



Identification of cytokine release syndrome and indicators of severity in retrospective databases among patients receiving immunotherapy

Scott F. Huntington¹ | Dee Lin² | Marie-Hélène Lafeuille³ | Philippe Thompson-Leduc³  | Aditi Shah³ | Nina Kim² | Laura Hester² | Anabelle Tardif-Samson³ | Bronwyn Moore³ | Jessica Fowler² | Alexander Marshall⁴ | Xinke Zhang² | Dina Gifkins⁵  | Bingcao Wu²

¹Yale School of Medicine, New Haven, Connecticut, USA

²Janssen Scientific Affairs, LLC, a Johnson and Johnson Company, Horsham, Pennsylvania, USA

³Analysis Group, Inc., Montreal, Quebec, Canada

⁴Janssen Global Services, LLC, a Johnson and Johnson Company, Raritan, Pennsylvania, USA

⁵Janssen Research and Development, LLC, a Johnson and Johnson Company, Raritan, Pennsylvania, USA

Correspondence

Philippe Thompson-Leduc, Analysis Group, Inc. 1190 avenue des Canadiens-de-Montréal, Suite 1500, Montréal, QC H3B 0G7, Canada.

Email: philippe.thompson-leduc@analysisgroup.com

Funding information

Janssen Scientific Affairs

Abstract

Cytokine release syndrome (CRS) can occur following cancer immunotherapies, but is most often mild and of limited duration. International Classification of Diseases (ICD)-10 codes allowing identification of CRS were introduced in 2020 but may be underutilized. We evaluated the performance of a published claims-based algorithm to detect CRS (any grade) and high-grade CRS (HG, grades 2-5), as well as identified indicators of HG CRS in retrospective data. Adults with low-grade and HG CRS during an encounter coinciding with administrations of blinatumomab or chimeric antigen receptor-T therapy were identified in three types of retrospective databases (hospital chargemaster data, electronic health records, and administrative claims). The algorithm's sensitivity in detecting any CRS and HG CRS was reported. A least absolute shrinkage and selection operator (LASSO) regression model was developed to identify indicators of HG CRS. Performance of the model was evaluated using area under the curve (AUC). The sensitivity of the algorithm to detect any grade CRS ranged between 77%–100% and between 8%–80% for HG CRS, depending on the type of database. The LASSO model identified hypotension, positive pressure (including mechanical ventilation), tocilizumab, and vasopressors as indicators of HG CRS. AUC varied between 60% and 75%. The algorithm accurately detected any grade CRS for over three-quarters of instances, but was not as reliable for HG CRS. Results varied based on database attributes. Hypotension, vasopressors, positive pressure, and tocilizumab were associated with HG CRS and may be methodologically helpful signals of CRS severity in retrospective data.

Abbreviations: ASTCT, American Society for Transplantation and Cellular Therapy; AUC, area under the curve; BsAbs, bispecific antibodies; CAR T-cell, chimeric antigen receptor T cell; CPT, Common Procedural Terminology; CRS, cytokine release syndrome; EHR, electronic health record; FDA, Food and Drug Administration; FN, false negative; FP, false positive; HCPCS, Healthcare Common Procedure Coding System; HG, high-grade; HIPAA, Health Insurance Portability and Accountability Act; ICD, International Classification of Diseases; ICD-10-CM, International Classification of Diseases, 10th Revision, Clinical Modification; ICD-10-PCS, ICD-10 Procedure Coding System; IL-6, interleukin 6; IP, inpatient; LASSO, least absolute shrinkage and selection operator; NPV, negative predictive value; Optum® EHR, Optum® de-identified Electronic Health Record data set; PHD, PINC AI™ Healthcare Database; PPV, positive predictive value; Quan-CCI, Quan Charlson Comorbidity Index; SD, standard deviation; TN, true negative; TP, true positive.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 Janssen Scientific Affairs, LLC. *Pharmacology Research & Perspectives* published by British Pharmacological Society and American Society for Pharmacology and Experimental Therapeutics and John Wiley & Sons Ltd.

KEYWORDS

administrative claims, bispecific antibodies, chimeric antigen receptors, cytokine release syndrome, drug-related side effects and adverse reactions, electronic health records, healthcare, systemic inflammatory response syndrome

1 | INTRODUCTION

Cytokine release syndrome (CRS) is a systemic inflammatory response caused by a high-level activation of the immune system following infection (e.g., COVID-19) or as an adverse event of certain immunotherapies, including bispecific antibodies (BsAbs) and chimeric antigen receptor T cell (CAR T-cell) therapies.¹⁻³ Increased serum cytokine levels, particularly interleukin 6 (IL-6), can induce symptoms characteristic of inflammation, including fever, rash, headache, arthralgia, hypoxia, hypotension, myalgia, and fatigue. Severe cases may result in uncontrolled systemic inflammatory response.³

The American Society for Transplantation and Cellular Therapy (ASTCT) has established consensus criteria to define CRS and categorized its severity based on the progression of hypotension and hypoxia.⁴ Grade 1 CRS is characterized by fever only and cases often resolve with minor intervention.^{4,5} Increasingly severe grades (i.e., grades 2 to 5) are associated with more therapeutic requirements (e.g., hypoxia requiring positive pressure airway support and/or hypotension requiring treatment with multiple vasopressors is indicative of “grade 4”).⁴ Though not a component of consensus grading, treatment with an IL-6 receptor inhibitor (i.e., tocilizumab) or corticosteroids may also be required.⁵

Rapid advancements in the treatment of hematologic malignancies have led to the Food and Drug Administration (FDA) approvals for numerous immunotherapies, which have been associated with CRS.⁶⁻⁸ Clinical trials of BsAbs and CAR T-cell therapies for hematologic malignancies have shown CRS as a prevalent side effect, especially at the initiation of the therapy, with incidence rates dependent on the therapeutic agent, dose, and patient population (though only up to a quarter of these cases were classified as severe CRS).^{7,9}

Understanding the rate and severity of CRS using retrospective databases is important to estimate the real-world safety outcomes of novel cancer therapies, including BsAbs and CAR T-cell therapies.^{9,10} To date, claims-based studies aiming at identifying CRS are limited to Keating et al. (the “Keating algorithm”), which used administrative claims among patients recently treated with CAR T-cell therapy to evaluate the incidence of CRS. That study preceded the addition of an International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) code (D89.83) in October 2020.¹¹ However, due to its novelty, this ICD-10-CM code may be underutilized. Moreover, the code for an “unspecified” grade poses challenges when evaluating CRS severity and determining the impacts of mild and severe adverse events on outcomes in a real-world setting.

With the recent introduction of CRS diagnostic coding in retrospective data, this study aimed to measure the performance of the Keating algorithm to identify cases of any CRS and high-grade (grade ≥ 2) CRS among patients with a diagnosis of CRS per

ICD-10-CM coding. The second objective was to develop and validate a least absolute shrinkage and selection operator (LASSO) regression model to identify severity indicators of high-grade CRS among patients receiving treatment associated with CRS in retrospective databases.

2 | METHODS**2.1 | Data source**

For the first study objective, the performance of the Keating algorithm was evaluated across three data sources. The first data source was the PINC AI™ Healthcare Database (“PHD,” 01/01/2020 to 11/04/2021). PHD is a large health system-based, geographically diverse, all-payer administrative database that contains diagnosis, procedures, healthcare utilization, and financial data on inpatient (IP) discharges and health system-based outpatient visits and covers over 244 million unique patients.¹² The second data source was Merative™ MarketScan® Research Databases (04/01/2020–04/11/2021) which represents nearly 240 million covered lives including employees and their dependents, self-insured employers, and Medicare-eligible retirees with employer-provided Medicare supplemental plans. The third data source was the Optum® de-identified Electronic Health Record data set (“Optum® EHR,” 04/01/2020–03/31/2022). Optum® data acquisition model aggregates de-identified Electronic Health Record (EHR) data from providers across the continuum of care.

For the second study objective, the LASSO model was developed using the PHD data and subsequently applied in the Merative™ MarketScan® and Optum® EHR databases to assess model performance in external data sources. All data were de-identified and complied with the patient requirements of the Health Insurance Portability and Accountability Act (HIPAA).

2.2 | Study design

A retrospective cohort study design was employed. In PHD, unique hospitalizations with an ICD-10-CM code for CRS (D89.83x) were identified. Patients were followed from their admission date to their discharge date. If a given patient had multiple hospitalizations with an ICD-10-CM code for CRS, only the first hospitalization was retained for analysis. Hospitalizations were included in the analysis if they met the following criteria: an ICD-10-CM code for CRS was recorded, information during the hospitalization was complete (i.e., no missing data due to data cut-off date, 11/04/2021), and the patient was ≥ 18 years old as of the admission date.

Hospitalizations were assigned a severity based on ICD-10-CM codes (i.e., “low-grade”: grade 1 [ICD-10-CM code: D89.831], “high-grade”: grade ≥ 2 [ICD-10-CM code: D89.832/D89.833/D89.834/D89.835] or “unspecified grade”: ICD-10-CM code: D89.839). Hospitalizations with more than one ICD-10-CM code observed were assigned to the highest severity achieved during the hospitalization. To identify patients with an ICD-10-CM diagnosis of CRS associated with a treatment (as opposed to other causes, such as COVID-19), hospitalizations during which blinatumomab or CAR T-cell therapies were administered formed the “Treated sample.”

In Merative™ MarketScan® and Optum® EHR, patients were followed for 14 days on and following the first observed ICD-10-CM code for CRS. Adult patients were included in the analysis if they had an ICD-10-CM code for CRS (severity assigned as above), ≥ 6 months of continuous enrollment before ≥ 14 days after the diagnosis, and had received a treatment associated with CRS (i.e., blinatumomab or CAR T-cell therapy) within 14 days prior to the diagnosis. Since patients in Merative™ MarketScan® and Optum® EHR had all received a treatment of interest, these were also considered “Treated samples.”

2.3 | Variables and analyses

2.3.1 | First objective: Performance of the Keating algorithm among hospitalizations/patients with a CRS diagnosis

In the Keating algorithm, a patient is identified as having “any CRS” if they had CRS symptoms (i.e., fever, fatigue, malaise, headaches [with a fever], arthralgias, tachycardia, hypotension, or hypoxia) within 14 days of a CAR T-cell infusion. In addition, “severe CRS” is further characterized by a combination of fever with hypotension/hypoxia and CRS management (i.e., use of tocilizumab, corticosteroids, or vasopressors [see [Table S1](#) for algorithm]).

Based on the Keating algorithm, the proportion of hospitalizations (in PHD) or patients (in Merative™ MarketScan® and Optum® EHR) with CRS symptoms and management, and those with “any CRS” and “severe CRS” were reported. CRS symptoms were identified through ICD-10-CM codes and text search strings. Fever was defined either through an ICD-10-CM code or use of acetaminophen (without concurrent antihistamines on the same day to avoid falsely classifying other conditions such as allergy, cold, sleep issues or pre-treatment prophylaxis) since ICD-10-CM codes for fever may be underused in real-world data. CRS management was identified through Common Procedural Terminology (CPT), Healthcare Common Procedure Coding System (HCPCS), and ICD-10 Procedure Coding System (ICD-10-PCS) codes.

The sensitivity of the Keating algorithm to detect CRS (any grade) was calculated as the number of hospitalizations/patients with CRS per algorithm and confirmed by ICD-10-CM codes (true positives) divided by the number of hospitalizations/patients with CRS as per ICD-10-CM codes (true positives + false negatives). Similarly, the

sensitivity to detect high-grade CRS was calculated as the number of hospitalizations/patients with severe CRS per algorithm and high-grade CRS by ICD-10-CM codes (true positives) divided by the number of hospitalizations/patients with CRS grade ≥ 2 as per ICD-10-CM codes (true positives + false negatives).

2.3.2 | Second objective: Indicators of high-grade (grade ≥ 2) CRS in the Treated sample

For the second objective, a LASSO logistic regression model was developed and trained to identify indicators of high-grade CRS (grade ≥ 2) observed during hospitalizations among the Treated sample in PHD with a known grade of CRS based on ICD-10-CM codes (i.e., excluded patients with unspecified grade). The LASSO regression is a penalized regression that uses shrinking as feature selection approach to identify the most relevant variables likely associated with the outcome of interest. This approach was chosen for its ability to identify parsimonious models with fewer parameters. Five-fold cross validation was conducted, and performance metrics of the model (i.e., area under the curve [AUC], accuracy, sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]) were reported.

Candidate indicators of high-grade CRS arising from cancer treatments considered for the LASSO regression were informed by descriptive analyses and clinical advice, and included age, gender, length of stay, Quan-CCI comorbidities, CRS symptoms, and CRS management during the hospitalization ([Table S2](#)). To validate the model in external databases, the LASSO model developed in PHD was applied in the Treated samples in Merative™ MarketScan® and Optum® EHR, and model performance was reported.

3 | RESULTS

A total of 7388 hospitalizations had an ICD-10-CM diagnosis for CRS and met selection criteria in PHD. Among them, 860 hospitalizations (12%) had an ICD-10-CM code for grade 1 CRS, 2014 (27%) had an ICD-10-CM code for grade ≥ 2 CRS, and 4514 (61%) had an ICD-10-CM code for CRS of unspecified grade ([Figure 1](#)). Of these, 146 hospitalizations (2%) were for patients who received treatment for blinatumomab or CAR T-cell therapies and formed the Treated sample (81 grade 1, 52 grade ≥ 2 , 13 unspecified grade CRS).

The Treated sample in the Optum® EHR database consisted of 37 patients (11 grade 1, 15 grade ≥ 2 , 11 unspecified grade CRS). In Merative™ MarketScan®, the Treated sample consisted of 35 patients (16 grade 1, 12 grade ≥ 2 , 7 unspecified grade CRS).

Characteristics of hospitalizations in PHD are presented in [Table 1](#). In all hospitalizations with an ICD-10-CM code of CRS, mean age was 61.0 years old, and 44% of patients were female. The mean length of stay was 11.2 days and 93% reported a diagnosis of COVID-19. In the Treated sample, mean age was 56.9 years old, and 34% of patients were female. The mean length of stay was 15.2 days and 5% had a concomitant diagnosis of COVID-19.

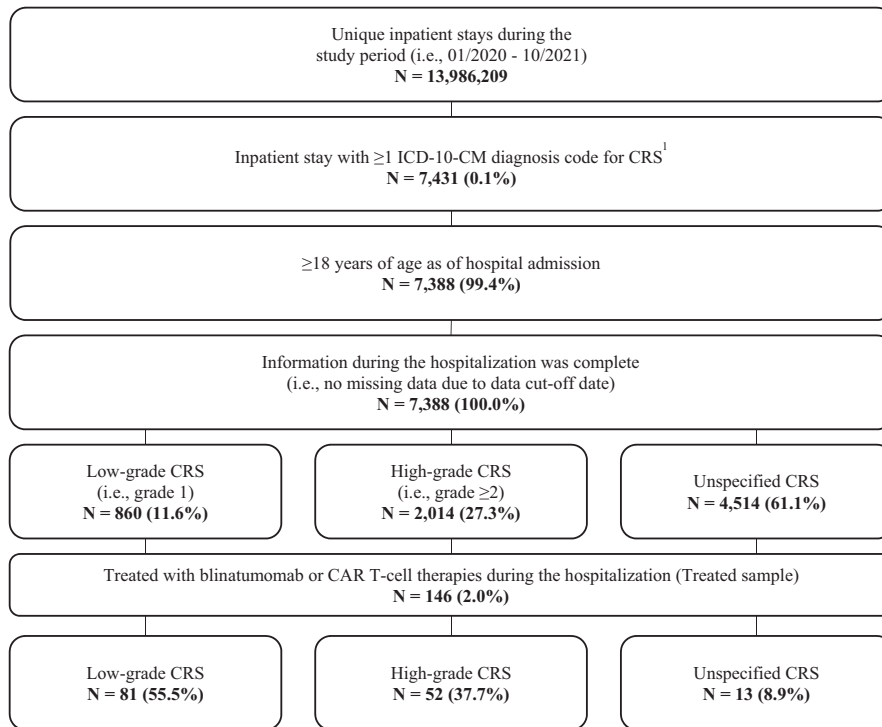


FIGURE 1 Sample selection in the PINC AI™ Healthcare Database. CAR T-cell: Chimeric antigen receptor T-cell; CRS: Cytokine release syndrome; ICD-10-CM: International Classification of Disease, Tenth Revision, Clinical Modification. Source: PINC AI™ Healthcare Database (From 01/2020 through 10/2021). ¹CRS was identified using ICD-10-CM codes D89.83 (Grade 1: D89.831; Grade 2: D89.832; Grade 3: D89.833; Grade 4: D89.834; Grade 5: D89.835; Grade unspecified: D89.839).

3.1 | First objective: Performance of the Keating algorithm among hospitalizations associated with a CRS diagnosis

Across all hospitalizations included in PHD, the most common CRS symptoms were hypoxia (74%), fever (64%), and tachycardia (24%). In the Treated sample, the most common CRS symptoms were fever (92%), hypotension (26%) and tachycardia (24%; [Table 2](#)). Most patients used corticosteroids (89% overall, 65% in the Treated sample), while tocilizumab was used in 29% of hospitalizations (49% in the Treated sample).

In the two additional data sources, the frequency of symptoms and management observed differed. In the Optum® EHR database, frequent symptoms included fever (97%), arthralgia (84%), tachycardia (76%), fatigue (73%), and headache with a fever (57%). The use of corticosteroids (81%), tocilizumab (62%), and vasopressors (46%) was frequently reported. In Merative™ MarketScan®, symptoms were less frequent (fever [60%], tachycardia [23%], hypoxia [23%]) as were management strategies (corticosteroids [23%], tocilizumab [3%], vasopressors [0%]; [Table S3](#)).

In PHD, the algorithm's sensitivity to detect CRS of any grade was 95% both in the overall and the Treated sample ([Figure 2](#)). Among hospitalizations with high-grade CRS, 52% of the overall sample and 44% of the Treated sample were correctly identified as grade ≥2 CRS.

Among patients in Optum® EHR, the algorithm's sensitivity in detecting any CRS was 100%, and 80% were correctly identified as grade ≥2 CRS. Among patients in Merative™ MarketScan®, the algorithm's sensitivity in detecting any CRS was 77%, and 8% were correctly identified as grade ≥2 CRS.

3.2 | Second objective: Indicators of high-grade (grade ≥2) CRS in the Treated sample

Among hospitalizations in the Treated sample of PHD, differences were noted in the prevalence of certain candidate indicators considered for the LASSO regression between low-grade and high-grade CRS, including hypotension (16% in low-grade vs. 42% in high-grade), positive pressure (including mechanical ventilation, 7% vs. 23%), tocilizumab (43% vs. 67%), and vasopressors (1% vs. 17%). Other candidate indicators, for example, acetaminophen, showed less discrimination between the severity of the CRS event. ([Table S2](#)).

The LASSO regression model identified hypotension, use of vasopressors, positive pressure (including mechanical ventilation), and tocilizumab as indicators of high-grade CRS ([Table 3](#)). The AUC of the holdout test dataset was 75%, which indicated acceptable performance of the model in identifying likely cases of high-grade CRS.¹³ ([Table 4](#)).

The performance of the LASSO model was assessed in the two external data sources ([Table 4](#)). Among patients in Optum® EHR, the AUC was 70% (acceptable model performance).¹³ Among patients in Merative™ MarketScan®, the AUC was 60% (less than acceptable model performance).¹³

4 | DISCUSSION

The Keating algorithm identified CRS of any grade with relatively high sensitivity in PHD (95%), Optum® EHR (100%), and Merative™ MarketScan® (77%), demonstrating the potential of the algorithm to identify any grade CRS when ICD-10-CM coding is inconsistent

TABLE 1 Patient characteristics in the PINCAI™ Healthcare Database.

	All hospitalizations ^a			Treated sample ^a			
	All CRS N = 7388	Low-grade (i.e., grade 1) N = 860	High-grade (i.e., grade ≥2) N = 2014	All CRS N = 146	Low-grade (i.e., grade 1) N = 81	High-grade (i.e., grade ≥2) N = 52	Unspecified CRS N = 13
Age, mean ± SD [median]	61.0 ± 15.7 [62.0]	58.5 ± 15.9 [60.0]	61.2 ± 15.7 [63.0]	56.9 ± 15.7 [61.0]	57.8 ± 14.9 [61.0]	56.8 ± 17.0 [63.0]	51.8 ± 15.6 [55.0]
Race, n (%)							
White	5242 (71)	719 (84)	1580 (78)	109 (75)	58 (72)	43 (83)	8 (62)
Black	1299 (18)	86 (10)	223 (11)	8 (5)	4 (5)	4 (8)	0 (0)
Other	632 (9)	51 (6)	199 (10)	25 (17)	17 (21)	3 (6)	5 (38)
Unknown	215 (3)	4 (0)	12 (1)	4 (3)	2 (2)	2 (4)	0 (0)
Female, n (%)	3234 (44)	423 (49)	894 (44)	49 (34)	22 (27)	19 (37)	8 (62)
Hospital census region, n (%)							
South	5025 (68)	640 (74)	1574 (78)	57 (39)	31 (38)	23 (44)	3 (23)
Midwest	1278 (17)	139 (16)	150 (7)	29 (20)	17 (21)	12 (23)	0 (0)
West	727 (10)	25 (3)	269 (13)	25 (17)	10 (12)	10 (19)	5 (38)
Northeast	358 (5)	56 (7)	21 (1)	35 (24)	23 (28)	7 (13)	5 (38)
Primary payer type, n (%)							
Medicare	3390 (46)	340 (40)	929 (46)	51 (35)	33 (41)	15 (29)	3 (23)
Commercial	2614 (35)	308 (36)	695 (35)	45 (31)	26 (32)	16 (31)	3 (23)
Medicaid	804 (11)	120 (14)	248 (12)	26 (18)	11 (14)	9 (17)	6 (46)
Other	580 (8)	92 (11)	142 (7)	24 (16)	11 (14)	12 (23)	1 (8)
Year of admission, n (%)							
2020	1588 (21)	46 (5)	122 (6)	43 (29)	20 (25)	17 (33)	6 (46)
2021	5800 (79)	814 (95)	1892 (94)	103 (71)	61 (75)	35 (67)	7 (54)
Quan-CCI ^{b,c} mean ± SD [median]	1.9 ± 2.1 [1.0]	1.8 ± 2.0 [1.0]	1.9 ± 2.1 [1.0]	2.8 ± 1.4 [2.0]	2.9 ± 1.6 [2.0]	2.7 ± 1.1 [2.0]	3.0 ± 1.2 [3.0]
COVID-19 ^d , n (%)	6900 (93)	678 (79)	1916 (95)	7 (5)	4 (5)	1 (2)	2 (15)
Length of stay (days), mean ± SD [median]	11.2 ± 11.5 [7.0]	6.2 ± 8.0 [4.0]	10.4 ± 9.5 [7.0]	15.2 ± 11.3 [14.0]	15.8 ± 12.9 [15.0]	14.0 ± 7.1 [14.0]	15.8 ± 14.9 [12.0]

Abbreviations: CRS, cytokine release syndrome; ICD-10-CM, International Classification of Disease, Tenth Revision, Clinical Modification; Quan-CCI, Quan Charlson Comorbidity Index; SD, standard deviation.

^aStays were classified as their highest known grade during the hospitalization. Classifications are mutually exclusive.

^bQuan-CCI was assessed over the hospitalization.

^cReference: Quan, H., Li, B., Couris, C. M., Fushimi, K., Graham, P., Hider, P., Januel, J. M., & Sundararajan, V. (2011). Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *American Journal of Epidemiology*, 173(6), 676–682.

^dCOVID-19 was identified using ICD-10-CM codes: U00, U07.1, U09, U49, U50, U85, U87.29, B34.2, J12.82, Z86.16 used during hospitalizations with admission dates on or after January 18th 2020 (i.e. first lab-confirmed COVID-19 case in the United States).

TABLE 2 CRS symptoms and management in the PINC AI™ Healthcare Database.

	All hospitalizations ^a			Treated sample ^a				
	All CRS N = 7388	Low-grade (i.e., grade 1) N = 860	High-grade (i.e., grade ≥2) N = 2014	Unspecified CRS N = 4514	All CRS N = 146	Low-grade (i.e., grade 1) N = 81	High-grade (i.e., grade ≥2) N = 52	Unspecified CRS N = 13
CRS symptoms, n (%)								
Hypoxia	5435 (74)	438 (51)	1562 (78)	3435 (76)	20 (14)	5 (6)	12 (23)	3 (23)
Fever ^b	4749 (64)	569 (66)	1311 (65)	2869 (64)	134 (92)	75 (93)	47 (90)	12 (92)
Diagnosis-based fever	378 (5)	105 (12)	108 (5)	165 (4)	88 (60)	51 (63)	31 (60)	6 (46)
Tachycardia	1768 (24)	156 (18)	491 (24)	1121 (25)	35 (24)	22 (27)	11 (21)	2 (15)
Hypotension	637 (9)	52 (6)	197 (10)	388 (9)	38 (26)	13 (16)	22 (42)	3 (23)
Malaise	136 (2)	12 (1)	25 (1)	99 (2)	3 (2)	1 (1)	2 (4)	0 (0)
Fatigue	69 (1)	9 (1)	25 (1)	35 (1)	4 (3)	2 (3)	2 (4)	0 (0)
Arthralgia	50 (1)	9 (1)	12 (1)	29 (1)	3 (2)	1 (1)	1 (2)	1 (8)
Headache with fever	21 (0)	9 (1)	4 (0)	8 (0)	11 (8)	7 (9)	3 (6)	1 (8)
CRS management, n (%)								
Corticosteroids	6594 (89)	681 (79)	1921 (95)	3992 (88)	95 (65)	53 (65)	31 (60)	11 (85)
Tocilizumab	2106 (29)	102 (12)	601 (30)	1403 (31)	72 (49)	35 (43)	35 (67)	2 (15)
Vasopressors	1714 (23)	25 (3)	424 (21)	1265 (28)	10 (7)	1 (1)	9 (17)	0 (0)

Abbreviation: CRS: cytokine release syndrome.

^aStays were classified as their highest known grade during the hospitalization. Classifications are mutually exclusive.^bFever was defined as a diagnosis of fever or a day with acetaminophen without antihistamines.

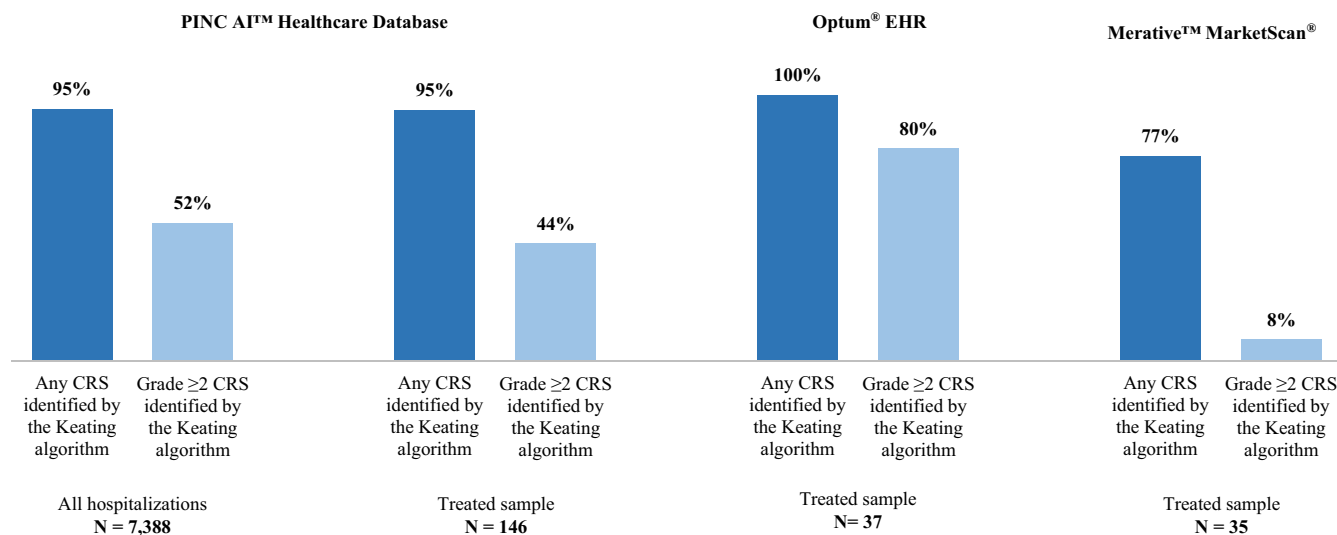


FIGURE 2 Sensitivity of Keating et al.¹ for identifying CRS in retrospective databases. CRS: Cytokine release syndrome; Optum® EHR: Optum® de-identified Electronic Health Record data set. ¹ Keating SJ, Gu T, Jun MP, McBride A. Health Care Resource Utilization and Total Costs of Care Among Patients with Diffuse Large B Cell Lymphoma Treated with Chimeric Antigen Receptor T Cell Therapy in the United States. *Transplant Cell Ther.* 2022 Jul;28(7):404.e1-404.e6. doi: [10.1016/j.jct.2022.03.021](https://doi.org/10.1016/j.jct.2022.03.021). Epub 2022 Mar 27. PMID: 35354101.

TABLE 3 Coefficient estimates for associations between indicators identified by the LASSO model and high-grade CRS among patients with treatment-related CRS.

	LASSO coefficient estimate	Standard errors ^a
Intercept	-0.55	-
Vasopressors	0.31*	0.94
Hypotension	0.30*	0.52
Positive pressure (inc. mechanical ventilation)	0.09*	0.73
Tocilizumab	0.07*	0.70

Abbreviations: CRS: cytokine release syndrome; LASSO: least absolute shrinkage and selection operator.

* $p < 0.05$.

^aStandard errors were generated using the following calculation: (Upper confidence interval-lower confidence interval)/3.92. (Reference: https://handbook-5-1.cochrane.org/chapter_7/7_7_2_obtaining_standard_errors_from_confidence_intervals_and.htm).

or unavailable. The relatively high sensitivity appeared driven by the high proportion of fever cases in all three datasets, a key parameter in CRS diagnosis and the only constant throughout grades 1-4.⁴ Relative to other databases, Merative™ MarketScan® had the lowest sensitivity, also driven by the lower proportion of fever cases observed. In PHD, supplementing the operational definition of fever as “acetaminophen without concurrent antihistamines on the same day,” in addition to using ICD-10-CM codes, was critical in identifying a higher and likely more accurate proportion of symptomatic patients.

Conversely, the Keating algorithm showed a relatively lower sensitivity to detect high-grade CRS in PHD (52%), indicating that symptoms and management strategies used as flags for high-grade

CRS were not often used concurrently with the ICD-10-CM diagnosis code in administrative datasets. Poor sensitivity to high-grade CRS was also shown in Merative™ MarketScan® (8%), where only one of 12 patients with high-grade CRS per ICD-10-CM was accurately identified by the Keating algorithm. This poor performance is likely due to the lack of details regarding the administration of CRS treatments and procedures in the IP setting that are not directly associated with a reimbursement (e.g., corticosteroids). By contrast, sensitivity of the Keating algorithm to identify high-grade CRS in Optum® EHR was higher (80%), potentially driven by the inclusion of comprehensive CRS symptom and management information in this type of data. All patients with high-grade CRS in Optum® EHR had mentions of althralgia and use of corticosteroids, and most had mentions of fever or fatigue. This suggests that the level of granularity available in Optum® EHR, as opposed to other types of retrospective data sources, improves the accuracy of the Keating algorithm. Together, findings indicate that the applicability of using the Keating algorithm to identify cases of high-grade CRS in retrospective data is limited, and other means to qualifying severity of CRS are warranted.

While ICD-10-CM codes for CRS include severity-level granularity, a large proportion of all CRS events (61% for all CRS hospitalizations, 9% for the Treated sample), were reported using the code for an unspecified grade. The use of an ICD-10-CM code with an unspecified grade may be appropriate when there is a lack of available clinical information about the encounter or when there is uncertainty about the definitive diagnosis. However, this creates challenges when evaluating the burden of CRS among patients receiving certain treatments, as CRS varies widely in severity and its implications cannot be inferred from the diagnostic code alone. While the Keating algorithm showed notable performance in detecting any grade CRS, it had relatively poor sensitivity to classify high-grade CRS. The LASSO algorithm developed in this study proposes an alternative to identify likely

TABLE 4 Performance metrics for identifying high-grade CRS among patients with treatment-related CRS.

Performance metrics	PINC AI™ Healthcare Database ^a	Optum® EHR	Merative™ MarketScan®
Model parameters		-	-
λ regularization parameter ^b	0.11	-	-
Cutoff associated with the maximum Youden index ^c	0.39	-	-
Maximum Youden index	0.39	-	-
Threshold-independent			
AUC (c-statistic) ^d	0.75	0.70	0.60
Log Loss	0.62	0.68	0.67
Threshold-dependent			
Accuracy ^e	0.69	0.69	0.68
Sensitivity ^f	0.67	0.80	0.25
Specificity ^g	0.71	0.55	1.00
PPV (precision) ^h	0.55	0.71	1.00
NPV ⁱ	0.80	0.67	0.64

Abbreviations: AUC, area under the curve; CRS, cytokine release syndrome; FN, false negative; FP, false positive; LASSO, least absolute shrinkage and selection operator; NPV, negative predictive value; Optum® EHR, Optum® de-identified Electronic Health Record data set; PPV, positive predictive value; TN, true negative; TP, true positive.

^aPerformance metrics for the LASSO models in the PINC AI™ Healthcare Database were evaluated in the holdout test dataset.

^bThe λ parameter was evaluated using 5-fold cross-validation within the training dataset.

^cThe cutoff was determined by selecting the probability cutoff associated with the maximum Youden Index value calculated in the training dataset. The Youden index is calculated as Sensitivity + Specificity - 1 based on the training dataset.

^dThe AUC, or c-statistic, is generated by plotting the true positive rate (i.e., sensitivity) versus the false positive rate (i.e., specificity) resulting from different threshold values.

^eAccuracy refers to the proportion of correct identifications (true positives and true negatives).

^fSensitivity is calculated as $TP \div (TP + FN)$, and refers to the proportion of actual positives correctly identified by the algorithm.

^gSpecificity is calculated as $TN \div (FP + TN)$, and refers to the proportion of actual negatives correctly identified by the algorithm.

^hPPV (precision) is calculated as $TP \div (TP + FP)$, and refers to the proportion of positive predictions that were actually positive.

ⁱNPV is calculated as $TN \div (TN + FN)$, and refers to the proportion of negative predictions that were actually negative.

cases of high-grade CRS within retrospective data. The model identified hypotension, use of vasopressors, positive pressure, and tocilizumab as variables likely associated with high-grade CRS. This aligned with the ASTCT consensus grading, whereby greater CRS severity is associated with increasing hypotension and hypoxia and progressive vasopressor and oxygen requirements.⁴ Though the preferred timing for administration of anti-cytokine therapy is not included within the grading criteria, tocilizumab is typically administered early

in the management of severe CRS and often preferred over the use of corticosteroids.^{4,5} While performance of the LASSO model was good in the PHD and Optum® EHR databases (AUC:75% and 70%, respectively), the LASSO model performed poorly in Merative™ MarketScan® (AUC:60%). This, once again, is likely due to the lack of details available in administrative claims regarding all components of a patient's presentation and received treatments during hospitalization. Symptoms or treatments that are not directly associated with a reimbursement may not be available in administrative-only IP claims, which represents a challenge when qualifying CRS severity.

Significant advancements in immunotherapies have provided great benefit to patients with hematologic malignancies, notably in B-cell malignancies and multiple myeloma. A better characterization of CRS in the real-world setting is important to help clinicians balance the risks of adverse events with benefits of treatment. This will serve as an important tool to supplement findings from clinical trials, as the applicability of clinical trials in the real world is limited to a carefully selected patient sample and a strict treatment protocol.¹⁴ The current findings detail a methodology that can be used to further inform the incidence and severity of CRS, especially in the context of new upcoming treatments, including CAR T-cell therapies and BsAbs, which represent important treatment options for patients with hematologic malignancies.⁹

4.1 | Limitations

Limitations of this study included the lack of timestamps for diagnoses and treatments in the PHD data, limiting visibility into the temporality of diagnoses posed and procedures received during the hospital patient journey. It is therefore possible that diagnoses or medications may have been administered prior to the occurrence of CRS but were assumed to be CRS-related in the analysis. Similarly, it was not possible to distinguish between prophylactic and therapeutic uses of some treatments (e.g., tocilizumab). Furthermore, results from this study should be interpreted with caution as the sample sizes of the Treated sample in all databases were fairly modest (<150 patients in PHD and <40 patients in Optum® EHR and Merative™ MarketScan® data).

Inherent to any study conducted using retrospective databases, data may be subject to coding errors or omissions. ICD-10-CM diagnostic codes were used as the "gold standard" against which to assess the Keating algorithm. However, it is possible that some cases of CRS were not recorded using ICD-10-CM coding. The analysis was also limited to variables that could be identified in the data source. Physician notes on patient presentation and other unobserved variables, which could have helped to further inform the analysis, were not available.

5 | CONCLUSION

Given the novelty of the ICD-10-CM diagnostic code for CRS, using the Keating algorithm to identify any grade of CRS in

retrospective real-world data continues to be a suitable option but may not be reliable in identifying high-grade CRS. Our research suggests that evidence of hypotension, vasopressors, positive pressure, and tocilizumab are signals of high-grade CRS and may be methodologically useful to inform CRS severity in retrospective data. A critical appraisal of the level of detail available in the data source is necessary, particularly when using administrative claims.

AUTHOR CONTRIBUTIONS

All authors have made substantial contributions to the conception or design of the study, or the acquisition, analysis, or interpretation of data, drafting the manuscript and revising it critically for important intellectual content, and have provided final approval of this version to be published and agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

The authors would like to acknowledge Iman Fakhri for her help with data analysis and methodological support. Medical writing assistance was provided by professional medical writer, Molly Gingrich, an employee of Analysis Group, Inc., a consulting company that has provided paid consulting services to Janssen Scientific Affairs, LLC, which funded the development and conduct of this study and manuscript.

FUNDING INFORMATION

This study was sponsored by Janssen Scientific Affairs, LLC.

CONFLICT OF INTEREST STATEMENT

Scott F. Huntington is an Associate Professor of Medicine (Hematology) at Yale School of Medicine and reports consulting fees from Beigene, Janssen, Pharmacyclics, AbbVie, AstraZeneca, Flatiron Health Inc., Novartis, SeaGen, Genetech, Merck, TG Therapeutics, ADC Therapeutics, Epizyme, Lilly, Ipsen, Servier, Arvinas, and Thyme Inc.; research funding from Celgene, DTRM Biopharm, and TG Therapeutics. Dee Lin, Nina Kim, Laura Hester, Jessica Fowler, Alexander Marshall, Xinke Zhang, Dina Gifkins, and Bingcao Wu are employees of Janssen and may hold stocks or stock options of Johnson & Johnson. Marie-Hélène Lafeuille, Philippe Thompson-Leduc, Aditi Shah, Anabelle Tardif-Samson, and Bronwyn Moore are employees of Analysis Group, Inc., a consulting company that has provided paid consulting services to Janssen Scientific Affairs, LLC.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study were used under license from the PINC AI™ Healthcare Database, Merative™ MarketScan® database, and the Optum® de-identified Electronic Health Record data set. Any researchers interested in obtaining the data used in this study can access these databases under a license agreement, including the payment of an appropriate license fee.

ETHICS STATEMENT

Data used in this study are de-identified and comply with the patient requirements of the Health Insurance Portability and Accountability Act (HIPAA); therefore, institutional review board approval was not sought.

PRIOR POSTINGS AND PRESENTATIONS

Portions of these results were presented at the International Society for Pharmacoepidemiology 39th Annual Meeting held on August 23–27, 2023, in Halifax, NS, Canada, and the Academy of Managed Care Pharmacy 2024 Conference held on April 15–18, 2024, in New Orleans, LA, USA.

ORCID

Philippe Thompson-Leduc  <https://orcid.org/0000-0001-9047-3941>

Dina Gifkins  <https://orcid.org/0000-0002-5645-6213>

REFERENCES

- Cosenza M, Sacchi S, Pozzi S. Cytokine release syndrome associated with T-cell-based therapies for hematological malignancies: pathophysiology, clinical presentation, and treatment. *Int J Mol Sci*. 2021;22(14):7652. doi:10.3390/ijms22147652
- Bobin A, Leleu X. Recent advances in the treatment of multiple myeloma: a brief review. *Fac Rev*. 2022;11:28. doi:10.12703/r/11-28
- Shimabukuro-Vornhagen A, Godel P, Subklewe M, et al. Cytokine release syndrome. *J Immunother Cancer*. 2018;6(1):56. doi:10.1186/s40425-018-0343-9
- Lee DW, Santomasso BD, Locke FL, et al. ASTCT consensus grading for cytokine release syndrome and neurologic toxicity associated with immune effector cells. *Biol Blood Marrow Transplant*. 2019;25(4):625-638. doi:10.1016/j.bbmt.2018.12.758
- Kotch C, Barrett D, Teachey DT. Tocilizumab for the treatment of chimeric antigen receptor T cell-induced cytokine release syndrome. *Expert Rev Clin Immunol*. 2019;15(8):813-822. doi:10.1080/1744666x.2019.1629904
- Kanate AS, Majhail N, DeFilipp Z, et al. Updated indications for immune effector cell therapy: 2023 guidelines from the American Society for Transplantation and Cellular Therapy. *Transplant Cell Ther*. 2023;29(10):594-597. doi:10.1016/j.jtct.2023.07.002
- Frey N, Porter D. Cytokine release syndrome with chimeric antigen receptor T cell therapy. *Biol Blood Marrow Transplant*. 2019;25(4):e123-e127. doi:10.1016/j.bbmt.2018.12.756
- Chen YJ, Abila B, Mostafa KY. CAR-T: what is next? *Cancers (Basel)*. 2023;15(3), 663. doi:10.3390/cancers15030663
- Markouli M, Ullah F, Unlu S, et al. Toxicity profile of chimeric antigen receptor T-cell and bispecific antibody therapies in multiple myeloma: pathogenesis, prevention and management. *Curr Oncol*. 2023;30(7):6330-6352. doi:10.3390/currenco30070467
- Stewart MD, McCall B, Pasquini M, et al. Need for aligning the definition and reporting of cytokine release syndrome (CRS) in immuno-oncology clinical trials. *Cytotherapy*. 2022;24(7):742-749. doi:10.1016/j.jcyt.2022.01.004
- Keating SJ, Gu T, Jun MP, McBride A. Health care resource utilization and total costs of care among patients with diffuse large B cell lymphoma treated with chimeric antigen receptor T cell therapy in the United States. *Transplant Cell Ther*. 2022;28(7):404e1-404 e6. doi:10.1016/j.jtct.2022.03.021
- Premier Applied Sciences® PI. Premier healthcare database white paper: data that informs and performs. 2020 <https://offers.premierinc.com/rs/381-NBB-525/images/PremierHealthcareDatabaseWhitepaper.pdf>

13. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol.* 2010;5(9):1315-1316.
14. Dang A. Real-world evidence: a primer. *Pharmaceut Med.* 2023;37(1):25-36. doi:[10.1007/s40290-022-00456-6](https://doi.org/10.1007/s40290-022-00456-6)

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Huntington SF, Lin D, Lafeuille M-H, et al. Identification of cytokine release syndrome and indicators of severity in retrospective databases among patients receiving immunotherapy. *Pharmacol Res Perspect.* 2024;12:e70024. doi:[10.1002/prp2.70024](https://doi.org/10.1002/prp2.70024)