·综述·

大语言模型助力药物流行病学研究

司书成[1,2#]，吴柳柳[1,2#]，王聪慧[3]，杨子铭[4]，杜　建[5]，王胜锋[2,6]，詹思延[1,2,6,7]

1. 北京大学第三医院临床流行病学研究中心（北京 100191）
2. 重大疾病流行病学教育部重点实验室（北京大学）（北京 100191）
3. 内蒙古自治区药物警戒中心药品监测与评价科（呼和浩特 010010）
4. 北京大学第一医院人力资源处（北京 100034）
5. 北京大学健康医疗大数据国家研究院（北京 100191）
6. 北京大学公共卫生学院流行病与卫生统计学系（北京 100191）
7. 北京大学人工智能研究院智慧公众健康研究中心（北京 100871）

【摘要】人工智能（AI）的出现无论是在研究数量还是研究范式上都对医学研究与实践产生了重大影响，成为促进药物流行病学发展的重要工具。然而，传统 AI 在助力药物流行病学研究的同时也面临着数据处理复杂、药物暴露与潜在结局识别困难、研究设计与实施耗时耗力等诸多挑战。以大语言模型（LLMs）为代表的生成式 AI 的快速发展，在提升研究效率、转变研究范式以及促进知识发现等方面展现出独特的潜力。LLMs 具备自然语言理解与生成能力，通过深度挖掘多维度数据资源，能够快速、准确地提取、分析、总结和呈现所需信息，不仅可助力药物发现、药物再利用、药物警戒等药物流行病学任务，还能为研究方案设计、数据分析、结果解读和论文发表等全过程提供有力支持。在 LLMs 的推动下，药物流行病学研究正逐渐迈入依托大数据与自动化分析的新阶段。然而 LLMs 在应用中也存在数据偏差、结果“幻觉”、伦理与法律监管空缺等问题。通过加强跨学科合作、建立规范化评价体系、完善伦理与法规指导、提高数据质量，加强从业者培训和能力建设、以及推进人机协同的研究模式等，有望充分释放 LLMs 在药物流行病学中的应用潜力，为药品监管与公共卫生决策提供更为科学、快速和高效的技术支撑。

【关键词】人工智能；大语言模型；药物流行病学；药物发现；药物再利用；药物警戒

【中图分类号】R 181.3+5　　**【文献标识码】**A

Large language models empowering pharmacoepidemiology research

SI Shucheng[1,2#], WU Liuliu[1,2#], WANG Conghui[3], YANG Ziming[4], DU Jian[5], WANG Shengfeng[2,6], ZHAN Siyan[1,2,6,7]

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

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

3. Drug Monitoring and Evaluation Division, Inner Mongolia Pharmacovigilance Center, Hohhot 010010, China

DOI: 10.12173/j.issn.1005-0698.202504033

共同第一作者

基金项目：国家自然科学基金重点基金项目（82330107）；国家自然科学基金青年基金项目（82404370）

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

4. Department of Human Resources, Peking University First Hospital, Beijing 100034, China
5. National Institute of Health Data Science, Beijing 100191, China
6. Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China
7. Center for Intelligent Public Health, Institute for Artificial Intelligence, Peking University, Beijing 100871, China
#Co-first authors: SI Shucheng and WU Liuliu
Corresponding author: ZHAN Siyan, Email: siyan-zhan@bjmu.edu.cn

【Abstract】The emergence of artificial intelligence (AI) has had a significant impact on medical research and practice, both in terms of the number of studies and research paradigms, and has become an important tool for the development of pharmacoepidemiology. However, traditional AI has faced many challenges, while facilitating pharmacoepidemiology research, such as complex data processing, difficulty in identifying drug exposures and potential outcomes, and time-consuming and laborious study design and implementation. The rapid development of generative AI, represented by large language models (LLMs), has demonstrated a unique potential to enhance research efficiency, shift research paradigms, and facilitate knowledge discovery. LLMs are equipped with natural language understanding and generation capabilities. Through deep mining of multi-dimensional data resources, LLMs can quickly and accurately extract, analyze, summarize, and present the required information, which can not only help drug discovery, drug repurposing, pharmacovigilance and other pharmacoepidemiological tasks, but also provide powerful support for the whole process of research protocol design, data analysis, result interpretation and paper publication. Driven by LLMs, pharmacoepidemiology research is gradually moving into a new stage based on big data and automated analysis. Of course, LLMs also have problems of data bias, "illusion" of results, and ethical and legal regulation. By strengthening interdisciplinary cooperation, establishing a standardized evaluation system, improving ethical and regulatory guidance, enhancing data quality, strengthening practitioner training and capacity building, and promoting human-machine collaborative research modes, it is expected that the potential of LLMs in pharmacoepidemiology will be fully released, and it will provide a more scientific, rapid, and efficient technological support for drug regulation and public health decision-making.

【Keywords】Artificial intelligence; Large language models; Pharmacoepidemiology; Drug discovery; Drug repurposing; Pharmacovigilance

人工智能（artificial intelligence，AI）从广义上讲是指数字计算机或计算机控制的机器人执行通常与智能生物相关任务的能力[1]。近十余年来，AI技术得到迅猛发展，特别是从基于传统规则基础的AI技术到以大语言模型（large language models，LLMs）为代表的生成式AI的技术转变，如ChatGPT（chat generative pre-trained transformer）和DeepSeek等AI模型正在深刻重塑各行各业的发展模式。在医学领域，AI的应用尤为广泛。通过高效处理和分析海量的生物医学数据，AI不仅可以自动学习并挖掘数据中的复杂规律，还能通过训练深度模型、LLMs等实现精准的预测、推断和解答。AI技术为疾病诊断、治疗决策、医疗管理和研究方法等提供了创新性的工具，也推动了医疗领域的全流程优化。

在药物流行病学研究中，AI的引入展现出重要的应用价值。药物流行病学作为研究人群中药物使用、疗效和安全性的学科，通常需要处理大量复杂、多维的患者药物效果数据。AI通过其强大的数据处理和建模能力，能够深入挖掘多源异构数据资源，可支持药物发现、药物再利用、上市后药品不良反应（adverse drug reaction，ADR）监测等多个方面的研究与应用[2–4]。尤其是LLMs凭借其卓越的自然语言理解与生成能力，为药物流行病学研究的全流程优化带来了前所未有的机遇[5–6]。本文旨在总结AI，特别是LLMs在医学研究中的角色，重点探索其在药物流行病学研究中的应用价值和潜力，并探讨这一领域可能面临的挑战和应对方式。

1 AI发展与医学研究

AI 在过去数十年间从最初的概念，发展到机器学习，再到深度学习，进而演变为现阶段的 LLMs，这些技术的发展对医学研究与实践，特别是研究数量和研究范式产生了重大影响。据统计[7]，2000—2022 年全球 AI 生命科学出版物数量年增长率达 20%，且主要集中在美国和中国（合计约占全球总量的 44%）。2010—2020 年，关于机器学习和深度学习的医学 AI 研究数量也呈指数级增长，仅 2019 年的累计数量就接近 2017 年的 3 倍[8]。此外，随着 AI 在科学研究中的广泛应用，其发展也跨越了传统上孤立的学科领域[9]，并通过增强和加速科学研究中的每一个环节，引发了研究范式的转变[10]。探索 AI 驱动的科学研究（AI for science，AI4S）新范式，对提高医学研究效率和质量具有重要意义。

随着 AI4S 的推进，类似于“机器人科学家”的生成式 AI 取得了快速发展。例如，计算生物学家 Greene 与 Pividori 曾测试使用 ChatGPT 协助改进研究论文[11]，结果表明，AI 模型可以捕捉学术文本中的概念，并生成高质量的修订，减少了撰写和修改学术文稿所需的时间。ChatGPT 等 LLMs 能够在大量文本数据集上训练，使其具备模仿人类回答的能力，在信息提取、命名实体识别以及情感分析等方面发挥重要作用[12]，并在多个医学研究领域均表现出优越的潜力。在药物流行病学方面，LLMs 能够从海量医学文献、病历记录、临床指南及研究论文等文本数据中自动识别并提取结构化信息，包括疾病症状、治疗方法、药物剂量、患者预后等。LLMs 还可以用于辅助诊断、问答系统和文档管理（如报告生成、ICD 编码管理等），例如将 LLMs 融入医学影像的计算机辅助诊断系统中[13]以及应用 GPT 等技术构建医疗聊天机器人等[14]。施呈昊等[15]总结了 LLMs 在临床实践中的应用方向与优势，包括诊断与治疗建议、数据分析与处理、风险预测与筛查、文书与行政工作处理、患者交流与健康宣教等。人们普遍预期，“机器人科学家”范式将在可预见的未来成为现实[16]。

2 传统AI在药物流行病学中的应用

以机器学习和深度学习等为代表的传统 AI 技术已经在药物流行病学研究中获得了广泛应用，涵盖多个研究方向。在基于医疗保健数据的相关研究中，AI 最常见的应用领域是健康诊断或治疗结果预测（研究数量占比约 32%），以及 ADR、药物相互作用（drug–drug interaction，DDI）和药物级联反应等（占比约 15%）[17]。在药物流行病学研究中，AI 的主要目的包括根据患者特征预测所需剂量、预测药物治疗后的临床反应、预测 ADR 发生或严重程度、预测倾向性评分、确定药物无效风险更高的亚群、预测药物消耗或预测药物引起的住院时间等，其中最常被提及的药物包括胰岛素、二甲双胍、维生素、对乙酰氨基酚和肝素等[17]。在 AI 方法应用方面，Maurizio 等[18]通过范围综述发现，随机森林、人工神经网络和支持向量机等 AI 技术是药物流行病学研究中最常用的知识发现技术。谈志远等[19]的研究也表明，AI 应用主要集中在 ADR 辅助诊断、信号预测、提取和报告偏倚分析等方面，主要采用决策树模型、反向机器学习、人工神经网络、关联规则和数据挖掘等技术。将 AI 技术与传统药物流行病学研究方法进行比较，50% 的 AI 表现优于传统药物流行病学方法，特别是随机森林和人工神经网络等方法表现突出[20]。Li 等[21]探索了机器学习和深度学习在从临床基准数据集中提取药品不良事件（adverse drug event，ADE）的有效性，表明对于命名实体识别任务，深度学习模型的性能优于机器学习模型；而对于关系分类任务，梯度提升、多层感知器和随机森林模型表现出色。这些研究表明，AI 技术已经成为促进药物流行病学研究的重要工具。

3 LLMs助力药物流行病学研究

ChatGPT、DeepSeek 等 AI 模型的核心是 LLMs，它们可利用来自互联网的大量文本和图像材料语料库进行训练，包括但不限于生物医学文献和代码。Wang 等[22]总结了 2023 年 ChatGPT 在生物信息学和生物医学领域的最新应用进展，其中包括药物发现等药物流行病学相关主题。大量研究表明，LLMs 在药物发现（drug discovery）、药物再利用（drug repurposing）、以及药物警戒（pharmacovigilance）等方面已经表现出巨大的潜力。

3.1 LLMs加速药物发现

药物发现是一个复杂且易失败的过程，需要大量的时间、精力和资金投入。由于 AI 在分析大型数据集和复杂生物网络方面的优势，在药物发现中发挥着越来越大的作用。最近一些研究[17-18]表明，LLMs 在处理药物发现相关任务方面表现出色，这些任务涉及了药物靶点识别、结构预测、药物－靶点相互作用估计、药物－靶点结合亲和力估计、从头药物设计、药物毒性预测、吸收、分布、代谢、排泄、毒性的估计以及 DDI 估计等过程。例如，截至 2024 年，一些利用 AI 识别的靶点正在通过实验进行验证，由 AI 衍生的药物也正在进入临床试验[23-25]。此外，GPT-3.5 在识别药物与疾病关联方面展现出较为可观的准确性[26]。GPT 系列模型在生成分子描述（如分子标题）等文本化学任务中同样表现优异[27]。LLMs 在预测和阐明 DDI 方面也取得了重要进展[28]。Al-Ashwal 等[29]研究指出，GPT 模型在 DDI 预测中的准确率为 50%~60%，结合 BING 搜索后准确率可进一步提升 20%~30%。但也有研究[30]指出，在使用 DDI 语料库进行评估时，ChatGPT 的表现一般，且在患者教育和识别 ADR 方面也表现出一定的局限性[31]。这表明，尽管 LLMs 在药物发现过程中提供了有价值的支持，其在应对复杂挑战方面的能力仍显不足，需要依赖人类专家的监督以提升研究效率和准确性。

既往研究[22]表明，在药物发现中使用 LLMs 提升研究效率主要取决于以下 3 个重要方面：人机协作、情境学习和指令微调。人机协作是指在应用 LLMs 进行药物开发的同时利用人类专家的专业知识来优化结果。例如，在开发抗可卡因成瘾药物的案例研究中，生成的响应经专家的严格评估后，被反馈给 LLMs 以进一步改进，通过这种迭代的人机协作方法，成功识别出 15 个有前景的多靶点抗可卡因成瘾候选药物[32]。此外，诸如 ChatDrug[33]、DrugChat[34] 和 DrugAssist[35] 等工具，利用交互式界面根据专家反馈进行迭代改进，也进一步提升了模型的准确性和实用性。情境学习是指通过利用领域知识库中的示例来增强 LLMs 的响应，使其回答更具相关性和准确性[36]。研究[32-33]表明，包含相似情境的示例能够显著提升模型的表现，且随着示例数量的增加，模型的性能逐步提升。除此之外，利用相关信息丰富语境（如疾病背景和同义词），也能有效提升 ChatGPT 在药物－疾病关联识别中的响应准确性[26]。指令微调是在特定数据集或特定任务上进一步训练模型以提高其在该领域的性能，可分为任务调优和指令调优。任务调优是针对特定任务微调预训练模型以提高其在该任务上的性能。例如，ChatMol 通过实验性质数据和分子空间知识进行调优，提升了其在描述和编辑目标分子方面的能力[37]。指令调优则通过使用“指令－输出对”进行 LLMs 训练，从而提高其在不同任务中的适用性[38]。例如，DrugAssist 通过指令调优后，在同时优化多个分子性质时表现较优；DrugChat 通过指令调优，能够有效回答关于图表示药物化合物的开放性问题[34]。任务调优与指令调优相结合的 ChemDFM[39] 在分子性质预测和反应预测等任务中表现优异。随着这些优化方案的逐步普及，有望显著推动 AI 在药物发现领域的发展。

3.2 LLMs促进药物再利用

除了药物发现之外，药物再利用是药物发现领域一种极具潜力的策略，因其高效性和成本效益而备受关注。通过挖掘已批准药物以前未被发现的治疗特性，LLMs 减少了与药物发现相关的时间和成本。近年来，LLMs 因其庞大的参数规模和广泛的训练数据，展现出了在药物再利用任务中的显著优势。LLMs 可用来辅助其他算法来提升药物再利用的能力。Picard 等[4]利用 GPT-4 来分析临床试验数据并识别经实验验证的阴性结果，由于机器学习算法需要具有标记的训练集将药物分类为阳性药物（表现出已知的治疗益处）或阴性药物（表明缺乏此类益处或存在不可接受的毒性），通过这些真阴性数据集训练机器学习算法，可以进一步提升药物再利用的准确性。LLM 不仅可以辅助其他算法，也能直接用于药物再利用任务。如 Wei 等[40]基于 LLMs 提出了一个结合多源提示技术的创新药物再利用方法——DrugReAlign 框架，旨在高效挖掘现有药物的潜力。该框架充分利用 LLMs 对靶点和药物进行综合分析的能力，从广泛的知识库中获取靶点和药物的通用知识，并结合多个数据库的靶点摘要和靶点－药物空间交互数据作为提示信息，显著提升了 LLMs 在药物再利用任务中的表现。为了应对 LLMs 在单一或有限数据源提示下

可能产生的“幻觉”和错误信息风险，该框架还引入了多源提示构建策略，整合了靶点结构信息和已知的靶点与小分子之间的空间交互数据等，避免了模型过度依赖语言层面的知识推测，确保其输出具有科学严谨性。实验验证结果表明，DrugReAlign 框架在靶点覆盖率和 Top 1 推荐成功率方面表现出较高的准确性，同时兼具一定的新颖性（对未知药物 – 靶点交互的探索能力）。LLMs 提升机器学习性能以及 DrugReAlign 的开发验证了 LLMs 在药物再利用领域的应用潜力，有望开启药物发现领域的新范式。在未来，随着 LLMs 模型能力和生物学数据库的进一步完善，将在药物再利用的高通量筛选、未知药物 – 靶点关系预测以及个性化治疗方案设计中展现更广阔的应用前景。

3.3 LLMs助力上市后药物监测

在上市后药物监测方面，LLMs 能够从电子健康记录等文本数据中高效识别潜在的 ADE 信号，从而加速流行病学调查，揭示医疗产品暴露与 ADE 之间的因果关系 [41]。相关综述研究 [42] 表明，生成式 LLMs 在减轻药物相关伤害方面主要面向 DDI 识别和预测、临床决策支持和药物警戒等关键应用，虽然不同的 LLMs 性能和效用各不相同，但它们在这些方面都显示出可观的前景。一些初步研究已经探索了 ChatGPT 识别 ADR[31]、药物警戒信号检测[43] 等。例如，对于信号检测和 ADE 分类，使用 LLMs 进行训练数据增强和降维，与现有模型相比，在药品上市后监测期间为自动评估药物风险类别而开发的分类模型显示出高达 97.9% 的准确率 [42]。在针对药物安全性的文献筛选任务方面，LLMs 也具有突出的表现。例如，药物性肝损伤（drug-induced liver injury，DILI）是药物安全性的重大问题之一，传统的文献检索方法常常依赖于关键词搜索，但由于 DDI 的复杂性，简单的搜索无法全面覆盖所有相关文献。LLMs 在此类任务中展现出了巨大的潜力。Ma 等 [44] 基于 LLaMA-2 开发了专门用于 DILI 分析的 LLMs，通过利用来自 CAMDA 2022 AI 挑战赛的 14 203 篇训练文献，成功实现了高效的 DILI 文献识别。与传统模型（如 BERT、GPT）相比，LLaMA-2 在三折交叉验证中显示了 97.19% 的准确率和 0.99 的受试者工作特征曲线下面积，表现出了优越的分类性能。这意味着 LLMs 能够大大加速药物安全性文献的筛选过程，并为药品监管提供更精确的支持。另外，LLMs 在社交媒体数据中的应用也为 ADE 的检测提供了新的方向。Deng 等 [45] 开发了多个 LLMs，对来自社交媒体平台 X（前 Twitter）的数据进行 ADE 分类，结果显示 LLMs 在社交媒体数据中对 ADE 的识别具有很大潜力，为药品上市后安全监测提供重要补充。美国食品药品管理局（Food and Drug Administration，FDA）对 AI 在药品上市后安全监测和评估中的应用也进行了总结，包括检测和评估来自医学文献和社交媒体的 ADE 关联、确定 ADE 以便提交个案安全报告、通过评估最低报告要求来确定案例的有效性并根据预期确定案例的优先级、将个案的安全性报告编码到结构化的医学词典中、对药物与 ADE 之间因果关系的概率进行分类、促进汇总的个案安全报告等。然而，LLMs 的推理过程可能会产生非事实的陈述，从而产生不可预见的后果，FDA 尚未授权 LLMs，需要进一步的 LLMs 评估工具（如 DeepEval 和 MLflow）和监管创新，以实现对这些信息来源的分析并整合到临床决策中。因此，为了确保 LLMs 能够公平、公正地用于药物监测，必须采取额外的措施来避免错误信息的影响，并确保其在实际应用中的安全性和有效性 [41]。

4 LLMs在药物流行病学研究中的挑战

尽管 LLMs 在药物流行病学研究中取得了巨大进步，但仍面临着多维度、系统性的挑战，这些挑战既包括技术和数据层面的局限性（如数据选择偏移、模型泛化能力、数据规范化、答案可靠性、用户接受度与依赖性等），也涉及伦理规范与法律监管等问题（如隐私与数据安全问题）。

从技术和数据角度来看，LLMs 模型的性能表现高度依赖于训练数据的质量和代表性，而现实中的药物流行病学数据往往存在显著的偏倚问题，特别是在真实世界研究中使用的真实世界数据表现尤为突出。例如，当前大多数训练数据集以欧美人群为主要来源，这可能导致模型在应用于其他种族群体时预测准确性下降，这一问题在肿瘤药物研发领域表现得尤为明显，可能造成患者选择偏差和疗效评估失真 [43-44]。虽然通过生成对抗网络等技术可以创建合成数据来弥补样本不

足，但过度依赖合成数据又可能引发模型过拟合和泛化能力下降的新问题[46]。相比较于“偏倚”，对研究准确性影响更为严重的情况是 LLMs 在产生结果时有时会出现“幻觉”，从而做出非事实的陈述且真假难辨。面对这一挑战，使用者需要从提高数据质量、优化模型架构、规范训练策略、加强实时监控和反馈等一系列途径来减少或避免这一现象。与此同时，药物流行病学研究需要整合不同来源的结构化和非结构化数据，包括实验室检测指标、电子病历文本、影像学资料等多模态数据，这些数据在采集标准、存储格式和质量控制方面存在显著差异，给数据整合与分析带来巨大挑战。未来亟需开发跨模态、跨平台的数据标准和整合算法，提升 LLMs 在药物流行病学研究过程中的实用价值。此外，由于计算资源、AI 专业知识等局限性限制了 LLMs 在药物流行病学领域的快速普及。为应对这些挑战，科研工作者和管理者应加强跨学科合作，整合药物流行病学领域相关资源、加强使用者的培训和能力建设，激发 LLMs 的应用潜力。当然，LLMs 虽然优势巨大，但过度依赖也可能带来一些潜在不良后果（如削弱研究人员的批判性思维能力、放大现有数据偏差）等，研究者在使用时仍然需要充分地权衡。

在伦理规范与法律监管层面，LLMs 的应用也面临着复杂的困境[47]。LLMs 在药物开发过程中往往涉及大量的个人敏感信息，如电子病历、模拟患者群体或生成蛋白质 – 药物相互作用数据。尽管目前世界各国尚未建立统一、系统的伦理法规，但仍需遵循当前严格的数据采集与使用规则，这也与 AI 技术发展之间产生了明显的张力：一方面 AI 模型的训练需要大量高质量数据，另一方面严格的隐私保护规定限制了数据的可获得性。监管适应性是另一个突出挑战，传统药物审批要求提供药物安全性和有效性的大量临床证据，而 LLMs 的输出具有典型的概率性特征（迭代、预测性等），这种差异导致监管机构在评估 AI 驱动的研究成果时面临诸多不确定性，可能会减缓药物审批流程。此外，算法公平性和责任归属问题也日益凸显，商业利益可能导致的治疗推荐偏倚、AI 决策过程缺乏透明度、以及 AI 应用对传统医患关系的重构等。总之，尽管将 AI 应用于早期药物发现和开发仍然存在挑战，但随着各种新的伦理规范和法律条文持续发布有望逐渐减轻这些限制。

5 LLMs在药物流行病学研究中的应用探索

针对上述 LLMs 在药物流行学研究中的潜在应用，本文以 DeepSeek 为例，测试了 LLMs 在相关研究各个环节中的辅助作用。

首先，在科研选题和研究设计方面，LLMs 可用于辅助寻找科学问题，并提供相应的研究方案和资料。例如，通过询问目前药物流行病学的研究热点，LLMs 能够清晰地给出一些该领域目前的主要研究方向供研究者参考（图 1–A）。在研究资料方面，LLMs 可协助制定基线调查问卷、随访问卷、知情同意书、随访计划、现场调查方案以及统计分析方案等相关研究材料。如图 1–B 中以糖尿病为例提供的基线调查和随访问卷，其条目和内容均较为符合糖尿病相关研究的特点。在确定具体研究选题后，LLMs 还可以辅助设计研究方案，例如，通过提供新型降糖药的药物再利用相关研究背景，LLMs 能够为研究设计出一个初步的方案框架（图 1–C）。当然，在不提供研究背景的前提下，LLMs 也可以基于前期的训练和网络资源，通过逐步提问的方式形成一份相对完整的研究设计方案。

在研究的分析阶段，LLMs 能够根据具体要求和研究类型提供所需的相应分析代码。例如，根据需要向 LLMs 询问队列研究的分析代码，LLMs 可以提供从数据处理、统计描述、统计分析、假设检验以及结果可视化等全流程的分析代码案例（图 2–A）。此外，LLMs 可以根据要求生成相应图表的分析代码，如提供一篇已发表的论文，要求其提供文中某个感兴趣的图片的分析代码，LLMs 可以直接生成模拟数据及绘图代码，多数情况下可以直接运行或经过微调后可顺利运行（图2–B）。在论文撰写和投稿阶段，LLMs 依然能提供有效支持，包括摘要提取、投稿信撰写、语言润色、语法纠正以及投稿选刊等任务。

这些探索表明，目前 LLMs 已经可以在药物流行病学研究的各个环节提供有效的辅助，显著提高研究的效率，并帮助研究人员在各个阶段顺利完成相关任务。尽管如此，测试中发现 LLMs 的反馈结果在不同的用户、不同时间、不同询问方式以及多次询问下均会有所不同，部分反馈可

A 大语言模型辅助寻找科学问题

序号	研究热点	意义/应用
1	抗生素耐药性与合理用药	优化抗生素使用策略，减少耐药性
2	真实世界数据（RWD）与药物安全	提升上市后药物监测效率，指导临床决策
3	老药新用与药物再评价	拓展药物适应症，降低研发成本
4	疫苗与抗生素管理协同	通过预防感染间接控制耐药性
5	新型抗病毒药物开发	应对病毒耐药性挑战
6	人畜共患病监测	预警新发传染病风险
7	药物不良反应信号挖掘	早期识别药物风险
8	个体化用药与精准医学	推动个性化治疗方案
9	药物经济学与卫生政策	优化医疗资源分配
10	多学科方法学创新	提升研究效率与准确性

B 大语言模型辅助研究资料——制定糖尿病队列研究的基线调查和随访问卷

糖尿病队列研究基线调查问卷

编号：_____ 调查日期：___年___月___日

A. 基本信息

1. 姓名：_____ 性别：□男 □女 □其他
2. 出生日期：年___月___日 联系电话：_____
3. 职业：_____ 婚姻状况：□未婚 □已婚 □离异 □丧偶

B. 糖尿病相关病史

4. 是否确诊糖尿病：□否 □是（类型：□1型 □2型 □妊娠糖尿病 □其他：_____）
5. 确诊时间：年___月 确诊医院：_____
6. 当前治疗方案（可多选）：
   □饮食控制 □口服降糖药（名称：） □胰岛素（类型：） □其他：_____

C. 健康状况与合并症

7. 其他疾病（可多选）：
   □高血压 □高血脂 □冠心病 □慢性肾病 □视网膜病变 □其他：_____
8. 家族糖尿病史：
   □无 □有（亲属关系：父母□ 兄弟姐妹□ 子女□ 其他：_____）

D. 生活方式

9. 吸烟史：
   □从不吸烟 □已戒烟（___年） □当前吸烟（___支/天，烟龄___年）
10. 饮酒习惯：
    □从不饮酒 □偶尔（<1次/周） □经常（1-3次/周） □每日饮酒（___两/天）
11. 运动情况：
    □几乎不运动 □1-2次/周 □3-5次/周 □每天运动（每次___分钟）

E. 体格检查（由医护人员填写）

| 身高：cm | 体重：kg | BMI：|
| 血压：/___mmHg | 空腹血糖：mmol/L | HbA1c：% |

受试者签名：_____ 调查员签名：_____ 备注：_____

糖尿病队列研究随访问卷

随访编号：_____ 患者基线ID：_____ 随访日期：___年___月___日

A. 基本信息确认

1. 姓名：_____ 性别：□男 □女
2. 出生日期：___年___月___日 本次随访类型：□3个月 □6个月 □年度

B. 疾病进展与治疗调整

3. 血糖控制情况：
   □达标（空腹<7.0mmol/L） □未达标 □不确定（未检测）
4. 治疗方案变更（对比基线）：
   □无变化 □新增药物（名称：） □停药（原因：） □胰岛素调整（剂量：_____）
5. 低血糖事件（过去3个月）：
   □无 □1-2次 □≥3次（最严重症状：_____）

C. 合并症与并发症

6. 新发合并症（可多选）：
   □高血压 □高血脂 □心血管事件 □糖尿病肾病 □视网膜病变 □其他：_____
7. 住院记录：
   □无 □有（原因：_____，住院天数：__）

D. 生活方式追踪

8. 吸烟/饮酒变化：
   - 吸烟：□戒烟（___个月） □减量（现___支/天） □无变化 □新增
   - 饮酒：□戒酒 □减量（现___两/周） □无变化 □新增
9. 运动依从性：
   □保持规律 □减少（现___次/周） □增加（现___次/周） □仍不运动

E. 本次随访检测结果

| 体重：___kg（较基线：□增↑kg □减↓kg） | BMI：|
| 血压：/___mmHg | 空腹血糖：mmol/L | HbA1c：% |

患者反馈：_____

随访员签名：_____ 下次随访时间：___年___月___日

C 大语言模型辅助研究方案设计——基于提供的研究背景，设计一项研究

新型降糖药物真实世界药物重定位研究方案

研究背景与意义

随着糖尿病全球流行趋势加剧，降糖药物的临床应用已超越单纯的血糖控制范畴。近年来，大量研究表明传统降糖药物如二甲双胍具有抗肿瘤、抗衰老等多重效应 [8] [9]，而GLP-1受体激动剂和SGLT-2抑制剂等新型降糖药也展现出对神经退行性疾病和心血管疾病的潜在治疗价值。然而，目前降糖药物重定位研究多集中于少数几种药物（如二甲双胍），缺乏对其他新型降糖药物（如GPR40激动剂、胰高血糖素受体拮抗剂等）系统性的重定位探索 [6] [10]。

药物重定位（Drug Repositioning）作为一种高效、低风险的药物开发策略，通过利用已有药物的安全性数据和临床实践经验，可大幅缩短新适应症的开发周期和成本 [1] [7]。特别是在大数据时代背景下，真实世界数据（RWD）与多组学分析技术的结合，为降糖药物的系统性重定位研究提供了前所未有的机遇 [3] [7]。

研究目的

本研究旨在基于真实世界医疗大数据和多组学分析，系统探索新型降糖药物在非糖尿病适应症中的潜在治疗价值，重点聚焦以下目标：

1. **建立新型降糖药物-疾病关联网络**：利用计算生物学方法预测新型降糖药物（如GPR40激动剂、胰高血糖素受体拮抗剂等）的潜在新适应症
2. **验证重定位信号**：通过真实世界临床数据分析，验证计算预测的药物-疾病关联
3. **探索作用机制**：结合基因组学、蛋白质组学等多组学数据，阐明降糖药物在新适应症中的潜在作用机制
4. **评估临床转化价值**：对具有前景的药物-适应症组合进行成本-效果分析和临床转化路径设计

研究方法

1. 计算预测阶段：药物-疾病关联网络构建

1.1 数据收集与处理

- **药物数据**：收集包括LY-2922470（GPR40激动剂） [10]、ISIS 449884（GCGR拮抗剂） [6]等新型降糖药物的化学结构、靶点信息、已知适应症等数据
- **疾病数据**：从OMIM、DisGeNET等数据库获取癌症、神经退行性疾病、心血管疾病等相关疾病的基因关联数据
- **生物网络数据**：整合蛋白质-蛋白质相互作用网络、代谢通路等数据

1.2 计算方法

采用**自适应图卷积网络（AdaDR）** [1] [9]进行药物-疾病关联预测，该方法通过深度集成药物和疾病的节点特征与拓扑结构，利用注意力机制学习嵌入的自适应重要性权重，显著提高了预测准确性。具体步骤包括：

1. **特征提取**：从药物化学结构、靶点信息、副作用谱等提取多维特征
2. **网络构建**：建立药物-靶点-疾病多层异质网络
3. **嵌入学习**：使用AdaDR模型学习药物和疾病的低维向量表示
4. **关联预测**：计算药物与疾病向量间的相似度，预测潜在关联

2. 真实世界验证阶段

2.1 研究设计与数据来源

采用**回顾性队列研究**设计，利用以下真实世界数据源验证计算预测结果：

- **电子健康记录(EHR)**：从合作医疗机构获取至少5年的糖尿病患者诊疗数据
- **医保索赔数据**：分析降糖药物使用与各类疾病发病率的关联
- **生物样本库**：对部分患者进行基因组学、蛋白质组学等分析

2.2 统计分析方法

1. **暴露定义**：以新型降糖药物使用为暴露因素，设立用药组与对照组
2. **结局指标**：关注癌症发病率、神经退行性疾病进展、心血管事件等终点
3. **统计分析**：
   - 采用多变量Cox比例风险模型评估药物暴露与结局的关联
   - 使用倾向评分匹配(PSM)减少混杂偏倚
   - 进行亚组分析探索不同人群中的效应异质性

3. 机制探索阶段

3.1 多组学整合分析

结合以下组学数据阐明药物作用机制：

- **基因组学**：分析药物靶点基因多态性与治疗响应的关联
- **转录组学**：通过RNA-seq检测药物处理前后基因表达变化
- **蛋白质组学**：质谱分析药物对关键信号通路蛋白的影响
- **代谢组学**：监测药物对全身代谢网络的重编程作用

3.2 实验验证

对最具前景的预测结果进行体外和动物实验验证：

- **细胞模型**：在疾病相关细胞系中测试药物的生物学效应
- **动物模型**：使用疾病动物模型评估药物的治疗潜力
- **分子对接**：模拟药物与潜在新靶点的结合模式

4. 临床转化评估

对验证有效的药物-适应症组合进行：

- **成本-效果分析**：评估与传统治疗相比的经济学优势
- **安全性评价**：系统回顾药物在目标人群中的安全性数据
- **临床路径设计**：制定适应症扩展的临床开发计划

研究计划与时间表

表：研究实施动态表

阶段	主要内容	时间节点	交付成果
第1年	数据收集与计算预测	2025.06-2026.05	新型降糖药物-疾病关联网络
第2年	真实世界数据分析	2026.06-2027.05	验证报告(3-5个有前景的重定位信号)
第3年	机制研究与实验验证	2027.06-2028.05	作用机制阐明、2-3篇高水平论文
第4年	临床转化评估	2028.06-2029.05	临床开发方案、专利申请

预期成果与创新点

预期成果

1. 建立**新型降糖药物重定位知识库**，包含至少10种药物的潜在新适应症预测
2. 验证3-5个具有临床转化价值的药物-疾病关联
3. 阐明2-3种降糖药物在新适应症中的作用机制
4. 发表3-5篇高水平研究论文，申请1-2项专利
5. 推动1-2个降糖药物的适应症扩展进入临床研究阶段

创新点

1. **方法学创新**：首次将自适应图卷积网络(AdaDR)系统应用于降糖药物重定位研究 [1] [9]
2. **数据整合创新**：结合真实世界数据与多组学分析，提高预测可靠性
3. **应用创新**：聚焦于研究不足的新型降糖药物，填补知识空白
4. **转化创新**：建立从计算预测到临床验证的完整研究链条

伦理考量与数据安全

1. **伦理审批**：研究方案将通过机构伦理委员会审查(IRB批准号：待申请)
2. **数据匿名化**：所有患者数据将进行去标识化处理
3. **知情同意**：回顾性数据使用将豁免知情同意，前瞻性研究将获取参与者书面同意
4. **数据安全**：建立符合HIPAA/GDPR标准的数据管理系统

研究团队与协作

本研究将组建跨学科研究团队，包括：

- **计算生物学专家**：负责药物-疾病关联预测
- **临床流行病学家**：主导真实世界数据分析
- **分子生物学家**：进行机制研究
- **临床医师**：提供临床见解和患者数据
- **生物信息学家**：负责多组学数据整合分析

图1 探索大语言模型DeepSeek在研究设计中的应用测试

Figure 1. Exploring the application of the LLM DeepSeek in study design: a test case

注：A. DeepSeek提供的药物流行病学领域十大研究热点；B. DeepSeek根据提问制作的糖尿病队列研究基线调查和随访问卷；C. DeepSeek根据提供的一段研究背景生成的简要研究设计方案。

能存在“幻觉”，在可重复性和真实性方面仍具有一定的局限性，因而现阶段的 LLMs 大多数情况下更适用于研究的辅助。整体而言，在研究中使用 LLMs 需要在遵循相关学术规范的前提下进行科学使用，并进行必要的披露。

6 结语与展望

综合而言，LLMs 为医学研究，特别是药物流行病学研究，提供了前所未有的优势和机会，主要体现在几个方面：强大的大型复杂数据的处理能力、对多种数据资源的深度挖掘与利用、

A 大语言模型辅助统计分析——提供一份队列研究的分析代码

1. 加载必要包

```r
library(survival)    # 生存分析核心功能
library(survminer)   # 生存曲线绘制
library(tableone)    # 描述性统计表格
```

2. 数据导入与处理

```r
data <- read.csv("cohort_data.csv") %>%
  mutate(
    time = as.numeric(end_date - start_date)/365.25,
    status = as.numeric(status)
  ) %>%
  na.omit()
```

3. 描述性统计

```r
# 创建表格
table1 <- CreateTableOne(
  vars = c("age", "gender", "group", "time"),
  strata = "group",
  data = data
)

# 打印结果
print(table1, showAllLevels = TRUE)
```

4. 生存分析

```r
# 创建生存对象
surv_obj <- Surv(data$time, data$status)

# Kaplan-Meier分析
km_fit <- survfit(surv_obj ~ group, data = data)

# 绘制生存曲线
ggsurvplot(km_fit, data = data,
           pval = TRUE,
           risk.table = TRUE,
           xlab = "Time (years)")
```

5. Cox回归分析

```r
# 构建Cox模型
cox_model <- coxph(surv_obj ~ group + age + gender,
                   data = data)

# 查看模型结果
summary(cox_model)
```

6. 比例风险假设检验

```r
# Schoenfeld残差检验
ph_test <- cox.zph(cox_model)

# 打印检验结果
print(ph_test)

# 绘制残差图（可视化检验）
ggcoxzph(ph_test)
```

7. 结果可视化

```r
# 森林图
ggforest(cox_model, data = data)
```

8. 保存结果

```r
# 保存比例风险检验结果
capture.output(print(ph_test),
               file = "proportional_hazards_test.txt")

# 保存Cox模型结果
capture.output(summary(cox_model),
               file = "cox_model_results.txt")
```

B 大语言模型辅助统计分析——上传一篇感兴趣的研究论文全文，令其提供论文中某个图的分析代码

1. 准备环境和数据

```r
# 加载必要的包
library(tidyverse)
library(survival)
library(survminer)
library(pheatmap)

# 假设数据已加载并预处理(模拟数据)
# 这里创建模拟数据集，实际分析应使用UK Biobank数据
set.seed(123)
n <- 344324

# 创建模拟数据
data <- tibble(
  id = 1:n,
  # 生存时间和状态
  time = rweibull(n, shape = 1, scale = 10),
  status = rbinom(n, 1, 0.0135),
  # 协变量
  age = rnorm(n, 56, 8),
  sex = sample(c("Male", "Female"), n, replace = TRUE, prob = c(0.458, 0.542)),
  apoe_e4 = rbinom(n, 1, 0.25),
  assessment_center = sample(1:22, n, replace = TRUE),
  # 模拟风险因素(实际分析中应为文章中的62个显著因素)
  disability = rbinom(n, 1, 0.1),
  stroke = rbinom(n, 1, 0.02),
  diabetes = rbinom(n, 1, 0.05),
  depression = rbinom(n, 1, 0.1),
  hand_grip = rnorm(n, 30, 10),
  lung_function = rnorm(n, 0, 1),
  income = sample(c("<18k", "18-31k", "31-52k", "52-100k", ">100k"),
                  n, replace = TRUE, prob = c(0.1, 0.2, 0.3, 0.3, 0.1))
)
```

2. 定义分析函数

```r
# 定义进行Cox回归的函数
run_cox <- function(data, var, subgroup = NULL) {
  # 构建公式
  if (is.null(subgroup)) {
    formula <- as.formula(paste("Surv(time, status) ~", var, "+ age + sex + apoe_e4 + factor(ass
essment_center)"))
  } else {
    # 对组分析时不需要重复协变量本身
    formula <- as.formula(paste("Surv(time, status) ~", var, "+ age + factor(assessment_cente
r)"))
  }

  # 运行Cox模型
  fit <- coxph(formula, data = data)

  # 提取结果
  tidy_fit <- broom::tidy(fit, conf.int = TRUE, exponentiate = TRUE) %>%
    filter(term == var) %>%
    select(term, estimate, conf.low, conf.high, p.value)

  return(tidy_fit)
}

# 定义分析所有变量和所有亚组的函数
analyze_all <- function(data, vars) {
  # 全样本分析
  full_sample <- map_dfr(vars, ~run_cox(data, .x), .id = "var") %>%
    mutate(subgroup = "Full sample")

  # 按年龄分组(<65 vs >=65)
  age_young <- map_dfr(vars, ~run_cox(filter(data, age < 65), .x, subgroup = "age"), .id = "va
r") %>%
    mutate(subgroup = "Age<65")
  age_old <- map_dfr(vars, ~run_cox(filter(data, age >= 65), .x, subgroup = "age"), .id = "var"
) %>%
    mutate(subgroup = "Age>=65")
```

3. 运行分析

```r
# 定义要分析的变量(实际应为文章中的62个显著因素)
vars_to_analyze <- c("disability", "stroke", "diabetes", "depression", "hand_grip", "lung_functi
on", "income")

# 运行分析
results <- analyze_all(data, vars_to_analyze)

# 查看结果
head(results)
```

4. 创建热图

```r
# 准备热图数据
heatmap_data <- results %>%
  select(var, subgroup, estimate) %>%
  pivot_wider(names_from = subgroup, values_from = estimate) %>%
  column_to_rownames("var")

# 创建显著性标记(文章中用星号表示Bonferroni校正后显著)
signif_data <- results %>%
  select(var, subgroup, significant) %>%
  pivot_wider(names_from = subgroup, values_from = significant) %>%
  column_to_rownames("var") %>%
  mutate_all(~ifelse(. == "Bonferroni-corrected", "*", ""))

# 定义颜色
color_palette <- colorRampPalette(c("blue", "white", "red"))(100)
```

```r
# 绘制热图
pheatmap(
  heatmap_data,
  color = color_palette,
  scale = "none",
  cluster_rows = FALSE,
  cluster_cols = FALSE,
  display_numbers = signif_data,
  number_color = "black",
  fontsize_number = 12,
  main = "Associations between modifiable risk factors and incident dementia",
  angle_col = 45,
  breaks = seq(0.5, 2, length.out = 101)
)
```

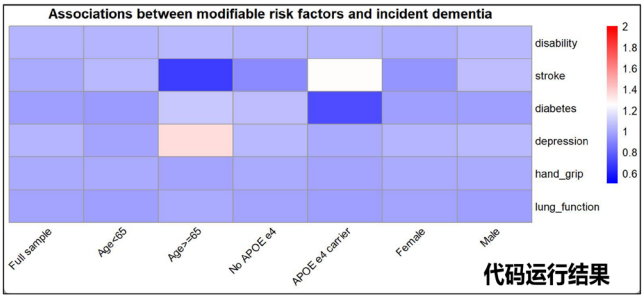

图2 探索大语言模型DeepSeek在研究分析中的应用测试

Figure 2. Exploring the application of the LLM DeepSeek in study analysis: a test case

注：A. DeepSeek按照常规队列研究的分析步骤提供的队列研究分析代码；B. 上传一篇研究论文全文，DeepSeek识别了文中相应图片的类型和主要内容（如包含亚组分析的热图），并分别提供了模拟数据、分析函数和绘图代码（空间限制仅展示部分结果），所提供的代码成功在R软件中运行并生成相应图片。

识别暴露与健康结局间复杂关系的能力、增强因果推理能力以及全流程助力研究与实践。当然，LLMs 的应用也面临诸多挑战。例如，“幻觉”问题可能导致结果偏差，缺乏标准化的基准数据集和评价指标限制了模型性能的评估与比较，伦理与法律问题亟需规范，计算资源和专业知识对广泛应用形成阻碍，以及过度依赖 AI 可能带来不良后果等。为应对这些挑战，科研工作者和管理者应加强跨学科合作，制定标准化指南和最佳实践，推动 LLMs 在公共卫生决策中的合理集成，

优先考虑公平和公正原则，以及加强相关领域的培训和能力建设等。

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 Copeland BJ. Artificial intelligence (AI). Definition, examples, types, applications, companies, & facts[EB/OL]. (2024-10-25) [2025-02-01]. https://www.britannica.com/technology/artificial-intelligence.

2 Gangwal A, Ansari A, Ahmad I, et al. Generative artificial intelligence in drug discovery: basic framework, recent advances, challenges, and opportunities[J]. Front Pharmacol, 2024, 15: 1331062. DOI: 10.3389/fphar.2024.1331062.

3 Zhang K, Yang X, Wang Y, et al. Artificial intelligence in drug development[J]. Nat Med, 2025, 31(1): 45-59. DOI: 10.1038/s41591-024-03434-4.

4 Picard M, Leclercq M, Bodein A, et al. Improving drug repositioning with negative data labeling using large language models[J]. J Cheminform, 2025, 17(1): 16. DOI: 10.1186/s13321-025-00962-0.

5 Rough K, Rashidi ES, Tai CG, et al. Core concepts in pharmacoepidemiology: principled use of artificial intelligence and machine learning in pharmacoepidemiology and healthcare research[J]. Pharmacoepidemiol Drug Saf, 2024, 33(11): e70041. DOI: 10.1002/pds.70041.

6 李佩芳，陈佳丽，宁宁，等. ChatGPT 在医学领域的应用进展及思考 [J]. 华西医学，2023, 38: 1456-1460. [Li PF, Chen JL, Ning N, et al. Application progress and thinking of ChatGPT in medical domain[J]. West China Medical Journal, 2023, 38: 1456-1460.] DOI: 10.7507/1002-0179.202309179.

7 Schmallenbach L, Bärnighausen TW, Lerchenmueller MJ. The global geography of artificial intelligence in life science research[J]. Nat Commun, 2024, 15(1): 7527. DOI: 10.1038/s41467-024-51714-x.

8 Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence[J]. NPJ Digit Med, 2020, 3: 126. DOI: 10.1038/s41746-020-00333-z.

9 Wang H, Fu T, Du Y, et al. Publisher correction: scientific discovery in the age of artificial intelligence[J]. Nature, 2023, 621(7978): E33. DOI: 10.1038/s41586-023-06559-7.

10 Li X, Guo Y. Paradigm shifts from data-intensive science to robot scientists[J]. Sci Bull(Beijing), 2025, 70(1): 14-18. DOI: 10.1016/j.scib.2024.09.029.

11 Pividori M, Greene CS. A publishing infrastructure for AI-assisted academic authoring[J/OL]. bioRxiv, 2023. (2023-01-21) [2025-02-15]. DOI: 10.1101/2023.01.21.525030.

12 何剑虎，王德健，赵志锐，等. 大语言模型在医疗领域的前沿研究与创新应用 [J]. 医学信息学杂志，2024, 45(9): 10-18. [He JH, Wang DJ, Zhao ZR, et al. The frontier research and innovative applications of large language models in the medical field[J]. Journal of Medical Informatics, 2024, 45(9): 10-18.] DOI: 10.3969/j.issn.1673-603 6.2024.09.002.

13 Wang S, Zhao Z, Ouyang X, et al. Interactive computer-aided diagnosis on medical image using large language models[J]. Commun Eng, 2024, 3(1): 133. DOI: 10.1038/s44172-024-00271-8.

14 Akilesh S, Sheik Abdullah A, Abinaya R, et al. A novel AI-based chatbot application for personalized medical diagnosis and review using large language models[C]. 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), 2023: 1-5.

15 施呈昊，屠馨怡，史佳伟，等. 大语言模型临床实践应用范围综述 [J]. 医学信息学杂志，2024, 45(9): 19-26. [Shi CH, Tu XY, Shi JW, et al. A scoping review of the application of large language models in clinical practice[J]. Journal of Medical Informatics. 2024, 45(9): 19-26.] DOI: 10.3969/j.issn.1673-6036.2024.09.003.

16 Wang H, Fu T, Du Y, et al. Scientific discovery in the age of artificial intelligence[J]. Nature, 2023, 620(7972): 47-60. DOI: 10.1038/s41586-023-06221-2.

17 Bukhtiyarova O, Abderrazak A, Chiu Y, et al. Major areas of interest of artificial intelligence research applied to health care administrative data: a scoping review[J]. Front Pharmacol, 2022, 13: 944516. DOI: 10.3389/fphar.2022.944516.

18 Sessa M, Khan AR, Liang D, et al. Artificial intelligence in pharmacoepidemiology: a systematic review. Part 1-overview of knowledge discovery techniques in artificial intelligence[J]. Front Pharmacol, 2020, 11: 1028. DOI: 10.3389/fphar.2020.01028.

19 谈志远，赵荣生. 人工智能技术在药物不良反应监测与上报中应用的研究进展 [J]. 临床药物治疗杂志，2019, 17(2): 23-27. [Tan ZY, Zhao RS. Progress of studies of artificial intelligence in surveillance and report of adverse drug reactions[J]. Clinical Medication Journal, 2019, 17(2): 23-27.] DOI: 10.3969/j.issn.1672-3384.2019.02.006.

20 Sessa M, Liang D, Khan AR, et al. Artificial intelligence in pharmacoepidemiology: a systematic review. Part 2-comparison of the performance of artificial intelligence and traditional pharmacoepidemiological techniques[J]. Front Pharmacol, 2021, 11: 568659. DOI: 10.3389/fphar.2020.568659.

21 Li Y, Tao W, Li Z, et al. Artificial intelligence-powered pharmacovigilance: a review of machine and deep learning in clinical text-based adverse drug event detection for benchmark datasets[J]. J Biomed Inform, 2024, 152: 104621. DOI: 10.1016/j.jbi.2024.104621.

22 Wang J, Cheng Z, Yao Q, et al. Bioinformatics and biomedical informatics with ChatGPT: year one review[J]. Quanti Bio, 2024, 12(4): 345-359. DOI: 10.1002/qub2.67.

23 Chakraborty C, Bhattacharya M, Lee SS, et al. The changing scenario of drug discovery using AI to deep learning: Recent advancement, success stories, collaborations, and challenges[J].

Mol Ther Nucleic Acids, 2024, 35(3): 102295. DOI: 10.1016/j.omtn.2024.102295.

24 Pun FW, Ozerov IV, Zhavoronkov A. AI-powered therapeutic target discovery[J]. Trends Pharmacol Sci, 2023, 44(9): 561-572. DOI: 10.1016/j.tips.2023.06.010.

25 Arnold C. Inside the nascent industry of AI-designed drugs[J]. Nat Med, 2023, 29(6): 1292-1295. DOI: 10.1038/s41591-023-02361-0.

26 Gao Z, Li L, Ma S, et al. Examining the potential of ChatGPT on biomedical information retrieval: fact-checking drug-disease associations[J]. Ann Biomed Eng, 2024, 52(8): 1919-1927. DOI: 10.1007/s10439-023-03385-w.

27 Guo T, Guo K, Nan B, et al. What can large language models do in chemistry? a comprehensive benchmark on eight tasks[A]. in:Proceedings of the 37th International Conference on Neural Information Processing Systems[C]. Red Hook, NY, USA: Curran Associates Inc, 2023: 59662-59688.

28 Juhi A, Pipil N, Santra S, et al. The capability of ChatGPT in predicting and explaining common drug-drug interactions[J]. Cureus, 2023, 15(3): e36272. DOI: 10.7759/cureus.36272.

29 Al-Ashwal FY, Zawiah M, Gharaibeh L, et al. Evaluating the sensitivity, specificity, and accuracy of ChatGPT-3.5, ChatGPT-4, Bing AI, and Bard against conventional drug-drug interactions clinical tools[J]. Drug Healthc Patient Saf, 2023, 15: 137-147. DOI: 10.2147/DHPS.S425858.

30 Chen Q, Sun H, Liu H, et al. An extensive benchmark study on biomedical text generation and mining with ChatGPT[J]. Bioinformatics, 2023, 39(9): btad557. DOI: 10.1093/bioinformatics/btad557.

31 Huang X, Estau D, Liu X, et al. Evaluating the performance of ChatGPT in clinical pharmacy: a comparative study of ChatGPT and clinical pharmacists[J]. Br J Clin Pharmacol, 2024, 90(1): 232-238. DOI: 10.1111/bcp.15896.

32 Wang R, Feng H, Wei GW. ChatGPT in drug discovery: a case study on anticocaine addiction drug development with Chatbots[J]. J Chem Inf Model, 2023, 63(22): 7189-7209. DOI: 10.1021/acs.jcim.3c01429.

33 Liu S, Wang J, Yang Y, et al. Conversational drug editing using retrieval and domain feedback[C]. The Twelfth International Conference on Learning Representations, 2023. https://openreview.net/forum?id=yRrPfKyJQ2.

34 Liang Y, Zhang R, Zhang L, et al. DrugChat: towards enabling ChatGPT-like capabilities on drug molecule graphs[J/OL]. arXiv, 2023. (2023-05-18) [2025-02-25]. http://arxiv.org/abs/2309.03907.

35 Ye G, Cai X, Lai H, et al. DrugAssist: a large language model for molecule optimization[J]. Brief Bioinform, 2024, 26(1): bbae693. DOI: 10.1093/bib/bbae693.

36 Cai X, Lai H, Wang X, et al. Comprehensive evaluation of molecule property prediction with ChatGPT[J]. Methods , 2024, 222: 133-141. DOI: 10.1016/j.ymeth.2024.01.004.

37 Zeng Z, Yin B, Wang S, et al. ChatMol: interactive molecular discovery with natural language[J]. Bioinformatics, 2024, 40(9): btae534. DOI: 10.1093/bioinformatics/btae534.

38 Wei J, Bosma M, Zhao V, et al. Finetuned language models are zero-shot learners[C]. International Conference on Learning Representations, 2021. https://openreview.net/forum?id=gEZrGCozdqR.

39 Zhao Z, Chen B, Li J, et al. ChemDFM-X: towards large multimodal model for chemistry[J]. Sci China Inform Sci, 2024, 67(12): 220109. https://arxiv.org/abs/2409.13194.

40 Wei J, Zhuo L, Fu X, et al. DrugReAlign: a multisource prompt framework for drug repurposing based on large language models[J]. BMC Biol, 2024, 22(1): 226. DOI: 10.1186/s12915-024-02028-3.

41 Matheny ME, Yang J, Smith JC, et al. Enhancing postmarketing surveillance of medical products with large language models[J]. JAMA Netw Open, 2024, 7(8): e2428276. DOI: 10.1001/jamanetworkopen.2024.28276.

42 Ong JCL, Chen MH, Ng N, et al. A scoping review on generative AI and large language models in mitigating medication related harm[J]. NPJ Digit Med, 2025, 8(1): 182. DOI: 10.1038/s41746-025-01565-7.

43 Wang X, Xu X, Liu Z, et al. Bidirectional encoder representations from transformers-like large language models in patient safety and pharmacovigilance: a comprehensive assessment of causal inference implications[J]. Exp Biol Med (Maywood), 2023, 248(21): 1908-1917. DOI: 10.1177/15353702231215895.

44 Ma C, Wolfinger RD. Toward an explainable large language model for the automatic identification of the drug-induced liver injury literature[J]. Chem Res Toxicol, 2024, 37(9): 1524-1534. DOI: 10.1021/acs.chemrestox.4c00134.

45 Deng Y, Xing Y, Quach J, et al. Developing large language models to detect adverse drug events in posts on x[J]. J Biopharm Stat, 2024: 1-12. DOI: 10.1080/10543406.2024.2403442.

46 Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy[J]. NPJ Digit Med, 2023, 6(1): 186. DOI: 10.1038/s41746-023-00927-3.

47 Ocana A, Pandiella A, Privat C, et al. Integrating artificial intelligence in drug discovery and early drug development: a transformative approach[J]. Biomarker Res, 2025, 13: 45. DOI: 10.1186/s40364-025-00758-2.

收稿日期：2025 年 04 月 08 日 修回日期：2025 年 07 月 02 日
本文编辑：冼静怡 周璐敏