Validation of Acute Outcomes Among Patients With Type 2 Diabetes Mellitus in the Clinical Practice Research Datalink

Heather E. Danysh,[1] Alicia Gilsenan,[2] Alejandro Arana,[3] Manel Pladevall-Vila,[3] J. Bradley Layton,[2] Ryan Ziemiecki,[2] Hungta Chen,[4] Phillip R. Hunt,[5] Catherine Johannes[1]

[1]RTI Health Solutions, Waltham, MA, United States; [2]RTI Health Solutions, Research Triangle Park, NC, United States; [3]RTI Health Solutions, Barcelona, Spain; [4]AstraZeneca, Medical and Payer Evidence Statistics, Medical Affairs, Biopharmaceuticals Medical, Gaithersburg, MD, United States; [5]AstraZeneca, Cardiovascular, Renal, & Metabolic Evidence, Medical Affairs, Biopharmaceuticals Medical, Gaithersburg, MD, United States

DISCLOSURE

This study was funded by AstraZeneca under a contract granting the research team independent publication rights. RTI Health Solutions receives institutional funding for projects from public and private entities. HD, AG, AA, MPV, JBL, RZ, and CJ are employees of RTI HS. HC and PH are employees of AstraZeneca, the marketing authorization holder of the study drug.

BACKGROUND

- The Clinical Practice Research Datalink (CPRD) is a primary care electronic health records (EHR) database in the United Kingdom (UK) with partial linkage to hospital health records.
- Outcome validation is often required in postauthorization drug safety studies conducted in EHR databases to evaluate and quantify possible outcome misclassification.
- In a postauthorization drug safety study, we conducted a validation assessment in CPRD to evaluate the positive predictive value of algorithms to identify acute outcomes among individuals with type 2 diabetes mellitus who initiated a glucose-lowering drug (GLD).

OBJECTIVE

- Estimate the positive predictive values (PPV) of diagnosis-coded algorithms for hospitalization for acute kidney injury (hAKI), hospitalization for acute liver injury (hALI), and severe complications of urinary tract infection (sUTI) in CPRD.

METHODS

- Eligible patients:
  - The study population included individuals registered in an up-to-standard general medical practice in the UK, aged ≥ 18 years, and initiating a GLD from November 13, 2012, to December 31, 2018.
  - Patients were prescribed the study drug, dapagliflozin (a sodium-glucose cotransporter-2 [SGLT2] inhibitor), or another GLD except other SGLT2 inhibitors, insulin monotherapy, metformin monotherapy, or sulfonylurea monotherapy.
- We used prespecified algorithms (Figure 1) to identify hAKI, hALI, and sUTI cases in the CPRD General Practitioner Online Database (GOLD) and the Hospital Episodes Statistics (HES) database. The validation process, including validation sample selection, patient profile review, GP questionnaire requests, and adjudication review, is outlined in Figure 2. The clinical case definitions for hAKI, hALI, and sUTI used during adjudication review are outlined in Table 1.
- After adjudication review, cases included in each outcome-specific validation sample were classified as either (1) a confirmed case, (2) a confirmed non-case, or (3) a postreview provisional case, which included all cases for which there was insufficient information to assign a definitive case status.
- PPVs and 95% confidence intervals (CI) were estimated in two ways depending on assumptions about postreview provisional cases.
  - **(1) PPV 1:** The proportion of confirmed cases among all cases included in the adjudication review (i.e., the validation sample), assuming all postreview provisional cases are non-cases (worst-case scenario)

$$PPV\ 1 = \frac{\text{Confirmed cases}}{\text{All cases in validation sample}}$$

  - **(2) PPV 2:** The proportion of confirmed cases among only cases where a definitive case status was assigned (i.e., confirmed cases and confirmed non-cases)

$$PPV\ 2 = \frac{\text{Confirmed cases}}{\text{Confirmed cases + Confirmed non-cases}}$$

Figure 1. Algorithms for Hospitalizations for Acute Kidney Injury, Acute Liver Injury, and Severe Complications of Urinary Tract Infection

Hospitalization for acute kidney injury

1. Any Read code for acute kidney injury plus hospitalization[a]

OR

2. Hospital record in HES with acute kidney injury as the first diagnosis (ICD-10 code) in the first episode[b] of the hospitalization stay

Hospitalization for acute liver injury

1. Any Read code for acute liver injury plus hospitalization[a]

OR

2. Hospital record in HES with acute liver injury as the first diagnosis (ICD-10 code) in the first episode[b] of the hospitalization stay

Severe complication of urinary tract infection

Pyelonephritis

1. Hospital record in HES with pyelonephritis as the first diagnosis in the first episode[b] of the hospitalization stay

OR

2. Any Read code for pyelonephritis with consultation type = ED visit[c]

OR

3. Any Read code for pyelonephritis with consultation type = GP plus hospitalization[a]

Urosepsis

***Urosepsis diagnosis** defined as*

a. GP diagnosis of urosepsis plus hospitalization[a]

OR

b. Read code for urosepsis with ED visit[b]

OR

AND

1. Urinary tract infection diagnosis (within 7 days before or after the sepsis hospitalization date) defined as

a. Any Read code for urinary tract infection with consultation type = GP or ED visit

OR

b. Hospital record in HES with urinary tract infection recorded as the first diagnosis in the first episode[b] of the hospitalization stay

2. Sepsis: hospitalization or ED visit for sepsis defined as

a. Hospital record in HES with sepsis recorded as the first diagnosis in the first episode[b] of the hospitalization stay

OR

b. Any Read code for sepsis with ED visit

OR

c. Any Read code for sepsis with consultation type = GP plus hospitalization[a]

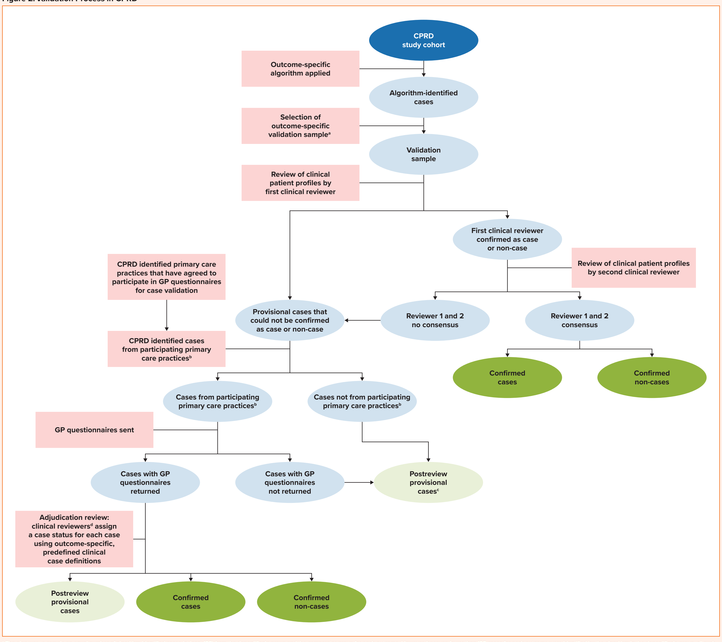
ED = emergency department; GP = general practitioner; ICD-10 = *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*.

[a] In CPRD GOLD data, hospitalization (within 30 days before or after the Read code of interest) is defined as (1) all Read codes where Death = 0 for hospitalization indicator variables; **OR** (2) a reported Clinical or Referral record with a consultation type as reported in the consultation data set = 12, 23, 47, 48; **OR** (3) inpatient variable from the referral data set = 1.

[b] In CPRD HES, the data are structured such that multiple "episodes" may be recorded for a single hospitalization, with one episode for each encounter within the hospital made by the patient (e.g., one episode may be the initial ED visit, one episode may be from another ward to which the patient is admitted). The first episode of the hospitalization stay is the first encounter that the patient has during the hospitalization.

[c] In CPRD GOLD data, accident or ED visits (on the same day as the outcome Read code of interest) are defined as (1) a Read code identified as related to an ED visit; **OR** (2) a reported clinical or referral record with a consultation type as r.

Figure 2. Validation Process in CPRD



[a] The plan was to sample an equal number of algorithm-identified cases (at least 125) from the dapagliflozin group and the comparator group for each outcome. However, because fewer than 125 cases for each outcome were identified by the algorithm in the dapagliflozin group, all such cases were selected into the validation sample.

[b] CPRD sent GP questionnaires only to primary care practices that have agreed to participate in questionnaires at the time of this study.

[c] Cases were classified as "postreview provisional cases" if (1) they could not be confirmed a case or a non-case during the patient profile review stage and for which GP questionnaires could not be requested because they were from a primary care practice that did not participate in GP questionnaires at the time of this study; or (2) all GP questionnaires were requested but not completed and returned by the GP.

[d] Two clinical reviewers independently reviewed information on each case to assign a case status. Disagreements between the two clinical reviewers were resolved through discussion among an adjudication committee consisting of three clinical adjudicators.

Table 1. Clinical Case Definitions for Hospitalizations for Acute Kidney Injury or Acute Liver Injury, and Severe Complications of Urinary Tract Infection

Outcome	Criteria for Confirmed Cases	Source
Hospitalization for acute kidney injury[a]	1. Hospital discharge diagnosis of acute kidney injury and 2. Increases in serum creatinine[b] at or within 72 h of hospital admission and 3. No recorded diagnosis of chronic kidney disease before cohort entry	Based on prior epidemiological research and on a subset of the RIFLE[c] criteria proposed by the Acute Dialysis Quality Initiative[d]
Hospitalization for acute liver injury[e]	1. Recorded hospitalization of acute liver injury and 2. ≥ 1 elevated liver enzyme test (ALT, AP, TBl)[f] within 26 wks before or within 48 h of hospital admission and 3. No chronic liver disease, hepatic or pancreatic cancer, or alcoholism before cohort entry	Based on guidance published by the FDA[g] and criteria proposed by Navarro et al., 2006[h]
Severe complications of urinary tract infection	Hospitalization or emergency department visit for pyelonephritis or urosepsis and met the other criteria for pyelonephritis or urosepsis	
Pyelonephritis	A confirmed case met Criterion 1 and either Criterion 2 or 3: 1. ≥ 2 symptoms: fever, dysuria, flank pain or costovertebral angle tenderness, leukocytosis or WBC count > 12,000/mm[3], abnormal urine 2. ≥ 1 imaging test (CT, MRI, or ultrasonography) indicating either renal inflammation, renal abscess, or hydronephrosis 3. ≥ 1 positive blood and/or urine culture test[i]	Patkar et al., 2009[a]
Urosepsis	1. Diagnosis of urinary tract infection and/or infection of male genital organs within 1 week of hospital admission for sepsis and 2. Either proof of bacteremia or clinical suspicion of sepsis and 3. ≥ 2 symptoms: fever, tachycardia, tachypnea, respiratory alkalosis, leucocytes ≥ 12,000 per µL or ≤ 4,000 per µL	Wagenlehner et al., 2008[i]

ALT = alanine transaminase; AP = alkaline phosphatase; CT = computed tomography; h = hours; hALI = hospitalization for acute liver injury; MRI = magnetic resonance imaging; TB = total bilirubin; ULN = upper limit of normal; WBC = white blood cell; wks = weeks.

[a] For hAKI, if information on serum creatinine levels for documenting a change in serum creatinine was not available, the clinical reviewers could use the clinical details in the EHR and GP questionnaire to determine that acute kidney injury was likely in order to confirm a case.

[b] An increase in serum creatinine was defined as follows: (1) in patients with normal baseline renal function (≤ ULN), at least a twofold increase from the baseline value to a value greater than the ULN or (2) in patients with baseline renal insufficiency (defined as > ULN), an increase from the baseline value to at least twice the ULN. Baseline was defined as the lowest recorded value during the 365 days before the outcome date.

[c] The components of the RIFLE classification system for acute renal failure includes Risk, Injury, Failure, Loss of kidney function, and End-stage kidney disease.

[d] For hALI, if information on laboratory tests and results was missing from the recorded information, the clinical reviewers could use clear evidence provided by the GP on the GP questionnaire (i.e., evidence of jaundice and the GP's confirmation of the acute liver injury diagnosis) to confirm a case.

[e] Elevated liver enzyme tests were defined as follows: (1) elevation of ALT > 3 times the ULN, (2) any increase of ALT and AP and an increase of TB > 2 times the ULN, or (3) elevation of AP > 2 times the ULN.

[f] A positive culture test includes any of the following: (1) blood cultures and urine cultures positive for the same organism; (2) blood cultures positive for gram-negative organisms, *Enterococcus* species, or *Staphylococcus saprophyticus*; (3) urine culture positive for > 100,000 gram-negative organisms, *Enterococcus* species, or *S. saprophyticus*; (4) urine culture positive for < 100,000 any organism and patient treated for ≥ 7 days with antibiotics.

RESULTS

Table 2. Disposition of Patient Profiles Reviewed, General Practitioner Questionnaires, and Adjudicated Cases

Cases	hAKI	hALI	sUTI[a]
Algorithm-identified cases, n	105	27	96
Validation sample			
Clinical patient profile review, n	105	27	96
Reviewer consensus to confirm as case or non-case, n (%)	30 (29)	13 (48)	< 5 (NR)[b]
GP questionnaires sent,[c] n	70	12	74
GP questionnaires completed and returned, n (%)[d]	34 (49)	6 (50)	37 (50)
Adjudication review			
Included in adjudication review, n	105	27	96
Sufficient information to assign case status, n (%)[e]	54 (51)	16 (59)	29 (30)
Confirmed cases, n (%)[f]	33 (61)	8 (50)	14 (48)
Confirmed non-cases, n (%)[f]	21 (39)[g]	8 (50)[g]	15 (52)
Insufficient information to assign case status (postreview provisional cases), n (%)[h]	51 (49)	11 (41)	67 (70)

Table 3. Positive Predictive Values of Algorithms for Adjudicated Cases

PPV Estimation Approach, % (95% CI)	hAKI	hALI	sUTI[a]
PPV 1	32.4 (23.6-42.2)	33.3 (16.5-54.0)	14.6 (8.2-23.3)
PPV 2	63.0 (48.7-75.7)	56.3 (29.9-80.2)	48.3 (29.4-67.5)

CI = confidence interval.

[a] Includes hospitalizations or emergency department visits for urosepsis and/or pyelonephritis after a diagnosis of urinary tract infection.

NR = not reportable to mask small cell count.

[a] Includes hospitalizations or emergency department visits for urosepsis and/or pyelonephritis after a diagnosis of urinary tract infection.

[b] Based on CPRD data privacy policies, frequency values of 1-4 must be suppressed.

[c] CPRD sent GP questionnaires only to primary care practices that have agreed to participate in questionnaires at the time of this study.

[d] Percentage among cases with GP questionnaires sent.

[e] Percentage among cases included in adjudication review.

[f] Percentage among cases included with definitive case status.

[g] For hAKI and hALI, includes cases (n < 5) found to be ineligible for study cohort inclusion during adjudication review.

DISCUSSION

- After excluding postreview provisional cases from the PPV calculation (i.e., PPV 2), the results for hAKI and hALI observed in our study are higher than other comparable published algorithms in the UK general practice health records.[4-8]
- To our knowledge, there are no published algorithms with validation results for our sUTI case definition of pyelonephritis or urosepsis.

LIMITATIONS

- A substantial proportion of each outcome-specific validation sample did not have sufficient information available to determine a case status (hAKI, 49%; hALI, 41%; sUTI, 70%).
  - A definitive case status could not be determined in some instances because of a lack of laboratory data (e.g., serum creatinine levels for hAKI, liver enzymes for hALI, and blood/urine cultures for pyelonephritis and urosepsis).
  - Of the questionnaires that were sent to GPs, only 50% were completed and returned.
  - Some cases were from primary care practices that were not participating in GP questionnaires at the time of this study. Therefore, a questionnaire could not be sent to acquire additional information about these cases.
- The low number of algorithm-identified cases, and subsequently the low number of cases included in the adjudication review, led to relatively imprecise PPV estimates with wide 95% CIs.

CONCLUSIONS

- In this validation study, diagnosis-coded algorithms applied to CPRD, a primary care EHR database, resulted in low to moderate validity for identifying hAKI, hALI, or sUTI in patients with type 2 diabetes mellitus.

REFERENCES

1. Bellomo R, et al. Crit Care. 2004;8(4):R204-12.
2. Food and Drug Administration. https://www.fda.gov/media/116737/download.
3. Navarro VJ, et al. N Engl J Med. 2006;354(7):731-9.
4. Patkar NM, et al. J Clin Epidemiol. 2009;62(3):321-7.e1-7.
5. Wagenlehner FM, et al. Eur J Clin Invest. 2008;38 Suppl 2:45-9.
6. Huerta C, et al. Am J Kidney Dis. 2005;45(3):531-9.
7. Lo Re V, et al. BMJ Open Diabetes Res Care. 2017;5(1):e000400.
8. Huerta, et al. Pharmacotherapy. 2002;22(9):1091-6.

CONTACT INFORMATION

Heather Danysh, MHS, PhD
Senior Research Epidemiologist

RTI Health Solutions
307 Waverley Oaks Road, Suite 101
Waltham, MA 02452-8413

Phone: +1.781.434.1772
Fax: +1.781.434.1701
Email: hdanysh@rti.org