

S1 Table. Performances of LR model, GBM model, and eCuraC-2 on the patients in ESD followed by surgery group (n=123) when the sensitivity is adjusted to that of eCuraC-2 (0.957)

	ESD followed by surgery group (n=123)		
	LR	GBM	eCuraC-2
Accuracy (95% CI)	0.317 (0.255-0.332)	0.333 (0.270-0.349)	0.268 (0.211-0.284)
Sensitivity (95% CI)	0.957 (0.791-0.998)	0.957 (0.788-0.998)	0.957 (0.803-0.998)
Specificity (95% CI)	0.170 (0.132-0.179)	0.190 (0.151-0.199)	0.110 (0.075-0.119)
PPV (95% CI)	0.210 (0.173-0.219)	0.214 (0.176-0.223)	0.198 (0.166-0.207)
NPV (95% CI)	0.944 (0.733-0.997)	0.950 (0.757-0.997)	0.917 (0.623-0.996)
AUROC (95% CI)	0.779 (0.667-0.886)	0.765 (0.653-0.869)	0.533 (0.478-0.582)

CI of each performance indicators are calculated with bootstrapping method. AUROC, area under receiver operating characteristic; CI, confidence interval; GBM, gradient boosting machine; LR, logistic regression; NPV, negative predictive value; PPV, positive predictive value.