

## 打破“双十”魔咒 助力产业变革

## AI为药物研发按下“快进键”

深瞳工作室出品  
采访：本报记者 张佳星  
策划：刘恕 李坤

不久前，国家卫生健康委、国家中医药局、国家药监局联合发布《卫生健康行业人工智能应用场景参考指引》，列出药物研发等84种具体应用场景，为人工智能(AI)技术赋能药物研发按下了“快进键”。

长期以来，在药物研发行业存在一个著名的“双十”魔咒，即新药研发需要花费10年时间、10亿美元。如何打破这一魔咒，AI被寄予厚望。在日前召开的学术会议上，中国科学院院士陈凯先表示，AI将为半导体行业带来5550亿美元的价值，而为制药行业带来的价值可达1.2万亿美元。

近年来，生成式AI在垂直产业持续发力。它给新药创制带来什么？记者了解到，一系列AI技术的应用和平台建设正不断提高药物研发效率，人工智能技术或将引发医药产业的颠覆性变革。

## 提升研发效率

在药物研发全链条发挥作用

2023年底，困扰业界60年的新抗生素发现难题被AI破题。《自然》杂志刊登了科学家运用AI技术首次发现抗耐甲氧西林金黄色葡萄球菌(MRSA)的新抗生素的历程。

以3.9万种化合物对MRSA的抗菌活性数据作为训练“脚本”，麻省理工学院研究团队获得了抗菌能力的评估预测模型。随后以3个深度学习模型为基础，团队又“塑造”出化合物人类细胞毒性的“鉴定师”，对1200万种化合物进行“筛选”，最终获得能对抗MRSA又对人体安全的化合物。

这样的实验，仅靠人力几乎难以完成。然而，AI凭借“硬核”实力，大大缩短了对每一个分子进行评价和验证的时间。

以往，针对靶点设计药物的成功率非常低。数据显示，即便是《细胞》《自然》等权威期刊上报道过的原始靶点，其成药几率也小于10%。

靶点有了，为什么设计药物还是难？以小分子化学药物为例，依据“钥匙开锁”的原理，针对靶点设计化合物，化合物库中的备选分子多达成百上千种，筛选耗时耗力。

“新药研究实践表明，要在老靶点上寻找更好的新药越来越不容易。”陈凯先说，与此同时，新靶点发现难度也越来越大，需要新思路新技术“破局”。

而人工智能在新靶点的发现和预测方

面能为人类提供前所未有的帮助。

陈凯先介绍，国外研究数据显示，人工智能技术应用可以使药物设计时间缩短70%、药物设计成功率提升10倍。

“理论上AI可以对药物研发的全链条发挥作用。”陈凯先告诉记者，在新药研发的整个链条中，一个新靶点的发现往往会带动一批新药产生，推动临床治疗的突破。

“我国尤其需要抓住AI助力医药原创发展的机遇期。”陈凯先表示，相关数据显示，过去十几年，我国每年发现的潜在药物靶标数量始终徘徊在6个以下，我国新药研究仍以追赶、跟随为主。

近年来，人工智能技术正成为发现新靶点的利器。例如，中国科学院上海药物研究所郑明月等研究团队发展“脸谱识别”新算法，通过提取化学结构特征、基因变化特征、药物活性特征来关联比对查明新靶点。这一技术成功用于寻找抗肿瘤老药甲氨蝶呤的免疫靶点。

AI在从文献知识中发现新靶点方面也颇有优势。陈凯先表示，过往基础和临床研究积累了大量数据，这些发现“互不关联”地分散存储在研究文献中，它们之间的潜在关联难以被人类发现。人工智能具有强大、高效的学习分析能力，能够将散布在大量文献中的关联关系挖掘出来，推动新机制、新靶点的识别。

“尽管我国现在还没有AI辅助研发的新药获批上市，但已经有不少新药在AI协助下快速进入临床试验阶段。”陈凯先说。

## 降低研发成本

药物试验费用不再高昂

2024年年底，斯坦福大学等研究团队在《细胞》杂志发文称，多尺度、多模态的大型神经网络模型已经具备表示和模拟分子、细胞和组织在不同状态下行为的能力。在此基础上，AI虚拟细胞有了高保真模拟、加速发现、指导研究的可信性。

此前，《美国国家科学院院刊》发文称，研究者用硅基“病人”替代生命形式的碳基“病人”，模拟结果与真实数据高度一致。

研究中，1635个“活在电脑里”的“虚拟病人”，患上了乳腺癌且癌细胞已发生转移。通过试验，研究者找到了生物标志物指导乳腺癌临床治疗的优化路径。

基于体外、体内、临床、人群水平和多组学等数据，研究者对患者药物反应进行“数字孪生”，生成“虚拟病人”丰富的药效药理数据，用于生物标志物、药物等方面的测试。

“人类对生命活动的想象和思考，能够以数据的方式传递给算力，这是实现虚拟生命或细胞的基础。”北京大学未来技术学院院长席建忠告诉科技日报记者，在半

个世纪的发展历程中，分子生物学通过不同层面的组学数据“解读”生命，如基因组学、蛋白质组学、转录组学等，积累了大量生命科学数据。

随着技术融合发展，人类的数据获取能力越来越强。“光学成像技术现在已达到纳米级别，可以看到细胞中的细胞器动态‘录像’。”席建忠说，大量新数据、新研究衍生出成像组学等新学科。这些突破在深入解读生命的同时，也成为数字生命的基础。

事实上，我国科研团队已在基础设施、科研课题等方面早有布局。例如在北京怀柔，耗资数十亿元的多模态跨尺度生物医学成像设施已初具规模，细胞成像楼、医学成像楼、全尺度整合中心等组成具有硬实力的科技“航母”，其中，全尺度数据处理中心将为相关研究提供强大算力支持。

“不同团队正在开展一些关键器官数字化的工作。我们希望实现肿瘤细胞数字化。”席建忠表示，肿瘤具有高度异质性和动态性，人人不同、时时在变，有效的药物筛选十分困难。

“虚拟肿瘤细胞能够告诉我们，在某一药物作用下，细胞内部信号通路如何变化。”席建忠说，要实现这样的目标，需要基于现有数据和基础模型进行肿瘤细胞“雏形”的构建，然后对其进行训练。

“现实生活中，要进行药物试验。如果一位患者吃一种药，几千种药至少需要几千位患者来试验。这样一来，实施困难、成本高。”席建忠说，虚拟细胞在一套模型中可以同时“吃”几千种药，获得几千套数据，实现高通量和高保真，将极大提升肿瘤药物的筛选效率。

生成式人工智能最令研究者着迷的是它的“出其不意”。席建忠表示，在科学探索方面，AI能够打破不同领域的边界限制。例如，对心血管、传染病的交叉研究，可能会发现抗病毒药物具有降血压潜力之类的“惊喜”。

## 缩短研发周期

啃下罕见病药物研发“硬骨头”

在药物研发领域，罕见病药物研发是块难啃的“硬骨头”。正因为如此，药物审评审批为它专门开设了“孤儿药”绿色通道。

4年到5年，这是当前罕见病诊断的平均耗资周期。患者少形成了“无米之炊”的困局——罕见病难以“被看见”，临床试验病例少是罕见病药物研发的难题之一。

现代医学发展至今，为什么确诊一个病还要耗时这么久？

“罕见病并不会带着‘铭牌’而来。”华

大基因AI专项负责人梁伦纲告诉记者，它会被当作一种常见病，如孩子发育明显落后同龄人时，往往会问营养师。

“罕见病诊断面临症状、基因变异两端都‘开放’的问题，而确诊是要通过各种方法实现两端‘收敛’，最终得到匹配的‘连线’。”梁伦纲说，患者经过各种试错仍无法找到病因后，才有可能被确诊患有罕见病。如今当人工智能迈入生成式大模型阶段，能够回答各种开放式问题之后，罕见病诊断有望跳过“试错”阶段。

日前，北京协和医院院长张抒扬在国家卫生健康委医药领域科技创新发布会上介绍，首个罕见病人工智能大模型GeneT上线，协助基层医生做罕见病的诊疗。

从回答“是”或“否”的判断题，到能够解答“这是什么，为什么”的复杂问答题，AI提升了罕见病的诊断能力。

“和其他应用领域一样，AI先是快速掌握罕见病领域的专业知识。”梁伦纲介绍，公开罕见病数据集和文献，以及华大在检测服务中的数据都会转化为AI的“知识”。最重要的是，华大基因与北京协和医院开展合作，及时应用罕见病临床诊疗中的一线经验，让AI具备临床“经验”。

“我们不仅给AI输入了罕见病的海量知识，还教会它如何像遗传专家一样思考。”梁伦纲介绍，华大基因通过将专家的思考过程转化为AI能理解的语言，让新模型GeneT学会精准筛选导致罕见病的基因变异，效率提升20倍，且在模拟和真实病例中的准确率分别达到99%和98%。

梁伦纲表示，目前GeneT在完成初步分析后，最终确诊还需要专家把关。

资料显示，在人工智能大模型的助力下，罕见病患者的确诊时间有望从数年缩短到4周以内，这使得绝大部分罕见病无特效药的“冰山”开始松动。

数据显示，我国2017—2022年期间的在研罕见病药物数量大幅增加，年均增长率达34%。然而，《中国临床药理学杂志》的一项研究显示，约43.9%罕见病药物临床试验实际入组人数小于目标入组人数。

在国家罕见病登记系统的支持下，罕见病临床队列相继建立，以推动相关领域药物研发。“这让罕见病患者得到早诊断。”梁伦纲表示，罕见病“被看见”，将缓解罕见病药物研发临床队列稀缺的问题，为罕见病药物研发提供有力支撑。

专家认为，未来3—5年，我国将进入AI药物研发快速发展阶段，人工智能技术将肩负起分子优化、合成路线设计以及自动生成、自动分析、自动筛选的全流程研究工作。

“我们期待第一个AI设计研发的药物能够尽快获批上市。成果转化和落地应用依旧是医药产业实现高质量发展的‘命门’。”席建忠说。

## 记者手记

从被拒门外到加快应用  
AI靠实力逆袭

◎张佳星

记得两年前，某计算学者告诉我，为了让智能算法走进药物筛选领域，自己经常自掏腰包，前往药理学学术会议进行交流。他想用自己的研究结果告诉人们，过去需要数千小时的筛选工作，借助新技术几十小时就能完成。

当时，相信者寥寥。他的经历，让我脑中闪过一个词——“布道者”。新技术诞生后，不乏“布道者”。

2024年底，在我参加的多个药学相关学术会议中，人工智能、算法、模型等作为关键词在学术报告中频频被提及。在综述性报告中，一些落地见成效的成果见诸知名期刊；在学术分享报告中，不少新药物的研发已经在人工智能辅助下进入临床试验阶段。

值得一提的是，这些成功的案例多来自联合研究团队，这些学术会议的嘉宾也愈加背景多样，他们中有算法工程师、临床医学学者、药学者、遗传学者……诚然，隔行依然隔山，不同学科的学术语言体系仍然存在隔阂。为了克服跨界融合的理解障碍，每一个时代的科研者都要不断学习新的研究体系。

一位受访者的感慨，可能代表着科技工作者的心声：在AI引发的技术革命浪潮中，研究者如果不行动起来，就是在等着被淘汰。

“日益复杂的科学问题需要探索跨学科团队合作的科研模式。”中国工程院院士、北京大学第三医院研究员董丹表示，尤其是在生物医学、健康等领域，要取得变革性研究成果，必须开展“有组织的科研”。

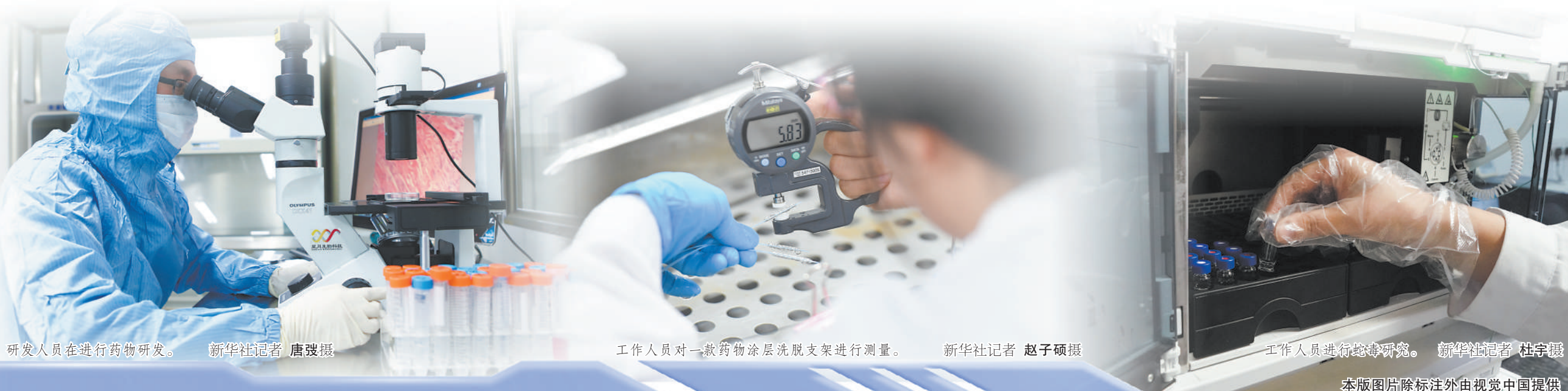
就像互联网打破了人类地域的概念，人工智能也在击碎“数据井”的壁垒。生成式人工智能能够处理的参数和数据以亿为单位。AI的能力正在倒逼不同来源的生物医学研究数据进行整合。如何提高数据的可访问性和互操作性，已经成为“有组织的科研”的课题之一。

在我们身边，医疗数据的互认、共享正在成为常态；在保障数据安全的前提下，数据“柏林墙”正在被推倒……这些，为人工智能+医药领域“有组织的科研”奠定了数据基础。

相关部门也通过政策法规积极引导。国家卫生健康委等部门日前发布的《卫生健康行业人工智能应用场景参考指引》，从4大领域给出了84个应用场景，以推进卫生健康行业“人工智能+”的应用创新发展。

如今，再同前述的数据学者聊起来，他告诉我，智能计算的应用早已开始，他们提供技术支持的罕见病药物已经在等待临床试验许可，他们还在为免疫治疗药物敏感性进行评估、为老药寻找新的适应症……

从被拒门外到被热情拥抱，AI靠实力成为药物学家的常规工具。生命数字化的研究已经在如火如荼开展，模拟生命活动的精度在持续提高，生命活动的模拟和预测将带来什么，或许下一个两年就有眉目。



研发人员在进行药物研发。新华社记者 唐茂摄

工作人员对一款药物涂层洗脱支架进行测量。新华社记者 赵子硕摄

新华社记者 赵子硕摄

工作人员进行病毒研究。新华社记者 杜宇摄