

三阴性乳腺癌新辅助化疗后病理完全缓解的影响因素及预测模型构建

杨柳¹, 季福庆², 张明坤¹, 王哲¹, 张聚良^{1*}

¹空军军医大学西京医院甲乳血管外科, 陕西西安 710032; ²西北大学附属医院/西安市第三医院甲乳外科, 陕西西安 710018

[中图分类号] R722.6 [文献标志码] A [DOI] 10.11855/j.issn.0577-7402.1519.2023.1120

[声明] 本文所有作者声明无利益冲突

[引用本文] 杨柳, 季福庆, 张明坤, 等. 三阴性乳腺癌新辅助化疗后病理完全缓解的影响因素及预测模型构建[J]. 解放军医学杂志, 2024, 49(8): 855-860.

[收稿日期] 2022-07-12 [录用日期] 2022-12-18 [上线日期] 2023-11-20

[摘要] **目的** 分析接受新辅助化疗的三阴性乳腺癌(TNBC)患者病理完全缓解(pCR)的影响因素, 并构建临床预测模型。**方法** 收集2018年5月—2021年5月空军军医大学西京医院收治的348例接受新辅助化疗的TNBC患者(作为建模集), 以及2018年5月—2021年5月西安市第三医院收治的69例接受新辅助化疗的TNBC患者(作为验证集)。分析并比较建模集与验证集的临床病理特征, 在建模集中采用LASSO回归模型分析TNBC患者新辅助化疗后pCR的独立影响因素, 并构建列线图模型。采用Bootstrap法对模型进行内部验证; 采用受试者工作特征(ROC)曲线评估模型的区分度, 校准曲线评估模型的准确性, 临床决策曲线分析(DCA)评价模型的临床获益和应用价值。**结果** 建模集与验证集的手术术式和T分期比较差异有统计学意义($P<0.05$)。LASSO回归模型分析结果显示, T分期、N分期、是否使用铂类药物及临床疗效评估是TNBC患者新辅助化疗后pCR的独立影响因素($P<0.05$), 将这些因素纳入并构建列线图预测模型。建模集中列线图模型预测TNBC患者新辅助化疗后pCR率的曲线下面积(AUC)为0.811(95%CI 0.763~0.859), 验证集中为0.801(95%CI 0.727~0.928)。Bootstrap法内部验证显示一致性指数(C-index)为0.79, 表明模型在建模集与验证集中均具有良好的区分度。校准曲线显示, 列线图模型的预测生存率与实际生存率接近; DCA显示列线图模型的临床获益及应用价值较高。**结论** 列线图模型能准确预测TNBC患者新辅助化疗后的pCR率, 可为临床诊疗提供指导依据。

[关键词] 三阴性乳腺癌; 新辅助化疗; 病理完全缓解; 列线图

The factors affecting pathological complete response of triple negative breast cancer patients after neoadjuvant chemotherapy and the construction of related model

Yang Liu¹, Ji Fu-Qing², Zhang Ming-Kun¹, Wang Zhe¹, Zhang Ju-Liang^{1*}

¹Department of Thyroid, Breast and Vascular Surgery, Xijing Hospital, Air Force Medical University, Xi'an, Shaanxi 710032, China

²Department of Thyroid and Breast, the Affiliated Hospital of Northwest University/Xi'an No.3 Hospital, Xi'an 710018, Shaanxi, China

*Corresponding author, E-mail: vascularzhang@163.com

This work was supported by the National Natural Science Foundation of China (81902677), and the Shaanxi Province Key Research and Development Plan (2018ZDXM-SF-066)

[Abstract] **Objective** To analyze the factors affecting pathological complete response (pCR) of triple-negative breast cancer (TNBC) patients after neoadjuvant chemotherapy, and construct a nomogram to forecast the pCR rate. **Methods** The clinical and pathological data of 348 TNBC patients who received neoadjuvant chemotherapy in the Air Force Medical University-Affiliated Xijing Hospital from May 2018 to May 2021 were collected and set as modeling set. The clinical and pathological data of 69 TNBC patients who received neoadjuvant chemotherapy in the Xi'an No.3 Hospital from May 2018 to May 2021 were collected and set as

[基金项目] 国家自然科学基金青年科学基金项目(81902677); 陕西省重点研发计划(2018ZDXM-SF-066)

[作者简介] 杨柳, 硕士研究生, 主治医师, 主要从事乳腺癌的基础及临床研究

[通信作者] 张聚良, E-mail: vascularzhang@163.com

validation set. The clinical and pathological characteristics were compared between the modeling set and the validation set. In the modeling set, the independent risk factors of pCR in TNBC patients after neoadjuvant chemotherapy were screened by LASSO regression model analysis, and the nomogram model was constructed. Internal validation of the model was conducted using Bootstrap method, and the discrimination of the model was assessed by receiver operating characteristic (ROC) curve. The accuracy of the model was evaluated by the calibration curve and the clinical benefits and application value of the model were evaluated by clinical decision curve analysis (DCA). **Results** There were significant differences in surgical method and T stage between the patients in modeling set and validation set ($P<0.05$). The results of analysis of LASSO regression model showed that T stage, N stage, the use of platinum drugs and clinical efficacy evaluation were independent risk factors of pCR in TNBC patients after neoadjuvant chemotherapy ($P<0.05$). Based on the above variables, the nomogram models were constructed. In modeling set, area under curve (AUC) was 0.811 (95%CI 0.763-0.859); in validation set, AUC was 0.801 (95%CI 0.727-0.928). The Bootstrap method showed the C-index for internal validation was 0.79, indicating the model has good discrimination in both the modeling and validation sets. The calibration curve analysis showed that model predicted pCR rates had a good consistency with the actual observed values, and the DCA showed that model can bring clinical benefit. **Conclusion** The nomogram can accurately predict the pCR rates of TNBC patients after neoadjuvant chemotherapy and provide scientific basis for clinical diagnosis and treatment.

[Key words] triple-negative breast cancer; neoadjuvant chemotherapy; pathological complete response; nomogram

国际癌症研究机构(International Agency for Research on Cancer, IARC)发布的全球癌症统计数据显示,2020年全球约有1930万新增癌症病例和1000万癌症死亡病例,其中发病人数最多的是女性乳腺癌,达226万例^[1]。三阴性乳腺癌(triple negative breast cancer, TNBC)是雌激素受体、孕激素受体及人表皮生长因子受体2均为阴性的一种特殊分子亚型的乳腺癌^[2],占有乳腺癌的10%~15%^[3],其侵袭性高,预后差,易发生转移和复发^[4]。TNBC对新辅助化疗反应敏感,客观有效率和病理完全缓解(pathological complete response, pCR)率较高,达到pCR的TNBC患者的预后与激素受体阳性乳腺癌患者类似^[5]。因此,术前准确预测能够达到pCR的患者对提高疗效和改善预后具有重要意义。本研究回顾性分析接受新辅助化疗的TNBC患者pCR的影响因素,并建立列线图预测模型预测pCR率,旨在为TNBC的临床诊疗提供依据。

1 资料与方法

1.1 一般资料 纳入2018年5月—2021年5月空军军医大学西京医院收治的348例接受新辅助化疗的TNBC患者(作为建模集),以及2018年5月—2021年5月西安市第三医院收治的69例接受新辅助化疗的TNBC患者(作为验证集)。纳入标准:(1)病理确诊为乳腺癌;(2)分子分型为TNBC,并接受新辅助化疗;(3)接受受环+紫杉为基础的化疗方案,可加用铂类药物。排除标准:(1)特殊类型的乳腺癌;(2)双侧乳腺癌;(3)分化程度、肿瘤分期及病理评估等重要信息缺失。

1.2 资料收集 收集患者的临床病理资料,包括年龄、T分期、N分期、分化程度、乳腺癌位置、Ki-67指数、体重指数(body mass index, BMI)、手术方式

及临床疗效评估等。其中Ki-67指数依据X-tile软件筛选的最佳截断值44%分为二分类变量。根据实体瘤疗效评价标准(response evaluation criteria in solid tumor, RECIST1.1)将临床疗效评估分为:(1)疾病进展(progressive disease, PD),以整个研究过程中所有测量的靶病灶直径之和的最小值为参照,直径和相对增加至少20%(若基线测量值最小则以基线值为参照)。此外,必须满足直径和的绝对值增加至少5 mm(出现一个或多个新病灶也视为疾病进展)。(2)疾病稳定(stable disease, SD),靶病灶减小的程度未达到部分缓解(partial response, PR),增加的程度也未达到PD,介于两者之间。(3)PR(所有可测量目标病灶的直径总和低于基线的30%)。(4)完全缓解(complete response, CR),所有目标病灶完全消失^[6]。417例中,402例接受EC-T(表柔比星100 mg/m²,环磷酰胺600 mg/m²,多西他赛100 mg/m²)方案化疗8个周期,每3周1次;15例接受EC-TP(表柔比星100 mg/m²,环磷酰胺600 mg/m²,多西他赛100 mg/m²,卡铂AUC=5)方案化疗8个周期,每3周1次。

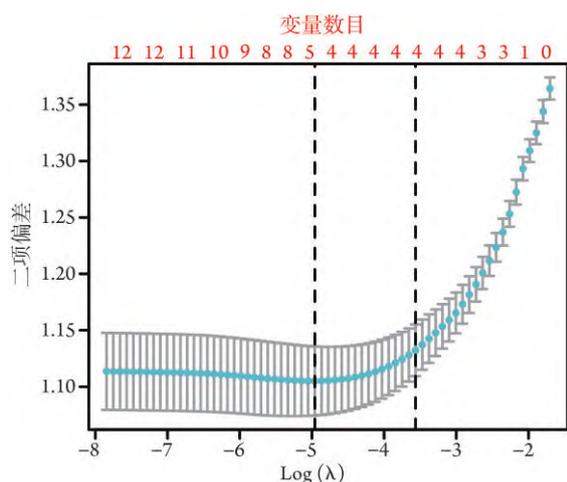
1.3 模型构建和验证 在建模集中,采用LASSO回归模型分析TNBC患者pCR的独立影响因素,并建立预测模型。采用Bootstrap法重抽样1000次进行内部验证;采用受试者工作特征(ROC)曲线评估模型的区分度,校准曲线评估模型的准确性,临床决策曲线分析(decision curve analysis, DCA)评价模型的临床获益和应用价值。

1.4 统计学处理 采用R软件(4.0.3)进行统计分析并绘制列线图。计量资料以 $\bar{x}\pm s$ 表示,两组间比较采用 t 检验;计数资料以率(%)表示,采用 χ^2 检验和Fisher精确检验比较建模集与验证集的临床病理特征,对于等级资料进行Mann-whitney U 秩和检验。 $P<0.05$ 为差异有统计学意义。

2 结 果

2.1 建模集与验证集中TNBC患者的临床病理特征比较 建模集中TNBC患者平均56.0岁，验证集中为53.0岁，两者比较差异无统计学意义($P=0.064$)；建模集与验证集中TNBC患者的手术术式和T分期比较，差异有统计学意义($P<0.05$)，其余临床病理特征差异均无统计学意义($P>0.05$ ，表1)。

2.2 LASSO回归模型筛选TNBC患者pCR的影响因素 在建模集中，将年龄、乳腺癌位置、Ki-67指数、绝经状态、病理类型、BMI指数、肿瘤分级、手术术式、是否使用铂类药物、临床疗效评估、T分期和N分期等12个因素纳入LASSO回归模型，通过10次折叠交叉验证确定最佳 λ 值为0.041(图1)；在最佳 λ 值下，筛选出T分期、N分期、是否使用铂类药物及临床疗效评估为TNBC患者新辅助化疗后pCR的独立影响因素($P<0.05$)。



通过10次折叠交叉验证，对于每一个 λ 值，计算其相应的二项式系数，最终筛选出最佳 λ 值，上横坐标红色数字代表不同的 $\log(\lambda)$ 下参与计算的变量数目

图1 通过LASSO算法确定最佳 λ 值

Fig.1 Determine the best λ value by LASSO algorithm

2.3 基于LASSO回归筛选出的最佳变量构建列线图模型 在建模集中用筛选出的4个最佳变量构建多因素logistic回归模型(图2)。依据多因素logistic回归模型分析的结果绘制列线图，将各个变量所得分值相加，得到的总分可预测TNBC患者新辅助化疗后的pCR率(图3)。

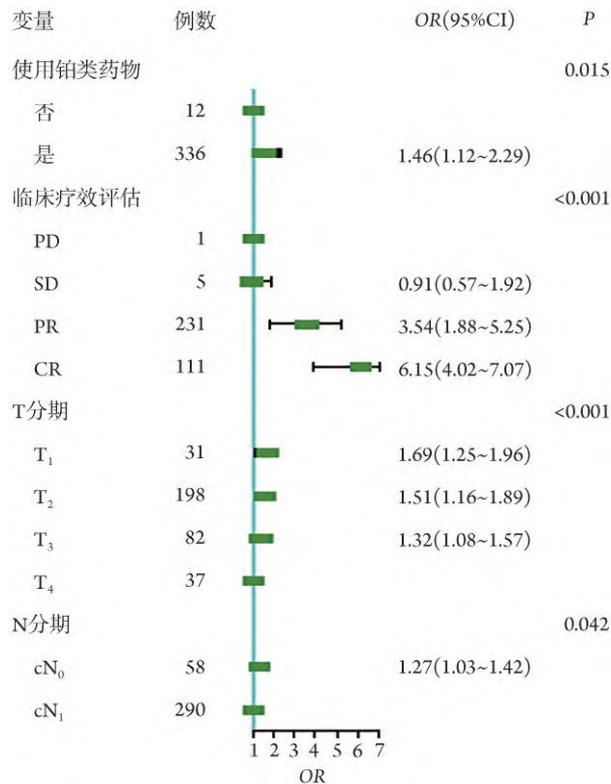
2.4 列线图模型的内部及外部验证 Bootstrap法重抽样1000次内部验证的C-index为0.79，表明列线图模型在建模集和验证集中均具有良好的区分度。ROC曲线分析显示，建模集与验证集中列线图模型预测TNBC患者新辅助化疗后pCR率的AUC分别为0.811(95%CI 0.763~0.859)、0.801(95%CI 0.727~0.928)

表1 建模集与验证集中TNBC患者的临床病理特征比较 [例(%)]

Tab. 1 Clinicopathologic features of TNBC patients in the training cohort and external validation cohort

临床病理特征	建模集(n=348)	验证集(n=69)	χ^2	P
是否pCR			0.567	0.451
是	96(27.6)	16(23.2)		
否	252(72.4)	53(76.8)		
乳腺癌位置			0.208	0.648
左侧	151(43.4)	32(46.4)		
右侧	197(56.6)	37(53.6)		
Ki-67指数			0.923	0.337
$\leq 44\%$	125(35.9)	29(42.0)		
$> 44\%$	223(64.1)	40(58.0)		
绝经状态			0.460	0.497
绝经前	202(58.0)	37(53.6)		
绝经后	146(42.0)	32(46.4)		
病理类型			2.826	0.093
浸润性导管癌	325(93.4)	68(98.6)		
非浸润性导管癌	23(6.6)	1(1.4)		
BMI(kg/m ²)			3.978	0.137
< 18.5	31(8.9)	6(8.7)		
18.5~25	189(54.3)	29(42.0)		
≥ 25	128(36.8)	34(49.3)		
肿瘤分级			3.005	0.223
I	15(4.3)	0		
II	119(34.2)	24(34.8)		
III	214(61.5)	45(65.2)		
手术术式			8.442	0.004
保乳术	92(26.4)	7(10.1)		
乳房全切术	256(73.6)	62(89.9)		
是否使用铂类药物			0.001	0.714
是	12(3.4)	3(4.3)		
否	336(96.6)	66(95.7)		
临床疗效			1.573	0.649
PD	1(0.3)	0		
SD	5(1.4)	1(1.4)		
PR	231(66.4)	50(72.5)		
CR	111(31.9)	18(26.1)		
T分期			31.210	< 0.001
T ₁	31(8.9)	0		
T ₂	198(56.9)	31(44.9)		
T ₃	82(23.6)	37(53.6)		
T ₄	37(10.6)	1(1.4)		
N分期			0.193	0.661
cN ₀	58(16.7)	13(18.8)		
cN ₁	290(83.3)	56(81.2)		

TNBC. 三阴性乳腺癌；pCR. 病理完全缓解；BMI. 体重指数；PD. 疾病进展；SD. 疾病稳定；PR. 部分缓解；CR. 完全缓解



TNBC. 三阴性乳腺癌; pCR. 病理完全缓解; PD. 疾病进展; SD. 疾病稳定; PR. 部分缓解; CR. 完全缓解

图2 TNBC患者新辅助化疗后pCR影响因素的森林图

Fig. 2 Forest map of factors affecting PCR in TNBC patients after neoadjuvant chemotherapy

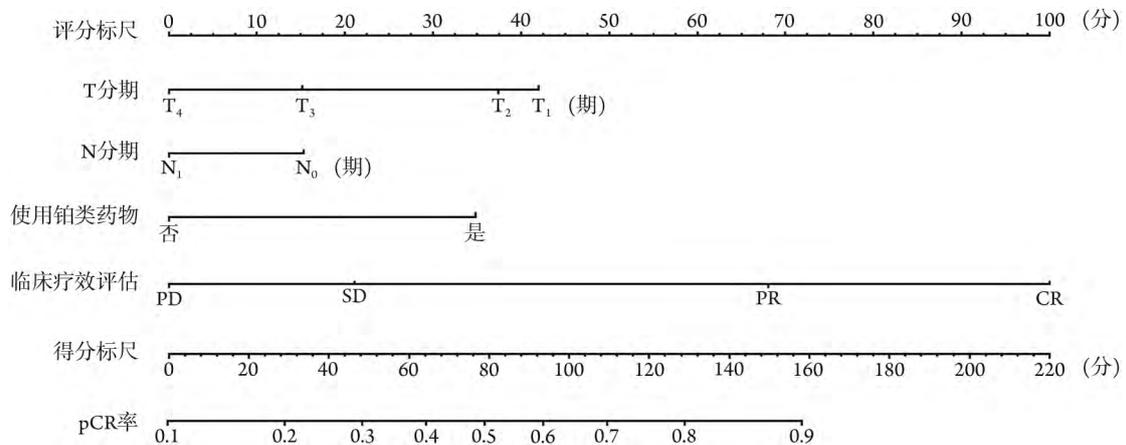
(图4); 校准曲线显示, 所有校正曲线均与理想曲线有较好的吻合度, 提示列线图模型的准确性较高(图5)。

2.5 列线图模型的临床效能分析 针对建模集和验证集的pCR率分别绘制DCA曲线, 结果如图6所示, 在建模集和验证集中, 列线图模型的临床净获益均较高, 显示模型的临床效能较好。

3 讨论

乳腺癌的新辅助化疗是指在手术前进行的全身系统性化疗, 主要应用于无远处转移的初治乳腺癌患者^[7]。新辅助化疗能够将不可手术的乳腺癌降期为可手术的乳腺癌, 使部分无法手术的患者获得手术治疗的机会。同时, 无法接受保乳手术的患者通过新辅助化疗使肿瘤降期后, 可获得保乳手术的机会。近年来, 随着新药物的出现和治疗理念的更新, 新辅助化疗的应用越来越广泛, 而新辅助化疗后能否达到pCR已成为后续辅助治疗方案的重要依据^[8]。目前多项研究表明, TNBC新辅助化疗后能否达到pCR受到多种因素的影响^[9-10], 但目前尚缺乏精准的临床预测模型。列线图作为一种新型预测模型, 相较传统预测方法准确性高、适应性广, 且易于推广^[11]。本研究通过LASSO回归模型分析, 构建列线图模型来预测TNBC患者的pCR率, 以为临床的诊疗提供指导依据。传统的诊断试验指标如敏感度、特异度及AUC等仅能反映模型的准确性, 不能反映在临床实践中模型的实际临床效用, 而DCA能够从临床决策的实际出发, 将患者或决策者的偏好整合到分析中^[12]。因此, 本研究同时使用DCA进行模型评价。

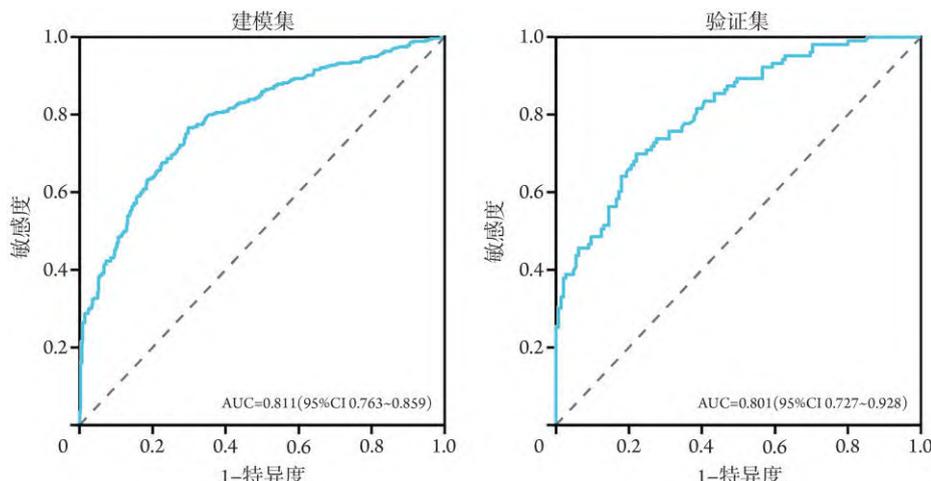
本研究中, 建模集的pCR率为27.6%, 验证集为23.2%, 与Huang等^[13]的研究结果一致, 其结果显示, TNBC患者新辅助化疗的pCR率为16.7%~67.0%。本研究建模集和验证集的保乳手术率分别为26.4%和10.1%, 差异有统计学意义。这可能与建模集和验证集的数据分别来自不同的医院有关。乳腺癌进行新辅助化疗的目的为使肿瘤降期、提高保乳率, 但由于新辅助治疗对肿瘤病灶产生的复杂作用可能对手术术式、瘤床的确定、切除范围的界定及切缘均有



TNBC. 三阴性乳腺癌; pCR. 病理完全缓解; PD. 疾病进展; SD. 疾病稳定; PR. 部分缓解; CR. 完全缓解

图3 TNBC患者新辅助化疗后pCR率的列线图预测模型

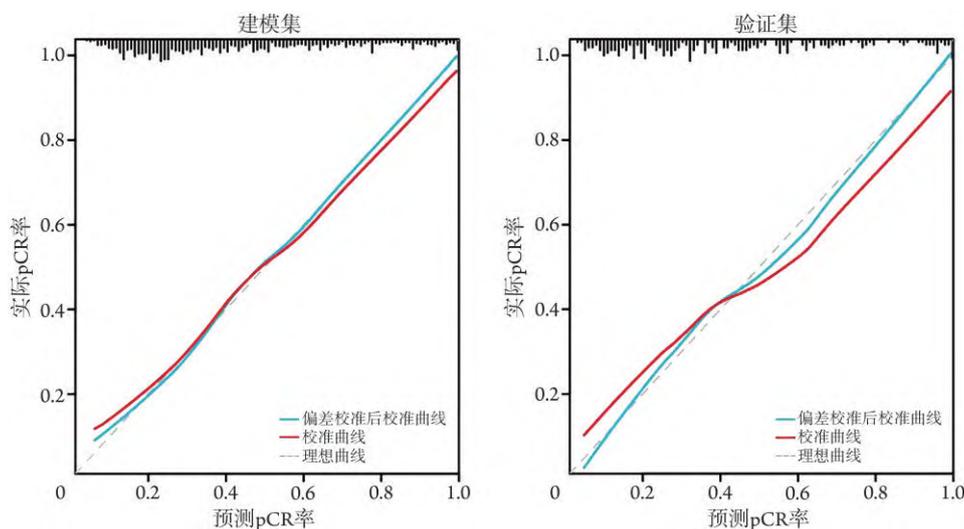
Fig. 3 Nomogram prediction model for pCR rate of TNBC patients after neoadjuvant chemotherapy



TNBC. 三阴性乳腺癌; pCR. 病理完全缓解

图4 建模集和验证集中列线图预测TNBC患者新辅助化疗后pCR率的ROC曲线

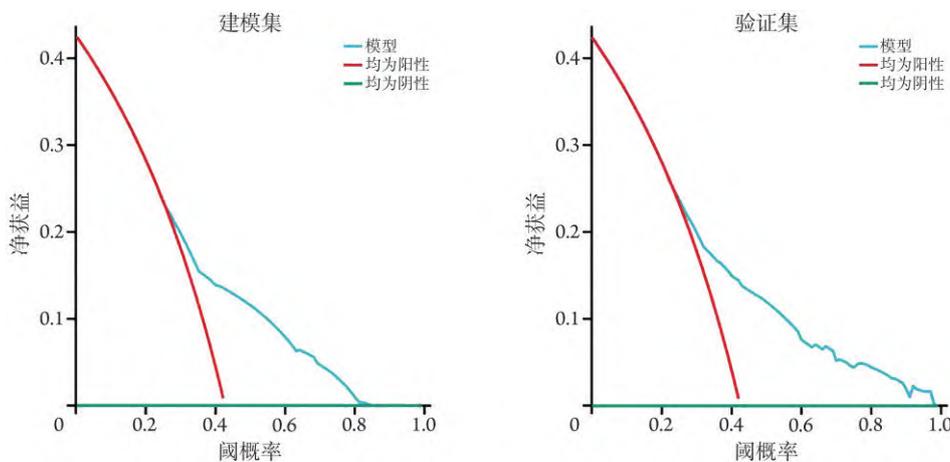
Fig.4 ROC curves of nomogram predicting pCR rate in TNBC patients after neoadjuvant chemotherapy using modeling set and validation set



pCR. 病理完全缓解; X轴为预测的pCR率, Y轴为实际的pCR率, 对角线为理想曲线表示两者完全一致

图5 建模集和验证集中的校准曲线

Fig.5 Calibration curves of modeling set and validation set



DCA. 临床决策曲线分析; 紫色横线代表所有样本均为阴性, 获益为0; 红色斜线表示所有样本均为阳性, 净获益为负值的反斜线; 蓝色曲线为模型的净获益情况

图6 建模集和验证集的DCA曲线

Fig.6 DCA curves of modeling set and verification set

影响,不同医院、不同医师对于乳腺癌新辅助治疗后的保乳手术理念可能也不尽相同,这可能导致不同医院的保乳手术率不同。本研究中,建模集和验证集的T分期差异有统计学意义。建模集中有更高比例的T₂期患者(56.9% vs. 44.9%),验证集中有更高比例的T₃期患者(53.6% vs. 23.6%),表明建模集纳入了更多肿瘤体积较小的患者,这也可能是建模集保乳手术率较高的原因之一。本研究结果显示,T分期、N分期、是否使用铂类药物及临床疗效评估是TNBC患者新辅助化疗后pCR的独立影响因素;随着T分期的增加,TNBC患者新辅助化疗后的pCR率逐渐降低,与Han等^[14]的研究结果一致。本研究中,N分期是pCR的独立影响因素(cN₀ vs. cN₁: OR=1.27, 95%CI 1.03~1.42, P=0.042)。但Zhang等^[15]发现,N分期与TNBC患者新辅助化疗后的pCR无关,可能是由于该研究仅纳入了10例cN₀患者,样本量较小。有研究发现,含有铂类药物的化疗方案可明显改善TNBC患者新辅助化疗的疗效^[16],与本研究结果一致。本研究中,使用铂类药物可明显提升TNBC患者新辅助化疗后的pCR率(使用铂类药物 vs. 未使用铂类药物: OR=1.46, 95%CI 1.12~2.29, P=0.015)。在验证集中进行外部验证,并通过Bootstrap方法(1000次重抽样)进行内部验证,采用C-index、ROC曲线和校正曲线评估模型的区分度和可靠性,最后通过DCA来评价模型的临床获益和应用价值。结果显示,列线图模型能准确地预测TNBC患者新辅助化疗后的pCR率。

综上所述,本研究结果表明,T分期、N分期、是否使用铂类药物及临床疗效评估是TNBC患者新辅助化疗后pCR的独立影响因素,据此构建的列线图模型能够准确预测TNBC患者新辅助化疗后的pCR率。但本研究仍存在不足之处:许多分子靶标如caveolin1/2、Survivin、乳腺癌易感基因1/2(breast cancer susceptibility gene 1/2, BRCA1/2)、PIK3CA (phosphatidylinositol-4, 5-bisphosphate 3-kinase catalytic subunit alpha)有可能是TNBC新辅助化疗疗效的预测因子^[17],本研究模型构建过程中未能纳入这些可能影响新辅助化疗疗效的分子靶标,未来可能需要对模型进行进一步的优化和验证。近年来,人工智能和机器学习也逐渐广泛应用于临床中,机器学习具有准确性高、可扩展性强的特点,将人工智能和机器学习应用于临床预测模型的构建中,可能会进一步提高模型的准确性。

【参考文献】

- [1] Ferlay J, Colombet M, Soerjomataram I, et al. Cancer statistics for the year 2020: an overview[J]. *Int J Cancer*, 2021. doi: 10.1002/ijc.33588.
- [2] Vagia E, Mahalingam D, Cristofanilli M. The landscape of targeted therapies in TNBC[J]. *Cancers (Basel)*, 2020, 12(4): 916.
- [3] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019[J]. *CA Cancer J Clin*, 2019, 69: 7-34.
- [4] da Silva JL, Cardoso Nunes NC, Izetti P, et al. Triple negative breast cancer: a thorough review of biomarkers[J]. *Crit Rev Oncol Hematol*, 2020, 145: 102855.
- [5] Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)[J]. *Eur J Cancer*, 2009, 45(2): 228-247.
- [6] Dieci MV, Griguolo G, Bottosso M, et al. Impact of estrogen receptor levels on outcome in non-metastatic triple negative breast cancer patients treated with neoadjuvant/adjuvant chemotherapy [J]. *NPJ Breast Cancer*, 2021, 7(1): 101.
- [7] 中国乳腺癌新辅助治疗专家组. 中国乳腺癌新辅助治疗专家共识(2019年版)[J]. *中国癌症杂志*, 2019, 29(5): 390-400.
- [8] von Minckwitz G, Huang CS, Mano MS, et al. Trastuzumab emtansine for residual invasive HER2-positive breast cancer[J]. *N Engl J Med*, 2019, 380(7): 617-628.
- [9] Lee JS, Yost SE, Yuan Y. Neoadjuvant treatment for triple negative breast cancer: recent progresses and challenges[J]. *Cancers (Basel)*, 2020, 12(6): 1404.
- [10] Tarantino P, Gandini S, Trapani D, et al. Immunotherapy addition to neoadjuvant chemotherapy for early triple negative breast cancer: a systematic review and meta-analysis of randomized clinical trials [J]. *Crit Rev Oncol Hematol*, 2021, 159: 103223.
- [11] Li G, Tian ML, Bing YT, et al. Nomograms predict survival outcomes for distant metastatic pancreatic neuroendocrine tumor: a population based STROBE compliant study[J]. *Medicine (Baltimore)*, 2020, 99(13): e19593.
- [12] Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis[J]. *JAMA*, 2015, 313(4): 409-410.
- [13] Huang M, O'Shaughnessy J, Zhao J, et al. Association of pathologic complete response with long-term survival outcomes in triple-negative breast cancer: a meta-analysis[J]. *Cancer Res*, 2020, 80(24): 5427-5434.
- [14] Han Y, Wang J, Xu B. Novel biomarkers and prediction model for the pathological complete response to neoadjuvant treatment of triple-negative breast cancer[J]. *J Cancer*, 2021, 12(3): 936-945.
- [15] Zhang F, Huang M, Zhou H, et al. A nomogram to predict the pathologic complete response of neoadjuvant chemotherapy in triple-negative breast cancer based on simple laboratory indicators [J]. *Ann Surg Oncol*, 2019, 26(12): 3912-3919.
- [16] Poggio F, Bruzzone M, Ceppi M, et al. Platinum-based neoadjuvant chemotherapy in triple-negative breast cancer: a systematic review and meta-analysis[J]. *Ann Oncol*, 2018, 29(7): 1497-1508.
- [17] Lee YM, Oh MH, Go JH, et al. Molecular subtypes of triplenegative breast cancer: understanding of subtype categories and clinical implication[J]. *Genes Genomics*, 2020, 42(12): 1381-1387.

(责任编辑:纪方方)