



OPEN Development and validation of interpretable machine learning models to predict distant metastasis and prognosis of muscle-invasive bladder cancer patients

Qian Deng^{1,6}, Shan Li^{2,6}, Yuxiang Zhang³, Yuanyuan Jia⁴✉ & Yanhui Yang⁵✉

Muscle-Invasive Bladder Cancer (MIBC) is a more aggressive disease than non-muscle-invasive bladder cancer (NMIBC), with greater chances of metastasis. We sought to develop machine learning (ML) models to predict metastasis and prognosis in MIBC patients. Clinical data of MIBC cases from 2000 to 2020 were sourced from the Surveillance, Epidemiology, and End Results (SEER) database. Clinical variables used to predict DM were identified through univariate and multivariate logistic regression, and Recursive Feature Elimination (RFE). Thirteen ML models predicting DM were evaluated based on AUC, PRAUC, accuracy, sensitivity, specificity, precision, cross-entropy, Brier score, balanced accuracy, and F-beta score. SHapley Additive exPlanations (SHAP) framework helped interpret the best model. Additionally, we utilized ML algorithm combinations to predict prognosis in MIBC patients with metastasis. A total of 43,951 T2-T4 MIBC patients aged over 18 years old from the SEER database were enrolled consecutively. Nine clinical variables were selected to predict DM. The CatBoost model was identified as the optimal predictor, with AUC values of 0.956 [0.933, 0.969] for the training set, 0.882 [0.857, 0.919] for the internal test set, and 0.839 [0.723, 0.936] for the external test set. The model achieved an accuracy of 0.875 [0.854, 0.896], sensitivity of 0.869 [0.851, 0.889], specificity of 0.883 [0.823, 0.912], and precision of 0.917 [0.885, 0.944]. SHAP analysis revealed that tumor size was the most influential factor in predicting distant metastasis. For prognosis, the “RSF + Enet[alpha = 0.8]” model emerged as the top performer, with C-index values of 0.683 in training, 0.688 in the internal test, and 0.666 in the external test sets. Our ML models provide high accuracy and dependability, delivering refined, individualized predictions for metastasis risk and prognosis in MIBC patients.

Keywords Machine learning, Muscle-Invasive bladder Cancer, Distant metastasis, Prognosis prediction, SEER

Bladder cancer (BLCA) is the leading malignancy affecting the urinary tract and ranks as the world's ninth most common cancer¹. Urothelial cell carcinoma (UCC), which generally originates in the lower urinary tract such as the bladder and urethra but can also develop in upper urinary structures like the renal pelvis and ureters, is the predominant subtype of BLCA. Clinically, bladder cancer is classified into non-muscle-invasive bladder cancer (NMIBC) and muscle-invasive bladder cancer (MIBC) according to tumor penetration depth. NMIBC includes early-stage lesions limited to the mucosa or submucosa, such as carcinoma in situ (CIS; Tis), non-

¹Luoyang Central Hospital Affiliated of Zhengzhou University, Henan, China. ²Department of Urology, Children's Hospital of Chongqing Medical University, Chongqing, China. ³Department of Urology Surgery, The First Affiliated Hospital, and College of Clinical Medicine of Henan University of Science and Technology, Luoyang, Henan, China. ⁴Department of Oncology, Huai'an Second People's Hospital, Affiliated to Xuzhou Medical University, Huai'an, Jiangsu, China. ⁵Department of Emergency Surgery (Trauma Center), The First Affiliated Hospital, and College of Clinical Medicine of Henan University of Science and Technology, Luoyang, Henan, China. ⁶Qian Deng and Shan Li contributed equally to this work. ✉email: 516791229@qq.com; hkdyh2024@sina.com

invasive papillary tumors (Ta), and tumors that extend into the lamina propria (T1)². In comparison, MIBC, identified by detrusor muscle invasion and classified as T2–T4 disease, carries a greater likelihood of lymph node spread and distant metastasis³. This underscores the importance of early detection of metastatic MIBC to improve treatment strategies and patient outcomes.

MIBC is highly aggressive, with a strong tendency for both local and distant spread, greatly elevating mortality risk in affected individuals. Around 50% of MIBC cases show micro metastases at the time of diagnosis, primarily involving lymph nodes, lungs, liver, and bones⁴. Such metastatic occurrences are linked to a poor outlook, as the five-year survival rate for metastatic MIBC patients remains low, between 5% and 15%⁵. Additionally, 20–30% of those initially diagnosed with NMIBC experience disease progression to MIBC during the disease course⁶. The primary treatment for MIBC includes cisplatin-based neoadjuvant chemotherapy (NAC), followed by radical cystectomy (RC) and pelvic lymph node dissection (PLND)³. However, RC is a highly invasive surgery with notable perioperative risks, leading a significant proportion of MIBC patients to be considered ineligible and unable to undergo the procedure^{7,8}. Furthermore, postoperative complications, especially those related to urinary diversion, can significantly diminish quality of life (QoL), making RC an unfavorable choice for some. Although recent urinary assays have enhanced early detection of bladder cancer, treatment options for MIBC remain limited, and patient prognosis often remains poor⁹. Trimodality therapy (TMT), which includes maximal transurethral resection of the bladder tumor (TURBT) followed by concurrent chemoradiotherapy, has emerged as a promising option for patients in T2N0M0 disease¹⁰. Previously, MIBC patients unfit for RC due to physical or psychological constraints would choose TMT, but it is generally limited to those who meet certain clinical criteria¹¹. This highlights the urgent need to further investigate the risk factors influencing MIBC metastasis and to advance therapeutic strategies that enhance patient survival.

Recently, nomograms have gained popularity for prognosis prediction, though their sensitivity and specificity can be limited. Machine learning (ML), a branch of artificial intelligence, offers a practical alternative due to its strength in managing complex, non-linear relationships, making it especially effective for large datasets and unstructured information¹². With advances in big data analytics, creating new tools to identify risk factors impacting metastasis and prognosis has become achievable. However, no ML-based model currently exists for predicting metastasis in MIBC, highlighting the need for an accurate model's development and validation. The Surveillance, Epidemiology, and End Results (SEER) database (<https://seer.cancer.gov/>) provides extensive clinicopathological data and follow-up records for cancer patients, making it an invaluable asset for developing and validating ML models in healthcare. Although ML is widely applied across numerous fields, there is still a significant lack of research dedicated to forecasting metastasis and prognosis in MIBC patients. This study seeks to fill this gap by creating and validating predictive models with various ML algorithms, offering clinicians personalized tools to more precisely evaluate metastasis risk and prognostic outcomes for individuals with MIBC.

Materials and methods

Data collection and patient demographics

Clinicopathological information of MIBC patients from 2000 to 2020 was sourced from the SEER database. Additionally, retrospective data from The First Affiliated Hospital of Henan University of Science and Technology (2010–2020) was gathered through its electronic medical records system. This study adhered to the Declaration of Helsinki and received approval from Ethics Committee of The First Affiliated Hospital of Henan University of Science and Technology (protocol code: 2024-03-K0144, approved on 2024-03-20), with written informed consent obtained from all participants. Inclusion criteria specified T2–T4 MIBC patients with an initial diagnosis of bladder cancer (C67.0–C67.9) per the Third Edition of the International Classification of Diseases for Oncology (ICD-O-3) and patients over 18 years old. Exclusion criteria included cases missing follow-up information on survival and cause of death, survival less than one month, lacking histological confirmation, absence of surgical data, and incomplete TNM staging or grading. Selecting one month as the time point for landmark analysis can exclude the interference of patients who died early, and more accurately evaluate the long-term effects of the clinical factors on those patients who were able to survive until that time point. Metastasis was defined as the spread to distant sites at initial diagnosis.

We gathered patient data covering demographics (age, gender, race, marital status, household location, and income), disease attributes (pathological grade, TNM stage, tumor size, primary tumor location, pathology, and metastasis details), treatment information (surgery, lymph node surgery, lymph node biopsy, radiotherapy, chemotherapy), and follow-up outcomes (overall survival (OS) and cancer-specific survival (CSS) status, survival duration in months). Age and tumor size, both continuous variables, were transformed into categorical variables based on established clinical practices and prior research findings. Age was grouped into “<40,” “40–59,” “60–69,” “70–79,” and “≥80” categories^{13,14}, while tumor size was categorized as “≤2 cm,” “2.1–4 cm,” “4.1–6 cm,” “>6 cm,” and “Unknown”^{15–17}. Tumor size was measured based on imaging and pathology reports according to SEER database guideline. Cases were classified as having “Distant Metastasis” if metastasis was identified in the brain, bone, liver, lung, or distant lymph nodes during follow-up, or if tumors were verified as M1 stage via pathology or imaging at diagnosis. Missing data from either the database or hospital records were categorized as “Unknown.” The minimum sample size required for the external validation cohort was calculated using Riley et al.'s formula¹⁸.

Model development and validation for predicting distant metastasis

To identify predictive variables, we employed univariate and multivariate logistic regression analyses in the training cohort, selecting those with P-values below 0.05 for the following feature selection. We then applied recursive feature elimination (RFE) technique to perform feature selection based on six ML algorithms, namely categorical boosting (CatBoost), random forest (RF), support vector machine (SVM), extreme gradient boosting

(XGB), decision tree (DT), and gradient boosting machine (GBM), with 10-fold cross-validation¹⁹. RFE operates by training an ML model and ranking features based on their predictive strength, iteratively evaluating and removing features until all are assessed and prioritized²⁰. We applied the Robust Rank Aggregation (RRA) method to consolidate the importance ranks of variables across the six ML algorithms in RFE process, yielding an integrated ranking of clinical variables²¹. The random seed was set to “123”. After feature selection, we developed the model by 13 ML algorithms, including CatBoost, RF, SVM, XGB, DT, GBM, k-nearest neighbor (KNN), logistic regression (LR), naive bayes classifier (NBC), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), neural network (NNET), and generalized linear model (GLM), instead of building a mixed model, using the “mlr3” R package²². This approach enabled us to compare each algorithm’s performance to identify the most effective predictive model for distant metastasis. To address potential class imbalance, which could bias performance metrics, we used the Synthetic Minority Over-sampling Technique (SMOTE) during model training²³. We refined the model development by applying nested resampling, which utilized a two-level k-fold cross-validation structure: an outer layer for hyperparameter optimization and an inner layer dedicated to model selection. Additionally, we conducted a random search with 1,000 evaluations within a 10-fold cross-validation framework, repeating the process five times for each model.

Subsequently, metrics including area under the curve (AUC), area under the precision-recall curve (PRAUC), accuracy, sensitivity, specificity, precision, cross-entropy, Brier score, balanced accuracy (bacc), and F-beta score (beta = 1) were calculated to identify the top-performing ML model. For internal validation, we employed 10-fold cross-validation. Precision-recall curves (PRC) were generated to evaluate each model’s handling of imbalanced data, while calibration curves assessed discrimination ability. Decision curve analysis (DCA) was used to gauge the clinical benefits of the ML model, implemented through the “runway” R package (<https://github.com/ML4LHS/runway/>). The best model was selected based on high AUC, high PRAUC, low Brier score, an accurate calibration curve, and balanced accuracy and F-beta score (beta = 1). To determine each variable’s significance, we calculated its mean contribution to the AUC as a percentage of the full model using the “DALEX” R package²⁴. SHapley Additive exPlanations (SHAP) values, visualized in a bee swarm plot via the “shapviz” R package (<https://github.com/ModelOriented/shapviz>)²⁵, were applied to interpret the black-box ML model.

Integrating ML algorithms for prognosis prediction

Univariate and multivariate Cox analyses were conducted to identify clinical variables with P-values below 0.05 that could predict overall survival (OS). We applied 10 ML algorithms, including random survival forest (RSF), elastic network (Enet), Lasso, Ridge, stepwise Cox, CoxBoost, partial least squares regression for Cox (plsRcox), supervised principal components (SuperPC), GBM, and survival support vector machine (survival-SVM), to estimate prognosis in terms of OS for MIBC patients with distant metastasis. In total, 101 ML algorithm configurations were tested on the training set, using a leave-one-out cross-validation (LOOCV) framework to identify the optimal prognostic model. Models utilizing fewer than three clinical variables were excluded from consideration. The concordance index (C-index) for each ML configuration was then computed across training, internal validation, and external validation sets. The top five ML configurations, achieving the highest average C-index across the three datasets, were selected for subsequent assessment using k-fold cross-validation to address overfitting and ensure model stability and generalizability. Logarithmic loss, recall, and decision calibration were calculated to identify the most effective prognostic ML model using the “mlr3proba” R package²⁶. Patient risk scores were derived through a linear combination for each prognostic ML configuration. The median risk score in the training set was used as a threshold, allowing us to categorize patients in the training, internal validation, and external validation sets into high- or low-risk groups. Kaplan-Meier (KM) survival analysis and log-rank testing were conducted with the “survival” and “survminer” R packages. Additionally, AUC, time-dependent ROC curves, calibration plots, and DCA were performed to evaluate the precision, discrimination capability, and clinical value of the optimal model.

Results

Demographic and clinical baseline characteristics

To develop a predictive model for distant metastasis, data from 43,951 MIBC patients in the SEER database and 125 MIBC patients from The First Affiliated Hospital of Henan University of Science and Technology (used as the external validation set) were collected. SEER patients were randomly divided into training and internal test sets with a 7:3 split. In the training cohort, 2,513 patients (8.17%) exhibited distant metastasis, while 28,253 patients (91.8%) did not. The internal test cohort included 1,048 patients (8.02%) with metastasis and 12,012 patients (92.0%) without it. Detailed clinical characteristics for both groups are provided in Table 1.

To construct a prognostic model for predicting outcomes in MIBC patients with distant metastasis, patients from the SEER database with distant metastasis were similarly divided into training and internal test sets in a 7:3 ratio. An additional 100 MIBC patients with distant metastasis from The First Affiliated Hospital of Henan University of Science and Technology served as the external validation group. In the training set, 2,394 patients (95.5%) had died by the last follow-up, while 114 (4.55%) remained alive (Table 2). For the internal test set, 1,016 patients (96.5%) had died, and 37 (3.51%) were alive (Table 2). Median follow-up durations for the training, internal, and external validation sets were 7.00 [3.00;15.0] months, 7.00 [3.00;14.0] months, and 7.00 [2.00;14.0] months, respectively (Table 2). Figure 1 outlines the selection process for MIBC patients from the SEER database.

Feature selection for the predictive model

We began by calculating the correlation coefficients of baseline clinical variables and visualizing the results with a lollipop chart, which indicated that “N stage” had the strongest association with metastasis (Fig. 2A). Based on clinical expertise, 19 variables were selected for logistic regression analysis (Table 3), while variables with correlation coefficients above 0.6 were excluded (Supplementary Fig. 1A). Next, univariate and multivariate

	Training cohort N = 30,766 (70.2%)	Internal validation cohort N = 13,060 (29.8%)	External validation cohort N = 125	p.overall
Sex				0.287
Male	22,625 (73.5%)	9509 (72.8%)	92 (73.6%)	
Female	8141 (26.5%)	3551 (27.2%)	33 (26.4%)	
Age				0.294
<40	159 (0.52%)	76 (0.58%)	1 (0.80%)	
40–59	4649 (15.1%)	1995 (15.3%)	28 (22.4%)	
60–69	7254 (23.6%)	3038 (23.3%)	34 (27.2%)	
70–79	9453 (30.7%)	4016 (30.8%)	32 (25.6%)	
≥80	9251 (30.1%)	3935 (30.1%)	30 (24.0%)	
Race				0.001
White	27,236 (88.5%)	11,466 (87.8%)	0 (0.00%)	
Black	2071 (6.73%)	972 (7.44%)	0 (0.00%)	
Asian and pacific islanders	1106 (3.59%)	508 (3.89%)	125 (100%)	
Other	353 (1.15%)	114 (0.87%)	0 (0.00%)	
Marital_Status				0.164
Married	17,551 (57.0%)	7374 (56.5%)	66 (52.8%)	
Unmarried	6475 (21.0%)	2723 (20.8%)	26 (20.8%)	
Widowed	5644 (18.3%)	2455 (18.8%)	24 (19.2%)	
Unknown	1096 (3.56%)	508 (3.89%)	9 (7.20%)	
Year_of_Diagnosis				0.259
2000–2005	9566 (31.1%)	3962 (30.3%)	41 (32.8%)	
2006–2010	10,355 (33.7%)	4412 (33.8%)	48 (38.4%)	
2011–2020	10,845 (35.2%)	4686 (35.9%)	36 (28.8%)	
Household_Location				0.890
Rural	4261 (13.8%)	1792 (13.7%)	16 (12.8%)	
Urban	26,505 (86.2%)	11,268 (86.3%)	109 (87.2%)	
Household_Income				0.838
<\$75,000	18,903 (61.4%)	8043 (61.6%)	74 (59.2%)	
≥\$75,000	11,863 (38.6%)	5017 (38.4%)	51 (40.8%)	
Primary_Site				0.117
Bladder wall	19,683 (64.0%)	8435 (64.6%)	70 (56.0%)	
Bladder base	3825 (12.4%)	1568 (12.0%)	15 (12.0%)	
Urachus/dome of bladder	1777 (5.78%)	703 (5.38%)	8 (6.40%)	
Overlapping lesion	5481 (17.8%)	2354 (18.0%)	32 (25.6%)	
Histology				0.887
Transitional cell carcinoma	17,860 (58.1%)	7664 (58.7%)	77 (61.6%)	
Papillary transitional cell carcinoma	9755 (31.7%)	4045 (31.0%)	35 (28.0%)	
Adenocarcinoma	693 (2.25%)	298 (2.28%)	4 (3.20%)	
Squamous cell carcinoma	1199 (3.90%)	519 (3.97%)	4 (3.20%)	
Other	1259 (4.09%)	534 (4.09%)	5 (4.00%)	
Grade				0.089
I	429 (1.39%)	168 (1.29%)	3 (2.40%)	
II	1993 (6.48%)	869 (6.65%)	16 (12.8%)	
III	11,504 (37.4%)	4830 (37.0%)	42 (33.6%)	
IV	16,840 (54.7%)	7193 (55.1%)	64 (51.2%)	
Laterality				0.438
Bilateral	321 (1.04%)	128 (0.98%)	0 (0.00%)	
Lateral	30,445 (99.0%)	12,932 (99.0%)	125 (100%)	
T_Stage				0.441
T2	18,930 (61.5%)	8004 (61.3%)	71 (56.8%)	
T3	6963 (22.6%)	2986 (22.9%)	37 (29.6%)	
T4	4873 (15.8%)	2070 (15.8%)	17 (13.6%)	
N_Stage				0.996
N0	26,000 (84.5%)	11,059 (84.7%)	107 (85.6%)	
N1	2343 (7.62%)	985 (7.54%)	9 (7.20%)	
Continued				

	Training cohort N= 30,766 (70.2%)	Internal validation cohort N= 13,060 (29.8%)	External validation cohort N= 125	p.overall
N2	2318 (7.53%)	969 (7.42%)	9 (7.20%)	
N3	105 (0.34%)	47 (0.36%)	0 (0.00%)	
M_Stage				0.687
M0	28,253 (91.8%)	12,012 (92.0%)	117 (93.6%)	
M1	2513 (8.17%)	1048 (8.02%)	8 (6.40%)	
Surgery_Type				0.353
No surgery	890 (2.89%)	357 (2.73%)	4 (3.20%)	
Local tumor destruction/excision	16,375 (53.2%)	6855 (52.5%)	57 (45.6%)	
Partial cystectomy	1236 (4.02%)	548 (4.20%)	7 (5.60%)	
Complete cystectomy	12,265 (39.9%)	5300 (40.6%)	57 (45.6%)	
Lymph_Nodes_Surgery				0.060
None	16,265 (52.9%)	6767 (51.8%)	55 (44.0%)	
Regional lymph nodes removed	10,435 (33.9%)	4488 (34.4%)	48 (38.4%)	
Unknown	4066 (13.2%)	1805 (13.8%)	22 (17.6%)	
Lymph_Node_Biopsy				0.526
No biopsy	18,766 (61.0%)	7889 (60.4%)	75 (60.0%)	
Negative	8448 (27.5%)	3661 (28.0%)	38 (30.4%)	
Positive	3483 (11.3%)	1476 (11.3%)	11 (8.80%)	
Unknown	69 (0.22%)	34 (0.26%)	1 (0.80%)	
Surgery_in_Other_Sites				0.342
None	25,098 (81.6%)	10,566 (80.9%)	97 (77.6%)	
Yes	1631 (5.30%)	701 (5.37%)	7 (5.60%)	
Unknown	4037 (13.1%)	1793 (13.7%)	21 (16.8%)	
Chemotherapy				0.532
None/unknown	19,458 (63.2%)	8253 (63.2%)	73 (58.4%)	
Yes	11,308 (36.8%)	4807 (36.8%)	52 (41.6%)	
Radiotherapy				0.136
None/unknown	25,496 (82.9%)	10,925 (83.7%)	104 (83.2%)	
Yes	5270 (17.1%)	2135 (16.3%)	21 (16.8%)	
Tumor_Size				0.660
0–2.0	2227 (7.24%)	936 (7.17%)	7 (5.60%)	
2.1–4.0	5258 (17.1%)	2186 (16.7%)	15 (12.0%)	
4.1–6.0	4539 (14.8%)	1970 (15.1%)	21 (16.8%)	
>6.0	3013 (9.79%)	1323 (10.1%)	15 (12.0%)	
Unknown	15,729 (51.1%)	6645 (50.9%)	67 (53.6%)	
Metastasis				0.687
No	28,253 (91.8%)	12,012 (92.0%)	117 (93.6%)	
Yes	2513 (8.17%)	1048 (8.02%)	8 (6.40%)	
Bone_Metastasis				0.178
No	12,569 (40.9%)	5363 (41.1%)	42 (33.6%)	
Yes	444 (1.44%)	172 (1.32%)	0 (0.00%)	
Unknown	17,753 (57.7%)	7525 (57.6%)	83 (66.4%)	
Brain_Metastasis				0.205
No	12,991 (42.2%)	5529 (42.3%)	42 (33.6%)	
Yes	31 (0.10%)	7 (0.05%)	0 (0.00%)	
Unknown	17,744 (57.7%)	7524 (57.6%)	83 (66.4%)	
Liver_Metastasis				0.267
No	12,782 (41.5%)	5447 (41.7%)	41 (32.8%)	
Yes	236 (0.77%)	88 (0.67%)	1 (0.80%)	
Unknown	17,748 (57.7%)	7525 (57.6%)	83 (66.4%)	
Lung_Metastasis				0.402
No	12,613 (41.0%)	5374 (41.1%)	41 (32.8%)	
Yes	385 (1.25%)	162 (1.24%)	1 (0.80%)	
Unknown	17,768 (57.8%)	7524 (57.6%)	83 (66.4%)	
Vital_Status				0.627
Continued				

	Training cohort N = 30,766 (70.2%)	Internal validation cohort N = 13,060 (29.8%)	External validation cohort N = 125	p.overall
Alive	5462 (17.8%)	2340 (17.9%)	26 (20.8%)	
Dead	25,304 (82.2%)	10,720 (82.1%)	99 (79.2%)	
Cancer_Specific_Death				0.323
Not cancer specific death	13,687 (44.5%)	5796 (44.4%)	59 (47.2%)	
Dead due to bladder cancer	16,834 (54.7%)	7181 (55.0%)	66 (52.8%)	
Unknown	245 (0.80%)	83 (0.64%)	0 (0.00%)	
Survival_Months	21.0 [8.00;73.0]	21.0 [8.00;75.0]	28.0 [12.0;84.0]	0.124

Table 1. Clinicopathological characteristics of MIBC patients in the training, internal validation and external validation cohorts.

logistic regression analyses were conducted on the training set, identifying “Age,” “Primary Site,” “Histology,” “Grade,” “T Stage,” “N Stage,” “Surgery Type,” “Lymph Node Surgery,” “Lymph Node Biopsy,” “Surgery in Other Sites,” “Radiotherapy,” “Chemotherapy,” and “Tumor Size” as significant predictors for metastasis ($P < 0.05$, Table 3; Fig. 2C). Further correlation analysis between model variables and metastasis highlighted “N stage” as the most influential predictor (Fig. 2B). We applied RFE with six ML algorithms (GBM, SVM, RF, DT, XGB, and CatBoost) to identify the relevant clinical features (Fig. 2D–I). RFE analysis revealed that RF was the optimal model, retaining 11 variables and achieving the highest AUC of 0.844 (Fig. 2F). The RRA method was then used to create an overall ranking of each clinical variable across the six ML algorithms, with “Age” emerging as the most influential factor (Supplementary Table 1). Given the strong associations between “Surgery Type,” “Lymph Nodes Surgery,” “Lymph Node Biopsy,” and “Surgery in Other Sites,” only “Surgery Type” and “Surgery in Other Sites” were retained for further analysis (Supplementary Fig. 1A). Ultimately, nine clinical variables with a frequency greater than three in RRA analysis, indicating substantial importance in most ML selection methods, were included in the subsequent model development steps (Supplementary Table 1). In summary, we selected the ultimate model variables by logistic regression analysis, RFE based on six ML algorithms, and RRA algorithm, guaranteeing the robustness of our feature selection process.

Design and assessment of the predictive model

To accurately predict metastasis, “Age,” “Chemotherapy,” “N Stage,” “Histology,” “Radiotherapy,” “Surgery in Other Sites,” “Surgery Type,” “Tumor Size,” and “T Stage” were selected through the feature selection process. Thirteen ML algorithms, including CatBoost, RF, SVM, XGB, DT, GBM, KNN, LR, NBC, LDA, QDA, NNET, and GLM, were employed to build the predictive model in the training set, with hyperparameters fine-tuned via 10-fold cross-validation and random search. Model performance was then evaluated in both internal and external validation sets. ROC curves (Figs. 3A, 4A and 5A) and AUC values (Figs. 3B, 4B and 5B) highlighted CatBoost as the leading model, achieving the highest AUC scores of 0.956 [0.933, 0.969], 0.882 [0.857, 0.919], and 0.839 [0.723, 0.936] across the training, internal test, and external test sets. To thoroughly evaluate model performance, we calculated accuracy, sensitivity, specificity, precision, cross-entropy, Brier scores, balanced accuracy (bacc), and F-beta score (beta = 1) across the 13 ML models. Results indicated that CatBoost was the most precise and reliable for predicting distant metastasis (Figs. 3C, 4C and 5C). CatBoost achieved outstanding metrics with an accuracy of 0.875 [0.854, 0.896], sensitivity of 0.869 [0.851, 0.889], specificity of 0.883 [0.823, 0.912], and precision of 0.917 [0.885, 0.944] (Fig. 3C). Grid search during hyperparameter tuning identified optimal CatBoost settings: depth of 4, learning rate of 0.03927546, iterations of 669, and leaf regularization of 6.098357. Precision-recall curves (PRC) demonstrated CatBoost’s effectiveness in managing imbalanced data (Figs. 3D, 4D and 5D). Calibration curves demonstrated that the CatBoost model excelled in both fit and predictive accuracy (Figs. 3E, 4E and 5E). DCA indicated that CatBoost provided the greatest benefit for predicting distant metastasis (Figs. 3F, 4F and 5F). 10-fold cross-validation further confirmed that CatBoost surpassed other ML algorithms (Supplementary Fig. 1B). Thus, CatBoost was identified as the most effective algorithm to build ML model for predicting distant metastasis, supporting clinical decision-making.

Model interpretation

To provide interpretability for the ML models, we calculated feature importance scores for each model, which were derived based on each algorithm’s intrinsic properties, showing that “N Stage” was the most influential risk factor for metastasis (Fig. 6A). We then applied the SHAP framework to further explain the top-performing CatBoost model. By evaluating the mean absolute SHAP values, it was evident that “Tumor size” had the greatest predictive power (Fig. 6B). Additionally, a bee swarm plot was used to illustrate the effect of each risk factor on metastasis (Fig. 6C). The y-axis indicates the strength of each risk factor, while the x-axis shows their impact on the model output, specifically the probability of metastasis, as quantified by SHAP values. The SHAP analysis revealed that factors such as absence of surgery, high N stage, advanced T stage, older age, lack of surgery in additional sites, and no radiotherapy are associated with a higher probability of metastasis (Fig. 6C). To demonstrate model interpretability, we examined two representative patients. SHAP values were employed to evaluate the influence of each feature on the model’s predictions. In this study, lower SHAP values corresponded to a reduced likelihood of metastasis, while higher values indicated an increased probability. We set the median SHAP value (0.0359) as the threshold for distinguishing between low and high metastasis risk. For example, the first patient, who did not experience metastasis, had a SHAP value and prediction score of -0.0143, suggesting a

	Training Cohort N=2508 (70.43%)	Internal Validation Cohort N=1053 (29.57%)	External Validation Cohort N=100	p-overall
Sex				0.351
Male	1824 (72.7%)	741 (70.4%)	71 (71.0%)	
Female	684 (27.3%)	312 (29.6%)	29 (29.0%)	
Age				0.081
<40	9 (0.36%)	8 (0.76%)	0 (0.00%)	
40–59	519 (20.7%)	213 (20.2%)	15 (15.0%)	
60–69	644 (25.7%)	289 (27.4%)	27 (27.0%)	
70–79	750 (29.9%)	311 (29.5%)	23 (23.0%)	
≥80	586 (23.4%)	232 (22.0%)	35 (35.0%)	
Race				0.044
White	2155 (85.9%)	908 (86.2%)	0 (0.00%)	
Black	238 (9.49%)	98 (9.31%)	0 (0.00%)	
Asian and pacific islanders	89 (3.55%)	36 (3.42%)	100 (100.00%)	
Other	26 (1.04%)	11 (1.04%)	0 (0.00%)	
Marital_Status				0.464
Married	1373 (54.7%)	561 (53.3%)	54 (54.0%)	
Unmarried	654 (26.1%)	284 (27.0%)	19 (19.0%)	
Widowed	406 (16.2%)	175 (16.6%)	23 (23.0%)	
Unknown	75 (2.99%)	33 (3.13%)	4 (4.00%)	
Year_of_Diagnosis				0.997
2000–2005	338 (13.5%)	144 (13.7%)	14 (14.0%)	
2006–2010	941 (37.5%)	397 (37.7%)	36 (36.0%)	
2011–2020	1229 (49.0%)	512 (48.6%)	50 (50.0%)	
Household_Location				0.014
Rural	319 (12.7%)	155 (14.7%)	5 (5.00%)	
Urban	2189 (87.3%)	898 (85.3%)	95 (95.0%)	
Household_Income				<0.001
<\$75,000	1628 (64.9%)	675 (64.1%)	15 (15.0%)	
≥\$75,000	880 (35.1%)	378 (35.9%)	85 (85.0%)	
Primary_Site				0.517
Bladder wall	1541 (61.4%)	664 (63.1%)	63 (63.0%)	
Bladder base	304 (12.1%)	112 (10.6%)	16 (16.0%)	
Urachus/dome of bladder	125 (4.98%)	46 (4.37%)	3 (3.00%)	
Overlapping lesion	538 (21.5%)	231 (21.9%)	18 (18.0%)	
Histology				0.331
Transitional cell carcinoma	1517 (60.5%)	640 (60.8%)	66 (66.0%)	
Papillary transitional cell carcinoma	628 (25.0%)	262 (24.9%)	25 (25.0%)	
Adenocarcinoma	65 (2.59%)	32 (3.04%)	0 (0.00%)	
Squamous cell carcinoma	109 (4.35%)	33 (3.13%)	1 (1.00%)	
Other	189 (7.54%)	86 (8.17%)	8 (8.00%)	
Grade				0.414
I	25 (1.00%)	9 (0.85%)	1 (1.00%)	
II	104 (4.15%)	39 (3.70%)	1 (1.00%)	
III	893 (35.6%)	374 (35.5%)	45 (45.0%)	
IV	1486 (59.3%)	631 (59.9%)	53 (53.0%)	
Laterality				0.561
Bilateral	31 (1.24%)	18 (1.71%)	1 (1.00%)	
Lateral	2477 (98.8%)	1035 (98.3%)	99 (99.0%)	
T_Stage				0.022
T2	1352 (53.9%)	512 (48.6%)	50 (50.0%)	
T3	413 (16.5%)	179 (17.0%)	13 (13.0%)	
T4	743 (29.6%)	362 (34.4%)	37 (37.0%)	
N_Stage				0.377
N0	1489 (59.4%)	585 (55.6%)	53 (53.0%)	
N1	423 (16.9%)	201 (19.1%)	22 (22.0%)	
N2	565 (22.5%)	255 (24.2%)	24 (24.0%)	
Continued				

	Training Cohort N=2508 (70.43%)	Internal Validation Cohort N=1053 (29.57%)	External Validation Cohort N=100	p-overall
N3	31 (1.24%)	12 (1.14%)	1 (1.00%)	
Surgery_Type				0.252
No surgery	212 (8.45%)	77 (7.31%)	4 (4.00%)	
Local tumor destruction/excision	1704 (67.9%)	700 (66.5%)	75 (75.0%)	
Partial cystectomy	43 (1.71%)	24 (2.28%)	1 (1.00%)	
Complete cystectomy	549 (21.9%)	252 (23.9%)	20 (20.0%)	
Lymph_Nodes_Surgery				0.219
None	1928 (76.9%)	789 (74.9%)	82 (82.0%)	
Regional lymph nodes removed	574 (22.9%)	258 (24.5%)	18 (18.0%)	
Unknown	6 (0.24%)	6 (0.57%)	0 (0.00%)	
Lymph_Node_Biopsy				0.710
No biopsy	1903 (75.9%)	787 (74.7%)	81 (81.0%)	
Negative	135 (5.38%)	57 (5.41%)	2 (2.00%)	
Positive	459 (18.3%)	204 (19.4%)	17 (17.0%)	
Unknown	11 (0.44%)	5 (0.47%)	0 (0.00%)	
Surgery_in_Other_Sites				0.351
None	2244 (89.5%)	936 (88.9%)	94 (94.0%)	
Yes	262 (10.4%)	117 (11.1%)	6 (6.00%)	
Unknown	2 (0.08%)	0 (0.00%)	0 (0.00%)	
Chemotherapy				0.412
None/unknown	1113 (44.4%)	465 (44.2%)	51 (51.0%)	
Yes	1395 (55.6%)	588 (55.8%)	49 (49.0%)	
Radiotherapy				0.030
None/unknown	1941 (77.4%)	845 (80.2%)	86 (86.0%)	
Yes	567 (22.6%)	208 (19.8%)	14 (14.0%)	
Tumor_Size				0.499
0-2.0	95 (3.79%)	38 (3.61%)	1 (1.00%)	
2.1-4.0	365 (14.6%)	176 (16.7%)	14 (14.0%)	
4.1-6.0	476 (19.0%)	202 (19.2%)	23 (23.0%)	
>6.0	455 (18.1%)	177 (16.8%)	22 (22.0%)	
Unknown	1117 (44.5%)	460 (43.7%)	40 (40.0%)	
Bone_Metastasis				0.420
No	976 (38.9%)	419 (39.8%)	41 (41.0%)	
Yes	453 (18.1%)	163 (15.5%)	15 (15.0%)	
Unknown	1079 (43.0%)	471 (44.7%)	44 (44.0%)	
Brain_Metastasis				0.931
No	1403 (55.9%)	574 (54.5%)	55 (55.0%)	
Yes	28 (1.12%)	10 (0.95%)	1 (1.00%)	
Unknown	1077 (42.9%)	469 (44.5%)	44 (44.0%)	
Liver_Metastasis				0.290
No	1215 (48.4%)	474 (45.0%)	47 (47.0%)	
Yes	215 (8.57%)	109 (10.4%)	10 (10.0%)	
Unknown	1078 (43.0%)	470 (44.6%)	43 (43.0%)	
Lung_Metastasis				0.873
No	1037 (41.3%)	423 (40.2%)	42 (42.0%)	
Yes	391 (15.6%)	156 (14.8%)	15 (15.0%)	
Unknown	1080 (43.1%)	474 (45.0%)	43 (43.0%)	
Vital_Status				0.377
Alive	114 (4.55%)	37 (3.51%)	4 (4.00%)	
Dead	2394 (95.5%)	1016 (96.5%)	96 (96.0%)	
Continued				

	Training Cohort N=2508 (70.43%)	Internal Validation Cohort N=1053 (29.57%)	External Validation Cohort N=100	p-overall
Cancer_Specific_Death				0.905
Not cancer specific death	365 (14.6%)	151 (14.3%)	18 (18.0%)	
Dead due to bladder cancer	2124 (84.7%)	894 (84.9%)	81 (81.0%)	
Unknown	19 (0.76%)	8 (0.76%)	1 (1.00%)	
Survival_Months	7.00 [3.00;15.0]	7.00 [3.00;14.0]	7.00 [2.00;14.0]	0.914

Table 2. Clinicopathological characteristics of MIBC patients with distant metastasis in the training, internal validation and external validation cohorts.

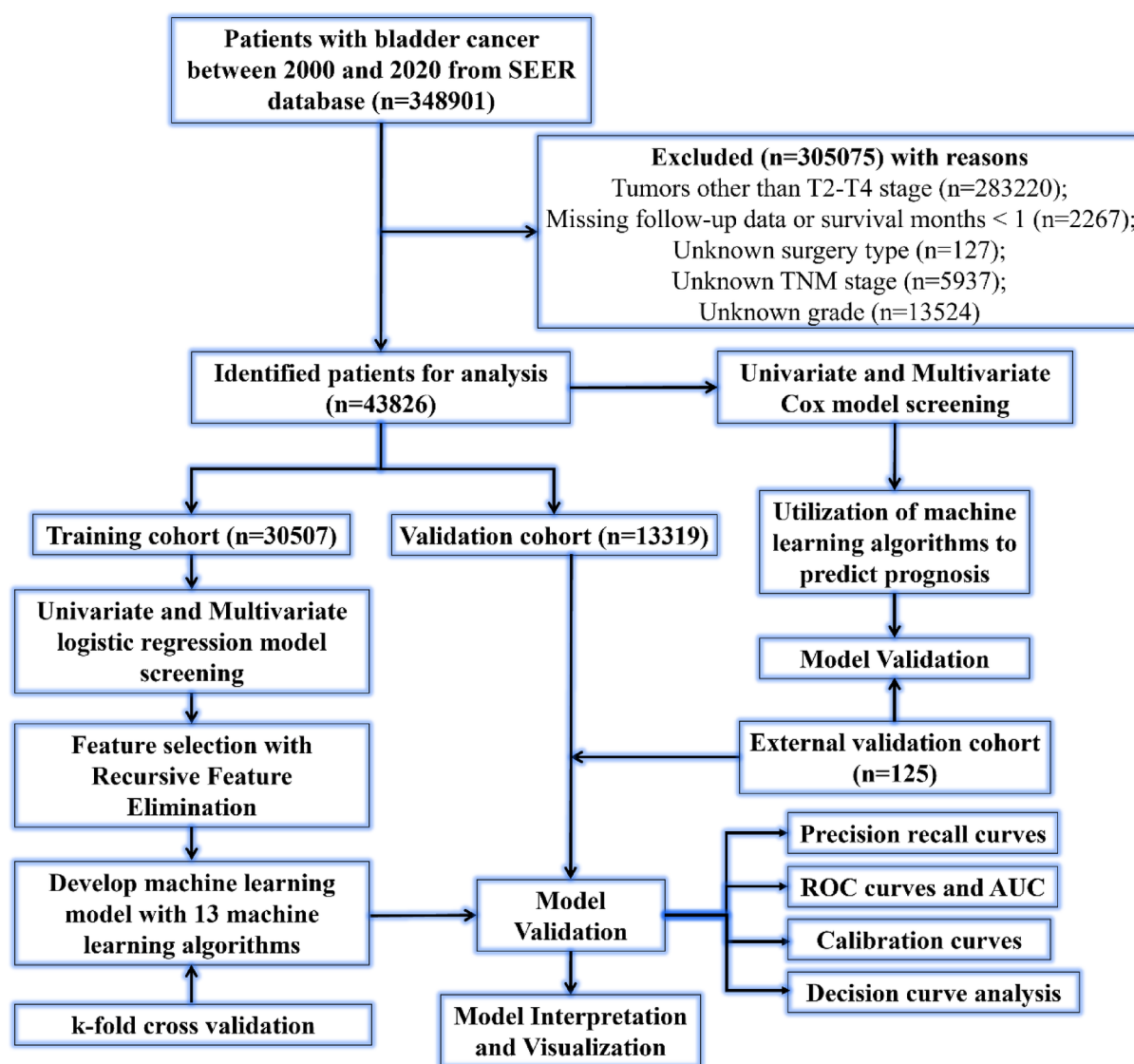


Fig. 1. The workflow diagram for study design and patient screening.

low metastasis risk (Fig. 6D). Conversely, the second patient, with confirmed metastasis, showed a SHAP value and prediction score of 0.0984, indicating a high probability of metastasis (Fig. 6E).

Prognostic model for metastatic MIBC patients

To build an accurate model for predicting OS in MIBC patients with metastasis, we performed univariate and multivariate Cox analyses. These analyses identified “Histology,” “T Stage,” “N Stage,” “Surgery Type,” “Lymph Node Surgery,” “Chemotherapy,” “Tumor Size,” “Bone Metastasis,” “Brain Metastasis,” “Liver Metastasis,” and “Lung Metastasis” as independent prognostic factors for OS in metastatic MIBC ($P < 0.05$, Table 4). Using the clinical variables identified, we generated 101 prognostic ML combinations within a LOOCV framework. The

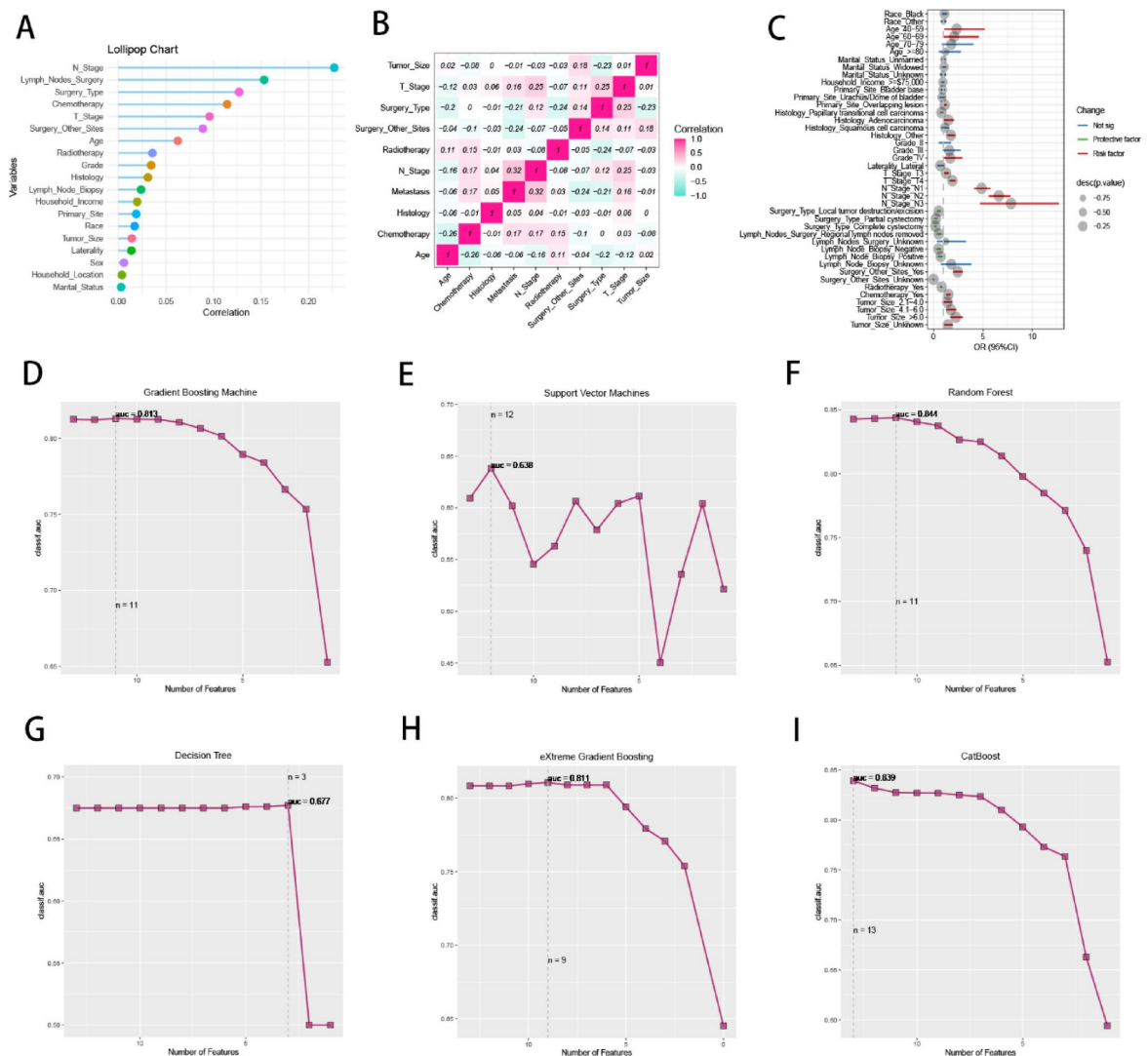


Fig. 2. The process of feature selection. (A) The correlation coefficients of the baseline characteristics with distant metastasis. (B) The heatmap of Spearman's correlation analysis among the clinical variables and distant metastasis. The correlation index ranges from -1.0 to 1.0, with a brighter color indicating a stronger correlation. (C) The forest plot visualized the feature selection with multivariate logistic regression analysis. (D-I) Feature selection process with Recursive Feature Elimination (RFE) method based on six ML algorithms (GBM, SVM, RF, DT, XGB and CatBoost).

C-index of each model was calculated across the training, internal validation, and external validation sets (Fig. 7A). Among the top five combinations with the highest C-index across all datasets, we assessed model performance through logarithmic loss, recall, and decision calibration, which highlighted the “RSF + Enet [alpha = 0.8]” model for its strong calibration and accuracy (Supplementary Fig. 1C). The optimal model was “RSF + Enet [alpha = 0.8],” employing RSF for feature selection (Fig. 7B) and an elastic network with alpha set to 0.8 for model construction (Fig. 7C), achieving the highest average C-index (0.679) across the three datasets (Fig. 7A). A prognostic ML model, “RSF + Enet [alpha = 0.8],” was developed to predict OS in metastatic MIBC patients, identifying “Chemotherapy” as the most influential factor in both the RSF and Enet [alpha = 0.8] algorithms (Fig. 7B, C). Using risk scores calculated by Enet[alpha = 0.8], we set the median risk score in the training cohort as a threshold to classify patients into high- and low-risk groups. Kaplan-Meier (K-M) curves validated the model's effectiveness in risk stratification, showing that low-risk patients had significantly longer OS than high-risk patients in the training, internal validation, and external validation cohorts (Fig. 7D). Additionally, ROC curves for 1-, 3-, and 5-year OS highlighted the model's strong specificity (Fig. 7E). Time-dependent ROC curves showed that the “RSF + Enet [alpha = 0.8]” model consistently outperformed others across most time points, demonstrating superior discrimination and predictive power over traditional clinical variables (Fig. 7F). Calibration curves (Fig. 7G) and DCA curves (Fig. 7H) further highlighted the model's accuracy and clinical value. Together, these evaluation metrics confirm that the “RSF + Enet [alpha = 0.8]” model provides a robust and superior tool for predicting OS in MIBC patients with metastasis, outperforming other ML models in reliability and predictive performance.

Variable	Univariable logistic analysis		Multivariate logistic analysis	
	OR (95%CI)	p.value	OR (95%CI)	p.value
Sex: Male	Reference			
Female	1.05 (0.96–1.15)	0.3	–	–
Age: <40	Reference			
40–59	1.99 (1.1–4.05)	0.037	2.35(1.19–5.12)	0.021
60–69	1.52 (0.84–3.1)	0.2	2.08(1.05–4.53)	0.048
70–79	1.28 (0.71–2.61)	0.448	1.82(0.92–3.97)	0.106
≥80	0.99 (0.55–2.03)	0.987	1.23(0.62–2.7)	0.575
Race: White	Reference			
Black	1.48 (1.28–1.7)	<0.001	1.12(0.95–1.31)	0.174
Asian and pacific islanders	1.05 (0.86–1.27)	0.611	1.02(0.83–1.26)	0.82
Other	/	/	/	/
Marital_Status: Married	Reference			
Unmarried	1.39 (1.26–1.53)	<0.001	1.02(0.91–1.15)	0.676
Widowed or divorced	0.97 (0.86–1.08)	0.568	1.08(0.95–1.23)	0.248
Unknown	0.92 (0.72–1.16)	0.475	0.96(0.73–1.23)	0.745
Household_Income: <\$70,000	Reference			
≥\$70,000	0.86 (0.79–0.94)	<0.001	0.95(0.87–1.05)	0.306
Household_Location: Rural				
Urban	1.04 (0.92–1.18)	0.509		
Tumor_Primary_Site: Bladder wall	Reference			
Bladder base	1.02 (0.9–1.16)	0.72	0.94(0.81–1.08)	0.357
Urachus/Dome of bladder	0.77 (0.63–0.94)	0.011	0.88(0.7–1.1)	0.266
Overlapping lesion	1.25 (1.13–1.39)	<0.001	1.13(1.01–1.27)	0.033
Histology: Transitional cell carcinoma	Reference			
Papillary transitional cell carcinoma	0.75 (0.68–0.82)	<0.001	0.81(0.73–0.9)	<0.001
Adenocarcinoma	1.06 (0.8–1.37)	0.681	1.48(1.07–2.02)	0.015
Squamous cell carcinoma	0.98 (0.79–1.21)	0.862	1.16(0.9–1.48)	0.254
Other	2.1 (1.79–2.46)	<0.001	1.76(1.47–2.11)	<0.001
Grade: Well differentiated I	Reference			
Moderately differentiated II	0.88 (0.55–1.47)	0.609	0.99(0.6–1.72)	0.976
Poorly differentiated III	1.55 (1.02–2.49)	0.054	1.61(1.01–2.7)	0.057
Undifferentiated anaplastic IV	1.73 (1.14–2.77)	0.016	1.7(1.07–2.85)	0.034
Laterality: Bilateral	Reference			
Lateral	0.65 (0.47–0.94)	0.017	0.69(0.47–1.03)	0.056
Surgery_Type: no surgery	Reference			
Local tumor destruction/excision	0.35 (0.3–0.42)	<0.001	0.53(0.44–0.65)	<0.001
Partial cystectomy	0.11 (0.08–0.16)	<0.001	0.25(0.17–0.36)	<0.001
Complete cystectomy	0.15 (0.12–0.18)	<0.001	0.19(0.15–0.25)	<0.001
T_Stage: T2	Reference			
T3	0.87 (0.77–0.97)	0.015	1.29(1.12–1.48)	<0.001
T4	2.54 (2.31–2.79)	<0.001	1.92(1.71–2.16)	<0.001
N_Stage: N0	Reference			
N1	4.21 (3.74–4.72)	<0.001	4.87(4.17–5.69)	<0.001
N2	5.95 (5.33–6.62)	<0.001	6.62(5.68–7.72)	<0.001
N3	6.42 (4.18–9.62)	<0.001	7.85(4.78–12.63)	<0.001
Tumor_Size: 0–2.0	Reference			
2.1–4.0	1.7 (1.35–2.16)	<0.001	1.4(1.1–1.8)	0.007
4.1–6.0	2.65 (2.11–3.35)	<0.001	1.76(1.39–2.26)	<0.001
>6.0	4.1 (3.27–5.2)	<0.001	2.27(1.78–2.92)	<0.001
Unknown	1.8 (1.46–2.25)	<0.001	1.46(1.16–1.84)	0.001
Lymph_Nodes_Surgery: None	Reference			
Regional lymph nodes removed	0.44 (0.4–0.49)	<0.001	0.54(0.39–0.74)	<0.001
Unknown	0.01 (0.01–0.03)	<0.001	1.24(0.42–3.21)	0.676
Lymph_Node_Biopsy: No biopsy	Reference			
Continued				

Variable	Univariable logistic analysis		Multivariate logistic analysis	
	OR (95%CI)	p.value	OR (95%CI)	p.value
Negative	0.13 (0.11–0.16)	< 0.001	0.51(0.37–0.7)	< 0.001
Positive	1.45 (1.3–1.61)	< 0.001	0.73(0.53–1)	0.048
Unknown	2.14 (1.15–3.72)	0.011	1.8(0.82–3.76)	0.127
Chemotherapy: none/unknown	Reference			
Yes	2.29 (2.11–2.49)	< 0.001	1.51(1.37–1.67)	< 0.001
Radiotherapy: none/unknown	Reference			
Yes	1.38 (1.25–1.52)	< 0.001	0.82(0.73–0.93)	0.001
Surgery_in_Other_Sites: No				
Yes	2.04 (1.77–2.33)	< 0.001	2.45(2.09–2.88)	< 0.001
Unknown	0 (0–0.01)	< 0.001	0(0–0.02)	< 0.001

Table 3. Univariate and multivariate logistics regression analyses of MIBC patients for distant metastasis in the training cohort.

Discussion

Muscle-invasive bladder cancer (MIBC) is a particularly aggressive form of bladder cancer, known for its ability to invade the detrusor muscle. Once this barrier is breached, the risk of metastasis increases significantly, leading to a poorer outlook for patients. Metastasis typically spreads through lymphatic and vascular pathways, often reaching organs like the lungs, liver, and bones. The development of metastatic disease is associated with sharply reduced survival rates, as treatment options become increasingly limited and less effective²⁷. Therefore, accurately predicting metastasis and identifying prognostic risk factors are crucial for managing MIBC. This study aimed to develop innovative ML models to predict metastasis at an early stage and to assess prognosis in MIBC patients with metastasis. By collecting clinical data on key variables and building ML models through a benchmark framework, we generated risk scores to aid in both metastasis prediction and prognosis estimation, supporting precise risk stratification. The model evaluates clinical and demographic factors to generate a risk score for metastasis and prognosis, assisting clinicians in decision-making and personalized treatment planning.

Previous research has identified various biomarkers, molecular subtypes, immune cell infiltration patterns, and protein expression levels as possible indicators of metastasis in MIBC. For example, combining CD8 + T cell infiltration with molecular subtype analysis has revealed that patients with the Basal-CD8 + T low subtype tend to have poorer OS and a higher risk of metastasis compared to other subtypes²⁸. Matrix metalloproteinase-14 (MMP14), associated with the basement membrane, has emerged as a critical marker linked to poor prognosis and heightened metastatic risk in MIBC. Elevated MMP14 expression correlates with aggressive tumor behavior and reduced treatment response²⁹. Additionally, recent studies using deep learning models have shown that lymphocytic inflammation within the tumor stroma may predict lymph node metastasis in MIBC, underscoring the significance of tumor microenvironment features in metastasis prediction³⁰. These findings imply that integrating molecular, immune, and histological biomarkers could improve the precision of metastasis predictions, assisting clinicians in creating more targeted treatment strategies.

Our research emphasizes the use of clinical variables for predicting metastasis, noting that older MIBC patients typically experience higher rates of metastasis and mortality compared to younger patients. This is partially due to their reduced likelihood of undergoing aggressive treatments like NAC and RC³¹. Additionally, the T stage, reflecting the degree of tumor invasion into the bladder wall and nearby tissues, plays a crucial role in influencing the risk of distant metastasis in MIBC³. As the T stage advances from T2 (invasion of the muscle layer) to T4 (invasion into nearby organs or tissues), the risk of metastasis rises. Higher T stages, particularly T3 and T4, are linked to an increased likelihood of cancer spreading to distant organs, including the liver, lungs, and bones. This underscores the importance of precise staging for prognosis and guiding treatment choices. The N stage, indicating regional lymph node involvement, is a critical factor to assess the risk of distant metastasis for MIBC patients. Those with lymph node metastasis (N1-N3) face a substantially higher chance of developing distant metastasis than patients without lymph node involvement (N0). This relationship between lymph node status and distant spread highlights the importance of precise lymph node staging to inform the selection of neoadjuvant or adjuvant therapies, ultimately aiming to improve patient outcomes³. These above three clinical variables, namely age, T stage and N stage, were not only significantly predictive for metastasis in our ML model, but also powerful in non-ML nomograms, confirming their huge impact on clinical decision-making³².

The choice of surgical treatment for MIBC has a major impact on the risk of metastasis and long-term survival outcomes. RC combined with PLND is considered the gold standard, but it is not a frequent combination during clinical practice. Research has indicated that RC could offer equivalent OS and CSS compared to bladder-preserving trimodal therapy (TMT) particularly in T2N0M0 disease¹⁰. Although short-term results (1–2 years) show minimal differences between TMT and RC, RC provides better outcomes at 5-year and 10-year follow-ups, especially in lowering the risk of metastasis^{33,34}. For patients unable to undergo RC due to health conditions or personal preference, bladder-preserving options like TMT remain feasible, although they may carry a greater risk of recurrence and metastasis^{35,36}. Tumor size is also a crucial factor in determining prognosis and the likelihood of metastasis in MIBC. Larger tumors are linked to higher metastasis risks and poorer OS. Research has shown that patients with larger tumors are often at more advanced disease stages, which correlates with an increased

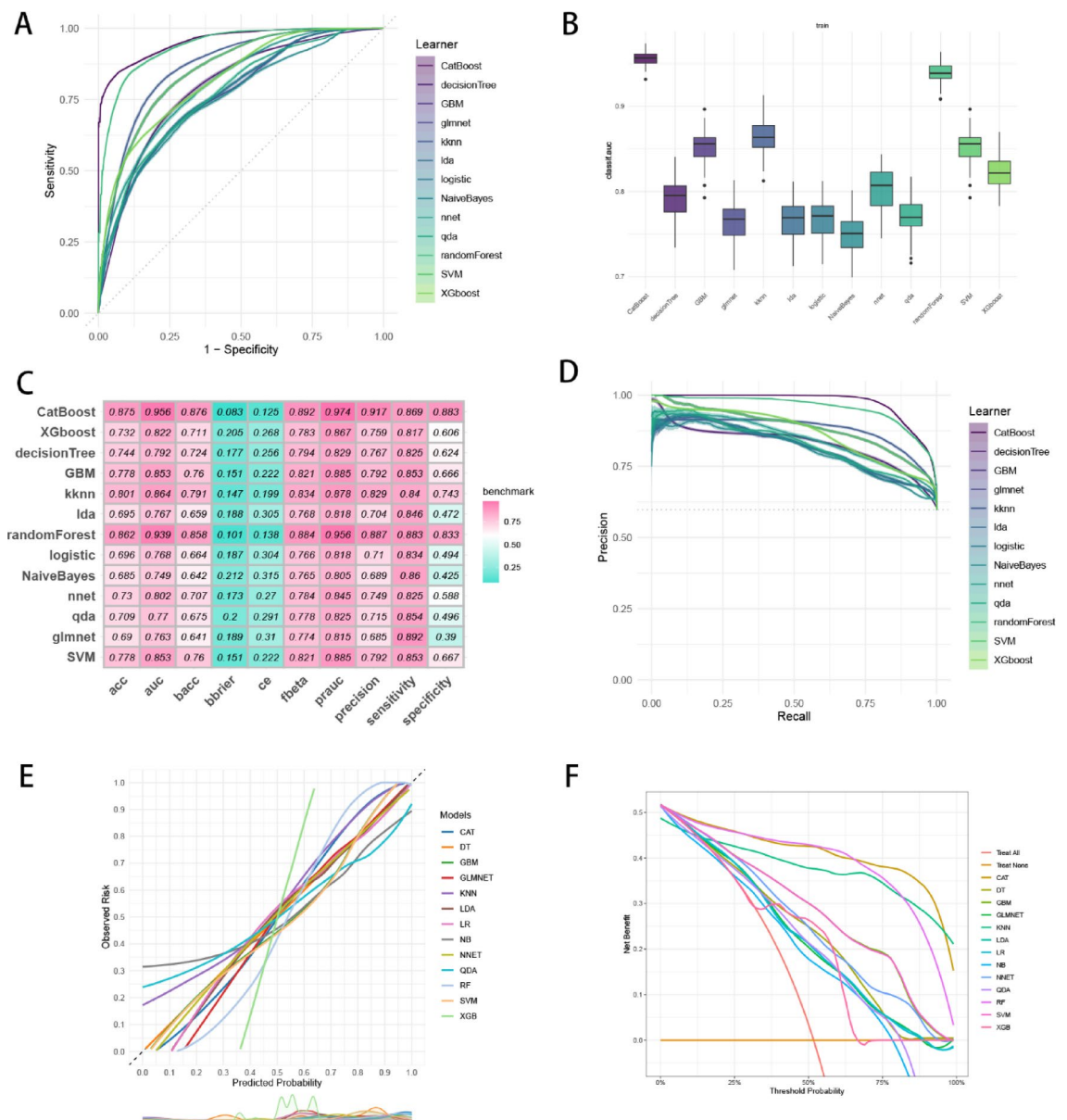


Fig. 3. Establishment and evaluation of the ML models in the training set. (A) ROC curves of different ML models in the training set. (B) AUC values of different ML models in the training set. (C) The performances of 13 ML models in terms of AUC, accuracy, sensitivity, specificity, precision, cross-entropy, Brier scores, Balanced Accuracy (bacc) and F1 Score in the training set. (D) PR curves of different ML models in the training set. (E) Calibration curves of different ML models in the training set. (F) DCA curves of different ML models in the training set. ML, machine learning; CAT, categorical boosting; LR, logistic regression; DT, decision tree; RF, random forest; XGB, extreme gradient boosting; GBM, gradient boosting machine; NB, Naive Bayes; LDA, linear discriminant analysis; QDA, quadratic discriminant analysis; NNET, neural network; GLMNET, generalized linear models with elastic net regularization; SVM, support vector machine; KNN, k-nearest neighbor.

chance of distant spread³⁷. Additionally, tumor size, along with factors like pathological response to treatment, is frequently used to estimate long-term survival and recurrence risks in patients undergoing RC or NAC³⁸.

The value of our study lies in its potential to enhance management and treatment strategies for patients with MIBC. By providing a dependable risk stratification tool, the model supports clinicians in making tailored and data-informed decisions for treatment. For example, patients deemed high-risk for metastasis could be prioritized for more intensive surgical options, supplemental therapies, and rigorous post-operative monitoring, which may contribute to improved survival rates. On the other hand, patients classified as low-risk could benefit from less aggressive treatments, reducing the likelihood of unnecessary side effects and complications linked to overtreatment. Additionally, the model's predictions can help tailor adjuvant therapy choices, set appropriate follow-up intervals, and determine when extra lab tests are needed. Incorporating this predictive tool into clinical

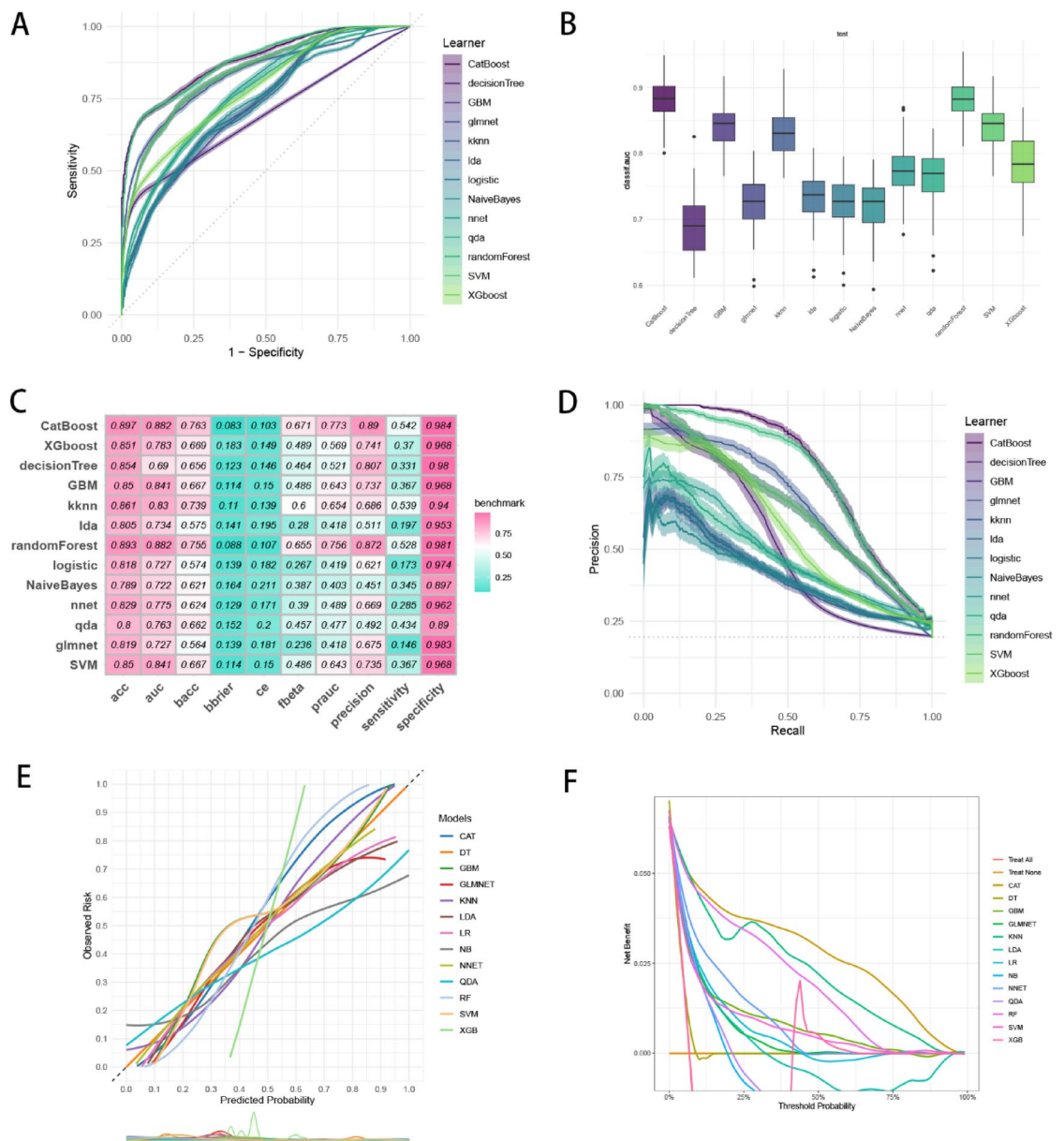


Fig. 4. Evaluation of the ML models in the internal validation set. (A) ROC curves of different ML models in the internal validation set. (B) AUC values of different ML models in the internal validation set. (C) The performance of 13 ML models in terms of AUC, accuracy, sensitivity, specificity, precision, cross-entropy, Brier scores, Balanced Accuracy (bacc) and F1 Score in the internal validation set. (D) PR curves of different ML models in the internal validation set. (E) Calibration curves of different ML models in the internal validation set. (F) DCA curves of different ML models in the internal validation set. ML, machine learning; CAT, categorical boosting; LR, logistic regression; DT, decision tree; RF, random forest; XGB, extreme gradient boosting; GBM, gradient boosting machine; NB, Naive Bayes; LDA, linear discriminant analysis; QDA, quadratic discriminant analysis; NNET, neural network; GLMNET, generalized linear models with elastic net regularization; SVM, support vector machine; KNN, k-nearest neighbor.

workflows is likely to enhance data-driven decision-making, improving both patient outcomes and resource efficiency. Furthermore, it can aid in standardizing care across providers and institutions, thereby minimizing variation in treatment practices and outcomes for MIBC patients.

Despite the innovative attempt of developing a ML model for predicting metastasis in our study, clinical utility depends on the interpretability and practicality of the selected variables should be noticed³⁹. One of the main strengths of our study is utilization of interpretable ML algorithms, particularly through SHAP values and the DALEX framework, to pinpoint crucial factors affecting metastasis predictions. The CatBoost algorithm, a gradient boosting method built on symmetric decision trees (oblivious trees), delivers high accuracy with minimal parameter tuning, excelling at handling categorical data⁴⁰. CatBoost's performance is on par with other advanced ML techniques, proving effective across various applications. However, the model's "black-box" nature

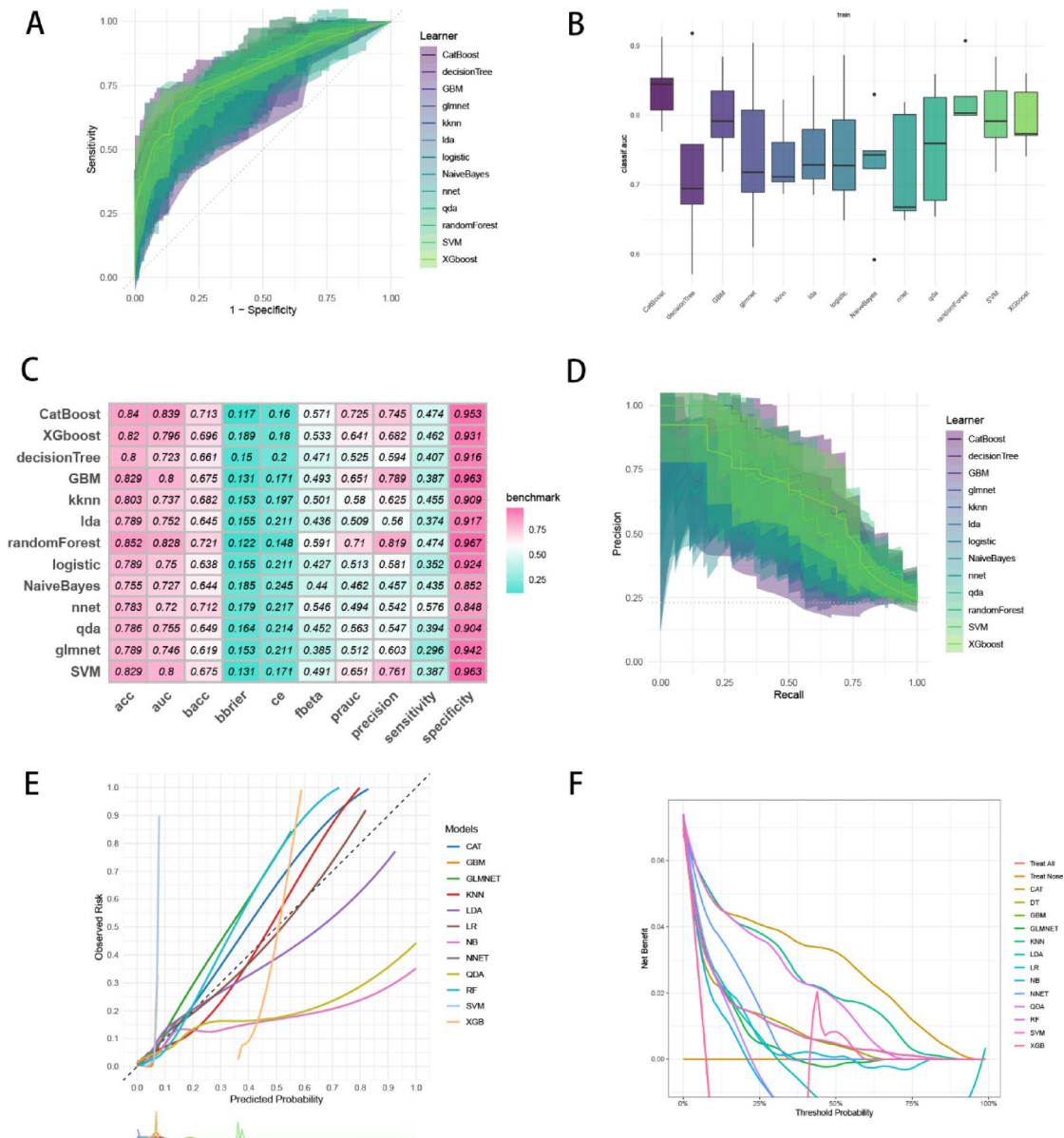


Fig. 5. Evaluation of the ML models in the external validation set. (A) ROC curves of different ML models in the external the external validation set. (E) Calibration curves of different ML models in the external validation set. (C) The performance of 13 ML models in terms of AUC, accuracy, sensitivity, specificity, precision, cross-entropy, Brier scores, Balanced Accuracy (bacc) and F1 Score in the external validation set. (D) PR curves of different ML models in the external validation set. (E) Calibration curves of different ML models in the external validation set. (F) DCA curves of different ML models in the external validation set. ML, machine learning; CAT, categorical boosting; LR, logistic regression; DT, decision tree; RF, random forest; XGB, extreme gradient boosting; GBM, gradient boosting machine; NB, Naive Bayes; LDA, linear discriminant analysis; QDA, quadratic discriminant analysis; NNET, neural network; GLMNET, generalized linear models with elastic net regularization; SVM, support vector machine; KNN, k-nearest neighbor.

calls for interpretability, which we provided using SHAP summary plots and force maps. These visualization tools give clinicians a straightforward way to understand the primary factors driving survival predictions, making the model's insights more accessible for informed clinical decisions. Furthermore, incorporating advanced techniques, such as RFE with cross-validation (RFEVCV), GridSearchCV for hyperparameter tuning, and SMOTE oversampling to address sample imbalance, has significantly boosted the model's accuracy in predicting metastasis. With this precise ML model, clinicians can create individualized treatment plans, allowing for timely adjustments in therapy that enhance diagnostic and prognostic outcomes for MIBC patients.

Although this study has several strengths, it also has some limitations. Firstly, approximately 8–9% of MIBC patients in SEER database developed metastasis, which is lower than anticipated based on the general survival trends. The primary reason for this discrepancy lies in the limitations of the retrospective database.

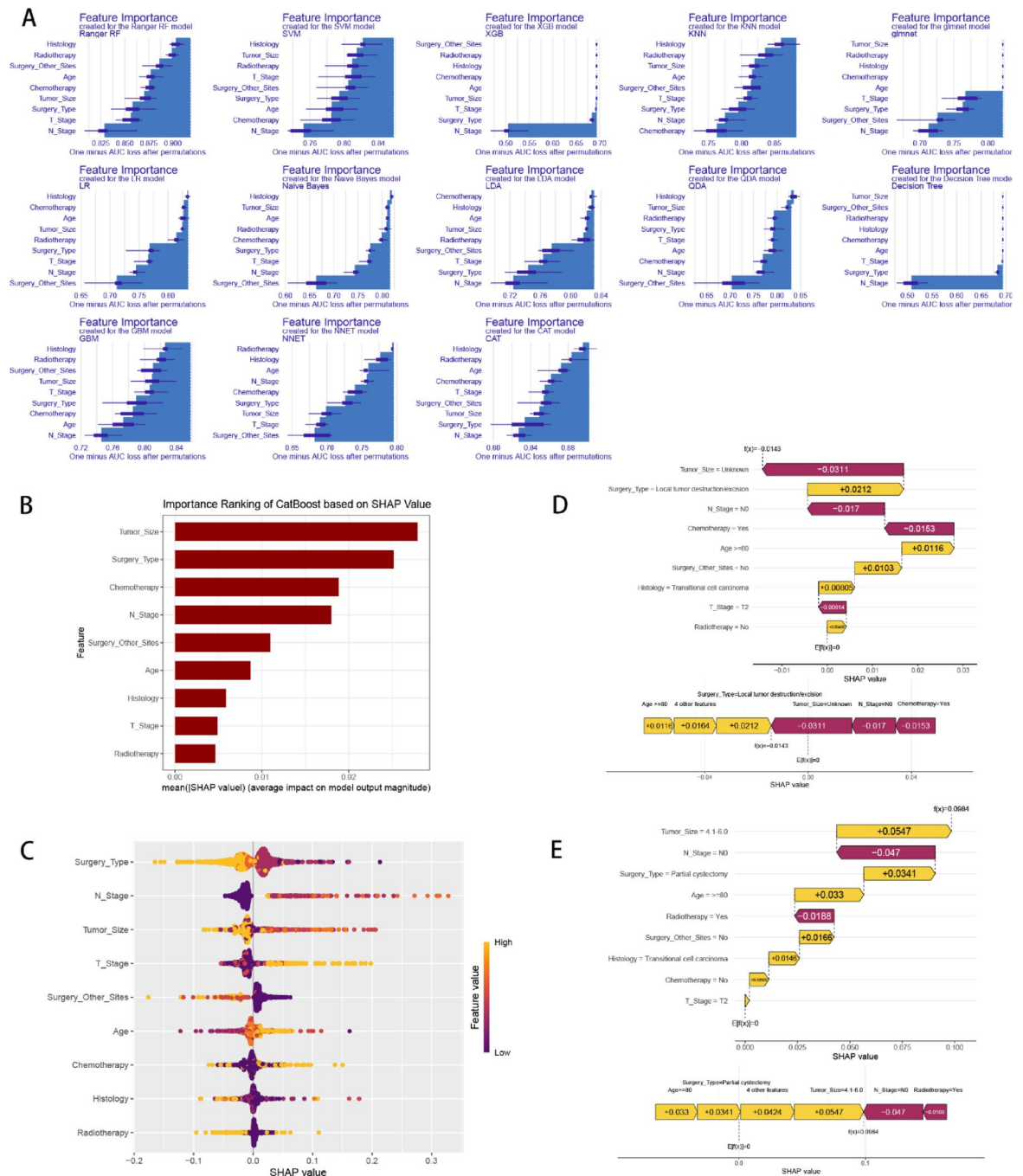


Fig. 6. ML model interpretation. (A) Importance ranking of features in 13 ML prediction algorithms. (B) The importance ranking of different variables according to the mean (|SHAP value|) in the optimal CatBoost model. (C) The importance ranking of different risk factors predicting metastasis with stability and interpretation in the optimal CatBoost model. The higher SHAP value of a feature is given, the higher risk of distant metastasis the patient would have. The yellow part in feature value represents higher value. (D) SHAP value explanation in a classical sample without distant metastasis. (E) SHAP value explanation in a classical sample with distant metastasis.

In retrospective studies, reporting of metastasis may be incomplete, particularly in patients who transition to palliative care. Although SEER database contains systematic follow-up data, metastatic screening and documentation were not consistently applied across all cases. Some patients may have been lost to follow-up or received treatment at other institutions, which may have led to an underrepresentation of metastatic cases. We acknowledge that this limitation may impact the generalizability of our results, particularly when assessing overall prognosis and metastasis-related risks. Future studies should aim to include a more representative sample of metastatic patients, or explore ways to improve metastatic screening and follow-up processes to minimize

Variable	Univariable cox analysis		Multivariate cox analysis	
Term	HR (95%CI)	p.value	HR (95%CI)	p.value
Sex: Male	Reference			
Female	1.11 (1.03–1.2)	0.006	1.026 (0.948–1.111)	0.525
Age: <40	Reference			
40–59	1.19 (0.73–1.96)	0.49	1.058 (0.641–1.745)	0.826
60–69	1.21 (0.74–1.98)	0.457	1.113 (0.674–1.838)	0.674
70–79	1.47 (0.9–2.41)	0.128	1.284 (0.778–2.12)	0.328
≥80	2.01 (1.23–3.3)	0.006	1.354 (0.817–2.243)	0.24
Race: White	Reference			
Black	1.07 (0.95–1.2)	0.25		
Asian and pacific islanders	0.9 (0.76–1.06)	0.207		
Other	/	/		
Marital_Status: Married	Reference			
Unmarried	1.11 (1.03–1.21)	0.008	1.069 (0.983–1.161)	0.117
Widowed or divorced	1.3 (1.18–1.43)	<0.001	0.971 (0.875–1.077)	0.576
Unknown	1.2 (0.98–1.46)	0.075	1.112 (0.911–1.357)	0.297
Household_Income: <\$70,000	Reference			
≥\$70,000	0.98 (0.91–1.05)	0.495		
Household_Location: Rural				
Urban	0.94 (0.85–1.03)	0.201		
Tumor_Primary_Site: Bladder wall	Reference			
Bladder base	1.06 (0.96–1.18)	0.26	1 (0.896–1.115)	0.997
Urachus/Dome of bladder	0.85 (0.72–0.99)	0.041	0.976 (0.826–1.152)	0.772
Overlapping lesion	1.04 (0.96–1.13)	0.351	1.067 (0.979–1.163)	0.141
Histology: Transitional cell carcinoma	Reference			
Papillary transitional cell carcinoma	0.82 (0.76–0.89)	<0.001	0.869 (0.801–0.943)	0.001
Adenocarcinoma	0.7 (0.57–0.86)	0.001	0.704 (0.565–0.877)	0.002
Squamous cell carcinoma	1.94 (1.63–2.3)	<0.001	1.906 (1.599–2.273)	<0.001
Other	1.19 (1.05–1.35)	0.007	1.091 (0.959–1.242)	0.186
Grade: Well differentiated I	Reference			
Moderately differentiated II	0.77 (0.53–1.12)	0.169		
Poorly differentiated III	0.85 (0.6–1.19)	0.346		
Undifferentiated anaplastic IV	0.76 (0.54–1.06)	0.104		
Laterality: Bilateral	Reference			
Lateral	1.4 (1.04–1.89)	0.026	1.172 (0.867–1.584)	0.302
Surgery_Type: No Surgery	Reference			
Local tumor destruction/excision	0.82 (0.72–0.93)	0.002	0.88 (0.772–1.002)	0.054
Partial cystectomy	0.7 (0.54–0.92)	0.01	1 (0.744–1.343)	0.999
Complete cystectomy	0.5 (0.44–0.58)	<0.001	0.637 (0.52–0.781)	<0.001
T_Stage: T2	Reference			
T3	0.76 (0.69–0.84)	<0.001	1.044 (0.937–1.163)	0.436
T4	1.05 (0.98–1.14)	0.174	1.162 (1.069–1.264)	<0.001
N_Stage: N0	Reference			
N1	0.92 (0.84–1.01)	0.067	1.047 (0.947–1.157)	0.373
N2	0.86 (0.8–0.94)	0.001	1.085 (0.987–1.192)	0.09
N3	1.34 (0.99–1.81)	0.059	1.508 (1.108–2.053)	0.009
Tumor_Size: 0–2.0	Reference			
2.1–4.0	1.15 (0.94–1.4)	0.177	1.179 (0.965–1.44)	0.108
4.1–6.0	1.39 (1.15–1.69)	0.001	1.278 (1.05–1.557)	0.015
>6.0	1.79 (1.47–2.18)	<0.001	1.62 (1.329–1.975)	<0.001
Unknown	1.72 (1.43–2.07)	<0.001	1.332 (1.102–1.61)	0.003
Lymph_Nodes_Surgery: None	Reference			
Regional lymph nodes removed	0.61 (0.56–0.66)	<0.001	0.764 (0.607–0.962)	0.022
Unknown	1.07 (0.61–1.88)	0.818	0.743 (0.311–1.773)	0.503
Lymph_Node_Biopsy: No biopsy	Reference			
Negative	0.61 (0.52–0.71)	<0.001	0.96 (0.755–1.22)	0.737
Continued				

Variable	Univariable cox analysis		Multivariate cox analysis	
Term	HR (95%CI)	p.value	HR (95%CI)	p.value
Positive	0.64 (0.58–0.7)	<0.001	1.096 (0.885–1.358)	0.399
Unknown	1.83 (1.12–2.99)	0.016	1.745 (0.829–3.674)	0.143
Chemotherapy: none/unknown	Reference			
Yes	0.43 (0.41–0.46)	<0.001	0.446 (0.413–0.48)	<0.001
Radiotherapy: none/unknown	Reference			
Yes	1.08 (1–1.17)	0.058		
Surgery_in_Other_Sites: No				
Yes	0.84 (0.75–0.93)	0.001	1.008 (0.899–1.131)	0.888
Unknown	0.42 (0.1–1.67)	0.216	0.568 (0.134–2.403)	0.442
Bone_Metastasis: No	Reference			
Yes	1.51 (1.37–1.66)	<0.001	1.41 (1.275–1.559)	<0.001
Unknown	1.25 (1.16–1.34)	<0.001	1.022 (0.644–1.622)	0.927
Brain_Metastasis: No	Reference			
Yes	1.67 (1.21–2.3)	0.002	1.735 (1.25–2.406)	0.001
Unknown	1.12 (1.05–1.2)	0.001	0.93 (0.485–1.786)	0.829
Liver_Metastasis: No	Reference			
Yes	1.62 (1.44–1.83)	<0.001	1.423 (1.256–1.612)	<0.001
Unknown	1.19 (1.11–1.28)	<0.001	1.124 (0.615–2.054)	0.704
Lung_Metastasis: No				
Yes	1.42 (1.29–1.57)	<0.001	1.219 (1.099–1.353)	<0.001
Unknown	1.22 (1.13–1.31)	<0.001	1.116 (0.772–1.612)	0.56

Table 4. Univariate and multivariate Cox regression analyses of MIBC patients with distant metastasis for overall survival in the training cohort.

such biases. Secondly, over 50% of MIBC patients in SEER database received only ‘Local tumor destruction/excision,’ with a subset also undergoing radiotherapy (17%). We acknowledge that this high proportion may influence overall dataset composition and outcomes, as these patients generally have a poor prognosis and shorter follow-up. However, we deliberately included these patients to ensure our study reflects real-world clinical practice, where certain patients, due to age, comorbidities, or treatment constraints, may receive limited interventions rather than standard radical treatment. Excluding them entirely could introduce selection bias and limit the generalizability of our findings. Nonetheless, we recognize that the inclusion of these patients may impact survival-related analyses and should be carefully considered when interpreting the results. Future studies may benefit from stratified analyses or alternative modeling approaches to account for the potential influence of treatment heterogeneity on outcomes.

Besides, while we determined the required sample size for external validation, patient availability with complete follow-up data was limited, resulting in a smaller validation cohort than ideal. Acknowledging that larger sample sizes provide more reliable evaluations, we gathered the maximum feasible number of samples given current constraints. To address this limitation, a 10-fold cross-validation approach was used and improved the model’s generalizability. Future studies will focus on enlarging the external validation cohort to further validate the model’s reliability. Moreover, differences in data collection methods across various hospitals in SEER database, along with the study’s retrospective design, led to some missing clinical variables. Additionally, essential clinicopathological details like imaging type and scope, comorbidities, inflammatory markers, genetic information, and occupational exposures were excluded due to their absence in the SEER database. Although we included a wide array of features to enhance predictive accuracy, this complexity might limit the model’s practicality in real-world clinical applications. Finally, the model has not yet been tested in clinical practice, highlighting the need for prospective, multicenter, large-scale validation studies to thoroughly evaluate its generalizability and suitability in real-world settings. Such future research will be essential for confirming the model’s effectiveness across varied patient groups and clinical contexts.

Conclusions

In this study, we identified key risk factors for predicting metastasis in MIBC patients and developed a CatBoost-based prediction model within an ML benchmark framework, outperforming both traditional clinical metrics and other ML models in accuracy and effectiveness. Additionally, we created an RSF + Enet [alpha = 0.8] prognostic model to reliably forecast outcomes for MIBC patients with metastasis. Our work provides a foundation for future advancements in metastasis prediction and prognosis assessment, aiding clinicians in making informed treatment decisions and personalizing therapy plans.

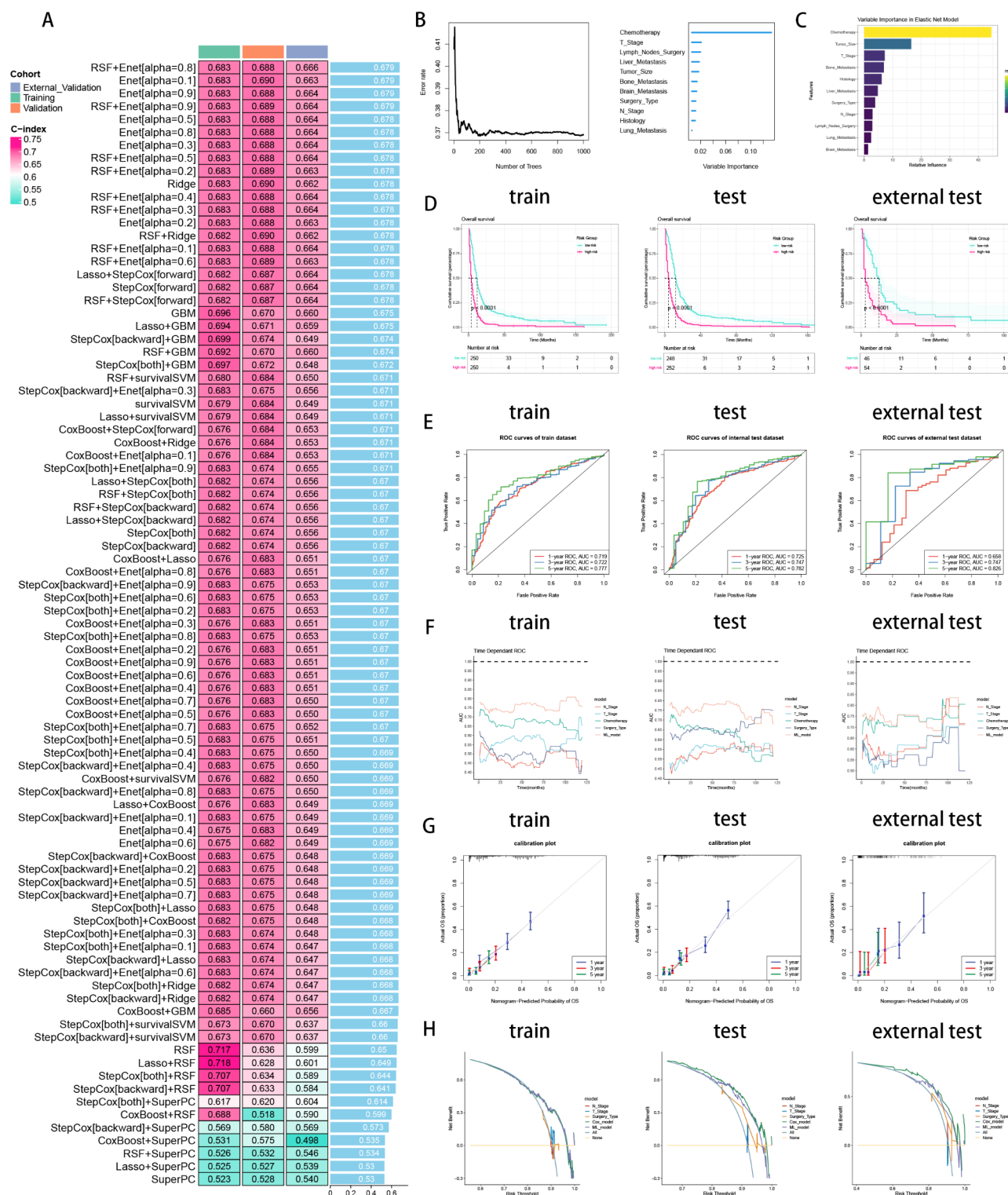


Fig. 7. Establishment and validation of prognostic model for MIBC patients with distant metastasis. (A) A total of 101 kinds of prognostic models via a leave-one-out cross-validation framework and further calculated the C-index of each model. (B) Feature selection process by RSF algorithm. (C) Model construction by Enet[alpha=0.8] algorithm and visualization of feature importance. (D) K-M curves of low-risk and high-risk groups divided by ML model in training, internal validation and external validation cohorts. (E) ROC curves of ML model in training, internal validation and external validation cohorts. (F) Time dependent AUC values of ML model in training, internal validation and external validation cohorts. (G) Calibration curves of ML model in training, internal validation and external validation cohorts. (H) DCA curves of ML model in training, internal validation and external validation cohorts. Left: training cohort, Middle: internal validation cohort, Right: external validation cohort.

Data availability

The original contributions presented in the study are included in the article or Supplementary Material. Further inquiries about the raw data and code can be directed to the corresponding authors.

Received: 4 November 2024; Accepted: 26 March 2025

Published online: 06 April 2025

References

- Antoni, S. et al. Bladder cancer incidence and mortality: a global overview and recent trends. *Eur. Urol.* **71**(1), 96–108 (2017).
- Knowles, M. A. & Hurst, C. D. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. *Nat. Rev. Cancer.* **15**(1), 25–41 (2015).
- Alfred Witjes, J. et al. Updated 2016 EAU guidelines on muscle-invasive and metastatic bladder cancer. *Eur. Urol.* **71**(3), 462–475 (2017).
- Stein, J. P. et al. Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *J. Clin. Oncol.* **41**(22), 3772–3781 (2023).
- Galsky, M. D. et al. Real-world effectiveness of chemotherapy in elderly patients with metastatic bladder cancer in the United States. *Bladder Cancer.* **4**(2), 227–238 (2018).
- Cumberbatch, M. G. K. et al. Epidemiology of bladder cancer: a systematic review and contemporary update of risk factors in 2018. *Eur. Urol.* **74**(6), 784–795 (2018).
- Smith, A. B. et al. Muscle-invasive bladder cancer: evaluating treatment and survival in the National Cancer Data Base. *BJU Int.* **114**(5), 719–726 (2014).
- Kang, N. W. et al. Comparative survival analysis of bladder preservation therapy versus radical cystectomy in muscle-invasive bladder cancer. *Cancer Med.* **13**(2), e6972 (2024).
- Chen, X. et al. Urine DNA methylation assay enables early detection and recurrence monitoring for bladder cancer. *J. Clin. Invest.* **130**(12), 6278–6289 (2020).
- Zlotta, A. R. et al. Radical cystectomy versus trimodality therapy for muscle-invasive bladder cancer: a multi-institutional propensity score matched and weighted analysis. *Lancet Oncol.* **24**(6), 669–681 (2023).
- Jiang, D. M., Chung, P., Kulkarni, G. S. & Sridhar, S. S. Trimodality therapy for muscle-invasive bladder cancer: recent advances and unanswered questions. *Curr. Oncol. Rep.* **22**(2), 14 (2020).
- Collins, G. S. et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open.* **11**(7), e048008 (2021).
- Wen, P., Wen, J., Huang, X. & Wang, F. Development and validation of nomograms predicting the 5- and 8-year overall and cancer-specific survival of bladder cancer patients based on SEER program. *J. Clin. Med.* **12**(4), (2023).
- Fang, W., Yang, Z. Y., Chen, T. Y., Shen, X. F. & Zhang, C. Ethnicity and survival in bladder cancer: a population-based study based on the SEER database. *J. Transl. Med.* **18**(1), 145 (2020).
- Karakiewicz, P. I. et al. Precystectomy nomogram for prediction of advanced bladder cancer stage. *Eur. Urol.* **50**(6), 1254–60; discussion 1261–2. (2006).
- Li, S. et al. Establishment and validation of nomograms to predict the overall survival and cancer-specific survival for non-metastatic bladder cancer patients: A large population-based cohort study and external validation. *Med. (Baltim).* **103**(11), e37492 (2024).
- Sylvester, R. J. et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. *Eur. Urol.* **49**(3), 466–5; discussion 475–7. (2006).
- Riley, R. D. et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Stat. Med.* **40**(19), 4230–4251 (2021).
- Li, S., Wang, J., Zhang, Z., Ren, C. & He, D. Individual risk and prognostic value prediction by interpretable machine learning for distant metastasis in neuroblastoma: A population-based study and an external validation. *Int. J. Med. Inform.* 105813. (2025).
- Guyon, I., Weston, J., Barnhill, S. & Vapnik, V. Gene selection for cancer classification using support vector machines. *Mach. Learn.* **46**(1), 389–422 (2002).
- Kolde, R., Laur, S., Adler, P. & Vilo, J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* **28**(4), 573–580 (2012).
- Lang, M. et al. mlr3: A modern object-oriented machine learning framework in R. *J. Open. Source Softw.* **4**, p1903 (2019).
- Wang, K. et al. Improving risk identification of adverse outcomes in chronic heart failure using SMOTE + ENN and machine learning. *Risk Manag. Healthc. Policy.* **14**, 2453–2463 (2021).
- Biecek, P. Dalex: explainers for complex predictive models in R. *J. Mach. Learn. Res.* **19**, (2018).
- Lo, Y. T., Liao, J. C., Chen, M. H., Chang, C. M. & Li, C. T. Predictive modeling for 14-day unplanned hospital readmission risk by using machine learning algorithms. *BMC Med. Inf. Decis. Mak.* **21**(1), 288 (2021).
- Sonabend, R., Király, F. J., Bender, A., Bischl, B. & Lang, M. mlr3proba: An R package for machine learning in survival analysis. *Bioinformatics.* **37**(17), 2789–2791. (2021).
- Soukup, V. et al. Follow-up after surgical treatment of bladder cancer: a critical analysis of the literature. *Eur. Urol.* **62**(2), 290–302 (2012).
- Li, B. et al. Integrating molecular subtype and CD8(+) T cells infiltration to predict treatment response and survival in muscle-invasive bladder cancer. *Cancer Immunol. Immunother.* **73**(4), 66 (2024).
- Zhang, X. et al. Basement membrane-related MMP14 predicts poor prognosis and response to immunotherapy in bladder cancer. *BMC Cancer.* **24**(1), 746 (2024).
- Zheng, Q. et al. Predicting lymph node metastasis status from primary muscle-invasive bladder cancer histology slides using deep learning: a retrospective multicenter study. *Cancers.* **15**(11), (2023).
- Posielski, N., Koenig, H., Ho, O., Porter, C. & Flores, J. P. Use of neoadjuvant chemotherapy in elderly patients with muscle-invasive bladder cancer: a population-based study, 2006–2017. *Oncology* **36**(1), 21–33 (2022).
- Shi, S., Peng, G., Luo, L. & Li, D. Predictive nomograms for risk and prognostic factors in metastatic bladder cancer: a population-based study. *Transl. Cancer Res.* **12**(12), 3284–3302 (2023).
- Matsukawa, A. et al. Trimodality therapy versus radical cystectomy for muscle-invasive bladder cancer: a systematic review and meta-analysis of matched cohort studies. *Eur. Urol. Focus* (2024).
- Qiu, J. et al. Comparing long-term survival outcomes for muscle-invasive bladder cancer patients who underwent with radical cystectomy and bladder-sparing trimodality therapy: a multicentre cohort analysis. *J. Oncol.* **2022**, 7306198. (2022).
- Fong, K. Y. et al. Trimodality therapy versus radical cystectomy for muscle-invasive bladder cancer: A systematic review and meta-analysis. *Urol. Oncol.* (2025).
- Su, X., Dong, C., Liao, W. & Liu, W. Oncological effectiveness of bladder-preserving trimodal therapy versus radical cystectomy for the treatment of muscle-invasive bladder cancer: a system review and meta-analysis. *World J. Surg. Oncol.* **21**(1), 271 (2023).
- Yin, M. et al. Neoadjuvant chemotherapy for muscle-invasive bladder cancer: a systematic review and two-step meta-analysis. *Oncologist* **21**(6), 708–715 (2016).

38. van Ginkel, N. et al. Survival outcomes of patients with muscle-invasive bladder cancer according to pathological response at radical cystectomy with or without neo-adjuvant chemotherapy: a case-control matching study. *Int. Urol. Nephrol.* **54**(12), 3145–3152 (2022).
39. Yan, M. et al. Comparative analysis of nomogram and machine learning models for predicting axillary lymph node metastasis in early-stage breast cancer: a study on clinically and ultrasound-negative axillary cases across two centers. *Ultrasound Med. Biol.* **51**(3), 463–474 (2025).
40. Zhang, C. et al. Using catboost algorithm to identify middle-aged and elderly depression, National health and nutrition examination survey 2011–2018. *Psychiatry Res.* **306**, 114261 (2021).

Acknowledgements

We are grateful to the Surveillance, Epidemiology, and End Results (SEER) database for providing data. We especially thank Ting Wei and Chunqian Yang from Zhujiang Hospital of Southern Medical University for their generous help in methodology.

Author contributions

QD and SL: Collection and assembly of data, Investigation, Visualization, Validation, Data analysis and interpretation, Data curation, Methodology, Writing original manuscript. YZ: Collection and assembly of data. YJ and YY: Conception and design, Administration support, Manuscript edition, Funding acquisition. All authors contributed to this article and approved the submitted version.

Declarations

Ethics statement

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of The First Affiliated Hospital of Henan University of Science and Technology (protocol code: 2024-03-K0144, approved on 2024-03-20).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-96089-1>.

Correspondence and requests for materials should be addressed to Y.J. or Y.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025