Article

Accurate Diagnosis and Survival Prediction of Bladder Cancer Using Deep Learning on Histological Slides

Qingyuan Zheng [1,2,†] , Rui Yang [1,2,†], Xinmiao Ni [1,2], Song Yang [1,2], Lin Xiong [3], Dandan Yan [3], Lingli Xia [3], Jingping Yuan [3], Jingsong Wang [1,2], Panpan Jiao [1,2], Jiejun Wu [1,2], Yiqun Hao [4], Jianguo Wang [5], Liantao Guo [6], Zhengyu Jiang [1,2], Lei Wang [1,2], Zhiyuan Chen [1,2,*] and Xiuheng Liu [1,2,*]

1 Department of Urology, Renmin Hospital of Wuhan University, Wuhan 430060, China
2 Institute of Urologic Disease, Renmin Hospital of Wuhan University, Wuhan 430060, China
3 Department of Pathology, Renmin Hospital of Wuhan University, Wuhan 430060, China
4 Division of Nephrology, Renmin Hospital of Wuhan University, Wuhan 430060, China
5 Department of Hepatic-Biliary-Pancreatic Surgery, Renmin Hospital of Wuhan University, Wuhan 430060, China
6 Department of Breast and Thyroid Surgery, Renmin Hospital of Wuhan University, Wuhan 430060, China
* Correspondence: chenzhiyuan163@163.com (Z.C.); drxhliu@163.com (X.L.)
† These authors contributed equally to this work.

Simple Summary: Early diagnosis and treatment are essential to reduce the mortality rate of bladder cancer. However, current techniques of diagnosis are susceptible to pathologist variability, and histopathological prognostic methods are insufficient to cover all features of muscle-invasive bladder cancer. In this work, we developed weakly supervised models based on deep learning for the diagnosis of bladder cancer and prediction of overall survival in muscle-invasive bladder cancer patients using whole slide digitized histological images in two cohorts. Encouragingly, results showed that our models can not only assist clinicians in the accurate diagnosis of bladder cancer, but also facilitate differential risk stratification in patients with muscle-invasive bladder cancer and improve personalized treatment decisions accordingly. Furthermore, the regions most relevant for diagnosis or prognosis can be further analyzed to increase the amount of information extracted from pathological images. Finally, we identified six genes closely related to cancer progression based on the predicted risk scores, which potentially led to new biomarker discoveries.

Abstract: (1) Background: Early diagnosis and treatment are essential to reduce the mortality rate of bladder cancer (BLCA). We aimed to develop deep learning (DL)-based weakly supervised models for the diagnosis of BLCA and prediction of overall survival (OS) in muscle-invasive bladder cancer (MIBC) patients using whole slide digitized histological images (WSIs). (2) Methods: Diagnostic and prognostic models were developed using 926 WSIs of 412 BLCA patients from The Cancer Genome Atlas cohort. We collected 250 WSIs of 150 BLCA patients from the Renmin Hospital of Wuhan University cohort for external validation of the models. Two DL models were developed: a BLCA diagnostic model (named BlcaMIL) and an MIBC prognostic model (named MibcMLP). (3) Results: The BlcaMIL model identified BLCA with accuracy 0.987 in the external validation set, comparable to that of expert uropathologists and outperforming a junior pathologist. The C-index values for the MibcMLP model on the internal and external validation sets were 0.631 and 0.622, respectively. The risk score predicted by MibcMLP was a strong predictor independent of existing clinical or histopathologic indicators, as demonstrated by univariate Cox (HR = 2.390, $p < 0.0001$) and multivariate Cox (HR = 2.414, $p < 0.0001$) analyses. The interpretability of DL models can help in the analysis of critical regions associated with tumors to enrich the information obtained from WSIs. Furthermore, the expression of six genes (ANAPC7, MAPKAPK5, COX19, LINC01106, AL161431.1 and MYO16-AS1) was significantly associated with MibcMLP-predicted risk scores, revealing possible potential biological correlations. (4) Conclusions: Our study developed DL models for accurately diagnosing BLCA and predicting OS in MIBC patients, which will help promote the precise pathological diagnosis of BLCA and risk stratification of MIBC to improve clinical treatment decisions.

1. Introduction

Bladder cancer (BLCA) is one of the most common tumors worldwide, with approximately 573,000 new cases and 213,000 deaths in 2020 [1]. According to the IDENTIFY study, the largest international prospective study of patients with suspected urinary tract cancer, BLCA is the most prevalent cancer diagnosis in patients with hematuria [2]. The five-year survival rate for patients with non-muscle-invasive bladder cancer (NMIBC) is estimated to be around 90%, but the five-year survival rate for patients with muscle-invasive bladder cancer (MIBC) decreases dramatically as the tumor invades different layers of the bladder [3]. Recently, the IDENTIFY study presents a multivariable prediction model for the detection of urinary tract cancers, which assists clinicians in early risk assessment, but only in patients with hematuria [4]. Despite the advances in surgery and other diagnosis and treatment techniques over the past 30 years, clinical outcomes of BLCA have not substantially improved [5]. Given that early detection, accurate diagnosis, and appropriate therapeutic intervention are critical for reducing the mortality of BLCA, precise and consistently effective methods of pathologic assessment are essential.

Currently, the diagnosis of BLCA is carried out by pathologists through biopsy, which serves as the gold standard. This typically requires pathologists to manually review each pathological slide and rely on personal expertise to make an accurate diagnosis. Such manual analysis is not only time-consuming, labor-intensive, and tedious, but is also subject to observer variability. Moreover, the shortage of expert pathologists has become a global problem, which, to a certain extent, causes an increase in the workload of available pathologists [6]. Hence, it is necessary to develop a convenient, efficient, and accurate method to diagnose BLCA using pathological slide images.

The pathological type of BLCA can be low-grade or high-grade; high-grade BLCA should be treated more aggressively and is more likely to result in death [7]. It turns out that in most human cancers, including BLCA, prognosis is closely related to pathological criteria [8]. Histopathological classification through the tumor-node-metastasis (TNM) staging system jointly developed and established by the American Joint Committee on Cancer (AJCC) and the Union International Committee on Cancer has some prognostic and therapeutic value, but is insufficient to cover all clinical characteristics of BLCA patients and the heterogeneity of patient outcomes [9]. Furthermore, risk stratification relying on histopathological staging is susceptible to variability in observation and judgment among pathologists. Accordingly, there is an urgent need to develop robust and reproducible methods to identify predictive markers consistently associated with survival in MIBC patients.

In recent years, the use of artificial intelligence has been greatly beneficial in pathology and has tremendously facilitated the growth of digital pathology. The advent of deep learning (DL) and the availability of thousands of digitized whole slide images (WSIs) may provide new opportunities for the diagnosis and prediction of disease outcomes [10]. DL can adaptively extract relevant image features according to learning objectives for tasks such as classification, segmentation, and detection [11–13]. It has been reported that an algorithm based on DL can identify bladder tumors with a specificity of up to 98.6% [14]. A previous study [15] used DL to successfully predict the molecular subtypes of MIBC by processing hematoxylin and eosin (H&E) slides with a similar or superior performance compared to that of pathologists.

There is growing evidence [16–19] indicating that automated analysis of WSIs can improve disease diagnosis and prognosis prediction, thereby enhancing treatment options and maximizing efficacy. To solve the time-consuming and laborious problem of manual annotation, unsupervised or weakly supervised DL models are gaining popularity.

Courtiol et al. [20,21] used DL to develop models that could accurately predict patient survival in hepatocellular carcinoma and mesothelioma, respectively, without the need for local annotated regions provided by any pathologist. Lu et al. [22] reported an interpretable weakly supervised DL method for binary and multi-class WSI classification using only slide-level labels without any additional manual annotations. Furthermore, DL has been shown to predict the expression of differential genes or molecular biomarkers from pathological images of BLCA, such as PD-L1 [23] and FGFR [24,25], which is cheaper, more effective, and more robust than next-generation sequencing or immunohistochemical staining methods. Therefore, a possible solution is to utilize DL to extract potential clinical and/or biological information in WSIs for diagnosis and prediction of overall survival (OS) in MIBC patients.

In this study, we developed a DL-based diagnostic model for BLCA patients and a prognostic model for MIBC patients, named BlcaMIL and MibcMLP, respectively, and demonstrated their effectiveness in tumor diagnosis and prognosis prediction on two cohorts. The results showed that BlcaMIL can accurately distinguish between tumor and normal tissues, and MibcMLP is more accurate than the use of most clinical information and pathological features in predicting the OS in MIBC patients. By visualizing the region of interest (ROI), it is possible to gain insights into the most relevant features of DL-models for diagnosing and predicting patient outcomes.

2. Materials and Methods

2.1. Patient Cohorts

We retrospectively analyzed two cohorts of BLCA patients for this study. The first cohort was from The Cancer Genome Atlas (TCGA) public dataset consisting of a total of 926 H&E-stained WSIs from 412 BLCA patients, of which 887 were tumor and 39 were normal. Given that the uneven distribution of tumor and normal images in TCGA cohort, the data augmentation technique was used to address this imbalance issue. An independent external dataset was obtained from the Renmin Hospital of Wuhan University (RHWU; Wuhan, Hubei, China) comprising 250 H&E-stained WSIs obtained from 150 BLCA patients from 2017 to 2022, of which 150 were tumor and 100 were normal.

For the development of diagnostic model, the inclusion criteria for both cohorts were (a) specific pathological diagnosis of BLCA and (b) availability of clear H&E-stained pathological slides.

The TCGA cohort provides two types of H&E-stained WSIs: tissue slides and diagnostic slides. Tissue slides are sections of frozen tumor specimens that are often used to determine whether tumor boundaries are clean. Diagnostic slides are formalin-fixed paraffin-embedded sections, which generally preserve cell morphology better and are more situable for computational analysis.

We adopted the following inclusion criteria for developing the prognostic model once the criteria for the diagnostic model were met: (a) availability of clinicopathological information, (b) availability of follow-up data, (c) specific pathological diagnosis of MIBC and (d) use of diagnostic slides rather than tissue slides.

In addition, we collected the corresponding clinical data as well as biological and pathological characteristics of the patients (including age, gender, lymphovascular invasion, tumor size, OS, survival status, pathological grade and TNM staging (according to AJCC 8th Edition Staging Manual [26]) for survival analysis of the prognostic model. The patient data for the TCGA cohort were collected through the UCSC Xena database (https://xenabrowser.net/datapages/, accessed on 22 October 2022), and patient data for the RHWU cohort were obtained through the hospital information management system.

2.2. WSI Preprocessing

The WSIs from the two cohorts had different magnifications. Specifically, WSIs from the TCGA cohort had an original magnification of 40× (without fixed size, the image size could be larger than 100,000 × 100,000 pixels), whereas those from the RHWU cohort

had an original magnification of 20× (64,000 × 58,000 pixels). Therefore, we uniformly processed these WSIs to 20× magnification and used the resulting WSIs in the next step.

Since the WSIs contained extremely rich content (up to a resolution of 100,000 × 100,000 pixels) and could not be directly used for model training, they were pre-processed first. We loaded WSIs at 10× magnification (0.5 µm/pixel), segmented out tissue regions using an area threshold filtering-based approach, and then used the openslide-python toolkit (https://openslide.org/, accessed on 22 October 2022) to divide WSIs into small images of the same fixed-out size (448 × 448 pixels), each of which is called a patch. Each patch had to contain 80% of the tissues to be included, and had the same label as the pathological diagnosis of the WSI itself. A color threshold was used to exclude possible background images from patches. Due to the heterogeneity in staining protocols used to obtain the WSIs by different centers, we unified the colors of all patches using the structure-preserving color normalization method proposed by Vahadane [27] and Anand [28].

2.3. Feature Extraction and Reduction

We used Resnet-50 to extract 2048 relevant features for each patch. The network was pre-trained on the ImageNet dataset (over one million images) and had been shown to accurately identify over a thousand categories [29]. At that point, we had constructed a 2000 × 2048 vector for each WSI. Because the data dimensions were too high and prone to overfitting during model training, we used the trained autoencoder for dimensionality reduction. This autoencoder included a hidden layer of 512 neurons, which reduced the input dimension of the prediction part from 2048 to 512 to avoid potential problems of overfitting, long processing time, and heavy usage of computational memory. We randomly selected 200 patches (66,000 patches in total) from each WSI to train the autoencoder, and the MSE loss was reduced to 0.0052 after 100 epochs.

2.4. Development of Diagnostic Model

BlcaMIL is an end-to-end weakly supervised neural network, which combines the attention strategy with multiple instance learning (MIL) algorithm, and can be used for the task of binary classification of the entire WSI. The theoretical basis of the MIL algorithm is that, assuming a WSI is a bag, then all patches it contained are instances of the bag. When a WSI is marked positive, at least one patch is positive; when a WSI is marked negative, then all patches are negative. Based on the assumptions of the MIL algorithm, which had been widely used for weakly supervised positive/negative binary classification, we added an additional attention mechanism to it.

During training and inference, the attention network in the BlcaMIL assigned an attention score to each patch, representing its importance to the overall WSI classification. We input the extracted patch-level features into the BlcaMIL, and aggregated them into a WSI-level representation through an average pooling function, which was used to predict the probability score for the final diagnosis.

2.5. Development of Prognostic Model

MibcMLP is also a weakly supervised neural network consisting of a two-step algorithm. To generate a risk score for each patch, we used a one-dimensional dense layer. The dense layer was composed of 512 neurons. The 512 features extracted from each patch were weighted and summed (the weights were obtained after model learning), and the risk score was calculated. Subsequently, we sorted the risk scores of the 2000 patches, selected the 25 highest and 25 lowest scoring patches, formed a vector of size 50, and used it as the input for the final step. This operation allowed us to clearly understand which patches were finally used as input for the risk prediction step, thereby facilitating subsequent interpretation of the ROIs for the DL-model.

The final step was a multilayer perceptron consisting of two fully connected layers with 200 and 100 neurons, each with a sigmoid activation function. This was a critical step

in predicting the prognosis of MIBC patients, and its function was to convert the 50-patch risk scores into a survival-related risk score that was representative of the entire WSI.

The loss functions for the diagnostic and prognostic models used smooth top1 SVM loss and the negative log-likelihood function of the Cox proportional hazards model, respectively. The training set was repeatedly validated using a five-fold cross-validation strategy, using internal and external validation. The layouts of the two DL models are shown in Figure 1.
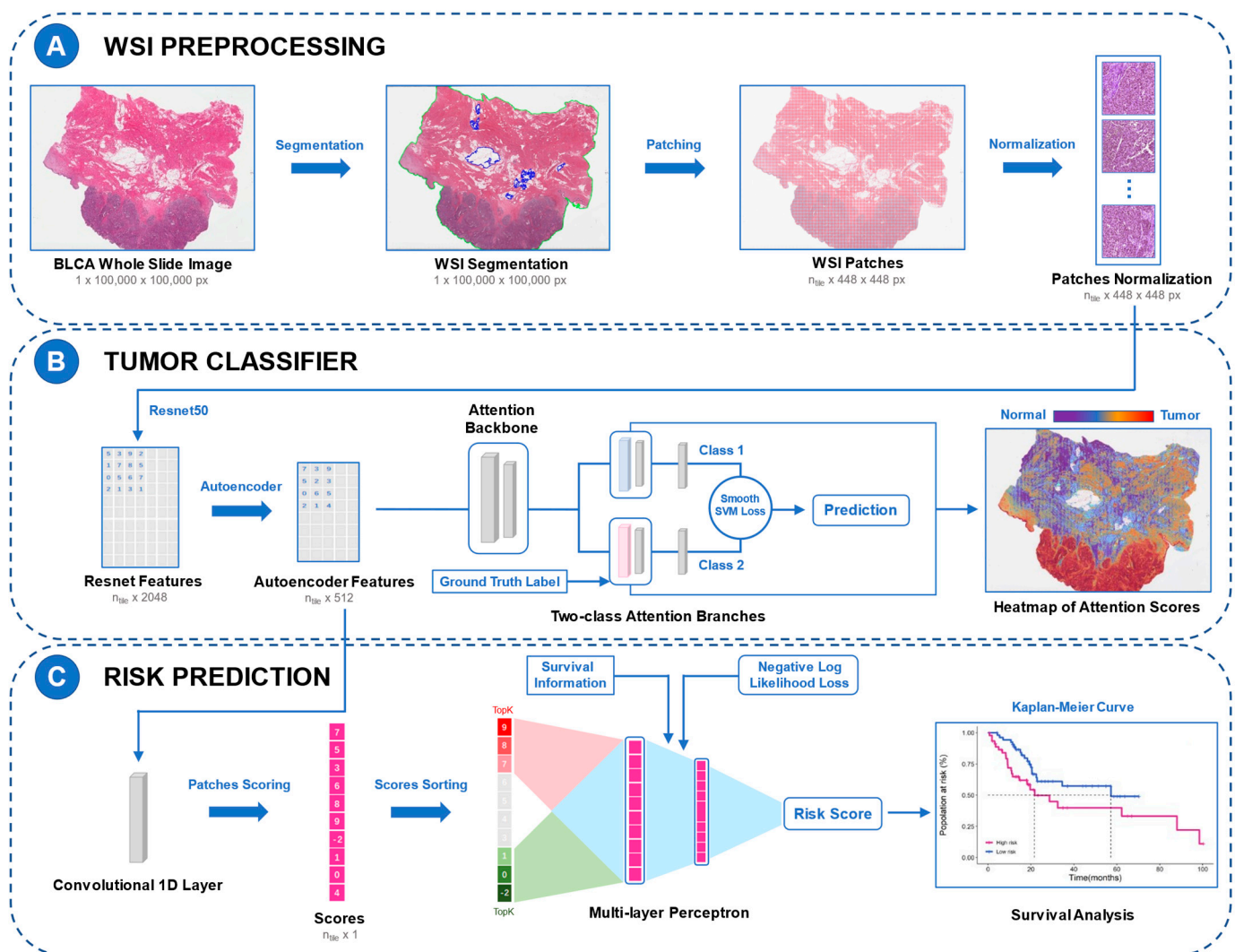

Figure 1. Study flow chart and the layouts of the DL models. The framework of BlcaMIL is shown in A and B and the framework of MibcMLP is shown in A and C. (**A**) Each WSI was first segmented into tissue-containing regions (green border) and empty regions inside the tissue (blue border), and then patches with 448×448 pixels were generated. (**B**) Feature extraction was performed on all patches using ResNet-50, and dimensionality reduction was performed with Autoencoder. Through the MIL model with attention mechanism, the extracted patch-level features were input into the BlcaMIL model, the attention scores of these patches were output, and the average pooling function was used to aggregate them into the WSI-level to make the final diagnosis. Heatmaps visualize ROIs for the model. (**C**) Patch-level features were fed into the network along with survival information, and each patch was assigned a risk score through an iterative learning process. Then, the 50 patches with the highest and lowest scores were selected to be input to the MLP model to predict patient survival. Finally, MIBC patients were stratified using the resulting risk scores. DL, deep learning; WSI, whole slide image; MIL, multiple instance learning; ROI, region of interest; MLP, multi-layer perceptron; MIBC, muscle invasive bladder cancer.

2.6. Model Interpretability and Visualization

For the interpretability of the diagnostic model, the attention scores of BlcaMIL-predicted categories were converted to percentiles and values between 0 and 1 (1 for the most contribution, 0 for the least). The normalized scores were then converted to RGB colors using a divergence colormap, mapped to their corresponding spatial locations in the WSI to generate an attention heatmap. The color red represented areas to which the neural network paid close attention, while the color blue represented areas receiving less attention. Furthermore, the BlcaMIL also indicated some patches with high attention scores, which was convenient for review, and helped in understanding the ROIs of the DL-model and explaining the tumor-related pathological features.

The interpretability of the prognostic model was high because we knew exactly which patches were used to create the risk scores. We extracted the scores of all patches, selected the 100 patches with the highest and 100 with the lowest scores after tiling and sorting, and invited two expert uropathologists for the analysis of tumor-related pathological characteristics. The uropathologists did not know the risk scores assigned to these patches in advance, and then made statistical records of tumor-related pathological characteristics for these 200 patches.

2.7. Statistical Analysis

The classification performance of the diagnostic model was assessed by the area under the receiver-operator curve (AUC), as well as the accuracy, sensitivity, and specificity. A two-sided McNemar's test was performed to compare the differences in accuracy, sensitivity, and specificity between the diagnostic model and the pathologists. Cohen's kappa coefficient was calculated to assess the diagnostic agreement between the diagnostic model and the pathologists. We used Harrell's concordance index (C-index) as an indicator to evaluate the performance of the prognostic model in predicting OS. The Kaplan–Meier survival curve was plotted using R software (Version 3.5.1, R Core Team, Vienna, Austria) to evaluate the correlation between the risk scores generated by the prognostic model and the OS of MIBC patients, and the Log-Rank test was performed. The R software package was obtained from CRAN (https://cran.r-project.org, accessed on 22 October 2022). Pearson's correlation test was performed to assess the significance of the correlation between the two covariates. Differences with p values lower than 0.05 were considered statistically significant (two-tailed). Python (Version 3.8.13, CreateSpace, Scotts Valley, CA, USA) and Pytroch (Version 1.10.0, Curran Associates, Inc., Vancouver, BC, Canada) were used for model building and development.

3. Results

3.1. Patient Characteristics

After screening for the inclusion criteria of the diagnostic model, we included 412 BLCA patients from the TCGA cohort and 150 BLCA patients from the RHWU cohort. A total of 926 WSIs from the TCGA cohort were obtained for training the BlcaMIL model. Through data augmentation, 1627 WSIs (tumor: normal = 887:740) were finally used for the development of the BlcaMIL model. Of those, 80% ($N = 1302$) were randomly selected as the training set while the remaining 20% ($N = 325$) were used as the internal validation set. 250 WSIs from the RHWU cohort were used for independent external validation. The detailed data distribution is shown in Supplementary Table S1.

From the TCGA cohort, 326 patients were eligible according to the inclusion criteria for the prognostic model to participate in the construction of MibcMLP. The TCGA cohort comprised 326 WSIs, which were randomly assigned to the training set ($N = 190$) and internal validation set ($N = 136$). The external validation set included 144 WSIs from the RHWU cohort. Table 1 exhibits the baseline characteristics of the two cohorts used for the prognostic model.

Table 1. Clinical, biological, and pathological features of the MIBC patients included in the prognostic model (MibcMLP).

	TCGA (*N* = 326)	RHWU (*N* = 144)
Age (years)	68 (57, 79)	66 (26, 87)
Sex		
female	87 (26.69%)	21 (14.58%)
male	239 (73.31%)	123 (85.42%)
pT stage		
pT2	99 (30.37%)	58 (40.28%)
pT3	158 (48.47%)	67 (46.53%)
pT4	42 (12.88%)	19 (13.19%)
pTx	27 (8.28%)	0 (0%)
pN stage		
pN0	179 (54.91%)	71 (49.31%)
pN1	38 (11.66%)	37 (25.69%)
pN2	67 (20.55%)	20 (13.89%)
pN3	6 (1.84%)	16 (11.11%)
pNx	36 (11.04%)	0 (0%)
pM stage		
pM0	138 (42.33%)	140 (97.22%)
pM1	8 (2.45%)	4 (2.78%)
pMx	180 (55.22%)	0 (0%)
pTNM stage		
Stage II	106 (32.52%)	41 (28.47%)
Stage III	104 (31.90%)	81 (56.25%)
Stage IV	115 (35.28%)	22 (15.28%)
Missing	1 (0.31%)	0 (0%)
Lymphovascular invasion		
No	100 (30.67%)	91 (63.19%)
Yes	121 (37.12%)	53 (36.81%)
Missing	105 (32.21%)	0 (0%)
Survival status		
Alive	180 (55.21%)	79 (54.86%)
Dead	146 (44.79%)	65 (45.14%)
OS time (months)	18.1 (0, 165.6)	16.0 (1.9, 66.0)

MIBC, Muscle-invasive bladder cancer.

3.2. Performance of the Diagnostic Model

A previous study has shown that the discriminative ability of the MIL model for images was optimal at $10\times$ magnification compared to other magnifications [17]. Accordingly, we loaded all WSIs at $10\times$ magnification and extracted a total of 13,115,687 patches (448×448 pixels). The labels of these patches were consistent with the corresponding WSI labels. Relevant features were extracted for each WSI using Resnet-50, a pre-trained convolutional neural network based on ImageNet, before training the model.

In the diagnostic model, the accuracy of the training set and that of the internal validation set were both 0.998 (AUC, 1.000). Even in the external validation set, the generalization ability of BlcaMIL was still strong with an accuracy of 0.987 (AUC, 0.993) (Table 2a). In addition, we invited two expert uropathologists A and B, who were chief or associate chief uropathologists, and a junior pathologist C who was under training to judge 250 WSIs from the external validation set. Our diagnostic model BlcaMIL was better than that of the junior pathologist C (Accuracy = 0.876) ($p < 0.0001$, paired Chi-squared test). There was no significant difference when using BlcaMIL compared to expert uropathologist A (Accuracy = 0.991) ($p > 0.05$, paired Chi-squared test) or expert uropathologist B (Accuracy = 0.993) ($p > 0.05$, paired Chi-squared test). Moreover, BlcaMIL achieved decent inter-observer agreement with the expert uropathologists (kappa = 0.909 and 0.925, respectively) (Table 2b).

Table 2. Accuracy, sensitivity, specificity, and AUC of the diagnostic model (BlcaMIL) and human pathologists.

a. Accuracy, Sensitivity, and Specificity in the Diagnostic Model (BlcaMIL)				
	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUC (95% CI)
Training set	0.998 (0.996, 0.999)	0.999 (0.998, 1.000)	0.998 (0.996, 0.999)	1.000 (1.000, 1.000)
Internal validation set	0.998 (0.996, 1.000)	1.000 (1.000, 1.000)	0.996 (0.992, 1.000)	1.000 (1.000, 1.000)
External validation set	0.987 (0.981, 0.994)	0.984 (0.971, 0.998)	0.986 (0.979, 0.993)	0.993 (0.990, 0.997)
b. Comparison of the BlcaMIL model and human pathologists in the external validation set				
	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	*p*-Value *	Kappa #
BlcaMIL Model	0.987 (0.981, 0.994)	0.984 (0.971, 0.998)	0.986 (0.979, 0.993)	-	-
Expert Uropathologist A	0.991 (0.987, 0.995)	0.988 (0.981, 0.995)	0.996 (0.989, 1.000)	1.000	0.909
Expert Uropathologist B	0.993 (0.991, 0.995)	0.991 (0.987, 0.995)	0.996 (0.989, 1.000)	1.000	0.925
Junior Pathologist C	0.876 (0.852, 0.900)	0.834 (0.811, 0.858)	0.940 (0.904, 0.976)	<0.0001	0.711

* A paired Chi-squared test (McNemar's test) was used to examine differences in accuracy between the BlcaMIL model and each uropathologist. # Inter-observer agreement between the BlcaMIL model and each uropathologist assessed by the Cohen kappa coefficient. CI, Confidence Interval.

3.3. Performance of the Prognostic Model

We used the C-index metric to assess the ability of the model to predict OS. We found that the MibcMLP model performed well on both the training set and the internal validation set, with C-index values of 0.744 and 0.631, respectively. Based on the input WSI, the MibcMLP model we trained was able to assign each patient a risk score. In contrast to the classification of histopathology, the score was a continuous numerical value rather than a discrete classification. We divided MIBC patients into high-risk and low-risk score groups using the median risk score in the training set as a cut-off point. We then adopted Kaplan-Meier plots and univariate and multivariate Cox models to assess the association between the risk scores and survival outcomes in MIBC patients. In the training set, the risk score predicted by MibcMLP was a strong predictor of OS in univariate analysis (HR = 3.896, $p < 0.001$, Cox analysis; Supplementary Table S2, Supplementary Figure S1). After retaining significant prognostic indicators in univariate analysis (pT stage, pN stage, pM stage, pTNM stage, and Lymphovascular invasion), the risk score remained strongly predictive in multivariate analysis (HR = 3.557, $p < 0.001$, Cox analysis; Supplementary Table S2).

In the internal validation set, the MibcMLP model stratified the MIBC population more accurately than any other clinical or pathological variable in Cox univariate analysis (HR = 3.274, $p < 0.001$, Figure 2A). Even after adjusting for significant prognostic indicators in univariate analysis (pT stage, pN stage, pM stage, and Lymphovascular invasion), the risk score predicted survival outcomes in the multivariate analysis (HR = 3.157, $p < 0.001$, Cox analysis; Figure 2C). The risk score predicted by MibcMLP had independent prognostic value ($p = 2 \times 10^{-5}$, Log-rank test; Figure 2B), even between among other subgroups (such as age, gender, pT stage, pN stage, pM stage, pTNM stage, histological grade and lymphovascular invasion; Figure 2D–L). These results indicated that the pathological features captured by MibcMLP were not redundant with the existing clinicopathological features, and constituted an effective prognostic method independent of current AJCC TNM staging.
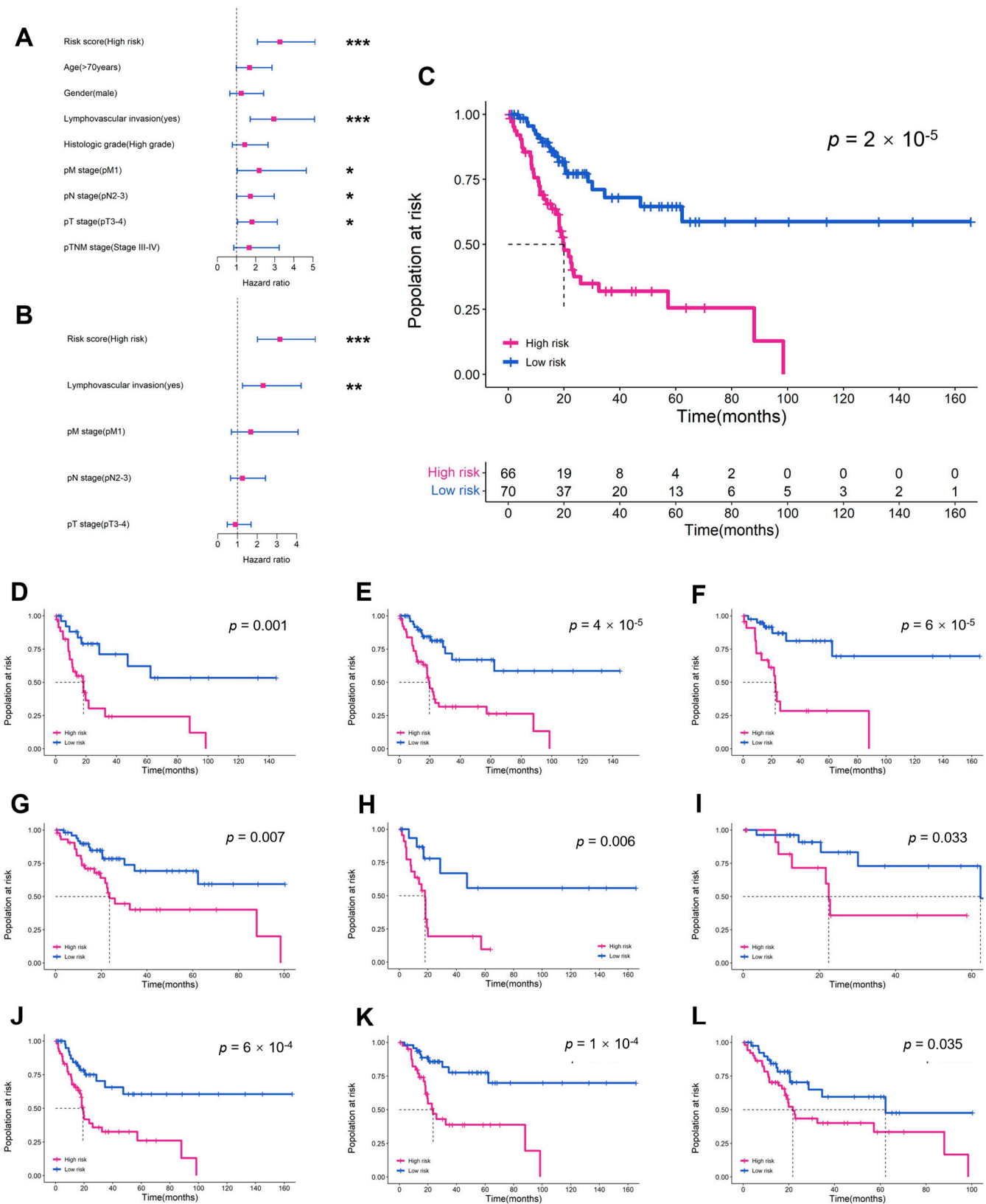
Figure 2. Prognostic value of MibcMLP-generated risk scores in the internal validation set. HR and 95% CI for MibcMLP and other clinicopathological features to predict survival in (**A**) univariate Cox and (**B**) multivariate Cox analyses. MibcMLP model scores were converted to binary scores (high risk or low risk) using the median risk score of the training set as a cut-off. K-M survival curves for (**C**) the

entire internal validation set and the following subgroups: (**D**) age $\geq$70; (**E**) male; (**F**) pT stage 3–4; (**G**) pN stage 0–1; (**H**) pN stage 2–3; (**I**) pTNM stage 1–2; (**J**) pTNM stage 3–4; (**K**) no lymphovascular invasion; (**L**) high histologic grade. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; HR, hazard ratio; CI, confidence interval.

We assessed the robustness of the MibcMLP model by testing it on an independent RHWU cohort. A total of 144 BLCA patients met the inclusion criteria (Table 1). Of these, 65 patients died, and the OS was lower than that in the TCGA cohort (median 16.0 vs. 18.1, $p = 0.330$, Kruskal-Wallis nonparametric test). The following variables were collected and included in subsequent analyses: age, gender, tumor size, pT stage, pN stage, pM stage, pTNM stage, lymphovascular invasion and histological grade. The MibcMLP model extracted and processed patches of 144 WSIs corresponding to 144 patients with a predicted survival C-index of 0.622. We divided the RHWU cohort patients into subclass of high-risk and low-risk using stratification cut-off point from the training set. The results showed that the risk score predicted by MibcMLP was significantly better than other clinicopathological features in the Cox univariate analysis (HR = 2.390, $p < 0.001$, Cox analysis; Figure 3A) and was an independent prognostic factor ($p = 1 \times 10^{-4}$ Log-rank test; Figure 3B). Multivariate Cox regression analysis showed that MibcMLP was also an independent prognostic variable (HR = 2.414, $p < 0.001$, Cox analysis; Figure 3C) after adjustment for important prognostic indicators. MibcMLP predicted survival well even after stratification for other common clinicopathological features (such as age, gender, pT stage, pN stage, pM stage, pTNM stage, histological grade, tumor size and lymphovascular invasion; Figure 3D–L).

3.4. Visualization of DL Models

The BlcaMIL model assigned an attention score to each patch, which represents the degree of contribution to the prediction result. We converted the attention scores into heatmaps to visualize the ROIs of the model. As shown in Figure 4, the high-attention areas mostly consisted of tumor cells with a dense arrangement, hyperchromatic nuclei, and high atypia, while the low-attention areas comprised mostly normal tissue or background. This demonstrated a high degree of concordance with annotations of the pathologists regardless of the pathological stage of BLCA, which was consistent with the established experience of detecting tumor regions, in line with human pathology expertise. This simple and intuitive interpretability and visualization technique allowed us to gain insight into the morphological patterns predicted by the model.

For the MibcMLP model, we aggregated all patches together, obtained the risk score of each patch through MibcMLP, picked out the patches with the highest scores ($n = 100$) and the lowest scores ($n = 100$), and tried to interpret their histopathological features. The patches were independently reviewed by two expert uropathologists who were unaware of the scores assigned, and the associated pathological features were recorded. The results showed that most of the patches associated with poor survival were mainly located in the stromal region, not within the tumor region (ratio, low [resp. high] survival patches in stroma = 94/100 (resp. 10/100, Chi-squared test, $p = 1.4 \times 10^{-32}$). Among the pathological features recorded, the features most predictive of high risk were the presence of vascular space (Chi-squared test, $p = 3.1 \times 10^{-20}$), high cytological atypia (Chi-squared test, $p = 4.8 \times 10^{-34}$), and high nuclear pigmentation (Chi-squared test, $p = 2.7 \times 10^{-33}$). The feature most predictive of low risk was immune cell infiltration (Chi-squared test, $p = 2.0 \times 10^{-17}$) (Figure 5). Taken together, the above results demonstrated that the BlcaMIL and MibcMLP models can detect histopathological features related to diagnosis and survival in BLCA.
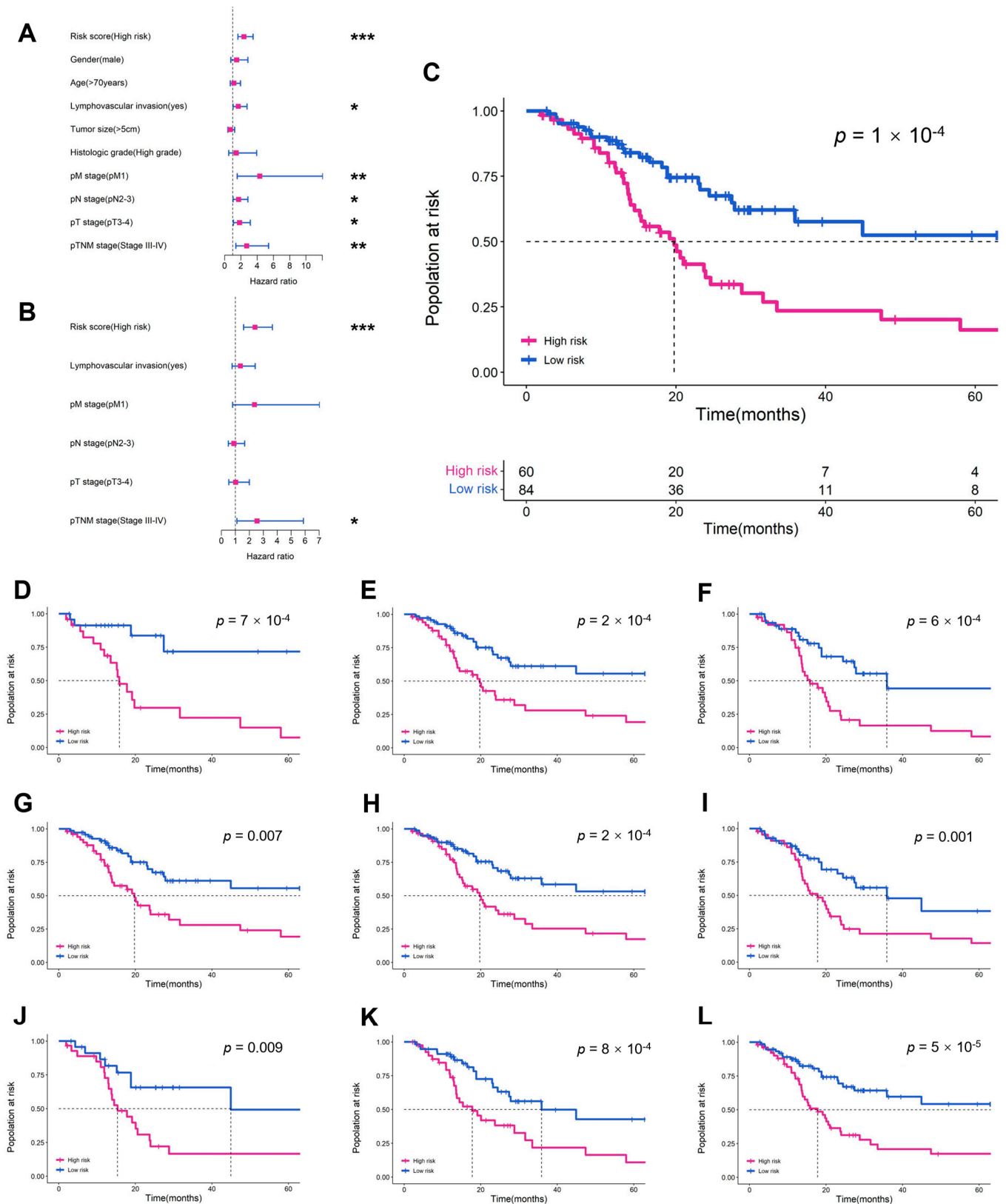
Figure 3. The performance of MibcMLP in predicting prognosis in the external validation set. (**A**) univariate Cox and (**B**) multivariate Cox analyses are exhibited. Using the median risk score in the training set as the cut-off point, (**C**) K-M survival curves for the entire external validation set and

the following subgroups: (**D**) age ≥70; (**E**) male; (**F**) pT stage 3–4; (**G**) pN stage 0–1; (**H**) pM stage 0; (**I**) pTNM stage 3–4; (**J**) lymphovascular invasion; (**K**) tumor size <5cm; (**L**) high histologic grade. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.
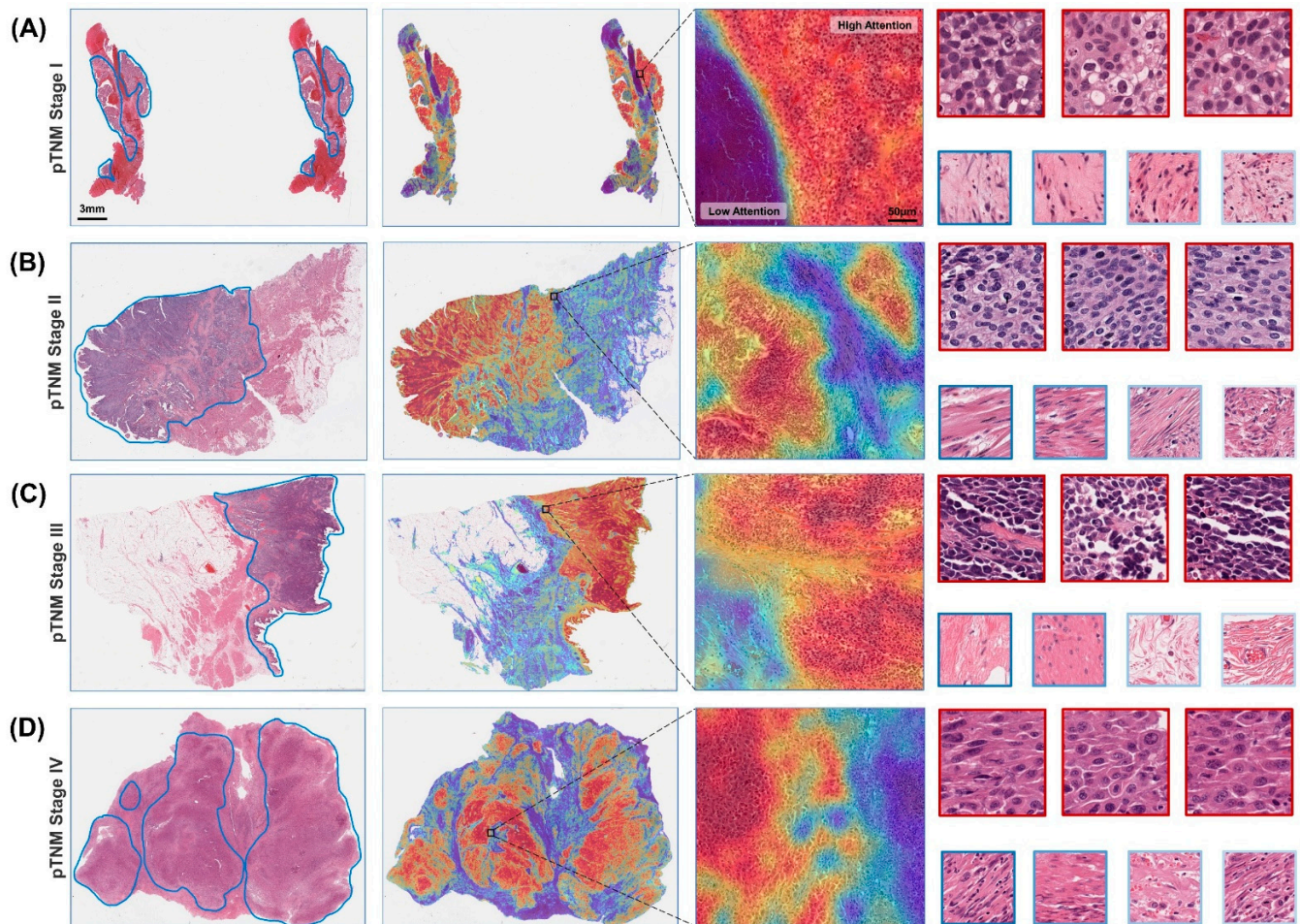
Figure 4. Heatmaps of the diagnostic model on WSIs at different pathological stages. A representative WSI for each pathological stage was annotated by a uropathologist who roughly delineated the tumor tissue area (first column), including (**A**) AJCC TNM stage I, (**B**) stage II, (**C**) stage III and (**D**) stage IV. The attention scores of the predicted categories of patches are calculated by the model, and the attention heatmap corresponding to each WSI was generated and overlaid onto it (second column). It is then further zoomed in to show the heatmap of the ROI, highlighting the tumor and normal borders (third column). Patches with the highest attention (red border) often exhibit well-known tumor morphology, while patches of low interest (blue border) tend to be normal tissue or background (fourth column).

3.5. Gene Expression Correlation with Risk Scores

We obtained gene expression data of TCGA MIBC patients from the UCSC Xena database, with up to 56,536 genes. The association between risk scores and gene expression levels in each patient was examined by Pearson correlation analysis, revealing possible potential biological correlations. The expression of six genes was significantly associated with MibcMLP-predicted risk scores: ANAPC7 (correlation = −0.407; Pearson's correlation test, $p = 6.9 \times 10^{-5}$), MAPKAPK5 (correlation = −0.443; Pearson's correlation test, $p = 1.2 \times 10^{-5}$), COX19 (correlation = −0.412; Pearson's correlation test, $p = 5.4 \times 10^{-5}$), LINC01106 (correlation = −0.410; Pearson's correlation test, $p = 6.2 \times 10^{-5}$), AL161431.1 (correlation = 0.393; Pearson correlation test, $p = 3.4 \times 10^{-4}$) and MYO16-AS1 (correlation = 0.497; Pearson correlation test, $p = 7.1 \times 10^{-4}$) (Figure 6).
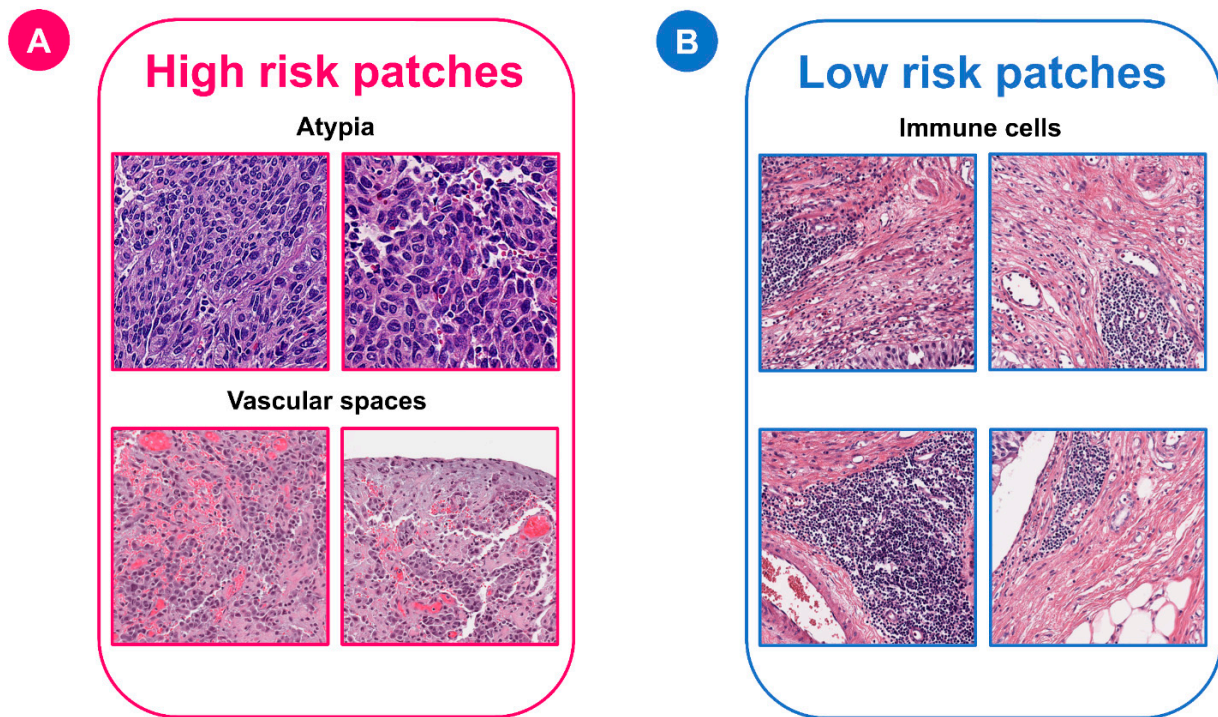
Figure 5. Representative examples of patches classified as high or low risk by the MibcMLP model. The top 200 most predictive patches were analyzed by expert uropathologists who was unaware of the risk scores. (**A**) Features predicting a high mortality risk included cellular atypia and vascular space. (**B**) Features predicting a low risk of death was the presence of immune cells.
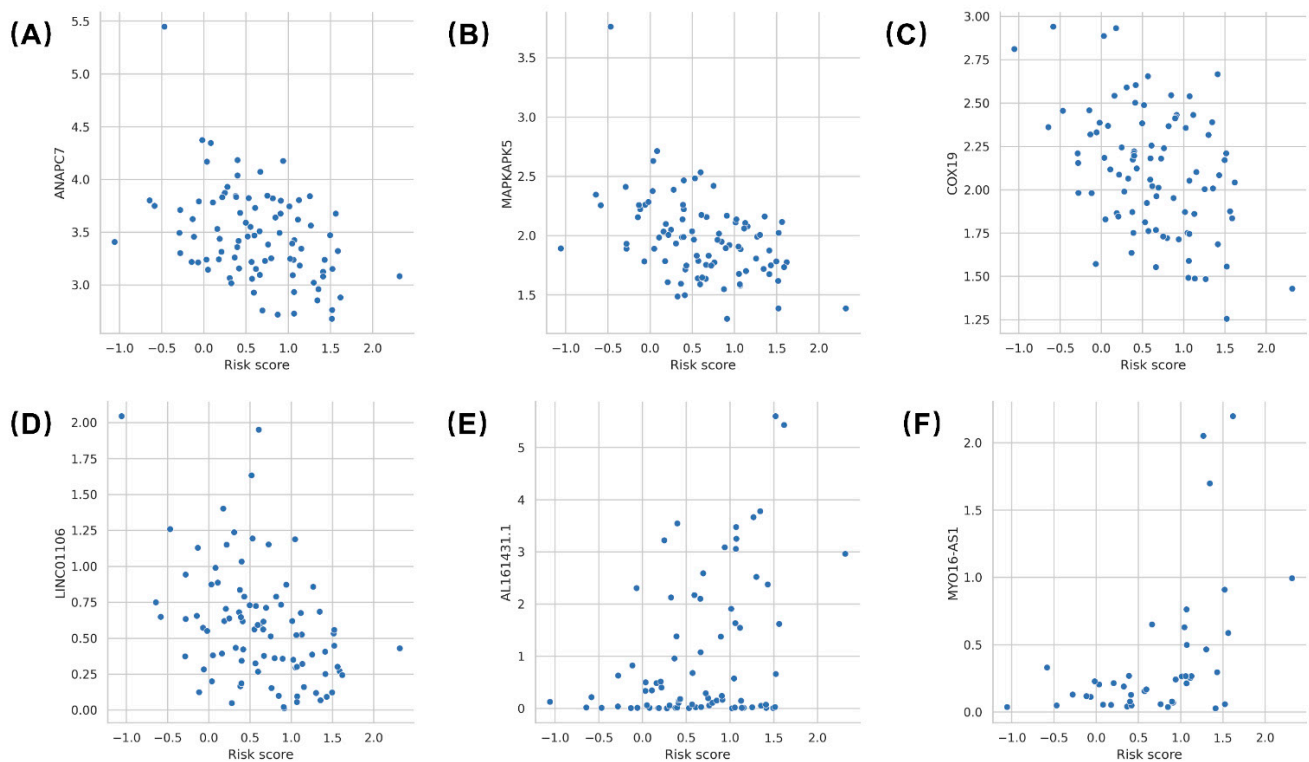

Figure 6. Correlates between risk scores and gene expression levels. Biological correlation between MibcMLP risk scores and (**A**) the ANAPC7 expression (N = 90 samples), (**B**) the MAPKAPK5 expression

(N = 90 samples), (**C**) the COX19 expression (N = 90 samples), (**D**) the LINC01106 expression (N = 90 samples), (**E**) the Al161431.1 expression (N = 78 samples), and (**F**) the MYO16-AS1 expression (N = 43 samples) available for the TCGA dataset.

4. Discussion

BLCA is a disease with complex molecular features, severe morbidity, and high mortality. Mining of potentially robust clinical and/or biological features will aid in the diagnosis and risk stratification of BLCA patients to facilitate personalized treatment. A recent series of studies has explored the impact of immunohistochemical assays [30], conventional serum and histological biomarkers [31–35], and adjuvant therapy [36,37] on longitudinal monitoring and prognosis definition in BLCA patients. In this study, we utilized DL to develop a diagnostic model, BlcaMIL, and a prognostic model, MibcMLP, using WSIs for accurate diagnosis of BLCA and prognosis prediction of MIBC patients. Encouragingly, our BlcaMIL accurately differentiated BLCA from normal pathological images (AUC close to 1), with a performance comparable to that of expert uropathologists and better than that of a junior pathologist. Our MibcMLP not only had excellent performance in the training set (C-index = 0.744) and internal validation set (C-index = 0.631), but also exhibited robust performance on the independent external validation set (C-index = 0.622). Furthermore, univariate and multivariate Cox analyses demonstrated that the risk score predicted by MibcMLP was an independent prognostic factor, a beneficial complement to existing markers, and more accurate than classical clinical, biological, and pathological features in predicting OS.

In recent years, data-driven machine learning and DL have been widely used in the processing and analysis of medical images, providing new tools for disease diagnosis and prognosis prediction. A novel approach combining radiomics and machine learning has brought encouraging results in the diagnosis and prediction of urological cancers [38–40]. Previously, our team developed a DL-model based on cystoscopy for clinical real-time recognition of bladder tumors with an accuracy comparable to that of experienced clinical experts [41]. In this study, we adopted DL to analyze digitized H&E-stained BLCA pathological images. There have been some DL studies based on WSIs where the diagnosis and survival prognosis of tumor patients have been successfully analyzed in soft tissue sarcoma [42], brain glioma [43], gastric cancer [17], and rectal cancer [19]. This type of research has been emerging also in the field of BLCA. Wetteland et al. [44] proposed a DL pipeline that identified diagnostic-relevant regions in WSI and predicted the grade. Fuster et al. [45] analyzed pathological images of NMIBC patients and proposed a multi-scale DL model that detected cancerous areas patterns across WSIs for accurate T1 staging. Woerl et al. [15] successfully identified molecular subtypes of MIBC patients based on convolutional neural networks, and visualized the consequent ROIs. Lucas et al. [46] performed an analysis of relapse-prone NMIBC and confirmed that using DL in conjunction with WSIs and clinical data could improve relapse prediction in BLCA patients. However, these recent studies did not attempt to make an accurate diagnosis using the WSIs, nor did they directly analyze the association between WSIs and survival outcomes in MIBC patients.

An increasing number of studies have used BLCA gene expression profiling or genetic testing approaches to predict OS in BLCA patients [47–50]. Next-generation sequencing approaches can provide the wealth of information required to molecularly classify BLCA patients, thereby identifying potential therapeutic targets [51]. Recently, Lindskrog et al. [52] conducted a comprehensive multi-omics analysis of 834 NMIBC patients from the URO-MOL project. Their results demonstrated the independent prognostic value of transcriptomic subtyping and chromosomal instability levels over clinicopathological features and were confirmed in 1228 validation samples. However, implementation of these high-throughput gene expression profiling/RNA-sequencing technologies in clinical practice is currently hindered by high costs, the requirement of nucleic acid extraction, and issues of standardization and reproducibility. In contrast, our diagnostic and prognostic models require only H&E-stained digitized slide images as inputs to make a diagnosis or output a risk score associated with survival. Such slides are readily obtained in a surgical treatment

setting due to the abundance of histological material available during surgery. Moreover, the collected WSIs do not require professional pathologists for the annotation of ROIs, as the trained model can automatically identify specific regions associated with tumors in WSIs, greatly reducing the need for pathologists, as well as lowering the time and effort to annotate ahead of time. Furthermore, WSIs contain a wealth of potential information that is often difficult to detect by pathologists. Traditional histological stage and pathological grade prognostic methods are likely to include inter-observer variation, whereas the DL-based method reduces human intervention and potentially improves reproducibility. Consequently, we believe that our methods are more helpful for the diagnosis and risk prediction of BLCA patients in economically underdeveloped areas with a shortage of pathologists.

Although DL tools have produced extremely reliable results so far, the inference process of these models is often highly opaque, making it difficult or impossible for us, and even for a domain expert, to understand. This so-called "black box" model undermines the credibility of the results and limits the practical application of DL in pathology [53]. In our models, we extracted the patches that were identified as the most relevant to diagnosis or prognosis for pathological interpretation. This transparent and interpretable process allowed us to further submit these patches to expert uropathologists for review and analysis. Our results indicated that histopathological features obtained by DL-models, such as high cytological atypia and high nuclear pigmentation, are currently used features for diagnosis and prognosis by pathologists. Moreover, the presence of vascular space is strongly associated with high-risk patients with low OS, which is consistent with the current findings [54]. For low-risk patches, immune cell infiltration is an extremely important feature. The tumor microenvironment plays a crucial role in tumor progression, and massive immune cell infiltration is often associated with molecular subtypes with better prognosis, not only in BLCA but also in other cancers [55–57]. Therefore, we believe that the novel features extracted by the DL-model not only enable reliable prognostic analysis of MIBC patients, but also make essential contributions to dissect tumor molecular subtypes.

We also further investigated the association between risk scores and gene expression levels in MIBC patients, in which ANAPC7, MAPKAPK5, COX19, and LINC01106 were negatively correlated with predicted risk scores, and AL161431.1 and MYO16-AS1 were positively correlated with risk scores. Previous studies have shown that these six genes play important roles in promoting or inhibiting the occurrence and development of tumors. ANAPC7 has been identified as a sponge of miR-373 that inhibits tumor growth in vitro and in vivo by regulating the cell cycle [58]. MAPKAPK5 upregulation is associated with high expression of the transcriptional regulator YAP and poor prognosis in clinical tumor samples, and its positive regulation of YAP activity plays an important role in human cancers [59]. COX19 may affect the assembly of cytochrome C oxidase (COX) subunits, which in turn affects the activity of COX and apoptosis [60]. COX19 has been confirmed to be a key factor in the inhibition of tumor cell apoptosis by microRNA-21, and inhibition of COX19 expression may enhance apoptosis and reduce tumor cell proliferation [61]. LINC01106 has been shown to promote the progression of BLCA and its expression is enhanced in BLCA. Knockdown of LINC01106 results in the inhibition of proliferation, migration, and invasion of BLCA cells, making LINC01106 a potential target for the treatment of BLCA patients [62]. AL161431.1 is a long non-coding RNA associated with tumor hypoxia and autophagy [63,64]. It exhibits an elevated level in pancreatic cancer [65], endometrial cancer [66] and lung cancer [67], and may promote the migration and invasion of tumor cells. MYO16-AS1 is a strong prognostic factor in MIBC, and its upregulation is associated with longer OS in MIBC patients, suggesting that it plays a cancer-promoting role in MIBC [68]. Therefore, some genes closely related to the progression of MIBC can be identified through the predicted risk score, which can provide a reference for the development of new therapeutic targets.

There are still some limitations in our study. First, the datasets we used to train and validate the two DL-models are not sufficiently large. Further incorporation of multicen-

ter data to improve the generalization performance of the models will be an important consideration in the future before the models are widely used in clinical practice. Second, since our study is a retrospective one, our diagnostic and prognostic DL-models need to be further validated based on prospective, randomized, multicenter clinical trial formulated by SPIRIT-AI and CONSORT-AI [69], to improve clinicopathological information related to treatment and prognosis [7], including variant histology, adjuvant chemotherapy, type of surgery and comorbidities. Third, pathological slides obtained from different laboratories show variations due to differences among labs in sample collection, fabrication techniques, staining materials, and digital scans. Although we used a staining normalization approach and demonstrated decent performance for both models on an external validation set, challenges related to data normalization remain. Hence, it is necessary to establish a standardized procedure for the production of pathological slide images to improve the quality of the dataset.

5. Conclusions

We developed weakly supervised models for diagnosing BLCA and predicting OS based on DL using digital H&E-stained images of BLCA patients. The performance of the diagnostic model was comparable to that of expert uropathologists, and the prognostic model showed better prognostic value than any other clinical, biological, and pathological features. Our models can not only assist clinicians in recognizing BLCA, but also help stratify MIBC patients and improve subsequent personalized treatment decisions. Finally, further analysis of the critical tumor-related regions and MibcMLP-predicted risk scores increased the amount of knowledge garnered from WSIs and potentially led to new biomarker discoveries.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef] [PubMed]

2. Khadhouri, S.; Gallagher, K.M.; Mackenzie, K.R.; Shah, T.T.; Gao, C.; Moore, S.; Zimmermann, E.F.; Edison, E.; Jefferies, M.; Nambiar, A.; et al. The IDENTIFY study: The investigation and detection of urological neoplasia in patients referred with suspected urinary tract cancer-a multicentre observational study. *BJU Int.* **2021**, *128*, 440–450. [CrossRef] [PubMed]

3. Patel, V.G.; Oh, W.K.; Galsky, M.D. Treatment of muscle-invasive and advanced bladder cancer in 2020. *CA Cancer J. Clin.* **2020**, *70*, 404–423. [CrossRef] [PubMed]

4. Khadhouri, S.; Gallagher, K.M.; Mackenzie, K.R.; Shah, T.T.; Gao, C.; Moore, S.; Zimmermann, E.F.; Edison, E.; Jefferies, M.; Nambiar, A.; et al. Developing a diagnostic multivariable prediction model for urinary tract cancer in patients referred with haematuria: Results from the IDENTIFY collaborative study. *Eur. Urol. Focus* **2022**. [CrossRef] [PubMed]

5. Zehnder, P.; Studer, U.E.; Skinner, E.C.; Thalmann, G.N.; Miranda, G.; Roth, B.; Cai, J.; Birkhäuser, F.D.; Mitra, A.P.; Burkhard, F.C.; et al. Unaltered oncological outcomes of radical cystectomy with extended lymphadenectomy over three decades. *BJU Int.* **2013**, *112*, E51–E58. [CrossRef]

6. Metter, D.M.; Colgan, T.J.; Leung, S.T.; Timmons, C.F.; Park, J.Y. Trends in the US and Canadian Pathologist Workforces From 2007 to 2017. *JAMA Netw. Open* **2019**, *2*, e194337. [CrossRef]

7. Witjes, J.A.; Bruins, H.M.; Cathomas, R.; Compérat, E.M.; Cowan, N.C.; Gakis, G.; Hernández, V.; Linares, E.E.; Lorch, A.; Neuzillet, Y.; et al. European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. *Eur. Urol.* **2021**, *79*, 82–104. [CrossRef]

8. Rosai, J. *Rosai and Ackerman's Surgical Pathology e-Book*; Elsevier Health Sciences: USA, 2011.

9. Rouprêt, M.; Babjuk, M.; Burger, M.; Capoun, O.; Cohen, D.; Compérat, E.M.; Cowan, N.C.; Dominguez-Escrig, J.L.; Gontero, P.; Hugh, M.A.; et al. European Association of Urology Guidelines on Upper Urinary Tract Urothelial Carcinoma: 2020 Update. *Eur. Urol.* **2021**, *79*, 62–79. [CrossRef]

10. Jiang, Y.; Yang, M.; Wang, S.; Li, X.; Sun, Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Commun.* **2020**, *40*, 154–166. [CrossRef]

11. Wu, S.; Chen, X.; Pan, J.; Dong, W.; Diao, X.; Zhang, R.; Zhang, Y.; Zhang, Y.; Qian, G.; Chen, H.; et al. An artificial intelligence system for the detection of bladder cancer via cystoscopy: A multicenter diagnostic study. *J. Natl. Cancer Inst.* **2022**, *114*, 220–227. [CrossRef]

12. Zou, Y.; Cai, L.; Chen, C.; Shao, Q.; Fu, X.; Yu, J.; Wang, L.; Chen, Z.; Yang, X.; Yuan, B.; et al. Multi-task deep learning based on T2-Weighted Images for predicting Muscular-Invasive Bladder Cancer. *Comput. Biol. Med.* **2022**, *151*, 106219. [CrossRef]

13. Freitas, N.R.; Vieira, P.M.; Cordeiro, A.; Tinoco, C.; Morais, N.; Torres, J.; Anacleto, S.; Laguna, M.P.; Lima, E.; Lima, C.S. Detection of bladder cancer with feature fusion, transfer learning and CapsNets. *Artif. Intell. Med.* **2022**, *126*, 102275. [CrossRef] [PubMed]

14. Shkolyar, E.; Jia, X.; Chang, T.C.; Trivedi, D.; Mach, K.E.; Meng, M.Q.; Xing, L.; Liao, J.C. Augmented bladder tumor detection using deep learning. *Eur. Urol.* **2019**, *76*, 714–718. [CrossRef]

15. Woerl, A.C.; Eckstein, M.; Geiger, J.; Wagner, D.C.; Daher, T.; Stenzel, P.; Fernandez, A.; Hartmann, A.; Wand, M.; Roth, W.; et al. Deep learning predicts molecular subtype of muscle-invasive bladder cancer from conventional histopathological slides. *Eur. Urol.* **2020**, *78*, 256–264. [CrossRef] [PubMed]

16. Shi, J.Y.; Wang, X.; Ding, G.Y.; Dong, Z.; Han, J.; Guan, Z.; Ma, L.J.; Zheng, Y.; Zhang, L.; Yu, G.Z.; et al. Exploring prognostic indicators in the pathological images of hepatocellular carcinoma based on deep learning. *Gut* **2021**, *70*, 951–961. [CrossRef]

17. Huang, B.; Tian, S.; Zhan, N.; Ma, J.; Huang, Z.; Zhang, C.; Zhang, H.; Ming, F.; Liao, F.; Ji, M.; et al. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: A retrospective multicentre study. *EBioMedicine* **2021**, *73*, 103631. [CrossRef]

18. Jiao, Y.; Li, J.; Qian, C.; Fei, S. Deep learning-based tumor microenvironment analysis in colon adenocarcinoma histopathological whole-slide images. *Comput. Methods Programs Biomed.* **2021**, *204*, 106047. [CrossRef]

19. Skrede, O.J.; de Raedt, S.; Kleppe, A.; Hveem, T.S.; Liestøl, K.; Maddison, J.; Askautrud, H.A.; Pradhan, M.; Nesheim, J.A.; Albregtsen, F.; et al. Deep learning for prediction of colorectal cancer outcome: A discovery and validation study. *Lancet* **2020**, *395*, 350–360. [CrossRef]

20. Courtiol, P.; Maussion, C.; Moarii, M.; Pronier, E.; Pilcer, S.; Sefta, M.; Manceron, P.; Toldo, S.; Zaslavskiy, M.; Le Stang, N.; et al. Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat. Med.* **2019**, *25*, 1519–1525. [CrossRef]

21. Saillard, C.; Schmauch, B.; Laifa, O.; Moarii, M.; Toldo, S.; Zaslavskiy, M.; Pronier, E.; Laurent, A.; Amaddeo, G.; Regnault, H.; et al. Predicting survival after hepatocellular carcinoma resection using deep learning on histological slides. *Hepatology* **2020**, *72*, 2000–2013. [CrossRef] [PubMed]

22. Lu, M.Y.; Williamson, D.; Chen, T.Y.; Chen, R.J.; Barbieri, M.; Mahmood, F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat. Biomed. Eng.* **2021**, *5*, 555–570. [CrossRef]

23. Shamai, G.; Livne, A.; Polónia, A.; Sabo, E.; Cretu, A.; Bar-Sela, G.; Kimmel, R. Deep learning-based image analysis predicts PD-L1 status from H&E-stained histopathology images in breast cancer. *Nat. Commun.* **2022**, *13*, 6753.

24. Loeffler, C.; Ortiz, B.N.; Jung, M.; Seillier, L.; Rose, M.; Laleh, N.G.; Knuechel, R.; Brinker, T.J.; Trautwein, C.; Gaisa, N.T.; et al. Artificial intelligence-based detection of FGFR3 mutational status directly from routine histology in bladder cancer: A possible preselection for molecular testing? *Eur. Urol. Focus* **2022**, *8*, 472–479. [CrossRef]

25. Velmahos, C.S.; Badgeley, M.; Lo, Y.C. Using deep learning to identify bladder cancers with FGFR-activating mutations from histology images. *Cancer Med.* **2021**, *10*, 4805–4813. [CrossRef]

26. Paner, G.P.; Stadler, W.M.; Hansel, D.E.; Montironi, R.; Lin, D.W.; Amin, M.B. Updates in the eighth edition of the Tumor-Node-Metastasis staging classification for urologic cancers. *Eur. Urol.* **2018**, *73*, 560–569. [CrossRef]

27. Vahadane, A.; Peng, T.Y.; Albarqouni, S.; Baust, M.; Steiger, K.; Schlitter, A.M.; Sethi, A.; Esposito, I.; Navab, N. Structure-preserved color normalization for histological images. In Proceedings of the IEEE 12th International Symposium on Biomedical Imaging (ISBI), Brooklyn, NY, USA, 16–19 April 2015; pp. 1012–1015.

28. Anand, D.; Ramakrishnan, G.; Sethi, A. Fast GPU-Enabled color normalization for digital pathology. In Proceedings of the International Conference on Systems, Signals and Image Processing (IWSSIP), Osijek, Croatia, 5–7 June 2019; pp. 219–224.

29. He, K.M.; Zhang, X.Y.; Ren, S.Q.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

30. Mertens, L.S.; Claps, F.; Mayr, R.; Bostrom, P.J.; Shariat, S.F.; Zwarthoff, E.C.; Boormans, J.L.; Abas, C.; van Leenders, G.; Götz, S.; et al. Prognostic markers in invasive bladder cancer: FGFR3 mutation status versus P53 and KI-67 expression: A multi-center, multi-laboratory analysis in 1058 radical cystectomy patients. *Urol. Oncol.* **2022**, *40*, 110–111. [CrossRef]

31. Claps, F.; Mir, M.C.; van Rhijn, B.; Mazzon, G.; Soria, F.; D'Andrea, D.; Marra, G.; Boltri, M.; Traunero, F.; Massanova, M.; et al. Impact of the controlling nutritional status (CONUT) score on perioperative morbidity and oncological outcomes in patients with bladder cancer treated with radical cystectomy. *Urol. Oncol.* **2022**. [CrossRef]

32. Mori, K.; Miura, N.; Mostafaei, H.; Quhal, F.; Motlagh, R.S.; Lysenko, I.; Kimura, S.; Egawa, S.; Karakiewicz, P.I.; Shariat, S.F. Prognostic value of preoperative hematologic biomarkers in urothelial carcinoma of the bladder treated with radical cystectomy: A systematic review and meta-analysis. *Int. J. Clin. Oncol.* **2020**, *25*, 1459–1474. [CrossRef]

33. Schuettfort, V.M.; David, D.; Quhal, F.; Mostafaei, H.; Laukhtina, E.; Mori, K.; Sari, M.R.; Rink, M.; Abufaraj, M.; Karakiewicz, P.I.; et al. Impact of preoperative serum albumin-globulin ratio on disease outcome after radical cystectomy for urothelial carcinoma of the bladder. *Urol. Oncol.* **2021**, *39*, 235. [CrossRef]

34. Claps, F.; Rai, S.; Mir, M.C.; van Rhijn, B.; Mazzon, G.; Davis, L.E.; Valadon, C.L.; Silvestri, T.; Rizzo, M.; Ankem, M.; et al. Prognostic value of preoperative albumin-to-fibrinogen ratio (AFR) in patients with bladder cancer treated with radical cystectomy. *Urol. Oncol.* **2021**, *39*, 835–839. [CrossRef]

35. Claps, F.; van de Kamp, M.W.; Mayr, R.; Bostrom, P.J.; Boormans, J.L.; Eckstein, M.; Mertens, L.S.; Boevé, E.R.; Neuzillet, Y.; Burger, M.; et al. Risk factors associated with positive surgical margins' location at radical cystectomy and their impact on bladder cancer survival. *World J. Urol.* **2021**, *39*, 4363–4371. [CrossRef] [PubMed]

36. Mir, M.C.; Campi, R.; Loriot, Y.; Puente, J.; Giannarini, G.; Necchi, A.; Rouprêt, M. Adjuvant systemic therapy for high-risk muscle-invasive bladder cancer after radical cystectomy: Current options and future opportunities. *Eur. Urol. Oncol.* **2021**. [CrossRef] [PubMed]

37. Afferi, L.; Lonati, C.; Montorsi, F.; Briganti, A.; Necchi, A.; Mari, A.; Minervini, A.; Tellini, R.; Campi, R.; Schulz, G.B.; et al. Selecting the best candidates for cisplatin-based adjuvant chemotherapy after radical cystectomy among patients with pN+ bladder cancer. *Eur. Urol. Oncol* **2022**. [CrossRef] [PubMed]

38. Beşler, M.S.; Koç, U. A new approach to predict the histological variants of bladder urothelial carcinoma: Machine Learning-Based radiomics analysis. *Acad. Radiol.* **2022**. [CrossRef]

39. Cuocolo, R.; Stanzione, A.; Faletti, R.; Gatti, M.; Calleris, G.; Fornari, A.; Gentile, F.; Motta, A.; Dell'Aversana, S.; Creta, M.; et al. MRI index lesion radiomics and machine learning for detection of extraprostatic extension of disease: A multicenter study. *Eur. Radiol.* **2021**, *31*, 7575–7583. [CrossRef]

40. Yang, G.; Nie, P.; Yan, L.; Zhang, M.; Wang, Y.; Zhao, L.; Li, M.; Xie, F.; Xie, H.; Li, X.; et al. The radiomics-based tumor heterogeneity adds incremental value to the existing prognostic models for predicting outcome in localized clear cell renal cell carcinoma: A multicenter study. *Eur. J. Nucl. Med. Mol. Imaging* **2022**, *49*, 2949–2959. [CrossRef]

41. Yang, R.; Du, Y.; Weng, X.; Chen, Z.; Wang, S.; Liu, X. Automatic recognition of bladder tumours using deep learning technology and its clinical application. *Int. J. Med. Robot* **2021**, *17*, e2194. [CrossRef]

42. Foersch, S.; Eckstein, M.; Wagner, D.C.; Gach, F.; Woerl, A.C.; Geiger, J.; Glasner, C.; Schelbert, S.; Schulz, S.; Porubsky, S.; et al. Deep learning for diagnosis and survival prediction in soft tissue sarcoma. *Ann. Oncol.* **2021**, *32*, 1178–1187. [CrossRef]

43. Jin, L.; Shi, F.; Chun, Q.; Chen, H.; Ma, Y.; Wu, S.; Hameed, N.; Mei, C.; Lu, J.; Zhang, J.; et al. Artificial intelligence neuropathologist for glioma classification using deep learning on hematoxylin and eosin stained slide images and molecular markers. *Neuro Oncol.* **2021**, *23*, 44–52. [CrossRef]

44. Wetteland, R.; Kvikstad, V.; Eftestøl, T.; Tøssebro, E.; Lillesand, M.; Janssen, E.A.; Engan, K. Automatic diagnostic tool for predicting cancer grade in bladder cancer patients using deep learning. *IEEE Access* **2021**, *9*, 115813–115825. [CrossRef]

45. Fuster, S.; Khoraminia, F.; Kiraz, U.; Kanwal, N.; Kvikstad, V.; Eftestøl, T.; Zuiverloon, T.C.; Janssen, E.A.; Engan, K. Invasive cancerous area detection in Non-Muscle invasive bladder cancer whole slide images. In Proceedings of the 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Nafplio, Greece, 26–29 June 2022; pp. 1–5.

46. Lucas, M.; Jansen, I.; van Leeuwen, T.G.; Oddens, J.R.; de Bruin, D.M.; Marquering, H.A. Deep learning-based recurrence prediction in patients with non-muscle-invasive bladder cancer. *Eur. Urol. Focus* **2022**, *8*, 165–172. [CrossRef]

47. Chen, F.; Wang, Q.; Zhou, Y. The construction and validation of an RNA binding protein-related prognostic model for bladder cancer. *BMC Cancer* **2021**, *21*, 244. [CrossRef]

48. Lin, J.T.; Tsai, K.W. Circulating miRNAs act as diagnostic biomarkers for bladder cancer in urine. *Int. J. Mol. Sci.* **2021**, *22*, 4278. [CrossRef]

49. Wang, Z.; Tu, L.; Chen, M.; Tong, S. Identification of a tumor microenvironment-related seven-gene signature for predicting prognosis in bladder cancer. *BMC Cancer* **2021**, *21*, 692. [CrossRef]

50. Zhang, P.; Liu, Z.; Wang, D.; Li, Y.; Xing, Y.; Xiao, Y. Scoring system based on RNA modification Writer-Related genes to predict overall survival and therapeutic response in bladder cancer. *Front. Immunol.* **2021**, *12*, 724541. [CrossRef]

51. Claps, F.; Mir, M.C.; Zargar, H. Molecular markers of systemic therapy response in urothelial carcinoma. *Asian J. Urol.* **2021**, *8*, 376–390. [CrossRef]

52. Lindskrog, S.V.; Prip, F.; Lamy, P.; Taber, A.; Groeneveld, C.S.; Birkenkamp-Demtröder, K.; Jensen, J.B.; Strandgaard, T.; Nordentoft, I.; Christensen, E.; et al. An integrated multi-omics analysis identifies prognostic molecular subtypes of non-muscle-invasive bladder cancer. *Nat. Commun.* **2021**, *12*, 2301. [CrossRef]

53. Yang, G.; Ye, Q.; Xia, J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. *Inf. Fusion* **2022**, *77*, 29–52. [CrossRef]

54. Muppa, P.; Gupta, S.; Frank, I.; Boorjian, S.A.; Karnes, R.J.; Thompson, R.H.; Thapa, P.; Tarrell, R.F.; Herrera, H.L.; Jimenez, R.E.; et al. Prognostic significance of lymphatic, vascular and perineural invasion for bladder cancer patients treated by radical cystectomy. *Pathology* **2017**, *49*, 259–266. [CrossRef]

55. Acs, B.; Ahmed, F.S.; Gupta, S.; Wong, P.F.; Gartrell, R.D.; Sarin, P.J.; Rizk, E.M.; Gould, R.B.; Saenger, Y.M.; Rimm, D.L. An open source automated tumor infiltrating lymphocyte algorithm for prognosis in melanoma. *Nat. Commun.* **2019**, *10*, 5440. [CrossRef]

56. He, Y.; Jiang, Z.; Chen, C.; Wang, X. Classification of triple-negative breast cancers based on Immunogenomic profiling. *J. Exp. Clin. Cancer Res.* **2018**, *37*, 327. [CrossRef] [PubMed]

57. Shi, S.; Ma, T.; Xi, Y. Characterization of the immune cell infiltration landscape in bladder cancer to aid immunotherapy. *Arch. Biochem. Biophys.* **2021**, *708*, 108950. [CrossRef] [PubMed]

58. Shi, X.; Yang, J.; Liu, M.; Zhang, Y.; Zhou, Z.; Luo, W.; Fung, K.M.; Xu, C.; Bronze, M.S.; Houchen, C.W.; et al. Circular RNA ANAPC7 inhibits tumor growth and muscle wasting via PHLPP2-AKT-TGF-β signaling axis in pancreatic cancer. *Gastroenterology* **2022**, *162*, 2004–2017. [CrossRef] [PubMed]

59. Seo, J.; Kim, M.H.; Hong, H.; Cho, H.; Park, S.; Kim, S.K.; Kim, J. MK5 regulates YAP stability and is a molecular target in YAP-Driven cancers. *Cancer Res.* **2019**, *79*, 6139–6152. [CrossRef]

60. Leary, S.C.; Cobine, P.A.; Nishimura, T.; Verdijk, R.M.; de Krijger, R.; de Coo, R.; Tarnopolsky, M.A.; Winge, D.R.; Shoubridge, E.A. COX19 mediates the transduction of a mitochondrial redox signal from SCO1 that regulates ATP7A-mediated cellular copper efflux. *Mol. Biol. Cell* **2013**, *24*, 683–691. [CrossRef]

61. Guo, Q.; Zhang, H.; Zhang, L.; He, Y.; Weng, S.; Dong, Z.; Wang, J.; Zhang, P.; Nao, R. MicroRNA-21 regulates non-small cell lung cancer cell proliferation by affecting cell apoptosis via COX-19. *Int. J. Clin. Exp. Med.* **2015**, *8*, 8835–8841.

62. Meng, L.; Xing, Z.; Guo, Z.; Liu, Z. LINC01106 post-transcriptionally regulates ELK3 and HOXD8 to promote bladder cancer progression. *Cell Death Dis.* **2020**, *11*, 1063. [CrossRef]

63. Jiang, H.; Xu, A.; Li, M.; Han, R.; Wang, E.; Wu, D.; Fei, G.; Zhou, S.; Wang, R. Seven autophagy-related lncRNAs are associated with the tumor immune microenvironment in predicting survival risk of nonsmall cell lung cancer. *Brief. Funct. Genomics* **2021**, *21*, 177–187. [CrossRef]

64. Shao, J.; Zhang, B.; Kuai, L.; Li, Q. Integrated analysis of hypoxia-associated lncRNA signature to predict prognosis and immune microenvironment of lung adenocarcinoma patients. *Bioengineered* **2021**, *12*, 6186–6200. [CrossRef]

65. Ma, G.; Li, G.; Fan, W.; Xu, Y.; Song, S.; Guo, K.; Liu, Z. The role of long noncoding RNA AL161431.1 in the development and progression of pancreatic cancer. *Front. Oncol.* **2021**, *11*, 666313. [CrossRef]

66. Gu, Z.R.; Liu, W. The LncRNA AL161431.1 targets miR-1252-5p and facilitates cellular proliferation and migration via MAPK signaling in endometrial carcinoma. *Eur Rev Med Pharmacol Sci* **2020**, *24*, 2294–2302.

67. Ju, Q.; Zhao, Y.J.; Ma, S.; Li, X.M.; Zhang, H.; Zhang, S.Q.; Yang, Y.M.; Yan, S.X. Genome-wide analysis of prognostic-related lncRNAs, miRNAs and mRNAs forming a competing endogenous RNA network in lung squamous cell carcinoma. *J. Cancer Res. Clin. Oncol.* **2020**, *146*, 1711–1723. [CrossRef]

68. Shen, D.; Zhang, Y.; Zheng, Q.; Yu, S.; Xia, L.; Cheng, S.; Li, G. A competing endogenous RNA network and an 8-lncRNA prognostic signature identify MYO16-AS1 as an oncogenic lncRNA in bladder cancer. *DNA Cell Biol.* **2021**, *40*, 26–35. [CrossRef]

69. Cruz, R.S.; Liu, X.; Chan, A.W.; Denniston, A.K.; Calvert, M.J. Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. *Nat. Med.* **2020**, *26*, 1351–1363. [CrossRef]