

**Supplementary Table 6.** Comparison of diagnostic metrics between specialists, physicians, and medical students according to AI prediction results

Group	Set	Accuracy	Precision	Recall	F1	Specificity	NPV	FPR	MCC	TS
AI predicts AKI occurrence (40 cases, 11 with AKI occurrence)										
Specialist	SET1	0.738	0.533	0.364	0.432	0.879	0.785	0.121	0.278	0.276
	SET2	0.612	0.395	0.773	0.523	0.552	0.865	0.448	0.291	0.354
Physician	SET1	0.600	0.391	0.818	0.529	0.517	0.882	0.483	0.303	0.360
	SET2	0.469	0.322	0.841	0.465	0.328	0.844	0.672	0.167	0.303
Medical student	SET1	0.590	0.361	0.636	0.461	0.572	0.806	0.428	0.187	0.299
	SET2	0.485	0.321	0.782	0.455	0.372	0.818	0.628	0.146	0.295
AI predicts no AKI occurrence (34 cases, 3 with AKI occurrence)										
Specialist	SET1	0.868	0.200	0.167	0.182	0.935	0.921	0.065	0.111	0.100
	SET2	0.897	0.333	0.167	0.222	0.968	0.923	0.032	0.186	0.125
Physician	SET1	0.507	0.101	0.583	0.173	0.500	0.925	0.500	0.047	0.095
	SET2	0.559	0.100	0.5	0.167	0.565	0.921	0.435	0.037	0.091
Medical student	SET1	0.653	0.121	0.467	0.192	0.671	0.929	0.329	0.082	0.106
	SET2	0.706	0.182	0.667	0.286	0.710	0.957	0.290	0.228	0.167
AI provides no prediction (26 cases, 1 with AKI occurrence)										
Specialist	SET1	0.981	0.667	1.000	0.800	0.980	1.000	0.020	0.808	0.667
	SET2	0.846	0.125	0.500	0.200	0.860	0.977	0.140	0.192	0.111
Physician	SET1	0.644	0.098	1.000	0.178	0.630	1.000	0.370	0.248	0.098
	SET2	0.596	0.068	0.750	0.125	0.590	0.983	0.410	0.132	0.067
Medical student	SET1	0.800	0.138	0.800	0.235	0.800	0.990	0.200	0.277	0.133
	SET2	0.677	0.089	0.800	0.160	0.672	0.988	0.328	0.191	0.087

AI, artificial intelligence; AKI, acute kidney injury; F1, F1 score; FPR, false-positive rate; MCC, Matthews correlation coefficient; NPV, negative predictive value; TS, threat score.