

急性中毒临床预测模型需要关注的几个问题

石齐芳^{1,2} 张劲松^{1,2,3}

¹ 南京医科大学第一附属医院急诊医学中心, 南京 210029; ² 南京医科大学中毒研究所, 南京 211166; ³ 南京医科大学现代毒理学教育部重点实验室, 南京 211166

通信作者: 张劲松, Email: zhangjso@njmu.edu.cn

基金项目: 国家自然科学基金 (82172184)

DOI: 10.3760/cma.j.issn.1671-0282.2024.11.001

近几年来临床预测模型的研究在医学的各个领域广泛的开展, 其类型包括诊断模型和预后模型。国际权威期刊 BMJ 杂志陆续发布了一系列临床预测模型的指南性报告, 旨在规范这类研究的方法学。这些指南性报告包括: 2015 年的个人预后或诊断多变量预测模型 (TRIPOD) 声明^[1], 2019 年的临床环境中临床预测模型的呈现形式^[2], 2020 年的计算开发临床预测模型所需的样本量^[3], 2023 年的 TRIPOD-Cluster 报告, 用于指导使用聚类数据集的预测模型^[4], 2023 年的 TRIPOD-SRMA 报告规范了预测模型的系统综述和 Meta 分析的质量标准^[5], 2024 年的临床预测模型的评估指南 (外部验证)^[6-8], 基于人工智能的临床预测模型的声明 (TRIPOD-AI)^[9] 和开发临床预测模型的分步指南^[10] (图 1)。

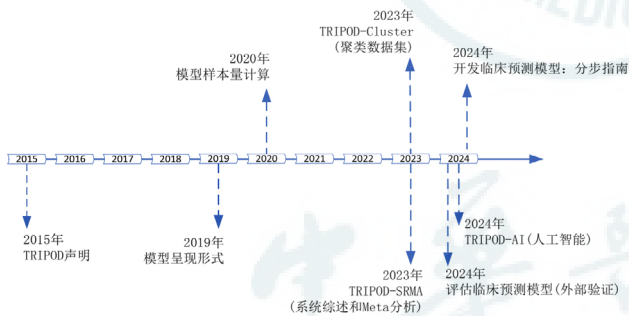


图 1 BMJ 杂志颁布的临床预测模型相关的指南性报告的时间线

急性中毒是临床医学中的一种特殊类型, 与常见疾病最大的区别在于涉及外源性物质, 具有异质性和散发性的特点^[11-12]。急性中毒的临床预测模型面临一些独特的挑战。研究表明传统的中毒严重程度评分 (poisoning severity score, PSS) 因其局限性, 在评估急性中毒的临床应用受到了限制^[12]。而新开发模型较少^[13], 并且在方法学上存在一定的争议^[14-16]。尽管已有指南性报告, 但预测模型的开发和验证仍面临诸多方法学挑战, 可能存在不少陷阱^[17]。本文基于前期指南性论文^[1-10]和重要综述^[17], 对急性中毒临床预测模型的几个特殊问题进行了深入思考和补充阐述。

1 预测模型的研究设计的问题

TRIPOD 声明指出临床预测模型的四种研究类型^[1] (图 2)。开发模型理想的类型是类型 3 (Type 3), 即独立的数据进行外部验证。需要指出的内部验证是指来自推导数据相同的数据源或基础人群, 外部验证包括时间验证 (推导数据集和验证数据集之间仅存在时间差异) 和空间验证 (不同地点进行验证)。时间验证是弱的外部验证形式, 空间验证是强的外部验证形式^[18]。

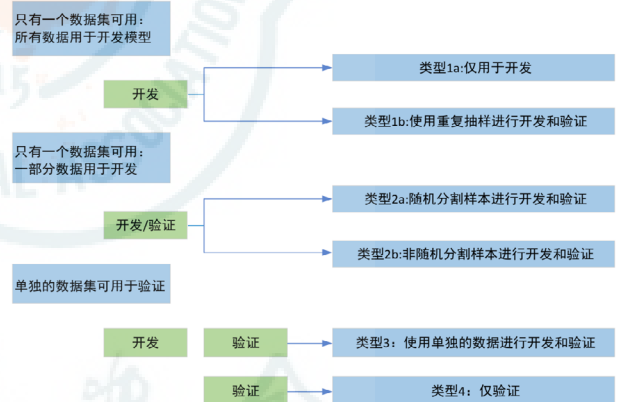


图 2 临床预测模型四种研究类型的中文版 (来源于 Collins GS, BMJ, 2015. PMID: 25569120^[1])

由于急性中毒的散发性, 样本量多数不大, 未来开展多中心合作是解决样本量小的方案之一。目前在样本量不大的情况下, 不建议拆分训练集 (开发集) 和测试集 (验证集), 因为减少了训练集的大小是对数据的低效利用, 并且增加模型过度拟合的风险^[6-8]。以评估急性中毒患者是否收入 ICU 的预测模型为例, 大多数是类型 1a (Type 1a)^[16]。然而 Type 1a 只是进行了模型的推导, 没有验证, 至少内部验证是可行的, 目前内部验证的主流方法是自举法 (Bootstrap), 其优于交叉验证法 (cross validation)^[19]。建议当样本量不大时, 选择自举法内部验证的 Type 1b 较为合适。另外, 区分度、校准度和临床适用度三个维度是模型的常用的评价方法, 以评估急性中毒患者是否收入

ICU 的预测模型为例, 大部分研究仅关注区分度, 而忽视校准度和临床适用度^[16,19]。

2 急性中毒临床预测模型的变量选择问题

模型变量的选择包括数据驱动 (data-driven) 和知识驱动 (knowledge-driven), 单独使用数据驱动或知识驱动都会导致模型指定错误^[16,20]。近几年随着大量预测模型被开发, 学者发现重新推导可能会产生具有相似变量的预测模型, 会导致多余的模型竞争解决相同的临床问题^[18]。数据驱动与知识驱动相结合的方法通常有助于构建更为稳定和有用的模型^[16,20]。知识驱动包括专家意见、专业知识和既往文献报道^[20], 目前急性中毒预测模型的变量基本大多数是基于数据驱动, 基于知识驱动的变量也应该被考虑纳入模型。

知识驱动的变量绕不开毒物的种类和中毒量, 与其他疾病不同, 急性中毒的物质是外源性物质, 不同的毒物种类毒性大小不一样, 靶器官损伤不同, 预后差异很大。发达国家的毒物暴露种类以药物多见, 总体中毒病死率不高。而国内仍以农药中毒为主^[11-12], 病死率较高。国外急性中毒模型可能在国内不能生搬硬套, 毒物的种类作为变量仍是模型开发的需要考虑的因素。另外中毒量在模型中非常重要, 百草枯中毒早期就使用百草枯中毒严重指数 (severity index of paraquat poisoning, SIPP) 评估, 其计算公式为入院时百草枯的血浆浓度 × 服毒后至入院的时间。近期开发的新型简单风险模型也包含摄入量, 另外七个变量是年龄、肌酸激酶同工酶、血小板、白细胞、中性粒细胞计数、 γ -谷氨酰转氨酶和血清肌酐^[21]。口服的剂量可能会不准确, 近几年毒物检测在临床开展持续推广, 毒物检测可以准确反应体内毒物的量^[22]。有研究已将毒物检测结果纳入为模型的变量, 如百草枯中毒患者急性肾损伤风险模型就纳入百草枯的血浆浓度, 另外四个变量是中毒至就诊时间、白细胞、胱抑素 C、天门冬氨酸氨基转移酶^[23]。建议急性中毒临床预测模型需要重视毒物检测, 尽量创造条件送毒物检测。

3 传统逻辑回归和机器学习的问题

机器学习是人工智能的一个子集, 近年来获得了相当大的普及, 在医学领域蓬勃发展。机器学习包括监督学习和无监督学习两种方法。监督学习的例子包括决策树、随机森林、朴素贝叶斯、极端梯度提升、支持向量机等^[24-25]。随着机器学习的预测模型的逐年增多, 学者发现基于机器学习的预测模型研究方法学质量较差, 偏倚风险较高^[24-25]。对此 2024 年 BMJ 杂志颁布了 TRIPPOD-AI 声明, 以提高的研究报告的质量和可用性^[9]。

机器学习构建的临床预测模型显示出比传统方

法更准确的预测, 然而仍然有一些问题需要关注。一是过拟合 (Overfitting) 的风险, 机器学习模型会在训练集显示出优越的预测性能, 但在测试集 (特别是外部测试集) 中并不稳定。近期的系统综述显示目前机器学习模型存在误导性报告 (Spin), 即选择性地报告, 将研究中的正面结果过分夸大^[25]。二是模型的可解释性, 传统的临床预测模型能以评分系统、列线图等方式呈现^[2,12,16,19], 提供了模型的可解释性和透明度。机器学习模型认为是“黑箱” (black box)^[26], 难以解释其决策过程。近几年 SHAP 方法的引入 (SHapley Additive exPlanations) 提供了机器学习模型的可解释性。SHAP 能够提供直观的可视化方法, 使得模型的决策过程更加透明^[27]。三是模型的临床应用, 机器学习模型通常通过嵌入临床决策支持系统 (clinical decision support system, CDSS) 来实现其在临床环境中的具体应用^[28]。

机器学习模型与传统预测模型在某些概念和指标上需进一步对标与融合^[29]。例如: 机器学习的验证集 (validation set) 用于调整模型参数, 类似于传统模型的内部验证集 (internal validation); 测试集 (test set) 对应外部验证集 (external validation)。此外, 机器学习中的召回率 (Recall) 对应传统模型的敏感度 (Sensitivity), 精确率 (Precision) 则对应阳性预测值 (Positive Predictive Value)。

相比于其他医学领域, 机器学习在急性中毒临床预测模型中的应用报道较少。近期 Li 等^[30]使用逻辑回归、随机森林、支持向量机和梯度提升四种机器学习算法预测敌草快中毒患者的死亡风险, 并提供了 SHAP 解释。结果显示四种模型的受试者工作特征曲线下面积分别为 0.91、0.98、0.96 和 0.94, 且净收益相似。四种机器学习模型中重要性排名前五共同包含的变量有动脉血二氧化碳分压、敌草快剂量和乳酸。机器学习在急性中毒临床预测模型中逐渐受到关注, 需要进一步探索和研究。

4 总结

本文主要探讨了急性中毒临床预测模型在开发阶段所面临的问题, 同时, 模型验证阶段也面临很多挑战, 这与其他模型验证的问题相似^[6-8]。在整个过程中, 模型的开发阶段被认为是起始和重要的环节^[16], 关注开发阶段的这些问题为后续验证^[6-8]、优化和模型系统评价奠定了基础^[5], 有助于建立更准确、可靠且具有实用性的急性中毒临床预测模型。

利益冲突 所有作者声明无利益冲突

参 考 文 献

[1] Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a

- multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement[J]. *BMJ*, 2015, 350: g7594. DOI: 10.1136/bmj.g7594.
- [2] Bonnett LJ, Snell KIE, Collins GS, et al. Guide to presenting clinical prediction models for use in clinical settings[J]. *BMJ*, 2019, 365: 1737. DOI: 10.1136/bmj.1737.
- [3] Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model[J]. *BMJ*, 2020, 368: m441. DOI: 10.1136/bmj.m441.
- [4] Debray TPA, Collins GS, Riley RD, et al. Transparent reporting of multivariable prediction models developed or validated using clustered data: TRIPOD-Cluster checklist[J]. *BMJ*, 2023, 380: e071018. DOI: 10.1136/bmj-2022-071018.
- [5] Snell KIE, Levis B, Damen JAA, et al. Transparent reporting of multivariable prediction models for individual prognosis or diagnosis: checklist for systematic reviews and meta-analyses (TRIPOD-SRMA) [J]. *BMJ*, 2023, 381: e073538. DOI: 10.1136/bmj-2022-073538.
- [6] Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation[J]. *BMJ*, 2024, 384: e074819. DOI: 10.1136/bmj-2023-074819.
- [7] Riley RD, Archer L, Snell KIE, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study[J]. *BMJ*, 2024, 384: e074820. DOI: 10.1136/bmj-2023-074820.
- [8] Riley RD, Snell KIE, Archer L, et al. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study[J]. *BMJ*, 2024, 384: e074821. DOI: 10.1136/bmj-2023-074821.
- [9] Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods[J]. *BMJ*, 2024, 385: e078378. DOI: 10.1136/bmj-2023-078378.
- [10] Efthimiou O, Seo M, Chalkou K, et al. Developing clinical prediction models: a step-by-step guide[J]. *BMJ*, 2024, 386: e078276. DOI: 10.1136/bmj-2023-078276.
- [11] Shi Q, Ba G, Xia Z, et al. The value of toxicological analysis in acute poisoning patients with uncertain exposure histories: a retrospective and descriptive study from an institute of poisoning[J]. *World J Emerg Med*, 2024, 15(2): 98-104. DOI: 10.5847/wjem.j.1920-8642.2024.022.
- [12] 石齐芳, 张劲松. 急性中毒相关评分系统的研究现状和进展 [J]. *中华急诊医学杂志*, 2023, 32(3): 434-437. DOI: 10.3760/cma.j.issn.1671-0282.2023.03.032.
- [13] Shi Q, Zhang J. Clinical prediction models for acute poisoning: not much development and even less verification[J]. *Aust Crit Care*, 2023, 36(2): 177-178. DOI: 10.1016/j.aucc.2022.11.006.
- [14] Shi Q, Zhang J. Methodological concerns about score of delayed neuropsychiatric sequelae after carbon monoxide poisoning[J]. *Am J Emerg Med*, 2023, 64: 188. DOI: 10.1016/j.ajem.2022.10.042.
- [15] Shi Q, Zhang J. Methodological concerns about the poisoning early warning score[J]. *Aust Crit Care*, 2023, 36(2): 174. DOI: 10.1016/j.aucc.2022.10.002.
- [16] Shi Q, Zhang J. Clinical prediction models for intensive care unit admission in patients with acute poisoning: is it time for a comprehensive evaluation of their utility? [J]. *Toxicol Res (Camb)*, 2024, 13(2): tfae031. DOI: 10.1093/toxres/tfae031.
- [17] Jin Y, Kattan MW. Methodologic issues specific to prediction model development and evaluation[J]. *Chest*, 2023, 164(5): 1281-1289. DOI: 10.1016/j.chest.2023.06.038.
- [18] Binuya MAE, Engelhardt EG, Schats W, et al. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review[J]. *BMC Med Res Methodol*, 2022, 22(1): 316. DOI: 10.1186/s12874-022-01801-8.
- [19] Shi Q, Dai H, Ba G, et al. Development and internal validation of a predictive model for prolonged intensive care unit stays in patients with psychotropic drug poisoning[J]. *Heart Lung*, 2024, 68: 350-358. DOI: 10.1016/j.hrtlng.2024.09.003.
- [20] Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling[J]. *Fam Med Community Health*, 2020, 8(1): e000262. DOI: 10.1136/fmch-2019-000262.
- [21] Gao Y, Liu L, Li T, et al. A novel simple risk model to predict the prognosis of patients with paraquat poisoning[J]. *Sci Rep*, 2021, 11(1): 237. DOI: 10.1038/s41598-020-80371-5.
- [22] 夏振瑜, 毛征生, 张劲松, 等. 毒物检测技术在 895 例临床中毒患者中的应用价值研究 [J]. *中华急诊医学杂志*, 2023, 32(2): 186-191. DOI: 10.3760/cma.j.issn.1671-0282.2023.02.010.
- [23] 杨莹, 蒋臻, 唐郭, 等. 百草枯中毒患者急性肾损伤风险预测 [J]. *中华急诊医学杂志*, 2023, 32(3): 332-338. DOI: 10.3760/cma.j.issn.1671-0282.2023.03.010.
- [24] Andaur Navarro CL, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review[J]. *BMJ*, 2021, 375: n2281. DOI: 10.1136/bmj.n2281.
- [25] Andaur Navarro CL, Damen JAA, Takada T, et al. Systematic review finds “spin” practices and poor reporting standards in studies on machine learning-based prediction models[J]. *J Clin Epidemiol*, 2023, 158: 99-110. DOI: 10.1016/j.jclinepi.2023.03.024.
- [26] Watson DS, Krutzinna J, Bruce IN, et al. Clinical applications of machine learning algorithms: beyond the black box[J]. *BMJ*, 2019, 364: 1886. DOI: 10.1136/bmj.1886.
- [27] Hu J, Xu J, Li M, et al. Identification and validation of an explainable prediction model of acute kidney injury with prognostic implications in critically ill children: a prospective multicenter cohort study[J]. *EClinicalMedicine*, 2024, 68: 102409. DOI: 10.1016/j.eclinm.2023.102409.
- [28] Lee KH, Choi GH, Yun J, et al. Machine learning-based clinical decision support system for treatment recommendation and overall survival prediction of hepatocellular carcinoma: a multi-center study[J]. *NPJ Digit Med*, 2024, 7(1): 2. DOI: 10.1038/s41746-023-00976-8.
- [29] Long A, Haggerty CM, Finer J, et al. Deep Learning for echo analysis, tracking, and evaluation of mitral regurgitation (DELINEATE-MR) [J]. *Circulation*, 2024, 150(12): 911-922. DOI: 10.1161/CIRCULATIONAHA.124.068996.
- [30] Li H, Liu Z, Sun W, et al. Interpretable machine learning for the prediction of death risk in patients with acute diquat poisoning[J]. *Sci Rep*, 2024, 12: 14101. DOI: 10.1038/s41598-024-67257-6.

(收稿日期: 2023-12-26)

(本文编辑: 张斯龙)