

专栏: 人工智能与新药研发 (解决方案和案例)

· 专家评述 ·

人工智能助力新药研发

姜正羽, 尤启冬*

(中国药科大学药学院药物化学系, 江苏省药物分子设计和成药性优化重点实验室, 南京 210009)



[专栏主编] 尤启冬, 中国药科大学药学院教授, 博士生导师。国家“万人计划”教学名师, 国家教学名师。获国家科技进步二等奖, 教育部霍英东优秀青年教师奖、江苏省“五一劳动奖章”等。担任国家药典委员会顾问委员; 中国药学会理事、药物化学专业委员会委员; 江苏省药学会副秘书长、常务理事、药物化学专业委员会主任委员。担任《中国药科大学学报》常务副主编。



[专栏副主编] 姜正羽, 中国药科大学药学院教授、博士生导师。入选国家重大人才工程青年学者、中国科协青年人才托举工程, 江苏省杰出青年基金, 江苏省“青蓝工程”青年骨干教师, 江苏省科协青年科技人才托举工程, 江苏省“科技副总”。主持包括国家新药创制重大专项、国家自然科学基金等科研项目 10 余项。获得第 21 届中国药学会-施维雅青年药物化学奖。担任《中国药科大学学报》青年编委。

中图分类号 TP181;R914 文献标志码 A 文章编号 1000-5048(2024)03-0281-03

doi: [10.11665/j.issn.1000-5048.2024060302](https://doi.org/10.11665/j.issn.1000-5048.2024060302)

新药研发一直是前沿科学和技术最重要的应用领域。近年来, 人工智能(AI)飞速发展, 已经从前沿的概念成为了赋能各行各业的重要技术。同样的, AI 在新药研发领域也正在引领着一场新的研究模式的变革。

AI 技术在新药研发领域的应用贯穿了新药研究的全流程。在靶标层面, AI 技术为新靶标的发现提供了全新途径。运用 AI 技术, 可以更好地从海量数据中挖掘疾病-靶标-药物的关联, 构建知识库, 从而建立数据驱动的靶标发现和确证的新模式, 为新药研究立项提供效率更高的决策方式。同时, AI 技术提供了可信度更高的生物大分子结构预

测方法, 有效辅助了针对复杂生物大分子复合物的药物发现。而在候选物发现方面, AI 技术在包括分子设计、虚拟筛选和化合物优化等方面均有良好的运用潜力。特别是深度生成模型能够直接生成具有类药属性的新分子, 为快速获得更安全、更有效的药物提供了新的工具。而机器学习等技术则提供了高可信度的药物相关属性预测模型, 能够在早期阶段预测药物的安全性和有效性, 降低研发过程中的风险和成本。AI 技术和自动化技术结合, 能够实现药物分子的高效合成和连续制造, 大幅降低候选物发现的人力资源消耗。而在药物的临床开发阶段, AI 同样能够辅助临床试验的优化。AI 技术

已经用于改善试验设计、加速患者招募以及提高试验效率等方面。同样, AI 在精准和个性化医疗领域也大有可为。在医学影像分析领域, AI 能够提高医学影像解读的准确性和效率。同时, AI 能够有效利用大数据, 研究个性化医疗模型, 制定个体化治疗方案。随着 AI 相关技术在药物研发领域的深化应用, 未来新药研发的模式将会不断迭代优化。由于期刊篇幅有限, 本专栏聚焦于 AI 提升新药发现效率和 AI 助力新模式药物研究两个方面。以期通过一些代表性的综述和研究论文, 更好地帮助读者理解和运用 AI 技术, 开展创新药物研究。

1 AI 提升新药发现效率

新药研发是一个典型的长周期系统工程, 其具有周期长、投入高、风险大和失败率高的特点。特别是近些年来, 新药研究的研发周期和研发成本不断增加, 而成功率却未见上升。传统的药物研发方法依赖于大量的临床前候选物筛选实验和耗费巨大的临床试验以验证其安全性和有效性, 药物研发资源和时间大量消耗。伴随计算科学快速发展, 计算机辅助药物设计(CADD)技术应运而生, 带来了合理药物设计和虚拟筛选等药物设计领域的巨大进步。而随着 AI 技术的发展, 人工智能药物设计(AIDD)也快速崛起, 为新药研发开辟了新路径。AIDD 和 CADD 都通过采用先进的算法和数据处理方法, 优化了传统药物研发流程, 提高了研发效率和成功率。包括自然语言处理、图像识别、深度学习和机器学习等 AI 技术, 不仅加速了药物靶点的识别和候选药物的筛选, 还在药理评估、质量检验和制剂设计等多个环节发挥关键作用。复旦大学药学院付伟教授团队撰写的《AIDD 与 CADD 提升药物成功率的思考》深入探讨了 AIDD 和 CADD 技术的区别和联系, 及其在提升药物研发效率的作用, 为新药研究中更好地利用相关技术提供了思路和经验。

化学结构是化学小分子药物的核心。人工智能技术的独特魅力在于其具有直接生成药物结构的潜力和可能性。小分子生成模型已成为 AI 药物发现领域的重要研究方向。该类模型包括生成对抗网络(GANs)、变分自编码器(VAEs)和扩散模型等多种方法, 已被初步证明在生成复杂分子结构和优化类药性质等方面具有广阔的发展潜力。浙江大学药学院董晓武教授团队撰写的《基于人工智能

的小分子生成模型在药物发现中的研究进展》综合分析了上述小分子生成模型技术在药物发现过程中的应用, 展示了这类新技术和新方法在如何补充和改进传统药物设计方法中的作用。同时, 该论文还提出了一些当前小分子生成模型在数据质量、模型复杂性、计算成本及泛化能力等方面的挑战, 并对未来的研究方向进行了展望。

生物活性测试筛选和化学结构迭代优化是候选物发现阶段的主要工作。利用计算方法预测生物活性在 CADD 时代已经被证明是一种有效提升药物研发效率的路径。AI 技术通过模型和算法的创新, 能够有效地提升预测精度, 提升先导物发现和优化的效率。针对抗肿瘤药物研究的热点靶标 KRAS, 中国药科大学理学院赵鸿萍教授团队开展研究, 撰写了《基于机器学习的 KRAS 抑制剂活性预测模型研究》。该团队将机器学习应用于 KRAS 抑制剂的定量构效关系模型的构建。研究从 ChEMBL、BindingDB、PubChem 3 个数据库中收集了大量 KRAS 小分子抑制剂的活性和结构数据, 采用 3 种不同的特征筛选方式结合随机森林、支持向量机、极端梯度提升机 3 种机器学习模型, 构建了 9 个不同的分类器, 研究发现支持向量机(SVM)模型结合 MI 信息筛选显示出最佳性能。该研究为使用人工智能方法进行 KRAS 抑制剂筛选提供了新的方法。

AI 技术在知识库的构建中具有独特优势。利用 AI 模型构建相关知识库可为靶标选择、新药设计和药物重定位等提供有力支撑。这一模式可以进一步地拓展至利用 AI 构建知识库辅助精神活性物质鉴定中。该方法对结构未知精神活性物质的鉴别具有独特优势。司法鉴定科学研究院陈航副主任法医师团队针对这一应用场景, 选择社会热点的合成大麻素, 进行了总结, 撰写了《机器学习在合成大麻素识别鉴定中的应用进展》。合成大麻素是一种人工合成的可以引起公共健康风险的精神活性物质, 且合成大麻素结构多变, 容易被结构修饰, 结构未知的合成大麻素的快速出现使得对其鉴识面临了新的挑战。机器学习为结构未知合成大麻素的鉴识以及可能的来源推断提供了新的策略。该文章阐述了常用机器学习方法的原理以及机器学习技术在合成大麻素类物质的质谱分析、拉曼光谱分析、代谢组学以及定量构效关系等方面的应用。同时, 文章进行了相关分析和展望, 以期为未知合成大麻素的鉴识提供新的思路。

由于 AI 技术的显著优势, 其在创新药物研究中的实际应用正在迅猛开展。杭州碳硅智慧科技发展有限公司联合国内多家研究机构带来了一个利用 AI 技术开展药物研究实例《**基于人工智能驱动分子工厂技术的 Menin 抑制剂优化**》。研究团队利用自主提出的一种基于人工智能技术的创新药物分子设计和优化工作流程“分子工厂”开展了 Menin 抑制剂的优化工作。“分子工厂”流程融合了自主研发的智能分子生成模型、高性能分子对接算法以及高精度亲和力预测方法, 已作为核心模块被整合进一站式药物设计软件平台 DrugFlow, 为先导化合物发现和优化提供了一整套成熟的解决方案。利用“分子工厂”模块, 研究团队针对 Menin 蛋白开展了抗耐药第二代抑制剂的研发。通过计算和实验的结合, 快速获得多个潜力化合物, 其中化合物 RG-10 与已进入 II 期临床的阳性参照分子 SNDX-5613 相比, 其对 M327I 和 T349M 突变体的抑制活性显著提升。上述研究充分展现了“分子工厂”技术在新药研发项目中的独特优势。该论文为相关研究者给出了一个很好的示范。

2 AI 助力新模式药物研究

近年来, 多肽药物、蛋白药物、核酸药物等药物的新模式快速发展。人工智能在这些新模式的药物研究中同样具有独特的作用。随着核酸药物领域蓬勃发展, 正逐步成为小分子和抗体类药物后的药物新模式。中国药科大学生命科学与技术学院余文颖教授团队针对 AI 在核酸药物研究中的应用, 总结撰写了《**人工智能在核酸药物研发中的应用和进展**》。该文章概述了核酸药物研发领域的人工智能算法、数据库、表征等基础, 重点阐述了人工智能在核酸结构预测、小核酸药物设计等核酸药物研发环节中的应用和进展, 旨在为人工智能和核酸药物交叉学科发展提供参考。

华东理工大学药学院唐贇教授团队针对 AI 在多肽药物研究中的应用进行了相关讨论。抗癌肽 (ACPs) 因其高效低毒和高选择性优势成为多肽药物研究中的一个重要方向。基于人工智能的 ACPs 识别和设计方法较传统实验方法成本低廉、成功率高且能够探索更广阔的序列空间。唐教授团队撰写的《**人工智能在抗癌肽研发中的应用与挑战**》

重点介绍了人工智能技术在 ACPs 生成和识别过程中的应用, 包括深度生成模型探索新型 ACPs 设计以及基于机器学习和深度学习的 ACPs 识别方法。同时, 文章作者还讨论了当前研究中存在的模型可复现性和可解释性不足、缺乏经过实验验证的阴性数据等挑战, 并对未来研究方向提出展望, 以期 ACPs 的研发提供新思路。

3 展 望

当前生物医药创新面临研发效率下降带来的巨大挑战, 制药公司对于能够提升研发效率的新技术存在巨大需求。AI 技术的快速发展和新药研究领域的巨大需求共同促进了 AI 技术在创新药物研究中的快速应用。全球范围内已有数百项制药公司与 AI 技术公司的合作, AI 与药物发现正在深度融合。AI 在新药研发中的技术、方法和应用方面都取得了突破性的进展。

算力、算法和数据是 AI 的几个核心要素。随着 AI 技术突飞猛进, 算力和算法的瓶颈正在逐渐被攻克。云计算、大数据、机器学习等工智能相关技术发展迅速, 为研究者提供了处理庞大研发数据的新手段, 为新药研发带来了新的视角。而数据则成为 AI 在生物医药领域大显身手的主要障碍。传统的药物研发以实验科学为主。药物研发数据的记录、储存和分析方式都以实验为核心, 以新药研究上市为终极目标, 根据药物研发流程而调整。生物制药公司的海量数据出于知识产权保护的角度并未公开。而无论是学术界和工业界, 公开的数据都以阳性数据为主, 大量阴性结果均属于沉睡的状态。AI 作为虚拟科学、计算科学和数据科学范畴内的方法, 数据是 AI 技术的核心资源, 从数据中提取知识, 对于数据有严格的内在要求。在实际应用中, AI 直接使用传统药物研发模式的数据往往遇到困难, AI 制药的“数据基础设施”尚待完备。

无论 AI 是否能够真正改变药物研发过程, AI 在药物发现这种工作量巨大, 耗时费力的长流程研发活动中的巨大作用定将日益凸显。AI 技术将有助于新药研究人员从大量的重复性工作中解放出来, 有更多的时间来思考新药研究中的科学问题、集中精力投入到创造性的工作当中。

引用本文 姜正羽, 尤启冬. 人工智能助力新药研发 [J]. 中国药科大学学报, 2024, 55(3): 281–283.

Cite this article as: JIANG Zhengyu, YOU Qidong. Assistance of artificial intelligence in new drug development[J]. J China Pharm Univ, 2024, 55(3): 281–283.