

基于 DEC-MRI 联合深度迁移学习和影像组学构建列线图模型预测浸润性乳腺癌组织学分级:一项多中心回顾性研究

徐静雅, 王国荣, 吴树剑, 李秀辉, 魏逸, 张虎, 翟建

【摘要】 目的:探讨基于 DCE-MRI 第二期(DCE2)图像的深度迁移学习(DTL)特征联合影像组学特征和临床-常规影像特征构建的列线图模型在预测浸润性乳腺癌组织学分级中的临床价值。**方法:**回顾性搜集 2019 年 1 月—2023 年 12 月在皖南医学院弋矶山医院(中心 A, n=356)和临汾市中心医院(中心 B, n=176)经手术病理学检查确诊的乳腺癌患者的临床和 DCE-MRI 资料。其中,低级别(I~II 级)368 例,高级别(III 级)164 例。在 DCE2 图像上手动分割获得全肿瘤容积 ROI,使用 PyRadiomics 软件提取影像组学特征,并使用预训练过的 ResNet-50 模型基于 DTL 方法提取肿瘤的特征。随后采用 Pearson 相关性分析和最小绝对收缩和选择算子(LASSO)回归分析进行特征降维,保留最具预测价值的特征。使用逻辑回归分析分别构建临床模型(基于临床指标和常规影像特征)、影像组学评分(Rad-score)、DTL 评分(DTL-score)以及列线图模型(临床独立危险因素+Rad-score+DTL-score)。采用 ROC 曲线评估各模型的预测效能,并采用 DeLong 检验比较各模型 AUC 的差异。采过校准曲线、决策曲线分析(DCA)和临床影响曲线(CIC)分别评估各模型的拟合优度、临床效益和临床适用性。**结果:**多因素 Logistic 回归分析结果显示,肿瘤直径(OR=1.046, 95%CI:1.014~1.080; P=0.005)、毛刺征(OR=0.415, 95%CI:0.219~0.755; P=0.005)及 MRI 提示腋窝淋巴结转移(OR=2.449, 95%CI:1.169~5.189, P=0.018)三个变量是肿瘤高病理分级的临床独立危险因素。经特征筛选后最终保留 17 个最优影像组学特征和 20 个最优 DTL 特征,并基于这些非零系数特征分别构建 Rad-score 和 DTL-score。列线图模型对浸润性乳腺癌组织学分级的预测效能,在训练集、验证集和外部测试集中分别为 0.907、0.919 和 0.898,诊断效能优于临床模型、Rad-score 及 DTL-score。校准曲线显示列线图的预测值与实际观察值之间的一致性较好。DCA 和 CIC 分析结果表明列线图模型在大部分阈值概率下具有较好的临床获益。**结论:**基于 Rad-score、DTL-score 及独立临床危险因素构建的列线图模型能在术前较准确地预测浸润性乳腺癌的组织学分级,具有一定的临床应用潜力。

【关键词】 乳腺肿瘤;影像组学;深度迁移学习;列线图;动态对比增强;磁共振成像;组织学分级

【中图分类号】 R445.2;R737.9 **【文献标志码】** A **【文章编号】** 1000-0313(2026)04-0420-09

DOI:10.13609/j.cnki.1000-0313.2026.04.011

开放科学(资源服务)标识码(OSID):



A nomogram model for predicting the histological grade of invasive breast cancer based on deep transfer learning and radiomics; a multicenter retrospective study XU Jing-ya, WANG Guo-rong, WU Shu-jian, et al. Department of Radiology, the First Affiliated Hospital of Wannan Medical College (Yijishan Hospital of Wannan Medical College), Anhui 241000, China

【Abstract】 Objective: To investigate the clinical value of a nomogram integrated by deep transfer learning (DTL) features extracted from second-phase DCE-MRI (DCE2), radiomics features, and clinicoradiological features for predicting the histological grade of invasive breast cancer. **Methods:** Clinical and imaging data of breast cancer patients with diagnoses confirmed by surgical pathology were retrospectively collected from Yijishan Hospital of Wannan Medical College (Center A, n=356) and Linfen Central Hospital (Center B, n=176) between January 2019 and December 2023. Among the 532 patients, 368 were classified as low histological grade (grades I~II), whereas 164 were classified as high histological grade (grade III). DCE2 images were analyzed, and the tumors were manually segmented to generate regions of interest (ROIs). Subsequently, radiomics features were extracted using PyRadiomics, and DTL features were extracted by applying transfer learning based on a pre-trained ResNet-50 model. Feature dimensionality reduction was performed using Pearson correlation analysis and least absolute shrinkage and selection operator (LASSO) regression to select the most predictive features. Using logistic regression analysis, a clinical model (based on clinical parameters and conventional imaging features), a radiomics score (Rad-score), a deep transfer learning score (DTL-score), and a nomogram model integrating independent clinical risk factors, Rad-score, and DTL-score were constructed. The predictive performance of each model was assessed by receiver operating characteristic (ROC) curve analysis, and differences in area under the curve (AUC) values among the models were compared using the DeLong test. The goodness of fit of each model was assessed through the calibration curve,

作者单位: 241000 安徽芜湖, 皖南医学院第一附属医院(皖南医学院弋矶山医院)放射科(徐静雅、吴树剑、李秀辉、魏逸、翟建); 041000 山西临汾, 临汾市中心医院放射科(王国荣); 241000 安徽芜湖, 芜湖市第二人民医院放射科(张虎)

作者简介: 徐静雅(1991-), 女, 安徽阜阳人, 硕士研究生, 主治医师, 主要从事基于人工智能的乳腺癌 MRI 诊断。

通讯作者: 翟建, E-mail: zhajjian@yjsyy.com

基金项目: 皖南医学院 2024 年度校中青年科研基金资助项目(WK2024ZQNZ57)

and clinical applicability was evaluated using decision curve analysis (DCA) and clinical impact curve (CIC). **Results:** Multivariate logistic regression analysis demonstrated that tumor diameter (OR=1.046, 95% CI: 1.014~1.080; $P=0.005$), spiculation (OR=0.415, 95% CI: 0.219~0.755; $P=0.005$) and MRI-axillary lymph node metastasis (MRI-ALNs, OR=2.449, 95% CI: 1.169~5.189; $P=0.018$) were independent clinical risk factors for high histological grade. Following feature selection, 17 optimal radiomics features and 20 optimal DTL features were retained. The Rad-score and DTL-score were subsequently constructed based on these non-zero coefficient features, respectively. The nomogram model demonstrated good predictive performance for the histological grading of invasive breast cancer, with AUCs of 0.907, 0.919, and 0.898 in the training, validation, and external test sets, respectively. Moreover, its diagnostic performance was superior to that of the clinical model, Rad-score, and DTL-score. The calibration curve showed that the predicted values of the nomogram model were consistent with the actual observed values. DCA and CIC indicated that the nomogram model provided good clinical benefits. **Conclusion:** The nomogram model incorporating the Rad-score, DTL-score, and independent clinical risk factors can accurately predict the histological grade of invasive breast cancer, thereby demonstrating promising clinical utility.

【Key words】 Breast neoplasms; Radiomics; Deep transfer learning; Nomogram; Dynamic contrast enhanced; Magnetic resonance imaging; Histological grade

乳腺癌(breast cancer, BC)已成为中国女性发病率第二高的恶性肿瘤,且是导致女性死亡的主要原因之一^[1]。乳腺癌的组织学分级在临床实践中是关键的预后因素,影响治疗方案的制订及临床预后^[2]。研究发现,乳腺癌的组织学分级不仅可作为独立预后预测指标,还与远处转移风险密切相关^[3-5]。尽管术前粗针穿刺活检是诊断乳腺癌组织学分级的主要手段,但其具有侵入性且取材有限。因此,开发一种能在术前无创性预测乳腺癌组织学分级的影像学方法具有重要的临床意义。

近年来,随着人工智能技术的发展,尤其是影像组学(radiomics)和深度学习(deep learning, DL)技术的开发和应用,为乳腺癌的诊断和治疗带来了新的机遇^[6-7]。影像组学通过高通量提取影像信息中病变的定量特征,为疾病的诊断和预后评估提供了新的视角。已有研究表明,基于 MRI 的影像组学可用于预测乳腺癌淋巴管侵犯、新辅助化疗反应、乳腺癌亚型及 Ki-67 表达水平等^[8-10]。DL 在医学影像分析中展现了显著优势,能够自动学习数据中的特征,减少对手工特征提取的依赖,从而能更全面地捕捉影像中的复杂信息^[11-12]。如,有研究中利用基于 DL 技术的 ResNeXt50 模型分析 5 种模态的 MRI 图像并构建预测模型对乳腺癌分子亚型进行预测,结果显示:5 种 DL 模型的 AUC 在 0.859~0.912 区间内,明显优于传统模型^[13]。然而,深度学习技术在乳腺癌组织学分级中的应用仍较为少见,尤其是在 MRI 方面的应用。已有研究者在乳腺癌组织学分级预测方面进行了初步探索,如郭亚欣等^[14]基于单中心 DWI 数据构建的瘤内-瘤周影像组学模型预测乳腺癌组织学分级,在测试集中 AUC 范围为 0.723~0.782;吴佩琪等^[15]虽引入 ADC 和 DCE-MRI 双模态数据,但样本量仅 71 例,所构建的 DL 模型的 AUC 最大值为 0.858 且稳定性受

限(同一队列 DCE-MRI 单模态的 AUC 仅 0.691)。相比之下,本研究纳入双中心患者数据,基于信息丰度更高的 DCE-MRI 第二期图像(DCE2)提取深度迁移学习特征和肿瘤区域放射组学特征,结合关键临床预测因子,构建可解释列线图模型,旨在探索一种能在术前精准预测浸润性乳腺癌组织学分级的影像学标志物,为个体化治疗决策提供直观和可靠的量化依据。

材料与方法

1. 研究对象

回顾性将 2019 年 1 月—2023 年 12 月皖南医学院第一附属医院(中心 1)和临汾市中心医院(中心 2)经术后病理证实的乳腺癌患者纳入本研究。根据诺丁汉组织学分级,Ⅰ级和Ⅱ级合并为低级别组,Ⅲ级归为高级别组,进行二分类分析。病例纳入标准:①MRI 检查前未行穿刺活检或新辅助治疗;②术前 14 天内完成常规 MRI 及 DCE-MRI 检查;③明确的术后病理诊断并获得肿瘤组织学分级。排除标准:①既往有其它恶性肿瘤史,或为乳腺癌复发;②MRI 图像质量差,无法满足分析要求(如明显运动伪影或肿瘤边界无法勾画);③临床或术后病理资料不全;④男性乳腺癌。最终中心 1 有 356 例、中心 2 有 176 例乳腺癌患者被纳入本研究。将中心 1 的患者按 8:2 的比例划分为训练集(284 例)和验证集(72 例),以中心 2 作为外部测试集(176 例)。

本研究已通过医院伦理委员会审批(202473);系回顾性研究,因此豁免患者知情同意。

2. MRI 扫描方法及影像特征分析

使用 Siemens Avanto 1.5T(中心 1)和 Signa Creator 1.5T(中心 2)磁共振扫描仪及 4 通道乳腺专用相控阵线圈。患者取俯卧位,双侧乳腺自然下垂于线圈内,在矢状位上定位线平行于乳腺的长轴,在横轴

位上调整视野定位中心于乳腺的中心,扫描范围包括整个乳腺及双侧腋下,需包括整个病变范围,常规行横轴面 T_1 WI 和矢状面 T_2 WI 脂肪抑制序列扫描。扫描参数:TR 5.05 ms,TE 1.94 ms,TI 16 ms,层厚 1.5 mm,层间距 0 mm,视野 340 mm \times 340 mm,矩阵 348 \times 348。增强扫描前使用双筒高压注射器经前臂静脉注射对比剂钆双胺,注射流率 2.5 mL/s,剂量 0.1 mmol/kg,对比剂注射完成后即行 LAVA T_1 WI 扫描,扫描参数同上,共扫描 9 期,每期扫描时间 60 s,总扫描时间 540 s。

3. 临床模型的构建

本研究纳入分析的临床和常规影像学特征:年龄、肿瘤最长径、位置(左乳/右乳)、形状(规则/不规则)、边界(清晰/不清晰)、毛刺征(有/无)、强化类型(均匀/不均匀)、瘤周水肿(有/无)、TIC 类型(I~III型)、周围血管征(有/无)及边缘强化(有/无)。淋巴结状态的判定标准:若淋巴结短径 ≥ 1.0 cm,或出现淋巴门消失、移位和皮质偏心性增厚中的任一征象,即判为阳性^[16]。由两位放射科医师(8 和 10 年放射诊断经验)在盲法下独立判读,意见存在分歧时则由另一位高级职称医师(12 年放射诊断经验)进行仲裁。连续变量按原值纳入,分类变量按预设分层进行编码。上述指标经单因素和多因素二元 Logistic 回归分析(逐步回归法)确定独立临床预测因子,并构建临床预测模型。

4. 病灶分割及特征提取

将所有患者的 MRI 图像以 DICOM 格式自 PACS 工作站导入 ITK-SNAP 3.8.0 软件,以第 2 期增强(DCE2)图像作为分析序列。对训练集中所有患者的图像进行图像分割和特征提取,验证集和外部测试集仅用于效能评估,不参与任何参数拟合、特征筛选或模型训练。

首先,对图像进行预处理,包括 N4 偏置场校正、重采样(1 mm \times 1 mm \times 1 mm)和灰度归一化,以确保数据的一致性和可比性。在 DCE2 图像上,先由一位具有 8 年诊断经验的放射科医师独立完成 2D-ROI 的勾画(逐层沿肿瘤边界勾画),最后由软件生成三维感兴趣区(volume of interest,VOI)。随后,由一位具有 15 年经验的放射科医师对全部分割结果逐一审核,存在分歧时经讨论达成一致后形成最终 VOI。随后,使用 Pyradiomics 3.0 软件自动提取每个 VOI 的影像组学特征,共提取了 1864 个组学特征:形态特征(Shape) 14 个;一阶统计特征(First Order)360 个;灰度共生矩阵(gray level co-occurrence matrix,GLCM)特征 440 个;灰度依赖矩阵(gray level dependence matrix, GLDM)特征 280 个;灰度游程长度矩阵(gray level run length matrix, GLRLM)特征 320 个;灰度区域大

小矩阵(gray level size zone matrix, GLSZM)特征 320 个;邻域灰度差异矩阵(neighboring gray tone difference matrix, NGTDM)特征 100 个。

为评估病变分割后提取的组学特征的可重复性,采用组内相关系数(intraclass correlation coefficient, ICC)对观察者内一致性进行评价,具体步骤如下。首先,由医师 A(8 年乳腺影像诊断经验)在完成全部乳腺病变初次分割后,间隔 1 周随机选取 50 例病变,按照原方法进行二次分割,以评估观察者内一致性(ICC_intra)。整个分割过程在盲法下进行,医师均不知晓患者的临床病理信息,以避免主观偏倚。最终,仅保留 ICC_intra 大于 0.75 的组学特征,确保所选特征具有良好的可重复性,用于后续特征筛选和建模分析。

在 DCE2 序列上基于每个病灶的 VOI,软件可自动定位肿瘤 VOI 最大横截面所在层面并提取该层面的 2D-ROI,以最小外接矩形截取图像块后等比缩放/零填充至 224 \times 224 大小的矩阵中。本研究利用 Onekey AI(20240616)科研平台,基于预训练的 ResNet-50 卷积神经网络作为深度迁移学习(deep transfer learning, DTL)特征的基础计算模型。ResNet-50 由 4 个层级(layer)组成:Layer1~Layer4 分别含 3、4、6 和 3 个瓶颈残差块,每个瓶颈块包含 1 \times 1、3 \times 3、1 \times 1 卷积并配备批归一化和修正线性单元(rectified linear unit, ReLU),通过跳跃连接实现残差学习。该模型在包含数百万张图像的大规模 ImageNet 数据集中进行了预训练,能够捕获丰富的图像特征。本研究中去除了 ResNet-50 的全连接层,仅保留卷积主干;将每例患者的 ROI 图像数据输入 ResNet-50 网络,在全局平均池化层提取 DTL 特征,共提取了 2048 个特征,用于后续特征筛选和建模分析。

5. 影像组学和 DTL 特征筛选及模型构建

对训练集中 ICC >0.75 的影像组学特征及 2048 维 DTL 特征进行 Z-score 归一化;然后以 Pearson 相关系数阈值 0.90 去除高度冗余特征后,采用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)回归分析(10 折交叉验证)确定惩罚参数并筛选非零系数特征;最后以筛选后的影像组学特征和 DTL 特征分别构建逻辑回归模型,计算相应的 Rad-score 和 DTL-score 值。

将 DCE2 序列的 Rad-score、DTL-score 及独立临床预测因子纳入多因素 Logistic 回归模型,构建联合模型,并绘制其列线图实现模型的可视化。在模型效能评估方面,校准曲线、分别评估列线图的拟合优度。

6. 统计学方法

使用 R4.4.2 和 SPSS 27.0 软件进行统计分析。连续变量在进行正态性检验后,符合正态分布者采用

独立样本 t 检验进行组间比较,不符合正态分布者则使用 Mann-Whitney U 检验。分类变量以频数和百分比形式描述,组间比较采用 χ^2 检验或 Fisher 精确概率法。利用 ROC 曲线分析评估模型在训练集、验证集和外部测试集中的判别能力;模型之间 AUC 的比较使用 DeLong 检验。模型拟合优度通过 Hosmer-Lemeshow(HL)检验进行评估,并绘制校准曲线以反映模型预测概率与实际结果之间的一致性,采用决策曲线分析(decision curve analysis, DCA)和临床影响曲线(clinical impact curve, CIC)分别评估模型的临床适用性及净获益。所有统计检验为双侧检验,以 $P < 0.05$ 为差异具有统计学意义。

结果

1. 临床资料分析

训练集中低级别和高级别乳腺癌组间临床资料的比较结果见表 1。MRI 评估的腋窝淋巴结状态(MRI-reported axillary lymph node status, MRI-ALNs)及毛刺征在两组间的差异有统计学意义($P < 0.05$)。

对训练集中的临床指标进行单因素和多因素 Logistic 回归分析,结果见表 2~3。单因素 Logistic 回归分析结果显示瘤周水肿、肿块最大直径、毛刺和 MRI-ALNs 是乳腺癌病理分级的影响因素。多因素 Logistic 回归分析结果显示肿瘤最大直径、边缘毛刺

及 MRI-ALNs 为独立临床预测因子;基于这 3 项指标构建临床模型,其在训练集、验证集和外部测试集中预测乳腺癌组织学分级的 AUC 分别为 0.710、0.725 和 0.774(表 4)。

表 2 训练集中临床资料的单因素逻辑回归分析

指标	单因素逻辑回归分析	
	OR(95%CI)	P 值
年龄	0.979(0.953~1.005)	0.122
肿瘤直径	1.055(1.024~1.089)	<0.001
位置	1.12(0.680~1.842)	0.656
形状	0.932(0.487~1.843)	0.836
边界	0.938(0.561~1.579)	0.807
毛刺征	0.354(0.189~0.633)	<0.001
瘤周水肿	1.649(1.002~2.724)	<0.05
TIC 类型 [#]		
II 型	3.220(0.882~11.752)	0.077
III 型	3.363(0.957~11.823)	0.059
周围血管征	1.247(0.753~2.083)	0.359
病灶强化方式	0.944(0.323~2.477)	0.911
MRI-ALNs	2.997(1.475~6.183)	<0.001
BI-RADS 分类 [*]		
4c 类	0.809(0.392~1.669)	0.566
4b 类	0.407(0.156~1.064)	0.407
4a 类	0.391(0.108~1.412)	0.391
3 类	—	0.999
病理组织学类型	0.818(0.116~3.873)	0.812

注: [#] 以 I 型作为参考值; ^{*} 以 MRI-BI-RADS 5 类作为参考值。

2. 影像组学特征和 DTL 特征的筛选

基于 DCE2 序列从全瘤 VOI 提取了 1864 个影像组学特征,从肿瘤最大截面经预训练的 ResNet50 模

表 1 训练集中低级别组与高级别组临床资料的比较

特征	训练集		统计量	P 值
	低级别(n=191)	高级别(n=93)		
年龄/岁	54.21±9.96	52.30±9.20	1.557	0.121
肿瘤最长径/mm	18.00 (15.00, 23.00)	22.00 (18.00, 27.00)	-32.885	<0.001
ADC/ $\times 10^{-3}$ mm ² /s	0.89(0.77, 0.99)	0.90(0.79, 0.99)	-28.490	0.795
右乳/例	85(44.50%)	44(47.31%)	0.102	0.750
形态不规则/例	160(83.77%)	77(82.80%)	0.001	0.970
边界不清晰/例	124(64.92%)	59(63.44%)	0.013	0.910
毛刺征/例	74(38.74%)	17(18.28%)	11.107	0.001
不均匀强化/例	153(80.10%)	83(89.25%)	3.100	0.078
瘤周水肿/例	46(24.08%)	17(18.28%)	3.395	0.341
TIC 类型/例			3.980	0.137
I 型	19(9.95%)	3(3.23%)		
II 型	59(30.89%)	30(32.26%)		
III 型	113(59.16%)	60(64.52%)		
周围血管征/例	109(57.07%)	58(62.37%)	0.522	0.470
边缘无强化/例	69(36.13%)	40(43.01%)	0.979	0.322
MRI-ALNs 阳性/例	16(8.38%)	20(21.51%)	— [*]	0.003
MRI-BI-RADS/例			— [*]	0.191
3 类	2(1.05%)	0(0.00%)		
4a 类	15(7.85%)	4(4.30%)		
4b 类	36(18.85%)	10(10.75%)		
4c 类	116(60.73%)	64(68.82%)		
5 类	22(11.52%)	15(16.13%)		
病理类型/例			— [*]	1.000
浸润性导管癌	186(97.38%)	91(97.85%)		
其它 [#]	5(2.62%)	2(2.15%)		

注: ^{*} Fisher 检验; [#] 包括实体型乳头状癌、浸润性大汗腺癌等病理类型。

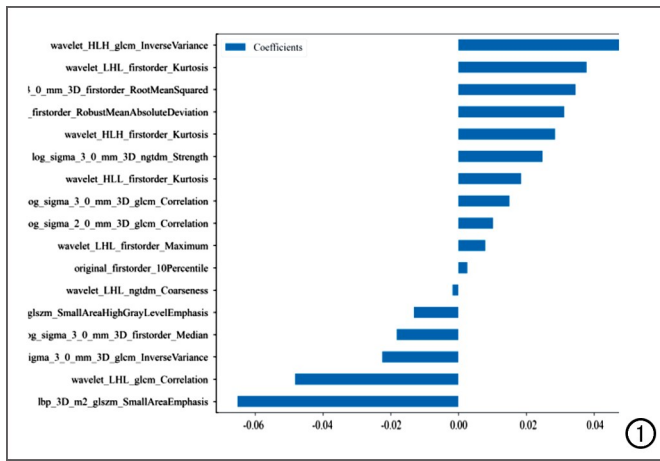


图1 通过 LASSO 回归筛选出最优影像组学特征及其权重系数。

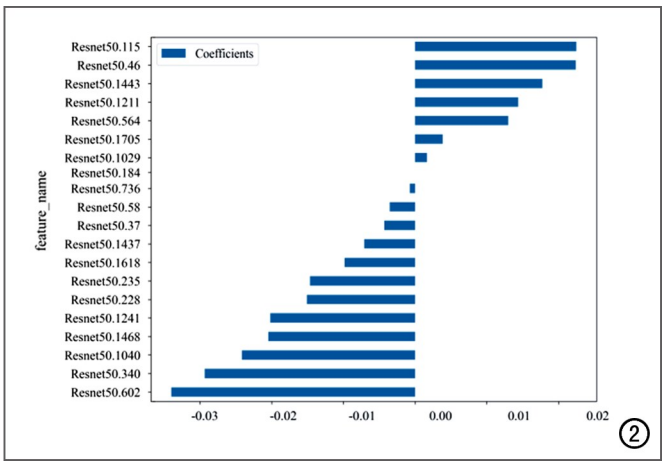


图2 通过 LASSO 回归筛选出最优 DTL 特征及其权重系数。

表3 训练集中临床资料的多因素逻辑回归分析

指标	多变量逻辑回归分析	
	OR(95%CI)	P 值
肿瘤直径	1.046(1.014~1.080)	0.005
毛刺征	0.415(0.219~0.755)	0.005
MRI-ALNs	2.449(1.169~5.189)	0.018

型中提取了 2048 个 DTL 特征。影像组学特征经一致性评估(ICC>0.75)保留了 1563 个稳定特征;随后在训练集中进行 Z-score 归一化后以 Pearson 相关系数阈值 0.90 去除冗余特征,并通过 LASSO 回归(10 折交叉验证)筛选非零系数特征,最终筛选出 17 个最佳影像组学特征(图 1)。DTL 特征不进行 ICC 评估,在训练集中完成 Z-score 归一化和 Pearson 相关系数法去冗余后,经 LASSO 回归筛选得到 20 个最优 DTL

特征用于建模(图 2)。

3.模型构建和评估

基于影像组学模型和 DTL 模型可计算相应的 Rad-score 和 DTL-score。对这 2 个指标预测乳腺癌病理学分级的效能进行 ROC 曲线分析,结果显示 Rad-score 在训练集、验证集和外部测试集中的 AUC 分别为 0.792、0.664 和 0.864;DTL-score 的 AUC 分别为 0.802、0.740 和 0.763。

联合 Rad-score、DTL-score 和 3 个临床独立影响因素(最大径线、毛刺征及 MRI-ALNs)构建联合模型并绘制其列线图(图 3)。在训练集、验证集及外部测试集中,联合模型预测乳腺癌组织学分级的效能指标见表 4。结果表明,联合模型在 3 个数据集中均展现

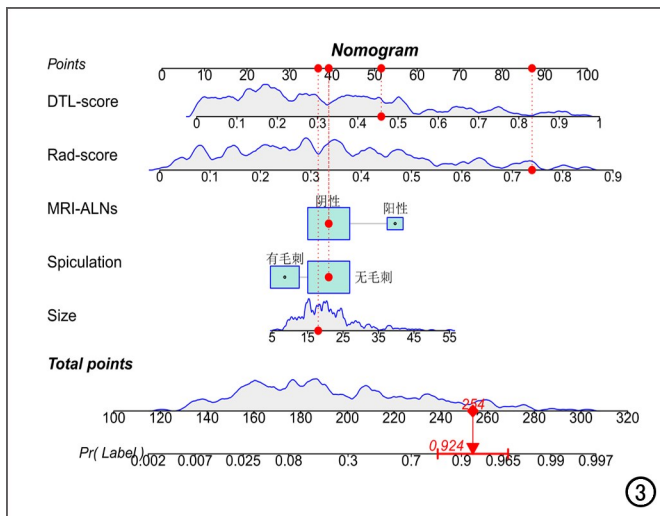
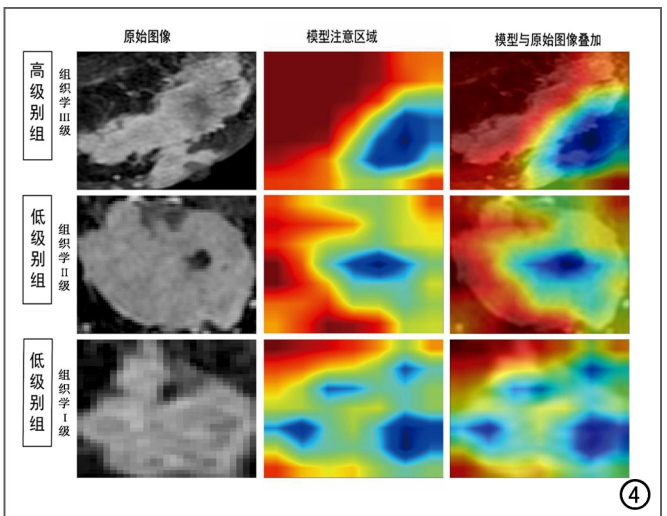


图3 基于临床-影像组学-DTL 特征构建的预测乳腺癌高病理分级的联合模型的列线图(含个体化示例)。示例患者对应取值以红点标注,总分约为 254,对应的预测高病理分级的概率值约为 0.924。图4 不同组织级别的深度学习模型可视化,关注区域与原始图像叠加对比。展示了随着组织学分级升高(I 到 III 级),模型注意力逐渐由分散转为集中,高级别肿瘤的关注区域更多集中在边界不规则区域,而低级别肿瘤则表现为关注区域分散,突出了高级别组织和低级别组织在模型决策中的差异。热力图颜色表示模型关注程度,蓝色区域为高关注区域,红/黄色区域为低关注区域。



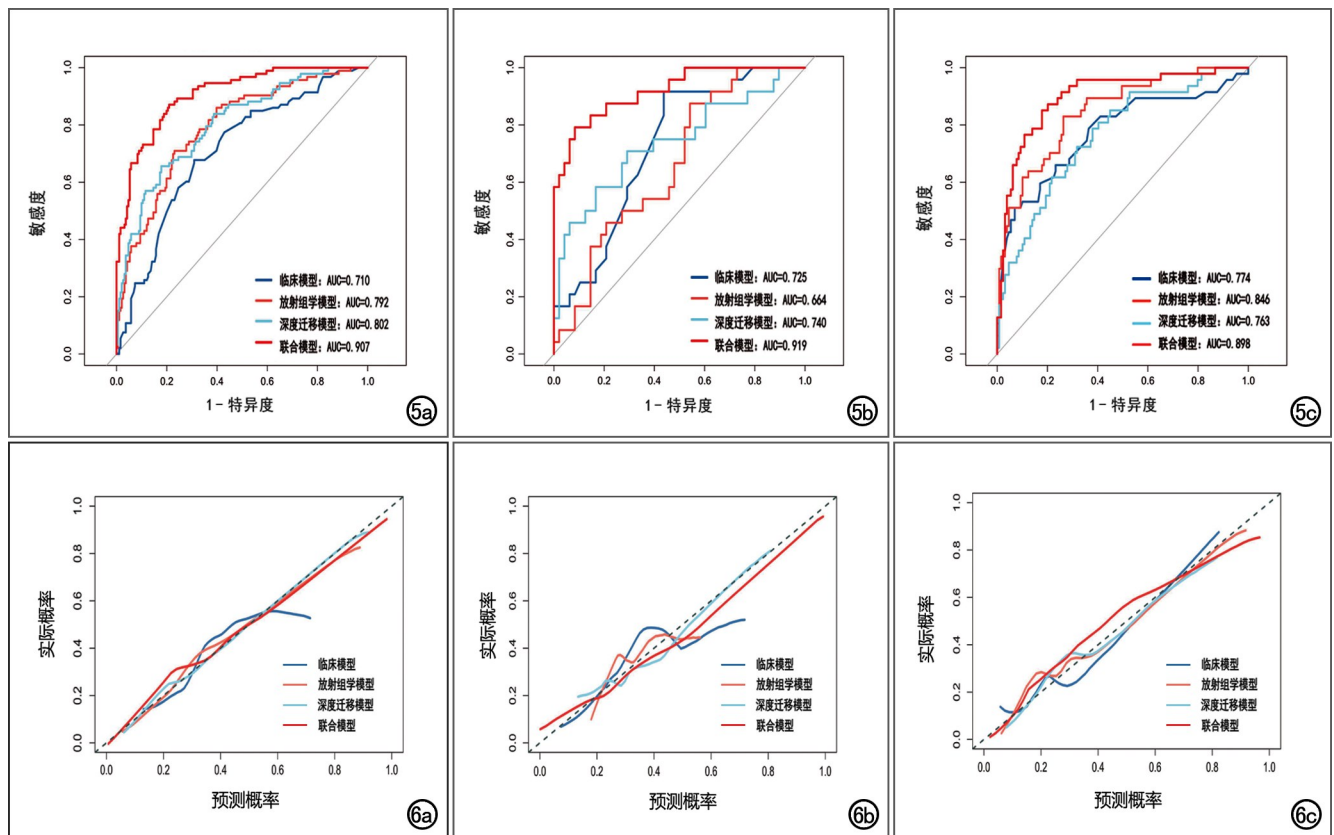


图5 各模型的ROC曲线,显示各数据集中均以列线图模型的AUC最大,提示其诊断效能最佳。a)训练集;b)验证集;c)测试集。图6 各模型的校准曲线,显示各数据集中均以联合模型的曲线更接近对角线(理想曲线),表明该模型的预测概率与实际观察概率的一致性更高,校准表现更佳。a)训练集;b)验证集;c)测试集。

出良好的预测效能,并在不同数据集中保持了相对稳定的性能。

为了理解模型的关注区域,本研究利用梯度加权类激活映射(gradient-weighted class activation mapping, Grad-CAM)可视化方法展示不同组织级别组中模型的决策依据,结果表明,高级别组图像中,模型能够集中关注关键特征区域;而在低级别组图像中,模型的关注区域则较为分散(图3)。

4.各模型效能比较及临床价值分析

本研究构建了4种乳腺癌病理学分级的预测模型:临床模型、Rad-score、DTL-score和联合模型,它们在3个数据集中的预测效能见表4和图5。在3个数据集中,联合模型的AUC均高于其它3种模型。DeLong检验结果显示,联合模型与其它3个模型间AUC的差异均具有统计学意义(表5),表明其判别能力更强。校准曲线显示联合模型在3个数据集中的预

表4 不同模型在训练集、验证集和外部测试集上的性能比较

Model	AUC(95%CI)	符合率	敏感度	特异度	PPV	NPV	F1分数
训练集							
临床模型	0.710(0.647~0.774)	0.687	0.691	0.677	0.815	0.516	0.748
放射组学模型	0.792(0.737~0.847)	0.750	0.770	0.710	0.845	0.600	0.805
深度迁移学习模型	0.802(0.747~0.856)	0.768	0.822	0.656	0.831	0.642	0.831
联合模型	0.907(0.872~0.942)	0.817	0.791	0.871	0.926	0.669	0.853
验证集							
临床模型	0.725(0.607~0.843)	0.681	0.562	0.917	0.931	0.512	0.701
放射组学模型	0.664(0.537,0.791)	0.597	0.458	0.875	0.880	0.447	0.603
深度迁移学习模型	0.740(0.608,0.871)	0.750	0.792	0.667	0.826	0.615	0.809
联合模型	0.919(0.849~0.979)	0.875	0.792	0.912	0.826	0.898	0.809
外部测试集							
临床模型	0.774(0.687~0.862)	0.818	0.930	0.511	0.839	0.727	0.839
放射组学模型	0.846(0.782~0.910)	0.761	0.736	0.830	0.922	0.534	0.819
深度迁移学习模型	0.763(0.684~0.843)	0.665	0.620	0.787	0.889	0.430	0.731
联合模型	0.898(0.844~0.953)	0.830	0.822	0.851	0.938	0.635	0.876

测概率与实际观察结果均拟合良好,HL 检验 P 值均大于 0.05(0.710, 0.468, 0.304),提示模型的校准性能良好(图 6)。DCA 结果显示,联合模型在阈值范围(0.20~0.80)内提供了最大的临床净获益(图 7)。CIC 进一步证实联合模型在较低阈值下可识别更多真阳性患者,具有临床实用性(图 8)。

讨论

本研究基于 DCE2 序列的放射组学特征与 Res-Net50 的 DTL 特征,结合独立临床预测因子,通过 Logistic 回归构建了基于临床、影像组学和 DTL 特征

的联合模型,用于术前评估乳腺癌组织学分级(I~III 级)。结果表明,联合模型的预测效能优于其它模型,在验证集和测试集中均显示出较好的预测效果。该模型相较于传统方法,能够显著提高乳腺癌组织学分级的预测精度,有望成为乳腺癌临床决策的新型生物标志物,为个性化治疗提供科学依据。

本研究中利用多中心样本构建了基于 DCE2 序列的联合模型用于预测浸润性乳腺癌的组织学分级,其在训练集、内部验证集和外部测试集中的 AUC 分别为 0.907、0.919 和 0.898,明显优于单一临床模型、Rad-score 和 DTL-score。匡静等^[17]研究开发的影像

表 5 在 3 个数据集中联合模型与其它 3 种模型 AUC 的比较

对比模型	AUC(95%CI)	Z 值	P 值
训练集			
联合模型 vs. 临床模型	0.196(0.134~0.259)	6.175	<0.001
联合模型 vs. 组学模型	0.114(0.068~0.161)	4.786	<0.001
联合模型 vs. DTL 模型	0.105(0.060~0.150)	4.562	<0.001
验证集			
联合模型 vs. 临床模型	0.194(0.082~0.307)	3.377	0.001
联合模型 vs. 组学模型	0.255(0.130~0.380)	4.001	<0.001
联合模型 vs. DTL 模型	0.180(0.068~0.292)	3.143	0.002
外部测试集			
联合模型 vs. 临床模型	0.124(0.052~0.195)	3.400	0.001
联合模型 vs. 组学模型	0.052(0.000~0.104)	1.977	0.048
联合模型 vs. DTL 模型	0.135(0.055~0.214)	3.317	0.001

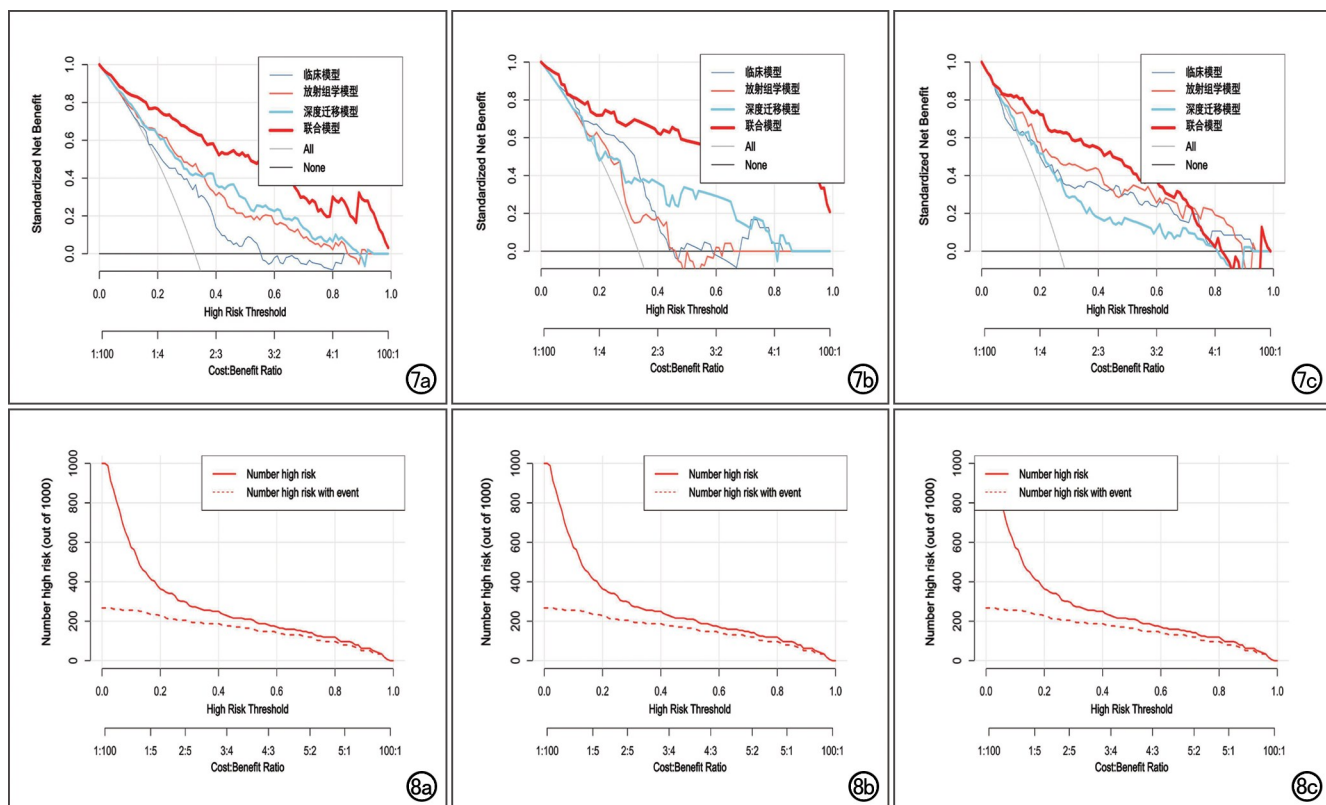


图 7 各模型的决策曲线,显示各数据集中均以联合模型在较宽阈值范围内获得更高净获益。a)训练集;b)验证集;c)测试集。

图 8 各模型的临床影响曲线,显示不同阈值下预测为高风险的人数及其中真实事件数。各数据集中均以联合模型的临床净获益更高,对高风险患者的识别表现更优。a)训练集;b)验证集;c)测试集。

征象-临床特征-影像组学特征联合模型预测 BC 组织学分级,其 AUC 在训练集和验证集中分别为 0.807 (95%CI: 0.723 ~ 0.891) 和 0.890 (95%CI: 0.795 ~ 0.984)。韩剑剑等^[18]利用数字化 X 线构建了术前预测 BC 组织学分级的放射组学模型,其在训练集、验证集和外部测试集中的 AUC 分别为 0.811、0.825 和 0.803。本研究表明,基于 DCE2 序列的放射组学模型在预测乳腺癌组织学分级方面与其它影像学方法相当,且首次引入深度迁移学习特征显著提升了模型性能。多中心样本数据分析显示,联合模型在验证和测试集中表现良好,具有较强的临床实用性和良好的泛化能力。DCA 结果表明,联合模型在组织学分级为 III 级的发病率阈值范围内具有较高的净临床收益。

影像组学能够从医学图像中提取肉眼无法识别的高通量特征,通过特定的算法得到定量信息,进行客观的影像学评估。在本研究中,筛选出了 17 个最优组学特征,其中有 9 个高阶纹理特征(如 log_sigma_2_0_mm_3D_glcmm_Correlation 和 lbp_3D_m2glszm_SmallAreaEmphasis)。灰度共生矩阵(GLCM)用于刻画相邻像素灰度强度之间的差异,从而反映图像的局部均匀性。该指标值越大,说明相邻像素灰度差异越明显,局部均匀性越差。本研究结果显示,组织学 III 级的 wavelet_HLH_glcmm_InverseVariance 特征值高于 I ~ II 级,提示浸润性乳腺癌分级越高,图像均匀性越差、纹理不均一性越强。在细胞层面,高级别的乳腺癌可能导致肿瘤细胞增殖活跃,肿瘤区域容易发生缺血、缺氧,导致肿瘤内部信号不均匀,导致灌注与组织结构在多尺度上的不均一,从而在影像层面表现为局部均匀性下降^[18-19]。邻域灰度差矩阵(neighborhood gray tone difference matrix, NGTDM)与灰度尺寸区域矩阵(gray level size zone matrix, GLSZM)分别从“局部灰度与其邻域平均值的差异”与“同灰度区域的尺寸与灰度分布”两条维度刻画纹理表型。前者的粗糙度(coarseness)值越大表示纹理越平滑,越小则更纹理越细腻且起伏更频繁;结构强度(strength)值越大提示条带/块状等组织化结构更显著,组织结构间的对比更强。后者的小区域强调(small area emphasis, SAE)值越大说明小尺寸区域占比更高、纹理更细碎。例如,研究发现高分级乳腺癌的 lbp_3D_m2glszm_SmallAreaEmphasis 特征值高于低级别 I ~ II 级, LBP-GLSZM 的 SAE 偏低通常代表细粒度斑块减少、粗大片化增强和多尺度异质性上升,符合高分级肿瘤的侵袭性与不均一生物学特征^[20-21]。在小波 LHL 子带的 NGTDM 粗糙度特征(wavelet_LHL_ngtdm_Coarseness),高级别(III 级)乳腺癌的取值低于低级别(I ~ II 级)。Coarseness 值降低提示局部起伏更频

繁、纹理更细而不均一,反映更强的微观异质性,与高级别肿瘤的侵袭性表型相一致。本研究表明,影像组学特征与乳腺癌组织学分级密切相关:高分级(III 级)病灶普遍呈现局部均匀性下降和纹理不均一增强、邻域起伏更显著、结构对比更突出、以及尺度构成由细碎转向粗大。上述影像表型与高级别肿瘤的生物基础(快速增殖、异常血管生成、缺血/坏死与间质重塑)相一致,支持影像组学特征有可能作为组织学分级的有效量化线索。

近年来,深度学习在医学影像分析,尤其是乳腺癌诊断和辅助治疗中取得了显著成果。Witowski 等^[22]使用 DL 模型提高乳腺癌诊断的准确性,其 AUC 为 0.920。在不同的数据集中,Chen 等^[23]的研究中构建的 CNN 模型在不同数据集中预测前哨淋巴结转移的 AUC 为 0.728 ~ 0.899。深度学习模型能够从医学图像中提取更加抽象高维度数据特征,丰富预测模型的信息输入,进一步提高列线图模型的预测准确性和稳定性。本研究利用 ResNet50 网络提取的深度迁移学习特征,结合其残差块结构有效整合全局特征信息,解决了深度网络训练困难的问题。本研究的结果与以往的研究一致^[24],表明多种影像特征的融合提升了预测准确性。在本研究中,基于临床-影像组学-DTL 特征的联合模型展现了更强的预测效果,验证了多特征联合在乳腺癌组织学分级预测中的重要性。此外,临床指标如肿瘤大小、边缘毛刺和 MRI-ALNs 的结合显示,较大肿瘤通常伴随较高的组织学分级,可能与肿瘤细胞增殖活跃相关^[20]。边缘毛刺在 I 级和 II 级乳腺癌中较为常见,这与梁园园等^[25]和韩剑剑等^[18]的研究结果一致,其病理机制是肿瘤细胞向周围纤维组织浸润,进而引发纤维组织的增生。组织学分级较低时,肿瘤的侵袭性较弱,导致周围纤维组织的增生程度大于其破坏程度^[18]。随着组织学分级的升高,肿瘤的侵袭性和淋巴结转移风险增加,这与众多研究结果一致^[18,25]。本研究首次基于双中心较大样本量构建联合模型,更全面地评估了模型对乳腺癌组织学分级的预测效能。该方法结合深度学习的特征提取优势与传统影像组学特征互补,为术前评估提供了可靠工具。未来研究将优化模型,探索特征组合对预测效能的影响,并推进临床应用。

本研究存在一定的局限性。首先,本研究为回顾性研究,样本量有限,可能存在一定的选择偏倚,泛化能力受限;未开展更多中心和跨设备的外部验证。其次,深度迁移学习流程仅基于肿瘤最大层面 2D-ROI 的信息进行分析和建模,未利用肿瘤全体积信息,可能低估或高估肿瘤内异质性,遗漏有价值的影像特征,影响模型的准确性和全面性。第三,病灶 ROI 勾画主要

依据 DCE2 单序列,未进行跨序列配准与多序列一致性标注;同时,未评估观察者间和观察者内一致性,测量不确定性和可重复性证据有限。第四,未整合钼靶与超声等多模态影像,限制了模型的综合判别与校准能力。第五,外部验证仅限单中心样本,跨区域迁移能力证据不足,同时缺乏前瞻性评估;模型的校准性能和临床净效益尚待进一步系统评估。有待在未来的研究中采用前瞻性设计,进行更系统的数据采集和验证,并探索多模态融合策略,以提升模型的鲁棒性和临床实用性。

综上所述,本研究构建的基于 DCE2 的影像组学特征和 ResNet50 深度迁移学习特征的列线图模型,在术前预测浸润性乳腺癌组织学分级方面表现出良好的预测效能。作为非侵入性和可重复高的工具,该模型有望为乳腺癌患者的个体化治疗和预后评估提供有力支持,具有重要的临床应用前景。

参考文献:

- [1] Zheng RS, Chen R, Han BF, et al. Cancer incidence and mortality in China, 2022[J]. *Zhonghua Zhong Liu Za Zhi*, 2024, 46(3): 221-231.
- [2] Rakha EA, Reis-Filho JS, Baehner F, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade [J/OL]. *Breast Cancer Res*, 2010, 12(4): e207 [2010-07-30]. DOI: 10.1186/bcr2607.
- [3] 崔军威, 刘晓岭, 胡艺冰, 等. 乳腺癌骨转移患者临床病理特征及预后影响因素分析: 基于 SEER 数据库的回顾性研究[J]. *中华乳腺病杂志(电子版)*, 2020, 14(5): 274-279.
- [4] 王俊男, 徐拯, 林健, 等. 乳腺癌肝转移患者的临床病理特征及预后影响因素分析: 基于 SEER 数据库的回顾性研究[J]. *上海交通大学学报*, 2022, 42(11): 1562-1568.
- [5] Guo Y, Liao J, Li S, et al. Preoperative prediction of breast cancer histological grade using intratumoral and peritumoral radiomics features from T₂ WI and DWI MR sequences [J/OL]. *Breast Cancer (Dove Med Press)*, 2024, 16: e981-e991 [2024-12-09]. DOI: 10.2147/BCTT.S487988.
- [6] Mzoughi H, Njeh I, Slima MB, et al. Deep efficient-nets with transfer learning assisted detection of COVID-19 using chest X-ray radiology imaging [J/OL]. *Multimed Tools Appl*, 2023: e1-e23 [2023-03-27]. DOI: 10.1007/s11042-023-15097-3.
- [7] Wang X, Yu J, Zhu Q, et al. Potential of deep learning in assessing pneumoconiosis depicted on digital chest radiography [J]. *Occup Environ Med*, 2020, 77(9): 597-602.
- [8] Zhang J, Wang G, Ren J, et al. Multiparametric MRI-based radiomics nomogram for preoperative prediction of lymphovascular invasion and clinical outcomes in patients with breast invasive ductal carcinoma [J]. *Eur Radiol*, 2022, 32(6): 4079-4089.
- [9] Huang Y, Wang X, Cao Y, et al. Multiparametric MRI model to predict molecular subtypes of breast cancer using Shapley additive explanations interpretability analysis [J]. *Diagn Interv Imaging*, 2024, 105(5): 191-205.
- [10] Li HE, Cheng C. Development and assessment of a predictive model for Ki-67 expression using ultrasound indicators and non-morphological magnetic resonance imaging parameters before breast cancer therapy [J]. *Ultrason Imaging*, 2024, 46(6): 332-341.
- [11] Preuss K, Thach N, Liang X, et al. Using quantitative imaging for personalized medicine in pancreatic cancer: a review of radiomics and deep learning applications [J/OL]. *Cancers (Basel)*, 2022, 14(7): e1564 [2022-3-24]. DOI: 10.3390/cancers14071654.
- [12] McBee MP, Awan OA, Colucci AT, et al. Deep learning in radiology [J]. *Acad Radiol*, 2018, 25(11): 1472-1480.
- [13] Ren W, Xi X, Zhang X, et al. Predicting molecular subtypes of breast cancer based on multi-parametric MRI dataset using deep learning method [J/OL]. *Magn Reson Imaging*, 2025, 117: e110305 [2024-12-14]. DOI: 10.1016/j.mri.2024.110305.
- [14] 郭亚欣, 王贇霞, 尚怡研, 等. 弥散加权成像瘤内及瘤周影像组学预测乳腺癌组织学分级 [J]. *中国介入影像与治疗学*, 2024, 21(3): 160-165.
- [15] 吴佩琪, 刘再毅, 梁长虹. 基于 MRI 的影像组学特征在鉴别乳腺浸润性导管癌病理分级中的价值 [J]. *医学研究生学报*, 2018, 31(9): 938-942.
- [16] Han L, Zhu Y, Liu Z, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer [J]. *Eur Radiol*, 2019, 29(7): 3820-3829.
- [17] 匡静, 黄松涛, 黄小华, 等. 磁共振高分辨率延迟期的可解释性机器学习模型术前预测浸润性乳腺癌组织学分级 [J]. *磁共振成像*, 2025, 16(05): 164-169, 216.
- [18] 韩剑剑, 马培旗, 王小雷, 等. 基于数字化乳腺 X 线影像组学列线图预测浸润性乳腺癌组织学分级的多中心研究 [J]. *放射学实践*, 2024, 39(05): 634-640.
- [19] Luo W Q, Huang Q X, Huang X W, et al. Predicting breast cancer in breast imaging reporting and data system (BI-RADS) ultrasound category 4 or 5 lesions: a nomogram combining radiomics and BI-RADS [J/OL]. *Sci Rep*, 2019, 9(1): e11921 [2019-08-15]. DOI: 10.1038/s41598-019-48488-4.
- [20] Wibmer A, Hricak H, Gondo T, et al. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores [J]. *Eur Radiol*, 2015, 25(10): 2840-2850.
- [21] van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype [J]. *Cancer Res*, 2017, 77(21): 104-107.
- [22] Witowski J, Heacock L, Reig B, et al. Improving breast cancer diagnostics with deep learning for MRI [J/OL]. *Sci Transl Med*, 2022, 14(664): eabo4802 [2022-09-28]. DOI: 10.1126/scitranslmed.abo4802.
- [23] Chen M, Kong C, Lin G, et al. Development and validation of convolutional neural network-based model to predict the risk of sentinel or non-sentinel lymph node metastasis in patients with breast cancer: a machine learning study [J/OL]. *E Clin Med*, 2023, 63: e102176 [2023-08-24]. DOI: 10.1016/j.eclinm.2023.102176.
- [24] 袁权, 吴树剑, 范莉芳, 等. 基于 MRI 深度迁移学习影像组学术前预测直肠癌脉管浸润的研究 [J]. *磁共振成像*, 2025, 16(1): 54-60.
- [25] 梁园园, 李昌师, 赵文静, 等. 多模态影像学术前评估浸润性乳腺癌组织学分级 [J]. *中国介入影像与治疗学*, 2021, 18(10): 617-621.