Am J Cancer Res 2026;16(1):173-187
www.ajcr.us /ISSN:2156-6976/ajcr0169218

Original Article

Construction and validation of deep learning-based pathomics signature model for predicting postoperative recurrence of patients with clear cell renal cell carcinoma

Yiren Yang[1*], Jinxin Li[1*], Yisha Gao[2*], Wenqiang Liu[1], Baohua Zhu[1], Jie'an Ding[1], Linhui Wang[1#], Xiaolei Shi[1#], Wei Zhang[1#]

[1]Department of Urology, Changhai Hospital, Naval Medical University, Shanghai, China; [2]Department of Pathology, Changhai Hospital, Naval Medical University, Shanghai, China. [*]Equal contributors. [#]Co-corresponding authors.

Received October 10, 2025; Accepted December 19, 2025; Epub January 15, 2026; Published January 30, 2026

Abstract: This study developed an attention-based multi-instance deep convolutional neural network (CRPNet) for the early prediction of postoperative recurrence in clear cell renal cell carcinoma (ccRCC). The model was trained on 183 whole slide images (WSIs) from ccRCC patients and validated on an internal cohort of 75 WSIs. Its prognostic performance was evaluated using Kaplan-Meier analysis, AUC, accuracy, precision, recall, false positive/negative rates (FPR/FNR), C-index, and hazard ratio (HR), and was compared against established tools including the UISS, SSIGN, and Karakiewicz nomograms. Results demonstrated that CRPNet-stratified high-risk groups had significantly poorer prognosis in both training and validation sets ($P < 0.001$), with consistency across subgroups based on T-stage, WHO/ISUP grade, and necrosis. In the training cohort, CRPNet achieved an AUC of 0.994 (95% CI: 0.974-1.000), accuracy of 97.70%, precision/recall of 95.56%, FPR of 1.55%, and FNR of 4.45%. In the validation cohort, it maintained an AUC of 0.879 (95% CI: 0.783-0.943), accuracy of 88.00%, precision of 85.71%, recall of 63.16%, FPR of 0%, and FNR of 36.84%, outperforming all comparator models. CRPNet also yielded a superior C-index compared to clinical parameters and traditional nomograms, and exhibited the highest HR (12.078, 95% CI: 1.611-90.539). In conclusion, CRPNet surpasses conventional prognostic models in recurrence prediction accuracy, AUC, precision, C-index, and risk stratification, while demonstrating lower FPR and FNR, thereby offering improved prognostication for metastatic ccRCC.

Keywords: Clear cell renal cell carcinoma, recurrence, deep learning, pathomics, CRPNet

Introduction

Renal cell carcinoma (RCC) is one of the most common tumors in the urinary system [1] and clear cell renal cell carcinoma (ccRCC) is the most frequently encountered pathological subtype of RCC [2]. One thirds of patients diagnosed with ccRCC exhibit regional or distant metastatic spread, and among patients with localized ccRCC who underwent nephrectomy, approximately one in four experienced recurrences at distant sites [3]. Despite significant advancements in the systemic management of recurrent and metastatic ccRCC over the past two decades, the mortality rate remains unacceptably high, with more than 179,000 patients succumbing to RCC-related causes [4]. Currently, postoperative recurrence detection in ccRCC patients heavily relies on imaging examinations, however, this approach lags. Furthermore, there remains controversy regarding which patients require postoperative adjuvant therapy, which may lead to potential complications. Therefore, the development of an early warning model is essential for postoperative precise classification of ccRCC patients, especially for those with locally advanced tumors, facilitating individualized treatment strategies and follow-up protocols [5]. Although imaging examinations are non-invasive and simple, the high cost and radiation risk of enhanced CT and bone scintigraphy limit their

https://doi.org/10.62347/ZCHX8002

clinical utility during the frequent postoperative follow-up [6, 7]. Currently, there is a lack of effective predictive model to determine the risk of postoperative recurrence in patients with ccRCC.

In recent years, the development of artificial intelligence technology has facilitated the emergence of numerous deep learning-based pathomics studies, which involve utilizing artificial intelligence technology to extract imaging features from hematoxylin-eosin (HE) staining pathological images obtained after surgery and employing deep learning models to analyze and construct predictive models based on these features [8-10]. Deep neural networks have the ability to detect fine structures that are beyond the naked eye of pathologists, discern even subtle differences, and thereby improve the accuracy of distinguishing patients with different prognosis [11]. The application of deep learning analysis to pathological sections has been extended to a range of cancer types, including breast cancer, colorectal cancer, lung cancer, liver cancer, and prostate cancer. Deep learning has rapidly advanced in clear cell renal cell carcinoma (ccRCC) research, primarily focusing on extracting prognostic information from multi-source data that surpasses traditional pathological assessments. For instance, Mahootiha et al. [12] developed a multimodal deep learning method based on a 3D CNN architecture, which showed promising results in estimating the survival probability of renal cell carcinoma patients using CT imaging and clinical data. Kawahara et al. [13] correlated imaging features with molecular information to achieve non-invasive preoperative prediction of two-year disease-free survival. Lu et al. [14] reported an interpretable weakly supervised deep learning method for data-efficient whole-slide image (WSI) processing and learning, applied to the subtyping of renal cell carcinoma and non-small cell lung cancer as well as the detection of lymph node metastasis. This method outperformed standard weakly supervised classification algorithms and demonstrated applicability to independent validation cohorts, smartphone microscopy, and varying tissue content.

In this study, we analyzed the complete pathological HE staining images of primary tumors from 258 ccRCC patients, developing a multi-instance deep convolutional neural network based on a multi-scale attention mechanism. The network was then trained to predict the postoperative recurrence status of ccRCC patients. The effective predicting performance of our pathomics signature model provides new ideas for personalized treatment of ccRCC patients in the future.

Methods

Workflow of the study

The process of model construction is shown in **Figure 1A**. Pathological slices are collected from all patients, diagnosed and confirmed under a microscope, and then scanned into digital slices to obtain whole slide imaging (WSI).

A total of 258 patients with ccRCC are included in this study. The pathological slices of each patient are diagnosed and confirmed by three pathologists, and the most typical WSIs are selected for inclusion in the study. We randomly split the dataset of WSIs from 258 patients into a training cohort (183, 71%) and validation cohort (75, 29%) (**Figure 1B**).

Patients and samples

The pathological specimens were collected from 258 patients with renal cancer who underwent surgical treatment in Changhai Hospital, Naval Medical University from February 2013 to August 2021. All the patients were diagnosed with ccRCC postoperatively. All slides are digitally scanned using the Hamamatsu Nano-zoomer S60 scanner to produce WSI in our study, as per the protocol described. The labels for WSIs are derived from pathological reports and postoperative follow-ups, and categorized as “recurrence” or “non-recurrence”. Each case of ccRCC is accompanied by a representative hematoxylin and eosin (HE)-stained digital slide in the Nuclear Digital Pathology Imaging (NDPI) format. These slides encompass both the ccRCC tissue and the adjacent normal kidney tissue for comparative analysis. All the 258 patients were segregated into a training cohort, comprising 183 patients, and a validation cohort, comprising 75 patients. The inclusion criteria for the two cohorts are delineated as follows: patients who have undergone surgical resection but have not received antitu-

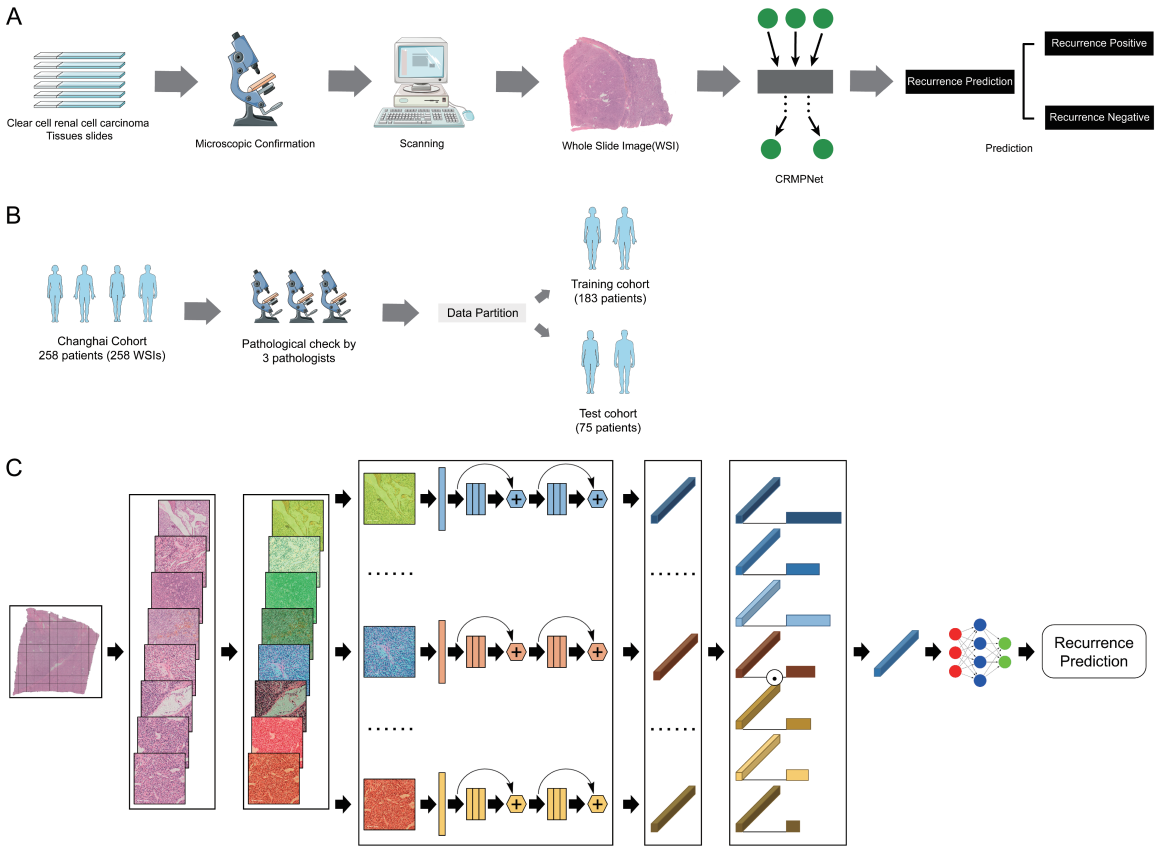


Figure 1. Overall flow chart of the study. (A) The process of model construction, (B) Patient cohorts grouping, (C) Training process of a multi-instance deep neural network (CRPNet) based on a multi-scale attention mechanism.

mor therapy; patients with a postoperative pathological diagnosis of ccRCC; patients without distant metastasis at the time of surgery; patients with available follow-up data; patients with available formalin-fixed paraffin-embedded tissue sections; and patients with a clinical focus on postoperative recurrence. The protocol for this investigation was granted approval by the Ethics Committee of Changhai Hospital (No. CHEC2021-275), and each participant provided documented, informed consent. Conduct of the study conformed strictly to the ethical tenets enshrined within the Declaration of Helsinki.

The core of this study established an integrative pathomics signature via deep learning, which serves as a strong predictor (P = 1). Based on the significant predictive signal observed in the training set (Cox-Snell R^2 = 0.57), calculations indicate that only approximately 10 samples are required to robustly fit this signature while preventing overfitting (shrinkage factor S ≥ 0.9). Even under the conventional '10 Events Per Variable (EPV)' rule, only 39 samples are needed. Our current training cohort of 174 samples exceeds this minimum requirement by a factor of 4 to 17. The selection of representative FFPE tissue blocks and corresponding HE stained slides for digital scanning was performed by an experienced uropathologist (10 years of experience) who was blinded to the clinical outcomes of the patients. Further calculations demonstrate that the current sample size (N = 174, Events = 45) is sufficient to support the simultaneous modeling of up to 17 independent predictors while still maintaining a shrinkage factor (S) ≥ 0.9. This indicates that the sample size is adequate not only for validating the single pathomics signature but also for constructing a multifactorial model that incorporates clinical variables. Power analysis shows that for a strong predictor similar to the pathomics signature in this study (Odds Ratio, OR ≥ 2.0), the current sample size achieves a high statistical power of

97.95% (alpha = 0.05), significantly exceeding the conventional standard of 80%.

Data collection and comparative analysis

Retrieval of clinical data from patients' records encompasses age, sex, body mass index (BMI), annotations regarding tobacco and alcohol use history, along with concomitant medical conditions including hypertension, diabetes mellitus, and coronary artery disease. Additionally, a comprehensive pathological profile was compiled which detailed tumor dimensions, T staging, WHO/ISUP histological grade, and particulars of tumor necrosis. For each patient, the SSIGN (Staging, Size, Grade, Necrosis) score, UISS (University of California, Los Angeles Integrated Staging System), and Karakiewicz nomogram were computed utilizing methodologies from extant scholarly inquiries [15-17].

Convolutional neural networks to predict recurrence outcome

The network's testing phase abstains from the use of data augmentation, directly outputting the risk classification for each test WSI. Our computational framework is powered by Python (Version 3.7) and the PyTorch deep learning platform (Version 1.7.1), ensuring a robust and efficient training regimen. The model benefits from the Adam optimizer and cross-entropy loss, with a total of 300 training iterations, a learning rate of 1e-4, a weight decay coefficient of 1e-5, and a batch size of 64. The hardware setup for training includes the 11th Gen Intel(R) Core (TM) i7-11700K CPU and an Nvidia 3090 GPU, ensuring high-performance computation throughout the model development process.

Statistical analysis

Statistical analysis and chart creation are conducted using SPSS 26.0 and GraphPad Prism 9.5 software. For measurement data, a test of normal distribution is first performed. Data that conformed to a normal distribution are expressed as "mean ± standard deviation", while data that do not conform to a normal distribution were expressed as "median ± interquartile range". Unpaired t-tests are used to compare differences between two groups. For comparisons among multiple groups, different statistical methods were selected based on whether the data conformed to a normal distribution.

The predictive performance of the model is evaluated separately for the validation cohort and the validation cohort, including false positive rate, false negative rate, accuracy, precision, recall and Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC). Progression-free survival (PFS) curves are estimated using the Kaplan Meier method (survival and survminer R packages). We use the concordance index (C-index) to evaluate the predictive performance of our model, and compare it with clinical baseline data and other scoring models through hazard ratios (HRs). The p value < 0.05 is considered to indicate a statistically significant difference.

Results

Dataset overview

In the training cohort, the highest proportion of patients with recurrence was observed in T3a (n = 12) and T1b (n = 14) subgroups. Conversely, in the non-recurrence group, T1a accounted for the highest proportion with 100 patients, and the disparity between the two groups was statistically significant (P < 0.001). Similarly, there were also notable differences in tumor diameter and N stage between the two groups (P < 0.001). Interestingly, although disparities in WHO/ISUP grading system were observed between the recurrence and non-recurrence groups, grade II was the most prevalent in both groups. However, in the validation cohort, there was no statistically significant difference in N stage (P = 1.000) and WHO/ISUP grade (P = 0.069) between the two patient groups. Nevertheless, there were still notable statistical differences in tumor diameter and T stage (both P < 0.001). More baseline characteristics of patients in the two datasets are presented in Supplementary Tables 1 and 2.

Model construction

For the preprocessing of the WSI dataset, we cropped each WSI at 20× magnification into multiple small image patches using fixed-size windows of 224×224 pixels without overlap. To remove regions that contained little to no tissue information, we calculated the image entropy for each patch and discarded those with an entropy value less than 5, as such patches typically correspond to slide background or large blank areas.

Subsequently, we incorporat the WSIs from the training cohort into the neural network training and construct the model, naming it Changhai RCC Recurrence Prediction Net (CRPNet). In this study, we introduce CRPNet, an advanced architecture based on the principles of attention mechanism-driven multi-instance deep convolutional neural networks, specifically designed to enhance the accuracy of prognostic predictions. The methodology underpinning the network's training is depicted in **Figure 1C**, illustrating a structured approach to model development. Initially, the process involves creating a 'bag' comprising a collection of patches extracted from a single WSI. Prior to each training iteration, these patches undergo a series of randomized data augmentation techniques, including the introduction of random noise, rotation, and color adjustments, to enhance the model's robustness and ability to generalize.

Feature extraction is conducted using a ResNet network pre-trained on the ImageNet dataset, serving as the primary mechanism for analyzing each patch. This step is critical for capturing the nuanced features within each patch, which are pivotal for accurate prognostic assessments. Following feature extraction, an attention module is employed to assign learnable weights to each patch's features, culminating in a comprehensive feature representation of the bag through an attention pooling process. The attention module innovatively reduces the dimensionality of each feature vector to 128 using a linear fully connected layer, followed by the application of a pixel-level hyperbolic tangent function (tanh(.)), normalizing the outputs between -1 and 1. The process is further refined by multiplying the output with another linear layer, and computing the attention weights for each patch through a softmax function. The culmination of this process is the generation of bag-level features through the inner product of the feature matrix of image patches with the corresponding attention weight matrix.

To predict the prognostic risk associated with each WSI, we construct a bag-level fully connected network (Multilayer Perceptron, MLP), which acts as a sophisticated classifier. This model innovatively translates the continuous variable of survival times into a binary classification of high and low-risk groups, enabling a streamlined training process via the cross-entropy loss function. Optimization of the network parameters is achieved through the application of stochastic gradient descent.

Predictive performance of CRPNet

The predictive performance of the CRPNet model in both training cohort and validation cohort for predicting the recurrence status of ccRCC was evaluated using metrics such as AUC, accuracy, precision, recall, false positive rate (FPR), and false negative rate (FNR).

The AUCs of CRPNet model are 0.994 (95% CI 0.974-1.000) and 0.879 (95% CI 0.783-0.943) in training and validation cohorts, respectively. The accuracy (97.70% vs. 85.71%), precision (95.56% vs. 85.71%), and FPR (1.55% vs. 3.57%) are consistent between training and validation cohorts. However, compared with training cohort, the recall decreases (95.56% vs. 63.16%) while the FNR increases (4.45% vs. 36.84%) in validation cohort.

Subsequently, we conduct survival analysis in both training and validation cohort using CRPNet, as well as T stage, WHO/ISUP grade, and necrosis status, constructing PFS curves for each model. Based on the median score of CRPNet, we divide all the patients into high-risk and low-risk groups. In the training cohort, high-risk group has shorter PFS than low-risk group. After stratifying patients based on characteristics such as T1-2 stage, T3-4 stage, WHO/ISUP I-II grade, WHO/ISUP III-IV grade, and necrosis, CRPNet was still able to predict survival outcomes (**Figure 2**). Furthermore, the same results could be seen in the validation cohort (**Figure 3**).

Besides, we also evaluate the performances of several commonly used prognostic models including UISS, SSIGN score and Karakiewicz nomogram, and compare them with the CRPNet. As **Table 1** shows, the AUC (0.879, 95% CI 0.783-0.943), accuracy (88.00%), precision (85.71%), recall (63.16%), FPR (3.57%), and FNR (36.84%) of CRPNet model significantly outperform other prognostic models. The aforementioned performance indicators demonstrate that the CRPNet model, based on pathological slices of primary lesions stained with HE, can accurately predict the postoperative recurrence risk of ccRCC patients. However, the Karakiewicz nomogram demonstrated the

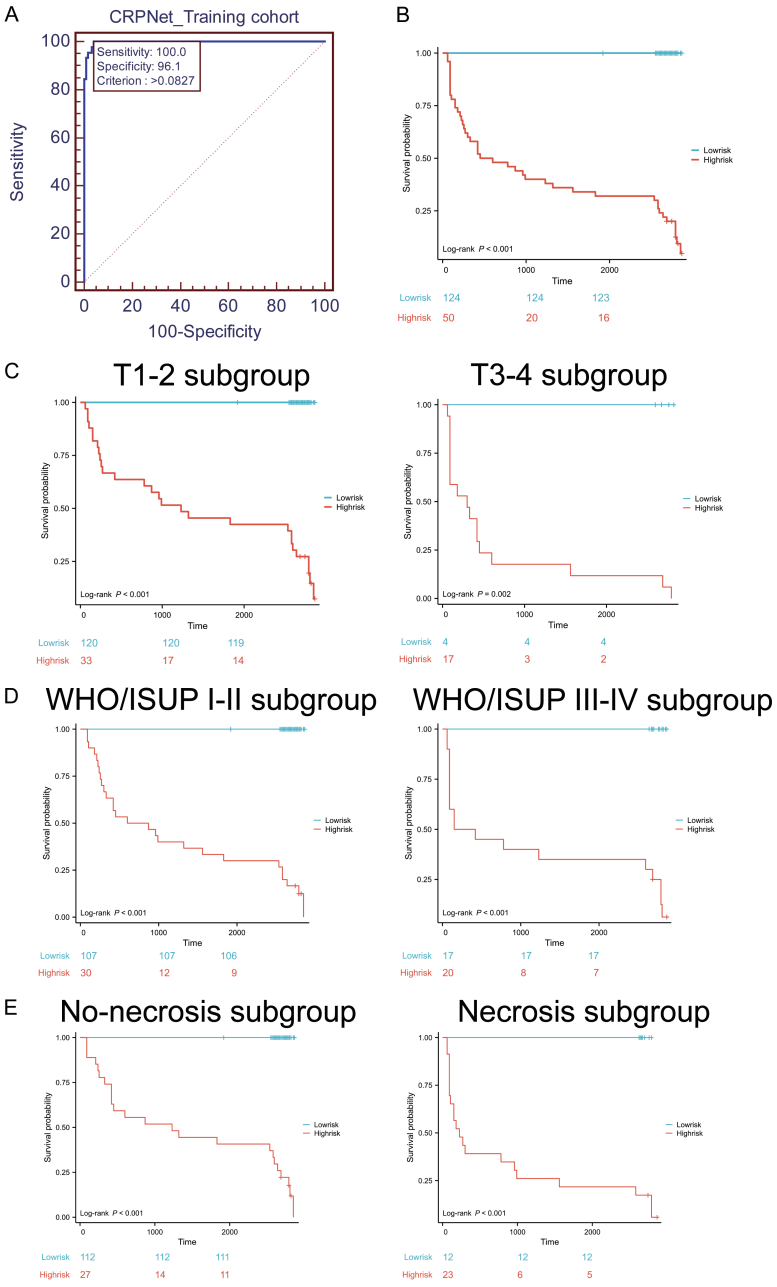


Figure 2. Kaplan-Meier survival analysis of recurrence events in training cohort using the CRPNet-predicted risk value. The CRPNet grade was calculated by dividing the CRPNet-predicted risk value into two groups using median of predicted risk value. A. ROC curve of the training cohort for CRPNet. B. Survival analysis of Lowrisk group and highrisk group. C. Survival analysis of pathological T1-2 stage and pathological T3-4 stage subgroup. D. Survival analysis of WHO I-II stage and WHO III-IV subgroup. E. Survival analysis of no-necrosis and necrosis subgroup.

highest AUC (0.786, 95% CI 0.676-0.872) among the three commonly used prognostic models, with higher accuracy (82.67%) and precision (75.00%) compared to the other two models. Furthermore, it exhibited the lowest FPR (5.36%), and its FNR (52.63%) was only second to SSIGN score. Surprisingly, in the validation cohort, we observed a FNR of 100% for UISS, and a FPR of 0%, suggesting that UISS demonstrates outstanding predictive capability for positive cases.

In addition, we perform confusion matrix to further validate the predictive performance of CRPNet (**Figure 4A**, **4B**). In the training cohort, CRPNet exhibited a notably low number of cases where it predicted a negative outcome but the actual result was positive, with only 3 instances. This is in stark contrast to UISS, SSIGN score, and Karakiewicz nomogram, which had 24, 18, and 22 such cases, respectively. On the other hand, CRPNet had only 2 cases where it predicted a positive outcome but the actual result was negative, a performance that was comparable to UISS, SSIGN score, and Karakiewicz nomogram. In the validation cohort, CRPNet again showed fewer cases of incorrect predictions than UISS, SSIGN score, and Karakiewicz nomogram, with only 7 positive instances being predicted a negative outcome. Similarly, there was 2 negative cases which CRPNet predicted a positive outcome, similar to the performance of other prediction models. Overall, CRPNet demonstrated a comparable FPR while a significantly lower FNR than UISS, SSIGN score, and Karakiewicz nomogram.

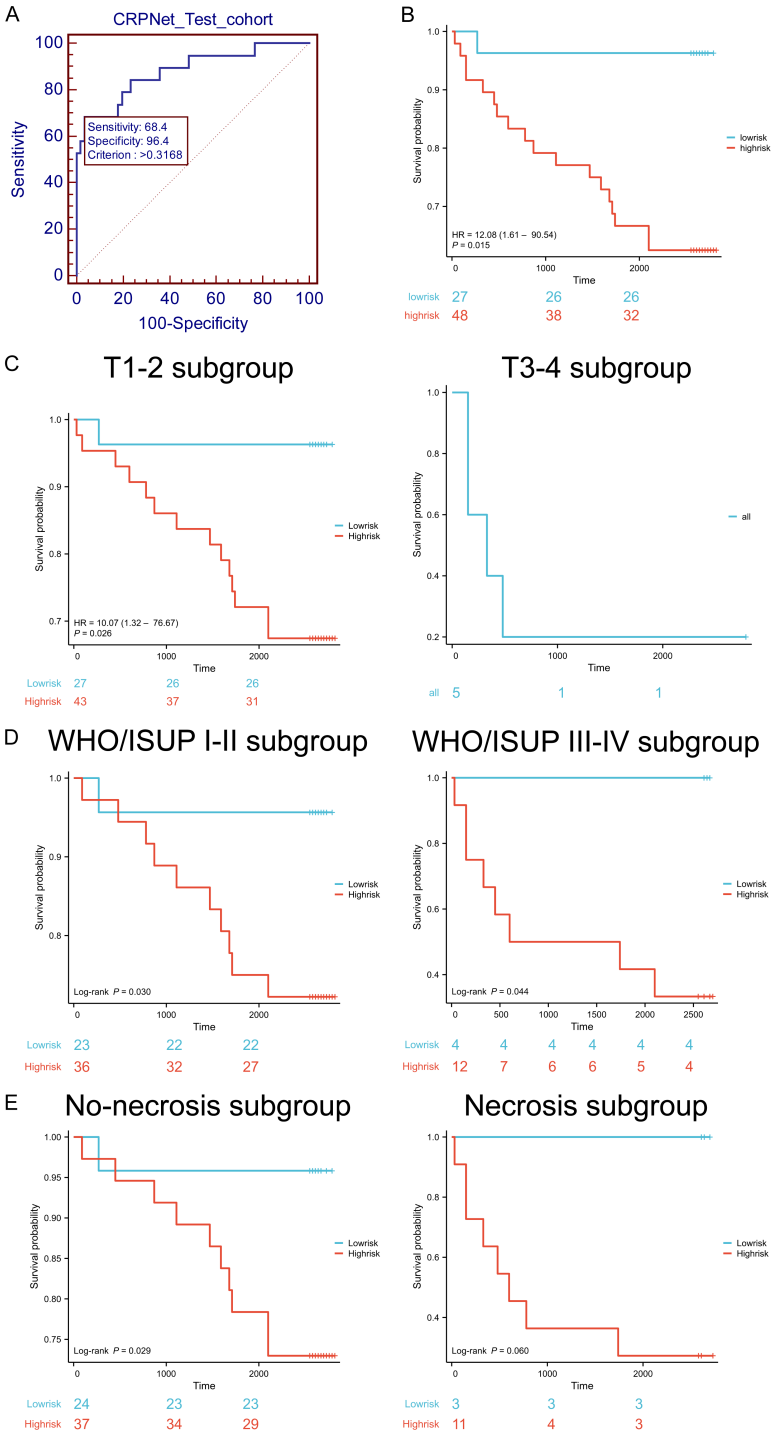


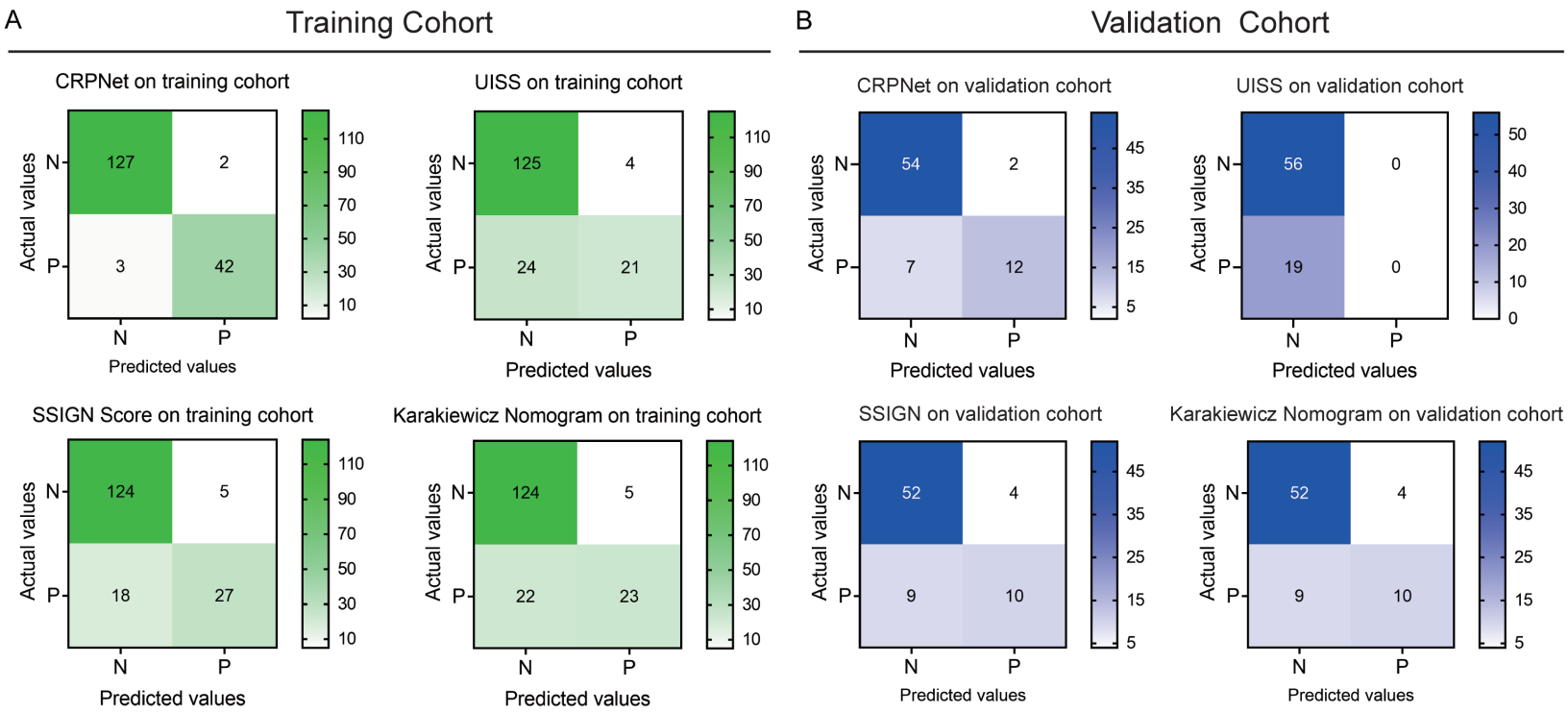
Figure 3. Kaplan-Meier survival analysis of recurrence events in validation cohort using the CRPNet-predicted risk value. The CRPNet grade was calculated by dividing the CRPNet-predicted risk value into two groups using median of predicted risk value. A. ROC curve of the validation cohort for CRPNet. B. Survival analysis of Lowrisk group and highrisk group. C. Survival analysis of pathological T1-2 stage and pathological T3-4 stage subgroup. D. Survival analysis of WHO I-II stage and WHO III-IV subgroup. E. Survival analysis of no-necrosis and necrosis subgroup.

The C-index, a key indicator for assessing predictive accuracy of prediction models, enables researchers to comprehend the model performance in guiding clinical decision-making. Thus, we conduct a C-index comparison for the T stage, WHO/ISUP grade, necrosis status, UISS, SSIGN score, Karakiewicz nomogram, and CRPNet in both training and validation cohort (**Figure 4C**, **4D**). In the training cohort, CRPNet exhibited the highest C-index of 0.937 (95% CI 0.926-0.949), followed by SSIGN score, T stage, Karakiewicz nomogram, UISS, necrosis status, and WHO/ISUP grade. Similarly, in the validation cohort, CRPNet demonstrated a C-index of 0.821 (95% CI 0.775-0.867), followed by Karakiewicz nomogram, T stage, SSIGN score, necrosis status, WHO/ISUP grade, and UISS. In summary, the C-index of CRPNet is significantly higher than that of T stage, WHO/ISUP grading, and other commonly used prognostic models.

Subsequently, we calculate the HR for T stage, WHO/ISUP grade, age, gender, BMI, UISS, SSIGN score, Karakiewicz nomogram, necrosis, and CRPNet, and presented them in a forest plot (**Figure 4E**). As shown in figure, the HRs of T stage (HR 9.170, 95% CI 2.929-28.711), WHO/ISUP grade (HR 3.513, 95% CI 1.410-8.750), SSIGN score (HR 3.689, 95% CI 1.327-10.256), Karakiewicz nomogram (HR 3.456, 95% CI 1.247-9.633), necrosis (HR 4.727, 95% CI 1.891-11.820), and CRPNet (HR 12.078,

Table 1. Performance metrics of CRPNet, UISS, SSIGN score, and Karakiewicz nomogram in predicting recurrence events in training and validation cohorts

	Training cohort				Validation cohort			
	CRPNet	UISS	SSIGN	Karakiewicz	CRPNet	UISS	SSIGN	Karakiewicz
AUC (95% CI)	0.994 (0.974-1.000)	0.785 (0.716-0.843)	0.847 (0.784-0.897)	0.799 (0.731-0.856)	0.879 (0.783-0.943)	0.563 (0.444-0.678)	0.767 (0.655-0.857)	0.786 (0.676-0.872)
Accuracy	97.70%	83.91%	86.78%	84.48%	88.00%	74.67%	82.67%	82.67%
Precision	95.56%	84.00%	84.38%	82.14%	85.71%	0	71.43%	75.00%
Recall	95.56%	46.67%	60.00%	51.11%	63.16%	0	52.63%	47.37%
False-positive rate (FPR)	1.55%	3.10%	3.88%	3.88%	3.57%	0	7.14%	5.36%
False-negative rate (FNR)	4.45%	53.33%	40.00%	48.89%	36.84%	100.00%	47.37%	52.63%



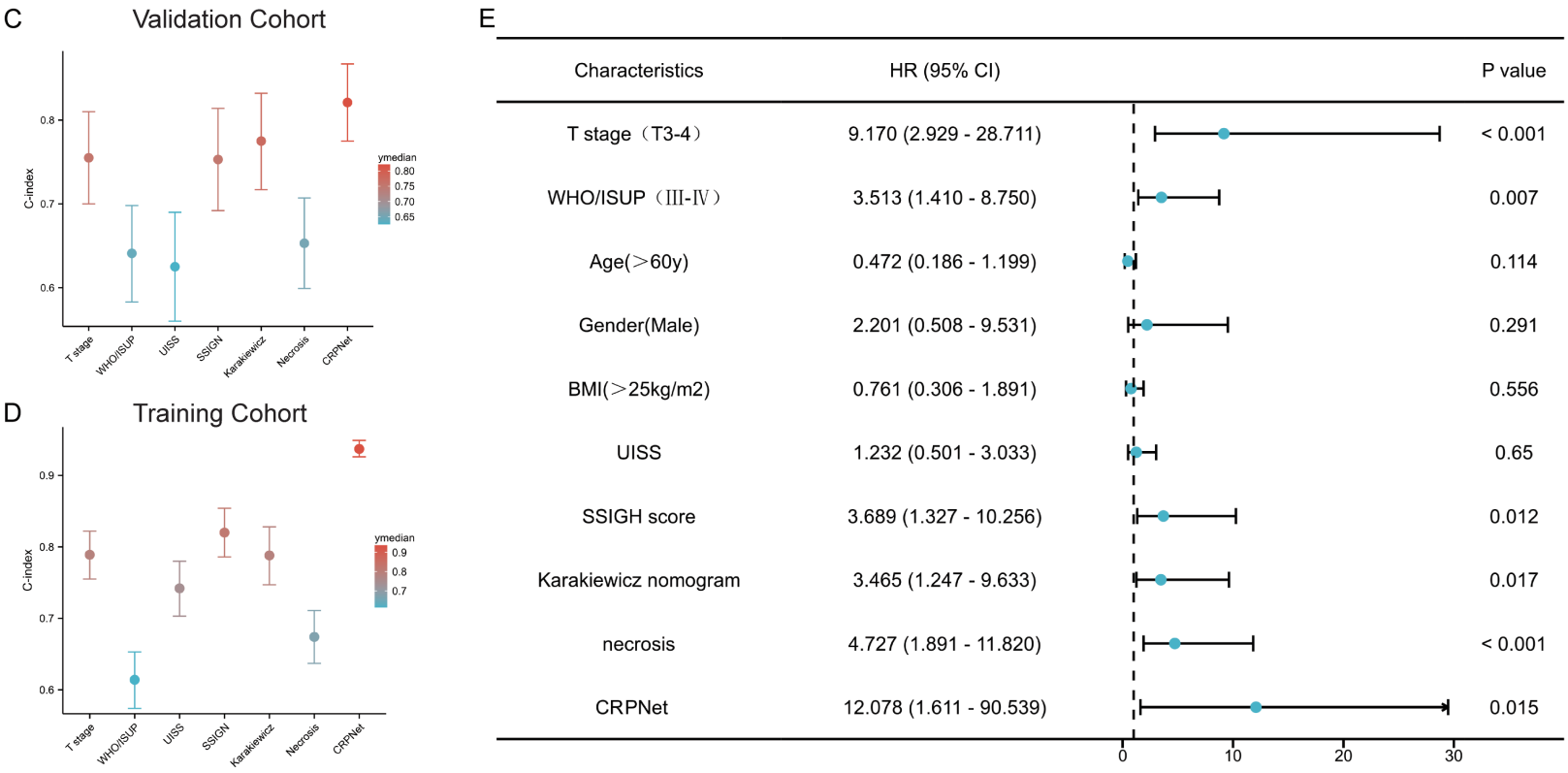


Figure 4. Comparison of CRPNet performance between the other prognostic parameters and models. (A, B) Confusion matrices of CRPNet, UISS, SSIGN score, and Karakiewicz nomogram for recurrence events in training and validation cohorts, (C, D) The C-index of various prognostic parameters with CRPNet-predicted risk value for ccRCC recurrence in training (D) and validation cohort (C), (E) Log hazard ratio (HR) of various prognostic parameters with CRPNet-predicted risk value in relation to recurrence events in validation cohort.

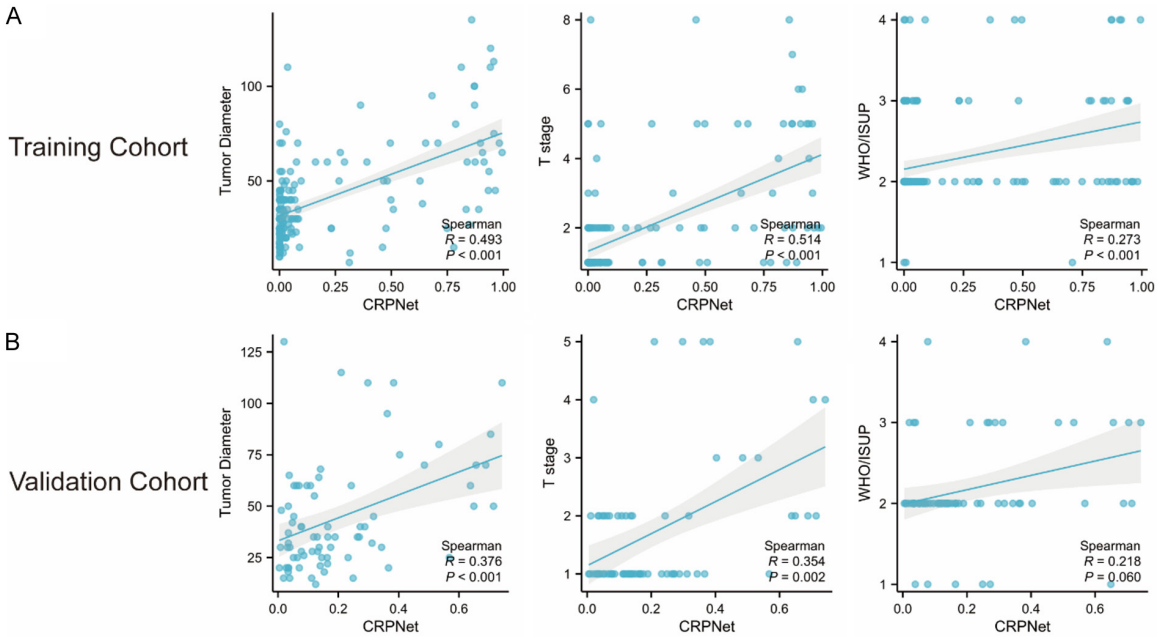


Figure 5. Correlation analysis of CRPNet with WHO/ISUP grade, T stage, and tumor diameter. A. Correlation analysis in the training cohort. B. Correlation analysis in the validation cohort.

95% CI 1.611-90.539) are statistically significant. Additionally, the CRPNet achieved the highest HR, which showed that it better reflected each patient's occurrence probability of recurrence event compared with the other prognostic parameters.

Pattern analysis of ccRCC recurrence status

We performed correlation analysis between the key deep learning pathomics signature (CRPNet score) and established clinicopathological factors (including WHO/ISUP grade, T stage and tumor diameter). The results are presented in **Figure 5**. We found that CRPNet showed significant but moderate correlations with these factors, confirming it captures related yet distinct information. Specifically, in the training cohort, the CRPNet score showed positive correlations with the WHO/ISUP grade, T stage, and tumor diameter of ccRCC, suggesting that CRPNet may have captured these established high-risk factors. Similarly, in the valiadation cohort, the CRPNet score was positively correlated with T stage and tumor diameter, while its correlation with WHO/ISUP grade, although positive, was not statistically significant, which is likely attributable to the smaller sample size of the validation cohort. Therefore, we conducted a detailed slide review of the HE images from both cohorts for further analysis.

In the light microscope, the tumor cells of ccRCC are arranged in compact nests, sheets, alveolar, or acinar structures. In addition, those cells show rounded, polygonal and nucleus round or oval. We checked whether CRPNet reflected the local features related to prognosis by checking the superpatches in the high and low score groups. In the subgroups with low CRPNet scores, cells often show a low-grade morphology (**Figure 6A**, **6B**). Tumor cells have clear cytoplasm due to loss of cytoplasmic lipid and glycogen during tissue processing and slide preparation. However, in the subgroups with high CRPNet scores, cells acquire granular eosinophilic cytoplasm (**Figure 6C-F**). According to the CRPNet, this is a similar local feature shared by ccRCC patients with different prognoses. It is worth noting that, in the subgroup with high CRPNet scores and recurrence, cells show strongly eosinophilic and presented to be long-spindle (**Figure 6D-F**). This group also shows high degree of atypia, and poorly differentiated tumors.

Discussion

RCC outcomes, treatment options, and surveillance regimens are highly dependent upon accurate staging, including the determination of nuclear grade and histologic subtype [18]. Other prognostic models for predicting RCC

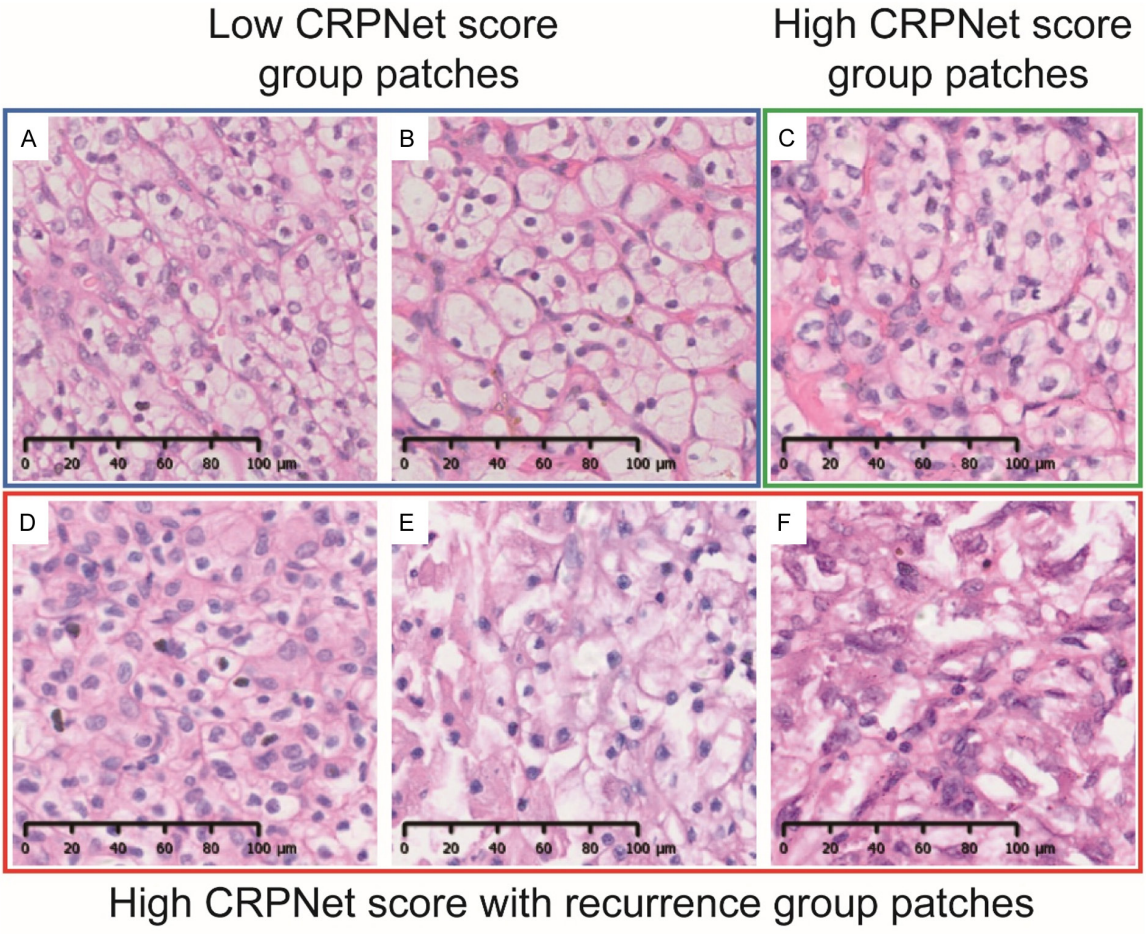


Figure 6. Visualization of high- and low-risk pathological patterns suggestive of recurrence (20×). (A, B) Clear cell renal cell carcinoma with low AI scores, (C) Clear cell renal cell carcinoma with high AI scores but without recurrence, (D-F) Clear cell renal cell carcinoma with high AI scores and recurrence.

recurrence after surgery such as UISS, SSIGN score, Leibovich, which also include nuclear grade as one parameter. However, the histopathologic analysis of RCC is subject to the poor interobserver reproducibility [19].

With the development of artificial intelligence, deep learning techniques have been intended to be applied in the analysis of pathological images. The benefit of deep learning in pathomics lies in its capacity to reveal nuanced and subtle evidence that may be easily overlooked by pathologists when examining samples with the naked eye [20]. Therefore, the introduction of deep learning technology to improve the accuracy and efficiency of pathological diagnosis has become an important research direction. Deep learning-based pathomics is capable of providing multi-parameter morphological information, playing an increasingly significant role in precision medicine [21]. Furthermore, deep learning algorithms can automatically extract complex features from images, reducing the influence of human factors and enhancing the accuracy and efficiency of diagnosis [22]. By automating the analysis of pathological images, deep learning algorithms can alleviate the workload of doctors, enabling them to focus more on complex cases and clinical decision-making [23].

Existing machine learning models related to pathomics have found numerous applications in urological tumors. Wessels et al. [24] applied convolutional neural network to predict the lymph node involvement in prostate cancer WSIs. Gao et al. [25] reported a deep learning-based pathological prediction of lymph node metastasis for patient with RCC from primary WSIs. Chen et al. [26] reported a deep learning-

based on multi-model (WSIs, CT images, and clinical data) prediction for disease-free survival status of patients with ccRCC after surgery. Employing the rich information of WSIs, computational pathology provided insights into the tumor environment and resulted in changes in the histopathological analysis method [27]. Currently, the segmentation and classification of the different cell types in a WSI are highly accurate, and predicting the oncogenic variant, gene expression or even origin of recurrence and metastasis is possible [28, 29].

To demonstrate the advantages of CRPNet's predictive performance across dimensions such as model architecture, sample size, predictive metrics, and clinical application scenarios compared to similar studies, we selected several studies of the same type for comparison. He et al. [30] designed and validated a prediction model that employs CT radiomics and a deep learning approach to predict synchronous distant metastasis in ccRCC. Different from our study, they utilized CT imaging data rather than pathological images, and after identifying radiomic features, they applied LASSO regression to select predictive features before model construction - a methodological approach fundamentally different from our deep learning framework. Furthermore, their model achieved an AUC of 0.863, whereas CRPNet reached 0.879, indicating that the model developed in this study possesses superior predictive performance. Additionally, Gao et al. [31] obtained WSIs from formalin-fixed and paraffin-embedded tissues across three cohorts - Shanghai General Hospital (SGH), the Clinical Proteomic Tumor Analysis Consortium (CPTAC), and The Cancer Genome Atlas (TCGA) - as well as frozen-section WSIs from the TCGA dataset. Based on these WSIs, they developed a deep learning-based strategy for predicting lymph node metastasis using a cluster-constrained attention-based multiple instance learning method and validated it across the three cohorts. However, the AUCs for lymph node metastasis prediction performance in the TCGA, SGH, and CPTAC cohorts were 0.836, 0.865, and 0.812, respectively, all of which are lower than CRPNet's AUC of 0.879.

To summarize, the CRPNet model constructed in this study is specifically designed for postoperative recurrence prediction, enabling it to directly support clinical decision-making. It employs a multi-scale regional feature aggregation strategy to simultaneously capture local cellular characteristics and overall architectural patterns, integrating these features into a clinically interpretable predictive nomogram. We believe these attributes form the cornerstone of the model's core competitiveness and translational potential. However, in comparison with similar studies, the framework used in this research is relatively conventional, the sample size is limited, and an independent external validation cohort is lacking. Therefore, further validation with an expanded sample size remains a necessary next step.

Our deep learning model can not only accurately predict the postoperative recurrence status of ccRCC patients based on HE-stained pathological slides of the primary tumor, but also provide pathologists with more definitive interpretable pathological patterns. In our research, through the adaptation of traditional survival analysis, deep learning-based pathology extracted the histopathological features including strongly eosinophilic and long-spindle, but not granular eosinophilic cytoplasm, to be recurrence risk-related. It allowed pathologists to review these histopathological markers associated with prognosis. Several other histological features have shown prognostic importance. Sarcomatous degeneration, present in less than 5% of ccRCC cases, is characterized by spindled elements and automatically classifies the tumor as nuclear grade 4 [32]. The reported five-year survival rate of those with and without sarcomatoid change was 22% and 79%, respectively [33]. Tumor necrosis is another histological feature that affects prognosis. In addition, the tumor necrosis was another histological feature that affects prognosis and associated with a four to five-fold increased risk of death among ccRCC patients [34]. In this study, we performed Cox regression analysis on most established factors that may influence postoperative metastasis in ccRCC. The results showed that CRPNet exhibited the highest hazard ratio, surpassing currently widely recognized predictive factors such as T stage, WHO/ISUP grade, UISS, SSIGN, and the Karakiewicz nomogram. This indicates that CRPNet incorporates the strengths of these predictive models and integrates them. However, this also suggests that in practice, constructing a model using a multimodal approach may yield even better predictive performance, which represents a direction for future research.

Certainly, this study has some limitations. First, there were large differences in staining and scan quality between slices, which might affect the performance of our pathomics signature model. Second, the study solely focused on pathomics and did not integrate features from genetic sequencing or clinical data. As certain information that cannot be inferred from tissue images such as tumor size, the performance of CRPNet would be limited. While pathological slide selection was performed by a blinded pathologist, the deep learning model development was not conducted in a fully blinded fashion to the dataset partition. Although we employed rigorous cross-validation to mitigate overfitting, future prospective studies with a completely blinded design from slide selection to model evaluation would further strengthen the evidence. Last, as a retrospective study, the constructed model still requires prospective validation across multiple centers.

Conclusion

By analyzing the complete WSIs of primary tumors from 258 ccRCC patients, we developed a multi-instance deep convolutional neural network and trained it to predict the postoperative recurrence status of ccRCC patients. The morphologies of strongly eosinophilic and long-spindle, which were related to high CRPNet score and postoperative recurrence, would assist pathologists in classifying patients. The effective predicting performance of our pathomics signature model CRPNet provides strong support for precision medicine and personalized treatment of ccRCC patients.

Acknowledgements

This study is funded by Natural Science Foundation of Shanghai (no. 25ZR1401397), National Natural Science Foundation of China (82203134, 82372883), Shanghai Shenkang Hospital Development Center Renal Cancer COC Project (no. SHDC22025204, SHDC220-24222); Naval Medical University Sailing Program, and Changfeng Talent Program of Shanghai Changhai Hospital.

Disclosure of conflict of interest

None.

Address correspondence to: Drs. Wei Zhang, Xiaolei Shi and Linhui Wang, Department of Urology, Changhai Hospital, Naval Medical University, 168 Changhai Road, Shanghai 200433, China. E-mail: zhweismmu@163.com (WZ); neilsxl@foxmail.com (XLS); wanglinhui@smmu.edu.cn (LHW)

References

[1] Jonasch E, Gao J and Rathmell WK. Renal cell carcinoma. BMJ 2014; 349: g4797.
[2] Rini BI, Campbell SC and Escudier B. Renal cell carcinoma. Lancet 2009; 373: 1119-1132.
[3] Choueiri TK and Motzer RJ. Systemic therapy for metastatic renal-cell carcinoma. N Engl J Med 2017; 376: 354-366.
[4] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021; 71: 7-33.
[5] Roussel E, Capitanio U, Kutikov A, Oosterwijk E, Pedrosa I, Rowe SP and Gorin MA. Novel imaging methods for renal mass characterization: a collaborative review. Eur Urol 2022; 81: 476-488.
[6] Yin Q, Xu H, Zhong Y, Ni J and Hu S. Diagnostic performance of MRI, SPECT, and PET in detecting renal cell carcinoma: a systematic review and meta-analysis. BMC Cancer 2022; 22: 163.
[7] Kowalewski KF, Egen L, Fischetti CE, Puliatti S, Juan GR, Taratkin M, Ines RB, Sidoti Abate MA, Mühlbauer J, Wessels F, Checcucci E and Cacciamani G; Young Academic Urologists (YAU)-Urotechnology-Group. Artificial intelligence for renal cancer: from imaging to histology and beyond. Asian J Urol 2022; 9: 243-252.
[8] Mousavi SM and Beroza GC. Deep-learning seismology. Science 2022; 377: eabm4470.
[9] Jiang Y, Yang M, Wang S, Li X and Sun Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. Cancer Commun (Lond) 2020; 40: 154-166.
[10] van der Velden BHM, Kuijf HJ, Gilhuijs KGA and Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Med Image Anal 2022; 79: 102470.
[11] Oprea S, Martinez-Gonzalez P, Garcia-Garcia A, Castro-Vargas JA, Orts-Escolano S, Garcia-Rodriguez J and Argyros A. A review on deep learning techniques for video prediction. IEEE Trans Pattern Anal Mach Intell 2022; 44: 2806-2826.
[12] Mahootiha M, Qadir HA, Bergsland J and Balasingham I. Multimodal deep learning for personalized renal cell carcinoma prognosis: integrating CT imaging and clinical data. Comput Methods Programs Biomed 2024; 244: 107978.
[13] Kawahara D, Kishi M, Kadooka Y, Hirose K and Murakami Y. Integrating radiomics and gene

expression by mapping on the image with improved DeepInsight for clear cell renal cell carcinoma. Cancer Genet 2025; 292-293: 100-105.
[14] Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M and Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat Biomed Eng 2021; 5: 555-570.
[15] Zisman A, Pantuck AJ, Dorey F, Said JW, Shvarts O, Quintana D, Gitlitz BJ, deKernion JB, Figlin RA and Belldegrun AS. Improved prognostication of renal cell carcinoma using an integrated staging system. J Clin Oncol 2001; 19: 1649-1657.
[16] Zigeuner R, Hutterer G, Chromecki T, Imamovic A, Kampel-Kettner K, Rehak P, Langner C and Pummer K. External validation of the Mayo Clinic stage, size, grade, and necrosis (SSIGN) score for clear-cell renal cell carcinoma in a single European centre applying routine pathology. Eur Urol 2010; 57: 102-109.
[17] Chun FK, Graefen M, Briganti A, Gallina A, Hopp J, Kattan MW, Huland H and Karakiewicz PI. Initial biopsy outcome prediction--head-to-head comparison of a logistic regression-based nomogram versus artificial neural network. Eur Urol 2007; 51: 1236-1240.
[18] Delahunt B, Cheville JC, Martignoni G, Humphrey PA, Magi-Galluzzi C, McKenney J, Egevad L, Algaba F, Moch H, Grignon DJ, Montironi R and Srigley JR; Members of the ISUP Renal Tumor Panel. The international society of urological pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters. Am J Surg Pathol 2013; 37: 1490-1504.
[19] Fenstermaker M, Tomlins SA, Singh K, Wiens J and Morgan TM. Development and validation of a deep-learning model to assist with renal cell carcinoma histopathologic interpretation. Urology 2020; 144: 152-157.
[20] Niikura H, Okamoto S, Yoshinaga K, Nagase S, Takano T, Ito K and Yaegashi N. Detection of micrometastases in the sentinel lymph nodes of patients with endometrial cancer. Gynecol Oncol 2007; 105: 683-686.
[21] Santo H, Samejima M, Sugano Y, Shi B and Matsushita Y. Deep photometric stereo networks for determining surface normal and reflectances. IEEE Trans Pattern Anal Mach Intell 2022; 44: 114-128.
[22] Osawa K, Tsuji Y, Ueno Y, Naruse A, Foo CS and Yokota R. Scalable and practical natural gradient for large-scale deep learning. IEEE Trans Pattern Anal Mach Intell 2022; 44: 404-415.
[23] Kiyasseh D, Zhu T and Clifton D. The promise of clinical decision support systems targetting low-resource settings. IEEE Rev Biomed Eng 2022; 15: 354-371.
[24] Wessels F, Schmitt M, Krieghoff-Henning E, Jutzi T, Worst TS, Waldbillig F, Neuberger M, Maron RC, Steeg M, Gaiser T, Hekler A, Utikal JS, von Kalle C, Fröhling S, Michel MS, Nuhn P and Brinker TJ. Deep learning approach to predict lymph node metastasis directly from primary tumour histology in prostate cancer. BJU Int 2021; 128: 352-360.
[25] Gao F, Jiang L, Guo T, Lin J, Xu W, Yuan L, Han Y, Yang J, Pan Q, Chen E, Zhang N, Chen S and Wang X. Deep learning-based pathological prediction of lymph node metastasis for patient with renal cell carcinoma from primary whole slide images. J Transl Med 2024; 22: 568.
[26] Chen S, Gao F, Guo T, Jiang L, Zhang N, Wang X and Zheng J. Deep learning-based multi-model prediction for disease-free survival status of patients with clear cell renal cell carcinoma after surgery: a multicenter cohort study. Int J Surg 2024; 110: 2970-2977.
[27] Lee Y, Park JH, Oh S, Shin K, Sun J, Jung M, Lee C, Kim H, Chung JH, Moon KC and Kwon S. Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning. Nat Biomed Eng 2022; 6: 1452-1466.
[28] Lu MY, Chen TY, Williamson DFK, Zhao M, Shady M, Lipkova J and Mahmood F. AI-based pathology predicts origins for cancers of unknown primary. Nature 2021; 594: 106-110.
[29] Campanella G, Hanna MG, Geneslaw L, Miraflor A, Werneck Krauss Silva V, Busam KJ, Brogi E, Reuter VE, Klimstra DS and Fuchs TJ. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat Med 2019; 25: 1301-1309.
[30] He WB, Zhou C, Yang ZJ, Zhang YF, Zhang WB, He H, Wang J and Zhou FH. The predictive value of radiomics and deep learning for synchronous distant metastasis in clear cell renal cell carcinoma. Discov Oncol 2025; 16: 86.
[31] Gao F, Jiang L, Guo T, Lin J, Xu W, Yuan L, Han Y, Yang J, Pan Q, Chen E, Zhang N, Chen S and Wang X. Deep learning-based pathological prediction of lymph node metastasis for patient with renal cell carcinoma from primary whole slide images. J Transl Med 2024; 22: 568.
[32] Lam JS, Klatte T, Kim HL, Patard JJ, Breda A, Zisman A, Pantuck AJ and Figlin RA. Prognostic factors and selection for clinical studies of patients with kidney cancer. Crit Rev Oncol Hematol 2008; 65: 235-262.

[33] de Peralta-Venturina M, Moch H, Amin M, Tamboli P, Hailemariam S, Mihatsch M, Javidan J, Stricker H, Ro JY and Amin MB. Sarcomatoid differentiation in renal cell carcinoma: a study of 101 cases. Am J Surg Pathol 2001; 25: 275-284.

[34] Sengupta S, Lohse CM, Leibovich BC, Frank I, Thompson RH, Webster WS, Zincke H, Blute ML, Cheville JC and Kwon ED. Histologic coagulative tumor necrosis as a prognostic indicator of renal cell carcinoma aggressiveness. Cancer 2005; 104: 511-520.

Supplementary Table 1. Baseline of Training cohort

Characteristics	Non-metastasis	Metastasis	*P* value
n	135	48	
Age, mean ± sd	61.37 ± 13.335	64.042 ± 9.0834	0.128
Gender, n (%)			0.061
Male	97 (53%)	41 (22.4%)	
Female	38 (20.8%)	7 (3.8%)	
BMI (kg/m^2), median (IQR)	24.221 (22.258, 26.724)	24.433 (22.885, 26.618)	0.980
Smoking History, n (%)			1.000
NO	122 (66.7%)	43 (23.5%)	
YES	13 (7.1%)	5 (2.7%)	
Drinking History, n (%)			0.928
NO	127 (69.4%)	46 (25.1%)	
YES	8 (4.4%)	2 (1.1%)	
Hypertension, n (%)			0.177
NO	72 (39.3%)	31 (16.9%)	
YES	63 (34.4%)	17 (9.3%)	
Coronary Heart Disease, n (%)			0.145
NO	132 (72.1%)	44 (24%)	
YES	3 (1.6%)	4 (2.2%)	
Diabetes, n (%)			0.294
NO	118 (64.5%)	39 (21.3%)	
YES	17 (9.3%)	9 (4.9%)	
Pathological T stage, n (%)			< 0.001
T1a	100 (54.6%)	11 (6%)	
T2a	3 (1.6%)	4 (2.2%)	
T3a	3 (1.6%)	12 (6.6%)	
T1b	27 (14.8%)	14 (7.7%)	
T4	1 (0.5%)	2 (1.1%)	
T2b	1 (0.5%)	2 (1.1%)	
T3c	0 (0%)	1 (0.5%)	
T3b	0 (0%)	2 (1.1%)	
Pathological N stage, n (%)			< 0.001
N0	135 (73.8%)	42 (23%)	
N1	0 (0%)	6 (3.3%)	
Tumor side, n (%)			0.350
Left	71 (38.8%)	29 (15.8%)	
Right	64 (35%)	19 (10.4%)	
Tumor diameter (mm), median (IQR)	30 (23.5, 40)	60 (35, 71.25)	< 0.001
WHO/ISUP stage, n (%)			< 0.001
Stage II	113 (61.7%)	27 (14.8%)	
Stage I	3 (1.6%)	1 (0.5%)	
Stage III	14 (7.7%)	12 (6.6%)	
Stage IV	5 (2.7%)	8 (4.4%)	

Supplementary Table 2. Baseline of Validation cohort

Characteristics	Non-metastasis	Metastasis	*P* value
n	56	19	
Age, mean ± sd	59.036 ± 11.954	59.053 ± 5.2965	0.993
Gender, n (%)			0.388
Male	43 (57.3%)	17 (22.7%)	
Female	13 (17.3%)	2 (2.7%)	
BMI (kg/m^2), mean ± sd	24.94 ± 3.1616	24.514 ± 2.6147	0.599
Smoking History, n (%)			0.205
NO	52 (69.3%)	15 (20%)	
YES	4 (5.3%)	4 (5.3%)	
Drinking History, n (%)			0.507
NO	52 (69.3%)	16 (21.3%)	
YES	4 (5.3%)	3 (4%)	
Hypertension, n (%)			0.318
YES	25 (33.3%)	6 (8%)	
NO	31 (41.3%)	13 (17.3%)	
Coronary Heart Disease, n (%)			0.253
NO	56 (74.7%)	18 (24%)	
YES	0 (0%)	1 (1.3%)	
Diabetes, n (%)			1.000
NO	52 (69.3%)	18 (24%)	
YES	4 (5.3%)	1 (1.3%)	
Pathological T stage, n (%)			< 0.001
T1a	41 (54.7%)	5 (6.7%)	
T1b	12 (16%)	6 (8%)	
T2a	1 (1.3%)	2 (2.7%)	
T3a	1 (1.3%)	4 (5.3%)	
T2b	1 (1.3%)	2 (2.7%)	
Pathological N stage, n (%)			1.000
N0	55 (73.3%)	19 (25.3%)	
N1	1 (1.3%)	0 (0%)	
Tumor side, n (%)			0.318
Right	25 (33.3%)	11 (14.7%)	
Left	31 (41.3%)	8 (10.7%)	
Tumor diameter (mm), median (IQR)	32 (24.25, 45)	65 (45, 82.5)	< 0.001
WHO/ISUP stage, n (%)			0.069
Stage II	43 (57.3%)	10 (13.3%)	
Stage IV	1 (1.3%)	2 (2.7%)	
Stage III	7 (9.3%)	6 (8%)	
Stage I	5 (6.7%)	1 (1.3%)	