ORIGINAL ARTICLE

CANCER
INNOVATION

Radiomics models to predict bone marrow metastasis of neuroblastoma using CT

Xiong Chen[1,2] 🟢 | Qinchang Chen[3] | Yuanfang Liu[4] | Ya Qiu[5] | Lin Lv[6] |
Zhengtao Zhang[1,2] | Xuntao Yin[7] 🟢 | Fangpeng Shu[1,2]

[1]Department of Paediatric Urology, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

[2]Department of Paediatric Surgery, Guangzhou Institute of Paediatrics, Guangzhou Medical University, Guangzhou, China

[3]Department of Pediatric Cardiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangdong Cardiovascular Institute, Guangdong Provincial Key Laboratory of Structural Heart Disease, Guangzhou, China

[4]Department of Radiology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

[5]Department of Radiology, the First People's Hospital of Kashi Prefecture, Kashi, China

[6]Medical School, Sun Yat-sen University, Guangzhou, China

[7]Department of Radiology, Guangzhou Women and Children's Medical Center, Guangzhou, China

Correspondence
Xuntao Yin, Department of Radiology, Guangzhou Women and Children's Medical Center, Guangzhou 510000, China.
Email: xuntaoyin@gmail.com

Fangpeng Shu, Department of Paediatric Urology, Guangzhou Women and

Abstract

Background: Bone marrow is the leading site for metastasis from neuroblastoma and affects the prognosis of patients with neuroblastoma. However, the accurate diagnosis of bone marrow metastasis is limited by the high spatial and temporal heterogeneity of neuroblastoma. Radiomics analysis has been applied in various cancers to build accurate diagnostic models but has not yet been applied to bone marrow metastasis of neuroblastoma.

Methods: We retrospectively collected information from 187 patients pathologically diagnosed with neuroblastoma and divided them into training and validation sets in a ratio of 7:3. A total of 2632 radiomics features were retrieved from venous and arterial phases of contrast-enhanced computed tomography (CT), and nine machine learning approaches were used to build radiomics models, including multilayer perceptron (MLP), extreme gradient boosting, and random forest. We also constructed radiomics-clinical models that combined radiomics features with clinical predictors such as age, gender, ascites, and lymph gland metastasis. The performance of the models was evaluated with receiver operating characteristics (ROC) curves, calibration curves, and risk decile plots.

Results: The MLP radiomics model yielded an area under the ROC curve (AUC) of 0.97 (95% confidence interval [CI]: 0.95–0.99) on the training set and 0.90 (95% CI: 0.82–0.95) on the validation set. The radiomics-clinical model using an MLP yielded an AUC of 0.93 (95% CI: 0.89–0.96) on the training set and 0.91 (95% CI: 0.85–0.97) on the validation set.

Children's Medical Center, Guangzhou
Medical University, Guangzhou 510000,
China & Department of Paediatric Surgery,
Guangzhou Institute of Paediatrics,
Guangzhou Medical University,
Guangzhou 510000, China.
Email: 15625060053@163.com

Conclusions: MLP-based radiomics and radiomics-clinical models can precisely predict bone marrow metastasis in patients with neuroblastoma.

KEYWORDS
bone marrow metastasis, machine learning, neuroblastoma, radiomics

1 | INTRODUCTION

Neuroblastoma, which is derived from the sympathetic nervous system, is one of the most common solid tumors in children and a leading cause of cancer deaths in children [1]. Ultimately, it accounts for 15% of cancer-related deaths in early childhood [2].

The clinical outcomes of patients with neuroblastoma vary dramatically. Some patients undergo spontaneous tumor regression and only need active surveillance, whereas others lose their lives within a short period, even after an aggressive treatment plan that may include five to six cycles of chemotherapy, surgery, chemoradiotherapy, autologous human stem cell transplant, and immunotherapy [3]. This diversity of clinical outcomes is associated with the high spatial and temporal heterogeneity of neuroblastoma [4]. To address this, a widely accepted grading system, the International Neuroblastoma Risk Group Staging System (INRGSS), was built [5]. According to INRGSS, patients are divided into low, intermediate, and high-risk groups based on patient clinical factors (age and stage) and tumor genomics, among which, metastasis is regarded as a main considering factor. Those patients with distant metastasis are assigned to the high-risk group, with a 5-year overall survival rate of less than 40% [6]. Among patients with metastasis, bone marrow is the leading site, accounting for 90% of cases [7]. Internationally agreed-upon treatment options for high-risk neuroblastoma include chemotherapy, surgery, high-dose chemotherapy followed by autologous stem cell transplantation, external-beam radiotherapy, radionuclide therapy, differentiation therapies, and immunotherapy [8]. Hence, bone marrow metastasis should be precisely diagnosed.

Clinically, the diagnosis of bone marrow metastasis relies on bone marrow aspiration and biopsy. However, current diagnostic biopsy procedures may miss metastatic tissues because of the high spatial heterogeneity of neuroblastoma, indicating that current diagnostic biopsy strategies may miss or underestimate therapeutically actionable alterations [4, 7, 9]. In addition, neuroblastoma shows highly heterogeneous temporal characteristics. Therefore, physicians have to repeat bone marrow aspiration and biopsies, which in turn brings more pain and suffering to children. Therefore, a noninvasive predictive method should be urgently developed.

Considering that CT is necessary during follow-up, whether we can predict bone marrow metastasis with CT is of great interest. Xinxian et al. reported that a calcification sign on CT correlated with bone marrow metastasis in neuroblastoma, indicating that CT has the potential to reflect neuroblastoma characteristics likely to be associated with metastasis [10].

Radiomics, which involves the extraction of large amounts of high-dimensional quantitative features from medical images, has been rapidly developed in the hope that it will advance precision diagnostics and cancer treatment [11, 12]. Built on the hypothesis that CT images of cancer can reflect the nature of the cancer, radiomics and machine learning approaches have the potential to elucidate the mystery of cancer and assist in clinical decision-making. Radiomics provides a good strategy to study tumor heterogeneity by mapping the region of interest (ROI) of tumors, which theoretically means that all features of the cancer are taken into consideration [13]. Previously, researchers have reported numerous radiomics nomograms for accurately predicting metastasis in various cancers [14, 15]. However, the application of radiomics to the prediction of bone marrow metastasis in patients with neuroblastoma has not been reported. Furthermore, previous literature reported the good predictive ability of a radiomics nomogram for MYCN status in neuroblastoma, another factor relating to prognosis [16, 17]. Therefore, we proposed that a machine learning-based radiomics nomogram using CT images could predict bone metastasis in patients with neuroblastoma.

2 | METHODS

2.1 | Study design and participants

We retrospectively collected information from patients with neuroblastoma treated at Guangzhou Women and Children's Medical Center from 2015 to 2020. Ethical approval was obtained from the institutional review board and the requirement for informed consent was waived. Inclusion criteria included the following: (a) pathologically confirmed abdominal neuroblastoma; (b) standard abdominal CT performed <30 days before bone marrow aspiration and biopsy. Exclusion criteria consisted of the following: (a) neuroblastoma originating from the chest; (b) presence of imaging artifacts on the CT images. The label "bone marrow metastasis" was applied to those who were detected with bone marrow metastasis on either bone marrow aspiration, biopsy, or PET/CT scan within 6 months after the CT scan. In this way, we could discover those who already had metastasis as well as those with a high tendency to develop metastasis, and the nomogram built according to this categorization strategy should be more valuable for early prevention as well as early diagnosis.

The patients were divided into training and validation sets in a ratio of 7:3 based on the date of the CT. The training set was used to develop radiomics and radiomics-clinical models using multiple machine learning approaches. The training and validation sets were completely independent. The study design is presented in Figure 1.

2.2 | Imaging acquisition, volume-of-interest segmentation, and radiomic feature extraction

All patients underwent an abdominal contrast-enhanced CT examination at Guangzhou Women and Children's Medical Center on a 64-slice CT scanner. All patients were scanned from thoracic inlet to pubic symphysis. The imaging parameters were as follows: 0.675 mm reconstructed slice thickness, 3.0 mm slice thickness, 3.0 slice interval, and 80 kV tube potential with automatic tube current modulation.

Nonionic iodinated contrast medium was used at a dose of 1.5 mL/kg (body weight) and a flow rate of 1.8–2.0 mL/s via power injection. Pre-contrast acquisitions were followed by two post-contrast scans performed in the arterial (20 s after injection) and venous phases (60 s after injection). The arterial and venous phase CT images were used in the following analysis. ROIs were delineated using 3D Slicer software version 4.7.0 by two skilled radiologists (YFL and YQ). Then, a large set of quantitative radiomics features was extracted using the PyRadiomics platform implanted in the 3D Slicer software [18]. The features can be divided into the following four categories: (a) first-order statistics features, (b) shape-based features, (c) statistics-based textural features, and (d) wavelet features. More detailed information about the radiomics features and their extraction can be found on the official website for PyRadiomics (https://pyradiomics.readthedocs.io/en/latest/).
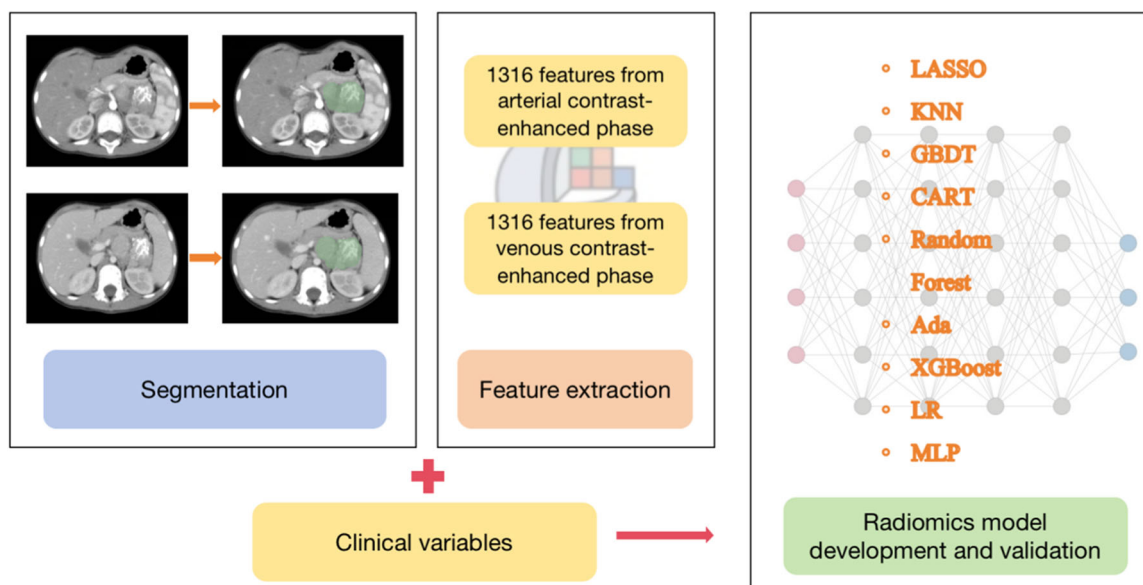

FIGURE 1 Study design. We extracted radiomic features from arterial and venous contrast-enhanced phase CT of neuroblastoma patients and then used multiple machine learning approaches to develop and validate the performance of radiomic models for predicting bone marrow metastasis. CT, computed tomography.

2.3 | Feature selection and radiomics signature construction

First, all radiomics features were standardized into the range [0,1] using a z-score normalization method. The z-score normalization process involves subtracting the mean value of the data from each individual data point and then dividing it by the standard deviation of the data. The formula for calculating the z-score of a data point (X) is

$$Z = (X - \mu)/\sigma, \tag{1}$$

where Z represents the standardized value, X is the original data point, μ is the mean of the data, and σ is the standard deviation of the data.

The z-score transformed data will have a mean of 0 and a standard deviation of 1. Data points below the mean will have negative z-scores, while data points above the mean will have positive z-scores. This normalization technique allows for easier interpretation and comparison of data across different variables because they are all measured in terms of their deviation from the mean in standard deviation units.

Then, we applied t tests to all the radiomics features and removed insignificant features ($p < 0.05$). After that, we oversampled the training set using the Synthetic Minority Oversampling Technique (SMOTE) approach, which involves generating new samples and adding them to the training queue to equalize the number of positive and negative examples [19]. No oversampling was performed on the validation set. Finally, we built the radiomics signatures using multiple machine learning approaches, including least absolute shrinkage and selection operator (LASSO) logistic regression, k-nearest-neighbor (KNN), gradient boosting decision tree (GBDT), classification and regression tree (CART), random forest (RF), adaptive boosting (Ada), XGBoost (XGB), logistic regression (LR), and multilayer perception (MLP) [20–22].

2.4 | Construction of the radiomics-clinical complex models

The radiomics signature represents the diagnostic performance of the CT features. Previous studies suggested that clinical factors may also have a guiding role in bone marrow metastasis [23]. Incorporating clinical factors into the model construction process is expected to result in a better-performing predictive model [24]. Therefore, we took clinical factors into consideration and constructed bone marrow metastasis radiomics-clinical complex models. We selected the bone marrow metastasis-related clinical factors using univariate and multivariate logistic regression and then constructed the radiomics-clinical complex nomogram.

Taking both radiomics features and clinical factors as texture variables, we also constructed complex models with the above machine learning approaches.

2.5 | Performance of the radiomics and radiomics-clinical models

For both radiomics and radiomics-clinical models, we only compared the diagnostic performances of the top three performance models in depth. The top three approaches were identified by comparing areas under the curves (AUCs) on the validation set. As the AUC increased, the performance improved. T tests were performed to evaluate the differences between the radiomics scores and radiomics-clinical scores of the bone marrow metastasis group and nonmetastasis group. Two-sided $p < 0.05$ was considered significant. The calibration curves of the models were constructed, and Brier scores were calculated to evaluate the fitting of the models. As the Brier score decreased, the calibration improved. Receiver operating characteristics (ROC) curves of the top three radiomics and radiomics-clinical models were constructed, and the AUCs were calculated to compare their discrimination accuracy on the training and validation sets. The accuracy, sensitivity, specificity, negative predictive value (NPP), positive predictive value (PPV), and F1 score of the prediction models were calculated. The formulas used for calculating these are presented in Supporting Information S1. All analyses were independently tested.

2.6 | Ethics approval and consent to participate

The relevant institutional review board approved this retrospective study and waived the need to obtain informed consent from the patients.

2.7 | Statistical analysis

All statistical analyses were performed using Python and R statistical software (version 3.4.2). A two-sided p value < 0.05 was considered significant.

3 | RESULTS

3.1 | Baseline characteristics and bone metastasis-correlated clinical variables

In total, 187 patients were enrolled in our study. The demographic and clinical information of the patients in the training and validation sets are listed in Table 1. Among the

187 patients, 81 (43.3%) were diagnosed with bone marrow metastasis, which is a rate consistent with other studies.

3.2 | Radiomics signature development and performance

In total, 1316 radiomics features from venous phase images and 1316 from arterial phase images were obtained. After removing features with a t test-derived p value > 0.05, 897 features were identified. We tested nine machine learning approaches to develop a radiomics signature on the training set with these 897 features. The diagnostic performance values of these radiomics signatures are shown in Table 2. Among the approaches, the performances on the validation set of the radiomics signatures built using an MLP, RF, and

XGB ranked in the top three. The accuracy, sensitivity, specificity, NPV, PPV, and F1 scores of the top three radiomics models are listed in Supporting Information S1: Table S1. The AUCs of these three approaches were 0.97 (95% CI: 0.95–0.99), 0.86 (95% CI: 0.81–0.91), and 0.87 (95% CI: 0.82–0.92), respectively, on the training set and 0.90 (95% CI: 0.82–0.95), 0.88 (95% CI: 0.80–0.94), and 0.82 (95% CI: 0.71–0.91) on the validation set (Supporting Information S1: Figure S1a,b). The calibration curves of these approaches on the training and validation sets are presented in Supporting Information S1: Figure S1c–h. The Brier scores of these approaches ranged from 0.097 to 0.186, indicating that these approaches showed a good fit in the prediction of bone marrow metastasis. The radiomics features included in these approaches are shown in Supporting Information S1: Figure S1i–k as radar plots. The main features varied between the models because of the principles of the different approaches, which is the reason why we tried different approaches to develop the most suitable prediction model. The feature importance of the top three models is shown in Supporting Information S1: Table S2. We divided all participants in the training and validation sets into 10 groups based on the estimated risk deciles of the top three models (Supporting Information S1: Figure S2). With an increase in prediction probability, the observed probability also increases.

3.3 | Radiomics-clinical model development and performance

The results of the univariate and multivariate logistic regression are presented in Table 3. MLP ($p < 0.001$) and ascites ($p = 0.03$) were independently associated with bone

TABLE 1 Characteristics of the 187 patients enrolled in this study.

Variables	Training set ($n = 130$)	Validation set ($n = 57$)
Sex		
Boy (%)	74 (56.9)	32 (56.1)
Girl (%)	56 (43.0)	25 (43.8)
Age (month)	35.5 (14–48)	35 (17–48)
Calcification (%)	101 (77.6)	46 (80.7)
Cross midline (%)	30 (23.0)	15 (26.3)
Ascites (%)	26 (20.0)	7 (12.2)
Lymph gland metastasis (%)	61 (46.9)	36 (63.1)
Bone marrow metastasis (%)	58 (44.6)	23 (40.3)

TABLE 2 Performance of the radiomics and radiomics-clinical models built using different machine learning approaches.

Machine learning approaches	Radiomics models		Radiomics-clinical models	
	AUC in the validation set	AUC in the training set	AUC in the validation set	AUC in the training set
LASSO	0.813	0.873	0.850	0.988
KNN	0.804	0.785	0.838	1.000
GBDT	0.785	0.906	0.841	0.999
CART	0.748	0.742	0.721	0.799
RF	0.880	0.864	0.845	0.998
Ada	0.795	0.826	0.840	0.949
XGB	0.815	0.875	0.866	0.898
LR	0.811	0.951	0.853	0.989
MLP	0.898	0.973	0.913	0.929

Abbreviations: Ada, adaptive boosting; AUC, area under the curve; CART, classification and regression tree; GBDT, gradient boosting decision tree; KNN, k-nearest-neighbor; LASSO, least absolute shrinkage and selection operator; LR, logistic regression; MLP, multilayer perception; RF, random forest; XGB, XGBoost.

marrow metastasis. The nomogram is shown in Figure 2. The AUCs on the training and validation sets were 0.98 (95% CI: 0.96–0.99) and 0.90 (95% CI: 0.84–0.96), respectively. When the radiomics feature and clinical factors were incorporated into the machine learning approaches in the radiomics-clinical model development, MLP, XGB, and LR were the top three approaches. The accuracy of these models and comparisons with the other radiomics models are shown in Table 1. The AUCs of the top three approaches (MLP, XGB, and LR) were 0.93 (95% CI: 0.89–0.96), 0.90 (95% CI: 0.85–0.94), and 0.99 (95% CI: 0.98–1.00), respectively, on the training set and 0.91 (95% CI: 0.85–0.97), 0.86 (95% CI: 0.77–0.94), and 0.85 (95% CI: 0.76–0.93) on the validation set (Figure 3a,b). The accuracy, sensitivity, specificity, NPV, PPV, and F1 scores of the top three radiomics-clinical models and nomograms are listed in Supporting Information S1: Table S3. The calibration curves and radar plots for the included features are presented in Figure 3c–k. From the above results, we can conclude that the prediction accuracy of most approaches will increase when clinical variables are included. The feature importance of the top three radiomics-clinical models is shown in Supporting Information S1: Table S4. Figure 4 shows that with an increase in prediction probability, the observed probability also increases. We also performed t tests on the model-predicted probabilities of metastasis and nonmetastasis patients in the training and validation sets (Figure 5). All top three radiomics and radiomics-clinical models yielded a p value < 0.001.

4 | DISCUSSION

In this study, we developed and validated multiple radiomics and clinical-radiomics models to personalize the prediction of bone marrow metastasis in neuroblastoma. These models effectively stratify patients according to the likelihood of bone marrow metastasis and provide guidance for more frequent follow-ups or aggressive preventive protocols. By using these machine learning prediction models, we could potentially reduce the frequency of invasive examinations without compromising the accuracy of detecting bone marrow metastasis, thereby improving the quality of life of children with neuroblastoma.

Among the various models evaluated, the MLP demonstrated the highest diagnostic accuracy in both the radiomics and radiomics-clinical models. In the radiomics model, the AUCs were 0.97 (0.95–0.99) on the training set and 0.90 (0.82–0.95) on the validation set. In the radiomics-clinical model, the AUCs were 0.93 (0.89–0.96) and 0.91 (0.85–0.97), respectively. Despite achieving satisfactory AUC values, the accuracy scores were 0.807 for the MLP radiomics model and 0.754 for the radiomics-clinical model. Thus, further efforts are required to improve the accuracy of this prediction model before its clinical application, such as increasing the size of the training data set.

Notably, the sensitivity of the radiomics-clinical models was only 0.609, although the specificity was 0.853. Consequently, our model can only identify a small portion of patients with metastasis. However, given its high specificity, we are confident that it could be used to reduce the frequency of bone marrow biopsies or aspirations in patients predicted to not have metastasis.

Several strategies exist for diagnosing bone marrow metastasis in cancer, including radiomics, clinical prediction nomograms, and molecular profiling [25, 26]. However, in the case of neuroblastoma, intratumor heterogeneity must be taken into consideration. Intratumor heterogeneity is the main cause of cancer progression and treatment resistance, which poses challenges for both physicians and patients. For

TABLE 3 Univariate and multivariate logistic regression analyses of the MLP probability and clinical candidate predictors in the training set.

Variables	Univariate logistic regression			Multivariate logistic regression		
	OR	OR (95% CI)	p value	OR	OR (95% CI)	p value
MLP probability	383,080	5902–24,861,724	<0.001*	20,606,765	27,793–15,278,328,143	<0.001*
Gender	1.46	0.72–2.96	0.29			
Age of month	1.02	1.00–1.03	0.01*	0.98	0.95–1.01	0.23
Calibration	1.42	0.61–3.32	0.41			
Cross midline	0.36	0.15–0.89	0.03*	1.12	0.15–8.59	0.92
Ascites	2.90	1.18–7.13	0.02*	4.44	1.28–84.85	0.03*
Lymph gland metastasis	3.07	1.50–6.31	0.002*	2.73	0.51–15.40	0.24

Abbreviations: CI, confidence interval; MLP, multilayer perception; OR, odds ratio.
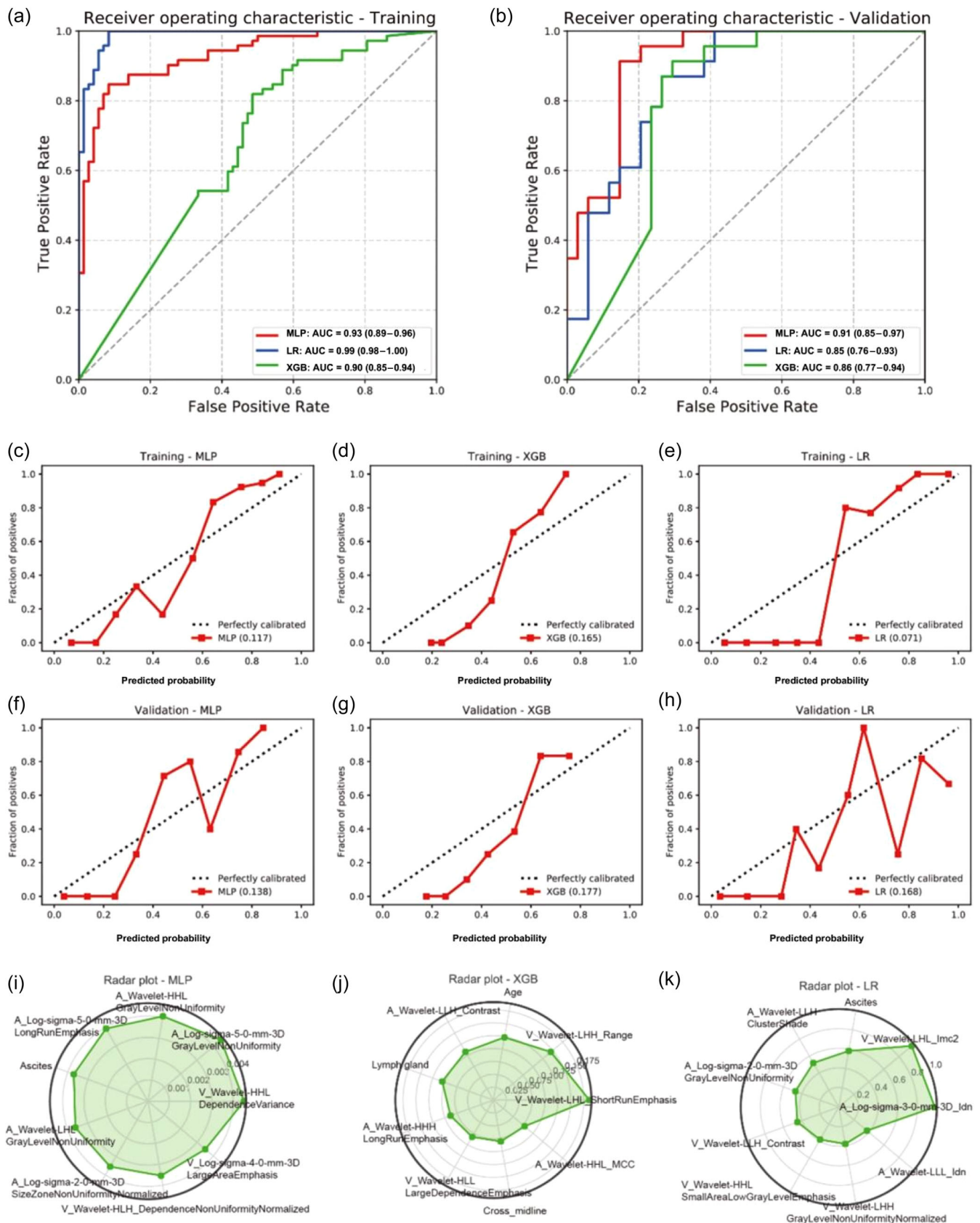
*p < 0.05, statistically significant.

FIGURE 2 Diagnostic performance of radiomics-clinical models. (a) Receiver operating characteristics (ROC) curves and area under the curves (AUCs) of the top three radiomics-clinical models applied to the training set. (b) ROC curves and AUCs of the top three radiomics-clinical models applied to the validation set. (c–h) Calibration curves of the top three radiomics-clinical models for the training and validation sets. (i–k) Radar plots of the top three radiomics-clinical models showing the most important features and their coefficients.
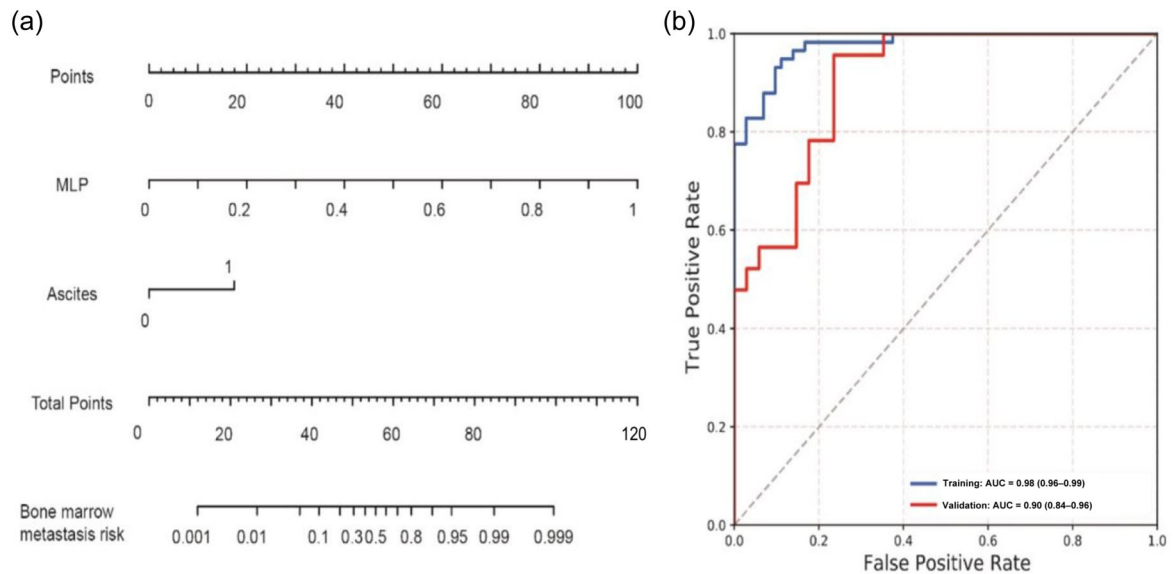
FIGURE 3 Nomogram for bone marrow metastasis in neuroblastoma (a) and ROC curves for the nomogram applied to training and validation sets (b). ROC, Receiver operating characteristics.
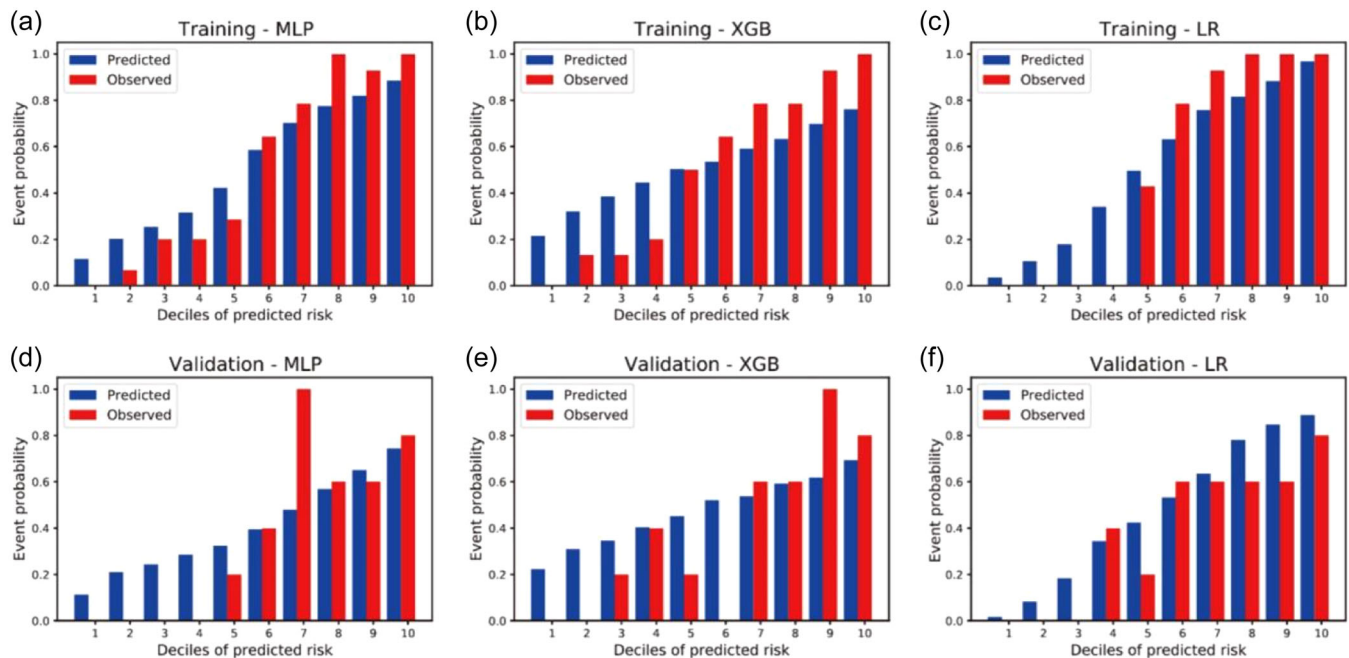

FIGURE 4 Risk deciles of the radiomics-clinical models represented as bar plots (observed vs. predicted risk) for the training and validation sets. Observed risk means the true bone marrow status of patients, the results of bone marrow aspiration, biopsy, or PET/CT scan. Predicted risk means the probability of bone marrow metastasis predicted by the machine learning approaches. (a) MLP in the training set, (b) XGB in the training set, (c) LR in the training set, (d) MLP in the validation set, (e) XGB in the validation set, and (f) LR in the validation set. CT, computed tomography; LR, logistic regression; MLP, multilayer perception; PET, positron emission tomography; XGB, XGBoost.

instance, Schmelz and colleagues reported extensive spatial and temporal intratumor heterogeneity within neuroblastoma [4]. As a result, molecular profiling based on bulk sequencing approaches from single-

tumor biopsies often proves inefficient or only exhibits short-term relevance. Therefore, radiomics or clinical prediction nomograms may be more suitable methods for prediction in this context.
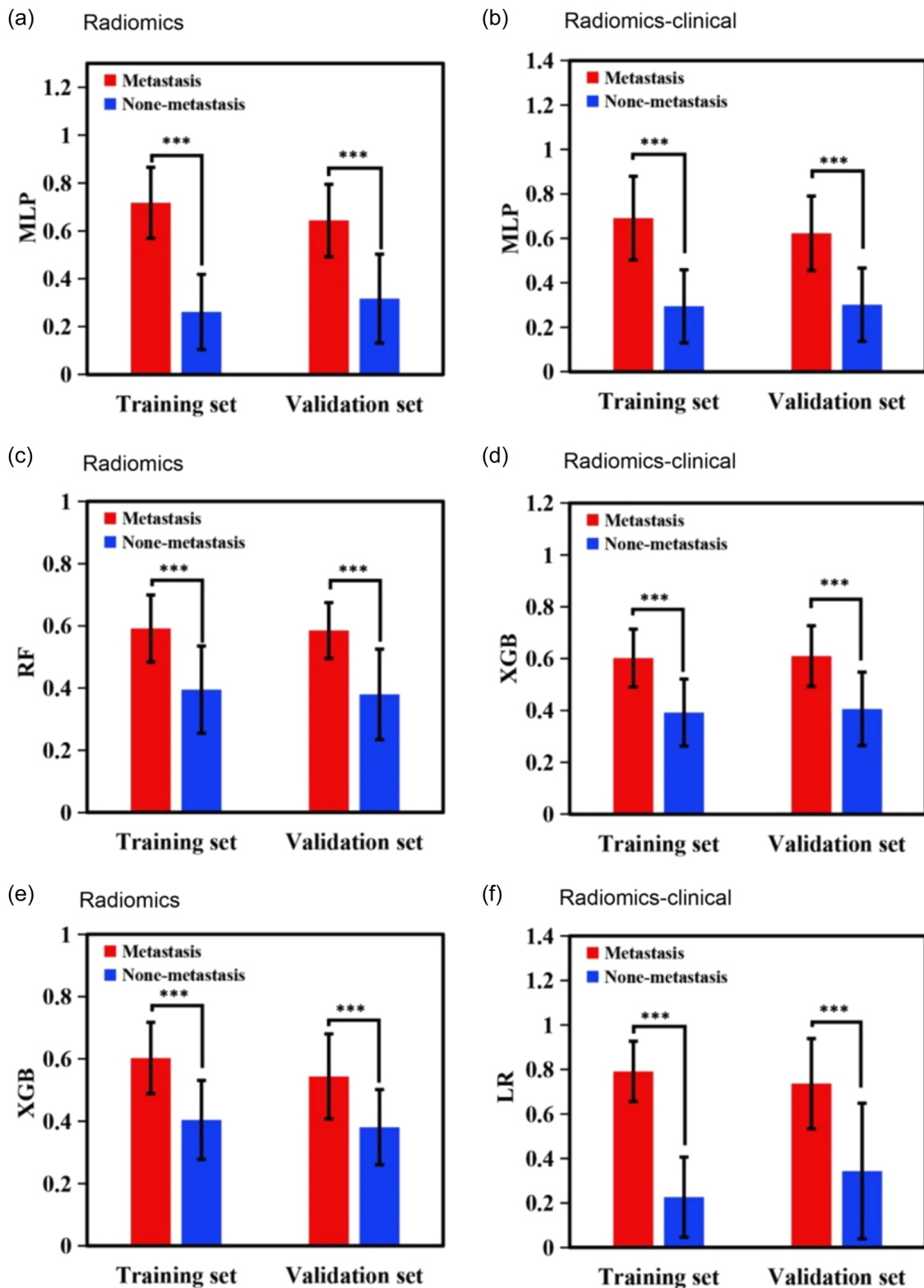
FIGURE 5 Results of t tests of the models' prediction of metastasis and nonmetastasis neuroblastoma patients in the training and validation sets. (a) MLP-based radiomics model; (b) MLP-based radiomics-clinical model; (c) RF-based radiomics model; (d) XGB-based radiomics-clinical model; (e) XGB-based radiomics model; and (f) LR-based radiomics-clinical model. LR, logistic regression; MLP, multilayer perception; RF, random forest; XGB, XGBoost. ***$p < 0.001$.

In the past, many researchers have dedicated themselves to achieving more precise diagnoses of cancers, prediction of cancer outcomes (including metastasis), and assessment of treatment responses using deep learning and machine learning methods. For example, we previously developed a deep learning model for the detection of bladder cancer that has been successfully implemented in practice [27]. Ji et al. successfully constructed a CT radiomics-based model to predict lymph node metastasis in biliary tract cancer [28]. Dercle et al. developed a radiomics response signature using CT to identify metastatic colorectal cancer sensitive to therapies targeting the EGFR pathway [29].

In our study, we combined both radiomics and clinical variables associated with bone marrow metastasis in our nomogram and applied various feature selection algorithms. This approach has the potential to provide individualized precision in predicting the probability of bone marrow metastasis, without adding any risks or costs. Patients with a high tendency for bone marrow metastasis will be recommended to undergo bone marrow aspiration, biopsy, or positron emission tomography (PET)/CT, while those at low risk will be advised to undergo active surveillance only. In this way, we can strike a balance between the adverse effects of more aggressive examinations and treatments and the risks posed by cancer progression.

Our study demonstrated that most models combining radiomics features and clinical factors outperformed models that only included radiomics features. Unsurprisingly, the accuracy of some models decreased when different machine learning algorithms were employed because the effectiveness of algorithms can vary.

Several studies also focused on detecting bone marrow involvement using machine learning approaches. For example, Mayerhoefer et al. developed a PET/CT radiomics model with an MLP to predict bone marrow involvement in mantle cell lymphoma in 97 patients [30]. The AUC of their model reached 0.81, whereas the AUC of our model reached 0.91. Our higher value is possibly due to our larger sample size. We compared nine different machine learning approaches and selected the best model, whereas Marius only used one machine learning approach. We believe that different machine learning approaches may perform differently on different datasets and in different clinical scenarios.

Wennmann et al. combined deep learning and radiomics to automatically detect disseminated bone marrow involvement in multiple myeloma [31]. In this study, the regions of interest were the whole skeleton, whereas in our study, the regions of interest were neuroblastomas. Using the skeleton as the region of interest might lead to better performance in predicting bone marrow involvement because the skeletal region directly represents the bone marrow. However, as mentioned earlier, neuroblastoma exhibits high temporal and spatial heterogeneity, and repeat whole-body CT scans expose children to high doses of radiation. Many studies have shown that analysis of specific regions of cancer can reveal its heterogeneity and propensity for metastasis [32, 33].

Our in-depth analysis suggested that venous phase features may be as important as arterial phase features. Both phases were included in both the radiomics and radiomics-clinical models. This is consistent with clinical knowledge that both arterial and venous phases present signs associated with disease progression. Recently, Chen et al. reported a radiomics nomogram to predict MYCN status in neuroblastoma [17]. Their study indicated that venous phases outperformed arterial phases. We compared the features included in our nomogram with those in their study but found no common features, suggesting that the important radiomics features for MYCN status and bone marrow metastasis may be different.

Multiple machine learning approaches were used to construct the radiomics and radiomics-clinical models. Our results indicate that MLP, RF, and XGB performed well in the construction of the radiomics model, whereas LR, XGB, and MLP performed well in the construction of the radiomics-clinical model. Many previous studies constructed radiomics models using LASSO regression but, in our study, LASSO was not the most suitable feature selection algorithm. Therefore, we encourage researchers to compare different machine learning approaches to obtain a more accurate prediction model.

Data independence is crucial in the model development process. In our study, CT data from the same patient were only included in either the training set or the validation set. After dividing the data into training and validation sets, we stored them as two independent files to avoid data leakage.

The reproducibility of radiomics models is an ongoing concern [34]. A reliable prediction model should be based on stable and reproducible predictors. In our study, the regions of interest were manually delineated or semi-automatically generated, making complete elimination of interobserver heterogeneity challenging. To address this, the regions of interest in the training and validation sets were independently delineated by two radiologists. The high AUC on the training set indicates that interobserver heterogeneity during delineation did not significantly affect the results. Therefore, our model has the potential to be generalized to other institutions.

Our model has the following limitations. First, only the arterial and venous contrast phases were included in our study, not the plain phase. We excluded the plain

phase because we observed that tumor boundaries were not well distinguished in the images. Tumor border characteristics can greatly influence the aggressiveness of the tumor. Because we did not have an accurate method to obtain clear boundaries on the plain phase, we believe that a model constructed solely using plain-phase images would be less reliable. Therefore, we opted to include only the arterial and venous contrast phases. Second, our study lacks multicenter validation. Different CT scanners and imaging environments may introduce variations in the original CT images, which could impact the predictive efficacy of our model in other medical centers. It is important to validate the performance of our model across multiple centers to ensure its generalizability. Third, we did not directly compare the diagnostic accuracy of our machine learning model with that of radiologists. In clinical practice, radiologists typically rely on bone marrow aspiration, biopsy, or PET/CT scans to detect bone marrow metastasis. These methods are considered the reference standards. However, our aim was to develop a noninvasive diagnostic model that could complement or assist the existing protocols. We evaluated and compared the accuracy of our model with clinically used protocols using metrics such as ROC and 10% risk deciles. We acknowledge these limitations and believe that future research should address these issues to further improve the reliability and applicability of our model.

5 | CONCLUSION

In this study, we constructed multiple machine learning radiomics and radiomics-clinical models using venous and arterial enhanced-phase CT scans to predict bone marrow metastasis of neuroblastoma. The MLP-based radiomics and radiomics-clinical models outperformed other machine learning approaches, with AUCs exceeding 0.90 on the validation sets.

AUTHOR CONTRIBUTIONS

Xiong Chen: Conceptualization; data curation; formal analysis; methodology; visualization; writing—original draft. **Qinchang Chen**: Conceptualization; investigation; validation; writing—review and editing. **Yuanfang Liu**: Data curation; formal analysis; investigation; validation. **Ya Qiu**: Data curation; formal analysis; investigation. **Lin Lv**: Formal analysis; methodology; visualization. **Zhengtao Zhang**: Data curation; formal analysis; investigation. **Xuntao Yin**: Conceptualization; project administration; writing—review and editing. **Fangpeng Shu**: Conceptualization; funding acquisition; project administration; writing—review and editing.

ORCID
Xiong Chen http://orcid.org/0000-0003-1765-3051
Xuntao Yin http://orcid.org/0000-0002-8228-6908

REFERENCES

1. Butler E, Ludwig K, Pacenta HL, Klesse LJ, Watt TC, Laetsch TW. Recent progress in the treatment of cancer in children. CA Cancer J Clin. 2021;71(4):315–32. https://doi.org/10.3322/caac.21665

2. Matthay KK, Maris JM, Schleiermacher G, Nakagawara A, MacKall CL, Diller L, et al. Neuroblastoma. Nat Rev Dis Primers. 2016;2:16078. https://doi.org/10.1038/nrdp.2016.78

3. Smith V, Foster J. High-risk neuroblastoma treatment review. Children. 2018;5(9):114. https://doi.org/10.3390/children5090114

4. Schmelz K, Toedling J, Huska M, Cwikla MC, Kruetzfeldt LM, Proba J, et al. Spatial and temporal intratumour heterogeneity has potential consequences for single biopsy-based neuroblastoma treatment decisions. Nat Commun. 2021;12(1):6804. https://doi.org/10.1038/s41467-021-26870-z

5. Cohn SL, Pearson ADJ, London WB, Monclair T, Ambros PF, Brodeur GM, et al. The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. J Clin Oncol. 2009;27(2):289–97. https://doi.org/10.1200/JCO.2008.16.6785

6. Colon NC, Chung DH. Neuroblastoma. Adv Pediatr. 2011;58(1):297–311. https://doi.org/10.1016/j.yapd.2011.03.011

7. Kushner BH, Kramer K, Modak S, Cheung NK. Sensitivity of surveillance studies for detecting asymptomatic and unsuspected relapse of high-risk neuroblastoma. J Clin Oncol. 2009;27(7):1041–6. https://doi.org/10.1200/JCO.2008.17.6107

8. Ladenstein R, Lambert B, Pötschger U, Castellani MR, Lewington V, Bar-Sever Z, et al. Validation of the mIBG skeletal SIOPEN scoring method in two independent high-risk neuroblastoma populations: the SIOPEN/HR-NBL1 and COG-A3973 trials. Eur J Nucl Med Mol Imaging. 2018;45(2):292–305. https://doi.org/10.1007/s00259-017-3829-7

9. Beiske K, Burchill SA, Cheung IY, Hiyama E, Seeger RC, Cohn SL, et al. Consensus criteria for sensitive detection of

minimal neuroblastoma cells in bone marrow, blood and stem cell preparations by immunocytology and QRT-PCR: recommendations by the International Neuroblastoma Risk Group Task Force. Br J Cancer. 2009;100(10):1627–37. https://doi.org/10.1038/sj.bjc.6605029

10. Zhang X, Li C, Xu C, Hao X, Yu X, Li Q. Correlation of CT signs with lymphatic metastasis and pathology of neuroblastoma in children. Oncol Lett. 2018;16(2):2439–43. https://doi.org/10.3892/ol.2018.8959

11. Bera K, Braman N, Gupta A, Velcheti V, Madabhushi A. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. Nat Rev Clin Oncol. 2022;19(2):132–46. https://doi.org/10.1038/s41571-021-00560-7

12. Guiot J, Vaidyanathan A, Deprez L, Zerka F, Danthine D, Frix AN, et al. A review in radiomics: making personalized medicine a reality via routine imaging. Med Res Rev. 2022;42(1):426–40. https://doi.org/10.1002/med.21846

13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77. https://doi.org/10.1148/radiol.2015151169

14. 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. https://doi.org/10.1016/j.ejrad.2020.109020

15. Hong JH, Jung JY, Jo A, Nam Y, Pak S, Lee SY, et al. Development and validation of a radiomics model for differentiating bone islands and osteoblastic bone metastases at abdominal CT. Radiology. 2021;299(3):626–32. https://doi.org/10.1148/radiol.2021203783

16. Wu H, Wu C, Zheng H, Wang L, Guan W, Duan S, et al. Radiogenomics of neuroblastoma in pediatric patients: CT-based radiomics signature in predicting MYCN amplification. Eur Radiol. 2021;31(5):3080–9. https://doi.org/10.1007/s00330-020-07246-1

17. Chen X, Wang H, Huang K, Liu H, Ding H, Zhang L, et al. CT-based radiomics signature with machine learning predicts MYCN amplification in pediatric abdominal neuroblastoma. Front Oncol. 2021;11:687884. https://doi.org/10.3389/fonc.2021.687884

18. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 2017;77(21):e104–7. https://doi.org/10.1158/0008-5472.CAN-17-0339

19. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321–57. https://doi.org/10.1613/jair.953

20. D'Ascenzo F, De Filippo O, Gallone G, Mittone G, Deriu MA, Iannaccone M, et al. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets. Lancet. 2021;397(10270):199–207. https://doi.org/10.1016/S0140-6736(20)32519-8

21. Viana Dos Santos Santana Í, CM da Silveira A, Sobrinho Á, Chaves E Silva L, Dias da Silva L, Santos DFS, et al. Classification models for COVID-19 test prioritization in Brazil: machine learning approach. J Med Internet Res. 2021;23(4):e27293. https://doi.org/10.2196/27293

22. Adamidi ES, Mitsis K, Nikita KS. Artificial intelligence in clinical care amidst COVID-19 pandemic: a systematic review. Comput Struct Biotechnol J. 2021;19:2833–50. https://doi.org/10.1016/j.csbj.2021.05.010

23. Su Y, Ma XL, Wang HM, Qin H, Qin MQ, Zhang FQ, et al. Clinical characteristics and prognostic analysis of 458 children with high-risk neuroblastoma in a single center. Zhonghua Er Ke Za Zhi. 2020;58(10):796–801. https://doi.org/10.3760/cma.j.cn112140-20200525-00540

24. Wang R, Dai W, Gong J, Huang M, Hu T, Li H, et al. Development of a novel combined nomogram model integrating deep learning-pathomics, radiomics and immunoscore to predict postoperative outcome of colorectal cancer lung metastasis patients. J Hematol Oncol. 2022;15(1):11. https://doi.org/10.1186/s13045-022-01225-3

25. Lazic D, Kromp F, Rifatbegovic F, Repiscak P, Kirr M, Mivalt F, et al. Landscape of bone marrow metastasis in human neuroblastoma unraveled by transcriptomics and deep multiplex imaging. Cancers. 2021;13(17):4311. https://doi.org/10.3390/cancers13174311

26. Hochheuser C, van Zogchel LMJ, Kleijer M, Kuijk C, Tol S, van der Schoot CE, et al. The metastatic bone marrow niche in neuroblastoma: altered phenotype and function of mesenchymal stromal cells. Cancers. 2020;12(11):3231. https://doi.org/10.3390/cancers12113231

27. Wu S, Chen X, Pan J, Dong W, Diao X, Zhang R, et al. An artificial intelligence system for the detection of bladder cancer via cystoscopy: a multicenter diagnostic study. J Natl Cancer Inst. 2022;114(2):220–7. https://doi.org/10.1093/jnci/djab179

28. Ji GW, Zhang YD, Zhang H, Zhu FP, Wang K, Xia YX, et al. Biliary tract cancer at CT: a radiomics-based model to predict lymph node metastasis and survival outcomes. Radiology. 2019;290(1):90–8. https://doi.org/10.1148/radiol.2018181408

29. Dercle L, Lu L, Schwartz LH, Qian M, Tejpar S, Eggleton P, et al. Radiomics response signature for identification of metastatic colorectal cancer sensitive to therapies targeting EGFR pathway. J Natl Cancer Inst. 2020;112(9):902–12. https://doi.org/10.1093/jnci/djaa017

30. Mayerhoefer ME, Riedl CC, Kumar A, Dogan A, Gibbs P, Weber M, et al. [18F]FDG-PET/CT radiomics for prediction of bone marrow involvement in mantle cell lymphoma: a retrospective study in 97 patients. Cancers. 2020;12(5):1138. https://doi.org/10.3390/cancers12051138

31. Wennmann M, Klein A, Bauer F, Chmelik J, Grözinger M, Uhlenbrock C, et al. Combining deep learning and radiomics for automated, objective, comprehensive bone marrow characterization from whole-body MRI: a multicentric feasibility study. Invest Radiol. 2022;57(11):752–63. https://doi.org/10.1097/RLI.0000000000000891

32. Tharmaseelan H, Hertel A, Tollens F, Rink J, Woźnicki P, Haselmann V, et al. Identification of CT imaging phenotypes of colorectal liver metastases from radiomics signatures-towards assessment of interlesional tumor heterogeneity. Cancers. 2022;14(7):1646. https://doi.org/10.3390/cancers14071646

33. Liu X, Zhang D, Liu Z, Li Z, Xie P, Sun K, et al. Deep learning radiomics-based prediction of distant metastasis in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy: a multicentre study. EBioMedicine. 2021;69:103442. https://doi.org/10.1016/j.ebiom.2021.103442

34. Ziegelmayer S, Reischl S, Harder F, Makowski M, Braren R, Gawlitza J. Feature robustness and diagnostic capabilities of convolutional neural networks against radiomics features in computed tomography imaging.

Invest Radiol. 2022;57(3):171–7. https://doi.org/10.1097/RLI.0000000000000827

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.