·论 著·

基于超声造影的影像组学联合模型预测甲状腺结节良恶性的价值

刘安信，陈成彩*，罗婷婷，卢 焦，黄秀林，刘桂伶
（右江民族医学院附属医院超声医学科，广西 百色 533000）

[摘要] **目的** 建立基于超声影像组学结合临床参数、常规超声特征的联合模型，并验证其预测甲状腺结节良恶性的价值。**方法** 回顾性纳入经病理证实的患者209例，甲状腺结节237个，按8∶2随机分为训练集（190个结节）和验证集（47个结节）。使用Pyradiomics提取影像组学特征，应用最小绝对收缩和选择算子（least absolute shrinkage and selection operator，LASSO）进行特征筛选，应用k近邻（k-nearest neighbor，KNN）机器学习算法构建影像组学模型。采用单因素和多因素Logistic回归分析筛选临床参数与常规超声特征，将多因素Logistic回归分析出的危险因素纳入k近邻机器学习算法构建临床预测模型。使用筛选后的影像组学特征与多因素Logistic回归分析出的危险因素构建联合模型。通过受试者工作特征（receiver operating characteristic，ROC）曲线比较各模型的诊断效能。**结果** 多因素Logistic回归结果显示，微钙化、垂直位、年龄为鉴别甲状腺良恶性结节的独立危险因素（OR=6.082、13.761、0.938，均$P<0.001$），使用上述危险因素构建临床预测模型，其训练集与测试集曲线下面积（area under the curve，AUC）分别为0.881、0.704。使用LASSO算法降维筛选到的6个关键影像组学特征建立超声影像组学模型，其训练集与测试集AUC分别为0.830、0.819。使用多因素Logistic回归分析出的危险因素与影像组学评分构建联合诊断模型，其训练集及验证集的AUC分别0.939、0.854，诊断效能优于单一临床预测模型或超声影像组学模型。决策曲线图显示联合模型具有较好的临床应用价值。校准曲线显示联合模型与理想模型的拟和度较好。**结论** 基于超声造影的影像组学联合临床参数、超声特征所构建的联合模型可有效的鉴别甲状腺良恶性结节。

[关键词] 甲状腺结节；影像组学；列线图；超声；人工智能 doi：10.3969/j.issn.1007-3205.2025.06.013
[中图分类号] R581 [文献标志码] A [文章编号] 1007-3205(2025)06-0701-09

The value of radiomics combined model based on contrast-enhanced ultrasound in predicting benign and malignant thyroid nodules

LIU An-xin, CHEN Cheng-cai*, LUO Ting-ting, LU Jiao, HUANG Xiu-lin, LIU Gui-ling
(*Department of Ultrasound Medicine, the Affiliated Hospital of Youjiang Medical University for Nationalities, Guangxi Zhuang Autonomous Region, Baise* 533000, *China*)

[**Abstract**] **Objective** To establish a combined model based on ultrasound radiomics combined with clinical parameters and conventional ultrasound features, and to verify its value in predicting benign and malignant thyroid nodules. **Methods** A total of 209 patients with 237 thyroid nodules confirmed by pathology were retrospectively included and randomly divided into training set (190 thyroid nodules) and validation set (47 thyroid nodules) according to an 8∶2 ratio. Pyradiomics was used to extract radiomics features, least absolute shrinkage and selection operator (LASSO) was used for feature screening, and k-Nearest Neighbor (KNN) machine

[收稿日期]2024－03－06
[基金项目]广西研究生教育创新计划项目（YCSW2023493）；自治区卫生健康委西医类自筹经费科研课题（Z-L20230880）
[作者简介]刘安信（1996－），男，广西百色人，右江民族医学院附属医院医师，医学硕士，从事医学超声诊断研究。
* 通信作者。E-mail：108363047@qq.com

learning algorithm was used to construct radiomics model. Univariate and multivariate Logistic regression were used to screen clinical parameters and conventional ultrasound features. The risk factors analyzed by multivariate Logistic regression were included in the k-nearest neighbor machine learning algorithm to construct a clinical prediction model. A combined model was constructed using the screened radiomics features and the risk factors analyzed by multivariate Logistic regression. The diagnostic efficacy of each model was compared by receiver operating characteristic (ROC) curve. **Results** Multivariate Logistic regression results showed that microcalcification, vertical position and age were independent risk factors for differentiating benign and malignant thyroid nodules (OR=6.082, 13.761, 0.938, all $P<0.001$). The clinical prediction model was constructed using the above risk factors, and the area under the curve (AUC) of the training set and the validation set was 0.881 and 0.704, respectively. The ultrasound radiomics model was established by using the six key radiomics features screened by LASSO algorithm. The AUC of the training set and the test/validation set was 0.830 and 0.819, respectively. The risk factors analyzed by multivariate Logistic regression and radiomics score were used to construct a combined diagnostic model. The AUC of the training set and the validation set was 0.939 and 0.854, respectively. The diagnostic efficiency was better than that of a single clinical prediction model or ultrasound radiomics model. The decision curve showed that the combined model had good clinical application value. The calibration curve showed good fit between the combined model and the ideal model. **Conclusion** The combined model constructed by radiomics based on contrast-enhanced ultrasound combined with clinical parameters and ultrasound features can effectively identify benign and malignant thyroid nodules.

[**Key words**] thyroid nodule; radiomics; nomogram; ultrasound; artificial intelligence

甲状腺结节是常见的临床疾病，19%～68%的普通人群经超声检查可检出甲状腺结节，但大多数是不具有干预价值的良性结节[1-2]，术前明确甲状腺结节性质对患者的预后及治疗具有重要的指导作用，但目前超声检查存在主观依赖性、经验依赖性等缺点，如何精准的定量、定性诊断甲状腺结节的性质成为了亟需解决的问题[3]。超声造影通过评估甲状腺结节的血供特征定量的判断结节恶性风险[4-6]，对甲状腺良恶性结节的诊断效能达到了较高水平，但其仍然无法避免传统影像诊断所存在的弊端，如主观依赖性等。影像组学(radiomics)通过计算机高通量地提取大量影像信息，对目标影像信息进行更深、更全面的信息挖掘，实现了目标疾病分割、特征提取、模型建立与评估的目的，最终辅助医生作出更精确的诊断[7-10]。目前联合使用上述定量两种方法预测甲状腺结节良恶性的研究相对较少，本研究旨在使用超声造影图像建立超声影像组学模型，结合临床参数和超声特征建立联合模型，探讨其在鉴别甲状腺良恶性结节方面的临床价值。

1 资料与方法

1.1 一般资料 回顾性纳入2022年6月—2023年9月右江民族医学院附属医院具有甲状腺结节术后或细针穿刺活检病理结果的患者209例，男性36例，女性173例，年龄12～86岁，平均(46.3±12.0)岁，总计237个结节，其中116个甲状腺结节为良性(结节性甲状腺肿107个、桥本甲状腺炎3个、亚急性甲状腺炎6个)，121个甲状腺结节为恶性(乳头状癌120个、转移瘤1个)，随机将甲状腺结节按照8∶2比例分为训练集(n=190，恶性结节97个、良性结节93个)和测试集(n=47，恶性结节24个、良性结节23个)。纳入标准：①术后具有明确病理结果的结节；②患者既往无甲状腺穿刺、消融等介入史。排除标准：①超声图像质控不合格；②病理结果缺失或临床资料不完整。

本研究经医院伦理委员会批准通过，伦理审批号为：YYFY-LL-2023-013。

1.2 临床资料收集 收集纳入患者的基本临床信息及术前甲状腺激素水平，包括性别、年龄、总三碘甲状腺原氨酸(total triiodothyronine，TT3)、总甲状腺素(total thyroid hormone，TT4)、促甲状腺激素(thyroid stimulating hormone，TSH)、游离三碘甲状腺原氨酸(free triiodothyronine，FT3)、游离甲状腺素(free thyroid hormone，FT4)，抗甲状腺球蛋

白抗体(anti-thyroglobulin antibodies，TGAb)，甲状腺过氧化物酶抗体(anti-thyroid peroxidase antibody，Anti-TPO)，血清甲状腺球蛋白(thyroglobulin，TG2)的具体临床数值。

1.3　图像采集　使用富士 Arietta 850 超声诊断仪，配备 2～12 MHz 线阵探头，所使用的超声造影剂为声诺维(博莱科)，获得完整甲状腺结节病灶的动态灌注过程，观察时间不少于 1 min。既往研究[11-13]表明，在甲状腺结节的超声造影定量分析中，甲状腺结节内造影剂所能达到的峰值强度存在显著差异性，并且目前大多数影像组学算法是对静态图像进行分析，因此本研究先使用机带软件 TIC Analysis 画出时间强度曲线(time signal intensity curve，TIC)，然后截取目标结节内造影剂灌注造影峰值时的静态图，以此作为代表用于影像组学分析。以上操作均由具有丰富甲状腺造影检查经验的高年资医师完成，医师在超声检查前并不知晓患者的病史信息。

1.4　常规超声影像学特征判读　本研究应用 2020 甲状腺结节超声恶性危险分层中国指南(chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules，C-TIRADS)[14]系统对甲状腺结节进行良恶性风险评估，评估内容为甲状腺结节的位置、方位、边缘、声晕、结构、回声、回声质地、局灶性强回声、后方回声特点、大小，以上特征由两位副主任医师分析得出，结果一致则记录下来，若仍然存在分歧，则请示更有经验的主任医师讨论并最终达成一致。

1.5　分割图像　首先将患者的术前超声造影图像以 DICOM 格式转换 nii.gz 格式，随后由一名高年资的超声诊断医师在使用开源软件(ITK-SNAP 3.8.0；www.itksnap.org)在患者甲状腺结节超声造影峰值图像上手动勾画感兴趣区(regionof interest，ROI)，再由高年资的超声诊断医师对勾画结果进行审核及修改，见图 1。

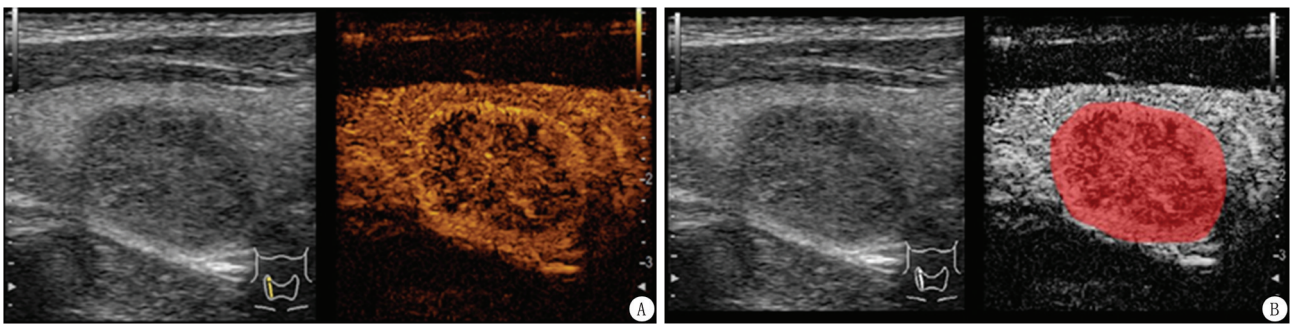


图 1　甲状腺结节及勾画 ROI

A.病灶的超声造影声像图；B.使用 ITK-SNAP 勾画病灶的 ROI 用于特征提取

Figure 1　Thyroid nodules and ROI delineation

1.6　提取并筛选影像组学特征　使用 Pyradiomics 自动提取超声造影峰值图像的影像组学特征，特征包括一阶特征(first order features)、形状特征(shape features)、灰度共生矩阵特征(gray level co-occurrence matrix features，GLCM)、灰度游程矩阵特征(gray level run length matrix features，GLRLM)、灰度区域大小矩阵特征(gray level size zone matrix features，GLSZM)、灰度依赖矩阵特征(gray level dependence matrix features，GLDM)和邻域灰度差分矩阵特征(neighbouring gray tone difference matrix features，NGTDM)。

1.7　构建预测模型　LASSO 算法特征筛选完成后，按照 8∶2的比例将患者随机化分为训练集和验证集，将最终的特征输入到逻辑回归(logistic regression，LR)、支持向量机(support vector machine，SVM)、k 近邻、随机森林(random forest，RF)、极度随机树(extra trees)、极度梯度提升树(extreme gradient boosting，XGBoost)、轻量梯度提升机(light gradient boosting machine，LightGBM)、多层感知器(multi-layer perceptron，MLP)机器学习算法中建立模型，选取结果最好的机器学习算法构建超声影像组学模型。

1.8　统计学方法　应用 Python3.0 和 SPSS26.0 统计软件分析数据。符合正态分布的计量资料比较采用 t 检验，不符合正态分布的计量资料比较采用秩和检验；计数资料比较采用 χ^2 检验；使用受试者工作特征曲线(receiver operating characteristic curve，ROC)评估各模型的诊断性能，使用 Pearson 检验计算特征相关性，使用单因素及多因素 Logistic 回归分析患者的临床参数和传统超声影像学特征，使用 Delong 检验评价各模型间诊断效能，绘制决策曲线评估各模型的临床实用价值，绘制校

准曲线分析联合模型预测结果与实际结果的一致性。$P<0.05$ 为差异有统计学意义。

2 结　果

2.1 一般资料比较　训练集中，良心组和恶心组血清甲状腺球蛋白、年龄、实性、微钙化、极低回声、边缘、垂直位、结节大小比较差异有统计学意义（$P<0.05$），见表1。测试集中，良心组和恶心组年龄、实性、微钙化、边缘、垂直位、结节大小比较差异有统计学意义（$P<0.05$），见表2。

表1　训练集甲状腺结节患者临床参数和超声特征比较

Tablel 1　Comparison of clinical parameters and ultrasound features of patients with thyroid nodules in training set

组别	例数	性别(例数,%)		年龄	TT3	TT4	FT3	FT4
		男性	女性	($\bar{x}\pm s$,岁)	($\bar{x}\pm s$,nmol/L)	($\bar{x}\pm s$,nmol/L)	($\bar{x}\pm s$,pmol/L)	[M(QR),pmol/L]
良性组	93	13(13.98)	80(86.02)	50.46±10.45	1.47±0.43	112.37±19.38	5.07±0.76	11.47(3.19)
恶性组	97	16(16.49)	81(83.51)	42.20±10.71	1.54±0.56	107.08±30.90	5.22±1.39	11.87(4.96)
$\chi^2/t/Z$ 值		0.232		5.381	1.004	1.418	0.909	1.733
P 值		0.779		<0.001	0.316	0.162	0.365	0.088

组别	例数	抗甲状腺球蛋白抗体	TSH	Anti-TPO	TG2	实性(例数,%)	
		[M(QR),kU/L]	[M(QR),mU/L]	[M(QR),kU/L]	[M(QR),μg/L]	无	有
良性组	93	1.43(19.00)	1.59(1.92)	16.60(12.35)	25.40(94.17)	51(54.84)	42(45.16)
恶性组	97	15.00(63.49)	1.92(1.43)	16.30(18.21)	10.90(33.73)	12(12.37)	85(87.63)
$\chi^2/t/Z$ 值		1.206	1.233	0.389	3.922	38.635	
P 值		0.228	0.218	0.697	<0.001	<0.001	

组别	例数	微钙化(例数,%)		极低回声(例数,%)		边缘(例数,%)	
		无	有	无	有	光整	不规则或模糊或甲状腺外侵犯
良性组	93	89(95.70)	4(4.30)	93(100.00)	0(0.00)	75(80.65)	18(19.35)
恶性组	97	61(62.89)	36(37.11)	91(93.81)	6(6.19)	31(31.96)	66(68.04)
$\chi^2/t/Z$ 值		30.756		4.090		45.629	
P 值		<0.001		0.043		<0.001	

组别	例数	垂直位(例数,%)		彗星尾征(例数,%)		结节大小(例数,%)		结节位置(例数,%)	
		<1	>1	无	有	<10 mm	>10 mm	上级	非上级
良性组	93	90(96.77)	3(3.23)	88(94.62)	5(5.38)	25(26.88)	68(73.12)	53(56.99)	40(43.01)
恶性组	97	52(53.61)	45(46.39)	97(100.00)	0(0.00)	52(53.61)	45(46.39)	67(69.07)	30(30.93)
$\chi^2/t/Z$ 值		46.856		3.463		14.071		2.979	
P 值		<0.001		0.063		<0.001		0.115	

表2　测试集甲状腺结节患者临床参数和超声特征比较

Tablel 2　Comparison of clinical parameters and ultrasound features of patients with thyroid nodules in test set

组别	例数	性别(例数,%)		年龄	TT3	TT4	FT3	FT4
		男性	女性	($\bar{x}\pm s$,岁)	($\bar{x}\pm s$,nmol/L)	($\bar{x}\pm s$,nmol/L)	($\bar{x}\pm s$,pmol/L)	[M(QR),pmol/L]
良性组	23	3(13.04)	20(86.96)	56.13±12.82	1.40±0.37	107.73±21.90	5.07±0.69	10.42(4.51)
恶性组	24	6(25.00)	18(75.00)	44.42±9.18	1.38±0.28	106.27±23.06	5.06±0.93	11.09(3.84)
$\chi^2/t/Z$ 值		0.450		3.614	0.146	0.222	0.013	0.936
P 值		0.502		<0.001	0.885	0.825	0.990	0.349

组别	例数	抗甲状腺球蛋白抗体	TSH	Anti-TPO	TG2	实性(例数,%)	
		[M(QR),kU/L]	[M(QR),mU/L]	[M(QR),kU/L]	[M(QR),μg/L]	无	有
良性组	23	16.00(20.20)	1.54(1.11)	16.90(56.45)	19.32(37.53)	12(52.17)	11(47.83)
恶性组	24	14.00(199.20)	1.78(1.70)	17.20(49.78)	20.93(8.19)	1(4.17)	23(95.83)
$\chi^2/t/Z$ 值		0.796	0.553	0.800	0.384	13.528	
P 值		0.426	0.580	0.424	0.701	<0.001	

组别	例数	微钙化(例数,%)		极低回声(例数,%)		边缘(例数,%)	
		无	有	无	有	光整	不规则或模糊或甲状腺外侵犯
良性组	23	23(100.00)	0(0.00)	22(95.65)	1(4.35)	16(69.57)	7(30.43)
恶性组	24	18(75.00)	6(25.00)	21(87.50)	3(12.50)	5(20.83)	19(79.17)
$\chi^2/t/Z$ 值		4.538		0.229		11.284	
P 值		0.033		0.632		0.002	

表2　（续）

组别	例数	垂直位（例数，%）		彗星尾征（例数，%）		结节大小（例数，%）		结节位置（例数，%）	
		＜1	＞1	无	有	＜10 mm	＞10 mm	上级	非上级
良性组	23	23(100.00)	0(0.00)	22(95.65)	1(4.35)	4(17.39)	19(82.61)	11(47.83)	12(52.17)
恶性组	24	18(75.00)	6(25.00)	24(100.00)	0(0.00)	14(58.33)	10(41.67)	16(66.67)	8(33.33)
$\chi^2/t/Z$ 值		4.538		＜0.001		8.331		1.705	
P 值		0.033		0.983		0.010		0.312	

2.2　筛选影像组学特征与构建影像组学评分　通过 Pyradiomics 从患者的超声造影图像中提取到特征影像组学特征 107 个，其中包括一阶特征 18 个、形状特征 14 个、灰度共生矩阵特征 24 个、灰度依赖矩阵特征 14 个、灰度游程矩阵特征 16 个、灰度区域大小矩阵特征 16 个、邻域灰度差分矩阵特征 5 个。经 LASSO 算法降维筛选特征，最终基于原始图像这种图像类型提取到 6 个与甲状腺结节良恶性显著相关的影像组学特征，上述关键影像组学特征具体为：original_firstorder_Kurtosis、original_glcm_Imc1、original_gldm_GrayLevelNonUniformity、original_glszm_SizeZoneNonUniformity、original_ngtdm_Strength、original_shape_Surface Volume Ratio，见图 2。根据每个参数的权重系数，用以下公式计算每个患者的影像组学评分。Radscore＝0.510 5485232067509-0.039829（original_firstorder_Kurtosis）－0.040847（original_glcm_Imc1）－0.055966（original_gldm_Gray Level Non Uniformity）－0.046964（original_glszm_Size Zone Non Uniformity）＋0.049596（original_ngtdm_Strength）＋0.001145（original_shape_SurfaceVolumeRatio）。

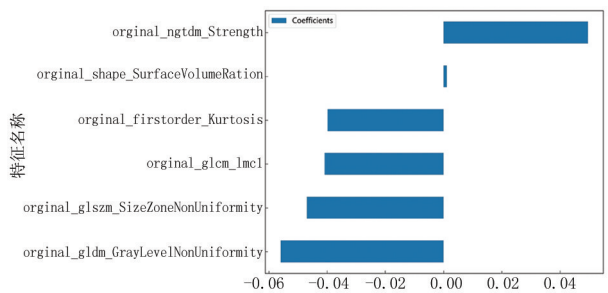


图2　超声影像组学标签特征权重图

注：Kurtosis，峰度，是图像 ROI 中值分布的“峰值”的量度；Surface Volume Ratio，表面积体积比；Imc1，相关性信息度量（IMC）1；Gray Level Non Uniformity，灰度非均匀性，测量图像中灰度级强度值的相似性；Size Zone Non Uniformity，尺寸区域非均匀性，测量图像中大小区域体积的可变性，值越低表示大小区域体积越均匀；Strength，强度

Figure 2　Weight map of ultrasound radiomics label features

2.3　预测模型构建与评价　以甲状腺结节良恶性（良性＝0，恶性＝1）为因变量，将单因素分析中差异有统计学意义的结果纳入多因素 Logistic 回归分析，各因素赋值见表 3。多因素分析显示，微钙化、垂直位、年龄、影像组学评分是预测甲状腺良恶性结节的独立危险因素，见表 4。将多因素 Logistic 回归分析出的临床独立危险因素（微钙化、垂直位、年龄）纳入 k 近邻机器学习建立临床预测模型，临床预测模型训练集 AUC 为 0.881（95% *CI*：0.835～0.925），正确率为 0.758，敏感度为 0.536，特异度为 0.989；测试集 AUC 为 0.704（95% *CI*：0.553～0.854），正确率为 0.551，敏感度为 0.208，特异度为 0.826，见表 5。LASS0 算法特征筛选完影像组学特征后，使用 k 近邻机器学习算法构建影像组学模型，其训练集 AUC 为 0.830（95% *CI*：0.776～0.883），正确率为 0.695，敏感度为 0.412，特异度为 0.989；测试集 AUC 为 0.819（95% *CI*：0.701～0.934），正确率为 0.681，敏感度为 0.458，特异度为 0.913，见表 5。

表3　Logistic 回归分析法赋值

Table 3　Logistic regression analysis assignments

变量类别	变量赋值
自变量	
实性	无＝0；有＝1
微钙化	无＝0；有＝1
极低回声	无＝0；有＝1
边缘	光整＝0；不规则或模糊或甲状腺外侵犯＝1
垂直位	＜1＝0；＞1＝1
彗星尾伪像	无＝0；有＝1
结节大小	＜10 mm＝0；＞10 mm＝1
结节位置	上极＝0；非上极＝1
TT3	连续变量
TT4	连续变量
TSH	连续变量
FT3	连续变量
FT4	连续变量
TGAb	连续变量
Anti-TPO	连续变量
TG2	连续变量
因变量	
甲状腺结节良恶性	良性＝0；恶性＝1

表4　多因素Logistic回归分析筛选预测甲状腺结节良恶性的独立危险因素

Table 4　Multivariate Logistic regression analysis of screening independent risk factors for predicting benign and malignant thyroid nodules

变量	回归系数	标准误	Wald χ^2 值	P 值	OR值	95%CI
实性	0.474	0.547	0.752	0.386	1.606	0.550～4.691
微钙化	1.805	0.726	6.177	0.013	6.082	1.465～25.256
边缘	0.993	0.530	3.513	0.061	2.700	0.956～7.626
垂直位	2.622	0.758	11.951	0.001	13.761	3.112～60.844
结节大小	0.480	0.576	0.696	0.404	1.617	0.523～4.999
TG2	0.000	0.002	0.016	0.900	1.000	0.996～1.004
年龄	−0.064	0.022	8.291	0.004	0.938	0.898～0.980
影像组学评分	5.917	1.372	18.610	<0.001	371.473	25.253～5 464.367

表5　三种预测模型的诊断性能参数

Table 5　Diagnostic performance parameters of three prediction models

模型	AUC	准确度	敏感度	特异度	阳性预测值	阴性预测值	95%CI
临床模型训练集	0.881	0.758	0.536	0.989	0.981	0.672	0.835～0.925
临床模型测试集	0.704	0.511	0.208	0.826	0.556	0.500	0.553～0.854
影像组学模型训练集	0.830	0.695	0.412	0.989	0.976	0.617	0.776～0.883
影像组学模型测试集	0.819	0.681	0.458	0.913	0.846	0.618	0.701～0.934
联合模型训练集	0.939	0.858	0.773	0.937	0.937	0.800	0.908～0.970
联合模型测试集	0.854	0.809	0.833	0.783	0.800	0.818	0.743～0.964

本研究使用多因素Logistic回归分析中差异有统计学意义的指标：微钙化、垂直位、年龄与影像组学评分联合建立影像－临床联合模型，该模型表现出良好的预测效能，其训练集AUC为0.939(95%CI：0.908～0.970)，正确率为0.858，敏感度为0.773，特异度为0.937；测试集AUC为0.854(95%CI：0.743～0.964)，正确率为0.809，敏感度为0.833，特异度为0.783，见表4。同时，为了联合模型的可视化，本研究建立了列线图，见图3。

三种模型在训练集和验证集中诊断效能比较见图4、图5及表5。在训练集及测试集中，联合诊断模型AUC均高于单独的临床预测模型或影像组学模型，影像组学模型优于临床预测模型。决策曲线图显示，在三种预测模型中，联合模型在较大的阈值范围内比单一的临床预测模型或者影像组学模型能更好的预测甲状腺结节良恶性，这表明联合模型具有较好的临床应用价值，见图6。校准曲线显示联合模型在训练集与测试集中与理想模型的拟和度较好，见图7。

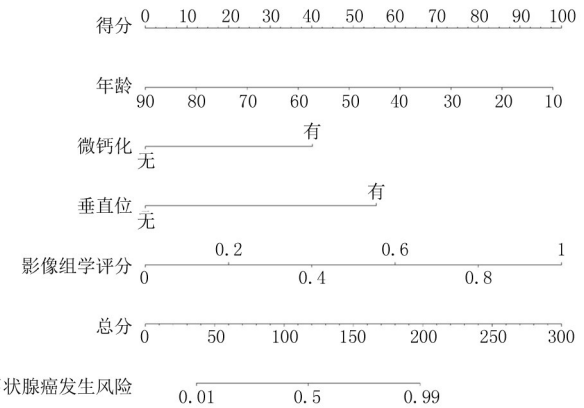


图3　联合模型鉴别甲状腺良恶性结节的列线图

注：每个变量的风险概率根据第一个尺度“点”定量确定每一个变量的评分，将各项变量得分相加后得到一个预测变量的总分，由总得分再对应相应的线性预测值以及恶性风险程度

Figure 3　Nomogram of combined model for differentiating benign and malignant thyroid nodules

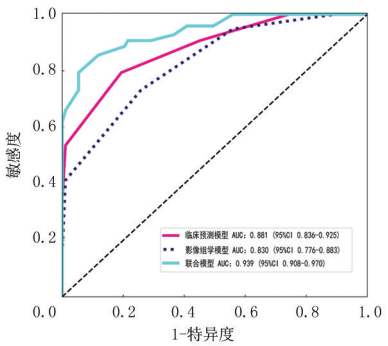


图4　三种预测模型训练集的ROC曲线

Figure 4　The ROC curves of the training sets of the three

prediction models

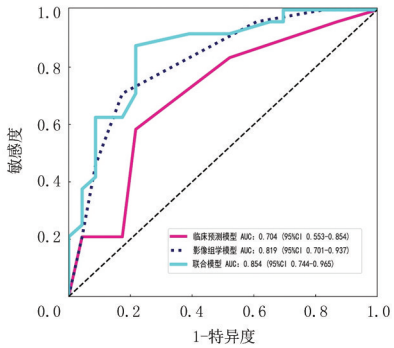


图5　三种预测模型测试集的ROC曲线

Figure 5　The ROC curves of the test sets of the three prediction models

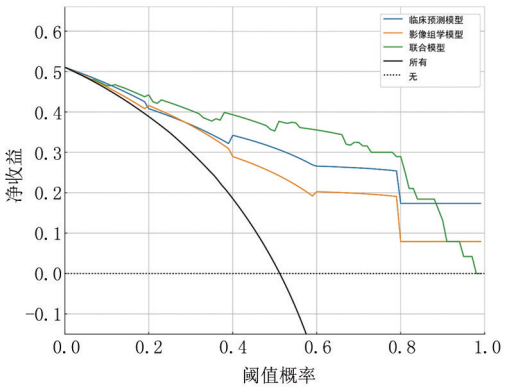


图6　三种模型预测甲状腺结节良恶性的决策曲线图

Figure 6　The decision curve of three models for predicting benign and malignant thyroid nodules

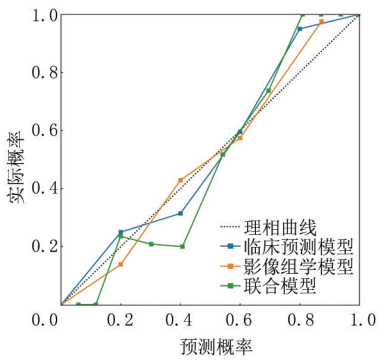


图7　三种模型预测甲状腺结节良恶性的校准曲线

Figure 7　The calibration curves of three models for predicting benign and malignant thyroid nodules

3　讨　论

术前及时并准确诊断甲状腺结节的良恶性对患者的治疗及其预后尤为重要，超声造影检查是当前诊断甲状腺结节病理性质的重要方法，但其依赖于个人操作手法及临床诊断经验，这可能为临床诊疗决策带来不确定性，从而导致甲状腺结节的过度诊断和过度治疗。影像组学通过提取超声造影图像的海量特征，筛选出关键特征，建立模型帮助诊断医师提高准确率。本研究的意义正是探索超声影像组学在鉴别甲状腺结节良恶性方面的价值。

本研究中将甲状腺结节患者的临床参数、C-TIRADS指南分析出的超声特征纳入Logistic回归分析，显示年龄是甲状腺结节的独立临床危险因素，恶性结节患者多为中青年患者，这与既往研究一致[15－16]。究其原因可能与随着年龄增长甲状腺功能逐渐减低，甲状腺其对碘的调节能力下降，TG2分泌下降，部分甲状腺组织细胞出现纤维化或退行性改变，从而容易造成甲状腺结节的形成，因此，对于中青年女性的甲状腺结节临床诊断需更加谨慎。本研究多因素Logistic回归分析显示，甲状腺结节声像图微钙化、垂直位是甲状腺恶性结节风险增加的独立危险因素。其原因可能为甲状腺癌生长迅速，乳头尖端发生局灶性、进行性梗死，引起甲状腺癌的血管及纤维组织增生、钙盐沉积于死亡细胞而形成，这与既往研究相似[17]。除此之外，既往研究中认为甲状腺激素水平与甲状腺癌相关[18－19]，但目前仍未形成统一定论，本研究中经多因素Logistic回归分析后，认为甲状腺激素水平不是甲状腺癌的独立危险因素，但甲状腺激素与甲状腺癌的相关性仍需进一步验证。

超声影像组学具有良好的客观分析和量化功能，通过标准化的流程可以获得稳定声学特征，这些特征肉眼无法显示，从而客观的量化甲状腺结节的病理性质，提高诊断准确性。本研究使用Pyrdiomics工具包进行特征提取，提取特征的图像类型包括原始图像、小波滤波、高斯滤波等，这提高了研究结果的可重复性[20－22]，最终用于建立模型的特征都是基于原始图像所提取到的特征，表明原始图像类型在预测甲状腺结节良恶性方面具有较大的潜能。目前大多数关于甲状腺超声造影的文献表明[23－25]，结节开始增强时间晚于正常甲状腺实质、不均匀低增强、结节消退早于周围正常甲状腺实质实质是恶性结节的表现；结节早于周围实质增强、无增强、等增强或高增强、均匀增强、周围环状增强是良性结节的表现。本研究最终用于建立影像组学模型的一阶特征（峰度）、纹理特征[相关性信息度量1、灰度非均匀性、尺寸区域非均匀性、测量图像中大小区域体积的可变性]反映了图像内部的不均匀的强度可能与甲状腺结节的良恶性有关，这与超声造影检查通过评估甲状腺结节中血管灌注和血流动力学的强度，从而提供结节的实时血流特征去鉴别甲

状腺结节的良恶性是一致的，同时也说明甲状腺结节在超声造影过程中表现出不同的血流微灌注与血流动力学的强度对于鉴别甲状腺结节的良恶性具有巨大的潜力。

本研究采用超声造影图像提取特征并建立影像组学模型，其对甲状腺结节良恶性预测具有较好的性能，其在训练集及验证集 AUC 分别为 0.830、0.819，诊断效能优于既往基于二维超声图像特征所建立的影像组学模型[26]，这可能是因为超声造影能反映病变组织各个区域血流灌注情况，相较于二维超声提供了更丰富的影像诊断信息，这表明基于超声造影图像的影像组学能够显示隐藏在超声造影图像中的大量数据，对甲状腺结节病理性质具有较好的预测效能。

本研究将临床参数、超声特征与影像组学特征结合建立联合模型并以列线图方式呈现，联合模型（训练集 AUC 为 0.939，测试集 AUC 为 0.854）比影像组学模型（训练集 AUC 为 0.830、验证集 AUC 为 0.819）或临床模型（训练集 AUC 为 0.881，测试集 AUC 为 0.704）具有更好的鉴别诊断效能，这与既往研究结果一致[27-28]，这说明综合了多个模态信息的列线图具有更高的预测效能，并且列线图能将模型可视化，对于往后的研究或临床应用具有更强的实用性。本研究决策曲线分析显示，相对于其他单一模型，联合模型的在较宽阈值概率范围临床净收益最高，表明联合模型在预测甲状腺结节良恶性中具有显著的临床应用价值。

本研究的局限性：①回顾性收集数据，难免存在潜在的选择偏倚；②本研究数据来源于单中心，未来需要前瞻性、多中心研究进一步验证；③使用单一超声造影图像分析，未涉及多模态超声图像分析。

综上所述，基于超声造影的影像组学模型在鉴别甲状腺结节良恶性方面具有较高的诊断效能，超声影像组学与临床参数、超声特征结合所建立的联合模型对甲状腺结节良恶性具有较好的鉴别能力，其鉴别能力优于单一临床预测模型或超声影像组学模型，可以辅助临床诊疗决策。

[参考文献]

[1] Durante C, Grani G, Lamartina L, et al. The diagnosis and management of thyroid nodules[J]. JAMA, 2018, 319(9): 914-924.

[2] 郑斌，高峰，韩若凌.甲状腺结节超声恶性风险评估相关国内外指南的对比与解读[J].河北医科大学学报，2019，40(8)：872-876.

[3] Grani G, Sponziello M, Filetti S, et al. Thyroid nodules: diagnosis and management[J]. Nat Rev Endocrinol, 2024, 20(12): 715-728.

[4] Ruan J, Xu X, Cai Y, et al. A Practical CEUS thyroid reporting system for thyroid nodules[J]. Radiology, 2022, 305(1): 149-159.

[5] Xiao F, Li JM, Han ZY, et al. Multimodality US versus thyroid imaging reporting and data system criteria in recommending fine-needle aspiration of thyroid nodules[J]. Radiology, 2023, 307(5): e221408.

[6] 邵春晖，李培英，罗永科，等.应变弹性成像参数与超声造影参数诊断甲状腺结节的价值及相关性分析[J].中国临床实用医学，2024，15(2)：1-7.

[7] Cao Y, Zhong X, Diao W, et al. Radiomics in differentiated thyroid cancer and nodules: explorations, application, and limitations[J]. Cancers (Basel), 2021, 13(10): 2436.

[8] Sorrenti S, Dolcetti V, Radzina M, et al. Artificial intelligence for thyroid nodule characterization: Where are we standing?[J]. Cancers (Basel), 2022, 14(14): 3357.

[9] Horvat N, Papanikolaou N, Koh DM. Radiomics beyond the hype: a critical evaluation toward oncologic clinical use[J]. Radiol Artif Intell, 2024, 6(4): e230437.

[10] Chang L, Liu J, Zhu J, et al. Advancing precision medicine: the transformative role of artificial intelligence in immunogenomics, radiomics, and pathomics for biomarker discovery and immunotherapy optimization[J]. Cancer Biol Med, 2025, 22(1): 33-47.

[11] Xiao F, Li JM, Han ZY, et al. Multimodality US versus thyroid imaging reporting and data system criteria in recommending fine-needle aspiration of thyroid nodules[J]. Radiology, 2023, 307(5): e221408.

[12] Li HJ, Sui GQ, Teng DK, et al. Incorporation of CEUS and SWE parameters into a multivariate Logistic regression model for the differential diagnosis of benign and malignant TI-RADS 4 thyroid nodules[J]. Endocrine, 2024, 83(3): 691-699.

[13] Yang W, Zhou J, Yue C, et al. Clinical value of contrast-enhanced ultrasound quantitative analysis for differentiating thyroid lesions in Hashimoto's thyroiditis patients[J]. Quant Imaging Med Surg, 2024, 14(1): 944-957.

[14] 中华医学会超声医学分会浅表器官和血管学组，中国甲状腺与乳腺超声人工智能联盟.2020 甲状腺结节超声恶性危险分层中国指南：C-TIRADS[J].中华超声影像学杂志，2021，30(3)：185-200.

[15] 侯迎迎，陈策，都晓英，等.基于 C-TIRADS 指南的阳性超声征象对甲状腺结节恶性风险的预测价值[J].中华解剖与临床杂志，2023，28(12)：782-786.

[16] 干露，张超学，邓大同，等.合肥地区 58094 例体检者甲状腺结节流行病学及患病特征分析[J/CD].中华医学超声杂志（电子版），2020，17(2)：164-168.

[17] 宋青，罗渝昆，田晓琦，等.常规超声联合超声造影诊断部分囊性甲状腺结节的 Logistic 回归分析[J/CD].中华医学超声杂志（电子版），2020，17(11)：1121-1124.

[18] 李晓会，常莹，郭颖.弹性成像技术联合甲状腺激素对甲状腺结节的诊断价值[J].临床与病理杂志，2020，40(12)：3127-3133.

[19] Hu MJ，Zhang C，Liang L，et al. Fasting serum glucose，thyroid-stimulating hormone，and thyroid hormones and risk of papillary thyroid cancer：A case-control study[J]. Head Neck，2019，41(7)：2277-2284.

[20] 张斯佳，祁佩红，杨新焕，等.基于增强门静脉期CT提取影像组学特征构建Nomogram评估肝纤维化分级[J].临床放射学杂志，2021，40(12)：2319-2325.

[21] Santinha J，Pinto DSD，Laqua F，et al. ESR Essentials：radiomics-practice recommendations by the European Society of Medical Imaging Informatics[J]. Eur Radiol，2025，35(3)：1122-1132.

[22] Ren JY，Lv WZ，Wang L，et al. Dual-modal radiomics nomogram based on contrast-enhanced ultrasound to improve differential diagnostic accuracy and reduce unnecessary biopsy rate in ACR TI-RADS 4－5 thyroid nodules[J]. Cancer Imaging，2024，24(1)：17.

[23] Zhu T，Chen J，Zhou Z，et al. Differentiation of thyroid nodules（C-TIRADS 4）by combining contrast-enhanced ultrasound diagnosis model with Chinese thyroid imaging reporting and data system[J]. Front Oncol，2022，12(30)：840819.

[24] Cheng H，Zhuo SS，Rong X，et al. Value of contrast-enhanced ultrasound in adjusting the classification of chinese-TIRADS 4 nodules[J]. Int J Endocrinol，2022，2022(6)：5623919.

[25] Cao H，Fan Q，Zhuo S，et al. The value of chinese thyroid imaging report and data system combined with contrast-enhanced ultrasound scoring in differential diagnosis of benign and malignant thyroid nodules[J]. J Ultrasound Med，2022，41(7)：1753-1761.

[26] Luo P，Fang Z，Zhang P，et al. Radiomics score combined with ACR TI-RADS in discriminating benign and malignant thyroid nodules based on ultrasound images：a retrospective study[J]. Diagnostics（Basel），2021，11(6)：1011.

[27] Lu W，Zhang D，Zhang Y，et al. Ultrasound radiomics nomogram to diagnose sub-centimeter thyroid nodules based on ACR TI-RADS[J]. Cancers（Basel），2022，14(19)：4826.

[28] Park VY，Lee E，Lee HS，et al. Combining radiomics with ultrasound-based risk stratification systems for thyroid nodules：an approach for improving performance[J]. Eur Radiol，2021，31(4)：2405-2413.

（本文编辑：何祯）