



## The 65th ASH Annual Meeting Abstracts

## POSTER ABSTRACTS

**623.MANTLE CELL, FOLLICULAR, AND OTHER INDOLENT B CELL LYMPHOMAS: CLINICAL AND EPIDEMIOLOGICAL****A Predictive Model Based on Machine Learning for Early Progression/Relapse of Follicular Lymphoma**Yingyue Liu<sup>1</sup>, Xiangxiang Zhou<sup>2</sup>, Linqun Zhan<sup>1</sup>, Mengyuan He<sup>1</sup>, Tiange Lu<sup>1</sup>, Xin Wang<sup>2,1</sup><sup>1</sup>Department of Hematology, Shandong Provincial Hospital, Shandong University, Jinan, China<sup>2</sup>Department of Hematology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China

**Introduction:** Management of follicular lymphoma (FL) improved dramatically in the last decade, and progression/relapse of disease within 24 months (POD24) is now established as a predictor for poor survival. Risk prediction of early POD (E-POD) prior to the start of therapy may enhance clinical management and research strategies. We aimed to establish a comprehensive model based on machine learning (ML) that could predict the E-POD of FL patients for earlier identify such patients earlier.

**Methods:** In this study, patients with a diagnosis of FL at the Shandong Provincial Hospital between January 1, 2010, to May 1, 2023, were identified retrospectively. All patients met the following criteria: pathologically confirmed with grades 1-3A FL and no previous treatment for FL. Model building was performed on a 70% split sample based on ML algorithms, and with 30% held for internal testing. Model performance was evaluated using the ROC curve, accuracy, precision, and f1-score. Furthermore, we compared the features of FL patients with early (<2 years) and late POD (relapsing after 5 years of starting treatment, L-POD). This study was in accordance with the Declaration of Helsinki and it was approved by the Medical Ethical Committee of the Shandong Provincial Hospital. All patients were obtained for the collection of informed consent at their first visit.

**Results:** Of the included 224 patients, 45 (20.1%) had at least one POD. E-POD occurred in of the 29 (64.4%) evaluable patients and 6 (13.3%) had late POD (relapsing after 5 years of starting treatment, L-POD), In our center, the median age at the time of the FL diagnosis was 49 years (range 19-82), and the median length of follow-up was 28 months (range 3-120). Most patients were diagnosed with advanced disease (80.8%) and had grade 1-2 disease (75.9%). 75 cases (33.9%) had bone marrow participation and 63 cases (28.1%) with Ki-67 $\geq$ 30%.

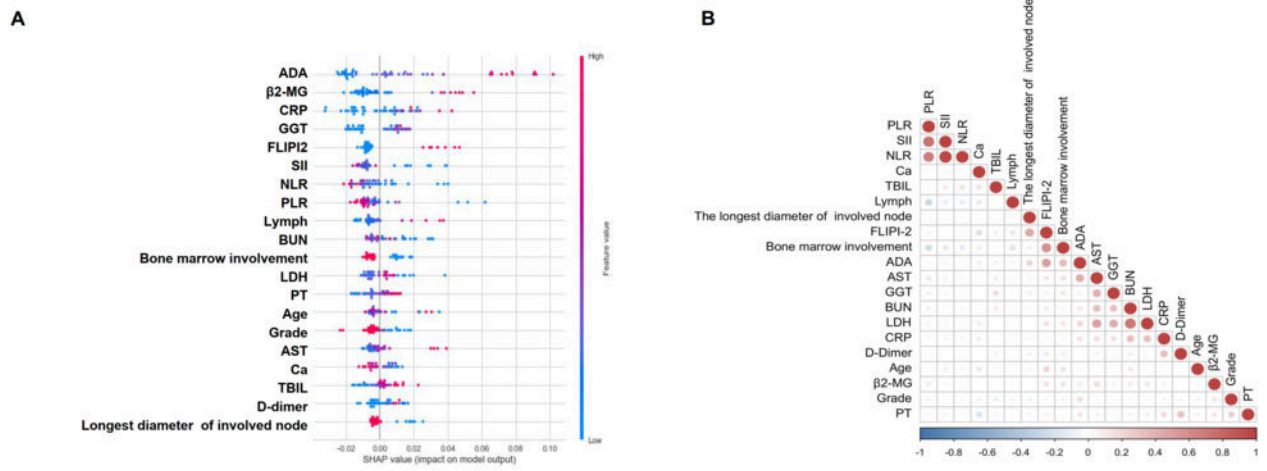
To establish a scoring system for predicting E-POD, we selected 34 variables for served as the input candidate features, including sociodemographic, clinical, and immunohistochemistry data, and used lymphoma E-POD (yes or no) as the output variable. In 14 models based on ML algorithms, the best models were RF Classifier (AUC: 0.83, Accuracy: 89.5%). To improve the performance, it was optimized after the complete construction of the model. Finally, the RF Classifier demonstrated AUC and accuracy as follows: 0.96, and 92.3% and we calculated the recall, precision, and F1 scores based on the confusion matrix as follows: 84.1%, 97.4%, and 90.3%, respectively. We evaluated the performance of prediction models built with data in a validation cohort and found the performance of the model was similar in both cohorts.

Based on this RF Classifier, we selected relative variables important out of the multiple clinical and laboratory blood markers to predict E-POD ( **Fig. 1A**). Meanwhile, we analyzed the Pearson correlation coefficients (*r*) between 20 statistically meaningful biomarkers to reduce the complexity of the model ( **Fig. 1B**). To further screen for prognostic biomarkers, the top ten risk factors from the three best-performing models were integrated. The five biomarkers were completely involved in these three modules, including adenosine deaminase (ADA),  $\beta$ 2-microglobulin ( $\beta$ 2-MG), C-reactive protein (CRP), aspartate transaminase (AST), and D-dimer. Additionally, we also found a significant difference in OS between the clusters among FL patients ( $p < 0.05$ ). Among 224 patients with FL, 45 (20.1%) patients experienced recurrence and progression. Besides 23 (60.5%) cases with E-POD, there are 6 (15.8%) patients who had L-POD. In comparison with those who experienced E-POD, L-POD patients showed a trend toward a better outcome, and a more favorable risk profile at presentation: lower pathological grade (1-2), anemia, and no mutation of the CARD11.

**Conclusion:** We established a prognostic model to identify a subgroup of patients at high risk of E-POD based on AI analysis. This model utilizes objective variables easily and reliably measured at diagnosis which makes it easier for clinicians to predict the E-POD risk of FL and dynamically monitor changes during treatment. Further validation will hopefully facilitate the incorporation of these simple biomarkers into clinical decision-making tools to prevent adverse outcomes and promote personalized treatment options.

**Disclosures** No relevant conflicts of interest to declare.

**Figure 1**



**Figure 1**

<https://doi.org/10.1182/blood-2023-185742>