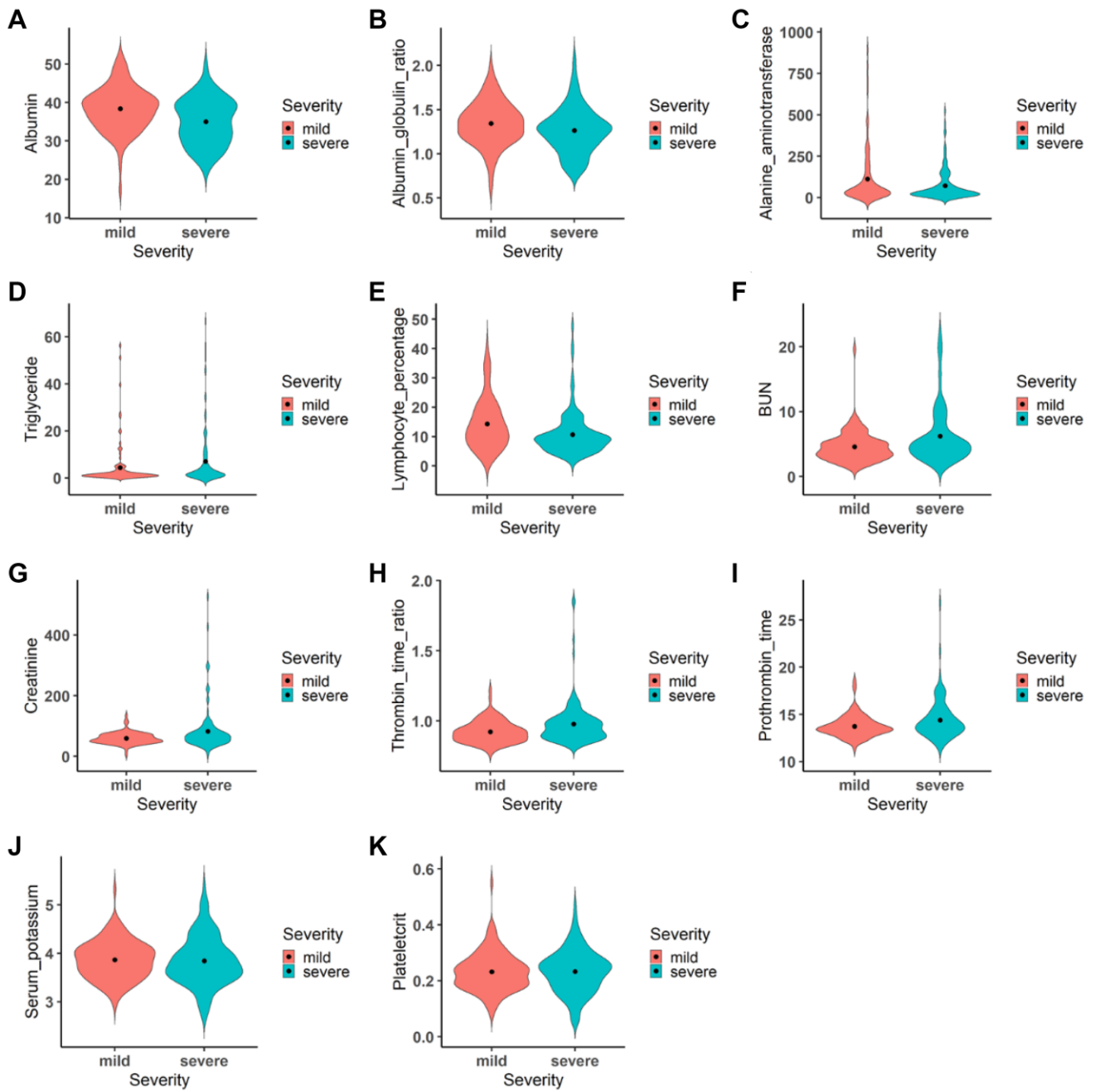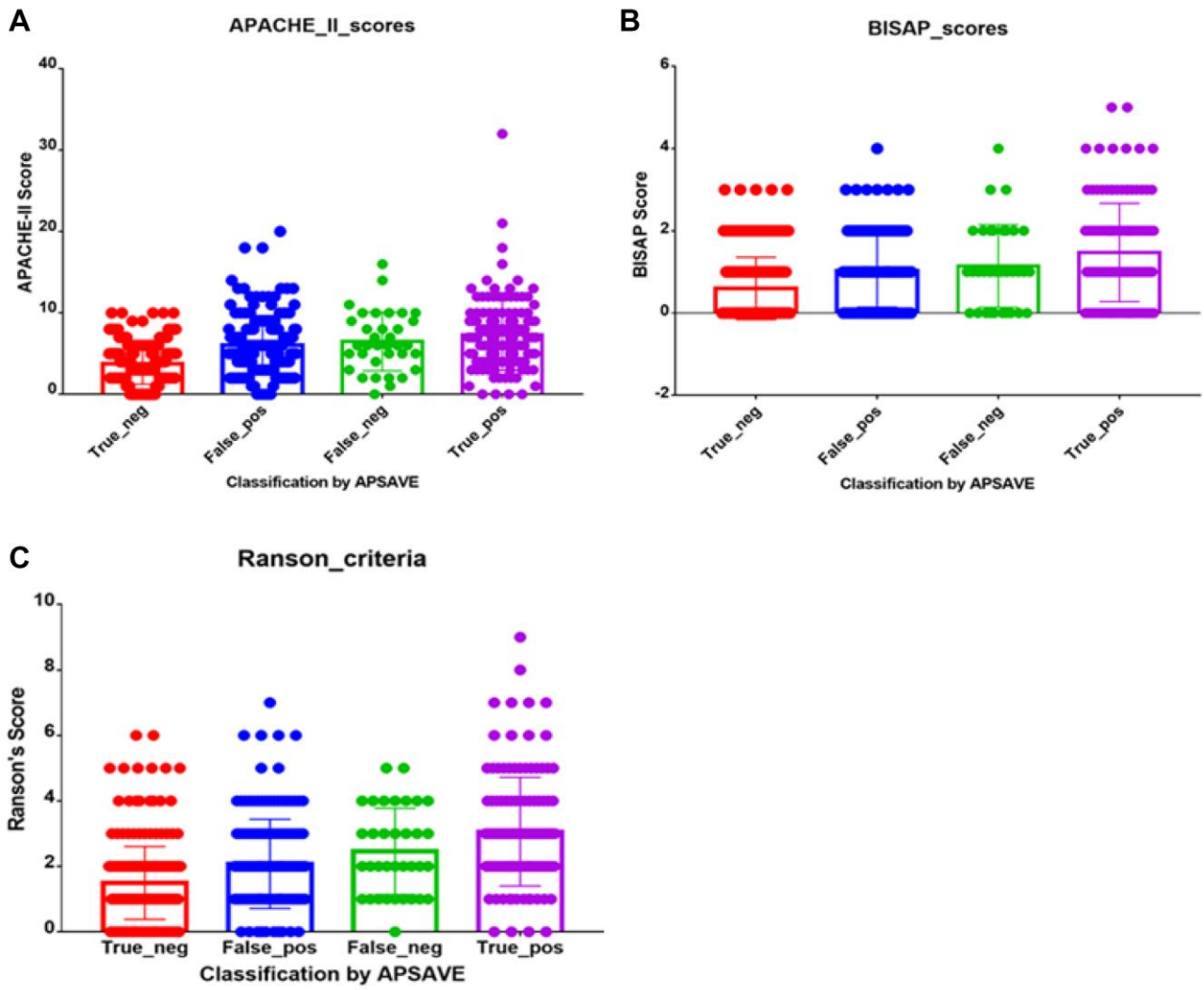**Supplementary Figure 1.** AUC curves of combinatorial models of APSAVE plus APACHE II, BISAP or Ranson's criteria classifying AP cases of the validation cohort.
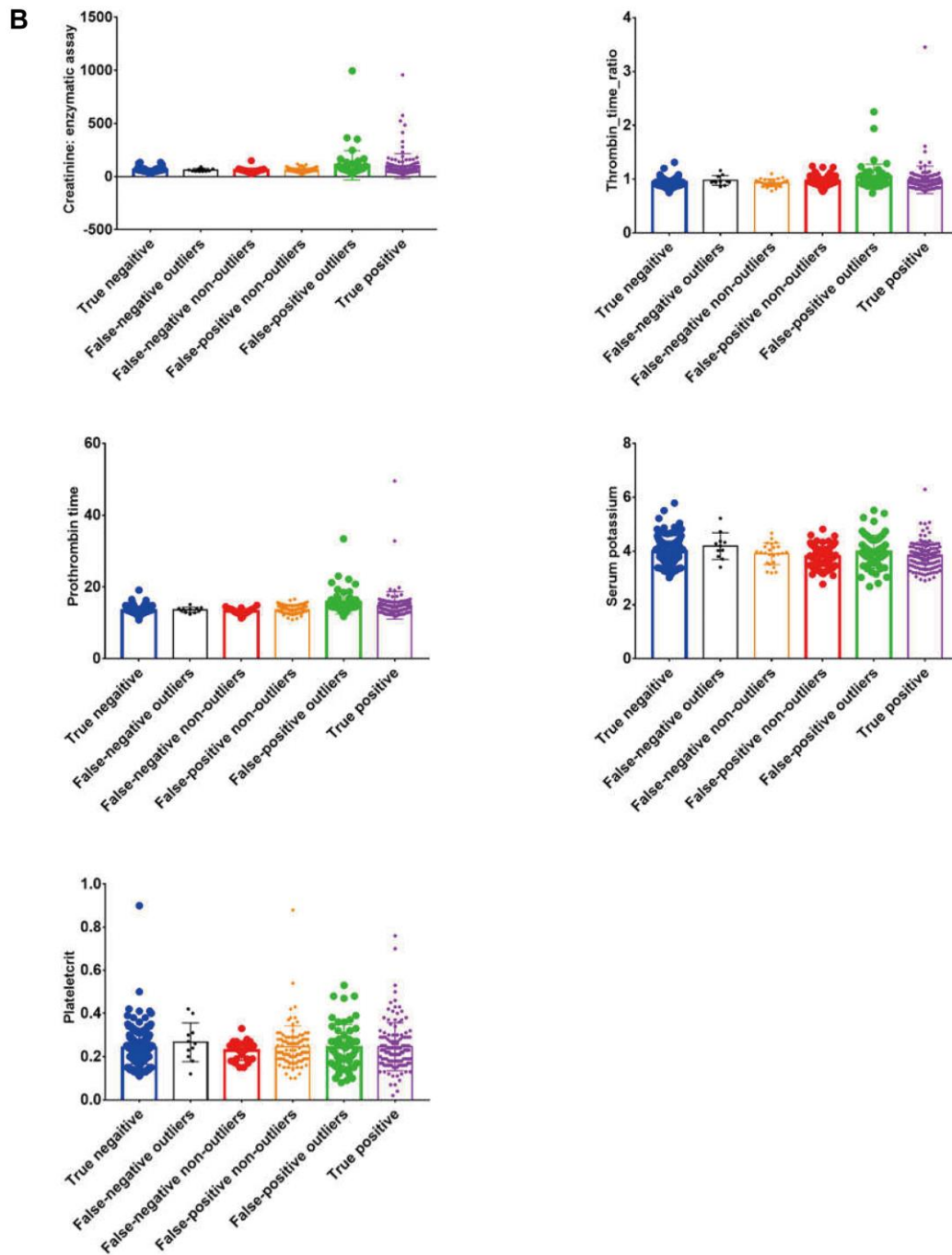
Supplementary Figure 2. Distributions of test results of the 11 biomarkers consisting the APSAVE model from the validation cohort are visualized in violin plots.

A APACHE_II_scores

B BISAP_scores

C Ranson_criteria

Supplementary Figure 3. False_neg cases (MAP cases mis-classified by APSAVE as SAP) were generally scored higher by either of the 3 stratification systems than the Ture_neg cases (Man-Whitney test, *p* < 0.0001 for APACHE II and Ranson's Criteria, and *p* = 0.0092 for BISAP).
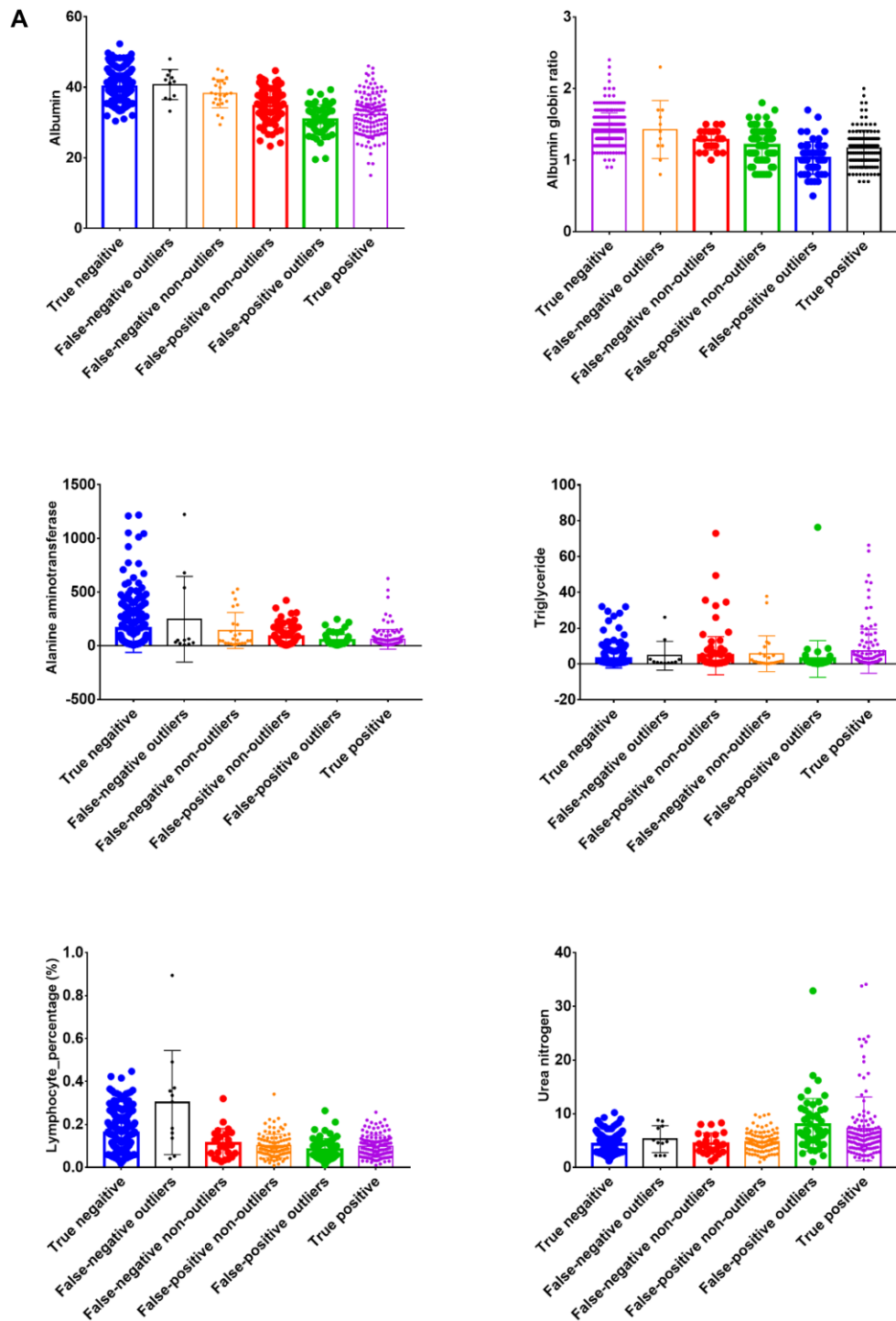
Supplementary Figure 4A. The case distributions of true negative, false-negative outliers, false-negative non-outliers, false-positive non-outliers, false-positive outliers and true positive of 11 biomarkers in the APSAVE model.

Supplementary Figure 4B. The case distributions of true negative, false-negative outliers, false-negative non-outliers, false-positive non-outliers, false-positive outliers and true positive of 11 biomarkers in the APSAVE model.