AI-Powered Pharmacovigilance for Oncology

Shenyao Zhang, Mengqi Liu, Yiming Jia, Yuheng

Center for Data Science, New York University, Genmab Medicines
Mentors: Sky Cheung, Mina Ebeid, Nicole Schmid Davis

Background & Motivation

Clinical Context: Cytokine Release Syndrome (CRS) is an important safety concern for immunotherapies such as CD3×CD20 bispecific antibodies, including epcoritamab. Large pharmacovigilance databases (FAERS, EudraVigilance, JADER) capture real-world safety signals at scale, but the data are often heterogeneous, incomplete, and noisy, and therefore are primarily suited for exploratory signal detection requiring careful clinical interpretation.

Problem: Traditional drug–AE analysis pipelines are slow, difficult to scale across oncology agents, and often generate outputs that are not easily interpretable by clinicians. As a result, rare but clinically important CRS signals may be overlooked.

Goal: Develop a generalized and explainable pharmacovigilance pipeline capable of analyzing any drug–AE pair, using epcoritamab and CRS as the motivating case study. Our system integrates anomaly detection, disproportionality statistics, and interpretable ML to produce clinician outputs.

Primary Databases: FAERS, EudraVigilance, JADER

Causal Inference and NLP Methods for Multi-Database CRS Pharmacovigilance in Epcoritamab

Overview: In this task, we developed unified framework combining causal inference and NLP-driven feature extraction to characterize CRS risk across multiple pharmacovigilance databases. The goal is to quantify the effect of key treatments and clinical factors while leveraging unstructured narratives to improve severe CRS detection.

Methods:

1. Causal Inference

- DAG separating exposure (dose), confounders (age, disease stage, prior therapies), effect modifiers (steroids, tocilizumab), and colliders.
- Association tests: odds ratios, p-values for dose, steroids, tocilizumab, co-medications.
- Propensity score modeling to estimate causal effect of steroids on severe CRS.
- Sensitivity analysis using E-values for robustness to unmeasured confounding.

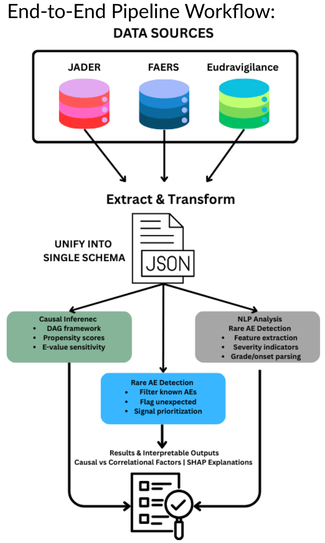
2. NLP & Narrative Features

- Rule-based extraction: ICU/hypotension, intubation, vasopressors.
- Identify steroid/tocilizumab administration, CRS grade, time-to-onset.
- Classifier predicts severe CRS using narrative-derived features.

3. Key Data Summary

- **Propensity score (steroids).** Treated (n=269) vs. controls (n=80); analysis suggests a lower risk of severe CRS among steroid-treated patients, with a wide confidence interval consistent with a protective association.
- **NLP (FAERS narratives).** Analysis of 536 narratives shows limited text detail; severe CRS is identified in 31% of reports, and classifier performance remains near baseline due to sparse narrative information.

End-to-End Pipeline Workflow:



Var	p	OR/SD	Interp
Steroids	0.0024	Sig.	Prot. (causal)
Tocilizumab	0.0267	Sig.	Prot. (IL-6)
Max dose (mg)	0.0277	–	Dose-resp.
Weight	0.0010	Sig.	Conf.
Rituximab	0.0007	Sig.	Corr. (line marker)
Data source	0.0043	–	Heterog.
Age	0.053	1.09	Borderline
Sex (M)	0.065	0.66	Borderline
Co-meds count	0.116	1.78	Corr. (NS)
# doses	0.227	1.28	Likely causal (NS)

Table 1. Statistical Association Analysis (Severe CRS)

Scalable Survival Analysis for Epcoritamab-Associated CRS

Objective: We developed a generalized, parameterized survival analysis pipeline that can evaluate *any* drug–adverse event combination. Using epcoritamab–CRS as the case study, we identify real-world CRS risk factors and characterize temporal onset patterns.

Methods: Cox proportional hazards models were fitted on FAERS 2022–2024 epcoritamab reports with time-to-CRS as the outcome; predictors included age, weight, sex, polypharmacy, and hospitalization, and weight-based subgroups were compared descriptively.

Variable	HR	95% CI	p-value
Weight (per kg)	0.992	0.985–1.000	0.037*
Age (per year)	0.995	0.984–1.006	0.347
Polypharmacy (≥3)	0.616	0.153–2.482	0.495
Prior hospitalization	1.123	0.892–1.413	0.321

Table 2. Cox Proportional Hazards Model Results

Weight Category	Patients	CRS Rate	Clinical Action
<60 kg	148	42.6%	High risk
60–80 kg	404	30.7%	Moderate risk
>80 kg	252	28.6%	Lower risk

Table 3. Weight-Based CRS Risk Stratification

Key Findings & Clinical Implications:

- Body weight is statistically associated with CRS risk (HR = 0.992 per kg, p = 0.037), with lower weight associated with higher observed CRS risk.
- Patients <60 kg show a higher observed CRS rate than those >80 kg (42.6% vs. 28.6%); this finding should be interpreted cautiously given potential confounding.
- All reported CRS events occur within the first 24 hours of epcoritamab, defining a short, high-risk monitoring window.

Rare & Unexpected Signal Detection

Overview: Traditional pharmacovigilance methods often miss rare adverse events (AEs) that emerge post-market, particularly in oncology where novel drug-AE relationships may be unexpected, while we developed an automated pipeline that proactively identifies statistically significant yet clinically unexpected drug-AE signals from large-scale surveillance data, using machine learning anomaly detection with multi-step statistical filtering.

Key Finding: 1,386 validated rare signals with associated risk factors were discovered across 37 oncology drugs

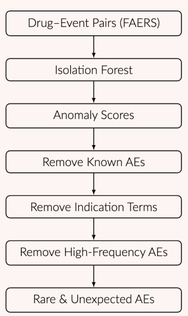


Figure 1. 4-Step Filtering Pipeline: From raw data to rare signals

Workflow: Our pipeline combines Isolation Forest anomaly detection with 4-step filtering (FDA label removal, indication term removal, frequency filtering) and validates signals using disproportionality metrics.

Drug–Event: Epcoritamab + Haemorrhagic gastroenteritis
Status: RARE & UNEXPECTED
Report Count: 1
Statistical Metrics:
PRR (Proportional Reporting Ratio): 111.69 (threshold: >2)
IC025 (Information Component): 3.602 (threshold: >0)
Chi-square: 41.14 (threshold: >4)
Clinical Impact:
Death Rate: 100.0%
Hospitalization Rate: 100.0%
Serious Rate: 100.0%
Assessment:
In No-Cap Results: Yes
Passes PRR Test (>2): Yes
Passes IC025 Test (>0): Yes
Passes Chi² Test (>4): Yes
Passes All 3 Tests: Yes
CONCLUSION:
RARE & UNEXPECTED
(passed IF + all 3 statistical tests)

(a) Rare & Unexpected

Drug–Event: Epcoritamab + Renal impairment
Status: NOT RARE/UNEXPECTED
Report Count: 8
Statistical Metrics:
PRR (Proportional Reporting Ratio): 2.84 (threshold: >2)
IC025 (Information Component): 0.562 (threshold: >0)
Chi-square: 7.68 (threshold: >4)
Clinical Impact:
Death Rate: 137.5%
Hospitalization Rate: 175.0%
Serious Rate: 100.0%
Assessment:
In No-Cap Results: No
Passes PRR Test (>2): Yes
Passes IC025 Test (>0): Yes
Passes Chi² Test (>4): Yes
Passes All 3 Tests: Yes
CONCLUSION:
NOT RARE/UNEXPECTED
(did not pass Isolation Forest)

(b) Not Rare/Unexpected

Risk Factor	AE Group	Control Group	Difference	Conclusion
Age	66.2 years	66.9 years	−0.7 years	Not a risk factor
Sex	Female 56.5%	Female 66.7%	−10.1%	Male associated (OR=1.54)
Medical History: DLBCL refractory	95.8%	15.4%	+80.4%	High risk factor
Medical History: Prophylaxis	116.7%	76.9%	+39.7%	Risk factor
Concomitant Drug: Prednisolone Oral	33.3%	0.0%	+33.3%	Risk factor

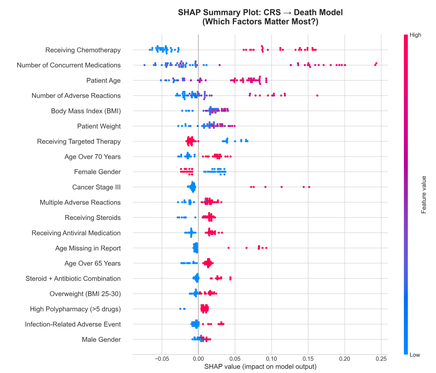
(c) Risk Factor Analysis

Figure 2. Three Example Outputs from Task 3 Analysis

CRS Mortality Modeling for Epcoritamab Using FAERS

Overview: We built a parameterized FAERS pipeline to identify CRS cases for Epcoritamab and predict CRS related mortality using machine learning. The workflow converts raw safety reports into structured clinical features and produces interpretable outputs through SHAP. Because drug and adverse event are passed as inputs, the pipeline can be reused for other therapies and safety signals.

Method: CRS cases were extracted from FAERS and enriched with engineered features including age, BMI, comorbidities, chemotherapy exposure, polypharmacy and reaction profiles. A Random Forest model was trained to predict death among CRS patients and evaluated using ROC AUC, PR AUC and F1 score. SHAP analysis was applied to quantify how each feature increases or decreases predicted mortality risk.



Summary: The model achieved a PR AUC of 0.897 and an F1 score of 0.833. The SHAP summary plot shows that chemotherapy exposure, high numbers of concurrent medications and older age have the strongest positive impact on predicted mortality. Features such as lower BMI, fewer comorbidities and limited polypharmacy generally push predictions toward survival. These patterns provide interpretable, clinically meaningful insights into CRS outcomes.

Conclusion & Future Work

We present a scalable AI-driven pharmacovigilance pipeline for proactive safety signal detection, integrating anomaly detection, survival analysis, and interpretable machine learning. Applied to epcoritamab-associated CRS, the framework identifies rare adverse event signals and factors associated with CRS risk. All findings are hypothesis-generating and intended to support clinical review and signal prioritization rather than causal inference. The pipeline is generalizable to other drug–adverse event pairs.

- **Biomarker integration:** Incorporate biomarker data (e.g., IL-6, ferritin) to enrich risk modeling and improve clinical context.
- **Prospective validation:** Evaluate findings using prospective clinical datasets to assess robustness and generalizability.
- **Automated surveillance:** Extend the framework toward near–real-time safety signal monitoring to support earlier signal prioritization.

Data Limitations & Disclaimer

FAERS, EudraVigilance, and JADER are spontaneous reporting systems with important limitations. Reports are subject to substantial underreporting, missing or incomplete clinical information (e.g., dose, time to onset, comorbidities, laboratory data), reporting bias, and duplicate submissions from multiple sources. These databases do not provide exposure or denominator data, precluding incidence estimation and comparative risk assessment. Reported associations reflect suspected relationships and do not establish causality. All findings should therefore be interpreted as hypothesis-generating and require careful clinical review and validation in controlled studies.