MRI-Based Radiomics Nomogram for Preoperative Differentiation Between Ocular Adnexal Lymphoma and Idiopathic Orbital Inflammation

Lijuan Yang, MM,[1] Huachen Zhang, MM,[2] Xiaoyang Xie, MM,[2] Shijie Jiang, MM,[1] Hui Zhang, MM,[1] Xin Cao, PhD,[2] Yuqing Hou, MM,[2] Xiaowei He, PhD,[2] Junming Wang, MM,[1] Tao Zhang, PhD,[2*] and Fengjun Zhao, PhD[2*]

Background: Ocular adnexal lymphoma (OAL) and idiopathic orbital inflammation (IOI) are malignant and benign lesions for which radiotherapy and corticosteroids are indicated, but similar clinical manifestations make their differentiation difficult.
Purpose: To develop and validate an MRI-based radiomics nomogram for individual diagnosis of OAL vs. IOI.
Study Type: Retrospective.
Population: A total of 103 patients (46.6% female) with mean age of 56.4 ± 16.3 years having OAL ($n = 58$) or IOI ($n = 45$) were divided into an independent training ($n = 82$) and a testing dataset ($n = 21$).
Field Strength/Sequence: A 3-T, precontrast T1-weighted imaging (T1WI), T2-weighted imaging (T2WI), and postcontrast T1WI (T1 + C).
Assessment: Radiomics features were extracted and selected from segmented tumors and peritumoral regions in MRI before-and-after filtering. These features, alone or combined with clinical characteristics, were used to construct a radiomics or joint signature to differentiate OAL from IOI, respectively. A joint nomogram was built to show the impact of the radiomics signature and clinical characteristics on individual risk of developing OAL.
Statistical Tests: Area under the curve (AUC) and accuracy (ACC) were used for performance evaluation. Mann–Whitney *U* and Chi-square tests were used to analyze continuous and categorical variables. Decision curve analysis, kappa statistics, DeLong and Hosmer–Lemeshow tests were also conducted. $P < 0.05$ was considered statistically significant.
Results: The joint signature achieved an AUC of 0.833 (95% confidence interval [CI]: 0.806–0.870), slightly better than the radiomics signature with an AUC of 0.806 (95% CI: 0.767–0.838) ($P = 0.778$). The joint and radiomics signatures were comparable to experienced radiologists referencing to clinical characteristics (ACC = 0.810 vs. 0.796–0.806, $P > 0.05$) or not (AUC = 0.806 vs. 0.753–0.791, $P > 0.05$), respectively. The joint nomogram gained more net benefits than the radiomics nomogram, despite both showing good calibration and discriminatory efficiency ($P > 0.05$).
Data Conclusion: The developed radiomics-based analysis might help to improve the diagnostic performance and reveal the association between radiomics features and individual risk of developing OAL.
Evidence Level: 3
Technical Efficacy: 3

J. MAGN. RESON. IMAGING 2022.

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.28404

Received Mar 11, 2022, Accepted for publication Aug 8, 2022.

*Address reprint requests to: T.Z. and F.Z., No. 1 Xuefu Avenue, Xi'an, Shaanxi 710127, China. E-mail: zhangtao129@nwu.edu.cn, or fjzhao@nwu.edu.cn.
Lijuan Yang and Huachen Zhang contributed equally to this work.

From the [1]Department of Radiology, Xi'an People's Hospital (Xi'an Fourth Hospital), Xi'an, Shaanxi, China; and [2]Xi'an Key Lab of Radiomics and Intelligent Perception, School of Information Science and Technology, Northwest University, Xi'an, Shaanxi, China

Additional supporting information may be found in the online version of this article

Orbital lymphoproliferative disorders (OLPDs) are the most common primary orbital tumors, representing a wide spectrum of lesions ranging from benign to malignant entities.[1] Ocular adnexal lymphoma (OAL) is a malignant OLPD that often involves the conjunctiva, lacrimal gland, eyelid, or orbit,[2–4] with mucosa-associated lymphoid tissue (MALT) lymphoma as the most common subtype.[5,6] OAL generally manifests as a low-grade B-cell non-Hodgkin lymphoma that occurs mainly in adults.[7,8] It accounts for ~34% of orbital malignancies.[9] Idiopathic orbital inflammation (IOI) belongs to a benign OLPD, representing a non-granulomatous inflammatory process in the orbit,[9,10] mainly involving the extraocular muscles and the lacrimal gland.[11,12] It is crucial to distinguish between different OLPDs because they relate to different treatment options: OAL is amenable to low-dose radiation therapy,[13,14] whereas IOI usually exhibits a positive response to oral corticosteroids.[15–17]

MRI plays an indispensable role for noninvasive characterization of different OLPDs and for assessing their severity.[18–20] To date, MRI including T1-weighted imaging (T1WI), T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced MRI (DCE-MRI) has been used for orbital lesion diagnosis.[21–25] Specifically, OAL was more likely to be homogenous and isointense on T1WI,[21,22] isointense or hyperintense on T2WI,[18,22–24] and mostly located unilateral[21,22] with molding around normal structures without deforming them.[22] OAL showed homogeneous contrast enhancement,[22] high DWI signal, low apparent diffusion coefficient (ADC) values,[1,22,25] and a low DCE-MRI-derived area under the curve (AUC).[25] In contrast, IOI was isointense or hypointense on T2WI,[19,22–24] smaller in size,[20] round or oval in shape[20] and associated with hyperostosis.[20] Also, it has a "flow void sign" in T2WI,[1,25] intermediate DWI and ADC values.[22] However, due to similar appearance of OAL and IOI on MRI, including the laterality, shape, location, and signal intensity between these two diseases,[25] some findings were not consistent and even contradictory. For example, Haradome et al and Valvassori et al reported that the well-defined margin and infiltration (or thickening of ocular muscles) favored the diagnosis of IOI lesions,[1,5] but Khan et al suggested the well-defined or infiltrative appearance supported OAL lesions.[22] In addition, qualitative assessment of MRI features was a subjective process with limited reproducibility between raters.[26,27]

Radiomics is a synergistic approach between machine learning and medical imaging, which performs high-throughput extraction of quantitative features from medical images to assist radiologists in making clinical decisions.[28–30] Unlike qualitative assessments by humans, these quantitative features could provide robust alternatives for automated diagnosis, prognosis, and prediction of tumors.[29,31,32] Several studies investigated the effectiveness of radiomics for diagnosing different orbital lesions. Guo et al extracted first-order gray-level statistics, gray-level run length matrix (GLRLM), and grey-level co-occurrence matrix (GLCM) features from contrast-enhanced T1WI (T1 + C) and T2WI, and then selected five representative features to build a radiomics model to discriminate OAL from IOI.[33] Hou et al developed bag-of-features radiomics encoding texture features (GLRLM, GLCM, laws and statistical features) extracted from T1 + C, with diagnostic performance comparable to a radiologist with 13 years of experience.[34] Duron et al further improved the diagnostic performance using the radiomics features selected from shape, intensity histogram, and texture analyses (GLRLM, GLCM, gray-level size zone matrix [GLSZM], gray-level differential matrix [GLDM] and neighborhood gray-tone difference matrix [NGTDM] features) that were extracted from six MRI sequences including T1WI, T2WI, DWI, and T1 + C.[35]

Despite promising diagnostic performance, these radiomics analyses seldomly considered MRI features in the transformed domain (such as Wavelet and Laplacian of Gaussian [LoG]) and performed the diagnosis without considering clinical information.[33–35] Moreover, they ignored some features used by human experts in the clinical diagnosis, such as tumor boundaries. On the other hand, previous studies only assessed the overall performance through statistical analyses, without assessing individual risk of malignancy.[21,33–35] Recently, radiomics nomograms with radiomics features for risk analysis have been used to assess lymph node metastasis,[36,37] microvascular invasion,[38] survival of patients,[39] and response to induction chemotherapy.[40] However, radiomics nomograms have not yet been investigated for the differential diagnosis of OLPDs.

Therefore, the aim of this study was to develop a radiomics nomogram that integrates the radiomics features extracted from both raw and filtered MRI data as well as clinical information and to assess whether this nomogram allows for an individual preoperative diagnosis of OAL vs. IOI.

Materials and Methods

This retrospective study was approved by the institutional review board and the requirement for written informed consent was waived due to retrospective design.

Population

Between July 2014 and June 2021, 103 consecutive patients (46.6% female) with OLPDs (58 IOIs and 45 OALs) were enrolled in the study, with mean ages of 50.8 ± 16.7 and 63.6 ± 12.7 years, respectively (Table 1). The inclusion criteria were as follows: 1) primary IOI or OAL diagnosis as confirmed by histopathological examination, or primary IOI that was sensitive to oral corticosteroid therapy; 2) MRI examination performed within 7 days prior to biopsy or surgery; 3) complete clinical characteristics including proptosis, eyelid swelling, eye pain, vision loss and eye movement disorder. The exclusion criteria were as follows: 1) patients with a history

TABLE 1. Demographic and Clinical Characteristics of Patients With Different OLPDs

Characteristics	IOI ($n = 58$)	OAL ($n = 45$)	P value
Male	25	30	0.017*
Female	33	15	
Age (mean ± SD)	50.8 ± 16.7	63.6 ± 12.7	0.001*
Proptosis	16	15	0.679
Eyelid swelling	42	22	0.025*
Eye pain	16	4	0.033*
Vision loss	10	6	0.592
Eye movement disorder	4	4	0.911

IOI = idiopathic orbital inflammation; OAL = ocular adnexal lymphoma; SD = standard deviation.
*$P < 0.05$.

of orbital tumor surgery before MR imaging; 2) incomplete MRI acquisition (i.e. missing T1WI, T2WI, or T1 + C); 3) lack of clinical or histopathological data; and 4) poor image quality caused by motion or susceptibility artifacts. Details of the included and excluded patients are shown in Fig. S1.

All patients were divided into a training dataset with 82 patients (46 IOIs and 36 OALs) and testing dataset with 21 patients (12 IOIs and 9 OALs), according to the ratio of 8:2. The analysis workflow is shown in Fig. 1.

MR Imaging

All patients underwent an MRI examination using a 3.0 T scanner (Signa HDxt, GE Healthcare, Milwaukee, WI, USA) with an eight-channel high-resolution head coil. Standard MRI protocols were performed including precontrast axial fast spin-echo (FSE) T1WI with repetition time (TR)/echo time (TE) of 400/10 msec; axial fast spin-echo fat-saturated T2WI with TR/TE of 2800/70 msec; and axial fast spin-echo fat-saturated T1 + C with TR/TE of 500/10 msec. For all sequences, the slice gap and thickness were 3 and 2 mm and the matrix and field of view were 320 × 192 and 180 × 180 mm^2, respectively. The T1 + C sequence was acquired after intravenous injection of 0.2 mL/kg (0.1 mmol/kg) gadolinium diethylenetriamine pentaacetic acid hydrate (DTPA) (Magnevist, Bayer AG, Mullerstrasse, Berlin-Wedding, Germany).

Manual Segmentation

Two radiologists with no prior knowledge of the histopathological diagnosis were involved in the segmentation of three-dimensional (3D) orbital lesions to obtain regions of interest (ROIs). Specifically, the first radiologist with 9 years of experience in head-and-neck radiology loaded the T1 + C sequence into the Medical Imaging Interaction Toolkit (MITK) Workbench (version 2015.5.0; http://www.mitk.org) and manually delineated the orbital lesion slice by slice and obtained the volume of the entire lesion. The segmentation on precontrast T1WI and T2WI followed the same approach and used the segmented ROI on T1 + C sequence as a reference. The second radiologist with 15 years of experience in head-and-neck radiology reviewed and modified the segmentation result if there was over- or under-segmentation. Segmenting the ROI from three MRI sequences took approximately 5–8 minutes per patient. In addition, the peritumoral region was automatically obtained by expanding the boundary of the segmented ROI by three pixels using Python (version 3.6.2; https://www.python.org).

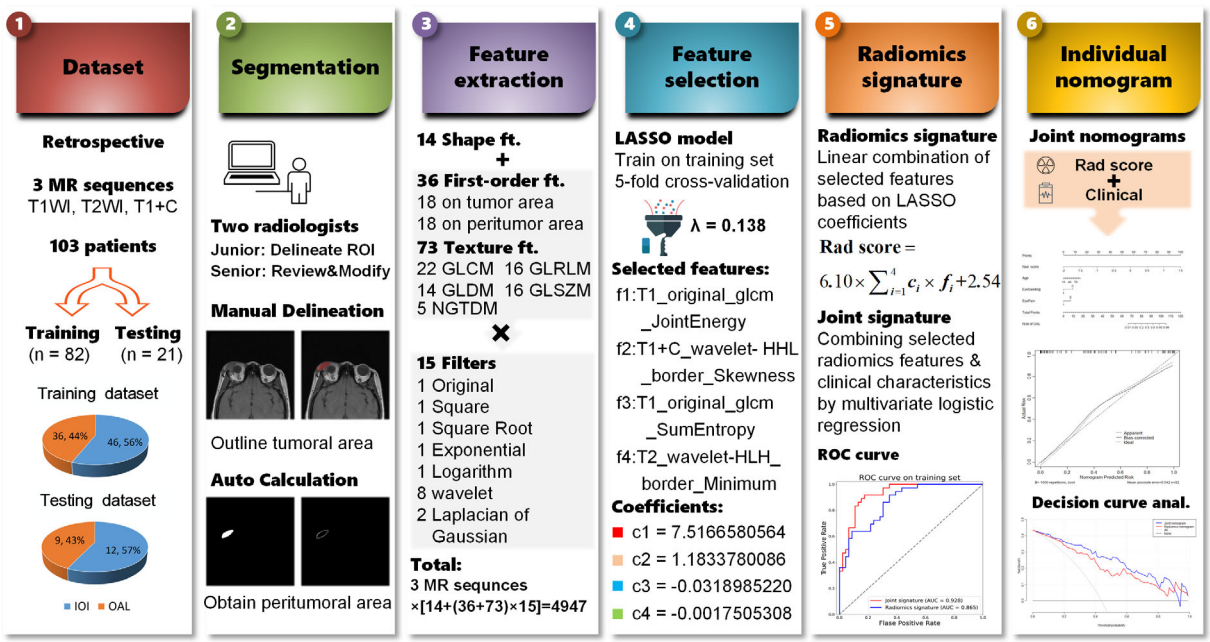

FIGURE 1: Framework of building the radiomics signature and individual nomogram.

Radiomics Feature Extraction

Radiomics features were extracted from MRI before and after filtering using the Pyradiomics software (version 2.2.0; https://pyradiomics.readthedocs.io). First, we extracted 105 features from the segmented ROI of each raw (original) MRI including 14 shape features, 18 first-order histogram features, and 73 texture features (22 GLCM, 16 GLRLM, 16 GLSZM, 14 GLDM, and 5 NGTDM features). Second, from the peritumoral region we extracted 18 first-order features to represent the boundary information. Third, we performed Square, Square Root, Logarithm, Exponential, Wavelet or LoG filtering on the raw MRI to generate the transformed (filtered) MRI from which 109 features were extracted, including first-order features ($n = 18$), texture features ($n = 73$), and peritumoral first-order features ($n = 18$). The Wavelet filter yielded eight decompositions per level, representing all possible combinations of applying either a high- or a low-pass filter in each of the three dimensions, and the LoG filter used two parameters, a low sigma emphasizing fine textures and a high sigma emphasizing coarse textures. Finally, 1649 features were extracted from each MRI, and a total of 4947 features were acquired for the three sequences. The radiomics features from each sequence are detailed in Table S1. It took approximately 2 minutes to extract features from each sequence and ~6 minutes for all three sequences per patient.

Feature Selection and Radiomics Signature Building

To alleviate overfitting caused by small sample sizes and large variable (feature) dimensions, we used the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm to reduce the feature dimensions. Specifically, we determined the tuning parameter λ that minimized the binomial deviances by means of 5-fold cross-validation on the training samples, which could select the optimal number of features that contributed more to OLPD diagnoses. These selected features were weighted by their respective LASSO coefficients to generate a radiomics signature for differentiating OAL from IOI with different rad scores. In addition, a joint signature combining selected radiomics features and clinical characteristics including age, eyelid swelling, and eye pain was also created by fitting another multivariate logistic regression model. The LASSO model was implement using the "glmnet" package in R software (version 3.5.1; https://www.R-project.org).

Individual Risk Nomogram Building

An individual risk model called joint nomogram was built in the training dataset to graphically show the impact of the radiomics signature and different clinical characteristics on the individual risk of malignancy (OAL). Taking a patient as an example (Fig. 6), the value of each risk factor was converted into corresponding points, and the sum of these points can ultimately determine the probability of occurrence of OAL. Calibration curves were plotted to evaluate the diagnostic performance of the joint nomogram in both training and validation datasets using Hosmer–Lemeshow test. Decision curve analysis was performed to quantify the net benefits at different thresholds in order to assess the clinical usefulness of the nomogram, compared to treat-all-patients and treat-none schemes.[36,37]

Imaging Analysis by Radiologists

To compare the radiomics analysis to human diagnosis, five independent radiologists (radiologists A–C, D: H.Z., and E: S.J.) with 2, 5, 7, 10, 13 years of experience separately performed visual assessment of all MRI data. They had no prior knowledge of the histopathological results and clinical findings. Specifically, they rated the lesions' laterality, location, shape, boundary, involvement of the orbital area, signal intensity on T1WI and T2WI sequences, and degree and model of enhancement in T1 + C. Independent assessment was scored in a 5-point scale by each radiologist, with the range from 1 to 5 points representing the likelihood of each lesion going from IOI to OAL. In other words, points 1–5 indicate a lesion identified as IOI, tending to be IOI, unidentifiable, tending to be OAL, and identified as OAL, respectively. The score of each radiologist was normalized to the interval [0, 1] by the formula (score-1)/4 as the probability of assessing the malignancy of each lesion. Furthermore, we also compared the joint signature with the five radiologists who performed the diagnosis as benign or malignant using both imaging and clinical features including age, eyelid swelling, and eye pain.

Statistical Analysis

All statistical analyses were performed using SPSS (version 22.0; IBM, Armonk, NY, USA), R software (version 4.1.2; https://www.R-project.org), and Python (version 3.6.2; https://www.python.org). The AUC from receiver operating characteristic (ROC) analysis was used to evaluate the diagnostic result of each model (radiomics signature and joint signature) and radiologist (A–E). Furthermore, accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), and net reclassification improvement (NRI) were also calculated based on the threshold determined by the maximum Youden index. Mann–Whitney U and Chi-square tests were used to analyze continuous (age, rad score) or categorical variables (gender, clinical characteristics, and diagnostic performances of models and radiologists), respectively. DeLong test was used to compare the ROC curves between the radiomics signature and joint signature. Kappa statistics was used to evaluate the consistencies between models and radiologists and between different radiologists. The calibration curve under Hosmer–Lemeshow test was used to assess the goodness of fit on each nomogram, with the net benefit evaluated by decision curve analysis (DCA). In all statistical analyses, a P value < 0.05 was considered statistically significant.

Results

Demographic and Clinical Characteristics

Demographic and clinical characteristics of patients were summarized in Table 1. Of the 58 IOI patients, 42 and 16 patients suffered from eyelid swelling and eye pain, compared to 22 and 4 of the 45 OAL patients, respectively. OAL more often occurred in elderly and male patients who were less likely to experience eyelid swelling and eye pain. There was no significant difference between OAL and IOI regarding occurrence of proptosis ($P = 0.679$), visual loss ($P = 0.592$), or eye movement disorders ($P = 0.911$).

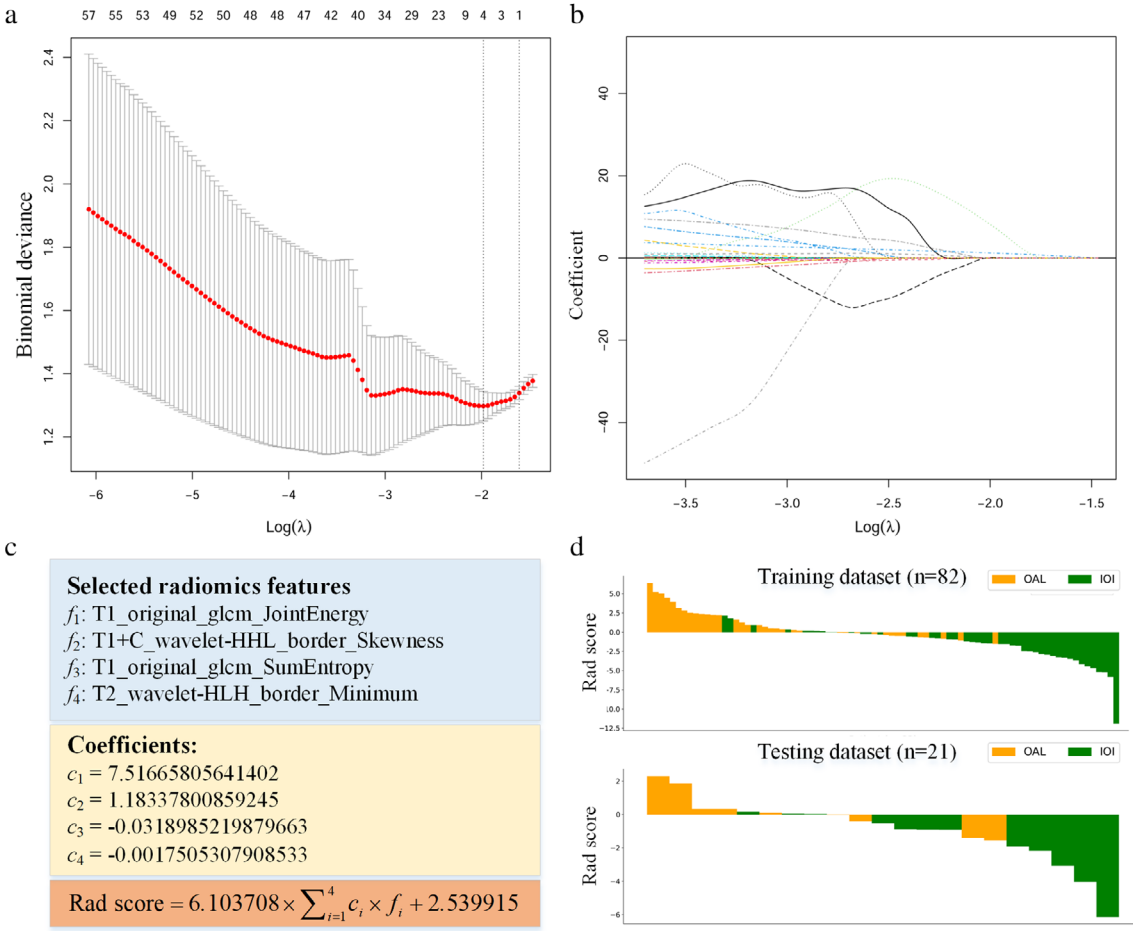

FIGURE 2: Results of feature selection and radiomics signature building. (a) The binomial deviance of the least absolute shrinkage and selection operator (LASSO) with respect to λ. (b) The coefficients of each feature in LASSO with respect to λ. (c) Selected radiomics features, corresponding coefficients, and the building of the radiomics signature quantified by the rad score. (d) The rad score value of each patient in the training and testing datasets.

Feature Selection and Radiomics Signature Building

Binomial deviances and coefficients with different tuning parameters (λ) are shown in Fig. 2a,b, respectively. At the optimal λ value of 0.138 ($\log(\lambda) = -1.981$), four radiomics features with non-zero coefficients were selected from 4947 features, including two from T1WI, one from T2WI, and one from T1 + C (Fig. 2c). Each selected radiomics feature was significantly different between OAL and IOI patients (Fig. 3). The radiomics signature (rad score) was built with these four features and corresponding coefficients. Figure 2d shows the rad score of each patient from both the training and testing datasets. There was a significant difference between the rad scores [median (interquartile range)] for the OAL and IOI patients in the training dataset [0.601 (−0.276 to 2.405) vs. −1.348 (−3.043 to −0.192)], which was confirmed in the testing dataset [0.112 (−0.409 to 0.329) vs. −0.911 (−2.401 to −3.087)].

Performance of Radiomics Signature With/Without Clinical Characteristics

The diagnostic performances of the radiomics signature and joint signature are given in Table 2 and Fig. 4. The radiomics signature achieved AUC values of 0.865 (95% confidence interval [CI]: 0.827–0.890) and 0.806 (95% CI: 0.767–0.838) in the training and testing datasets, respectively. Integrating clinical characteristics, the joint signature performed slightly better than the radiomics signature, with AUC values for the training and testing datasets of 0.928 (95% CI: 0.908–0.948) and 0.833 (95% CI: 0.806–0.870), respectively. However, the ROC curves in Fig. 4 did not show significant differences between the joint and radiomics signatures, with $P = 0.058$ and $P = 0.778$ in the training and testing datasets, respectively. Similar results were observed for the NRI in the testing dataset (0.222, $P = 0.367$), where the NRI values for event (OAL) and non-event (IOI) groups were 0.222 and 0, respectively.

Comparison Between Radiomics Analysis and Human Diagnosis

The comparison between the radiomics model and human diagnoses is given in Table 3. Using only imaging features, the radiomics signature was superior to radiologists A–C with an AUC of 0.806 vs. [0.574, 0.618, 0.626], and comparable

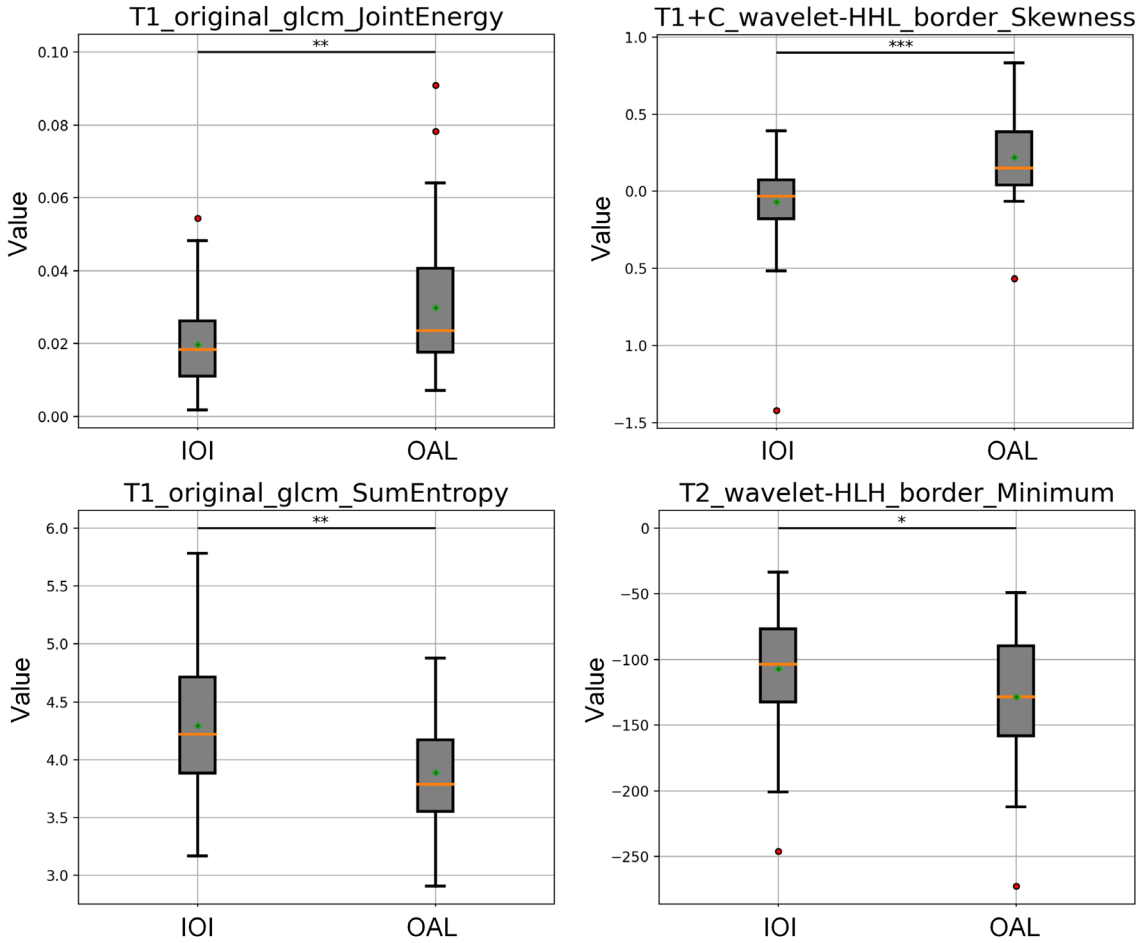

FIGURE 3: Comparison of the four selected radiomics features in patients with idiopathic orbital inflammation (IOI) and ocular adnexal lymphoma (OAL).

TABLE 2. Performance of Different Models in Training and Testing Datasets

	AUC (95% CI)	ACC	SEN	SPE	PPV	NPV
Training dataset						
Radiomics signature	0.865 (0.827–0.890)	0.768	0.917	0.652	0.673	0.909
Joints signature	0.928 (0.908–0.948)	0.866	0.917	0.826	0.805	0.927
Testing dataset						
Radiomics signature	0.806 (0.767–0.838)	0.762	0.778	0.750	0.700	0.818
Joint signature	0.833 (0.806–0.869)	0.810	0.889	0.750	0.727	0.900

AUC = area under curve; CI = confidence interval; ACC = accuracy; SEN = sensitivity; SPE = specificity; PPV = positive predictive value; NPV = negative predictive value.

to radiologists D and E (H.Z. and S.J.) with an AUC of 0.806 vs. [0.751, 791], P = [0.521, 0.870]. In the clinical characteristics taken into account, the joint signature was better than three radiologists A–C with an ACC of 0.810 vs. [0.621, 0.641, 0.709], although not significantly different from other two radiologists (D: H.Z. and E: S.J.) with an ACC of 0.810 vs. [0.796, 0.806], P = [0.271, 0.354].

MR images of four patients with OLPDs are shown in Fig. 5, from which we note that the diagnostic performance of either the radiomics model or human diagnosis or both is improved by combination with clinical characteristics. In addition, the kappa values between the model and radiologists and between different radiologists increased from 0.059–0.516 to 0.206–0.618 (Tables S2 and S3). The detailed

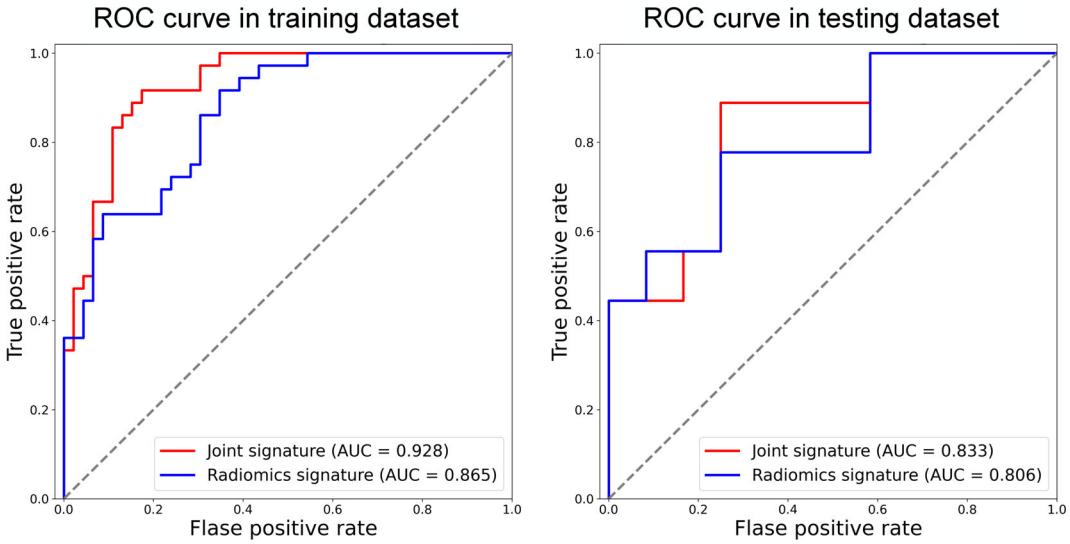

FIGURE 4: Diagnostic ROC curves of two models in the training and testing datasets.

TABLE 3. Comparison Between Radiomics Analysis and Human Diagnosis

	Experience	AUC	ACC	SEN	SPE	PPV	NPV
Diagnosis with only imaging information							
Radiologist A	2 years	0.574	0.544	0.578	0.517	0.481	0.612
Radiologist B	5 years	0.618	0.602	0.600	0.603	0.540	0.660
Radiologist C	7 years	0.626	0.631	0.622	0.638	0.571	0.685
Radiologist D	10 years	0.753	0.728	0.822	0.655	0.649	0.826
Radiologist E	13 years	0.791	0.757	0.600	0.879	0.794	0.739
Radiomics signature	--	0.806	0.762	0.778	0.750	0.700	0.818
Diagnosis with imaging and clinical information							
Radiologist A	2 years	--	0.621	0.644	0.603	0.558	0.686
Radiologist B	5 years	--	0.641	0.667	0.621	0.577	0.706
Radiologist C	7 years	--	0.709	0.733	0.690	0.647	0.769
Radiologist D	10 years	--	0.796	0.889	0.724	0.714	0.894
Radiologist E	13 years	--	0.806	0.689	0.897	0.838	0.788
Joint signature	--	0.833	0.810	0.889	0.750	0.727	0.900

AUC = area under curve; ACC = accuracy; SEN = sensitivity; SPE = specificity; PPV = positive predictive value; NPV = negative predictive value.

comparison between the model and human diagnosis on all patients in the testing dataset was given in Table S4.

Validation of Individual Risk Nomogram

With the rad score and clinical characteristics including age, eyelid swelling, and eye pain as independent risk factors, the joint nomogram built with logistic regression analysis is given in Fig. 6a. The calibration curves indicated good agreement between the assessed and actual risk of malignancy (Fig. 6b; training dataset, $P = 0.230$; testing dataset, $P = 0.358$). In addition, the nomogram with the rad score as a single risk factor (named radiomics nomogram) also passed the Hosmer–Lemeshow test (training dataset, $P = 0.521$; testing dataset, $P = 0.067$), but the calibration did not perform as well as the joint nomogram.

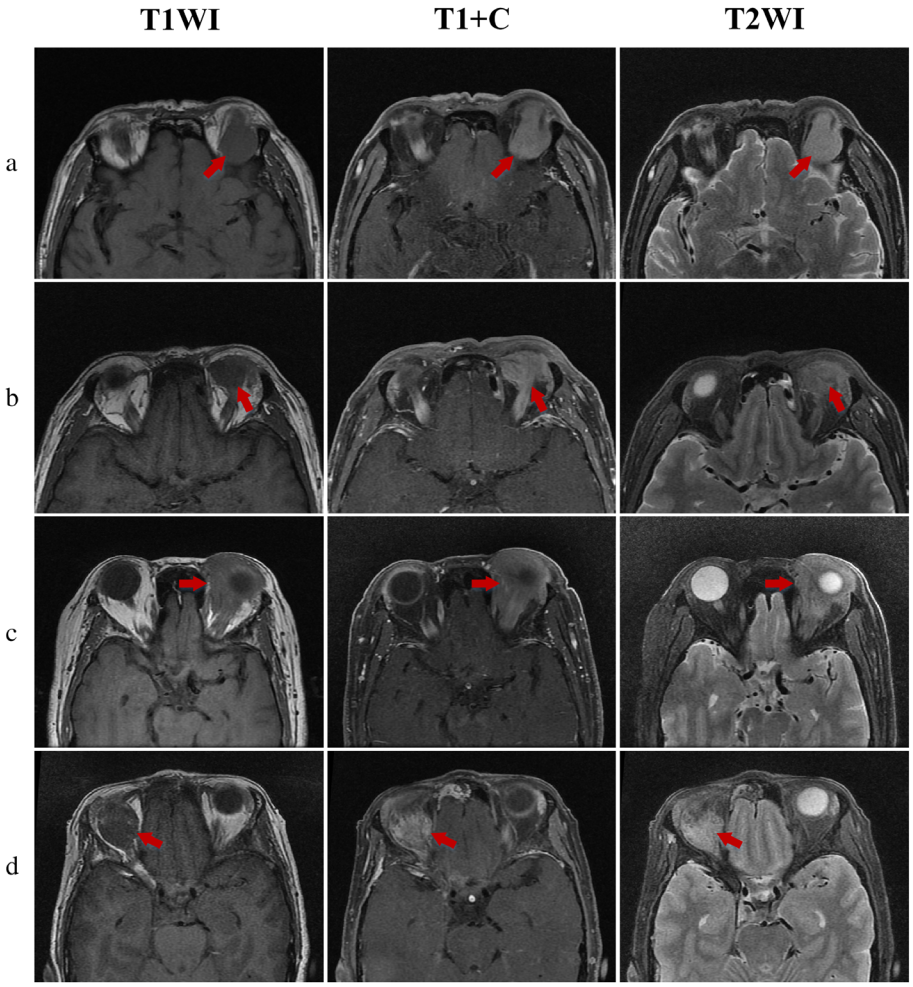

FIGURE 5: MR images of four typical patients. (a) A 72-year-old man with OAL, misclassified as IOI by the radiomics signature and four radiologists (B–E with scores of 2, 1, 1, 1 points, respectively), was correctly identified by the joint signature and all five radiologists after combining clinical characteristics. (b) A 47-year-old man with IOI, misclassified as OAL by radiomics signature and one radiologist (A with the score of 4 points), both of which were corrected by integrating clinical characteristics. (c) A 68-year-old woman with OAL, correctly identified by the radiomics signature and four radiologists (A, B, D, E with scores of 4, 3, 4, 5 points, respectively), and the misclassification of radiologist C (with the score of 2 points) was corrected by combining clinical characteristics. (d) A 60-year-old woman with IOI, correctly identified by four radiologists (B-E with scores of 2, 3, 2 and 3 points, respectively), while the misclassification of radiomics signature was corrected by the joint signature incorporating clinical characteristics.

Discriminatory Efficiency of Different Nomograms

Decision curve analyses of the joint and radiomics nomograms are shown in Fig. 7. On the combined training–testing dataset, both the joint and radiomics nomograms added more benefits for identifying OAL than treat-all-patients and treat-none schemes when the threshold probability of a patient or physician is less than 0.9 or above 0.95. Within this interval, the net benefit of the joint nomogram was superior or comparable to that of the radiomics nomogram. In the interval from 0.9 to 0.95, the radiomics signature performed slightly better than the joint nomogram as well as treat-all-patients and treat-none schemes.

Discussion

We assessed a radiomics-based diagnostic signature and individual nomogram for preoperatively distinguishing between OAL and IOI based on routine MRI data. The joint signature was built with four radiomics features and three clinical characteristics, and it achieved better performance than the radiomics signature that was built with only the radiomics features. The results suggest that this radiomics analysis might help to improve diagnostic performance and reveal the association between radiomics features and individual risk of developing OAL.

Different from previous orbital lesion diagnoses that extracted features only from the lesion area in raw MRI data,[21,33–35] this study extracted features from both the tumoral and peritumoral areas in the raw as well as filtered MRI data. Owing to this process, the testing AUC value achieved by the radiomics signature was superior or similar to the radiomics analyses reported by Guo et al and Hou et al, with AUC values of 0.73 (95% CI: 0.65–0.88) and 0.803

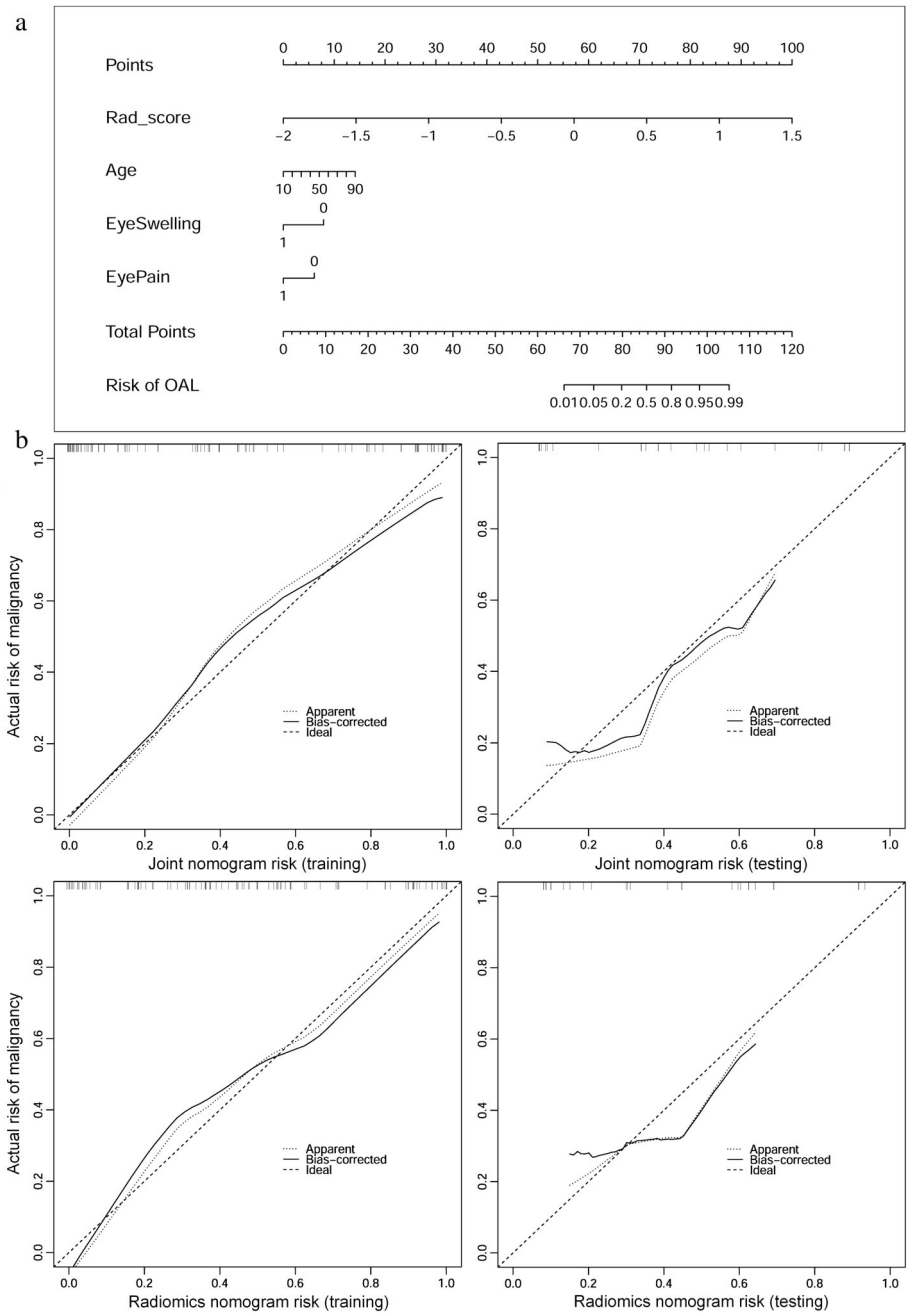

FIGURE 6: (a) Individual risk model (joint nomogram) that can assess the risk of developing ocular adnexal lymphoma (OAL) for each patient with the rad score and clinical characteristics including age, eyelid swelling, and eye pain as independent risk factors. (b) Calibration curves of the joint nomogram and radiomics nomogram in the training and testing phases.

(95% CI: 0.725–0.880), respectively.[33,34] Nevertheless, the performance of the radiomics signature was not as good as the one shown by the study of Duron et al, who obtained an AUC of 0.869 (95% CI: 0.834–0.898).[35] This might be caused by consideration of more sequences including T1WI, T2WI, T1 + C, and DWI. Yet, by incorporating three clinical characteristics, the joint signature in this study improved the diagnostic AUC to 0.833 (95% CI, 0.806–0.870), which was comparable to the previous study of Duron et al,[35] but we only utilized routine sequences excluding DWI.

Of the four selected radiomics features, two were extracted from T1WI (T1_original_glcm_JointEnery, T1_original_glcm_SumEntropy), one from T2WI (T2_wavelet-HLH_bordor_Minimum), and one from T1 + C (T1 + C_wavelet-HHL_border_Skewness), suggesting that multiparametric MRI could provide complementary information. Similar observations were made in previous work to distinguish OAL from benign lesions such as IOI[35] or IgG4-related ophthalmic diseases.[21] Specifically, T1_original_glcm_JointEnery was a measure of homogeneous patterns, while T1_original_glcm_SumEntropy represented a sum of neighborhood intensity-value differences, both belonging to the

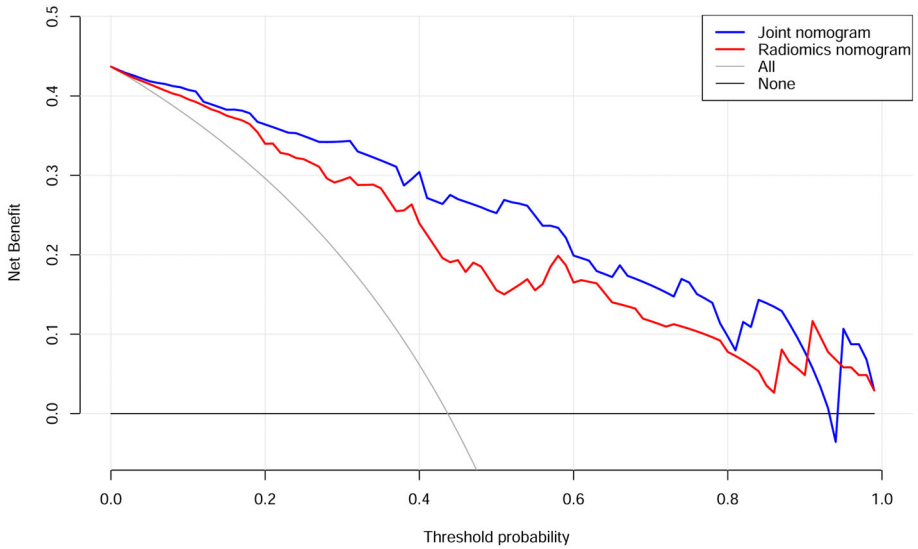

FIGURE 7: Decision curve analysis for the individual risk nomogram.

texture feature (GLCM). OAL had higher joint-energy value and low sum-entropy value compared to IOI, indicating a more homogeneous pattern of OAL in T1WI, which was consistent with the clinical imaging diagnosis reported earlier.[25,26] T1 + C_wavelet-HHL_border_Skewness and T2_wavelet-HLH_bordor_Minimum were the first-order statistical histogram features acquired from the peritumoral areas of T1 + C and T2WI, respectively, which demonstrated the usefulness of boundary information in the Wavelet transformed domain.

It is worth noting that both the radiomics signature and five radiologists can correctly diagnose the patients of OAL with high confidence. For the patients with high confidence of IOI and low confidence of OAL, the radiomics signature still performed well, which was not the case for all radiologists. For the patients with moderate confidence of OAL or IOI, the radiomics signature performed slightly better than or comparable to that of radiologists. Nevertheless, it was these patients who gained more benefit after integrating clinical characteristics (age, eyelid swelling, and eye pain), as demonstrated by improved assessments by both the joint model and radiologists who referenced the clinical information.

Limitations

First, the sample size of this retrospective study was limited because of the lower prevalence of OLPDs compared to lung and breast tumors and the lack of histopathological examination of some patients. The inclusion of large number of samples from multiple centers is warranted to further evaluate the radiomics nomogram in a subsequent study. Second, ROIs were manually delineated by radiologists, which not only increases the tediousness of the analysis process but could also introduce subjective bias. Automatic segmentation techniques based on machine learning and especially deep learning might be used in the future to improve the efficiency and reproducibility of segmented results. Third, although our findings may show the promise of the radiomics analysis with routine MR images (T1WI, T2WI, and T1 + C), previous studies have demonstrated the effectiveness of further sequences including DCE-MRI and DWI in identifying different OLPDs.[25,35] Nevertheless, the quality and quantity of these images are currently not sufficient to identify different OLPDs since only a fraction of patients have undergone these further examinations. The performance of radiomics analyses may be further improved if further sequences such as DCE-MRI and DWI are collected and integrated in the models.

Conclusion

We assessed MRI-based radiomics signatures with and without the clinical information for preoperatively differentiating OAL from IOI. The joint signature performed slightly better than the radiomics signature, and both signatures were comparable to superior compared to evaluations by radiologists. With the rad score and clinical characteristics as independent risk factors, the joint nomogram could gain more net benefits than the radiomics nomogram. Thus, combining the radiomics signature and clinical characteristics including age, eyelid swelling, and eye pain as independent factors to create an easy-to-use nomogram facilitates uncovering the association between radiomics features and individual risk of OAL.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61971350), Shaanxi International Science and Technology Cooperation Program (2021KW-55), Shaanxi Key R&D Plan (2020SF-036), China

Postdoctoral Science Foundation (2019M653717), and Xi'an Science and Technology Plan (201805060ZD11CG44).

References

1. Haradome K, Haradome H, Usui Y, et al. Orbital lymphoproliferative disorders (OLPDs): Value of MR imaging for differentiating orbital lymphoma from benign OPLDs. Am J Neuroradiol 2014;35(10):1976-1982.
2. Freeman C, Berg JW, Cutler SJ. Occurrence and prognosis of extranodal lymphomas. Cancer 1972;29(1):252-260.
3. Shields JA, Shields CL, Scartozzi R. Survey of 1264 patients with orbital tumors and simulating lesions: The 2002 Montgomery Lecture, Part 1. Ophthalmology 2004;111(5):997-1008.
4. Lee S-E, Paik J-S, Cho W-K, et al. Feasibility of the TNM-based staging system of ocular adnexal extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma). Am J Hematol 2011;86(3):262-266.
5. Valvassori GE, Sabnis SS, Mafee RF, Brown MS, Putterman A. Imaging of orbital lymphoproliferative disorders. Radiol Clin North Am 1999; 37(1):135-150.
6. Jenkins C, Rose GE, Bunce C, et al. Histological features of ocular adnexal lymphoma (REAL classification) and their association with patient morbidity and survival. Br J Ophthalmol 2000;84(8):907-913.
7. Watkins LM, Carter KD, Nerad JA. Ocular adnexal lymphoma of the extraocular muscles: Case series from the University of Iowa and review of the literature. Ophthalmic Plast Reconstr Surg 2011;27(6):471-476.
8. Ponzoni M, Govi S, Licata G, et al. A reappraisal of the diagnostic and therapeutic management of uncommon histologies of primary ocular adnexal lymphoma. Oncologist 2013;18(7):876-884.
9. Sniegowski MC, Roberts D, Bakhoum M, et al. Ocular adnexal lymphoma: Validation of American joint committee on cancer seventh edition staging guidelines. Br J Ophthalmol 2014;98(9):1255-1260.
10. Mendenhall WM, Lessner AM. Orbital pseudotumor. Am J Clin Oncol 2010;33(3):304-306.
11. Weber AL, Romo LV, Sabates NR. Pseudotumor of the orbit. Clinical, pathologic, and radiologic evaluation. Radiol Clin North Am 1999; 37(1):151-168.
12. Gordon LK. Orbital inflammatory disease: A diagnostic and therapeutic challenge. Eye 2006;20(10):1196-1206.
13. Akansel G, Hendrix L, Erickson BA, et al. MRI patterns in orbital malignant lymphoma and atypical lymphocytic infiltrates. Eur J Radiol 2005; 53(2):175-181.
14. Raderer M, Kiesewetter B, Ferreri AJ. Clinicopathologic characteristics and treatment of marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma). CA Cancer J Clin 2016;66(2):153-171.
15. Gordon LK. Diagnostic dilemmas in orbital inflammatory disease. Ocul Immunol Inflamm 2003;11(1):3-15.
16. Brannan P. A review of sclerosing idiopathic orbital inflammation. Curr Opin Ophthalmol 2007;5(18):402-404.
17. Glass LD, Freitag SK. Orbital inflammation: Corticosteroids first. Surv Ophthalmol 2016;61(5):670-673.
18. Xian J, Zhang Z, Wang Z, et al. Value of MR imaging in the differentiation of benign and malignant orbital tumors in adults. Eur Radiol 2010; 20(7):1692-1702.
19. Tailor TD, Gupta D, Dalley RW, Keene CD, Anzai Y. Orbital neoplasms in adults: Clinical, radiologic, and pathologic review. Radiographics 2013;33(6):1739-1758.
20. Ben Simon GJ, Annunziata CC, Fink J, Villablanca P, McCann JD, Goldberg RA. Rethinking orbital imaging establishing guidelines for interpreting orbital imaging studies and evaluating their predictive value in patients with orbital tumors. Ophthalmology 2005;112(12): 2196-2207.
21. Yuan Y, Chu G, Gong T, et al. To explore MR imaging radiomics for the differentiation of orbital lymphoma and IgG4-related ophthalmic disease. Biomed Res Int 2021;2021:6668510-6668518.
22. Khan SN, Sepahdari AR. Orbital masses: CT and MRI of common vascular lesions, benign tumors, and malignancies. Saudi J Ophthalmol 2012;26:373-383.
23. Polito E, Galieni P, Leccisotti A. Clinical and radiological presentation of 95 orbital lymphoid tumors. Graefes Arch Clin Exp Ophthalmol 1996;234(8):504-509.
24. Cytryn AS, Putterman AM, Schneck GL, Beckman E, Valvassori G. Predictability of magnetic resonance imaging in differentiation of orbital lymphoma from orbital inflammatory syndrome. Ophthal Plast Reconstr Surg 1997;13(2):129-134.
25. Xu XQ, Hu H, Liu H, et al. Benign and malignant orbital lymphoproliferative disorders: Differentiating using multiparametric MRI at 3.0T. J Magn Reson Imaging 2017;45(1):167-176.
26. Li YA, Jian JM, Pickhardt PJ, et al. MRI-based machine learning for differentiating borderline from malignant epithelial ovarian tumors: A multicenter study. J Magn Reson Imaging 2020;52(3):897-904.
27. Mayer R, Simone CB, Turkbey B, Choyke P. Development and testing quantitative metrics from multi-parametric magnetic resonance imaging that predict Gleason score for prostate tumors. Quant Imaging Med Surg 2022;12(3):1859-1870.
28. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: The bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14(12):749-762.
29. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48(4):441-446.
30. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures they are data. Radiology 2016;278(2):563-577.
31. Kumar V, Gu Y, Basu S, et al. Radiomics: The process and the challenges. Magn Reson Imaging 2012;30(9):1234-1248.
32. Visvikis D, Le Rest CC, Jaouen V, Hatt M. Artificial intelligence, machine (deep) learning and radio(geno)mics: Definitions and nuclear medicine imaging applications. Eur J Nucl Med Mol Imaging 2019; 46(13):2630-2637.
33. Guo J, Liu Z, Shen C, et al. MR-based radiomics signature in differentiating ocular adnexal lymphoma from idiopathic orbital inflammation. Eur Radiol 2018;28(9):3872-3881.
34. Hou Y, Xie X, Chen J, et al. Bag-of-features-based radiomics for differentiation of ocular adnexal lymphoma and idiopathic orbital inflammation from contrast-enhanced MRI. Eur Radiol 2021;31(1):24-33.
35. Duron L, Heraud A, Charbonneau F, et al. A magnetic resonance imaging radiomics signature to distinguish benign from malignant orbital lesions. Invest Radiol 2021;56(3):173-180.
36. Huang Y-q, Liang C-h, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016;34(18):2157-2164.
37. Wu S, Zheng J, Li Y, et al. A radiomics nomogram for the preoperative prediction of lymph node metastasis in bladder cancer. Clin Cancer Res 2017;23(22):6904-6911.
38. Ma X, Wei J, Gu D, et al. Preoperative radiomics nomogram for microvascular invasion prediction in hepatocellular carcinoma using contrast-enhanced CT. Eur Radiol 2019;29(7):3595-3605.
39. Yang L, Yang J, Zhou X, et al. Development of a radiomics nomogram based on the 2D and 3D CT features to predict the survival of non-small cell lung cancer patients. Eur Radiol 2019;29(5):2196-2206.
40. Zhao L, Gong J, Xi Y, et al. MRI-based radiomics nomogram may predict the response to induction chemotherapy and survival in locally advanced nasopharyngeal carcinoma. Eur Radiol 2020;30(1):537-546.