

Project Summary Report

YODA Project Protocol #: 2022-5013

Identification of the Need for Hematopoietic Stem Cell Transplantation in Crohn's Disease Patients: A Machine Learning Approach

Status: Analysis completed and article has been prepared for dissemination, publication remains pending.

Abstract from article prepared by investigators:

Objective

To study the feasibility of using explainable machine learning algorithms to predict the adherence of Crohn's Disease (CD) patients to the Brazilian protocol for hematopoietic stem cell transplantation.

Methods

Three widely used interpretable machine learning models from the medical literature were evaluated: Decision Trees, Naïve Bayes, and Logistic Regression.

Results

Among the three models, the Decision Tree model showed the best performance, with an average accuracy of 70% and 60% macro-F1, standing out as a promising tool to support healthcare professionals.

Conclusions

The average results across 500 repetitions indicated that the Decision Tree achieved the best performance in terms of Accuracy and Macro-F1, with no significant impact from class balancing. Our conclusions, however, are limited by the adopted dataset, which included only patients with active Crohn's disease, refractory to at least one biologic medication, thus introducing a bias in the data distribution, as well as its size, which limited the choice of methods.