

# From Data to Decisions: FAERS-Based Machine Learning for Predicting Severe Outcomes in Breast Cancer Treatment Reports

**Mohammadhossein Amini\***

*Olin Business School, Washington University in St. Louis*

**Zihang Shi\*\***

*McKelvey School of Engineering, Washington University in St. Louis*

**Arefeh Esmailpour\*\*\***

*School of Medicine, Saint Louis University,*

**Abstract:** This study uses the FDA Adverse Event Reporting System (FAERS) to develop and evaluate a machine-learning model that predicts severe outcomes in breast cancer treatment reports. By combining patient demographics, treatment plans, and adverse event reports, the model detects high-risk cases through machine learning models. Unlike traditional methods that depend on clinical trials, this paper uses the FAERS database, which provides large-scale real-world safety signals that reflect heterogeneous patients, settings, and reporting behavior. We describe a transparent data preparation pipeline from raw FAERS extracts to an analysis cohort, define severity using standardized FAERS outcome codes, and evaluate multiple classification algorithms using cross-validation testing. Results show that a supervised learning approach using routinely available report features can meaningfully discriminate fatal from non-fatal outcomes, supporting real-world risk stratification as an input to pharmacovigilance and decision support, while highlighting important limitations of spontaneous reporting data, including missingness, reporting bias, and confounding.

**Keywords:** FAERS, Breast Cancer, Machine Learning, Pharmacovigilance, Adverse Events

---

\* Corresponding author. Email: [h.amini@wustl.edu](mailto:h.amini@wustl.edu), \*\* Email: [s.zihang@wustl.edu](mailto:s.zihang@wustl.edu)

\*\*\* Email: [arefeh.es@gmail.com](mailto:arefeh.es@gmail.com)