

Artificial Intelligence-Driven Drug Molecule Design and Innovative Antidiabetic Drug Development Strategies

Xiufeng Zhang

Maternal and Child Health Hospital, Inner Mongolia Autonomous Region, Hohhot, Inner Mongolia, 010020, China

Abstract

Diabetes mellitus, a chronic disease characterized by glucose metabolism disorders, has become a major global public health challenge. Traditional drug development methods, marked by lengthy cycles, high costs, and low success rates, have constrained pharmaceutical innovation in diabetes treatment. The application of artificial intelligence (AI) has opened new pathways for drug molecule design. Through deep learning, molecular generative models, virtual screening, and multi-omics data integration, AI can significantly enhance the efficiency and accuracy of new drug development. This paper reviews key AI technologies in drug design, explores their synergistic roles in target discovery, lead compound screening, pharmacological evaluation, and preclinical research, and proposes an "AI + multi-omics + high-throughput experimentation" intelligent R&D model. It analyzes the technical and ethical challenges in AI-driven drug development, providing insights for establishing an efficient and intelligent new drug innovation system.

Keywords

artificial intelligence; drug molecule design; antidiabetic drugs; molecular generation

基于人工智能的药物分子设计及抗糖尿病创新药物研发策略

张秀峰

内蒙古自治区妇幼保健院, 中国·内蒙古 呼和浩特 010020

摘要

糖尿病是一种以血糖代谢紊乱为特征的慢性疾病,已成为全球重大公共卫生问题。传统药物研发周期长、成本高、成功率低,限制了抗糖尿病药物创新。人工智能的应用为药物分子设计开辟新途径,通过深度学习、分子生成模型、虚拟筛选及多组学数据融合,可显著提升新药研发效率与准确性。本文综述AI在药物设计中的关键技术体系,探讨其在靶点发现、先导化合物筛选、药效评价及临床前研究中的协同作用,并提出“AI+多组学+高通量实验”智能研发模式,分析AI药物研发的技术与伦理挑战,为构建高效、智能化新药创新体系提供参考。

关键词

人工智能; 药物分子设计; 抗糖尿病药物; 分子生成

1 引言

糖尿病的全球患病率持续上升,世界卫生组织数据显示,截至2024年,全球糖尿病患者已超过5亿人,并呈年轻化趋势。现有的抗糖尿病药物包括胰岛素类、双胍类、磺脲类、GLP-1受体激动剂等,但多数药物存在副作用、耐药性及疗效下降等问题,亟需开发新机制、高安全性的新型药物。传统药物研发平均耗时10年以上,成本可达20亿美元,而临床成功率不足10%。面对药物创新的“研发瓶颈”,人工智能以其强大的数据挖掘与预测能力,成为新药发现领域的颠覆性工具。人工智能通过分子结构建模、药理机制推演与虚拟筛选,可在短时间内完成数百万种化合物的性质预测,大幅降低实验成本与时间。

【作者简介】张秀峰(1984-),女,中国内蒙古乌兰察布人,本科,从事儿科、妇产科药学研究。

近年来,人工智能在药物分子设计中的应用范围不断扩大,从早期的QSAR(定量构效关系)建模到如今的深度生成模型与多模态药理网络,AI技术已从“辅助分析”转向“主导创新”。在抗糖尿病药物研发领域,AI不仅可用于识别新的作用靶点,如AMPK、GLP-1、SGLT2等,还可通过多目标优化算法设计具备多重药理作用的化合物^[1]。本文旨在系统分析人工智能在药物分子设计与抗糖尿病创新药物研发中的应用机制,探讨技术路径、实验验证及未来趋势,为实现精准、高效的药物创新提供系统参考。

2 人工智能在药物分子设计中的理论基础与技术框架

2.1 药物分子设计的传统模式与AI革命

传统药物设计基于实验化学与药理学理论,主要依赖高通量筛选、分子对接与结构优化。其过程线性、周期长、失败率高。人工智能的介入实现了从经验驱动到数据驱动的

范式转变。深度学习算法可在大量化学与生物数据中识别隐藏规律，通过分子特征编码、图结构分析与性质预测实现药物设计自动化。AI系统能从已有药物数据库中学习化学空间特征，进而生成具有特定药理性质的新分子，实现“逆向设计”^[2]。例如，AlphaFold在蛋白质结构预测领域的突破，使药物靶点三维结构推断的准确性显著提升，为精准设计配体提供了基础。

2.2 分子表示与特征学习模型的构建

在AI药物设计中，分子表示是核心问题。传统的SMILES线性符号表达难以捕捉空间结构特征，而图神经网络(GNN)通过将分子表示为节点和边的图结构，可有效学习原子间的拓扑关系与化学键类型，实现分子性质的高维映射。近年来，Message Passing Neural Network(MPNN)与Graph Attention Network(GAT)被广泛应用于分子溶解度、毒性、结合能等特征预测。结合Transformer结构的ChemBERTa、MolFormer等预训练模型，能够在百万级分子数据上进行特征迁移学习，为药物分子生成提供强大的语义基础。

2.3 生成式模型与强化学习在分子创新中的应用

生成模型使AI具备了“创造”药物分子的能力。变分自编码器(VAE)、生成对抗网络(GAN)和扩散模型被用于分子结构的自动生成与优化。强化学习(RL)机制进一步引入“奖励函数”体系，根据分子的药理活性、合成可行性和毒理预测动态调整生成策略。例如，REINVENT与MolDQN算法可实现多目标分子设计，在满足药效需求的同时优化药代动力学参数。这种AI驱动的“智能迭代”设计，使药物研发从被动筛选走向主动创造。

3 人工智能在抗糖尿病药物研发中的关键应用场景

3.1 AI在靶点发现与作用机制预测中的应用

糖尿病病理网络跨越胰岛 β 细胞应激、胰岛素信号转导、慢性低度炎症与肠道菌群代谢轴，传统单通路解析难以解释异质性^[3]。以转录组、蛋白组、代谢组与表观组为基础，构建疾病—基因—药物异构图，利用GCN/注意力机制与多模态Transformer完成跨模态对齐，并以因果表示学习与可干预推理区分相关与因果。模型在AMPK、PPAR γ 、GLP-1等经典靶点上可复现实验已知通路，并识别复合调控节点(如短链脂肪酸/胆汁酸经肠促胰素轴调控)，据此判定协同多靶干预的最优组合，为候选靶点优先级与机制假说提供可验证的定量证据。

3.2 AI在先导化合物筛选与优化中的作用

面向海量化学空间，级联式AI筛选以生成模型/扩散模型+强化学习产生满足药效、合成可达与知识库相容性的母核，再经粗粒度打分与集成对接淘汰低优先级分子，最终以分子动力学与自由能扰动精修亲和力与构象稳定性。针

对GLP-1R、DPP-4、SGLT2等靶点，模型并行优化脂溶性、极性表面积、P-gp外排敏感性与CYP介导代谢脆弱位点，结合合成路线预测与可达性评分降低实验成本。主动学习在“设计—合成—测定—反馈”闭环中动态挑选信息增益最大的化合物，显著压缩从虚拟命中到先导确证的迭代周期。

3.3 AI辅助的药效与安全性预测

药效与安全评估通过多任务网络联学ADMET端点、受体偏向信号与离体—在体桥接参数，并以不确定性量化输出置信区间，在分子生成阶段即提前剔除高风险结构。基于PBPK/PK-PD的机理模型与群体贝叶斯框架，可预测暴露—效应关系与人群间变异，联用hERG阻断、DILI、基因毒性等毒理子模型完成多目标权衡。结合电子病历与真实世界数据的外部校准与子群敏感性分析，校正年龄、性别、肾功能与并用药带来的系统性偏差；可解释方法(如注意力归因/SHAP)明确导致风险的结构基序与代谢位点，提高结论的可追溯性与临床转化可行性。

4 基于AI的抗糖尿病药物研发体系构建

4.1 AI与多组学数据的融合建模机制

多组学融合的要义在于以统一本体与标准化元数据打通基因变异、转录调控、蛋白互作与代谢通路的层级关联，并以可微结构嵌入实现跨模态对齐。以GCN/GCNII在蛋白—化合物—通路异构图上的消息传递为骨架，引入多模态Transformer对序列、图结构与连续谱信号进行联合编码，通过对比学习与门控注意力抑制批次效应与平台偏移。在知识图谱上融入因果边与约束正则，采用因果表示学习与工具变量推断，区分混杂与真实靶向作用；利用结构方程与可干预推理开展“do-算子”级别的反事实评估，量化靶点扰动对胰岛素信号与葡萄糖稳态的因果贡献^[4]。将代谢组通量与稳定同位素示踪数据并入模型，配合深度强化学习对代谢网络进行策略搜索，可给出对胰岛素抵抗状态下关键节点的药物匹配建议，并以外部纵向队列与独立实验室数据实现稳健验证与可迁移性评估。

4.2 AI驱动虚拟实验与高通量模拟平台

虚拟实验平台以分子动力学与自由能扰动为计算内核，以图注意力势能面与等变神经力场近似量子级相互作用，结合不确定性量化提供可置信的结合能区间。上亿级规模的虚拟筛选通过级联式代理模型与分层过滤完成：生成模型产出合成可达分子，粗粒度打快速淘汰低优先级结构，精细对接与增强采样解析关键氢键与疏水口袋，最终以FEP/ λ —动力学精修亲和力。主动学习在“设计—合成—测试—反馈”闭环中基于信息增益挑选最具价值实验点，云原生并行计算与机器人合成表征实现连续流式验证；贝叶斯优化在ADMET多目标下进行帕累托搜索，自动权衡口服暴露、代谢稳定性与安全边际。平台级审计记录数据谱系与模型版本，确保不同批次与不同站点间结果的可复现与可追溯，为

候选物从先导到临床前的加速迭代提供工程化保障。

4.3 AI 在临床前研究与个体化药物开发中的作用

临床前阶段，AI辅助的系统药理学与群体PK/PD模型将离体、离体器官片与动物数据进行层级映射，利用迁移学习外推至人群参数分布，并以生理学机理模型校准关键吸收与代谢环节。在个体化开发中，基于基因型—表型—暴露反应的整合模型联合电子病历与真实世界数据，构建“数字孪生胰岛—外周组织”仿真体，模拟不同代谢亚型对剂量—疗效—安全的动态响应，给出人群分层与剂量自适应策略。药物—药物相互作用与并发症风险通过图因果网络进行前瞻评估；公平性与稳健性以子群灵敏度分析与校准曲线量化，防止在年龄、性别与遗传背景上的系统性偏差。结合自适应试验与模型知情药物研发 workflow，形成从生物标志物伴随诊断、剂量优化到长期结局预测的闭环证据链，提升抗糖尿病候选药物的临床可转化性与决策透明度。

5 AI 赋能抗糖尿病创新药物研发的挑战与前景

5.1 算法可靠性与数据质量的双重挑战

AI模型在抗糖尿病药物设计中的预测力取决于数据充分性、代表性与标注一致性。真实世界药理数据跨多平台与多实验条件采集，存在批次效应、测量噪声与异构标准，易诱发样本分布偏移与标签偏差，削弱模型外推能力。为提高鲁棒性，应构建覆盖化学空间与人群异质性的多源基准库，实施严格的数据溯源与元数据记录，采用分层抽样与困难样本重加权缓解类别不平衡；在建模层面引入不确定性量化、集成学习与分布外检测，联合迁移学习与领域自适应对抗批次差异，辅以可解释神经网络与注意力归因识别关键基团与相互作用位点。通过因果表示学习与知识图谱约束，将先验机理嵌入模型结构，减少“相关性陷阱”；以独立时空外部验证、前瞻性盲测与跨实验室重复性评估作为硬性门槛，方可确保算法在先导筛选、ADMET预测与多靶点优化中的稳定可复现。

5.2 伦理与合规问题的审慎应对

AI驱动研发涉及患者隐私、临床数据使用与算法生成分子的权属界定，需在合规与创新间取得平衡^[5]。数据治理层面，应确立最小可用原则与明确同意机制，采用联邦学习、差分隐私与安全多方计算在不暴露原始数据的条件下完成建模；建立数据使用审计、访问控制与撤回机制，确保跨境流动与二次利用符合监管要求。方法学层面，推行模型卡与数据卡，详列训练语料、适用边界与已知风险，设置持续

监测与失效预案。知识产权方面，应区分算法、训练数据与生成分子的权利归属，明确共同研发与平台服务的专利与许可路径，防止“数据—模型—产物”之间的权利错配。伦理委员会与监管沙盒可为新范式提供阶段性评估与动态准入，结合透明可解释与可追责设计，降低“黑箱决策”在药物安全性审查与临床转化中的阻碍。

5.3 未来发展方向与技术集成趋势

抗糖尿病药物研发将迈向多模态、端到端与自动化协同的系统工程。在模型层面，面向蛋白—化学—多组学—临床表型的基础模型将以跨模态对齐与结构先验驱动，从靶点结构预测到分子生成与ADMET一体化优化；量子化学近似与可微分物理约束用于精化结合能与反应路径。实验层面，闭环“设计—合成—表征—反馈”的自驱实验室通过主动学习选择最具信息增益的实验，显著压缩迭代周期。临床与人群层面，结合真实世界数据的数字孪生“虚拟胰岛—个体代谢网络”可进行体内动态模拟与人群分层敏感性分析，支持剂量—疗效—安全的个体化优化。标准化数据接口、共享基准与可复现实验协议将成为跨机构协作的基础设施，促成从基础发现到临床转化的高通量、低成本与可追溯的新范式。

6 结语

人工智能为抗糖尿病药物研发注入了前所未有的创新动力。通过深度学习、生成模型与多组学融合，AI正在重塑药物分子设计的范式，使“精准—高效—低成本”的新药开发成为可能。尽管AI药物研发仍面临算法透明性、数据安全与监管规范等挑战，但其在靶点发现、分子生成与个体化治疗中的潜力已得到广泛验证。未来，构建以人工智能为核心、以多源数据驱动、以实验验证为支撑的智能药物研发体系，将成为应对糖尿病全球化挑战的关键路径。科技与医学的深度融合，必将推动抗糖尿病药物从经验探索走向理性设计，实现疾病管理的智能化与精准化。

参考文献

- [1] 李星雨, 武潼, 徐启兰, 等. 机器学习在药物分子结构优化中的研究进展[J/OL]. 药学学报, 1-41[2025-10-31].
- [2] 谭相端, 张妞妞, 马小盼, 等. 人工智能背景下研究生药物分子设计课程“五维一体”教学改革探索[J]. 科教文汇, 2025, (12): 118-123.
- [3] 潘晓航. 基于深度学习的药物分子生成研究[D]. 兰州大学, 2025.
- [4] 袁素素, 叶秀云, 鄢仁祥. 基于蛋白质受体的药物分子计算机辅助设计策略[J]. 生物信息学, 2024, 22(03): 159-173.
- [5] 董靖鑫. 基于深度学习的药物分子图像生成与识别研究[D]. 湖南大学, 2023.