

S8 Table. Area under the receiver operating characteristics of each independent variables using light gradient boosting algorithm with Bayesian optimization using internal validation SEER dataset

	Internal validation (18-Fold-CV)					
	1-Year	2-Year	3-Year	5-Year	8-Year	10-Year
Age	68.9±1.6	66.1±1.4	65.1±1.5	65.7±1.7	68.7±1.7	70.9±2.1
Sex	50.8±1.0	50.4±0.9	49.9±0.6	50.6±0.8	50.9±1.2	51.0±1.5
Site	55.9±1.0	55.1±0.5	54.4±0.6	53.7±0.6	53.6±1.5	53.5±1.3
Histology	52.6±0.8	53.0±0.5	52.9±0.4	52.5±0.6	52.5±0.9	52.5±1.1
Stage	67.7±1.3	69.6±1.0	70.1±1.0	69.6±1.3	67.7±1.6	66.7±1.5
Grade	60.5±1.5	59.9±1.3	59.2±1.3	58.3±1.9	57.7±1.6	57.2±1.9
Tumor size	61.7±1.6	61.8±1.4	61.4±1.3	61.0±1.3	60.1±1.8	59.8±2.1
Examined LNs	56.1±1.2	55.3±1.2	54.9±1.8	55.3±1.3	55.7±1.6	55.7±1.4
Positive LNs	63.6±1.4	65.2±1.2	65.2±1.3	64.7±1.2	62.7±1.7	61.8±1.2
Radiation	54.1±0.8	52.9±0.6	52.1±0.5	51.5±0.7	50.9±0.6	50.8±0.7
Chemotherapy	57.5±1.4	52.5±1.3	49.7±1.2	51.5±1.3	51.1±1.5	50.4±1.9
CEA level	59.8±1.5	60.5±1.3	60.8±1.5	60.4±1.4	59.3±1.2	59.0±1.3

Values are presented as mean±SD. CEA, carcinoembryonic antigen; LN, lymph node; SD, standard deviation; SEER, Surveillance, Epidemiology, and End Results.