;130:53-60.

**Publisher's Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.