

人工智能结合生理药代动力学模型的研究进展

李龙杰¹, 计佩影², 郑澳乐¹, 穆耶赛尔·阿里甫¹, 相小强^{1*}

(1. 复旦大学药学院, 上海 200120; 2. 上海市杨浦区控江医院, 上海 200000)

摘要: 生理药代动力学 (physiologically based pharmacokinetic, PBPK) 模型已经被广泛用于预测药物的吸收、分布、代谢和排泄等特性, 而基于机器学习 (machine learning, ML) 和人工智能 (artificial intelligence, AI) 可以和 PBPK 模型进行深度融合, 从而加快 PBPK 的预测速度和提高其预测质量, 进一步加快药物研发进展。本文介绍了机器学习和人工智能在药代动力学中的应用, 对基于机器学习和人工智能的生理药代动力学模型研究进展进行综述, 并分析了机器学习和人工智能应用的局限性以及其应用前景和展望。

关键词: 生理药代动力学模型; 人工智能; 机器学习; 药代动力学; 药物毒理学; 药物相互作用

中图分类号: R917 文献标识码: A 文章编号: 0513-4870(2024)09-2491-08

Research progress of artificial intelligence combined with physiologically based pharmacokinetic models

LI Long-jie¹, JI Pei-ying², ZHENG Ao-le¹, ALIFU MUYESAIER¹, XIANG Xiao-qiang^{1*}

(1. School of Pharmacy, Fudan University, Shanghai 200120, China; 2. Kong Jiang Hospital of Yangpu District, Shanghai 200000, China)

Abstract: Physiologically based pharmacokinetic (PBPK) models have been widely used to predict various stages of drug absorption, distribution, metabolism and excretion. Models based on machine learning (ML) and artificial intelligence (AI) can provide better ideas for the construction of PBPK models, which can accelerate the prediction speed and improve the prediction quality of PBPK. ML and AI can complement the advantages of PBPK model to accelerate the progress of drug research and development. This review introduces the application of machine learning and artificial intelligence in pharmacokinetics, summarizes the research progress of physiological pharmacokinetic models based on machine learning and artificial intelligence, and analyzes the limitations of machine learning and artificial intelligence applications and their application prospects and prospects.

Key words: physiologically based pharmacokinetic model; artificial intelligence; machine learning; pharmacokinetics; pharmaceutical toxicology; drug-drug interaction

1 机器学习和人工智能在药代动力学中的应用

传统药代动力学和药效学研究通过动物实验和人体临床试验等方法进行, 然而这些方法存在诸多局限性, 例如伦理问题、样本量大小问题和个体间的变异性问题等, 在特定人群例如老年人和儿童中也存在临床

试验的困难性。此外, 这些研究往往不能准确预测药物在人体内的药代动力学和药效学的特性。近年来, 人们利用机器学习 (machine learning, ML) 和人工智能 (artificial intelligence, AI) 以更快、更经济、更准确的方式预测药物的药代动力学和药效学, 这些技术方法已被广泛用于预测药物的吸收、分布、代谢和排泄 (absorption, distribution, metabolism, and excretion, ADME) 特性^[1]。

机器学习是人工智能的子领域, 它使用统计方法,

收稿日期: 2024-03-05; 修回日期: 2024-04-19.

基金项目: 国家自然科学基金面上项目 (82373952).

*通讯作者 Tel: 86-21-51980024, E-mail: xiangxq@fudan.edu.cn

DOI: 10.16438/j.0513-4870.2024-0195

能够在有或没有明确编程的情况下进行学习,它已广泛用于研究药代动力学特性、识别药物靶点、标准化合成方法、构建临床试验和预测药物毒性。机器学习可分为监督学习、无监督学习、强化学习。监督学习的算法需要一组输入数据(通常称为向量 X)和已知的数据响应(即输出数据,也称为标签 Y),即模型在已提供标签的数据集上进行训练,监督学习的主要任务包括分类(预测标签)和回归(预测数量),主要有贝叶斯模型(Bayesian model)、随机森林(random forest, RF)、支持向量机(support vector machine, SVM)决策树(decision tree, DT)等。无监督学习是指一种不向算法提供带有标签的数据集的机器学习,通常任务包括聚类、降维、密度估计等。强化学习是一种基于反馈的学习方法,可通过奖励和有效惩罚的模式来优化模型。机器学习的另一个子领域称为深度学习(deep learning, DL),即使用人工神经网络(artificial neural networks, ANNs)来适应大量实验数据并从中学习,作为人工智能的一个更深层次的子集,它可以处理数据并创建决策模式,主要用于监督学习任务。而深度神经网络(deep neural network, DNN)指具有数个隐藏层的人工神经网络,与ANNs不同,它可以应用于监督和非监督任务,而且可以自动提取输入数据中有用的特征^[1-3]。

1.1 利用机器学习和人工智能预测药物吸收和分布

基于AI/ML的模型可以通过分析药物的溶解度、渗透性和制剂特性等因素来预测药物吸收参数,如口服生物利用度和胃肠吸收速率常数。这些模型可以分析药物理化性质,并将它们与吸收数据关联起来以估计血药浓度。基于AI/ML的模型是预测药物释放和吸收参数的有效而快速的工具之一。

近年来基于统计学利用AI/ML开发定量结构-活性关系(quantitative structure-activity relationship, QSAR)模型或者定量结构-性质关系(quantitative structure-property relationship, QSPR)模型越来越普遍, QSAR和QSPR模型可使用各种化学计量方法找出一组分子的结构特征(称为描述符)和已知活性或性质之间的定量相关性^[4]。这些模型可基于不同的AI/ML的方法构建,已经有研究通过ANNs^[5]、分类和回归模型^[6]、SVM^[7]、多元自适应回归样条(multivariate adaptive regression splines, MARS)^[8]等构建QSAR或QSPR模型来预测药物在人体肠道的吸收率,这些模型在预测药物吸收方面显示出了良好的预测性能。Turón等^[9]开发了一种基于AI/ML的QSAR/QSPR平台ZairaChem,利用了RF和神经网络的方法,可根据药物物理化学参数进行药物渗透性预测。

此外,药物在体内分布的平衡时间与药物的血浆/组织平衡分布比例系数 K_p 有着密切联系,目前已经开发了几种基于AI/ML的计算机模型,通过分析人体组织的组成信息和药物的物理化学特征(如亲脂性、药物血浆游离分数 F_{up} 、辛醇-水的分配系数 K_{ow})以及组织中蛋白质、磷脂的比例来预测不同组织或器官中的 K_p ^[10]。

1.2 利用机器学习和人工智能预测药物代谢和排泄

AI/ML模型可以识别与特定代谢转化相关的结构特征,预测潜在的代谢产物并深入了解参与药物代谢的主要酶。从广义上讲,人工智能在药物代谢预测领域的应用可分为三大类:①预测代谢位点;②预测负责代谢的酶或转运体;③预测酶促动力学^[2]。

药物代谢位点的预测可为开发更安全、更有效的药物提供策略。目前,已经开发了几种基于AI/ML的代谢位点预测模型,并显示出良好的预测性能。例如,Zaretski等^[11-13]开发了一种名为“RS-Predictor”的工具,即使用分层描述符与多实例学习方法的组合来识别CYP450介导的药物分子代谢的潜在位点,其形式与SVM类似,但在使用期间可以为每个实例确定误差范围。

除此之外,基于AI/ML的预测模型可用于CYP底物的特异性,例如Mishra等^[14]使用基于SVM的QSAR模型预测了主要CYP同工酶(如CYP3A4、2D6、1A2、2C9和2C19)的底物特异性。研究^[15]报道建立了基于更多描述符数量的SVM预测器,根据化学结构计算的分子描述符来预测141种已批准药物在人体中的主要代谢途径(例如CYP酶、有机阴离子转运多肽OATP等),并显示出较好的预测性能。近年来,国内外出现了很多ADMET预测平台,例如Daina等^[16]开发的SwissADME平台使用基于SVM的底物/非底物或抑制剂/非抑制剂的大数据集,能够估计一种药物是否为P-gp的底物或CYP同工酶的抑制剂。Xiong等^[17]开发的ADMETlab 2.0平台,可预测药物是否为CYP1A2、2C19、3A4等代谢酶的底物或抑制剂,并成功应用于预估奥拉帕尼不是CYP1A2和CYP2D6的抑制剂,而是CYP3A4和CYP2C9的抑制剂,这与阿斯利康报道的体外细胞色素P450评估结果一致^[18]。

最后,基于AI/ML的模型可以计算酶促动力学,例如反应速率和酶-底物相互作用,以预测药物的代谢过程。通过考虑酶表达水平、遗传变异和药物相互作用等因素,人工智能模型可以评估代谢对药物清除和疗效的潜在影响。这些信息对于优化药物剂量方案和预测潜在的药物相互作用非常有价值^[19]。

2 机器学习和人工智能在生理药代动力学模型中的应用

2.1 生理药代动力学模型的局限性

PBPK 模型广泛用于模拟药物在体内的分布和清除,但这些模型通常很复杂,而且 PBPK 模型利用体外实验数据的外推来进行体内预测,由于体内条件与体外不同,使用体外数据构建的模型可能无法准确预测体内条件,例如在某些情况下,一些低溶解度药物在肠道中沉淀的可能性会被高估,因为药物辅料可增加其溶解性从而避免沉淀,药物实际溶解度会更高^[20]。不仅如此, PBPK 模型中所需要的参数可能难以测量或者没有相关数据,例如儿童的单克隆抗体 (monoclonal antibodies, mAbs) 的 PBPK 模型的建立中没有 Fc 受体 (Fc receptor, FcRn) 的绝对丰度^[21]。如果药物和酶的信息较少,所建立的 PBPK 模型可能表现不佳^[22]。如果缺乏对酶-转运蛋白相互作用的体内相关性的了解,这些限制可使得转运蛋白和非 CYP 介导的药物的相互作用 (drug-drug interaction, DDI) 更难以预测^[23]。

总的来说,建立新药的 PBPK 模型是耗时且消耗大量资源的。很多输入参数需要通过将模型与动物研究的观察数据进行拟合来估计。然而,动物研究耗时昂贵,需要考虑动物伦理问题,故进行动物研究来帮助开发每种药物的 PBPK 模型是不现实的。因此,将 ML 和 AI 与 PBPK 模型相结合的新方法,不仅在很大程度上加快了药物 PBPK 模型的构建,并且节省了大量的资源^[10]。

2.2 机器学习和人工智能的应用优势

目前 AI/ML 已被用于活性药物设计、药物活性评

分、药物虚拟筛选以及药代动力学的计算机评估等^[3,24]。近年来,药物设计的传统方法已被计算机辅助药物设计所取代。AI/ML 被广泛用于改进药物设计技术和缩短药物发现所需时间。此外,利用人工智能可以方便地识别靶蛋白,从而提高设计药物的成功率。

总的来说, AI/ML 的发展增加了药物成功批准上市的可能性,而且可以优化研发流程来降低经济成本,可以减少动物实验的需求,促进动物福利事业的发展,同时也可以为患者提供更有效的治疗方案并提高患者的依从性。

2.3 机器学习和人工智能在生理药代动力学模型中的应用进展

2.3.1 简化 PBPK 模型和预测重要 PK 参数与 PD 结果

AI/ML 与基于生理的药代动力学模型相结合可以用来预测 ADME 等重要 PK 参数。Chou 等^[10]研究了一种新兴的研究范式,他们首先利用文献和体外测定结果建立研究药物的数据库,然后利用 SVM、RF 等 AI 算法来估计药物的 ADME 参数,将其作为 PBPK 模型的输入参数,随后将 PBPK 的输出参数进行评估,评估有效的参数又可整合到数据库中形成闭环。具体见图 1。

基于此种范式, AI/ML 和 PBPK 结合可用于预测一些难以获得的参数。Chou 等^[25]按照以上流程,预测了纳米颗粒 (nanoparticles, NPs) 在肿瘤微环境中的几个关键动力学参数,例如纳米颗粒的最大摄取速率常数 ($KTRES_{max}$)、释放速率常数 ($KTRES_{rel}$)、希尔系数 ($KTRES_n$) 和达到 50% 最大摄取速率的时间 ($KTRES_{50}$),这些数据通过实验测量非常困难,所以通

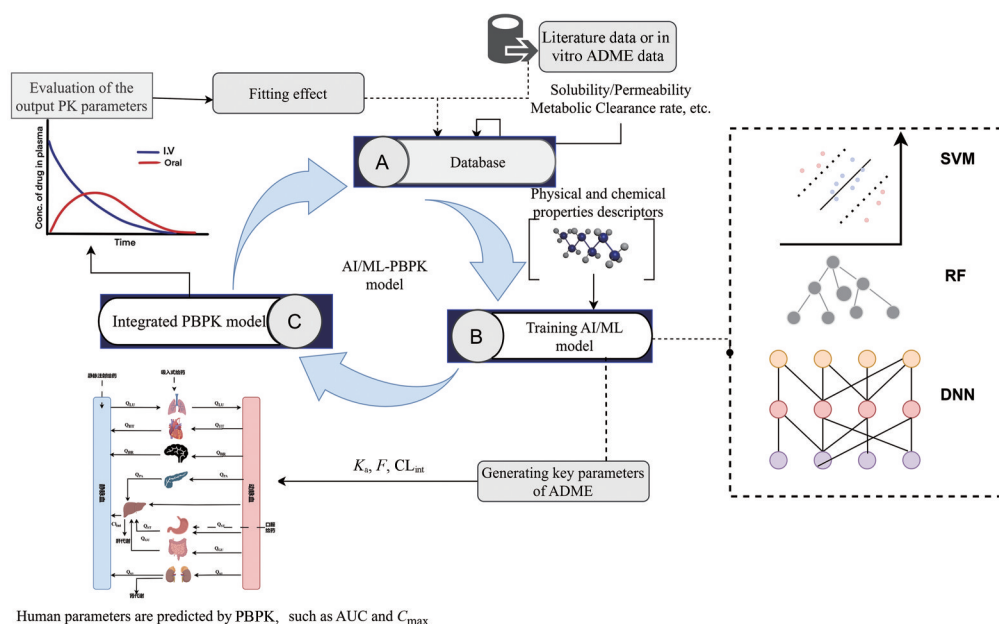


Figure 1 The research paradigm to integrate AI/ML with PBPK, reprinted with permission from reference^[10] research paradigm. Copyright © 2022, © The Author (s) 2022. Published by Oxford University Press on behalf of the Society of Toxicology

常通过拟合体内动物实验数据来获得,但是该模型的构建严重依赖动物研究,并且模型的外推能力有限,而AI与PBPK结合的新模型可以在传统的机器学习算法(例如SVM)、集成学习模型(例如RF)和DNN基础上开发基于AI的QSAR模型,其不需要依赖动物研究来训练数据集,可以直接根据NP的理化特性以及与肿瘤相关的生理特征来预测其关键动力学参数,然后再选择性能最好的AI-QSAR模型与PBPK模型组合成AI-PBPK模型,预测关键参数,以模拟不同类型NPs在小鼠肿瘤中的内吞和释放。有研究^[26]也开发了类似此范式的方法预测PK参数,首先生成了大量虚拟药物属性的数据集,利用这些数据集建立PBPK模型生成PK参数,然后使用了RF算法通过数据集和生成PK参数训练机器学习模型。

AI/ML可以生成PBPK建模过程中重要的吸收参数从而简化模型的构建。比如Kamiya等^[27]开发了一种简化的PBPK模型,首先用于大鼠的血药浓度推测,他们使用轻量级梯度增强框架(light gradient boosting framework, LightGBM)的机器学习算法预测三个PBPK输入参数(吸收速率常数、体循环血容量和肝脏内在清除率),然后将此参数纳入大鼠PBPK模型来预测了246种化合物在血浆、肝脏和肾脏的AUC和 C_{max} ,最终结果显示出了良好的相关性($r \geq 0.83$)。随后Kamiya等^[28]又报道了预测人体血药浓度的相关研究,通过使用机器学习算法岭回归(ridge regression)和LightGBM预测PBPK模型的所需输入参数,预测了212种化合物口服后的血浆药物浓度。除此之外,Schneckener等^[29]利用各种建模技术,其中包括ML和PBPK模型的组合,对大鼠的静脉和口服药物血药浓度和口服生物利用度进行预测,研究结果有近70%的准确度和精确度。

AI/ML可以帮助加快PBPK模型的构建,缩短建模时间。例如有研究者^[30]开发了一种新型高通量PBPK(high-throughput PBPK, HT-PBPK)方法,该方法使用了ML预测了PBPK的相关输入参数,其产生的结果与完整的PBPK建模相当,但将模拟时间从几小时缩短到几秒钟。

AI/ML与PBPK结合也可以帮助更好预测特殊给药状态下的PK参数。例如透皮给药,BIOiSIM是一种基于AI/ML的PK-PD建模平台,其包含16个对应人体组织和器官的隔室,该平台首先被报道成功预测了人类受试者透皮给药,例如吗啡、丁丙诺啡、尼古丁的PK输出参数,为其制定更合理有效的给药方案提供了参考^[31]。随后,研究^[32]报道使用BIOiSIM平台预测了丁丙诺啡和羟考酮透皮给药后的血浆暴露情况来优化和

调整给药策略,研究结果显示两种化合物的三种PK输出(AUC_{0-t}, C_{max} , t_{max})的模拟结果都非常准确,拟合的平均误差小于1.3倍。

AI/ML还可以融合至PBPK模型中,然后与PD模型相结合预测药物效应。例如Wu等^[33]建立了AI-PBPK平台,首先使用ML预测PBPK建模参数,再使用PBPK输出PK参数,将此参数构建基于机制的PD模型来预测药物浓度的变化如何影响胃酸分泌和胃pH值,成功比较了五种钾竞争性酸阻断剂的治疗效果和确定其最佳治疗剂量。

2.3.2 改进药物毒理学模型 近年来,AI/ML方法越来越多地应用于毒理学的不同学科领域,包括毒物动力学^[34]、纳米毒理学^[35,36]等。在毒理学领域,PBPK模型是进行人类健康风险评估的重要工具,特别是在剂量-效应分析、暴露评估、体外到体内外推(*in vitro* to *in vivo* extrapolation, IVIVE)以及毒性和剂量学数据的种间外推中,目前已经开发了许多针对不同药物的PBPK模型,其中许多已被用于药物的风险评估^[37,38],但建立模型时需要的体内PK数据必须从动物实验中外推获得,这种传统方法无法满足数千种化学品毒理学评估的需求。将AL和ML的方法结合可以改进PBPK的毒理学模型,为改善PBPK毒理学模型的局限性提供新的思路。

AI/ML方法可以帮助药物毒理学研究中预测重要的代谢动力学参数。Pradeep等^[32]使用了多种机器学习算法(例如SVM和RF等),构建了基于1487种环境化学品数据集的结构特性以确定预测2个基本毒代动力学参数[血浆中未结合的化学物质的百分数(F_{ub})和内在清除率(CL_{int})]的QSAR模型。该模型可用于预测无法获得实验数据的其他化学品的毒代动力学参数。这也表明为大量化合物开发通用的用于毒性剂量预测评估的PBPK模型是可行的。

纳米毒理学评估是纳米材料安全应用于人体的前提。AI和ML可凭借其强大的数据挖掘能力,可以大大节省药物毒理学评估的时间和资源^[39]。ML已成功应用于模拟纳米材料的生物动力学和相互作用,研究最多的便是纳米定量构效模型(quantitative structure-activity relationships at nanoscale, nano-QSAR),这些模型可以用来为复杂的PBPK模型提供输入参数;除此之外,当PBPK用于识别纳米材料的毒性作用机制时,如果靶基因位于相关通路中,PBPK模型便无法预测其是否调控其他基因及内部的作用机制,这时候可以运用ML算法例如DL、RF等来弥补这一不足^[36]。

2.3.3 辅助药物相互作用的评估 PBPK模型已经广泛用于DDI预测^[40-42],并被美国食品药品监督管理局

(FDA) 和欧洲药监局 (EMA) 推荐用于临床 DDI 试验和群体 PK 研究的设计。此外, PBPK 甚至可以提供专门临床试验的替代方案, 以找到相互作用的药物联合给药或在特殊人群 (例如肾功能不全人群等) 中的给药方案的建议^[43]。DDI 预测模型包括基础模型、静态机制模型和动态机制模型 (如 PBPK 模型)。另一种新方法则可采用监督式学习的 ML, 以体外和体内 ADME 数据为标签变量, 预测结果变量血浆暴露情况, 通过利用变量之间的统计关系对 DDI 进行预测。

与 PBPK 相比, ML 预测 DDI 的关键优势在于, ML 模型在每次数据更新时自动重新建立训练样本之间的数学关系, 而 PBPK 需要手动推导微分方程来关联不同的 PK 参数。ML 模型可以使用其已知的 DDI 作为标签数据, 以及药物开发周期早期可用的相关特征, 对现有候选药物进行训练, 例如 Ryu 等^[44]将深度学习模型应用于 DDI 预测中, 开发了 DeepDDI 模型, 以药物结构信息和 DrugBank 的 DDI 数据集为变量进行训练, 该模型应用了 DNN 的 ML 算法, 可以针对给定的药物对同时预测多种 DDI 类型, 也可以用于药物和食物的相互作用预测。除此之外, AI/ML 方法如卷积神经网络可用来预测蛋白质和配体复合物的结合常数, 例如抑制常数 K_i 或者半数最大抑制浓度 (IC_{50})^[45]。这表明构建 PBPK 模型预测 DDI 时, AI/ML 模型可以帮助生成难以获得的 DDI 输入参数。除此之外, ML 也可以帮助筛选具有高 DDI 风险的药物, 在使用更耗时的 PBPK 模型之前, 可以将实验后得到的额外的 PK 参数和给药数据应用于基于回归的 ML 模型, 这样更准确地估计 DDI 风险。在药物上市后, 对 PBPK 建模所需参数的了解有限的情况下, ML 可用于预测病例中已建立的药物之间的 DDI 风险^[22]。

总的来说, ML 和 PBPK 可以相辅相成, PBPK 模型可用于生成额外的参数, 这些参数可用作 ML 模型的输入特征, 然后利用算法将模型进行训练, 从而用于做出更好的预测, 而 ML 可以帮助生成 PBPK 预测 DDI 所需要的关键参数^[22]。

3 机器学习和人工智能应用的局限性和解决思路

3.1 缺乏透明度

AI 模型使用复杂的算法, 通常被称为“黑匣子”, 也称为黑箱性质 (black box), 即人们很难理解模型如何得出预测, 因为 ML 可以基于算法输出预测结果, 但是无需提供得到结果的原理^[46]。尤其是性能较好的无监督学习式的模型例如 DL 的结果是很难解释的, 这可能会让临床医生或者医药行业的监管部门对此方法产生质疑从而限制了 AI 模型的应用^[47,48]。由于 ML 算法的黑箱性质, 基于 ML 的模型不能完全解释药物物

理化学性质的贡献, 也不能提取用于预测结果的关键特征, 透明度高、可解释性强的机器学习模型仍有待人们开发研究。

要解决这一局限性, 首先要将输入数据结构化以便机器能够分析, 其次, 机器必须能够识别输入数据和结果相关的关键向量, 然后研究人员还需要设计内在可解释的模型^[49]。例如 Ciallella 等^[50]开发了一个基于知识的深度神经网络来识别雌激素模拟物。由于模型的构建是基于可解释的由雌激素受体 α 启动的不良结果信号通路, 他们的模型可以做出可解释的输入端到输出端的预测。这种基于知识和机制的方法是开发可解释的 ML 模型的潜在的通用策略。

3.2 数据可用性有限

AI/ML 模型往往需要大量的数据才能进行准确的预测。例如深度学习需要大量数据, 因此该技术受数据范围和数据质量的限制^[51]。另外, 在某些情况下, 可用于特定药物或人群的数据可能有限, 导致预测不太准确或结果有偏差。例如, 罕见疾病的可用数据可能有限, 或者某些类型的数据, 如纵向数据或真实世界证据, 可能并不容易获得, 这可能会限制 AI/ML 模型的实用性。而且数据的质量也对模型的准确预测有很大的影响。此外, 用于训练 AI 模型的数据可能不能代表研究的目标人群, 这可能导致结果存在偏倚, 例如在临床试验中, 患者群体的同质性是药理学领域的一个重要问题。如果特定的人口统计或疾病状态在训练数据集中没有得到充分的表示, 模型对药物在该特定人群中的疗效做出准确预测的能力可能会受到影响。另外, 在大分子药物领域中, AI 的应用也是有限的, 目前对生物疗法的 PK 参数进行建模的研究有限, 将 AI/ML 方法应用于抗体和蛋白质的障碍之一是与小分子疗法相比数据集较小^[52]。而且由于商业数据的机密性, 某些药物的数据可能有限。

为了解决这一限制, 人们可以设计新的模型来简化所需要的数据, 例如, Altae-Tran 等^[53]开发了一种基于卷积神经网络开发了一次性学习方法 (one-shot learning), 可以用来显著降低 AI 模型所需的数据量, 该方法可以在小型生物数据集上提供强大的性能。还有一种是使用数据加强的方法, 通过生成合成数据来补充现有数据, 以此提高预测的质量和可靠性^[54]。

3.3 整合新数据能力有限

新的药物或者临床试验数据信息可能决定了模型预测能力的优劣, 因此在获得新信息时对模型进行更新是至关重要的。但是往往训练出了 AI/ML 模型后, 很难根据新生成的数据和新算法持续快速地更新模型。例如, 随着新药的推出或临床试验产生更多数据,

AI/ML 模型可能需要更新以纳入这些新的信息^[51]。

可以用的解决方法包括设计易于更新的模型, 或者将模型集成到一个更大的框架中, 该框架可以随着时间的推移不断改进^[1]。

3.4 过拟合问题

ML 模型在训练过程中, 由于输入特征太多, 训练数据集太小或者代表性数据少都可能出现过拟合问题, 使模型预测准确率降低。过度拟合问题将导致模型无法准确预测训练数据之外的数据集, 并可能在将误导性信息应用于毒性终点预测时而误报潜在危险^[55]。

为了解决这一限制, 首先应该在模型开发过程中, 应用简约原则选择少量必要的参数以拟合最佳结果, 同时避免过度参数化^[56]。针对噪声太多问题, 一些研究使用早期停止 (early-stopping) 和网络约简 (network-reduction) 的策略, 以确定适当的停止训练的时间, 从而减少噪音或者修剪噪音。对于输入特征太多的问题, 可以使用正则化 (regularization) 的策略, 删除无用的特征, 只选择有用的特征^[57]。

4 总结和展望

AI/ML 目前已被广泛用于药学领域, 包括识别分子靶点、寻找先导化合物、合成类药物化合物和预测 ADME 参数、辅助药物临床试验设计等^[58], 而将 AI/ML 与 PBPK 建模相结合是进一步促进药物开发的新方法。首先, 基于 AI/ML 的方法可以通过使用机器学习算法来识别模型的相关特征, 从而简化 PBPK 模型的开发; 其次, 基于 AI/ML 的计算方法还可以优化 PBPK 模型的 PK 参数, 更好预测药物在人体中的暴露量和清除率等, 从而减少动物实验和人类临床试验的需要; 最后, AI 和 ML 能大大节省药物 PBPK 毒理学评估的时间和资源, 而在 DDI 的预测中, AI 和 ML 可以预测 PBPK 的 DDI 动态建模所需要的抑制参数 K_i , 从而加速模型的构建。

近年来, 随着神经-常微分方程 (neural-ODE algorithms) 算法的发展。神经常微分方程模型可以集成不同的特征, 使得 ML 方法具备了超越训练数据的预测能力, 这将增强该模型在外推到具有不同药物性质和实验设置的通用 PBPK 模拟方面的能力。Lu 等^[59]首次将神经常微分方程应用于 PK 建模, 表现出了出色的通用预测性能, 这可能使在一种治疗方案上训练的模型能外推应用于不同治疗方案。

总体而言, AI/ML 在医药领域的发展仍然存在巨大的挑战, 需要制定具有严格措施的法律法规来防止滥用, 例如数据所有权和患者的隐私权是一个重要的伦理问题, 需要确保合理正当收集和使用患者数据的方式并且保护他们的隐私和尊重他们的权利, 考虑患

者的安全和利益^[60,61]。因此, 将 AI/ML 充分合理应用到医药领域, 需要政府部门、科研人员、药企研发成员的共同努力。

作者贡献: 李龙杰负责对文章结构逻辑进行构思、文献检索和文章撰写; 计佩影、郑澳乐负责对文章进行文献补充和部分内容修改; 穆耶赛尔·阿里甫负责文章绘图和文章部分内容润色; 相小强指导论文写作和修改, 把控文章整体结构和质量。

利益冲突: 无任何利益冲突。

References

- [1] Vora LK, Gholap AD, Jetha K, et al. Artificial intelligence in pharmaceutical technology and drug delivery design [J]. *Pharmaceutics*, 2023, 15: 1916.
- [2] Yang X, Wang Y, Byrne R, et al. Concepts of artificial intelligence for computer-assisted drug discovery [J]. *Chem Rev*, 2019, 119: 10520-10594.
- [3] Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects [J]. *Drug Discov Today*, 2019, 24: 773-780.
- [4] De P, Kar S, Ambure P, et al. Prediction reliability of QSAR models: an overview of various validation tools [J]. *Arch Toxicol*, 2022, 96: 1279-1295.
- [5] Talevi A, Goodarzi M, Ortiz EV, et al. Prediction of drug intestinal absorption by new linear and non-linear QSPR [J]. *Eur J Med Chem*, 2011, 46: 218-228.
- [6] Ghafourian T, Freitas AA, Newby D. The impact of training set data distributions for modelling of passive intestinal absorption [J]. *Int J Pharm*, 2012, 436: 711-720.
- [7] Yan A, Wang Z, Cai Z. Prediction of human intestinal absorption by GA feature selection and support vector machine regression [J]. *Int J Mol Sci*, 2008, 9: 1961-1976.
- [8] Deconinck E, Ates H, Callebaut N, et al. Evaluation of chromatographic descriptors for the prediction of gastro-intestinal absorption of drugs [J]. *J Chromatogr A*, 2007, 1138: 190-202.
- [9] Turon G, Hlozek J, Woodland JG, et al. First fully-automated AI/ML virtual screening cascade implemented at a drug discovery centre in Africa [J]. *Nat Commun*, 2023, 14: 5736.
- [10] Chou WC, Lin Z. Machine learning and artificial intelligence in physiologically based pharmacokinetic modeling [J]. *Toxicol Sci*, 2023, 191: 1-14.
- [11] Zaretski J, Bergeron C, Rydberg P, et al. RS-predictor: a new tool for predicting sites of cytochrome P450-mediated metabolism applied to CYP 3A4 [J]. *J Chem Inf Model*, 2011, 51: 1667-1689.
- [12] Zaretski J, Rydberg P, Bergeron C, et al. RS-predictor models augmented with SMARTCyp reactivities: robust metabolic regioselectivity predictions for nine CYP isozymes [J]. *J Chem*

- Inf Model, 2012, 52: 1637-1659.
- [13] Zaretski J, Bergeron C, Huang TW, et al. RS-WebPredictor: a server for predicting CYP-mediated sites of metabolism on drug-like molecules [J]. *Bioinformatics*, 2013, 29: 497-498.
- [14] Mishra NK. Computational modeling of P450s for toxicity prediction [J]. *Expert Opin Drug Metab Toxicol*, 2011, 7: 1211-1231.
- [15] Toshimoto K, Wakayama N, Kusama M, et al. *In silico* prediction of major drug clearance pathways by support vector machines with feature-selected descriptors [J]. *Drug Metab Dispos*, 2014, 42: 1811-1819.
- [16] Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules [J]. *Sci Rep*, 2017, 7: 42717.
- [17] Xiong G, Wu Z, Yi J, et al. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties [J]. *Nucleic Acids Res*, 2021, 49: W5-W14.
- [18] McCormick A, Swaisland H, Reddy VP, et al. *In vitro* evaluation of the inhibition and induction potential of olaparib, a potent poly (ADP-ribose) polymerase inhibitor, on cytochrome P450 [J]. *Xenobiotica*, 2018, 48: 555-564.
- [19] Li YH, Meng Q, Yang MB, et al. Current trends in drug metabolism and pharmacokinetics [J]. *Acta Pharm Sin B*, 2019, 9: 1113-1144.
- [20] Santos LGA, Jaiswal S, Chen KF, et al. Real-world application of PBPK in drug discovery [J]. *Drug Metab Dispos*, 2023. DOI: 10.1124/dmd.122.001036.
- [21] Gill KL, Jones HM. Opportunities and challenges for PBPK model of mAbs in paediatrics and pregnancy [J]. *AAPS J*, 2022, 24: 72.
- [22] Gill J, Moullet M, Martinsson A, et al. Comparing the applications of machine learning, PBPK, and population pharmacokinetic models in pharmacokinetic drug-drug interaction prediction [J]. *CPT Pharmacometrics Syst Pharmacol*, 2022, 11: 1560-1568.
- [23] Peters SA, Dolgos H. Requirements to establishing confidence in physiologically based pharmacokinetic (PBPK) models and overcoming some of the challenges to meeting them [J]. *Clin Pharmacokinet*, 2019, 58: 1355-1371.
- [24] Sahu A, Mishra J, Kushwaha N. Artificial intelligence (AI) in drugs and pharmaceuticals [J]. *Comb Chem High Throughput Screen*, 2022, 25: 1818-1837.
- [25] Chou WC, Chen Q, Yuan L, et al. An artificial intelligence-assisted physiologically-based pharmacokinetic model to predict nanoparticle delivery to tumors in mice [J]. *J Control Release*, 2023, 361: 53-63.
- [26] Habiballah S, Reisfeld B. Adapting physiologically-based pharmacokinetic models for machine learning applications [J]. *Sci Rep*, 2023, 13: 14934.
- [27] Kamiya Y, Handa K, Miura T, et al. *In silico* prediction of input parameters for simplified physiologically based pharmacokinetic models for estimating plasma, liver, and kidney exposures in rats after oral doses of 246 disparate chemicals [J]. *Chem Res Toxicol*, 2021, 34: 507-513.
- [28] Kamiya Y, Handa K, Miura T, et al. Machine learning prediction of the three main input parameters of a simplified physiologically based pharmacokinetic model subsequently used to generate time-dependent plasma concentration data in humans after oral doses of 212 disparate chemicals [J]. *Biol Pharm Bull*, 2022, 45: 124-128.
- [29] Schneckener S, Grimbs S, Hey J, et al. Prediction of oral bioavailability in rats: transferring insights from *in vitro* correlations to (deep) machine learning models using *in silico* model outputs and chemical structure parameters [J]. *J Chem Inf Model*, 2019, 59: 4893-4905.
- [30] Naga D, Parrott N, Ecker GF, et al. Evaluation of the success of high-throughput physiologically based pharmacokinetic (HT-PBPK) modeling predictions to inform early drug discovery [J]. *Mol Pharm*, 2022, 19: 2203-2216.
- [31] Maharao N, Antontsev V, Hou H, et al. Scalable *in silico* simulation of transdermal drug permeability: application of BIOiSIM platform [J]. *Drug Des Devel Ther*, 2020, 14: 2307-2317.
- [32] Pradeep P, Patlewicz G, Pearce R, et al. Using chemical structure information to develop predictive models for *in vitro* toxicokinetic parameters to inform high-throughput risk-assessment [J]. *Comput Toxicol*, 2020, 16: 100136.
- [33] Wu K, Li X, Zhou Z, et al. Predicting pharmacodynamic effects through early drug discovery with artificial intelligence-physiologically based pharmacokinetic (AI-PBPK) modelling [J]. *Front Pharmacol*, 2024, 15: 1330855.
- [34] Bhattarai B, Walters WP, Hop C, et al. Opportunities and challenges using artificial intelligence in ADME/Tox [J]. *Nat Mater*, 2019, 18: 418-422.
- [35] Ji Z, Guo W, Wood EL, et al. Machine learning models for predicting cytotoxicity of nanomaterials [J]. *Chem Res Toxicol*, 2022, 35: 125-139.
- [36] Singh AV, Ansari MHD, Rosenkranz D, et al. Artificial intelligence and machine learning in computational nanotoxicology: unlocking and empowering nanomedicine [J]. *Adv Healthc Mater*, 2020, 9: e1901862.
- [37] Tan YM, Worley RR, Leonard JA, et al. Challenges associated with applying physiologically based pharmacokinetic modeling for public health decision-making [J]. *Toxicol Sci*, 2018, 162: 341-348.
- [38] Chen Q, Riviere JE, Lin Z. Toxicokinetics, dose-response, and risk assessment of nanomaterials: methodology, challenges, and future perspectives [J]. *Wiley Interdiscip Rev Nanomed Nanobio-technol*, 2022, 14: e1808.
- [39] Singh AV, Varma M, Laux P, et al. Artificial intelligence and machine learning disciplines with the potential to improve the nanotoxicology and nanomedicine fields: a comprehensive

- review [J]. *Arch Toxicol*, 2023, 97: 963-979.
- [40] Hanke N, Gómez-Mantilla JD, Ishiguro N, et al. Physiologically based pharmacokinetic modeling of rosuvastatin to predict transporter-mediated drug-drug interactions [J]. *Pharm Res*, 2021, 38: 1645-1661.
- [41] Bhatnagar S, Mukherjee D, Salem AH, et al. Dose adjustment of venetoclax when co-administered with posaconazole: clinical drug-drug interaction predictions using a PBPK approach [J]. *Cancer Chemother Pharmacol*, 2021, 87: 465-474.
- [42] Wen HN, He QF, Xiang XQ, et al. Predicting drug-drug interactions with physiologically based pharmacokinetic/pharmacodynamic modelling and optimal dosing of apixaban and rivaroxaban with dronedarone co-administration [J]. *Thrombosis Res*, 2022, 218: 24-34.
- [43] Hanke N, Frechen S, Moj D, et al. PBPK models for CYP3A4 and P-gp DDI prediction: a modeling network of rifampicin, itraconazole, clarithromycin, midazolam, alfentanil, and digoxin [J]. *CPT Pharmacometrics Syst Pharmacol*, 2018, 7: 647-659.
- [44] Ryu JY, Kim HU, Lee SY. Deep learning improves prediction of drug-drug and drug-food interactions [J]. *Proc Natl Acad Sci U S A*, 2018, 115: E4304-E4311.
- [45] Karlov DS, Sosnin S, Fedorov MV, et al. GraphDelta: MPNN scoring function for the affinity prediction of protein-ligand complexes [J]. *ACS Omega*, 2020, 5: 5150-5159.
- [46] Poon AIF, Sung JY. Opening the black box of AI-medicine [J]. *J Gastroenterol Hepatol*, 2021, 36: 581-584.
- [47] Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence [J]. *BMC Med*, 2019, 17: 195.
- [48] Kiseleva A, Kotzinos D, De Hert P. Transparency of AI in healthcare as a multilayered system of accountabilities: between legal requirements and technical limitations [J]. *Front Artif Intell*, 2022, 5: 879603.
- [49] Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead [J]. *Nat Mach Intell*, 2019, 1: 206-215.
- [50] Ciallella HL, Russo DP, Aleksunes LM, et al. Revealing adverse outcome pathways from public high-throughput screening data to evaluate new toxicants by a knowledge-based deep neural network approach [J]. *Environ Sci Technol*, 2021, 55: 10875-10887.
- [51] Tripathi MK, Nath A, Singh TP, et al. Evolving scenario of big data and artificial intelligence (AI) in drug discovery [J]. *Mol Divers*, 2021, 25: 1439-1460.
- [52] Akbar R, Bashour H, Rawat P, et al. Progress and challenges for the machine learning-based design of fit-for-purpose monoclonal antibodies [J]. *mAbs*, 2022, 14: 2008790.
- [53] Altae-Tran H, Ramsundar B, Pappu AS, et al. Low data drug discovery with one-shot learning [J]. *ACS Cent Sci*, 2017, 3: 283-293.
- [54] Blanco-González A, Cabezón A, Seco-González A, et al. The role of AI in drug discovery: challenges, opportunities, and strategies [J]. *Pharmaceuticals*, 2023, 16: 891.
- [55] Luechtefeld T, Rowlands C, Hartung T. Big-data and machine learning to revamp computational toxicology and its use in risk assessment [J]. *Toxicol Res*, 2018, 7: 732-744.
- [56] Basak SC, Vracko MG. Parsimony principle and its proper use/application in computer-assisted drug design and QSAR [J]. *Curr Comput Aided Drug Des*, 2020, 16: 1-5.
- [57] Ying X. An overview of overfitting and its solutions [J]. *J Phys*, 2019, 1168: 022022.
- [58] Gallego V, Naveiro R, Roca C, et al. AI in drug development: a multidisciplinary perspective [J]. *Mol Divers*, 2021, 25: 1461-1479.
- [59] Lu J, Deng K, Zhang X, et al. Neural-ODE for pharmacokinetics modeling and its advantage to alternative machine learning models in predicting new dosing regimens [J]. *iScience*, 2021, 24: 102804.
- [60] Guan J. Artificial intelligence in healthcare and medicine: promises, ethical challenges and governance [J]. *Chin Med Sci J*, 2019, 34: 76-83.
- [61] Naik N, Hameed BMZ, Shetty DK, et al. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? [J]. *Front Surg*, 2022, 9: 862322.