

《药物流行病学研究方法学指南（第2版）》系列解读（17）：预测模型在药物疗效与安全性评价中的应用



聂晓璐¹, 吴昉^{2,3}, 赵厚宇⁴, 詹思延^{2,3,4,5,6}, 孙 凤^{2,3,5,6,7,8,9}

1. 国家儿童医学中心, 首都医科大学附属北京儿童医院临床流行病与循证医学中心 (北京 100045)
2. 北京大学公共卫生学院流行病与卫生统计学系 (北京 100191)
3. 重大疾病流行病学教育部重点实验室 (北京大学) (北京 100191)
4. 北京大学第三医院临床流行病学研究中心 (北京 100191)
5. 北京大学医学部药品上市后安全性研究中心 (北京 100191)
6. 北京大学人工智能研究院智慧公众健康研究中心 (北京 100191)
7. 北京大学第三医院眼科 (北京 100191)
8. 新疆医科大学中医学院 (乌鲁木齐 830017)
9. 新疆石河子大学公共卫生学院 (新疆石河子 832000)

【摘要】 预测模型是实现个体化用药与精准医学的关键工具。本文以《药物流行病学研究方法学指南（第2版）》为基础，系统阐述预测模型在药物疗效与安全性评价中的四大核心应用场景：治疗效应异质性识别（谁应用药）、生物标志物发现（为何有效）、用药患者疗效预测（效果如何）及药物安全性评价（是否安全）。在此基础上，进一步探讨四方面的特殊方法学考量：预测模型与因果模型的本质区别、类不平衡与罕见事件预测的处理策略、从预测到决策的三层级转化路径（内部验证、外部验证、临床效用评估），以及TRIPOD+AI与PROBAST+AI所构建的报告规范与偏倚风险评价闭环。本文旨在为药物流行病学研究者在开展预测模型研究时提供方法学指引，推动该领域从“模型开发”向“临床可用”的范式转变。

【关键词】 预测模型；药物流行病学；精准医学；疗效评价；安全性评价；治疗效应异质性；生物标志物

【中图分类号】 R181.3+5 **【文献标志码】** A

Guide on Methodological Standards in Pharmacoepidemiology (2nd edition) and their series interpretation (17): application of predictive models in drug effectiveness and safety evaluation

NIE Xiaolu¹, WU Yunxiao^{2,3}, ZHAO Houyu⁴, ZHAN Siyan^{2,3,4,5,6}, SUN Feng^{2,3,5,6,7,8,9}

1. Center for Clinical Epidemiology & Evidence-based Medicine, Beijing Children's Hospital, Capital

DOI: 10.12173/j.issn.1005-0698.202605004

基金项目：国家自然科学基金面上项目（82204149、72474008）；国家自然科学基金国际（地区）合作与交流项目（72361127500）；海南省科学技术厅重点研发专项（ZDYF2024LCLH002）；国家药品监督管理局药品评价中心开放课题（CDR2024R01001）；药物不良反应杂志《2025年药物安全研究项目》（ADR2025MS01）

通信作者：孙凤，研究员，博士研究生导师，Email: sunfeng@bjmu.edu.cn

詹思延，教授，博士研究生导师，Email: siyan-zhan@bjmu.edu.cn

<https://ywlxwb.whuzhmedj.com/>

Medical University, National Center for Children's Health, Beijing 100045, China

2. Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China

3. Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing 100191, China

4. Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing 100191, China

5. Center of Postmarketing Safety Evaluation, Peking University Health Science Center, Beijing 100191, China

6. Center for Intelligent Public Health, Institute for Artificial Intelligence, Peking University, Beijing 100191, China

7. Department of Ophthalmology, Peking University Third Hospital, Beijing 100191, China

8. School of Traditional Chinese Medicine, Xinjiang Medical University, Urumqi 830017, China

9. School of Public Health, Shihezi University, Shihezi 832000, Xinjiang Uygur Autonomous Region, China

Corresponding authors: SUN Feng, Email: sunfeng@bjmu.edu.cn; ZHAN Siyan, Email: siyan-zhan@bjmu.edu.cn

【Abstract】 Predictive models are essential methodological tools for individualized therapy and precision medicine. Based on the *Guide on Methodological Standards in Pharmacoepidemiology (2nd Edition)*, this paper systematically elucidates four core application scenarios of predictive models in drug effectiveness and safety evaluation: heterogeneity of treatment effects identification (who should receive the drug), biomarker discovery (why it is effective), effectiveness prediction for patients receiving drugs (how effective it is), and drug safety evaluation (whether it is safe). Furthermore, four special methodological considerations are discussed: the essential distinction between predictive and causal models, strategies for handling class imbalance and rare event prediction, the three-level translational pathway from prediction to decision-making (internal validation, external validation, and clinical utility assessment), and the closed-loop system for reporting specification and bias risk assessment formed by TRIPOD+AI and PROBAST+AI. This paper aims to provide methodological guidance for pharmacoepidemiologists and facilitate paradigm shift from "model development" to "clinical utility" in this field.

【Keywords】 Predictive models; Pharmacoepidemiology; Precision Medicine; Effectiveness evaluation; Safety evaluation; Heterogeneity of treatment effects; Biomarkers

个体化用药与精准医学是当代医学的核心追求，而预测模型正是将这一愿景转化为临床实践的关键工具。药物流行病学作为运用流行病学方法研究人群中药物利用与效应的应用学科，其核心使命是在真实世界环境中评估药物的疗效与安全性。在这一背景下，中国药学会于2024年启动了《药物流行病学研究方法学指南（第2版）》（以下简称“指南第2版”）的制订/修订工作^[1]。相比第1版，在研究设计、数据源、具体应用场景等方面进行了重要更新，并增加了人工智能应用等全新内容。本文以指南第2版的方法学框架为基础，系统阐述预测模型在药物疗效与安全性评价中的应用。需要说明的是，本文所述预测模型特指以个体化预测为目的的建模方法——即在单臂或双臂数据框架下，基于患者特征估计结局概率或风险分层；而对于以因果效应估计为目的的因果模型，如基于效应建模的治疗效应异质性

（heterogeneity of treatment effect, HTE）分析，虽有涉及但仅作方法学界定，具体方法细节不在本文详细展开。尽管二者在方法学上有重叠（如均需调整混杂、处理缺失数据等问题），但本文重点关注预测模型（主要为预后预测模型）的开发、验证与报告规范。文章从四大核心应用场景——HTE识别、生物标志物发现、用药患者疗效预测、药物安全性评价——分别展开方法学论述与案例剖析，进而系统梳理预测模型应用于药物疗效与安全性评价的特殊方法学考量，以期为该领域的规范化研究提供参照。

1 预测模型在药物疗效与安全性评价中的核心应用场景

预测模型在药物流行病学中的应用，围绕药物疗效与安全性评价两大核心任务，可归纳为四个既相互区别又彼此关联的核心场景（图1）：



图 1 预测模型在药物疗效与安全性评价中的核心应用场景

Figure 1. Core application scenarios of predictive models in drug efficacy and safety evaluation

HTE 识别、生物标志物发现、用药患者疗效预测、药物安全性评价。这四个场景贯穿药物从上市前安全性筛选到上市后个体化用药管理的全生命周期，分别对应了个体化用药决策中的四个关键问题——“谁应用药”（HTE 识别）、“为何有效”（生物标志物发现）、“效果如何”（用药患者疗效预测）以及“是否安全”（药物安全性评价）。

1.1 HTE 识别

随机对照试验（randomized controlled trial, RCT）作为评价药物疗效的金标准，其所得的疗效实则为药物的平均治疗效应。然而，这一群体层面的证据在支持为患者制定个体化治疗决策过程中难以直接映射到个体患者^[2-3]。传统的亚组分析试图通过检验治疗效应在特定患者亚群中的异质性来解决这一问题。然而，基于单个或少量基线变量进行的亚组分析面临多重局限：统计功效不足、多重比较导致的假阳性风险、以及难以整合多维预后信息。更为关键的是，个体患者往往同时属于多个不同的亚组，而不同亚组在获益与否及获益方式上可能存在差异^[4]。单变量亚组分析无法反映患者整体风险状态对治疗效应的影响，而这一影响恰恰是临床决策的核心依据。

精准医学的兴起，进一步提高了对识别重要 HTE 的关注度^[5]。基于患者的特征组合精准识别 HTE，判断其从治疗中获益的预期幅度。这超越了传统“一刀切”的治疗指南，是实现精准用药的核心方法。2020 年，由美国以患者为中心的结

局研究中心资助并发布了《治疗效应异质性预测方法声明》（Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement)^[6-7]。该声明系统阐述了如何在 RCT 研究设计中采用能够同时整合多个患者特征的预测建模方法，以识别 HTE 并预测个体化的治疗效应。识别 HTE 的两种建模路径：基于风险建模（risk-based modeling）与基于效应建模（effect-based modeling）。需要重点说明的是，从方法学上，基于风险建模属于预测模型用于 HTE 识别的重要应用场景，而基于效应建模则属于因果模型，与风险建模在建模目标、适用场景等方面有较大差别。由于篇幅所限，本文在 HTE 识别场景中的详细介绍将基于风险建模的方法及其应用案例，而对于因果机器学习应用于 HTE 效应建模的方法将另有后续解读文章进行深入解析。

根据数据来源不同（包括 RCT 与观察性数据），将风险建模识别 HTE 的研究分为如下两类。

1.1.1 基于 RCT 数据开展风险建模识别 HTE

风险建模方法不要求识别具体的效应修饰因子（若研究存在很明确的效应修饰因子，则应使用基于效应建模的方法更为准确），而是通过预后风险评估实现“风险分层”基础上的效应检验。风险建模的核心逻辑建立在“风险放大”这一数学关系之上：若治疗在相对尺度上效应同质（即相对风险比恒定），则患者的基线预后风险越高，其从治疗中获得的绝对获益越大。这一关系可形式化表达为：绝对获益=基线风险×（相对获益）。基于这一原理，风险建模通过两个步骤实现从预后评估到治疗决策的转化^[6]：首先，基于多变量回归模型，利用患者的多维基线特征（用药前特征）构建预后风险模型，估算其在对照条件下的主要结局风险；第二步根据所计算的基线风险将患者划分为不同风险层级（通常可采用分位数），在各风险层级中检验治疗效应的异质性，从而识别出哪些患者群体绝对获益最大或哪些患者可能遭受净伤害。需要说明的是，在第一步利用基线特征构建预后风险模型时，通常采用外部模型（external model），即采用非本次研究所用 RCT 数据所构建、已经过验证后发表的既有风险评估模型；若外部模型不可用，可直接构建内部模型（internal model），即基于本次 RCT 人群治疗

分配前的基线信息构建风险模型。阿哌沙班预防癌症患者静脉血栓栓塞 (venous thromboembolism, VTE) 的 AVERT 试验^[8], 纳入 Khorana 临床风险评分 (Khorana Risk Score, KRS) ≥ 2 分的具有潜在发生静脉血栓栓塞的较高风险癌症患者, 为探讨能否通过生物标志物 (D-二聚体) 进一步识别出“更受益”和“不受益”的患者亚群而开展二次分析。研究者使用经外源数据验证发表的 Vienna CATScore 预测评分, 结合肿瘤风险分类以及 D-二聚体连续变量预测个体 6 个月 VTE 风险 (%), 并通过决策曲线分析确定了 8% 为最佳 VTE 风险阈值。该研究还进一步通过计算需治疗人数 (number needed to treat, NNT) 来量化不同风险亚组的治疗效果, 从而识别 HTE。研究结果证实高风险患者 (6 个月 VTE 风险 $\geq 8\%$) 从阿哌沙班血栓预防中获益巨大 [HR (95%CI): 0.33 (0.14, 0.81), NNT=6], 而低风险患者 (<8%) 无任何获益 [HR (95% CI): 0.89 (0.30, 2.65), NNT=167]。该研究为“精准血栓预防”提供了方法学范例, 通过风险建模实现从“一刀切”到“风险靶向”的治疗策略优化。除采用单个 RCT 数据开展 HTE 研究外, 还可使用个体参与者数据的 Meta 分析 (individual participant data Meta) 数据^[9]。

1.1.2 基于观察性数据开展风险建模识别 HTE

目标试验仿真 (target trial emulation, TTE) 框架的提出和推广, 有效降低了传统观察性研究的设计偏倚, 使其成为连接观察性研究数据与因果推断的桥梁。将 TTE 框架与 HTE 识别结合应用于观察性研究数据, 可为个体化用药决策提供参考依据。Xu 等^[10] 为探讨胰高血糖素样肽-1 受体激动剂 (glucagon-like peptide-1 receptor agonists, GLP-1RA) 是否能降低成年 2 型糖尿病患者发生心力衰竭住院 (hospitalization for heart failure, HHF) 风险, 以及治疗效应是否为类别效应或因具体药物而异, 研究使用瑞典斯德哥尔摩肌酐测量项目 (Stockholm Creatinine Measurements project) 的医疗保健数据 (2010-2021 年) 开展了 TTE 研究。该研究纳入新使用 GLP-1RA 与二肽基肽酶-4 抑制剂 (dipeptidyl peptidase-4 inhibitors, DPP-4i) 的成年 2 型糖尿病患者, 使用 Cox 回归估计意向治疗的 HHF 风险

比, 在进一步探讨 HTE 时, 根据糖尿病心肌梗死溶栓治疗心力衰竭风险评分 (Thrombolysis in Myocardial Infarction Risk Score for Heart Failure in Diabetes, TRS-HF_{DM}) 所量化的心衰基线风险进行分层^[10]。结果发现, 与初始使用 DPP-4i 相比, 初始使用 GLP-1RA 的患者 3 年 HHF 绝对风险更低 (3.4% vs. 4.3%), 对应的加权风险比为 0.77 (95% CI: 0.66, 0.91)。GLP-1RA 与 DPP-4i 在 HHF 方面的绝对风险差 (risk difference, RD) 在基线心衰风险预测较高的患者中最大: 即 TRS-HF_{DM} 评分为 0 的患者, 3 年的 RD 为 -0.06% (95% CI: -0.41%, 0.35%), 而 TRS-HF_{DM} 评分 ≥ 2 的患者, 3 年的 RD 为 -1.77% (95% CI: -3.23%, -0.13%)。

1.2 生物标志物发现

生物标志物是实现精准医学的核心工具, 可反映或部分解释个体化治疗异质性产生的生物学基础, 而预测模型正成为识别药物疗效/安全性生物标志物的关键方法学路径。基因组水平的生物标志物主要有单核苷酸多态性 (single nucleotide polymorphism, SNP)、拷贝数变异 (copy number variation, CNV)、结构变异 (structural variation, SV) 等几种, 最终在表型上均体现为个体差异。其中, 最为普遍的多态性为 SNP^[11]。UK Biobank 分子药物流行病学项目是该方向的典型代表。该项目利用大规模基因组数据与电子健康记录的链接, 聚焦于心血管代谢药物的遗传药理学评价^[12]。目前, 药物疗效相关的生物标志物识别已从基因组层级发展到多组学 (蛋白质组、代谢组等), 通过整合真实世界队列构建预测模型来筛选预测药物反应的生物标志物。

基于现有研究, 预测模型在药物疗效和安全性生物标志物的发现可归纳为以下三种方法路径: 生物学知识嵌入、多模态数据融合以及网络驱动预测。

1.2.1 生物学知识嵌入

在基因数量 (~20 000) 远大于样本量 (通常数百至数千) 的情况下, 模型可能学习到统计上显著但生物学上荒谬的关联, 所发现的标志物可能是统计噪音, 缺乏生物学合理性。在多组学数据丰富、先验知识可靠的前提下, 通过将先验知识 (如 KEGG 通路、GO 注释) 作为结构约束嵌入模型, 引导模型关注具有生物学意义的特征组合。K-net 模型是该路径的代表性工作。K-net 本

质上是一个生物学信息引导的稀疏神经网络，利用KEGG通路信息对网络结构施加约束——只有KEGG中注释的基因-基因相互作用才被允许建立连接，同时采用Lasso特征筛选，使得每个通路中只有少量“代表性基因”被保留，其余权重被压缩至零。上述双重稀疏化机制有效控制了模型的复杂度，使其快速学习生物学上合理的特征组合，而非捕捉训练集中的虚假相关，最终在输出层输出药物反应（如敏感或耐药）概率。应用该模型，成功在奥希替尼耐药性预测中识别出KRAS突变、TP53突变和AKT3过表达等关键生物标志物，并首次揭示LUAD以KRAS突变为主要耐药机制、SCLC以AKT3过表达为主要耐药机制的亚型特异性差异^[13]。这一框架代表了从“黑箱预测”到“机制揭示”的方法学进阶——通过将生物学知识从“辅助验证工具”提升为“模型架构设计原则”，为预测模型在药物疗效与安全性评价中的生物标志物发现提供了可复制的范式。

1.2.2 多模态数据融合

在药物疗效与安全性评价中，传统的生物标志物发现主要依赖单模态数据（如仅基因组学或仅临床指标）。然而，疾病的发生发展是遗传、环境、分子和临床因素相互作用的结果，单一数据源存在信息碎片化、生物学异质性无法表征等根本局限性。多模态数据融合通过整合基因组学、转录组学、蛋白质组学、医学影像和临床表型等多源异构数据，在统一的特征空间中学习跨模态的互补信息，从而识别出单模态分析无法发现的药物反应生物标志物。例如，部分局部晚期直肠癌患者对新辅助化疗（neoadjuvant chemotherapy, NAC）反应不佳，不仅延误了治疗时机，还需承受不必要的药物毒性。为了在治疗前识别出这些患者，一项发表于2025年的研究构建了一个创新的多模态深度学习框架^[14]：将磁共振成像（magnetic resonance imaging, MRI）影像与蛋白质组学数据进行整合，使用3D ResNet152网络从MRI图像中提取肿瘤的宏观病理学特征（如肿瘤周围基质异质性），同时利用图神经网络分析组织样本中的蛋白表达谱，并结合KEGG/GO通路与蛋白互作网络信息。该综合模型在预测NAC反应方面表现优异，测试集的受试者工作特征曲线下面积（area under the curve, AUC）显著优于仅使用MRI影像或仅使用蛋白质组数据的单

模态模型。决策曲线分析也证实了其临床实用性。SHAP分析进一步显示，MRI来源的特征贡献了57.7%的预测能力，而蛋白质组学则识别出了10种关键的化疗耐药相关蛋白，如CYBA、GUSB、ATP6AP2等。这些发现共同揭示了影像表型与分子机制间的关联。

1.2.3 网络驱动预测

网络驱动预测建模的核心技术特点在于超越传统的“单基因-药物”关联分析，通过整合蛋白质-蛋白质相互作用网络、基因调控网络或信号通路等生物学信息，构建能够反映复杂分子间互作模式的预测模型。该方法具有较高的计算复杂性，利用网络可解释性使得模型预测结果可溯源至具体的分子通路机制，另一方面也可通过推断网络发现未知网络关系。这种“生物学先验编码+网络拓扑分析”的范式标志着药物基因组学生物标志物发现正从统计关联向机制驱动的方向演进。例如，急性髓系白血病（acute myeloid leukemia, AML）是一种髓系祖细胞的克隆性恶性肿瘤，患者对当前治疗方案的反应高度异质，凸显了精准医学的迫切需求。然而，由于细胞异质性的存在，可靠的精准治疗生物标志物仍然难以获得。传统的基于Bulk RNA测序和体外药敏实验的模型，往往无法捕捉治疗反应和耐药性背后的复杂分子通路及基因互作网络。Wang等^[15]使用520例AML患者的RNA-Seq数据以及87种临床药物的体外药敏数据，开发出NetAML网络驱动精准医学平台，该平台不将基因视为孤立的预测因子，而是在基因互作网络的框架下，通过推断共表达网络，识别能够反映复杂分子交互模式的基因特征。基于NetAML平台，研究发现*C19ORF59*基因与*FLT3*基因的共表达，与AML患者对FLT3抑制剂类药物的耐药性显著相关。

1.3 用药患者疗效预测

已使用药物的患者，特别是长期用药患者的疗效预测，是预测模型应用于药物流行病学的重要场景之一。根据预测信息是否随治疗过程更新，可划分为静态预测与连续动态预测两类，二者在数据结构、方法学假设和临床决策支持模式上存在本质差异。

1.3.1 静态预测

静态预测仅使用患者治疗前的基线特征（如人口学变量、基线生物标志物、基线影像学指

标) 构建模型, 预测固定时间窗口内的疗效结局, 其核心优势在于数据采集成本低、实施简便、结果可在治疗开始前预先获知。静态预测的价值在于为初始治疗方案选择提供依据。若预测为“无反应”, 临床医生可在治疗开始前考虑替代方案, 避免无效试错。然而, 静态预测无法利用治疗过程中陆续获得的反应信息(如早期症状变化、药物浓度监测、生物标志物轨迹), 适用于“治疗前决策”场景, 如RCT入组筛选、一线方案选择, 但对治疗过程中的动态调整缺乏指导能力。

Morikubo 等^[16]开展的一项多中心回顾性研究, 纳入活动期溃疡性结肠炎患者作为训练队列, 采集治疗前49项临床特征(包括背景因素、临床活动度、内镜活动度及血液检查结果), 采用随机森林(random forest, RF)算法筛选预测因子, 并构建了基于支持向量机多项式核函数的预测模型。研究结果显示, 血清白蛋白、单核细胞百分比、身高、平均红细胞体积等10项基线特征被识别为治疗6个月后达到无激素临床缓解的重要预测因子。该模型在独立外部验证队列中表现出稳健的预测效能(阳性预测值68.8%, 阴性预测值71.4%)。值得注意的是, 研究者发现既往针对维得利珠单抗开发的预测模型在应用于乌司奴单抗时预测效能显著下降(阳性预测值仅56.3%), 这一发现有力论证了药物特异性预测模型开发的必要性, 即不同作用机制的生物制剂需要独立建立预测工具, 而不能简单外推。

1.3.2 连续动态预测

与仅使用基线信息的静态预测不同, 动态预测的核心特征是: 在患者治疗过程中, 利用不断积累的纵向数据(包括多次测量的生物标志物、患者报告结局、累计用药信息等)实时更新对长期预后的判断, 为长期用药管理提供“随数据积累而进化的风险预警”——临床医生可在决策时间点(如治疗2周期后、6个月后), 基于历史全部数据判断“当前方案是否仍合适”, 支持“继续、调整还是终止治疗”的个体化决策, 从而触发方案调整(如更换二线药物), 避免不可逆的临床恶化。常用的动态预测模型方法有: landmark分析、联合模型(joint model)以及近年来应用较多的基于深度学习的建模。

Landmark分析是一种处理时依协变量的动态生存预测方法。其核心思想是在预设的某个时间

点(如治疗后第6个月), 利用该时间点及之前收集的全部信息构建预测模型, 预测该时间点之后的生存概率。Landmark方法可处理高维纵向数据, 不受参数模型假设约束。该方法因为需选定时间点, 因此时间点的选择可能会引入偏倚。另外, 无法对纵向测量与事件时间之间的内生性关联进行建模。联合模型是一类同时拟合纵向子模型和生存子模型的统计框架, 通过共享随机效应关联纵向生物标志物轨迹与事件发生时间, 实现对个体化动态风险的精准预测。该方法的优势在于可量化生物标志物变化速率对风险的影响, 可处理纵向测量与事件时间之间的内生性关联。但由于联合模型计算复杂度高, 同时对纵向数据的测量方法(时间点、频率)敏感, 因此, 通常需要较大的样本量支持模型收敛。在药物疗效的动态预测领域, 深度学习模型(循环神经网络、神经微分方程及端到端深度学习架构等不同技术路径), 凭借其处理复杂时间序列数据的能力, 展现出优于传统方法(landmark分析和联合模型)的性能。传统药代动力学模型在预测重症患者万古霉素血药浓度时, 常因患者生理状态的剧烈波动而失效。为解决这一问题, Park等^[17]开发了一个名为DeepTDM的综合决策支持系统, 凸显了深度学习在个体化用药监测中的应用价值, 即通过持续利用电子病历中的实时数据, 实现对药物浓度的动态校准与精准预测。该模型核心是一个门控循环单元(GRU)集成的联合多层感知器网络(JointMLP)。该模型由三个模块构成, 通过整合回归与分类预测的双重机制, 利用患者的人口学特征、肾功能指标、用药剂量及既往TDM结果等纵向时序数据, 预测后续万古霉素谷浓度。在内部验证队列及包含不同种族的MIMIC-IV数据集中, 该模型均展现出优于传统PK模型及其他机器学习方法的预测精度。

1.4 药物安全性评价

药物安全性评价是药物流行病学的核心任务之一, 其意义在于识别和量化药物在真实世界人群中的不良反应风险, 为临床安全用药、药品监管决策及上市后风险控制提供关键证据。根据建模时机分为上市前化合物毒性预测及上市后基于真实世界数据的风险评估。

1.4.1 上市前化合物毒性预测

上市前化合物毒性预测研究的意义在于, 在

药物进入人体试验或上市之前, 基于化学结构等分子特征提前识别潜在的毒性风险 (如肝肾毒性、血液毒性、心脏毒性、致癌性等), 从而优化化合物筛选、降低临床开发后期失败率, 并从根本上提升药物研发效率与安全性。同时也可以预测已上市药品, 特别是为新药提供潜在安全性的风险提示。其核心优势体现在: 无需等待临床数据积累, 可在研发早期对大量候选化合物进行快速、高通量的虚拟筛选, 显著减少对动物实验的依赖, 并规避跨物种差异带来的预测偏差。在此背景下, “冷启动” 场景——即预测那些在训练数据中未出现的药物 (或称“未见药物”) 的毒性风险——以及对于上市药物风险评估具有关键意义, 通过评估能够检验模型对未知化学空间的泛化能力, 避免模型仅记忆已知药物的毒性模式, 从而为监管机构和药物开发者提供更可靠、更贴近实际应用场景的安全性预测依据, 构建以“发现为导向”的药物安全评价体系^[18]。

包含药品种类丰富且经过准确标注的药物毒性基准数据库, 是支撑上述研究最为关键的数据基础^[19-22]。这类数据库系统整合了多种药物在上市前后对特定或多类不良事件的发生情况、是否存在药源性毒性等重要信息, 同时还可涵盖药物结构、化学分类等多维度的丰富内容。随着精准医学理念的深入推进以及高通量技术的迅猛发展, 高质量的基准数据库不仅为未来融合蛋白组学、基因组学等多维数据进行个体化药物毒性评估研究奠定了坚实的数据基础, 也为全球范围内不同研究方法之间的横向比较与持续优化提供了统一的标准与平台。美国食品药品监督管理局的 SafetAI 计划是由药物评价与研究发展中心与国家毒理学研究中心联合发起的一项合作倡议, 旨在开发一套基于深度学习的定量构效关系 (Quantitative Structure-Activity Relationship, QSAR) 模型, 用于在候选药物进入临床试验前的研究性新药审评阶段评估其安全性^[23]。该计划聚焦五大关键毒理学终点: 肝毒性、致癌性、致突变性、肾毒性和心脏毒性, 其核心创新在于根据化合物的化学特征进行个体化毒性预测, 初步研究显示该方法在上述毒性终点上相较于传统深度学习和 QSAR 方法有显著提升^[21, 24-25]。SafetAI 旨在为新药审评过程中的安全性评估提供关键信息, 减少对动物实验的依赖, 提高药物安全性评价的精准度。

1.4.2 药品上市后安全性评估

药品上市后风险预警是药物警戒体系的重要内容, 其意义在于可识别临床试验中未能发现的药品不良反应 (adverse drug reaction, ADR), 发现用药数年后才显现的迟发性毒性。另外, 通过监测特殊人群 (如老年人、儿童、孕妇、肾功能不全者) 的用药安全, 为完善药品说明书修订和风险管理措施提供重要的证据支持。

传统的药品安全信号检测方法 (如比例报告比、报告比值比、信息成分等) 虽然广泛应用于 ADR 自发报告系统, 但需要在数据积累到足够数量后才能发现信号, 且无法实现个体化风险分层。预测模型的引入, 正是为了弥补这一缺陷, 实现从“被动响应”向“主动预警”的范式转变, 有效提升药物警戒效率, 辅助早期识别高风险 ADR, 优化医疗资源分配。刘溪等^[26]基于 ADR 自发报告数据, 系统构建严重 ADR [根据国家常见严重 ADR 评价标准^[27]及国外 CTCAE 不良事件通用术语评价标准 (V5.0)^[28]进行综合判定] 的智能预测模型, 比较了伯努利朴素贝叶斯 (Bernoulli Naive Bayes, BNB)、RF 算法等 10 种机器学习算法, 并评估 SMOTE、ADASYN、TomekLinks 等采样技术对数据不平衡 (严重 ADR 占 10.4%) 的处理效果。结果发现将 DeepSeek 大语言模型应用于结构化 ADR 数据标准化处理, 针对高维稀疏特征, 欠采样技术 TomekLinks 优于过采样方法。BNB 算法作为贝叶斯定理的经典分类算法, 凭借其高效的分类能力, 在众多算法中表现优异。

除使用 ADR 自发呈报数据开展安全信号挖掘以外, 医院电子病历数据、医疗保险索赔数据等电子健康记录 (electronic health record, EHR) 也是开展药品安全预警监测的重要数据来源^[29]。与自发呈报仅记录“可疑药物-不良事件”的相对较粗略信息不同, EHR 整合了患者的全方位、多维度诊疗数据 (包括结构化数据, 如诊断、医嘱、检验结果、生命体征, 以及非结构化数据, 如病程记录、影像报告、出院小结等), 贯穿药品从采购、处方、调配、给药到监测的整个链条, 实现闭环追溯。利用 EHR 大数据, 可以计算出特定人群 (如使用某种药物的患者) 发生某一事件 (如急性肝损伤) 的“背景发生率”。当实际发生频率显著高于背景率时, 系统可自动触发

预警^[30]。EHR中的丰富信息(如基因检测结果、既往病史、合并用药)使得风险预警不再是“一刀切”,同时也可开展用药安全风险的动态预测。药物洗脱支架植入后,患者需接受双联抗血小板治疗(dual anti-platelet therapy, DAPT)以预防缺血事件(支架血栓),但DAPT同时增加出血风险。临床面临的困境是:DAPT的最佳持续时间因人而异,而传统风险评估工具基于基线特征,无法随治疗过程动态更新。Rasmy等^[31]利用Cerner Health Facts数据库和Optum Clinformatics数据库的回顾性数据,开发并验证了名为“AI-DAPT”的AI驱动动态预测模型。研究将36个月的随访期划分为6个顺序预测窗口(每6个月一个窗口),在每个窗口内利用截至该时间点的患者临床数据更新后续缺血和出血风险的预测。研究者评估了5种AI算法,最终加权LightGBM表现最优。基于此,AI-DAPT模型在30~36个月窗口达到最佳预测性能:缺血事件AUC为90%(95%CI: 88%, 92%),出血事件AUC为84%(95%CI: 82%, 87%)。

重点人群用药安全监测也同样是预测模型应用于药品安全评价的重要任务之一。在老年人群中,同时服用多种药物的情况十分常见,不合理的用药方案容易诱发药品不良事件(adverse drug event, ADE)。准确预测老年人多重用药相关的不良事件,对于降低再住院率、减少医疗资源消耗、提高老年患者生活质量以及实现个体化用药管理具有重要的临床与公共卫生意义。某些ADE在单独使用某一种药物时发生风险较低,往往只有在多种药物联合使用的特定组合条件下才会出现,因此具有罕见性特征。传统统计方法由于受到模型假设和数据运算能力的限制,预测效果并不理想;龙昱军等^[32]总结了机器学习技术在多重用药ADE预测中的应用现状,研究发现,RF算法和图神经网络等凭借其强大的数据建模与处理能力,能够显著提高预测的准确性,尤其在药物相互作用风险的识别方面展现出明显优势。

2 预测模型应用于药物疗效与安全性评价的特殊方法学考量

2.1 预测模型与因果模型的区别与联系

在药物流行病学研究中,“预测”与“因果”是两个经常被混淆的概念。预测模型与因果模型

在目标、方法和解读上存在本质区别,但在实际研究中常被混淆。预测模型旨在基于个体特征(X)准确估计结局(Y)的条件概率分布,回答“会发生什么”或“谁风险更高”的预测性问题。其核心是最小化预测误差,因此可纳入中介变量或非结构化特征以提升准确性,而不要求变量间存在因果解释。相比之下,因果模型则关注干预或暴露(A)对结局(Y)的效应估计,回答“如果改变某个因素,结局会如何变化”的干预性问题^[33]。二者在较多方面存在差异(表1),特别是对协变量的处理策略截然不同——预测模型可纳入中介变量以改善预测准确性,而因果模型若纳入中介变量则会阻断因果路径,导致过度校正偏倚。因果模型必须处理混杂偏倚,遵循有向无环图(directed acyclic graph, DAG)原则调整变量——不可纳入中介变量或碰撞变量(否则会引入选择偏倚),且需满足可交换性、正性、一致性等假设。研究者若不明确自身研究属于预测还是因果框架,可能导致模型选择错误、变量调整不当,进而影响结论的可信度。在实际应用中,二者可形成互补——预测模型负责高风险人群的筛选,因果模型在该人群中验证治疗的净获益——但研究者必须在方法学上明确其研究归属,避免将关联信号误解为因果证据,或将因果外推建立在预测模型之上^[34]。

2.2 类不平衡与罕见事件预测

在药物安全性评价中,不良反应通常属于罕见事件(发生率常低于1%甚至0.1%),导致建模数据呈现严重的类别不平衡。这种不平衡会直接影响传统机器学习算法的学习效率——模型倾向于将所有样本预测为“多数类”(无事件),虽然宏观准确率看似很高,但真正需要被识别出的少数类(不良反应事件)却几乎无法被检出,从而失去临床预警价值。由于受限于样本量,临床试验难以发现发生率低于1/1 000的罕见不良反应;上市后真实世界数据虽然样本量大,但阳性样本的极端稀疏性依然是核心分析障碍。此外,自发报告系统还将面对报告偏倚(漏报、迟报、选择性报告),模型实际学习到的是“被报告的ADR模式”,而非真实的ADR发生风险。

需要采取特殊的设计与评估策略以应对上述挑战。首先,评估指标不应再依赖传统的准确率或AUC,而应优先采用精确率-召回率曲线下面积

表 1 预测模型与因果模型的区别对比

Table 1. Comparison of differences between prediction models and causal models

维度	预测模型	因果模型
核心问题	结局是什么?	为什么发生? 干预会改变什么?
典型提问	“谁会得病?” “谁会对药物有反应?”	“药物 A 是否引起结局 B?” “如果将剂量从 X 提高到 Y, 结局会变化多少?”
建模目标	最小化预测误差	无偏估计因果效应
关注焦点	关联的强度与方向	关联的因果解释
对混杂的态度	可纳入混杂变量以改善预测	必须控制混杂, 否则估计有偏
对中介变量的处理	可纳入, 它们携带预测信息	不可纳入, 会阻断因果路径, 导致“过度校正偏倚”
对碰撞变量的处理	不可纳入, 引入选择偏倚	不可纳入, 同样引入选择偏倚
变量筛选原则	以预测准确性为导向	以因果结构 (DAG) 为指导
模型输出	结局概率、风险评分、分类标签	因果效应估计值 (ATE、ATT、CATE)、风险比、率差
验证标准	区分度 (AUC)、校准度、预测误差	因果识别的有效性 (无未测量混杂、正性假设等)
可解释性要求	可选 (黑箱可接受, 只要预测准确)	必需 (需明确因果假设和识别策略)
对数据的要求	可来自任意分布 (包括非随机化数据)	需满足因果识别条件 (如可交换性、正性、一致性)

(area under the precision-recall curve, AUPRC)、F1-score、灵敏度与阳性预测值组合等对不平衡数据更为稳健的指标, 因 AUC 可能因真阴性率高而被夸大, 无法反映模型对少数类的识别能力。其次, 可通过重采样技术 (如对多数类欠采样的 TomekLinks 方法、对少数类过采样的 SMOTE/ADASYN 方法) 或代价敏感学习 (为少数类赋更高权重) 来平衡类别分布^[35]; 但在高维稀疏特征场景 (如自发报告数据) 中, 欠采样方法通常优于过采样, 因其不易放大噪声^[26]。最后, 由于罕见事件的低发生率特性, 外部验证格外重要——模型需在独立、多源数据的测试集中验证其泛化能力, 避免模型过度拟合特定数据库的报告偏倚模式; 缺乏外部验证的模型对真实临床场景的推广价值极为有限^[36]。

2.3 从预测到决策的转化路径

预测模型输出的“风险概率”或“反应概率”与临床医生的实际决策之间存在根本性的“转化鸿沟”: 前者是统计学上的条件概率, 而后者是综合了获益、风险、患者意愿和资源可及性的复杂权衡。一个 AUC 达到 0.90 的模型, 若无法回答“当概率超过多少时应采取何种行动”, 便难以被临床采纳。跨越这一鸿沟, 需要在传统技术验证的基础上, 建立三个递进的“转化验证层级”^[37]。

内部验证 (技术可靠性), 回答“模型可重复吗?” 的问题。这是转化的技术基础, 要求模型在同一数据集的交叉验证或自助法中表现出稳定的区分度和校准度。其中, 校准度尤为关键——一个区分度很高但校准不佳的模型 (如将

10% 的实际风险高估为 50%), 将误导临床决策。外部验证 (场景可迁移性), 回答“模型可泛化吗?” 的问题。要求模型在独立的外部数据 (不同地域、不同机构、不同时间段的人群) 中保持稳定的预测性能。外部验证是连接“研究环境”与“真实临床环境”的桥梁。由于不同医疗中心的患者构成、诊断标准、数据采集方式存在系统性差异, 缺乏外部验证的模型在实际部署中常出现性能衰减, 即“验证鸿沟”。临床效用评估 (决策价值), 回答“模型有用吗?” 的问题。这是转化的最终环节, 要求在前瞻性研究中证明: 模型指导下的临床决策能够改善患者结局或优化医疗资源配置。具体方法包括: ①决策曲线分析, 评估模型在不同概率阈值下的临床净获益; ②整群随机试验, 比较“模型预警组”与“常规管理组”在不良事件发生率、治疗调整时效性等结局上的差异; ③卫生经济学评价, 量化模型应用的成本-效果与成本-效用比^[38]。

以上三个层级构成了“技术可行性→场景可迁移性→临床价值”的完整证据链。当前的大部分研究仍止步于第一层级, 而第二、三层级的缺失正是预测模型“发表多、落地少”的结构根源。研究者应在模型开发阶段即明确定义临床行动路径 (如“当预测概率>0.3 时, 建议加强每周一次的血常规监测”; “当预测概率>0.7 且经因果模型确认后, 建议考虑替代疗法”), 并通过前瞻性研究验证该路径的实际临床价值。经过上述三个步骤的转化, 预测模型才能从“统计学意义上的成功”真正转化为“临床决策中可信赖的利器”。

2.4 报告规范及偏倚风险评价工具

预测模型的报告规范及偏倚风险评价工具，是确保模型研究透明性、可复现性与临床可信度的方法学基石。在报告规范方面，TRIPOD+AI声明（2024年更新版）是当前核心指南^[39]，其将传统回归方法与机器学习方法统一纳入包含27个主条目的报告框架，要求研究者完整报告研究设计、数据处理（如缺失值插补、类别不平衡处理）、模型开发（如超参数调优）、开放科学（如代码共享）及算法公平性评估等关键信息，旨在解决预测模型研究普遍存在的报告不完整、不透明问题。该规范取代了2015年版TRIPOD，是预测模型研究的新报告标准^[40]。另外，需要说明的是，目前尚无专门针对HTE预测研究的报告规范发布，已有学者正在开展聚焦生物医药领域的因果和反事实预测模型报告规范的建议^[41]。

在偏倚风险评价方面，PROBAST工具（Prediction model Risk Of Bias ASsessment Tool）于2019年由国际专家团队开发^[42]，通过参与者、预测因子、结局和分析四个领域共20个信号问题系统评估预测模型开发或验证研究的偏倚风险与适用性。随着人工智能与机器学习技术的普及，原工具对AI模型特有的技术问题（如类别不平衡处理、超参数调优、数据泄露、算法公平性等）覆盖不足。为此，PROBAST+AI于2025年发布^[43]，将评估框架分别细化为模型开发的16个质量评估条目（侧重方法学质量）和模型评估的18个偏倚风险评估条目（侧重性能验证的可靠性），并首次纳入公平性与伦理风险评估维度。该工具适用于所有基于回归或AI/ML方法构建的诊断/预后模型研究，是系统评价及证据质量分级的标准依据。

上述两项工具共同构成了预测模型研究的方法学保障闭环：TRIPOD+AI确保“报告得清”（研究设计、执行、结果全透明），PROBAST+AI确保“评价得准”（偏倚风险、适用性系统评估）。这一闭环体系有助于缓解当前预测模型领域“发表多、落地少”的结构性问题——报告不完整导致读者无法判断模型可信度，偏倚风险高限制模型临床转化。研究者应在模型开发阶段即参照TRIPOD+AI设计研究方案，在成果发表时遵循其完成报告，在证据合成时应用PROBAST+AI进行质量审查，从而从源头提升预测模型研究的

方法学质量。

3 结语

在药物流行病学中，预测模型的核心价值不在于追求AUC的数值提升，而在于能否真正回答个体化用药决策中的四个关键问题——“谁应用药、为何有效、效果如何、是否安全”。本文围绕HTE识别、生物标志物发现、用药患者疗效预测及药物安全性评价四大核心场景，系统梳理了预测模型的方法学路径与特殊考量，强调从“群体证据”向“个体决策”的转化需依次跨越内部验证（技术可靠性）、外部验证（场景可迁移性）与临床效用评估（决策价值）三个层级，并遵循TRIPOD+AI与PROBAST+AI所构建的报告与评价闭环。唯有将方法学严谨性、报告透明性与临床需求紧密结合，预测模型才能真正成为精准医学时代药物流行病学研究中值得信赖的工具。

利益冲突声明：作者声明本研究不存在任何经济或非经济利益冲突。

参考文献

- 1 颜济南, 吴昉, 聂晓璐, 等. 《中国药物流行病学研究方法学指南(第2版)》的制订/修订过程[J]. 药物流行病学杂志, 2025, 34(2): 121-135. [Yan JN, Wu YX, Nie XL, et al. Revision process of the Guide on Methodological Standards in Pharmacoepidemiology in China (2nd edition)[J]. Chinese Journal of Pharmacoepidemiology, 2025, 34(2): 121-135.] DOI: [10.12173/j.issn.1005-0698.202502028](https://doi.org/10.12173/j.issn.1005-0698.202502028).
- 2 Mant D. Can randomised trials inform clinical decisions about individual patients? [J]. Lancet, 1999, 353(9154): 743-746. DOI: [10.1016/S0140-6736\(98\)09102-8](https://doi.org/10.1016/S0140-6736(98)09102-8).
- 3 Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification[J]. JAMA, 2007, 298(10): 1209-1212. DOI: [10.1001/jama.298.10.1209](https://doi.org/10.1001/jama.298.10.1209).
- 4 Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation[J]. Lancet, 2005, 365(9454): 176-186. DOI: [10.1016/S0140-6736\(05\)17709-5](https://doi.org/10.1016/S0140-6736(05)17709-5).
- 5 Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects[J]. BMJ, 2018, 363: k4245. DOI: [10.1136/bmj.k4245](https://doi.org/10.1136/bmj.k4245).
- 6 Kent DM, Paulus JK, van Klaveren D, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement[J]. Ann Intern Med, 2020, 172(1): 35-45. DOI: [10.7326/M18-3667](https://doi.org/10.7326/M18-3667).

- 7 Kent DM, van Klaveren D, Paulus JK, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement: Explanation and Elaboration[J]. *Ann Intern Med*, 2020, 172(1): W1–W25. DOI: [10.7326/m18-3668](https://doi.org/10.7326/m18-3668).
- 8 Kumar V, Shaw JR, Key NS, et al. D-dimer enhances risk-targeted thromboprophylaxis in ambulatory patients with cancer [J]. *Oncologist*, 2020, 25(12): 1075–1083. DOI: [10.1002/onco.13540](https://doi.org/10.1002/onco.13540).
- 9 Rekkas A, van Klaveren D, Ryan PB, et al. A standardized framework for risk-based assessment of treatment effect heterogeneity in observational healthcare databases[J]. *NPJ Digit Med*, 2023, 6(1): 58. DOI: [10.1038/s41746-023-00794-y](https://doi.org/10.1038/s41746-023-00794-y).
- 10 Xu Y, Huang T, Zhang Y, et al. Risk of heart failure hospitalization for GLP-1 receptor agonists versus DPP-4 inhibitors or SGLT-2 inhibitors in patients with type 2 diabetes: a target trial emulation[J]. *Circulation*, 2026, 153(15): 1086–1100. DOI: [10.1161/CIRCULATIONAHA.125.075157](https://doi.org/10.1161/CIRCULATIONAHA.125.075157).
- 11 张伟霞. 基于药物基因组学的药物流行病学研究设计概述[J]. *药物流行病学杂志*, 2021, 30(12): 787–792. [Zhang WX. Overview of pharmacoepidemiological research design based on pharmacogenomics[J]. *Chinese Journal of Pharmacoepidemiology*, 2021, 30(12): 787–792.] DOI: [10.19960/j.cnki.issn1005-0698.2021.12.001](https://doi.org/10.19960/j.cnki.issn1005-0698.2021.12.001).
- 12 Uk Biobank. Molecular pharmacoepidemiology: the assessment of the effectiveness and safety of commonly prescribed medications in routine clinical practice according to genetic variability [EB/OL]. (2025-03-18) [2026-04-12]. <https://www.ukbiobank.ac.uk>.
- 13 Ran M, Zhang SL, Tam KY. Identifying meaningful drug response biomarkers from public pharmacogenomic datasets with biologically informed interpretable neural networks[J]. *Comput Biol Chem*, 2026, 120(Pt 1): 108669. DOI: [10.1016/j.combiolchem.2025.108669](https://doi.org/10.1016/j.combiolchem.2025.108669).
- 14 Li Y, Ding J, Du F, et al. Deep learning-based multimodal prediction of NAC response in LARC by integrating MRI and proteomics[J]. *Cancer Res Treat*, 2025, DOI: [10.4143/crt.2025.707](https://doi.org/10.4143/crt.2025.707).
- 15 Wang Y, Liu R, Zhang Y, et al. A network-driven framework for drug response precision prediction of acute myeloid leukemia[J]. *Adv Sci (Weinh)*, 2025, 12(36): e06447. DOI: [10.1002/adv.202506447](https://doi.org/10.1002/adv.202506447).
- 16 Morikubo H, Tojima R, Maeda T, et al. Machine learning using clinical data at baseline predicts the medium-term efficacy of ustekinumab in patients with ulcerative colitis[J]. *Sci Rep*, 2024, 14(1): 4386. DOI: [10.1038/s41598-024-55126-1](https://doi.org/10.1038/s41598-024-55126-1).
- 17 Park J, Kim D, Lee D, et al. DeepTDM: deep learning-based prediction of sequential therapeutic drug monitoring levels of vancomycin[J]. *IEEE J Transl Eng Health Med*, 2025, 13: 493–506. DOI: [10.1109/JTEHM.2025.3623605](https://doi.org/10.1109/JTEHM.2025.3623605).
- 18 聂晓璐, 赵厚宇, 霍东辉, 等. 基于真实世界数据与定量构效关系的儿童药源性免疫性血小板减少症研究[J]. *中国药物警戒*, 2025, 22(1): 16–22. [Nie XL, Zhao HY, Huo DH, et al. Pediatric drug-induced immune thrombocytopenia based on real world data and quantitative structure-activity relationship[J]. *Chinese Journal of Pharmacovigilance*, 2025, 22(1): 16–22.] DOI: [10.19803/j.1672-8629.20240807](https://doi.org/10.19803/j.1672-8629.20240807).
- 19 聂晓璐, 孙凤, 阎爱侠, 等. 基于多源信息构建药品不良反应基准数据库[J]. *中国药物警戒*, 2025, 22(1): 10–15, 36. [Nie XL, Sun F, Yan AX, et al. Methodology for constructing benchmark database of adverse drug reactions based on multi-source information[J]. *Chinese Journal of Pharmacovigilance*, 2025, 22(1): 10–15, 36.] DOI: [10.19803/j.1672-8629.20240911](https://doi.org/10.19803/j.1672-8629.20240911).
- 20 姚克宇, 张舒琪, 金锐, 等. 药物致肝损伤数据集的比较研究[J]. *中国药物警戒*, 2023, 20(5): 568–573. [Yao KY, Zhang SQ, Jin R, et al. Comparative study on data set of drug-induced liver injury[J]. *Chinese Journal of Pharmacovigilance*, 2023, 20(5): 568–573.] DOI: [10.19803/j.1672-8629.20220447](https://doi.org/10.19803/j.1672-8629.20220447).
- 21 Connor S, Li T, Qu Y, et al. Generation of a drug-induced renal injury list to facilitate the development of new approach methodologies for nephrotoxicity[J]. *Drug Discov Today*, 2024, 29(4): 103938. DOI: [10.1016/j.drudis.2024.103938](https://doi.org/10.1016/j.drudis.2024.103938).
- 22 Qu Y, Li T, Liu Z, et al. DICTrank is a reliable dataset for cardiotoxicity prediction using machine learning methods[J]. *Chem Res Toxicol*, 2025, 38(4): 647–655. DOI: [10.1021/acs.chemrestox.4c00428](https://doi.org/10.1021/acs.chemrestox.4c00428).
- 23 US Food and Drug Administration. SafetAI Initiative[EB/OL]. (2025-09-05) [2026-04-13]. <https://www.fda.gov/about-fda/nctr-research-focus-areas/safetai-initiative>.
- 24 Li T, Tong W, Roberts R, et al. DeepCar: deep learning-powered carcinogenicity prediction using model-level representation[J]. *Front Artif Intell*, 2021, 4: 757780. DOI: [10.3389/frai.2021.757780](https://doi.org/10.3389/frai.2021.757780).
- 25 Li T, Liu Z, Thakkar S, et al. DeepAmes: a deep learning-powered Ames test predictive model with potential for regulatory application[J]. *Regul Toxicol Pharmacol*, 2023, 144: 105486. DOI: [10.1016/j.yrtph.2023.105486](https://doi.org/10.1016/j.yrtph.2023.105486).
- 26 刘溪, 李晨, 田源, 等. 基于自发报告数据驱动机器学习模型预测严重药品不良反应[J]. *中国药物警戒*, 2025, 22(12): 1410–1417. [Liu X, Li C, Tian Y, et al. Prediction of severe adverse drug reactions based on spontaneous reporting data-driven machine learning models[J]. *Chinese Journal of Pharmacovigilance*, 2025, 22(12): 1410–1417.] DOI: [10.19803/j.1672-8629.20250512](https://doi.org/10.19803/j.1672-8629.20250512).
- 27 NMPA. Measures for the administration of reporting and monitoring of adverse drug reactions [EB/OL]. (2011-05-04) [2026-04-25]. <https://www.nmpa.gov.cn/yaopin/yypfgwj/yypfgbmgh/20110504162501325.html>.
- 28 National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0[EB/OL]. (2017-11-27) [2026-04-26]. <https://dctd.cancer.gov/research/ctep-trials/for-sites/adverse-events/ctcae-v5-5x7.pdf>.
- 29 江永贤, 赵珊, 李张悦, 等. 人工智能技术在药物不良事件预测中的应用现状及模型性能评估: 一项系统文献分析[J]. *药品不良反应杂志*, 2025, 27(12): 735–744. [Jiang YX, Zhao S, Li ZY, et al. Application status of artificial intelligence technology in adverse drug event prediction and model performance

- evaluation: a systematic literature analysis[J]. *Adverse Drug Reactions Journal*, 2025, 27(12):735–744.] DOI: [10.3760/ema.j.cn114015-20241211-00200](https://doi.org/10.3760/ema.j.cn114015-20241211-00200).
- 30 Kim JV, Davis SE, Matheny ME, et al. Integrating electronic health records with other data sources for postmarket drug safety signal identification: a review[J]. *Front Drug Saf Regul*, 2024, 4: 1428831. DOI: [10.3389/fdsfr.2024.1428831](https://doi.org/10.3389/fdsfr.2024.1428831).
- 31 Li F, Rasmy L, Xiang Y, et al. Dynamic prognosis prediction for patients on DAPT after drug-eluting stent implantation: model development and validation[J]. *J Am Heart Assoc*, 2024, 13(3): e029900. DOI: [10.1161/JAHA.123.029900](https://doi.org/10.1161/JAHA.123.029900).
- 32 龙昱军, 刘剑钊, 杨智荣. 机器学习在多重用药不良事件预测中的应用: 典型方法与案例的综述[J]. *药物流行病学杂志*, 2026, 35(3): 327–335. [Long YJ, Liu JZ, Yang ZR. Machine learning for predicting adverse drug events in polypharmacy: a review of methods and case studies[J]. *Chinese Journal of Pharmacoepidemiology*, 2026, 35(3): 327–335.] DOI: [10.12173/j.issn.1005-0698.202505005](https://doi.org/10.12173/j.issn.1005-0698.202505005).
- 33 Abécassis J, Dumas É, Alberge J, et al. From prediction to prescription: machine learning and causal inference for the heterogeneous treatment Effect[J]. *Annual Review Biomedical Data Science*, 2025, 8: 381–404. DOI: [10.1146/annurev-biodatasci-103123-095750](https://doi.org/10.1146/annurev-biodatasci-103123-095750).
- 34 Feuerriegel SAO, Frauen D, Melnychuk V, et al. Causal machine learning for predicting treatment outcomes[J]. *Nat Med*, 2024, 30: 958–968. DOI: [10.1038/s41591-024-02902-1](https://doi.org/10.1038/s41591-024-02902-1).
- 35 周江杰, 王予童, 冯天, 等. 大数据驱动的罕见事件非均衡数据分析方法研究进展[J]. *药物流行病学杂志*, 2025, 34(8): 952–961. [Zhou JJ, Wang YT, Feng T, et al. Research progress on big-data-driven analysis strategies for imbalanced data of rare events[J]. *Chinese Journal of Pharmacoepidemiology*, 2025, 34(8): 952–961.] DOI: [10.12173/j.issn.1005-0698.202411080](https://doi.org/10.12173/j.issn.1005-0698.202411080).
- 36 Chalabianloo N, Ahmadi F, Omrani MA, et al. Machine learning methods for predicting adverse drug events: a systematic review[J]. *Br J Clin Pharmacol*, 2026, 92(2):422–444. DOI: [10.1002/bcp.70377](https://doi.org/10.1002/bcp.70377).
- 37 Collins GAO, 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](https://doi.org/10.1136/bmj-2023-074819).
- 38 Amusa T, Okunola D, Izinyon O, et al. Strategies for embedding prediction models in clinical decision-making workflows[J]. *Cureus*, 2026, 18(1): e101185. DOI: [10.7759/cureus.101185](https://doi.org/10.7759/cureus.101185).
- 39 Collins GAO, Moons KAOX, Dhiman PAO, 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](https://doi.org/10.1136/bmj-2023-078378).
- 40 Cohen JF, Bossuyt PMM. TRIPOD+AI: an updated reporting guideline for clinical prediction models[J]. *BMJ*, 2024, 385: q824. DOI: [10.1136/bmj.q824](https://doi.org/10.1136/bmj.q824).
- 41 Xu J, Guo Y, Wang F, et al. Protocol for the development of a reporting guideline for causal and counterfactual prediction models in biomedicine[J]. *BMJ Open*, 2022, 12(6): e059715. DOI: [10.1136/bmjopen-2021-059715](https://doi.org/10.1136/bmjopen-2021-059715).
- 42 Moons KGM, Wolff RF, Riley RD, 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–W33. DOI: [10.7326/M18-1377](https://doi.org/10.7326/M18-1377).
- 43 Moons KAOX, Damen JAO, Kaul TAO, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods[J]. *BMJ*, 2025, 388: e082505. DOI: [10.1136/bmj-2024-082505](https://doi.org/10.1136/bmj-2024-082505).

收稿日期: 2026年05月06日 修回日期: 2026年05月11日

本文编辑: 李绪辉 桂裕亮