Equity in Post-HCT Survival Predictions

Sergio Nicolás Mendivelso Martínez - 20231020227
Sergio Leonardo Moreno Granado - 20242020091
Juan Manuel Otálora Hernández - 20242020018
Juan Diego Moreno Ramos - 20242020009

Professor: Eng. Carlos Andrés Sierra, M.Sc.

Understanding the Medical Challenge

What is HCT?

Allogeneic Hematopoietic Cell Transplantation replaces damaged bone marrow with healthy stem cells from a compatible donor. It treats serious blood diseases like leukemia, lymphomas, and severe aplastic anemia.

Why Prediction Matters

Accurate survival prediction helps physicians make better transplant decisions, enables personalized treatment protocols, and improves healthcare resource utilization while reducing health inequalities.

The Dual Challenge: Accuracy and Equity

Prediction Accuracy

Models must correctly order risk among patient pairs using comprehensive clinical, genetic, and demographic data to forecast survival outcomes.

Fairness Requirement

The stratified C-index ensures consistent performance across all ethnic subgroups, preventing models from disadvantaging patients based on demographics.

Equity Impact

Fair predictions ensure all patients receive accurate care regardless of race, gender, or socioeconomic status, directly addressing healthcare disparities.

System Complexity

Multiple interacting factors–age, disease stage, genetic compatibility, comorbidities–create nonlinear relationships where small parameter changes significantly impact predictions.

Sensitive Parameters

- Patient age variations
- Disease risk indices
- HLA matching degree
- Comorbidity presence

Inherent Randomness

Biological diversity, incomplete records, and hidden genetic variables introduce stochasticity that even advanced models struggle to capture fully.

Module Functions & Data Flow

01

Preprocessing (M1)

Receives medical charts, documents, and RAM data. Performs data preprocessing and feeds M2 and M3.

02

Equity Analysis (M2)

Analyzes equity from M1 output, providing fairness insights to the modeling core.

03

Feature Selection (M3)

Identifies important features from preprocessed data for predictive modeling.

04

Modeling Core (M4)

Central hub performing predictive modeling, fairness calibration, and uncertainty quantification with feedback loops.

05

Fairness Calibration (M5)

Calibrates model outputs for fairness, feeding adjustments back to M4.

06

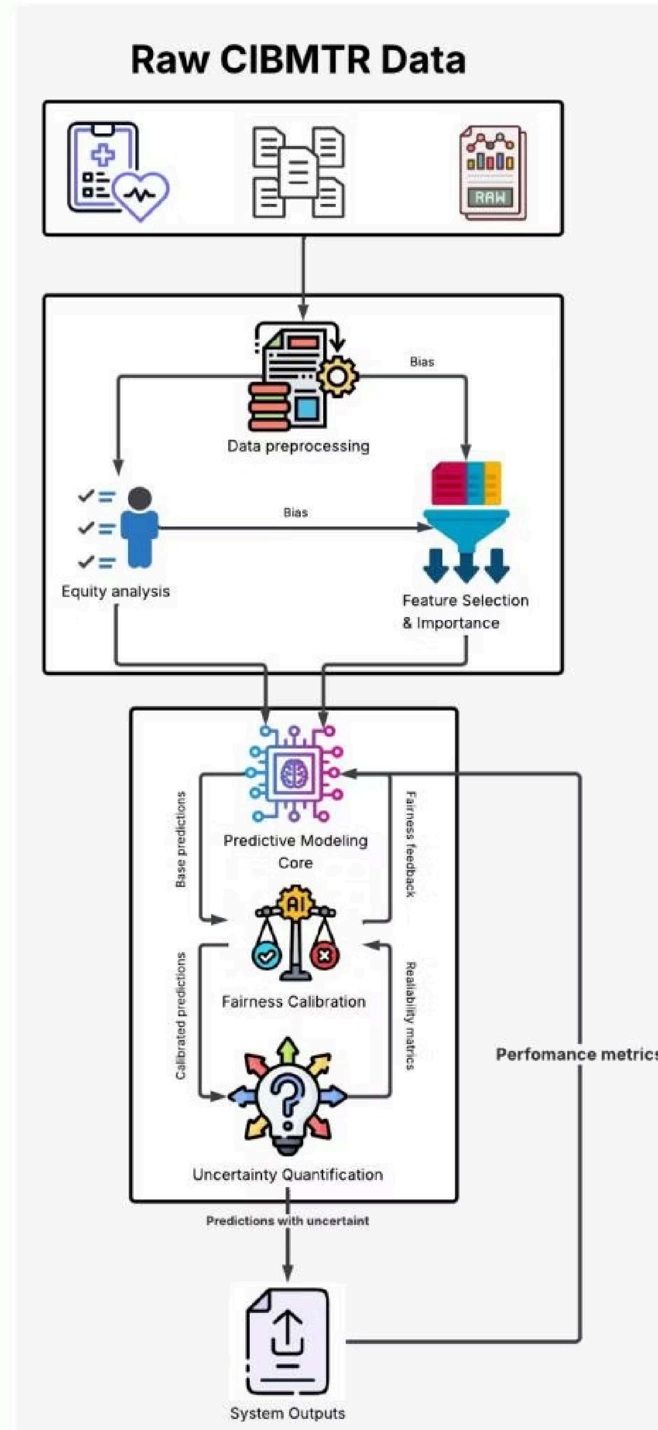
Uncertainty Quantification (M6)

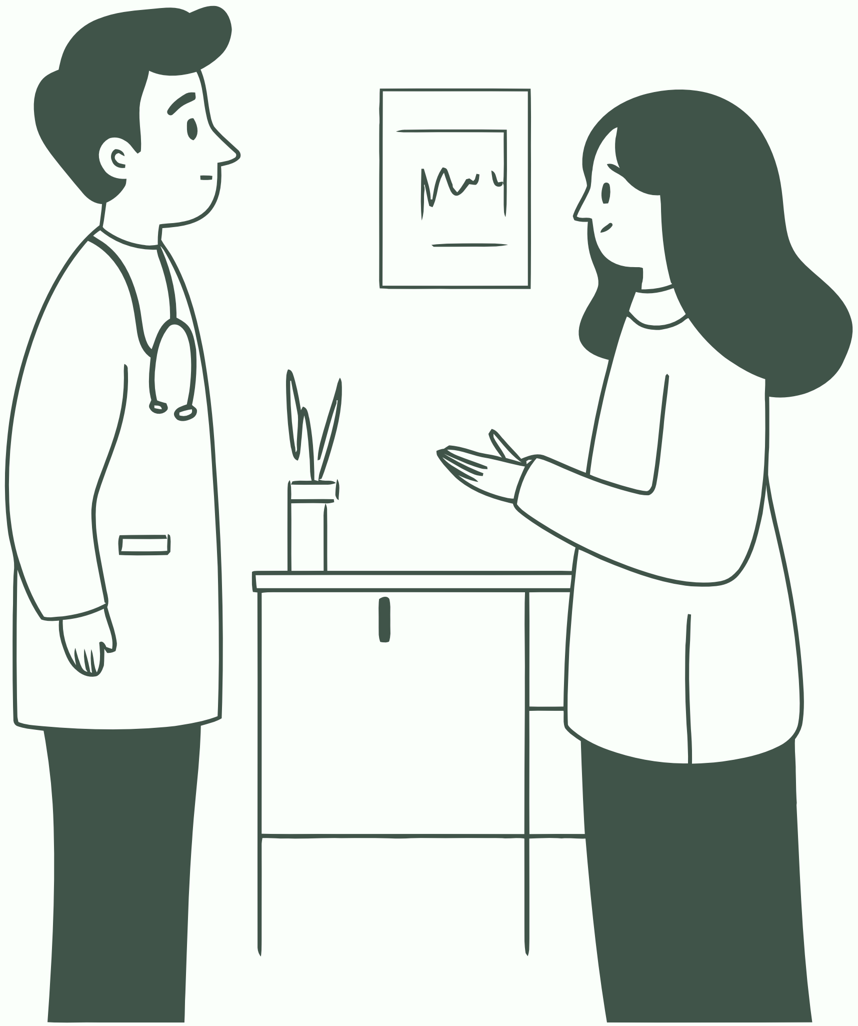
Quantifies prediction reliability and uncertainty from M4 outputs.

07

Final Outputs (M7)

Delivers predictions with uncertainty metrics, creating feedback loops to M4 and M6.





Data Preparation & Success Metrics

CIBMTR Dataset

28,803 patient records with 59 clinical features from Kaggle competition. Event-Free Survival (EFS) binary classification with 53.12% event rate.

- Missing values: median/mode imputation
- Categorical variables: label-encoded
- Numerical features: z-score normalized

Target Metrics

Defined success criteria for validation:

- Accuracy ⩾ 0.70
- AUC-ROC ⩾ 0.70
- Coefficient of Variation ⩽ 0.15
- Emergent pattern observation

Scenario 1: Machine Learning Simulation

Gradient Boosting Machine

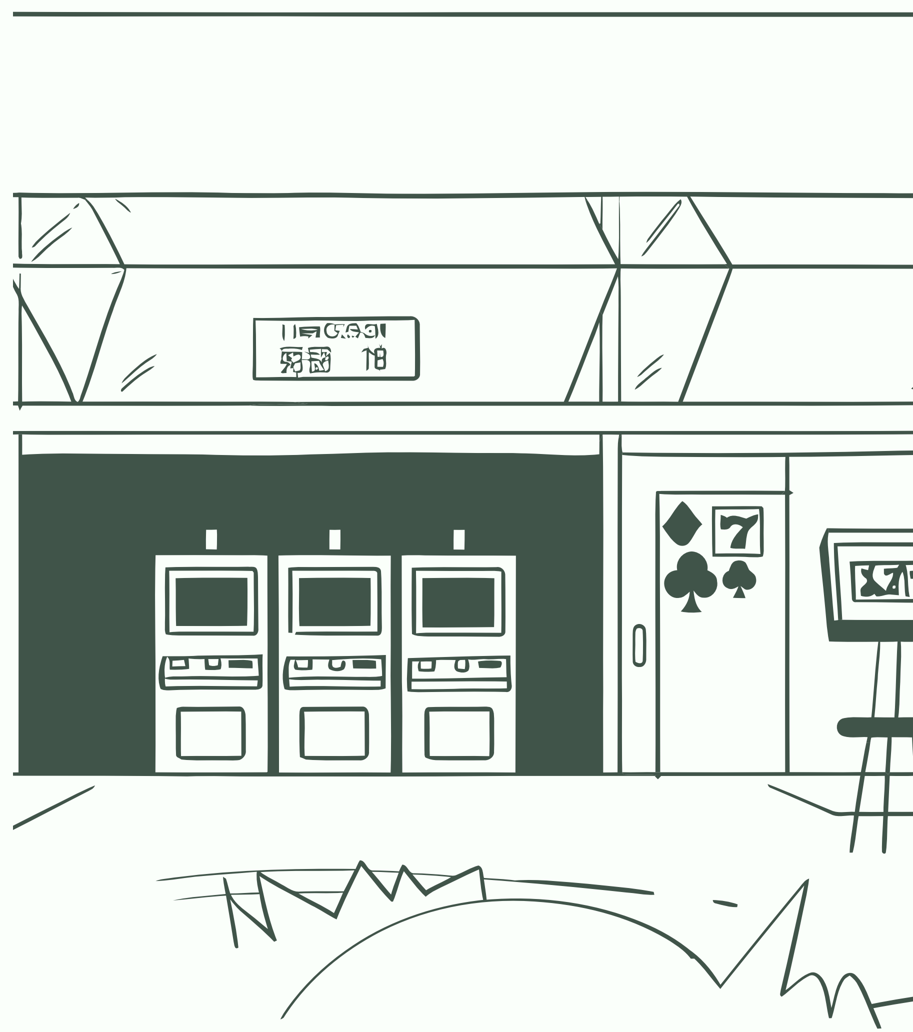
Selected for handling mixed data types, robustness to overfitting, and clinical interpretability through feature importance rankings.

Performance Results

Mean accuracy: 67.84%, AUC: 0.7391 across 5 iterations. Stable behavior (CV = 0.012) but below 70% target.

Chaos Sensitivity

Graceful degradation under 15% noise—only 4.7% predictions changed. Critical for clinical data quality issues.



Scenario 2: Cellular Automata Simulation

Model Design

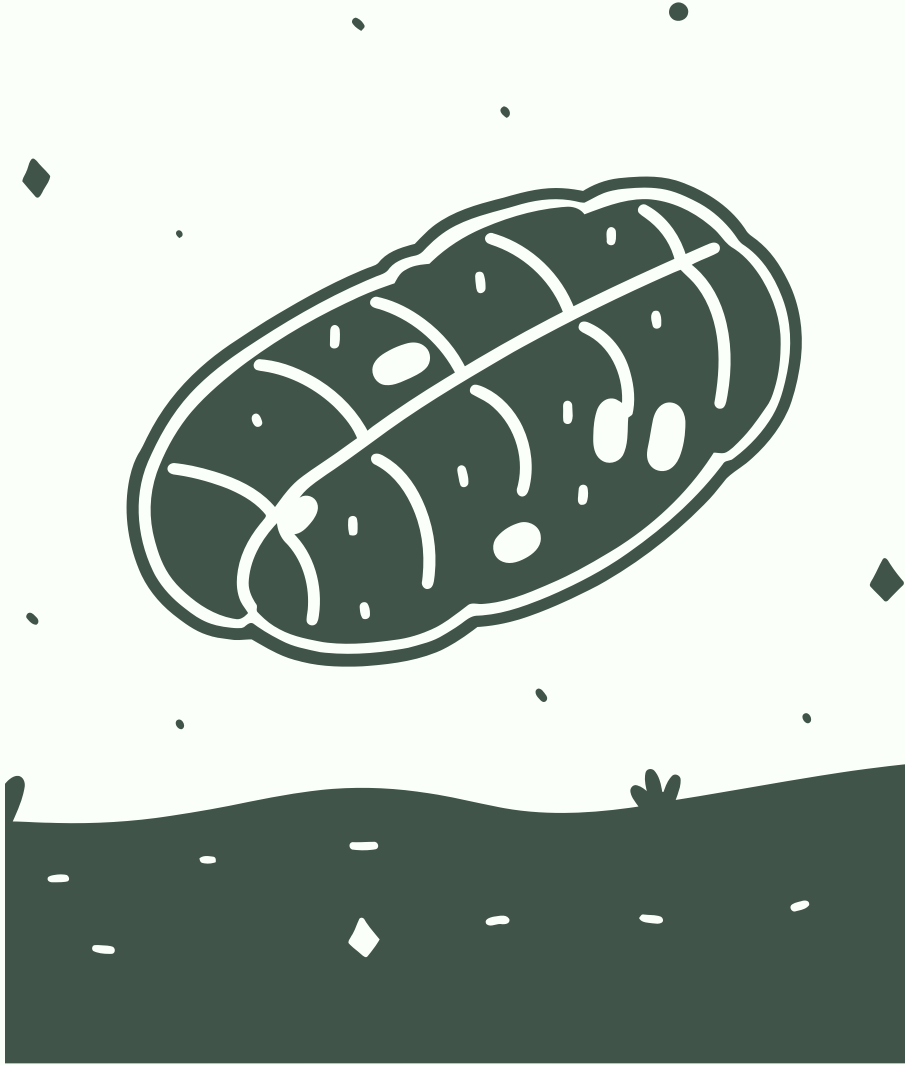
40×40 toroidal grid (1,600 cells) with three patient states:

- **Stable:** No complications
- **At Risk:** Requires monitoring
- **Event:** Death/relapse (absorbing)

Critical Finding

Complete system collapse to 100% event rate by Step 20, demonstrating phase transition from mixed states to full absorption.

All four parameter scenarios (Baseline, High Recovery, High Progression, High Chaos) converged identically–indicating **supercritical regime** where initial conditions dominate dynamics.



Architecture Validation & Equity Considerations

Module M1: Data Pre-processing

Successfully handled missing values and feature encoding for 28,803 records. Extracted 53.1% event rate for CA initialization.

Module M4: Predictive Core

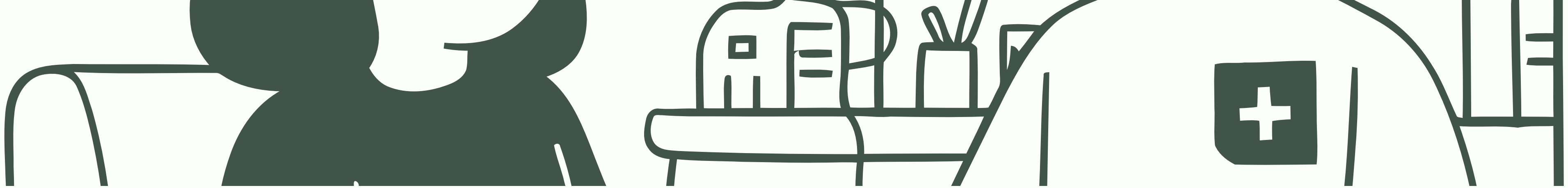
GBM achieved AUC = 0.7391 with stable performance. SHAP analysis revealed conditioning intensity, sex match, and year of HCT as top predictors.

Module M6: Uncertainty Quantification

Chaos sensitivity showed 4.7% prediction change at 15% noise. CA demonstrated high uncertainty with inevitable collapse despite parameter variations.

Equity Analysis

Race group appears in top 20 influential features. Future work requires demographic parity evaluation using Fairlearn framework to prevent systematic bias.



Regime Analysis: Clinical Implications

Supercritical (>50%)

Rapid, inevitable collapse. Observed in simulation with real clinical data. Prevention and early intervention essential.

Critical (20-50%)

Delayed collapse with moderate parameter sensitivity. Extended time horizons before full absorption.

Subcritical (<20%)

Mixed equilibrium possible. Highest parameter sensitivity makes intervention strategies meaningful.

Transplant programs managing high-risk cohorts face fundamentally different challenges than lower-risk populations. Once cascade begins in supercritical regime, parameter modifications cannot reverse trajectory.