Machine Learning for Frailty Stratification in Hematologic Malignancies

A Systematic Review of Predictive Models for Treatment Toxicity and Survival

Youmna Ismail[1] • Sara Anwar Ali[1] • Lena Mohamed[1] • Ahmed Faisal[1] • Ahad Magzoub[1] • Alaa Mohamed[1] • Mohamed Elmahi[1] • Malaz Salah[1] • Nadir Abdelrahman[1]

[1]*Tele-Geriatric Research Fellowship, Okemos, MI, USA*

18
Studies Included in Final Synthesis

0.64–0.90
AUC range for toxicity & survival

100%
Studies with Internal Validation

61%
Studies with External Validation

Background

Frailty is a key predictor of adverse outcomes in older patients with hematologic malignancies, including treatment-related toxicity and reduced overall survival.

Machine learning offers a robust way to stratify frailty and predict adverse events beyond traditional geriatric assessment.

Objective

Evaluate the performance of ML models in predicting treatment toxicity and survival in frail older patients (age 60+) with hematologic malignancies ,focusing on predictive accuracy, validation, and risk of bias .

Population of Interest

60+

Older adults with hematologic malignancies undergoing frailty-informed treatment decisions.

Data Inputs to ML Models

- Comprehensive geriatric assessment domains
- Laboratory and hematologic biomarkers
- Comorbidity and functional-status indices
- Treatment and disease-specific clinical variables

Design & Methods

A systematic review conducted in accordance with PRISMA-2020 guidelines. The literature search spanned PubMed, Cochrane Library, and Google Scholar from database inception through 2026.

Outcomes Assessed

- Primary: predictive performance of ML models for overall survival
- Secondary: model validation approach (internal vs. external)
- Secondary: explainability via SHAP and LIME techniques
- Secondary: real-world clinical applicability of models

Quality Appraisal

Study quality and risk of bias were assessed using the PROBAST-AI tool.

Records Identified (PubMed, Cochrane Library, Google Scholar)

Screened for Eligibility

18 Studies Included in Synthesis

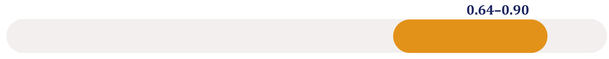
Results

18 studies met inclusion criteria for the final qualitative and quantitative synthesis, applying a range of ML techniques to frailty stratification.

ML Techniques Used

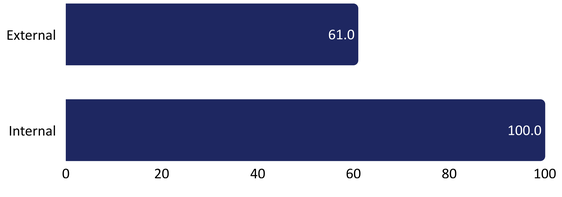
- Random Forest models
- Gradient boosting approaches (e.g., XGBoost)
- Neural networks
- Regression-based ML algorithms

Predictive Performance (AUC)



Reflects moderate-to-excellent discrimination for treatment toxicity and survival outcomes.

Model Validation Approach



Conclusions

ML-based frailty stratification models show strong potential for predicting treatment toxicity and survival in older adults with hematologic malignancies, with moderate-to-excellent performance (AUC up to 0.90).

Why It Matters

- Explainable AI, particularly SHAP, improves transparency and supports clinical trust in model outputs
- Findings support a growing role for ML in precision geriatric hematology and risk-adapted care
- Heterogeneity remains in ML methodologies and frailty definitions across studies

Future Directions

- Standardize frailty definitions across ML studies
- Expand external validation in independent cohorts
- Prospective testing of ML tools in real clinical workflows

Contact information

Email: saraanwar99@gmail.com

ACCURACY

ML models discriminate treatment toxicity and survival risk with AUCs up to 0.90, outperforming conventional geriatric scoring alone.

TRANSPARENCY

SHAP and LIME translate model output into clinician-interpretable risk drivers, supporting trust at the bedside.

READINESS

With low predominant risk of bias, ML frailty tools are approaching readiness for risk-adapted hematology care pathways.