

21 天！AI 新药研发头部公司创纪录，“图片造假术”正加速救命药问世


原创：王新凯、张静 [DeepTech 深科技](#) 9 月 3 日

从选择一个靶点，到形成潜在的新药候选分子，这个过程需要多长的时间？过去的回答可能是数月，乃至数年。但在昨日，来自人工智能药物发现公司 Insilico Medicine、药明康德以及多伦多大学的科学家们在《自然》子刊 *Nature Biotechnology* 上给出了不同的答案。**他们用一种称为生成张力强化学习 (GENTRL) 的新 AI 系统，将这一数字缩短到了仅仅 21 天！**

完成新药候选分子设计和筛选后，GENTRL 系统共设计出 6 种新的 DDR1 抑制剂，并完成临床前生物学验证，整个过程耗时 46 天。**这一速度，要比传统制药公司的药物研发过程，快 15 倍。**

在 6 种新药候选分子中，4 种化合物在生化分析中具有活性，2 种化合物在体外细胞实验中展现了预期的 DDR1 抑制能力，其中 1 号化合物进入临床前的动物实验。

Deep learning enables rapid identification of potent DDR1 kinase inhibitors

Alex Zhavoronkov , Yan A. Ivanenkov, Alex Aliper, Mark S. Veselov, Vladimir A. Aladinskiy, Anastasiya V. Aladinskaya, Victor A. Terentiev, Daniil A. Polykovskiy, Maksim D. Kuznetsov, Arip Asadulaev, Yury Volkov, Artem Zholus, Rim R. Shayakhmetov, Alexander Zhebrak, Lidiya I. Minaeva, Bogdan A. Zagribelnyy, Lennart H. Lee, Richard Soll, David Madge, Li Xing, Tao Guo & Alán Aspuru-Guzik

Nature Biotechnology (2019) | [Download Citation](#) 

(来源：Nature Biotechnology)

众所周知，新药研发之路并非坦途。在高达近 9 成的失败率面前，每一款上市新药的背后，都是平均 10 多年的漫长道路，以及 20 多亿美元的高昂开支。其中，光是早期药物发现，就占去了将近一半的成本。

这项新的研究可能会促成新药研发困境的改变。“这篇论文是我们在人工智能驱动药物发现之路上的一个重要里程碑。现在，这项技术正在成为主流，当这些模型被整合到全面的药物发现管线中时，它们适用于许多靶点，我们与领先的生物技术公司合作，将进一步推动生成化学和生成生物学的极限。”论文通讯作者、Insilico Medicine 创始人兼 CEO Alex Zhavoronkov 博士表示。

Insilico Medicine 被美国知名创投研究机构 CB Insights 誉为 2018 全球人工智能百强公司。值得注意的是，药明康德与 Insilico Medicine 关系密切。后者的最近一次融资是 2018 年 6 月由药明康德风险投资基金领投的战略融资，从那以后，双方开始了紧密合作。这种亲密关系，也让 Insilico Medicine 更接近中国的新药研发，随着数千家中国制药公司从仿制药向创新药过渡，中国的新药研发预计会有爆发性增长，市场前景广阔。

从靶点到候选分子，仅用 21 天

在这次发表的文章中，研究团队选择了 DDR1 作为靶点。DDR1 是一种在上皮细胞中表达的酪氨酸激酶，与组织纤维化疾病进展密切相关。

为了寻找到潜在的 DDR1 抑制剂，研究人员们开发了一种使用了强化学习技术的新 AI 系统，初步得到大约 3 万个不同的分子结构。

具体来看，这个算法的训练用到了多个不同的数据库。其中最大的一个数据库里包含海量的分子结构，其他的数据库则分别为已知的 DDR1 抑制剂及其 3D 结构、具有激酶抑制剂活性的常见分子（作为正对照）、无法靶向激酶结构的分子（作为负对照）、以及已被医药企业申请专利的分子。

之后，研究团队根据反应基团与化学空间等指标对这些分子结构进一步筛选，并在实验室中合成了 6 个潜在新药分子。

经过初步生化测试，研究人员筛选出 4 个具有活性的潜在分子进入体外细胞实验。而在体外实验中，结果显示其中 2 个化合物对于 DDR1 具有抑制性，展现了显著的 DDR1 抑制能力，且能有效降低与纤维化进程有关的标志物。

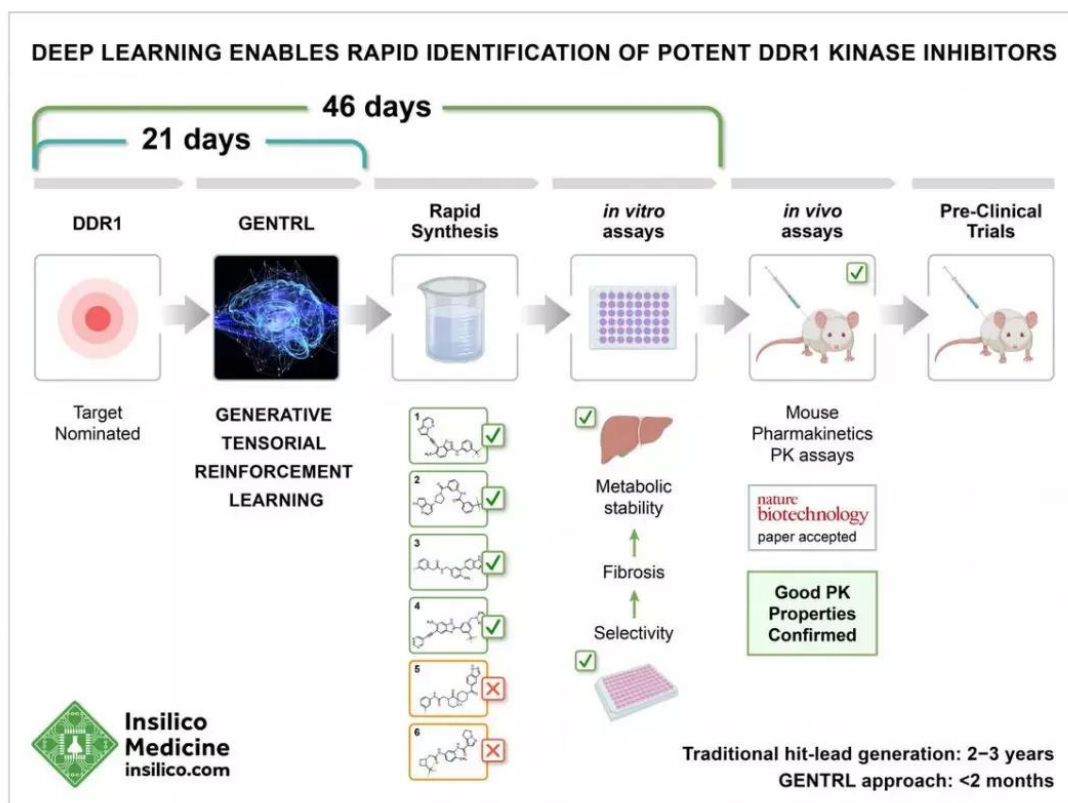


图 | AI 系统设计筛选全新 DDR1 激酶抑制剂流程 (来源: Insilico Medicine)

减少研发周期和总成本，对药物研发的未来至关重要。在这些最新成果中，Insilicon 强调了一种基于人工智能的新技术(GAN-RL)，该技术使他们能够在很短的时间内在动物模型中筛选出有效的潜在药物分子。如果这项技术被广泛证明是有用的，它很可能在未来的药物研发中具有转化潜力。” 高点大学 (High Point University) 药学院兼职教授 Stevan Djuric 说道。

从最初的靶点确定，到完成潜在新药分子结构筛选，仅用时 21 天；而到完成初步的生物学验证，用时仅 46 天。在之后的体内动物实验中，筛选出的新药分子的药代动力学特征均达到预期结果。此外，科学家们也决定公开该算法的源代码，供产业更多研发人员使用。

2013 年诺贝尔化学奖得主、斯坦福大学结构生物学教授 Michael Levitt 评价道，“这篇论文无疑是一个令人印象深刻的进步，很可能适用于药物设计中的许多其他问题。基于最先进的强化学习，我也对这项研究的广度印象深刻，因为它涉及到分子建模、亲和力测量和动物研究。”

AI 直击药物研发痛点

药物和作用靶点就如同钥匙和锁的关系。在药物研发过程中，选择和确定有效的药物靶点也是新药开发的首要任务。确定靶点之后，通过研究靶标生物大分子的结构、构效关系进行药物分子的筛选和设计，寻找适合这把锁的钥匙，至关重要。

大多数情况下，科学家只能通过不断扩大筛选对象，在自然界无数种物质中苦苦寻觅一种能有效结合靶点的“钥匙”。这一过程，即使是利用由机器自动完成的高通量筛选（HTS）进行试验，也往往耗时耗力，就如同在黑夜里蒙着眼睛开枪打靶。

在传统药物研发过程中，需要经过寻找靶点、筛选先导化合物、临床前动物实验、

临床试验等过程。一款新药的诞生，往往意味着无数科学家十多年的时间、数十亿美元的资金投入，更重要的是这一过程中失败率极高。

2015年，FDA报告了60种获批药物，每种获批药物的研发成本平均高达6.98亿美元，并且有将近420亿美元用在了失败药物研发上。

在传统药物研发过程中，对数千个小分子的测试，最终只为寻找几个有希望的候选分子，而这些分子中只有大约十分之一能够进入临床试验。不仅如此，传统药物的研发，越来越难以取得突破，这是因为大多数可以被使用的化合物已经被发现，新的化合物开发难度逐渐加大，另一方面则是由于科学成果的数量增长速度很快，人类个体不可能完全理解这些数据。



(来源：Pixabay)

而 AI 技术的出现和发展，为解决药物研发痛点，提供了新的希望。人工智能可以从海量论文中摄取所需的分子结构等信息，并且可以自主学习，建立其中的关联。

可以想象，即使在发现新药的时间上，或是成功的可能性上有一个微小的改进，也将给药企带来显著的成本优势，并为病人带来诸多福利。

微软和谷歌前高管、创新工场创始人李开复博士表示，“本文提出的生成张力强化学习方法大大提高了生物化学在药物发现中的应用效率。虽然这种方法还有待进一步的大规模实验，但已经标志着 AI 制药在产业层面的突破，可能会给我们带来重大的社会和经济影响。”

里程碑，也是新起点

去年 12 月，DeepMind 首次推出 AlphaFold，这是一种预测蛋白质折叠的算法。它轻而易举地击败了制药行业长期以来的竞争对手，尽管如此，一些专家仍然怀疑人工智能是否能够创造出既有效又真正实用的分子。

目前的市场上也还没有出现人工智能生产的药物。像 AlphaFold 这样的系统虽然炫酷，能展示很多令研究人员兴奋的前沿技术，但是却不太可能产生可以迅速转化为新药的结果。

今年 4 月，一项重要的药物 AI 项目折戟，IBM 停止销售其药物 AI 工具“沃森药物发现系统”（Watson for Drug Discovery），该系统曾设想被用来搜索医学文献和基因数据以寻找被人类忽视的治疗方法。有报道称，该系统并不能实现对其首批客户的预期交付。

尽管 AI 药物研发项目屡屡失败，但投资者仍在奋力前行。**有数据显示，去年有超过 10 亿美元的资金涌入人工智能药物研发初创企业，制药行业也开始关注人工智能药物的研发。**

斯坦福大学孵化的 Insitro 公司于今年 4 月与吉利德科学公司签署了一项协议，开发针对肝脏疾病的药物分子，包括税费在内项目价值高达 10 亿美元。同月，英国初创公司 Exscientia 宣布，该公司已生产出一种有望治疗慢性阻塞性肺病(COPD)的分子，即将与葛兰素史克(GSK)进行临床试验。

Insilico 的方法基于两种形式的人工智能：生成对抗网络（GAN，该技术最初作为一种以假乱真的图片生成技术进入大众视野）和强化学习。它通过查找过去的研究和专利，发现其中对特定的药物靶点和其他结构有效的分子结构。查找的思路是优先考虑最新的、合乎逻辑的、可以在实验室合成的分子结构。这和药物化学家在阅读文献和拼凑分子成分时所做的事情是一样的。

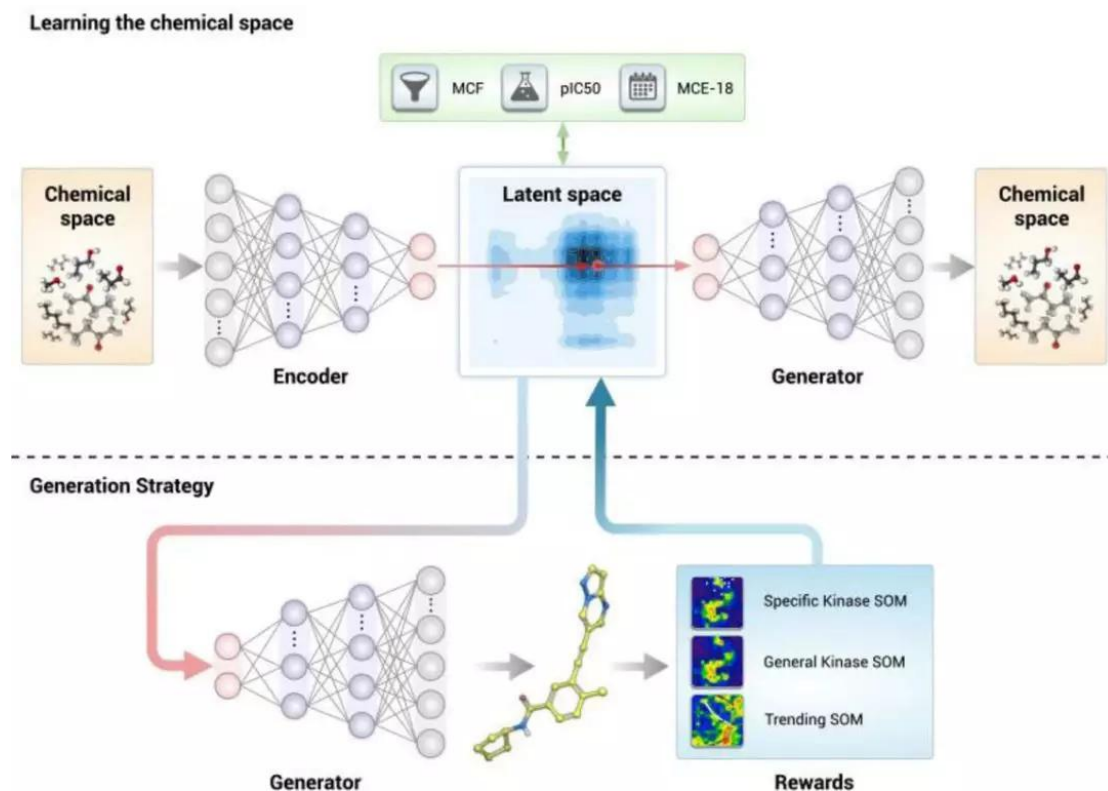


图 | 生成张力强化学习模型 (来源: Insilico Medicine)

人类基因组计划首席科学家、波士顿大学教授 Charles Cantor 表示, “关于人工智能 AI 在改善医疗保健和开发新医疗工具方面的前景, 存在着许多夸张的说法。然而, 最近发表在 *Nature Biotechnology* 上的这篇论文确实具有重要意义, 原因有二, 一是人工智能取代了通常由药物化学家扮演的角色, 二是药物开发速度的加快, 意味着专利覆盖范围的延长, 从而提高了药物开发的经济性。如果这种方法可以推广, 它将成为制药工业中广泛采用的方法。”

哈佛大学系统生物学家 Mohammed AlQuraishi 表示, 即使算法还不能战胜化学家团队, 但是这个研究证实了 AI 是可以快速挖掘出有用的线索, 并引导人类研究者去探索的。这仍然是一个起点。

-End-

参考：

<https://www.nature.com/articles/s41587-019-0224-x>

https://www.eurekalert.org/pub_releases/2019-09/im-nmd083019.php

<https://www.drugtargetreview.com/news/48404/ai-designed-synthesised-and-validated-new-drug-in-46-days/>

<https://www.wired.com/story/molecule-designed-ai-exhibits-druglike-qualities/>