

## RESEARCH ARTICLE

# Machine learning to improve prognosis prediction of metastatic clear-cell renal cell carcinoma treated with cytoreductive nephrectomy and systemic therapy

Wenjie Yang , Lin Ma , Jie Dong , Mengchao Wei , Ruoyu Ji , Hualin Chen , Xiaoqiang Xue , Yingjie Li , Zhaoheng Jin , Weifeng Xu , and Zhigang Ji 

Cytoreductive nephrectomy (CN) combined with systemic therapy is commonly used to treat metastatic clear-cell renal cell carcinoma (mccRCC). However, prognostic models for these patients are limited. In the present study, the clinical data of 782 mccRCC patients who received both CN and systemic therapy were obtained from the Surveillance, Epidemiology, and End Results (SEER) database (2010–2016), and patients were divided into training and internal test cohorts. A total of 144 patients who met the same criteria from our center (Peking Union Medical College Hospital) were placed in the external test cohort. The cancer-specific survival rate (CSS) at 1, 3, and 5 years was set as the research outcome. Then, four ML models, i.e., a gradient boosting machine (GBM), support vector machine (SVM), random forest (RF), and logistic regression (LR), were established. Fifteen potential independent features were included in this study. Model performance was evaluated using the area under the receiver operating characteristic curves (AUC), calibration plots, and decision curve analysis (DCA). Seven clinical features, namely, pathological grade, T stage, N stage, number of metastatic sites, brain or liver metastases, and metastasectomy, were selected for subsequent analysis via the recursive feature elimination (RFE) algorithm. In conclusion, the GBM model performed best at 1-, 3- and 5-year CSS prediction (0.836, 0.819, and 0.808, respectively, in the internal test cohort and 0.819, 0.805, and 0.786, respectively, in the external cohort). Furthermore, we divided the patients into three strata (high-, intermediate-, and low-risk) via X-tile analysis and concluded that clinically individualized treatment can be aided by these practical prognostic models.

**Keywords:** Metastatic clear-cell renal cell carcinoma (mccRCC), cytoreductive nephrectomy (CN), systemic therapy, machine learning (ML), prognosis prediction model, Surveillance, Epidemiology, and End Results (SEER) database.

## Introduction

Renal cell carcinoma (RCC) is one of the most common genitourinary cancers, with an increasing incidence and morbidity rate worldwide. Clear-cell RCC (ccRCC) remains the most prevalent histological subtype of renal cancer (accounting for 80%–85%). Although an ever-growing number of renal cancer patients can be detected at an early stage, 25%–30% of patients have metastasized at the time of diagnosis, and over 20% of patients will develop metastases after curative surgery, which contributes to a poor prognosis in patients [1–3].

Standard treatment strategies for metastatic ccRCC (mccRCC) have progressed substantially over the past decades. The majority of studies have shown that cytoreductive nephrectomy (CN), also termed radical nephrectomy for primary lesions, provides a significant survival benefit for patients with mccRCC. Theoretically, removing the primary lesion can effectively reduce the tumor burden and create favorable conditions for subsequent systemic therapy [4–7]. In an era where cytokine

therapy has been the only systemic therapy option for patients with mccRCC, a combined treatment regimen of CN and interferon-based adjunctive therapy has shown to be more effective than interferon therapy alone [8, 9]. In the targeted therapy age, several rigorous randomized clinical trials have also demonstrated the promising effectiveness of CN combined with systemic therapy in patients with mccRCC. Owing to the spectacular advantages in relieving local symptoms, such as chronic pain and hematuria, a substantial number of mccRCC patients have received both CN and systemic therapy in the real world [10, 11].

Building a prognostic prediction model is an effective method for identifying the patients who will benefit the most from these treatments. However, a model for predicting the survival of mccRCC patients treated with CN and systemic therapy is still lacking.

As an important subfield of artificial intelligence, machine learning (ML) algorithms involve multiple disciplines and

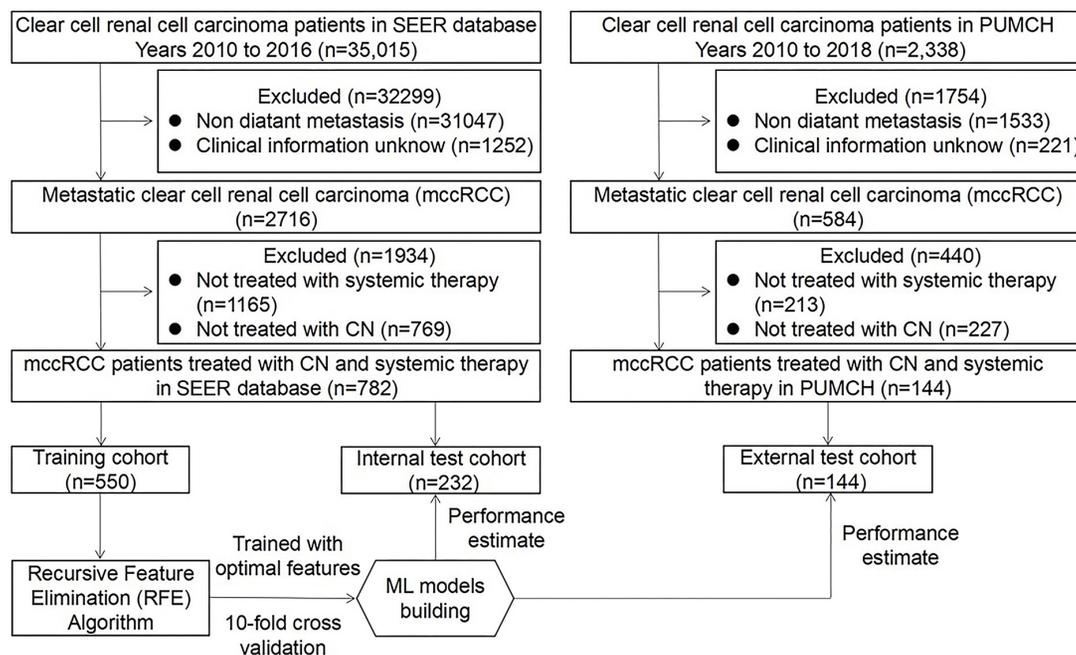
Department of Urology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Dongcheng, Beijing, China.

\*Correspondence to Weifeng Xu: [weifengxu2008@sina.com](mailto:weifengxu2008@sina.com); Zhigang Ji: [jzgjmc@163.com](mailto:jzgjmc@163.com)

\*Wenjie Yang and Lin Ma contributed equally to this work.

DOI: 10.17305/bjbm.2022.8047

© 2022 Yang et al. This article is available under a Creative Commons License (Attribution 4.0 International, as described at <https://creativecommons.org/licenses/by/4.0/>).



**Figure 1. Strategies for study population selection and establishment of predictive model.** mccRCC: Metastatic clear-cell renal cell carcinoma; ML: Machine learning; CN: Cytoreductive nephrectomy; PUMCH: Peking Union Medical College Hospital.

demonstrate advancement compared with traditional tools. Mounting evidence has revealed that ML models can provide a more accurate prognosis prediction by comprehensively integrating and analyzing the complex connections between clinical features and outcomes [12–14]. For example, one of the most prevalent ML algorithms, the gradient boosting machine (GBM), has shown excellent performance in terms of speed and accuracy in both classification and regression models [15]. Nevertheless, the benefit of an ML model in predicting the prognosis of patients with mccRCC receiving both CN and systemic therapy has yet to be fully explored.

In this study, we developed several clinical ML models to predict cancer-specific survival (CSS) for patients in this cohort according to data available from the Surveillance, Epidemiology and End Results (SEER) database. Although the specific drug regimens were not available in the SEER database, considering the widespread use of targeted drugs since 2005, systemic therapy in this study mainly refers to angiogenesis therapies and mammalian rapamycin (mTOR) therapies [16]. Furthermore, we used an external test cohort from the Peking Union Medical College Hospital (PUMCH) to test the validity of the models we developed.

## Materials and methods

### Study population

Data concerning mccRCC patients treated with CN combined with systemic therapy were retrospectively extracted from two sources: (I) the SEER database between 2010 and 2016 (<https://seer.cancer.gov/>) and (II) the PUMCH medical records between 2008 and 2018. The inclusion criteria were as follows: (I) age  $\geq 18$  years; (II) confirmation of ccRCC by histology; (III) distant

metastases based on the 8th American Joint Committee on Cancer staging systems; (IV) CN treatment; and (V) a history of receiving systemic therapy.

The exclusion criteria were as follows: (I) incomplete information, including unknown age, sex, race, laterality, pathological grade, T stage, N stage, tumor size, metastatic organ sites (bone, brain, liver, and lung), number of metastatic sites, metastasectomy, or radiotherapy; (II) diagnosis of malignant tumors other than ccRCC. The study population selection process is illustrated in Figure 1.

Research involving human participants was reviewed and approved by Peking Union Medical College's Ethical Committee and Institutional Review Board. All patients have signed an informed consent declaration.

### Outcome and data collection

Considering the relatively high mortality rates of mccRCC patients, we selected CSS as the primary endpoint in this study to avoid discrepancies in deaths. CSS is the interval between the date of treatment and the date of death caused by the tumor. Deaths caused by any factors unrelated to cancer or intervention were identified as non-cancer-specific and censored at the date of death. Patients in the external test (PUMCH) cohort underwent regular physical examinations, laboratory tests, urological ultrasound scans, bone scintigraphy, enhanced computerized tomography (CT) or magnetic resonance imaging every six months. The follow-up was terminated on April 30, 2021.

### ML model establishment and performance evaluation

We employed a recursive feature elimination (RFE) algorithm to select important features for model building. In short,

RFE starts with a model that covers all features and gradually removes the features that have the least impact on model performance until the retained features exceed a set performance threshold. The feature subset with the highest accuracy is then selected as the optimal feature combination. To determine the optimal hyperparameters, all the models were trained using 10-fold cross-validation in the training cohort. The data were split into ten parts: one part was assigned to the validation cohort and the other nine parts were used for training. The cross-validation process was repeated ten times, with each part validated once, and the average accuracy of the ten validations gave the final accuracy. Then, four ML algorithms, i.e., the GBM, support vector machine (SVM), random forest (RF), and logistic regression (LR) algorithms, were employed to construct ML models. Receiver operating characteristic (ROC) curves were employed to estimate the ML model accuracy by calculating the area under the curve (AUC). The ROC curve is calculated for all possible cut points (thresholds) and shows the correlation between sensitivity and specificity, thus providing a dynamic and objective response to the model's performance. Higher AUC values indicated better accuracy of the predictive model. The model fit was evaluated using calibration plots. The decision curve analysis (DCA) method was used to visualize the net benefits and usefulness of the prediction models.

### Statistical analysis

To estimate the differences between groups, categorical variables were expressed as numbers and percentages, and comparisons were made using Chi-square tests or Fisher's exact tests. The patients were divided into three risk groups: low, intermediate, and high, using the X-tile software (version 3.6.1). All statistical analyses in this study were performed using R software. Statistical significance was defined as  $P < 0.05$ .

## Results

### Baseline characteristics

This study examined clinical data from 782 patients with mCCRCC treated with CN combined with systemic therapy from the SEER database. Overall, 70% ( $n = 550$ ) of the patients were randomly divided into the training cohorts, while the remaining 30% ( $n = 232$ ) were assigned to the internal test cohort. The median follow-up was 25 (17–31) months in the SEER database. A total of 144 patients were included in the external test (PUMCH) cohort, and the median follow-up for this cohort was 37 (24–52) months. In the SEER cohort, 567 patients (72.5%) had single organ metastases (lung, bone, brain, or liver), 184 (23.5%) had double organ metastases, and 31 (4.0%) had three or more organ metastases. In the PUMCH cohort, 110 patients (76.4%) had single organ metastases, 30 (20.8%) had double organ metastases, and 4 (2.8%) had three or more organ metastases. The lungs were the most common site of metastases at a total of 589 SEER cases (75.3%) and 119 PUMCH cases (82.6%). This was followed by 283 SEER (36.2%) and 43 PUMCH cases (29.9%) of bone metastases, 83 SEER (10.6%) and 13 PUMCH (9.0%) cases of liver metastases, and 74 SEER (9.5%)

and 8 PUMCH (5.5%) cases of brain metastases. In addition, 154 SEER (19.7%) and 35 PUMCH (24.3%) patients underwent metastasectomy. Other characteristics of the clinical population and demographics are summarized in [Table 1](#).

### Feature selection

Incorporating redundant features may degrade the performance of an ML model [17]. The RFE algorithm was employed as a feature selection method to identify the optimal feature subset among all features. After RFE screening, seven important features, namely, pathological grade, T stage, N stage, the number of metastatic sites, brain or liver metastases, and metastasectomy, were determined. These features were then included in all our ML models in both the training and testing cohorts ([Figure S1](#)).

### ML models accurately predicted patient prognosis

We established prognosis prediction models for mCCRCC patients treated with CN and systemic therapy using four ML algorithms (GBM, SVM, RF, and LR). To evaluate the discriminatory abilities of these models, ROC curves for 1-, 3-, and 5-year CSS were constructed. In the training cohort, the AUC values of these ML models for the prediction of 1-, 3-, and 5-year CSS were 0.878, 0.832, and 0.828 for GBM; 0.860, 0.803, and 0.816 for SVM; 0.854, 0.810, and 0.806 for RF; and 0.848, 0.792, and 0.796 for LR, respectively. Compared with the other three ML models, the GBM model had the highest accuracy in predicting 1-, 3-, and 5-year CSS. Similar results were found in the internal and external test cohorts ([Figure 2A–2C](#)). In the internal test cohort, the AUC values of the GBM model for the prediction of 1-, 3-, and 5-year CSS were 0.836, 0.819, and 0.808, respectively ([Figure 2D–2F](#)). In the external test cohort, the AUC values of the GBM model for the prediction of 1-, 3-, and 5-year CSS were 0.819, 0.805, and 0.786, respectively ([Figure 2G–2I](#)). In addition, calibration curves for the four ML models are demonstrated in [Figure 3](#). Compared to the other three ML models, the calibration curves of the GBM model were the closest to the ideal lines.

### Clinical value of the ML models

DCA is a novel method for visualizing whether the use of a prediction model in clinical practice will benefit decision-making. The percentage of threshold probability is displayed on the X-axis, and the net benefit is indicated on the Y-axis [18, 19]. In our study, we hypothesized that patients whose predicted probability exceeds a set threshold would benefit from CN combined with systemic therapy. DCA indicated that all ML models achieved a net benefit. The DCA of the GBM has higher net benefits in the majority of the cohort subgroups, indicating that it had better clinical outcome values ([Figure 4](#)). Finally, the varying importance of the features for predicting CSS in each ML model is shown in [Figure S2](#).

### Risk stratification

As GBM was the optimal ML model based on the performance evaluation above, we set two optimal cut-off values (−5.4 and −4.8) depending on the GBM prediction score and divided patients into high-, intermediate-, and low-risk groups via

Table 1. Demographics and clinicopathologic features of the training, internal, and external validation

Characteristic		Training cohort (n = 550)	Internal validation cohort (n = 232)	External validation (PUMCH) cohort (n = 144)	P value
Age (years)	<65	362 (65.8%)	160 (69.0%)	86 (59.7%)	0.184
	≥65	188 (34.2%)	72 (31.0%)	58 (40.3%)	
Sex	Male	397 (72.2%)	158 (68.1%)	105 (72.9%)	0.461
	Female	153 (27.8%)	74 (31.9%)	39 (27.1%)	
Race	White	481 (87.5%)	196 (84.5%)	-	<0.001
	Black	27 (4.9%)	14 (6.0%)	-	
	Other	42 (7.6%)	22 (9.5%)	144 (100%)	
Laterality	Left	292 (53.1%)	110 (47.4%)	80 (55.6%)	0.229
	Right	258 (46.9%)	122 (52.6%)	64 (44.4%)	
T stage	T1	47 (8.5%)	25 (10.8%)	9 (6.2%)	0.255
	T2	80 (14.5)	22 (9.5)	15 (10.4)	
	T3	370 (67.3)	158 (68.1)	107 (74.3)	
	T4	53 (9.6%)	27 (11.6%)	13 (9.0%)	
Tumor size (cm)	≤4	22 (4.0%)	6 (2.6%)	1 (0.7%)	0.404
	4~7	108 (19.6%)	52 (22.4%)	33 (22.9%)	
	7~10	188 (34.2%)	85 (36.6%)	49 (34.0%)	
	>10	232 (42.2%)	89 (38.4%)	61 (42.4%)	
N stage	N0	395 (71.8%)	185 (79.7%)	107 (74.3%)	0.069
	N1	155 (28.2%)	47 (20.3%)	37 (25.7%)	
Mets.Lung	No	126 (22.9%)	67 (28.9%)	25 (17.4%)	0.033
	Yes	424 (77.1%)	165 (71.1%)	119 (82.6%)	
Mets.Bone	No	365 (66.4%)	134 (57.8%)	101 (70.1%)	0.024
	Yes	185 (33.6%)	98 (42.2%)	43 (29.9%)	
Mets.Liver	No	500 (90.9%)	199 (85.8%)	131 (91.0%)	0.084
	Yes	50 (9.1%)	33 (14.2%)	13 (9.0%)	
Mets.Brain	No	495 (90.0%)	213 (91.8%)	134 (94.5%)	0.242
	Yes	55 (10.0%)	19 (8.2%)	8 (5.5%)	
Metastatic surgery	Not Performed	445 (80.9%)	183 (78.9%)	109 (75.7%)	0.803
	Performed	105 (19.1%)	49 (21.1%)	35 (24.3%)	
Multiple organ metastasis	Single	407 (74.0%)	160 (69.0%)	110 (76.4%)	0.653
	Two	123 (22.4%)	61 (26.3%)	30 (20.8%)	
	Three-Four	20 (3.6%)	11 (4.7%)	4 (2.8%)	
Histological grade	I-II	91 (16.5%)	38 (16.4%)	24 (16.7%)	0.997
	III-IV	459 (83.5%)	194 (83.6%)	120 (83.3%)	
Radiotherapy	No	376 (68.4%)	150 (64.7%)	105 (72.9%)	0.244
	Yes	174 (31.6%)	82 (35.3%)	39 (27.1%)	

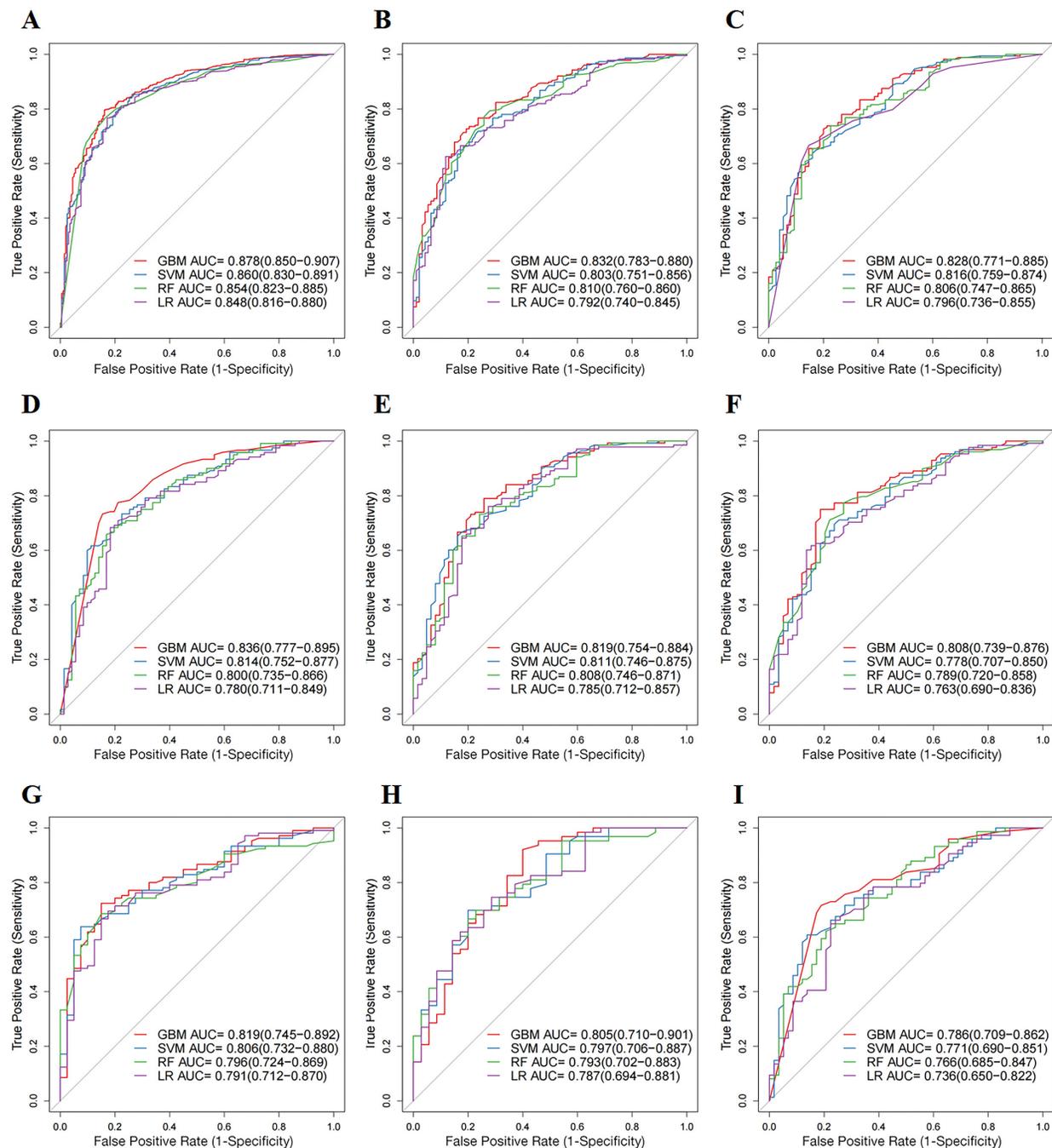
PUMCH: Peking Union Medical College Hospital; Mets: Metastasis organ site.

X-tile analysis (Figure S3). In the training cohort, the long-term survival (5-year CSS) of each group was as follows: high-risk, 2.3%; intermediate-risk, 18.7%; and low-risk, 23.7%. We then verified the actual performance of the aforementioned three risk strata in the internal and external test cohorts. In the internal test cohort, the 5-year CSS was 4.0% in the high-risk group, 15.1% in the intermediate-risk group, and 26.2% in the low-risk group. In the PUMCH cohort, the 5-year CSS was 5.9% in the high-risk group, 21.2% in the intermediate-risk group, and 22.8% in the low-risk group (Table 2). According to these results, the ML model provided excellent prognostic stratification in the training, internal test, and external test cohorts. The survival curves of patients with different risk

stratifications also showed that the GBM model provided excellent prognostic stratification (Figure 5A–5C).

## Discussion

At present, positive associations between cytoreductive resection of primary tumors and superior overall survival rates have been demonstrated in a variety of solid metastatic tumors, such as advanced-stage endometrial cancer, cohesive gastric cancer, and ovarian cancer [20–22]. As for mcrRCC, a significant survival benefit associated with CN and systemic therapy was found in a study by Chakiryan et al., who constructed an analysis containing 5005 mcrRCC patients using the National Cancer



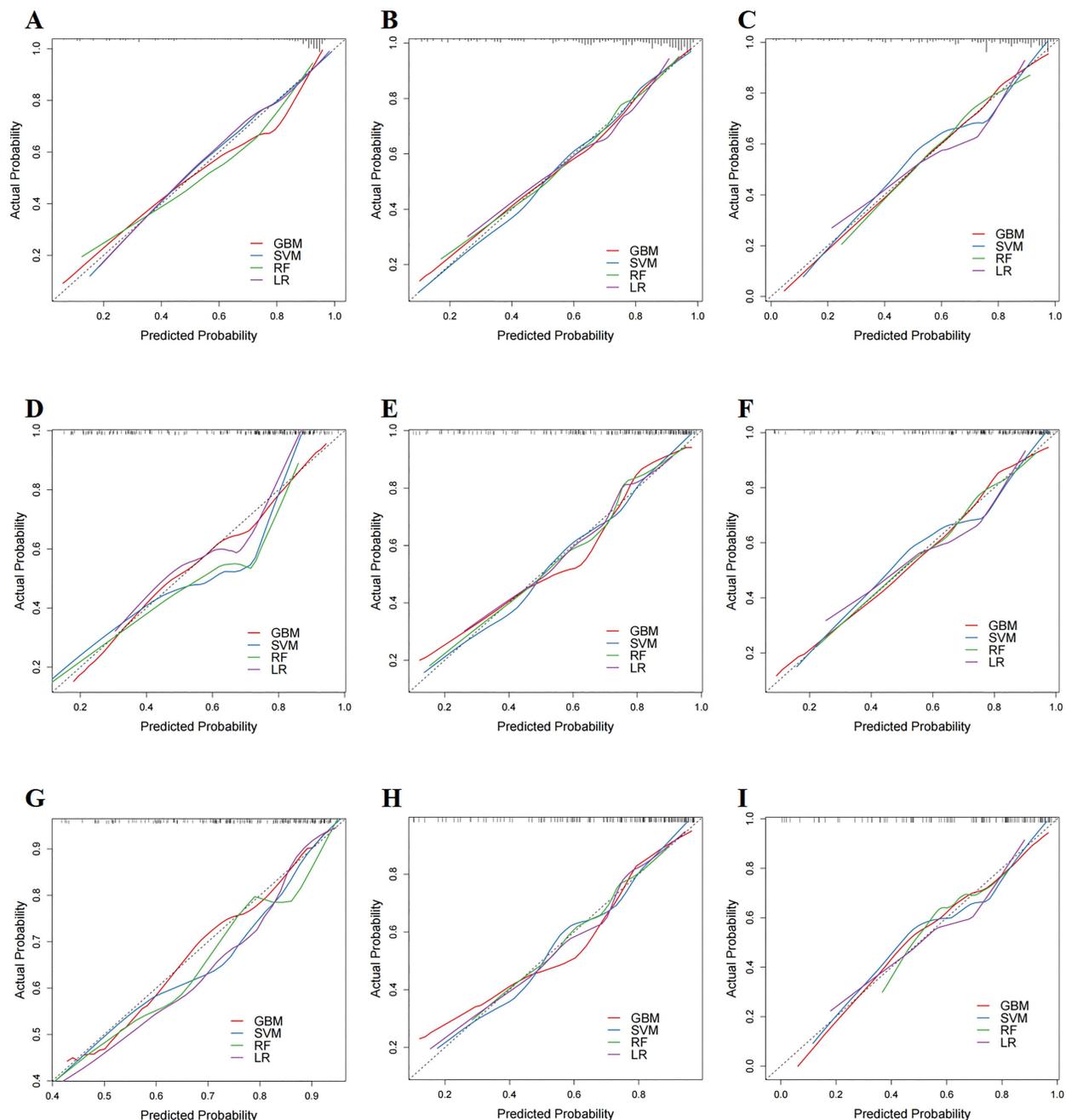
**Figure 2.** Receiver operating characteristic (ROC) curves of four ML models (GBM, SVM, RF, and LR). (A)–(C) AUC values of predictive models at 1-, 3-, and 5-year CSS in the training cohorts, respectively; (D)–(F) AUC values of predictive models at 1-, 3-, and 5-year CSS in the internal validation cohorts, respectively; (G)–(I) AUC values of predictive models at 1-, 3-, and 5-year CSS in the external validation cohorts, respectively. ML: Machine learning; GBM: Gradient boosting machine; SVM: Support vector machine; RF: Random forest; LR: Logistic regression; AUC: Area under the receiver operating characteristic curves.

Database registry data as the instrumental variable. Similar results were also reported by Zhang et al. and McIntosh et al. based on clinical data from the SEER database [23–25].

However, given the potential surgical complications and toxic effects of systemic therapy, a corresponding prognosis prediction model needs to be developed to improve patient selection and outcome prediction. Several prognostic prediction models have been used for mcrcc patients, such as

the International Metastatic Renal Cell Carcinoma Database Consortium model and the Memorial Sloan-Kettering Cancer Center model. However, models that specialize in predicting the prognosis of mcrcc patients receiving both CN and systemic therapy, especially ML models, have yet to be developed [26, 27].

Therefore, we aimed to develop a practical survival prediction model to accurately predict the individualized survival

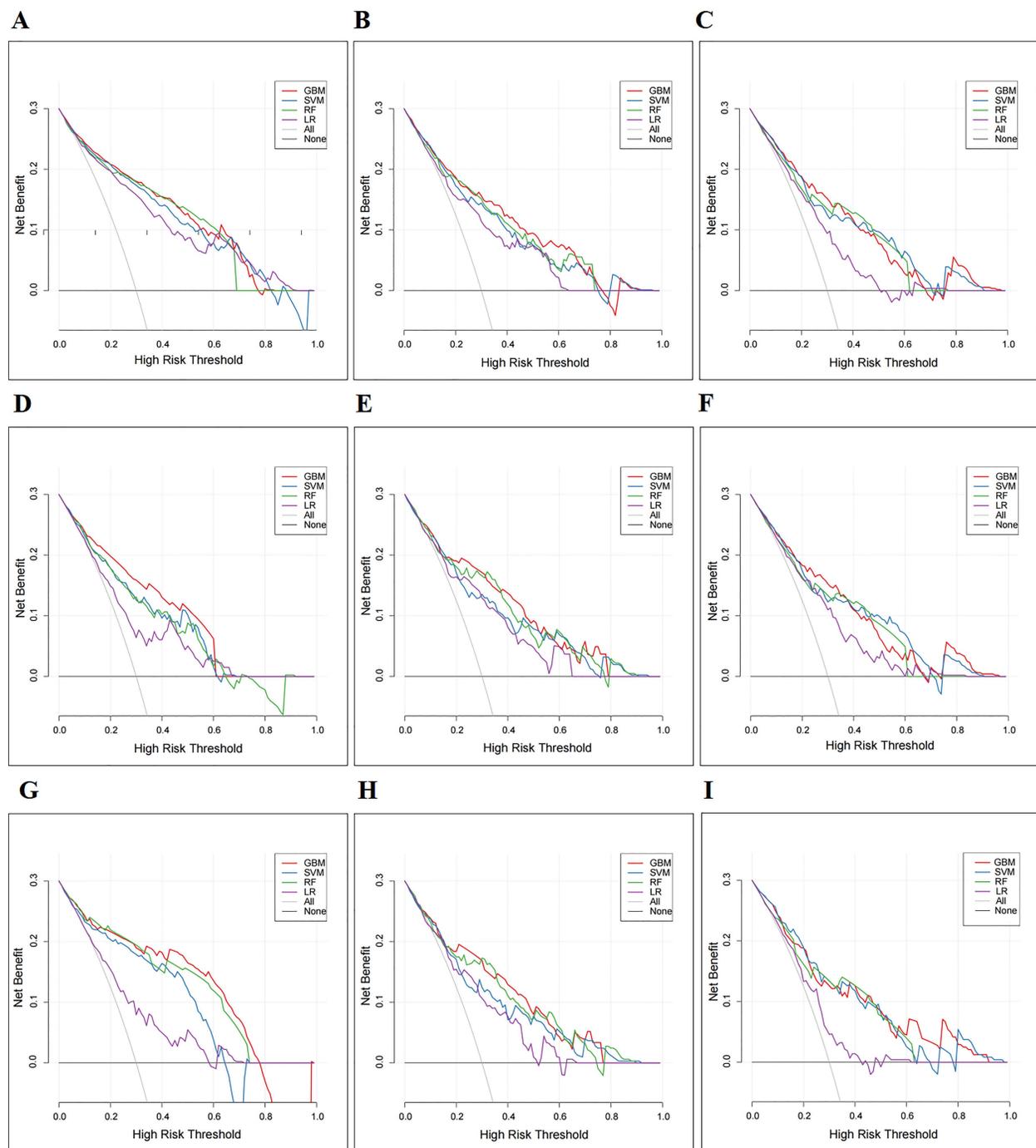


**Figure 3.** Calibration curves for predicting CSS of patients with four ML models (GBM, SVM, RF, and LR) at 1, 3, and 5 years, respectively. (A)–(C) The calibration curves of predictive models in the training cohorts; (D)–(F) The calibration curves of predictive models in the internal validation cohorts; (G)–(I) The calibration curves of predictive models in the external validation cohorts. CSS: Cancer-specific survival rate; ML: Machine learning; GBM: Gradient boosting machine; SVM: Support vector machine; RF: Random forest; LR: Logistic regression.

of mcrCC patients using ML algorithms. In this study, we established four ML models to predict the prognosis of these patients. Among them, GBM was the best model in terms of accuracy, fitness, and clinical application. To our knowledge, this is the first study to apply ML algorithms to predict mcrCC patient survival in such a large patient cohort.

One of the most significant advantages of ML models over traditional predictive models is their ability to analyze feature

importance and provide optimal feature subsets without requiring manual processing, resulting in models with high accuracy and stability. To date, multiple ML models have been developed and validated for predicting the prognosis of ccRCC patients based on medical imaging, gene expression data, or clinical information. For instance, Nazari et al. constructed a radiomics-based predictor using several ML algorithms to analyze CT images, which could accurately predict the 5-year survival of ccRCC patients [28]. However, a specific ML model



**Figure 4.** Decision curve analysis demonstrating the clinical benefit of four ML models (GBM, SVM, RF, and LR). (A)–(C) Training cohorts; (D)–(F) Internal validation cohorts; (G)–(I) External validation cohorts. ML: Machine learning; GBM: Gradient boosting machine; SVM: Support vector machine; RF: Random forest; LR: Logistic regression.

for predicting the prognosis of patients with mccRCC remains undeveloped.

In our study, the clinical data of over 900 mccRCC patients treated with CN and systemic therapy were included. We then constructed four ML models (GBM, SVM, RF, and LR) for 1-, 3-, and 5-year CSS prediction. Model performance was evaluated using ROC curves, calibration plots, and DCA. In the training, internal validation, and external validation cohorts, the

GBM model demonstrated the highest level of prediction accuracy and more favorable correlations. Furthermore, according to the DCA results, the GBM model can effectively assess the advantages and disadvantages of clinical decisions. In clinical research, GBM models are gaining increasing traction. We are the first to apply the GBM model to mccRCC survival prediction.

Selecting the most effective features from the original variables to reduce the dimensionality of the datasets is a key

Table 2. Cancer-specific survival (CSS) according to risk stratification

Risk group	1-year CSS, % (95% CI)	3-year CSS, % (95% CI)	5-year CSS, % (95% CI)	Hazard ratio % (95% CI)	P value
Training cohorts (n = 550)					
Low-risk	80.2 (76.0–84.5)	44.4 (38.7–51.1)	23.7 (18.0–31.2)	Ref	
Intermediate-risk	57.7 (49.1–67.9)	24.9 (16.9–36.7)	18.7 (11.3–31.1)	1.82 (1.33–2.48)	<0.001*
High-risk	32.4 (23.0–45.6)	6.9 (2.6–18.3)	2.3 (0.4–15)	2.02 (1.41–2.90)	<0.001 <sup>†</sup>
Internal validation cohorts (n = 232)					
Low-risk	83.1 (77.2–89.5)	48.4 (39.6–59.3)	26.2 (17.6–38.9)	Ref	
Intermediate-risk	53.1 (40.0–70.6)	20.0 (9.6–42.2)	15.1 (6.0–38.3)	2.24 (1.33–3.80)	<0.001*
High-risk	36.3 (23.5–55.9)	8.1 (2.4–27.1)	4.0 (0.6–25.4)	1.79 (1.08–2.96)	<0.001 <sup>†</sup>
External validation cohorts (n = 144)					
Low-risk	78.8 (70.8–87.8)	45.8 (34.6–57.9)	22.8 (13.2–39.3)	Ref	
Intermediate-risk	52.4 (38.3–71.7)	25.2 (15.4–46.2)	21.2 (10.9–41.2)	1.72 (1.02–2.91)	<0.001*
High-risk	11.8 (3.2–43.2)	5.9 (0.9–39.4)	5.9 (0.9–39.4)	2.14 (1.03–4.41)	<0.001 <sup>†</sup>

\*P value versus low-risk; <sup>†</sup>P value versus intermediate-risk.

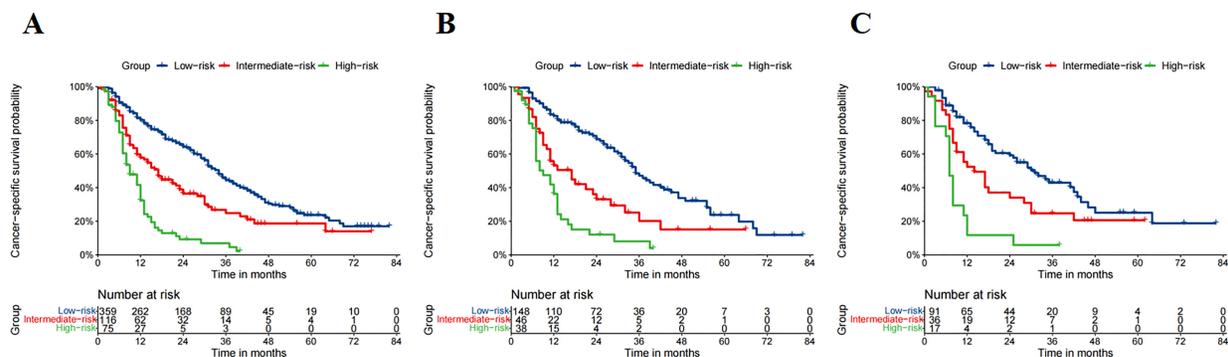


Figure 5. Kaplan-Meier showing disparities between groups. (A) CSS stratified by GBM model in the training cohorts; (B) CSS stratified by GBM model in the internal validation cohorts; (C) CSS stratified by GBM model in the external validation cohorts. CSS: Cancer-specific survival rate; GBM: Gradient boosting machine.

step in improving the performance of an ML model [17]. Using RFE methods, seven features, including pathological grade, T stage, N stage, number of metastatic sites, brain or liver metastases, and metastasectomy, were selected, and the optimal feature subset was chosen for further analysis. In the next step, the relative importance of each input feature was ranked using ML models. Despite the slight differences in feature importance ranking, histological grade, tumor stage, lymph node stage, number of metastatic sites, and metastasectomy were ranked in the top five in all ML models. These results revealed that mcrCC patients with lower histological grade, earlier tumor stage, no indication of lymph node metastases, fewer metastatic sites, and who received metastasectomy may have better prognoses after treatment with CN and systemic therapy.

Previous studies have also indicated that radiotherapy is significantly associated with better prognosis in patients with mcrCC [29, 30]. However, in this study, radiotherapy was not selected by the RFE algorithm for the optimal feature subset. Further studies on the influence of radiotherapy on the prognosis of patients with mcrCC treated with CN and systemic therapy are needed.

There have been dramatic changes in the treatment of mcrCC over the past few decades. The role of CN has become increasingly unclear with the advent of new treatment options. According to the results of the CARMENA trial and SURTIME trials, in comparison with the CN, sunitinib alone did not show any inferiority. However, all of the above studies were deemed to be underpowered, limited by insufficient study subjects, slow to accrue and lack of homogeneity in patients selection. Based on these limitations, it is important to select patients appropriately and publish prospective studies that contain high levels of evidence [31].

After the era of cytokine and targeted therapies, mcrCC treatment has gradually entered the era of immunotherapy. Nirmish et al. compared the prognosis of mcrCC patients treated with immune checkpoint inhibitors (ICIs) alone or CN combined with ICIs. The result showed that CN plus immunotherapy had a longer OS than immunotherapy alone based on the NCDB datasets [32]. Thus, despite the uncertainty of efficacy, CN remains an important treatment option, especially for alleviating hematuria or pain. However, optimal candidates for CN need to be carefully screened. Recent perspectives supported that CN could be performed in a patient

with a kidney that is in place and a disease of favorable or intermediate risk [33].

In this study, the GBM model identified approximately 15% of patients with mcrRCC in the high-risk group who experienced an extremely poor 5-year survival rate after receiving CN and systemic therapy. Conversely, more than half of the mcrRCC patients were classified as low-risk patients. The low-risk subset displayed relatively satisfactory long-term CSS, indicating that CN combined with systemic therapy is particularly suitable for this population. With the above risk stratification, overtreatment can be largely avoided.

Present study has several limitations: First, because the training set data used for building models were from the SEER database, the treatment regimen and information on systemic therapy were not available. The latest studies have reported that deferred CN may be more beneficial for patients who respond favorably to systemic therapy compared with upfront CN [6, 34, 35]. Second, the SEER database lacks data on some vital clinical characteristics, such as basic diseases, surgical complications, and biochemical indicators, which may influence the accuracy of the model prediction. Third, there is inter-group heterogeneity in the external validation cohort compared with the data from the SEER database, which may be related to the relatively small amount of data in the external validation cohort. Therefore, a multi-center study must be conducted to further validate the model performance.

Despite mentioned limitations, these ML models, built on a large population database and validated with data from external groups, provide the first targeted and practical survival prediction tools for patients with mcrRCC receiving both CN and systemic therapy, which have a high potential for use in clinical practice.

## Conclusion

We developed and validated four ML models based on significant clinicopathological characteristics for predicting CSS in patients with mcrRCC treated with CN and systemic therapy. These models will not only be used in CSS prediction, patient risk stratification, and clinical decision making but also encourage further research on the use of ML algorithms to improve personalized prognostic prediction.

**Conflicts of interest:** The authors declare no conflicts of interest.

**Funding:** This study was supported by grant from the Research Foundation of Youth, Peking Union Medical College (2020zlgc0103).

Submitted: 24 August 2022

Accepted: 13 October 2022

Published online: 2 November 2022

## References

[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* 2022;72(1):7–33. <https://doi.org/10.3322/caac.21708>.

[2] Choueiri TK, Motzer RJ. Systemic therapy for metastatic renal-cell carcinoma. *N Engl J Med* 2017;376(4):354–66. <https://doi.org/10.1056/NEJMra1601333>.

[3] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394–424. <https://doi.org/10.3322/caac.21492>.

[4] Bhindi B, Abel EJ, Albiges L, Bensalah K, Boorjian SA, Daneshmand S, et al. Systematic review of the role of cytoreductive nephrectomy in the targeted therapy era and beyond: an individualized approach to metastatic renal cell carcinoma. *Eur Urol* 2019;75(1):111–28. <https://doi.org/10.1016/j.eururo.2018.09.016>.

[5] Motzer RJ, Russo P. Cytoreductive nephrectomy—patient selection is key. *N Engl J Med* 2018;379(5):481–2. <https://doi.org/10.1056/NEJMe1806331>.

[6] Bhindi B, Graham J, Wells JC, Bakouny Z, Donskov F, Fraccon A, et al. Deferred cytoreductive nephrectomy in patients with newly diagnosed metastatic renal cell carcinoma. *Eur Urol* 2020;78(4):615–23. <https://doi.org/10.1016/j.eururo.2020.04.038>.

[7] Bhindi B, Habermann EB, Mason RJ, Costello BA, Pagliaro LC, Thompson RH, et al. Comparative survival following initial cytoreductive nephrectomy versus initial targeted therapy for metastatic renal cell carcinoma. *J Urol* 2018;200(3):528–34. <https://doi.org/10.1016/j.juro.2018.03.077>.

[8] Flanigan RC, Salmon SE, Blumenstein BA, Bearman SI, Roy V, McGrath PC, et al. Nephrectomy followed by interferon alfa-2b compared with interferon alfa-2b alone for metastatic renal-cell cancer. *N Engl J Med* 2001;345(23):1655–9. <https://doi.org/10.1056/NEJMoa003013>.

[9] Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, Kapoor A, et al. Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. *N Engl J Med* 2007;356(22):2271–81. <https://doi.org/10.1056/NEJMoa066838>.

[10] Méjean A, Ravaud A, Thezenas S, Colas S, Beauval JB, Bensalah K, et al. Sunitinib alone or after nephrectomy in metastatic renal-cell carcinoma. *N Engl J Med* 2018;379(5):417–27. <https://doi.org/10.1056/NEJMoa1803675>.

[11] Bex A, Mulders P, Jewett M, Wagstaff J, van Thienen JV, Blank CU, et al. Comparison of immediate vs deferred cytoreductive nephrectomy in patients with synchronous metastatic renal cell carcinoma receiving sunitinib: the SURTIME randomized clinical trial. *JAMA Oncol* 2019;5(2):164–70. <https://doi.org/10.1001/jamaoncol.2018.5543>.

[12] Deo RC. Machine learning in medicine. *Circulation* 2015;132(20):1920–30. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>.

[13] Rajkomar A, Dean J, Kohane I. Dean, I. Kohane. Machine learning in medicine. *N Engl J Med* 2019;380(14):1347–58. <https://doi.org/10.1056/NEJMra1814259>.

[14] Peterson ED. Machine learning, predictive analytics, and clinical practice: can the past inform the present? *JAMA* 2019;322(23):2283–4. <https://doi.org/10.1001/jama.2019.17831>.

[15] Ji GW, Jiao CY, Xu ZG, Li XC, Wang K, Wang XH. Development and validation of a gradient boosting machine to predict prognosis after liver resection for intrahepatic cholangiocarcinoma. *BMC Cancer* 2022;22(1):258. <https://doi.org/10.1186/s12885-022-09352-3>.

[16] Pal SK, Ghate SR, Li N, Swallow E, Peebles M, Zichlin ML, et al. Real-world survival outcomes and prognostic factors among patients receiving first targeted therapy for advanced renal cell carcinoma: a SEER–medicare database analysis. *Clin Genitourin Cancer* 2017;15(4):e573–82. <https://doi.org/10.1016/j.clgc.2016.12.005>.

[17] Chen X, Yang J, Lu Z, Ding Y. A 70-RNA model based on SVR and RFE for predicting the pancreatic cancer clinical prognosis. *Methods* 2022;204:278–85. <https://doi.org/10.1016/j.ymeth.2022.02.011>.

[18] Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. *JAMA* 2015;313(4):409–10. <https://doi.org/10.1001/jama.2015.37>.

[19] Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. Reporting and interpreting decision curve analysis: a guide for investigators. *Eur Urol* 2018;74(6):796–804. <https://doi.org/10.1016/j.eururo.2018.08.038>.

[20] Shikama A, Minaguchi T, Takao W, Hosokawa Y, Nishida K, Tasaka N, et al. Predictors of favorable survival after secondary cytoreductive surgery for recurrent endometrial cancer. *Int J Clin Oncol* 2019;24(10):1256–63. <https://doi.org/10.1007/s10147-019-01469-z>.

[21] Bonnot PE, Lintis A, Mercier F, Benzerdjeb N, Passot G, Pocard M, et al. Prognosis of poorly cohesive gastric cancer after complete cytoreductive surgery with or without hyperthermic intraperitoneal chemotherapy (CYTO-CHIP study). *Br J Surg* 2021;108(10):1225–35. <https://doi.org/10.1093/bjs/znab200>.

- [22] Nica A, Covens A, Parra-Herran C, May T, et al. Does timing of intraperitoneal chemotherapy initiation following primary cytoreductive surgery with bowel resection impact outcomes in patients with advanced ovarian cancer? *Gynecol Oncol* 2020;158(3):622–30. <https://doi.org/10.1016/j.ygyno.2020.06.478>.
- [23] Chakiryan NH, Gore LR, Reich RR, Dunn RL, Jiang DD, Gillis KA, et al. Survival outcomes associated with cytoreductive nephrectomy in patients with metastatic clear cell renal cell carcinoma. *JAMA Netw Open* 2022;5(5):e2212347. <https://doi.org/10.1001/jamanetworkopen.2022.12347>.
- [24] Zhang Y, Hu J, Yang J, Xie Y, Chen Z, Shanguan W, et al. Selection of optimal candidates for cytoreductive nephrectomy in patients with metastatic clear cell renal cell carcinoma: a predictive model based on SEER database. *Front Oncol* 2022;12:814512. <https://doi.org/10.3389/fonc.2022.814512>.
- [25] McIntosh AG, Umbreit EC, Holland LC, Gu C, Tannir NM, Matin SF, et al. Optimizing patient selection for cytoreductive nephrectomy based on outcomes in the contemporary era of systemic therapy. *Cancer* 2020;126(17):3950–60. <https://doi.org/10.1002/cncr.32991>.
- [26] Ciccavese C, Iacovelli R, Porta C, Procopio G, Bria E, Astore S, et al. Efficacy of VEGFR-TKIs plus immune checkpoint inhibitors in metastatic renal cell carcinoma patients with favorable IMDC prognosis. *Cancer Treat Rev* 2021;100:102295. <https://doi.org/10.1016/j.ctrv.2021.102295>.
- [27] Fiala O, Finek J, Poprach A, Melichar B, Kopecky J, Zemanova M, et al. Outcomes according to MSKCC risk score with focus on the intermediate-risk group in metastatic renal cell carcinoma patients treated with first-line sunitinib: a retrospective analysis of 2390 patients. *Cancers (Basel)* 2020;12(4):808. <https://doi.org/10.3390/cancers12040808>.
- [28] Nazari M, Shiri I, Zaidi H. Radiomics-based machine learning model to predict risk of death within 5-years in clear cell renal cell carcinoma patients. *Comput Biol Med* 2021;129:104135. <https://doi.org/10.1016/j.combiomed.2020.104135>.
- [29] Hsieh PY, Hung SC, Li JR, Wang SS, Yang CK, Chen CS, et al. The effect of metastasectomy on overall survival in metastatic renal cell carcinoma: a systematic review and meta-analysis. *Urol Oncol* 2021;39(7):422–30. <https://doi.org/10.1016/j.urolonc.2021.02.026>.
- [30] Cheung P, Patel S, North SA, Sahgal A, Chu W, Soliman H, et al. Stereotactic radiotherapy for oligoprogression in metastatic renal cell cancer patients receiving tyrosine kinase inhibitor therapy: a phase 2 prospective multicenter study. *Eur Urol* 2021;80(6):693–700. <https://doi.org/10.1016/j.eururo.2021.07.026>.
- [31] Lichtbroun BJ, Srivastava A, Doppalapudi SK, Chua K, Singer EA. New paradigms for cytoreductive nephrectomy. *Cancers (Basel)* 2022;14(11):2660. <https://doi.org/10.3390/cancers14112660>.
- [32] Singla N, Hutchinson RC, Ghandour RA, Freifeld Y, Fang D, Sagalowsky AI, et al. Improved survival after cytoreductive nephrectomy for metastatic renal cell carcinoma in the contemporary immunotherapy era: an analysis of the national cancer database. *Urol Oncol* 2020;38(6):604.e9–17. <https://doi.org/10.1016/j.urolonc.2020.02.029>.
- [33] Rathmell WK, Rumble RB, Van Veldhuizen PJ, Al-Ahmadie H, Emamekhoo H, Hauke RJ, et al. Management of metastatic clear cell renal cell carcinoma: ASCO guideline. *J Clin Oncol* 2022;40(25):2957–95. <https://doi.org/10.1200/JCO.22.00868>.
- [34] Poprach A, Holanek M, Chloupkova R, Lakomy R, Stanik M, Fiala O, et al. Cytoreductive nephrectomy and overall survival of patients with metastatic renal cell carcinoma treated with targeted therapy—data from the national renal registry. *Cancers (Basel)* 2020;12(10):2911. <https://doi.org/10.3390/cancers12102911>.
- [35] Ghatalia P, Handorf EA, Geynisman DM, Deng M, Zibelman MR, Abbosh P, et al. The role of cytoreductive nephrectomy in metastatic renal cell carcinoma: a real-world multi-institutional analysis. *J Urol* 2022;208(1):71–9. <https://doi.org/10.1097/JU.0000000000002495>.

## Related articles published in BJBMS

1. Development and validation of radiomics machine learning model based on contrast-enhanced computed tomography to predict axillary lymph node metastasis in breast cancer

Jieqiu Zhang et al., BJBMS, 2022

2. Machine learning as the new approach in understanding biomarkers of suicidal behavior

Alja Videtič Paska et al., BJBMS, 2021

Supplemental Data

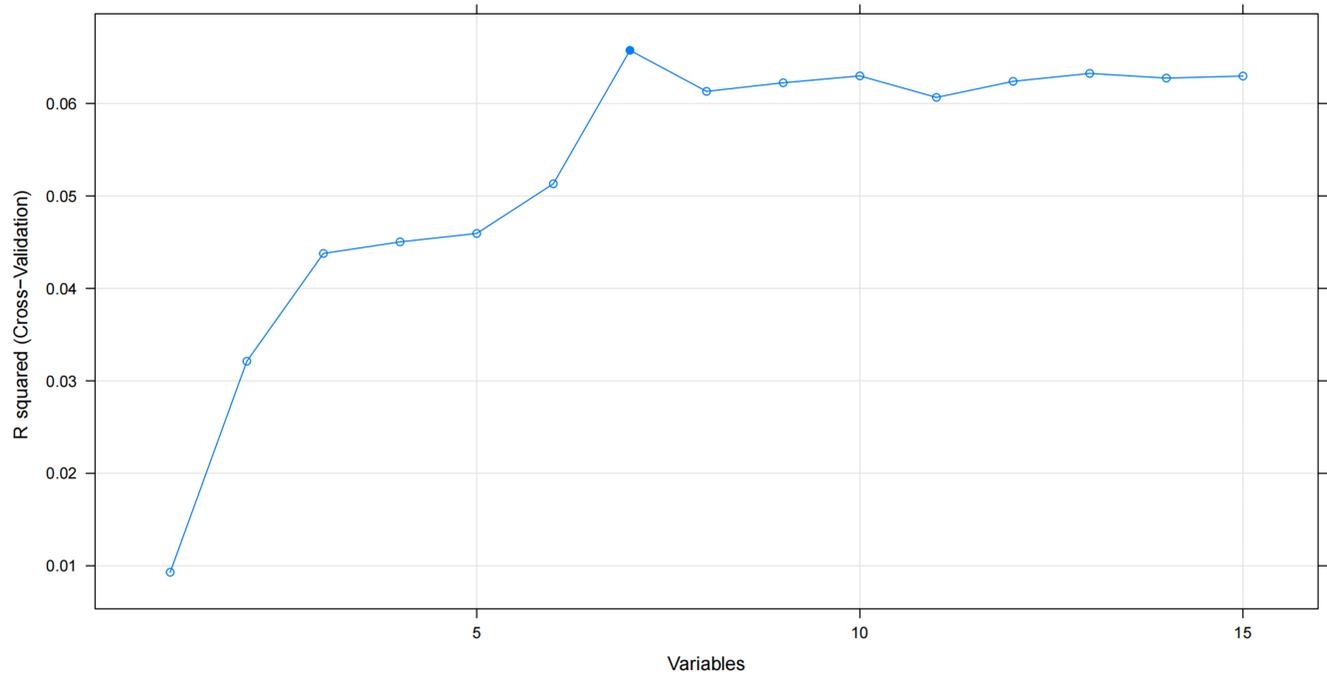
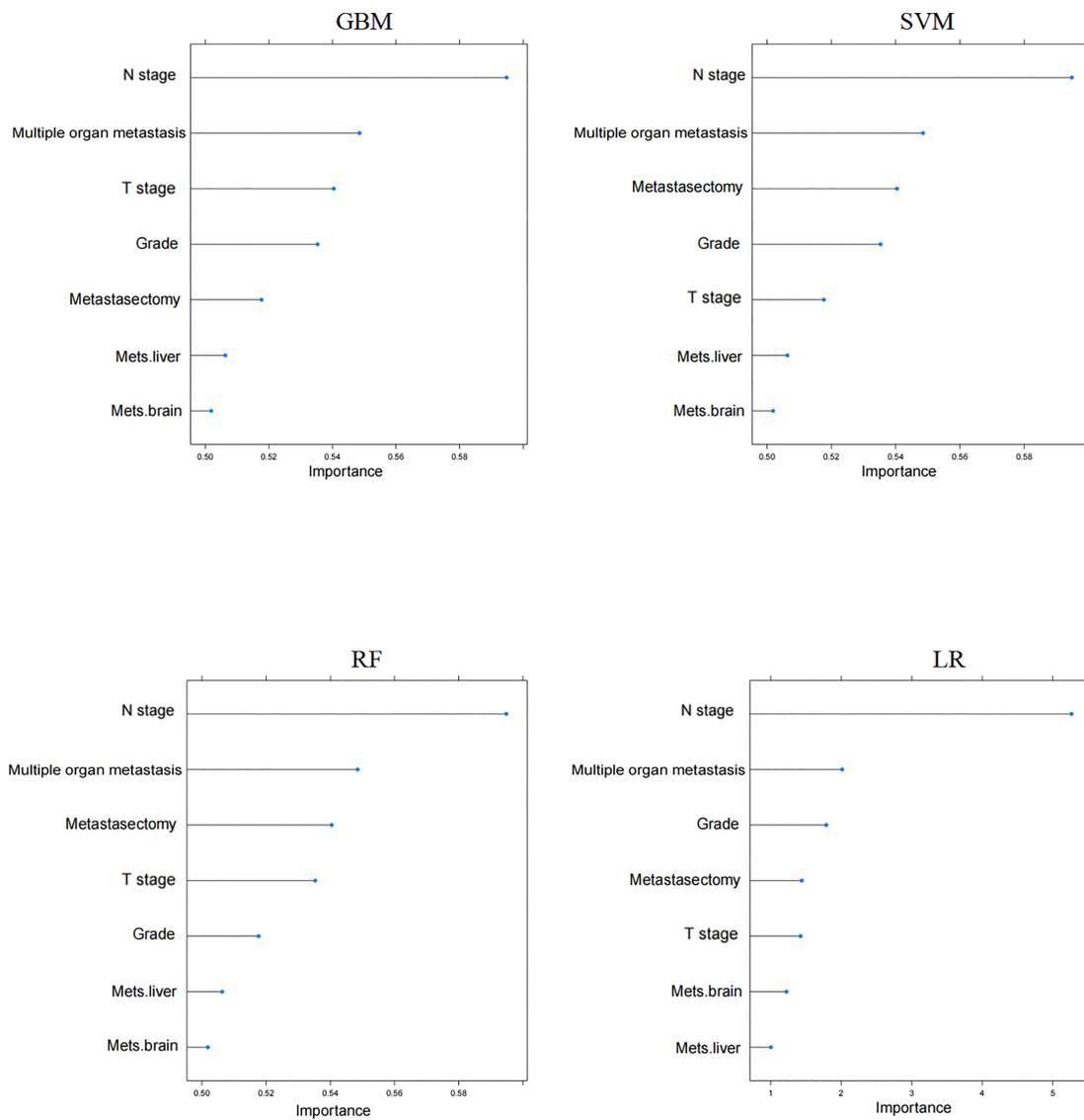
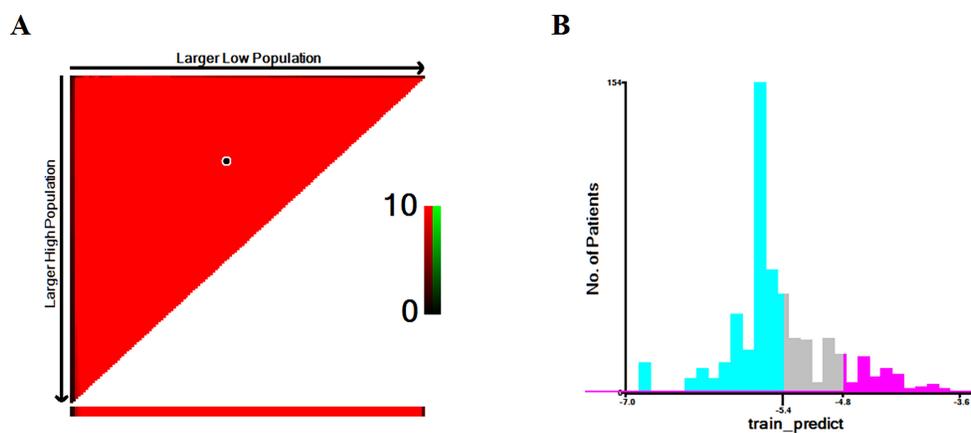


Figure S1. The correlation between included features numbers and the prediction fitness in the recursive feature elimination (RFE) algorithm.



**Figure S2.** The relative importance of features in four ML models (GBM, SVM, RF, and LR). GBM: Gradient boosting machine; SVM: Support vector machine; RF: Random forest; LR: Logistic regression; Mets: Metastasis organ site.



**Figure S3.** Determine the optimal cutoffs for GBM via X-tile analysis. (A) The optimal cutoffs highlighted by black circle; (B) Numbers of patients in high, intermediate, and low mortality risk subsets. GBM: Gradient boosting machine.