

# Noncontrast MRI-based machine learning and radiomics signature can predict the severity of primary lower limb lymphedema

Jie Ren, MA,<sup>a</sup> Xingpeng Li, MD,<sup>b</sup> Mengke Liu, MD,<sup>b</sup> Tingting Cui, MA,<sup>a</sup> Jia Guo, MA,<sup>b</sup> Rongjie Zhou, BS,<sup>a</sup> Kun Hao, MD,<sup>c</sup> Rengui Wang, MD,<sup>b</sup> and Yunlong Yue, MD,<sup>a,b</sup> Beijing, China

## ABSTRACT

**Objective:** According to International Lymphology Society guidelines, the severity of lymphedema is determined by the difference in volume between the affected limb and the healthy side divided by the volume of the healthy side. However, this method of measuring volume is time consuming, laborious, and has certain errors in clinical applications. Therefore, this study aims to explore whether machine learning radiomics features based on noncontrast magnetic resonance imaging (MRI) can predict the severity of primary lower limb lymphedema.

**Methods:** A retrospective analysis of 119 patients with primary lower limb lymphedema. The enrolled patients were divided into a nonsevere group (mild and moderate) and a severe group. Using the semiautomatic threshold method in ITK-snap software on the patient's noncontrast MRI, we filled the area between the subcutaneous tissue and muscle of the edematous site. The PyRadiomics software package was used to extract radiomic features. The radiomic features were analyzed using the *t* test or Mann-Whitney test. Subsequently, Pearson correlation testing and least absolute shrinkage and selection operator screening were performed. Using Scikit-learn, the remaining features were used to construct five models: logistic regression, support vector machine, random Forest, ExtraTrees, and light gradient boosting machine. The predictive performance were evaluated by the receiver operating characteristic curve, and the sensitivity and specificity of these measures were calculated. The predictive curve was used to evaluate the performance of the predictive model in guiding decisions for nonsevere and severe lymphedema patients.

**Results:** The enrolled patients including 28 patients with mild lymphedema (grade I), 38 patients with moderate lymphedema (grade II), and 53 patients with severe lymphedema (grade III) was conducted. A total of 1196 features were extracted, and after Pearson correlation testing and least absolute shrinkage and selection operator screening, 21 nonzero features were selected. The ExtraTree model performed the best, with an area under the curve of 0.974 (95% confidence interval, 0.9437-1.0000) in the training set, a sensitivity of 89.2%, and a specificity of 95.7%. In the test set, these values were 0.938 (95% confidence interval, 0.8539-1.0000), 75%, and 100%, respectively. The decision curve showed that when the predicted probability was between 16% and 78%, the net benefit of the ExtraTree model was greater than that of the two extreme curves, indicating strong clinical value in guiding decisions for nonsevere and severe lymphedema patients.

**Conclusions:** All five models performed well in distinguishing between the nonsevere group and the severe group. Noncontrast MRI-based machine learning radiomics signature can predict the severity of primary lower limb lymphedema. (*J Vasc Surg Venous Lymphat Disord* 2025;■:102161.)

**Keywords:** Lower extremity; Lymphedema; Magnetic resonance imaging; Radiomics

Primary lymphedema (PLE) is a type of lymphatic system dysfunction or disorder caused by congenital abnormalities in the structure and/or function of lymph nodes

and/or lymphatic vessels, leading to inadequate lymphatic drainage and accumulation of protein-rich fluid in the subcutaneous tissues of the lower extremities. Prolonged stimulation of adipose tissue results in irreversible fibrosis.<sup>1</sup> According to International Lymphology Society guidelines, the severity of lymphedema is determined by the difference in volume between the affected limb and the healthy side divided by the volume of the healthy side. A difference of 5%-20% is considered mild, 20%-40% is considered moderate, and of >40% is considered severe.<sup>2</sup> Early diagnosis and timely intervention of primary lower extremity lymphedema can effectively slow disease progression, making it crucial to accurately diagnose and assess the severity of the condition.<sup>3</sup>

In practical clinical work, commonly used methods for measuring the severity of lymphedema include (1) the

From the Department of MRI,<sup>a</sup> Department of Radiology,<sup>b</sup> and Department of Lymphatic Surgery,<sup>c</sup> Beijing Shijitan Hospital, Capital Medical University.

Additional material for this article may be found online at [www.jvsvenous.org](http://www.jvsvenous.org).

Correspondence: Yunlong Yue, MD, Department of MRI, Beijing Shijitan Hospital, Capital Medical University, Yangfangdian Tieyiyuan Rd No.10, Haidian District, Beijing, China (e-mail: [yueyunlong@bjsjth.cn](mailto:yueyunlong@bjsjth.cn)).

The editors and reviewers of this article have no relevant financial relationships to disclose per the Journal policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

2213-333X

Copyright © 2024 The Author(s). Published by Elsevier Inc. on behalf of the Society for Vascular Surgery. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

<https://doi.org/10.1016/j.jvs.2024.102161>

water displacement method for volume measurement and (2) the conical method for volume measurement. The conical method for volume measurement involves first using a flexible tape measure to measure the circumference at different levels and then applying a volume calculation formula to obtain the final result. Different research institutions or hospitals use varying volume calculation formulas.<sup>4,5</sup> However, the aforementioned two methods of volume measurement are time consuming and labor intensive in clinical application and have certain errors. The current problems include the following. (1) There is an urgent need for clinical diagnosis and treatment methods that can quickly and accurately grade patients with PLE. (2) A method is needed for determining the severity of lymphedema in each limb in bilateral cases.

Radiomics, a novel quantitative analysis method, extracts a large amount of high-throughput information, which radiologists cannot independently recognize with the naked eye, from images in a noninvasive manner and transforms these data into higher-order quantitative features. The routine process includes image acquisition, image segmentation, feature extraction and selection, and model establishment. Machine learning, as a method of radiomics, has already shown potential in disease diagnosis, classification, and prognosis prediction.<sup>6-10</sup> However, according to our literature review, the application of this method in determining the severity of lymphedema has not yet been reported.

Therefore, this study aimed to explore whether machine learning radiomics features based on noncontrast magnetic resonance imaging (MRI) can predict the severity of primary lower limb lymphedema.

## METHODS

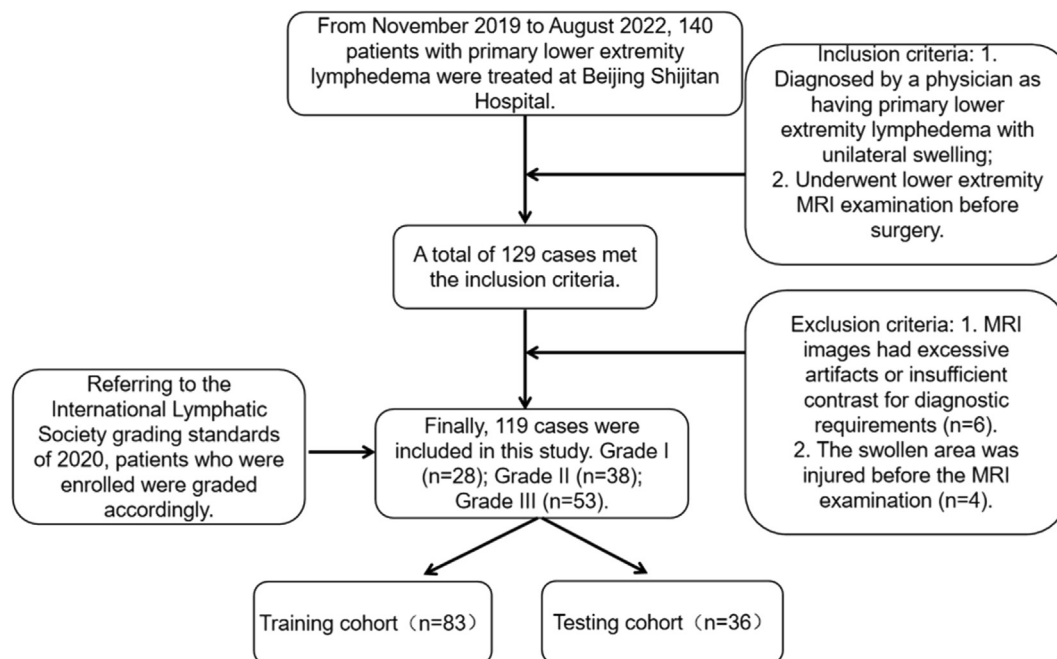
This was a retrospective analysis of 140 patients with primary lower extremity lymphedema at Beijing Shijitan Hospital from November 2019 to August 2022. The inclusion criteria were (1) diagnosis by a physician as having primary lower extremity lymphedema with unilateral swelling; (2) having undergone lower extremity MRI examination before surgery. Exclusion criteria were (1) MRI with excessive artifacts or insufficient contrast for diagnostic requirements and (2) injury of the swollen area before the MRI examination. A total of 119 patients were ultimately included (Fig 1). According to the International Lymphatic Society grading standards of 2020, this study selected six planes: the foot and ankle, the lower one-third of the calf, the upper one-third of the calf, the knee joint, the lower one-third of the thigh, and the upper one-third of the thigh. The following formula was used for calculating limb volume:  $\pi h(R1^2 + r1^2 + R1 \times r1) + \pi h(R2^2 + r2^2 + R2 \times r2) + \dots + \pi h(Rn^2 + rn^2 + Rn \times rn)$ , where R is one-half of the measured segment's lower landmark point corresponding with the limb circumference/ $2\pi$ ; r is one-half of the measured

## ARTICLE HIGHLIGHTS

- **Type of Research:** A single-center, registry-based retrospective cohort study
- **Key Findings:** We stratified 119 patients with primary lower extremity lymphedema into groups based on varying degrees of edema severity and established models using noncontrast magnetic resonance imaging-derived parameters. All five models demonstrated robust performance in distinguishing between nonsevere and severe groups, with area under the curve (AUC) values exceeding 0.850 in the test sets. Among them, ExtraTree exhibited the best performance, achieving an AUC of 0.974 (95% confidence interval, 0.9437-1.0000) in the training set, with a sensitivity of 89.2% and a specificity of 95.7%. In the test set, the AUC was 0.938 (95% confidence interval, 0.8539-1.0000), accompanied by a sensitivity of 75% and a specificity of 100%. Decision curve analysis revealed that the net benefit of the ExtraTree model surpassed both extreme curves when the prediction probability ranged from 16% to 78%.
- **Take Home Message:** Models established using radiomic features derived from noncontrast magnetic resonance imaging significantly aid in diagnosing the severity of edema in patients with primary lower extremity lymphedema. The results of this study herald further potential for advancements in the research of lower extremity lymphedema using machine learning and radiomics.

segment's upper landmark point corresponding with the limb circumference/ $2\pi$ ; and h is the distance between the two landmark points of the measured segment) for limb volume calculation. Grading was based on the formula  $[(\text{affected limb volume} - \text{healthy limb volume})/\text{healthy limb volume}] \times 100\%$ . Patients were graded as mild (>5% to <20%), moderate (>20% to <40%), or severe (>40%) and were divided into a non-severe group (mild and moderate) and a severe group. Clinical characteristics of the enrolled patients were collected through the Health Information System of Beijing Shijitan Hospital. The patient characteristics included sex, congenital onset, age of onset, duration of disease, age at admission, nonpitting edema, increased skin temperature, skin redness, significant skin thickening, and pain. The ethics committee of our hospital approved this study and waived the requirement for informed consent to review the medical records and images owing to the retrospective study design.

**MR data acquisition and parameters.** All MRI examinations in this study were performed using a 1.5 T MR scanner (Ingenia, Philips, Best, the Netherlands) with an eight-



**Fig 1.** Patient recruitment and selection workflow. *MRI*, Magnetic resonance imaging.

channel body coil. The patients were first placed in a supine position with the feet first. The scanning range extended from the level of the femoral head to the entire foot. Short-time inversion recovery sequences were used for axial and coronal scanning with the following parameters: TR, 5298 ms/5649 ms; TE, 80 ms; field of view, 400 mm × 220 mm/360 mm × 400 mm; slice thickness, 6 mm/7 mm; interslice gap, 0.6 mm; voxel size, 1.5 mm<sup>3</sup>; IR, 165 ms/160 ms; and NEX, 1 time.

**Radiomics and radiomics signature building.** Case images were exported from the Picture Archiving and Communication System in a DICOM format and imported into ITK-snap (<https://www.itksnap.org>). On the noncontrast MRI, 10 consecutive transaxial images of the affected side where the edema was most severe were selected while avoiding the joint surface as much as possible; these images were saved in the Nifti format. A semiautomatic threshold method in the application software was used to fill the area between the subcutaneous tissue and muscle, and the region of interest spreading to the skin and muscle was erased layer by layer, with manual filling of any unfilled areas between the subcutaneous tissue and muscle. This information was then reviewed by two physicians with >15 years of experience in diagnosing lymphedema. During the review phase, if the two physicians disagreed, they reached a consensus through consultation. A volume of interest was generated, and a mask file was saved with dimensions consistent with those of the original image.

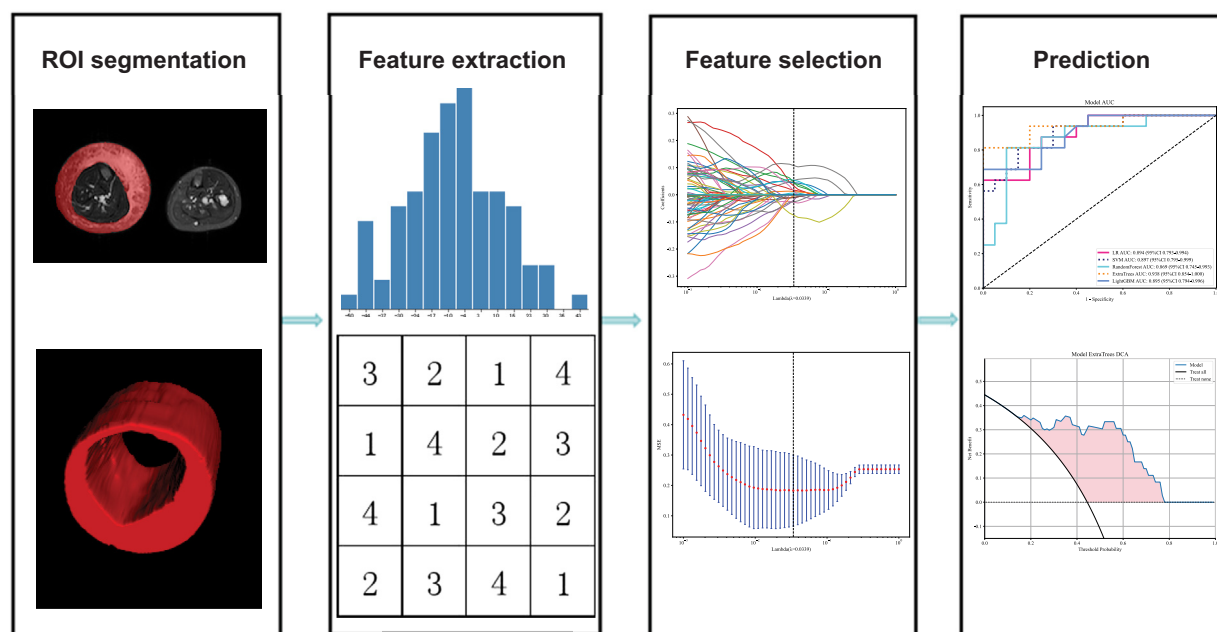
PyRadiomics version 3.0.1 was used for radiomic feature extraction. The extracted radiomic features are

composed of several categories: geometric features, first-order features, and texture features, including gray level co-occurrence matrix (GLCM), gray level run length matrix, gray level size zone matrix, gray level dependence matrix, and neighborhood gray tone difference matrix, on the original image and the images after various filters (Table 1). The description of derived image types is shown in the Supplementary Table (online only). After feature extraction, Z score normalization was performed. The extracted radiomic features were subjected to statistical analysis. Only features with a *P* value of <.05 were retained. Highly correlated similar features were excluded using Pearson's correlation coefficient. A Pearson's correlation coefficient of >0.9 was considered to indicate highly correlated features, and only one of the two parameters was retained. To maximize the retention of the described features, a greedy recursive elimination strategy was used for feature filtering, whereby the most redundant feature in the current set was removed each time. Subsequently, the least absolute shrinkage and selection operator (LASSO) method was used to select an optimized subset of features for building the final model. LASSO first adjusted the parameter  $\lambda$  through 10-fold cross-validation to obtain an optimized  $\lambda$  and then calculated the coefficients of each feature using the optimal parameter  $\lambda$  and selected features with nonzero coefficients. After LASSO feature selection, the model was constructed using Scikit-learn version 1.0.2. Five types of radiomics signatures were established using the remaining features, including logistic regression, support vector machine, random Forest, ExtraTrees, and light gradient boosting machine (Fig 2).

**Table I.** Description of radiomic features

Radiomic features categories	Radiomic features	Description
3D Geometric features (shape) (n = 17)	VoxelVolume, Sphericity, Compactness1, MajorAxisLength, et al.	Describes the morphological characteristics within the volume of interest
First-order features (n = 19)	Skewness, kurtosis, mean value, root mean square deviation, median, et al.	Describes the distribution of grayscale intensity
Texture features: gray level co-occurrence matrix (GLCM) (n = 24)	Autocorrelation, JointAverage, ClusterProminence, Correlation, et al.	Describes the how combinations of discretized intensities (grey levels) of neighboring pixels, or voxels in a 3D volume
Texture features: Gray Level Run Length Matrix (GLRLM) (n = 16)	ShortRunEmphasis, GrayLevelNonUniformity, RunEntropy, et al.	Assesses the distribution of discretized grey levels in an image or in a stack of images
Texture features: gray level size zone matrix (GLSZM) (n = 16)	SmallAreaEmphasis, GrayLevelNonUniformity, ZonePercentage, et al.	Counts the number of groups (or zones) of linked voxels, which are linked if the neighboring voxel has an identical discretized grey level
Texture features: neighborhood gray level dependence matrix (NGLDM) (n = 14)	SmallDependenceEmphasis, GrayLevelVariance, HighGrayLevelEmphasis, et al.	Describes the relations between intensities of every central pixel and all its neighbors within a window
Texture features: Neighborhood Gray Tone Difference Matrix (NGTDM) (n = 5)	Coarseness, Contrast, Busyness, Complexity and Strength	Evaluates the differences between pixel intensity and the average intensities of their neighborhoods within a window

Reference: <https://pyradiomics.readthedocs.io/en/latest/features.html#radiomic-features>.

**Fig 2.** The steps in the workflow of this study were image acquisition, region of interest (ROI) delineation, feature extraction, feature selection, and radiomics signature building.

**Statistical analysis.** All data analyses were performed using Python 3.7.12 on the OnekeyAI platform version 3.5.28. For the statistical analyses, we used Statsmodels

version 0.13.2. using the Shapiro-Wilk test. For continuous variables, significance was determined using either the *t* test or the Mann-Whitney *U* test. Based on the data

**Table II.** Clinical data table

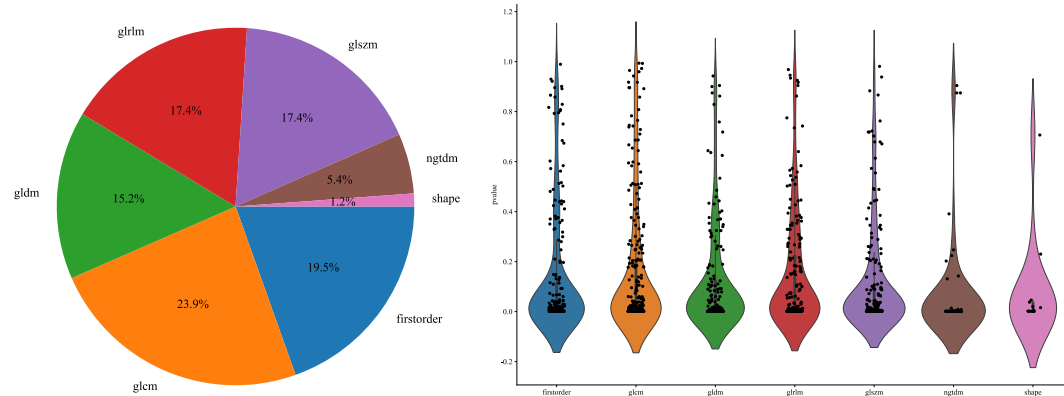
Clinical data	Train-label = ALL	Train-label = 0	Train-label = 1	P value	Test-label = ALL	Test-label = 0	Test-label = 1	P value
Age at onset, years	19.01 ± 16.27	18.33 ± 14.91	19.86 ± 17.98	.905	20.53 ± 15.52	18.95 ± 10.39	22.50 ± 20.45	.848
Length of course	11.75 ± 10.82	10.17 ± 10.44	13.71 ± 11.10	.101	9.19 ± 10.31	3.92 ± 4.30	15.79 ± 11.88	<.001
Age at admission, years	30.87 ± 17.39	28.70 ± 16.28	33.57 ± 18.54	.219	29.69 ± 18.25	22.85 ± 12.26	38.25 ± 21.12	.029
Sex				.84				.553
0	47 (56.63)	27 (58.70)	20 (54.05)		24 (66.67)	12 (60.00)	12 (75.00)	
1	36 (43.37)	19 (41.30)	17 (45.95)		12 (33.33)	8 (40.00)	4 (25.00)	
Onset at birth				.444				.91
0	67 (80.72)	39 (84.78)	28 (75.68)		35 (97.22)	20 (100.00)	15 (93.75)	
1	16 (19.28)	7 (15.22)	9 (24.32)		1 (2.78)	Null	1 (6.25)	
Nonpitting edema				.614				.605
0	66 (79.52)	38 (82.61)	28 (75.68)		29 (80.56)	15 (75.00)	14 (87.50)	
1	17 (20.48)	8 (17.39)	9 (24.32)		7 (19.44)	5 (25.00)	2 (12.50)	
High skin temperature				.1				.742
0	70 (84.34)	42 (91.30)	28 (75.68)		29 (80.56)	17 (85.00)	12 (75.00)	
1	13 (15.66)	4 (8.70)	9 (24.32)		7 (19.44)	3 (15.00)	4 (25.00)	
Skin red				.106				.022
0	58 (69.88)	36 (78.26)	22 (59.46)		26 (72.22)	18 (90.00)	8 (50.00)	
1	25 (30.12)	10 (21.74)	15 (40.54)		10 (27.78)	2 (10.00)	8 (50.00)	
Skin is markedly thickened				.779				.964
0	18 (21.69)	11 (23.91)	7 (18.92)		8 (22.22)	5 (25.00)	3 (18.75)	
1	65 (78.31)	35 (76.09)	30 (81.08)		28 (77.78)	15 (75.00)	13 (81.25)	
Pain				.46				.91
0	79 (95.18)	45 (97.83)	34 (91.89)		35 (97.22)	20 (100.00)	15 (93.75)	
1	4 (4.82)	1 (2.17)	3 (8.11)		1 (2.78)	Null	1 (6.25)	

distribution, the data are presented herein as the mean ± standard deviation. For categorical variables, analysis was performed using the chi-square test, and the results are expressed as counts and percentages. A *P* of value <.05 was considered to indicate statistical significance. Receiver operating characteristic curves were used to compare the diagnostic performance of different radiomics signatures between the nonsevere and severe groups and to calculate the area under the curve (AUC), accuracy, sensitivity, and specificity of the receiver operating characteristic curves.

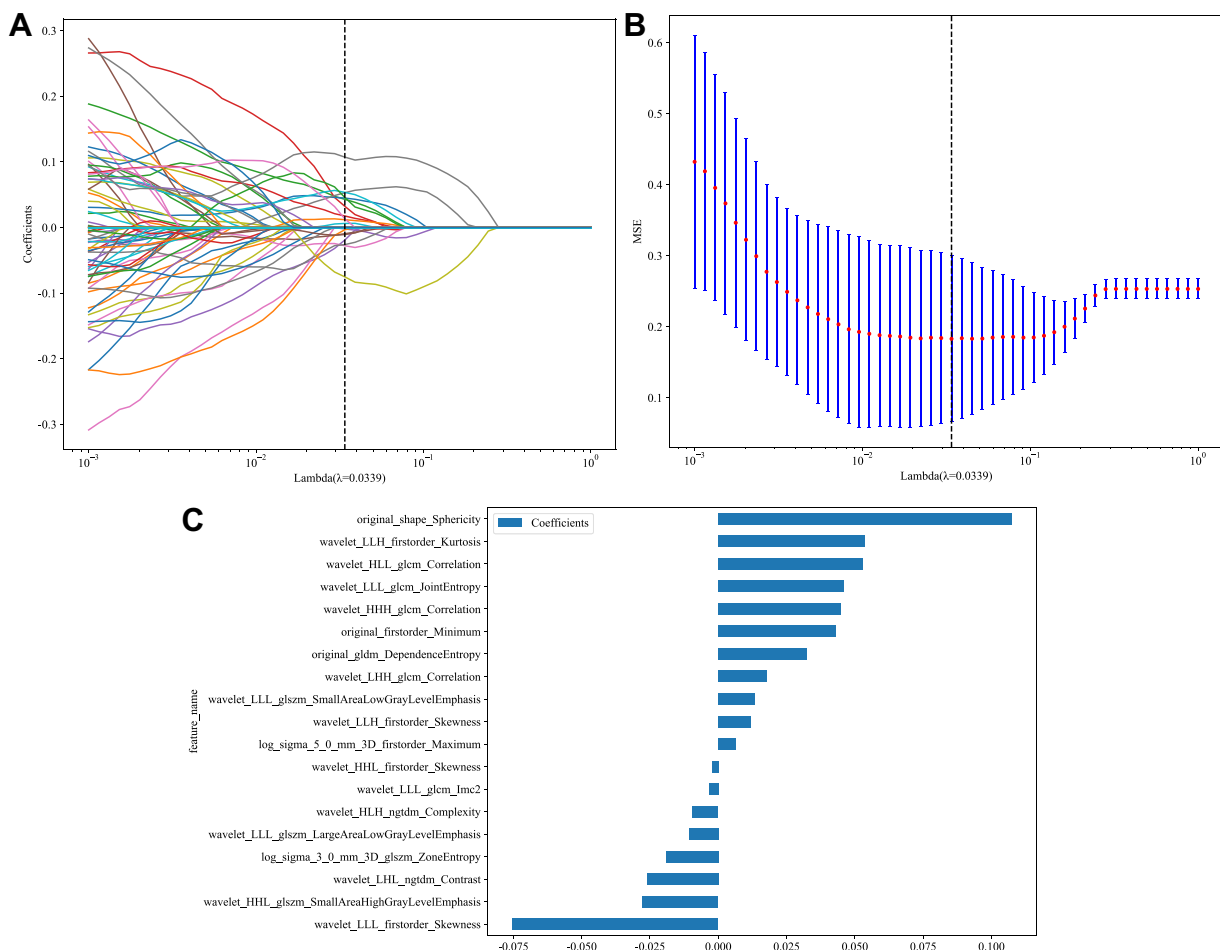
## RESULTS

The 119 patients assessed in this study were divided randomly into a training cohort and a testing cohort at a ratio of 7:3, with 83 patients in the training cohort (46 in the nonsevere group and 37 in the severe group) and 36 patients in the testing cohort (20 in the nonsevere

group and 16 in the severe group). In the training and testing sets, the age at onset was 19.01 ± 16.27 years and 11.75 ± 10.82 years; the length of course was 20.53 ± 15.52 months and 9.19 ± 10.31 months; and the age at admission was 30.87 ± 17.39 years and 29.69 ± 18.25 years. In the training set, there were 36 males (43.37%) and 47 females (56.63%). Among the patients, 17 (20.48%) had nonpitting edema, 13 (15.66%) had high skin temperature, 25 (30.12%) had skin redness, 65 (78.31%) had markedly thickened skin, and 4 (4.82%) reported pain. In the testing set, there were 12 males (33.33%) and 24 females (66.67%). Among these patients, 7 (19.44%) had nonpitting edema, 7 (19.44%) had high skin temperature, 10 (27.78%) had skin redness, 28 (77.78%) had markedly thickened skin, and 1 (2.78%) reported pain. In the testing cohort, differences in Length\_of\_course, Age\_at\_admission, and Skin\_redness (*P* < .05) were statistically significant. The remaining clinical data



**Fig 3.** The proportion, distribution and p of various radiomics features. *glcm*, gray level co-occurrence matrix; *glgm*, gray level dependence matrix; *glrlm*, gray level run length matrix; *glszm*, gray level size zone matrix; *ngldm*, neighborhood gray level dependence matrix.

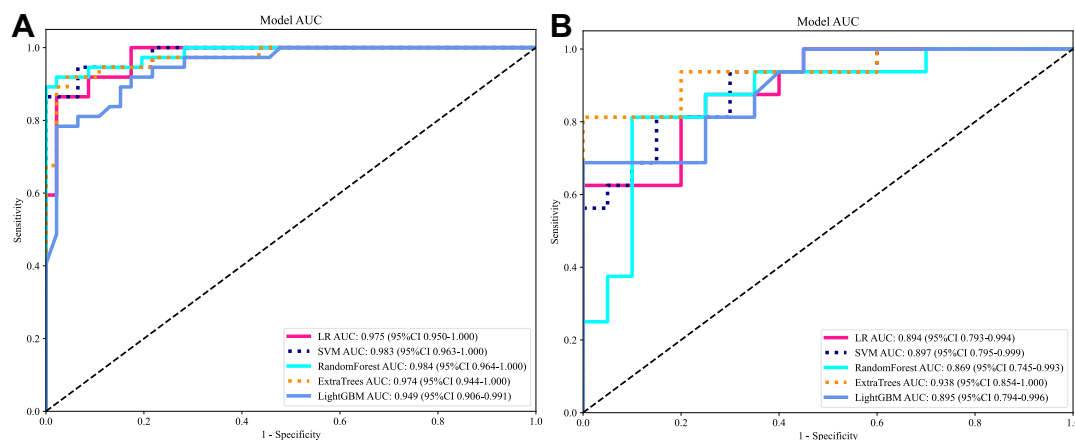


**Fig 4.** Least absolute shrinkage and selection operator (LASSO)-based radiomic feature selection. **(A)** Ten-fold cross-validated coefficients. **(B)** Mean square error. **(C)** Histogram of the radiomics score based on the selected features. The description of radiomic features is shown in [Table I](#), and the description of image types is shown in the [Supplementary Table](#) (online only).

**Table III.** Comparison of the discriminative power between different radiomics signatures

model_name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV
LR_Train	0.916	0.975	0.9497-1.0000	0.838	0.978	0.969	0.882
LR_Test	0.806	0.894	0.7932-0.9943	0.562	1.000	1.000	0.741
SVM_Train	0.928	0.983	0.9631-1.0000	0.919	0.935	0.919	0.935
SVM_Test	0.806	0.897	0.7952-0.9985	0.750	0.850	0.800	0.810
RandomForest_Train	0.940	0.984	0.9640-1.0000	0.892	0.978	0.971	0.918
RandomForest_Test	0.833	0.869	0.7446-0.9929	0.750	0.900	0.857	0.818
ExtraTrees_Train	0.928	0.974	0.9437-1.0000	0.892	0.957	0.943	0.917
ExtraTrees_Test	0.889	0.938	0.8539-1.0000	0.750	1.000	1.000	0.833
LightGBM_Train	0.880	0.949	0.9058-0.9914	0.757	0.978	0.966	0.833
LightGBM_Test	0.833	0.895	0.7943-0.9963	0.625	1.000	1.000	0.769

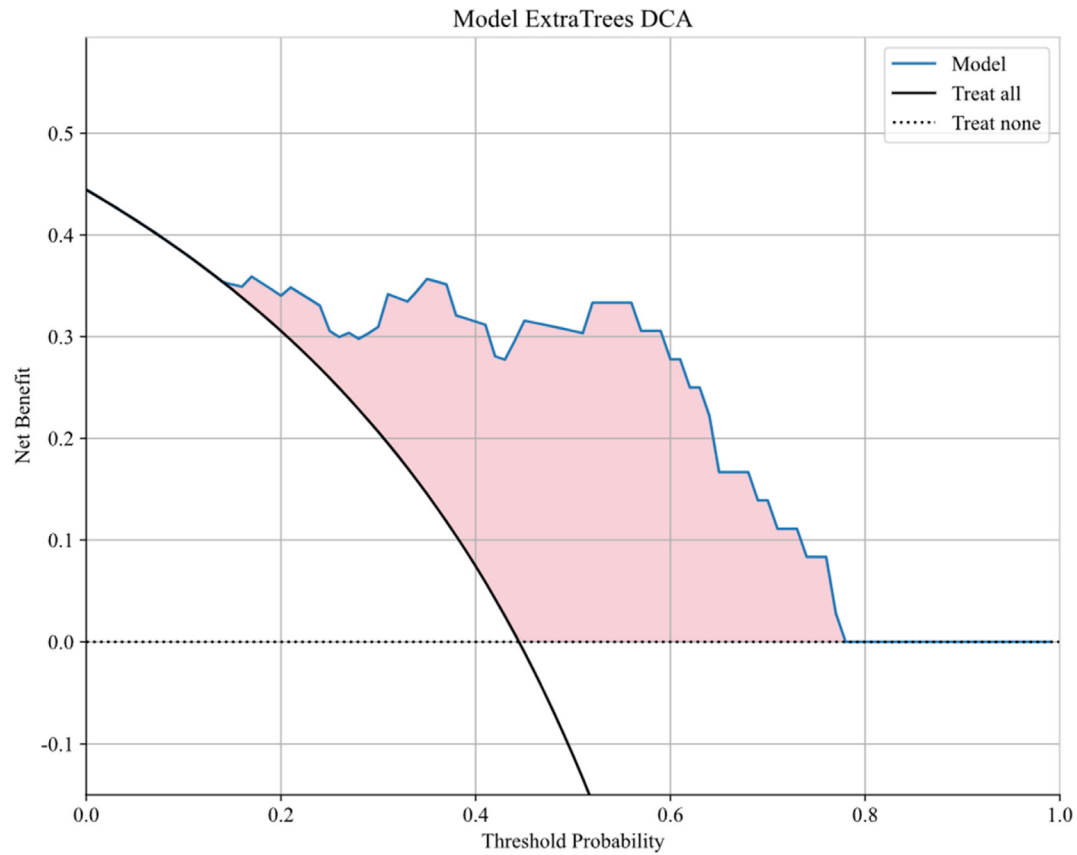
AUC, Area under curve; CI, confidence interval; LR, logistic regression; LightGBM, light gradient boosting machine; NPV, negative predictive value; PPV, positive predictive value; SVM, support vector machine.

**Fig 5.** Receiver operating characteristic (ROC) curve of the five radiomics signatures discriminating between the nonsevere group and the severe group. **(A)** Radiomics signature training cohort ROC curves. **(B)** Radiomics signature testing cohort ROC curve. AUC, Area under the curve; LR, logistic regression; SVM, support vector machine.

in the testing cohort and all clinical data in the training cohort were not significantly different (Table II).

A total of 1196 radiomic features were extracted. Geometric features (shape), 14; first-order features (first-order), 233; GLCM, 286; GLDM, 182; gray level run length matrix, 208; GLSZM, 208; neighborhood gray tone difference matrix, 64 (Fig 3). The LASSO was further used for feature selection. The regularization parameter  $\lambda$  was adjusted through 10-fold cross-validation, obtaining an optimized  $\lambda$  of 0.0339, and 21 features with nonzero coefficients were ultimately selected (Fig 4). All five models performed well in distinguishing between the nonsevere group and the severe group. The ExtraTree model performed the best, with an AUC of 0.974 (95% confidence interval [CI], 0.9437-1.0000), a sensitivity of 89.2%, and a specificity of 95.7% on the training set. On the test set, these values were 0.938 (95% CI, 0.8539-1.0000), 75%, and 100%, respectively (Table III and Fig 5).

**Decision curve analysis.** This analysis was used to evaluate the clinical usefulness of the model in guiding treatment decisions for patients with severe PLE (Fig 6), with treat-all and treat-none as two extreme curves. The results indicated that, when the predicted probability was between 16% and 78%, the net benefit of this prediction model was greater than that of the two extreme curves, demonstrating that the prediction model had strong clinical value in guiding decisions for both nonsevere and severe lymphedema patients. Therefore, when the probability of severe lymphedema occurring was 16%, aggressive intervention measures needed to be taken, benefiting 35 of every 100 patients using the prediction model without harming others. However, when the probability of severe edema occurring was 78%, none of the 100 patients included in this prediction model would have benefitted from preventive measures.

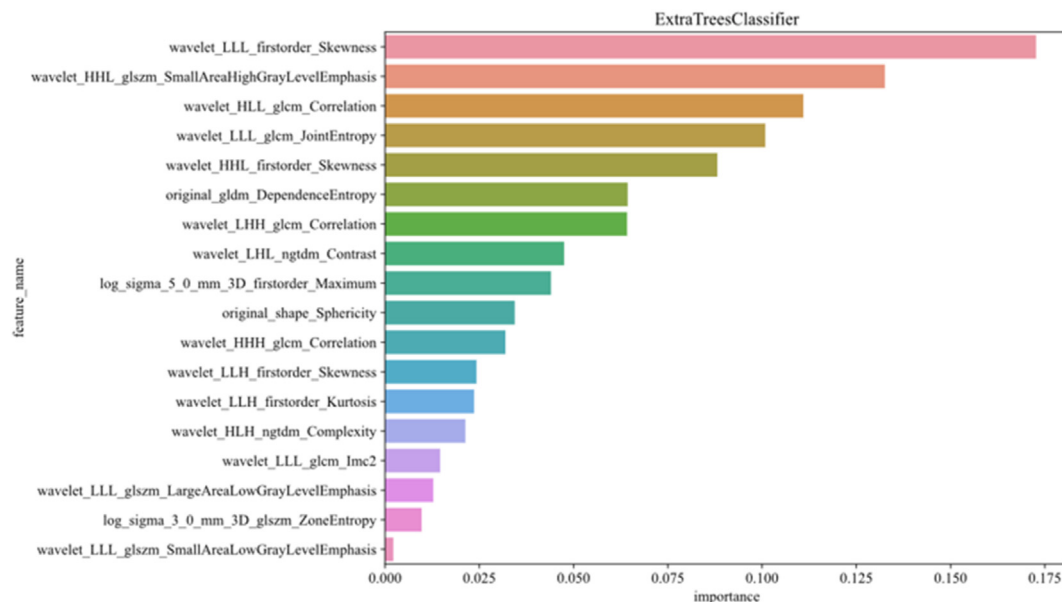


**Fig 6.** DCA of the testing cohort. The results indicated that, when the predicted probability was between 16% and 78%, the net benefit of this prediction model was greater than that of the two extreme curves, demonstrating that the prediction model had strong clinical value in guiding decisions for both nonsevere and severe lymphedema patients. DCA, decision curve analysis.

Based on the model's performance, we selected the ExtraTree model for further analysis of variable importance (Fig 7). The risk of patients developing severe lymphedema is determined by the importance of variables in the ExtraTree model. Among the 18 variables of the ExtraTree model (with 3 of 21 variables having an importance of 0), wavelet LLL first-order skewness was found to be the most important, followed by wavelet HHL glszm SmallAreaHighGrayLevelEmphasis and wavelet HLL glcm correlation. Wavelet LLL first-order skewness among the first-order variables reflects the asymmetry degree of the feature values relative to the mean value. The greater the skewness is the more the distribution deviates from the normal distribution. HHL glszm SmallAreaHighGrayLevelEmphasis reflects the fineness of the image in high signal areas of the image in the gray-level size area matrix variables. Wavelet HHL glcm correlation is a value between uncorrelated and perfectly correlated, displaying the linear dependence of grayscale values in GLCM on their respective voxels. Original\_shape\_Sphericity is a variable that describes the volume and sphericity of the lesion.

## DISCUSSION

Unlike secondary lymphedema, PLE is caused by congenital or idiopathic factors, and its pathogenesis is not yet understood fully. Moreover, the disease progression of PLE cannot be predicted. Some patients experience rapid spontaneous resolution after symptom onset, whereas others exhibit a gradual worsening of symptoms in an irreversible chronic progressive process. Therefore, the current research directions and focuses for PLE are its qualitative diagnosis, severity assessment, and personalized treatment for patients at different stages and grades.<sup>11,12</sup> MRI, with its nonionizing radiation, multiple imaging sequences, and high soft tissue resolution, plays a significant role in the qualitative and quantitative diagnosis of lymphedema. The short-time inversion recovery sequence is a nonselective fat suppression technique that essentially works by exploiting the difference in T1 relaxation times between water and fat tissues. This technique is minimally affected by magnetic field inhomogeneities and has the advantage of allowing large-scale fat suppression with good visualization of complex anatomical structures and lesions rich



**Fig 7.** Importance of features in the ExtraTree. *log\_sigma\_5\_0\_mm\_3D\_firstorder\_Maximum*, first-order features—maximum extracted from derived images using Laplacian of Gaussian filter with sigma of 5, 0 mm; *original\_gldm\_DependenceEntropy*, texture features of gray level dependence matrix—DependenceEntropy extracted from original images; *wavelet\_HHL\_glcM\_correlation*, texture features of gray level co-occurrence matrix—correlation extracted from derived images using wavelet transform with High Low Low pass filter; *wavelet\_HHL\_glszm\_SmallAreaHighGrayLevelEmphasis*, texture features of gray level size zone matrix—SmallAreaHighGrayLevelEmphasis extracted from derived images using wavelet transformed with High High Low pass filter; *wavelet\_LLL\_firstorder\_skewness*, first-order features—skewness extracted from derived images using wavelet transform with Low Low Low-pass filter in the three dimensions. Other features refer to the above expression. The description of radiomic features is shown in [Table I](#), and the description of image types is shown in [Supplementary Table](#) (online only).

in water content. Currently, the nonenhanced MRI is widely used in the clinical application and scientific research of MRI imaging in lymphedema.<sup>13</sup>

The radiomics signature and radiomics have been preliminarily applied in the study of lymphedema, mainly focusing on three aspects: (1) lymphedema prediction models; (2) differential diagnostic models for lymphedema and lipedema; and (3) lymphedema component analysis. Among these variables, predicting postoperative lymphedema in patients with cancer is the hottest topic and involves establishing various machine learning models based on clinical data, laboratory tests, and imaging indicators to predict the risk of early lymphedema in patients with breast cancer or patients with gynecological patients with cancer.<sup>14-18</sup> Effectively distinguishing the components of lymphedema is a popular subject in the clinical surgical treatment of lymphedema and a major challenge faced by clinicians. Therefore, research on lymphedema component segmentation models is also a hot topic.<sup>19</sup> It has long been challenging to differentiate between lymphedema and limb swelling caused by other factors. Therefore, a differential diagnostic model for lymphedema and lipedema based on radiomics has emerged.<sup>20</sup> Trinh et al<sup>17</sup> successfully developed a machine learning model that uses complete blood

counts and treatment data to predict lymphedema. The authors believe that their study marks the first step in applying machine learning to detect the stages of lymphedema, and future research will focus on developing prediction models for lymphedema of varying severity. However, according to our literature review, no researchers have yet used the radiomics signature or radiomics to distinguish the severity of PLE. This study combined the nonenhanced MRI, which is most commonly used for lymphedema, with five common models to establish an automatic recognition model for patients with severe and nonsevere PLE. Among them, the ExtraTree model performed the best, with AUC values of 0.974 (95% CI, 0.9437-1.0000) and 0.938 (95% CI, 0.8539-1.0000) in the training and test sets, respectively.

Current research suggests that PLE lesions are located in the subcutaneous fat layer of the lower limbs, and their composition includes adipose tissue, fibrosis, lymph fluid, dilated lymphatic vessels, and so on.<sup>19,21-23</sup> The reason for outlining all components of the subcutaneous fat layer in this study, rather than a single component within the fat layer, is threefold. First, preliminary research by our group revealed that skin thickness, lymphatic fluid thickness (band sign), and fibrosis

(honeycomb sign) are all indicators for assessing the severity of PLE. Second, some researchers believe that subcutaneous fat hypertrophy is an indicator of PLE severity.<sup>24</sup> Third, currently, it is still challenging for imaging technologies to distinguish these components accurately.

The results of this study show that the five models (logistic regression, support vector machine, random Forest, ExtraTrees, and light gradient boosting machine) all perform well in distinguishing between nonsevere and severe cases of PLE. These findings establish confidence in the future research directions of radiomics and radiomics signatures in PLE. Based on the good performance of the ExtraTree model, we selected this model to further analyze the importance of the variables. The risk of developing severe edema in PLE is determined based on the importance of its variables in the ExtraTree model. Among the 18 variables in the ExtraTree model (of 21 variables, 3 have an importance of 0), wavelet LLL first-order skewness is the most important, followed by wavelet HHL glszm SmallAreaHighGrayLevelEmphasis and wavelet HHL glcm correlation. The importance of the wavelet LLL first-order skewness variable is significantly greater than that of the other variables. The skewness among the first-order variables reflects the asymmetry degree of the feature values relative to the mean value. The greater the skewness is, the more the distribution deviates from the normal distribution. LASSO regression showed that skewness is negatively correlated with the severity of lymphedema, indicating that in severe PLE, although the lesion range is large, the complexity of the components within the lesion is low. HHL glszm SmallAreaHighGrayLevelEmphasis, the second most important variable, reflects the fineness of the image in high-signal areas of the image in the gray-level size area matrix variables. LASSO regression also revealed that this variable is negatively correlated with the severity of lymphedema. This finding indicates that, in severe PLE, the lymphatic fluid (high-signal areas) appears more blurred, which may be related to the fact that patients with severe PLE have less fat content and more lymphatic fluid components, and the fat suppression in the nonenhanced MRI actually shows poor results. Furthermore, this result suggests that physicians may overestimate the severity of PLE assessed through imaging signs. Notably, although the morphological index `original_shape_Sphericity` in the original image showed a strong positive correlation with the severity of lymphedema according to LASSO regression, its variable importance was far lower than that of the top three variables mentioned above. These results also illustrate from another perspective that the area and volume of subcutaneous soft tissue swelling observed in general imaging indicators can reflect the severity of edema. However, the indices obtained from radiomics have greater value in predicting the risk of severe PLE. This study

demonstrated that MR-based radiomic parameters have better potential for application in assessing PLE.

### Limitations

1. This study is a retrospective single-center study, and the results only report internal validation, but the model still needs external validation. In the future, we may conduct a multicenter study.
2. This study is a preliminary investigation into the severity of PLE, focusing only on a binary classification between patients with severe and nonsevere disease. Therefore, the mild to moderate degrees in the guideline classification were categorized as nonsevere. Future studies will further explore three-way classification determination in PLE using machine learning.
3. This study did not include treatment data; therefore, future research will integrate clinical therapy.

### CONCLUSIONS

In this study, we successfully developed a radiomics signature based on the Noncontrast MRI combined with radiomics to predict the severity of primary lower limb lymphedema. Compared with the observable general imaging signs, the extracted radiomic data are larger in volume, more objective and authentic, more important for identifying genomic indicators, faster in differential diagnosis, and more suitable for determining lymphedema severity. The ExtraTrees algorithm model in this study exhibited good performance in predicting the severity of PLE, indicating a further step in the research on radiomics in PLE.

We appreciate the excellent medical care, diagnostic techniques, surgery, and rehabilitation plans provided by the lymphatic surgery team of Beijing Shijitan Hospital during the patients' hospitalization. We would like to express our gratitude to patients and their family for their excellent communication and cooperation between patients and the doctor, as well as for their permission in drafting and publishing the final manuscript.

### AUTHOR CONTRIBUTIONS

Conception and design: JR, XL, TC, RW, YY  
 Analysis and interpretation: JR, XL, ML, YY  
 Data collection: JR, XL, ML, TC, JG, RZ, KH  
 Writing the article: JR, XL, TC, JG, RZ, KH, RW, YY  
 Critical revision of the article: JR, XL, ML, TC, RW, YY  
 Final approval of the article: JR, XL, ML, TC, JG, RZ, KH, RW, YY  
 Statistical analysis: JR, TC, JG  
 Obtained funding: JR, YY  
 Overall responsibility: YY  
 JR and XL contributed equally to this article and share co-first authorship.

### FUNDING

This work was funded by National Natural Science Foundation of China (61876216).

**DISCLOSURES**

None.

**REFERENCES**

- Warren AC, Brorson H, Borud LJ, Slavin SA. Lymphedema: a comprehensive review. *Ann Plast Surg.* 2007;59:464–472.
- Executive Committee of the International Society of Lymphology. The diagnosis and treatment of peripheral lymphedema: 2020 consensus document of the International Society of Lymphology. *Lymphology.* 2020;53:3–19.
- Grada AA, Phillips TJ. Lymphedema: pathophysiology and clinical manifestations. *J Am Acad Dermatol.* 2017;77:1009–1020.
- Brunelle CL, Taghian AC. Lymphoedema screening: setting the standard. *Br J Cancer.* 2020;123:1–2.
- Sitzia J. Volume measurement in lymphoedema treatment: examination of formulae. *Eur J Cancer Care.* 1995;4:11–16.
- Conti A, Duggento A, Indovina I, Cuerrisi M, Toschi N. Radiomics in breast cancer classification and prediction. *Semin Cancer Biol.* 2021;72:238–250.
- shin J, Seo N, Baek SE, et al. MRI Radiomics model predicts pathologic complete response of rectal cancer following chemoradiotherapy. *Radiology.* 2022;303:351–358.
- Anconina R, Ortega C, Metser U, et al. Combined 18 f-FDG PET/CT radiomics and sarcopenia score in predicting relapse-free survival and overall survival in patients with esophagogastric cancer. *Clin Nucl Med.* 2022;47:684–691.
- Wang T, Gao T, Guo H, et al. Preoperative prediction of parametrial invasion in early-stage cervical cancer with MRI-based radiomics nomogram. *Eur Radiol.* 2020;30:3585–3593.
- Fang J, Zhang B, Wang S, et al. Association of MRI-derived radiomic biomarker with disease-free survival in patients with early-stage cervical cancer. *Theranostics.* 2020;10:2284–2292.
- Xin J, Sun Y, Xia S, et al. Application of liposuction in treating the primary end-stage lymphedema of lower extremities. *Chin J Plast Surg.* 2019;35:142–147.
- Pappalardo M, Cheng MH. Lymphoscintigraphy for the diagnosis of extremity lymphedema: current controversies regarding protocol, interpretation, and clinical application. *J Surg Oncol.* 2020;121:37–47.
- Salehi BP, Sibley RC, Friedman R, et al. MRI of lymphedema. *J Magn Reson Imaging.* 2023;57:977–991.
- Li MM, Wu PP, Qiang WM, et al. Development and validation of a risk prediction model for breast cancer-related lymphedema in post-operative patients with breast cancer. *Eur J Oncol Nurs.* 2023;63:102258.
- Gross JP, Whelan TJ, Parulekar WR, et al. Development and validation of a nomogram to predict lymphedema after axillary surgery and radiation therapy in Women with breast cancer from the NCIC CTG MA.20 randomized trial. *Int J Radiat Oncol Biol Phys.* 2019;105:165–173.
- Bevilacqua JL, Kattan MW, Changhong Y, et al. Nomograms for predicting the risk of arm lymphedema after axillary dissection in breast cancer. *Ann Surg Oncol.* 2012;19:2580–2589.
- Trinh XT, Chien PN, Long NV, et al. Development of predictive models for lymphedema by using blood tests and therapy data. *Sci Rep.* 2023;13:19720.
- Nomura Y, Hoshiyama M, Akita S, et al. Computer-aided diagnosis for screening of lower extremity lymphedema in pelvic computed tomography images using deep learning. *Sci Rep.* 2023;13:16214.
- Son H, Lee S, Kim K, Koo KI, Hwang CH. Deep learning-based quantitative estimation of lymphedema-induced fibrosis using three-dimensional computed tomography images. *Sci Rep.* 2022;12:15371.
- Nowak S, Henkel A, Theis M, et al. Deep learning for standardized, MRI-based quantification of subcutaneous and subfascial tissue volume for patients with lipedema and lymphedema. *Eur Radiol.* 2023;33:884–892.
- Oliszewski WL, Ambujam PJ, Zaleska M, Cakala M. Where do lymph and tissue fluid accumulate in lymphedema of the lower limbs caused by obliteration of lymphatic collectors? *Lymphology.* 2009;42:105–111.
- Li X, Li B, Hao K, et al. Value of the short time inversion recovery sequence of magnetic resonance imaging in the staging of Klippel-Trenaunay syndrome complicated with lymphedema. *J Vasc Surg Venous Lymphat Disord.* 2024;12:101746.
- Arrivé L, Derhy S, Dahan B, et al. Primary lower limb lymphoedema: classification with non-contrast MR lymphography. *Eur Radiol.* 2018;28:291–300.
- Karlsson T, Karlsson M, Ohlin K, Olsson G, Brorson H. Liposuction of breast cancer-related arm lymphedema reduces fat and muscle hypertrophy. *Lymphat Res Biol.* 2022;20:53–63.

Submitted Jul 5, 2024; accepted Dec 8, 2024.

Additional material for this article may be found online at [www.jvsvenous.org](http://www.jvsvenous.org).