

550 consecutive patients diagnosed with CAD who were admitted to the Beijing Friendship Hospital, Capital Medical University, between January 2015 and December 2023.

To identify latent subgroups, we employed the unsupervised K-means clustering algorithm.

550 CAD patients were classified into three distinct and clinically coherent phenotypes.

The identified phenotypes were characterized based on demographic, lifestyle, and biochemical data.

Phenotype 1 (n=117): "Metabolism-Driven" Type

Phenotype 2 (n=194): "Age- and Inflammation-Related" Type.

Phenotype 3 (n=239): "Lifestyle- and Inflammation-Driven" Type.

The association between these phenotypes and severe CAC (Agatston score > 400) was assessed using multivariable logistic regression.

Phenotype 3 consistently demonstrated a significantly higher risk of severe calcification when compared to Phenotype 1. Patient subgroup is an independent predictor of severe calcification.

To identify the most robust predictors for classifying patient subgroups, we implemented a multi-pronged feature selection strategy: LASSO, SHAP, Random Forest.

A Venn diagram synthesizes the results from all three methods, revealing a core set of 23 features.

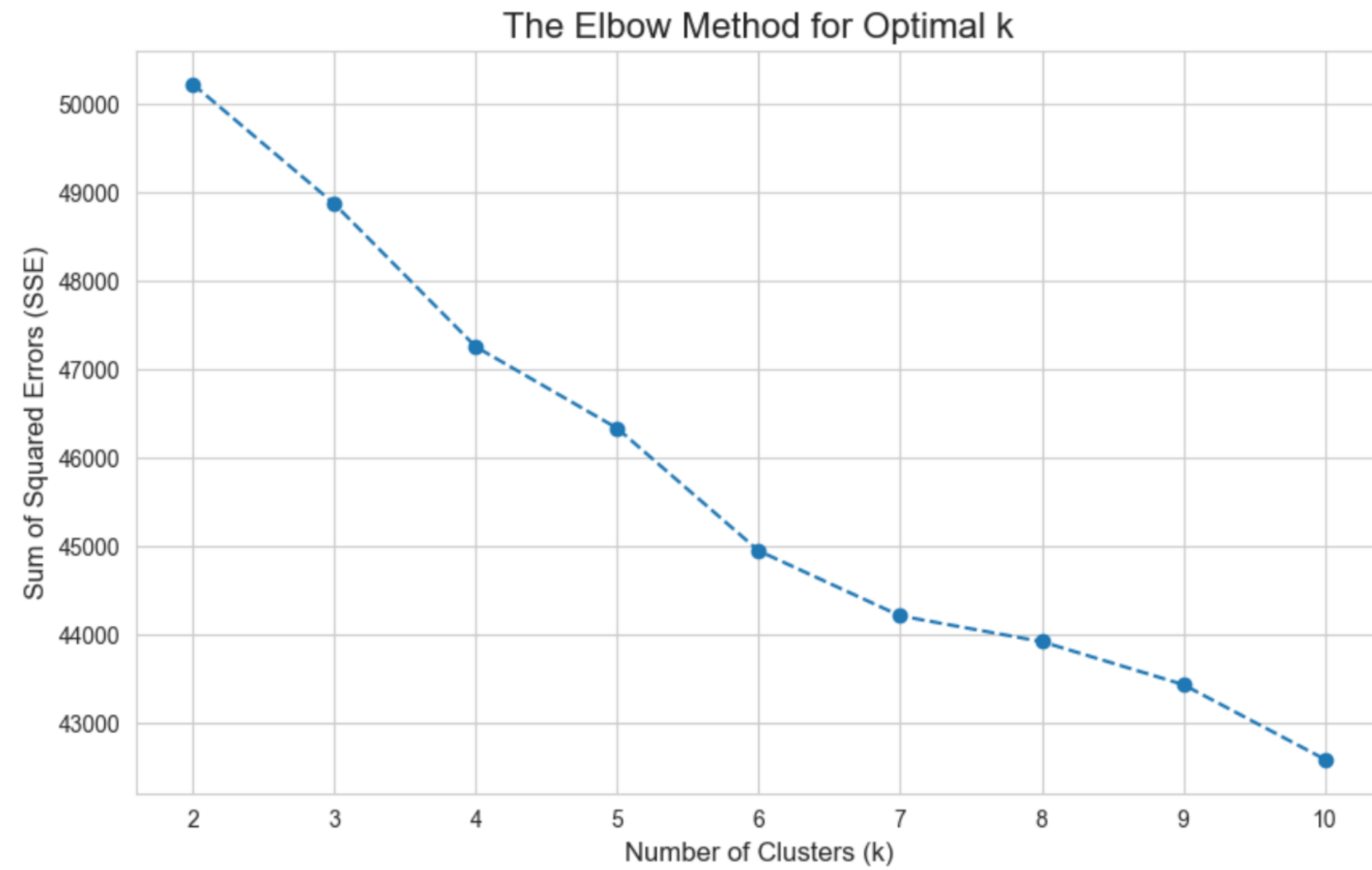
Using the 23 selected features, we developed and compared four machine learning models for multiclass classification: multinomial Logistic Regression, Support Vector Machine (SVM), Random Forest, and XGBoost.

All four models demonstrated excellent discriminative ability.

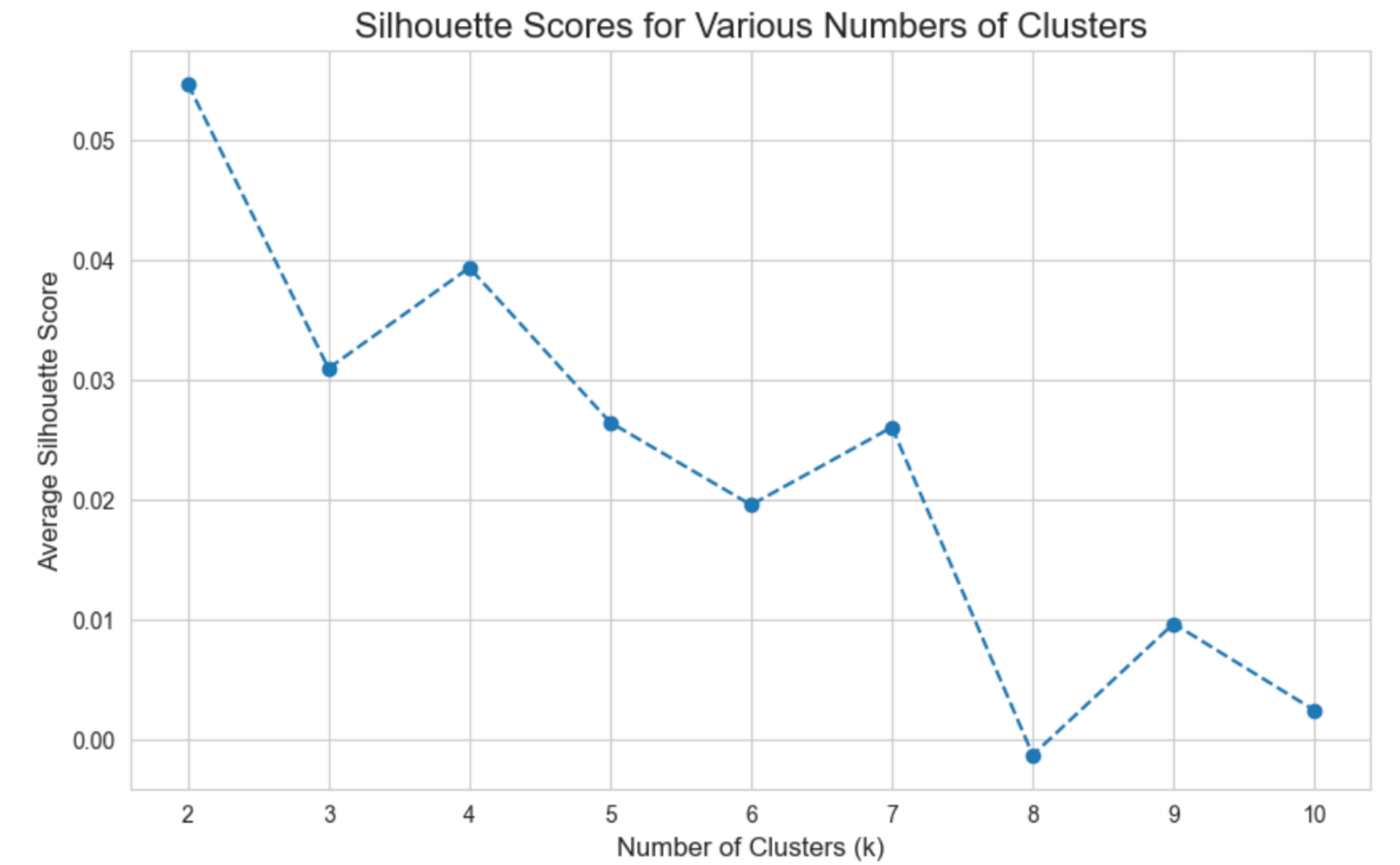
The performance of the four candidate classification models was evaluated using the micro-averaged Area Under the Curve (Micro-AUC) and calibration curves to identify the most suitable algorithm.

Given its strong discriminative power and superior calibration, the Logistic Regression model was chosen for the development of the final clinical risk scoring system.

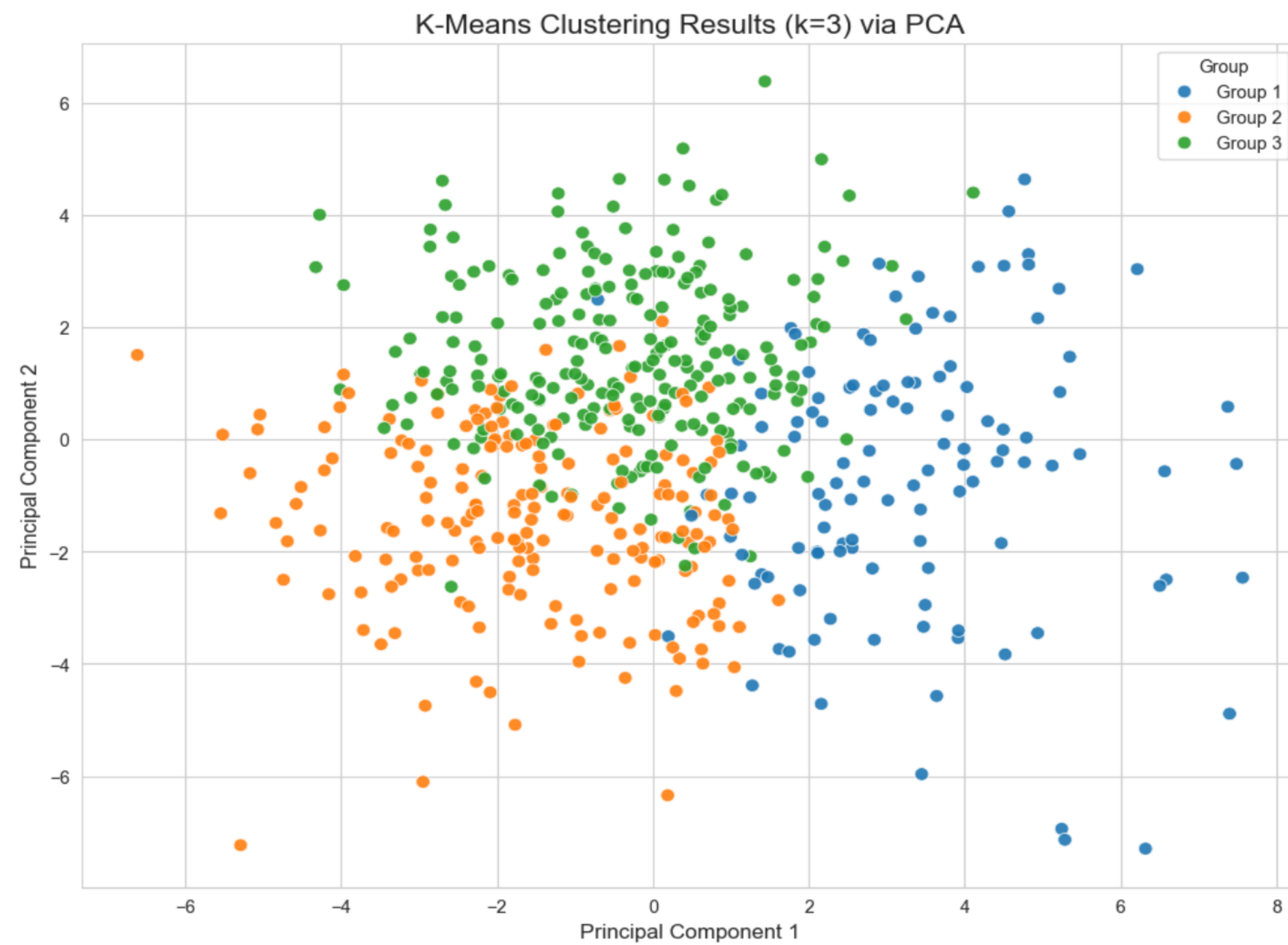
A



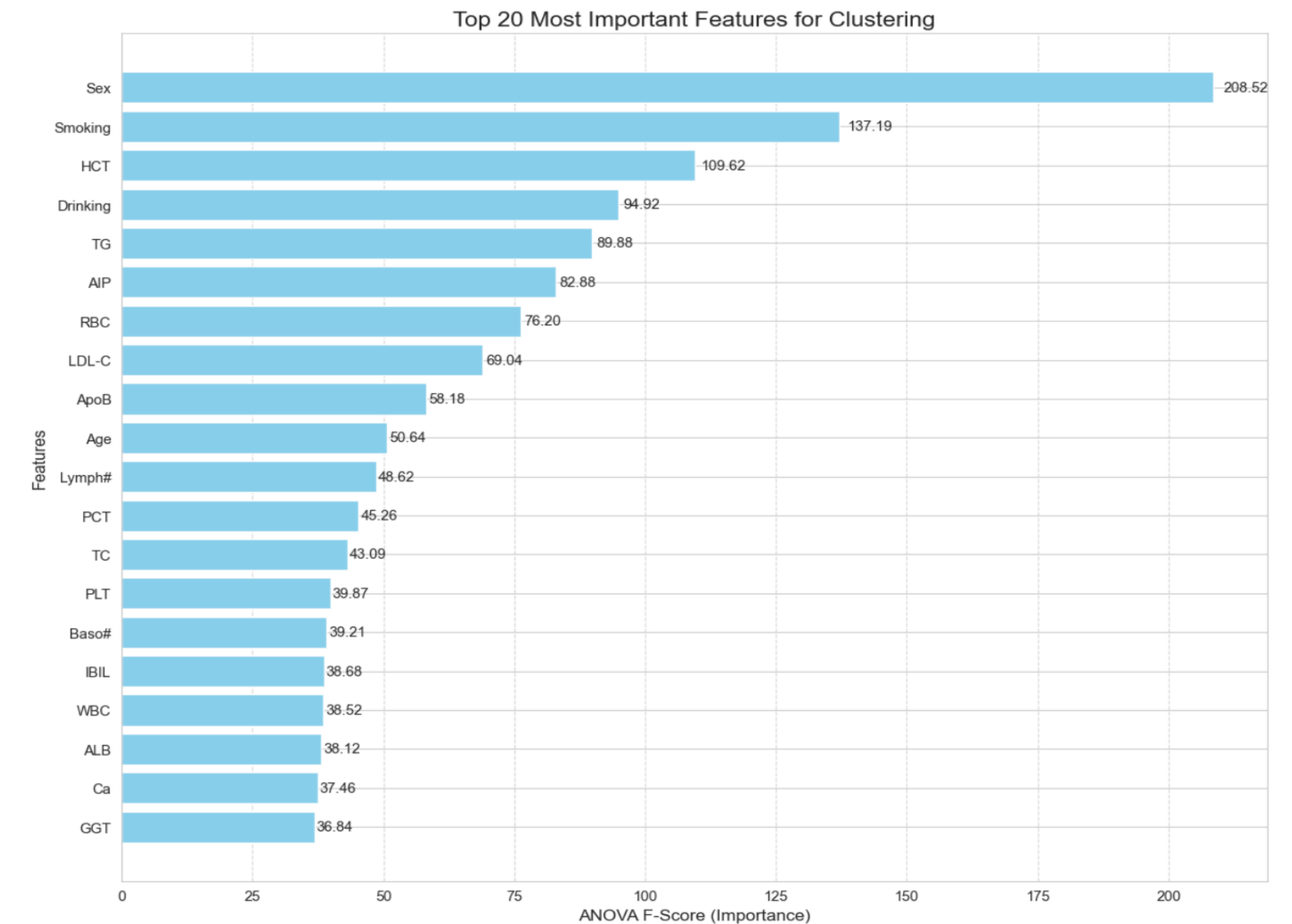
B



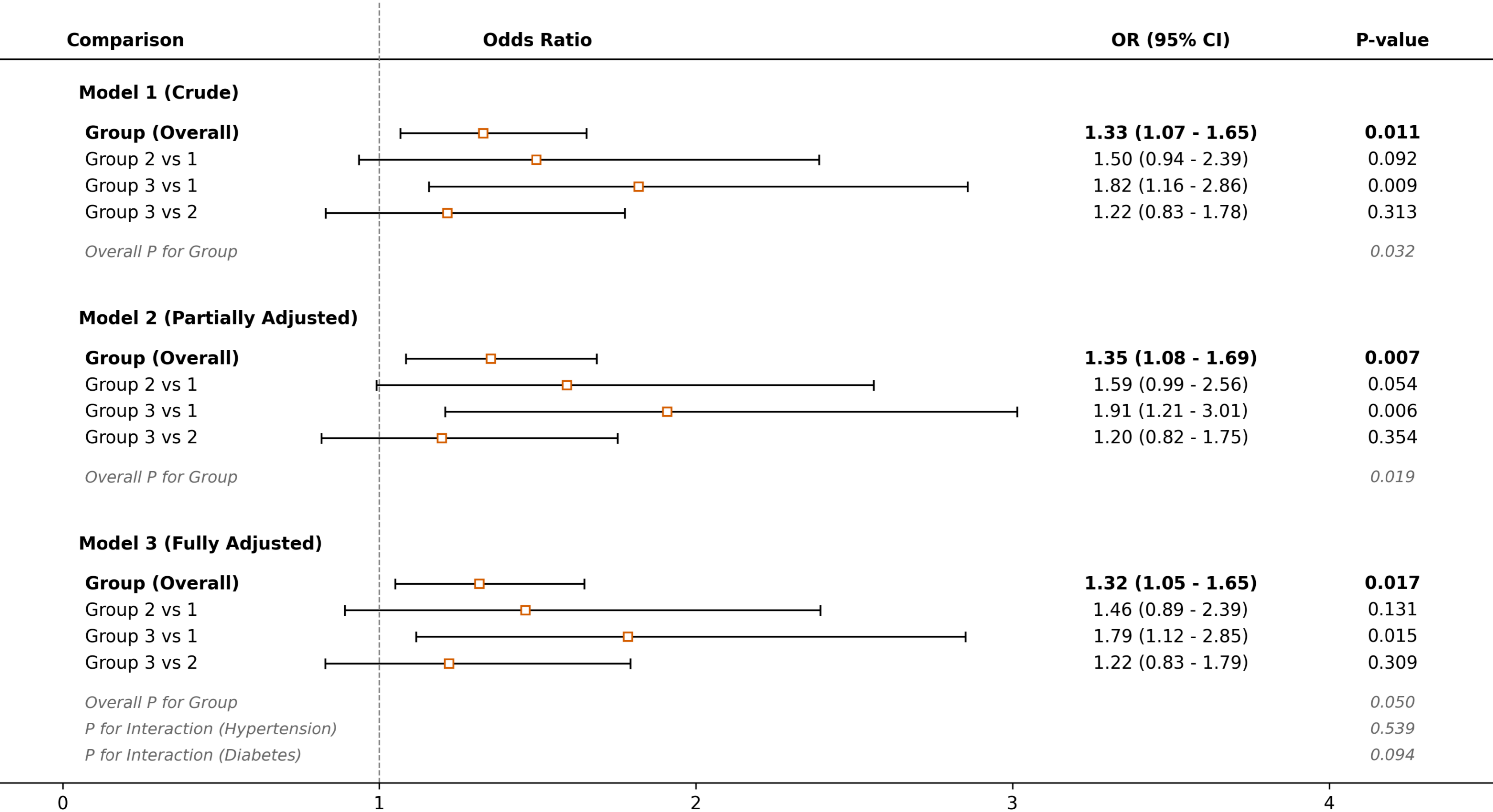
C



D

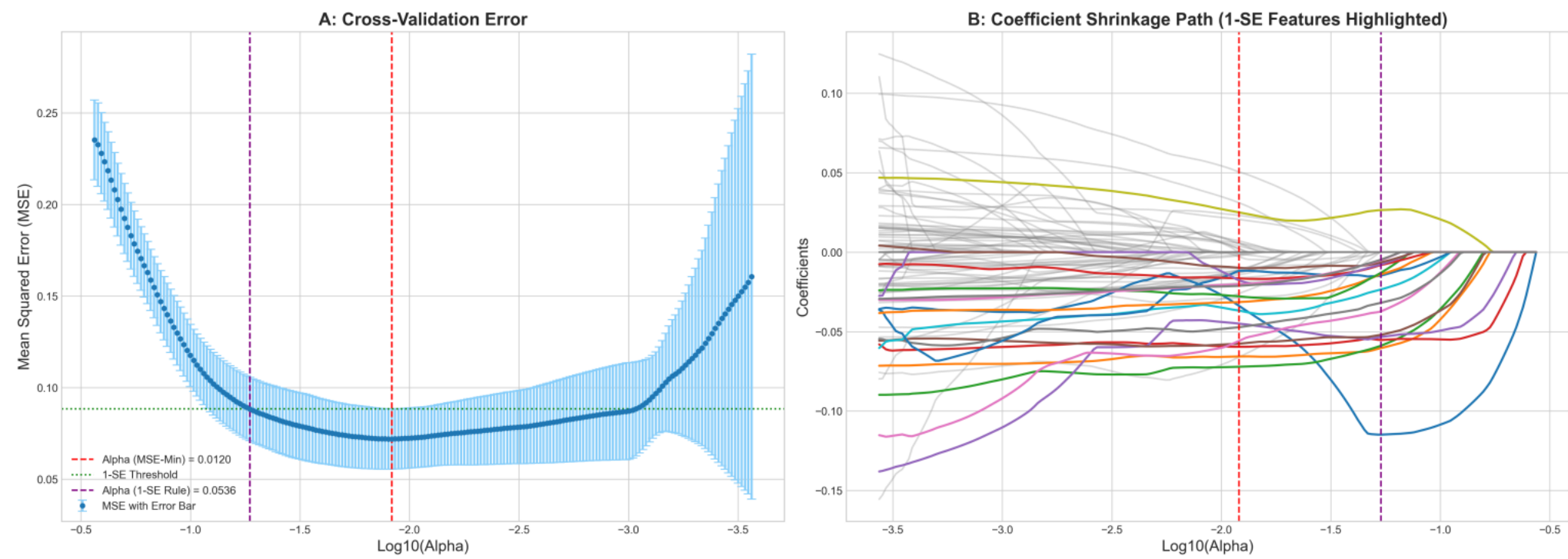


Association of Patient Subgroups with Severe Calcification

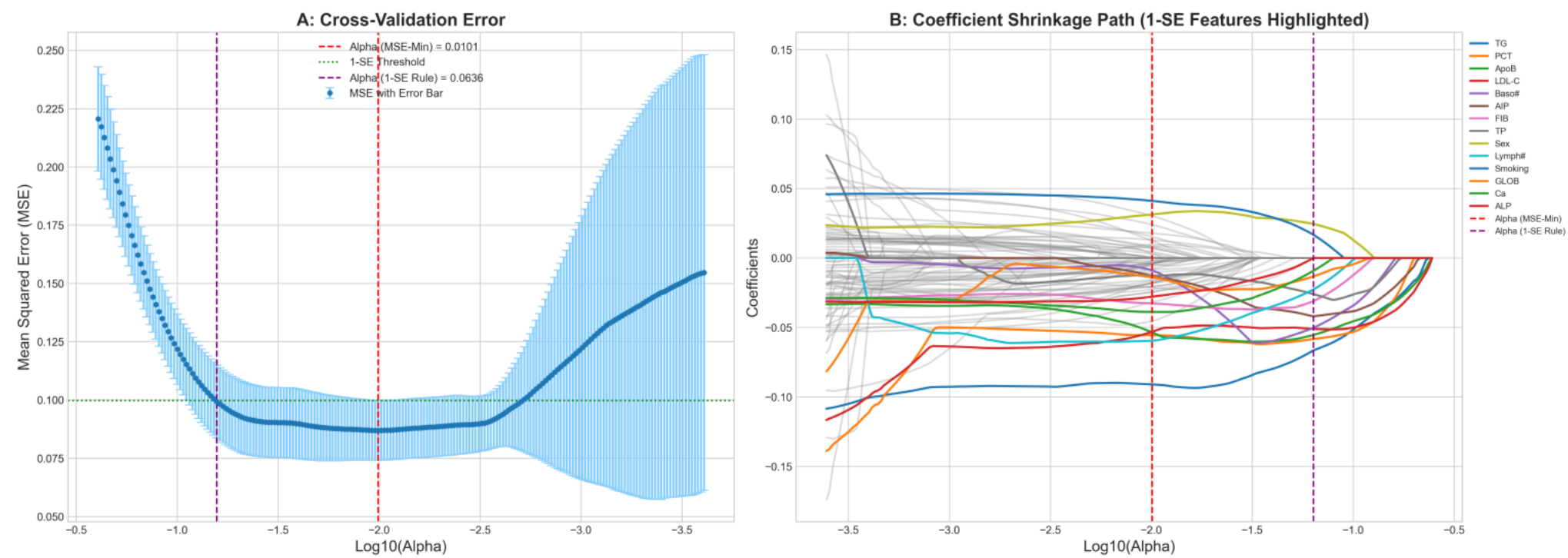


A

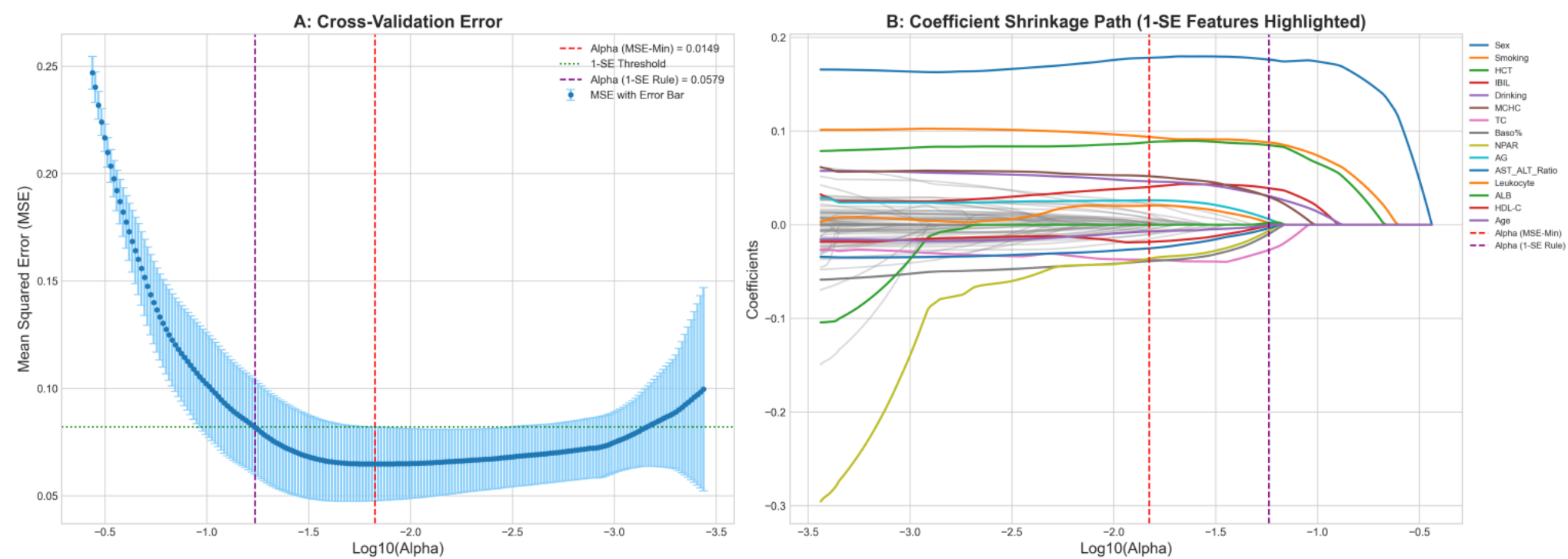
LASSO Analysis: Group_1_vs_2



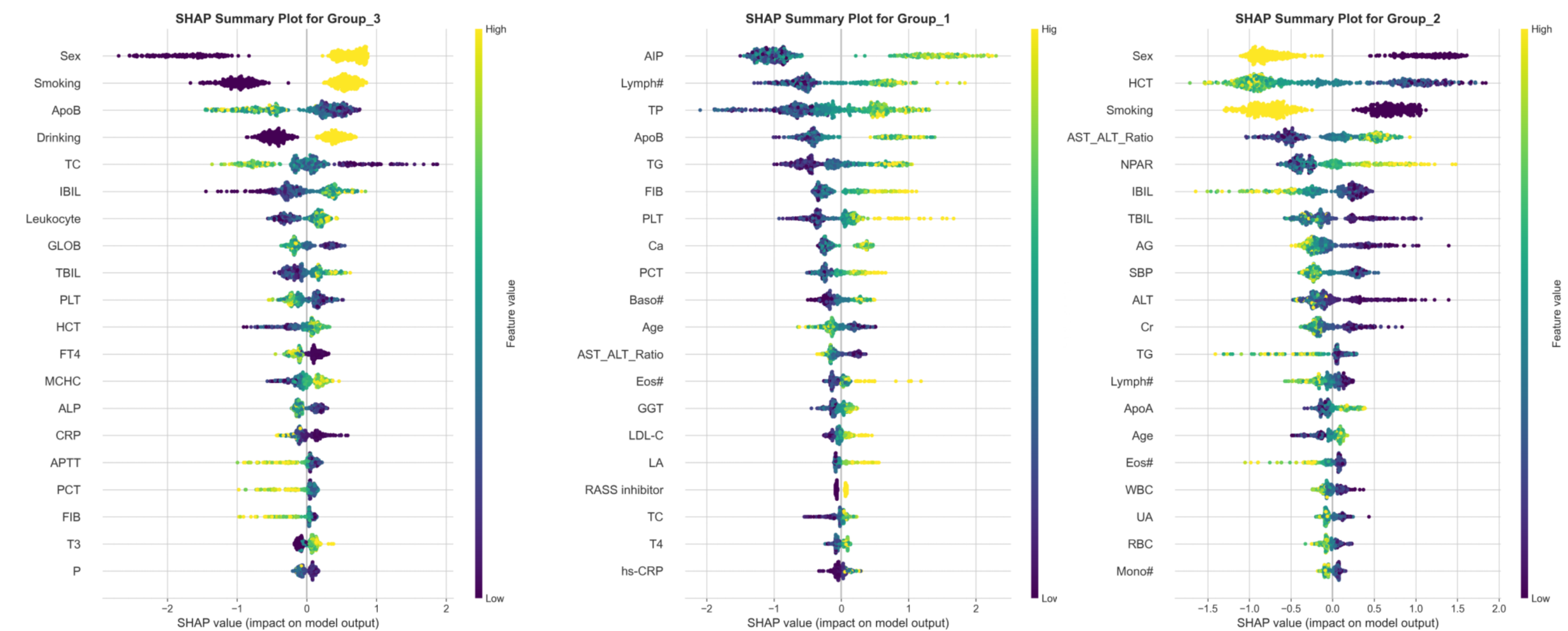
LASSO Analysis: Group_1_vs_3



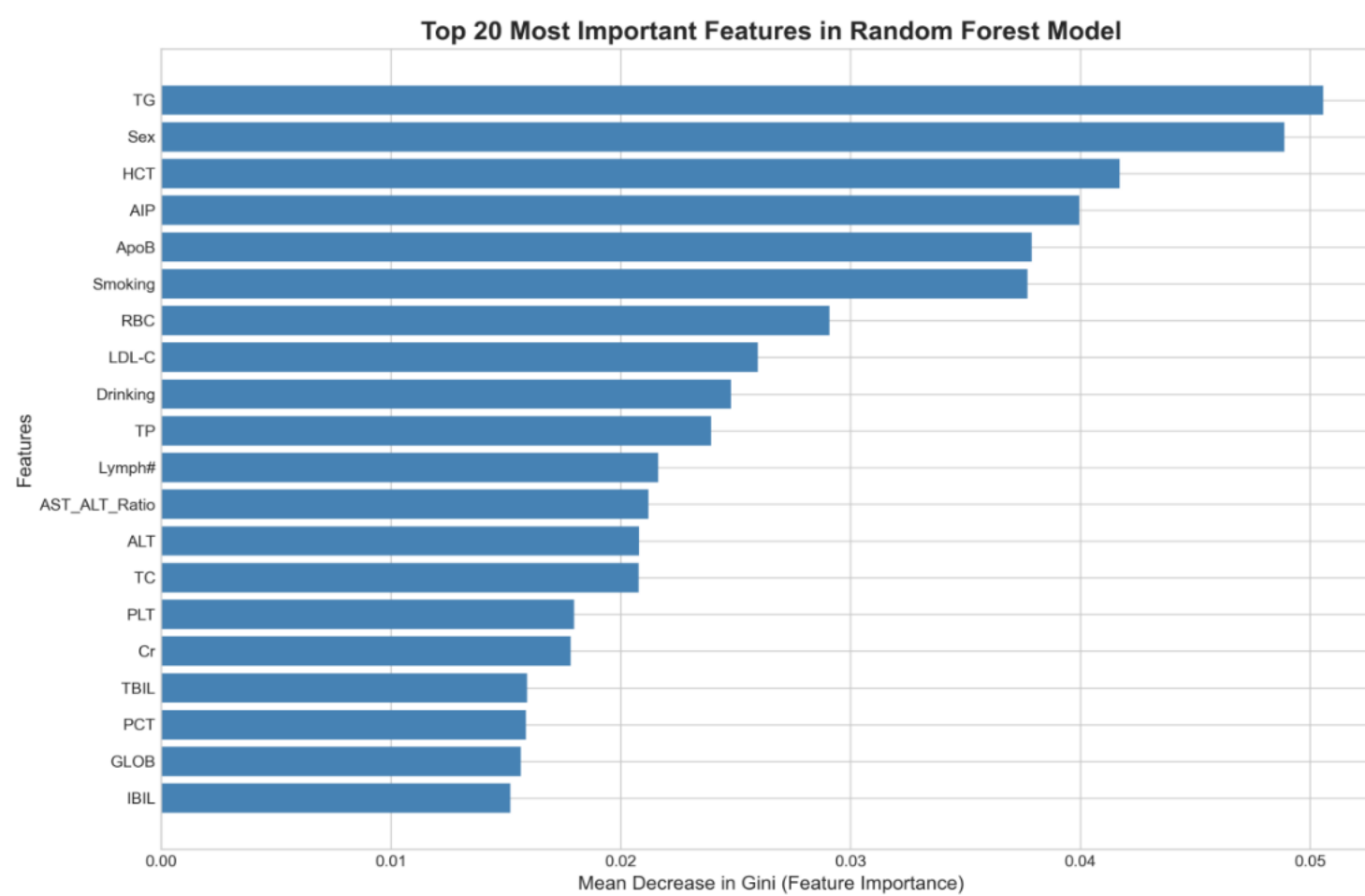
LASSO Analysis: Group_2_vs_3



B

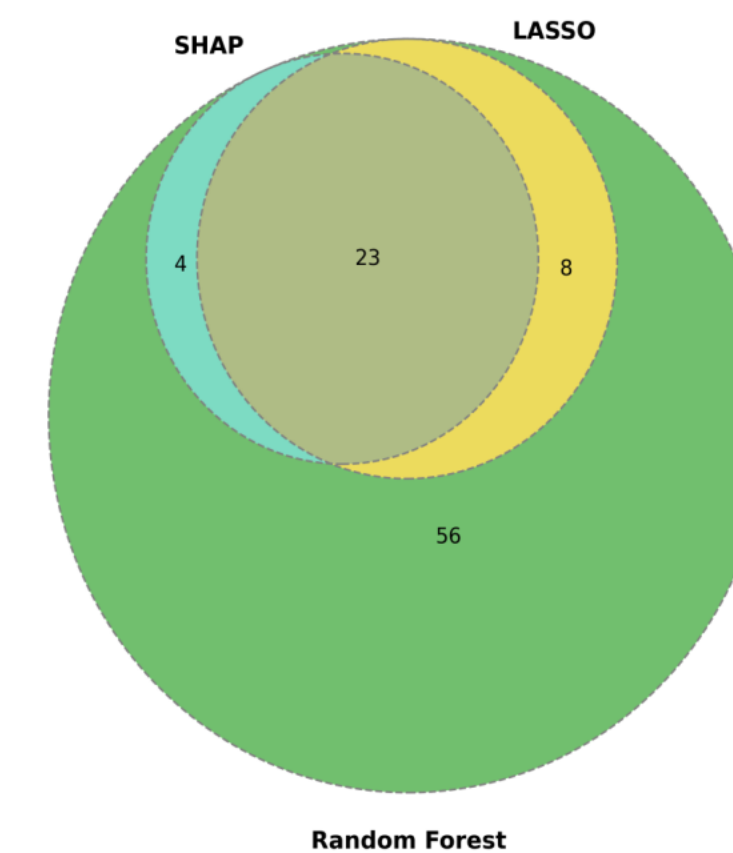


C

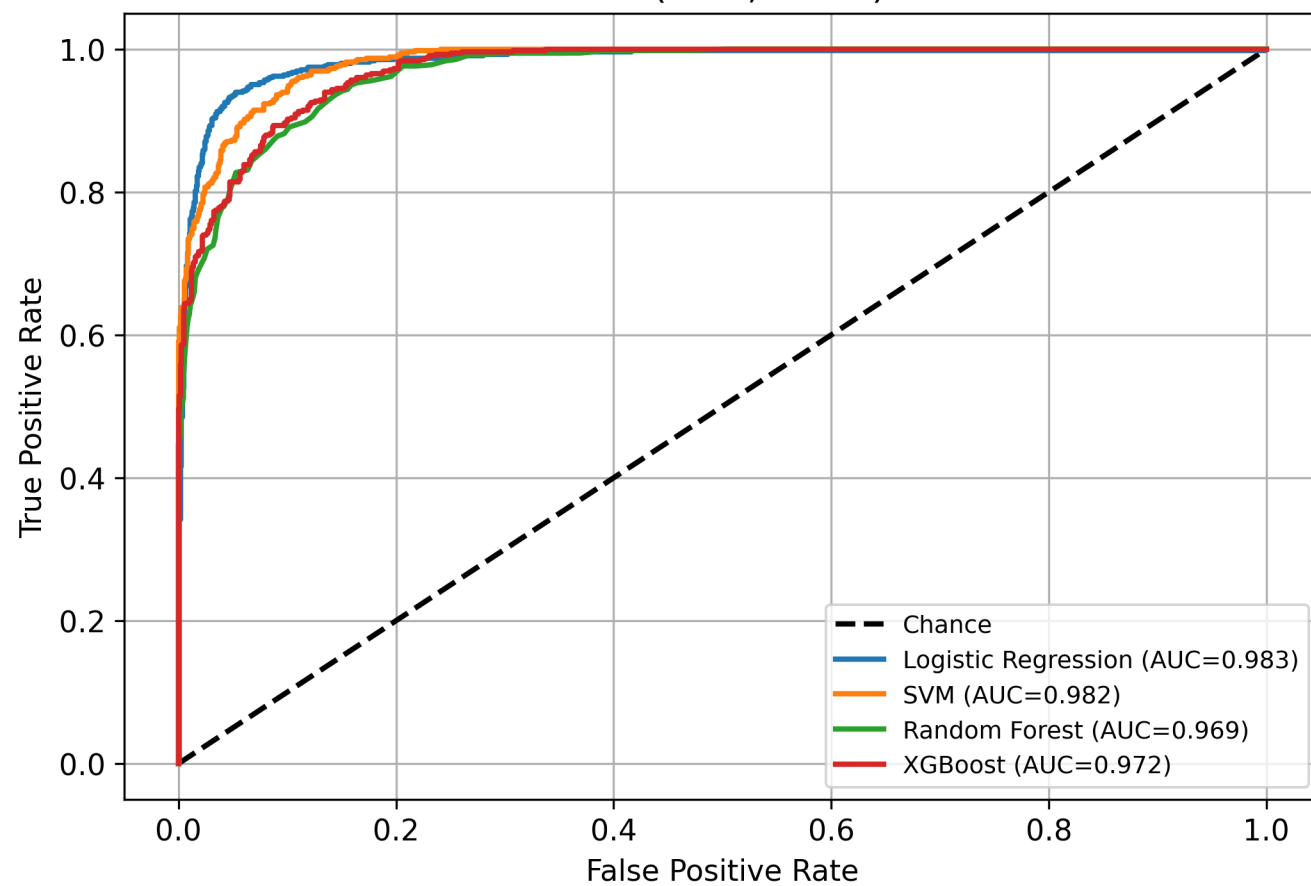


D

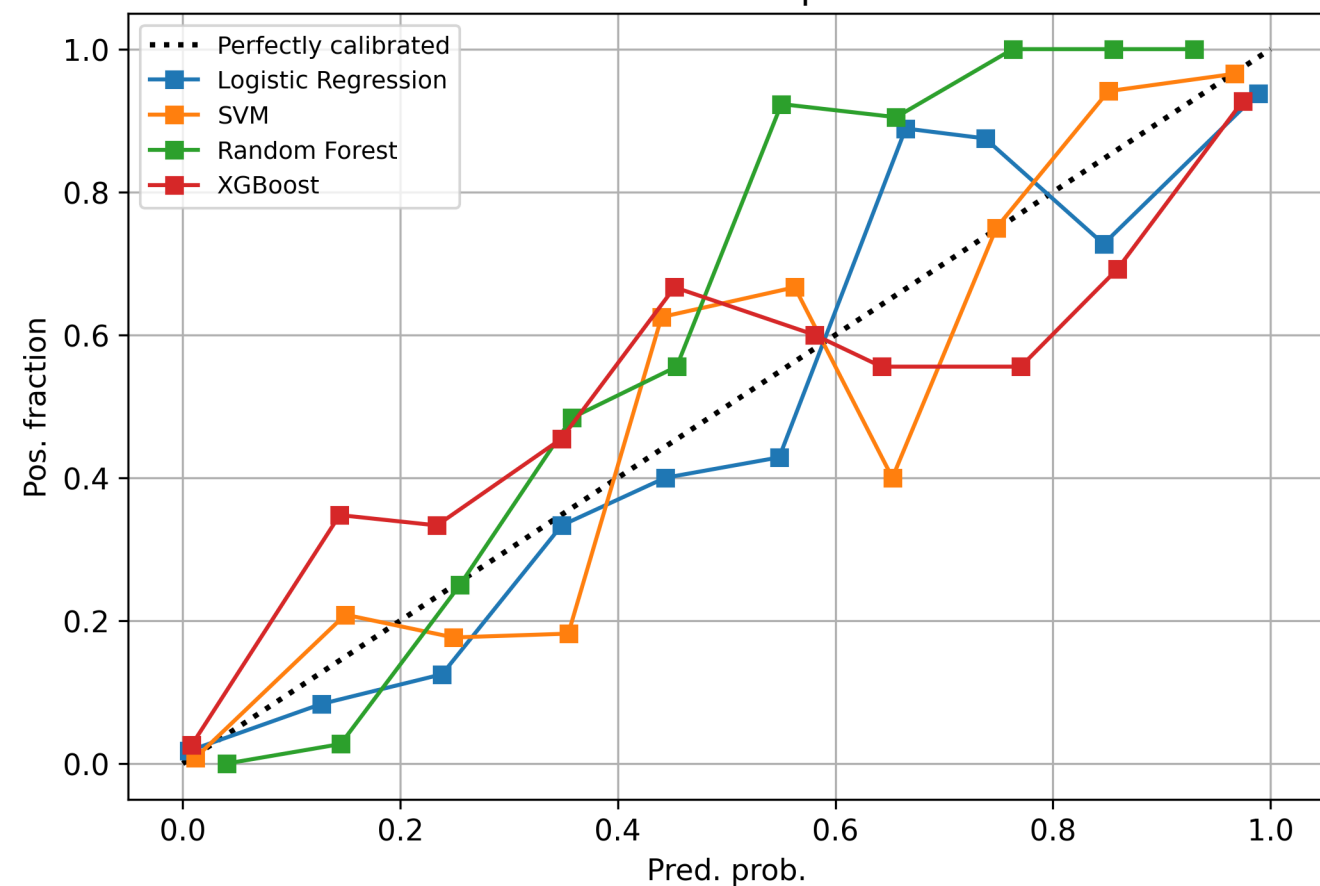
Comparison of Features Selected by SHAP, LASSO, and Random Forest



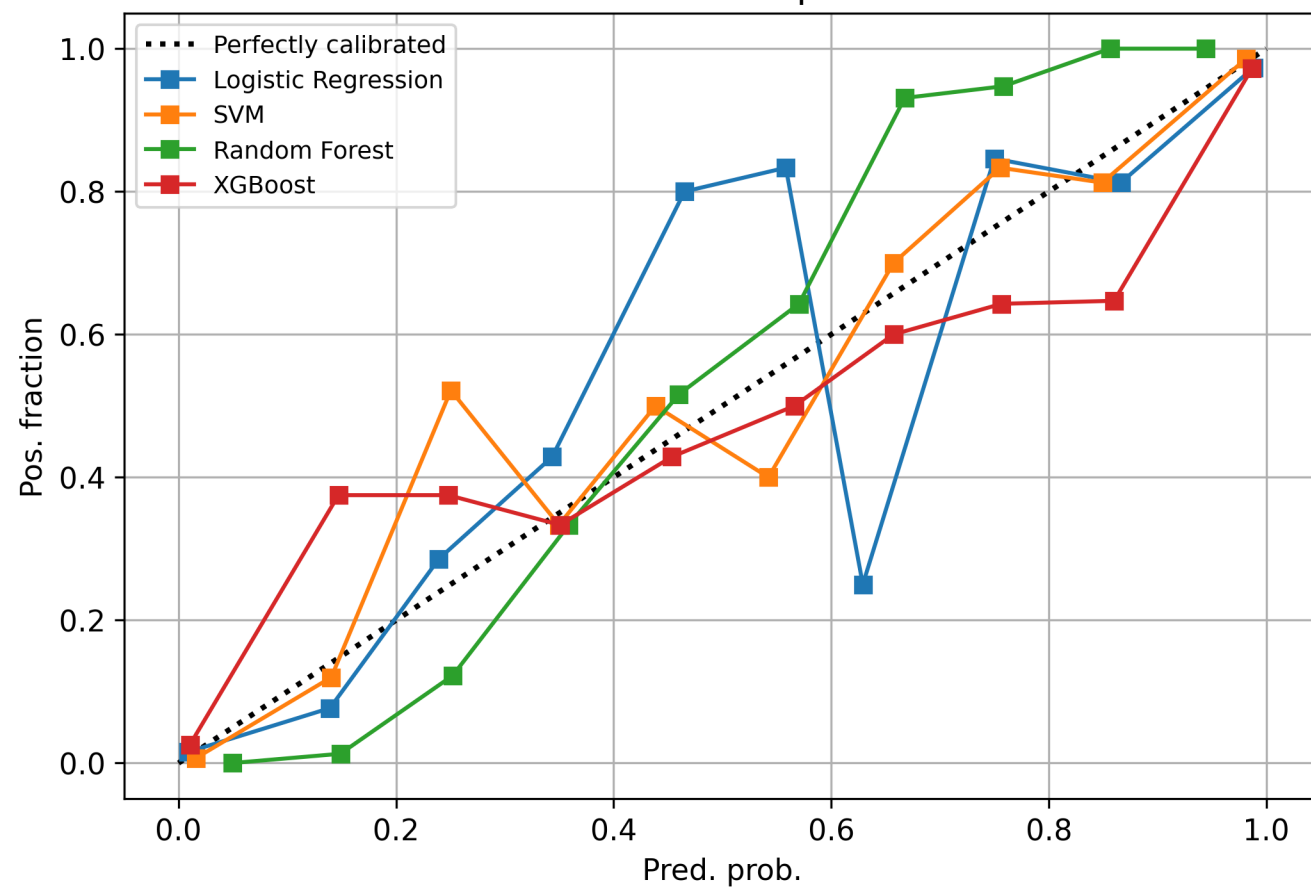
ROC (OOF, Micro)



Calib. Group 1



Calib. Group 2



Calib. Group 3

