A semi-automatic deep learning model based on biparametric MRI scanning strategy to predict bone metastases in newly diagnosed prostate cancer patients

Song Xinyang[1,2], Shen Tianci[1], Hu Xiangyu[1], Zhang Shuang[3], Wang Yangyang[4], Du Mengying[1], Xu Tonghui[2]*, Zhou Jingran[1]* and Yang Feng[1]*

[1]Department of Radiology, Xiangyang No. 1 People's Hospital, Hubei University of Medicine, Xiangyang, China, [2]Department of Radiology, The People's Hospital of Zouping City, Zouping, China, [3]Department of Radiology, Xiangyang Central Hospital, Affiliated Hospital of Hubei University of Arts and Science, Xiangyang, China, [4]Department of Orthopedics, Xiangyang No. 1 People's Hospital, Jinzhou Medical University Union Training Base, Xiangyang, China

Objective: To develop a semi-automatic model integrating radiomics, deep learning, and clinical features for Bone Metastasis (BM) prediction in prostate cancer (PCa) patients using Biparametric MRI (bpMRI) images.

Methods: A retrospective study included 414 PCa patients (BM, n=136; NO-BM, n=278) from two institutions (Center 1, n=318; Center 2, n=96) between January 2016 and December 2022. MRI scans were confirmed with BM status via PET-CT or ECT pre-treatment. Tumor areas on bpMRI images were delineated as tumor's region of interest (ROI) using auto-delineation tumor models, evaluated with Dice similarity coefficient (DSC). Samples were auto-sketched, refined, and used to train the ResNet BM prediction model. Clinical, radiomics, and deep learning data were synthesized into the ResNet-C model, evaluated using receiver operating characteristic (ROC).

Results: The auto-segmentation model achieved a DSC of 0.607. Clinical BM prediction's internal validation had an accuracy (ACC) of 0.650 and area under the curve (AUC) of 0.713; external cohort had an ACC of 0.668 and AUC of 0.757. The deep learning model yielded an ACC of 0.875 and AUC of 0.907 for the internal, and ACC of 0.833 and AUC of 0.862 for the external cohort. The Radiomics model registered an ACC of 0.819 and AUC of 0.852 internally, and ACC of 0.885 and AUC of 0.903 externally. ResNet-C demonstrated the highest ACC of 0.902 and AUC of 0.934 for the internal, and ACC of 0.885 and AUC of 0.903 for the external cohort.

Conclusion: The ResNet-C model, utilizing bpMRI scanning strategy, accurately assesses bone metastasis (BM) status in newly diagnosed prostate cancer (PCa) patients, facilitating precise treatment planning and improving patient prognoses.

1 Introduction

Prostate cancer (PCa) ranks as the second most common cancer in men globally (1). While endocrine therapy for PCa boasts a 5-year and 10-year survival rate of 98% and 78% (2), the onset of bone metastasis (BM) marks a decline in patient prognosis. BM can trigger a range of skeletal complications, such as hypercalcemia, pathological fractures, spinal compression, and bone pain, all of which exacerbate prognosis and heighten mortality in advanced cases (2, 3). Managing BM patients poses challenges for clinicians, with only systemic palliative care options currently available (4). Hence, precise BM prediction in PCa patients is crucial for enhancing survival quality.

Research indicates that tailored interventions for bone metastases can extend patient survival (5–9). However, the European Urological Association (EUA) guidelines classify patients with prostate-specific antigen (PSA) levels exceeding 10 ng/ml as high-risk for BM, yet their BM positive rate at follow-up is a mere 7% (10). This hints at the inadequacy of clinical features as BM risk predictors. Some retrospective studies have employed PSA-dominant clinical risk prediction models to forecast the occurrence of BM in PCa. However, the results of these studies exhibit instability (AUC=0.68–0.88) (11–16), underscoring the need for a reliable diagnostic method.

Radiomics, introduced by Lambin in 2012 (17), shows potential in predicting BM in PCa by quantitatively analyzing target regions in images. Wang, Zhang et al. first used radiomics modeling for BM prediction predictions with an attempted area under the curve (AUC) value of 0.91 (18, 19). Numerous studies have unequivocally affirmed the value of MRI in the diagnosis of prostate cancer, markedly enhancing the accuracy of PCa diagnosis (20, 21). Yet, recent evidence suggests that Dynamic Contrast Enhance (DCE) sequences may not significantly impact clinical decision-making or patient benefit (21–25). The biparametric magnetic resonance imaging (bpMRI), as an emerging scanning strategy, showcases its advantages primarily by simplifying the selection of imaging parameters while retaining sufficient information for effective medical diagnosis. It not only reduces the invasiveness and potential risks of allergic reactions to patients but also brings about higher socio-economic benefits. For patients with contraindications to contrast agents (such as renal failure), it offers more possibilities. Simultaneously, due to the alteration in imaging sequences, novel predictive models are required to adapt to this change (19). Leveraging artificial intelligence technologies such as Deep learning (DL) to autonomously derive quantitative representations from medical images is an evolving direction in radiomic research (26–28). DL technology is gaining popularity as a creative tool due to its hierarchical network structure, which possesses a high capacity for memory and the ability to analyze abstract features. This capability makes it feasible to identify tumor regions. Concurrently, transfer learning and pre-trained DL networks facilitate the execution of new tasks, enabling the application of small datasets (25–27, 29, 30). DL methods based on convolutional neural networks are an emerging approach for predicting tumor metastasis with great potential (31–35). Until now, no authoritative statement or consensus has suggested that DL is inherently superior and replaces radiomics. Therefore, the combination of DL with radiomic features may lead to exceptional performance in predicting BM.

While using the mp-MRI imaging radiomics method to differentiate between NO-BM and BM states has been examined in several articles (18, 19), the development of predictive models based on bpMRI images remains unexplored. In this study, we endeavor to distinguish BM states in PCa patients non-invasively using DL combined with Radiomics, exclusively employing bpMRI images.

2 Materials and methods.

2.1 Patient information and clinical data

This study was conducted in accordance with the Helsinki Declaration and was approved by the ethics committees of both institutions. Due to retrospective study, the requirement for patient informed consent was waived.

Center 1 collected 318 patients which, including 95 patients with BM and 223 patients with PCa without BM (Non-BM patients), divided into an internal test cohort, validation cohort, and a training cohort according to a ratio of 3:1:6 between January 2016 and December 2022. The Center 2 collected data from 96 patients between March 2018 and December 2022, including 41 BM patients and 55 Non-BM patients as an external test cohort.

Inclusion criteria were as follows: (a) medical information is complete; (b) PCa confirmation through radical surgery or puncture biopsy; (c) BM status ascertained using bone scan or PET-CT (13); (d) Image scanning time and diagnostic time not exceeding two weeks. Exclusion criteria were: (a) Metastasis from other tumor sites; (b) Undergoing surgical or hormonal therapy; (c) poor image quality.

Patient age and clinical symptoms were acquired from medical histories. Clinical data encompassed: (a) Age; (b) PSA levels before intervention therapy; (c) hematuria and Leukocyturia results; (d) GS and ISUP classification; Immunohistochemistry and imaging data. PET-CT or bone scans were employed to distinguish the BM state.

Pathology sampling and grading involved sending patients for pathological examination with 3–6 biopsy pieces via rectal ultrasound puncture. GS and ISUP classifications were determined following PCa diagnosis confirmation. The comprehensive screening procedure is illustrated in Figure 1.

2.2 Experimental equipment

2.2.1 Imaging parameters

Images of patients within the first institution underwent scanning using a 3.0T MRI (750 W 3 T; GE Medical Systems). Using an eight-channel abdominal coil, imaging parameters included: T2-weighted imaging (T2WI) Sequence type: FSE, TR: 2500–5400: msec, TE: 60–120 msec, matrix: 512×512, slice thickness: 4 mm, FOV: 20 cm, resolution:0.39 mm/pixel, flip angle: 90°. Diffusion-weighted imaging (DWI): b: 0,800and 1400 s/mm^2, TR: 5000–6000 msec, TE: 80–120 msec, matrix: 256×256, slice thickness: 4 mm, FOV: 20 cm, resolution: 0.78 mm/pixel, flip angle: 90°.

Images of patients within the second institution underwent scanning using a 3.0T MRI system (Siemens Magnetom Verio 3.0T; Siemens Healthcare). With an eight-channel abdominal coil, the imaging parameters were: T2WI Sequence type: TSE, TR: 5000–8000 ms, TE: 80–100 ms,matrix: 320 x 320, slice thickness: 4 mm, FOV: 20 cm, resolution: 0.63 mm/pixel, flip angle: 150°. DWI imaging: b: 0,800 and 1400 s/mm^2, TR: 3000–5500 ms, TE: 55–85 ms, matrix: 256 x 256, slice thickness: 4 mm, FOV: 20 cm, resolution:0.78 mm/pixel, flip angle: 180°.

2.3 Experimental methods

2.3.1 Data processing and augmentation

Prior to the extraction of radiomics features, T2-weighted imaging (T2WI) and diffusion-weighted imaging (DWI) were normalized using the Z-score method. Image outliers were removed by applying the 3σ method (36). The tumor's region of interest (ROI) on DWI (b: 1400 s/mm2) and T2WI axial images was manually delineated layer by layer by a radiologist with three years of experience, using the 3Dslicer software (version 5.0.2). This delineation was conducted without prior knowledge of the study's objectives or the BM status, under the supervision of a senior radiologist with 15 years of experience. The ROI for the Apparent Diffusion Coefficient (ADC) sequence was sourced directly from the DWI region.

For the development of auto-sketch models, all images were normalized to a resolution of 224x224x64 using linear interpolation methods, employing the SimpleITK package in Python (version 3.6.2). Due to the different tumor sizes, the ROI images of the tumor regions in the T2WI and DWI sequences were resampled to a resolution of 64x64x32 using the difference method. Pixel consistency was ensured for training the ResNet tumor classification model. Due to the limited patient sample size, the number of images was increased by using data enhancement strategies. These included random rotations of 90, 180, and 270 degrees around the Z-axis of the image pixels and random intensity scaling and offsetting.

FIGURE 1
Diagram of the experimental inclusion-exclusion criteria.
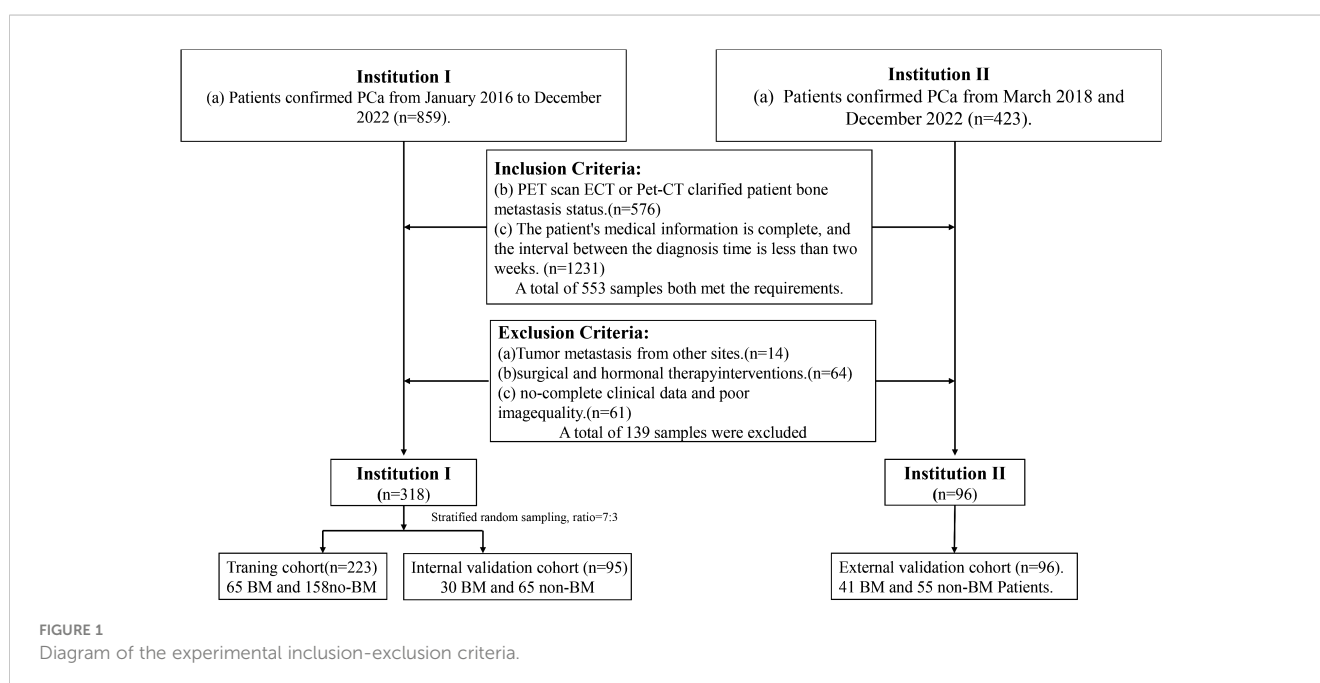
2.3.2 Automatic tumor outline model

UNETR model architecture was used to accurately identify regions impacted by Prostate Cancer (PCa). As the top-performing medical classification model of 2021 (37), UNETR model was trained on image from Center 1 and validated by Center 2 image. DWI, T2WI, and ADC images were trained through the three-channel model structure, and Dice similarity coefficient (DSC) was used to evaluate the model prediction efficiency. To optimize our training, we implemented a learning rate scheduler and early stopping criteria, which reduced the learning rate based on validation accuracy and halted training when validation accuracy had not improved for 10 epochs, respectively.

2.3.3 BM prediction model

Radiomics features were extracted from T2WI, DWI, and ADC sequences using pyradiomics. These features were then combined using the XGboost algorithm to create radiomics models. In choosing the structure of a DL model, four different 3D-ResNet network structures, three learning rates, and five optimizers were validated to identify the most effective DL model for predicting BM status. The experimental results can be found in the Supplementary Document. Tumor region images were obtained from the UNETR model and then were used to train the ResNet model after radiologist supervision and image enhancement. The validation cohort ROI was utilized to update network parameters, and the internal test cohort was used to assess the model's performance. The softmax function was employed to produce predictions from the ResNet model. The model with the highest accuracy in the internal test cohort was used to train stacking model (referenced in Supplementary Tables 1, 2). DL features, obtained from the last fully connected layer of all samples, were used to establish a composite model. Finally, The XGboost algorithm was utilized to construct a composite ResNet model (ResNet-C) by integrating the clinical, radiomics, and DL features of the training cohort and validation cohort (Figure 2). The model efficiency was subsequently validated through internal and external test cohorts.

2.3.4 Statistical analysis

R (version 4.1.3) and Python (version 3.6.2) were used to conduct statistical analysis and construct models. The Kolmogorov-Smirnov KS test was used to assess the distribution of continuous variables. T-test and Mann-Whitney U test (version 1.7.0)and multi-factor logistic regression analysis were performed to compare the differences between two cohorts. The adjusted p-value was calculated by the Benjamini-Hochberg correction, and Chi-square test was used to analyze the categorical variables. The 95% confidence interval (CI) of the AUC was determined using the bootstrapping method (1000 intervals). From these selected clinical features, we constructed a predictive model using the XGboost algorithm. The receiver operating characteristic (ROC) curve was used to visually demonstrate the prediction ability of ResNet, Radiomics, and ResNet-C models. Delong test was to validate and distinguish differences between the models.

3 Results

3.1 Clinical model

In our cohort, a mean age of 71.4 years was observed in BM patients, marginally higher than the 70.9 years recorded for the non-BM group; however, this difference was not statistically significant (P=0.687). Additionally, a pronounced 78.3% of BM
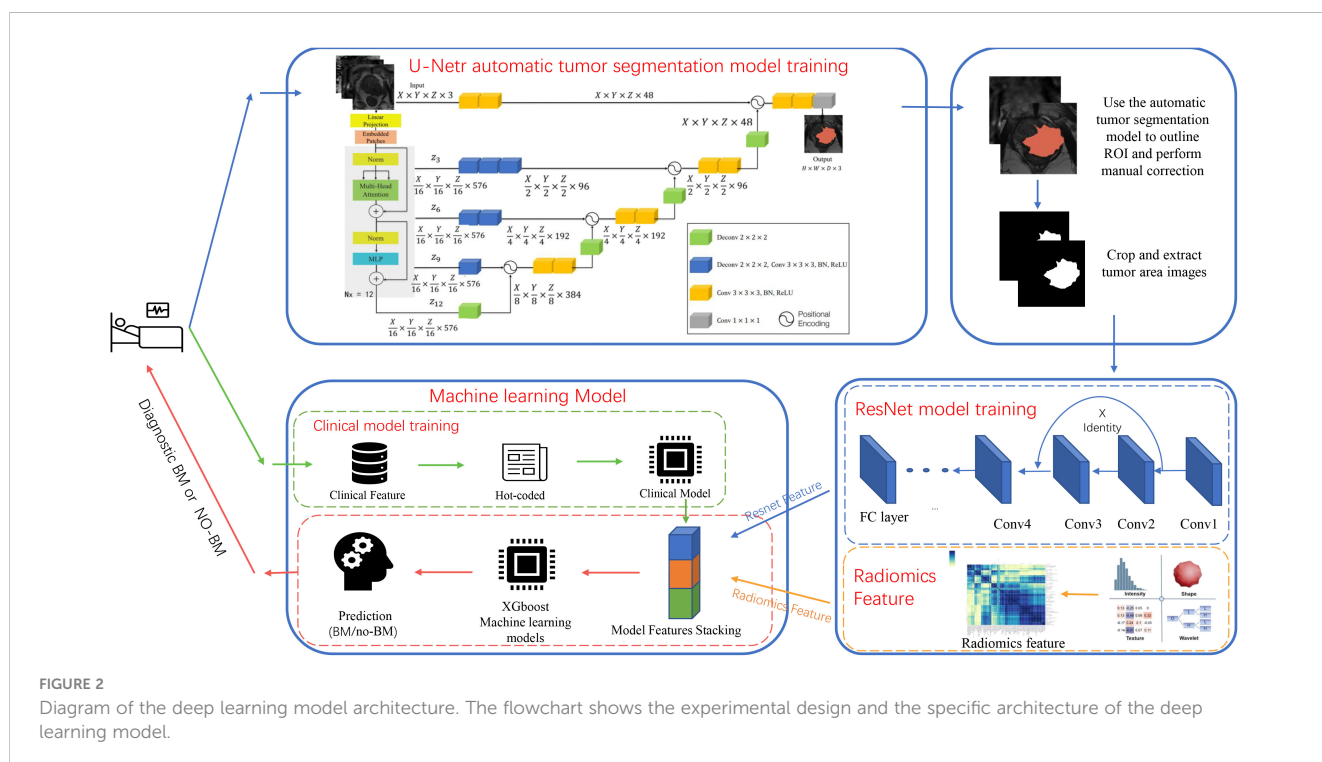

FIGURE 2
Diagram of the deep learning model architecture. The flowchart shows the experimental design and the specific architecture of the deep learning model.

patients were noted to have Gleason scores of ≥8, in contrast to 46.8% in the non-BM counterparts. Notably, PSA levels exceeding 100 ng/ml were found in 44.1% of BM patients, compared to only 13.5% in the non-BM group. A significant difference in PSA levels between BM and non-BM groups was observed. No significant difference in hematuria and urine leukocyte levels was found between the groups. Through univariate and multivariate analyses, the GS score, ISUP score, and PSA were identified as high-risk factors for predicting PCa's BM.

3.2 Model performance

3.2.1 PCa segmentation model

The model progressively converged with the increase in training epochs during the training process. The peak accuracy was achieved at epoch 331, recording DSC of 0.607. Figures 3 and 4 display the High-performing and low-performing samples of the model, respectively.

3.3 BM predictive model performance

In the clinical model, an accuracy (ACC) of 0.650 (95% CI 0.558–0.742) was observed in the internal test cohort and 0.668 (95% CI 0.584–0.782) in the external test cohort. Area Under the Curve (AUC) values of 0.713 (95% CI 0.641–0.785) and 0.757 (95% CI 0.661–0.853) were recorded for these cohorts, respectively. For the Radiomics model, an ACC of 0.857 (95% CI 0.796–0.918) was achieved in the internal test cohort and 0.819 (95% CI 0.773–0.865) in the external test cohort. The AUC values were 0.899 (95% CI 0.852–0.946) and 0.852 (95% CI 0.791–0.913) for the internal and external test cohorts, respectively.

Our model was trained using data from the training cohort, and its predictive performance was evaluated with the internal validation cohort. To determine the most effective baseline model, 48 combinations were explored (detailed results are available in the Supplementary Material). In the internal validation cohort, ResNet 101 (Adam, LR=0.001) emerged as the top performer, achieving an ACC of 0.875 (95% CI 0.818–0.932) for the internal test cohort and 0.833 (95% CI 0.791–0.875) for the external test cohort. Corresponding AUC values of 0.907 (95% CI 0.860–0.954) and 0.862 (95% CI 0.799–0.925) were recorded.

The clinical feature model was then combined with ResNet 101 and radiomics feature to develop the ResNet-C model which model posted an ACC of 0.902 (95% CI 0.867–0.937) in the internal test cohort and 0.885 (95% CI 0.832–0.938) in the external test cohort. The AUC values were 0.934 (95% CI 0.906–0.963) for the internal and 0.903 (95% CI 0.864–0.942) for the external test cohorts, respectively. In the final analysis, Delong validation was used to compare the clinical, ResNet 101, radiomics, and ResNet-C models. The results indicated that the diagnostic efficacy of the Clinic model differs significantly from the other models (refer to Figure 5, Table 1).

4 Discussion

Accurate assessment of BM status is crucial, as the status of BM is a decisive factor for undertaking radical prostatectomy and a key determinant affecting patient prognosis. However, conducting PET-CT/bone scan examinations for all newly diagnosed PCa patients is impractical. Therefore, the development of a convenient and accurate method to identify BM status and provide newly diagnosed patients with precise, personalized BM prediction models is a key objective in current clinical research. This aims to
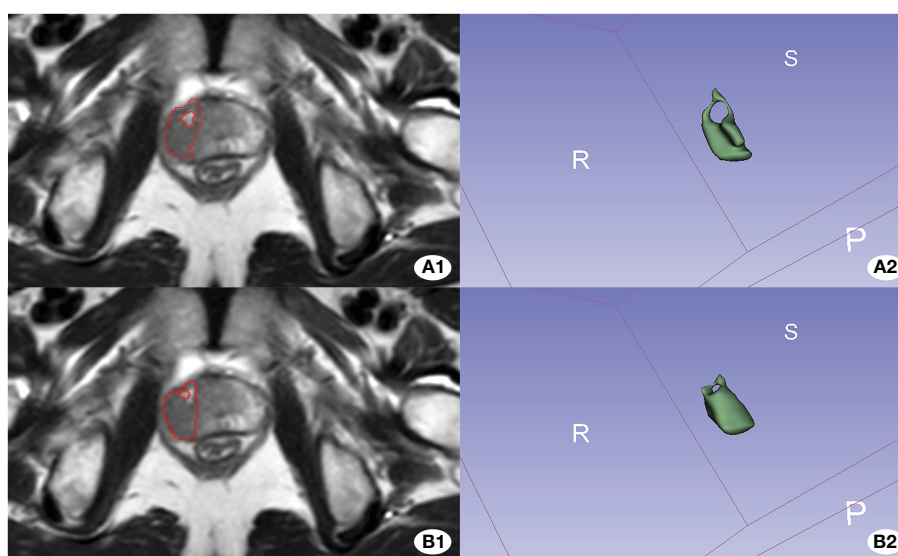

FIGURE 3
Illustration of the worse effective automatic delineation. In this figure, **(A1)** represents the manual delineation region, while **(A2)** showcases its 3D shape. **(B1)** denotes the automatically recognized region, and **(B2)** displays its 3D form. Notably, the automatic delineation model struggled to accurately identify and circumvent the bleeding foci.
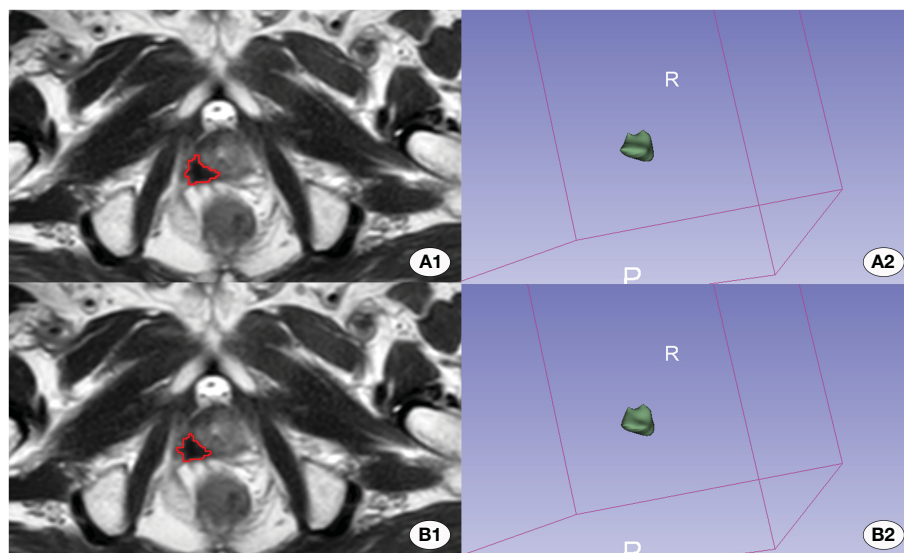
FIGURE 4
Depiction of the better effective automatic delineation. Here, **(A1)** signifies the manual delineation region, and **(A2)** presents its 3D shape.
(B1) highlights the automatically recognized region, and **(B2)** reveals its 3D configuration. The automatic delineation model adeptly identified the tumor region in this instance.

prevent patients eligible for surgical intervention from missing their opportunity for surgery and to assist patients with BM in receiving timely anti-metastasis treatment for their benefit. Recent studies in radiomics and DL suggest that integrated learning with multimodal MRI can facilitate the assessment of a patient's tumor metastatic status (31–33). Nonetheless, previous studies rarely reported on predicting BM status in PCa using radiomics and DL methods. In our study, to meet the demands of individualized treatment, we developed a more precise model based on bp-MRI scanning strategy. Our stacked model differentiated BM status in internal and external test cohorts with an average AUC value of 0.903–0.934. Additionally, this scanning strategy, reducing contrast agent

involvement, is also applicable to patients with contraindications to contrast agents and those with hepatic or renal insufficiency. In this study, we also sought an end-to-end approach to simplify the application of the model in terms of time and manpower costs.

The BM proportions in newly diagnosed PCa patients in our study was 32.8%, higher than the 10% reported in related studies (38). This is due to the fact that patients undergoing PET-CT and bone scan examinations were predominantly high-risk or suspected of having a metastatic status, with inclusion criteria leading to a higher BM detection rate. However, this distortion in BM proportions does not impact the quality of MRI imaging, which is critical for establishing our bp-MRI-based predictive model.
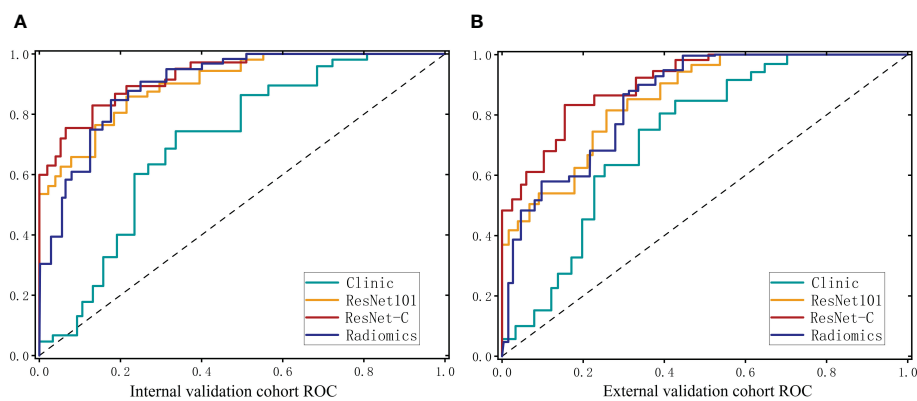

FIGURE 5
ROC curves for the clinical model. The deep learning model and the integrated model in the internal test cohort **(A)** and the external test cohort **(B)**.

TABLE 1 Predictive model performance effectiveness.

Model	Accuracy		AUC		Delong test*	
	Internal test cohort	External test cohort	Internal test cohort	External test cohort	Internal test cohort	External test cohort
ResNet-C	0.902 (0.867,0.937)	0.885 (0.832,0.938)	0.934 (0.906,0.963)	0.903 (0.864,0.942)	0.006	0.010
ResNet101	0.875 (0.818,0.932)	0.833 (0.791,0.875)	0.907 (0.860,0.954)	0.862 (0.799,0.925)	0.011	0.015
Radiomics	0.852 (0.791,0.913)	0.813 (0.767,0.859)	0.899 (0.852,0.946)	0.852 (0.791,0.913)	0.024	0.035
Clinic	0.650 (0.558,0.742)	0.668 (0.584,0.782)	0.713 (0.641,0.785)	0.757 (0.661,0.853)	–	–

*DeLong test is performed with Clinic as the benchmark, and the 95% confidence interval is listed for AUC and ACC, respectively.

Currently, PSA levels are widely used in clinical practice to screen patients at high risk for BM, with some studies indicating a higher risk of BM in populations with high PSA (12–15, 19). Therefore, our study also considers the research value of related indicators.

Univariate and multivariate logistic regression analyses showed significant statistical differences in GS score, PSA, and ISUP grading (P < 0.05, Table 2), consistent with previous research. However, modeling analysis of clinical features did not provide

TABLE 2 Characteristics of patients in the training and test cohort.

	BM	Non-BM	Univariable P	Multivariable analysis	
	(n=136)	(n=278)		Odds ratio	P
Age	71.4*	70.9*	0.786	NA	0.989
Gleason			<0.001	0.348	0.004
6	2 (1.4%)	60 (21.6%)			
7	28 (20.6%)	90 (32.4%)			
≥8	106 (77.9%)	128 (46.0%)			
PSA	169.59*	55.16*	<0.001	1.003	0.005
<10	17 (12.5%)	47 (16.9%)			
10–100	59 (43.4%)	194 (69.8%)			
>100	60 (44.1%)	37 (13.3%)			
ISUP			<0.001	4.45	<0.001
≤3	30 (22.0%)	150 (54.0%)			
4	45 (33.1%)	65 (23.4%)			
5	61 (44.9%)	63 (22.7%)			
Urine					
hematuria	39 (28.6%)	71 (25.5%)	0.498	NA	0.060
Leukocyt-uria	54 (39.7%)	93 (33.6%)	0.212	NA	0.859

P-value is derived from the univariable association analyses between the clinicopathologic variables and Bone status. The data marked with * are averaged.

optimum accuracy. This limitation may be attributed to the unreliability of PSA and GS scores in measuring tumor behavior, as well as the susceptibility of PSA measurements to external influences such as prostatitis, age, endocrine, or metabolic disorders. Additionally, GS scores calculated based on cytological puncture measurements may lead to significant errors (11).

Radiomics research based on high-throughput data and the advancements in convolutional neural network-based DL have provided new insights into accurately predicting the biological behavior of tumors (39). Traditional radiomics methods have achieved significant success in this field. Radiomics offers the potential for non-invasive prediction of tumor biological behavior, Wang and Zhang, among others (18, 19), proposed a DCE-MRI BM prediction model using radiomics, achieving an impressive AUC of 0.91. Compared to previous studies, our research has several notable advantages: the imaging strategy of bpMRI limits the utility of existing models. In contrast, our model, based on the bp-MRI scanning strategy, did not perform DCE sequence feature analysis, expanding the model's audience. Additionally, previous studies did not consider the abstract feature information offered by convolutional neural networks. Deep neural networks can capture nonlinear, interrelated spatial structural features within the target area through convolution and pooling, as well as analyze relationships between distant pixels, all of which are key to accurately predicting BM. The AUC of the ResNet 101 model was 0.907 and 0.862. Thirdly, we conducted ensemble learning of radiomics and DL features, establishing a combined prediction model, ResNet-C, using the XGboost algorithm. The AUCs for Centers 1 and 2 were 0.934 and 0.903, respectively. Machine learning algorithms, compared to linear models, can accurately process nonlinear features. Lastly, our study underwent multicenter validation, providing higher reliability. The DeLong test (all p < 0.05) indicated that MRI image models and clinical risk models have different diagnostic efficacies, proving the necessity of stacking different modal models.

Radiomics and DL have achieved encouraging results, but the analysis of tumor regions largely depends on manual delineation. Extracting tumor information from MRI multiple sequences and multiple planes using manual methods is a time-consuming task. Accurate segmentation of images aids in advancing the application of radiomic models in clinical settings. Moreover, the automatic recognition of tumor regions is a crucial component of an end-to-end model. Therefore, this study has made some attempts in this area. After validation with an external test cohort, our established tumor segmentation model yielded a DSC value of 0.607. This indicates that the model can only be semi-automated; while it can more accurately locate tumors, the precise determination of tumor extent and boundaries still requires manual intervention by radiologists. Our initial goal was full automation, but the model struggles to differentiate PCa from other abnormalities (such as inflammation and hemorrhage), challenges that may stem from

overlapping signal characteristics between lesions and PCa. To ensure accuracy in tumor regions, radiologists must also examine these differences. Deep learning networks are commonly referred to as "black box" structures due to their complex and often non-transparent nature. Therefore the Grad-CAM technique is employed to attempt to uncover and interpret the decision-making process of convolutional neural networks. This is achieved through the analysis of the feature weight maps of the final convolutional layer. In observing the ResNet model, it was noted that the decision-making process relies on the tumor margin areas (Figure 6). We posit that this reliance may be attributed to the interactions between the microenvironment of the active tumor cells at the periphery and the surrounding normal tissue (40).

The results of our study's model, validated through internal and external verification sets, demonstrate significant potential. However, the establishment of end-to-end models also presents numerous challenges. Concerns regarding data privacy and security, coupled with insufficient interpretability leading to difficulties for physicians in understanding the model's decision-making process, pose substantial obstacles. Additionally, the validation and adaptability testing in real clinical settings demands considerable time and resources. Furthermore, issues related to regulatory compliance, physician acceptance, and model usability are critical. Successful integration of the medical end-to-end model into clinical practice necessitates interdisciplinary collaboration and continuous improvement of model performance. The visualization application of the model requires the design of a software platform, and we are actively collaborating with software companies to seamlessly integrate the model into software modules embedded in the medical record diagnostic system, ultimately benefiting patients.

Despite promising findings, our study has limitations. Firstly, the DL and radiomics models were established using retrospective data. Prospective data from more clinical trials in the future will improve the clinical evidence of our models. Secondly, differences in examination equipment might affect the model's performance. In the external test cohort, a deeper network model, ResNet152, showed better diagnostic performance (see Supplementary Tables 1, 2), We believe this is attributed to deeper model architectures, which aid in better understanding the inherent relationships between different regions of the tumor and assist the model in extracting more discriminative abstract features from the imaging data, necessitating an expansion of the number of centers in future research to confirm this phenomenon. Additionally, our study neglected 1.5T images, due to the challenge of accurately delineating tumor regions on T1 sequence images, the exploration of the value of T1 sequence images has been omitted in this study. Due to the limited sample size, this study analyzed images of delineated tumor regions. However, this introduces a drawback: normalizing tumor sizes during training prevents our model from assessing the correlation between tumor size and bone metastasis. In fact, with a sufficiently large training set, CNNs may not require highly accurate tumor
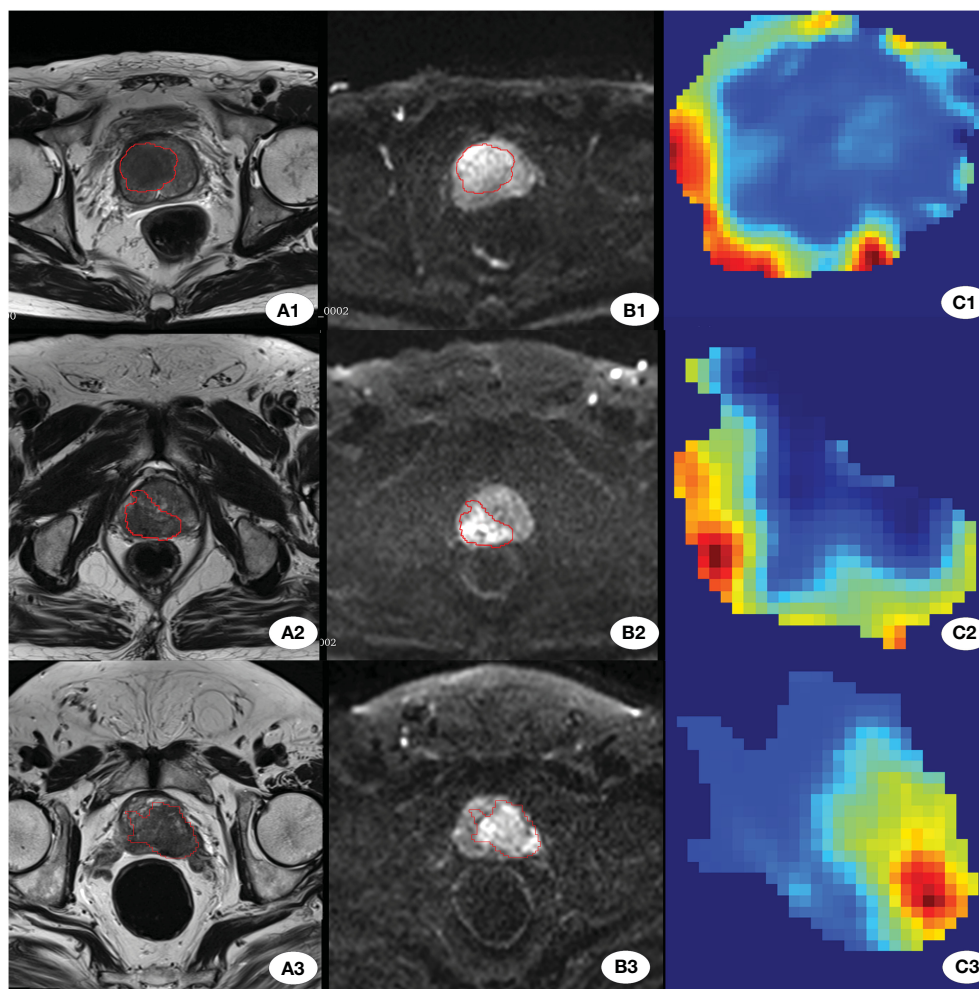
FIGURE 6
Grad_GAM. **(A1-A3)** show T2-weighted images from three different patients with BM. **(B1-B3)** display the corresponding DWI images from the same three patients. **(C1-C3)** present the Grad-GAM images of the largest cross-sectional area of the tumor for each patient.

masks as input. Therefore, our future research goal is to broaden the sample scope, and increase its scale and diversity, thus further enhancing the performance of our model.

In summary, we have developed an integrated learning model using bp-MRI images to accurately predict the BM status of PCa patients. For PCa patients, this model can assist urologists in deciding whether there is an opportunity for radical surgery, which is positively significant for the patient's treatment approach and prognosis.

5 Conclusions

Our study combined clinical parameters with prostate bpMRI, offering clinicians a non-invasive tool to inform treatment decisions for prostate patients. Additionally, we developed an automatic tumor delineation model to streamline the process and augment efficiency, increasing its potential for clinical adoption.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

Author contributions

SX: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization. ST: Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Resources, Methodology, Investigation, Data curation. HX: Writing – review & editing, Validation, Methodology, Investigation, Data curation. ZS: Writing – review & editing, Validation, Supervision, Methodology, Investigation. WY: Writing – review & editing, Supervision, Investigation, Data curation. DM: Writing – review & editing, Validation, Project administration, Investigation. XT: Writing – review & editing, Supervision, Software, Resources, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. ZJ: Writing – review & editing, Visualization, Supervision, Resources, Methodology, Funding acquisition, Data curation, Conceptualization. YF: Writing – review & editing, Validation, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1298516/full#supplementary-material

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J Clin*. (2021) 71:209–49. doi: 10.3322/caac.21660

2. Mohammed Nawi A, Masdor NA, Othman R, Kandayah T, Ahmad N, Safian N. Survival rate and prognostic factors of localised prostate cancer in southeast Asian countries: A systematic review with meta-analysis. *Asian Pac J Cancer Prev*. (2023) 24:2941–095. doi: 10.31557/APJCP.2023.24.9.2941

3. Muñoz-Rodríguez J, Domínguez A, Rosado MA, Centeno C, Parejo V, Costa-Trachsel I, et al. Effect of muscle density in patients with metastatic prostate cancer administered androgen deprivation therapy. *Endocrinol Diabetes Nutr (Engl Ed)*. (2021) 68:92–8. doi: 10.1016/j.endinu.2020.03.014

4. Gillessen S, Attard G, Beer TM, Beltran H, Bjartell A, Bossi A, et al. Clarke N et al: Management of Patients with Advanced Prostate Cancer: Report of the Advanced Prostate Cancer Consensus Conference 2019. *Eur Urol*. (2020) 77:508–47. doi: 10.1016/j.eururo.2020.01.012

5. Cornford P, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, et al. Gillessen S et al: EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer. Part II-2020 Update: Treatment of Relapsing and Metastatic Prostate Cancer. *Eur Urol*. (2021) 79:263–82. doi: 10.1016/j.eururo.2020.09.046

6. Kyriakopoulos CE, Chen YH, Carducci MA. Chemohormonal therapy in metastatic hormone-sensitive prostate cancer: long-term survival analysis of the randomized phase III E3805 CHAARTED trial. *J Clin Oncol*. (2018) 36(11):1080–7. doi: 10.1200/JCO.2017.75.3657

7. Sweeney CJ, Chen Y-H, Carducci M, Liu G, Jarrard DF, Eisenberger M, et al. Cooney MM et al: Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer. *New Engl J Med*. (2015) 373:737–46. doi: 10.1056/NEJMoa1503747

8. James ND, De Bono JS, Spears MR, Clarke NW, Mason MD, Dearnaley DP, et al. Jones RJ et al: Abiraterone for Prostate Cancer Not Previously Treated with Hormone Therapy. *New Engl J Med*. (2017) 377:338–51. doi: 10.1056/NEJMoa1702900

9. Fizazi K, Tran N, Fein L, Matsubara N, Rodriguez-Antolin A, Alekseev BY, et al. Protheroe A et al: Abiraterone plus Prednisone in Metastatic, Castration-Sensitive Prostate Cancer. *New Engl J Med*. (2017) 377:352–60. doi: 10.1056/NEJMoa1704174

10. Briganti A, Passoni N, Ferrari M, Capitanio U, Suardi N, Gallina A, et al. Salonia A et al: When to perform bone scan in patients with newly diagnosed prostate cancer: external validation of the currently available guidelines and proposal of a novel risk stratification tool. *Eur Urol*. (2010) 57:551–8. doi: 10.1016/j.eururo.2009.12.023

11. Wei RJ, Li TY, Yang XC, Jia N, Yang XL, Song HB. Serum levels of PSA, ALP, ICTP, and BSP in prostate cancer patients and the significance of ROC curve in the diagnosis of prostate cancer bone metastases. *Genet Mol Res*. (2016) 15(2):10.4238/gmr.15027707. doi: 10.4238/gmr.15027707

12. Kikuchi E, Nakashima J, Ishibashi M, Ohigashi T, Oya M, Nakagawa K, et al. Usefulness of alpha1-antichymotrypsin-PSA complex for predicting bone metastases of prostate cancer. *Urology*. (2006) 68:371–5. doi: 10.1016/j.urology.2006.02.031

13. Nieder C, Haukland E, Pawinski A, Norum J. (2006). Seven-month prostate-specific antigen (PSA) is prognostic in patients with prostate cancer initially diagnosed with distant metastases. *Med Oncol*. (2018) 35(4):46. doi: 10.1007/s12032-018-1110-y

14. Johnston AW, Longo TA, Davis LG, Zapata D, Freedland SJ, Routh JC. Bone scan positivity in non-metastatic, castrate-resistant prostate cancer: external validation study. *Int Braz J urol*. (2020) 46:42–52. doi: 10.1590/s1677-5538.ibju.2019.0225

15. Chen S, Wang L, Qian K, Jiang W, Deng H, Zhou Q, et al. Establishing a prediction model for prostate cancer bone metastasis. *Int J Biol Sci*. (2019) 15:208–20. doi: 10.7150/ijbs.27537

16. Kamiya N, Suzuki H, Yano M, Endo T, Takano M, Komaru A, et al. Implications of serum bone turnover markers in prostate cancer patients with bone metastasis. *Urology*. (2010) 75:1446–51. doi: 10.1016/j.urology.2009.11.049

17. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RGPM, Granton P, et al. Dekker A et al: Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036

18. Wang Y, Yu B, Zhong F, Guo Q, Li K, Hou Y, et al. MRI-based texture analysis of the primary tumor for pre-treatment prediction of bone metastases in prostate cancer. *Magn Reson Imaging*. (2019) 60:76–84. doi: 10.1016/j.mri.2019.03.007

19. Zhang W, Mao N, Wang Y, Xie H, Duan S, Zhang X, et al. A Radiomics nomogram for predicting bone metastasis in newly diagnosed prostate cancer patients. *Eur J Radiol*. (2020) 128:109020. doi: 10.1016/j.ejrad.2020.109020

20. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2. *Eur Urol*. (2019) 76:340–51. doi: 10.1016/j.eururo.2019.02.033

21. Massanova M, Vere R, Robertson S, Crocetto F, Barone B, Dutto L, et al. Clinical and prostate multiparametric magnetic resonance imaging findings as predictors of general and clinically significant prostate cancer risk: A retrospective single-center study. *Curr Urol*. (2023) 17:147–52. doi: 10.1097/CU9.0000000000000173

22. Heidenreich A, Bastian PJ, Bellmunt J, Bolla M, Joniau S, van der Kwast T, et al. EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *Eur Urol*. (2014) 65:124–37. doi: 10.1016/j.eururo.2013.09.046

23. Brancato V, Di Costanzo G, Basso L, Tramontano L, Puglia M, Ragozzino A, et al. Assessment of DCE utility for PCa diagnosis using PI-RADS v2.1: effects on diagnostic accuracy and reproducibility. *Diagnostics (Basel)*. (2020) 10(3):164. doi: 10.3390/diagnostics10030164

24. Boesen L, Nørgaard N, Løgager V, Balslev I, Bisbjerg R, Thestrup KC, et al. Assessment of the diagnostic accuracy of biparametric magnetic resonance imaging for prostate cancer in biopsy-naive men: the biparametric MRI for detection of prostate cancer (BIDOC) study. *JAMA Netw Open*. (2018) 1:e180219. doi: 10.1001/jamanetworkopen.2018.0219

25. Poirion OB, Jing Z, Chaudhary K, Huang S, Garmire LX. DeepProg: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data. *Genome Med*. (2021) 13(1):112. doi: 10.1186/s13073-021-00930-x

26. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. (2016) 278:563–77. doi: 10.1148/radiol.2015151169

27. Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, et al. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut*. (2019) 68:729–41. doi: 10.1136/gutjnl-2018-316204

28. Zhou J, Zhang Y, Chang KT, Lee KE, Wang O, Li J, et al. (2019). Chow D et al: Diagnosis of benign and malignant breast lesions on DCE-MRI by using radiomics and deep learning with consideration of peritumor tissue. *J Magn Reson Imaging*. (2020) 51:798–809. doi: 10.1002/jmri.26981

29. Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: Current applic ations and future directions. *PloS Med*. (2018) 15(11):e1002707.

30. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. (2018) 18(8):500–10. doi: 10.1038/s41568-018-0016-5

31. Liang HY, Yang SF, Zou HM, Hou F, Duan LS, Huang CC, et al. Deep learning radiomics nomogram to predict lung metastasis in soft-tissue sarcoma: A multi-center study. *Front Oncol*. (2022) 12:897676. doi: 10.3389/fonc.2022.897676

32. Zhou L-Q, Wu X-L, Huang S-Y, Wu G-G, Ye H-R, Wei Q, et al. Cui X-W et al: Lymph Node Metastasis Prediction from Primary Breast Cancer US Images Using Deep Learning. *Radiology*. (2020) 294:19–28. doi: 10.1148/radiol.2019190372

33. Dong D, Fang MJ, Tang L, Shan XH, Gao JB, Giganti F, et al. Palumbo D et al: Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. *Ann Oncol*. (2020) 31:912–20. doi: 10.1016/j.annonc.2020.04.003

34. Wang J, Zeng J, Li H, Yu X. A deep learning radiomics analysis for survival prediction in esophageal cancer. *J Healthc Eng*. (2022) 2022:4034404. doi: 10.1155/2022/4034404

35. Zeng Q, Li H, Zhu Y, Feng Z, Shu X, Wu A, et al. Xiong J et al: Development and validation of a predictive model combining clinical, radiomics, and deep transfer learning features for lymph node metastasis in early gastric cancer. *Front Med (Lausanne)*. (2022) 9:986437. doi: 10.3389/fmed.2022.986437

36. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. (2000) 29:722–9. doi: 10.1093/ije/29.4.722

37. Liu D, Kuai Y, Zhu R, Zhou C, Tao Y, Han W, et al. Prognosis of prostate cancer and bone metastasis pattern of patients: a SEER-based study and a local hospital based study from China. *Sci Rep*. (2020) 10:9104. doi: 10.1038/s41598-020-64073-6

38. Karimi D, Dou H, Gholipour A. Medical image segmentation using transformer networks *IEEE Access*. (2022) 10:29322–32. doi: 10.1109/access.2022.3156894

39. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med*. (2021) 13(1):152. doi: 10.1186/s13073-021-00968-x

40. Rycaj K, Li H, Zhou J, Chen X, Tang DG. Cellular determinants and microenvironmental regulation of prostate cancer metastasis. *Semin Cancer Biol*. (2017) 44:83–97. doi: 10.1016/j.semcancer.2017.03.009