BMC Medicine

RESEARCH

Open Access

Unsupervised learning-based quantitative analysis of CT intratumoral subregions predicts risk stratification of bladder cancer patients

Ying Wang[1,2†], Hexiang Wang[3†], Na Li[4], Siyu Wu[5], Rongchao Shi[6], Kui Sun[7*] and Ximing Wang[1*]

Abstract

Background Preoperative diagnosis of muscle invasion and American Joint Committee on Cancer (AJCC) stage plays a crucial role in guiding treatment strategies for bladder cancer (BCa). Utilizing quantitative analysis of tumor subregions via CT imaging holds promise in identifying high-risk populations. Developing and evaluating the performance of an unsupervised clustering algorithm-based intratumoral subregion radiomics model for distinguishing between bladder muscle invasion and AJCC stage.

Methods This retrospective study included BCa patients who underwent CT imaging prior to transurethral resection of bladder tumor (TURBT) between January 2006 and December 2022, with all cases being histologically confirmed. Utilizing an unsupervised clustering learning method, the tumor region was classified into three intratumoral subregions, and radiomics features were extracted from both the whole tumor region and each intratumoral subregion. Following feature selection and nested cross-validation, seven predictive models were developed: clinicopathological model, whole-tumor model, intratumoral subregion 1 model, intratumoral subregion 2 model, intratumoral subregion 3 model, merged intratumoral subregion model, and fusion model. Predictive performance was evaluated using the area under the receiver operating characteristic curve (AUROC).

Results A total of 1017 patients were included in this study, with 778 from center A (training cohort; median age, 67 years [IQR, 69–75 years]) and 239 from center B (external validation cohort; median age, 58.5 years [IQR, 66–74 years]). In the external validation cohort, the fusion model AUROC was highest for predicting muscle invasion and AJCC stage tasks, at 0.884 (95% CI, 0.842, 0.926) and 0.832 (95% CI, 0.768, 0.897), respectively. The merged intratumoral subregion model outperformed the whole-tumor model (AUROC, muscle invasion, 0.871 vs 0.804, $p=0.004$; AJCC stage, 0.832 vs 0.804, $p=0.4$).

Conclusions The subregional radiomics model based on preoperative CT shows potential value in predicting muscle invasion and AJCC stage. This noninvasive risk stratification tool may offer supportive insights into surgical approaches and treatment decisions for BCa patients.

Keywords Bladder cancer, Muscle invasion, American Joint Committee on Cancer, Subregion, Radiomics

†Ying Wang and Hexiang Wang contributed equally to this work.

*Correspondence:
Kui Sun
sksunkui@163.com
Ximing Wang
wxming369@163.com
Full list of author information is available at the end of the article

Background

Bladder cancer (BCa) ranks as the second most prevalent urological malignancy worldwide, with approximately 549,000 new cases and 200,000 deaths annually [1, 2]. In China alone, there are over 82,900 new cases and 41,400 deaths reported each year [3]. Accurate assessment of muscle invasion and American Joint Committee on Cancer (AJCC) stage is pivotal in guiding treatment strategies and prognostic evaluation, significantly impacting patient survival [4, 5]. Multiphase computed tomography (MPCT) imaging, owing to its remarkable spatial resolution, widespread availability, and rapid performance, has become a routine diagnostic procedure. It serves as the primary method for evaluating muscle invasion, lymph nodes, and distant metastases in BCa [6–8]. Consequently, CT has emerged as a cornerstone in the preoperative assessment of BCa patients.

Despite its many advantages, CT is limited in its ability to accurately assess muscle invasion in BCa. This imaging assessment shows variability among observers with different levels of experience and across different institutions. While preoperative biopsy serves as the gold standard for diagnosing muscle invasion, it often underestimates the extent of tumor invasion into the bladder wall due to the inability to sample every part of the tumor. Additionally, cystoscopy does not provide visualization of peripheral invasion and metastasis of the tumor [9]. Given the high variability of biopsy sampling, there is a need to explore new methods for the preoperative evaluation of BCa.

Intratumoral heterogeneity (ITH) is a significant mechanism influencing tumor survival and evolution and has been linked to patient survival and prognosis across various cancer types, including BCa [10, 11]. Therefore, comprehending the underlying biological mechanisms of BCa and the heterogeneity during its progression is crucial for the prognosis and precise treatment of patients with this condition. Radiomics techniques, capable of automatically extracting high-throughput quantitative information from medical images, have demonstrated utility in disease diagnosis and prognostic assessment [12–14]. Furthermore, segmenting tumor lesions into subregions using unsupervised cluster analysis allows for a detailed characterization of cancer lesions by extracting additional radiomics features from these subregions, rather than solely from the entire lesion [15, 16]. This radiomics analysis based on different subregions can provide insights into the diversity of ITH from multiple perspectives [17]. For instance, Shi et al. [18] developed a radiomics predictive model based on MR breast tumor subregions, establishing a quantitative measure of ITH that predicted the area under the receiver operating characteristic curves (AUROCs) range of 0.83–0.87 for pathologic complete remission after neoadjuvant chemotherapy.

Therefore, the objective of this study was to develop a subregional radiomics predictive model utilizing unsupervised cluster analysis to unveil ITH in BCa. Additionally, we assessed whether a model integrating clinicopathological variables and subregional radiomics could accurately predict both muscle invasion and AJCC stage based on CT imaging.

Methods

Ethics

All participating centers' Institutional Review Board approved the study protocols (Center A, No. QYFY WZLL 28166; Center B, No.2023–526). The informed consent was waived because of the two-center retrospective nature of this study and due to all procedures were part of routine care. After matching imaging and clinical data, our data was completely anonymized to delete any connections between data and patients' identity.

Patient enrollment

This retrospective study was conducted across two centers. Center A included 778 eligible patients who underwent MPCT scans from January 2006 to December 2021, serving as the training cohort. Center B enrolled 239 eligible patients who underwent MPCT scans from January 2018 to December 2022, constituting the external validation cohort. The inclusion criteria across all centers were as follows: (1) histologically confirmed uroepithelial carcinoma; and (2) standard pelvic MPCT scan performed within 20 days prior to transurethral resection of bladder tumor (TURBT). Exclusion criteria comprised the following: (1) patients with other concomitant tumors; (2) patients with missing or poor-quality CT scan results; (3) tumors not detected on preoperative enhanced CT images; (4) missing clinicopathological variables; and (5) patients who are currently receiving or have previously received neoadjuvant chemotherapy or preoperative immunotherapy (Fig. 1).

Clinicopathological variables

The demographic data encompassed age and gender. Laboratory assessments comprised neutrophil count, lymphocyte count, neutrophil-to-lymphocyte ratio (NLR), cytokeratin-20 (CK20) levels, GATA binding protein 3 (GATA3) levels, presence of urinary infection, hematuria, and proteinuria. CK20 and GATA3 were done by immunohistochemistry (IHC), and the intensity were recorded. Additionally, a morphological indicator, the presence of a stalk [19], was considered.

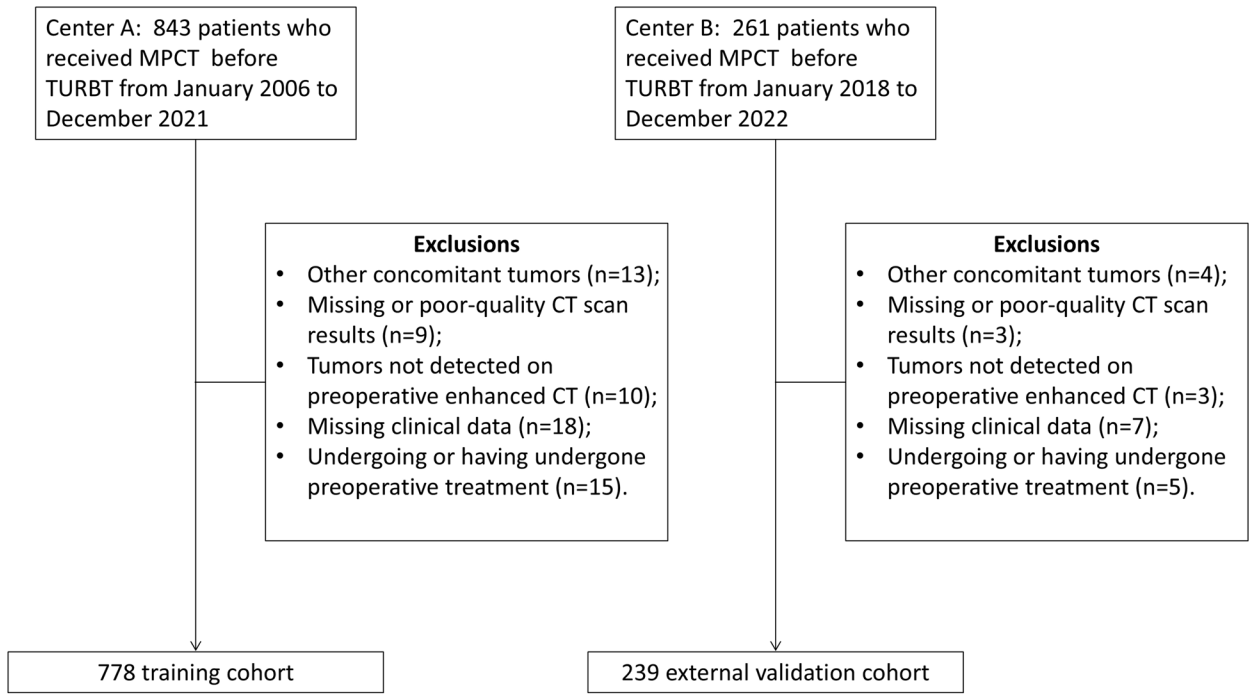

Fig. 1 Inclusion and exclusion workflow

Outcome definition

This study included two analysis endpoints. The first aimed to predict the presence of muscular invasion, distinguishing between non-muscle invasive BCa (NMIBC) and muscle invasive BCa (MIBC). This endpoint relied on postoperative pathology for assessment. The second endpoint focused on predicting AJCC stage, categorized into two groups: one comprising stages 0 A, 0is, I, and II (indicating the absence of extravesical bladder invasion), and the other comprising stages IIIA, IIIB, IVA, and IVB (indicating the presence of extravesical bladder invasion). The surgical modalities used for pathological T-staging of surgical specimens included TURBT, partial cystectomy, and radical cystectomy.

MPCT image acquisition protocol

The acquisition parameters for MPCT are detailed in Additional file 1: Table 1. Following the injection of contrast agent, arterial phase, venous phase, and delayed phase images were obtained at 25 s, 75 s, and 5 min, respectively, at an abdominal aortic Hounsfield Units (HUs) threshold of 120 (Fig. 2).

Volume of interests (VOIs) delineation, image preprocessing, best number of subregion determination, and subregion generated

VOIs encompassing the entire BCa tumor were manually delineated on the venous phase CT images by two radiologists (Y.W and N.L, with 5 and 10 years of experience in pelvic CT, respectively) using ITK-SNAP version 4.0.2 software (http://www.itksnap.org). Any discrepancies were resolved through consensus-based discussion, with appropriate adjustments made accordingly.

Before generating subregions using unsupervised learning algorithms, image preprocessing methods were applied. Firstly, all images were resampled to a voxel size of $1\times 1\times 1$ mm^3 using B-spline interpolation to ensure consistency. Subsequently, image intensities were normalized using min–max normalization (Equation).

$$x_{\text{new}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

The elbow method was employed to determine the optimal number of intratumoral subregions. Initially, a random subset comprising 1/5 of the samples ($n=160$) from the training cohort was selected to calculate the sum of squared errors (SSE). The optimal number of clusters was identified at the point where the SSE exhibited a significant decrease, indicating an inflection point or leveling off in the SSE decline, within a cluster range from 1 to 10 (Fig. 3). Upon determining the optimal number of clusters, the K-Means + + algorithm was applied to cluster the intratumoral subregions. Based on the SSE results, the optimal number of clusters was determined to be 3.

Table 1 Comparison of clinicopathological characteristics between the training cohort and the external validation cohort

Characteristics	Training cohort ($n = 778$)	External validation cohort ($n = 239$)	*P*
Age (y)[a]	67 (60–75)	58.5 (66–74)	0.222
Male (*n*, %)	626 (80.5)	200 (83.7)	0.265
Urinary infection (*n*, %)	188 (24.2)	40 (16.7)	0.016
Hematuria (*n*, %)	499 (64.1)	189 (79.1)	< 0.001
Proteinuria (*n*, %)	314 (40.4)	118 (49.4)	0.014
Lymphocyte (10^9/L)[a]	1.73 (1.26–2.16)	1.65 (1.25–1.98)	0.055
Neutrophil (10^9/L)[a]	4.19 (3.11–5.98)	3.52 (2.86–4.75)	0.027
NLR[a]	2.45 (1.68–3.95)	2.17 (1.59–3.11)	0.147
CK20 ≥ 5% (*n*, %)	254 (32.6)	47 (19.7)	< 0.001
GATA3 ≥ 10% (*n*, %)	253 (32.5)	43 (17.9)	< 0.001
Stalk (*n*, %)	157 (20.2)	145 (60.7)	< 0.001
AJCC stage (*n*, %)			0.025
≤ I	421 (54.1)	154 (64.4)	
II	238 (30.6)	47 (19.7)	
III	109 (14.0)	36 (15.1)	
IV	10 (1.3)	2 (0.8)	
Muscle invasion (*n*, %)	354 (45.5)	85 (35.6)	0.007

NLR Neutrophil-to-lymphocyte ratio, *CK20* Cytokeratin-20, *AJCC* American Joint Committee on Cancer

[a] Data are medians, with IQRs in parentheses

Subregion quantitative feature generation

Quantitative radiomics feature extraction was implemented using PyRadiomics version 3.1.0 (https://pyradiomics.readthedocs.io), which complies with the image biomarker standardization initiative. Before process, gray-level was discretized using a fixed bin width value of 25 with absolute discretization from −1000 to 3000 HUs. The feature extraction was performed separately for the whole tumor region and well-clustered intratumoral subregions. A total of 1218 features (14 shape-based, 252 Gy histogram-based, and 952 texture-based) were generated in each region separately, including the properties of the original filter, Laplacian of Gaussian (LoG) filter, and wavelet filter (Additional file 1: Fig. S1).

Reproducibility test and redundancy feature elimination

Forty patient images were randomly assigned for delineation by another radiologist (Rc.S, with 5 years of experience in Pelvic CT), and the extracted features were utilized to assess interobserver agreement. Features exhibiting inter-observer correlation coefficients (ICCs) > 0.75 were considered to possess good reproducibility and were consequently included in the subsequent analysis.

Feature selection comprised two steps: (1) Elimination of potential multicollinearity between features through correlation analysis, retaining features with a correlation coefficient < 0.9; (2) Implementation of the filtered feature screening method, whereby the k best features were selected by calculating the coefficient of determination, also known as R-square. The parameter k ranged from 1 to 10 to prevent overfitting.

For clinicopathological variables, the stepwise variable selection procedure for generalized linear model (GLM) was employed to optimize variable selection. Initially, the GLM was applied to the candidate variables using both stepwise forward and backward methods until the best model was achieved. The optimal model was determined based on a dual assessment of the p-value and variance inflation factor (VIF), meeting criteria of a p-value less than 0.05 and a VIF less than 2.5 for continuous variables and less than 10 for categorical variables.

Machine learning prediction model development

A range of machine learning algorithms were applied to fit the discovery dataset, including support vector machine (SVM), decision tree (DT), random forest (RF), gradient boosting (GB), logistic regression (LR), and K-Nearest Neighbors (KNN).

Training of the discovery dataset, hyperparameter tuning of the models and selection of the optimal model were conducted using a nested cross-validation algorithm. This approach comprised two components: inner five-fold cross-validation and outer five-fold cross-validation. Internal cross-validation was utilized to optimize the hyperparameters of the models by employing a grid search method to identify the optimal hyperparameters for each model. External cross-validation was then

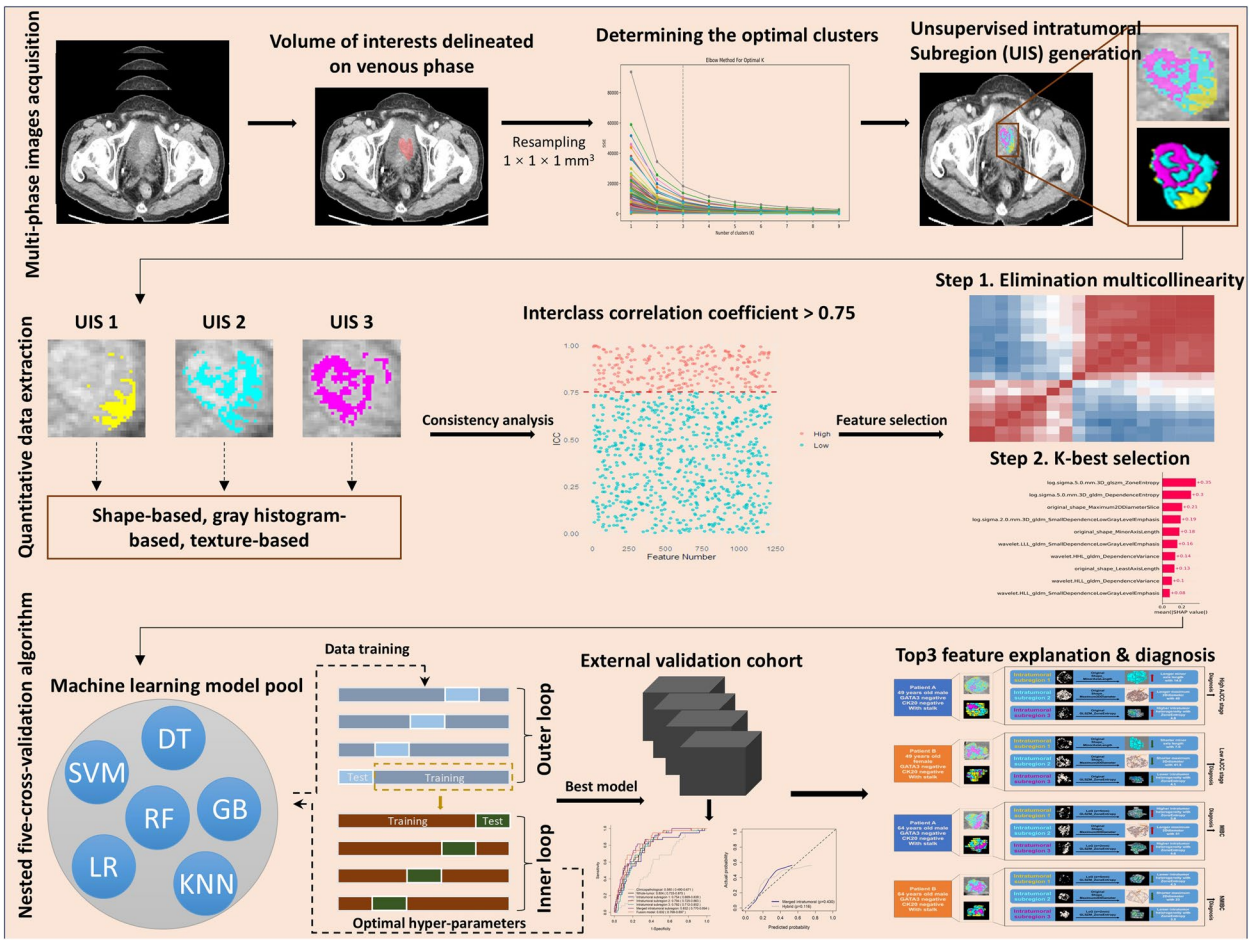

Fig. 2 The flowchart of this study. Section 1, image acquisition and delineation. Section 2, segmentation of intratumor subregions based on unsupervised learning. Section 3, quantitative radiomics data extraction and selection. Section 4, multiple machine learning model construction and external evaluation. Section 5, feature interpretation and diagnosis. DT, decision tree; SVM, support vector machine; RF, random forest; GB, gradient boosting; LR, logistic regression; KNN, k-nearest neighbor

employed to train the data on the model configured with the optimal hyperparameters. Following the completion of the training process, an external validation cohort was utilized to assess the generalization of the model and determine the best predictive model.

Ultimately, seven predictive models were generated: the clinicopathological model (comprising optimal clinicopathological variables), the whole-tumor model, intratumoral subregion 1 model, intratumoral subregion 2 model, intratumoral subregion 3 model, merged intratumoral subregion model, and a fusion model (combining clinicopathological, whole-tumor, and intratumoral subregion 1–3 models).

Statistical analysis

Continuous data were presented as median with interquartile range (IQR), while categorical data were presented as frequency with percentage. The predictive performance of the models was assessed using the AUROC and its 95% confidence interval, sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV). The calibration curve efficacy was evaluated using the Hosmer–Lemeshow goodness-of-fit test, with $p > 0.05$ indicating a good fit. Differences in predictive ability among models were assessed using the DeLong test, with $p < 0.05$ considered statistically significant. The SHapley Additive exPlanations (SHAP) method was employed to reveal the contribution of radiomics features. All analyses were conducted using R version 4.2.2 (https://www.r-project.org) and Python version 3.9.1 (https://www.python.org).

Results

Patient baseline

A total of 1017 patients diagnosed with BCa were included in this study, with 826 (81.2%) being male. In

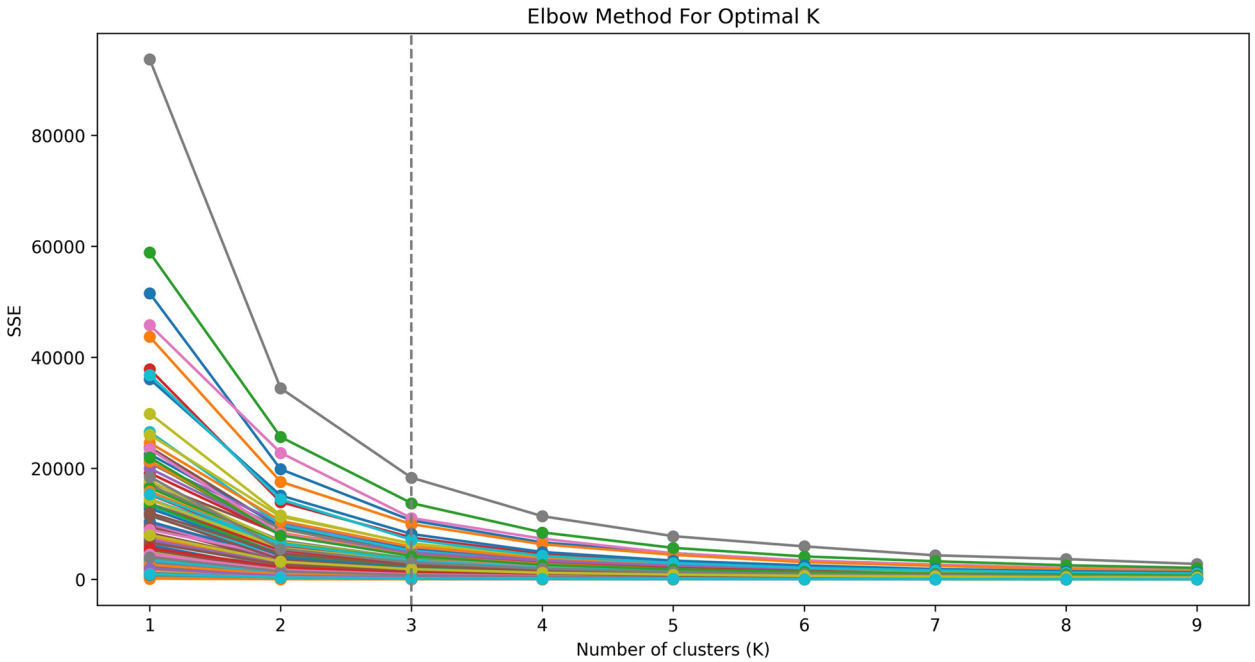

Fig. 3 The plot generated by the elbow method displays the number of clusters on the *x*-axis and the sum of squared errors (SSE) on the *y*-axis. A dashed vertical line is included to mark the inflection point where the SSE begins to level off, indicating the optimal number of clusters

the training cohort, the median age was 67 years (IQR, 60–75), with 354 (45.5%) of the 778 patients confirmed to have MIBC, and 119 (15.3%) presenting with an AJCC stage ≥ 3. The external validation cohort had a median age of 58.5 years (IQR, 66–74), with 85 (35.6%) of the 239 patients confirmed to have MIBC, and 38 (15.9%) presenting with an AJCC stage ≥ 3 (Table 1).

Optimal clinicopathological variables

Based on the results of the stepwise GLM, for the muscle invasion outcome, four clinicopathological variables were retained, namely CK20 (VIF: 1.27), GATA3 (VIF: 1.27), hematuria (VIF: 1.43), and stalk (VIF: 1.01). For the AJCC stage, six clinicopathological variables were retained, including age (VIF: 1.05), lymphocyte (VIF: 1.04), GATA3 (VIF: 1.01), hematuria (VIF: 1.31), proteinuria (VIF: 1.32), and stalk (VIF: 1.00). Further details are provided in Table 2.

Table 2 Optimal clinicopathological variables selected based on generalized linear model

Outcome	Characteristic	Training cohort ($n = 778$)			
		β coefficient	Wald	*P* value	VIF
Muscle invasion	CK20	− 0.56	8.93	0.003	1.27
	GATA3	1.48	63.15	< 0.001	1.27
	Hematuria	0.78	22.67	< 0.001	1.43
	Stalk	− 0.47	5.63	0.018	1.01
AJCC stage	Age	0.02	4.62	0.032	1.05
	Lymphocyte	− 0.36	5.13	0.024	1.04
	GATA3	0.70	11.30	< 0.001	1.01
	Hematuria	0.52	4.12	0.042	1.31
	Proteinuria	− 0.48	3.93	0.047	1.32
	Stalk	− 1.17	10.34	0.001	1.00

VIF Variance inflation factor, *AJCC* American Joint Committee on Cancer

Optimal radiomics features of predictive models

For the muscle invasion task, the distribution of features across different models was as follows: five features in the whole-tumor model, ten features in intratumoral subregion 1 model, one feature in intratumoral subregion 2 model, and ten features in intratumoral subregion 3 model. In the whole-tumor model, the majority of contributing features were textural (3 out of 5), while the remaining features were shape-related (2 out of 5). Similarly, in intratumoral subregion 1 model, textural features dominated (7 out of 10), with the rest being shape-related (3 out of 10). Conversely, only one shape-related feature (Original_Maximum2DDiameterSlice) contributed to intratumoral subregion 2 model. In intratumoral subregion 3 model, textural features were predominant (8 out of 10), with the remaining features being shape-related (2 out of 10).

For the AJCC stage task, the distribution of features across different models was as follows: eight features in the whole-tumor model, six features in intratumoral subregion 1 model, two features in intratumoral subregion 2 model, and five features in intratumoral subregion 3 model. In the whole-tumor model, texture-type features contributed the most (6 out of 8), while the remaining features were shape-related (2 out of 8). In intratumoral subregion 1 model, texture-type features (3 out of 6) were equally distributed with shape-type features (3 out of 6). In intratumoral subregion 2 model, both features were shape-related, namely, Original_MinorAxisLength and Original_Maximum2DDiameterSlice. In intratumoral subregion 3 model, texture-type features were predominant (4 out of 5), with one feature being shape-related (1 out of 5).

Model discriminability for muscle invasion

In discriminating between NMIBC and MIBC, the clinicopathological model achieved an AUROC of 0.695 (95% CI, 0.629, 0.761) in the external validation cohort. The whole-tumor model demonstrated an AUROC of 0.804 (95% CI, 0.745, 0.862). The predictive efficacy of the intratumoral subregion 1–3 models yielded AUROCs of 0.797 (95% CI, 0.738, 0.857), 0.813 (95% CI, 0.757, 0.870), and 0.810 (95% CI, 0.753, 0.868), respectively. Moreover, the merged intratumoral subregion model exhibited robust predictive performance, with an AUROC of 0.871 (95% CI, 0.826, 0.917). However, the predictive performance of the fusion model, incorporating multiple scale variables, did not significantly improve compared to the merged intratumoral subregion model, yielding an AUROC of 0.884 (95% CI, 0.842, 0.926; Fig. 4a). Calibration curve analysis revealed good agreement between predicted and actual probabilities for both the merged intratumoral subregion and fusion models (Fig. 4b).

Model discriminability for AJCC stage

The clinicopathological model maintained the lowest predictive performance, with an AUROC of 0.580 (95% CI, 0.490, 0.671). Conversely, the whole-tumor model demonstrated superior performance compared to each individual intratumoral subregion model (AUROC, 0.804 versus 0.754, 0.794, 0.782). Notably, the merged intratumoral subregion model exhibited improved predictive performance, with an AUROC of 0.832 (95% CI, 0.770, 0.894). However, similar to previous observations, the fusion model did not demonstrate significant improvement, yielding an AUROC of 0.832 (95% CI, 0.768, 0.897; Fig. 4c, Table 3). Calibration curve analysis indicated reasonable agreement between predicted and actual probabilities for both the merged intratumoral subregion and fusion models (Fig. 4d).

Comparison of predictive effectiveness

To assess whether the AUROC values of different prediction models were statistically different, the Delong test was employed. Regarding the muscle invasion task, the merged intratumoral subregion model outperformed other individual models, demonstrating robust predictive capability with statistically significant differences in AUROC comparisons. Specifically, the p-values were <0.001, 0.003, 0.006, 0.001, and 0.004 for the merged intratumoral subregion model versus the clinicopathological model, intratumoral subregion 1–3 models, and whole-tumor model, respectively.

For the AJCC stage task, the merged intratumoral subregion model exhibited statistically significant differences in AUROC comparisons with the clinicopathological model, intratumoral subregion 1, and 3 models, with p-values of <0.001, 0.04, and 0.02, respectively. However, no statistically significant differences were observed in comparisons between the merged intratumoral subregion model and the whole-tumor model (p-value, 0.4), as well as with intratumoral subregion 2 (p-value, 0.05).

Explanatory description

We employed the SHAP method to calculate the contributions of radiomic features from each intratumoral subregion for both tasks. In the muscle invasion task, the feature LoG ($\sigma=$ 5 mm) GLZSM_ZoneEntropy exhibited the highest contribution in intratumoral subregion 1, with a mean absolute Shapley value of 0.35. In intratumoral subregion 2, the feature Original_shape_Maximum2DDiameterSlice was the sole significant contributor, with a value of 0.43. Similarly, in intratumoral subregion 3, the feature LoG ($\sigma=$ 2 mm) GLZSM_ZoneEntropy demonstrated the highest contribution, with a value of 0.36 (Fig. 5a–c). For the AJCC task, the feature Original_shap_MinorAxisLength showed the highest contribution in intratumoral subregion 1, with an absolute Shapley value of 0.29. In intratumoral subregion 2, Original_shape_Maximum2DDiameterSlice was the most significant contributor, with a value of 0.20. In intratumoral subregion 3, the feature Original_GLZSM_ZoneEntropy had the highest contribution, with a value of 0.28 (Fig. 5d–f).

Figure 6 illustrates the radiomic interpretation of the intratumoral subregions in two patients with similar demographics, laboratory results, and morphological descriptions but differing diagnostic outcomes. The study results indicate that the thresholds for the top features—LoG ($\sigma=$ 5 mm) GLSZM_ZoneEntropy, Original Shape_Maximum2DDiameter, and LoG ($\sigma=$ 2 mm) GLSZM_ZoneEntropy—in intratumoral subregions 1, 2, and 3 were 4.5, 36.5, and 3.79, respectively. ZoneEntropy

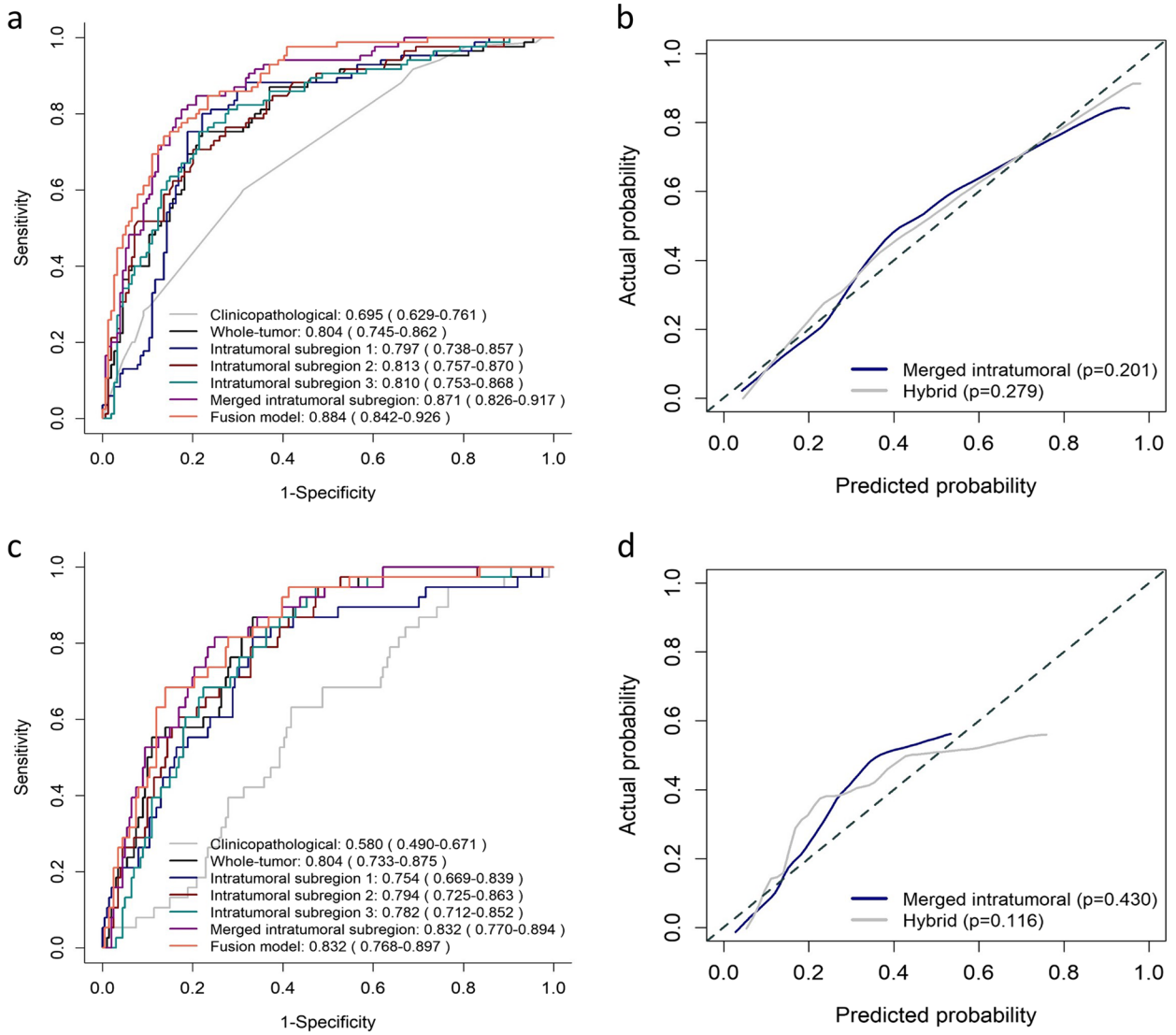

Fig. 4 The area under the receiver operating characteristic (AUROC) and calibration curve on the external validation set. AUROC comparison of multiple predictive models on the task of muscle invasion (**a**). Merged intratumoral subregion model and fusion model calibration curve plot for the muscle invasion task (**b**). AUROC comparison of multiple predictive models on the task of American Joint Committee on Cancer (AJCC) stage (**c**). Merged intratumoral subregion model and fusion model calibration curve plot for the AJCC stage task (**d**)

reflects the uncertainty or randomness in the distribution of zone sizes and gray levels, with higher values signifying greater heterogeneity. This suggests that, in patients with MIBC, subregional analysis indicates a trend toward increased intratumoral heterogeneity and larger maximum 2D diameters.

Figure 7 illustrates the radiomic interpretation of the intratumoral subregions for two patients with similar laboratory findings and morphological descriptions but differing AJCC risk stratifications. The findings indicate that the thresholds for the top features—Original Shape_MinorAxisLength, Original Shape_Maximum2DDiameter, and Original GLSZM_ZoneEntropy—in intratumoral subregions 1, 2, and 3 were 13.8, 42.1, and 4.14, respectively. The figure suggests that patients in the high-risk group tended to exhibit longer minor axis lengths, larger maximum 2D diameters, and greater intratumoral heterogeneity.

Discussion

In clinical practice, preoperative determination of MIBC and AJCC stage holds significant sway over treatment decisions for BCa. Muscle invasion status aids in selecting the appropriate surgical approach, while AJCC stage

Table 3 Diagnostic performance of different models for muscle invasion and AJCC stage tasks

Outcome	Model	AUROC (95%CI)	Sensitivity	Specificity	Accuracy	PPV	NPV
Muscle invasion	Clinicopathological	0.694 (0.629, 0.761)	0.600	0.689	0.657	0.515	0.757
	Whole-tumor	0.804 (0.745, 0.862)	0.753	0.779	0.770	0.653	0.851
	Intratumoral subregion 1	0.797 (0.738, 0.857)	0.800	0.779	0.787	0.667	0.876
	Intratumoral subregion 2	0.813 (0.757, 0.870)	0.706	0.799	0.766	0.659	0.831
	Intratumoral subregion 3	0.811 (0.753, 0.868)	0.753	0.786	0.774	0.660	0.852
	Merged intratumoral subregion	0.871 (0.826, 0.917)	0.847	0.792	0.812	0.692	0.904
	Fusion model	0.884 (0.842, 0.926)	0.847	0.766	0.795	0.667	0.901
AJCC stage	Clinicopathological	0.580 (0.490, 0.671)	0.632	0.582	0.590	0.222	0.893
	Whole-tumor	0.804 (0.733, 0.875)	0.868	0.667	0.699	0.330	0.964
	Intratumoral subregion 1	0.754 (0.669, 0.839)	0.816	0.667	0.690	0.316	0.950
	Intratumoral subregion 2	0.794 (0.725, 0.863)	0.947	0.522	0.590	0.272	0.981
	Intratumoral subregion 3	0.782 (0.712, 0.852)	0.842	0.637	0.669	0.305	0.955
	Merged intratumoral subregion	0.832 (0.770, 0.894)	0.816	0.751	0.762	0.383	0.956
	Fusion model	0.832 (0.768, 0.897)	0.684	0.861	0.833	0.481	0.935

AUROC Area under the receiver operating characteristic curve, *PPV* Positive Predictive value, *NPV* Negative predictive value, *AJCC* American Joint Committee on Cancer

influences subsequent treatment choices, such as the administration of adjuvant therapies like cisplatin [20, 21]. In this study, we conducted a radiomics analysis of tumor subregions generated through an unsupervised learning algorithm applied to CT images, leading to the establishment of multiple prediction models. Our findings revealed that the optimal fusion model accurately predicted MIBC and AJCC stage, achieving an AUROC of 0.884 and 0.832, respectively. Furthermore, the intratumoral subregion radiomics model demonstrated robust predictive performance for both muscle invasion and AJCC stage in BCa.

The gold standard for evaluating the extent of bladder cancer invasion remains pathohistology following TURBT. Clinical staging, as well as lymph node involvement or organ metastasis, can be assessed through imaging modalities. While current guidelines have adopted the Vesical Imaging–Reporting and Data System (VI-RADS), based on multiparametric MRI, to estimate the likelihood of detrusor muscle invasion (VI-RADS scores of 1 and 2 indicate low probability of MIBC, 3 indicates equivocal likelihood, 4 indicates MIBC is likely to be present, and 5 indicates MIBC or invasion beyond the bladder wall), visual assessment of the depth of invasion in the bladder wall remains a significant challenge. In recent years, studies have increasingly focused on developing artificial intelligence models to predict muscle invasion in BCa using radiomics [22]. Zhang et al. [23] employed contrast-enhanced CT-based radiomics to predict MIBC in 441 BCa patients, achieving an AUROC of 0.784 in the external test set. Wei et al. [24] gathered data from 323 BCa patients and constructed a nomogram integrating deep learning features and radiomics features, yielding an AUROC of 0.862 in the external validation set. Wang et al. [25] recruited 106 patients with BCa who underwent preoperative MRI scanning. Their study utilized T2 and multi-b-value DWI images for radiomics analysis, achieving an AUROC of 0.813 for the predictive model in the validation set. When merged with MRI-determined tumor stalk, the fusion model attained an AUROC of 0.877. Despite these studies utilizing radiomics features to predict MIBC, none have analyzed the heterogeneity of different subregions within a BCa tumor. In our study, an unsupervised clustering learning method was utilized to categorize the tumor region into three subregions for analysis, demonstrating the potential predictive value of subregions. In diagnosing MIBC, the AUROCs of the subregions ranged from 0.797 to 0.813. The predictive efficacy of the merged model of the three subregions surpassed that of the whole tumor and was statistically significant, underscoring the importance of subregion analysis.

Regarding diagnostic AJCC stage, the AUROCs for the three subregions ranged from 0.754 to 0.782, with the merged intratumoral subregion model still exhibiting a higher AUROC (0.832) than the whole tumor model (0.804). Previous studies have primarily focused on radiomics analysis of muscle invasion to identify MIBC, which serves as a predictive risk factor [23–25]. However, categorizing lesions without muscle invasion but presenting with other risk factors (e.g., lymph node or distant metastasis) as low-risk can be inaccurate, even though in a minority of cases. Distinguishing AJCC stages 0–2 from higher stages may hold more clinical significance than studies solely assessing the presence or absence of muscular invasion without explicitly considering AJCC

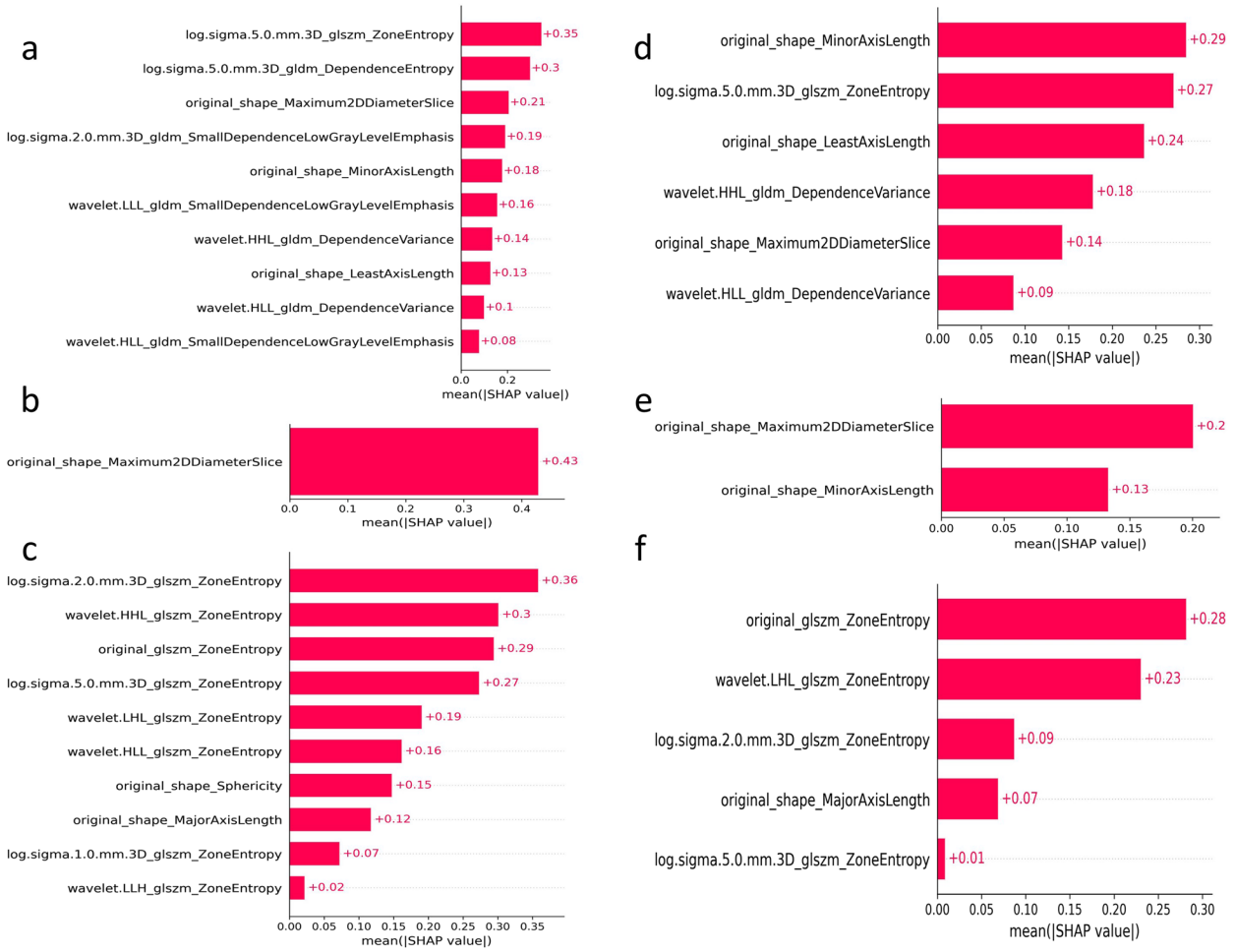

Fig. 5 Feature contribution maps based on SHapley Additive exPlanations (SHAP) are presented. The *x*-axis represents SHAP feature importance, measured as the mean absolute Shapley values, while the *y*-axis indicates the radiomic features of the model. For the muscle invasion task, the contribution maps for intratumoral subregions 1 (**a**), 2 (**b**), and 3 (**c**) are displayed. For the AJCC task, the contribution maps for intratumoral subregions 1 (**d**), 2 (**e**), and 3 (**f**) are shown

stage. For instance, Wu et al. [26] analyzed MRI images of 103 BCas and developed a radiomics nomogram to predict preoperative lymph node metastasis, achieving an AUROC of 0.844 in the validation set. Similarly, Sun et al. [27] collected data from 239 BCa patients and constructed a fusion model incorporating deep learning, radiomics, and clinical features, yielding an AUROC of 0.834. However, current studies are predominantly focused on the status of muscular invasion or lymph node metastasis [28, 29], with very few analyzing AJCC stage.

The differences among the intratumoral subregions can be attributed to the distinct radiomic information they each reflect. Specifically, intratumoral subregion 1 appears to be distributed at the edges in a wrapped pattern for both the muscle invasion and AJCC tasks. The radiomic features associated with this subregion are characterized as "mixed," with the top three contributing features being both texture-based and shape-based. This mixed composition may partly explain its slightly lower predictive performance in similar models, with AUC values of 0.797 for the muscle invasion task and 0.754 for the AJCC task. In contrast, intratumoral subregion 2 is located closer to the center and consists of fewer radiomic features—only one for the muscle invasion task and two for the AJCC task—all of which are shape-related measures (diameter or axis length). It is reasonable to suspect that this may correlate with the degree and depth of tumor invasion. The resulting "homogeneity" contributes to its high predictive performance, with AUC values of 0.813 for the muscle invasion task and 0.794 for the AJCC task. We note that the distribution of intratumoral subregion 3 is relatively centralized and encapsulated, with features that are almost entirely comprised of ZoneEntropy. As mentioned above, larger values of ZoneEntropy indicate

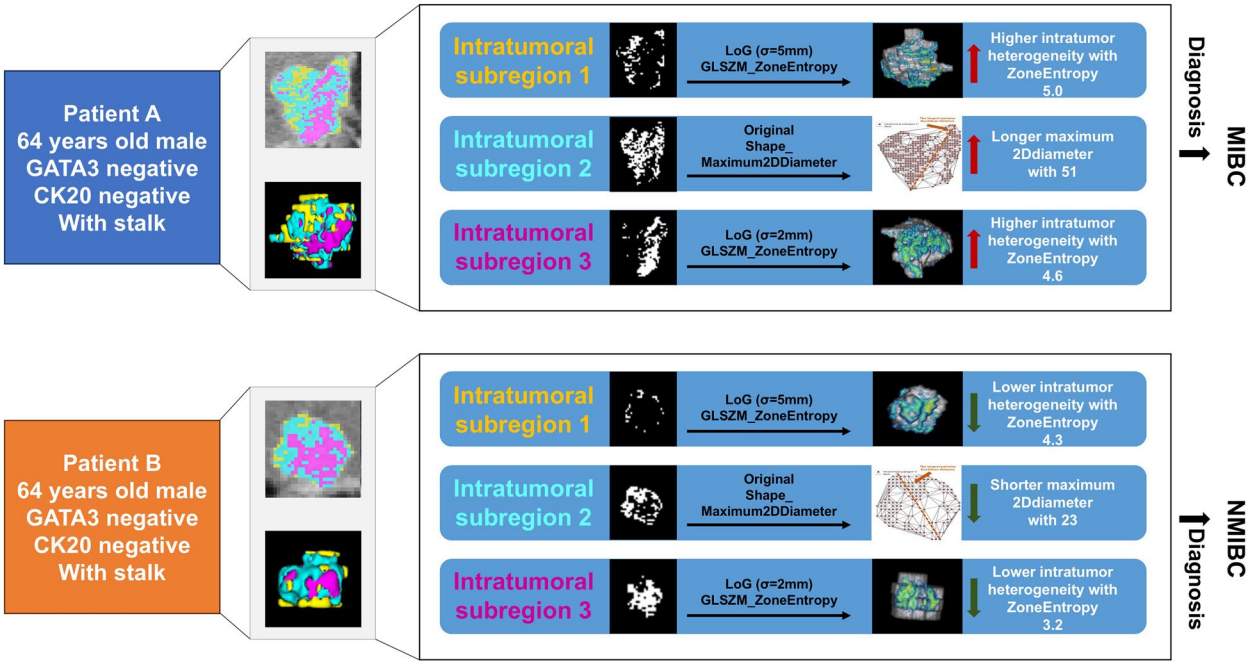

Fig. 6 Interpretation plot of intratumoral subregion radiomics of muscle invasion task. The figure features two patients with comparable age, gender, laboratory test results, and morphological descriptions but differing outcomes. Comparing a patient diagnosed with non-muscular invasive bladder cancer (NMIBC) to one with muscular invasive bladder cancer (MIBC), the analysis of MIBC subregions indicates a tendency towards increased intratumoral heterogeneity in subregions 1 and 3, alongside longer maximum 2D diameters observed in subregion 2

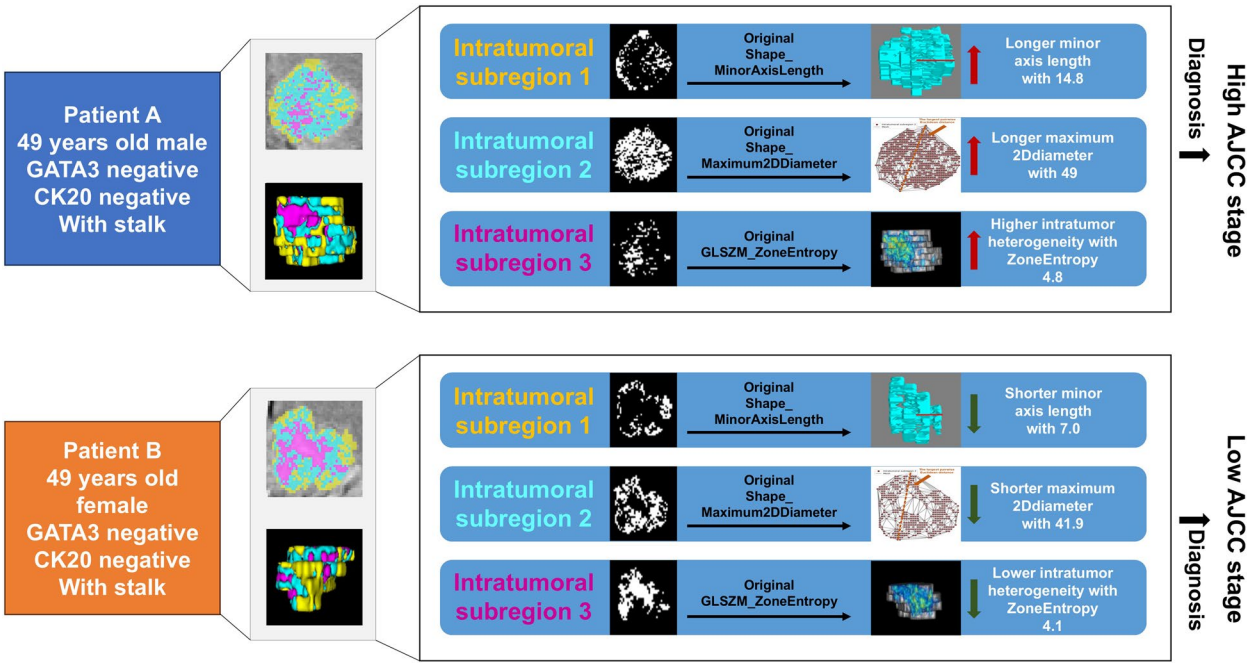

Fig. 7 The interpretation plot depicts the intratumoral subregion radiomics analysis concerning the American Joint Committee on Cancer (AJCC) stage task. Contrasting patients with low AJCC stage to those with high AJCC stage, the analysis reveals that high AJCC stage subregions tend to exhibit a longer minor axis length in subregion 1, an increased maximum 2D-diameter in subregion 2, and higher intratumoral heterogeneity in subregion 3

greater heterogeneity in texture patterns. Therefore, intratumoral subregion 3 may be particularly sensitive to heterogeneity within bladder cancer, potentially linked to the tumor microenvironment, which warrants further investigation. Considering that differences in clinicopathological characteristics between cohorts could potentially impact the validation process, subgroup analyses were conducted to assess the credibility of the predictive model. The results showed that for the muscle invasion task, the model's predictive performance remained robust, with an AUROC range of 0.801–0.930. For the AJCC task, the model continued to perform stably, with an AUROC range of 0.779–0.989 (Additional file 1: Fig. S2).

This study has several limitations. Firstly, considering the possible limitations in comparability between the two cohorts, arising from the retrospective nature of this study and the unavoidable selection bias, prospective studies are necessary to further validate the reliability of the findings. Secondly, relying solely on venous phase CT images may not fully capture the ITH of BCa. However, the venous phase is widely regarded as optimal for detecting bladder tumors, ensuring the stability of the extracted features [30]. Thirdly, considering that preoperative treatment may result in inaccurate staging or misdiagnosis of the depth of invasion into the bladder wall, patients who received neoadjuvant therapy and preoperative immunotherapy prior to CT scanning were excluded from this study. This exclusion may limit the model's applicability in clinical practice. Lastly, manual delineation of VOIs by different radiologists is time-consuming and labor-intensive. Implementing automated and precise tumor segmentation methods would enhance the efficiency and consistency of quantitative imaging analysis.

Conclusions

This study investigated the potential of quantitative analysis of intratumoral subregions, segmented by an unsupervised learning algorithm on preoperative CT scans, to predict muscle invasion and AJCC stage in BCa patients. The findings underscored the predictive capability of individual intratumoral subregion models across both tasks. Particularly noteworthy was the superior predictive performance of the merged intratumoral subregion model compared to the single whole-tumor model.

Abbreviations

BCa	Bladder cancer
AJCC	American joint committee on cancer
MPCT	Multi-phase computed tomography
ITH	Intratumoral heterogeneity
AUROC	Area under the receiver operating characteristic curve
TURBT	Transurethral resection of bladder tumor
NMIBC	Non-muscle invasive bladder cancer
MIBC	Muscle invasive bladder cancer

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-025-04163-2.

Additional file 1: Table 1 - [CT acquisition parameters]. Figures S1–S2. Fig S1 –[Schematic representation of the CT image preprocessing and radiomic feature extraction pipeline]. Fig S2 –[Subgroup analysis was performed based on optimal clinicopathologic characteristics].

Acknowledgements

This work was supported by the National Natural Science Foundation of China.

Authors' contributions

YW: Methodology, Writing – original draft. HW: Writing – review & editing. NL: Data curation. SW: Investigation, Formal analysis. RS: Formal analysis. KS: Visualization, Writing – review & editing. XW: Funding acquisition. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (No. 82271993 and 82471978).

Data availability

The data generated in this study are available upon request from the corresponding author. The underlying code for this study is available in https://github.com/Leon-skops/Unsupervised-learning-based-quantitative-analysis-for-bladder-cancer.

Declarations

Ethics approval and consent to participate

All participating centers' Institutional Review Board approved the study protocols (The Affiliated Hospital of Qingdao University, No. QYFY WZLL 28166; Shandong Provincial Hospital, No.2023–526). The informed consent was waived because of the two-center retrospective nature of this study and due to all procedures being part of routine care.

Consent for publication

N/A.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Radiology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jing Wu Road, No. 324, Jinan 250021, China. [2]School of Radiology, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian 271016, China. [3]Department of Radiology, The Affiliated Hospital of Qingdao University, 16 Jiangsu Road, Qingdao, Shandong, China. [4]Department of Radiology, The People's Hospital of Zhangqiu Area, Jinan 250200, China. [5]Department of Radiology, Shandong Provincial Hospital, No. 324, Jingwu Road, Jinan 250021, Shandong, China. [6]Department of Radiology, Beijing Friendship Hospital, Capital Medical University, Beijing, China. [7]Department of General Surgery, Peking University Third Hospital, 49 North Garden Road, Haidian District, Beijing 100191, China.

Received: 11 August 2024 Accepted: 26 May 2025
Published online: 02 June 2025

References

1. Compérat E, Amin MB, Cathomas R, et al. Current best practice for bladder cancer: a narrative review of diagnostics and treatments. Lancet. 2022;400(10364):1712–21.
2. Lenis AT, Lec PM, Chamie K, Mshs MD. Bladder cancer: a review. JAMA. 2020;324(19):1980–91.

3. Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent. 2024;4(1):47–53.
4. Patel VG, Oh WK, Galsky MD. Treatment of muscle-invasive and advanced bladder cancer in 2020. CA Cancer J Clin. 2020;70(5):404–23.
5. Powles T, Bellmunt J, Comperat E, et al. Bladder cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann Oncol. 2022;33(3):244–58.
6. Sadow CA, Silverman SG, O'Leary MP, Signorovitch JE. Bladder cancer detection with CT urography in an Academic Medical Center. Radiology. 2008;249(1):195–202.
7. Hensley PJ, Panebianco V, Pietzak E, et al. Contemporary staging for muscle-invasive bladder cancer: accuracy and limitations. Eur Urol Oncol. 2022;5(4):403–11.
8. Lobo N, Afferi L, Moschini M, et al. Epidemiology, screening, and prevention of bladder cancer. Eur Urol Oncol. 2022;5(6):628–39.
9. Alfred Witjes J, Max Bruins H, Carrión A, et al. European Association of Urology guidelines on muscle-invasive and metastatic bladder cancer: summary of the 2023 guidelines. Eur Urol. 2024;85(1):17–31.
10. Griffin JL. Devil in the detail: intratumour heterogeneity and personalised medicine for bladder cancer. Eur Urol. 2019;75(1):23–4.
11. Beckabir W, Wobker SE, Damrauer JS, et al. Spatial relationships in the tumor microenvironment demonstrate association with pathologic response to neoadjuvant chemoimmunotherapy in muscle-invasive bladder cancer. Eur Urol. 2024;85(3):242–53.
12. Huisman M, Akinci DT. What a radiologist needs to know about radiomics, standardization, and reproducibility. Radiology. 2024;310(2):e232459.
13. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441–6.
14. Sun K, Shi L, Qiu J, Pan Y, Wang X, Wang H. Multi-phase contrast-enhanced magnetic resonance image-based radiomics-combined machine learning reveals microscopic ultra-early hepatocellular carcinoma lesions. Eur J Nucl Med Mol Imaging. 2022;49(8):2917–28.
15. Cai Z, Xu Z, Chen Y, et al. Multiparametric MRI subregion radiomics for preoperative assessment of high-risk subregions in microsatellite instability of rectal cancer patients: a multicenter study. Int J Surg. 2024;3(18):10–1097.
16. Xie C, Yang P, Zhang X, et al. Sub-region based radiomics analysis for survival prediction in oesophageal tumours treated by definitive concurrent chemoradiotherapy. EBioMedicine. 2019;44:289–97.
17. Zhang X, Su GH, Chen Y, Gu YJ, You C. Decoding intratumoral heterogeneity: clinical potential of habitat imaging based on radiomics. Radiology. 2023;309(3):e232047.
18. Shi Z, Huang X, Cheng Z, et al. MRI-based quantification of intratumoral heterogeneity for predicting treatment response to neoadjuvant chemotherapy in breast cancer. Radiology. 2023;308(1):e222830.
19. Kang KA, Kim MJ, Kwon GY, Kim CK, Park SY. Computed tomography-based prediction model for identifying patients with high probability of non-muscle-invasive bladder cancer. Abdom Radiol (NY). 2024;49(1):163–72.
20. Lopez-Beltran A, Cookson MS, Guercio BJ, Cheng L. Advances in diagnosis and treatment of bladder cancer. BMJ. 2024;384:e076743.
21. Flaig TW, Spiess PE, Abern M, et al. NCCN Guidelines® Insights: Bladder Cancer, Version 2.2022. J Natl Compr Canc Netw. 2022;20(8):866–878.
22. Kozikowski M, Suarez-Ibarrola R, Osiecki R, et al. Role of radiomics in the prediction of muscle-invasive bladder cancer: a systematic review and meta-analysis. Eur Urol Focus. 2022;8(3):728–38.
23. Zhang G, Wu Z, Zhang X, et al. CT-based radiomics to predict muscle invasion in bladder cancer. Eur Radiol. 2022;32(5):3260–8.
24. Wei Z, Liu H, Xv Y, et al. Development and validation of a CT-based deep learning radiomics nomogram to predict muscle invasion in bladder cancer. Heliyon. 2024;10(2):e24878.
25. Wang H, Xu X, Zhang X, et al. Elaboration of a multisequence MRI-based radiomics signature for the preoperative prediction of the muscle-invasive status of bladder cancer: a double-center study. Eur Radiol. 2020;30(9):4816–27.
26. Wu S, Zheng J, Li Y, et al. Development and validation of an MRI-based radiomics signature for the preoperative prediction of lymph node metastasis in bladder cancer. EBioMedicine. 2018;34:76–84.
27. Sun R, Zhang M, Yang L, et al. Preoperative CT-based deep learning radiomics model to predict lymph node metastasis and patient prognosis in bladder cancer: a two-center study. Insights Imaging. 2024;15(1):21.
28. Starmans MPA, Ho LS, Smits F, et al. Optimization of preoperative lymph node staging in patients with muscle-invasive bladder cancer using radiomics on computed tomography. J Pers Med. 2022;12(5):726.
29. Gresser E, Woźnicki P, Messmer K, et al. Radiomics signature using manual versus automated segmentation for lymph node staging of bladder cancer. Eur Urol Focus. 2023;9(1):145–53.
30. Kim JK, Park SY, Ahn HJ, Kim CS, Cho KS. Bladder cancer: analysis of multi-detector row helical CT enhancement pattern and accuracy in tumor detection and perivesical staging. Radiology. 2004;231(3):725–31.

Publisher's Note