

## · 智能医疗专题 · 论著 ·

DOI: 10.12464/j.issn.0253-9802.2025-0023

# 超声影像组学特征联合 miRNA-34a 表达水平对乳腺癌新辅助化疗患者病理完全缓解的预测能力

张伟娜<sup>1</sup>, 钟李长<sup>2</sup>, 师琳<sup>2</sup>, 来金玉<sup>2</sup>, 陈杰桓<sup>1</sup>, 谷丽萍<sup>2</sup>✉

(1. 东莞市滨海湾中心医院健康管理科, 广东 东莞 523000; 2. 上海市第六人民医院 / 上海交通大学医学院附属第六人民医院超声医学科, 上海 200233)

**【摘要】目的** 探讨超声影像组学特征联合外周血 miRNA-34a 表达水平对乳腺癌新辅助化疗 (NAC) 患者病理完全缓解 (pCR) 的预测能力。**方法** 选取 2019 年 1 月至 2023 年 12 月确诊为乳腺癌并接受 NAC 的 93 例女性患者为研究对象进行回顾性分析, 其中来自上海市第六人民医院 53 例, 东莞市滨海湾中心医院 40 例。采用实时荧光定量逆转录聚合酶链式反应 (qRT-PCR) 检测外周血中 miRNA-34a 的表达水平, 并采用 Pyradiomics 软件从术前超声图像中提取 107 个影像组学特征。经 Spearman 秩相关检验、Z-score 归一化处理及 LASSO 回归分析后, 筛选出 5 个关键影像组学特征。分别构建基于 miRNA-34a 的临床模型、基于影像组学特征的超声影像组学模型以及结合两者的联合模型, 并通过 K 近邻算法 (KNN) 分类器进行诊断性能评估。**结果** 单因素分析显示, pCR 组 miRNA-34a 表达水平高于 Non-pCR 组 ( $P < 0.001$ )。多因素 Logistic 分析显示, miRNA-34a 表达水平升高是乳腺癌 NAC 患者 pCR 的独立危险因素 ( $P = 0.015$ )。临床模型在训练组和验证组中的 AUC 分别为 0.787 (95%CI 0.547~1.000) 和 0.764 (95%CI 0.640~0.888), 超声影像组学模型分别为 0.806 (95%CI 0.605~1.000) 和 0.806 (95%CI 0.711~0.901), 联合模型的 AUC 提高至 0.875 (95%CI 0.712~1.000) 和 0.875 (95%CI 0.792~0.959)。DeLong 检验结果显示, 联合模型的性能优于单一临床模型 ( $P = 0.015$ )。决策曲线分析进一步证实了联合模型的临床实用性。**结论** 通过整合 miRNA-34a 表达水平与超声影像组学特征的联合模型可有效提高对乳腺癌患者 NAC 后 pCR 的预测能力, 超声影像组学特征与分子标志物之间具有潜在生物学关联, 可为个体化治疗提供新的工具和理论依据。

**【关键词】** 乳腺癌; miRNA-34a; 新辅助化疗; 病理完全缓解; 超声影像组学; 联合模型

## Predictive ability of ultrasound radiomics features combined with miRNA-34a expression levels for pathological complete response in breast cancer patients receiving neoadjuvant chemotherapy

ZHANG Weina<sup>1</sup>, ZHONG Lichang<sup>2</sup>, SHI Lin<sup>2</sup>, LAI Jinyu<sup>2</sup>, CHEN Jiehuan, GU Liping<sup>2</sup>✉

(1. Department of Health Management, Dongguan Binhaiwan Central Hospital, Dongguan 523000, China; 2. Department of Ultrasound Medicine, the Sixth People's Hospital of Shanghai / the Sixth People's Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai 200233, China)

Corresponding author: GU Liping, E-mail: guliping666@126.com

**【Abstract】Objective** To investigate the predictive ability of ultrasound radiomics features combined with peripheral blood miRNA-34a expression levels for pathological complete response (pCR) in breast cancer patients receiving neoadjuvant chemotherapy (NAC). **Methods** A retrospective analysis was conducted on 93 breast cancer female patients diagnosed with breast cancer and treated with NAC from January 2019 to December 2023. Among them, 53 patients were from the Sixth People's Hospital of Shanghai, and 40 patients were from Dongguan Binhaiwan Central Hospital. The expression level of miRNA-34a in peripheral blood was detected using real-time fluorescence quantitative reverse transcription polymerase chain reaction (qRT-PCR). 107 radiomics features were extracted from preoperative ultrasound images using Pyradiomics software. After Spearman rank correlation test, Z-score normalization and LASSO regression analysis, five key radiomics features were selected. Clinical models based on miRNA-34a, ultrasound radiomics models based on radiomics features, and a combined model integrating both were constructed. The diagnostic performance was

收稿日期: 2025-02-05

基金项目: 上海市浦东新区科技发展基金事业单位民生科研专项 (医疗卫生) (PKJ2020-Y09); 东莞市滨海湾中心医院内科科研项目 (2021001)

作者简介: 张伟娜, 副主任医师, 研究方向: 超声诊断, E-mail: 342109263@qq.com; 谷丽萍, 通信作者, 副主任医师, 研究方向: 妇产超声诊断, E-mail: guliping666@126.com

evaluated using the K-nearest neighbor (KNN) classifier. **Results** Univariate analysis showed that the expression levels of miRNA-34a in the pCR group were higher than those in the Non-pCR group ( $P < 0.001$ ). Multivariate Logistic analysis revealed that elevated miRNA-34a expression was an independent risk factor for pCR in breast cancer patients receiving NAC ( $P = 0.015$ ). The clinical model showed an AUC of 0.787 (95%CI 0.547-1.000) in the training group and 0.764 (95%CI 0.640-0.888) in the validation group. The ultrasound radiomics model showed an AUC of 0.806 (95%CI 0.605-1.000) in the training group and 0.806 (95%CI 0.711-0.901) in the validation group. The combined model significantly improved the AUC to 0.875 (95%CI 0.712-1.000) in the training group and 0.875 (95%CI 0.792-0.959) in the validation group. DeLong test results showed that the performance of the combined model was superior to the clinical model ( $P = 0.015$ ). Decision curve analysis further confirmed the clinical utility of the combined model. **Conclusions** The combined model significantly improves the predictive ability for pCR after NAC in breast cancer patients through integrating miRNA-34a expression levels and ultrasound radiomics features. The potential biological associations were observed between ultrasound radiomics features and molecular markers, providing new tools and theoretical support for personalized treatment.

**【Key words】** Breast cancer; miRNA-34a; Neoadjuvant chemotherapy; Pathologic complete response; Ultrasound radiomics; Combined model

乳腺癌是女性中最为常见的恶性肿瘤之一，其发病率和病死率呈现持续增长趋势<sup>[1,2]</sup>。新辅助化学治疗（neoadjuvant chemotherapy, NAC）作为治疗局部晚期乳腺癌的重要策略，能够缩减肿瘤体积，增加进行保乳手术的可能性，并为评估肿瘤对化学治疗反应提供有价值的信息<sup>[3]</sup>。然而，NAC的疗效在不同患者中表现出明显差异，部分患者对化学治疗缺乏敏感性，导致不良反应并延迟了最佳的治疗时机。因此，如何在治疗前准确预测患者对NAC的反应，对于实现个体化治疗具有重要意义。

影像学技术在乳腺癌的诊断、分期及疗效评估中发挥了重要作用，尤其是超声成像，由于其无创性、实时性和经济性，已被广泛应用于乳腺癌的筛查与监测<sup>[4]</sup>。然而，传统的影像学分析方法受限于分辨率、主观性等缺点，难以全面反映肿瘤内部的生物学特性，且其预测准确性有限<sup>[5]</sup>。近年来，影像组学（radiomics）作为一种新兴的研究方向，通过从医学图像中提取大量高通量的定量特征，影像组学能够反映肿瘤的异质性和微环境特征，从而用于疾病的诊断、分期和疗效预测<sup>[6-7]</sup>。例如，Wan等<sup>[8]</sup>利用超声影像组学特征构建乳腺癌NAC疗效预测模型，取得了较高的预测准确性，曲线下面积（area under the curve, AUC）为0.92。此外，Wang等<sup>[9]</sup>结合深度学习算法优化了超声影像组学模型，提高对病理完全缓解（pathologic complete response, pCR）的预测能力。尽管超声影像组学在乳腺癌的研究中取得重要进展，但它主要反映肿瘤的表型特征，无法直接揭示肿瘤的分子机制。近年来的分子水平研究表明，微小RNA（microRNA, miRNA）在乳腺癌的发生、发展及

治疗反应中起着关键调控作用<sup>[10]</sup>。其中，miRNA-34a作为一种肿瘤抑制因子，参与细胞周期调控、侵袭行为及凋亡机制，其表达水平与乳腺癌的预后密切相关<sup>[11]</sup>。研究证实，miRNA-34a通过调控Notch通路抑制肿瘤干细胞活性，可能增强化疗敏感性<sup>[12]</sup>。Li等<sup>[13]</sup>发现血清miRNA-34a高表达与乳腺癌NAC后pCR率呈正相关，但其与影像特征的关联性及其协同作用尚未被深入研究。

值得注意的是，影像学特征与分子标志物之间可能存在潜在关联。研究表明，肿瘤的影像异质性反映了其内部的基因表达模式和代谢状态<sup>[14]</sup>。有研究通过整合磁共振成像（magnetic resonance imaging, MRI）及增强MRI影像组学与转录组学数据，揭示了乳腺癌影像特征与特定基因表达谱之间的相关性<sup>[15-16]</sup>，为影像学与分子生物学的结合提供了理论依据。然而，目前关于超声影像组学与miRNA的关联及协同作用的研究仍处于初步阶段，尤其在乳腺癌NAC疗效预测中的应用尚未得到充分探索。

基于此，本研究分析超声影像组学特征与外周血miRNA-34a表达水平之间的生物学关联，评估二者在预测乳腺癌NAC患者疗效中的潜在价值。此外，本研究将进一步构建一个融合超声影像组学特征与外周血miRNA-34a表达水平的机器学习模型，为临床提供高效、可靠的辅助决策工具，以推动乳腺癌个体化治疗的精准实施与发展。

## 1 对象与方法

### 1.1 研究对象

本研究选取2019年1月至2023年12月确诊

为乳腺癌并接受 NAC 的 93 例女性患者为研究对象进行回顾性分析, 其中来自上海市第六人民医院 53 例, 东莞市滨海湾中心医院 40 例。纳入标准: ①经病理证实为乳腺癌; ②在 NAC 前接受超声检查并获得完整的影像学资料; ③有完整的临床和病理随访信息。排除标准: ①既往接受过乳腺手术或放射治疗、化学治疗; ②影像资料不完整或质量不佳; ③合并其他恶性肿瘤。所有患者均根据美国国家综合性癌症网络 (National Comprehensive Cancer Network, NCCN) 指南接受 6 个或 8 个周期的 NAC<sup>[17]</sup>, 根据 Miller&Payne 分级法将 NAC 病理反应 4~5 级判定为 pCR, 而 1~3 级为非病理完全缓解 (Non-pCR)<sup>[18]</sup>。采用随机抽样策略, 将 2 家医院的病例完全混合后, 按 4 : 1 比例进行随机分组。本研究获上海市第六人民医院伦理委员会 (批件号: 2024-KY-117K) 及东莞市滨海湾中心医院伦理委员会批准 (批件号: 2023041), 并已获患者知情同意。

## 1.2 研究方法

### 1.2.1 超声影像组学特征提取与筛选

所有患者在 NAC 前 1 周接受高频超声检查, 保存目标病灶最大径切面图像, 对超声图像进行统一的质量控制和标准化处理, 灰度范围标准化至 0~600, 确保后期影像组学特征提取的准确性和可比性。由 1 位拥有超过 10 年经验的超声科医师对肿瘤内的感兴趣区域 (region of interest, ROI) 进行手动绘制, 对于边界不清的乳腺癌结节的瘤内 ROI 则与另一位具有 15 年以上乳腺超声诊断经验的医师共同商讨后进行勾画。为确保 ROI 勾画的可靠性, 两位医师在正式分析前完成 50 例训练集的一致性校准。对存在分歧的边界判定, 采用协商讨论统一意见后进行勾画。最终所有纳入 ROI 均通过 Dice 相似系数  $\geq 0.85$  的质量控制。采用 Pyradiomics 软件提取包括形状特征、纹理特征和高阶特征等在内的影像组学参数。对提取的数据进行归一化处理, 留取组间相关系数 (intraclass correlation coefficient, ICC)  $> 0.75$  的影像组学特征。为减少冗余并筛选出最具预测能力的特征, 首先进行 Spearman 秩相关检验以排除高度相关的特征, 随后对数据进行 Z-score 归一化处理以消除量纲影响, 最后通过最小绝对收缩和选择算子 (least absolute shrinkage and selection operator, LASSO) 回归分析进一步降维, 筛选出关键的影像组学特征。

### 1.2.2 外周血 miRNA-34a 表达水平检测

为评估 miRNA-34a 在外周血中的表达水平, 采用实时荧光定量逆转录聚合酶链式反应 (quantitative reverse transcription-polymerase chain reaction, qRT-PCR) 技术进行检测。具体流程如下: 采用德国 Qiagen miRNeasy Serum/Plasma Kit 提取血浆总 RNA, 并加入合成 cel-miR-39 ( $5.6 \times 10^8$  copies/mL) 作为外源性对照以监控提取效率。逆转录采用日本 TaKaRa PrimeScript RT 试剂盒, 定量聚合酶链式反应 (quantitative polymerase chain reaction, qPCR) 采用瑞士 Roche LightCycler 480 SYBR Green I Master, 引物序列 (正向: 5'-TGGCAGTGTCTTAGCT-3', 反向: 5'-GTGCAGGTCGAGGT-3') 由生工生物工程 (上海) 股份有限公司合成。内参基因为 U6 snRNA 和 miR-16, 数据标准化采用  $2^{-\Delta\Delta Ct}$  法, 以健康志愿者混合血浆样本为校准组, 所有实验遵循定量实时荧光聚合酶链式反应实验发表的最低信息标准 (Minimum Information for Publication of Quantitative Real-Time PCR Experiments, MIQE) 指南。最终, miRNA-34a 表达水平被纳入 KNN 分类器, 构建临床模型以预测 pCR 发生概率。

### 1.2.3 模型建立与评价

本研究采用严格的建模流程构建乳腺癌 NAC 患者 pCR 与非 pCR 分类的预测模型, 所有特征选择及模型训练均严格限定于训练集数据范围内进行。基于收集的乳腺癌 NAC 患者临床及超声影像特征数据, 通过单因素分析及多因素 Logistic 分析筛选关键预测变量后, 构建了乳腺癌 NAC 患者 pCR 与非 pCR 分类的机器学习临床模型。影像组学模型的特征筛选及建模步骤如下: 首先, 对提取的影像组学特征进行 Z-score 标准化处理, 并通过 *t* 检验 ( $P < 0.05$ ) 筛选出具有差异的特征。接着, 利用 Pearson 相关性分析剔除高度相似的特征, 将相关系数  $r > 0.9$  的特征视为冗余特征, 每对特征中仅保留 1 个。随后, 按照 4 : 1 的比例随机划分训练组和验证组。在训练组中, 采用 LASSO 回归进一步筛选出最具预测能力的特征。基于这些优化后的特征, 构建影像组学模型。在建模过程中, 采用 K-最近邻算法 (K-nearest neighbor, KNN) 进行模型训练, 并通过 5 倍交叉验证评估模型性能。最后, 在独立的验证组中对模型进行测试以验证其泛化能力。基于上述数据分析, 分别构建了 3 种模型。①临床模型: 以 miRNA-34a 表达水

平为核心变量构建；②超声影像组学模型：以筛选出的关键影像组学特征为核心变量构建；③联合模型构建与评估：整合临床模型与超声影像组学模型2个模型的结果构建新的联合模型。

模型的诊断性能通过 AUC、灵敏度、特异度、准确度、精确度及 F1 值等指标进行评估。DeLong 检验用于比较联合模型与单一模型之间的差异，决策曲线分析评估模型在临床实践中的实用性。

### 1.3 统计学方法

采用 SPSS 26.0 和 Python 3.10 进行统计学分析。计量资料以  $\bar{x} \pm s$  表示，2 组比较采用  $t$  检验。计数资料以  $n (%)$  或相对数表示，2 组比较采用  $\chi^2$  检验。采用 Python 3.10 进行临床变量筛选、Spearman 秩相关检验、Z-score 归一化处理 and LASSO 回归分析。绘制预测模型受试者操作特征 (receiver operating characteristic, ROC) 曲线，比较模型的 AUC、敏感度、特异度、准确度、精确度、F1 值等。以双

侧  $P < 0.05$  为差异有统计学意义。

## 2 结果

### 2.1 患者基本特征

本研究共纳入 93 例接受 NAC 的乳腺癌患者，其中 pCR 组 54 例，Non-pCR 组 39 例。患者年龄 ( $52 \pm 5$ ) 岁。2 家医院的病例完全混合后，按 4 : 1 比例随机分为训练组 ( $n = 75$ ) 和验证组 ( $n = 18$ )。2 组在年龄、肿瘤大小、雌激素受体状态 (estrogen receptor, ER)、孕激素受体状态 (progesterone receptor, PR)、人类表皮生长因子受体 2 状态 (human epidermal growth factor receptor 2, HER2)、血流信号、乳腺影像报告与数据系统 (breast imaging-reporting and data system, BI-RADS) 分级及钙化情况等方面差异均无统计学意义 (均  $P > 0.05$ )，表明分组具有可比性，见表 1。

表 1 训练组和验证组乳腺癌 NAC 患者的临床特征

Table 1 Training and validation group clinical characteristics of breast cancer NAC patients

临床特征	训练组 ( $n=75$ )	验证组 ( $n=18$ )	$t/\chi^2$ 值	$P$ 值
miRNA-34a	0.64 ± 0.20	0.63 ± 0.24	0.755	0.452
年龄 / 岁	52.29 ± 4.35	49.56 ± 7.33	1.520	0.145
肿瘤最大径 / mm	16.75 ± 8.49	16.83 ± 7.87	0.010	0.989
ER/ $n (%)$			0.640	0.425
阴性	38 (50.67)	11 (61.11)		
阳性	37 (49.33)	7 (38.89)		
PR/ $n (%)$			1.450	0.229
阴性	41 (54.67)	7 (38.89)		
阳性	34 (45.33)	11 (61.11)		
HER2/ $n (%)$			0.010	0.931
阴性	30 (40.00)	7 (38.89)		
阳性	45 (60.00)	11 (61.11)		
钙化 / $n (%)$			0.320	0.571
无	55 (73.33)	12 (66.67)		
有	20 (26.67)	6 (33.33)		
彩色多普勒血流成像 / $n (%)$			0.100	0.752
无	47 (62.67)	12 (66.67)		
有	28 (37.33)	6 (33.33)		
BI-RADS/ $n (%)$			0.120	0.724
4b 及以上	69 (92.00)	17 (94.44)		
4a	6 (8.00)	1 (5.56)		

### 2.2 miRNA-34a 表达水平分析

单因素分析显示，pCR 组 miRNA-34a 表达水平高于 Non-pCR 组 ( $P < 0.001$ )。多因素 Logistic 分析显示，miRNA-34a 表达水平升高是乳腺癌 NAC

患者 pCR 的独立危险因素 ( $P = 0.015$ )，见表 2。

### 2.3 临床模型构建及诊断表现

单因素和多因素分析均显示，miRNA-34a 为乳腺癌 NAC 患者 pCR 的独立预测因子。因此，临

床模型仅纳入 miRNA-34a 表达水平作为输入变量, 采用 KNN 分类器进行建模。结果显示, 该模型在训练组中的 AUC 为 0.787 (95%CI 0.547~1.000), 灵敏度 0.600, 特异度 0.875, 准确度 0.722, 精确

度 0.857, F1 值 0.706; 在验证组中, AUC 为 0.764 (95%CI 0.640~0.888), 灵敏度 0.886, 特异度 0.710, 准确度 0.813, 精确度 0.812, F1 值 0.848, 见表 3。

表 2 训练组和验证组乳腺癌 NAC 患者 pCR 及非 pCR 的临床特征比较

Table 2 Comparison of clinical characteristics between pCR and non-pCR breast cancer NAC patients in the training and validation groups

临床特征	训练组		验证组		P <sub>单因素分析</sub>	P <sub>多因素分析</sub>
	Non-pCR (n=31)	pCR (n=44)	Non-pCR (n=8)	pCR (n=10)		
miRNA-34a	0.44 ± 0.10	0.77 ± 0.13	0.43 ± 0.17	0.80 ± 0.15	0.001	0.015
年龄	53.16 ± 5.85	51.68 ± 2.77	49.00 ± 10.88	50.00 ± 2.98	0.170	
肿瘤最大径 /mm	17.84 ± 7.10	15.98 ± 9.34	22.35 ± 4.52	12.41 ± 7.22	0.357	
ER/n (%)					0.613	
阴性	12 (38.71)	28 (63.64)	3 (3/8)	6 (6/10)		
阳性	19 (61.29)	16 (36.36)	5 (5/8)	4 (4/10)		
PR/n (%)					0.019	0.719
阴性	19 (61.29)	17 (38.64)	6 (6/8)	6 (6/10)		
阳性	12 (38.71)	27 (61.36)	2 (2/8)	4 (4/10)		
HER2/n (%)					0.080	
阴性	14 (45.16)	15 (34.09)	4 (4/8)	4 (4/10)		
阳性	17 (54.84)	29 (65.91)	4 (4/8)	6 (6/10)		
钙化 /n (%)					0.080	
无	14 (45.16)	15 (34.09)	4 (4/8)	4 (4/10)		
有	17 (54.84)	29 (65.91)	4 (4/8)	6 (6/10)		
彩色多普勒血流成像 /n (%)					0.014	0.741
0	19 (61.29)	16 (36.36)	4 (4/8)	7 (7/10)		
1	12 (38.71)	28 (63.64)	4 (4/8)	3 (3/10)		
BI-RADS/n (%)					0.280	
4a	1 (3.23)	5 (11.36)	1 (1/8)	0		
4b 及以上	30 (96.77)	39 (88.64)	7 (7/8)	10 (10/10)		

注: 单因素和多因素 Logistic 分析基于训练组。

表 3 不同模型的诊断表现

Table 3 Diagnostic performance of different models

组别	模型名称	AUC (95%CI)	灵敏度	特异度	准确度	精确度	F1 值
训练组	临床模型	0.787 (0.547~1.000)	0.600	0.875	0.722	0.857	0.706
	超声影像组学模型	0.806 (0.605~1.000)	0.300	1.000	0.611	1.000	0.462
	联合模型	0.875 (0.712~1.000)	0.600	1.000	0.778	1.000	0.750
验证组	临床模型	0.764 (0.640~0.888)	0.886	0.710	0.813	0.812	0.848
	超声影像组学模型	0.806 (0.711~0.901)	0.727	0.677	0.707	0.762	0.744
	联合模型	0.875 (0.792~0.959)	0.864	0.774	0.827	0.844	0.854

### 2.4 超声影像组学模型构建及诊断表现

采用 Pyradiomics 软件从患者的超声图像中共提取 107 个影像组学特征, 包括灰度区尺寸矩阵、灰度依赖矩阵、灰度共生矩阵、灰度游程矩阵、邻域灰度差矩阵、一阶直方图特征及形态特征等。经过 Spearman 秩相关检验、Z-score 归一化处

理和 LASSO 回归分析后, 筛选出 5 个关键的影像组学特征纳入超声影像组学模型 (Score= 0.580 645 161 290 322 5 + 0.025 709 \* original\_glem\_Correlation -0.044 043 \* original\_glrml\_ShortRunEmphasis +0.065 530 \* original\_glszm\_GrayLevelNonUniformity-0.033 418 \* original\_ngtdm\_Contrast +0.189 714 \* original\_

shape\_Elongation)。该模型在训练组中的 AUC 为 0.806 (95%CI 0.605~1.000)，灵敏度 0.300，特异度 1.000，准确度 0.611，精确度 1.000，F1 值 0.462；在验证组中，AUC 为 0.806 (95%CI 0.711~0.901)，灵敏度 0.727，特异度 0.677，准确度 0.707，精确度 0.762，F1 值 0.744，见表 3。

### 2.5 联合模型诊断表现

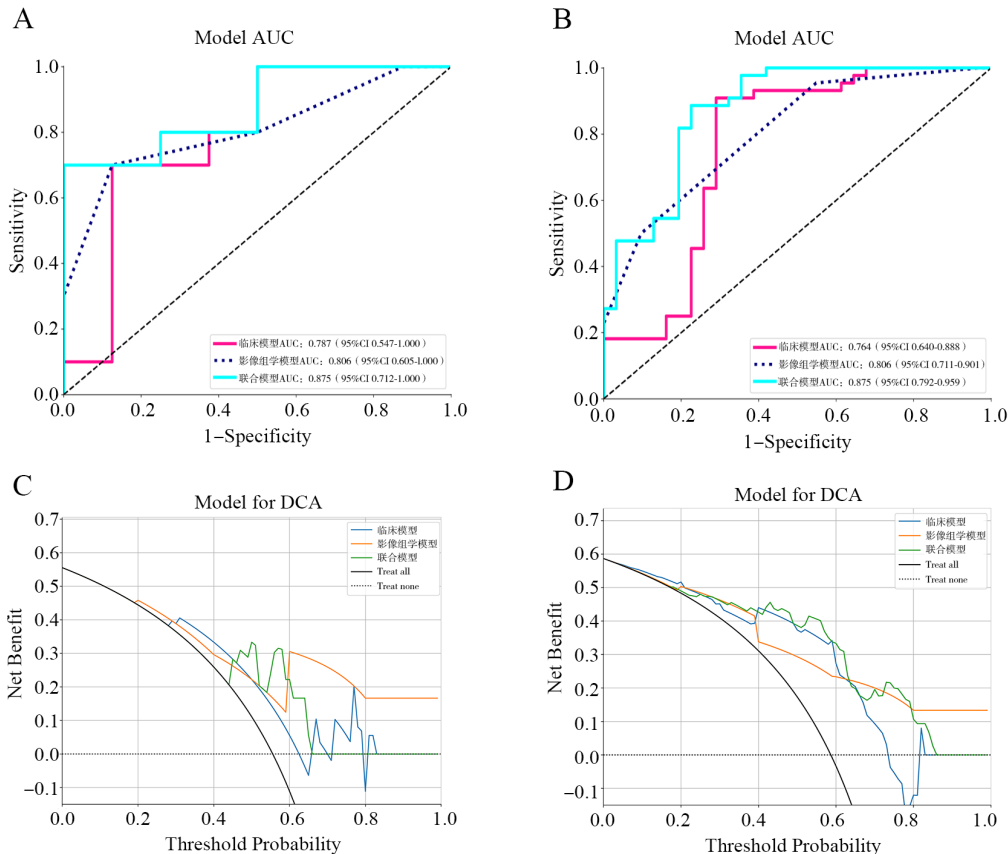
将 miRNA-34a 与上述筛选出的 5 个关键影像组学特征同时输入 KNN 分类器，构建新的联合模型，并通过 ROC 曲线对其进行评价。结果显示，联合模型在训练组和验证组中的 AUC 分别为 0.875 (95%CI 0.712~1.000) 和 0.875 (95%CI 0.792~0.959)，明显优于单一的临床模型和超声影像组学模型，见图 1A、B。DeLong 检验结果显示，在验证组中，联合模型的性能优于单一临床模型 ( $P = 0.015$ )，高于超声影像组学模型，但差异无统计学意义 ( $P = 0.057$ )。决策曲线分析进一步表明，联合模型在预测乳腺癌患者 NAC 后 pCR 方面表现出更高的临床实用性，见图 1C、D，图 2。

### 2.6 生物学关联初步分析

根据  $Contrast = \sum_{i,j} |i_j - j_i|^2 \cdot P(i, j)$  其中  $P(i, j)$  为归一化后的灰度共生矩阵 (gray-level co-occurrence matrix, GLCM) 元素，计算 GLCM 对比度。通过 Pearson 相关性分析发现，超声影像组学特征中的“GLCM 对比度”与 miRNA-34a 表达水平呈正相关 ( $r = 0.38, P < 0.01$ )，提示影像异质性可能与肿瘤的分子调控机制存在潜在联系。研究表明，miRNA-34a 可能通过调低 B 淋巴细胞瘤 -2 (B-cell lymphoma, Bcl-2) 等抗凋亡基因的表达<sup>[19]</sup>，增强乳腺癌细胞对化疗药物的敏感性。

## 3 讨论

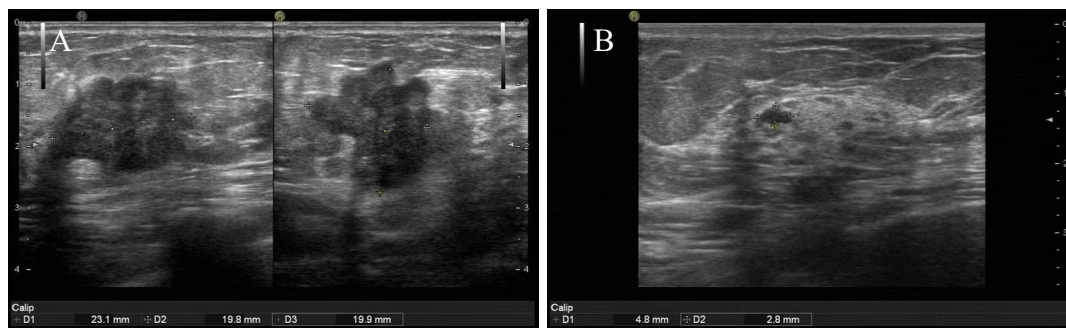
本研究通过整合超声影像组学特征与外周血 miRNA-34a 表达水平，构建一个用于预测乳腺癌 NAC 后 pCR 的联合模型，结果显示联合模型在训练组和验证组中的 AUC 均优于单一的临床模型和超声影像组学模型。这一结果表明，结合影像学



注：A、C 为训练组，B、D 为验证组。

图 1 临床模型、超声影像组学模型、联合模型诊断表现

Figure 1 Diagnostic performance of the clinical model, ultrasound radiomics model and joint model



注: A 为 NAC 前乳腺癌病灶, B 为 NAC 3 个周期后乳腺癌病灶; 患者女, 55 岁, NAC 前病灶大小约 23.1 mm × 19.8 mm × 19.9 mm, 联合模型预测达到 pCR 概率为 0.857, 在接受 NAC 3 个周期后, 病灶大小 4.8 mm × 2.8 mm, 达到 pCR。

图 2 一例联合模型预测乳腺癌 NAC 患者 pCR

Figure 2 Joint model for predicting pathological complete response in one case of breast cancer patient after neoadjuvant chemotherapy

与分子生物学信息能更全面地反映肿瘤的生物学特性, 从而提高对 NAC 疗效的预测准确性。

超声影像组学通过提取高通量定量特征, 能够无创地反映肿瘤内部的异质性和生物学行为<sup>[20-21]</sup>。本研究筛选出的 5 个关键影像特征 (如 GLCM 对比度) 在预测 NAC 疗效中表现出较高的诊断效能 (验证组 AUC=0.806), 这与 Wan 等<sup>[8]</sup> (AUC=0.92)、冯晓丹等<sup>[22]</sup> (AUC=0.844)、杜瑶等<sup>[23]</sup> (AUC=0.828) 的研究结果基本一致。值得注意的是, GLCM 对比度与肿瘤异质性密切相关, 高对比度可能提示肿瘤内部结构复杂、代谢活跃。本研究发现, 该特征与 miRNA-34a 表达呈正相关 ( $P < 0.01$ ), 进一步支持了影像异质性与分子调控机制存在关联的假设<sup>[15-16]</sup>。然而, 与基于 MRI 的影像组学研究相比<sup>[15, 24]</sup>, 超声影像组学在分辨率上的局限性可能导致部分微环境特征未被充分捕捉, 但 MRI 的高成本和较长的成像时间在临床推广上仍存在障碍, 而超声影像组学具有较低的成本和较短的扫描时间, 且对软组织的对比度较好, 能较为准确地提取肿瘤特征, 未来可通过多模态影像融合提升特征提取的全面性。

miRNA-34a 作为肿瘤抑制因子, 其高表达与 NAC 后 pCR 相关, 本研究结果与 Li 等<sup>[13]</sup> 的发现一致。机制上, miRNA-34a 通过靶向调控 Notch 信号通路抑制肿瘤干细胞活性, 同时下调 Bcl-2 等抗凋亡基因的表达<sup>[12]</sup>, 从而增强化疗诱导的细胞凋亡。此外, miRNA-34a 还可通过调控程序性死亡配体 -1 表达影响肿瘤免疫微环境<sup>[25]</sup>, 可能间接提升化疗敏感度。然而, miRNA-34a 的表达受表观遗传修饰 (如 DNA 甲基化) 调控<sup>[26]</sup>, 其动态变化

可能影响预测模型的稳定性。因此, 未来需结合多时间点采样及表观遗传学分析, 以更全面地评估其临床价值。

单一的影像组学或分子标志物模型在预测效能上均存在局限性。本研究中, 临床模型 (miRNA-34a) 和超声影像组学模型的验证组 AUC 分别为 0.764 和 0.806, 而联合模型的 AUC 提升至 0.875 ( $P = 0.015$ ), 优于单一模型。这一结果与 Wang 等<sup>[9]</sup> 提出的深度学习与影像组学多模态数据融合优于单一维度信息一致。近年来, 机器学习模型在肿瘤影像与临床数据分析中的广泛应用, 极大地推动了乳腺癌诊断精度的提升。研究表明, 通过特征选择与模型优化, 常见算法如支持向量机 (support vector machine, SVM)、Logistic 回归和随机森林等均能在乳腺癌分类任务中获得超过 90% 的准确率<sup>[7]</sup>。尤其是在整合多维度数据 (如影像组学与临床指标) 后, 模型性能可进一步增强。例如, Rahman 等<sup>[27]</sup> 比较了 6 种机器学习算法发现, Logistic 回归和 SVM 在多指标评估下表现最优, 其准确率可达 97% 以上。同时, 优化算法在提升诊断模型表现方面亦发挥关键作用。Bezdan 等<sup>[28]</sup> 提出将正余弦算法与对立学习机制结合进行超参数调整, 有效提升 5 类主流算法的诊断性能。Poornajaf 和 Yosefi<sup>[29]</sup> 研究发现, Logistic 回归与 Extra Tree 模型在乳腺肿瘤预测中 AUC 值超过 99%, 显示其在特征维度优化后的诊断潜力。此外, Rasool 等<sup>[30]</sup> 基于数据探索性技术, 构建了多个模型并在两个公开数据集上进行验证, 最优模型 (多项式 SVM) 在 Wisconsin Breast Cancer Database (WDBC) 数据集上准确率高达 99.3%。这些成果表明, 借助机器

学习算法<sup>[31]</sup>,尤其是在结合特征选择、超参数优化和多模态数据融合后,诊断模型的效能可获得显著提升。本研究表明,影像组学反映肿瘤表型特征,miRNA-34a揭示肿瘤的分子机制,二者的整合能从宏观与微观层面更全面评估肿瘤特性。

然而,本研究也存在一定的局限性。首先,因客观因素导致样本量相对较小、数据来源单一(仅2家医院),研究结论可能存在选择偏倚和外部效度受限问题,未来应开展多中心、大样本研究,以验证模型的稳定性与推广性。其次,超声影像组学特征的提取可能受设备和成像参数的影响,可能影响纹理特征的生物学解释,需进一步标准化提取流程,确保特征的可重复性和一致性。再次,本研究未纳入超声功能成像技术(如超声造影、弹性成像及超微血流成像)对NAC疗效的功能性评估指标。这些技术可提供微循环动态变化、组织弹性及残余病灶活性等信息,未能与超声影像组学或miRNA-34a建立关联。此外,本研究仅纳入一种分子标志物(miRNA-34a),未来可考虑引入更多的miRNA或其他分子标志物,如miRNA-21、miRNA-155等,以进一步提高模型的预测性能。

综上所述,超声影像组学与miRNA-34a的联合模型显著提升了乳腺癌NAC疗效预测的准确性,其生物学互补性为个体化治疗提供了可靠工具,这一多维度整合策略有望推动乳腺癌精准医学的临床转化。

**利益冲突声明:** 本研究未受到企业、公司等第三方资助,不存在潜在利益冲突。

## 参 考 文 献

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. *CA Cancer J Clin*, 2021, 71 (3): 209-249. DOI: 10.3322/caac.21660.
- [2] 左艳玲, 何广敏, 李涛, 等. 高频超声彩色像素密度在BI-RADS 4类乳腺肿块鉴别诊断中的价值 [J]. *新医学*, 2022, 53 (6): 431-435. DOI: 10.3969/j.issn.0253-9802.2022.06.009.
- [3] ZUO Y L, HE G M, LI T, et al. Value of high-frequency ultrasound color pixels per inch in differential diagnosis of BI-RADS 4 breast masses [J]. *J New Med*, 2022, 53 (6): 431-435. DOI: 10.3969/j.issn.0253-9802.2022.06.009.
- [4] 顾依涵, 毛晓韵. 乳腺癌新辅助化疗疗效评估 [J]. *中华肿瘤防治杂志*, 2024, 31 (2): 118-124. DOI: 10.16073/j.cnki.cjcp.2024.02.08.
- [5] GU Y H, MAO X Y. Efficacy evaluation of neoadjuvant chemotherapy in breast cancer [J]. *Chin J Cancer Prev Treat*, 2024, 31 (2): 118-124. DOI: 10.16073/j.cnki.cjcp.2024.02.08.
- [6] LI H, ZHAO J, JIANG Z. Deep learning-based computer-aided detection of ultrasound in breast cancer diagnosis: a systematic review and meta-analysis [J]. *Clin Radiol*, 2024, 79 (11): e1403-e1413. DOI: 10.1016/j.crad.2024.08.002.
- [7] CUI H, ZHAO D, HAN P, et al. Predicting pathological complete response after neoadjuvant chemotherapy in advanced breast cancer by ultrasound and clinicopathological features using a nomogram [J]. *Front Oncol*, 2021, 11: 718531. DOI: 10.3389/fonc.2021.718531.
- [8] CONTI A, DUGGENTO A, INDOVINA I, et al. Radiomics in breast cancer classification and prediction [J]. *Semin Cancer Biol*, 2021, 72: 238-250. DOI: 10.1016/j.semcancer.2020.04.002.
- [9] ZHONG L, SHI L, ZHOU L, et al. Development of a nomogram-based model combining intra-and peritumoral ultrasound radiomics with clinical features for differentiating benign from malignant in Breast Imaging Reporting and Data System category 3-5 nodules [J]. *Quant Imaging Med Surg*, 2023, 13 (10): 6899-6910. DOI: 10.21037/qims-23-283.
- [10] WAN C F, JIANG Z Y, WANG Y Q, et al. Radiomics of multimodal ultrasound for early prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer [J]. *Acad Radiol*, 2025, 32 (4): 1861-1873. DOI: 10.1016/j.acra.2024.11.012.
- [11] WANG Z, LI X, ZHANG H, et al. Deep learning radiomics based on two-dimensional ultrasound for predicting the efficacy of neoadjuvant chemotherapy in breast cancer [J]. *Ultrason Imaging*, 2024, 46 (6): 357-366. DOI: 10.1177/01617346241276168.
- [12] IORIO M V, CROCE C M. microRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review [J]. *EMBO Mol Med*, 2017, 9 (6): 852. DOI: 10.15252/emmm.201707779.
- [13] 沈丽丽, 李倩倩, 陈玥彤, 等. miRNA-34a在乳腺癌患者中的表达水平及临床意义 [J]. *系统医学*, 2022, 7 (5): 164-168. DOI: 10.19368/j.cnki.2096-1782.2022.05.164.
- [14] SHEN L L, LI Q Q, CHEN Y T, et al. Expression level and clinical significance of miRNA-34a in breast cancer patients [J]. *Sys Med*, 2022, 7 (5): 164-168. DOI: 10.19368/j.cnki.2096-1782.2022.05.164.
- [15] ZHENG H, WANG W, LI S, et al. The effect of Zbxz23ir-21 NANO (nanomaterials) delivery vector on apoptosis and PTEN (phosphatase and tensin homolog deleted on chromosome ten) / PI3K (Intracellular phosphatidylinositol kinase) / AKT (related to the A and C kinase) in children with CHOLESTEATOMA in middle ear [J]. *Bioengineered*, 2021, 12 (1): 8809-8821. DOI: 10.1080/21655979.2021.1984718.
- [16] LI Z H, YU N S, DENG Q, et al. LncRNA SNHG7 mediates the chemoresistance and stemness of breast cancer by sponging

- miR-34a [J]. *Front Oncol*, 2020, 10 : 592757. DOI: 10.3389/fonc.2020.592757.
- [14] TAÏEB D, HICKS R J, PACAK K. Radiopharmaceuticals in paraganglioma imaging: too many members on board [J]. *Eur J Nucl Med Mol Imaging*, 2016, 43 ( 3 ): 391-393. DOI: 10.1007/s00259-015-3213-4.
- [15] KIM G R, KU Y J, KIM J H, et al. Correlation between MR image-based radiomics features and risk scores associated with gene expression profiles in breast cancer [J]. *Taehan Yongsang Uihakhoe Chi*, 2020, 81 ( 3 ): 632-643. DOI: 10.3348/jksr.2020.81.3.632.
- [16] YEH A C, LI H, ZHU Y, et al. Radiogenomics of breast cancer using dynamic contrast enhanced MRI and gene expression profiling [J]. *Cancer Imaging*, 2019, 19 ( 1 ): 48. DOI: 10.1186/s40644-019-0233-5.
- [17] GRADISHAR W J, ANDERSON B O, ABRAHAM J, et al. Breast Cancer, Version 3.2020, NCCN Clinical Practice Guidelines in Oncology [J]. *J Natl Compr Canc Netw*, 2020, 18 ( 4 ): 452-478. DOI: 10.6004/jnccn.2020.0016.
- [18] OGSTON K N, MILLER I D, PAYNE S, et al. A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival [J]. *Breast*, 2003, 12 ( 5 ): 320-327. DOI: 10.1016/s0960-9776(03)00106-1.
- [19] TOKUMARU Y, KATSUTA E, OSHI M, et al. High expression of miR-34a associated with less aggressive cancer biology but not with survival in breast cancer [J]. *Int J Mol Sci*, 2020, 21 ( 9 ): 3045. DOI: 10.3390/ijms21093045.
- [20] ELSAYED B, ALKSAS A, SHEHATA M, et al. Exploring neoadjuvant chemotherapy, predictive models, radiomic, and pathological markers in breast cancer: a comprehensive review [J]. *Cancers*, 2023, 15 ( 21 ): 5288. DOI: 10.3390/cancers15215288.
- [21] 程辰, 李雪平, 赵红艳. 超声影像组学在乳腺癌新辅助化疗中的应用进展 [J]. *医学影像学杂志*, 2023, 33 ( 5 ): 878-881, 885.  
CHENG C, LI X P, ZHAO H Y. Application progress of ultrasonography in neoadjuvant chemotherapy of breast cancer [J]. *J Med Imag*, 2023, 33 ( 5 ): 878-881, 885.
- [22] 冯晓丹, 刘菲菲, 杜瑶, 等. 基于超声影像组学及临床特征的可解释性机器学习模型预测乳腺癌新辅助化疗疗效 [J]. *中国超声医学杂志*, 2024, 40 ( 3 ): 278-282.  
FENG X D, LIU F F, DU Y, et al. Predicting the effect of interpretable machine learning model based on ultrasound radiomics and clinical features on neoadjuvant chemotherapy for breast cancer [J]. *Chin J Ultrasound Med*, 2024, 40 ( 3 ): 278-282.
- [23] 杜瑶, 冯晓丹, 吴萌, 等. 基于超声影像组学构建联合模型预测乳腺癌新辅助化疗疗效 [J]. *中国超声医学杂志*, 2023, 39 ( 12 ): 1353-1356. DOI: 10.3969/j.issn.1002.0101.2023.12.010.  
DU Y, FENG X D, WU M, et al. To construct a combined radiomics model based on ultrasound to predict the efficacy of neoadjuvant chemotherapy for breast cancer [J]. *Chin J Ultrasound Med*, 2023, 39 ( 12 ): 1353-1356. DOI: 10.3969/j.issn.1002-0101.2023.12.010.
- [24] 赵玲, 陈正国, 罗锐, 等. 基于 SHAP 可解释性分析和 MRI 影像组学构建机器学习模型预测乳腺癌新辅助化疗疗效 [J]. *现代肿瘤医学*, 2024, 32 ( 15 ): 2839-2844. DOI: 10.3969/j.issn.1672-4992.2024.15.023.  
ZHAO L, CHEN Z G, LUO R, et al. Building a machine learning model based on SHAP interpretable analysis and MRI imageology to predict the efficacy of neoadjuvant chemotherapy for breast cancer [J]. *J Mod Oncol*, 2024, 32 ( 15 ): 2839-2844. DOI: 10.3969/j.issn.1672-4992.2024.15.023.
- [25] HONG W, XUE M, JIANG J, et al. Circular RNA circ-CPA4/let-7 miRNA/PD-L1 axis regulates cell growth, stemness, drug resistance and immune evasion in non-small cell lung cancer ( NSCLC ) [J]. *J Exp Clin Cancer Res*, 2020, 39 ( 1 ): 149. DOI: 10.1186/s13046-020-01648-1.
- [26] STRUTZ J, BAUMANN K, WEISS E, et al. Transient hyperglycemia and hypoxia induce memory effects in AngiomiR expression profiles of feto-placental endothelial cells [J]. *Int J Mol Sci*, 2021, 22 ( 24 ): 13378. DOI: 10.3390/ijms222413378.
- [27] RAHMAN M S, RAHMAN M A, NIPA T A, et al. A machine learning approach to predictive modeling for breast cancer prediction [C]//2024 3rd International Conference on Advancement in Electrical and Electronic Engineering ( ICAEEE ). Gazipur, Bangladesh. IEEE, 2024 : 1-6. DOI: 10.1109/ICAEEE62219.2024.10561811.
- [28] BEZDAN T, STRUMBERGER I, TUBA M. Optimizing machine learning for breast cancer detection by hybrid metaheuristic approach [C]//2024 12th International Symposium on Digital Forensics and Security ( ISDFS ). San Antonio, TX, USA. IEEE, 2024 : 1-6. DOI: 10.1109/ISDFS60797.2024.10527334.
- [29] POORNAJAF M, YOSEFI S. Improvement of the performance of machine learning algorithms in predicting breast cancer [J]. *Front Health Inform*, 2023, 12 : 132. DOI: 10.30699/fhi.v12i0.400.
- [30] RASOOL A, BUNTERNGCHIT C, TIEJIAN L, et al. Improved machine learning-based predictive models for breast cancer diagnosis [J]. *Int J Environ Res Public Health*, 2022, 19 ( 6 ): 3211. DOI: 10.3390/ijerph19063211.
- [31] 程兆瑞, 王彤. 人工智能技术在肝细胞癌诊断、复发及预后预测研究进展 [J]. *中山大学学报 ( 医学科学版 )*, 2023, 44 ( 6 ): 903-909. DOI: 10.13471/j.cnki.j.sun.yat-sen.univ ( med. sci ).2023.0602.  
CHENG Z R, WANG T. Research progress in artificial intelligence for diagnosis, prediction of recurrence and prognosis in hepatocellular carcinoma [J]. *J Sun Yat Sen Univ ( Med Sci )*, 2023, 44 ( 6 ): 903-909. DOI: 10.13471/j.cnki.j.sun.yat-sen.univ ( med. sci ).2023.0602.

( 责任编辑: 谢汝莹 洪悦民 )