

Figure S1. Density map of candidate predictors. Only predictors with at least 60% completeness were included in the model. Predictors with less than 40% available data were excluded. eGFR: estimated glomerular filtration rate.

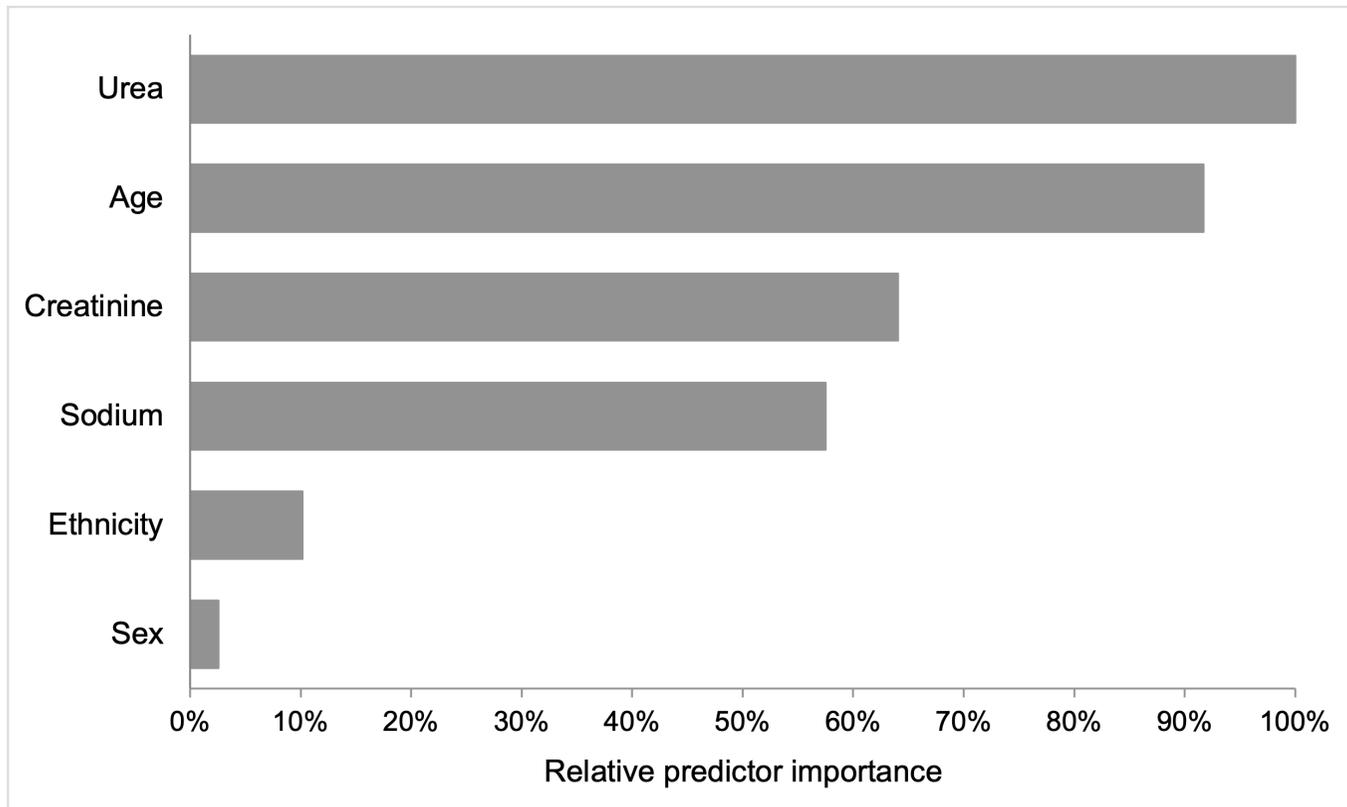


Figure S2. Relative importance of the predictors in the Random Forest model. The importance was relative to the most important predictor 'Urea'.

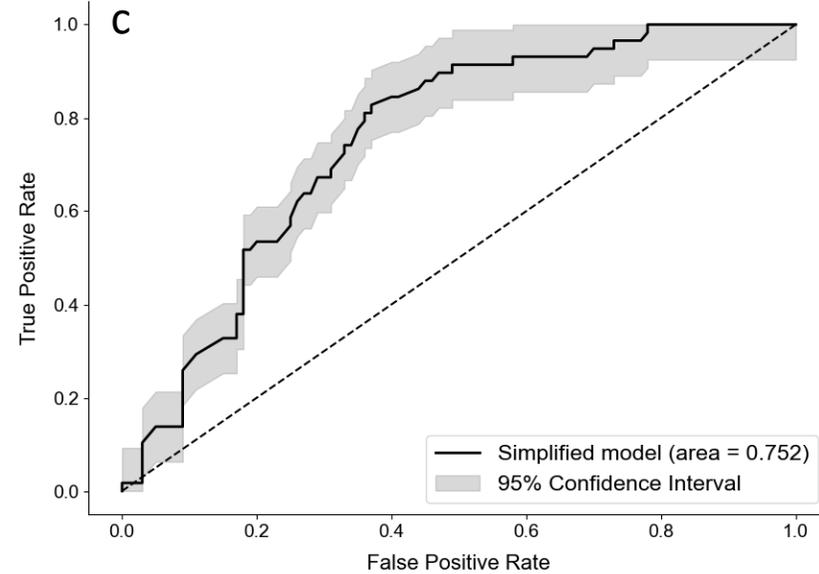
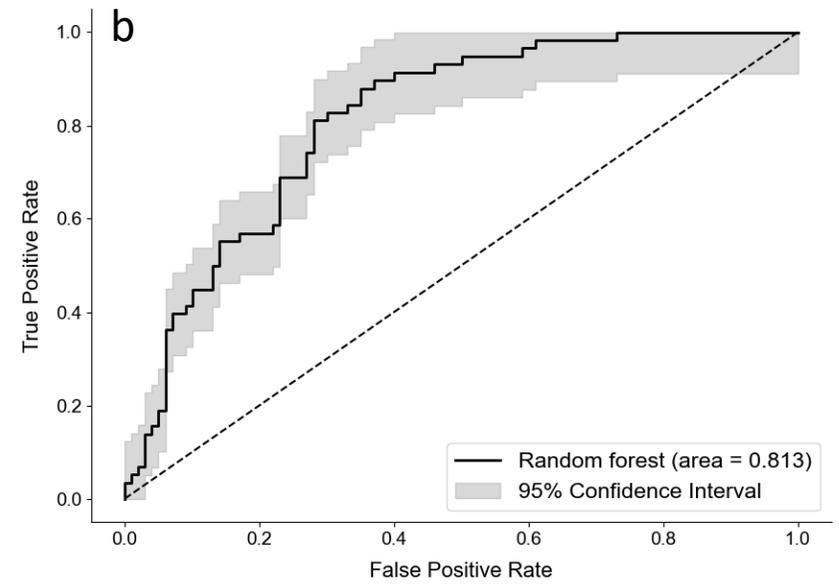
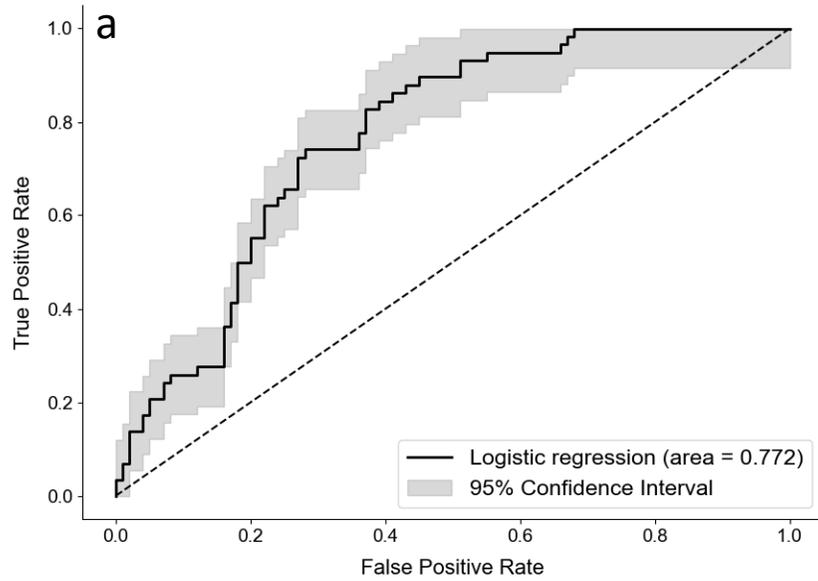


Figure S3. Area under receiver operating characteristic curve in the development cohort for (a) Logistic Regression, (b) Random Forest model, and (c) simplified model.

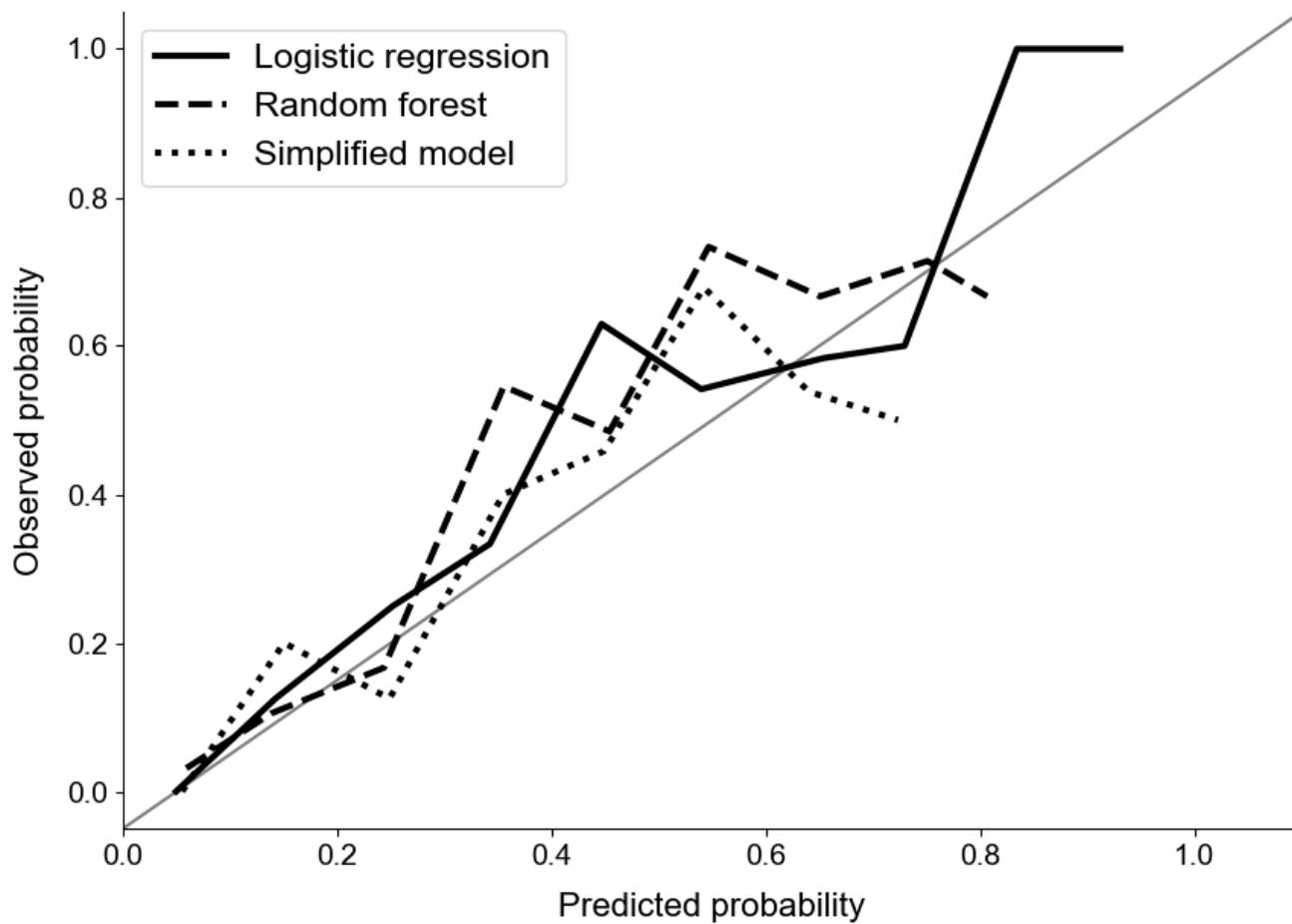


Figure S4. Calibration curves for the Logistic Regression, Random Forest, and simplified model in the development cohort.

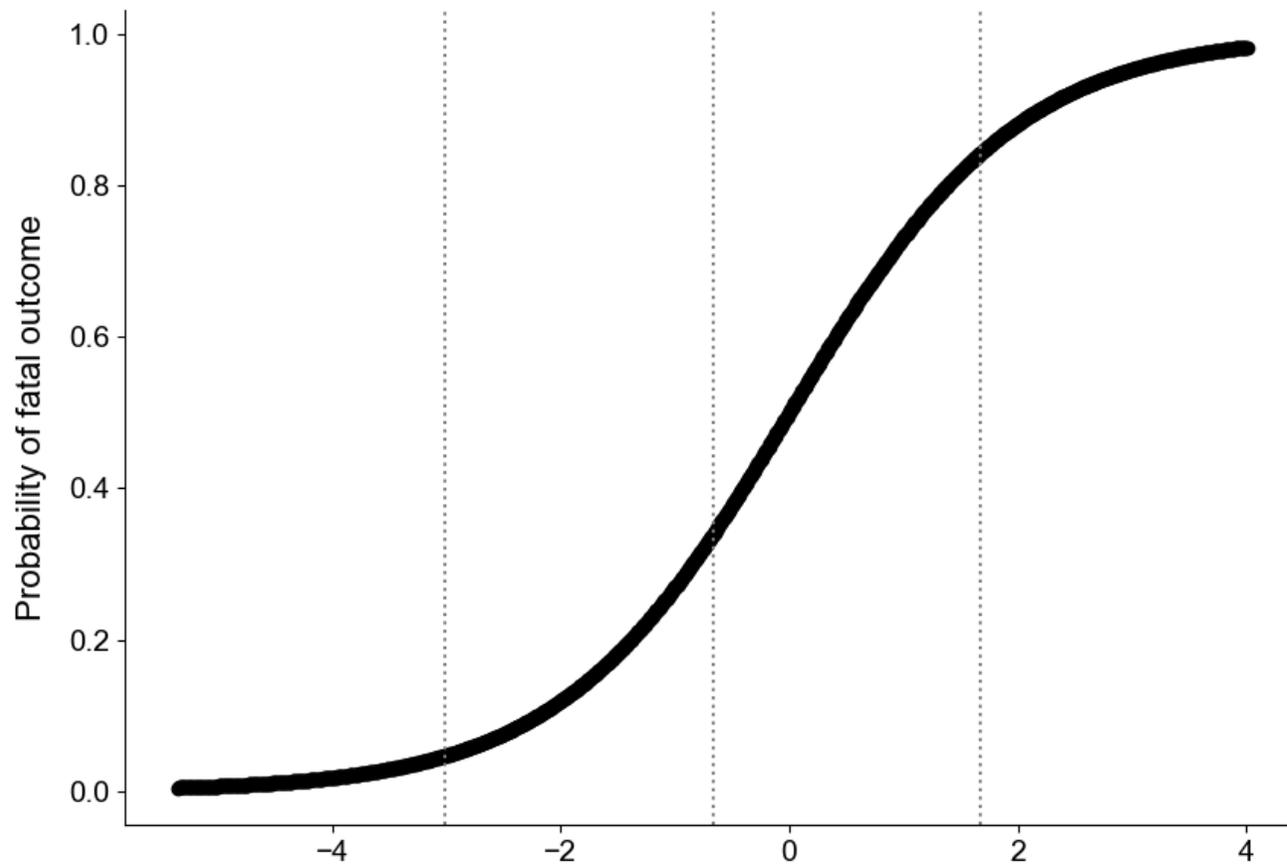


Figure S5. Sigmoid curve of logistic regression model with predictors sex, age, ethnicity, sodium, and urea. The grey dotted vertical lines show the cut-off values for the four risk groups.