

## · 糖尿病心脏病专题研究 ·

## 2型糖尿病患者发生心力衰竭风险预测模型的系统评价

扫描二维码  
查看更多杨玉涵<sup>1</sup>, 刘岩<sup>1</sup>, 袁如月<sup>1</sup>, 胡超越<sup>1</sup>, 张晔<sup>1</sup>, 张力<sup>2</sup>, 杨晓晖<sup>1</sup>

**【摘要】** 目的 系统评价2型糖尿病(T2DM)患者发生心力衰竭的风险预测模型, 以期为临床医生选择合适的预测模型提供参考。方法 计算机检索PubMed、Embase、Cochrane Library、中国知网、万方数据知识服务平台、维普网及中国生物医学文献数据库中有关T2DM患者发生心力衰竭风险预测模型的文献, 检索时间从建库至2022-04-30。根据CHARMS清单提取文献资料, 应用预测模型偏倚风险和适用性评估工具(PROBAST)评估模型的偏倚风险和适用性。结果 最终纳入12篇文献, 共构建了14个风险预测模型。仅1个模型未提及AUC, 13个模型的AUC为0.72~0.87; 仅1个模型未提及校准方法, 13个模型报告了校准方法; 8个模型采用Bootstrap法进行内部验证, 5个模型采用分割样本法进行内部验证, 1个模型采用交叉验证法进行内部验证; 8个模型是作者或其他研究者进行外部验证, 6个模型未进行外部验证; 14个模型包含3~16个模型变量, 其中最常见模型变量是年龄(8个模型)、糖化血红蛋白(HbA<sub>1c</sub>)(8个模型)及BMI(7个模型); 模型最常见的呈现形式为评分分级(6个模型), 其次为方程(5个模型)。14个模型整体均存在高偏倚风险, 但整体适用性高。结论 目前构建的T2DM患者发生心力衰竭风险预测模型的AUC为0.72~0.87, 具有一定区分度, 但部分模型缺乏外部验证, 且所有模型存在高偏倚风险。

**【关键词】** 2型糖尿病; 心力衰竭; 预测模型; 系统评价**【中图分类号】** R 587.1 R 541.6 **【文献标识码】** A DOI: 10.12114/j.issn.1008-5971.2023.00.045

### Risk Prediction Models for Heart Failure in Patients with Type 2 Diabetes Mellitus: a Systematic Review YANG Yuhui<sup>1</sup>, LIU Yan<sup>1</sup>, YUAN Ruyue<sup>1</sup>, HU Chaoyue<sup>1</sup>, ZHANG Ye<sup>1</sup>, ZHANG Li<sup>2</sup>, YANG Xiaohui<sup>1</sup>

1. Department of Nephrology and Endocrinology, Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing 100700, China

2. Scientific Research Department, Dongfang Hospital, Beijing University of Chinese Medicine, Beijing 100078, China

Corresponding author: YANG Xiaohui, E-mail: yxh0616@126.com

**【Abstract】 Objective** To systematically evaluate the risk prediction models for heart failure in patients with type 2 diabetes mellitus (T2DM), in order to provide a reference for clinicians to choose the appropriate prediction model. **Methods** PubMed, Embase, Cochrane Library, CNKI, Wanfang Data, VIP and CBM were searched to collect the literature on the risk prediction model of heart failure in patients with T2DM. The retrieval time was from the establishment of the database to 2022-04-30. The literature data was extracted according to the CHARMS checklist, and the risk of bias and applicability of the model were evaluated by the Predictive Model Risk of Bias and Suitability Assessment Tool (PROBAST). **Results** Finally, 12 literature were included, and 14 risk prediction models were constructed. Only 1 model did not mention the AUC, and the AUC of the other 13 models was 0.72-0.87; only 1 model did not mention the calibration method, and 13 models reported the calibration method; 8 models were internally validated by Bootstrap method, 5 models were internally validated by split-sample method, and 1 model was internally validated by cross-validation method; 8 models were externally validated by authors or other researchers, and 6 models were not externally validated; 14 models had 3-16 model variables, the most common model variables were age (8 models), glycosylated hemoglobin (HbA<sub>1c</sub>) (8 models) and BMI (7 models); the most common presentation of the models was score grading (6 models), followed by equations (5 models). Fourteen models had high risk of bias and high applicability. **Conclusion** The AUC of the current risk prediction model for heart failure in T2DM patients was 0.72-0.87, with a certain degree of discrimination, but some models lack external validation and all models have a high risk of bias.

**【Key words】** Type 2 diabetes mellitus; Heart failure; Predictive model; Systematic review

基金项目: 国家自然科学基金资助项目(81974541); 北京市自然科学基金资助项目(7212176)

1.100700北京市, 北京中医药大学东直门医院肾病内分泌二区 2.100078北京市, 北京中医药大学东方医院科研处

通信作者: 杨晓晖, E-mail: yxh0616@126.com

心力衰竭是各种心脏病的终末阶段,其具有患病率高、病死率高、再住院率高等特点<sup>[1]</sup>。2017年,全球心力衰竭患者约有6 430万例<sup>[2]</sup>,心力衰竭发病后5年内死亡率约为50%<sup>[3]</sup>。在我国,心力衰竭患者人数约为890万,约占全球心力衰竭患者总数的13.8%<sup>[4]</sup>。2型糖尿病(type 2 diabetes mellitus, T2DM)是心力衰竭的独立危险因素,T2DM患者心力衰竭发生风险较非T2DM患者增加2~4倍<sup>[5]</sup>,心力衰竭住院风险较非T2DM患者增加2倍<sup>[6]</sup>。T2DM患者合并心力衰竭后临床症状、体征更明显,心功能、生活质量更差,死亡风险较未合并心力衰竭患者增加10~12倍<sup>[7]</sup>,故T2DM患者心力衰竭风险防控形势依然严峻。近年随着研究不断深入,T2DM患者发生心力衰竭的危险因素被不断发现,且国内外研究者根据危险因素构建了多个T2DM患者发生心力衰竭的风险预测模型,但鲜有研究对这些风险预测模型进行评价,这在一定程度上影响了风险预测模型在医疗决策中的应用。因此,本研究对T2DM患者发生心力衰竭风险预测模型进行系统评价,以期临床医生选择合适的预测模型提供参考。

## 1 资料与方法

**1.1 检索策略** 计算机检索PubMed、Embase、Cochrane Library、中国知网、万方数据知识服务平台、维普网及中国生物医学文献数据库中有关T2DM患者发生心力衰竭风险预测模型的文献,检索时间从建库至2022-04-30。采用主题词和关键词相结合的检索方式,其中英文检索词:“diabetes/diabetes mellitus/type 2 diabetes mellitu/T2DM”“heart failure/cardiac dysfunction/incident heart failure/congestive heart failure”“risk prediction/risk score/risk assessment model/risk prediction model”;中文检索词:“糖尿病/2型糖尿病”“心力衰竭/心功能不全”“风险预测/风险评分/风险评估模型”<sup>[8]</sup>。并手动检索纳入文献的参考文献。

**1.2 文献纳入与排除标准** 纳入标准:(1)研究对象:T2DM患者,种族、国籍、T2DM病程不限;(2)研究类型:开发、更新或验证风险预测模型的研究;(3)结局指标:心力衰竭,包括心力衰竭、再次发生心力衰竭、因心力衰竭住院、因心力衰竭死亡或残疾。排除标准:(1)以包括心力衰竭在内的复合终点为结局指标的文献;(2)会议摘要、灰色文献等非正式发表的文献;(3)综述、述评及基于系统评价/Meta分析建立风险预测模型的文献;(4)基础实验;(5)无法获取全文的文献。

**1.3 资料提取** 根据CHARMS清单制定资料提取表<sup>[9]</sup>,提取内容包括纳入文献的基本特征(包括第一作者、发表时间、研究类型、数据来源、模型类型、研究对象、随访时间、结局指标)、模型构建情况〔包括样本量、心力衰竭发生例数、候选变量情况(包括数量和连续变量处理方法)、缺失数据情况(包括数量和处理方法)、建模方法及变量选择方法〕、模型预测效能〔包括AUC、校准方法、内部验证和外部验证(包括作者或其他研究者进行外部验证)〕情况、模型变量及模型呈现形式。

**1.4 偏倚风险与适用性评价** 应用预测模型偏倚风险和适用性评估工具(Prediction Model Risk of Bias Assessment Tool,

PROBAST)评估模型构建和验证的偏倚风险<sup>[10-11]</sup>,其中偏倚风险评价内容包括研究对象、预测因素、结局、分析4个领域共20个信号问题,任意一个领域评为高风险或不清楚则为整体偏倚风险高;适用性评价内容包括研究对象、预测因素和结局3个领域,任意一个领域适用性低则为整体适用性低。

**1.5 质量控制与数据分析** 文献筛选、资料提取、偏倚风险与适用性评价均由两位经过培训的研究者平行完成,并交叉核对结果。若两人意见存在分歧且经过讨论不能达成共识,则寻求第3方意见并达成共识。数据以描述性分析为主,计量资料以 $M(QR)$ 表示,计数资料以绝对数表示。

## 2 结果

**2.1 检索结果** 初步检索相关文献3 420篇,根据文献纳入与排除标准,最终纳入12篇文献<sup>[12-23]</sup>,均为英文文献。文献筛选流程见图1。

**2.2 文献基本特征** 12篇文献<sup>[12-23]</sup>共构建了14个风险预测模型。研究类型为随机对照试验5篇<sup>[13, 17-19, 22]</sup>,队列研究4篇<sup>[12, 14, 16, 21]</sup>,医疗机构电子数据2篇<sup>[20, 23]</sup>,注册研究1篇<sup>[15]</sup>;数据来源为欧洲4篇<sup>[13-16]</sup>,美洲4篇<sup>[17-18, 20-21]</sup>,亚洲2篇<sup>[12, 23]</sup>,全球26个国家2篇<sup>[19, 22]</sup>;随访时间为2.1~17.6年。纳入文献的基本特征见表1。

**2.3 模型构建情况** 14个模型的样本量( $M=8 756$ ,  $QR=22 168$ )和心力衰竭发生例数( $M=402.5$ ,  $QR=976$ )间的差异较大,候选变量为6~109个。12个模型未提及缺失数据,6个模型未提及缺失数据处理方法。14个模型中建模方法为Cox比例风险回归模型10个,Weibull比例风险回归模型2个,机器学习1个,参数比例风险模型1个;14个模型中变量选择方法为逐步选择法9个,弹性网正则化1个,随机生存森林选择1个,未提及变量选择方法3个,见表2。

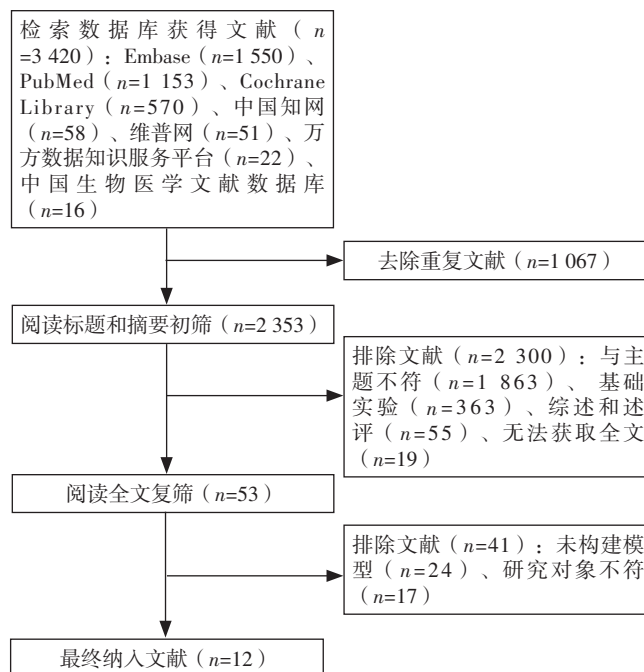


图1 文献筛选流程

Figure 1 Literature screening process

表1 纳入文献的基本特征  
Table 1 Basic characteristics of included literature

第一作者	发表时间(年)	研究类型	数据来源	模型类型	研究对象	随访时间(年)	结局指标
YANG <sup>[12]</sup>	2008	前瞻性队列研究	中国	开发模型	排除心力衰竭及冠心病史的T2DM患者	5.5	因心力衰竭住院
PFISTER <sup>[13]</sup>	2013	随机对照试验	欧洲19个国家	开发模型	T2DM合并大血管病变患者	2.9	因心力衰竭住院、死亡或残疾
HAYES <sup>[14]</sup>	2013	队列研究	英国	开发模型	新诊断的T2DM患者	17.6	心力衰竭
AHMAD KIADALIRI <sup>[15]</sup>	2013	注册研究	瑞典	开发模型	T2DM患者并排除T2DM后发生心脑血管事件(AMI、心力衰竭、IHD、卒中)者	4	首次发生心力衰竭
AHMAD KIADALIRI <sup>[15]</sup>	2013	注册研究	瑞典	开发模型	T2DM患者并排除T2DM后发生心脑血管事件(AMI、心力衰竭、NAIHD、卒中)者	4	再次发生心力衰竭
HIPPISLEY-COX <sup>[16]</sup>	2015	队列研究	英国	开发、外部验证模型	未发生心力衰竭的糖尿病患者	15	心力衰竭
BASU <sup>[17]</sup>	2017	随机对照试验	美国、加拿大	开发、外部验证模型	具有心血管疾病高风险的T2DM患者	4.7	心力衰竭
SEGAR <sup>[18]a</sup>	2019	随机对照试验	美国、加拿大	开发、外部验证模型	具有心血管疾病高风险的T2DM患者	4.9	因心力衰竭住院或死亡
BERG <sup>[19]</sup>	2019	随机对照试验	全球26个国家	开发、外部验证模型	具有心血管疾病高风险的T2DM患者	2.1	因心力衰竭住院
WILLIAMS <sup>[20]</sup>	2020	医疗机构电子数据	美国	开发模型	无心力衰竭病史的T2DM患者	6.6	因心力衰竭住院
PANDEY <sup>[21]</sup>	2021	前瞻性队列研究	美国	开发模型	T2DM及糖尿病前期患者	17	心力衰竭
BERG <sup>[22]</sup>	2021	随机对照试验	全球26个国家	开发、外部验证模型	具有心血管疾病高风险的T2DM患者	2.1	因心力衰竭住院
QUAN <sup>[23]</sup>	2021	医疗机构电子数据	中国	开发模型	T2DM及糖尿病前期患者	4.1	心力衰竭

注: T2DM=2型糖尿病, AMI=急性心肌梗死, IHD=缺血性心脏病, NAIHD=非急性缺血性心脏病; <sup>a</sup>表示该研究采用不同建模方法构建模型

表2 模型构建情况  
Table 2 Model construction

第一作者	样本量(例)	心力衰竭发生例数(例)	候选变量		缺失数据		建模方法	变量选择方法
			数量(个)	连续变量处理方法	数量(个)	处理方法		
YANG <sup>[12]</sup>	7 067	274	20	未处理	-	-	Cox比例风险回归模型	向后逐步选择法
PFISTER <sup>[13]</sup>	4 951	233	34	转换为分类变量	-	-	Cox比例风险回归模型	逐步选择法
HAYES <sup>[14]</sup>	5 102	351	19	未处理	-	-	Cox比例风险回归模型	向后逐步选择法
AHMAD KIADALIRI <sup>[15]</sup>	29 034	1 366	10	未处理	1 132-22 597	末次检测数据填补	Weibull比例风险回归模型	向后逐步选择法
AHMAD KIADALIRI <sup>[15]</sup>	29 034	947	6	未处理	1 132-22 597	末次检测数据填补	Weibull比例风险回归模型	向后逐步选择法
HIPPISLEY-COX <sup>[16]</sup>	437 806	25 480	16	未处理	-	多重插补法	Cox比例风险回归模型	-
BASU <sup>[17]</sup>	9 635	454	32	未处理	-	多重插补法	Cox比例风险回归模型	弹性网正则化
SEGAR <sup>[18]</sup>	8 756	319	109	未处理	-	随机森林插补法	Cox比例风险回归模型	向前向后逐步选择法
SEGAR <sup>[18]</sup>	8 756	319	109	未处理	-	随机森林插补法	机器学习	随机生存森林选择
BERG <sup>[19]</sup>	8 212	228	25	转换为分类变量	-	-	Cox比例风险回归模型	向前向后逐步选择法
WILLIAMS <sup>[20]</sup>	54 452	1 884	80	转换为分类变量	-	插补法	Cox比例风险回归模型	向前逐步选择法
PANDEY <sup>[21]</sup>	6 799	891	-	转换为分类变量	-	-	Cox比例风险回归模型	-
BERG <sup>[22]</sup>	6 106	168	7	转换为分类变量	-	-	Cox比例风险回归模型	-
QUAN <sup>[23]</sup>	97 628	2 159	32	未处理	-	完整案例分析	参数比例风险模型	向后逐步选择法

注: -表示未提及

2.4 模型预测效能、模型变量及模型呈现形式 仅1个模型未提及AUC, 13个模型的AUC为0.72~0.87; 仅1个模型未提及校准方法, 13个模型报告了校准方法; 8个模型采用Bootstrap法进行内部验证, 5个模型采用分割样本法进行内部验证, 1个模型采用交叉验证法进行内部验证; 8个模型是作者或其他研究者进行外部验证, 6个模型未进行外部验证; 14个模型包含3~16个模型变量, 其中最常见模型变量是年龄(8个模型)、糖化血红蛋白(glycosylated hemoglobin, HbA<sub>1c</sub>)(8个模型)及BMI(7个模型); 模型最常见的呈现形式为评分分级(6个模型), 其次为方程(5个模型), 见表3。

2.5 偏倚风险及适用性评价 分别有8、13、3、2个模型在研究对象、预测因素、结局和分析领域被评为低偏倚风险, 偏倚风险方面存在的主要问题如下: (1)未说明预测因素的采集时间以致无法确定是否是在不清楚结局数据的情况下评估预测因素(1个模型)及预测因素评估和结局确定的时间间隔是否合理(9个模型); (2)所有研究对象结局的定义可能不相同(2个模型)。14个模型整体均存在高偏倚风险, 但整体适用性高, 见表4~5。

### 3 讨论

2015年, WANG等<sup>[24]</sup>进行的Meta分析共纳入31篇文献,

**表3** 模型预测效能、模型变量及模型呈现形式  
**Table 3** Model prediction efficiency, model variable and model presentation form

第一作者	AUC	校准方法	内部验证(方法)	外部验证	模型变量	模型呈现形式
YANG <sup>[12]</sup>	0.85	校准曲线+Hosmer-Lemeshow拟合优度检验	有(分割样本法)	无	年龄、性别、BMI、HbA <sub>1c</sub> 、尿白蛋白/肌酐比、Hb、随访期间是否发生冠心病	方程
PFISTER <sup>[13]</sup>	0.75	校准曲线+Hosmer-Lemeshow拟合优度检验	有(Bootstrap法)	无	年龄、肌酐、利尿剂使用情况、HbA <sub>1c</sub> 、糖尿病病程、LDL、心率、左束支传导阻滞、右束支传导阻滞、心肌梗死病史、微量白蛋白、吡格列酮使用情况	评分分级
HAYES <sup>[14]</sup>	-	校准曲线	有(Bootstrap法)	无	诊断T2DM年龄、心房颤动、BMI、eGFR、LDL、微量或大量蛋白尿、周围动脉阻塞性疾病、截肢病史及足部感染病史	方程
AHMAD KIADALIRI <sup>[15]</sup>	0.84	校准曲线+Hosmer-Lemeshow拟合优度检验	有(分割样本法)	其他研究者进行外部验证	性别、确诊糖尿病年龄、HbA <sub>1c</sub> 、收缩压、BMI、TC/HDL、微量蛋白尿、大量蛋白尿、吸烟史、T2DM诊断前有心力衰竭病史	方程
AHMAD KIADALIRI <sup>[15]</sup>	0.84	校准曲线+Hosmer-Lemeshow拟合优度检验	有(分割样本法)	其他研究者进行外部验证	性别、确诊糖尿病年龄、HbA <sub>1c</sub> 、吸烟史、糖尿病病程、距离首次发生心力衰竭间隔时间是否>1年	方程
HIPPISLEY-COX <sup>[16]</sup>	男性: 0.76; 女性: 0.77	校准曲线	有(分割样本法)	作者进行外部验证	年龄、BMI、收缩压、TC/HDL、HbA <sub>1c</sub> 、Townsend剥夺指数、种族、吸烟史、糖尿病病程、心房颤动、心血管疾病、慢性肾脏病病史	网页计算器
BASU <sup>[17]</sup>	0.75	校准曲线+其他方法	有(交叉验证法)	作者进行外部验证+其他研究者进行外部验证	年龄、性别、种族、吸烟史、收缩压、心血管疾病病史、降压药使用情况、降脂药使用情况、抗凝药使用情况、HbA <sub>1c</sub> 、TC、HDL、肌酐及尿白蛋白/肌酐比	方程、网页计算器
SEGAR <sup>[18]</sup>	0.72	校准曲线+Hosmer-Lemeshow拟合优度检验	有(Bootstrap法)	作者进行外部验证	年龄、BMI、收缩压、舒张压、空腹血糖、肌酐、HDL、QRS间期时长、心肌梗死病史、冠状动脉旁路移植术史	评分分级
SEGAR <sup>[18]</sup>	0.77	校准曲线+Hosmer-Lemeshow拟合优度检验	有(Bootstrap法)	作者进行外部验证	年龄、BMI、收缩压、舒张压、空腹血糖、肌酐、HDL、QRS间期时长、心肌梗死病史、冠状动脉旁路移植术史	网页计算器
BERG <sup>[19]</sup>	0.81	校准曲线+其他方法	有(Bootstrap法)	作者进行外部验证	心力衰竭病史、心房颤动病史、冠状动脉疾病病史、eGFR、尿白蛋白/肌酐比	评分分级
WILLIAMS <sup>[20]</sup>	0.78	校准曲线	有(分割样本法)	无	年龄、冠状动脉疾病、BUN、心房颤动、HbA <sub>1c</sub> 、血清白蛋白、收缩压、慢性肾脏病病史、吸烟史	评分分级
PANDEY <sup>[21]</sup>	0.74	-	有(Bootstrap法)	无	hs-TnT、NT-proBNP、hs-CRP、心电图判定为左心室肥厚	评分分级
BERG <sup>[22]</sup>	0.87	校准曲线+其他方法	有(Bootstrap法)	作者进行外部验证	hs-TnT、NT-proBNP、心力衰竭病史	评分分级
QUAN <sup>[23]</sup>	0.80	校准曲线	有(Bootstrap法)	无	年龄、糖尿病病程、HbA <sub>1c</sub> 、收缩压、舒张压、TG、HDL、BMI、WBC、Hb、吸烟史、慢性肾脏病分期、胰岛素使用情况、缺血性心脏病病史、心肌梗死病史、心房颤动病史	网页计算器

注: -表示未提及; BMI=体质指数, HbA<sub>1c</sub>=糖化血红蛋白, Hb=血红蛋白, LDL=低密度脂蛋白, eGFR=估算肾小球滤过率, TC=总胆固醇, HDL=高密度脂蛋白, hs-TnT=超敏肌钙蛋白T, NT-proBNP=N末端脑钠肽前体, hs-CRP=超敏C反应蛋白, WBC=白细胞计数

但其文献纳排标准宽泛, 主要研究目的是分析T2DM患者发生心力衰竭的风险。2022年, RAZAGHIZAD等<sup>[25]</sup>考虑模型的实际应用情况, 要求模型有严格的纳排标准, 即提供模型的截距、预测因素系数等关键数据, 最终纳入15篇文章, 该研究有助于研究人员进一步校准、验证T2DM患者发生心力衰竭的风险预测模型, 但临床医生仍无法仅凭借相关风险预测模型指导T2DM患者的预防用药及制定护理措施。为了补充既往Meta分析的不足, 本研究评价了T2DM患者发生心力衰竭风险预测模型的偏倚风险及适用性, 同时延长了文献检索时间, 最终纳入12篇文章, 共构建了14个T2DM患者发生心力衰竭风险预测模型, 其AUC为0.72~0.87, 提示T2DM患者发生心力衰竭风险预测模型具有一定区分度。

外部验证是验证风险预测模型临床适用性的必要步骤, 但开发模型的研究者在进行外部验证时可能会倾向于选择与开发模型的队列具有相似特征的人群, 进而导致对模型的预测效能过于乐观<sup>[26]</sup>。本研究中6个模型由开发模型的研究者完成外部验证, 3个模型由其他研究者完成外部验证。COLLINS等<sup>[27]</sup>研究发现, 由开发模型的研究者进行外部验证的平均AUC为0.78, 高于其他研究者进行外部验证的0.72。另一方面, 考虑实际应用模型的人群可能在基线特征、基线

发病风险等方面存在较大差异<sup>[28]</sup>, 临床应用模型前需要多次进行外部验证以明确其是否具有可推广性。因此, 研究者应注重对现有模型的外部验证, 确定模型的可重复性及可推广性。

本研究14个模型整体均存在高偏倚风险, 偏倚风险方面存在的主要问题如下: (1) 未说明预测因素的采集时间以致无法确定是否是在不清楚结局数据的情况下评估预测因素(1个模型)及预测因素评估和结局确定的时间间隔是否合理(9个模型); (2) 所有研究对象结局的定义可能不相同(2个模型)。分析主要偏倚风险来源为缺失数据处理方法不当(8个模型)、样本量不足(6个模型)、连续变量处理方法不当(3个模型)。首先, 本研究中12个模型未提及缺失数据, 6个模型未提及缺失数据处理方法。在临床研究中, 缺失数据多为非随机缺失, 通常与纳入研究的人群或疾病特征相关, 故直接排除缺失数据不仅会影响模型的预测效能, 还会增加模型的偏倚风险<sup>[29-30]</sup>。其次, 样本量不足可能会错误估计模型的预测效能<sup>[31]</sup>。样本量大小通常以结局事件数与协变量个数比(the number of events per variable, EPV)进行衡量。一般认为, 模型的EPV<10提示存在过度拟合的可能性较大, 偏倚风险高。许多研究者推荐EPV至少应>20以降低模型的

表4 模型的研究对象、预测因素、结局、分析领域偏倚风险评价结果

Table 4 Bias risk assessment results of the research objects, predictors, results and analysis fields of the model

第一作者	条目 1	条目 2	条目 3	条目 4	条目 5	条目 6	条目 7	条目 8	条目 9	条目 10	条目 11	条目 12	条目 13	条目 14	条目 15	条目 16	条目 17	条目 18	条目 19	条目 20
YANG <sup>[12]</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	NI	NI	Y	PY	Y	N	Y
PFISTER <sup>[13]</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	Y	PN	Y	NI	Y
HAYES <sup>[14]</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NI	Y	Y	Y	NI	Y	PY	N	Y	Y
AHMAD KIADALIRI <sup>[15]</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NI	Y	Y	Y	N	Y	Y	Y	N	Y
AHMAD KIADALIRI <sup>[15]</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NI	Y	Y	Y	N	Y	Y	Y	N	Y
HIPPISLEY-COX <sup>[16]</sup>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	PN	Y	Y	Y
BASU <sup>[17]</sup>	Y	PN	Y	Y	Y	Y	PY	Y	Y	Y	NI	N	Y	Y	Y	Y	PY	Y	Y	Y
SEGAR <sup>[18]</sup>	Y	Y	PY	Y	Y	Y	Y	Y	Y	Y	NI	N	Y	Y	Y	Y	Y	Y	Y	Y
SEGAR <sup>[18]</sup>	Y	Y	PY	Y	Y	Y	Y	Y	Y	Y	NI	N	Y	Y	Y	Y	Y	Y	Y	Y
BERG <sup>[19]</sup>	Y	PN	Y	Y	Y	Y	Y	Y	Y	Y	NI	N	PY	NI	NI	Y	Y	Y	Y	Y
WILLIAMS <sup>[20]</sup>	PN	Y	PY	Y	Y	Y	PY	Y	PN	Y	Y	Y	N	Y	Y	Y	PY	Y	N	Y
PANDEY <sup>[21]</sup>	PY	PN	PY	NI	Y	PN	PN	PN	PN	PN	NI	Y	PN	Y	NI	Y	NI	PN	Y	PY
BERG <sup>[22]</sup>	Y	PN	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	PY	NI	NI	Y	Y	Y	Y	Y
QUAN <sup>[23]</sup>	PN	Y	Y	Y	Y	Y	Y	Y	Y	Y	NI	Y	Y	Y	Y	Y	PY	Y	Y	Y

注: Y表示“是”, PY表示“可能是”, N表示“否”, PN表示“可能否”, NI表示“未提供相关信息”; 条目1: 采取的数据来源是否合适? 条目2: 纳排标准是否合适? 条目3: 所有研究对象预测因素的定义和评估是否相同? 条目4: 是否是在不清楚结局数据的情况下评估预测因素? 条目5: 预测模型包含的预测因素是否有效? 条目6: 结局的分类方法是否合理? 条目7: 结局的定义是否合理? 条目8: 结局的定义是否排除了预测因素? 条目9: 所有研究对象的结局定义是否相同? 条目10: 确定结局时是否不清楚预测因素的信息? 条目11: 预测因素评估和结局确定的时间间隔是否合理? 条目12: 样本量是否合理? 条目13: 对连续自变量和分类自变量的处理方法是否合适? 条目14: 所有研究对象是否包含在统计分析中? 条目15: 缺失数据是否进行了合适处理? 条目16: 是否避免采用单因素分析法筛选预测因素? 条目17: 对数据的复杂性是否进行了考虑? 条目18: 是否对预测模型性能进行评估? 条目19: 是否对预测模型过度拟合、欠拟合和最优拟合情况进行考虑? 条目20: 预测因素及其权重是否与报告结果一致?

表5 模型的偏倚风险及适用性评价结果

Table 5 Evaluation results of bias risk and applicability of the model

第一作者	偏倚风险				适用性			总体	
	研究对象	预测因素	结局	分析	研究对象	预测因素	结局	偏倚风险	适用性
YANG <sup>[12]</sup>	+	+	+	-	+	+	+	-	+
PFISTER <sup>[13]</sup>	+	+	+	-	+	+	+	-	+
HAYES <sup>[14]</sup>	+	+	?	-	+	+	+	-	+
AHMAD KIADALIRI <sup>[15]</sup>	+	+	?	-	+	+	+	-	+
AHMAD KIADALIRI <sup>[15]</sup>	+	+	?	-	+	+	+	-	+
HIPPISLEY-COX <sup>[16]</sup>	+	+	+	-	+	+	+	-	+
BASU <sup>[17]</sup>	-	+	?	+	+	+	+	-	+
SEGAR <sup>[18]</sup>	+	+	?	-	+	+	+	-	+
SEGAR <sup>[18]</sup>	+	+	?	-	+	+	+	-	+
BERG <sup>[19]</sup>	-	+	?	-	+	+	+	-	+
WILLIAMS <sup>[20]</sup>	-	+	-	-	+	+	+	-	+
PANDEY <sup>[21]</sup>	-	?	-	-	+	+	+	-	+
BERG <sup>[22]</sup>	-	+	-	-	+	+	+	-	+
QUAN <sup>[23]</sup>	-	+	?	+	+	+	+	-	+

注: +表示低偏倚风险/适用性高, -表示高偏倚风险/适用性低, ?表示不清楚

潜在偏倚<sup>[32-34]</sup>，而对于使用机器学习建立的预测模型则需要更高的EPV (>200)以减少过度拟合情况<sup>[35]</sup>。本研究纳入研究中近一半研究EPV<20，尤其是SEGAR等<sup>[18]</sup>采用机器学习建立风险预测模型，但其EPV仅为2.9，导致了高偏倚风险。最后，针对连续变量应首先分析变量分布情况及其与结局指标的关系，最佳处理方法是线性或非线性拟合后代入方程<sup>[30]</sup>。有时考虑实际需求，研究者会将连续变量根据切点转换为分类变量，但若直接转换为分类变量，由于数据信息丢失严重，会导致模型性能大幅下降且偏倚风险较高；当采用3个或以上切点分割连续变量时，尤其是采用共识或指南中提供的切点时，模型的偏倚风险相对较低<sup>[36-37]</sup>。本研究中PANDEY等<sup>[21]</sup>和BERG等<sup>[22]</sup>将相同的生物学指标作为预测因素，前者选择单一切点分割连续变量，偏倚风险较高；后者采用4个切点分割连续变量，偏倚风险相对较低。

综上所述，目前构建的T2DM患者发生心力衰竭风险预测模型的AUC为0.72~0.87，具有一定区分度，但部分模型缺乏外部验证，且所有模型存在高偏倚风险。未来应积极开展模型的外部验证及更新，明确模型的可重复性及可推广性；建立或验证模型时注重报告的完整性及方法学规范，可参考个体预后或诊断的多变量预测模型透明报告(Transparent Reporting of A Multivariable Prediction Model for Individual Prognosis or Diagnosis, TRIPOD)<sup>[38]</sup>及PROBAST<sup>[10]</sup>以降低模型偏倚风险。但本研究仍存在一定局限性：本研究仅纳入中文、英文文献，未检索灰色文献，且纳入的模型数量较少；因纳入文献的数据来源不同且存在统计学异质性，故未对研究结果进行定量分析。

作者贡献：杨玉涵、刘岩、杨晓晖进行文章的构思及设计、可行性分析及统计学处理；袁如月、胡超越、张晔负责研究的实施、资料收集；杨玉涵进行资料整理，论文撰写；刘岩、张力、杨晓晖负责文章的质量控制及审校；杨晓晖对文章整体负责、监督管理。

本文无利益冲突。

#### 参考文献

- [1] HEIDENREICH P A, BOZKURT B, AGUILAR D, et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines [J]. *Circulation*, 2022, 145 (18): e895-1032. DOI: 10.1161/CIR.0000000000001063.
- [2] GROENEWEGER A, RUTTEN F H, MOSTERD A, et al. Epidemiology of heart failure [J]. *Eur J Heart Fail*, 2020, 22 (8): 1342-1356. DOI: 10.1002/ejhf.1858.
- [3] ROGER V L. Epidemiology of heart failure: a contemporary perspective [J]. *Circ Res*, 2021, 128 (10): 1421-1434. DOI: 10.1161/CIRCRESAHA.121.318172.
- [4] 国家心血管病医疗质量控制中心专家委员会心力衰竭专家工作组. 2020中国心力衰竭医疗质量控制报告 [J]. *中国循环杂志*, 2021, 36 (3): 221-238.
- [5] NICHOLS G A, HILLIER T A, ERBEY J R, et al. Congestive heart failure in type 2 diabetes: prevalence, incidence, and risk factors [J]. *Diabetes Care*, 2001, 24 (9): 1614-1619. DOI: 10.2337/diacare.24.9.1614.
- [6] 中华医学会糖尿病学分会. 中国2型糖尿病防治指南(2020年版) [J]. *中华糖尿病杂志*, 2021, 13 (4): 315-409. DOI: 10.3760/ema.j.cn115791-20210221-00095.
- [7] SEFEROVIC P M, PETRIE M C, FILIPPATOS G S, et al. Type 2 diabetes mellitus and heart failure: a position statement from the Heart Failure Association of the European Society of Cardiology [J]. *Eur J Heart Fail*, 2018, 20 (5): 853-872. DOI: 10.1002/ejhf.1170.
- [8] CHOWDHURY M Z I, YEASMIN F, RABI D M, et al. Prognostic tools for cardiovascular disease in patients with type 2 diabetes: a systematic review and meta-analysis of C-statistics [J]. *J Diabetes Complications*, 2019, 33 (1): 98-111. DOI: 10.1016/j.jdiacomp.2018.10.010.
- [9] MOONS K G, DE GROOT J A, 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. DOI: 10.1371/journal.pmed.1001744.
- [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. DOI: 10.7326/M18-1376.
- [11] MOONS K G M, WOLFF R F, RILEY R D, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration [J]. *Ann Intern Med*, 2019, 170 (1): W1-33. DOI: 10.7326/M18-1377.
- [12] YANG X L, MA R C, SO W Y, et al. Development and validation of a risk score for hospitalization for heart failure in patients with type 2 diabetes mellitus [J]. *Cardiovasc Diabetol*, 2008, 7: 9. DOI: 10.1186/1475-2840-7-9.
- [13] PFISTER R, CAIRNS R, ERDMANN E, et al. A clinical risk score for heart failure in patients with type 2 diabetes and macrovascular disease: an analysis of the PROactive study [J]. *Int J Cardiol*, 2013, 162 (2): 112-116. DOI: 10.1016/j.ijcard.2011.05.056.
- [14] HAYES A J, LEAL J, GRAY A M, et al. UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82 [J]. *Diabetologia*, 2013, 56 (9): 1925-1933. DOI: 10.1007/s00125-013-2940-y.
- [15] AHMAD KIADALIRI A, GERDTHAM U G, NILSSON P, et al. Towards renewed health economic simulation of type 2 diabetes: risk equations for first and second cardiovascular events from Swedish register data [J]. *PLoS One*, 2013, 8 (5): e62650. DOI: 10.1371/journal.pone.0062650.
- [16] HIPPLISLEY-COX J, COUPLAND C. Development and validation of risk prediction equations to estimate future risk of heart failure in patients with diabetes: a prospective cohort study [J]. *BMJ Open*, 2015, 5 (9): e008503. DOI: 10.1136/bmjopen-2015-008503.
- [17] BASU S, SUSSMAN J B, BERKOWITZ S A, et al. Development

- and validation of risk equations for complications of type 2 diabetes (RECODE) using individual participant data from randomised trials [J]. *Lancet Diabetes Endocrinol*, 2017, 5 (10): 788–798. DOI: 10.1016/S2213-8587(17)30221-8.
- [18] SEGAR M W, VADUGANATHAN M, PATEL K V, et al. Machine learning to predict the risk of incident heart failure hospitalization among patients with diabetes: the WATCH-DM risk score [J]. *Diabetes Care*, 2019, 43 (2): 2298–2306. DOI: 10.2337/dc19-0587.
- [19] BERG D D, WIVIOTT S D, SCIRICA B M, et al. Heart failure risk stratification and efficacy of sodium–glucose cotransporter-2 inhibitors in patients with type 2 diabetes mellitus [J]. *Circulation*, 2019, 140 (19): 1569–1577. DOI: 10.1161/CIRCULATIONAHA.119.042685.
- [20] WILLIAMS B A, GEBA D, CORDOVA J M, et al. A risk prediction model for heart failure hospitalization in type 2 diabetes mellitus [J]. *Clin Cardiol*, 2020, 43 (3): 275–283. DOI: 10.1002/clc.23298.
- [21] PANDEY A, VADUGANATHAN M, PATEL K V, et al. Biomarker-based risk prediction of incident heart failure in pre-diabetes and diabetes [J]. *JACC Heart Fail*, 2021, 9 (3): 215–223. DOI: 10.1016/j.jchf.2020.10.013.
- [22] BERG D D, WIVIOTT S D, SCIRICA B M, et al. A biomarker-based score for risk of hospitalization for heart failure in patients with diabetes [J]. *Diabetes Care*, 2021, 44 (11): 2573–2581. DOI: 10.2337/dc21-1170.
- [23] QUAN J C, NG C S, KWOK H H Y, et al. Development and validation of the CHIME simulation model to assess lifetime health outcomes of prediabetes and type 2 diabetes in Chinese populations: a modeling study [J]. *PLoS Med*, 2021, 18 (6): e1003692. DOI: 10.1371/journal.pmed.1003692.
- [24] WANG Y, NEGISHI T, NEGISHI K, et al. Prediction of heart failure in patients with type 2 diabetes mellitus—a systematic review and meta-analysis [J]. *Diabetes Res Clin Pract*, 2015, 108 (1): 55–66. DOI: 10.1016/j.diabres.2015.01.011.
- [25] RAZAGHIZAD A, OULOUSIAN E, RANDHAWA V K, et al. Clinical prediction models for heart failure hospitalization in type 2 diabetes: a systematic review and meta-analysis [J]. *J Am Heart Assoc*, 2022, 11 (10): e024833. DOI: 10.1161/JAHA.121.024833.
- [26] RAMSPEK C L, JAGER K J, DEKKER F W, et al. External validation of prognostic models: what, why, how, when and where? [J]. *Clin Kidney J*, 2021, 14 (1): 49–58. DOI: 10.1093/ckj/sfaa188.
- [27] COLLINS G S, DE GROOT J A, DUTTON S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting [J]. *BMC Med Res Methodol*, 2014, 14: 40. DOI: 10.1186/1471-2288-14-40.
- [28] RILEY R D, ENSOR J, SNELL K I, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges [J]. *BMJ*, 2016, 353: i3140. DOI: 10.1136/bmj.i3140.
- [29] BOUWMEESTER W, ZUTHOFF N P A, MALLETT S, et al. Reporting and methods in clinical prediction research: a systematic review [J]. *PLoS Med*, 2012, 9 (5): 1–12. DOI: 10.1371/journal.pmed.1001221.
- [30] 刘岩, 李新龙, 商洪才, 等. 针刺治疗功能性便秘研究中校正填补值的模式混合模型方法 [J]. *世界中医药*, 2017, 12 (6): 1258–1260, 1267. DOI: 10.3969/j.issn.1673-7202.2017.06.011.
- [31] VAN DER PLOEG T, AUSTIN P C, STEYERBERG E W. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints [J]. *BMC Med Res Methodol*, 2014, 14: 137. DOI: 10.1186/1471-2288-14-137.
- [32] VAN SMEDEEN M, DE GROOT J A, MOONS K G, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis [J]. *BMC Med Res Methodol*, 2016, 16 (1): 163. DOI: 10.1186/s12874-016-0267-3.
- [33] OGUNDIMU 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. DOI: 10.1016/j.jclinepi.2016.02.031.
- [34] COURVOISIER D S, COMBESCURE C, AGORITSAS T, et al. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure [J]. *J Clin Epidemiol*, 2011, 64 (9): 993–1000. DOI: 10.1016/j.jclinepi.2010.11.012.
- [35] VAN DER PLOEG T, AUSTIN P C, STEYERBERG E W. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints [J]. *BMC Med Res Methodol*, 2014, 14: 137. DOI: 10.1186/1471-2288-14-137.
- [36] 谷鸿秋, 王俊峰, 章仲恒, 等. 临床预测模型: 模型的建立 [J]. *中国循证心血管医学杂志*, 2019, 11 (1): 14–16, 23. DOI: 10.3969/j.issn.1674-4055.2019.01.04.
- [37] COLLINS G S, OGUNDIMU 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. DOI: 10.1002/sim.6986.
- [38] COLLINS G S, REITSMA J B, ALTMAN D G, et al. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement [J]. *J Clin Epidemiol*, 2015, 68 (2): 134–143. DOI: 10.1016/j.jclinepi.2014.11.010.

(收稿日期: 2022-10-21; 修回日期: 2023-02-02)

(本文编辑: 谢武英)