Additional file 1 — Supplementary methods and definitions

Article title: A Payload- and Target-Informed Toxicity Atlas of Antibody–Drug Conjugates: A Disproportionality Analysis of the FDA Adverse Event Reporting System

Journal: BMC Cancer

Authors: Engin Çeşmeci, Berkay Yeşilyurt, Mesut Yılmaz, Emre Çakır, İlhan Hacıbekiroğlu

Affiliations: 1 Department of Medical Oncology, Faculty of Medicine, Sakarya University, Sakarya, Türkiye; 2 Medical Oncology Clinic, Ankara Bilkent City Hospital, University of Health Sciences, Ankara, Türkiye

Corresponding author: Engin Çeşmeci – engincesmeci@sakarya.edu.tr

Supplementary methods and definitions

A Payload- and Target-Informed Toxicity Atlas of Antibody-Drug Conjugates

A disproportionality analysis of the FDA Adverse Event Reporting System

S1. Data source and study period

Quarterly FDA Adverse Event Reporting System (FAERS) data files covering 2012Q4 through 2026Q1 were used. The data were first accessed on 1 May 2026 and downloaded in XML format from the US Food and Drug Administration FAERS Quarterly Data Extract Files interface. Drug, reaction, patient, seriousness, outcome, and reporting-quarter information were linked within a local SQL analytical database.

The report was the analytical unit and was identified by safetyreportid. Because FAERS distributes successive follow-up versions of the same case, each case was represented once by the record carrying its highest numeric safetyreportversion; where the same version occurred in more than one quarterly file, the record was chosen by the latest source quarter and then the receipt date, with the source file name and report ordinal as deterministic tie-breakers. Child drug, reaction, patient and outcome records were attached to the selected report through a unique report-instance identifier rather than by case number alone, so that rows belonging to superseded versions could not inflate report-level counts. Report-level numerators and denominators were therefore based on distinct retained reports, preventing multiple drug–reaction rows belonging to the same report from being counted repeatedly within an analysis cell.

Cases appearing in the FDA deletion (nullification) files, available from 2019Q1 onward, were excluded under a historical-union rule applied through the 2026Q1 cutoff; 96,109 such cases were removed. After case selection and this exclusion the analysis dataset contained 17,223,612 reports, which formed the denominator for all disproportionality estimates. The executed database-integrity controls and the complete case-selection and deletion rules are documented in Additional file 3.

S2. ADC identification and study population

A curated antibody–drug conjugate dictionary was used to identify eligible reports. The dictionary incorporated generic names, brand names, active-substance terms, and standardized ADC keys. Candidate ADCs observed from 2012Q4 onward were screened, and agents with at least 100 primary-suspect reports were included in the main toxicity landscape. The threshold was

prespecified to reduce instability caused by sparse reporting; agents below the threshold remained documented in Supplementary Table S1.

Analyses were restricted to reports in which the ADC was designated as the primary suspect drug. This restriction increased exposure specificity and reduced confounding from malignancy-related treatment and supportive-care polypharmacy, while potentially decreasing sensitivity when a relevant ADC was coded as secondary suspect or concomitant.

S3. Target and payload classification

Each included ADC was assigned to a target-antigen category, payload class, and specific payload name using the curated project dictionary. Drug-level estimates constituted the most specific analytical layer. Target- and payload-level analyses pooled reports across ADCs sharing the relevant category and were interpreted as platform-level reporting patterns rather than isolated molecular effects. The complete ADC–target–payload mapping is provided in Supplementary Table S1.

S4. Toxicity-cluster definitions

Clinically defined toxicity clusters were constructed from MedDRA Preferred Terms for interstitial lung disease/pneumonitis, neuropathy, neutropenia/hematologic toxicity, thrombocytopenia/bleeding, hepatotoxicity, gastrointestinal toxicity, skin toxicity, ocular toxicity, and hyperglycemia. Complete cluster membership is provided in Supplementary Table S2.

The neutropenia/hematologic and thrombocytopenia/bleeding dictionaries were maintained as analytically distinct term sets without shared Preferred Terms. Conceptual overlap remains possible at report level because marrow suppression, cytopenias, and bleeding manifestations may coexist. A report could contribute to more than one toxicity cluster when distinct qualifying Preferred Terms were recorded.

S5. Clinical Preferred Term filtering

Administrative, utilization-related, disease-context, intercurrent-infection, non-toxicity outcome, and other contextual Preferred Terms were excluded from the reader-facing clinical adverse-event profile using a prespecified project dictionary. The excluded term, exclusion category, and rationale are reported in Supplementary Table S3. Unfiltered Preferred Term outputs were retained separately during the analysis but were not used to define the filtered clinical profile.

S6. Disproportionality analyses

Reporting odds ratios (RORs) and proportional reporting ratios (PRRs) were calculated from 2 × 2 contingency tables. For each ADC–toxicity combination the exposed-event cell represented reports in which the ADC was recorded as primary suspect and the toxicity cluster was present; the comparator comprised all remaining eligible FAERS reports in the study period, not reports for other ADCs alone. Drug-, payload-, and target-level estimates are presented in Supplementary Tables S4–S6.

Combinations with fewer than three exposed event reports were not interpreted as signals. Positive disproportionality was defined by at least three exposed event reports and an ROR lower 95% confidence bound greater than 1. Because the study was designed for exploratory signal detection

across correlated toxicity phenotypes, the principal presentation emphasises effect estimates and confidence intervals; to quantify the influence of multiple testing, a Benjamini–Hochberg false-discovery-rate procedure was applied within each analytical family (drug, payload, target and serious-only), with two-sided p values obtained from Wald tests on the natural logarithm of the ROR using the same standard error as the reported confidence intervals. Adjusted q values are reported in Supplementary Table S11. Lower-magnitude findings were interpreted cautiously and in conjunction with clinical plausibility, confidence-interval precision and the sensitivity analyses. ROR and PRR values quantify disproportionate reporting and do not estimate incidence, prevalence, attributable risk, causality, or comparative safety.

S7. Direct comparisons and sensitivity analyses

Biologically prespecified ADC pairs were compared to determine whether agents sharing a target, payload class, or clinically relevant toxicity hypothesis retained distinguishable reporting profiles. Comparative RORs greater than 1 indicate enrichment in the first ADC named in the contrast, whereas values below 1 indicate enrichment in the comparator. The complete comparison set is presented in Supplementary Table S8.

A serious-only sensitivity analysis repeated drug-level disproportionality calculations after restricting the dataset to reports coded as serious. Concordance with the primary analysis was interpreted as support for robustness to the seriousness restriction rather than causal confirmation. Complete estimates are presented in Supplementary Table S7.

Two further sensitivity analyses assessed the report-handling rules described in S1. In the first, every source version of each retained case contributed to its drug and reaction content instead of the selected version alone; in the second, cases listed in the FDA deletion files were retained rather than excluded. Both used the same ADC dictionary, toxicity-cluster definitions, contingency-table structure and signal criterion as the primary analysis and were restricted to the drug level. Complete estimates are presented in Supplementary Table S14.

S8. Outcome-reporting summaries

Fatal outcome and hospitalization summaries were calculated as the proportions of selected ADC–toxicity reports containing the corresponding FAERS outcome indicator. Cells with fewer than 20 ADC–toxicity reports were omitted from the heatmaps to reduce instability from sparse denominators. These measures are outcome-reporting proportions within selected FAERS reports and must not be interpreted as mortality, case-fatality, hospitalization incidence, or hospitalization risk. Seriousness outcomes are not mutually exclusive, and a report may contribute to more than one outcome category. Supporting counts and proportions are reported in Supplementary Table S9.

S9. Temporal analysis

Quarterly reporting proportions were calculated as toxicity-cluster reports divided by all eligible reports for the corresponding ADC and quarter. For each toxicity cluster, the three ADCs with the largest cumulative cluster-report counts were displayed. No minimum quarterly denominator was imposed and no formal trend test was performed; temporal findings are descriptive, and early or low-volume quarters may yield unstable proportions.

S10. Software and reporting framework

Data linkage and report-level aggregation were performed in DuckDB version 1.5.2 accessed through DBI version 1.3.0; statistical summaries and visualizations were produced using R version 4.6.0 (R Foundation for Statistical Computing, Vienna, Austria). Reporting and interpretation follow the READUS-PV recommendations for disproportionality analyses of individual case safety reports; the completed checklist is provided as Additional file 4.