

Infections were the cause of death in 9.8% of patients and in 39% of those with severe infections ($p < 0.001$). **Discussion:** Infectious complications were associated with decreased OS of patients treated with azacitidine. Infection rates in the first three cycles were consistent with existing literature. Despite bacterial prophylaxis, no bacterial resistance was found. Thrombocytopenia was the main variable associated with higher infection risk, as previously described in previous studies. **Conclusion:** In this unicentric outpatient study, patients with MDS and thrombocytopenia are at greater risk of developing infection, especially in the first cycles and should be monitored cautiously. Although infection is associated with reduced survival, death generally occurs due to other causes, such as disease progression.

<https://doi.org/10.1016/j.htct.2024.09.762>

NOVEL KEY VARIABLES IN THE SURVIVAL OF PATIENTS WITH MYELOYDYSPLASTIC NEOPLASMS: A PRACTICAL APPROACH USING MACHINE LEARNING

PRC Passos^a, RDB Dias^{a,b}, SCC Carneiro^a, IB Nogueira^a, JMGF Lima^a, RC Venâncio^a, ACG Lavor^a, JVG Gama^a, RF Pinheiro^{a,b,c}, SMM Magalhães^{a,c}

^a Center for Research and Drug Development (NPDM), Universidade Federal do Ceará (UFC), Fortaleza, Brazil

^b Postgraduate Program in Pathology, Universidade Federal do Ceará (UFC), Fortaleza, Brazil

^c Department of Clinical Medicine, Universidade Federal do Ceará (UFC), Fortaleza, Brazil

Introduction: Prognostic models like the IPSS-R play a crucial role in assessing outcomes for patients with myelodysplastic neoplasms (MDS). However, recent advancements in machine learning (ML) offer the potential to uncover novel predictive variables and enhance prognostic accuracy. Models like ElasticNet are particularly adept at handling multidimensional data, thereby expanding the scope beyond the variables considered in IPSS-R. **Objectives:** Assessing the performance of ML in predicting overall survival in MDS patients by incorporating clinical and hematological variables not traditionally included in prognostic models. **Methods:** We conducted a retrospective cohort study at a single reference center involving patients diagnosed with MDS between 2004 and 2024. We included patients with available clinical outcomes, missing data was handled using the 'cart' multiple imputation, following confirmation of non-random missingness through Little's test. The dataset was then randomly split into a training group (70%) and a testing group (30%). Utilizing group elastic net machine learning, an artificial intelligence model capable of selecting relevant variables and assessing their discriminative power, we constructed 3 receiver operating characteristic (ROC) curves to predict 1, 3, and 5-year survival, extracting the area under the curve (AUC) and identifying variables with non-zero coefficients. Based on these coefficients, we categorized our dataset into "High Risk" and "Low Risk" groups. Subsequently, we

conducted a multivariate Cox proportional hazard regression analysis, adjusting for the new risk variable, age at diagnosis, sex, and transfusion burden. All statistical analyses were performed using R, with the involvement of packages such as 'mice', 'ggplot', 'glmnet', and 'survfit'. **Results:** 162 patients were included in this study. Using the group ElasticNet model, we identified 10 critical variables with notable predictive power: hemoglobin count, mean corpuscular volume, platelet count, presence of dysgranulopoiesis, presence of dysmegakaryopoiesis, serum iron, transferrin saturation, bone marrow cellularity, percentage of blasts in bone marrow, and percentage of ring sideroblasts. ROC curve analysis utilizing these variables' coefficients yielded AUCs of 0.863, 0.822, and 0.719 for predicting 1-year, 3-year, and 5-year survival, respectively. The coefficients were then extracted and used for risk stratification. 5-year survival rates were 24.3% for the High Risk group and 70.4% for the Low Risk group ($p < 0.001$, log-rank test). In multivariable Cox regression analysis, the risk group variable was the most discriminative predictor (HR = 3.56, $p < 0.001$), with sex, age at diagnosis, and transfusion burden also being significant ($p = 0.01$, 0.009, and 0.002, respectively). **Discussion:** The strong performance of the model, as evidenced by the ROC curve analysis, suggests that the selected variables offer substantial discriminative power. ML models offer more refined risk stratification than traditional methods, which may be useful in identifying occult relationships between variables. Validation in independent datasets may be necessary to strengthen the relationships herein exhibited. **Conclusion:** ML is a valuable tool for risk classification and survival prediction, offering significant insights for clinical decision-making and patient management that may be overlooked by other methods.

<https://doi.org/10.1016/j.htct.2024.09.763>

BLASTING BEYOND THE OBVIOUS: IDENTIFICATION OF OVERLOOKED PREDICTIVE FACTORS FOR LEUKEMIC CONVERSION IN MYELOYDYSPLASTIC NEOPLASMS

PRC Passos^a, RDB Dias^{a,b}, SCC Carneiro^a, IB Nogueira^a, JMGF Lima^a, RC Venâncio^a, ACG Lavor^a, JVG Gama^a, RF Pinheiro^{a,b,c}, SMM Magalhães^{a,c}

^a Center for Research and Drug Development (NPDM), Universidade Federal do Ceará (UFC), Fortaleza, Brazil

^b Postgraduate Program in Pathology, Universidade Federal do Ceará (UFC), Fortaleza, Brazil

^c Department of Clinical Medicine, Universidade Federal do Ceará (UFC), Fortaleza, Brazil

Introduction: Myelodysplastic neoplasms (MDS) are a group of clonal hematopoietic disorders characterized by dysplasia in one or more myeloid cell lines, ineffective hematopoiesis, and an increased risk of progression to acute myeloid leukemia (AML). The transition from MDS to AML is a pivotal event that significantly impacts patient outcomes. While the blast