International Orthopaedics (2023) 47:2497–2505
https://doi.org/10.1007/s00264-023-05875-x

ORIGINAL PAPER

Risk stratification by nomogram of deep learning radiomics based on multiparametric magnetic resonance imaging in knee meniscus injury

Tao Zhen[1] · Jing Fang[2] · Dacheng Hu[1] · Mei Ruan[1] · Luoyu Wang[1] · Sandra Fan[1] · Qijun Shen[1]

Received: 26 January 2023 / Accepted: 19 June 2023 / Published online: 29 June 2023

Abstract

Purpose To construct and validate a nomogram model that integrated deep learning radiomic features based on multiparametric MRI and clinical features for risk stratification of meniscus injury.

Methods A total of 167 knee MR images were collected from two institutions. All patients were classified into two groups based on the MR diagnostic criteria proposed by Stoller et al. The automatic meniscus segmentation model was constructed through V-net. LASSO regression was performed to extract the optimal features correlated to risk stratification. A nomogram model was constructed by combining the Radscore and clinical features. The performance of the models was evaluated by ROC analysis and calibration curve. Subsequently, the model was simulated by junior doctors in order to test its practical application effect.

Results The Dice similarity coefficients of automatic meniscus segmentation models were all over 0.8. Eight optimal features, identified by LASSO regression, were employed to calculate the Radscore. The combined model showed a better performance in both the training cohort (AUC = 0.90, 95%CI: 0.84–0.95) and the validation cohort (AUC = 0.84, 95%CI: 0.72–0.93). The calibration curve indicated a better accuracy of the combined model than either the Radscore or clinical model alone. The simulation results showed that the diagnostic accuracy of junior doctors increased from 74.9 to 86.2% after using the model.

Conclusion Deep learning V-net demonstrated great performance in automatic meniscus segmentation of the knee joint. It was reliable for stratifying the risk of meniscus injury of the knee by nomogram which integrated the Radscores and clinical features.

Keywords Knee · Meniscus · Deep learning · Nomogram · Magnetic resonance imaging

Introduction

Osteoarthritis and meniscus tears are two crucial disorders that may occur in the knee and cause knee pain [1]. Meniscus tears are a common injury, with an estimated annual incidence as high as 60–70 per 100,000 [2]. Meniscus tears lead to the stress anomaly of the knee joint and compositional and environmental changes of the joint, which can cause joint cartilage injury and articular interlocking. It is one of the major risk factors for the pathogenesis and progression of knee osteoarthritis [3, 4]. Therefore, the early diagnosis of meniscus tears is conducive to treatment strategies and prognosis.

MRI is the primary diagnostic modality for evaluating the injury of the meniscus. However, the identification of meniscus injury imaging features has been mostly based on qualitative or semi-quantitative scoring systems [5]. Thus, as a traditional evaluation method, MRI is unable to assess the risk of meniscus injury. With the development of artificial intelligence, computer-assisted diagnostic systems have been proposed for the detection of abnormalities in the knee joint for early diagnosis and treatment purposes [1]. Deep learning methods of artificial intelligence have also played an initial role in the detection of osteoarthropathy [6–8]. The automatic segmentation function of deep learning has shown

✉ Qijun Shen
shenqijun80@163.com

[1] Department of Radiology, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, No. 261, Huansha Road, Zhejiang 310006, Hangzhou, China

[2] Zhejiang Provincial Hospital of Chinese Medicine, Hangzhou 310006, China

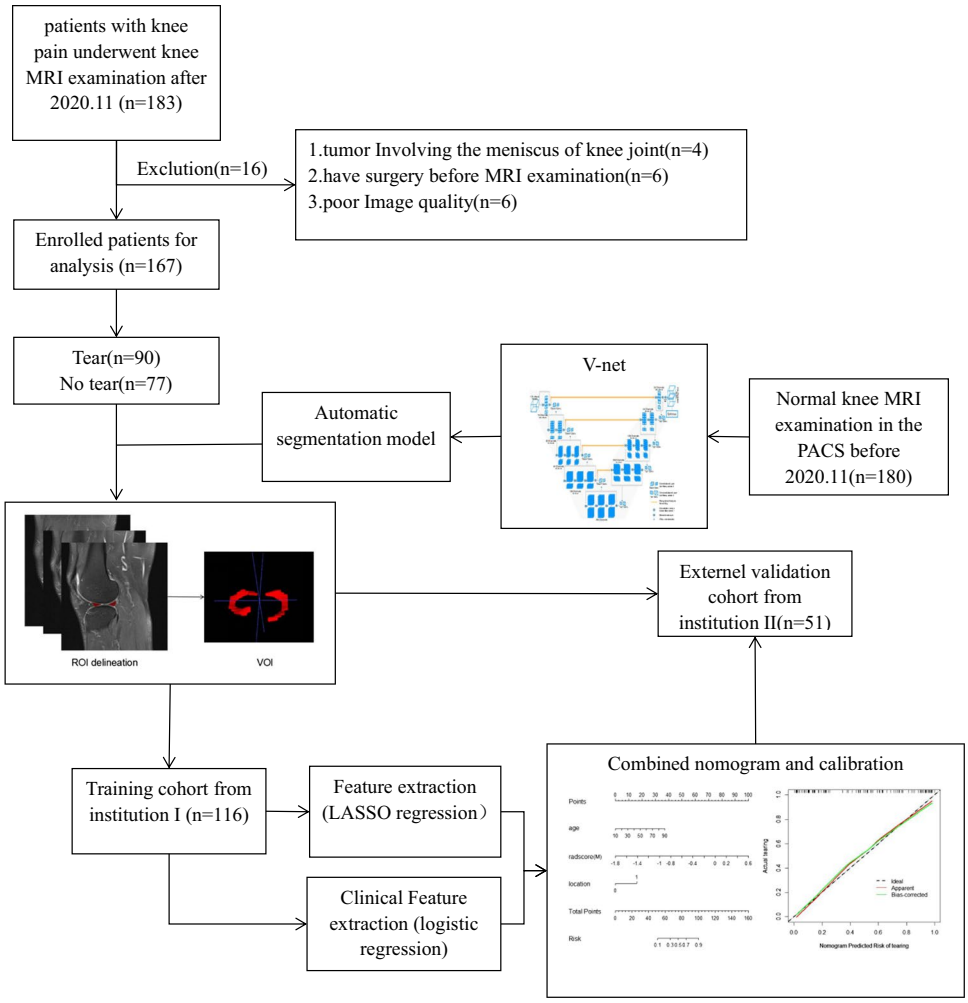

Fig. 1 The workflow of the study. The V-net architecture diagram comes from the study of Milletari et al. [18]

good reliability and practicability in many fields, such as the segmentation of various tumours, as well as the segmentation of some large organs, such as the brain, lung, and bone [9–15]. Most previous studies on meniscus injury are qualitative assessments, and the studies regarding deep learning models were less explanatory and contained no visualization [5, 7, 16]. Radiomics has better interpretation and efficiency than deep learning for the classification of lesions, due to human supervision [17]. However, image annotation of the radiomic analysis was a tedious task. Artificial segmentation was not only time-consuming, but also poorly repeatable, while automatic segmentation based on deep learning could solve this limitation.

In this study, we hypothesized that radiomic analysis could identify associations between the quantitative imaging features and the pathophysiology of knee meniscus tears and effectively and precisely stratify the risk of meniscus injury of the knee in MR images. The aim of the study was to establish an imaging model in order to stratify the risks of meniscus injury of the knee before therapy. We investigated a nomogram model combined with MR radiomic features and clinical features in order to improve the risk stratification of meniscus injury based on fully convolutional neural networks (CNNs) for volumetric medical image segmentation (V-net) for personalized precision therapy and prognosis.

Patients and methods

Our institutional review board approved this retrospective study, and the requirement for informed consent was waived. The workflow of this study was summarized in Fig. 1.

Patients

Clinical data were collected by an orthopaedic surgeon, including the patient's age, gender, location of the limb, and history of direct violent knee trauma in the last one month.

We applied the following inclusion criteria to determine eligibility: (1) patients with knee pain, activity disorder, unstable joint, or had clicking of knee joints, (2) patients who were suspicious for meniscus injury after physical examination, (3)

no surgery before MRI examination, (4) MRI scans were available for qualitative and radiomic analysis.

The exclusion criteria were (1) patients with severe osteoarthritis of the knee joint, rheumatoid arthritis, and other diseases which caused severe knee joint damage, (2) patients with knee joint infection or systemic disease, (3) patients with a meniscal cyst of the knee joint, (4) patients who had poor image quality because of prominent image artifacts or magic angle effect.

Magnetic resonance scanning protocols

MRI examination was performed with different scanners in two institutions (institution I: MAGNETOM Avanto, Siemens AG, 1.5T, Germany; institution II: GE Healthcare, 1.5T, China), fitted with a special magnetic resonance coil for the knee joint. Both applied the same protocol, including the sagittal fat-saturated (FS) proton density-weighted image (PdWI) and sagittal T1-weighted images (T1WI). The scanning parameters of the two institutions were as follows: T1WI (TR 418 ms, TE 11 ms), FS-PdWI (TR 2810 ms, TE 47 ms), layer thickness (3 mm), layer spacing (0.5 mm), and FOV (170 mm × 170 mm). The images were exported in DICOM format.

Evaluation of meniscus injury

The MRI data were independently evaluated by two senior musculoskeletal (MSK) radiologists, with ten and 12 years of diagnostic experience respectively, according to the MRI diagnostic criteria for the knee joint proposed by Stoller et al. [19]. In case of disagreement, the diagnosis was made with another senior MSK radiologist who had 20 years of diagnostic experience. The final diagnosis was determined by the consensus of radiologists. The MR grading system was based on the distribution of meniscal signals in relation to an articular surface (Table 1). Grade 3 of meniscus injury required operation, while grade 0, grade 1, or grade 2 did not. Grade 3 injury was defined as a tear, while the other three were not considered to be a tear.

Image processing, feature extraction, and screening

One radiologist with six years of experience in MSK imaging manually segmented the normal MR T1WI and PdWI images of 180 patients in the Picture Archiving and Communication Systems (PACS) of institution I before November 2020 using the uAI Research Portal. The ROIs delineated along the margin of the meniscus on each slice of the image avoided the image surrounding the joint capsule, cartilage, and ligaments and were extracted for analysis.

The delineation results were reviewed and determined by a MSK radiologist with ten years of experience in MSK imaging diagnosis. The final result was used as the reference ROI. The automatic meniscus segmentation model was constructed through V-net, and the relevant parameters were described in Supplementary I. The Dice similarity coefficient (DSC) was used to evaluate the segmentation efficiency of the model, and the formula was provided in Supplementary II. Subsequently, the model was used to automatically segment the images of 167 patients enrolled after November 2020 from both institutions. The radiomic feature module on the uAI Research Portal was used to extract the features of ROI, and the *Z*-score normalization algorithm was used for standardization. The LASSO regression using 5-fold cross-validation was conducted to choose the optimized subset of features in order to construct models based on radiomic features of T1WI, PdWI, and features combined with T1WI and PdWI. Radscore, defined by corresponding non-zero coefficients of features selected by LASSO, was created by a linear combination of selected features weighted by their coefficients.

Nomogram building, calibration, and external validation

The independent risk factors of the clinical features and the Radscore were integrated to build the combined model through multivariate logistic regression in the training cohort. A nomogram was constructed for risk stratification of meniscus tear by R software. Along with the Hosmer-Lemeshow test measuring for the goodness of fit of the nomogram, the classification accuracy was assessed via calibration curves. ROC analysis was used to evaluate the classification efficacy of each model. Finally, the models were validated by the external validation cohort using the same process.

The simulation application of the model

A junior doctor with less than three years of diagnostic experience in MSK imaging reviewed all the MRIs and was blind to the clinical data and imaging diagnosis. Afterward, the

Table 1 MR grading system of the knee meniscus

Grade 0	There was no abnormality in the shape or signal of the meniscus
Grade 1	One or several punctate signal intensities not contiguous with an articular surface
Grade 2	A linear intrameniscal signal intensity without articular surface extension
Grade 3	Signal intensity extended to at least one articular surface

doctor gave an independent diagnosis of a knee meniscus injury, and one week later, a second diagnosis was made with the assistance of a prediction result from the nomogram. The differences in the capability stratification between the two diagnostic results were compared by Pearson's chi-square test.

Statistical analysis

Statistical analysis was conducted by MedCalc (version 19.1) and R software (version 4.1.2). Binary logistic regression analysis was performed to screen out the clinically independent risk factors. Variables of normal distribution were shown as mean ± SD. Variables of non-normal distribution were shown as median [iqr]. For continuous clinical variables, Student's *t*-tests or Mann-Whitney *U* tests were conducted. For categorical clinical factors, Pearson's chi-square tests or Fisher's exact tests were conducted. $P < 0.05$ was considered statistically significant. The Wilcoxon test was used to compare the evaluation efficiency of the Radscores in the training and validation cohorts regarding the risk degree of meniscus injury. The Hosmer-Lemeshow test was used to analyze the fit degree of the model, and $P > 0.05$ indicated that the model fit was good. The sensitivity, specificity, and AUC of the ROC were used to evaluate the classification efficacy by DeLong's test.

Results

Automatic segmentation model of meniscus

A total of 180 knee MRIs were used to construct the automatic meniscus segmentation model. One hundred forty-four cases from institution I were taken as the training set, and 36 cases from institution II were taken for validation. The DSCs of the automatic segmentation model on PdWI were medial (0.88, 95%CI:0.87–0.89) and lateral (0.86, 95%CI:0.81–0.91) and on T1WI were medial (0.89, 95%CI:0.88–0.89) and lateral (0.88, 95%CI:0.87–0.88).

Patient characteristics

A total of 167 patients admitted to the hospital after November 2020 were recruited in this study, including 62 males and 105 females (ages: 11–90 years, mean age: 53.1±16.4 years). Among them, 116 patients from institution I were used as the training cohort, and 51 patients from institution II were used as the external validation cohort. According to the MRI diagnostic criteria proposed by Stoller et al., all subjects were divided into either the tear group ($n = 90$) or the non-tear group ($n = 77$). Table 2 shows the clinical characteristics and Radscores of these two cohorts. No significant difference was observed in the clinical data for gender and injury history between the training and validation cohorts ($P > 0.05$). However, the age and location showed significant differences in training cohorts ($P < 0.05$). The Radscores had significant differences in both cohorts ($P < 0.001$). The age (OR=1.073, $P< 0.001$) and location (OR=3.432, $P < 0.05$) were screened as independent risk factors by binary logistic regression, and the clinical feature model equation is shown in Supplementary Table S1

Radiomic analysis

In this study, a total of 2600 radiomic features in eight categories were extracted, and a total of twelve T1WI and nine PdWI features were obtained by dimensionality reduction of all features using LASSO regression, respectively (Supplementary III, Table S2, S3). Radscore(T1) and Radscore(Pd) were obtained by summing the correlation coefficients, respectively. After the LASSO regression, eight features were finally selected, which included three T1WI and five PdWI features (supplementary Table S4). Heatmap of feature correlation coefficients is shown in Fig. 2, and clustering analysis was provided

Table 2 Clinical characteristics of training cohort and the validation cohort

Variable		Training cohort (*n*=116)		*P* value	Validation cohort (*n*=51)		*P* value
		Tear (58)	No tear (58)		Tear (32)	No tear (19)	
Age	Mean ± (SD)	60.7±14.3	45±17.1	0.000*	56.1±13.5	49.3±12.3	0.08
Trauma [*n* (%)]	No	49 (84)	46 (79)	0.469	27 (84)	18 (95)	0.509
	Yes	9 (16)	12 (21)		5 (16)	1 (5)	
Gender [*n* (%)]	Female	37 (64)	33 (57)	0.448	22 (69)	13 (68)	0.98
	Male	21 (36)	25 (43)		10 (31)	6 (32)	
Location [*n* (%)]	Right knee	36 (62)	23 (40)	0.016*	19 (59)	8 (42)	0.232
	Left knee	22 (38)	35 (60)		13 (41)	11 (58)	
Radscore(T1)	Median [iqr]	0.56 [−0.25, 1.17]	−0.60 [−0.85, −0.23]	0.000*	0.21 [−0.67, 0.86]	−0.58 [−1.04, −0.1]	0.027*
Radscore(Pd)	Median [iqr]	0.68 [−0.14, 1.16]	−0.58 [−1.19, 0.12]	0.000*	0.45 [−0.07, 0.95]	−0.071 [−1.09,0.31]	0.027*
Radscore(*M*)	Median [iqr]	0.55 [0.14, 1.13]	−0.71 [−1.23,−0.18]	0.000*	0.58 [0.04, −1.19]	−0.73 [−1.08, −0.19]	0.000*

The symbol "*" indicates asignificant statistical difference

in the supplementary (Fig. S1). Radscore (multiple sequence, *M*) =0.032819055*feature1+(−0.0322553255)*feature2++(−0.0421443656)*feature3+0.1094842*feature4+0.002099369*feature5+(−0.00498326868)*feature6+(−0.0263792761) *feature7+(−0.0357271321)*feature8+0.538922131

There were statistically significant differences in the Radscore(*M*) for both the training and validation cohorts of the tear and non-tear groups ($P<0.001$) (Fig. 3). Radscore(*M*), combined with two clinical features (age and location) as the risk factors to evaluate the risk degree of meniscus injury, was used to establish the combined model. The nomogram which demonstrated the visualization of the model is shown in Fig. 4. Meanwhile, the Hosmer-Lemeshow test ($P=0.51$) and the calibration curve showed that the nomogram had a good fit and mean absolute error=0.014 (Fig. S3). ROC curves showed that the AUC of the combined model was higher in the training cohort than that of the clinical model and the T1WI radiomic model ($z=2.136$, 2.228, $P=0.0327$, 0.0259, respectively). In the validation cohort, the AUC of the combined model and radiomic model was higher than that of the clinical model and T2WI radiomic model, ($z=2.353$, 2.722, 2.917, 2.453, $P = 0.0186$, 0.0065, 0.0035, and 0.0142 respectively) (Fig. S4, Table 3). The results of the simulation application of the nomogram are shown in Table 4. After the application of the nomogram, the missed diagnosis of junior doctors was significantly reduced. The diagnostic sensitivity was significantly improved (0.9 vs. 0.6, $P < 0.001$), and the diagnostic accuracy was also significantly improved (0.86 vs. 0.75, $P < 0.01$).

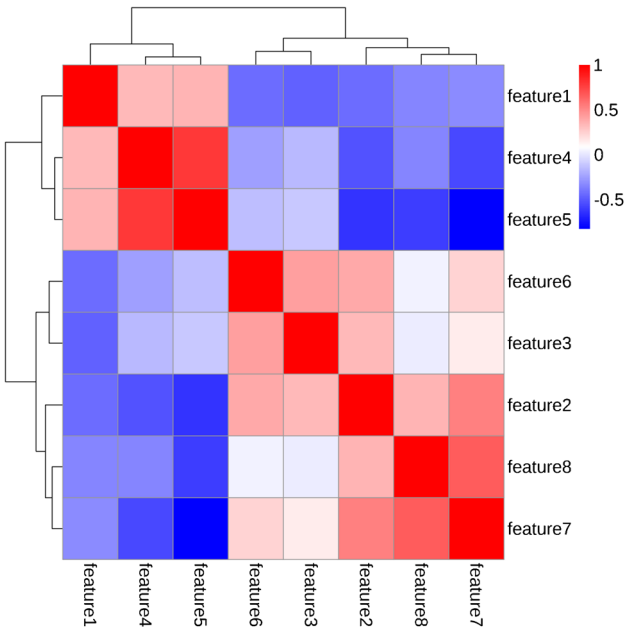

Fig. 2 Heatmap showing the correlation of the top eight tear-predictive radiomic features of the combined model

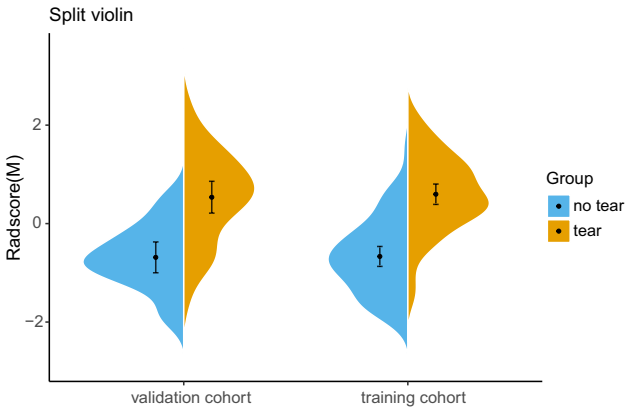

Fig. 3 Wilcoxon's analysis of Radscore(*M*). The Radscore(*M*) of the tear group was significantly higher than that of the non-tear group in both the training group and the validation group

Discussion

In this study, an automatic meniscus segmentation model of knee MRI based on deep learning V-net was constructed, and a combined clinical-radiomic model was established to evaluate the risk of knee meniscus injury. The results showed that the automatic segmentation model had a good segmentation efficiency. Both the combined and radiomic models were superior to the clinical model, and the T1WI and PdWI combined model was superior to the T1WI or

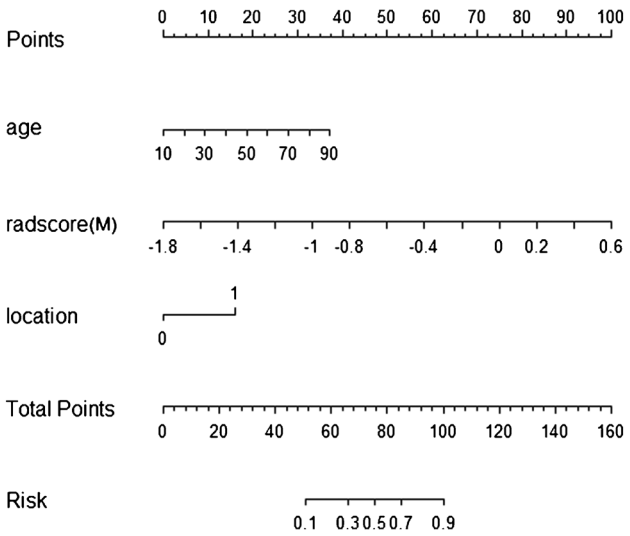

Fig. 4 Nomogram of the combined model in the training cohort (the location of 0 represents the left knee and 1 represents the right knee). For example, a 60-year-old man underwent magnetic resonance imaging because of discomfort in the right knee joint. After calculation by radiomic analysis, the Radscore(*M*) value of the meniscus was −0.8, so his total point was 16 + 22.5 + 42.5(red arrow) = 81, and the corresponding risk of meniscus injury was about 0.6 (blue arrow). For details, see Fig. S2

PdWI model alone. Meanwhile, the nomogram demonstrated the visualization of the model for meniscus injury risk, indicating that the nomogram could be used as a quantitative stratification tool for meniscus injury risk.

At present, the evaluation of MRI for the meniscus of the knee joint was mainly based on qualitative diagnosis, which can only interpret the presence or absence of a meniscus tear but cannot evaluate the risk of a meniscus tear. Roblot et al. developed a deep learning model based on CNNs, consisting of 1123 MRIs of knee joints with an AUC of 0.94. However, the dataset contained only two T2-weighted images per patient, and it had poor explanation [20]. The model in this study was based on a small sample size of 3D image information including the entire medial and lateral meniscus, and it combined the radiomic features of T1WI and PdWI. In addition to basic morphological features, it also included internal information such as texture features and wavelet features. Furthermore, we used the nomogram to realize the visualization of the model.

V-net is a network architecture based on CNNs that have been developed in recent years and have been mainly proposed for three-dimensional medical images. A series of recent studies have demonstrated its superior automatic segmentation performance [21–24], while the required sample size was not as large as expected. In this study, only 180 patients were used to construct the automatic segmentation models. Four automatic segmentation models were constructed in this study, and the DSCs of the models on T1WI and PdWI sequences were greater than 0.8. We tried to merge the medial and lateral meniscus as a combined model, but it was found that only one side of the meniscus underwent random automatic segmentation. On the other hand, the segmentation efficiency of the separate models was ideal. We speculated that, firstly, the medial and lateral menisci were anatomically two independent structures, and putting the two independent ROIs in the same model caused confusion in model identification; secondly, our sample size was too small to fuse the two independent ROIs by V-net. Therefore, we believed that it would be better to construct segmentation models separately for two or more disconnected ROIs based on small samples. We also calculated the time cost of manual and automatic segmentation. Manually segmenting a case took about 15 min, while an automatic segmentation model only required less than ten s for each sequence for one side of the meniscus, taking about 30 s to complete a case. Compared with single-case segmentation, batch segmentation did not require opening every case. Although it only saved about 10 s for each case, the cumulative time for a large number of data was considerably significant.

In our radiomic model, a total of eight features from two sequences were extracted, which included gray level size zone matrix (GLSZM), gray level dependence matrix (GLDM), and gray level cooccurrence matrix (GLCM) mainly focused on texture features, which demonstrated that the inherent heterogeneity of the meniscus was better than utilizing

Table 3 Comparison between the training and validation cohort models

Model	Training cohort				Validation cohort			
	AUC (95%CI)	YIJ	SEN	SPEC	AUC (95%CI)	YIJ	SEN	SPEC
Clinical	0.83 (0.75 to 0.90)	0.60	0.93	0.67	0.69 (0.55 to 0.81)	0.32	0.84	0.47
T1WI	0.82 (0.74 to 0.88)	0.59	0.90	0.69	0.71 (0.57 to 0.83)	0.43	0.53	0.89
PdWI	0.86 (0.79 to 0.92)	0.57	0.72	0.84	0.73 (0.58 to 0.84)	0.37	0.69	0.68
T1 and PdWI	0.87 (0.80 to 0.93)	0.66	0.84	0.81	0.86 (0.73 to 0.94)	0.62	0.78	0.84
Combined	0.90 (0.84 to 0.95)	0.69	0.86	0.82	0.84 (0.72 to 0.93)	0.60	0.81	0.79

AUC area under the ROC curve, *SEN* sensitivity, *SPEC* specificity, *YIJ* Youden's index J

Table 4 Comparison of diagnosis results of junior doctor before and after using the nomogram

Three dimensional Chi-square test		+ (Tear)			− (No tear)		
		Doctor			Doctor		
Doctor using nomogram		+	−	Total	+	−	Total
	+	52	29	81	3	11	14
	−	2	7	9	3	60	63
	Total	54	36	90	6	71	77
Group	SEN	SPEC			ACC		
Doctor	0.6	0.92			0.75		
Doctor using nomogram	0.9	0.82			0.86		
X^2, P	21.81, 0.000	3.5, 0.061			8.02, 0.005		

SEN sensitivity, *SPEC* specificity, *ACC* accuracy

morphological features to identify meniscus damage. Among these, PdWI features were the majority, also indicating that the PdWI sequence had more advantages in the identification of meniscus injury. This finding was also in line with clinical data, as we relied more on the PdWI sequence for the identification of meniscus injury. However, when the features of the two sequences were combined, the diagnostic efficiency was improved, indicating that a single sequence had its own limitations, while the combined advantages were complementary.

In this study, the model was visualized by nomogram, through which the risk of meniscus injury could be quantified. A higher calculated value suggested an increased risk of meniscus injury. For instance, a typical case was seen in a 57-year-old patient initially diagnosed with II meniscus injury; however, the nomogram showed a high risk of injury, and III injury was found in the follow-up review six months later (Fig. 5). If we had intervened in patients at high risk of injury before further development, we may have been able to delay the progression of injury. Therefore, the nomogram made it easier to understand the prognosis and helped make clinical decisions [25].

The incidence of asymptomatic meniscus tears has been reported to increase with age [26]. This was consistent with our study. We also found that the meniscus of the right knee was more prone to tear than the left knee (P <0.01), which was due to the fact that the dominant limb of most patients in this study was the right limb. However, to our surprise, a history of direct violent knee injury within one month was not a risk factor for meniscal tears. This could indicate that most meniscus damages was a chronic process, and an acute injury that directly resulted in a tear of the meniscus was relatively rare unless a blow-out fracture involved the articular surface.

The diagnostic accuracy of our combined model was 88.79% (103/116) for the training cohort and 84.31% (43/51) for the validation cohort, which was close to or even greater than that of the junior doctor. With the help of the nomogram, the diagnostic accuracy of the junior doctor was greatly improved from 0.75 to 0.85, and the rate of missed diagnosis significantly decreased from 0.4 to 0.1. Therefore, the nomogram was expected to help doctors save time for image evaluation and improve diagnostic efficiency.

There were some limitations in the current study that still needs to be further investigated. (1) The automatic segmentation model in this study needs to distinguish the medial and lateral meniscus, which would increase a certain

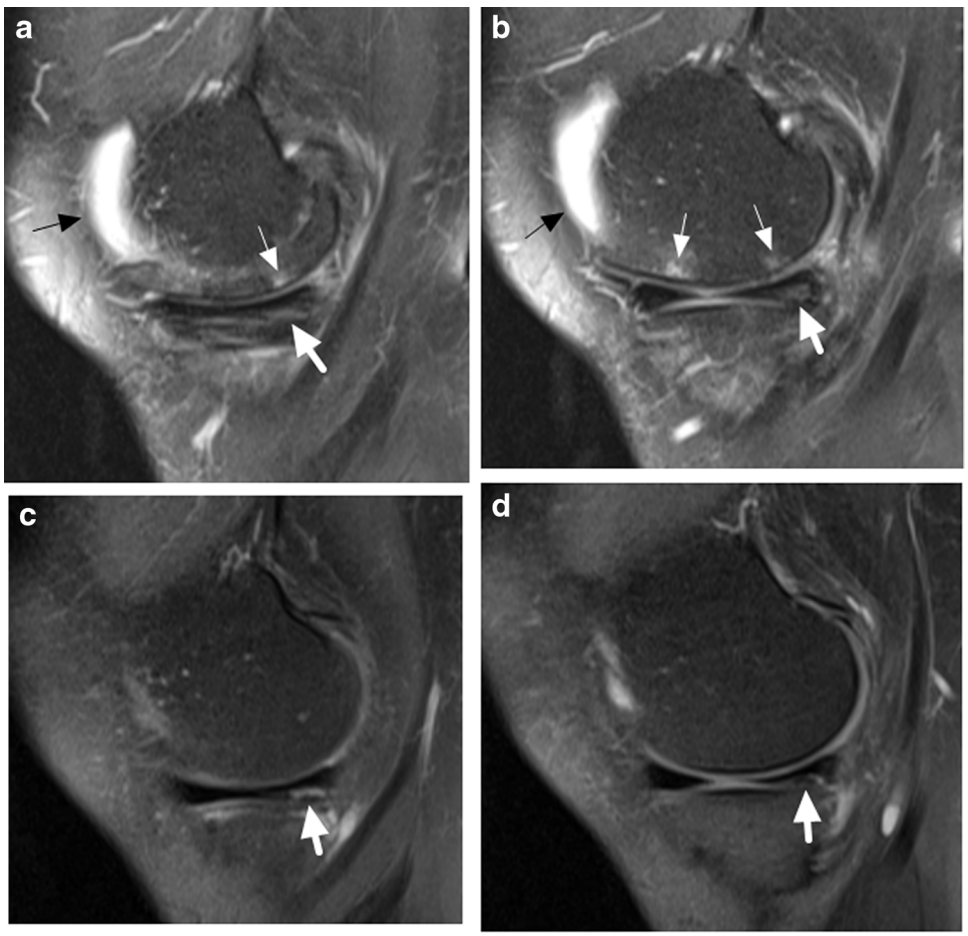

Fig. 5 A 57-year-old female patient underwent an initial MR examination of the knee joint (**a**, **b** two consecutive layers of sagittal PdWI), which showed fluid in the joint cavity (thin black arrow), femoral bone marrow oedema (thin white arrow), and II injury of the posterior angle of the medial meniscus (thick black arrow). Radiomic analysis showed Radscore = 0.634 when combined with the nomogram. The total score was about 120, indicating that the risk of tear was over 90%. The patient was not specially treated, and the knee MR was reexamined after half a year of rest (**c**, **d** two consecutive layers of sagittal PdWI): joint cavity effusion and bone marrow oedema absorption was seen, while the medial meniscus of the original II injury progressed to III injury (white thick arrow)

amount of work for practical application. (2) This study did not include discoid meniscus due to a small sample size and few incidences of cases. (3) This study did not separate the tear type from the tear group.

Conclusions

This study constructed an automatic segmentation model of knee joint meniscus based on V-net. A clinical-radiomic nomogram was established and could potentially be used as a reliable tool to evaluate the risk of knee meniscus injury, suggesting great future possibilities for clinical applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00264-023-05875-x.

Author contribution All authors contributed to either the conception, design, data collection, or analysis. Material preparation and data collection were performed by Tao Zhen, Jing Fang, Mei Ruan, and Dacheng Hu. The first draft of the manuscript was written by Tao Zhen, and all authors commented on the previous versions of the manuscript. Tao Zhen and Luoyu Wang contributed to data analysis. Qijun Shen contributed to the final manuscript and supervised all the data. All authors read and approved the final manuscript.

Funding This study has received funding by the Medical Health Science and Technology Commission of Zhejiang Province, China (No. 2021KY240, 2023KY162, 2023KY953), the Traditional Chinese Medicine Science and Technology Commission of Zhejiang Province, China (No. 2023ZL571), and the Hangzhou Biological Medicine and Health Industry Development Support Science and Technology Project (No. 2021WJCY028).

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability Not applicable.

Declarations

Ethics approval This study was approved by the local institutional review board of the Hangzhou First People's Hospital, Sichuan, China.

Consent to participate Written informed consent was waived by the Institutional Review Board.

Consent for publication We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed.

Conflict of interest The authors declare no competing interests.

References

1. Saygili A, Albayrak S (2020) Knee meniscus segmentation and tear detection from MRI: a review. Curr Med Imaging Rev 16(1):2–15. https://doi.org/10.2174/1573405614666181017122109
2. Fox AJ, Wanivenhaus F, Burge AJ, Warren RF, Rodeo SA (2015) The human meniscus: a review of anatomy, function, injury, and advances in treatment. Clin Anat 28(2):269–287. https://doi.org/10.1002/ca.22456
3. Kawahara T, Sasho T, Katsuragi J, Ohnishi T, Haneishi H (2017) Relationship between knee osteoarthritis and meniscal shape in observation of Japanese patients by using magnetic resonance imaging. J Orthop Surg Res 12(1):97. https://doi.org/10.1186/s13018-017-0595-y
4. Malanga GA, Chirichella PS, Hogaboom NS, Capella T (2021) Clinical evaluation of micro-fragmented adipose tissue as a treatment option for patients with meniscus tears with osteoarthritis: a prospective pilot study. Int Orthop 45(2):473–480. https://doi.org/10.1007/s00264-020-04835-z
5. Bien N, Rajpurkar P, Ball RL, Irvin J, Park A, Jones E, Bereket M, Patel BN, Yeom KW, Shpanskaya K, Halabi S, Zucker E, Fanton G, Amanatullah DF, Beaulieu CF, Riley GM, Stewart RJ, Blankenberg FG, Larson DB et al (2018) Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. Plos Med 15(11):e1002699. https://doi.org/10.1371/journal.pmed.1002699
6. von Schacky CE, Sohn JH, Liu F, Ozhinsky E, Jungmann PM, Nardo L, Posadzy M, Foreman SC, Nevitt MC, Link TM, Pedoia V (2020) Development and validation of a multitask deep learning model for severity grading of hip osteoarthritis features on radiographs. Radiology 295(1):136–145. https://doi.org/10.1148/radiol.2020190925
7. Subhas N, Li H, Yang M, Winalski CS, Polster J, Obuchowski N, Mamoto K, Liu R, Zhang C, Huang P, Gaire SK, Liang D, Shen B, Li X, Ying L (2020) Diagnostic interchangeability of deep convolutional neural networks reconstructed knee MR images: preliminary experience. Quant Imag Med Surg 10(9):1748–1762. https://doi.org/10.21037/qims-20-664
8. Garwood ER, Tai R, Joshi G, Watts VGJ (2020) The use of artificial intelligence in the evaluation of knee pathology. Semin Musculoskel R 24(01):21–29. https://doi.org/10.1055/s-0039-3400264
9. Balakrishnan R, Valdés HM, Farrall AJ (2021) Automatic segmentation of white matter hyperintensities from brain magnetic resonance images in the era of deep learning and big data - a systematic review. Comput Med Imaging Graph 88:101867. https://doi.org/10.1016/j.compmedimag.2021.101867
10. Barra D, Nicoletti G, Defeudis A, Mazzetti S, Panic J, Gatti M, Faletti R, Russo F, Regge D, Giannini V (2021) Deep learning model for automatic prostate segmentation on bicentric T2w images with and without endorectal coil. Annu Int Conf IEEE Eng Med Biol Soc 2021:3370–3373. https://doi.org/10.1109/EMBC46164.2021.9630792
11. Zhu J, Bolsterlee B, Chow B, Cai C, Herbert RD, Song Y, Meijering E (2021) Deep learning methods for automatic segmentation of lower leg muscles and bones from MRI scans of children with and without cerebral palsy. Nmr Biomed 34(12):e4609. https://doi.org/10.1002/nbm.4609
12. Grøvik E, Yi D, Iv M, Tong E, Rubin D, Zaharchuk G (2020) Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI. J Magn Reson Imaging 51(1):175–182. https://doi.org/10.1002/jmri.26766
13. Zhou J, Zhang Y, Chang KT, Lee KE, Wang O, Li J, Lin Y, Pan Z, Chang P, Chow D, Wang M, Su MY (2020) Diagnosis of benign and malignant breast lesions on DCE-MRI by using radiomics and deep learning with consideration of peritumor tissue. J Magn Reson Imaging 51(3):798–809. https://doi.org/10.1002/jmri.26981
14. Kavur AE, Gezer NS, Barış M, Şahin Y, Özkan S, Baydar B, Yüksel U, Kılıkçıer Ç, Olut Ş, Bozdağı AG, Ünal G, Dicle O, Selver MA (2020) Comparison of semi-automatic and deep learning-based automatic methods for liver segmentation in living liver transplant donors. Diagn Interv Radiol 26(1):11–21. https://doi.org/10.5152/dir.2019.19025

15. Kong Z, Li T, Luo J, Xu S (2019) Automatic tissue image segmentation based on image processing and deep learning. J Healthc Eng 2019:2912458. https://doi.org/10.1155/2019/2912458
16. Kim SH, Lee H, Jang Y, Chun K, Park Y (2021) Diagnostic accuracy of magnetic resonance imaging in the detection of type and location of meniscus tears: comparison with arthroscopic findings. J Clin Med 10(4):606. https://doi.org/10.3390/jcm10040606
17. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278(2):563–577. https://doi.org/10.1148/radiol.2015151169
18. Milletari F, Navab N, Ahmadi S (2016) V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). IEEE, pp 565–571
19. Stoller DW, Martin C, Crues JR, Kaplan L, Mink JH (1987) Meniscal tears: pathologic correlation with MR imaging. Radiology 163(3):731–735. https://doi.org/10.1148/radiology.163.3.3575724
20. Roblot V, Giret Y, Bou Antoun M, Morillot C, Chassin X, Cotten A, Zerbib J, Fournier L (2019) Artificial intelligence to diagnose meniscus tears on MRI. Diagn Interv Imag 100(4):243–249. https://doi.org/10.1016/j.diii.2019.02.007
21. Ma J, Deng Y, Ma Z, Mao K, Chen Y (2021) A liver segmentation method based on the fusion of VNet and WGAN. Comput Math Method M 2021:1–12. https://doi.org/10.1155/2021/5536903
22. Kuiper RJA, Sakkers RJB, Stralen M, Arbabi V, Viergever MA, Weinans H, Seevinck PR (2022) Efficient cascaded V-net optimization for lower extremity CT segmentation validated using bone morphology assessment. J Orthop Res 40(12):2894–2907. https://doi.org/10.1002/jor.25314
23. Kanakatte A, Bhatia D, Ghose A (2021) Heart region segmentation using dense VNet from multimodality images. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2021. IEEE, pp 3255–3258. https://doi.org/10.1109/EMBC46164.2021.9630303
24. Hua R, Huo Q, Gao Y, Sui H, Zhang B, Sun Y, Mo Z, Shi F (2020) Segmenting brain tumor using cascaded V-nets in multimodal MR images. Front Comput Neurosci 14:9. https://doi.org/10.3389/fncom.2020.00009
25. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP (2015) Nomograms in oncology: more than meets the eye. Lancet Oncol 16(4):e173–e180. https://doi.org/10.1016/S1470-2045(14)71116-7
26. Nguyen JC, De Smet AA, Graf BK, Rosas HG (2014) MR imaging-based diagnosis and classification of meniscal tears. Radiographics 34(4):981–999. https://doi.org/10.1148/rg.344125202