

Supplementary Table 2. Performance metrics of the PRIME Solution model on test set and external validation set at selected thresholds

Threshold	Accuracy	Precision	Recall	F1	Specificity
A					
0.3	0.80	0.07	0.84	0.13	0.79
0.5	0.89	0.10	0.73	0.18	0.89
0.7	0.94	0.16	0.56	0.25	0.95
0.9	0.98	0.30	0.24	0.26	0.99
B					
0.3	0.53	0.05	0.94	0.09	0.52
0.5	0.70	0.07	0.89	0.13	0.70
0.7	0.83	0.11	0.74	0.18	0.83
0.9	0.94	0.21	0.42	0.28	0.96

A: Model performance on test set. B: Model performance on external validation set. The table presents model performance at various thresholds (0.3, 0.5, 0.7, 0.9). The model was initially optimized for high recall to capture a wide range of potential acute kidney injury cases. Lower thresholds (e.g., 0.3) demonstrate this high-recall characteristic. For clinical implementation, a threshold of 0.9 was selected to balance the model's sensitive detection capabilities with the need for higher specificity in practical settings. This choice aims to minimize false positives while maintaining the model's ability to identify high-risk cases.