



2026 World Conference on Lung Cancer

SEPTEMBER 12 – 15, 2026 | SEOUL, REPUBLIC OF KOREA

Machine Learning Prediction of Survival Based on Metastatic Burden in Stage IV NSCLC

Kang Qin, MD Anderson Cancer Center; An A. Qin, Rutgers University

Control #3426 | E-Poster EP17.27 | SEER 2010-2022 cohort, n = 62,606

The analysis tests whether metastatic burden improves survival prediction

2010-2022

SEER diagnosis years

49,540

CSS events in Sheet1

52,780

OS events in Sheet2

12

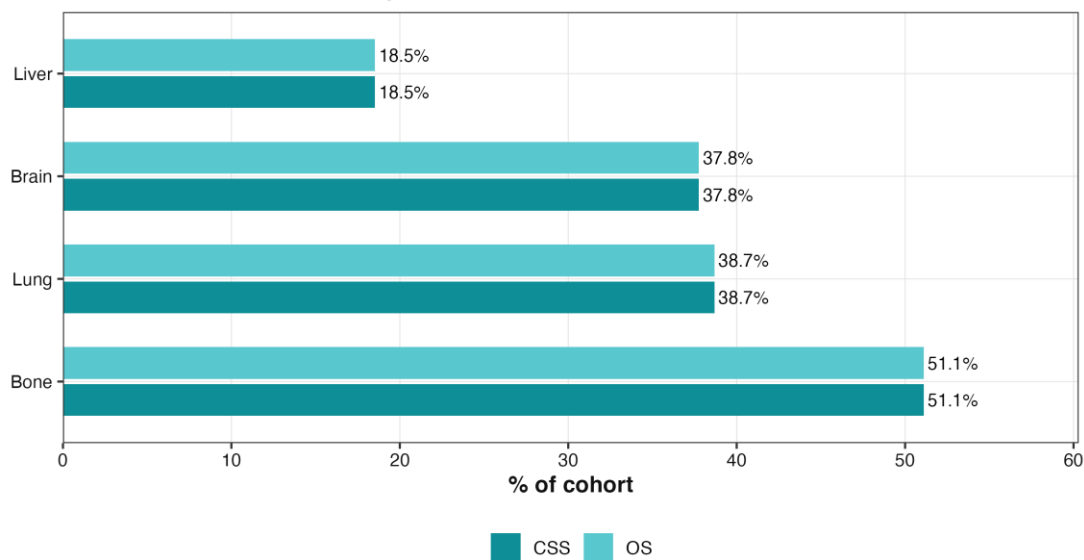
original Python classifiers

- Cox models estimate adjusted associations for metastatic extent and organ sites.
- Machine-learning models classify OS and CSS using clinicopathologic, treatment, and metastatic features.
- Model utility is evaluated with AU-ROC, AU-PRC, calibration, DCA, and SHAP interpretability.

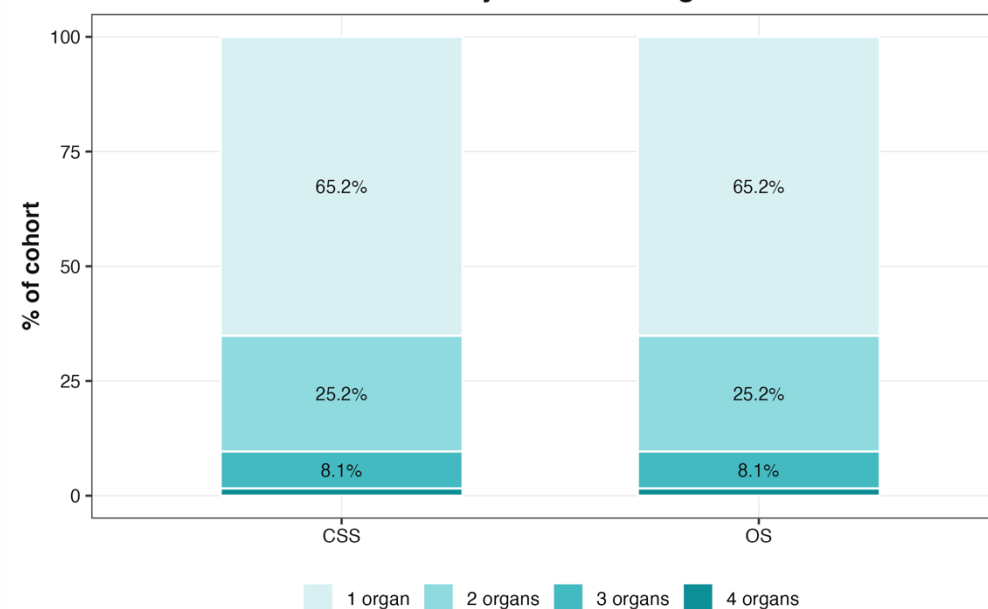
Metastatic organ burden is common and heterogeneous at diagnosis

Baseline metastatic organ distribution in SEER

Sheet1 CSS and Sheet2 OS only; real-world Sheet3 excluded



Baseline metastatic burden by number of organs

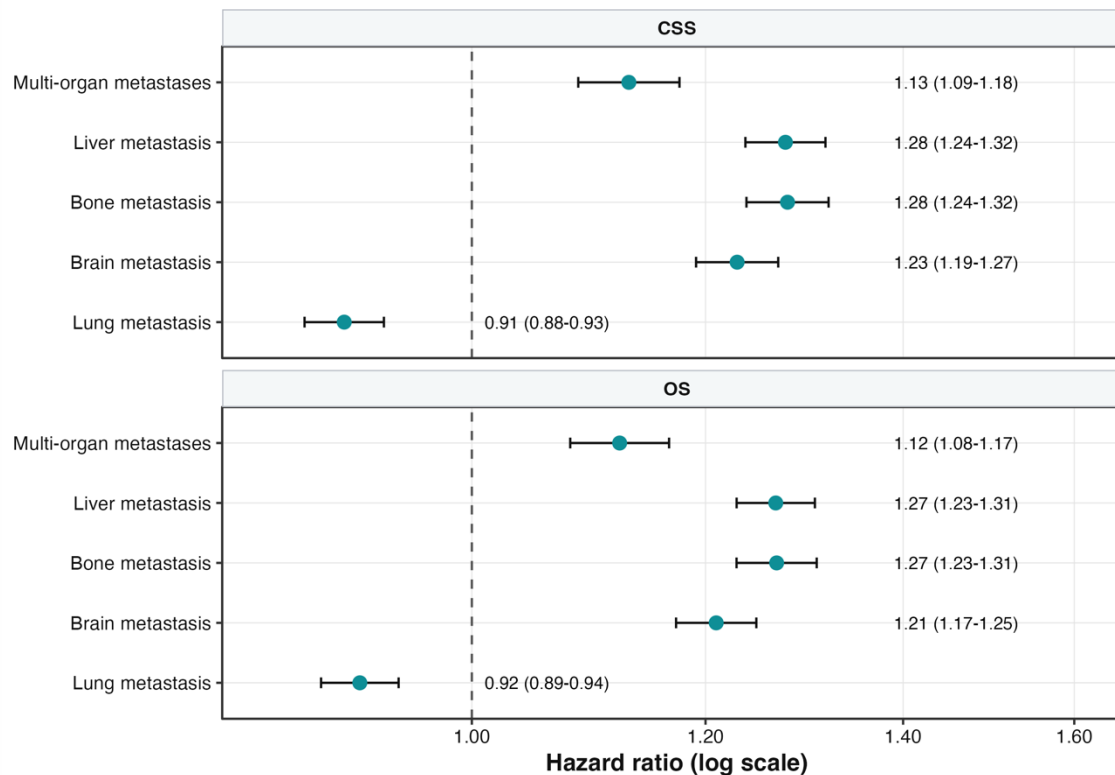


Metastatic burden includes both the number of involved organs and site-specific metastatic pattern.

Adjusted Cox models confirm prognostic separation by metastatic pattern

Adjusted association of metastatic burden with CSS and OS

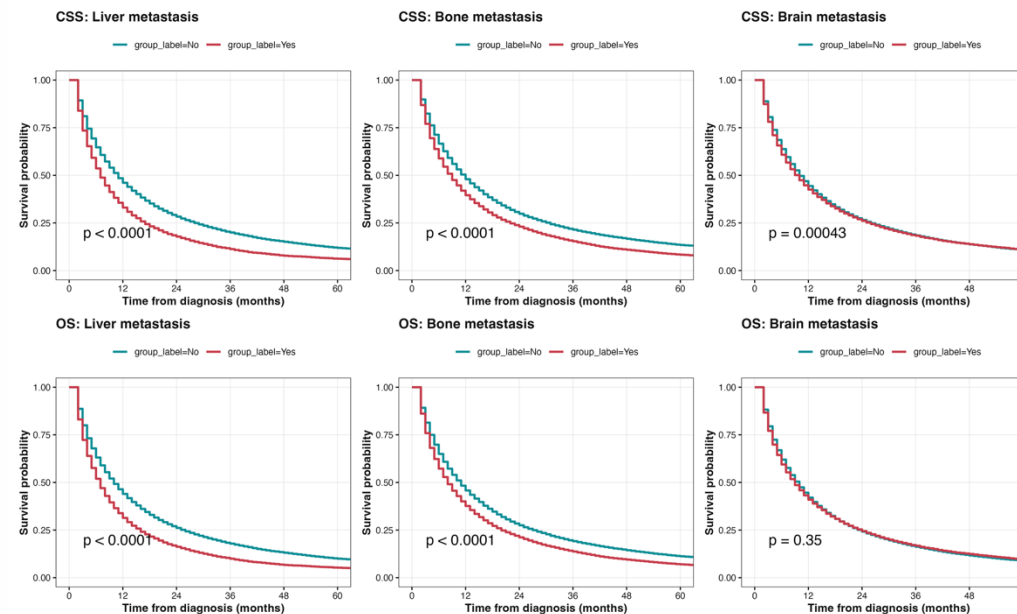
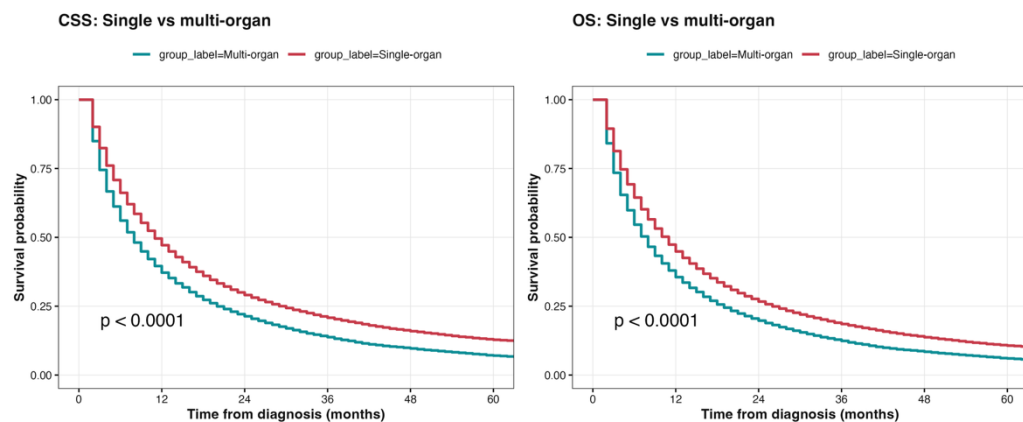
Multivariable Model C; HR > 1 indicates higher event hazard



Adjusted HRs

- Multi-organ burden remains adverse.
- Liver, bone, and brain metastases show higher risk.
- Lung metastasis is not a worse-survival signal after adjustment.

Kaplan-Meier curves show survival gradients by metastatic extent

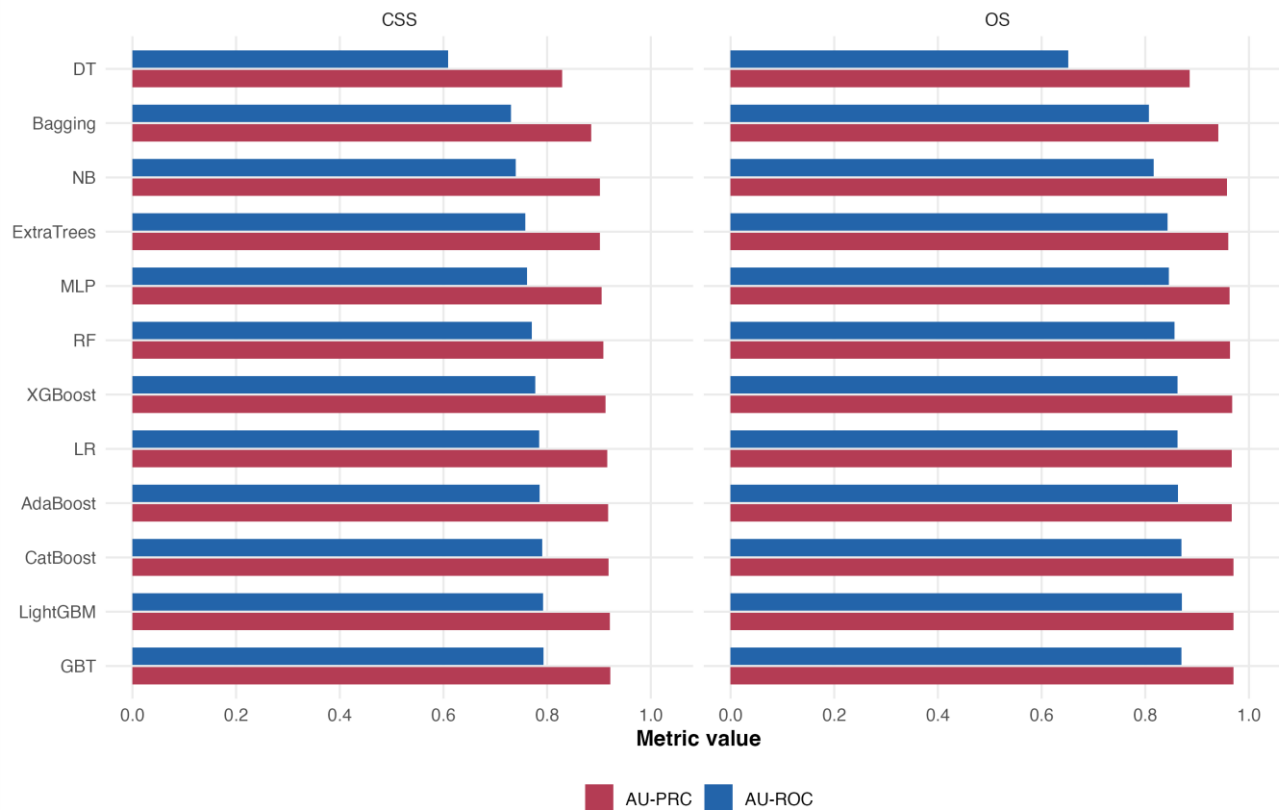


Visible survival separation before modeling remains directionally consistent after multivariable adjustment.

Original Python pipeline compares 12 classifiers

Original 12-Model Classification Performance

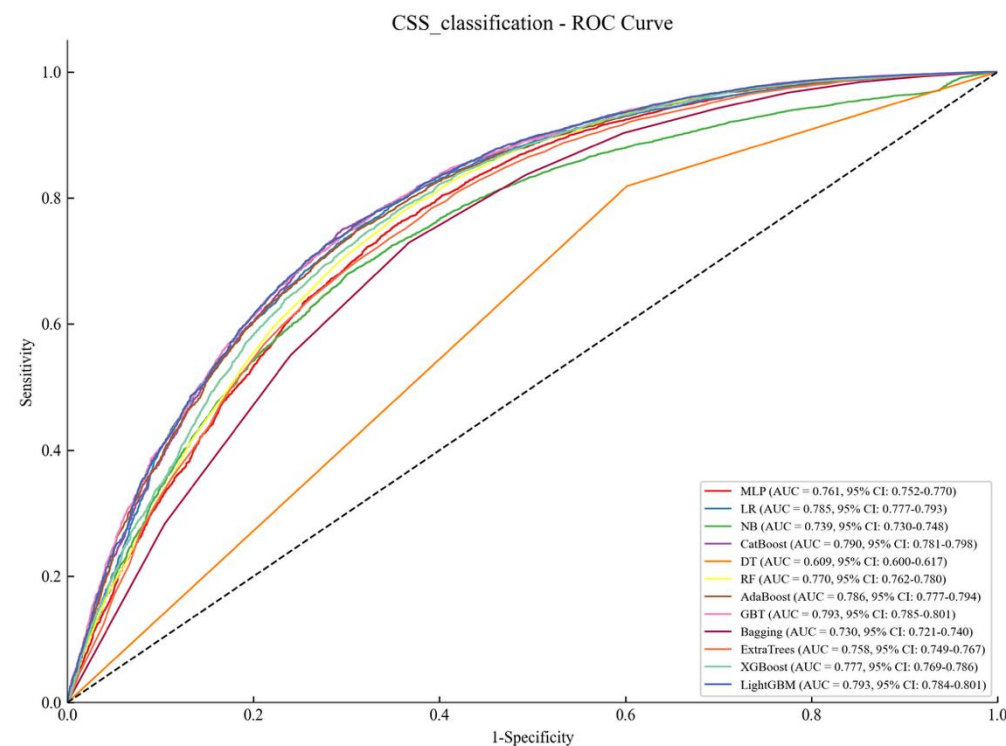
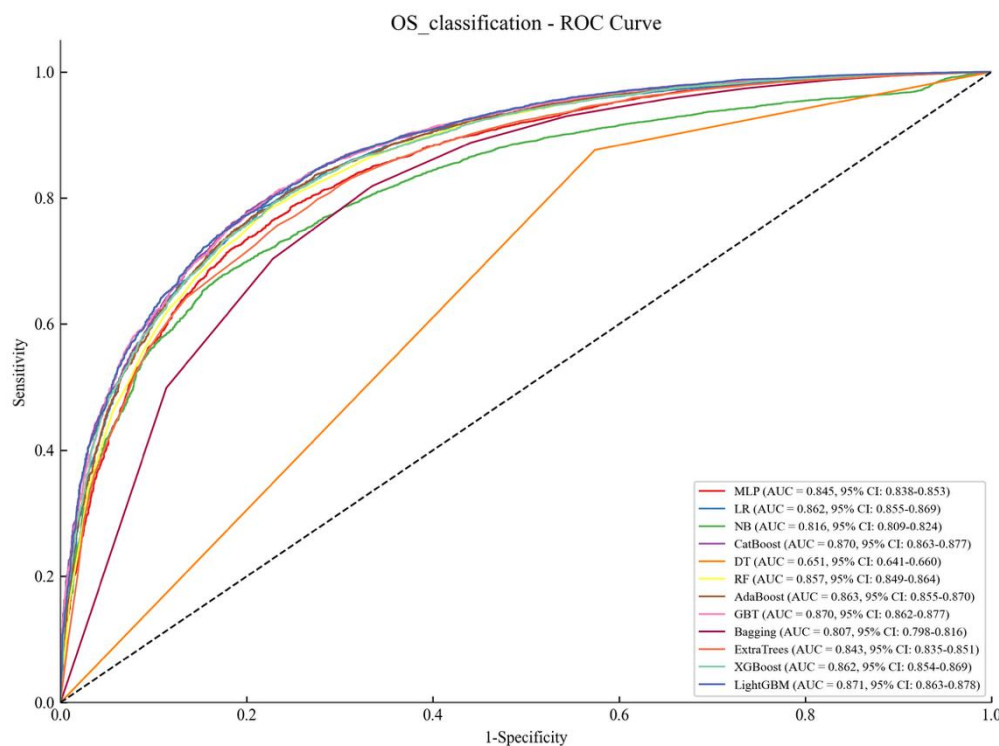
Test-set AU-ROC and AU-PRC from prior Python ML pipeline



OS
LightGBM
AUC 0.871
AP 0.970

CSS
CatBoost
AUC 0.790
AP 0.919

LightGBM and CatBoost provide the abstract-matched classification benchmarks

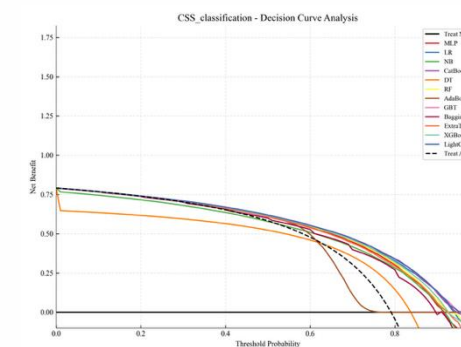
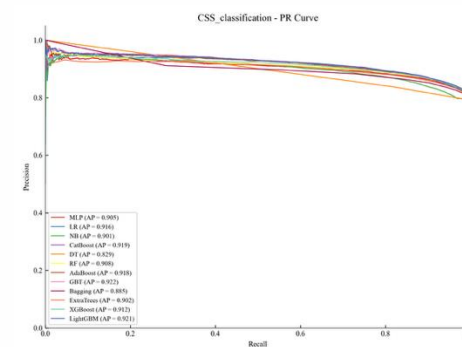
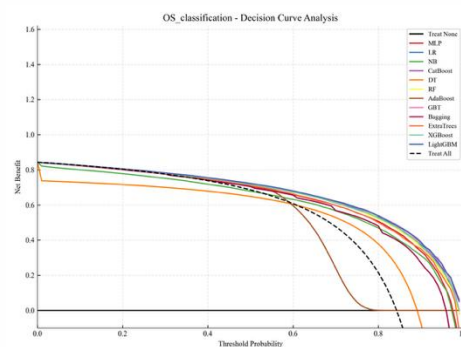
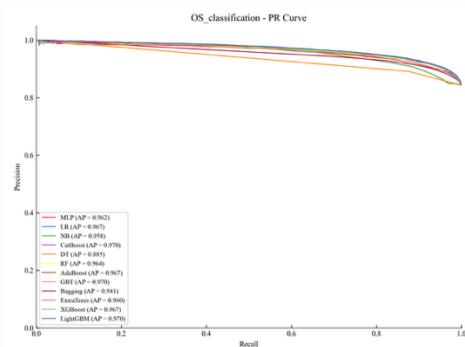


OS best model: LightGBM AU-ROC 0.871. CSS best model: CatBoost AU-ROC 0.790.

Precision-recall and DCA support clinical model utility

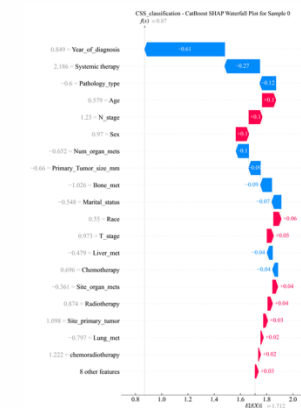
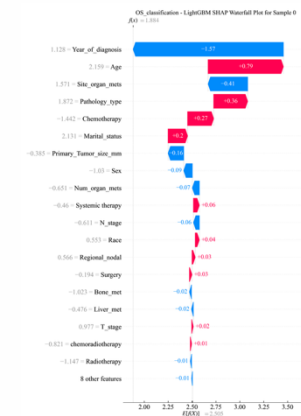
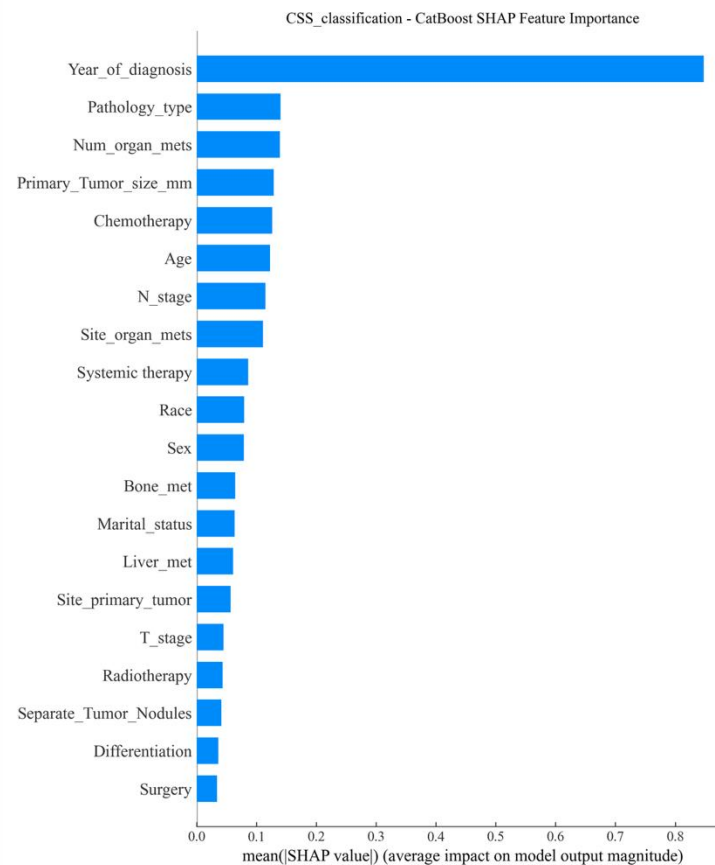
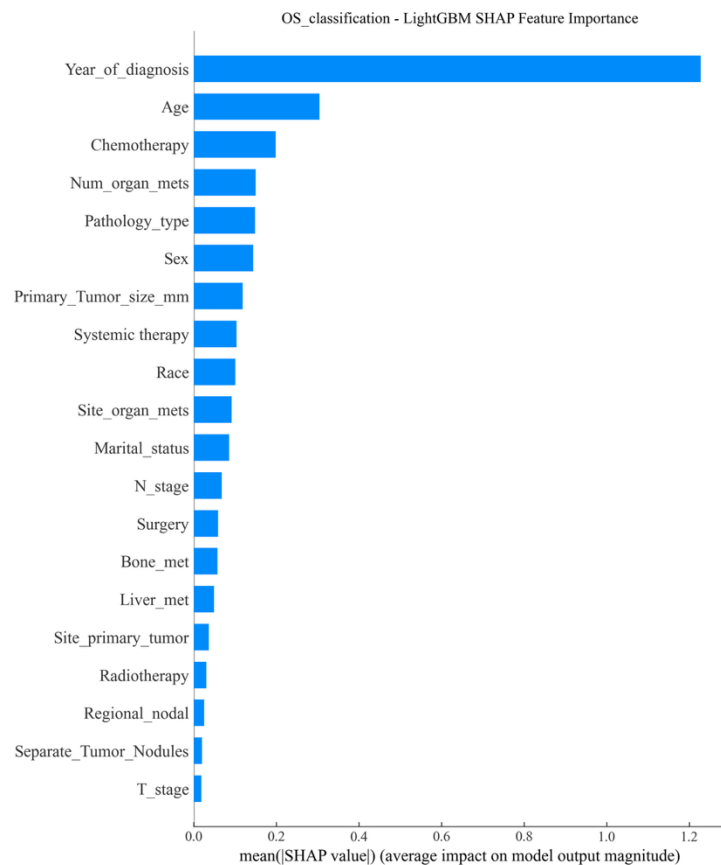
OS classification

CSS classification



Average precision: OS 0.970; CSS 0.919. Calibration curves are retained in the supplementary asset folder.

SHAP highlights metastatic, treatment, and tumor predictors



Take-home message and disclosure

- Metastatic burden and organ-specific sites stratify OS and CSS in 62,606 SEER stage IV NSCLC cases.
- Original ML classifiers show strong discrimination: OS LightGBM AU-ROC 0.871 and CSS CatBoost AU-ROC 0.790.
- SHAP interpretation links predictions to metastatic, treatment, tumor, and demographic domains.