

# 乳腺癌化疗致骨髓抑制风险预测模型的系统评价<sup>△</sup>

刘 阳<sup>1,2\*</sup>, 李红健<sup>1</sup>, 吴建华<sup>1</sup>, 刘学涛<sup>2</sup>, 焦 敏<sup>1</sup>, 于鲁海<sup>1#</sup> (1. 新疆维吾尔自治区人民医院药学部, 乌鲁木齐 830001; 2. 石河子大学药学院, 新疆 石河子 832000)

中图分类号 R979.1 文献标志码 A 文章编号 1001-0408(2025)05-0612-07

DOI 10.6039/j.issn.1001-0408.2025.05.19



**摘要** **目的** 系统评价乳腺癌化疗致骨髓抑制的风险预测模型,为临床医疗工作者选择或开发有效预测模型提供科学的参考依据。**方法** 系统检索中国知网、维普网、万方数据库、PubMed、Web of Science、Cochrane Library、Embase、Scopus数据库中有关乳腺癌化疗致骨髓抑制风险预测模型的研究,检索时限为建库至2024年5月7日。由2名研究者独立筛选文献,根据预测模型系统评价的严格评估和数据清单提取数据,并采用预测模型研究的偏倚风险评价工具分析纳入研究的偏倚风险和适用性。**结果** 共纳入7项研究,包含12个模型,其中11个模型报告了受试者工作特征曲线下面积,为0.600~0.908;2个模型报告了校准方法;纳入模型常见的预测变量为年龄、化疗前中性粒细胞计数、化疗前淋巴细胞计数、化疗前白蛋白含量。7项研究整体偏倚风险高(主要因为研究设计缺陷、样本量不足、变量处理方式不当、未报告缺失数据、模型评估指标缺乏等)但适用性良好。**结论** 乳腺癌化疗致骨髓抑制风险预测模型的预测性能有待进一步提升,且模型整体偏倚风险高。未来的研究应遵循模型开发与报告规范,并结合机器学习算法开发出预测性能好、稳定性强、偏倚风险低的风险预测模型,为临床提供决策依据。

**关键词** 乳腺癌;化疗;骨髓抑制;风险预测模型;系统评价

## Systematic review of risk predictive models for chemotherapy-induced myelosuppression in breast cancer

LIU Yang<sup>1,2</sup>, LI Hongjian<sup>1</sup>, WU Jianhua<sup>1</sup>, LIU Xuetao<sup>2</sup>, JIAO Min<sup>1</sup>, YU Luhai<sup>1</sup> (1. Dept. of Pharmacy, Xinjiang Uygur Autonomous Region People's Hospital, Urumqi 830001, China; 2. School of Pharmacy, Shihezi University, Xinjiang Shihezi 832000, China)

**ABSTRACT** **OBJECTIVE** To systematically evaluate risk prediction models for chemotherapy-induced myelosuppression in breast cancer, and provide a scientific reference for clinical healthcare workers in selecting or developing effective predictive models. **METHODS** A systematic search was conducted for studies on predictive models of the risk of chemotherapy-induced myelosuppression in breast cancer across the CNKI, VIP, Wanfang, PubMed, Web of Science, Cochrane Library, Embase, and Scopus databases, with a time frame of the establishment of the database to May 7, 2024. Literature was independently screened by 2 investigators, data were extracted according to critical appraisal and data extraction for systematic reviews of predictive model studies, and the risk of bias evaluation tool for predictive model studies was used to analyze the risk of bias and applicability of the included studies. **RESULTS** There were totally 7 studies, comprising 12 models. Among them, 11 models indicated an area under the subject operating characteristic curve of 0.600-0.908; 2 models indicated calibration. The common predictor variables of the included models were age, pre-chemotherapy neutrophil count, pre-chemotherapy lymphocyte count, and pre-chemotherapy albumin. The overall risk of bias of the 7 studies was high, which was mainly attributed to the flaws in the study design, insufficient sample sizes, inappropriate treatment of variables, non-reporting of missing data, and the lack of indicators for the assessment of the models, but the applicability was good. **CONCLUSIONS** The predictive performance of risk predictive models for chemotherapy-induced myelosuppression in breast cancer remains to be further enhanced, and the overall risk of model bias is high. Future studies should follow the specifications of model development and reporting, then combine machine learning algorithms to develop risk predictive models with good predictive performance, high stability, and low risk of bias, so as to provide a decision-making basis for the clinic.

**KEYWORDS** breast cancer; chemotherapy; myelosuppression; risk predictive model; systematic review

<sup>△</sup>基金项目“天山英才”医药卫生高层次人才培养计划项目(No. TSYC202301A028)

\* 第一作者 硕士研究生。研究方向:临床药学。E-mail: 2363901559@qq.com

# 通信作者 主任药师,硕士生导师,硕士。研究方向:医院药学、药物分析。E-mail: 1523264450@qq.com

乳腺癌发病日益趋向年轻化,其发病率及致死率均呈现逐年攀升的态势,对女性的生命健康构成了严峻威胁,已成为重大的社会性健康问题。调查显示,2022年全球新发乳腺癌病例数达到约230.9万例,发病率位列

全球癌症第二,并约有66.6万人因乳腺癌而死亡,在全球癌症死亡原因中位列第四<sup>[1]</sup>。

乳腺癌的治疗方法包括手术、化疗、放疗和内分泌激素治疗,其中化疗在乳腺癌治疗中占据核心地位。尽管化疗在控制疾病进展方面发挥着重要作用,但同时也伴随着一系列副作用,其中骨髓抑制是肿瘤化疗中最常见且严重的毒性反应<sup>[2]</sup>,通常表现为贫血、中性粒细胞减少、血小板减少和/或淋巴细胞减少<sup>[3-4]</sup>。由此产生的感染、贫血和潜在自发性出血会增加死亡风险,甚至因其相关并发症而导致患者死亡。化疗所致骨髓抑制会导致化疗剂量减少、治疗延迟或中断抗肿瘤治疗,从而影响治疗效果和患者的生存质量<sup>[5]</sup>。有研究报告指出,4级骨髓抑制相关病死率可达4%~12%<sup>[6]</sup>。

风险预测模型能够帮助医务工作者识别高风险患者,提前预测疾病发生的可能性,从而实现早期干预和预防,提高患者的生存质量<sup>[7]</sup>。因此,早期利用风险预测模型识别骨髓抑制高风险患者并及时采取相应的预防和治理措施,不仅能提高肿瘤治疗的有效性,还能显著减轻相关并发症带来的疾病负担<sup>[8]</sup>。在医疗大数据的推动下,国内外研究者开始关注如何利用临床数据构建预测模型来评估乳腺癌患者化疗所致骨髓抑制的风险。然而,目前缺乏对该类模型的系统性评价研究,且现有的预测模型在构建方法、预测效能以及泛化能力方面存在差异,缺乏系统梳理。为此,本研究拟系统地评价乳腺癌化疗所致骨髓抑制风险预测模型的偏倚风险及适用性,旨在为临床医疗工作者选择或开发有效预测模型提供参考依据。

## 1 资料与方法

### 1.1 研究问题的确立

根据PICOTS原则明确研究问题。P(patient):确诊为乳腺恶性肿瘤,并接受化疗的患者;I(intervention):所有已开发、可用于预测骨髓抑制的风险预测模型;C(comparator):不适用;O(outcome):骨髓抑制;T(timing):化疗前预测;S(setting):任何环境。本项研究已在国际系统综述平台注册,注册号为CRD42024607602。

### 1.2 文献纳入与排除标准

纳入标准包括:(1)研究对象为接受化疗的乳腺恶性肿瘤患者;(2)研究内容为骨髓抑制风险预测模型的构建、验证;(3)研究类型为队列研究、病例对照研究、横断面研究或随机对照研究;(4)干预措施为不同的化疗方案等;(5)结局指标包括白细胞计数、中性粒细胞计数、血小板计数、血红蛋白含量,其中之一即可;(6)对重复发表的文章,只纳入发表日期最近的1篇;(7)中文或英文发表的文章。

排除标准包括:(1)研究对象为同时接受化疗、放疗或内分泌治疗的乳腺恶性肿瘤患者;(2)只进行相关危险因素分析,未建立具体预测模型的研究;(3)会议摘

要、新闻报道等非正式发表的文献;(4)综述、述评等文献;(5)动物、细胞和分子水平的研究;(6)无法从公开渠道获取全文的文献;(7)病例报告。

### 1.3 文献检索策略

使用主题词与自由词相结合的方式对中国知网、维普网、万方数据库、PubMed、Web of Science、Cochrane Library、Embase、Scopus共8个数据库进行系统检索,检索时限均为建库至2024年5月7日。中文数据库以中国知网为例,检索策略如下:((主题:乳腺癌)OR(篇文摘:乳腺癌+乳腺肿瘤))AND((主题:骨髓抑制)OR(篇文摘:骨髓抑制+中性粒细胞减少+白细胞减少+血小板减少+贫血+血三系减少+血红蛋白减少+血液毒性))AND((主题:预测)OR(篇文摘:预测+风险评估+风险评估+模型+列线图+预测因子+风险预测+诊断模型+临床预测模型+风险因素+筛查模型+预测因素+筛查工具))。英文数据库以PubMed为例,检索策略如下:(“breast neoplasms”[MeSH Terms]OR“breast cancer”[Title/Abstract]OR“breast neoplasm\*”[Title/Abstract]OR“breast carcinoma”[Title/Abstract]OR“breast tumor\*”[Title/Abstract])AND(“myelosuppression”[Title/Abstract]OR“bone marrow suppression”[Title/Abstract]OR“neutropenia”[MeSH Terms]OR“neutropenia\*”[Title/Abstract]OR“leukopenia”[MeSH Terms]OR“leukopenia\*”[Title/Abstract]OR“leukocytopenia\*”[Title/Abstract]OR“thrombocytopenia”[MeSH Terms]OR“thrombocytopenia\*”[Title/Abstract]OR“thrombopenia\*”[Title/Abstract]OR“anemia”[MeSH Terms]OR“anemia\*”[Title/Abstract])AND(“predict\* model”[Title/Abstract]OR“prognos\* model”[Title/Abstract]OR“risk predict\*”[Title/Abstract]OR“risk assessment”[Title/Abstract]OR“predictive factor\*”[Title/Abstract]OR“nomogram”[Title/Abstract]OR“machine learning”[MeSH Terms]OR“deep learning”[Title/Abstract]OR“machine intelligence”[Title/Abstract])。

### 1.4 文献筛选与资料提取

由2名经过系统化循证伦理学习培训过的研究者独立筛选文献、提取资料并交叉核对,发生分歧时,寻求第三方仲裁裁决。资料提取依据预测模型系统评价的严格评估和数据清单(critical appraisal and data extraction for systematic reviews of prediction modeling studies, CHARMS)<sup>[9]</sup>制定标准化表格,主要包括第一作者、发表年份、国家、数据来源、预测变量数、结局信息等。

### 1.5 纳入文献的偏倚风险与适用性评价

由2名研究者独立使用预测模型研究的偏倚风险评估工具(prediction model risk of bias assessment tool, PROBAST)<sup>[10]</sup>对纳入的模型进行质量评价,若评价结果存在分歧,则咨询第三方研究人员。PROBAST旨在评

价诊断和预后预测等各类模型的偏倚风险和适用性。

### 1.5.1 偏倚风险评估

偏倚风险评估包括研究对象、预测变量、结局和统计分析4个领域,共20个信号问题。研究者可参照标准以“是”“可能是”“否”“可能否”“不清楚”对各领域风险进行评估。具体标准为:当某一领域所有信号问题的评判结果为“是”或“可能是”,则评定该领域为低风险偏倚;若该领域信号问题的评判结果出现1个或以上的“否”或“可能否”,则评定该领域为高风险偏倚;若该领域信号问题的评判结果出现1个或以上的“不清楚”,同时其他信号问题被评判为“是”或“可能是”,则评定该领域为偏倚风险不清楚。当所有领域均为低风险偏倚时,则评定该研究整体为低风险偏倚;当某一领域为高风险偏倚时,则评定该研究整体为高风险偏倚;当某一领域偏倚风险为不清楚,而其他领域为低风险偏倚时,则评定该研究整体的偏倚风险为不清楚。

### 1.5.2 适用性评估

适用性评估包括研究对象、预测变量和结局3个领域,研究者可采用“适用性好”“适用性差”或“适用性不清楚”对各领域的适用性进行评价。若所有领域均被评判为“适用性好”,则评定该研究整体适用性好;若任一领域被评判为“适用性差”,则评定该研究整体适用性差;若存在1个或以上的领域被评判为“适用性不清楚”,其他领域为“适用性好”,则评定该研究整体适用性不清楚。

## 2 结果

### 2.1 文献筛选结果

本研究通过初步检索数据库,共获取3 509篇文献,去除重复文献后剩余2 213篇文献。经过逐级筛选后,最终纳入7篇文献<sup>[11-17]</sup>。文献筛选流程见图1。

### 2.2 纳入研究的基本特征

7篇文献涉及7项研究<sup>[11-17]</sup>,发表时间为2008—2024年。其中,4项研究<sup>[14-17]</sup>在国内开展,另各有1项研究在加拿大<sup>[11]</sup>、英国<sup>[12]</sup>、韩国<sup>[13]</sup>开展;研究类型方面,5项研究<sup>[11-15]</sup>为回顾性队列研究,1项研究<sup>[16]</sup>为回顾性病例对照研究,1项研究<sup>[17]</sup>为前瞻性队列研究;研究对象分布方面,3项研究<sup>[11-12,16]</sup>对研究对象的化疗方案具有限制;在结局评估方面,2项研究<sup>[13-14]</sup>采用美国国立综合癌症网

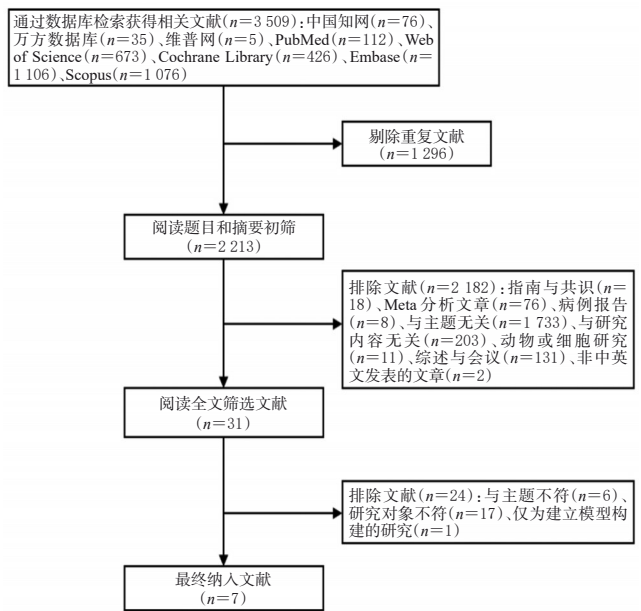


图1 文献筛选流程图

络临床实践指南,2项研究<sup>[15-16]</sup>采用世界卫生组织(World Health Organization, WHO)抗肿瘤药物急性及亚急性毒性反应分度标准,2项研究<sup>[11,17]</sup>采用美国国家癌症研究所的标准,1项研究<sup>[12]</sup>采用欧洲肿瘤协会的标准。纳入研究的基本特征见表1。

### 2.3 模型建立情况

纳入的7项研究<sup>[11-17]</sup>共构建了12个模型,模型的候选预测变量有9~27个;在连续变量处理方法中,4项研究<sup>[11,15-17]</sup>将连续变量处理成分类变量;各研究样本总量为119~933例,结局事件数为31~409例;各研究均未报告缺失数据和缺失数据处理方法,均未报告数据的复杂性处理情况。模型建立方面,3项研究<sup>[15-17]</sup>使用逻辑回归,1项研究<sup>[12]</sup>根据绝对淋巴细胞计数和绝对中性粒细胞计数自定义危险分层,1项研究<sup>[14]</sup>为改良Jenkins模型,1项研究<sup>[11]</sup>使用广义估计方程,1项研究<sup>[13]</sup>使用机器学习方法构建逻辑回归(logistic regression, LR)、极限梯度提升(extreme gradient boosting, XGB)、最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)、支持向量机(support vector machine, SVM)、人工神经网络(artificial neural network, ANN)、决策树(decision tree, DT)模型。变量筛选方面,5项研

表1 纳入研究的基本特征

纳入研究	发表年份	国家	研究类型	研究对象	结局评估方法	预测结局
Dranitsaris <sup>[11]</sup>	2008	加拿大	回顾性队列研究	接受阿霉素或多柔比星脂质体化疗的乳腺癌患者	美国国家癌症研究所的标准	中性粒细胞减少症≥Ⅱ级、发热性中性粒细胞减少症或中性粒细胞减少症伴感染
Jenkins <sup>[12]</sup>	2012	英国	回顾性队列研究	接受多西他赛和阿霉素联合环磷酰胺化疗方案的乳腺癌患者	欧洲医学肿瘤学会的标准	发热性中性粒细胞减少症
Cho <sup>[13]</sup>	2020	韩国	回顾性队列研究	乳腺癌术后接受化疗的患者	美国国立综合癌症网络临床实践指南	发热性中性粒细胞减少症
Chen <sup>[14]</sup>	2014	中国	回顾性队列研究	接受基于蒽环类或紫杉醇类化疗方案的乳腺癌患者	美国国立综合癌症网络临床实践指南	化疗第1个周期发热性中性粒细胞减少
李亚玲 <sup>[15]</sup>	2024	中国	回顾性队列研究	接受化疗的乳腺癌患者	WHO抗肿瘤药物急性及亚急性毒性反应分度标准	骨髓抑制
黄家良 <sup>[16]</sup>	2023	中国	回顾性病例对照研究	接受多柔比星联合环磷酰胺序贯紫杉醇化疗方案的三阴性乳腺癌患者	WHO抗肿瘤药物急性及亚急性毒性反应分度标准	骨髓抑制>Ⅱ度
Chang <sup>[17]</sup>	2013	中国	前瞻性队列研究	接受基于蒽环类或紫杉醇类化疗方案的乳腺癌患者	美国国家癌症研究所的标准	中性粒细胞减少≥Ⅳ级、发热性中性粒细胞减少、化疗剂量减少、化疗周期延迟、感染

究<sup>[11,14-17]</sup>采用先单因素分析后再进行多因素分析,1项研究<sup>[13]</sup>采用递归特征消除结合单因素与多因素分析,1项研究<sup>[12]</sup>仅关注先前发布模型的外部验证,不适用。模型建立情况见表2。

## 2.4 模型性能与预测变量

纳入的7项研究中,6项研究<sup>[11,13-17]</sup>报告了受试者工作特征曲线下面积(area under the receiver operating characteristic curve, AUC),为0.600~0.908,1项研究<sup>[12]</sup>采用灵敏度、特异度、阳性预测值和阴性预测值来评估模型性能;模型校准方法方面,5项研究<sup>[11-14,17]</sup>未报告校准方法,1项研究<sup>[15]</sup>采用H-L检验并绘制校准曲线,1项研究<sup>[16]</sup>只采用了H-L检验且 $P>0.05$ ,表示模型预测值与实际观测值之间的差异不具有统计学意义;模型验证方面,5项研究<sup>[11,13-15,17]</sup>采用内部验证方法,其中1项研究<sup>[13]</sup>采用了五倍交叉验证方法,2项研究<sup>[12,16]</sup>采用外部验证方法。模型开发中考虑的候选预测变量包括社会人口统计、实验室检查、基础疾病和治疗以及生活方式行为的相关因素,而最终预测模型中出现次数较多的预测变量为年龄、化疗前淋巴细胞计数、化疗前中性粒细胞计数等,以上预测变量容易获取和测量,在基层医疗机构也可进行评估。模型最终的呈现形式各有不同,2项研究<sup>[12,14]</sup>为风险分层,2项研究<sup>[11,17]</sup>为风险评分,1项研究<sup>[15]</sup>为列线图,2项研究<sup>[13,16]</sup>未报告模型最终呈现形式。模型性能与预测变量见表3。

## 2.5 偏倚风险和适用性评价

7项研究<sup>[11-17]</sup>的整体偏倚风险均被评为高风险;6项研究<sup>[11-12,14-17]</sup>总体适用性好,1项研究<sup>[13]</sup>总体适用性差。结果见表4。

### 2.5.1 研究对象领域

纳入的7项研究中,6项研究<sup>[11-16]</sup>被评为高偏倚风险。原因是其中5项研究<sup>[11-15]</sup>为回顾性队列研究,存在数据完整性和准确性的问题,并且存在一定的选择偏移,1项研究<sup>[16]</sup>的对象来源于非巢式病例对照研究(巢式病例对照研究较传统病例对照研究可以更好地确定风险因素与结局之间的时序关系以及便于控制混杂因素<sup>[18]</sup>)。

### 2.5.2 预测变量领域

2项研究<sup>[13,16]</sup>被评为高偏倚风险,其中1项研究<sup>[13]</sup>在应用模型时,无法获得全部预测变量信息;另外1项研究<sup>[16]</sup>为病例对照研究,而病例对照研究是在已发生结果的情况下,追溯性收集可能相关的预测变量信息,研究人员在数据收集时无法实现盲法操作,使得测量过程会受到结局的潜在影响。5项研究<sup>[11-12,14-15,17]</sup>被评为偏倚风险不清楚,其预测变量的评估可能受到了偏倚影响,原因在于文献中没有明确评估者是否在不了解结局指标的情况下进行评估,即研究人员在预测变量的评估过程中是否采用盲法尚不能确定。

表2 风险预测模型建立的基本情况

纳入研究	候选预测变量/个	连续变量处理方法	样本例数	结局事件例数	缺失数据	缺失数据处理方法	建模方法	变量筛选方法
Dranitsaris <sup>[11]</sup>	10	转换为分类变量	509	39	未报告	未报告	广义估计方程	单因素分析后多因素分析
Jenkins <sup>[12]</sup>	9	保持连续变量	263	31	未报告	未报告	根据绝对中性粒细胞计数和绝对淋巴细胞计数自定义危险分层	不适用
Cho <sup>[13]</sup>	27	保持连续变量	933	409	未报告	未报告	使用机器学习方法建立LR、XGB、LASSO、SVM、ANN、DT模型	递归特征消除结合单因素与多因素分析
Chen <sup>[14]</sup>	17	保持连续变量	428	55	未报告	未报告	改良Jenkins模型	单因素分析后多因素分析
李亚玲 <sup>[15]</sup>	18	转换为分类变量	270	190	未报告	未报告	逻辑回归	单因素分析后多因素分析
黄家良 <sup>[16]</sup>	9	转换为分类变量	313	106	未报告	未报告	逻辑回归	单因素分析后多因素分析
Chang <sup>[17]</sup>	15	转换为分类变量	119	35	未报告	未报告	逻辑回归	单因素分析后多因素分析

表3 风险预测模型的性能及预测变量

纳入研究	模型性能	校准方法	验证方法	最终预测变量	模型呈现形式
Dranitsaris <sup>[11]</sup>	AUC=0.74,灵敏度0.58,特异度0.787	未报告	内部验证	含阿霉素的化疗方案、年龄、WHO患者体能状况分级、前一个化疗周期中性粒细胞计数下降至 $\leq 2 \times 10^9 L^{-1}$ 、第1次接受化疗	风险评分
Jenkins <sup>[12]</sup>	灵敏度、特异度、阳性预测值、阴性预测值分别为0.23、0.94、0.33、0.90	未报告	外部验证	化疗前中性粒细胞计数、化疗前淋巴细胞计数	风险分层
Cho <sup>[13]</sup>	LR: AUC=0.87,准确率0.781,灵敏度0.878,特异度0.696;DT: AUC=0.855,准确率0.759,灵敏度0.707,特异度0.804;XGB: AUC=0.908,准确率0.816,灵敏度0.829,特异度0.804;LASSO: AUC=0.862,准确率0.805,灵敏度0.805,特异度0.804;SVM: AUC=0.88,准确率0.782,灵敏度0.829,特异度0.739;ANN: AUC=0.865,准确率0.782,灵敏度0.854,特异度0.717	未报告	内部五倍交叉验证	年龄、肿瘤分期、基于紫杉醇的化疗方案、化疗后5d的淋巴细胞计数	未报告
Chen <sup>[14]</sup>	ACU=0.60,灵敏度0.382,特异度0.812	未报告	内部验证	化疗前中性粒细胞计数、化疗前淋巴细胞计数、化疗前单核细胞计数	风险分层
李亚玲 <sup>[15]</sup>	AUC=0.755,灵敏度0.523,特异度1.0	H-L检验、校准曲线	内部验证	手术、化疗前白细胞计数、化疗前血红蛋白含量、化疗前血小板计数	列线图
黄家良 <sup>[16]</sup>	AUC=0.763,准确率0.84,灵敏度0.80,特异度0.88,阳性预测值0.87,阴性预测值0.82	H-L检验	外部验证	年龄、体重指数、肿瘤分期、糖尿病病史、化疗前白蛋白含量、化疗前淋巴细胞计数	未报告
Chang <sup>[17]</sup>	AUC=0.64,灵敏度0.49,特异度0.69	未报告	内部验证	化疗方案、基础疾病、年龄、肌酐含量、胆红素含量、肿瘤分期、既往发热中性粒细胞计数减少或中性粒细胞计数减少、化疗前淋巴细胞计数、化疗前白蛋白含量、化疗前血红蛋白含量、卡氏评分、肿瘤累及骨髓、前期广泛化疗、组织感染、开放性伤口	风险评分

表4 偏倚风险及适用性评价结果

纳入文献	偏倚风险				适用性			总体	
	研究对象	预测变量	结局	统计分析	研究对象	预测变量	结局	偏倚风险	适用性
Dranitsaris <sup>[1]</sup>	+	?	+	+	-	-	-	+	-
Jenkins <sup>[12]</sup>	+	?	?	+	-	-	-	+	-
Cho <sup>[13]</sup>	+	+	+	+	-	+	-	+	+
Chen <sup>[14]</sup>	+	?	?	+	-	-	-	+	-
李亚玲 <sup>[15]</sup>	+	?	?	+	-	-	-	+	-
黄家良 <sup>[16]</sup>	+	+	?	+	-	-	-	+	-
Chang <sup>[17]</sup>	-	?	+	+	-	-	-	+	-

“-”表示低偏倚风险/适用性好;“+”表示高偏倚风险/适用性差;“?”表示不清楚。

### 2.5.3 结局领域

3项研究<sup>[11,13,17]</sup>被评为高偏倚风险,其中2项研究<sup>[11,17]</sup>存在多个结局定义,存在偏倚风险;另外1项研究<sup>[13]</sup>的预测变量包含化疗后5 d的淋巴细胞计数,而模型的应用是在患者化疗前,即预测变量的测量与结局之间的时间间隔不合适;4项研究<sup>[12,14-16]</sup>被评为偏倚风险不清楚,原因是文献中未报告研究人员是否在不了解预测变量的情况下确定结局。

### 2.5.4 统计分析领域

所纳入的7项研究<sup>[11-17]</sup>均被评为高风险偏倚。主要原因是样本量不足,即预测变量的事件数(events per variable, EPV) < 20。4项研究<sup>[11,15-17]</sup>将连续变量转换为分类变量。7项研究<sup>[11-17]</sup>均未报道缺失数据与缺失数据处理方法。在模型构建方面,1项研究<sup>[12]</sup>为模型验证研究,未建立模型;6项研究<sup>[11,13-17]</sup>在筛选预测变量时采用单因素分析结合多因素分析,其中1项研究<sup>[13]</sup>还结合了递归特征消除;7项研究<sup>[11-17]</sup>均未报告数据的复杂性处理情况;6项研究<sup>[11,13-17]</sup>采用AUC评价模型区分度,1项研究<sup>[12]</sup>采用灵敏度、特异度等指标评价模型区分度;5项研究<sup>[11-14,17]</sup>未报告模型校准度。在模型过拟合方面,2项研究<sup>[12,16]</sup>进行了外部验证,5项研究<sup>[11,13-15,17]</sup>进行了内部验证,其中1项研究<sup>[13]</sup>采用的是五倍交叉验证。

### 2.5.5 适用性评价

在研究对象领域,7项研究<sup>[11-17]</sup>均适用性好;在预测变量领域,1项研究<sup>[13]</sup>适用性差,原因为预测变量评估时间与系统评价问题不匹配;在结局领域,7项研究<sup>[11-17]</sup>均适用性好。

## 3 讨论

乳腺癌患者数量庞大,且化疗后骨髓抑制高发<sup>[19]</sup>,给医生制订治疗方案带来了较大的挑战。目前,可通过制定和实施骨髓抑制的预防和管理策略降低化疗后骨髓抑制的发生率。风险评估可辅助医疗决策者对患者的骨髓抑制风险进行有效分层,针对高风险患者优化治疗策略,提高护理人员及患者对风险管理的认知,最终增强患者对风险管理行为的依从性。因此,建立一个科学、合理的风险预测模型对于乳腺癌患者的风险管理具有重要意义。

### 3.1 预测模型的预测性能有待提高,校准度及适用性需进一步验证

本研究最终纳入了7项研究,共报告了12个预测模型,包括11项模型开发研究和1项模型验证研究。11个模型的AUC值为0.600~0.908,表明各模型均具有良好的预测性能,但大部分研究<sup>[11-14,17]</sup>未报告模型校准度,无法评价模型预测的准确性。1项研究<sup>[16]</sup>使用H-L检验考察模型预测值与实际观测值之间的差异,但该方法容易受到样本量及数据分布的影响<sup>[20]</sup>。另外1项研究<sup>[13]</sup>使用H-L检验结合校准曲线评估模型校准度,校准曲线可直观观察预测概率与实际结果之间的关系且不受限于特定的数据分布。内部验证有助于减少模型过度拟合,而外部验证不仅能够防止模型过度拟合,还能够检验模型的泛化能力、提升研究结果的可靠性,从而增强预测模型的权威性和普适性<sup>[21]</sup>。本文纳入的研究中仅2项研究<sup>[12,16]</sup>进行了外部验证,而缺乏外部验证会影响预测模型的应用范围及决策信任度。因此,医务人员在选择或开发有效预测模型时需考虑模型的校准方法及模型内、外部验证问题。

### 3.2 预测模型总体偏倚风险高、总体适用性良好

本研究纳入的7项研究<sup>[11-17]</sup>均为高偏倚风险,主要原因为研究设计缺陷(研究类型多为回顾性研究)、未对研究人员实施盲法、应用模型时预测变量无法获取、多个结局定义、样本量不足、变量处理方式不当、未对缺失数据进行报告、模型评估指标缺乏等。由此得到对未来研究的启示:(1)回顾性研究虽然数据容易获取、收集成本低,但存在选择偏倚及数据质量问题<sup>[22]</sup>。因此,当研究时间、经济条件等允许时,应当优先考虑采用前瞻性队列研究、随机对照研究、巢式病例对照研究等方法所获取的数据作为构建风险预测模型的数据来源。(2)对收集数据的研究人员实施盲法,可进一步减少偏倚风险。(3)在筛选模型的预测变量时,必须考虑在实际应用该模型时是否能获取到所有预测变量所需的数据,以及预测变量评估与结局确定之间的时间间隔是否合适,否则将会由于缺少预测变量信息导致模型准确性降低、泛化能力及适用性受限等。(4)明确选择主要的结局定义。单一结局定义可避免多个结局定义带来的选择性报告偏倚及结局异质性,并可简化模型。(5)保证样本量充足。样本量不足(EPV < 20)会降低模型稳定性及增加模型过拟合风险<sup>[23]</sup>。(6)合理转换连续变量。不合理的分类会导致不必要的预测信息丢失,导致预测能力下降<sup>[24]</sup>,研究人员可根据预测变量的临床阈值、业内共识或采用限制性立方样条方式合理转换连续变量。(7)完善报告缺失数据。所纳入的研究中均未描述缺失数据及缺失数据的处理方法,这会导致选择偏倚<sup>[25]</sup>。(8)完善模型评估指标。在评估预测模型时,区分度和校准度至关重要,但本研究中仅有2项研究报告了校准度。未来

的模型开发和验证研究至少需要评估模型的区分能力和校准性能,以实现模型性能的全面评估。此外,未来的研究还应依据 PROBAST 来优化试验设计,减少潜在的偏倚风险,提高模型的可信度和适用性<sup>[10]</sup>。

本研究所纳入的模型总体适用性良好,其中 1 项研究适用性差的原因主要在于预测变量评估与结局时间间隔不合适。在未来的模型开发研究中,研究人员可在数据预处理阶段对预测变量进行适当清洗,以确保预测变量的评估时间点发生在结局发生之前。

### 3.3 预测模型的预测变量分析

本研究所纳入的预测模型最终筛选出的预测变量不一致,主要包括患者一般资料、基础疾病、化疗方案、实验室检查等。各模型同时出现频率较高的预测变量为年龄、化疗前中性粒细胞计数、化疗前淋巴细胞计数、化疗前白蛋白含量。Leon 等<sup>[26]</sup>研究显示,老年患者较青年患者更易发生中性粒细胞减少症,这可能是老年患者生理性免疫功能下降或机体对药物清除及排泄能力下降所致;淋巴细胞是免疫系统的重要组成部分,低淋巴细胞计数可能表明患者的免疫功能已经受损,而化疗会进一步抑制免疫系统,使得骨髓抑制的风险增加。有研究表明,通过人血白蛋白和淋巴细胞计数计算得出的预后营养指数(prognostic nutritional index, PNI)与乳腺癌患者的预后密切相关,较低的 PNI 值与乳腺癌患者较低的生存率呈现出正相关<sup>[27]</sup>。而 Chen 等<sup>[28]</sup>研究发现,乳腺癌化疗过程中低 PNI 患者较 PNI 患者更容易发生骨髓抑制。

### 3.4 预测变量筛选及预测模型建立方法

本研究所纳入的研究在预测变量筛选时多采用单因素分析结合多因素分析方法。随着人工智能技术的进步,预测变量筛选的方法日益增多,包括基于树的算法进行特征选择、自编码器技术以及应用正则化技术的模型等。与传统统计学筛选方法相比,机器学习特征选择方法在识别变量间的非线性联系和复杂的相互作用方面展现出了优势,并且在处理高维度数据时更有效<sup>[29]</sup>。在未来的模型开发研究中,研究人员可在预测变量筛选时将机器学习特征选择方法与传统统计学方法相融合,从而有助于规避依赖单一方法引入的偏倚或误差,并可提高模型的泛化能力。

本研究发现,现有的乳腺癌化疗致骨髓抑制风险预测模型多是基于传统 Logistic 回归方法建立,仅有 Cho 等<sup>[13]</sup>研究采用机器学习方法建立模型并取得了良好的预测效果。机器学习方法能够弥补传统逻辑回归和数学模型的不足之处,并且拥有出色的灵敏度、特异度和预测能力。机器学习技术已被广泛应用于开发各类疾病的预测模型,但在乳腺癌化疗所致骨髓抑制的风险预测模型研究领域,相关报道仍然不多。因此,建议在未来的研究中采用机器学习方法构建预测模型,以此弥补

传统模型存在的局限性。

### 3.5 本研究的局限性

本研究存在以下局限性:(1)研究只纳入了中、英文文献,存在发表偏倚;(2)由于纳入的研究在研究对象、模型建立方法、结局评估等方面存在异质性,故仅对纳入的研究进行了定性分析;(3)纳入的研究中,大部分研究未完整报告模型评估性能指标,影响了本研究对各模型性能的评价结果。

## 4 结语

本研究纳入了 7 项乳腺癌化疗致骨髓抑制的风险预测模型研究,共包含 12 个风险预测模型;采用 PROBAST 对模型各项特征进行系统评价的结果表明:模型预测性能有待提高,模型建立方法存在缺陷,整体偏倚风险较高,适用性良好。在未来的模型开发研究中,建议研究者们严格遵循 PROBAST 标准,确保模型构建的科学性和严谨性;同时,鼓励研究者在研究时间、经济条件等允许的情况下结合机器学习算法开展多中心、大样本的研究,开发出预测性能好、适用性广及稳定性强的预测模型,为临床提供强有力且高质量的决策依据。

## 参考文献

- [1] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. *CA Cancer J Clin*, 2024, 74(3):229-263.
- [2] LIN S Z, LEI S, LIU W, et al. Safety analysis of therapeutic drugs for breast cancer patients and construction of a predictive model for serious adverse drug reactions[J]. *Eur J Clin Pharmacol*, 2024, 80(2):249-259.
- [3] TAYLOR S J, DUYVESTYN J M, DAGGER S A, et al. Preventing chemotherapy-induced myelosuppression by repurposing the FLT3 inhibitor quizartinib[J]. *Sci Transl Med*, 2017, 9(402):eaam8060.
- [4] BARRETO J N, MCCULLOUGH K B, ICE L L, et al. Antineoplastic agents and the associated myelosuppressive effects: a review[J]. *J Pharm Pract*, 2014, 27(5):440-446.
- [5] HAVRILESKY L J, REINER M, MORROW P K, et al. A review of relative dose intensity and survival in patients with metastatic solid tumors[J]. *Crit Rev Oncol Hematol*, 2015, 93(3):203-210.
- [6] TALCOTT J A, SIEGEL R D, FINBERG R, et al. Risk assessment in cancer patients with fever and neutropenia: a prospective, two-center validation of a prediction rule[J]. *J Clin Oncol*, 1992, 10(2):316-322.
- [7] LI M, WANG Q, LU P, et al. Development of a machine learning-based prediction model for chemotherapy-induced myelosuppression in children with Wilms' tumor [J]. *Cancer (Basel)*, 2023, 15(4):1078.
- [8] EPSTEIN R S, WEERASINGHE R K, PARRISH A S,

- et al. Real-world burden of chemotherapy-induced myelosuppression in patients with small cell lung cancer: a retrospective analysis of electronic medical data from community cancer care providers[J]. *J Med Econ*, 2022, 25(1): 108-118.
- [9] MOONS K G M, DE GROOT J A H, BOUWMEESTER W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist[J]. *PLoS Med*, 2014, 11(10): e1001744.
- [10] WOLFF R F, MOONS K G M, RILEY R D, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies[J]. *Ann Intern Med*, 2019, 170(1): 51-58.
- [11] DRANITSARIS G, RAYSON D, VINCENT M, et al. Identifying patients at high risk for neutropenic complications during chemotherapy for metastatic breast cancer with doxorubicin or pegylated liposomal doxorubicin: the development of a prediction model[J]. *Am J Clin Oncol*, 2008, 31(4): 369-374.
- [12] JENKINS P, SCAIFE J, FREEMAN S. Validation of a predictive model that identifies patients at high risk of developing febrile neutropenia following chemotherapy for breast cancer[J]. *Ann Oncol*, 2012, 23(7): 1766-1771.
- [13] CHO B J, KIM K M, BILEGSAIKHAN S E, et al. Machine learning improves the prediction of febrile neutropenia in Korean inpatients undergoing chemotherapy for breast cancer[J]. *Sci Rep*, 2020, 10(1): 14803.
- [14] CHEN K, ZHANG X L, DENG H R, et al. Clinical predictive models for chemotherapy-induced febrile neutropenia in breast cancer patients: a validation study[J]. *PLoS One*, 2014, 9(6): e96413.
- [15] 李亚玲, 饶真真, 胡保玲, 等. 乳腺癌化疗患者骨髓抑制风险预测模型的构建与验证[J]. *军事护理*, 2024, 41(2): 6-10.  
LI Y L, RAO Z Z, HU B L, et al. Construction and validation of a risk prediction model for myelosuppression in breast cancer chemotherapy patients[J]. *Mil Nurs*, 2024, 41(2): 6-10.
- [16] 黄家良, 夏坤健, 郭伟, 等. 三阴性乳腺癌化疗致重度骨髓抑制的危险因素分析及其预测模型的构建[J]. *中山大学学报(医学科学版)*, 2023, 44(5): 886-892.  
HUANG J L, XIA K J, GUO W, et al. Risk factors and predictive model for severe myelosuppression due to chemotherapy in triple-negative breast cancer[J]. *J Sun Yat Sen Univ Med Sci*, 2023, 44(5): 886-892.
- [17] CHANG L L, SCHNEIDER S M, CHIANG S C, et al. Implementing an evidence-based risk assessment tool to predict chemotherapy-induced neutropenia in women with breast cancer[J]. *Cancer Nurs*, 2013, 36(3): 198-205.
- [18] ERNSTER V L. Nested case-control studies[J]. *Prev Med*, 1994, 23(5): 587-590.
- [19] CAO W, CHEN H D, YU Y W, et al. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020[J]. *Chin Med J (Engl)*, 2021, 134(7): 783-791.
- [20] NATTINO G, PENNELL M L, LEMESHOW S. Assessing the goodness of fit of logistic regression models in large samples: a modification of the Hosmer-Lemeshow test[J]. *Biometrics*, 2020, 76(2): 549-560.
- [21] PARK S H, HAN K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction[J]. *Radiology*, 2018, 286(3): 800-809.
- [22] TALARI K, GOYAL M. Retrospective studies: utility and caveats[J]. *J R Coll Physicians Edinb*, 2020, 50(4): 398-402.
- [23] OGUINDIMU E O, ALTMAN D G, COLLINS G S. Adequate sample size for developing prediction models is not simply related to events per variable[J]. *J Clin Epidemiol*, 2016, 76: 175-182.
- [24] COLLINS G S, OGUINDIMU E O, COOK J A, et al. Quantifying the impact of different approaches for handling continuous predictors on the performance of a prognostic model[J]. *Stat Med*, 2016, 35(23): 4124-4135.
- [25] PESKOE S B, ARTERBURN D, COLEMAN K J, et al. Adjusting for selection bias due to missing data in electronic health records-based research[J]. *Stat Methods Med Res*, 2021, 30(10): 2221-2238.
- [26] LEON RAPOPORT B, GARCIA-MORILLO M, FONT C, et al. A prospective, real-world, multinational study of febrile neutropenia (FN) occurrence in oncology patients receiving chemotherapy with intermediate risk of FN: a MASCC Neutropenia, Infection, and Myelosuppression Study Group initiative[J]. *Support Care Cancer*, 2023, 31(12): 628.
- [27] DAI Y, LIU M B, LEI L, et al. Prognostic significance of preoperative prognostic nutritional index in ovarian cancer: a systematic review and meta-analysis[J]. *Medicine (Baltimore)*, 2020, 99(38): e21840.
- [28] CHEN L, BAI P, KONG X Y, et al. Prognostic nutritional index (PNI) in patients with breast cancer treated with neoadjuvant chemotherapy as a useful prognostic indicator [J]. *Front Cell Dev Biol*, 2021, 9: 656741.
- [29] ZHOU Y L, LIU Y D, ZHANG Y, et al. Identifying non-linear association between maternal free thyroxine and risk of preterm delivery by a machine learning model[J]. *Front Endocrinol (Lausanne)*, 2022, 13: 817595.

(收稿日期:2024-08-29 修回日期:2025-01-07)

(编辑:刘明伟)