



**Figure S1** Evaluation of the five models in the Surveillance, Epidemiology, and End Results (SEER) training set. (A) Bar plot of Harrell's concordances (c-indices) with 95% confidence intervals (CIs) calculated by 1000-times bootstrapping. (B) Time-dependent area under the curve (TdAUC) curves in the 5<sup>th</sup>-95<sup>th</sup> percentile ranges of the follow-up time. (C) Time-dependent Brier score (BS) curves in the 5<sup>th</sup>-95<sup>th</sup> percentile ranges of the follow-up time. (D-H) Kaplan-Meier curves of low-risk and high-risk groups stratified by Cox proportional hazard (CoxPH, D), Cox with elastic net penalty (CoxNet, E), random survival forest (RSF, F), gradient boosting machine (GBM, G), and survival tree (ST, H). P values were calculated by log-rank test. mAUC, mean AUC; IBS, integrated BS; est., estimated.