人工智能在罕见病超说明书用药不良反应主动监测与预警中的应用

王权，许嘉越，付昱*（中山大学附属第三医院药剂科，广东 广州 510630）

摘要：超说明书用药是罕见病治疗的主要手段，但其不良反应风险缺乏系统评估，传统药物警戒体系因数据稀疏、报告不足和信号识别困难而难以有效覆盖。人工智能（Artificial Intelligence，AI）技术为构建主动、精准、实时的不良反应监测与预警体系提供了新路径。本文综述了AI在罕见病超说明书用药不良反应主动监测中的应用进展，涵盖隐匿患者识别、基于多源数据的智能风险预警以及面向临床的实用化系统（如DrugGPT、“协和·太初”“哪吒·灵童”）。尽管前景广阔，AI应用仍面临数据稀缺与异质性、模型可解释性不足、隐私伦理风险及监管框架缺失等挑战。未来需医学、药学、数据科学与政策制定者协同合作，推动建立以患者为中心、数据驱动、可信可控的罕见病用药安全智能监测新生态。

关键词：人工智能；罕见病；超说明书用药；不良反应

中图分类号：R969　　**文献标志码：**A　　**文章编号：**1674-229X（2026）07-0530-06

Doi：10.12048/j.issn.1674-229X.2026.07.007

Application of Artificial Intelligence in Active Surveillance and Early Warning of Adverse Drug Reactions Associated with Off-Label Drug Use in Rare Diseases

WANG Quan，XU Jiayue，FU Yu*（*Department of Pharmacy*，*The Third Affiliated Hospital of Sun Yat-sen University*，*Guangzhou*，*Guangdong* 510630，*China*）

ABSTRACT: Off-label drug use serves as the primary therapeutic approach for rare diseases，yet systematic evaluation of the risks of its associated adverse drug reactions （ADRs） is lacking. Conventional pharmacovigilance systems are often inadequate due to data sparsity，underreporting，and challenges in signal detection. Artificial intelligence （AI） enables the development of proactive，precise，and real-time ADR surveillance and early-warning frameworks. This review summarizes recent advances in applying AI for active surveillance of off-label ADRs in rare diseases. Key applications include identification of concealed patients，intelligent risk early warning based on multi-source data，and clinical decision support systems powered by AI，such as DrugGPT，“Xiehe-TaiChu” and “Nezha-LingTong”. Despite their promising prospects，significant challenges persist，including data scarcity and heterogeneity，limited model interpretability，privacy and ethical concerns，and the absence of a mature regulatory framework. Future efforts require multidisciplinary collaboration among clinicians，pharmacists，data scientists，and policymakers to establish a patient-centered，data-driven，trustworthy and controllable intelligent monitoring ecosystem for the medication safety of rare diseases.

KEYWORDS: artificial intelligence；rare diseases；off-label drug use；adverse drug reactions

罕见病（Rare disease，RDs），又称“孤儿病”，是一类发病率极低的疾病的统称。罕见病的诊疗是全球公共卫生领域的重大挑战。尽管单个罕见病的患病率较低，但全球已确认的罕见病种类约达7 000种，累计影响人群约占全球总人口的10%。然而，目前仅有不足6%的罕见病拥有经监管机构批准的治疗方案，凸显出该领域存在巨大的未满足医疗需求[1]。在此背景下，超说明书用药（Off-label drug use，OLDU）——即药品在临床实践中被用于未经药品监督管理部门核准的适应证、剂量、给药途径、用药频率、疗程或适用人群等情形——已成为罕见病诊疗中的常见策略[2]。由于获批疗法极度匮乏，临床医生往往不得不依赖超说明书用药以治疗患者的病情，但其疗效与安全性常缺乏充分的循证医学证据支持[3-5]。我国政府高度重视罕见病防治工作。自“十三五”以来，已陆续推出涵盖诊疗、研

基金项目：广东省药学会科学研究基金（2025HJB01004）

作者简介：王权，博士，研究方向：临床药学　　***通信作者：**付昱，主管药师，研究方向：临床药学，E-mail：fuyu28@mail.sysu.edu.cn

发、审评、生产及医保的全链条支持政策[6]。针对超说明书用药这一现实需求，相关部门积极推动《罕见病超说明书用药专家共识》的制定，旨在规范临床实践、提升用药合理性与安全性[7]。

然而，部分超说明书用药缺乏系统性安全评估，其潜在的不良反应（Adverse drug reaction，ADR）风险尚不明确。传统药物警戒（Pharmacovigilance，PV）体系主要依赖自发报告和人工分析，在应对罕见病这类低频、高异质性事件时存在灵敏度低、滞后性强等显著局限。《中国超药品说明书用药管理指南（2021）》明确指出，应加强对超说明书用药的不良反应/不良事件监测与评估，并倡导在被动监测基础上，开展主动监测与流行病学研究，以提升用药安全水平[8]。

近年来，人工智能（Artificial Intelligence，AI）技术的迅猛发展，特别是以机器学习（Machine learning，ML）和深度学习（Deep learning，DL）为核心，结合自然语言处理（Natural language processing，NLP）领域中大语言模型（Large language ，odel，LLM）等前沿突破，为罕见病超说明书用药的不良反应主动监测与早期预警提供了全新的技术路径与解决方案。本文旨在系统综述AI在该领域的固有困境、核心价值、应用现状、现存挑战及未来发展方向，以期为构建更智能、更高效、更精准的罕见病药物警戒体系提供理论参考与实践指引。

1 罕见病超说明书用药ADR监测与预警的固有困境与AI应用的核心价值

罕见病的临床实践与超说明书用药密不可分。鉴于超过90%的罕见病缺乏获批疗法，临床医生普遍依赖超说明书用药以应对患者迫切的治疗需求[3,4]。因此，罕见病的ADR监测问题，本质上可归结为其超说明书用药的安全性评估问题。然而，罕见病药物警戒本身所固有的多重障碍——如患者基数小、疾病异质性强、诊疗路径复杂等在超说明书用药这一特殊场景下被进一步放大，对传统药物警戒体系构成严峻挑战。基于此，本节将首先剖析传统药物警戒体系在此领域面临的具体挑战，进而阐述人工智能技术为应对这些挑战所提供的核心价值。

1.1 罕见病超说明书用药ADR监测与预警的固有困境

1.1.1 罕见病ADR数据稀疏与异质性导致信号识别困难　罕见病的低患病率直接导致其相关ADR数据极度稀疏。上市前临床试验因受试者数量有限（通常仅数十至数百例），难以充分捕捉低频但潜在严重的不良反应。根据统计学原理，在1个300例的队列中若未观察到某特定ADR，仅能以95%的置信水平排除其发生率高于1%的可能性[9]。这意味着大量非频发但具有临床意义的ADR可能在药品上市前被遗漏。

此外，罕见病常具有复杂的亚型分类和高度多样的临床表型，且重症患者常伴随与基础疾病进展高度重叠的症状（如疲劳、肝酶升高、神经系统异常等），使得ADR与疾病本身表现之间的因果关系难以厘清，存在显著的“混杂偏倚”[9]。数据稀疏性与高度异质性的双重制约，严重削弱了传统自发报告系统及基于频率的定量信号检测方法（如贝叶斯置信传播神经网络、比例失衡法等）在罕见病场景下的敏感性与特异性。

1.1.2 罕见病监测覆盖不足　罕见病普遍存在诊断延迟问题，其成因涉及多个层面：在疾病层面，临床表现非特异、表型谱宽泛，易被误诊为常见病；在医疗系统层面，基层医生对罕见病认知不足、缺乏专科经验，且特异性诊断工具与生物标志物严重匮乏[10,11]。在资源分布层面，优质诊疗资源高度集中，患者就医路径冗长、转诊链条断裂，进一步加剧延误[10,11]。

在此背景下，传统依赖医务人员或患者自愿上报的被动药物警戒体系面临结构性失效。一方面，报告流程繁琐、激励机制缺失，导致临床一线报告意愿低；另一方面，罕见病患者群体分散、教育程度与健康素养参差不齐，加之长期疾病负担沉重，使其主动参与ADR报告的积极性极低[12]。即便对于已确诊且需使用治疗窗窄的孤儿药患者，治疗药物监测也因样本量过小而难以建立可靠的药代动力学-药效学模型，同时缺乏标准化检测方法与统一临床指南，限制了该技术在真实世界中的普及与深度应用[13]。

1.2 AI在罕见病超说明书用药ADR监测与预警中的核心价值

1.2.1 整合多源数据并主动挖掘风险信号　AI技术凭借其强大的高维数据处理、模式识别与预测能力，正推动药物警戒从被动响应向主动预警范式转型。借助NLP和DL技术，AI可自动化整合电子健康记录（Electronic health records，EHR）、医学文献、

真实世界数据（Real-world evidence，RWE）等多源异构信息，有效打破“数据孤岛”，构建跨域关联的知识图谱[14,15]。

在此基础上，ML算法能够持续、动态地从稀疏、噪声密集的真实世界数据中识别微弱但具临床意义的风险信号，揭示人眼难以察觉的复杂药物-事件关联。这种“数据驱动”的主动监测模式显著提升了ADR信号检测的灵敏度、时效性与覆盖广度，为超说明书用药的安全性评估提供了前所未有的可能性[16]。

1.2.2 实现个体化风险预测与精准决策　AI可融合多维度个体特征——包括基因组信息、表型数据、既往用药史、合并症及实验室指标等——构建个性化ADR风险预测模型，实现对特定患者不良反应的早期预警[17]。

在缺乏权威指南支持的超说明书用药场景中，此类模型可为临床医生提供基于循证的决策支持，例如推荐更安全的替代药物、优化给药剂量或调整联合用药方案，从而在保障疗效的同时最大限度降低风险[18]。这种“以患者为中心”的智能辅助，有望显著提升罕见病治疗的安全边界与临床可行性。

1.2.3 增强系统可信度并赋能闭环管理生态　通过可解释人工智能（Explainable AI，XAI）技术，AI系统可清晰呈现其风险预测的关键依据（如特定基因变异、实验室异常值或用药时间窗），增强模型输出的透明性与临床可接受度。这种可解释性不仅是AI融入临床工作流的前提，也是满足监管机构对算法治理与责任追溯要求的关键[19]。

更重要的是，AI的价值不仅体现在单点预警，更在于其赋能构建“监测-预警-干预-反馈”的闭环管理生态。通过集成临床决策支持、实时用药提醒、可穿戴设备连续监测及患者自报症状等多模态数据流，AI系统可实现对超说明书用药全过程的动态追踪与智能调优。这不仅有助于及时识别并干预潜在ADR，还能持续优化治疗策略，最终改善患者预后。

随着多模态数据融合、联邦学习、隐私计算等技术的成熟，以及适应AI驱动药物警戒的法规与伦理框架逐步完善，AI有望成为构建下一代智能、高效、以患者为中心的罕见病药物安全主动监测体系的核心引擎。

2　AI在罕见病超说明书用药ADR监测与预警中的应用现状

构建有效的罕见病超说明书用药ADR主动监测体系，不能仅局限于对已报告不良反应的被动分析，而必须形成覆盖患者发现、风险预判、监测预警与决策干预的完整闭环。人工智能技术正通过赋能以上3个关键阶段，重塑这一闭环的每一个环节，从而系统性地提升监测能力。

2.1　罕见病隐匿患者识别与表型挖掘

实现主动监测的第一步，是解决“监测谁”的问题。大量罕见病患者因诊断延迟而隐匿于医疗系统，构成了监测的前置盲区。AI技术，特别是基于EHR的表型挖掘，正在破解这一难题，为后续监测锁定目标人群。

通过监督或半监督机器学习算法，AI可从已确诊患者的EHR中自动学习疾病特有的临床表型模式（如实验室异常组合、症状共现规律、用药轨迹等），并以此构建泛化能力强的筛查模型。例如，针对普通变异型免疫缺陷病（Common variable immunodeficiency，CVID）开发的PheNet算法，利用多维度疾病编码构建高维表型特征，在全院级EHR中进行大规模筛查[20]。外部验证显示，其识别出的高危人群中，74%经临床评估高度疑似CVID，其中超过半数患者的诊断时间可提前1年以上，从而使得相关超说明书用药的安全监测得以早期介入[20]。

更进一步，通用化深度学习模型正推动罕见病筛查从“单病种”向“多病种”跃迁。RarePT（Rare Phenotype Transformer）是一种基于Transformer架构的多任务学习模型，可在数百万份EHR中同步预测155种罕见病的患病风险，并成功识别出32种疾病的未确诊患者[21]。这种高通量、广谱覆盖的筛查能力，极大拓展了主动药物警戒的潜在受监人群基数，为后续个体化用药安全评估奠定了数据基础。

2.2　罕见病超说明书用药ADR智能风险预警

在识别目标患者群体后，监测的核心在于如何“早发现、早预警”。NLP技术可以自动从文献与报告中提取药物-事件关联，而ML模型能够从这些高维、稀疏的数据中，识别出人眼难以察觉的微弱不良反应信号[14]。基于自发报告系统（如FAERS和Vigibase）的不成比例分析（从传统的PRR到贝叶斯方法如MGPS）是当前识别潜在ADR信号的主要手段。然而，研究表明，未来的范式将转向利用ML与

DL算法，对更丰富的数据源（如电子健康记录）进行挖掘[22]。图神经网络（Graph neural networks，GNN）能够有效建模药物、靶点、基因、ADR等实体之间的复杂关系，捕捉中的多层次交互[23,24]。例如，HyperADRs框架利用分层超图来预测药物-基因-ADRs风险[24]。PreciseADR模型也利用异构图神经网络整合患者、疾病、药物和ADRs之间的关系，提高了患者个体水平ADRs预测的准确性[23]。结合多种药物特性（如分子结构、生物相互作用）的融合DL模型可以提高ADR预测的全面性和准确性。例如，BiMPADR模型整合药物基因表达特征，通过消息传递神经网络预测新药的ADR[25]。基于Transformer的神经网络可用于多标签ADR预测，处理药物分子结构数据[26]。另外，超说明书用药常涉及缺乏明确指南的复杂多药联用，其药物相互作用（Drug-Drug Interaction，DDI）风险未知。基于ML和知识图谱的AI模型，能够预测潜在的、未被文献记载的DDI，实现对联合用药安全性的前瞻性评估[27]。这标志着监测从“识别已知关联”向“预判未知风险”的跨越，是主动预警的高级形态。

2.3 面向临床的实用化AI系统

前沿风险预测算法的临床转化，核心在于构建可靠、可用的系统。为克服通用大模型（如ChatGPT）在医疗领域存在的“幻觉”与证据溯源难题，领域内催生了如DrugGPT[28]、PharmaGPT[29]、ShennongMGS[30]、DoctorGPT[31]等一系列专业化大模型。这些模型通过整合权威药学知识库与进行生物医学领域特定训练，显著提升了任务的准确性与可靠性，为临床实用化奠定了基础。其中西湖大学与牛津大学等机构开发的DrugGPT系统代表了该领域的前沿。它通过问诊分析、知识获取、证据生成三大协同模块，整合Drugs.com、NHS、PubMed等权威知识库，在药物推荐、剂量调整、不良反应识别等任务上，准确率较通用大模型有显著提升，这为医生处理超说明书用药的复杂决策提供了可溯源、可解释的辅助工具[28]。ShennongMGS基于LLM的中药指导系统，专为中医药领域的复杂多药方场景和药物不良反应预测而设计[30]。

更值得关注的是，国内领先医疗机构正加速推进AI能力的系统集成，打造专科化、全流程的罕见病智能平台。2025年2月16日，北京协和医院与中国科学院自动化研究所联合开发的“协和·太初”罕见病人工智能大模型正式投入临床使用[32]。该模型基于中国人群基因检测数据与罕见病知识库构建，是国际首个针对中国人群特点设计的罕见病辅助诊断工具。“协和·太初”具备渐进式推理能力，可帮助医生快速识别罕见病，显著缩短确诊时间，并缓解全国范围内诊疗水平不均衡的问题。而华中科技大学附属同济医院联合神州医疗科技股份有限公司，正式发布全球首个儿童罕见病AI大模型——“哪吒·灵童”，核心突破在于将患者识别、风险预警、决策支持等离散环节自动化串联，形成以患者为中心、数据驱动的全程化管理生态。笔者所在的中山大学附属第三医院于2026年1月9日首次推出“中山三院人工智能医生”。该系统深度融合了超130万条真实脱敏诊疗数据、70余万份临床病例报告等资料，旨在助力罕见病诊疗。

3 AI在罕见病超说明书用药ADR监测与预警中的挑战

尽管前景广阔，AI在罕见病超说明书用药不良反应主动监测与预警中仍面临严峻挑战。（1）数据稀缺性是核心挑战之一。尽管AI技术为应对数据稀疏困境提供了新工具，但在实际模型构建与应用中，数据本身的稀缺性与异质性，仍是其面临的首要技术挑战。罕见病患者群体规模小且分布分散，导致相关临床数据极为有限[33]。这种数据匮乏直接影响了AI模型训练的有效性。（2）数据的量和异质性构成了显著障碍[9]。即使有数据，EHR中的非结构化数据包含了大量重要的社会健康决定因素，将其结构化并用于AI分析仍具有挑战[14]。此外，整合自不同医疗机构的数据在结构、编码及所服务人群特征上存在固有差异[34]。若不加校正，这种“数据异质性”会在模型训练中被学习，进而导致“算法偏倚”，即模型预测性能在不同亚群（如来自不同医院、不同地域的患者）中出现不公正的差异，损害其泛化能力和公平性。（3）模型泛化能力和可解释性是AI在罕见病领域应用的关键难题。而XAI有助于克服这一问题，向可信赖的AI迈进，但这在小样本数据环境下仍面临巨大困难[19]。（4）伦理与隐私问题不容忽视。患者数据的高度敏感性要求在收集、存储和分析过程中严格遵守数据保护和伦理标准。特别是在罕见病领域，患者数量有限，数据泄露或滥用可能对个体造成更大的影响[35]。为实现

多中心数据协作以克服数据稀缺，必须在技术层面构建隐私增强的计算框架（如联邦学习、同态加密等），使数据分析能够在保护原始数据隐私的前提下进行[36]。同时，需应对此类技术自身存在的隐私-效用权衡等新挑战，构建涵盖技术、管理与伦理的多层防御体系[36]。(5)监管与合规性也是一大挑战。目前全球范围内缺乏针对AI用于药物警戒的明确法规和审评标准，关于AI模型“超说明书使用”的监管框架仍在完善中，即便AI模型在批准的用途之外使用，也必须进行人工审阅和验证以确保患者安全[37]。

4 展望

需要指出的是，与数据相对丰富的非罕见病超说明书用药监测不同，罕见病场景的核心矛盾在于数据的极端稀疏与异质性，这决定了AI应用的重点必须从常规的“海量数据中提纯信号”转向“从有限数据中挖掘与预警”，从而对算法的灵敏度与小样本学习能力提出了更高要求。尽管面临诸多挑战，AI在罕见病超说明书用药不良反应主动监测与预警中仍具有巨大的潜力。AI在重塑罕见病超说明书用药的安全性监测范式，使其从滞后、被动、模糊走向实时、主动、精准。通过隐匿患者挖掘、多源信号检测、个性化决策支持和全流程系统构建，AI为破解罕见病药物警戒的核心困境提供了切实可行的技术路径。当前的应用已展现出显著潜力，但模型的可靠性、可解释性以及数据、伦理、监管等系统性挑战的解决，是决定其能否从示范性应用走向广泛临床实践的关键。未来，需要医学、药学、数据科学、政策制定者等多方协同，共同推动以患者为中心、数据驱动、智能预警的罕见病用药安全新生态系统的建成。

参考文献

[1] Roessler H I, Knoers N V A M, van Haelst M M, *et al*. Drug repurposing for rare diseases[J]. Trends Pharmacol Sci, 2021, 42(4): 255-267.

[2] Zuo W, Sun Y J, Liu R J, *et al*. Management guideline for the off-label use of medicine in China (2021)[J]. Expert Rev Clin Pharmacol, 2022, 15(10): 1253-1268.

[3] Mazzucato M, Minichiello C, Vianello A, *et al*. Real-world use of orphan medicinal products (OMPs) in rare disease (RD) patients: a population-based registry study[J]. Front Pharmacol, 2022, 13: 940010.

[4] Fung A, Yue X M, Wigle P R, *et al*. Off-label medication use in rare pediatric diseases in the United States[J]. Intractable Rare Dis Res, 2021, 10(4): 238-245.

[5] Dooms M, Cassiman D, Simoens S. Off-label use of orphan medicinal products: a Belgian qualitative study[J]. Orphanet J Rare Dis, 2016, 11(1): 144.

[6] 中国药学会罕见病药物专委会. 医疗机构罕见病药学服务专家共识(2025)[J]. 协和医学杂志, 2025, 16(3): 659-672.

[7] 陈超阳, 段玥涵, 卓琳, 等. 罕见病超说明书用药专家共识计划书[J]. 药物流行病学杂志, 2025, 34(9): 1066-1073.

[8] 左玮, 刘容吉, 孙雅佳, 等. 《中国超药品说明书用药管理指南(2021)》推荐意见及要点解读[J]. 协和医学杂志, 2023, 14(1): 86-93.

[9] Sardella M, Belcher G. Pharmacovigilance of medicines for rare and ultrarare diseases[J]. Ther Adv Drug Saf, 2018, 9(11): 631-638.

[10] Phillips C, Parkinson A, Namsrai T, *et al*. Time to diagnosis for a rare disease: managing medical uncertainty. A qualitative study[J]. Orphanet J Rare Dis, 2024, 19(1): 297.

[11] Isono M, Kokado M, Kato K. Why does it take so long for rare disease patients to get an accurate diagnosis: a qualitative investigation of patient experiences of hereditary angioedema[J]. PLoS One, 2022, 17(3): e0265847.

[12] Crowe A L, McKnight A J, McAneney H. Communication needs for individuals with rare diseases within and around the healthcare system of northern Ireland[J]. Front Public Health, 2019, 7: 236.

[13] Tang S J, Yuan K, Chen L N. Molecular biomarkers, network biomarkers, and dynamic network biomarkers for diagnosis and prediction of rare diseases[J]. Fundam Res, 2022, 2(6): 894-902.

[14] Han S F, Zhang R F, Shi L Y, *et al*. Classifying social determinants of health from unstructured electronic health records using deep learning-based natural language processing[J]. J Biomed Inform, 2022, 127: 103984.

[15] Zolnoori M, Fung K W, Patrick T B, *et al*. A systematic approach for developing a corpus of patient reported adverse drug events: a case study for SSRI and SNRI medications[J]. J Biomed Inform, 2019, 90: 103091.

[16] Hurvitz N, Azmanov H, Kesler A, *et al*. Establishing a second-generation artificial intelligence-based system for improving diagnosis, treatment, and monitoring of patients with rare diseases[J]. Eur J Hum Genet, 2021, 29(10): 1485-1490.

[17] Spahn C, Toda N, Groat B, *et al*. Transforming pharmacovigilance with pharmacogenomics: toward personalized risk management[J]. Clin Pharmacol Ther, 2025, 118(6): 1286-1296.

[18] Elhaddad M, Hamam S. AI-driven clinical decision support systems: an ongoing pursuit of potential[J]. Cureus, 2024, 16(4): e57728.

[19] Markus A F, Kors J A, Rijnbeek P R. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies[J]. J Biomed Inform, 2021, 113: 103655.

[20] Johnson R, Stephens A V, Mester R, *et al*. Electronic health record signatures identify undiagnosed patients with common variable immunodeficiency disease [J]. Sci Transl Med, 2024, 16(745): eade4510.

[21] Jordan D M, Vy H M T, Do R. A deep learning transformer model predicts high rates of undiagnosed rare disease in large electronic health systems [J]. medRxiv Prepr Serv Health Sci, 2023. Doi: 10.1101/2023.12.21.23300393.

[22] Arku D, Yousef C, Abraham I. Changing paradigms in detecting rare adverse drug reactions: from disproportionality analysis, old and new, to machine learning [J]. Expert Opin Drug Saf, 2022, 21(10): 1235-1238.

[23] Gao Y, Zhang X, Sun Z Q, *et al*. Precision adverse drug reactions prediction with heterogeneous graph neural network [J]. Adv Sci (Weinh), 2024, 12(4): e2404671.

[24] Cai Z, Tang H, Gao S, et al. HyperADRs: a hierarchical hypergraph framework for drug-gene-ADR prediction [J]. Quantitative Methods, 2025.

[25] Li S, Zhang L C, Wang L Y, *et al*. BiMPADR: a deep learning framework for predicting adverse drug reactions in new drugs [J]. Molecules, 2024, 29(8): 1784.

[26] Yadav M, Ahlawat P, Singh V. A graph-based transformer neural network for multi-label ADR prediction [J]. Arab J Sci Eng, 2025, 50(10): 7323-7341.

[27] Asfand-E-Yar M, Hashir Q, Shah A A, *et al*. Multimodal CNN-DDI: using multimodal CNN for drug to drug interaction associated events [J]. Sci Rep, 2024, 14(1): 4076.

[28] Zhou H J, Liu F L, Wu J G, *et al*. A collaborative large language model for drug analysis [J]. Nat Biomed Eng, 2026, 10(5): 870-881.

[29] Chen L, Wang W, Bai Z, *et al*. PharmaGPT: Domain-specific large language models for bio-pharmaceutical and chemistry [J]. arXiv, 2024.

[30] Dou Y T, Huang Y W, Zhao X J, *et al*. ShennongMGS: An LLM-based Chinese medication guidance system [C]//ACM Transactions on Management Information Systems. ACM, 2025: 1-14.

[31] Li W Q, Yu L N, Wu M, *et al*. DoctorGPT: A large language model with Chinese medical question-answering capabilities [C]//2023 International Conference on High Performance Big Data and Intelligent Systems (HDIS). December 6-8, 2023, Macau, China. IEEE, 2023: 186-193.

[32] 徐鹏航, 宋晨. 全国首个！罕见病有了AI大模型 [N]. 新华每日电讯, 2025-02-21.

[33] Edvard Smith C I E, Bergman P, Hagey D W. Estimating the number of diseases – the concept of rare, ultra-rare, and hyper-rare [J]. iScience, 2022, 25(8): 104698.

[34] 吴敏敏, 王伟炳. 多源异构电子健康记录在观察性研究中的偏倚和控制 [J]. 复旦学报(医学版), 2025, 52(5): 765-770.

[35] Iyer A A, Barzilay J R, Tabor H K. Patient and family social media use surrounding a novel treatment for a rare genetic disease: a qualitative interview study [J]. Genet Med, 2020, 22(11): 1830-1837.

[36] Khalid N, Qayyum A, Bilal M, *et al*. Privacy-preserving artificial intelligence in healthcare: techniques and applications [J]. Comput Biol Med, 2023, 158: 106848.

[37] Krishnamoorthy M, Sjoding M W, Wiens J. Off-label use of artificial intelligence models in healthcare [J]. Nat Med, 2024, 30(6): 1525-1527.

（收稿日期：2026-01-21；在线出版日期：2026-05-26）