

## 新药开发前沿案例应用于药物设计学教学 ——以计算机辅助药物设计章节为例

王贯, 王欣, 帅雯\*, 欧阳亮\*

四川大学生物治疗全国重点实验室, 成都 610041

**摘要:** 药物设计学是一门以药物化学为基石的新兴学科, 它作为一门综合性交叉学科在综合素质药学人才培养过程中有着举足轻重的作用。针对药物设计学学科跨度大、知识覆盖面广、晦涩难懂、课本案例陈旧等特点, 我们以计算机辅助药物设计章节为例, 提出了构建新药研发前沿案例库的教学方法, 期望提高药物设计学教学的趣味性、创新性和高阶性。

**关键词:** 药物设计学; 计算机辅助药物设计; AI辅助药物研发; 案例教学

**中图分类号:** G64; O6

## Application of Cutting-Edge Case Studies in Drug Development to Drug Designology: A Focus on the Chapter of Computer-Aided Drug Design

Guan Wang, Xin Wang, Wen Shuai\*, Liang Ouyang\*

State Key Laboratory of Biotherapy, Sichuan University, Chengdu 610041, China.

**Abstract:** Drug designology, an emerging discipline rooted in medicinal chemistry, plays a pivotal role as an interdisciplinary subject in the holistic training of pharmaceutical professionals. Given the extensive scope, wide-ranging knowledge base, complexity, and outdated case studies in the current curriculum, this paper exemplifies the Computer-Aided Drug Design (CADD) section to introduce an innovative pedagogical approach, with the establishment of cutting-edge case database in new drug development. This approach aims to enhance the engagement, innovation and sophistication in the teaching process of drug designology.

**Key Words:** Drug designology; Computer-aided drug design; Artificial intelligence-driven drug design; Case-based teaching

2020年9月, 习近平总书记曾强调, “要把人民健康放在优先发展战略地位, 努力全方位全周期保障人民健康, 加快建立完善制度体系, 保障公共卫生安全, 加快形成有利于健康的生活方式、生产方式、经济社会发展模式和治理模式, 实现健康和经济社会良性协调发展。”十四五以来, 我国医药产业合理布局及创新资源配置得到里程碑式的突破, 摘掉了以往医药产业原料、仿制药、低水平重复等负面标签, 保障健康中国战略的顺利实施。医药教育是实施健康中国战略的重要基石之一, 面对时代赋予的机遇与挑战, 如何在医药领域大改革的背景下培养出顺应时代发展的高素质医药人

收稿: 2024-04-02; 录用: 2024-07-04; 网络发表: 2024-09-26

\*通讯作者, Emails: ouyangliang@scu.edu.cn (欧阳亮); shuaiwen05@163.com (帅雯)

基金资助: 国家自然科学基金(82273770, 22177083); 四川省自然科学基金(2023NSFSC1839); 1·3·5 卓越学科项目; 四川大学华西医院临床研究孵化项目(ZYJC21016)

才是各大高校需要思考的问题。

自21世纪以来, 药物化学得到了突飞猛进的发展。与药物化学相关的基础学科飞速发展, 理论知识日趋丰富, 使药物化学发展成为有雄厚理论基础的成熟学科。伴随着计算机辅助药物设计、组合化学、高通量筛选等技术的出现和药物化学学科知识的完善, 药物设计的面貌也发生了极大改变。药物设计学在1990年由复旦大学率先创建, 国内许多重点高校, 如四川大学、中国药科大学、北京大学、武汉大学及山东大学等均有开设本课程<sup>[1]</sup>。药物设计学涉及科目多、知识覆盖面广、内容抽象、枯燥难以理解, 学习效果不佳。目前药物设计学教学设计存在的问题有: ① 教材内容多, 涉及知识面广, 难以把握学习重点; ② 教材内容更侧重于基础知识, 缺乏实际应用; ③ 教学方法单一, 未能有效激发学生学习的积极性。

案例教学创新性强、时效性高, 有助于调动学生学习主动性, 提高学生的跨学科思考能力, 提高学生的综合素质并且有效缩短教材与科学前沿的差距。本文搜集了大量新药研发前沿案例以构建教学案例库, 帮助学生深入了解新药研发现状, 了解熟悉新药研发过程中的新技术和新策略, 有助于提升药物设计学教学的趣味性、创新性和时效性, 对培养具有创新意识和全面综合素质的药学人才具有重要的教育价值。在本文中, 笔者结合自身教学和科研经验, 以计算机辅助药物设计章节为例, 将基础知识、前沿案例和思政教育有机结合, 构建新药研发前沿教学案例库, 对药物设计学的教学改革进行了深入探讨。

## 1 计算机辅助药物设计(CADD)教学案例设计

计算机辅助药物设计(Computer-aided drug design, CADD)以计算机作为操作媒介, 利用计算化学、计算生物学、分子图形学、数理统计以及数据库等技术, 研究药物和受体的相互作用, 以期发现、设计和优化创新药物分子的方法学集合, 其技术核心在于分子模拟, 研究目的在于合理药物设计<sup>[2]</sup>。CADD已渗透到现代药物研发的各个环节, 如新靶点识别, 苗头化合物的发现, 药物-靶点相互作用解析, 苗头分子结构优化与改造等。在药物合理设计过程中, 许多药物设计新策略不断涌现, 如: 定量构效关系(quantitative structure activity relationship, QSAR)、药效团建模(pharmacophore model)、分子对接(molecular docking)、分子动力学模拟(molecular dynamics simulation)等(图1)。CADD大大提高了药物研发的成功率, 减少药物开发的成本, 缩短药物研发周期, 在药物合理设计方面具有极大的潜力。



图1 计算机辅助药物设计流程图

### 1.1 CADD应用于先导化合物的发现

高通量自动化药物筛选技术(**high-throughput-screening**, HTS)的出现推动先导化合物的快速发现。运用HTS技术可以在短时间内筛选大量化合物,是传统药物发现策略不能匹敌的。组合化学最初由R. B. Merrifield于1963年提出,是一种高效低耗合成大量化合物的方法,为满足高通量筛选对大量化合物的需求应运而生。有机小分子库和类药库的合成与高通量筛选技术结合,在先导化合物的发现过程中起着重要的作用。组合化学涉及到数学、生物化学、有机化学、药物化学等多门学科,如何将这这些学科结合起来,提高药物筛选的效率、降低经济成本,找到高效低毒的药物是未来发展的方向。

案例设计:阿哌沙班是由杜邦-默克公司开发的凝血因子Xa抑制剂,临床应用于患者术后静脉血栓栓塞症的预防(图2)。得益于阿哌沙班出色的“**best-in-class**”属性,阿哌沙班在2021年一举成为全球小分子销售冠军。杜邦-默克公司在GPIIb/IIIa拮抗剂中进行高通量筛选,得到具有单脒结构的苗头化合物(1)。当时的主流结构-双芳脒结构,研发人员大胆创新,将其改为联苯基团,抑制活性一举提高到纳摩尔级别。通过分子对接模拟,异恶唑基团和吡啶基团与FXa结合,因此,进一步将骨架优化为吡啶结构,这成为开发FXa抑制剂过程中里程碑式的发现。在此基础上,通过基于结构的药物设计(**structure-based drug design**, SBDD),考虑活性、药代动力学性质、安全性等方面,对结构进行全面、系统的优化,最终得到了阿哌沙班。阿哌沙班的开发过程中,研发人员精益求精有大量的工作积累。首先通过高通量筛选技术得到一个苗头化合物,通过SBDD对结构不断进行优化,并且研究人员基于计算机辅助药物设计大胆的进行了骨架结构的创新,最终得到了活性好、安全性高、结构新颖的阿哌沙班<sup>[3-6]</sup>。创新药开发的道路充满荆棘,通过本案例引导学生学习精益求精、勇于创新、不怕吃苦的科研精神。

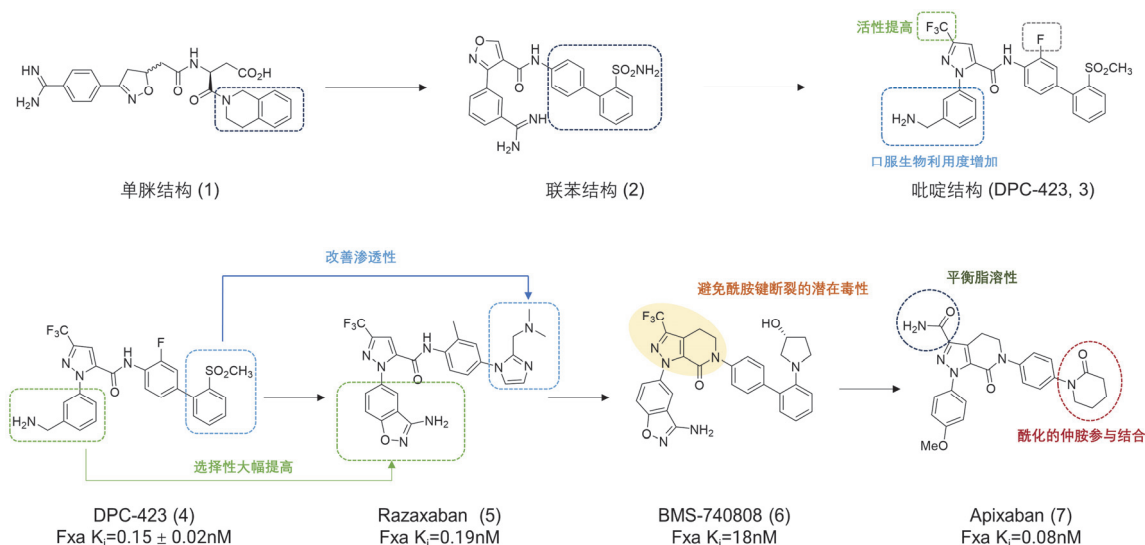


图2 阿哌沙班的开发过程

药物再利用(**Drug repurposing**)又称为药物再定位,是指对已获批准的药物进行再利用的方法,大大缩短了药物研发的周期并降低风险。许多药物可能具有多个靶点,不同的靶点会引起不同的作用。比如,被称为“神药”的二甲双胍是治疗II型糖尿病的口服降糖药,在全球范围内使用广泛,主要通过抑制肝糖输出发挥调控血糖的作用。近年来,二甲双胍更多新作用被陆续发现,如,治疗多囊卵巢综合征、保护心血管、抗肿瘤、抗衰老、减肥等作用,奠定了二甲双胍的“神药”地位。药物和疾病是与药物再利用相关的两个基本组成部分,就药物设计而言可以概括为九种关联网路:基

因调节网络、代谢网络、蛋白质-蛋白质网络、药物-靶点、药物-药物-药物、药物-疾病、靶点-疾病、药物不良反应和疾病-疾病网络。合并多个网络对药物再利用与新靶点预测结合至关重要。

案例设计：DTINet算法是基于网络的深度学习方法，用于识别已知药物的新分子靶点。DTINet专注于学习特征的低维向量表示，准确解释异构网络中各个节点的拓扑特性，然后通过向量空间投影方案基于这些表示进行预测。这种方法认为阿仑膦酸钠、氯丙酰胺和替米沙坦具有新的环氧化酶阻断作用。随后，通过评估促炎症成分的生成情况，证明了这些已鉴定的环氧化酶抑制剂在预防炎症性疾病方面的新潜在应用。查阅相关文献，简述基于药物-靶标预测手段的药物再利用案例<sup>[7]</sup>。(应用场景：课后扩展)

### 1.2 CADD加速Hit-to-lead优化过程

定量结构-活性关系(quantitative-activity relationship, QSAR)建模是一种成熟的化学数据分析计算方法，已被广泛应用于CADD。QSAR模型是通过根据分子结构计算出的化学描述符值与实验测得的这些分子的特性或生物活性之间建立经验、线性或非线性关系。然后应用这些模型来预测或设计所需特性的化学物质<sup>[8]</sup>。自1964年，Wilson、Hansch和Fujita首次提出QSAR的概念以来，QSAR的概念快速发展，相继推出了2D-QSAR方法、3D-QSAR方法以及多维(*n*D)-QSAR方法<sup>[9]</sup>。1964年，Hansch等人提出了著名的Hansch方程，该方程巧妙地利用线性回归技术描述二维构效关系，开启了QSAR评估的时代。同年，新提出了一种描述化学结构与生物活性之间关联的新技术——Free-Wilson技术。最近，Subramanian G.等人描述了一种利用深度神经网络(DNN)方法预测Hbace-1阻滞剂的研究，DNN方法表现出良好的回归能力。定量-构效关系已经成为药物开发中强大而有效的工具。我们相信，随着算法的不断进步，QSAR算法将在先导物优化的过程中发挥越来越重要的作用。

案例设计：QSAR将化合物结构变化与活性变化相关联，可用于新型候选药物的设计，药物活性及毒性预测以及阐明化合物与生物分子之间相互作用机制。请查阅相关文献，选定一个自己感兴趣的疾病靶点，通过QSAR模型设计一个针对该靶点的小分子化合物。(应用场景：课后扩展)

### 1.3 CADD用于从头药物设计

过去几年中，从头药物设计已广泛用于药物合成。每个潜在的候选药物需要满足许多标准，包括ADMET(吸收、分布、代谢、排泄和毒性)性质以及合成可行性。因此高精度预测分子特征在药物开发过程中至关重要。人工智能的出现取代了传统的从头药物设计。Bartosz Grzybowski和他的团队制造了一款合成软件——Chematica，它可以用于设计制定多种合成路线，并且可以全面评估包括成本、原料易得程度、反应步骤难易、反应操作难度等方面，最终决策出最优合成路线。后来Evgeny Putin等人提出一种名为RANC(强化对抗神经计算机)的原始深度神经网络(DNN)架构，用于基于生成对抗网络(GAN)范式和强化学习(RL)的新型小分子有机结构从头设计<sup>[10]</sup>。RANC被认为是开发对不同生物靶点或途径新型分子的起点。

案例设计：Bartosz Grzybowski教授及其团队开发了一款基于CADD的合成软件——Chematica。Chematica的数据库拥有约700多万个分子，通过有机反应将这些分子彼此连接形成化学网络。该研究团队通过手动录入超5万个有机反应规则，让Chematica适应小分子在反应中出现的各种变化。使用者只需要将目标分子结构式输入至软件中，Chematica就可以短时间内设计出最优合成路线(图3A)。奥美拉唑是第一代质子泵抑制剂，它可以有效抑制胃酸分泌快速治疗消化道溃疡。由于奥美拉唑疗效好、安全性高、复发率低和服用方便的优势，使它一经上市便得到广泛应用(图3B)。以奥美拉唑的分子结构为例，用Chematica设计多条合成路线，分析不同合成路线的优势和局限性。(应用场景：课后扩展)

### 1.4 CADD预测化学分子的类药性

不良的药物代谢动力学和动物毒性(即ADMET性质)是药物临床试验失败的重要原因。ADMET性质可以反映药物在生物体内吸收、分布、代谢、排泄的规律。评估药物早期ADMET性质可以提高药物安全性，提高药物研发的成功率，降低药物的开发的时间成本及经济成本，加快药物临床转化

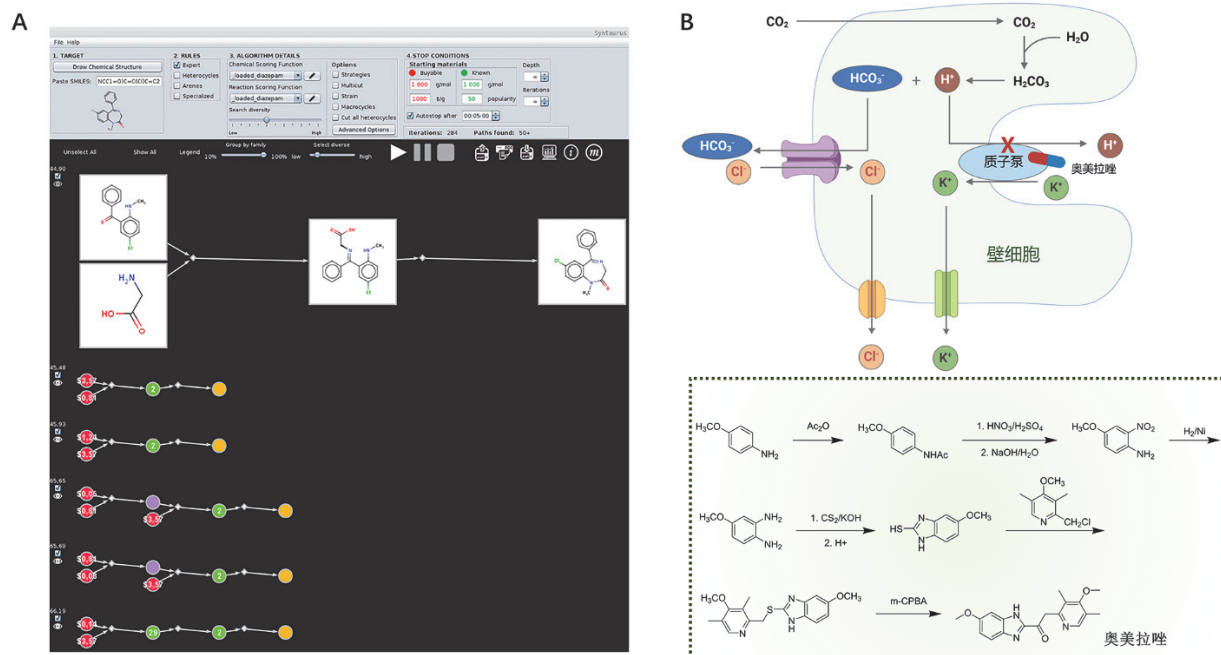


图3 Chematica的工作界面和奥美拉唑作用机制(A)及其经典合成路线(B)

的速度。由此可见,理论预测ADMET性质具有重要的现实意义。2011年报告了一种口服生物利用度数据库。他们收集了1014个在人类中确定的分子,对它们的理化性质与口服生物利用度之间的关系进行系统评估,研究这些性质对口服生物利用度之间的影响<sup>[11]</sup>。Franco Lombardo等人介绍了一种由1096种化合物组成的训练集上生成的三种稳态分布容积(VDss)计算机模型。但是该模型中只有大概约50%的分子在两倍误差之内。因此,仅从分子结构数据预测VDss值是有问题的<sup>[12]</sup>。Lombardo等人采用了主成分分析(principal components analysis, PCA)的方法来预测药物的主要清除途径,该算法表现出良好的判别结果,预测准确率为84%<sup>[13]</sup>。近年来,我国高校也不断推出多种ADMET性质预测工具,如上海交通大学洪亮课题组推出的Mat-ADMET,中南大学曹东升教授和浙江大学侯廷军教授联合推出的ADMETlab等。由此,可以设置一个课程作业,内容大致为:通过查阅相关文献,小组合作搜集我国高校开发的ADMET性质预测平台,并且总结不同预测平台包含的功能以及各自的优势。同时,使用这些平台预测某一药物的ADMET性质,并通过查阅文献或者数据库比较这些平台的准确性。通过这样的课后总结,可以让学生了解我国国内ADMET性质预测平台的发展现状,让学生主动思考这项技术目前的优势以及需要克服的问题。小组合作的方式还可以培养学生的团队合作能力。

## 2 学科前沿——人工智能(Artificial intelligence, AI)赋能新药研发

### 2.1 AI预测药物靶点结构

大多数疾病都与功能失调的蛋白质有关。通过研究蛋白质结构可以采用基于结构的药物设计来设计特定蛋白质靶标活性小分子化合物。计算蛋白质三维结构需要大量的时间和财力投入,深度学习(Deep learning, DL)的出现为预测蛋白质结构带来了便利。DL可以通过属性提取完美学习序列与结构之间的关系,预测蛋白质的二级结构、扭转角和残基接触。但是,准确预测三维结构仍是尚未实现的目标<sup>[14]</sup>。

案例设计: AlphaFold蛋白质结构数据库(AlphaFold protein structure database, AlphaFold DB)是一个可公开访问的广泛数据库,用于高精度蛋白质预测。AlphaFold以空前的准确性和速度创建大规

模结构预测的广泛数据库,这使得生物学家可以获得几乎任何蛋白的结构<sup>[15]</sup>。AlphaFold的出世即带来便利,但是也为结构分子生物学领域的科学家们蒙上了一层乌云。我们不禁开始思考一个问题:AI的出现会终结结构分子生物学?通过查阅相关资料,提出自己对这个问题的见解。(应用场景:课堂讨论)

## 2.2 AI用于药物研发全过程

截至2022年,已确定的成功药物靶点不到500个,仅占人类可药用靶点的一小部分。2009–2018年期间临床试验的平均失败率高达84.6%,主要原因是没有表现出良好的临床疗效,而适宜的靶标是提高临床试验成功率的关键。由于典型的基于实验的靶标识别即费力又消耗资源,因此计算方法已成为实现高效靶标筛选的有前途的替代方案。根据蛋白质结构的可用性和目标化合物的化学结构、药效团筛选、反向对接和结构相似性评估已被用于预测小分子的新生物学靶点。比如,Microsoft的BioGPT和Insilico Medicine的ChatPandaGPT,可以连接疾病、基因和生物过程,以快速识别疾病发展和进展中涉及的生物学机制,以及识别潜在的药物靶点和生物标志物(图4)<sup>[16]</sup>。但是,在人类生成的文本上训练的大型语言模型无法确定输入数据的准确性和适当性。此外,这些模型严重依赖于已发表的数据,在识别真正新颖的靶点方面潜力有限。

### AI在药物发现中的应用

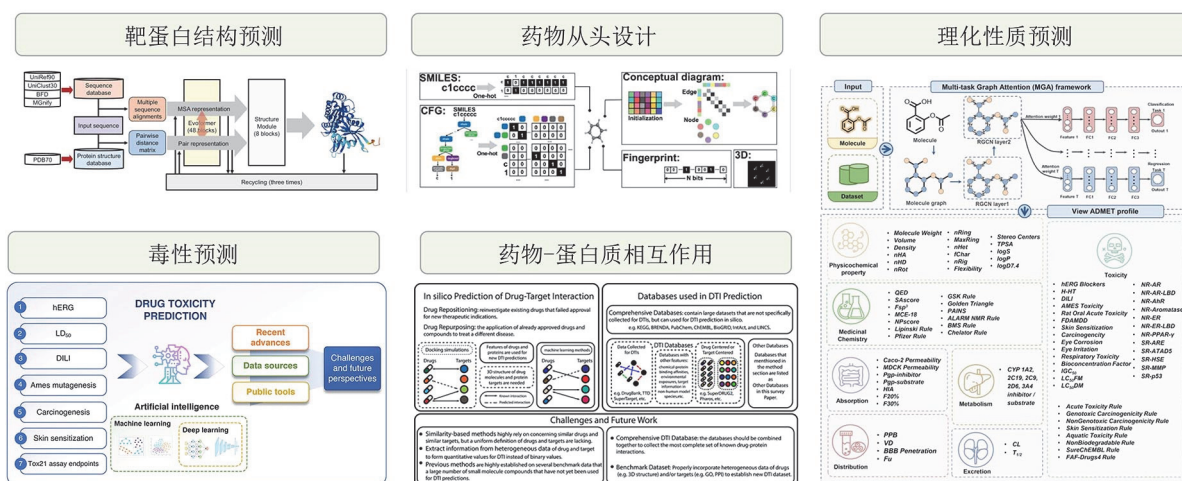


图4 人工智能在药物发现中的应用

案例设计:现在人工智能衍生药物越来越多地出现在临床研究中,例如英矽智能推出的用于治疗特发性肺纤维化的INS018\_055是第一个在I期临床试验中取得成功的人工智能衍生药物。2023年6月INS018\_055项目已成功推进至II期临床试验阶段。在INS018\_055的开发过程中,从靶点发现到临床前候选化合物(PCC)提名仅花费了18个月,远远快于传统药物发现过程。该公司的自有人工智能药物发现平台Pharm.AI包括三大部分: PandaOmics、Chemistry42和inClinico。PandaOmics采用自然语言处理模型和算法开发选项,生成知识图谱体现靶点和疾病的关联,已找到最佳的潜在治疗靶点。Chemistry42能基于生成式人工智能搭建从新颖靶点识别到分子从头生成的 workflow。inClinico可以实现临床试验结果预测和方案设计优化。请查阅相关资料,调研还有哪些人工智能衍生药物,并简述其用到的技术手段。(应用场景:课后扩展)

## 3 将思政教育融入课程设计中

习近平总书记在全国高校思想政治工作会议上强调,高校教师要坚持教育者先受教育,努力成

为先进思想文化的传播者、党执政的坚定支持者,更好担起学生健康成长指导者和引路人的责任<sup>[17]</sup>。为了响应国家战略需求,配合推进高等教育改革,改善专业教育和思政教育两张皮的现状,笔者在进行设计教学内容时,深度挖掘教学案例中的思政元素,通过现实的药物开发案例,培养学生的爱国主义精神、职业道德、科学精神、人文关怀精神。除此之外,通过授课教师言传身教,向学生传达以理性信念、实证方法、批判态度、试错模式为内涵的科学精神,真正做到春风化雨,润物无声(表1)。

**表1 教学案例中思政元素的挖掘**

知识点	课程思政元素
高通量筛选技术	通过介绍阿哌沙班的开发过程,教育学生作为药学工作者应该具有精益求精,大胆创新,不怕吃苦的科学精神
定量结构活性关系	通过介绍QSAR的发展历程,教育学生技术创新推动科学发展
ADMET性质预测	通过让学生搜索了解国内开发的ADMET预测平台,让学生感受科学精神的爱国主义底色
AI应用于药物开发	通过让学生思考AI用于药物开发的优点和局限性,教育学生科学批判精神和伦理道德问题

### 参 考 文 献

- [1] 黄增琼. 药学教育, **2019**, No. 1, 47.
- [2] 陈凯先, 蒋华良. 计算机辅助药物设计: 原理、方法及应用. 上海: 上海科学技术出版社, 2000.
- [3] Pinto, D. J.; Orwat, M. J. *J. Med. Chem.* **2001**, 44 (4), 566.
- [4] Quan, M. L.; Lam, P. Y. *J. Med. Chem.* **2005**, 48 (6), 1729.
- [5] Pinto, D. J.; Orwat, M. J. *J. Med. Chem.* **2007**, 50 (22), 5339.
- [6] Pinto, D. J.; Orwat, M. J. *Bioorg. Med. Chem. Lett.* **2006**, 16 (15), 4141.
- [7] Luo, Y.; Zhao, X. *Nat. Commun.* **2017**, 8 (1), 573.
- [8] Muratov, E. N.; Bajorath, J. *Chem. Soc. Rev.* **2020**, 49 (11), 3525.
- [9] Myint, K. Z.; Xie, X. Q. *Int. J. Mol. Sci.* **2010**, 11 (10), 3846.
- [10] Putin, E.; Asadulaev, A. *J. Chem. Inf. Model.* **2018**, 58 (6), 1194.
- [11] Tian, S.; Li, Y.; Wang, J. *Mol. Pharm.* **2011**, 8 (3), 841.
- [12] Lombardo, F.; Jing, Y. *J. Chem. Inf. Model.* **2016**, 56 (10), 2042.
- [13] Lombardo, F.; Obach, R. S. *J. Med. Chem.* **2014**, 57 (10), 4397.
- [14] Wang, S.; Sun, S. *PLoS Comput. Biol.* **2017**, 13 (1), e1005324.
- [15] Varadi, M.; Anyango, S. *Nucleic Acids Res.* **2022**, 50 (D1), D439.
- [16] Pun, F. W.; Ozerov, I. V. *Trends Pharmacol. Sci.* **2023**, 44 (9), 561.
- [17] 吴晶, 胡浩. 中国高等教育, **2016**, No. 24, 5.