

· 综 述 ·

心力衰竭风险预警模型研究进展

于 丽^{1△} 卫靖靖¹ 朱明军² 郑超楠¹ 贾梦楠¹ 常鑫迪¹

心力衰竭（简称心衰）是一种全球性疾病，全世界约 2 600 万人被影响^[1]。研究发现，1 年内心衰患者的病死率男性为 43%，女性为 36%，5 年内心衰患者病死率男性为 75%，女性为 62%^[2]。因此，心衰已成为世界范围内的主要健康威胁，其预测研究也成为医学研究的热点。心衰患者需要进行持续监测和调节，国内外多年实践证明，早预警、早诊断、早干预、早治疗可以有效地预防心血管疾病的发生，早期识别高危人群和预防心衰逐渐成为专业指南中强调的策略^[3-5]。临床预测模型（clinical prediction models, CPMs）是一种针对患者发病风险的预估工具，通过相关的临床变量，来提供重要结果概率，这些工具旨在使临床医生能够为单个患者“个性化”风险提供医疗决策^[6]。

1 早期的心衰预警模型建立 风险预测模型被认为是预防性模型，用于识别人群中更易患病的个体。早期的风险预警是使用传统的逻辑回归（logistic regression, LR）建立的，LR 是专门用于开发医学领域的风险预测模型^[7]。二十世纪五六十年代，美国 Framingham 心脏研究中心通过长期随访发现，某些危险因素如年龄、性别、总胆固醇水平、高密度脂蛋白水平、收缩压、吸烟、糖尿病等可以用来预测冠心病的发病事件，并建立了最早的 Logistic 回归模型和冠心病 10 年风险预警模型^[8]。随着 Framingham 流行病学不断研究的深入，心衰的独立预测因子逐渐被发现，如年龄、收缩压、左室肥厚（left

ventricular hypertrophy, LVH）、肺活量、心率、冠心病、心脏瓣膜杂音、糖尿病、扩张型心肌病等^[9, 10]。在此基础上，Framingham 心脏中心以左心室肥厚，心率等心电图特征为基准建立了充血性心衰 4 年危险预测模型^[11]。模型的建立是基于该中心 38 年随访期间的 486 例心衰病例，利用心衰的危险因素，构建了一个多变量的风险函数，有效地从易患人群中识别出心衰的主要候选因子，再使用混合逻辑回归模型 $f(x) = 1 / (1 + e^{-x})$ 生成心衰风险计算公式，计算 4 年发生心衰的概率 x 是由一个截距为 (a) 的方程估计的，是每个特定危险因素的 β 系数乘以它们的实际值的总和^[12]。Schnabel RB 等^[12]通过随访 Framingham 心脏研究中心的 725 例房颤患者，在 10 年的随访期内，有 161 例参与者发生心衰，其中女性占 45%，10 年心衰新事件的发生率为 22%。在多变量模型中，年龄、体重指数、心电图左心肥大、糖尿病、心脏杂音和心肌梗死病史与心衰事件呈正相关（C 统计量为 0.71，95% CI: 0.67-0.75）。该研究者开发了一种风险算法，以良好的模型拟合和校准来预测房颤患者发生心衰的绝对风险（校正调整 $\chi^2=7.29$, $P=0.61$ ）。Miao F 等^[13]随后通过建立两个多变量 Cox 模型来评估心衰的一年发病风险，一个模型仅具有常规临床风险因素，另一个模型结合实验室参数，包括血清肌酐（serum creatinine, SCr）、血尿素氮（blood urea nitrogen, BUN）、葡萄糖、凝血酶原时间（prothrombin time, PT）、活化部分凝血活酶时间（arterial pulse propagation time, APTT）和总胆红素（total bilirubin, TBIL），研究发现，除了已知的危险因素外，还有一些临床和实验室指标如肺循环疾病、周围血管疾病、慢性肺部疾病、甲状腺功能减退、电解质和液体紊乱、BUN 和 APTT 等被认为是心衰发生发展的独立预测因素，经内部验证的 C 统计量为 0.712，证明了结合临床和实验室因素的预测模型的有效性。袁斌等^[14]通过分析 433 例患者的数据，采用 Logistics 多因素回归分析的方法筛选出预测急性心力衰竭（acute heart failure, AHF）发作的 5 个指标如氧饱和度、尿量、心率、情

基金项目：国家重点研发计划重点专项课题（No. 2019YFC1710003, No. 2019YFC1710000）；国家中医药管理局中医药循证能力建设项目（No. 2019XZZX-XXG003）；河南省中医药科学研究专项课题（No. 2018JDZX009）；河南省创新型科技团队（No. C20130050）；河南省中原千人计划-中原名医项目 [No. 豫财行（2020）140]

作者单位：1. 河南中医药大学第一临床医学院（郑州 450000）；2. 河南中医药大学第一附属医院心脏中心（郑州 450000）

通讯作者：朱明军，Tel: 0371-66233478；E-mail: zhumingjun317@163.com

△现单位为河南省中医院（河南中医药大学第二附属医院）急诊科

DOI: 10.7661/j.cjim.20230227.037

绪状态、呼吸频率,初步建立了 AHF 发作的 Super-Score 评分模型,能够提前 2~6 h 预警 AHF 发作。早期的心衰风险预警模型大多是以 Framingham 心脏病研究提供的常规风险特征为基础,来预测心衰发病风险。然而,随着心血管疾病患病率和死亡率的持续增加,需要采取更好的数据挖掘方法来开发反映风险因素和结果真实关系的风险模型。

2 机器学习(machine learning, ML)下心衰预警模型的完善 在过去的十几年里,机器学习和深度学习技术在糖尿病、肝病、传染病和心脏病等多种疾病的管理中得到了快速发展^[15]。ML 是一项位于统计学和计算机科学交汇处的科学研究,它利用人工智能从数据中学习相互关系或模式,为风险预测模型的建立提供了替代方法^[16, 17]。常见的 ML 有支持向量机(support vector machine, SVM)、决策树(decision tree, DT)、k-最近邻(k-nearest neighbor, KNN)、随机森林(random forest, RF)、基于模糊逻辑的算法(Fuzzy analytic hierarchy process, FAHP)、人工神经网络(artificial neural network, ANN)和神经网络集成的各种决策支持系统等。多年来,研究人员不断开发完善基于机器学习的预测模型,并应用于心衰的早期预测。

Yang G 等^[18]提出了 SVM 的评分模型。该模型是一种基于临床特征的心衰检测系统,包括血液检测结果、心率变异性、超声心动图、心电图、胸片、6 分钟步行距离测试和体格测试。在他们的研究中,参与的个体根据是否有心功能不全被分为 3 组,即健康组(无心功能障碍)、心衰倾向组(心功能不全的无症状阶段)和心衰组(心功能不全的有症状阶段)。使用 SVM 将其分为 3 类,总体准确率达到 74.44%,对健康组、心衰倾向组和心衰组的识别准确率分别为 78.79%、87.5% 和 65.85%。与临床实践中报道的结果比较,该模型有助于提高心衰诊断的准确性。

为了预测心衰风险,ANN 被广泛提出。Gharehchopogh FS 等^[19]建立的神经网络模型是以 40 例患者的数据集为基础,使用年龄、性别、血压和吸烟习惯作为变量,并获得了 0.95 的受试者工作特征曲线下面积(area under curve, AUC),他们基于神经网络的模型实现了 95% 的精度。

Ali L 等^[20]在 SVM 的基础上研发了一个新的心衰预测模型专家系统,用于心衰的早期预测。该系统堆叠了两个 SVM 模型以有效预测心衰。第一个 SVM 模型是线性的,并且 L_1 正则化,它能够通过将系数缩小到零来消除不相关的特征。第二个 SVM 模型是

L_2 正则化,它用作预测模型。为了优化这两个模型,研究者提出了一种能够同时优化这两个模型的混合网格搜索算法(hybrid grid search algorithm, HGSA),该算法使用 6 种不同的评估指标:准确性、敏感性、特异性、马修斯相关系数(matthews correlation coefficient, MCC)、接受者操作特性曲线(receiver operating characteristic curve, ROC)和 AUC 来评估所提出方法的有效性。实验结果证明,该方法将传统 SVM 模型的性能提高了 3.3%,其准确度在 57.85%~91.83% 的范围内。此外,该方法的性能也优于其他先进的机器学习集成模型。

随后,在 χ^2 统计模型和高斯朴素贝叶斯(gauss naive bayes, GNB)的基础上,研究者们又提出了一个特征驱动的心衰疾病预测决策支持系统^[21]。该系统包括两个阶段,其中第一阶段是采用 χ^2 统计模型对 13 个常用的高频特征因素进行排序,再通过前向最佳优先搜索策略来找到这些特征的最优子集。在第二阶段,使用 GNB 分类器作为预测模型。该研究通过使用 297 例受试者的在线心脏病数据库来评估新提出的方法 χ^2 -GNB 的性能。实验结果表明,改进后的方法即 χ^2 -GNB 将 GNB 模型的预测性能提高了 3.33%,其模型预测准确率达到 93.33%。

Guo CY 等^[22]通过对杰克逊心脏研究数据库(Jackson Heart Study, JHS)中 3 327 例患者进行研究,采用 LASSO 回归、Logistic 回归、支持向量机、随机森林和极端梯度提升(extreme gradient boosting, XGBoost)来建立心衰风险预测模型。在 JHS 数据库建立心衰预测模型时,使用缺失率 < 30% 的变量,并采用随机森林插补。通过对比得出, XGBoost 在大多数情况下预测心衰的性能最好,最高预测准确率为 0.8409,建议用 XGBoost 构建心衰风险预测模型。

3 结合中医证候和“治未病”理论构建心衰预警模型 有研究报道, Framingham 研究预测模型会明显高估我国人群心血管疾病的发病风险,若要应用于我国人群尚需进行适当调整^[23]。因此,寻找一种适合我国国情的心衰风险预警模型迫在眉睫。中医药在心衰的防治中应用较广泛,中医学中虽未对“心力衰竭”这一名词做具体阐述,但心衰可以归属“喘证”“痰饮”“惊悸”“心水”“水肿”等范畴。现在大多数医家认为,心衰的病位在心,与肺、脾、肾三脏有关,为本虚标实之证。本虚为气虚阳微,或兼夹心阴不足,标实是血瘀、痰浊、水饮,其中心气亏虚是发病的内在基础,血瘀是中心病理环节,痰浊、水饮是

重要的病理产物^[24]。国家中医药管理局在“十一五”期间起草的《慢性心力衰竭中医诊疗专家共识》中指出,心衰的中医证型总体概括为以下 3 类:气虚血瘀证、气阴两虚血瘀证、阳气亏虚血瘀证,且 3 种证型中均可兼见痰饮证^[25]。

研究发现,心衰各证候与心功能受损程度之间有密切的关系^[26]。张松等^[27]研究发现,慢性心衰患者病性类证候要素分布前 6 位依次为气虚、血瘀、痰浊、阴虚、阳虚、水饮。刘镇等^[28]研究发现,中医证候积分能够充分反映慢性心衰患者的疾病严重程度,与红细胞分布宽度、氨基末端脑利钠肽前体(N-terminal pro-brain natriuretic peptide, NT-proBNP)、Bosdon 心衰积分呈正相关关系。刘德秀等^[29]探索心衰患者 BNP 与中医证候之间的关系发现,心肾阳虚证、阳虚水泛证、痰饮阻肺证的 BNP 明显升高。综上所述,中医证候与心衰的发病密切相关,是心衰发病的一项独立危险因素。

中医学“治未病”理论思想在中医学中占据重要的地位,其起源可追溯到春秋战国时期,主要包括未病先防、已病防变、瘥后防复三个方面,贯穿在疾病发生发展的各个环节,其重点是未病先防。《素问·四气调神大论》提出:“是故圣人不治已病治未病,不治已乱治未乱,此之谓也。夫病已成而后药之,乱已成而后治之,譬犹渴而穿井,斗而铸锥,不亦晚乎。”由此可见,未病先防在养生中占据重要地位^[30]。现代预防医学中,疾病预防被分为 3 个阶段:一级预防,亦称病因预防,是在疾病尚未发生时针对致病因素(或危险因素)采取的措施,相当于中医学“治未病”思想中的未病先防;二级预防,亦称临床前预防,即早发现、早诊断、早治疗,是在发病初期为防止或延缓疾病进展而采取的措施,相当于中医学“治未病”思想中的既病防变;三级预防,亦称临床期预防,是在疾病发生后为防止伤残、提高生活质量、降低病死率而采取的对症治疗和康复治疗措施,相当于中医学“治未病”思想中的瘥后防复^[31, 32]。由此看来,“治未病”思想与现代医学的三级预防理论是相通的,将现代预防医学与传统“治未病”思想相融合,将疾病的发病因素与中医证候相结合,构建具有中医特色的风险评估预警模型,是当今心脏病领域的重要方向。

以此为基础,我国学者应用多元数据统计和神经网络挖掘方法构建慢性心衰患者血瘀证基于理化指标信息的诊断模型,运用 Logistic 回归分析,筛选出 15 项理化指标如中性粒细胞、中性粒细胞比率、单核细胞、血小板、血细胞比容,电解质指标

中钠、氯、镁,肝肾功指标中肌酸、补体 C3、碱性磷酸酶、腺苷酸氨酶、总蛋白,甲状腺激素水平指标中 TT3,心动超声指标中左心室后壁厚度等与心衰血瘀证相关,其中理化指标信息为自变量,血瘀证为因变量。再通过整合上述指标,进行神经网络多层感知器(multi-layer perceptron, MLP)建模分析,反复验证达到较为稳定的状态。其中训练样本的准确率为 75.4%,测试样本的准确率为 82.4%,与血瘀证有相关性的理化指标依次为总蛋白、中性粒细胞比率、单核细胞等^[33]。朱泽豪^[34]以心衰急性发作反复入院风险预测为结局事件,通过单因素分析鉴别危险因素,再利用 R 软件绘制列线图,建立了回归方程: $\ln[p/(1-p)] = -4.509 + 0.617 \times \text{年龄} + 0.919 \times \text{肺部感染} + 1.447 \times \text{用药依从性} + 0.976 \times \text{心功能} + 0.712 \times \text{LVEF} + 1.030 \times \text{气虚} + 0.970 \times \text{阳虚} + 0.952 \times \text{痰浊} + 0.770 \times \text{血瘀}$ 。研究者们通过 C 指数、校正图、bootstrap 内部验证的方式检验模型的可行性,最终,内部验证结果显示 C 指数为 0.823,模型稳定性较好。

4 发展与展望 在既往的心衰风险预测模型中,常用的是传统心血管危险因素,如年龄、性别、收缩压、舒张血压、体重指数、低密度脂蛋白、高密度脂蛋白、吸烟、糖尿病病史和心肌梗死等。然而,独立于传统风险因素之外,有些危险因素与心衰的发生发展同样密切相关。有研究发现,腰围增大、尿白蛋白-肌酐比值升高、血清炎症因子、生物标志物如肾上腺髓质中段肽、C 反应蛋白、A 型利钠肽原中间部分等也被确定为新的危险预测因子^[35, 36]。

近年来,不断有国内外学者致力于心衰风险预警模型建模的研究,机器学习下心衰预警模型层出不穷。然而目前现有的预警模型研究仍有待改善:(1)国外的传统风险预警模型开发多年,随着时间的变化,模型存在校准度漂移,模型预测性能可能不准确。(2)我国的心衰预测模型相对匮乏,具有中医学特色的心衰预警模型更是寥寥无几,应用机器学习的建模方法仍有待探究。(3)现有的风险预测模型大多是使用传统的逻辑回归建立的,逻辑回归遵循概率模型的某些统计假设,这可能会限制其在复杂问题中的使用。(4)构建模型时主要是以传统心血管危险因素进行单因素及多因素分析,但随着新的危险预测因子的出现,既往模型的预测因素可能不适用于现在的临床需求。(5)构建模型时,预测因素之间可能相互影响,缺少对其交互作用的分析。(6)目前的模型大多是从单中心研究中收集的数据集开发的。虽然数据涉

及到不同社会和经济阶层的人群,但单中心研究的相关问题仍不容忽视。(7)大多数现有模型缺乏外部验证,模型的实际性能很难评估。

本文总结了国内外心衰预警的相关模型,发现深度学习不仅对于复杂问题的处理能力更强,而且预测的准确率更高,因此致力于找到心衰中医证候和西医危险因素的联系,以中医学“治未病”思想为理论基础,应用数据挖掘技术和基于机器学习建立的预警模型进行心衰早期防治,是当今心脏病领域研究的重要方向。

利益冲突:本研究与其他机构无利益冲突,作者署名排名无争议。

参 考 文 献

- [1] Ambrosy AP, Fonarow GC, Butler J, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries[J]. J Am Coll Cardiol, 2014, 63 (12): 1123–1133.
- [2] Ho KK, Anderson KM, Kannel WB, et al. Survival after the onset of congestive heart failure in Framingham Heart Study subjects[J]. Circulation, 1993, 88 (1): 107–115.
- [3] Koutkias VG, Chouvarda I, and Maglaveras N. Multi-agent system architecture for heart failure management in a home care environment[A]. Computers in cardiology 2003[C]. Piscataway NJ: IEEE, 2003: 383–386.
- [4] Hunt SA, Abraham WT, Chin MH, et al. 2009 focused update incorporated into the ACC/AHA 2005 Guidelines for the Diagnosis and Management of Heart Failure in Adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines: developed in collaboration with the International Society for Heart and Lung Transplantation[J]. Circulation, 2009, 119 (14): 391–479.
- [5] Yancy CW, Jessup M, Bozkurt B, et al. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines and the heart failure society of America[J]. J Card Fail, 2017, 23 (8): 628–651.
- [6] Moons KGM, Altman DG, Vergouwe Y, et al. Prognosis and prognostic research: application and impact of prognostic models in clinical practice BMJ 2009[J]. Br Med J, 2009, 338: 606.
- [7] Dimopoulos AC, Nikolaidou M, Caballero FF, et al. Georgousopoulou EN (2018) Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk[J]. BMC Med Res Methodol, 2018, 18 (1): 179.
- [8] Kannel WB, Feinleib M, McNamara PM, et al. An investigation of coronary heart disease in families: the Framingham Offspring Study[J]. Am J Epidemiol, 2017, 185 (11): 1903–1102.
- [9] McKee PA, Castelli WP, McNamara PM, et al. The natural history of congestive heart failure: the Framingham study[J]. N Engl J Med, 1971, 285 (26): 1441–1446.
- [10] Ho KKL, Pinsky JL, Kannel WB, et al. The epidemiology of heart failure: the Framingham Study[J]. J Am Coll Cardiol, 1993, 22 (4): A6–A13.
- [11] Kannel WB, D'agostino RB, Silbershatz H, et al. Profile for estimating risk of heart failure[J]. Arch Intern Med, 1999, 159 (11): 1197–1204.
- [12] Schnabel RB, Rienstra M, Sullivan LM, et al. Risk assessment for incident heart failure in individuals with atrial fibrillation[J]. Eur J Heart Fail, 2013, 15 (8): 843–849.
- [13] Miao F, Cai YP, Zhang YT. Risk prediction for heart failure incidence within 1-year using clinical and laboratory factors[A]. 2014 36th annual international conference of the IEEE engineering in medicine biology society (EMBC) [C]. Piscataway NJ: IEEE, 2014: 1790–1793.
- [14] Bian Y, Xu F, Lv RJ, et al. An early warning scoring system for the prevention of acute heart failure[J]. Int J Cardiol, 2015, 183: 111–116.
- [15] Fatima M, Pasha M. Survey of machine learning algorithms for disease diagnostic[J]. J Intel Learn Syst Appl, 2017, 9 (1): 1–16.
- [16] Suzuki S, Yamashita T, Sakama T, et al. (2019) Comparison of risk models for mortality and cardiovascular events between machine learning and conventional logistic regression analysis[J]. PLoS One, 2019, 14 (10): e0221911.

- [17] Domingos P. A few useful things to know about machine learning[J]. Commun ACM, 2012, 55 (10): 78–87.
- [18] Yang G, Ren Y, Pan Q, et al. A heart failure diagnosis model based on support vector machine[A]. 2010 3rd international conference on biomedical engineering and informatics (BMEI 2010) [C]. Piscataway NJ: IEEE, 2010: 1105–1108.
- [19] Gharehchopogh FS, Khalifelou ZA. Neural network application in diagnosis of patient: a case study[A]. International Conference on Computer Networks and Information Technology[C]. Abbottabad: IEEE, 2011, 245–249.
- [20] Ali L, Niamat A, Khan JA, et al. An optimized stacked support vector machines based expert system for the effective prediction of heart failure[J]. Access, 2019, 7: 54007–54014.
- [21] Ali L, Khan SU, Golilarz NA, et al. A feature-driven decision support system for heart failure prediction based on χ^2 statistical model and gaussian naive bayes[J]. Comput Math Methods Med, 2019, 2019: 6314328.
- [22] Guo CY, Wu MY, Cheng HM. The comprehensive machine learning analytics for heart failure[J]. Int J Environ Res Public Health, 2021, 18 (9): 4943.
- [23] Wu ZS, Yao CH, Zhao DA, et al. A prospective cohort study on cardiovascular disease incidence in 11 provinces of China, II: associations between individual risk factor aggregation and cardiovascular disease incidence[J]. Chin J Cardiol, 2001, 29: 246–250.
- [24] 倪靖怡, 蔡海荣, 姚红, 等. 慢性心力衰竭的中医药治疗进展 [J]. 中医临究, 2017, 9 (5): 144–146.
- [25] 毛静远, 朱明军. 慢性心力衰竭中医诊疗专家共识 [J]. 中医杂志, 2014, 55 (14): 1258–1260.
- [26] 刘镇, 古春花, 雷健, 等. 慢性心力衰竭患者红细胞体积分布宽度、心力衰竭计分及中医辨证的相关性研究 [J]. 中医学报, 2015, 30 (10): 1485–1487.
- [27] 张松, 廖宪勇, 李白雪. 中医治疗慢性心力衰竭的证候分析 [J]. 四川中医, 2020, 38 (2): 208–211.
- [28] 刘镇, 刘惠灵, 李建松, 等. 慢性心力衰竭患者 RDW、Bosdon 积分及证候积分的相关性分析 [J]. 现代诊断与治疗, 2019, 30 (19): 3387–3388.
- [29] 刘德秀, 李本志. 慢性心力衰竭患者的中医证候特征及其与 B 型利钠肽的关系 [J]. 中日友好医院学报, 2019, 33 (3): 186–187, 190.
- [30] 蒋银珂, 张琪. 从中医“治未病”探讨冠心病的防治 [J]. 河北中医, 2019, 41 (3): 459–462.
- [31] 黄建波, 张光霁. 论“治未病”理论体系建设 [J]. 中华中医药杂志, 2017, 32 (3): 911–914.
- [32] 宋敏, 宋永嘉, 王凯. 论中医“治未病”思想对现代预防医学的贡献与启示 [J]. 甘肃中医药大学学报, 2020, 37 (5): 31–35.
- [33] 王娟, 赵慧辉, 陈建新, 等. 基于神经网络技术分析的慢性心力衰竭血瘀证诊断模型研究 [J]. 世界中医药, 2018, 13 (9): 2122–2126.
- [34] 朱泽豪. 心力衰竭患者反复急性发作住院的影响因素分析及列线图预测模型构建 [D]. 广州: 广州中医药大学, 2020.
- [35] Wong YW, Thomas L, Sun JL, et al. Predictors of incident heart failure hospitalizations among patients with impaired glucose tolerance: insight from the nateglinide and valsartan in impaired glucose tolerance outcomes research study[J]. Circ Heart Fail, 2013, 6 (2): 203–210.
- [36] Smith JG, Newton-Cheh C, Almgren P, et al. Assessment of conventional cardiovascular risk factors and multiple biomarkers for the prediction of incident heart failure and atrial fibrillation[J]. J Am Coll Cardiol, 2010, 56 (21): 1712–1719.

(收稿: 2021-06-29 在线: 2023-05-24)

责任编辑: 白 霞