



OPEN

Fisher discriminant model based on LASSO logistic regression for computed tomography imaging diagnosis of pelvic rhabdomyosarcoma in children

Lu Tian¹, Xiaomeng Li², Helin Zheng², Longlun Wang², Yong Qin^{2,3}✉ & Jinhua Cai^{2,3}✉

Computed tomography (CT) has been widely used for the diagnosis of pelvic rhabdomyosarcoma (RMS) in children. However, it is difficult to differentiate pelvic RMS from other pelvic malignancies. This study aimed to analyze and select CT features by using least absolute shrinkage and selection operator (LASSO) logistic regression and established a Fisher discriminant analysis (FDA) model for the quantitative diagnosis of pediatric pelvic RMS. A total of 121 pediatric patients who were diagnosed with pelvic neoplasms were included in this study. The patients were assigned to an RMS group (n = 36) and a non-RMS group (n = 85) according to the pathological results. LASSO logistic regression was used to select characteristic features, and an FDA model was constructed for quantitative diagnosis. Leave-one-out cross-validation and receiver operating characteristic (ROC) curve analysis were used to evaluate the diagnostic ability of the FDA model. Six characteristic variables were selected by LASSO logistic regression, all of which were CT morphological features. Using these CT features, the following diagnostic models were established: (RMS group) $G_1 = -14.283 + 6.613x_1 + 5.333x_2 + 5.753x_3 + 12.361x_4 + 8.095x_5 - 0.715x_6$; (Non-RMS group) $G_2 = -2.008 + 3.539x_1 + 1.080x_2 + 1.154x_3 + 2.307x_4 + 1.656x_5 + 1.380x_6$, where x_1, x_2, \dots and x_6 are lower than normal muscle density (1 = yes; 0 = no), multinodular fusion (1 = yes; 0 = no), enhancement at surrounding blood vessels (1 = yes; 0 = no), heterogeneous progressive centripetal enhancement (1 = yes; 0 = no), ring enhancement (1 = yes; 0 = no), and hemorrhage (1 = yes; 0 = no), respectively. The calculated area under the ROC curve (AUC) of the model was 0.992 (0.982–1.000), with a sensitivity of 94.4%, a specificity of 96.5%, and an accuracy of 95.9%. The calculated sensitivity, specificity and accuracy values were consistent with those from cross-validation. An FDA model based on the CT morphological features of pelvic RMS was established and could provide an easy and efficient method for the diagnosis and differential diagnosis of pelvic RMS in children.

Abbreviations

FDA	Fisher discriminant model
LASSO	Least absolute shrinkage and selection operator
CT	Computed tomography
RMS	Rhabdomyosarcoma
ROC	Receiver operating characteristic
AUC	Area under the ROC curve
MRI	Magnetic resonance imaging
SD	Standard deviation

¹Department of Radiology, Children's Hospital, Chongqing Medical University and Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing International Science and Technology Cooperation Center for Child Development and Disorders and Key Laboratory of Pediatrics in Chongqing, Chongqing 400014, China. ²Children's Hospital of Chongqing Medical University, Chongqing, China. ³These authors contributed equally: Yong Qin and Jinhua Cai. ✉email: 308787364@qq.com; cai_jinhua@126.com

Rhabdomyosarcoma (RMS) is a soft tissue sarcoma that accounts for approximately 4.5% of pediatric cancer cases and is characterized by a high degree of malignancy, infiltration of adjacent tissues, and early lymph node and distant metastases¹. The pelvic cavity is one of the most common sites, second only to the head and neck². Accurate preoperative diagnosis is very important to provide appropriate planning for surgery and could lead to a better prognostic outcome. In recent years, medical imaging technology has made great progress and is playing an increasingly important role in the diagnosis of tumours. Of many imaging modalities, computed tomography (CT) has the advantages of noninvasiveness, high spatial and density resolution, and high scanning speed and has been widely applied for the assessment of tumours in the abdomen, pelvis and thorax. CT images allow radiologists not only to determine the extent of tumours but also to determine the absence or presence of bony destruction, calcification, haemorrhage and/or metastases³. However, in fact, during the course of routine radiology diagnosis, because RMS has the general radiological appearance of soft tissue tumors, it is difficult to distinguish it from other pelvic soft tissue malignancies. However, to our knowledge, there are currently few literature reports on the CT features of pelvic RMS^{4–9}, and some reports have indicated that the CT findings of pelvic RMS lack specificity^{10–12}. Therefore, it is imperative to investigate the CT features of pelvic RMS and establish an accurate diagnostic method for pelvic RMS in children.

Least absolute shrinkage and selection operator (LASSO), first proposed by Robert Tibshirani in 1996, is a regression analysis method to reduce the dimensionality of data. Serving as a regularized estimation, LASSO can shrink the coefficients of variables and force certain regression coefficients to 0 by constructing a penalty function, so it is often adopted for variable selection before establishing prediction and diagnostic models^{13,14}. Fisher discriminant analysis (FDA) is a classic method for identifying linear functions of variables to distinguish samples of different groups¹⁵. It was proposed by Fisher in 1936 and serves as a dimensionality reduction method that finds a linear combination of features that maximizes the between-class differences and minimizes the within-class variation¹⁶. FDA has been widely used in many fields, including medical research¹⁷. To date, a variety of studies have applied FDA to predict, diagnose or identify diseases^{18–20}. A study by Ni et al.¹⁸ applied FDA to predict clinicopathological subtypes of breast cancer based on the radiological features of diffusion-weighted magnetic resonance imaging (MRI) and suggested that FDA was a promising method for predicting clinicopathological subtypes of breast cancer. In a report by Zou et al.¹⁹, FDA was applied to classify autism spectrum disorders (ASDs) based on folate-related metabolic markers, and the results showed that the FDA model could effectively distinguish ASD patients from healthy controls. Hao et al.²⁰ used FDA to establish a discriminant formula to distinguish patients with gastric cancer and colorectal cancer from healthy controls and achieved good results.

In this study, we used LASSO logistic regression to evaluate and select valuable CT features. Based on these features, an FDA model for pelvic RMS diagnosis was established, and the diagnostic accuracy was validated. Our aim was to develop a simple and accurate diagnostic method for the quantitative diagnosis of pelvic RMS by means of mathematical statistics.

We present the following article in accordance with the Tripod Checklist.

Materials and methods

Clinical data and pathological results were obtained from patients' medical records. Imaging data were retrieved from our picture archiving and communication system.

Study subjects. A total of 121 pediatric patients who underwent abdominal contrast-enhanced CT scans and were diagnosed with pelvic tumors from January 2013 to January 2021 were included in this study. According to the pathological results, these patients were divided into RMS and non-RMS groups. The RMS group included 36 patients (14 males, 22 females) with ages ranging from 1 month to 16 years, and the non-RMS group included 85 patients (24 males, 61 females) with ages ranging from 1 month to 15 years. Those in the non-RMS group suffered from a variety of malignant tumors, including 19 patients with yolk sac tumors, 18 patients with malignant teratoma, 13 patients with neuroblastoma, 9 patients with lymphoma, 5 patients with Ewing sarcoma and 21 patients with other types of malignant reproductive tumors.

CT imaging and feature extraction. All pediatric patients in our study underwent pelvic plain and contrast-enhanced CT examinations on a 256-slice spiral CT system (Philips Brilliance iCT, Philips, Netherlands) with a low tube voltage of 80–100 kV, a low tube current of 100–200 mA (the current varied during the acquisition and according to the child's body weight), a rotation time of 0.4 s, a pitch of 0.925, a collimation of 128*0.625 mm, a slice thickness of 5.0 mm and a reconstruction layer thickness of 1 mm. To obtain arterial and venous phase contrast-enhanced images, dual-phase dynamic contrast-enhanced CT was performed at 30 and 65 s, respectively, after intravenous injection of the contrast agent (Omnipaque, 350 mg/mL, Amersham Healthcare, Shanghai, China). The contrast agent was administered at a dose of 2 mL/kg body weight, with a maximum of 80 mL.

During the course of routine radiology diagnosis, for tumor lesions, we mainly need to pay attention to their morphology, density, margin, enhancement mode and metastasis. Therefore, our study mainly selected image features from these five aspects to explore the differences between the two groups (RMS group and non-group). According to previous literature and diagnostic experience. The following CT features of the tumors were evaluated: (a) morphology (multinodular fusion/lobulated/round/orbicular); (b) density (lower or higher than normal muscle density/calcification/hemorrhage/necrosis); (c) margin (clear or unclear); (d) contrast enhancement modes (surrounding blood vessels/homogeneous progressive centripetal enhancement/ring enhancement/grape cluster reinforcement); (e) metastasis (lymphatic metastasis/bone erosion). The evaluation criteria for the above relatively special CT features are as follows: (1) multinodular fusion: in the CT images, multiple nodules of different sizes were observed in the pelvic cavity. Some of the nodules were fused together and fused into a

lobulated mass²¹; (2) surrounding blood vessels: multiple strip-like and punctate vascular shadows can be seen in the mass on CT enhanced scan images²¹; (3) homogeneous progressive centripetal enhancement: on dynamic contrast-enhanced CT images, the mass can be seen with peripheral annular inhomogeneous enhancement in the arterial phase, and gradually centripetal inhomogeneous enhancement in the venous and delayed phases²¹; (4) grape cluster reinforcement: when the mass is in a hollow structure (vaginal or bladder, etc.), a mass like a grape cluster will appear on the CT-enhanced image²²; (5) lymph node metastasis: cervical lymph nodes I, II and inguinal lymph nodes short diameter ≥ 1.5 cm, other cervical lymph nodes short diameter ≥ 1 cm, or the degree of enhancement was significantly higher than muscle tissue²³.

Prospective evaluation of the CT images for each patient was independently performed by two abdominal radiologists with 10–20 years of experience. The abdominal radiologists were blinded to patient characteristics and histologic results and evaluated the morphology, density, margin, enhancement modes and metastasis of the tumors. In case of disagreement between the two radiologists, a consensus was reached with a third senior abdominal radiologist and two other abdominal radiologists.

Statistical analysis. The patients were divided into an RMS group and a non-RMS group according to the pathological results. For quantitative variables, continuous variables that followed a normal distribution are described as the mean and standard deviation (SD), and a parametric t test was used to determine the statistical significance between the two groups. Otherwise, the variables are described as medians and interquartile ranges (IQRs), and a nonparametric Mann–Whitney U test was used for comparisons between the two groups. Categorical variables are expressed as the number of patients and respective percentage, and the χ^2 test or Fisher's exact test was used to compare the rates.

LASSO logistic regression was used to select the optimal characteristic features for diagnosing RMS from the basic and CT morphological features of the patients. The penalty parameter λ was optimized, and the resulting nonzero coefficient variables in the model were selected as the diagnostic variables. Based on these findings, FDA was established as a quantitative diagnostic model of pediatric pelvic RMS. The diagnostic ability of this model was evaluated by the receiver operating characteristic (ROC) curve. Additionally, the cumulative diagnostic ability of the features was analyzed.

A two-tailed $P < 0.05$ was considered to be statistically significant. All statistical analyses were performed using R 3.6.1 software and SPSS 23.0 software (IBM, Armonk, New York, USA).

Establishing the model. FDA is a classical approach to identify a linear function of variables to distinguish samples from different groups as much as possible¹³. In our study, patients with RMS and without RMS were set as the two groups: G_1 (RMS group) and G_2 (non-RMS group). A total of 6 CT features (x_i) were used as diagnostic variables to establish a linear discriminant function:

$$G_i = \sum_{j=1}^6 c_j x_{ij} \quad i = 1, 2$$

By using the discriminant rule, the result of the examination was found to belong to G_1 or G_2 .

(1) Raw data matrix: Two matrices (W^1 , W^2) were constructed for G_1 and G_2 , with CT features as the column vectors and pediatric patients as the row vectors.

Data matrix of G_1 :

$$W^1 = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 2 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 36 & 1 & \dots & 0 \end{bmatrix}$$

Data matrix of G_2 :

$$W^2 = \begin{bmatrix} 37 & 0 & \dots & 1 \\ 51 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 121 & 0 & \dots & 1 \end{bmatrix}$$

(2) The mean column distributions of matrices W^1 and W^2 are as follows:

$$\overline{x_j^1} = \frac{1}{36} \sum_{i=1}^{36} X_{ij}^1 \quad j = 1, 2, \dots, 6$$

$$\overline{x_j^2} = \frac{1}{85} \sum_{i=1}^{85} X_{ij}^2 \quad j = 1, 2, \dots, 6$$

(3) The coefficients c_i were calculated using the differential calculus method.

(4) The discriminant function is defined as:

$$G_i = c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5 + c_6 x_6$$

(5) The discriminant values represented by G_1 and G_2 were calculated.

$$\overline{G}_1 = \sum_{i=1}^{36} c_i \overline{x}_i^1$$

$$\overline{G}_2 = \sum_{i=1}^{85} c_i \overline{x}_i^2$$

(6) The value of the Fisher discrimination function at the centroids was obtained as follows:

$$G_0 = \frac{36\overline{G}_1 + 85\overline{G}_2}{121}$$

(7) The above calculation was conducted using SPSS software with the following discriminant rule: If $\overline{G}_1 > G_0$, the sample of G belongs to G_1 , which means that the sample belongs to the RMS group. Otherwise, the sample belongs to G_2 ; that is, the sample belongs to the non-RMS group.

(8) The resulting discriminant functions for classification were also calculated using SPSS.

$$G_1 = c_{11}x_1 + c_{21}x_2 + c_{31}x_3 + c_{41}x_4 + c_{51}x_5 + c_{61}x_6 (\text{RMSgroup})$$

$$G_2 = c_{12}x_1 + c_{22}x_2 + c_{32}x_3 + c_{42}x_4 + c_{52}x_5 + c_{62}x_6 (\text{non-RMSgroup})$$

The values of G_1 and G_2 were calculated by substituting the CT features into the function. By comparing the G_1 and G_2 values, the subjects were classified according to the following principle: if $G_1 > G_2$, the subjects were classified into the RMS group; if $G_1 < G_2$, the subjects were classified into the non-RMS group.

Notations:

G_i : population of disease, $i = 1, 2$;

W^i : data matrix of G_i ;

c_i : coefficients of Fisher's discriminant function, $i = 0, 1, 2, \dots, 6$;

G_0 : values of Fisher discrimination function at centroids;

x_{ij} : content of the j CT feature in the i patient.

Model validation. Leave-one-out cross-validation, in which each respective case is classified using all cases other than that case for deriving the classification formula, was used to validate the accuracy of the model. In addition, ROC curves were used to validate the accuracy of the model, where an area under the ROC curve (AUC) between 0.5 and 0.7 represented a low diagnostic value, that between 0.7 and 0.9 represented a medium diagnostic value, and that more than 0.9 represented a high diagnostic value²⁴.

Ethical statement. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Institutional Ethics Committee of Children's Hospital Affiliated to Chongqing Medical University and individual consent for this retrospective analysis was waived.

Results

Basic and CT morphological features of the patients. RMS (36/121, 29.7%) was the most common pelvic malignant tumor in our study, followed by yolk sac tumor (19/121, 15.7%), malignant teratoma (18/121, 14.9%), neuroblastoma (13/121, 10.7%), lymphoma (9/121, 7.4%) and Ewing sarcoma (5/121, 4.1%). There was no significant difference in age or sex between the two groups (both $P > 0.05$). A comparison of the basic and CT morphological characteristics between the two groups is shown in Table 1.

Selection of diagnostic features. LASSO logistic regression included a total of 16 basic factors (2/16) and CT morphological characteristics (14/16) of all patients to select potential diagnostic factors. The LASSO regression partial likelihood deviation and coefficient profiles against $\log(\lambda)$ are shown in Fig. 1. Six variables with nonzero coefficients were selected, all of which were CT morphological features, including lower than normal muscle density, multinodular fusion, enhancement at surrounding blood vessels, heterogeneous progressive centripetal enhancement, ring enhancement and hemorrhage. These variables then served as the preferred features for the diagnosis of RMS.

Construction of the diagnostic model. Through the construction of the fisher discriminant model, the discriminant functions of RMS and non-RMS were obtained as follows:

$$G_1 = -14.283 + 6.613x_1 + 5.333x_2 + 5.753x_3 + 12.361x_4 + 8.095x_5 - 0.715x_6;$$

$$G_2 = -2.008 + 3.539x_1 + 1.080x_2 + 1.154x_3 + 2.307x_4 + 1.656x_5 + 1.380x_6;$$

where x_1, x_2, \dots and x_6 represent lower than normal muscle density (1 = yes; 0 = no), multinodular fusion (1 = yes; 0 = no), enhancement at surrounding blood vessels (1 = yes; 0 = no), heterogeneous progressive centripetal enhancement (1 = yes; 0 = no), ring enhancement (1 = yes; 0 = no), and hemorrhage (1 = yes; 0 = no), respectively. Wilks' lambda of the model = 0.245, $\chi^2 = 163.237$, $df = 6$, and $P < 0.001$. The function value at the model centroid was 2.676 in the RMS group and -1.133 in the non-RMS group, and the critical value of the discriminant function was $G_0 = 0.800$.

	RMS group (n = 36)	Non-RMS group (n = 85)	Statistic	P value
Baseline				
Age (year) Q (Q1–Q3)	2.5 (1–6.5)	6 (1–9)	1.767 ^a	0.077
Sex (male/female)	14/22	24/61	1.332 ^b	0.248
Density N (%)				
Lower than normal muscle density	35 (97.2)	51 (60.4)	17.043 ^b	< 0.001
Calcification	1 (2.8)	23 (27.1)	9.377 ^b	0.002
Hemorrhage	5 (13.9)	28 (32.9)	4.628 ^b	0.031
Necrosis	35 (97.2)	71 (83.5)		0.066
Shape N (%)				
Multinodular fusion	23 (63.9)	10 (11.8)	34.641 ^b	< 0.001
Lobulated	28 (77.8)	44 (51.8)	7.102 ^b	0.008
Round/orbicular	6 (16.7)	22 (25.9)	1.208 ^b	0.272
Margin N (%)				
Unclear	8 (22.2)	26 (30.6)	0.876 ^b	0.349
Enhancement feature N (%)				
Surrounding blood vessels	29 (80.6)	16 (18.8)	41.257 ^b	< 0.001
Heterogeneous progressive centripetal enhancement	31 (86.1)	10 (11.8)	62.395 ^b	< 0.001
Ring enhancement	5 (13.9)	6 (7.1)	c	0.300
Grape cluster reinforcement	0 (0.0)	71 (0.0)	–	–
Metastasis N (%)				
Lymphatic metastasis	12 (33.3)	14 (16.5)	4.263 ^b	0.039
Bone erosion	2 (5.6)	5 (5.9)	c	1.000

Table 1. Comparison of the basic and CT morphological characteristics between the RMS group and the non-RMS group. Q is the median age, Q1–Q3 are 25–75% quantiles. CT computed tomography, RMS rhabdomyosarcoma. ^aUsing the M–U test. ^bUsing the Chi-square test. ^cUsing Fisher's exact probability test.

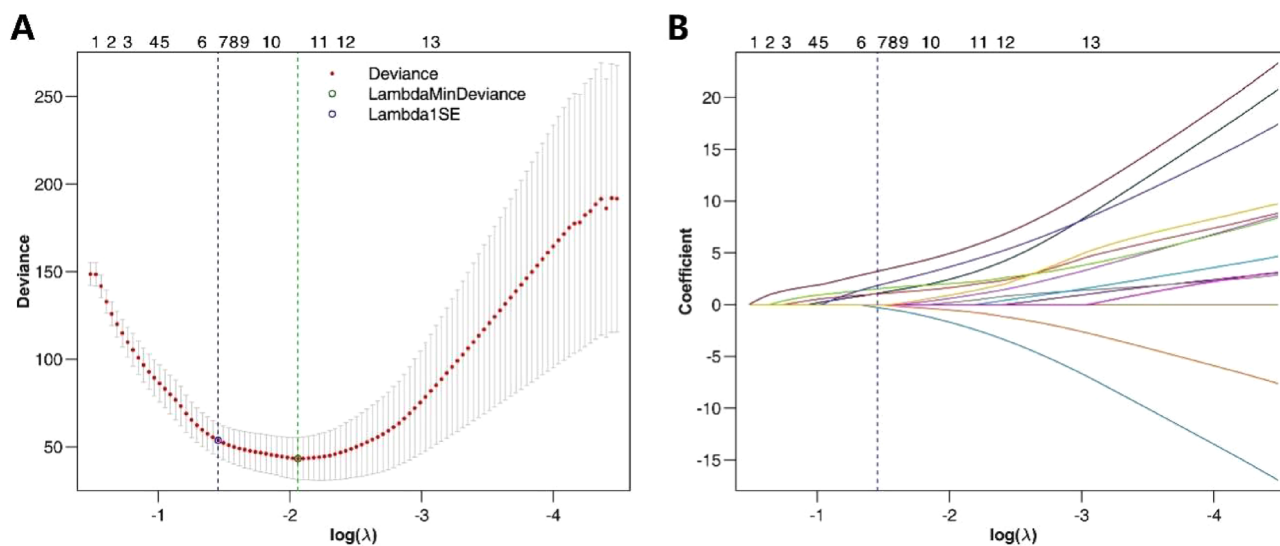


Figure 1. LASSO logistic regression plot. (A) Plot of partial likelihood deviance; (B) plot of LASSO coefficient profiles. Each colored curve represents the LASSO coefficient profile of a feature against the $\log(\lambda)$ sequence.

Accuracy and cross-validation of the diagnostic model. The ROC curve was used to verify the accuracy of the diagnostic model (Fig. 2). The AUC of the FDA score in this study was 0.992 (95% confidence interval: 0.982–1.000; sensitivity: 94.4%, specificity: 96.5%), and the accuracy of the model was 95.9%. The result of the cross-validation also showed an accuracy rate of 95.9%. The model correctly classified 34 patients (94.4%) in the RMS group, with a misclassification rate of 5.6%, and 82 patients (96.5%) in the non-RMS group, with a misclassification rate of 3.5% (Table 2).

Single feature analysis and cumulative FDA. The 6 selected CT morphological features were ordered by their importance according to the Fisher discriminant model as follows: heterogeneous progressive centrip-

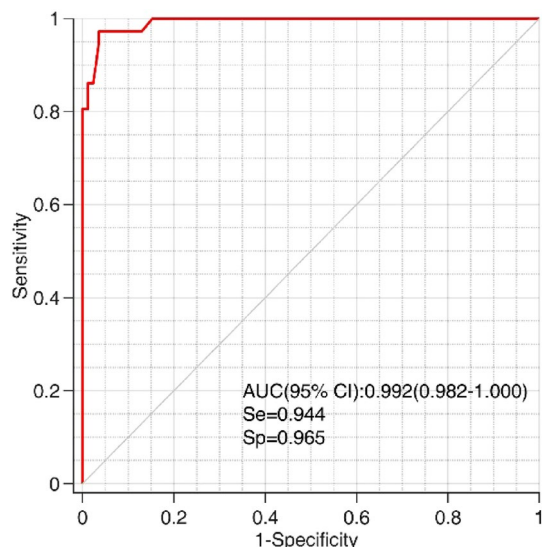


Figure 2. Receiver operating characteristic (ROC) curve of the quantitative diagnostic model for pelvic RMS in children using Fisher discriminant analysis.

	Predicted (n, %)	
	RMS	Non-RMS
Fisher model		
RMS group	33 (91.7)	3 (8.3)
Non-RMS group	2 (2.4)	83 (97.6)
Cross-validation		
RMS group	33 (91.7)	3 (8.3)
Non-RMS group	2 (2.4)	83 (97.6)

Table 2. Results of the Fisher model and cross-validation.

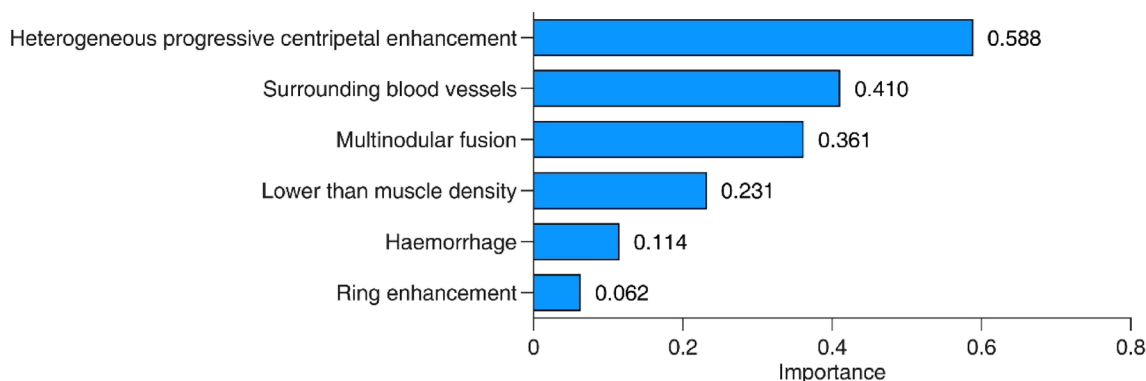


Figure 3. Importance of fisher discriminant model features.

etal enhancement, enhancement at surrounding blood vessels, multinodular fusion, lower than normal muscle density, hemorrhage and ring enhancement (Fig. 3). The sensitivity, specificity, AUC and 95% confidence interval of the 6 selected features are shown in Table 3 in order of importance. The sensitivity, specificity and AUC of the FDA score were significantly higher than those of each single CT characteristic feature.

Next, the feature importance was further evaluated using cumulative fisher discriminant models following the previously determined order of importance (Fig. 4). After the fourth cumulated feature, there was no significant improvement in the diagnostic ability of the discriminant models. Namely, there was no significant difference in the resulting AUCs (all above 0.96) for the first 4 indicators, the first 5 indicators and all 6 indicators (all

	Se (95% CI)	Sp (95% CI)	AUC (95% CI)
Heterogeneous progressive centripetal enhancement	0.861 (0.697–0.948)	0.882 (0.790–0.939)	0.872 (0.795–0.948)
Enhancement at surrounding blood vessels	0.806 (0.634–0.912)	0.812 (0.709–0.885)	0.809 (0.720–0.898)
Multinodular fusion	0.639 (0.462–0.787)	0.882 (0.790–0.939)	0.761 (0.658–0.864)
Lower than normal muscle density	0.972 (0.838–0.999)	0.400 (0.297–0.512)	0.686 (0.592–0.780)
Hemorrhage	0.861 (0.697–0.948)	0.329 (0.234–0.441)	0.595 (0.489–0.701)
Ring enhancement	0.139 (0.052–0.303)	0.929 (0.847–0.971)	0.534 (0.419–0.649)

Table 3. Sensitivity, specificity, AUC and 95% CI of the single characteristic features. *Se* sensitivity, *Sp* specificity, *CI* confidence interval, *AUC* area under the curve.

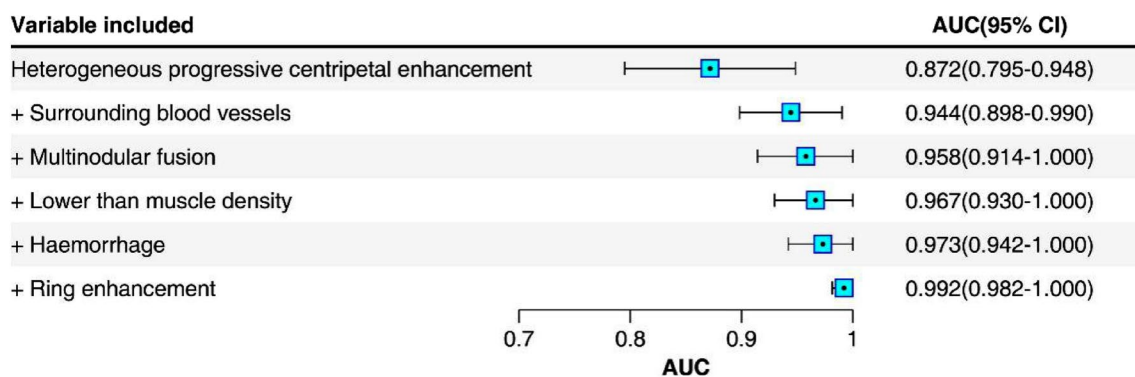


Figure 4. AUC of fisher's discriminant model and its 95% CI accumulated according to the importance of the features. *CI* confidence interval, *AUC* area under the curve.

$P > 0.05$). This result suggested that using fewer indicators can also accurately diagnose RMS, which is helpful to save resources.

Discussion

During routine radiological diagnosis, RMS is difficult to distinguish from other pelvic tumors except malignant teratoma, which involves characteristic calcification and fatty composition¹⁰. Previous studies have shown some CT features of RMS, such as tumoral necrosis, lower density than muscle, clear margin, and susceptibility to lymphatic node or/and bone metastasis^{25–37}. However, the specificity and sensitivity of these CT features have not been systematically analyzed due to the limited number of study samples. In this study, we extracted the important CT features of pelvic RMS by using LASSO logistic regression and established a quantitative diagnostic model for pelvic RMS by FDA. This diagnostic model is easy to operate, has a high diagnostic accuracy and can be applied in the diagnosis of pelvic RMS and differentiation from other pelvic tumors in children.

Our study showed that a total of 8 CT features of RMS were significantly different from those of other pelvic malignancies. The differences in these CT features may be related to their pathological characteristics. Studies have shown that pelvic RMS in children is highly invasive and characterized by multicentric growth^{35–38}. In the CT images, multiple nodules of different sizes were observed in the pelvic cavity, and some nodules fused together and formed a lobulated mass. Studies also found that some RMS tumor cells were distributed around the blood vessels using microscopes and that the blood vessels were gradually surrounded by tumor cells as they grew^{39,40}. This pathological phenomenon is consistent with multiple vascular shadows shown in the contrast-enhanced CT images of pelvic RMS. Moreover, it has been shown that RMS contains abundant fibrous tissue and that the contrast agent gradually permeates into the tumor center over time, persisting in the tumor fibrous tissue for a long time^{41,42}. This is possibly why RMS tends to have heterogeneous progressive centripetal enhancement. In addition, the rich mucus in RMS tumors could result in the CT feature of lower than normal muscle density, and rapid tumor growth with nutrient requirements exceeding the vessel supply could contribute to the CT feature of tumoral necrosis.

In this study, we selected 6 CT features with diagnostic value using LASSO logistic regression and established a quantitative diagnostic model of pelvic RMS in children through FDA.

The advantage of this model is that it is simple to operate and easy to implement in routine imaging diagnosis. It is actually a simple mathematical formula that is: (RMS group) $G_1 = -14.283 + 6.613x_1 + 5.333x_2 + 5.753x_3 + 12.361x_4 + 8.095x_5 - 0.715x_6$; (Non-RMS group) $G_2 = -2.008 + 3.539x_1 + 1.080x_2 + 1.154x_3 + 2.307x_4 + 1.656x_5 + 1.380x_6$, where x_1, x_2, \dots and x_6 are lower than normal muscle density (1 = yes; 0 = no), multinodular fusion (1 = yes; 0 = no), enhancement at surrounding blood vessels (1 = yes; 0 = no), heterogeneous progressive centripetal enhancement (1 = yes; 0 = no), ring enhancement (1 = yes; 0 = no), and hemorrhage (1 = yes; 0 = no), respectively. After inserting the CT feature values into the model and calculating G_1 and G_2 values. By comparing the G_1 and G_2 values, subjects with unknown

classification were classified according to the following principles: if $G_1 > G_2$, the subjects were classified into the RMS group; if $G_1 < G_2$, the subjects were classified into the non-RMS group. Therefore clinicians or radiologists can insert this mathematical formula into an excel document to use it, or make a small software to use it. The ROC curve suggested that the model had a high diagnostic value. In addition, cross-validation showed that the model had a high diagnostic value. Our study also found that the AUC of the overall FDA model was higher than that of each CT characteristic feature. Cumulative FDA was carried out based on the importance of the features. There was no significant improvement in the discrimination performance of the FDA model when using the first 4 features, the first 5 features and all 6 features (in the order of importance: heterogeneous progressive centripetal enhancement, enhancement at surrounding blood vessels, multinodular fusion, lower than normal muscle density, hemorrhage and ring enhancement). Therefore, the number of diagnostic CT features can be reduced to 4, and RMS can still be accurately diagnosed, which is beneficial for saving resources. The establishment of a diagnostic model allows us to go from image diagnosis based on human experiences to quantitative imaging diagnosis, which makes the diagnosis of pelvic RMS simpler and more accurate. In the future, we aim to use the CT diagnostic model of RMS to develop artificial intelligence diagnostic software for clinical practice.

There were some limitations in this study. First, this was a retrospective study, which may have inherent selection bias. Second, the sample size was small. RMS is not a common disease in children; therefore, in future research, we need to further expand the sample size to improve the accuracy of our model. Finally, our study summarized only CT features, and the diagnostic value of MRI for RMS needs to be further studied.

Conclusions

Our study showed that pelvic RMS in children has some specific CT features. Furthermore, LASSO logistic regression is a reliable method for selecting diagnostic features of RMS. The FDA model based on CT morphological features can accurately diagnose pelvic RMS in children, with promising cross-validation performance. This diagnostic model could provide an easy and efficient method for the diagnosis and differential diagnosis of pelvic RMS in children.

Data availability

All data generated or analysed during this study are included in this article and its supplementary information files.

Received: 24 April 2022; Accepted: 8 September 2022

Published online: 17 September 2022

References

- Allen, S. D., Moskovic, E. C., Fisher, C. & Thomas, J. M. Adult rhabdomyosarcoma: Cross-sectional imaging findings including histopathologic correlation. *AJR Am. J. Roentgenol.* **189**, 371–377. <https://doi.org/10.2214/ajr.07.2065> (2007).
- Gartrell, J. & Pappo, A. Recent advances in understanding and managing pediatric rhabdomyosarcoma. *F1000Res* **9**, 1685. <https://doi.org/10.12688/f1000research.22451.1> (2020).
- Soffer, S. *et al.* Rhabdomyosarcoma disease spread evaluation on CT scans: association with primary tumor size and Ki-67 proliferation marker. *Clin. Imaging* **56**, 41–46. <https://doi.org/10.1016/j.clinimag.2019.02.009> (2019).
- Jian, B. *et al.* Infectious lung masses misdiagnosed as metastatic rhabdomyosarcoma: Two case reports. *J. Med. Imaging Radiat. Oncol.* **65**, 731–734. <https://doi.org/10.1111/1754-9485.13138> (2021).
- Jokoji, R., Ikeda, J., Tsujimoto, M. & Morii, E. Epithelioid rhabdomyosarcoma; a case report with immunohistochemical and molecular study. *Diagn. Pathol.* **10**, 124. <https://doi.org/10.1186/s13000-015-0349-2> (2015).
- Burnette, J. O. *et al.* Staging paratesticular rhabdomyosarcoma in the “as low as reasonably achievable” age: The case for PET-CT. *Urology* **82**, 220–223. <https://doi.org/10.1016/j.urology.2012.11.051> (2013).
- Sparreboom, B., Litton, B. & Yaxley, J. A rare case of adult rhabdomyosarcoma. *Pol. J. Radiol.* **82**, 395–397. <https://doi.org/10.12659/pjr.901967> (2017).
- Luporsi, M., Cassou-Mounat, T., Amiot, H. M., Laurence, V. & Jehanno, N. Rhabdomyosarcoma revealed by a breast metastasis. *Clin. Nucl. Med.* **43**, e98–e100. <https://doi.org/10.1097/rlu.0000000000001971> (2018).
- Natarajan, A. *et al.* An infrequent case of adult alveolar rhabdomyosarcoma with pancreatic metastases detected in F-18 FDG PET/CT, Indian. *J. Nucl. Med.* **32**, 227–229. https://doi.org/10.4103/ijnm.IJNM_28_17 (2017).
- Chen, X. *et al.* CT-based radiomics to differentiate pelvic rhabdomyosarcoma from yolk sac tumors in children. *Front. Oncol.* **10**, 584272. <https://doi.org/10.3389/fonc.2020.584272> (2020).
- Lee, J. H. *et al.* Rhabdomyosarcoma of the head and neck in adults: MR and CT findings. *AJNR Am. J. Neuroradiol.* **17**, 1923–1928 (1996).
- Tateishi, U. *et al.* Comparative study of FDG PET/CT and conventional imaging in the staging of rhabdomyosarcoma. *Ann. Nucl. Med.* **23**, 155–161. <https://doi.org/10.1007/s12149-008-0219-z> (2009).
- Xie, J. *et al.* A predictive nomogram for predicting improved clinical outcome probability in patients with COVID-19 in Zhejiang Province, China. *Engineering (Beijing)* <https://doi.org/10.1016/j.eng.2020.05.014> (2020).
- Huang, Y. Q. *et al.* Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J. Clin. Oncol.* **34**, 2157–2164. <https://doi.org/10.1200/jco.2015.65.9128> (2016).
- Bian, W. & Tao, D. Asymptotic generalization bound of fisher’s linear discriminant analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 2325–2337. <https://doi.org/10.1109/tpami.2014.2327983> (2014).
- Kim, S. W. & Oommen, B. J. On using prototype reduction schemes to optimize kernel-based fisher discriminant analysis, *IEEE Trans. Syst. Man Cybern. B Cybern.* **38**, 564–570. <https://doi.org/10.1109/tsmcb.2007.914446> (2008).
- Shadvar, A. & Erfanian, A. Mutual information-based fisher discriminant analysis for feature extraction and recognition with applications to medical diagnosis. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2010**, 5811–5814. <https://doi.org/10.1109/iembs.2010.5627461> (2010).
- Ni, M. *et al.* Prediction of the clinicopathological subtypes of breast cancer using a fisher discriminant analysis model based on radiomic features of diffusion-weighted MRI. *BMC Cancer* **20**, 1073. <https://doi.org/10.1186/s12885-020-07557-y> (2020).
- Zou, M. *et al.* Fisher discriminant analysis for classification of autism spectrum disorders based on folate-related metabolism markers. *J. Nutr. Biochem.* **64**, 25–31. <https://doi.org/10.1016/j.jnutbio.2018.09.023> (2019).

20. Hao, S. *et al.* Fisher linear discriminant analysis for classification and prediction of genomic susceptibility to stomach and colorectal cancers based on six STR loci in a Northern Chinese Han population. *PeerJ* **7**, e7004. <https://doi.org/10.7717/peerj.7004> (2019).
21. Tian, L. *et al.* Computed tomography (CT) features of pelvic rhabdomyosarcoma (RMS) in children. *Curr. Med. Imaging* **18**, 299–304. <https://doi.org/10.2174/1573405617666210826160710> (2022).
22. Sun, F. *et al.* Computed tomography and magnetic resonance imaging appearances of malignant vaginal tumors in children: Endodermal sinus tumor and rhabdomyosarcoma. *J. Comput. Assist. Tomogr.* **44**, 193–196. <https://doi.org/10.1097/RCT.0000000000000954> (2020).
23. Gopinathan, A. & Tan, T. Y. Kimura's disease: Imaging patterns on computed tomography. *Clin. Radiol.* **64**, 994–999. <https://doi.org/10.1016/j.crad.2009.07.003> (2009).
24. Obuchowski, N. A. & Bullen, J. A. Receiver operating characteristic (ROC) curves: Review of methods with applications in diagnostic medicine. *Phys. Med. Biol.* **63**, 1. <https://doi.org/10.1088/1361-6560/aab4b1> (2018).
25. Jawad, N. & McHugh, K. The clinical and radiologic features of paediatric rhabdomyosarcoma. *Pediatr. Radiol.* **49**, 1516–1523. <https://doi.org/10.1007/s00247-019-04386-5> (2019).
26. Kralik, S. F. *et al.* Orbital infantile hemangioma and rhabdomyosarcoma in children: Differentiation using diffusion-weighted magnetic resonance imaging. *J. AAPOS* **22**, 27–31. <https://doi.org/10.1016/j.jaaapos.2017.09.002> (2018).
27. Kim, J. R. *et al.* Rhabdomyosarcoma in children and adolescents: Patterns and risk factors of distant metastasis. *AJR Am. J. Roentgenol.* **209**, 409–416. <https://doi.org/10.2214/ajr.16.17466> (2017).
28. Seitz, G. *et al.* Imaging of cell trafficking and metastases of paediatric rhabdomyosarcoma. *Cell Prolif.* **41**, 365–374. <https://doi.org/10.1111/j.1365-2184.2008.00520.x> (2008).
29. Kim, E. E., Valenzuela, R. F., Kumar, A. J., Raney, R. B. & Eftekari, F. Imaging and clinical spectrum of rhabdomyosarcoma in children. *Clin. Imaging* **24**, 257–262. [https://doi.org/10.1016/s0899-7071\(00\)00222-9](https://doi.org/10.1016/s0899-7071(00)00222-9) (2000).
30. Tannous, W. N., Azouz, E. M., Homay, Y. L., Kiruluta, H. G. & Grattan-Smith, D. CT and ultrasound imaging of pelvic rhabdomyosarcoma in children. A review of 56 patients. *Pediatr. Radiol.* **19**, 530–534. <https://doi.org/10.1007/bf02389565> (1989).
31. Chung, C. J. *et al.* Intraperitoneal rhabdomyosarcoma in children: Incidence and imaging characteristics on CT. *AJR Am. J. Roentgenol.* **170**, 1385–1387. <https://doi.org/10.2214/ajr.170.5.9574621> (1998).
32. Stevens, T. G. Rhabdomyosarcoma (grape-like sarcoma) of the cervix. *Specimen Proc. R. Soc. Med.* **25**, 570. <https://doi.org/10.1177/003591573202500494> (1932).
33. Pop, D. L. *et al.* Alveolar rhabdomyosarcoma in an adolescent male patient—Case report and current perspectives. *Rom. J. Morphol. Embryol.* **59**, 1247–1252 (2018).
34. Gilles, R. *et al.* Head and neck rhabdomyosarcomas in children: Value of clinical and CT findings in the detection of loco-regional relapses. *Clin. Radiol.* **49**, 412–415. [https://doi.org/10.1016/s0009-9260\(05\)81828-8](https://doi.org/10.1016/s0009-9260(05)81828-8) (1994).
35. Ramadan, F., Fahs, A., Ghayad, S. E. & Saab, R. Signaling pathways in rhabdomyosarcoma invasion and metastasis. *Cancer Metastasis Rev.* **39**, 287–301. <https://doi.org/10.1007/s10555-020-09860-3> (2020).
36. Bailey, K. A. & Wexler, L. H. Pediatric rhabdomyosarcoma with bone marrow metastasis. *Pediatr. Blood Cancer* **67**, e28219. <https://doi.org/10.1002/pbc.28219> (2020).
37. Deroy, Á. N., Batista, S. M., Cabrera, H. Á. A., Hernández, S. G. & Sánchez, J. M. R. Nasal embryonal rhabdomyosarcoma with bone marrow metastasis simulating acute leukemia: A case report and review of the literature. *Rev. Esp. Patol.* **52**, 87–91. <https://doi.org/10.1016/j.patol.2018.08.001> (2019).
38. Nishida, Y. *et al.* High incidence of regional and in-transit lymph node metastasis in patients with alveolar rhabdomyosarcoma. *Int. J. Clin. Oncol.* **19**, 536–543. <https://doi.org/10.1007/s10147-013-0571-4> (2014).
39. Kishimoto, T. E. *et al.* Identification of tumor-initiating cells derived from two canine rhabdomyosarcoma cell lines. *J. Vet. Med. Sci.* **79**, 1155–1162. <https://doi.org/10.1292/jvms.16-0412> (2017).
40. Basu, S., Baghel, N. S., Puri, A., Shet, T. & Merchant, N. H. 18 F-FDG avid lesion due to coexistent fibrous dysplasia in a child of embryonal rhabdomyosarcoma: source of false positive FDG-PET. *J. Cancer Res. Ther.* **6**, 92–94. <https://doi.org/10.4103/0973-1482.63564> (2010).
41. Ritchie, E. L., Gonzalez-Crussi, F. & Zaontz, M. R. Fibrous hamartoma of infancy masquerading as a rhabdomyosarcoma of the spermatic cord. *J. Urol.* **140**, 800–801. [https://doi.org/10.1016/s0022-5347\(17\)41817-9](https://doi.org/10.1016/s0022-5347(17)41817-9) (1988).
42. Iannella, G., De Vincentiis, M., Corsi, A., Greco, A. & Magliulo, G. A rare case of embryonal rhabdomyosarcoma of the parapharyngeal space. *Acta Otorhinolaryngol. Ital.* **37**, 346–349. <https://doi.org/10.14639/0392-100x-1450> (2017).

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by L.T., L.W. and H.Z. The first draft of the manuscript was written by L.T. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.Q. or J.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022