

# 中国医药

## AI 有望改变创新药研发范式

我们认为，当前 AI 制药的价值主要体现在早期研发效率的提升上。据弗若斯特沙利文，AI 可将早期药物发现时间压缩 30%-40%。当前，监管框架持续完善，传统药企与 AI 原生公司加速合作，AI 制药正处于从研发提效工具迈向创新药研发范式重构的转折期。我们认为，行业真正的分水岭不在于效率验证，而在于临床成功率验证，即 AI 制药公司管线的临床数据读出能否证明 AI 不仅能提升药物研发效率，更能提升临床成功率。此外，我们认为，AI 制药的核心壁垒正从算法转向数据资产，拥有高质量私域数据与闭环迭代能力的企业有望建立竞争优势。

- AI 重构药物研发全链条，早期效率红利已率先显现。** 传统创新药研发长期受困于高成本、长周期和低成功率，AI 技术正在改变靶点发现、分子筛选与优化等关键环节的研发范式。小分子领域，生成式化学及多模态属性预测已显著压缩研发周期，如英矽智能（3696 HK，未评级）利用 Pharma.AI 平台将靶点发现到 PCC（临床前候选化合物）提名的周期从行业平均 4.5 年压缩至 18 个月。大分子领域，蛋白质结构预测与抗体从头设计正推动药物发现从传统筛选转向设计-预测-验证。与此同时，美国 FDA 与欧洲药品管理局（EMA）联合发布 AI 实践十项原则，为技术落地提供制度支撑。我们认为，AI 制药在早期研发提效方面的价值已初步得到验证，但从分子到药品上市的完整闭环尚未走通，后续关键在于临床数据读出能否支撑更深层的范式变革。
- 数据是 AI 制药的重要护城河。** 我们认为，在公域数据对所有参与者均等可及，主流算法架构通过学术论文和开源代码快速扩散的情况下，真正决定 AI 平台性能上限的是高质量、标准化且可持续迭代的私域数据。公开数据虽构成行业共同起点，但其噪声、偏倚和阴性数据缺失等问题，制约了模型的泛化能力。相比之下，药企与 CXO 积累的 SAR 过程数据、ADMET 实验数据、临床数据，以及 AI 原生公司通过自动化实验室产出的高通量标准化数据，对模型训练的价值更高。因此，以 Recursion（RXRX US，未评级）为代表的 AI 原生公司正通过自建湿实验室，构建“模型生成假设-实验验证-数据反哺模型”的闭环飞轮，形成数据壁垒。我们认为，未来 AI 制药的核心竞争可能将是数据资产与闭环能力的竞争。
- 行业估值逻辑已从平台叙事切换至临床与 BD 兑现。** 全球尚无 AI 生成药物完成关键 III 期临床验证，行业仍处概念验证阶段，市场关注焦点正由平台能力本身转向管线资产、数据读出和商业化兑现。尽管 2025 年以来 AI 相关医药交易总金额快速增长，传统药企合作热情显著提升，但首付款占比仍较低，反映出传统药企对 AI 平台仍持审慎态度。此外，多家海外 AI 制药企业曾经历了管线的挫折，也说明临床结果仍是检验平台价值的金标准。2026 年起，多家全球 AI 制药公司的核心管线将迎来数据读出，若能够持续验证 AI 在临床获益，AI 制药行业有望从概念验证迈向价值兑现。

优于大市  
(维持)

中国医药行业

武煜, CFA

(852) 3900 0842  
jillwu@cmbi.com.hk

王云逸

(852) 3916 1729  
cathywang@cmbi.com.hk

## 目录

<b>AI 技术将带来创新药物研发的工业革命 .....</b>	<b>3</b>
跨越制药界的“反摩尔定律”.....	3
AI 在早期药物发现与靶点识别中的应用 .....	3
小分子药物研发：生成式化学与多模态属性预测引擎 .....	4
大分子与生物药研发：从蛋白质结构预测到抗体设计 .....	9
临床试验的 AI 重构：数字孪生、患者分层与成功率预测 .....	14
<b>FDA 积极推进 AI 监管体系建设.....</b>	<b>17</b>
FDA 的 AI 内生能力建设.....	17
基于风险的可信度评估框架：FDA 发布决策支持指南草案.....	18
监管共识：FDA 与 EMA 联合发布十项指导原则.....	18
临床前评价演变：新方法论（NAMs）加快应用 .....	19
临床试验的数字化重构：AI 辅助提升效率，监管保持开放谨慎.....	20
<b>AI 制药行业核心壁垒在于数据.....</b>	<b>20</b>
<b>AI 制药行业发展：传统药企的积极拥抱与 AI 制药公司的崛起 .....</b>	<b>22</b>
AI 制药公司商业模式的迭代 .....	27
全球资本从狂热走向理性验证 .....	28
<b>中国 AI 制药行业：从追赶到引领 .....</b>	<b>30</b>
<b>AI 制药代表企业 .....</b>	<b>30</b>
Recursion Pharmaceuticals (RXRX US, 未评级).....	30
Schrödinger, Inc. (SDGR US, 未评级).....	35
Absci (ABSI US, 未评级).....	37
Generate Biomedicines (GENB US, 未评级).....	39
英矽智能 (3696 HK, 未评级) .....	41
<b>估值表 .....</b>	<b>44</b>

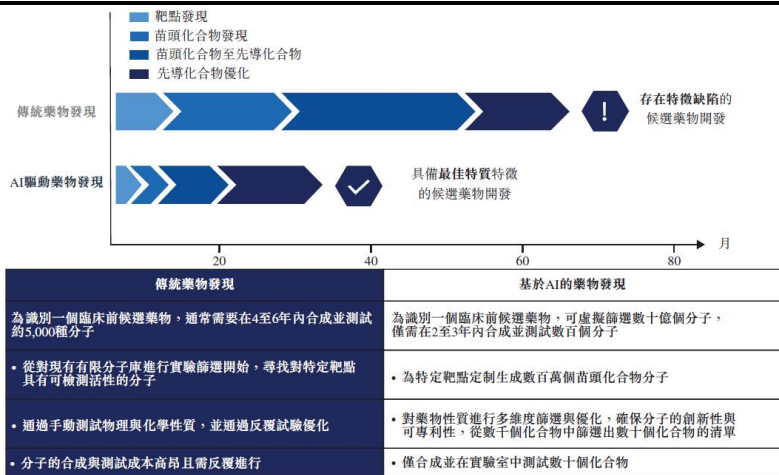
## AI 技术将带来创新药物研发的工业革命

### 跨越制药界的“反摩尔定律”

全球医药产业长期面临“反摩尔定律”（Eroom's Law）的效率困境：尽管基因组学、高通量筛选和结构生物学等底层技术持续进步，但单款创新药物的研发成本仍在持续攀升，获批上市的新药数量并未实现成比例的增长。在传统研发模式下，一款新药从早期发现到商业化上市通常需要10-15年，平均研发投入超20亿美元，且超过90%的候选药物进入临床后因疗效不足或安全性问题而失败。我们认为，传统模式的效率瓶颈集中于三个关键环节：1) 靶点发现与验证阶段，研发人员依赖组学数据与文献识别疾病相关靶点，但验证周期长、失败率高；2) 先导化合物发现阶段，高通量筛选虽可在短期内检测数十万化合物，但其覆盖的化学空间相对于理论可成药空间（约 $10^{60}$ 量级）仍极为有限；3) 临床开发阶段，患者招募困难、终点设计欠优、生物标志物缺失等问题进一步拉长开发周期并增加失败风险。

AI 与机器学习技术正从多个维度重构创新药研发流程。据弗若斯特沙利文，AI 可将早期药物发现阶段的研发周期压缩30%-40%。Absci (ABSI US, 未评级) 管理层表示，在早期发现至临床前阶段，AI 平均可节约~60%的研发成本。在临床成功率方面，据波士顿咨询发表在《Drug Discovery Today》的研究统计，传统方法发现的候选药物I期临床成功率通常在40%-65%，而AI驱动的候选药物同阶段成功率可达80%-90%；整体临床成功概率有望从5%-10%提升至9%-18%。我们认为，上述数据表明AI的价值已不局限于效率提升，而是开始触及研发成功率这一更核心的行业痛点，制药行业正逐步从AI赋能（AI-Enabled）向AI原生（AI-Native）的研发架构演进。

图 1: 基于 AI 与传统药物发现之间的效率比较



资料来源：弗若斯特沙利文，英矽智能招股书，招银国际环球市场

### AI 在早期药物发现与靶点识别中的应用

药物研发的起点是确定一个与疾病因果相关且可被药物调控的生物学靶点。传统的靶点发现高度依赖于基础生物学研究的积累，需进行耗时的体外细胞实验和体内动物实验，且往往局限于已知的生物学通路。从疾病关联基因的识别到靶点功能验证，通常需要数年的实验周期。AI 技术通过对海量多模态数据的降维与模式识别，加速实现对疾病机制的深度洞察。

## 知识图谱构建与多组学数据整合

传统研发模式中，科学家面临着信息孤岛，而自然语言处理（NLP）和计算机视觉技术，能够自主阅读并处理大量科学文献、专利、电子病历（EHR）和临床试验数据，并将这些碎片化的信息拼接成庞大的生物医学知识图谱。这些图谱不仅能够将疾病、基因、蛋白质、化合物、表型等实体及其关系进行结构化表征，还能利用图推理算法挖掘隐含但尚未被文献直接证实的潜在“靶标-疾病”关联，大幅减少早期研究所需时间。此外，深度学习模型能够从基因组学（如全基因组关联研究（GWAS））、转录组学（如转录组差异表达分析）和蛋白质组学（蛋白质互作网络）等多组学数据中识别出患者人群的关键差异表达分子和关键信号通路。

以英矽智能及其核心的 Pharma.ai 为例，其内部的 PandaOmics 应用程序平台将经典的生物信息学方法与其基于 iPANDA 算法的信号通路分析相结合，实现靶点发现。PandaOmics 4.0（2023 年 11 月 24 日发布）进一步整合了大语言模型（LLM），推出了自动化基因-疾病报告系统（Gene-Disease Report）。该系统不仅评估靶点与疾病之间的机制联系（Mechanistic Link），还可基于多模态组学、文献、财务指标和专家见解等在内的多种数据，从靶点与目标疾病的关联性、可成药性以及商业管线竞争力等维度进行量化评分和适应症优先级排序，从而大幅缩减靶点发现所需时间。比如，利用 PandaOmics，英矽智能仅用 18 个月便完成了 TNIK 这一纤维化相关的新型机制靶点发现到临床前候选化合物提名的全过程，远快于约 4.5 年的药物研发行业平均速度。

但目前阶段，AI 并不能完全替代湿实验，而更多是作为辅助工具通过缩小筛选范围，缩减湿实验所需时间。因为 AI 方法的本质是基于关联性而非因果性进行推理，即识别出与疾病高度关联的基因或蛋白质，因此仍需进行基因敲除/敲入实验、动物模型等湿实验手段来对该靶点进行因果验证，即证明调控该靶点能够产生治疗效果。

## 攻克“不可成药”靶点与表型组学筛选

在传统的结构生物学认知中，许多缺乏明确深层结合口袋的蛋白质被视为不可成药靶点。AI 技术，特别是基于高级图神经网络（GNNs）和分子动力学模拟的图表征学习，能够精准预测蛋白质的动态构象变化和隐藏的变构位点（Allosteric sites），从而为这些传统意义上的禁区重新打开成药之门。

与此同时，以 Recursion Pharmaceuticals 为代表的 AI 原生公司开创了另一条不依赖单一靶点假设的表型组学路径。该公司通过其自动化湿实验室，每周利用高通量机器人和计算机视觉技术捕获数百万次细胞实验的高内涵图像。深度学习模型对这些细胞在疾病状态下以及受化合物干预后的细微形态学变化进行学习，能够在不预先假设特定靶点的情况下，直接筛选出能够将患病细胞表型逆转为健康状态的化合物。这种方法不仅有望加速全新生物学机制的发现，还通过表型匹配提升了老药新用的成功率。

## 小分子药物研发：生成式化学与多模态属性预测引擎

在确定靶点后，便进入药物发现阶段，即寻找或合成能有效调控靶点功能的化学分子，再通过系统优化形成具有临床开发价值的临床前候选药物（PCC）。传统小分子药物发现的主要方法之一是高通量筛选（HTS），即从已有化合物库中逐一检测每个分子对靶蛋白的结合活性。然而，这一方法存在根本性局限：理论上满足类药性规则的小分子化学空间约为  $10^{60}$  量级，而人类已合成的化合物仅约  $10^8$  量级，这意味着高通量筛选仅能覆盖有限的化学空间，大量潜在的优质药物分子从未被探索过。

AI 技术在小分子药物发现阶段的应用主要可分为两类：

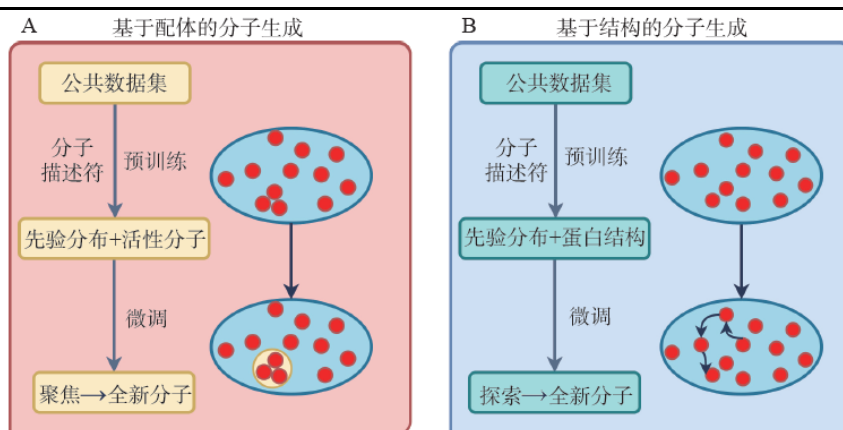
1) 虚拟筛选：利用 AI 模型，通过计算分子间的相互作用来预测化合物的生物活性，实现对大规模虚拟化合物库的快速评分与筛选，从而大幅压缩需湿实验验证的化合物数量，加速候选药物的发现过程。

2) 分子设计：通过深度生成式人工智能（Deep generative AI），实现小分子研发从“已知化合物筛选”向“全新分子生成”的转变。分子从头生成（de novo molecular generation）是通过模型从已知分子数据中学习化学空间的底层规律，基于靶点的蛋白结构，创造性地生成全新分子结构。而基于片段修饰的分子生成和先导化合物的优化则是分子设计阶段的延续。AI 性质预测与生成模型可生成修饰方案，如 R 基团替换、骨架跃迁、片段链接、片段生长等，并根据多维度性质预测结果对候选方案进行排序与筛选，从而压缩每轮优化循环的周期。根据 Cell Reports Medicine 上 2022 年发表的一篇系统性综述，生成式分子设计已在苗头化合物发现、先导物优化、骨架跃迁、基于结构的药物设计（SBDD）等多个环节展现出显著价值。

### 生成式 AI 架构在分子生成中的演进

过去十年间，AI 在计算化学中的应用经历了从早期基于规则的专家系统、定量构效关系（QSAR）模型，向深度生成模型和自动化智能体系统的演进。当前，分子生成方法主要分为基于配体的分子生成（LBMG）和基于结构的分子生成（SBMG）。基于配体的分子生成方法依据已知活性分子的骨架与属性生成新分子，但由于训练过程中缺乏靶点口袋信息，无法建模分子与靶蛋白的相互作用，且易受训练集骨架多样性不足的影响，生成分子的结构多样性有限，在突破专利壁垒方面存在局限。基于结构的分子生成（SBMG）方法则利用靶蛋白三维结构信息，如活性位点或结合口袋的几何与化学特征引导分子生成与优化，能够直接纳入分子-靶蛋白相互作用的约束，设计精度上优于配体方法。常见的数据来源主要包括 1) 各类小分子数据库如 ZINC 数据库、ChEMBL 数据库等，2) 基准数据集，如 MOSES 数据集和 QM9 数据集等，3) 提供序列信息的 Uniprot 数据库，4) 提供蛋白质-配体复合物数据的 CrossDock 数据集和 PDBbind 数据库等。

图 2: 分子生成方法示意图

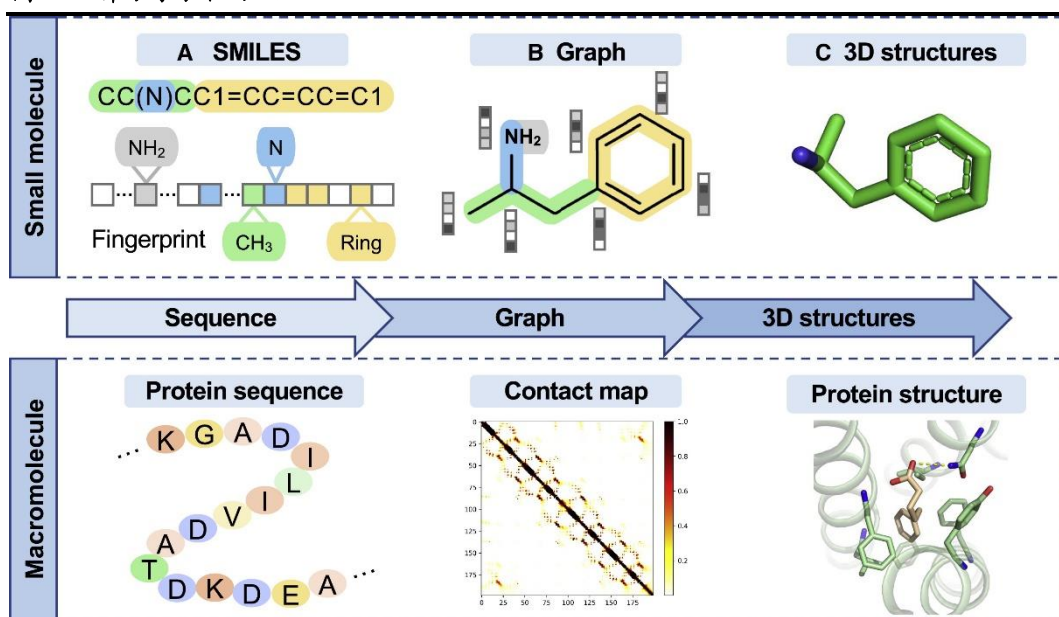


资料来源：中国药科大学学报（[基于人工智能的小分子生成模型在药物发现中的研究进展，2024](#)），招银国际环球市场

小分子 AI 药物研发的技术核心在于如何将分子的化学信息转化为机器可学习的数学表征，并在此基础上构建合适的生成模型架构来完成药物设计。目前，主要的分子表征方式按维度划分，涵盖一维的序列表征、二维的图表表征以及三维的空间表征。

一维表征以 SMILES 为代表，将分子结构编码为一个由字符组成的线性字符串。这种表征方式的优点在于可以直接利用自然语言处理领域成熟的序列模型，如循环神经网络 (RNN) 和 Transformer，且计算效率高。但缺乏拓扑和空间构象信息，随机生成的字符串序列可能无法对应合法的分子结构。二维图表征将分子建模为由原子 (节点) 和化学键 (边) 构成的图结构，更好地反映了分子的拓扑连接关系，但依旧损失空间立体关系。三维空间表征将分子视为三维空间中的原子坐标集合，是与物理现实最为贴近的表征方式，在涉及蛋白质-配体结合构象、分子对接等需要精确空间推理的任务方面更具优势。但三维表征的主要挑战在于数据的获取与计算成本：高精度的三维构象数据依赖于 X 射线晶体学、冷冻电镜等实验技术或分子动力学模拟，因此数据规模较小。

图 3: 3 种分子表征方法



资料来源: [Cell reports. Medicine vol. 3,12 \(2022\): 100794](https://doi.org/10.1016/j.celrep.2022.100794), 招银国际环球市场

生成模型通过学习已知药物分子的分布特征，并在此基础上生成具有特定期望性质的新分子。常见的分子生成领域的模型主要包括：RNNs、Transformer、VAEs、GANs 和扩散模型等。在实践中，这些生成模型通常与基于第一性原理的物理模拟深度整合。例如，英矽智能的 Chemistry42 平台以及 Schrödinger (SDGR US, 未评级) 的计算平台，将深度学习生成架构 (如英矽智能的 Generative Chemistry) 与底层的物理引擎 (如英矽智能的 Alchemy 和 Schrödinger 的 FEP+) 相结合。AI 负责在广阔的化学空间中提出具有极高新颖性的分子骨架，而物理方法则利用相对结合自由能的严谨计算，对先导化合物进行优先级排序，从而提高生成分子在真实热力学环境下实现高亲和力的概率。

值得关注的是，扩散模型自 2022 年以来在三维分子生成领域快速迭代。Hoogeboom 等人提出了一种用于三维分子生成的扩散模型，E(3) Equivariant Diffusion Model (EDM)，解决了三维空间中分子旋转和平移不变性的问题，能够无条件地生成三维分子构象，并在原子位置和原子类型上表现出极高的生成质量 (Emiel Hoogeboom et al., 2022)。Schneuing 等人的 DiffSBDD (Arne Schneuing et al., 2023) 以及 Guan 等人的 TargetDiff (Jiaqi Guan et al., 2023) 将蛋白质口袋作为条件输入，根据靶蛋白的几何与化学环境生成配体分子的三维构象。这一方法将结合构象的合理性直接融入生成环节。尽管扩散模型的采样速度相对较慢，生成分子的合成可行性仍需进一步优化，但其在生成质量和多样性方面的突出表现使其成为当前分子生成领域备受瞩目的技术方向之一。

图 4: 主要的生成模型算法

模型架构	核心技术机制	优势与应用场景	存在的技术局限性
<b>循环神经网络 (RNNs)</b>	将分子生成转化为序列建模问题，将 SMILES 字符串视为化学语言中的句子，通过自回归方式，基于前一个字符来预测下一个字符的概率分布。	在处理 <b>序列化的分子表征</b> 方面具备优势。模型结构成熟，应用广泛，工程实现门槛低，生成速度快。	对三维空间构象的表征能力较弱，容易生成无效或在三维空间中存在空间位阻的分子。
<b>自注意力模型 (Transformer)</b>	基于自注意力机制，通过并行处理分子序列中的所有原子，捕捉长距离的化学依赖关系。在分子生成中，它利用解码器的自回归属性，根据已生成的原子序列递归地预测下一个最可能的原子，从而生成完整的化学结构。	基于 <b>序列生成</b> ，计算效率高。尤其擅长捕捉长程依赖（如大环分子的闭合）。	依赖大量数据，成本高。容易出现化学幻觉，即生成语法正确但化学上不合理的结构。
<b>变分自编码器 (VAEs)</b>	由编码器 (Encoder) 和解码器 (Decoder) 组成。通过编码器将分子结构映射至连续的潜在空间并参数化为概率分布，再由解码器从潜在空间中采样并重构分子结构。	应用于 <b>序列化的分子表征</b> 以及 <b>分子图表征</b> 。平滑的连续潜在空间，适合进行分子性质的梯度优化。	存在后验坍塌 (Posterior collapse) 风险导致生成的分子与训练集中的分子高度相似，缺乏多样性和新颖性。存在“有效性-质量权衡”问题，易收敛于生成结构过于简单、缺乏成药复杂度的分子。潜在空间中经常存在大片的无效死区 (Dead region)，导致解码出的分子在化学上不合法。
<b>生成对抗网络 (GANs)</b>	由生成器 (生成新分子) 与判别器 (区分生成分子结构与真实分子结构) 组成的对抗博弈系统，通过隐式极小极大博弈来逼近真实的化学分布。	应用于 <b>序列化的分子表征</b> 以及 <b>分子图表征</b> 。可生成理化分布与训练集高度重合、具有极高保真度 (Fidelity) 的分子。	训练过程不稳定，容易出现模式崩溃 (Mode Collapse)，即生成器只生产高度重复的同质化化合物，无法提供筛选所需的多样性。
<b>扩散模型 (Diffusion Models)</b>	迭代去噪机制：正向扩散过程中向数据中逐步添加噪声，逆向扩散过程中从随机噪声中重构原始数据样本。	具有生成多样性高、可控性强和训练稳定性好的特点， <b>适合 3D 分子生成</b> ，可直接建模原子坐标和键合信息，在基于结构的药物设计中表现突出，能生成贴合蛋白质结合口袋的 3D 配体分子。	生成过程需经历多步迭代，其计算成本较高，且对预测结构稳定性的信心相对有限，若缺乏合成可及性约束，易生成难以化学合成的结构。

资料来源：[Cell reports. Medicine vol. 3,12 \(2022\): 100794](#), [ACS Chem. Neurosci. 17, 4, 666-680 \(2026\)](#), [中国药科大学学报](#), 55 (3): 295-305 (2024), [生命科学, 37\(12\): 1587-1604 \(2025\)](#), [Commun Chem 6, 249 \(2023\)](#), [ArXiv abs/2205.01592 \(2022\)](#), [Sci Rep 15, 35588 \(2025\)](#), 招银国际环球市场

## ADMET 预测：从经验规则到图神经网络模型

药物的成药性不仅取决于其对靶点的结合活性，更取决于药物分子在吸收 (Absorption)、分布 (Distribution)、代谢 (Metabolism)、排泄 (Excretion) 和毒性 (Toxicity) 等成药性质的研究，即 ADMET 属性。ADMET 问题是导致临床阶段药物失败的主要原因之一。据 *Acta pharmaceutica Sinica. B* 的一篇文献报道，药物研发失败的原因中，与成药性相关的缺乏临床疗效、不可控毒性和不良性质占比达约 80%-90% ([Sun, Duxin et al., 2022](#))。

传统的 ADMET 评估依赖体外实验，如 Caco-2 细胞渗透性实验、微粒体稳定性实验和动物药代实验等，较为耗时。而 AI 模型，如基于图神经网络和 Transformer 的分子性质预测模型，能够在早期阶段对化合物的 ADMET 属性进行快速的计算预测，从而在分子设计阶段便将成药性约束纳入优化目标，减少产生候选药物所需要的“设计-制造-测试-分析 (DMTA)”循环次数。多项研究表明，应用 *in silico* (计算机模拟) 预测性 ADMET 模型，能显著加速发现具有更佳药效、安全性和剂量表现的小分子药物。

早期 ADMET 预测依赖 Lipinski 五规则等经验法则和简单物理化学描述符，结合简单的线性模型，如 Free-Wilson 和 Hammett 分析。这些方法在粗粒度筛选中具有实用价值，但由于特征工程严重依赖人工设计，对复杂体内行为，如代谢位点、组织分布动态、药物间相互作用的预测能力十分有限。非线性模型可以捕捉更复杂的结构与活性关系，包括支持向量机 (SVM)、递归分割法，如随机森林 RF、Cubist 和 XGBoost 以及深度学习方法，如深度人工神经网络 (DNNs)。受益于算力的提升、标准化生成的大型数据集，以及图像和语言处理技术在化学领域的适配，基于深度学习的方法在 ADMET 预测中展现出日益增长的应用前景。据 *Future Medicinal Chemistry* 的文献报道，与随机森林 (RF) 模型相比，DNNs 在 15 个大型实验数据集中使  $R^2$  性能平均提升了 10%，意味着 DNNs 可有效提升预测值与实验数据的匹配性能。但 DNNs 与 RF 一样，均需要化合物的静态指纹表征，即表示化学碎片存在或数量的离散数字向量。研究人员提出采用图卷积神经网络 (GCNN) 方法，以动态学习并优化能够表征关键信息的分子指纹。一项针对 GCNN 对 31 个 ADMET 指标预测的研究表明，单任务 GCNN 在约三分之一的指标上表现出优于 RF 的强劲性能， $R^2$  提升了 0.15-0.31，增幅达 45%-133%。这些指标包括血浆蛋白结合率 (PPB)、HLM、RLM、CYP3A4/2D6 抑制、hERG 结合和动力学水溶性。安进 (AMGN US, 未评级) 和拜耳 (BAYRY US, 未评级) 的数据也证明了 GCNN 的竞争力，多任务 GCNN 在膜亲和力和 PPB 指标上的  $R^2$  提升了 0.26-0.28 ([Elena L Cáceres et al., 2020](#))。据 *J. Med. Chem* 的研究报道，研究人员引入一种名为 Attentive FP 的新型图神经网络架构，利用图注意力机制 (Graph Attention Mechanism) 从相关的药物研发数据集中学习分子表征。Attentive FP 在多种数据集上均取得了 SOTA (当前最优) 的预测性能，且其学习到的内容具有可解释性 ([Xiong, Z. et al., 2019](#))。

2024 年以来，多模态基础模型成为新的突破口。研究人员开发了诸如 SMI-TED289M 等基于 Transformer 的编码器-解码器模型，这类模型利用超过 9,100 万个 PubChem 化合物进行自监督预训练，提取跨越化学和生物领域的统一表征。这种跨模态框架极大提升了在训练数据稀缺和异质性条件下的预测泛化能力，有效降低了分子在后期被淘汰的风险。

## AI 生成药物的临床验证

生成式 AI 在小分子领域的临床验证证据正在快速累积，其中研发进度较快的分子是英矽智能的候选药物 TNIK 抑制剂 ISM001-055 (Rentosertib)，用于治疗特发性肺纤维化 (IPF)。该药物是全球首款药物靶点和分子结构均由生成式 AI 赋能发现的候选药物。依托 Pharma.AI (PandaOmics 用于靶点发现，Chemistry42 用于分子生成)，英矽智能仅用时 18 个月便完成了从识别新型靶点 TNIK 到临床前候选化合物 (PCC) 的提名，研发成本压缩至 260 万美元以内，仅为传统路径的 1/10。2024 年，英矽智能宣布 ISM001-055 达到 IIa 期主要终点，接受 60mg 每日一次 Rentosertib 治疗的患者用力肺活量 (FVC) 均值增加 98.4ml，而对照组 FVC 均值下降 20.3ml。我们认为，ISM001-055 的研发经验表明，AI 可在两方面提升药物发现能力：1) 通过加速靶点优先级排序和化合物设计等任务来简化工作流程。在多组学、结构和药理学数据上训练的模型可快速识别可行的靶点并生成候选分子，缩短早期发现时间线。2) AI 有望识别常规方法可能忽略的生物学见解。英矽智能确定

TNIK 这一与 IPF 关联的新型靶点，展示了 AI 整合多样化数据以揭示功能相关机制的能力。此外，Recursion、Schrödinger 和 Relay Therapeutics (RLAY US, 未评级) 等公司也均有多个管线进入临床阶段。

另一成功路径则是通过知识图谱进行的药物重新定位 (Drug Repurposing)。BenevolentAI 利用其知识图谱和虚拟筛选技术，通过 AI 分析药物-靶标相互作用网络，仅用 2 天预测出原用于治疗类风湿性关节炎的 JAK 抑制剂巴瑞替尼 (Baricitinib) 能够通过抑制 AAK1 减少病毒内吞作用，可用于 COVID-19 的候选药物。2020 年 11 月，巴瑞替尼获 FDA 紧急授权，批准与瑞德西韦联用治疗 COVID-19，并于 2022 年 5 月获全面批准用于治疗需要补充氧气、无创或有创机械通气或体外膜肺氧合 (ECMO) 的住院成人 COVID-19 患者。

但值得注意的是，绝大多数 AI 生成药物仍处于临床早期阶段 (I/II 期)，目前尚无 AI 生成的小分子药物完成关键性 III 期临床并获得监管批准。此前，多款 AI 药物在临床阶段宣告研发失败。比如，由 Exscientia 设计、住友制药 (4506 JP, 未评级) 推进的强迫症药物 DSP-1181 因效果不理想而终止研发，Schrödinger 的针对急性髓性白血病的药物 SGR-2921 也因毒性问题停止开发。因此，我们认为，AI 制药在药物发现阶段已证明了其强大的效率和成本优势，但在创新药研发最为关键且最需提高的临床成功率方面，尚未得到验证，AI 药物研发“从分子到药品”的闭环尚未完全走通。展望未来，尽管面临挑战，不断深入的研究仍在持续拓展 AI 的潜力。随着更多 AI 药物管线步入临床中后期，关键临床数据的陆续读出，若 AI 药物能展现显著临床优势，将成为检验 AI 在药物研发领域掀起底层范式革命的试金石。

## 大分子与生物药研发：从蛋白质结构预测到抗体设计

随着精准医疗的深入，单克隆抗体、抗体偶联药物 (ADCs)、双特异性抗体以及多肽等大分子生物药逐渐成为全球药物市场增长最快的品种之一，也是各大药企的重要创新研发管线。但其研发流程的复杂度远超小分子。将 AI 与实验方法相结合，使得从头设计 (de novo) 抗体成为可能。AI 还通过分析数据集、预测相互作用并指导分子修饰以提高疗效和安全性，从而加速了药物发现、靶点识别和候选药物的优先级排序。

但早期 AI 制药公司大都聚焦于小分子，如 Exscientia (现已被 Recursion 收购)、Recursion、Schrödinger、Relay Therapeutics 等。目前，全球的人工智能驱动的药物研发 (Artificial Intelligence in Drug Discovery, 简称 AIDD) 也更多聚焦于化药小分子的研发，达成授权交易的小分子药物数量也远超大分子生物药。据医药魔方的数据，2024 年至今，小分子药物达成交易的数量占比达到近 50%。但新一代平台正将 AI 应用于生物药，包括抗体、蛋白质和基于 RNA 的治疗药物，以 AI 驱动大分子设计为核心的公司相继涌现，如 Generate Biomedicines (GENB US) 和 Absci (ABSI US)。但整体而言，我们认为，AI 在大分子生物药研发中的应用尚处于起步阶段。我们认为，这可能主要由数据基础以及分子特性决定的研发成本和研发难度差异所致：

1) 小分子领域经过多年发展，积累了庞大的结构化数据资产，如 ChEMBL 数据库中包含超 200 万个化合物的活性数据，ZINC 数据库包含数十亿虚拟化合物的可生成结构。这些公开数据库，可提供丰富的分子结构、理化性质及生物活性数据，为 AI 模型提供了充足燃料，适合训练监督学习和生成式模型。而大分子领域由于实验数据获取成本高昂、周期漫长，如蛋白质必须以足够的数量生产并进行纯化，选择合适的样品制备条件，才可收集高质量的数据。因此大分子领域的结构数据积累相对滞后，如国际蛋白质结构数据库 (PDB) 中仅累积存储超 25 万条实验结构数据 (含蛋白质、核酸等结构数据)，SAbDab 包含 10,673 条抗体结构数据和 735 条抗体亲和力数据，AACDB 数据库仅包含 7,498 条抗体-抗原复合

物的结构数据。这些数据库的规模远小于小分子数据库，制约了需要大量数据进行学习的 AI 模型的开发与应用。

2) 小分子的化学结构相对简单，易被转化为机器可学习的数学表征，如 SMILES 序列表示和分子图表示，这使得图神经网络 (GNN) 和 Transformer 等模型可以直接应用于小分子生成和属性预测。但大分子的计算表征则更为复杂，比如，一个 IgG 抗体约有上千个氨基酸残基，其功能不仅取决于一级序列，还高度依赖三维折叠构象、糖基化修饰模式、二硫键配对等因素。要准确建模抗体的结合功能，需要在序列、结构和动力学三个层面上进行联合表征，对模型架构和计算资源的要求远高于小分子。

### 蛋白质结构预测与抗体的“零样本”从头设计

理解靶点蛋白和抗体本身的三维结构，是大分子药物理性设计的前提。传统实验手段如 X 射线晶体学、冷冻电镜 (Cryo-EM) 和核磁共振 (NMR) 等虽然检测精度极高，但单个蛋白结构的解析通常需要数月至数年，且对膜蛋白、无序蛋白等“难表达、难结晶”靶点的覆盖率极低。据 Nature 的文献报道，实验手段解析的结构仅覆盖人类蛋白组中约 17% 的氨基酸残基 ([Tunyasuvunakool, K. et al., 2021](#))。因此，蛋白质结构解析曾是阻碍生物药设计的难题，也是制约 AI 在大分子药物研发领域应用的重要原因。

2020 年，DeepMind 的 AlphaFold2 在第 14 届蛋白质结构预测关键评估竞赛 (CASP14) 中取得了中位 GDT 分数 92.4 的突破性成绩，其蛋白质结构预测结果的精确度与实验室方法近乎相当，实现了基于氨基酸序列精确预测蛋白质 3D 结构的重大突破，被 Science 评选为 2021 年度十大科学突破。2021 年，DeepMind 与欧洲生物信息学研究所 (EMBL-EBI) 合作推出了公开数据库 AlphaFold 蛋白结构数据库 (AlphaFold Protein Structure Database, 缩写 AlphaFold DB)，目前 AlphaFold DB 包含了由 AlphaFold 预测的超 2 亿个蛋白预测结构。同年，华盛顿大学的 David Baker 研究团队发布 RoseTTAFold 的蛋白结构预测系统，突破了 AlphaFold2 的局限，不仅可用于基于氨基酸序列预测单一蛋白质的结构，还可以用于预测蛋白复合物结构。研究显示，RoseTTAFold 系统利用 IL-12 和 IL-12 受体 (IL-12R) 的序列预测的 IL-12/IL-12R 复合体结构与此前利用冷冻电镜解析的结构非常类似 ([Minkyung Baek et al., 2021](#))。2024 年 AlphaFold 与 RoseTTAFold 的开发者分享了诺贝尔化学奖。

2024 年发布的 AlphaFold3 再次实现重要技术跃迁，它将预测范围从单一蛋白扩展至蛋白质与 DNA、RNA 以及小分子 (如配体和离子) 的相互作用的预测，在无需任何结构信息输入的情况下，AlphaFold3 在 PoseBusters 基准上的配体构象预测准确性较传统方法提升 49% ([Abramson, J., 2024](#))。其免费平台 AlphaFold Server 可建模模拟蛋白质相互作用，预测翻译后修饰及核酸变化对蛋白质功能的影响。这种利用计算机解析蛋白质与其他分子复杂相互作用的能力，为抗体-抗原复合物、双特异性抗体的多结构域空间排布、ADC 中抗体与连接子的界面构象等此前高度依赖实验的结构问题，带来了全新的见解。

得益于上述底层突破，基于结构的“从头设计 (de novo design)”开始具备可行性。传统抗体药物的发现依赖于杂交瘤技术或噬菌体/酵母展示文库筛选，本质上是在自然免疫系统或人工构建的有限序列空间中淘金，其可探索的抗体序列多样性往往受限于文库容量，同时筛选过程也较为繁琐，耗时较长。

### 1. 动物免疫：小鼠杂交瘤技术 (Mouse Hybridoma Technology)

该技术利用动物自身免疫系统实现抗体的体内亲和力成熟。其核心是将能分泌特异性抗体但寿命短暂的 B 细胞，与能够在体外无限增殖的骨髓瘤细胞融合。杂交瘤技术得到的抗体通常具有较高的天然亲和力，但人源化过程耗时且存在失败风险。

具体流程：

- 1) 抗原制备与动物免疫：将目标靶点蛋白（抗原）注射到小鼠体内，通常需要多次免疫（初免和加强免疫），以激发小鼠免疫系统产生强烈的免疫反应。
- 2) 获取脾脏 B 细胞：取出产生强免疫反应的小鼠脾脏，分离出其中能够分泌抗体的 B 淋巴细胞。
- 3) 细胞融合：在化学促融剂（如 PEG）或电融合技术的帮助下，将这些 B 细胞与小鼠骨髓瘤细胞融合，形成杂交瘤细胞。
- 4) HAT 培养基筛选：将融合后的细胞混合物放入特定的 HAT 培养基中。未融合的 B 细胞几天后会自然死亡，未融合的骨髓瘤细胞在 HAT 环境中无法存活。只有成功融合的杂交瘤细胞才能存活。
- 5) 克隆化与高通量筛选：将存活的杂交瘤细胞稀释，分配到多孔板中进行单克隆培养。然后使用 ELISA 或流式细胞术（FACS）检测每个孔的上清液，筛选出能分泌高亲和力且可特异性结合目标抗原的克隆。
- 6) 测序与人源化改造：获取目标抗体的基因序列。由于鼠源抗体直接注入人体会引起强烈的免疫排斥（HAMA 反应），因此必须通过基因工程手段，将鼠源抗体的互补决定区（CDR）移植到人源抗体骨架上，进行“人源化”改造，才能最终成为药物候选分子。

## 2. 体外展示文库：噬菌体展示技术（Phage Display Technology）

为了绕过动物免疫的限制，如动物对高同源性人类蛋白的免疫耐受，及人源化过程，体外展示技术应运而生。它的核心是将“基因型”（DNA 序列）和“表型”（抗体蛋白）在病毒或酵母表面直接关联起来。体外展示文库，特别是全人源文库可以直接获得全人源抗体，且能针对一些在动物体内难以产生免疫反应的“困难靶点”进行筛选，但通常需要后续进行体外的亲和力成熟改造。

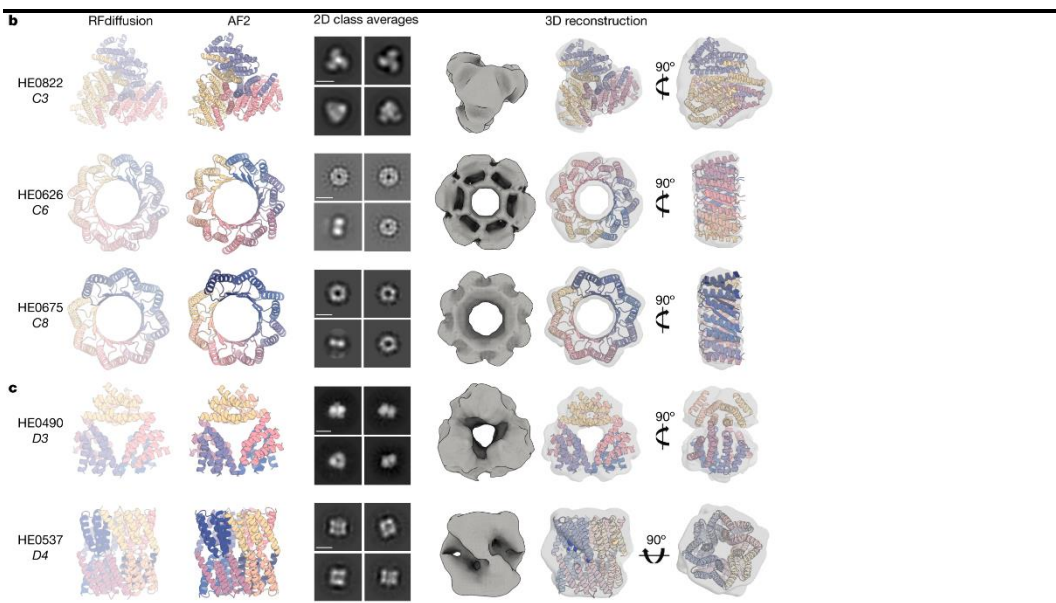
以应用最广泛的噬菌体展示为例，具体流程：

- 1) 构建基因文库：提取人类 B 细胞的 mRNA，逆转录成 cDNA，扩增出抗体可变区基因，通常是 scFv 或 Fab 的片段。将这些基因片段插入到丝状噬菌体的外壳蛋白基因中。这个文库的库容通常较大，一般会达到  $10^9$  到  $10^{11}$  量级。
- 2) 抗体表面展示：当噬菌体感染大肠杆菌并繁殖时，它们会将抗体片段展示在病毒外壳表面，而编码该抗体的 DNA 则包裹在噬菌体内部。
- 3) 生物淘选：
  - 结合：将目标抗原固定在固相载体（如磁珠或孔板）上，加入噬菌体文库。能与抗原结合的噬菌体会被留下。

- 洗脱与富集：将结合的噬菌体洗脱下来，去感染大肠杆菌进行扩增。
  - 循环：将扩增后的噬菌体用于下一轮淘选。通常经过 3-5 轮“结合-洗脱-扩增”的严苛筛选，高亲和力的抗体噬菌体会得到指数级富集。
- 4) 单克隆鉴定与测序：挑取单克隆进行 ELISA 鉴定，找到强结合的克隆后，测序噬菌体内部的 DNA，即可获得抗体的基因序列。
  - 5) 全长抗体表达与纯化：将获得的片段序列转换成完整的 IgG 抗体格式，在哺乳动物细胞（如 CHO 细胞）中表达出来，进行后续的成药性评估。

生成式 AI 正在将抗体发现从“筛选”范式推向“设计-预测-验证”范式。2023 年，David Baker 研究团队发布深度学习模型 RFdiffusion，从头生成全新的蛋白质骨架，再使用 ProteinMPNN 设计编码这些结构的序列。其设计的分子在无需进一步实验优化的情况下即可达到纳摩尔级的亲和力，整体成功率达 19%（定义为结合响应达到或超过阳性对照最大响应的 50%），相较于相同靶标上基于 Rosetta 的方法设计成功率提高了约两个数量级。就所设计的分子本身而言，实验结构与 RFdiffusion 设计结构的偏差可低至 0.6Å。这些结果证明了 RFdiffusion 能够以原子级精度生成新蛋白质，并能精确瞄准具有治疗重要性的蛋白质上与功能相关的位点（[Joseph L. Watson et al., 2023](#)）。

图 5: RFdiffusion 预测的结构与实际电镜下的结构对比



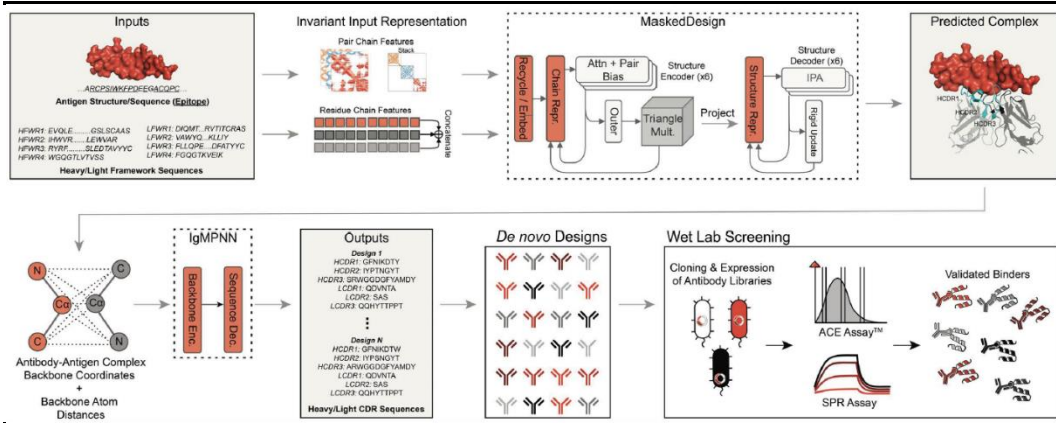
资料来源：[Nature 620, 1089-1100 \(2023\)](#)，招银国际环球市场

注：从左到右依次为 RFdiffusion 预测的结构；AF2 通过 ProteinMPNN 预测的序列，再预测结构得到的示意图；电镜下 2D 类平均值；电镜下的 3D 示意图。从这些图中可以看出实际结构与预测非常接近。

2024 年，AI 制药公司 Absci 发布研究成果，展示了 AI 在抗体“零样本”从头设计（zero shot de novo）和先导化合物优化方面具有提高研发速度、质量和可控性的潜力。利用在抗体-抗原相互作用上训练的深度学习模型，Absci 对抗体互补决定区（CDRs）进行“零样本（zero shot）”设计，即在没有任何已知结合物数据的情况下，仅凭靶点的结构信息，直接生成具备功能性的对 HER2 具有结合活性的全新抗体。再结合高通量湿实验，将 AI 设计的抗体方案转化为可表达的抗体分子。AI 生成的 HER2 抗体与已知抗体的序列一致性低，且在功能特性方面，如交叉反应性、结合力等，与曲妥珠单抗相似甚至更优。2026 年初，Absci 发布抗体设计平台 Origin-1。在缺乏靶点-抗体复合物结构数据的条件下，该平台成功

针对 4 个新靶点 (COL6A3、AZGP1、CHI3L2 和 IL36RA) 设计出具备特异性结合能力的全长单克隆抗体, 并通过冷冻电镜 (Cryo-EM) 验证了其设计结构达到原子级精度, 揭示了复合物结构与计算模型高度匹配, 具有极高的结构保真度 (分辨率为 3.0-3.1Å; DockQ 值为 0.73-0.83) ([链接](#))。

图 6: 生成式 AI 模型用于抗体从头设计 (De novo design)



资料来源: Absci, 招银国际环球市场

我们认为, 目前生成式 AI 应用于大分子药物研发领域的主要价值在于, 以 AI 预测作为实验设计的起点, 缩短从序列到可用结构的周期, 同时以实验数据反向校准和迭代模型。展望未来, 生成式 AI 更大的价值可能将体现在攻克孤儿靶点或不可成药靶点方面。传统的杂交瘤技术无法控制小鼠体内产生的抗体与靶点蛋白的结合表位, 若结合在无效表位, 则会导致缺乏治疗效力。而生成式 AI 模型通过对蛋白质语言的理解, 可基于靶点结构, 利用算法识别具备治疗潜力的结合表位, 并设计能与之结合的抗体结构。

### 深度学习驱动的抗体亲和力成熟

初步筛选的抗体通常不具备足够的靶标亲和力。为了达到治疗级别的效力, 抗体必须经历亲和力成熟 (Affinity Maturation) 以达到所需的生理特性。传统的亲和力成熟依赖于经验性实验, 通常需通过对互补决定区 (CDRs) 进行突变, 再通过展示技术进行筛选。AI 应用于抗体亲和力成熟的价值便在于将“突变+筛选”转变为“预测+定向验证”, 从而加速亲和力成熟过程和抗体研发。

基于深度学习的蛋白质语言模型 (pLM, 如 AntiBERTa) 和图神经网络 (如 AbMPNN、AbDiffuser), 可通过学习大量自然演化形成的蛋白质序列, 提取生物物理特征与进化保守性信息; 而基于结构的图模型则着眼于解决逆向折叠问题 (Inverse Folding), 即寻找能够折叠出最优结合界面构象的氨基酸序列组合。通过模拟抗原-抗体对接, AI 能够预测 CDR 区域不同突变引发的结合自由能改变, 从而在庞大的突变空间中锁定能够提升亲和力的热点 (Hotspots)。据 Nature Biotechnology 的研究报道, 使用蛋白质语言模型进行亲和力成熟, 仅通过两轮实验室进化及约 20 个抗体筛选, 就可将四种具有临床相关性且已经高度成熟的抗体亲和力提高多达 7 倍, 并将三种未成熟抗体的结合亲和力提高多达 160 倍 ([Brian L. Hie et al., 2024](#))。Absci 通过其抗体设计平台 Origin-1 得到的从头设计抗体 IL36RA 亲和力为微摩尔级别。利用 AI 指导的亲和力成熟技术, Absci 仅通过一轮优化就将亲和力提升了 68 倍, 最终将抗体优化为 EC<sub>50</sub> 达约 104nM 的功能性拮抗剂 ([链接](#))。

此外, AI 技术在预测和优化抗体的特异性、稳定性以及整体开发可行性方面, 也展现出巨大潜力, 如 1) 优化特异性以减少脱靶结合, 从而减少潜在的副作用; 2) 增强稳定性以提

升开发可行性 (developability) 和可生产性 (manufacturability) ; 3) 预测抗体的粘度, 因为高粘度可能影响制剂开发及给药; 4) 优化溶解度和表达水平、聚合倾向等与开发可行性相关的属性 ([Matsunaga, R., 2025](#)) 。

### 抗体偶联药物 (ADCs) 与双抗的优化设计

抗体偶联药物 (ADC) 是当前肿瘤创新药研发最活跃的赛道之一, 但其分子设计复杂度显著高于传统单抗。其研发进程受限于多重挑战: 包括抗体、连接子 (Linker) 与载荷 (Payload) 三者匹配产生的组合复杂性、非靶向毒副作用等。由于高效的 ADC 需要在循环系统中保持稳定并在肿瘤内特异性释放, 这种精密的设计逻辑导致研发过程既耗时又成本高昂。若 linker-payload 匹配或偶联策略优化不足, 还可能带来 DAR 异质性或药物在血液循环中的提前脱落, 引发脱靶毒性。

AI 正在推动 ADC 研发从传统的经验主义试错转向数据驱动的闭环工程。通过整合抗体序列特征、结构与分子动力学特征等多维数据, AI 模型可优化抗体筛选, 模拟 ADC 各组分的相互作用, 辅助筛选更优 linker-payload 组合, 并结合反应动力学模型优化 DAR 分布与偶联工艺条件。双特异性抗体也是 AI 可切入的重要方向。双抗天然面临重链异源二聚、轻重链错配、聚集倾向与稳定性不足等工程化挑战, 将直接影响表达产率、纯化复杂度与商业化可行性。计算分析和蛋白设计正有望优化链配对界面、降低错配风险。

### AI 技术降低免疫原性

免疫原性是大分子生物药开发中的关键系统性风险之一。治疗性蛋白在体内可能诱导抗药抗体 (ADA) 形成, 从而影响药物暴露、削弱疗效, 严重时还可带来超敏反应等安全性问题。AI 正逐步成为低免疫原性蛋白工程的重要工具, 通过结合序列、结构、生物物理与免疫学信息, 对潜在 T 细胞表位、表面暴露残基、聚集倾向及似天然性进行综合评估, 在尽量保留功能的前提下, 降低免疫识别与 ADA 风险。

Generate Biomedicines 创建了表面重构蛋白技术 (Re-surfaced proteins) 利用机器学习方法生成表面重构蛋白, 在维持蛋白功能的同时引入多达 58 个氨基酸残基突变, 因此可有效破坏 ADA 的结合表位。大肠杆菌来源的 L-天冬酰胺酶 (ASN) 是一种在儿童和成人急性淋巴细胞白血病 (ALL) 治疗方案中作为标准成分的酶。据 Generate 披露, 经过表面重构的 ASN 在小鼠和家兔模型以及 ALL 患者样本中, ADA 结合力显著降低, 且有效缓解了小鼠模型中的急性超敏反应。

### 临床试验的 AI 重构: 数字孪生、患者分层与成功率预测

临床试验是循证医学的金标准, 也是药物研发中最耗时且费用投入最高的阶段, 约占从药物发现到上市整体研发时间和成本的 60% 以上。同时, 临床研究在方案设计和落地实施阶段也面临着多重挑战。据 Clin. Trials 期刊的报道, III 期试验平均需要 1,900 万美元的资金投入, 且从启动到完成需耗时 6-7 年。更值得注意的是, 80% 的试验因入组挑战而导致延迟, 比如 37% 的研究中心无法招募到参与者, 招募失败将使试验时间线平均延长 6-8 个月。此外, 传统的数据收集方法严重依赖人工流程, 导致高错误率。研究表明, 高达 50% 的临床试验数据包含错误或不一致, 需要清理, 导致研究被进一步延迟。技术和科学的发展融合, 为通过 AI 技术应对这些挑战创造了机遇。比如, 电子病历 (EHRs) 的普及产生了庞大的临床数据库, 为 AI 驱动的学习和预测提供了此前无法获取的原始数据。同时, 机器学习, 特别是深度学习和自然语言处理的突破, 使算法能够处理复杂的数据集, 并识别出人类分析师无法检测到的细微规律。

但由于药物的临床研究受到严格监管，其流程受各种法规制度以及审批等因素约束，因此我们认为，AI 对临床阶段的潜在助力并非直接的流程加速或是研究范式的根本性迭代，而是有望通过 AI 驱动的方案优化、AI 风险分层框架、试验结果的预测建模以及“数字孪生”等底层技术的引入，为传统的临床试验提供具有前瞻性的动态模拟，从而提升临床试验的成功率。

### AI 在临床试验设计与可行性中的应用

我们认为，AI 驱动的临床试验设计与方案优化，本质上是从基于经验的试验设计向循证规划的转变：

- 1) 优化方案：机器学习算法可分析庞大的历史试验数据，识别出能够为最佳研究参数提供参考的模式。自然语言处理 (NLP) 工具能系统地从大量先前的研究方案、监管提交资料和已发表文献中提取信息，以提出基于证据的入组/排除标准。这些 AI 驱动的方法展示了优于传统统计方法的显著优势。据 Int J Med Inform 的文献报道，机器学习算法在试验方案优化方面达到 80% 的预测准确率，而传统的基于回归的方法预测准确率仅为 65%。此外，深度学习模型可以通过同时分析多个变量，包括目标人群特征、地理分布、季节性变化和竞争格局因素，来预测方案的可行性。这些预测模型在预测招募成功率方面的准确率超过 80%，显著优于传统的可行性评估 (David B. Olawade et al., 2026)。
- 2) 中心选择：通过整合多个数据源分析人口统计数据、疾病流行率、医疗基础设施、研究者经验和历史中心表现指标，机器学习算法还能对潜在的研究中心进行排名，从而实现对研究者表现和招募潜力的预测。通过在中心层面预测招募率，AI 工具有望实现更准确的时间表预测和资源分配，从而降低研究延误和成本超支的风险。

图 7: 试验设计与中心选择中的 AI 应用与传统方法的性能对比

应用领域	AI 技术	传统方法性能	AI 性能	主要优势
方案优化	NLP+ML	65% 预测准确率	80% 预测准确率	缩短设计时间，提高可行性
中心选择	预测分析	45% 入组预测准确率	75% 入组预测准确率	更好的入组预测
可行性评估	机器学习	8-12 周规划时间	4-6 周规划时间	风险缓解，资源优化
竞争情报	数据挖掘	60% 试验覆盖率	90% 相关试验覆盖率	市场格局分析
人群建模	深度学习	55% 招募准确率	75% 招募预测准确率	患者可用性预测

资料来源: [International journal of medical informatics vol. 206 \(2026\): 106141.](#), 招银国际环球市场

### AI 驱动患者招募与筛选

在患者招募阶段，自然语言处理 (NLP) 算法能够高效扫描电子病历 (EHRs) 和非结构化的临床记录文本，结合计算机视觉技术对病理学切片和医学影像进行特征提取，从而在多模态临床数据中识别潜在的合格参与者，大幅提高患者招募与筛选的效率与准确性。同时，深度学习模型还可以通过评估数千种临床与多组学 (如基因组、蛋白质组) 特征，实现高精度的患者分层。比如，通过识别出最有可能从特定干预措施中获益的患者亚群，从而在临床试验中开发个性化治疗策略。AI 驱动的患者分层有望使试验设计更高效，在维持统计效能的同时减少样本量需求。2026 年 1 月 6 日，阿斯利康 (AZN US, 未评级) 与 BostonGene 宣布达成战略合作，BostonGene 将为阿斯利康提供肿瘤与免疫生物学的基础模型及多模态分析能力，重点在早期临床试验中提供患者层面的安全性与疗效预测洞见，助力生物标志物开发与试验设计优化。

但值得注意的是，由于 AI 系统需利用各地的医院电子病历系统等真实世界数据来实现患者筛选，目前仍面临较大阻碍。一方面，由于数据存在非标准性和严格的隐私保护限制，企业需针对当地临床术语、实验室结果单位和影像报告风格进行大量的定制化清洗与对齐工作；另一方面，在跨机构集成 EMR 系统时往往需要耗费高昂的资金与时间成本。因此，AI 驱动的患者筛选在真实世界的规模化落地仍是一项极具挑战的系统性工程。

### 临床管线成功率预测

药企研发资金投向的决策通常是由临床、转化医学、统计、商业、注册和管理层共同参与。PoS (Probability of Success) 或预测性成功概率模型如 Bayesian PoS 本身并不是新概念，在临床试验中已被用于辅助临床决策，如试验的继续/停止判断和项目规划。AI 的增量价值主要体现在提升系统处理分子结构、靶点生物学、组学、方案文本、历史试验结果和外部知识库等异构数据的能力，从而使风险评估更标准化且可扩展。

以英矽智能的 inClinico 为例，公司将其定位为一个数据驱动的多模态预测单项临床试验成功概率 (PoS) 的平台。利用靶点、疾病、试验方案和结果等多模态数据，inClinico 可通过量化临床试验 II 期到 III 期的成功概率来预测临床试验结果。据公司招股书披露，inClinico 在 II 期至 III 期转化预测这一特定任务上取得了 0.86 的预测值 (以 ROC 曲线下面积 (AUC-ROC) 计量)，AUC-ROC 是衡量模型预测试验成功能力的指标，1.0 表示其具备完美预测能力。我们认为，这一结果说明 AI 在关键临床开发节点的风险分层及辅助决策方面具备潜力，特别是在临床方案比较和管线决策中有望提供有价值的量化参考，帮助药企更早识别缺乏生物学支撑或方案设计薄弱的项目，并将资源向机制更扎实、试验路径更优的管线资产倾斜。

### 数字孪生 (Digital Twins) 与合成对照臂 (SCA)

在罕见病、儿科疾病及部分肿瘤的临床研究中，传统随机对照试验 (RCT) 常常面临更大的伦理与执行障碍，尤其是在患者基数有限、疾病进展迅速、未满足需求显著的场景下，招募患者进入安慰剂组或疗效有限的标准治疗组往往具有较高难度。因此，外部对照臂、自然史研究等真实世界数据 (RWD) 和真实世界证据 (RWE) 正成为临床研究中的方法学补充。美国 FDA 近年来也在推进 RWD/RWE 的监管框架，比如在 2025 年发布 Rare Disease Evidence Principles (RDEP)，明确提出对患者规模极小、明确遗传驱动且存在重大未满足需求的罕见病，监管审评可纳入更广泛的支持性证据，而不再机械依赖传统证据路径，即审批在某些情形下基于一项充分且受控的研究+强有力确认性证据 (confirmatory evidence)。监管机构将接受多元化的确证性证据支持，如强有力的作用机制或生物标志物证据，相关非临床模型证据及临床药效学 (PD) 数据，及病例报告、扩大准入数据或自然史研究 (Natural History Studies)。同时，FDA 在官网上也列举多个利用 RWD 和 RWE 支持监管决策的案例，表明该路径具备现实监管基础。

而 AI 赋能的数字孪生 (DT) 与合成对照臂 (SCA) 技术也开始展现优化临床研究的潜力。

- **数字孪生 (DT)**：数字孪生是通过 AI 生成的针对受试者的预测模型。通过整合 RWD 与计算建模，DT 可以实现个体患者生理特征、疾病演变及治疗反应的动态模拟，从而优化资源配置，并解决受试者不愿被随机分配至对照组的伦理痛点。此外，在个体层面，由于整合了临床、基因组学、多组学及实时监测数据，DT 可预测患者对特定治疗方案的反应，从而支持个性化给药。而在群体层面，通过 RWD 和观察性研究数据构建的群体 DT，可在试验启动前模拟不同设计方案、识别潜在失败点并优化方案设计。这一预演机制有望降低因设计缺陷导致的高昂失败成本。以 Unlearn.AI 为代表

的公司，正尝试基于历史患者数据训练生成式或预测性模型，为每位受试者构建其在对照治疗或标准治疗下的个体化结果预测。我们认为，数字孪生的潜在价值是将患者个体的预后信息纳入统计分析框架，从而提升试验成功概率，并在特定场景下支持缩小对照组规模或增强单臂研究的可解释性。但据npj Digital Medicine的研究报道，目前，DT的应用主要是从回顾性验证（如阿尔茨海默症试验中预测轨迹与历史数据的一致性评估）向实时验证演进。由于电子健康档案不完整，可穿戴设备数据缺失以及临床报告异步，实时数据的缺口仍是阻碍数字孪生的主要因素。且DT的泛化性依然是核心挑战，需要建立透明的、针对具体疾病的验证管线进行大规模基准测试。

- **合成对照臂**：通过整合真实世界数据、历史试验数据和患者登记数据，AI可构建在关键基线特征上可比的对照人群，从而补充或替代实际对照组。这种干实验（In-silico）模式有望缩减对照组入组人数，从而缩短招募周期，甚至实现单臂试验。理想情况下，它还能剔除患者个体异质性对疗效评估造成的干扰，使得新药的真实疗效能够被更加清晰、客观地量化展现。

## FDA 积极推进 AI 监管体系建设

### FDA 的 AI 内生能力建设

随着生成式AI在生命科学领域的广泛应用，FDA自2024年起加快了AI治理与部署节奏。2024年3月，FDA发布《Artificial Intelligence and Medical Products: How CBER, CDER, CDRH, and OCP are Working Together》，阐述其下属的生物制品审评与研究中心（CBER）、药品审评与研究中心（CDER）、器械与放射健康中心（CDRH）以及组合产品办公室（OCP）四大中心在AI医疗产品全生命周期监管中的跨部门协同框架，明确药品、生物制品及医疗器械领域AI应用的审评路径与职责分工。同年，FDA成立CDER AI委员会（CDER AI Council），负责对CDER围绕AI的活动进行监督、协调与统筹整合，主要涵盖两大方面：一是CDER内部AI能力建设，包括人才、技术、数据、算法和模型；二是推进药物监管决策相关的AI政策。值得注意的是，FDA明确赋予该委员会确保CDER对AI相关沟通保持统一的职责，这意味着药企未来在与FDA就AI相关议题沟通时，将拥有更稳定的政策和审评口径。

图 8: FDA 与 AI 相关的关键职位

治理机构与关键职位	核心领导人物	战略定位与主要职责
首席 AI 官 (Chief AI Officer)	Jeremy Walsh	统筹 FDA 的 AI 战略部署，监督内部 AI 技术创新与外部政策制定的协同，推动医疗数据与 AI 架构的深度融合。
外部委员会 (AIPC)	Jeremy Walsh (主席) Shantanu Nundy	强化跨部门在 AI 政策制定方面的协作，针对 AI 相关监管挑战，指导跨领域解决方案的制定，协调受监管产品中新兴 AI 技术的试点项目。
内部委员会 (AICC)	Jeremy Walsh (主席) William Liu	制定明确的内部 AI 使用政策，推广 AI 能力建设，主导 AI 相关的员工培训。
CDER AI 委员会 (CDER AI Council)	跨部门高级别专家	作为药物评估与研究中心 (CDER) 的专属决策机构，整合原有的 AI 指导委员会和政策工作组，负责 CDER 内部及外部 AI 相关活动的一致性，提升在药物安全、有效性评估中 AI 考量的一致性。

资料来源：FDA, Citeline, 招银国际环球市场

从内部能力建设看，FDA正快速将AI技术嵌入日常监管流程。2025年5月，FDA宣布完成首个AI辅助科学审查试点（AI-assisted scientific review pilot），并计划在2025年6月30日在所有中心全面部署生成式AI，有望显著减少审查时间。2025年6月，FDA发布内

部的生成式 AI 工具 Elsa，用于加快临床方案审查，缩短科学评估所需的时间。2025 年 12 月，FDA 进一步面向全员部署 agentic AI capabilities（智能体 AI 能力），用于上市前审评、上市后检测、现场检查与合规性事务以及运营管理等。

### 基于风险的可信度评估框架：FDA 发布决策支持指南草案

从外部监管框架看，FDA 正逐步搭建针对药物与生物制品的 AI 合规体系框架。2025 年 1 月，FDA 发布《Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products》草案，首次提出七步风险可信度评估框架（Risk-Based Credibility Assessment Framework），为 AI 模型在药物全生命周期（涵盖非临床、临床、上市后及生产阶段）中生成支持安全性、有效性或质量监管决策的数据与信息提供可信度建立、评估及监管申报的系统性指导。我们认为，FDA 针对 AI 的监管核心原则包括：1) 风险分层：将 AI 模型风险拆解为模型影响力和决策后果两个维度，为药企在临床试验终点设计、CMC 质量控制、上市后不良反应监测等场景中使用 AI 提供了可量化的可信度证据要求。2) 情景化可信度：使用场景（Context of Use, COU）是可信度评估的锚点，同一 AI 模型，若用于早期非临床毒性预测（低后果），其可信度要求将低于用于替代 III 期临床主要终点。3) 数据质量为刚性约束：FDA 要求数据必须可溯源，具有代表性，且与模型预期用途在科学上匹配。

图 9: 七步风险可信度评估框架

评估步骤	监管要求与执行内涵
步骤 1: 定义问题	明确 AI 模型拟解决的具体问题、临床决策或关注点，避免模糊的全局性宣称。
步骤 2: 定义使用场景 (COU)	明确 AI 模型的具体使用场景 (Context of Use) 和角色，包括其输出是直接支持决策还是辅助参考，使用阶段和用户交互方式及决策流程。
步骤 3: 评估 AI 模型风险	通过模型影响力×决策后果二维矩阵确定模型的风险等级和可信度所需的严格程度。 <b>模型影响力 (Model Influence)</b> ：模型输出对最终决策的权重。 <b>决策后果 (Decision Consequence)</b> ：决策错误对患者安全/产品质量的潜在影响。
步骤 4: 制定可信度评估计划	基于步骤 3 确定的风险等级，制定一个 AI 模型输出可信度评估计划，包括数据治理策略、模型开发生文档、性能评估指标、不确定性量化方法等。在执行前与美国 FDA 讨论计划有助于明确适当的可信度评估活动，并识别潜在挑战及解决方案。
步骤 5: 执行可信度计划	按照计划开展模型模型训练、验证、不确定性分析及敏感性测试。需严格隔离训练集/验证集/测试集，监测数据漂移和分布偏移，并与金标准或现有方法的对比研究。
步骤 6: 记录结果与偏差讨论	记录评估计划的执行结果，并对任何偏离预定评估计划的情况进行透明的科学论证。
步骤 7: 确定模型对使用场景的适用性	基于可行度评估结果，明确模型适用范围与禁忌情境，阐述剩余不确定性及风险缓释措施，并提供充分信息支持监管审评。

资料来源：FDA，招银国际环球市场

### 监管共识：FDA 与 EMA 联合发布十项指导原则

2026 年 1 月，FDA 下属 CDER 和 CBER 与 EMA 联合发布《Guiding Principles of Good AI Practice in Drug Development》，就 AI 在药品全生命周期中的应用提出 10 项共同指导原则，初步搭建起面向 AI 制药的高层监管框架，也为 AI 技术在医药领域的规范化推广提供了重要制度基础。

我们认为，此次联合发文标志着欧美药品监管机构在 AI 治理核心逻辑上的趋同性正在增强。监管关注重点正由单一模型性能评价，逐步转向对 AI 系统在特定应用场景下可信用度、可追溯性、可验证性与生命周期可控性的综合审查。换言之，传统算法性能指标仍是必要基础，但已不足以单独构成监管可接受性的核心证据，企业还需同步证明其模型开发、验证、部署和迭代过程具备充分的治理能力与合规支撑。

从企业 AI 能力建设层面看，我们认为这 10 项原则可进一步被传统制药企业及 AI 制药企业内化为一套 AI 平台建设及治理的核心指引。一方面，企业需围绕具体应用场景建立覆盖模型立项、验证、变更控制及持续监测的全生命周期管理机制。另一方面，企业还需强化数据采集、治理与记录规范，将数据来源、处理流程、分析决策以及模型更新全过程纳入可追溯、可审计的管理框架。整体来看，这份指导原则有望提升行业应用质量与监管可接受度，推动 AI 制药技术快速且健康发展。

图 10: 十项 AI 实践原则

核心原则	深层监管启示与合规要求
1. 以人为本的设计	AI 技术的开发与使用须符合伦理要求，以患者和医护人员为中心，尊重人类自主决策权。
2. 基于风险的方法	AI 技术的开发和使用的遵循基于风险的方法，根据使用场景和确定的模型风险，进行相对应的验证、风险缓解和监督。
3. 遵守标准 (GxP)	AI 技术应遵守相关的法律、伦理、技术、科学、网络安全和监管标准，包括良好实践 (GxP)，如 GLP, GCP, GMP。
4. 明确的使用场景	AI 技术应具有定义明确的使用场景，即使用该技术的原因、角色和范围。
5. 多学科专业知识	涵盖 AI 技术及其使用场景的多学科专业知识应贯穿于技术的整个生命周期，即要求贯穿系统生命周期的整合算法专家、病理学家与合规专家的意见。
6. 数据治理与记录归档	数据来源出处、处理步骤和分析决策应以详细、可追溯且可验证的方式记录，并符合 GxP 要求。在技术的整个生命周期中，应维持适当的治理，包括对敏感数据的隐私和保护，这是消除算法“黑匣子”的关键。
7. 模型设计与开发实践	AI 技术的开发遵循模型、系统设计及软件工程的最佳实践，并在考虑可解释性、说明性和预测性能的前提下，利用符合用途 (fit-for-use) 的数据。良好的模型和系统开发能提升 AI 技术的透明度、可靠性、泛化能力和稳健性，从而助力患者安全。
8. 基于风险的性能评估	应利用符合用途的数据和适用于预期使用场景的指标，对包括人机交互在内的整个系统进行评估，并辅以通过适当设计的测试与评价方法对预测性能进行的验证。
9. 生命周期管理	AI 技术的整个生命周期中应实施基于风险的质量管理体系，包括支持问题的捕获、评估和解决。AI 技术须接受预定的监测和定期重新评估，以确保足够的性能，比如应对数据漂移问题。
10. 清晰的关键信息	强调透明度，要求以通俗易懂的语言向目标受众（包括用户和患者）提供关于 AI 技术的使用场景、性能、局限性、底层数据、更新情况，以及可解释性或说明性的清晰、易获取且与场景相关的信息。

资料来源：FDA，招银国际环球市场

### 临床前评价演变：新方法论 (NAMs) 加快应用

传统药物研发长期依赖动物模型对候选药物的人体毒性与疗效进行预测，但受限于跨物种生物学差异，动物实验的转化率存在天然局限。2025 年 4 月，FDA 宣布计划将从 IND 申请开始，逐步减少在单克隆抗体及部分其他药物开发中的动物实验要求，并鼓励以新方法论 (NAMs) 数据作为替代性证据。NAMs 包括 AI 驱动毒性计算模型、人源细胞/类器官平台及器官芯片 (Organs-on-chips) 等体外系统。我们认为，与传统动物模型相比，NAMs 的优势在于能够以更接近人体生理机制的方式开展毒性识别与药效评估，AI 模型可模拟单抗在人体内的分布与作用过程，而基于人源细胞和类器官的体外平台可以揭示在动物模型中难以被充分识别的毒性信号，从而提升对人体反应的预测相关性。此外，NAMs 有望通

过减少动物实验，压缩非临床研究周期，从而提升研发效率并降低研发成本。FDA对NAMs的政策支持反映出计算生物学、AI及人源体外证据在监管评价体系中的重要性正在提升。中长期看，随着监管框架、验证标准及产业基础设施逐步完善，NAMs有望逐步走向关键性证据来源，推动AI制药由研发提效走向研发范式重构。

## 临床试验的数字化重构：AI辅助提升效率，监管保持开放谨慎

我们认为，当前欧美药品监管机构对于AI在临床试验阶段的应用总体持开放但审慎的立场，监管讨论较多的应用方向主要为利用AI技术提高试验效率，而非直接替代随机对照试验。比如，2022年Unlearn.AI的PROCOVA获得EMA的资格认定（Qualification Opinion），明确将其认证为可提高II期和III期临床试验效率的统计学方法。PROCOVA利用AI数字孪生预测患者未接受治疗的结果（即预后评分），在新随机试验的分析中将预后评分作为协变量，通过降低治疗响应估计值的方差，进而减少实现预期置信水平所需的最低样本量。我们认为，这主要说明监管机构对AI方法用于完善临床试验设计，提升临床效率持开放态度，但并不涉及对AI生成临床证据的采纳。2025年，基于机器学习的肝脏活检数字化病理分析系统AIM-NASH获得EMA和FDA的资格认定，用于NASH/MASH临床试验中评估疾病活动度（炎症、纤维化、肝细胞气球样变）。据FDA，药物开发者可在IND、NDA和BLA申请时，将AIM-NASH用于已获资格认定的特定使用情境，从而无需重新提交DDT信息或接受额外的CDER复审。这是全球首个获得CHMP和FDA正式资格认定的AI药物研发工具，标志着AI工具在特定使用场景下，已开始被监管机构接受为药物开发支持证据的一部分。

但值得注意的是，AIM-NASH的认定是建立在“人在回路”（Human-in-the-Loop）的要求之上，即病理学家需审查AI输出，并对最终检测结果负责。同时重大算法变更需重新资格认定。我们认为，这体现了欧美监管机构拥抱创新但不放弃临床专家最终决策权的核心原则，即AI在现阶段仍仅作为有限的辅助手段而非替代。

## AI制药行业核心壁垒在于数据

我们认为，AI制药的模型质量与高质量、大规模的数据基础息息相关，数据质量越高，数据量越大，越能准确地进行建模训练和预测。而当前支撑AI制药的核心数据资源主要为以下几类公域数据：1) 化学数据：ChEMBL数据库收录了超过两百万个具有生物活性注释的化合物，涵盖靶点结合亲和力、细胞活性、ADMET属性等多维度信息；ZINC数据库则提供了超过十亿个可商业获取的化合物三维结构，为虚拟筛选提供了庞大的搜索空间。2) 结构生物学：蛋白质数据银行（PDB）中已沉积超过二十万个实验解析的蛋白质三维结构，而AlphaFold2的横空出世更将结构覆盖范围扩展至超过两亿个预测结构，拓展了结构基础药物设计的可及性。3) 基因组与转录组：TCGA、GTEx、UK Biobank等大型队列数据集为疾病靶点的系统性挖掘提供了庞大数据支撑。这些公开数据库构成了当前AI制药行业的公共数据底座，也是大部分AI制药公司模型训练和技术验证的主要数据来源。

然而，公开数据库中的数据在质量与一致性方面存在不足。比如，ChEMBL等数据库中的生物活性数据来源于不同实验室、不同时期、不同测定条件下的实验，测定体系的差异也引入了大量系统性噪声，直接影响模型训练的可靠性。此外，公开数据在靶点覆盖方面存在不均衡性，热门靶点积累了较多活性数据点，而大量新兴靶点或不可成药靶点的公开数据则相对有限，为训练高精度的深度学习模型带来挑战。更值得注意的是，公开数据库提供的是药物研发过程中经过高度筛选和编辑后的结果性数据，即最终被发表或公开的活性化合物及其性质，而非研发过程本身的过程性数据。因此，大量在筛选中被证明无活性或存在问题的化合物，即阴性数据几乎从不被公开，而这些阴性数据对于训练模型准确识别

“什么样的分子不行”同样重要。发表偏倚 (publication bias) 导致的正样本过度代表问题，是公域数据训练的 AI 模型在实际应用中泛化性不足的重要原因之一。

当公域数据的局限性日益显现，我们认为，真正决定 AI 制药模型性能上限的是创新药研发的私域数据，主要包括 1) 药物化学过程数据，如制药企业和 CXO 在先导化合物优化过程中积累的完整构效关系 (SAR) 数据集。一个先导优化项目可能涉及大量化合物的迭代设计、合成与测试，每一轮优化都系统地记录了结构修饰与活性、选择性、代谢稳定性等多维属性之间的精细关联。这类数据的信息密度远高于公开数据库中零散的活性数据点，因为它们记录的是有经验的药化团队在明确目标引导下进行系统性探索的完整轨迹，对于模型学习“如何优化分子”具有极高的训练价值。2) 临床前安全性与 ADMET 数据：如 CXO 和制药企业在 GLP 毒理学研究、药代动力学研究中产生的大量结构化实验数据。这些数据涉及化合物在动物体内的吸收、分布、代谢、排泄行为及器官毒性表现，是训练高精度 ADMET 预测模型的理想数据来源，但几乎完全处于企业专有的数据围墙之内。3) 临床试验数据：如临床试验中收集的患者疗效与安全性数据、生物标志物数据、电子病历 (EHR) 以及真实世界数据。这些数据不仅对训练临床试验优化模型至关重要，更是连接分子设计与患者获益这一最终目标的关键信息桥梁。

然而，这些高价值的私域数据处于数据孤岛状态，目前难以被大规模整合，且标准化不足。大型制药企业视其积累的多维度研发及临床数据为核心商业资产，公开共享可能将削弱其竞争优势。因此，数据共享对制药企业而言是一个需慎重考虑的决定。CXO 企业虽然在服务过程中接触到大量跨客户的实验数据，但受限于严格的客户保密协议和知识产权归属条款，也无法将这些数据用于自身或第三方的 AI 模型训练。医院和临床研究机构掌握的患者数据则受到更为严格的隐私法规限制，如欧盟的 GDPR、中国的《个人信息保护法》以及美国的 HIPAA，跨机构的数据共享面临伦理审查、患者知情同意、数据脱敏等多重合规壁垒。数据孤岛的本质并非技术问题，而是由商业利益、知识产权保护和隐私监管三重约束共同决定的结构性问题，短期内难以寻找到系统性的解决方案。

因此，我们认为，AI 制药行业的竞争主要是数据资产的比拼。在公域数据对所有参与者均等可及、主流算法架构通过学术论文和开源代码快速扩散的情况下，私域数据的规模与质量正在成为各家 AI 制药企业之间最核心的差异化壁垒。

Recursion 自建大规模自动化湿实验平台，通过自有实验体系持续生成标准化的高通量表型组学数据。目前，Recursion 已积累了超 50 PB 的高质量多模态数据，涵盖了干实验和湿实验领域。自建实验平台虽然前期资本投入较大，但产出的数据在标准化程度、标注质量和覆盖系统性方面远优于从不同来源的公开数据，且完全归企业所有，构成难以复制的竞争壁垒。新一代 AIDD 公司如 Absci 和 Generate 也均通过自建湿实验室，实现分子生成-测试-实验数据反哺模型的闭环。

另一条路径是通过与制药企业的大规模合作获取专有数据。赛诺菲 (SNY US, 未评级) 与 Exscientia (现被 Recursion 收购) 签署的超过 52 亿美元合作协议，安进与 Generate Biomedicines 达成长期合作框架。在这些合作中，AI 制药企业通常可以在一定范围内接触到合作方的专有实验数据，并将合作项目中产生的新数据纳入自身平台的训练体系。每一次合作项目的执行，不仅产生直接的商业收入，更为 AI 平台贡献了新的高质量训练数据，从而持续优化模型性能，形成“合作越多、数据越多、模型越好、吸引更多合作”的正向循环。从这个角度而言，大型制药企业在数据壁垒方面占据着天然的结构优势。辉瑞 (PFE US, 买入)、罗氏 (RHHBY US, 未评级)、诺华 (NVS US, 未评级) 等全球头部药企在数十年的药物研发历程中积累了大量化合物的完整 SAR 数据集、临床前研究数据以

及临床试验数据，形成了强大的数据壁垒。当前，MNC 都在积极建设内部 AI 研发体系，因此，AI 制药公司面临的竞争不仅来自同业，还来自拥有数据优势的传统药企。

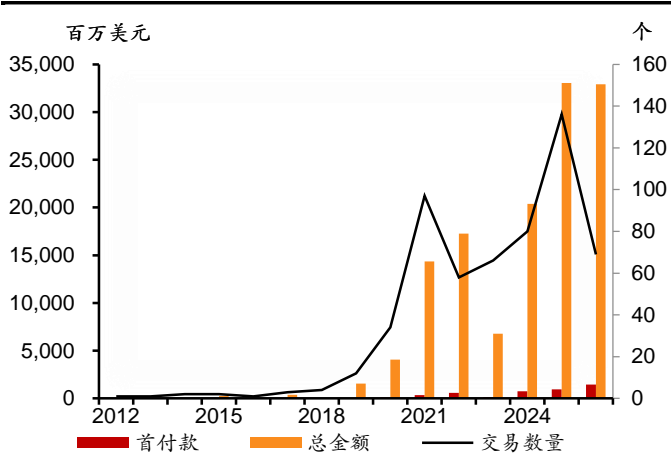
联邦学习 (Federated Learning) 通过模型参数共享，在不转移原始数据的前提下为实现跨机构的数据协同建模提供了可能。2019 年，包括安进、阿斯利康、拜耳、杨森等在内的十家制药企业联合 2 所大学、4 家初创公司和英伟达 (NVDA US, 未评级) 组成 MELLODDY 联盟，基于联邦学习技术在不共享化合物结构的前提下联合训练药物性质预测模型。但是，MELLODDY 项目的数据仅限于联盟内部使用，且仅持续了 3 年，于 2022 年终止。2025 年，礼来 (LLY US, 未评级) 宣布推出 Lilly TuneLab，基于联邦学习技术为生物技术公司提供根据礼来多年的研究数据训练的 药物发现模型。礼来估计，首次发布的 AI 模型包含超过价值 10 亿美元的专有数据。2026 年初，Schrödinger 宣布将 TuneLab 平台集成到其广泛使用的企业信息解决方案 LiveDesign 中，以解决数据稀缺这一常阻碍 AI 驱动药物发现的痛点。但我们认为，联邦学习在商业价值分配上仍存在难点，比如计算各参与方数据贡献度的方式，且此类合作的规模和深度目前仍非常有限，因此尚不足以从根本上改变数据孤岛的局面。

整体而言，我们认为，数据已成为 AI 制药领域最核心的竞争壁垒。在算法层面，学术界的快速迭代使得单纯的算法优势窗口通常较短，而高质量的专有数据资产则具有显著的时间壁垒和资本壁垒。庞大的标准化实验数据需要大量时间投入和资本支出，且难以通过捷径复制。

## AI 制药行业发展：传统药企的积极拥抱与 AI 制药公司的崛起

AI 制药产业格局正在经历快速变革。据医药魔方，2024 年及以前，AI 制药领域的合作尚处于尝试阶段，合作范围多局限于少数靶点或单一项目，平均前期付款多集中在 500-1,500 万美元，里程碑付款也极少突破 5 亿美元。而 2025 年以来，随着 AI 技术的持续突破，基于 AI 技术相关的医药交易数量和交易规模在 2025 年均实现大幅增长。

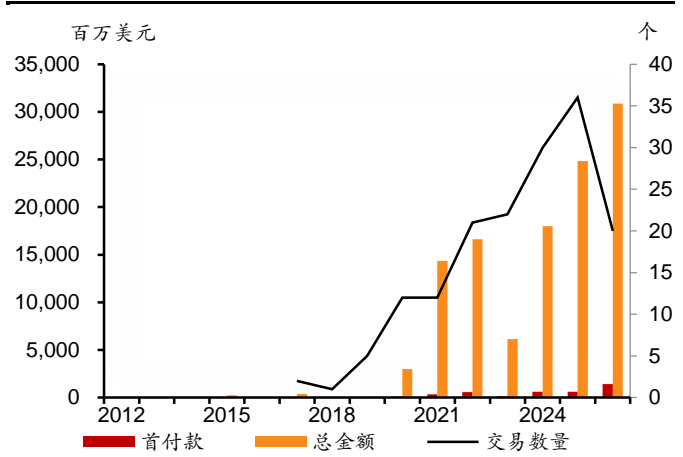
图 11: 基于 AI 技术相关的医药交易



资料来源：医药魔方，招银国际环球市场

注：按“AI 技术”检索，首付款和总金额均仅统计已披露情况。数据截至 2026 年 4 月 8 日

图 12: TOP20 MNC 基于 AI 技术相关的医药交易



资料来源：医药魔方，招银国际环球市场

注：按“AI 技术”、“引进方为 TOP20 MNC”检索，首付款和总金额均仅统计已披露情况。数据截至 2026 年 4 月 8 日。

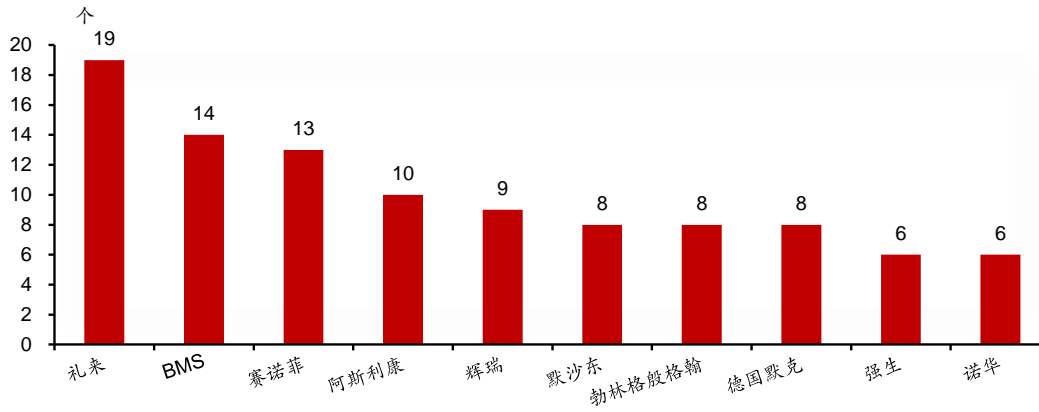
我们认为，面对专利悬崖的巨大压力，AI 技术凭借其在缩短研发周期的显著优势，正成为除外部资产引进外，MNC 加速管线扩张、对冲专利到期风险的又一核心战略。MNC 通过

内部建设+外部合作双轮驱动，加速构建以 AI 为基础的整体解决方案。因此，AI 制药领域的商业化合作正在快速发生结构性变化：

- 1) 单笔交易规模快速增长。据医药魔方，AI 制药领域的前期付款金额普遍增长至 2,000-6,500 万美元甚至更高，里程碑金额超 10 亿美元的交易越来越多。
- 2) 合作周期被显著拉长，从靶点发现延伸至临床前。
- 3) 合作领域随着研发热点和 AI 制药应用边界的扩大而逐渐丰富，从早期的小分子领域合作拓展至抗体等领域。比如，2025 年 4 月，赛诺菲以 1.25 亿美元预付款+最高达 18.45 亿美元的总交易额，引进了华深智药两款双抗 HXN-1002 ( $\alpha 4\beta 7/TL1A$ ) 与 HXN-1003 (靶向 TL1A/IL-23) 的全球独家权益。
- 4) 合作项目类型逐渐丰富。从较为单一的基于 AI 技术平台的研发合作，如基于特定靶点或适应症的药物发现，拓展至基于特定管线的授权合作和基于试验数据的合作等。

据医药魔方的数据，2020 年至今，礼来、BMS (BMY US, 未评级)、赛诺菲、阿斯利康及辉瑞位居 MNC 药企 AI 交易活跃度第一梯队。其中，礼来以显著领先的成交频次成为该领域的领跑者。

图 13: 2020 年至今 TOP20 MNC 达成 AI 相关交易合作的数量



资料来源: 医药魔方, 招银国际环球市场

注: 数据统计截至 2026 年 4 月 8 日。

因此, 我们以礼来为例, 可以看出 MNC 药企正通过“外部合作+内部建设”将 AI 从单点工具升级为数据+算力基建+平台技术+管线的 AI 生态整体解决方案。一方面, MNC 药企积极与 AI 原生公司达成合作, 利用他们的 AI 技术平台进行靶点和药物发现。另一方面, MNC 药企也重视内部 AI 能力的建设, 从依赖外部算法转向构建私有化模型与数据护城河。比如, 礼来投入算力基础设施建设, 并与 Chai Discovery 合作开发独家使用的定制 AI 模型。阿斯利康自主研发的 AI agent 系统 AIDA 有望将 CMC 开发周期缩短 50%。2026 年初, 阿斯利康还收购 Modella AI, 并计划将 Modella 的基础模型与 AI 智能体全面整合进自身肿瘤研发体系, 重点用于支撑临床开发与生物标志物发现工作。数据层面除过往积累的庞大药物研发数据外, 礼来还引入 BigHat Biosciences, 利用其平台快速生成多样化、高质量的数据集, 以反哺礼来内部的基础模型。此外, 礼来重视高潜力外部管线资产的收并购, 积极引进具有 BIC 或 FIC 潜力的 AI 药物。

图 14: 礼来具有代表性的 AI 相关合作或内部体系建设

年份	合作方	主要内容	意义
2019	Atomwise	利用 Atomwise 的 AI 技术支持礼来的临床前药物发现工作, 将开展最多十个由礼来选定的药物靶点项目, 旨在加速新药的发现和开发。	早期 AI 相关合作
2022	Genesis Therapeutics	利用 Genesis Molecular AI Platform 人工智能平台, 共同发现针对多达五个靶点的新型疗法。	
2023	Fauna Bio	利用其专有的 Convergence™ 人工智能平台, 成功识别出一个潜在的同类首创 (first-in-class) 肥胖治疗靶点。	布局 AI 赋能靶点发现, 获得创新管线资产
2024	Isomorphic Labs	利用专有的技术平台、下一代 AlphaFold 进行多个靶点的小分子发现。	布局 AI 赋能小分子药物发现, 攻克高难度药物靶点
2024	Genetic Leap	利用 Genetic Leap 的 RNA 靶向 AI 平台, 针对礼来在高优先级治疗领域选定的靶点, 开发寡核苷酸药物。	布局 AI 赋能基因药物开发
2024	OpenAI	利用生成式 AI 技术开发新型抗菌药物, 用于治疗耐药性病原体。	布局 AI 赋能药物发现
2025	内部建设 英伟达	基于 NVIDIA Blackwell 架构 DGX SuperPOD, 配备超过 1000 块 GPU, 部署全球制药行业规模最大、算力最强的自主运营 AI 工厂, 用于药物发现并缩短药物研发周期。	算力基础设施投入
2025	内部建设	推出 AI 和机器学习平台 TuneLab, 为生物技术公司提供基于礼来多年研究数据训练的药物发现模型访问权。	自建平台
2026	BigHat Biosciences	利用其在机器学习领域的专业知识和 Milliner™ 平台, 快速生成多样化、高质量的数据集, 用于通用抗体可开发性基础模型, 以支持 TuneLab 平台的持续发展。	布局数据资产

2026	Chai Discovery	采用 Chai Discovery 的前沿 AI 平台，为多个靶点设计新型生物治疗药物。Chai Discovery 还将专为礼来开发独家使用的定制 AI 模型，该模型将利用礼来大规模专有数据进行训练，并针对礼来的发现工作流程进行优化。	布局 AI 赋能大分子药物开发，及相关模型
2026	内部建设 英伟达	在未来五年内联合投入 10 亿美元，在美国旧金山建立 AI 联合创新实验室，通过生成大规模数据，并依托英伟达 BioNeMo™ 平台，构建强大的 AI 模型，从而加速新药研发。	布局模型开发与前沿技术创新
2026	英矽智能	获得了临床前阶段针对特定适应症的潜在同类最佳新型口服疗法的全球许可，并将利用 Pharma.AI 平台针对礼来选定的靶点进行合作开发。	布局 AI 赋能小分子药物发现，获得创新管线资产

资料来源：医药魔方，招银国际环球市场

整体而言，传统药企正成为 AI 制药行业的重要参与方。但对 AI 制药行业，尤其是 AI 药物发现等技术产生更大推动作用的，是一批从创立首日便基于计算逻辑和 AI 模型架构的“AI 原生 (AI-Native)”公司的崛起，即通常意义上的 AI 制药公司。

图 15: AI 原生与传统 AI 赋能型企业的比较

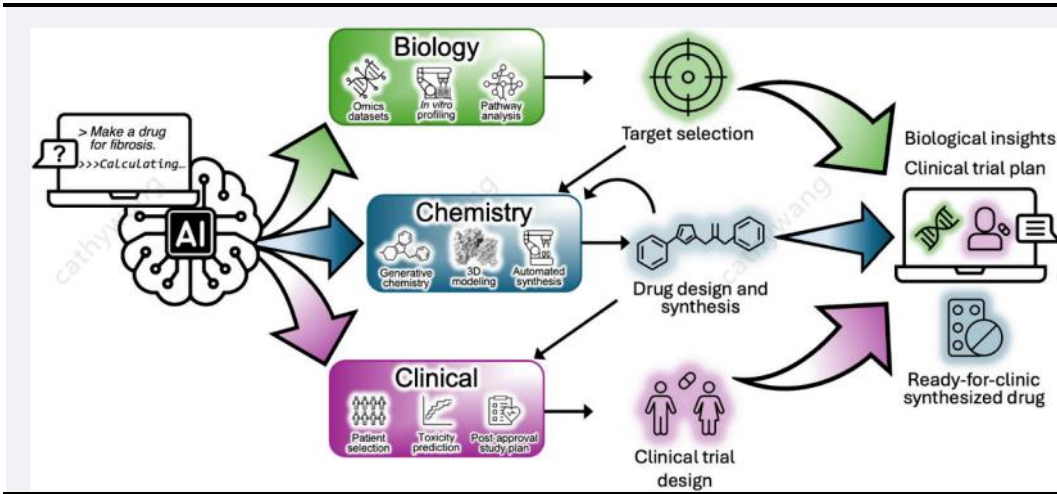
维度	AI 赋能型企业 (传统药企转型)	AI 原生型企业 (如 Recursion, Generate)
底层架构思维	<b>资产驱动</b> : AI 是解决管线瓶颈、应对专利悬崖、提升研发效率的重要强力工具。	<b>平台驱动</b> : 建立“干湿实验闭环”的操作系统 (OS) 是核心。药物管线是平台能力的验证。
数据护城河	私有研发和临床数据+真实世界研究: 在过往研发中积累了大量药物开发、临床及患者随访数据, 但数据质量参差。	通过自动化实验室产生从零开始标注、一致性极高的机器可读数据, 以专门喂养 AI, 但数据量需时间积累。
成本结构	传统湿实验占比大。AI 主要用于降低“失败成本”。	前期需重投资算力基建和自动化实验室, 但管线扩展的边际研发成本低于传统模式。
核心优势	临床+监管+商业化全栈能力, 以及资金优势。	算法更先进, 且在持续升级迭代 AI 平台。
主要劣势	AI 或 IT 与传统研发团队协作存在“语言鸿沟”; 私有数据清洗和标准化难度较大。	缺乏后期临床经验, 商业化策略上相对薄弱, 抗风险能力依赖融资或管线授权。
战略重心	外部合作+内生基建。	自有管线临床验证+对外授权获得收入。

资料来源：招银国际环球市场

以 Schrödinger 为代表的早期 AI 制药公司侧重于基于物理规则的分子对接，比如 Schrödinger 的核心技术是基于物理学第一性原理的分子模拟平台，FEP+ (Free Energy Perturbation Plus) 是当前精度最高的结合自由能预测工具之一。随后，以 Atomwise 等为代表的 AI 制药公司尝试通过干实验模式 (计算与算法) 预测药物活性，但因生物学系统的复杂性及公共数据库的“噪声”问题，面临临床转化率的严峻挑战。正如我们前文的观点，我们认为 AI 制药的核心壁垒在于数据，而以 Recursion 为代表 AI 制药企业不再被动依赖外部数据，而是构建了自动化湿实验室，将干实验与湿实验相结合。持续运行的湿实验，产生大量高质量且标准的专有数据集，可用于训练 AI 模型。而 AI 模型则可指导高通量湿实验，再产生新数据用于模型迭代，干湿闭环可形成自我强化的数据飞轮。

生成式人工智能与自动化实验室系统的融合带来的药物发现范式变革，有望在未来实现“从提示到药物 (prompt-to-drug)”的全自主药物开发管线愿景。据 ACS Central Science 的报道，这类 AI 药物发现集成系统的理想工作流是由 AI 模型进行靶点发现和药物发现，然后返回经合成和湿实验验证，确定进入临床研究的候选药物，同时提供临床试验和上市后监测计划 (Alex Zhavoronkov et al., 2026)。

图 16: 全自主药物开发示意图



资料来源: ACS Cent. Sci., 12, 3, 265-279.(2026), 招银国际环球市场

随着 Recursion、Schrödinger、Absci 等头部企业的核心管线均已陆续进入临床阶段并读出数据，我们认为 AI 制药公司的市场关注焦点已从 AI 平台技术本身转向管线资产及临床验证。换言之，当前 AI 制药公司更具备 Techbio 的特点，以专有的 AI 平台作为核心技术，但最终价值需通过自研管线资产体现。这也表明，市场对于 AI 制药的价值期待正从早期的“效率提升”转向“临床成功率提升”。

图 17: 全球主要 AI 制药企业的核心管线

公司	最高临床阶段	研发代码/通用名	靶点	创新属性	分子种类
Absci	I/II 期	ABS-201	PRLR	AI 生成分子	单抗
Recursion	I/II 期	REC-1245	RBM39	AI 靶点+AI 药物发现	降解型分子胶
Schrödinger	I 期	SGR-1505	MALT1	AI 优化分子	小分子抑制剂
Generate Biomedicines	III 期	GB-0895	TSLP	AI 优化筛选已知 TSLP 抗体	单抗
英矽智能	II 期	Rentosertib	TNIK	AI 靶点发现+AI 生成分子	小分子抑制剂

资料来源: 医药魔方, 招银国际环球市场

注: Nimbus 与 Schrödinger 共同开发了 TYK2 抑制剂 zasocitinib, 并于 2023 年授权给武田制药 (4502 JP, 未评级), 目前已取得针对中重度斑块状银屑病的临床 3 期顶线结果。但因为合作开发, 所以未纳入上表的统计。

此外, 多家 AI 大模型的领军企业也纷纷瞄准 AI 在药物研发领域的落地应用。比如 2026 年 4 月, OpenAI 推出专为生物学和药物研发打造的垂直领域推理模型 GPT-Rosalind, 用于提升早期研发效率, 包括靶点筛选、假说构建和实验方案设计。据报道, OpenAI 正与安进、Moderna (MRNA US, 未评级)、艾伦研究所、赛默飞 (TMO US, 买入) 等客户合作, 将 GPT-Rosalind 应用于加速研究和发现的工作流程。Anthropic 在 25 年 10 月也曾推出专为生命科学领域定制的 Claude for Life Sciences, 并在 2026 年 4 月以 4 亿美元收购 AI 生物技术初创公司 Coefficient Bio, 以快速补强药物研发专用模型, 包括实验设计、药物筛选、协助制定临床试验监管策略等。

## AI 制药公司商业模式的迭代

复盘过去 10 年，我们认为 AI 制药公司的商业模式经历了较为清晰的演进路径。

### 第一阶段（约 2018 年以前）：SaaS 服务时期。

早期 AI 制药企业普遍缺乏湿实验与临床推进能力，核心商业模式是以出售计算服务或软件为主。Schrödinger 是这一模式的传统标杆，即向全球制药、生物技术与材料企业销售计算软件，从买断模式逐步转变为以 SaaS 模式授权其计算平台，获得稳定订阅收入。这一阶段 Schrödinger 的业务模式具有轻资产、低风险但增长天花板受限的特征。在这一时期，部分头部公司也陆续尝试与药企合作，提供技术服务的商业模式。比如，2013 年，Exscientia 与杨森制药达成合作，应用其设计技术，助力杨森制药发现新型化合物。2015 年，Schrödinger 与赛诺菲达成合作，为最多十个药物发现项目提供计算药物设计支持。但此时，每年达成的 AI 领域交易或合作极少，还处于早期探索阶段。

### 第二阶段（2019 至 2021 年）：平台+合作开发时期。

2017-2018 年，Transformer 架构的提出，以及 BERT、GPT-2 等预训练模型刷新 NLP 纪录，AI 技术迎来跨越式发展。随着 AI 模型能力的提升和早期验证数据的积累，越来越多的 AI 制药公司与传统药企达成合作，基于特定适应症或靶点进行合作开发。2020 年初，Exscientia 与住友制药共同研发的 DSP-1181 进入临床 I 期，成为全球首个进入临床试验阶段的 AI 设计分子，成为 AI 制药发展历程的重要里程碑，也很大程度上引发了传统药企的合作热情，AI 相关的研发合作数量呈现持续增长趋势。这一阶段，AI 制药公司与传统药企的合作并不针对特定管线，业务模式更类似于 AI+CRO，但常见的收入结构为“首付款+里程碑付款+销售分成”，与传统 Biotech 的 BD 交易结构趋同。因此，我们认为 AI 制药公司区别于传统 CRO 公司的主要特点，在于 AI 制药公司通常作为合作伙伴，可参与并分享新药研发推进及商业化上市后的价值红利释放，而非仅仅作为技术供应商。

### 第三阶段（约 2021 年至今）：向 Techbio 模式过渡，并走向多元。

随着融资能力的增强和临床前数据的积累，AI 制药公司普遍开始投入自有管线，多款 AI 优化药物或 AI 生成药物陆续进入临床阶段。比如，2021 年英矽智能利用 AI 平台进行靶点发现并生成的 TNIK 抑制剂 rentosertib 进入 I 期临床，Exscientia 利用 AI 平台优化的 A2aR 拮抗剂 EXS-21546 也进入临床 I 期。截至目前，Absci、Recursion、英矽智能等领先的 AI 制药企业均有多个管线进入临床阶段。AI 制药公司与传统药企的合作开发开始进入管线兑现期，合作模式也逐步拓展至就某一特定药物的授权合作，即 biotech 常见的 BD 交易合作，收入构成为“首付款+里程碑付款+销售分成”。比如，2021 年 BMS 行使选择权，获得 Exscientia 涉及的一款免疫调节候选药物的许可。2024 年，英矽智能与 Menarini Group 达成合作，授予其新型小分子 KAT6 抑制剂候选药物 ISM5043 的全球独家开发和商业化权益。同时，也涌现了一批专注内部药物研发的新一代 AI 制药企业。我们认为，这一阶段的核心转变是 AI 公司开始承担药物研发风险，其估值逻辑也逐步开始从单一“技术平台价值”向“平台价值+管线临床价值”转变。

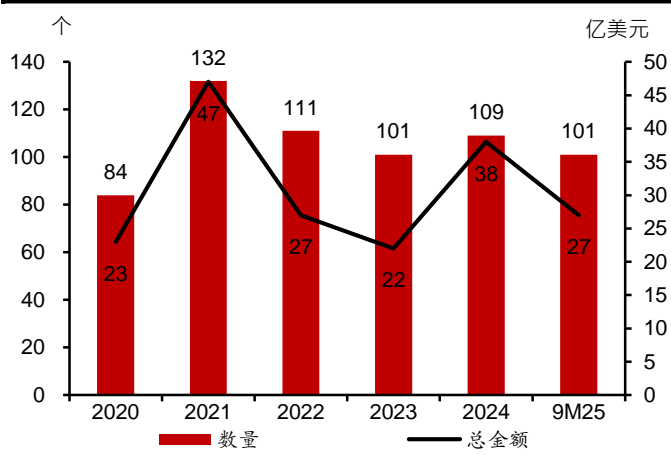
值得注意的是，全球尚无完全由 AI 生成的药物完成临床 III 期验证或实现商业化闭环。受此影响，传统药企对直接引进 AI 原生产资产仍持审慎态度。现阶段的对外授权交易的管线多集中于 PCC（临床前候选化合物）阶段，这导致了早期交易中首付款比例受限。鉴于创新药研发天然的高风险、长周期属性，一旦核心管线在临床阶段遇挫，依赖里程碑收入的 Techbio 企业将面临财务风险。

在此背景下，行业出现分化，部分企业选择“AI+CRO”模式，利用AI平台为传统药企提供技术服务，以期获得稳定收入。比如Verge Genomics在2026年初宣布转型，将转型为一家为大型药企提供AI技术服务的平台公司。而更多的企业则选择兼顾Techbio和“AI+CRO”模式，即兼顾1)内部管线的持续推进，2)在研管线的对外授权，以及3)与药企达成研发战略合作，主要聚焦早期靶点或药物发现。此外，2025-2026年，行业涌现了一批区别于传统BD授权的新型合作范式，即为传统药企提供AI模型的定制化开发。比如，Chai Discovery与礼来合作，帮助礼来开发前沿AI平台；Boltz与辉瑞达成合作，在开源基础模型的同时，为辉瑞提供融合其专有数据的专属版本。

### 全球资本从狂热走向理性验证

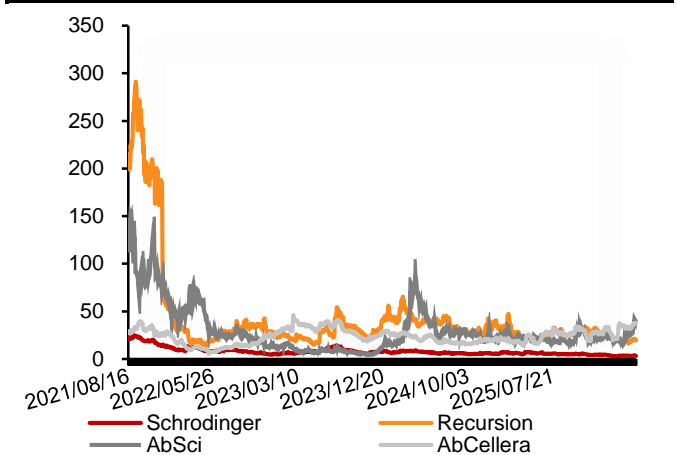
全球AI制药市场展现出强劲的增长潜力。据弗若斯特沙利文，2023年，全球AI赋能药物研发市场规模达119亿美元，预计到2032年将增长至746亿美元，年复合增长率达22.6%。过去十年间，深度学习和生成式人工智能的爆发吸引了巨额资本涌入全球人工智能驱动的药物研发（Artificial Intelligence in Drug Discovery，简称AIDD）行业，而当前全球市场正经历理性价值重估。据智药局的报道，全球AI药物研发领域的融资活跃度在2021年达到顶峰，全年融资笔数达132起，融资总金额达47亿美元，随后有所回落并进入震荡期。我们认为，这主要是由于2020年，以Schrödinger为代表的AI制药公司开始陆续上市，AI制药投资的退出渠道变得明确，一时间点燃了资本的热情。但从2022年开始，多家公司的在研管线出现挫折，AI制药领域的投资回归理性。

图 18:全球 AI 药物研发领域 VC 投资规模



资料来源: Morningstar, 智药局, 招银国际环球市场

图 19:美股 AI 制药公司动态市销率



资料来源: 彭博, 招银国际环球市场, 截至 2026 年 4 月 23 日收盘

我们认为，全球AI制药行业已跨越了以算法叙事驱动估值的早期阶段，进入了以管线后期临床数据为核心焦点的深水区。尽管据医药魔方统计，2025年全球已披露AI技术相关的医药交易总金额突破330亿美元，但首付款占比仍较低。若以已披露首付款金额和交易总金额粗略计算，2025年首付款占比仅不到3%。纵览医药魔方统计的全球AI相关合作项目，首付款规模排名前10的项目主要发生在近3年，但首付款占比仍仅约5%。这揭示了传统药企对AI产品临床交付的谨慎态度。

图 20: 全球 AI 相关合作或交易首付款 TOP10 的项目

时间	转让方	受让方	首付款金额 (百万美元)	总金额 (百万美元)	首付款占比
2026.1.30	石药集团	阿斯利康	1,200	18,500	6%
2022.5.10	Evotec	BMS	200	5,000	4%
2024.12.13	石药集团	百济神州	150	1,835	8%
2021.12.7	Recursion	罗氏	150	12,150	1%
2024.10.28	Monte Rosa Therapeutics	诺华	150	2,250	7%
2025.4.17	华深智药	赛诺菲	125	1,845	7%
2025.9.15	Monte Rosa Therapeutics	诺华	120	5,700	2%
2026.3.29	英矽智能	礼来	115	2,750	4%
2026.2.18	Unnatural Products	诺华	100	1,800	6%
2024.10.7	石药集团	阿斯利康	100	2,020	5%
平均					5%

资料来源: 医药魔方, 招银国际环球市场

注: 部分交易未披露总金额, 故未纳入上表统计。MeiraGTx 与 Hologen AI 达成的战略合作, 主要形式为成立合资公司, 并向该合资公司注册, 因此我们认为不是通常意义上的首付款与交易总金额, 未纳入上述统计。

此外, 多家企业在 2025 年因现金流紧张而宣布裁员并精简管线, 包括但不限于 1) Schrödinger 宣布裁员 60 人, 约占其员工总数的 7%; 2) Recursion 宣布暂停或取消优先开发 5 款临床药物, 并裁员 20%; 3) Relay 宣布裁员 70 人, 并终止 SHP2 变构抑制剂 RLY-1971 的开发。2024 年, 基因泰克曾终止与 Relay 就 RLY-1971 达成的合作并退回权益。

我们认为, 尽管在早期发现阶段, AI 制药已证明了其对于研发效率的提升, 但在推进至后期临床阶段时, AI 制药行业正面临着挑战。主要原因可能在于精准的三维结构预测能力并不完全等同于分子的成药性, 也不意味着该靶点在复杂人体病理网络中具有治疗意义。三期临床试验主要考察药物在真实、异质性极高的患者群体中的有效性, 本质上是一个复杂的系统生物学和病理学验证过程, 而远非单纯的计算化学或结构预测问题。因此, 在首个 AI 新药获批上市之前, 我们认为 AI 制药行业仍处于概念验证阶段, 研发速度和效率的提升不再是衡量 AI 平台价值的唯一标尺, 能够产出确切的临床获益数据, 才是 AI 技术在药物研发领域应用的“试金石”。但随着 AI 制药管线临床试验的推进, AI 制药的分水岭可能已经临近, 多家 AI 制药公司的管线将在 2026 年迎来数据读出, 将为 AI 制药行业提供重要的验证, 并成为推动行业发展的重要催化剂。

图 21: AIDD 领域未来重要催化剂

公司	里程碑
Recursion	REC-1245 (RBM39) 计划于 2026 年上半年公布 I 期临床早期数据。REC-4881 (MEK1/2) 预计 2027 年上半年公布 Ib/II 期的更新数据。此外, 预计 2027 年将有三款药物 (REC-617, REC-4539, REC-3565) 公布 I 期临床早期数据。
Absci	ABS-201 (PRLR) 计划于 2026 年上半年公布 I/IIa 期临床早期安全性、耐受性和药代动力学 (PK) 数据, 并在 2026 年下半年和 2027 年初公布中期及完整概念验证 (PoC) 数据。
Schrödinger	SGR-3515 (Wee1/Myt1) 计划于 2026 年公布 I 期临床早期数据。
英矽智能	ISM001-055 (TNIK 抑制剂) 计划于 2026 年启动中国临床 III 期试验。
AbCellera	ABCL635 (NK3R) 预计将于 2026 年第三季度公布针对中重度更年期血管舒缩症 (VMS) 的 I/II 期试验结果。

资料来源: 各公司财报及电话会, 招银国际环球市场

注: AbCellera (ABCL US, 未评级)

## 中国 AI 制药行业：从追赶到引领

在中国政府的政策鼎力支持、庞大且极具性价比的高级研发人才红利、以及香港联合交易所 18A 生物科技上市规则等全球多元化资本市场的强力助推下，中国 AI 制药企业呈现出有别于欧美的独特商业活力与估值溢价体系。凭借在软硬件一体化工程实施能力、极速的临床推进效率上的显著优势，中国头部企业已经从简单的技术追随者，正逐步成长为全球医药研发基础设施与源头创新的重要供给方。

### 结构性重塑与“数据引力”的战略崛起

据 WIPO 在 2025 年发布的《Patent Landscape Report: Generative Artificial Intelligence》，2014-2023 年，全球生成式 AI 在生命科学领域共有 5,346 个专利家族，其中中国拥有 3,181 个，占比达约 59.5%。而中国生物科技企业在全球对外授权（Out-licensing）交易总金额中的份额也快速攀升，据 SynBioBeta，已从 2021 年的约 10% 跃升 2025 年的约 50%。

展望未来，我们看好中国 AIDD 技术的长期发展潜力。中国拥有庞大的患者基数和日益完善的电子病历及多组学数据库，这为训练具有超高参数量的疾病生物学大模型提供了丰富的数据源。在临床试验规模方面，据 SynBioBeta 的数据，2021 年中国的临床试验总量便已超过美国，并在此后逐年扩大领先优势；2025 年，中国共开展了约 7,100-7,700 项临床试验，显著高于美国的约 6,200-6,300 项。如果对于临床试验数据加以有效利用，中国的海量临床数据有望加速 AI 算法的迭代。

### 出海授权 (License-out) 成为 AI 平台的初步验证

当前，中国 AIDD 企业不再局限于扮演计算外包服务商 (In-silico CRO) 的角色，而是以 First-in-Class 或 Best-in-Class 源头创新者的身份，将经过充分平台验证的高价值管线的海外开发与商业化权益授权给欧美跨国药企 (MNCs) 或生物制药公司。例如，晶泰控股 (XtalPi, 2228 HK, 买入) 在 2025 年与 DoveTree 达成了总潜在价值超过 60 亿美元的重大管线合作协议，覆盖肿瘤、神经系统等多个大病种领域；英矽智能 (Insilico Medicine) 则与赛诺菲、礼来 (Eli Lilly) 等产业巨头达成了多项预付款达数千万至上亿美元的战略药物发现合作。这些具有标志性意义的交易不仅为本土企业带来了充沛的现金流，更证明了中国企业的 AI 底层架构算法与新靶点发现能力已经获得了全球 MNC 以及 biotech 企业的背书。

## AI 制药代表企业

### Recursion Pharmaceuticals (RXRX US, 未评级)

图 22: Recursion 核心信息

股票代码	RXRX US
成立时间	2013
2025 年收入 (百万美元)	75 (+27% YoY)
2025 年净亏损 (百万美元)	645
截至 2025 年底在手现金 (百万美元)	754

资料来源: Recursion, 招银国际环球市场

注: 在手现金包括现金、现金等价物及受限现金。

Recursion Pharmaceuticals 是全球 Techbio 领域最具代表性的公司之一。公司围绕 Recursion OS 构建了一个原生 AI 驱动的端到端的药物发现与开发平台，将大规模自动化湿实验与 AI 算法深度整合，以表型优先策略推动药物重定位和全新药物发现。2024 年 11 月完成收购 Exscientia 后，公司进一步强化了小分子生成能力。

Recursion OS 平台将传统的药物研发流程重塑为三个阶段，且每个阶段都将自动化湿实验与内部计算分析相结合。湿实验室生成大量标准化的高质量生物和化学数据，而干实验室则运用机器学习、基于物理的建模和统计推断，从数据中提取有效信息。这种紧密集成使得湿实验结果能够直接训练模型，而模型输出又能够指导后续湿实验，从而加快迭代速度。与传统药物发现流程相比，Recursion OS 在成本和时间方面具备显著优势，据披露，Recursion 平均每个项目仅需合成功约 330 个化合物，即可在约 17 个月内获得候选化合物，而行业平均水平分别为超过 2,500 个化合物和 42 个月。

### 1. Biology to Insight (从生物学到洞察)

这一阶段的核心目标是将复杂的生物学信号转化为可操作的研发洞察。Recursion 通过整合大规模细胞表型组学、高通量转录组学、计算机模拟结合亲和力预测及真实世界患者数据，系统性地识别疾病相关机制、验证靶点并优先排序具有更强转化潜力的机会。

Recursion 拥有全球生物技术领域规模最大的自动化实验室之一，每周可进行数百万次实验，如将多种人类细胞用化合物或基因敲除处理后在高内涵显微镜下成像 (HCI)，获得疾病状态或细胞受到的扰动反应。高内涵显微镜是一种自动化、计算机辅助成像系统，可将先进显微镜与图像分析软件相结合，在近 50 种细胞类型上提取大量表型数据。通过比较这些表型数据，Recursion 的 AI 平台可以在作用机制未知的情况下，识别出某种干预措施/治疗手段可扭转/恢复特定疾病相关的细胞表型，从而使得平台能够发现具有治疗潜力的活性分子或遗传修饰因子，产生假设驱动方法可能遗漏的苗头化合物。据披露，Recursion 已积累了超 50 PB 的高质量多模态数据，涵盖了干实验 (计算) 和湿实验 (生物) 领域。

### 2. Insight to Molecule (从洞察到分子)

当系统锁定了具有潜力的生物学机制或靶点后，这一阶段的目标是进行分子设计，从而获得候选药物。Recursion 以 Centaur Chemist 为核心，可基于靶点特征谱构建新型化学结构并利用合成感知设计前置合成可行性评估。2025 年，Recursion 利用合成感知设计虚拟生成了超 1 亿个分子。此外，公司还拥有专有模型，可用于预测靶点-配体相互作用以及效力、选择性和 ADMET 性质。借助 Centaur Chemist 和化学自动化实验室，公司打造了 DMTL (设计-制造-测试-学习) 闭环，大幅提升从靶点到候选化合物的转化效率。截至 2025 年，Recursion 已设计出超 10 个开发候选物。

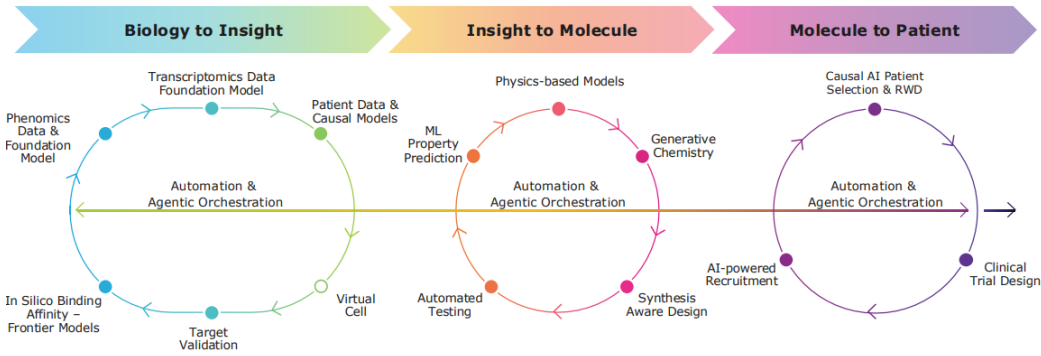
### 3. Molecule to Patient (从分子到患者)

药物进入临床阶段后，这一阶段的目标是提高临床试验的效率，扩大患者可及性，并最终提升临床成功的概率。通过 ClinTech 平台，Recursion 整合全球临床试验中心情报数据、大规模真实世界患者数据及因果推断、模拟和智能体自动化技术，优化试验设计并加速患者入组。

通过 AI 驱动的患者招募和临床试验设计，Recursion 在临床执行层面已经展现出显著的早期优势：1) 符合入组条件的患者群体增加了 10%-40%。2) 临床入组率提高了约 1.5 倍。3) 试验中心和国家的选择过程从过去的数月大幅缩短至数小时。

图 23: Recursion OS 平台示意图

### Recursion OS: An AI-native, full stack platform for drug discovery and clinical development



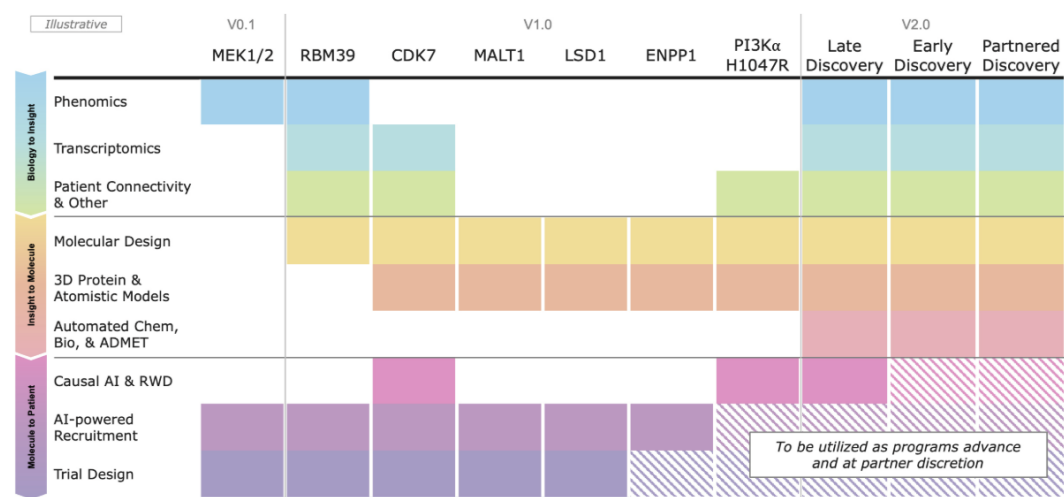
资料来源: Recursion, 招银国际环球市场

#### AI 技术落地与临床管线:

Recursion 的平台价值主要有三个方面: 1) 利用 AI 表型组学平台发现新的生物学机制和药物靶点; 2) 针对新兴的生物学机制, 利用 AI 优化药物设计和开发策略; 3) 针对已被验证的生物学靶点, 用 AI 优化药物设计以解决既往药物的局限性。但公司的管线推进并非一帆风顺。REC-994 曾是公司进展最快的管线, 由 Recursion 的 AI 平台识别出能够逆转脑海绵状血管瘤的细胞缺陷, 并推进至 II 期临床试验。然而, 2024 年 9 月, Recursion 公布了 REC-994 的二期临床 (SYCAMORE) 数据。虽然达到了安全性终点, 且在 MRI 影像上显示了病灶缩小的趋势, 但缺乏明确的患者或医生报告获益。2025 年 5 月, 公司宣布终止 REC-994 项目。后续的长期延伸试验数据显示, MRI 上的积极趋势未能维持, 且功能性结果与不接受任何治疗的安慰剂组无异。另一个用于治疗 2 型神经纤维瘤病的早期项目 REC-2282, 在 II 期试验后也因疗效不足而被终止。目前, 公司将核心管线精简至 7 个, 其中 5 个已推进至临床阶段。

图 24: Recursion 的主要管线以及 AI 平台在管线中的应用情况

Target	Disease Indication	Late Discovery	Preclinical	Phase 1/2	Phase 3	Potential Milestone
REC-4881	MEK1/2 Familial adenomatous polyposis (FAP)					Initiate FDA engagement - 1H26
REC-617	CDK7 Advanced solid tumors					Combination data - 1H27
REC-1245	RBM39 Biomarker-enriched solid tumors & lymphoma					Ph 1/2 data - 1H26
REC-3565	MALT1 B-cell malignancies					Ph 1 data - 1H27
REC-4539	LSD1 Solid tumors & hematology oncology					Ph 1 trial start - 1H26
REC-7735	PI3Kα H1047R Solid tumors (incl. HR+ breast cancer)					Go/no-go decision - 2H26 <sup>1</sup>
REC-102	ENPP1 Hypophosphatasia (HPP)					Go/no-go decision - 2H26 <sup>1</sup>



资料来源：Recursion，招银国际环球市场

REC-4881 (MEK1/2 口服抑制剂) 是 Recursion 目前临床进展最快的管线。虽然 MEK 是一个竞争较为激烈的已知靶点，但 Recursion 在与武田合作期间，采用 AI 驱动表型筛选方法从其化合物库中发现 MEK1/2 抑制剂，可以以浓度依赖性的方式逆转细胞的 APC 基因失活表型。因此，Recursion 从武田引进了 REC-4881，并将其重新定位用于治疗家族性腺瘤性息肉病 (FAP)。FAP 是由 APC 基因常染色体显性失活突变引起的罕见病，影响美国和欧盟超过 50,000 名患者。患者的胃肠道会长满息肉，如不反复通过切除干预，就会发展为结肠癌。目前尚无获批的药物可以阻止或减缓息肉生长，存在巨大未被满足医疗需求。2025 年 12 月，Recursion 公布 REC-4881 在针对 FAP 的 Ib/II 期临床试验 (TUPELO) 的更新数据。安全性方面，REC-4881 的安全性特征与 MEK1/2 抑制剂一致，不良事件主要为低级别，有 4 例停药。疗效方面，治疗 12 周时，在 12 例可评估患者中，75% 的患者出现结肠息肉总负荷减少，中位减少幅度为 43%；停药 12 周后，82% 可评估患者维持了持久的息肉总量减少，中位减少幅度达基线水平的 53%。我们认为，这在一定程度上为 AI 技术应用于新药筛选提供了临床验证。据公司披露，预计将于 2026 年上半年与 FDA 沟通以确定注册路径，并将在 2027 年上半年公布临床更新数据。

REC-1245 是具备 First-in-class (FIC) 潜力的 RBM39 分子胶降解剂，RBM39 是一种关键的 RNA 结合蛋白，参与选择性剪接和 DNA 损伤修复 (DDR) 通路。利用 Recursion OS，公司发现了全新的生物学洞察，即 RBM39 的基因敲除可以在表型上模拟已验证的 DDR 靶点 CDK12 的缺失，且不影响 CDK13。因此，REC-1245 更能体现 Recursion 的 AI 平台的“原创发现”能力，即发现 RBM39 与 CDK12 功能相关的新调控机制，并设计新分子。管理层预计将在 2026 年上半年公布 1 期单药剂量递增部分的早期安全性和 PK 数据。

图 25: Recursion 主要管线预期里程碑

研发代码	预期里程碑
REC-4881 (MEK1/2)	预计 2026 年上半年启动 FDA 沟通，就 REC-4881 潜在注册路径达成共识；2027 年上半年预计公布更多 Ib/II 期临床数据
REC-1245 (RBM39)	预计 2026 年上半年公布早期 I 期安全性及单药 PK 数据
REC-7735 (PI3Kα H1047R)	预计 2026 年下半年基于数据决定是否启动 I 期试验
REC-102 (ENPP1)	预计 2026 年下半年基于数据决定是否启动 I 期试验
REC-617 (CDK7)	预计 2027 年上半年公布早期 I 期安全性及联合用药药代动力学数据
REC-3565 (MALT1)	预计 2027 年上半年公布早期 I 期安全性及联合用药药代动力学数据
REC-4539 (LSD1)	预计 2027 年下半年公布早期 I 期安全性及单药治疗药代动力学数据

资料来源：Recursion，招银国际环球市场

除了内部推进创新药研发以外，Recursion 也通过 AI 平台技术达成了多项对外合作。截至 2026 年 2 月 25 日，Recursion 已获得超过 5 亿美元的预付款及里程碑付款，并有望在产品上市之前继续获得超过 200 亿美元的额外里程碑付款。2026 年 2 月，Recursion 达成与赛诺菲合作框架下的第五个里程碑，获得 400 万美元里程碑付款，标志着双方在历史性难攻且创新性强的生物学领域实现首创性肿瘤项目合作。罗氏/基因泰克也已采纳 6 张 Recursion 构建的 Phenomap 图谱，并基于此启动一项小分子药物项目。

图 26: Recursion 与 MNC 药企的合作梳理 (含 Exscientia 和 Cyclica)

时间	合作方	主要合作内容 (Recursion 及其子公司主要职责)	首付款	总金额
2013	杨森研发 (强生)	应用其设计技术，助力 Janssen 研发发现新型化合物。	未披露	未披露
2014	住友制药	双特异性药物发现。新药候选物 5-HT1A 血清素受体激动剂 DSP-1181 在 2020 年于日本启动治疗强迫症的 I 期临床研究。5-HT1A 激动剂/5-HT2A 拮抗剂 DSP-0038 在 2021 年于美国启动 I 期临床研究，用于改善精神类疾病。	未披露	未披露
2017	赛诺菲	共同开发针对代谢疾病领域的双特异性小分子药物。	赛诺菲提供研究资金	最高 2.5 亿欧元
2017	GSK	针对最多 10 个靶点发现/设计新型分子 (临床前候选药物)。Exscientia 还将获得一项未公开的激励机制，旨在通过减少化合物的合成和测定数量来实现先导化合物和候选药物目标。	GSK 提供研发款项	最高 4,300 万美元
2017	武田	在一年内提供针对罕见病的临床前候选药物。2019 年武田针对 2 款候选药物行使选择权，并延长合作期限。	未披露	最高超 9,000 万美元+个位数销售分成
2019	罗氏	设计临床前候选药物，罗氏将获得对此次合作产生的候选药物的独家开发和商业化权利。	首付款+研究资金	最高达 6700 万瑞士法郎+销售分成
2019	Celgene	负责 3 个小分子新药的药物发现。	2,500 万美元	未披露
2020	拜耳	参与三个项目，负责提高发现效率和优化新兴候选药物	未披露	最高达 2.4 亿欧元+销售分成
2020	Cerevel	对预先指定的神经科学疾病靶点进行化合物的筛选、识别、设计和评估	未披露	未披露
2020	拜耳	共同开发针对纤维化疾病的新疗法。2021 年，双方扩大在纤维化领域的合作。	3,000 万美元预付款+5,000 万美元股权投资	每个项目超 1 亿美元里程碑付款+销售分成，可启动超 10 个项目
2021	BMS	基于 2019 年与 Celgene 的合作扩大，将负责肿瘤学和免疫学等多个治疗领域的药物发现、AI 设计和实验工作。同年，BMS 行使选择权，获得一款免疫调节药物的许可。	5000 万美元+2000 万美元行权费	超 12 亿美元+销售分成
2021	罗氏/基因泰克	在神经科学和肿瘤学领域加速识别新靶点并推进创新药物的开发。	1.5 亿美元首付款	每个项目里程碑付款可达超 3 亿美元+销售分成。最多可启动 40 个项目
2022	赛诺菲	针对共同选择的靶点，主导小分子药物设计和先导化合物优化活动，直至开发候选药物提名，最多可达 15 种小分子候选药物。2024 年，两个药物发现项目取得进展并获得 1500 万美元里程碑付款，后续还有资格就这两个项目获得总额超过 6 亿美元的额外里程碑付款，以及商业分成。	1 亿美元	最高可达 52 亿美元+销售分成
2023	拜耳	将合作重点转向肿瘤领域，以识别和推进治疗高未满足需求的肿瘤适应症的新型靶点，Bayer AG 将获得对源自此次研究活动的创新疗法的独家许可选择权。最多可启动 7 个项目	未披露	最高可达 15 亿美元

资料来源：医药魔方，招银国际环球市场

我们认为，Recursion 的股价在过去承受了较大压力，主要反映出其估值受到临床管线验证的影响。公司当前的估值中枢不在于其数据和算力规模，而在于 AI 平台能力能否外推为更具价值的临床数据与更高质量的 BD 兑现。

## Schrödinger, Inc. (SDGR US, 未评级)

图 27: Schrödinger 核心信息

股票代码	SDGR US
成立时间	1990
2025 年收入 (百万美元)	256 (+23% YoY)
2025 年净亏损 (百万美元)	103
截至 2025 年底在手现金 (百万美元)	402

资料来源: Schrödinger, 招银国际环球市场

注: 在手现金包括现金、现金等价物、受限现金及有价证券。

Schrödinger 是一家在计算化学软件领域深耕多年的领军企业，后拓展至 AI 驱动的药物发现领域。Schrödinger 的核心技术是一个基于物理学的分子计算平台，可应用于基于结构的药物设计，例如虚拟筛选和结合亲和力的自由能计算。2018 年前后，Schrödinger 开始将这些基于物理的方法，如分子对接和分子动力学，与现代机器学习算法相结合，以增强预测能力。我们认为，Schrödinger 计算平台的核心能力在于能够实现高精度的分子关键性质预测，因此只要具备清晰的靶点结构，就能使药物发现团队设计和优化小分子。但其方法也存在局限性；它对于靶向特征明确的结合位点最为有效，而在新型靶点发现或表型探索方面适用性较弱。在这些领域，Recursion 等公司可能更具优势。因此，Schrödinger 主要攻克的是已经过遗传学或生物学验证的靶点，如 MALT1 和 WEE1，并通过优化化学获得同类最佳分子。

Schrödinger 的基石业务是软件授权，即为药企客户提供多种软件解决方案，加速分子发现、设计和优化的所有阶段。2025 年，公司软件收入达约 2 亿美元，同比增长 11%。按 2024 年收入计算的所有前 20 大制药公司均授权使用了公司的软件解决方案，占公司软件收入的 37%。目前，公司软件业务在商业策略上正经历转型，主动将传统的本地部署软件一次性买断模式向托管式云服务和订阅制许可 (SaaS) 转移。尽管这在短期内 GAAP 口径软件收入仍可能持续承压，但长期来看将带来更可预测的年度合同价值 (ACV) 和更高的客户黏性。

此外，公司还利用平台和能力，与生物制药公司达成药物发现合作 (BD 合作)，协助识别对一种或多种特定蛋白靶点具有活性的分子。2025 年，公司基于合作产生的预付款和里程碑付款，即药物发现业务收入达 5,640 万美元，同比增长 107%，收入占比达 22%。据公司披露，过去五年，BD 合作已累计为公司带来约 6 亿美元收入。

Schrödinger 与 Nimbus Therapeutics 早期合作发现的 TYK2 抑制剂 Zasocitinib (NDI-034858) 是最具代表性的合作项目之一。针对已知免疫学靶点，Schrödinger 利用 AI 和物理方法设计了一种差异化的变构结合剂，在银屑病治疗中展现出具有同类最佳潜力的 II 期临床结果，并于 2022 年被武田以 40 亿美元收购。2026 年 3 月，武田公布 Zasocitinib 的 2 项 III 期临床数据。在第 16 周，接受 Zasocitinib 治疗的患者中，71.4% 和 69.2% 达到静态医师整体评估 (sPGA) 0/1 分，而安慰剂组为 10.7% 和 12.6%，阿普米司特 (apremilast) 组为 32.1% 和 29.7% ( $p < 0.001$ )；接受 Zasocitinib 治疗的患者中，61.3% 和 51.9% 达到银屑病面积和严重程度指数 (PASI) 90，安慰剂组为 5.0% 和 4.0%，阿普米司特 (apremilast) 组为 16.8% 和 15.9% ( $p < 0.001$ )。武田计划于 2026 年正式向 FDA 提交新药上市申请，Zasocitinib 可能成为首批获批的 AI 药物之一。

图 28: Schrödinger 与 MNC 药企的合作梳理

时间	合作方	主要合作内容 (Schrödinger 及其子公司主要职责)	首付款	总金额
2015	赛诺菲	为最多十个药物发现项目提供计算药物设计支持, 从靶点分析和验证到先导化合物鉴定和优化等阶段。	未披露	最高可达 1.2 亿美元
2017	武田	自主领导多靶点药物发现工作, 利用其计算平台和武田提供的蛋白质晶体结构来指导新化学实体的设计, 并承担其发现成本, 武田拥有独家许可的选择权。2019 年, 双方扩大合作, 增加了可启动的药物发现项目的潜在数量, 并提高了每个项目的未来里程碑付款的累计潜在价值。且已实现首个重要的临床前里程碑。	未披露	每个项目最高可达 1.7 亿美元里程碑付款+销售分成
2020	BMS	发现针对五个指定生物靶点的小分子化合物。	5500 万美元	最高可达 27.55 亿美元+销售分成
2021	Centessa/Orexia	发现靶向食欲素-2 受体 (OX2R) 的新型疗法。	获得软件使用费	里程碑付款+销售分成
2024	诺华	推进多个未公开靶点的开发候选药物, 并签署一项为期三年的扩大软件协议。	1.5 亿美元	最高达 24.22 亿美元+销售分成

资料来源: 医药魔方, 招银国际环球市场

### AI 技术落地与临床管线:

自 2018 年起, 公司开始构建自有管线, 截至 2025 年末已有两个候选药物进入临床阶段。2022 年 6 月, SGR-1505 (MALT1 抑制剂) 的 IND 申请获美国 FDA 批准, 是公司首个进入临床阶段的自有管线, 目前已启动针对复发或难治性 B 细胞恶性肿瘤的临床 I 期试验。SGR-1505 是公司利用基于物理学的计算平台, 结合多参数优化、FEP+和机器学习筛选出的具有强效抗肿瘤活性的新型 MALT1 抑制剂。2025 年, FDA 授予 SGR-1505 快速通道认定, 用于治疗既往至少接受过两线治疗失败的华氏巨球蛋白血症 (WM) 成人患者, 并授予其用于治疗 WM 的孤儿药认定。2025 年 12 月, 公司公布 SGR-1505 的 I 期临床更新数据。安全性方面, 总体耐受性良好, 20% (12 例) 的患者在治疗期间出现严重不良事件 (SAE), 其中 3 例与治疗相关。疗效方面, 所有剂量下的总缓解率 (ORR) 为 25%。

2024 年 4 月, SGR-3515 (Wee1/Myt1 抑制剂) 的 IND 申请也获得 FDA 批准, 并启动其在晚期实体瘤患者中的 I 期临床试验给药。管理层预计将在 2026 年第二季度公布 SGR-3515 的 I 期临床早期数据。公司正积极为 SGR-1505 和 SGR-3515 寻求对外授权或合作机会。但公司的自有管线研发并非一帆风顺。2025 年 8 月, 因安全性不佳, 公司终止 SGR-2921 (CDC7 抑制剂) 的临床开发。

图 29: Schrödinger 的主要管线

PROGRAM	DISCOVERY	IND ENABLING	PHASE 1	PHASE 2	PHASE 3
<b>SGR-1505 (MALT1)</b> Relapsed/Refractory B-Cell Malignancies	██████████	██████████	██████████		
<b>SGR-3515 (Wee1/Myt1)</b> Advanced Solid Tumors	██████████	██████████	██████████		
<b>SGR-6016 (NLRP3)</b> Neurodegeneration	██████████	██████████			
<b>SGR-5573 (EGFR<sup>C797S</sup>)</b> Oncology	██████████	██████████			
<b>SGR-4174 (SOS1)</b> Oncology and Rare Disease	██████████	██████████			
<b>LRRK2</b> Neurology	██████████				
<b>Other Discovery Programs</b> Multiple Areas	██████████				

资料来源：Schrödinger，招银国际环球市场

我们认为，Schrödinger 软件业务的成功已验证了 AI 和计算平台可加速先导化合物的优化，有意义地增强传统药物设计工作流程。但利用其 AI 平台能否获得更优的药物并提高临床成功率，还有待于其自研管线如 SGR-1505 和 SGR-3515 的临床试验验证，比如能否显示出明确的患者获益。

### Absci (ABSI US, 未评级)

图 30: Absci 核心信息

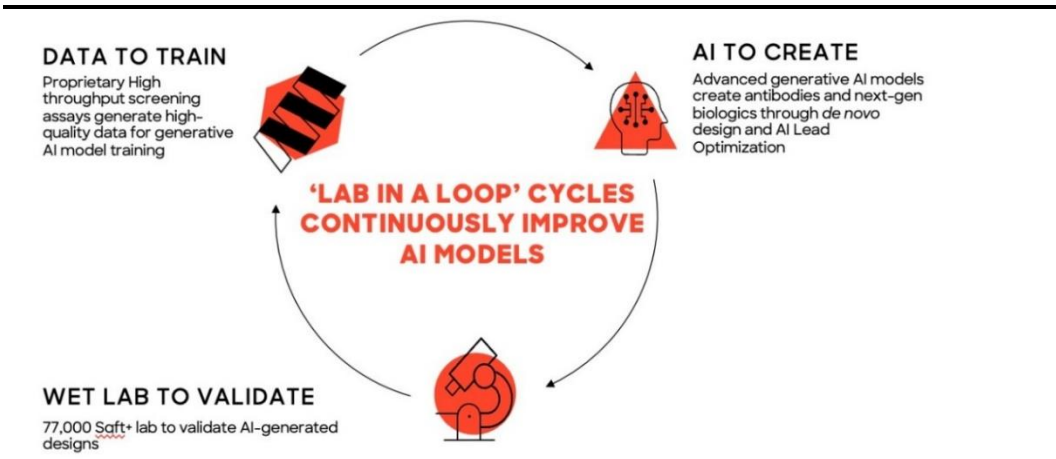
股票代码	ABSI US
成立时间	2011
2025 年收入 (百万美元)	2.8 (-38%)
2025 年净亏损 (百万美元)	115
截至 2025 年底在手现金 (百万美元)	144

资料来源：Absci，招银国际环球市场

注：在手现金包括现金、现金等价物及有价证券

Absci 是一家聚焦 AI 原生抗体药物发现的生物技术公司，利用生成式 AI 设计具备明确目标产品特征 (TPP) 的候选药物，并通过湿实验形成快速闭环验证。依托先进 AI 算法与湿实验验证之间的持续反馈循环，可不断优化公司模型并缩短研发周期。据公司披露，其实验室闭环 (Lab-in-a-loop) 能够将从 AI 设计到湿实验验证候选物的周期压缩至最短约 6 周。

图 31: Absci 的实验室闭环 (Lab-in-a-loop)



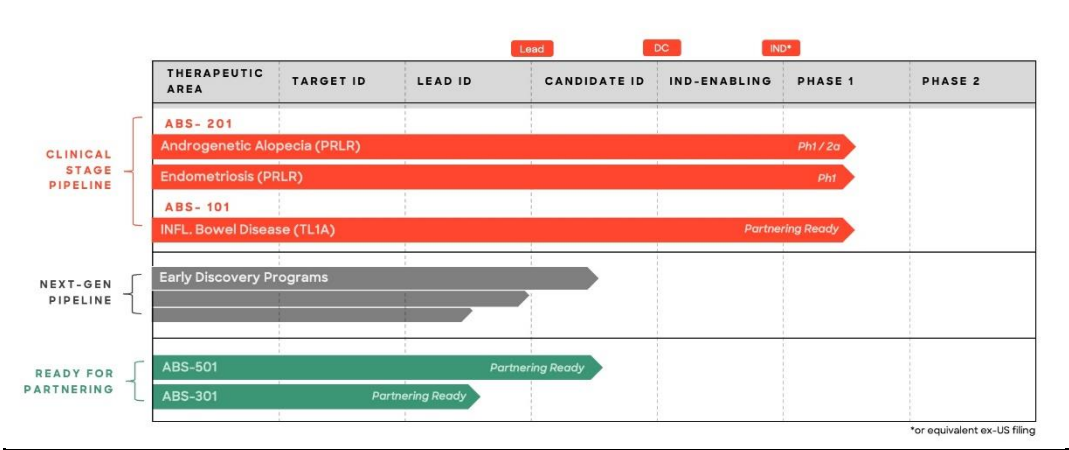
资料来源：Absci，招银国际环球市场

目前，公司已建成合计超 77,000 平方英尺的湿实验和干实验设施，可规模化生成用于 AI 模型训练和测试的高质量数据，包括涉及高通量 SPR 抗体亲和力测定、SoluPro/ACE 细菌合成生物学平台，以及酵母和噬菌体展示文库的富集、分选和滴定等。基于这些专有数据和公共生物学数据集，公司在 2026 年 1 月发布 Origin-1 平台，可通过 1) AbsciGen 生成抗体结构和序列；2) AbsciBind 在实验前对抗体候选物进行评分和筛选。与传统药物发现活动相比，Origin-1 每个靶点仅需筛选少于 100 个候选抗体设计。

**AI 技术落地与临床管线：**

截至 2025 年底，公司共有 2 个候选药物已推进至临床阶段。

**图 32: Absci 的在研管线**



资料来源：Absci，招银国际环球市场

ABS-201 是由 AI 设计生成的 PRLR 单抗，用于治疗雄激素性脱发 (AGA)。Absci 已启动 ABS-201 在 AGA 患者及健康受试者中的 I/IIa 期临床试验，并完成单次递增剂量 (SAD) 部分前三个队列给药。公司预计将于 2026 年上半年公布初步安全性、耐受性和药代动力学 (PK) 数据，并在 2026 年下半年和 2027 年初公布中期及完整概念验证 (PoC) 数据。此外，公司还计划将 ABS-201 的适应症拓展至子宫内膜异位症。若 AGA 项目早期安全性数据积极，预计将于 2026 年第四季度启动 ABS-201 针对子宫内膜异位症的 II 期临床试验。

除自研管线外，Absci 已与默沙东、阿斯利康等达成合作。目前，公司收入仍全部来源于合作项目，整体规模较小，因此平台商业化价值仍有待通过更多合作兑现及临床进展进一步验证。公司也在积极推进自研管线的对外授权，其中 ABS-101 是潜在 BD 项目之一。ABS-101 是公司利用 AI 从头设计并优化的 TL1A 单抗，可识别 TL1A 单体和三聚体形式。考虑到 TL1A 在体内以三聚体为主要活性形式，但同时也存在单体形式，因此 ABS-101 对两种构象的覆盖能力有望带来更充分的靶点抑制。此外，Absci 在药物发现阶段就利用 AI 技术主动进行低免疫原性的设计，有望降低高 ADA 带来的安全性和疗效风险。临床前数据显示，ABS-101 表现出更低的 TL1A 内吞作用，提示其潜在 ADA 风险可能相对更低。在非人灵长类 (NHP) 动物模型中，其半衰期比较同靶点药物 MK-7240 和 RVT-3101 延长了 2-3 倍，有望支持每 8 周甚至每 12 周 (Q8W-Q12W) 一次的皮下给药方案，从而提升了患者依从性和用药便利性。公司已于 2025 年 5 月启动 ABS-101 的 I 期临床试验，并取得了积极的中期结果。

图 33: Absci 与 MNC 药企的合作梳理

时间	合作方	主要合作内容 (Absci 主要职责)	首付款	总金额
2022	默沙东	将部署其“仿生蛋白” (Bionic Protein™) 非标准氨基酸技术, 以生产为默沙东的生物制造应用量身定制的酶。默沙东有选择权, 可提名最多三个靶点进入药物发现协议	未披露	未披露 若行使选择权, 则最高可获得 6.1 亿美元+研究资助+销售分成
2023	阿斯利康	针对特定肿瘤靶点提供治疗性候选抗体	未披露	预付款+研发资金+里程碑付款+销售分成

资料来源: 医药魔方, 招银国际环球市场

## Generate Biomedicines (GENB US, 未评级)

图 34: Generate Biomedicines 核心信息

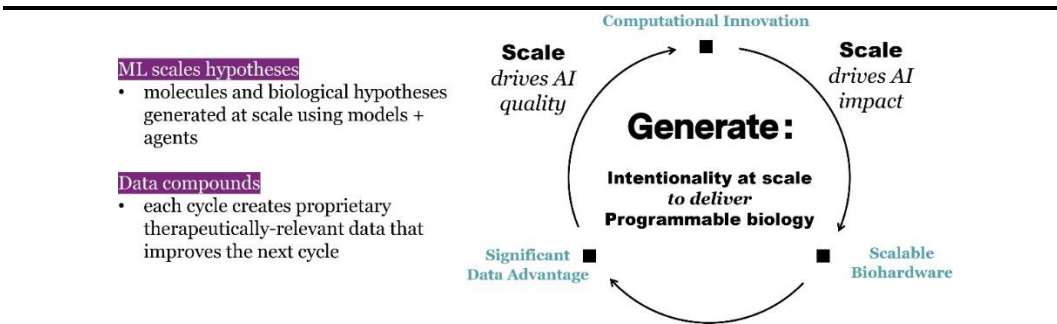
股票代码	GENB US
成立时间	2018
2025 年收入 (百万美元)	32
2025 年净亏损 (百万美元)	223
截至 2025 年底在手现金 (百万美元)	221

资料来源: Generate Biomedicines, 招银国际环球市场

注: 在手现金包括现金、现金等价物及有价证券

Generate Biomedicines 是新一代 AIDD 公司, 利用生成式 AI 平台 Generate 可实现大分子生物药的从头设计 (de novo)。公司将生成模型、性质预测模型与可拓展的高通量实验设施 (Scalable Biohardware) 深度整合, 围绕 Generate 平台构建了一个设计-构建-测试-学习的闭环体系。

图 35: Generate 平台示意图




资料来源: Generate Biomedicines, 招银国际环球市场

Generate 平台的核心是 AI 模型, 包括扩散蛋白质模型, 如 Chroma 模型和图神经网络等。这些模型从专有的蛋白质序列、结构和功能数据中学习通用的设计原则, 可在给定约束条件下生成全新蛋白及蛋白复合物。随后, 通过 DNA 组装和蛋白质生产技术可将设计转化为真实分子, 并利用高通量实验平台 (包括冷冻电子显微镜、多重生物检测等) 验证蛋白功能和特性。这些实验数据将反哺 AI 模型, 实现持续迭代优化。

### AI 技术落地与临床管线:

目前, Generate 已将 3 个 AI 优化或设计的候选药物推进至临床阶段, 说明平台已初步跨越了从 AI 药物发现到临床转化的关键门槛。

图 36: Generate Biomedicines 主要管线

	Proposed Indication(s)	Target	Modality	Preclinical	Phase 1	Phase 2	Phase 3	Next Milestone	Collaborations
<b>IMMUNOLOGY &amp; INFLAMMATION</b>									
GB-0895	Severe Asthma	TSLP	Antibody					Fully enrolled Ph 3 studies (2H27/1H28)	
GB-0895	COPD	TSLP	Antibody					Ph1b data (1H26)	
<b>ONCOLOGY</b>									
GB-4362 <sup>1</sup>	Various in combo with MMAE ADCs	Free MMAE	Antibody					Ph1 initiation (2026)	
GB-5267 <sup>1</sup>	Metastatic Ovarian Cancer	MUC-16	Armored CAR-T					Ph1 initiation (2026)	 50:50

资料来源: Generate Biomedicines, 招银国际环球市场

GB-0895 (长效抗 TSLP 单抗) 是公司当前临床进展最快的管线。GB-0895 并非从头设计的抗体, 而是利用 Generate 平台对已上市的 TSLP 抗体 tezepelumab 进行 AI 优化筛选。临床前数据显示, GB-0895 对人 TSLP 的结合亲和力达到 106 飞摩尔, 相比 tezepelumab 提高了约 20 倍。此外, 公司对 Fc 段进行了 YTE 突变工程化改造, 使抗体在酸性条件下与 FcRn 的结合亲和力大幅提升, 从而在很大程度上免于被溶酶体降解, 有效延长了半衰期。对 TSLP 的超高亲和力和 YTE 介导的半衰期延长为实现 Q26W 给药方案奠定了基础。针对重度哮喘的 I 期临床数据显示安全性良好, 未报告 3 级以上 TRAE 事件, 且半衰期约 89 天, 并在至少 6 个月内维持对多项下游炎症生物标志物如 FeNO、IL-5 和 IL-13 的抑制。2025 年 12 月, 公司启动了两项全球 III 期研究 SOLAIRIA-1/2, 合计拟纳入约 1,572 名严重哮喘患者, 主要终点为 52 周年化哮喘急性发作率。GB-0895 目前还在 COPD 中开展 I 期研究。我们认为, 若 GB-0895 的 III 期临床试验成功, 其差异化给药频次有望提升患者给药便利性, 成为现有 TSLP 通路产品的有力竞争者。

除呼吸管线外, 公司正推动两项肿瘤项目进入早期临床。GB-4362 是一款靶向游离 MMAE 的单抗, 通过选择性中和 ADC 释放后的游离毒素, 降低 MMAE 类 ADC 的系统毒性, 进而支持 MMAE 载荷的 ADC 药物的持续给药或更高剂量给药。临床前数据显示, GB-4362 可降低游离 MMAE 水平, 并在动物实验中改善中性粒细胞减少和皮肤毒性, 且未影响 ADC 药物的抗肿瘤活性。GB-4362 已于 2025 年 12 月在美获批 IND, 并在 2026 年初获得 FDA 快速通道认证, 公司预计将在 2026 年启动 I 期临床试验。GB-5267 是由 Generate 平台设计并优化的靶向 MUC16 的 IL-18 修饰的 CAR-T 细胞疗法, 用于铂耐药卵巢癌。2023 年 10 月, 公司与 Roswell Park 达成协议, 将合作开发包括 GB-5267 在内的多达三个肿瘤靶点的 CAR-T 细胞疗法和装甲化技术, 双方将分摊研发费用并共享商业化利润。2025 年 12 月, GB-5267 在美获批 IND, 预计将于 2026 年启动 I 期临床试验。

除自研管线外, Generate 也通过对外合作验证平台价值。公司与安进于 2021 年达成多靶点蛋白药物合作, 获得 5,000 万美元首付款及 2,500 万美元股权投资。安进已行使一个额外靶点的选择权, 并支付 500 万美元付款。公司与诺华则于 2024 年达成多靶点蛋白药合作, 交易首付款达 6,500 万美元, 其中包括 1,500 万美元股权投资, 并有资格获得超 10 亿美元的里程碑付款以及销售分成。

图 37: Generate Biomedicines 与 MNC 药企的合作梳理

时间	合作方	主要合作内容 (Generate Biomedicines 主要职责)	首付款	总金额
2022	安进	发现和创造针对五种临床靶点的蛋白质药物, 安进拥有提名最多五个额外项目的选择权, 需支付额外费用	5,000 万美元首付款	2,500 万美元股权投资, 每个项目最高可达 3.7 亿美元里程碑付款以及销售分成
2024	诺华	发现和开发蛋白质疗法	6,500 万美元 (含 1,500 万美元股权投资)	最高可达超 10.65 亿美元+销售分成

资料来源: 医药魔方, 招银国际环球市场

## 英矽智能 (3696 HK, 未评级)

图 38: 英矽智能核心信息

股票代码	3696 HK
成立时间	2014
2025 年收入 (百万美元)	56 (-34.5% YoY)
2025 年净亏损 (百万美元)	352
截至 2025 年底在手现金 (百万美元)	448

资料来源: 英矽智能, 招银国际环球市场。注: 在手现金包括银行结余及现金和按公允价值计入损益的金融资产

英矽智能是 AIDD 领域的代表性企业之一。公司自主开发的生成式 AI 平台 Pharma.AI, 具备从靶点发现到临床结果预测的端到端药物研发赋能能力。平台主要由四大模块构成: Biology42 用于新靶点发现与优先级排序, Chemistry42 用于新型小分子生成、优化及 ADMET 预测, Medicine42 通过 inClinico 支持临床试验结果预测, Science42 则依托多智能体研究助手 DORA 加速文献分析与科学发现。

依托 Pharma.AI, 英矽智能既可实现 FIC 靶点发现并生成候选药物, 也可针对已知靶点设计具备差异化优势的新型小分子, 并推进至临床阶段。据公司披露, 利用该平台可将项目从靶点发现推进至临床前候选药物 (PCC) 确认的时间缩短至约 12-18 个月, 显著短于传统研发平均约 4.5 年的周期。截至 2026 年 3 月 27 日, 公司已通过 Pharma.AI 产生 28 项 PCC, 其中 20 余项已推进至临床或 IND 阶段, 在 AIDD 公司中具备领先的临床转化能力。

图 39: Pharma.AI 平台构成



资料来源：英矽智能，招银国际环球市场

英矽智能目前主要的收入来源为对外授权及合作。利用 Pharma.AI，公司持续扩充产品管线，并在早期阶段（主要为 PCC 阶段）实现对外授权，如 ISM3091 (USP1) 授权给 Exelixis (EXEL US, 未评级)，ISM5043 (KAT6) 授权给 Menarini。同时，公司也与药企达成战略性药物发现及开发合作，以探索潜在靶点和候选药物。据公司披露，公司已与全球前 20 大制药公司中的 13 家建立了战略合作关系，体现了海外 MNC 对于英矽智能 AI 平台技术的认可。2025 年，公司药物发现和管线开发业务实现收入 4,884 万美元，收入占比达 87%。此外，公司还通过 Pharma.AI 的软件订阅及非医药领域的研发合作实现收入。

图 40: 英矽智能与主要 MNC 药企的合作梳理

时间	合作方	主要合作内容（英矽智能主要职责）	首付款	总金额
2020	勃林格 殷格翰	识别与多种疾病相关的潜在治疗靶点。英矽智能将提供额外的人工智能能力，帮助发现新靶点。	未披露	未披露
2020	安斯泰 来	加速针对传统上具有挑战性的靶点家族的候选药物开发	未披露	预付款+里程碑付款
2020	杨森制 药	为指定的多个靶点设计具有特定属性的小分子先导化合物。2023 年达到里程碑，即苗头化合物达到双方认可的类药性标准，英矽智能将获得相应的里程碑付款	未披露	首付款+里程碑付款
2022	赛诺菲	推进基于不超过 6 个创新靶点的候选药物研发。2024 年达成里程碑，发现了具有 FIC 潜力的先导化合物。	不超过 2150 万美元 (含已支付的 1250 万 美元首付款+指定额外三 个合作靶点的 900 万美 元首付款)	最高可达 12 亿美元+销 售分成。
2023	礼来	软件授权合作	未披露	未披露
2023	诺和诺 德	针对肝纤维化疾病识别潜在的新颖靶点并进行验证	未披露	未披露
2025	礼来	针对双方确定的创新靶点开展候选化合物的生成、设计与优化	未披露	最高超 1 亿美元
2026	礼来	获得了在临床前阶段针对特定适应症的潜在同类最佳新型口服疗法的独家全球许可，并将针对选定的靶点合作开展多个研发项目。	1.15 亿美元	最高可达 27.5 亿美元+ 销售分成

资料来源：医药魔方，招银国际环球市场

**AI 技术落地与临床管线：**

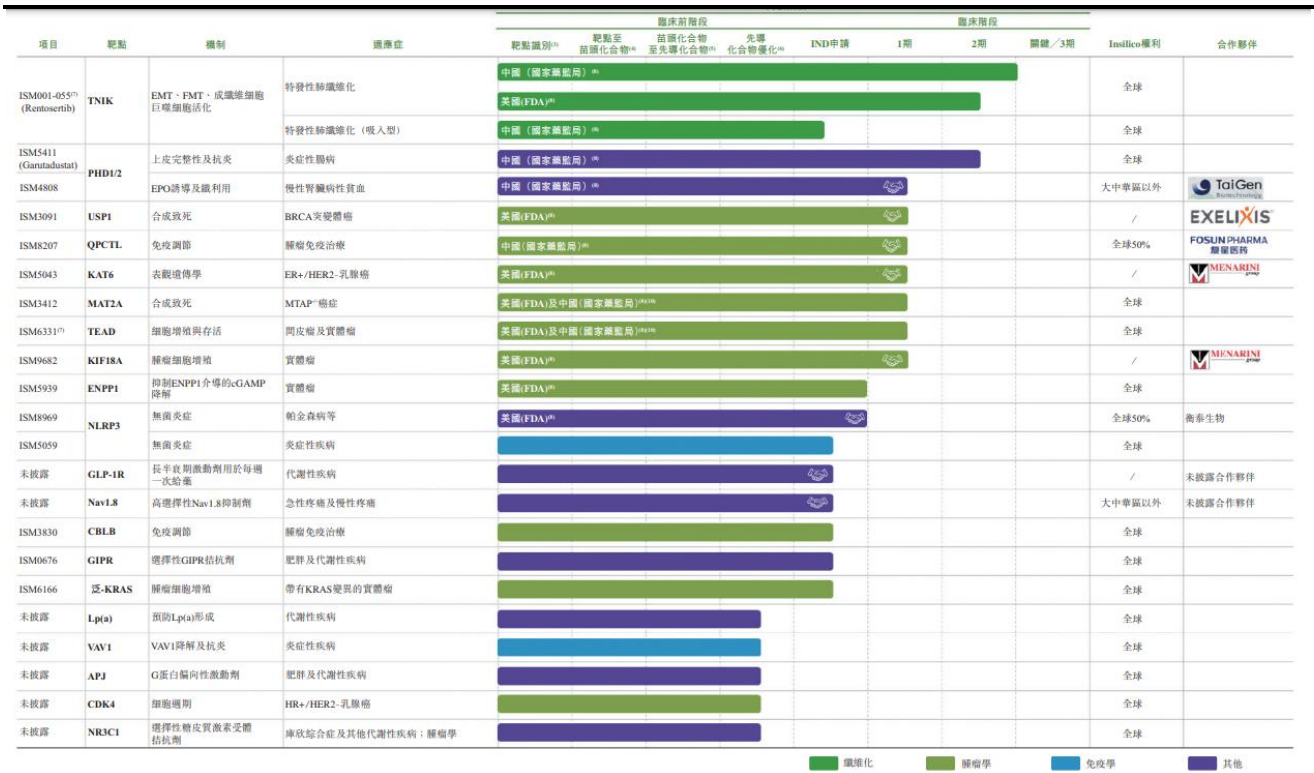
英矽智能当前临床进展最快的核心管线为 TNIK 抑制剂 ISM001-055 (Rentosertib)，是首个完全基于 AI 平台开发并推进至 II 期临床的药物。2024 年，公司公布 Rentosertib 用于治疗特发性肺纤维化 (IPF) 的中国 IIa 期临床试验顶线结果。

从安全性看，Rentosertib 总体可控，3 级及以上 TRAE 发生率为 16.7%，高于安慰剂组的 5.9%；导致治疗中止的不良事件发生率为 18.5%，略高于安慰剂组的 11.8%。此外，30mg BID 剂量组出现 1 例死亡事件，但研究者判断其与研究药物无关。

从疗效看，Rentosertib 在肺功能改善方面显示出积极信号。在 60mg QD 最高给药剂量组中，患者 FVC 较基线平均提升 98.4 mL，而安慰剂组较基线平均下降 20.3 mL，公司计划于 2026 年启动 Rentosertib 针对 IPF 的中国 III 期临床试验。

我们认为，Rentosertib 的 III 期临床研究及后续的数据读出将成为 AIDD 行业的重要催化事件。一方面，Rentosertib 是基于 AI 发现新靶点并生成候选分子，且临床进度比较领先；另一方面，若 Rentosertib 能在 III 期临床试验中取得成功，将有望验证 AI 平台在创新药早期发现与临床转化中的实际价值，推动 AIDD 向临床与商业价值验证迈进。中长期看，随着更多 AI 生成的候选分子进入临床后期阶段，AI 技术能否提升创新靶点及药物的研发成功率，将得到更充分的验证，并有望成为支撑公司估值的重要锚点。

**图 41: 英矽智能主要管线**



资料来源：英矽智能，招银国际环球市场

## 估值表

图 42: 同业公司估值

公司名称	代码	评级	市值 (百万美元)	股价 (LC)	PS (x)		ROE (%)		收入 (百万美元)	
					26E	27E	26E	27E	25A	26E
Recursion	RXRX US	未评级	1,553	2.93	23.4	13.9	(56.7)	(110.1)	74.7	66.4
Schrödinger	SDGR US	未评级	893	11.95	3.8	3.6	(30.6)	(19.9)	255.9	234.8
Absci	ABSI US	未评级	800	5.13	1249.4	73.9	(56.9)	(62.4)	2.8	3.2
Generate Biomedicines	GENB US	未评级	2,051	16.00	67.1	72.2	(89.2)	(57.0)	31.9	30.6
英矽智能	3696 HK	未评级	3,827	52.35	23.1	16.1	2.0	(0.1)	56.2	166.0
晶泰控股	2228 HK	买入	4,490	8.17	30.4	22.0	0.5	1.9	111.7	146.8

资料来源：彭博，招银国际环球市场

注：截至2026年5月18日。上述所有公司的盈利预测及估值数据均来自于彭博一致预期。

## 免责声明及披露

### 分析员声明

负责撰写本报告的全部或部分内容的分析员，就本报告所提及的证券及其发行人做出以下声明：（1）发表于本报告的观点准确地反映有关于他们个人对所提及的证券及其发行人的观点；（2）他们的薪酬在过往、现在和将来与发表在报告上的观点并无直接或间接关系。

此外，分析员确认，无论是他们本人还是他们的关联人士（按香港证券及期货事务监察委员会操作守则的相关定义）（1）并没有在发表研究报告 30 日前处置或买卖该等证券；（2）不会在发表报告 3 个工作日内处置或买卖本报告中提及的该等证券；（3）没有在有关香港上市公司内任职高级人员；（4）并没有持有有关证券的任何权益。

招银国际环球市场或其关联机构曾在过去 12 个月内与本报内所提及发行人有投资银行业务的关系。

### 招银国际环球市场投资评级

买入	: 股价于未来 12 个月的潜在涨幅超过 15%
持有	: 股价于未来 12 个月的潜在变幅在-10%至+15%之间
卖出	: 股价于未来 12 个月的潜在跌幅超过 10%
未评级	: 招银国际证券并未给予投资评级

### 招银国际环球市场行业投资评级

优于大市	: 行业股价于未来12个月预期表现跑赢大市指标
同步大市	: 行业股价于未来12个月预期表现与大市指标相若
落后大市	: 行业股价于未来12个月预期表现跑输大市指标

### 招银国际环球市场有限公司

地址: 香港中环花园道3号冠君大厦45楼 电话: (852) 3900 0888 传真: (852) 3900 0800

招银国际环球市场有限公司(“招银国际环球市场”)为招银国际金融有限公司之全资附属公司(招银国际金融有限公司为招商银行之全资附属公司)

### 重要披露

本报内所提及的任何投资都可能涉及相当大的风险。报告所载数据可能不适合所有投资者。招银国际环球市场不提供任何针对个人的投资建议。本报告没有把任何人的投资目标、财务状况和特殊需求考虑进去。而过去的表现亦不代表未来的表现，实际情况可能和报告中所载的大不相同。本报告中所提及的投资价值或回报存在不确定性及难以保证，并可能会受目标资产表现以及其他市场因素影响。招银国际环球市场建议投资者应该独立评估投资和策略，并鼓励投资者咨询专业财务顾问以便作出投资决策。

本报告包含的任何信息由招银国际环球市场编写，仅为本公司及其关联机构的特定客户和其他专业人士提供的参考数据。报告中的信息或所表达的意见皆不可作为或被视为证券出售要约或证券买卖的邀请，亦不构成任何投资、法律、会计或税务方面的最终操作建议，本公司及其雇员不就报告中的内容对最终操作建议作出任何担保。我们不对因依赖本报告所载资料采取任何行动而引致之任何直接或间接的错误、疏忽、违约、不谨慎或各类损失或损害承担任何的法律上责任。任何使用本报告信息所作的投资决策完全由投资者自己承担风险。

本报告基于我们认为可靠且已经公开的信息，我们力求但不担保这些信息的准确性、有效性和完整性。本报告中的资料、意见、预测均反映报告初次公开发布时的判断，可能会随时调整，且不承诺作出任何相关变更的通知。本公司可发布其它与本报告所载资料及/或结论不一致的报告。这些报告均反映报告编写时不同的假设、观点及分析方法。客户应该小心注意本报告中所提及的前瞻性预测和实际情况可能有显著区别，唯我们已合理、谨慎地确保预测所用的假设基础是公平、合理。招银国际环球市场可能采取与报告中建议及/或观点不一致的立场或投资决定。

本公司或其附属关联机构可能持有报告中提到的公司所发行的证券头寸并不时自行及/或代表其客户进行交易或持有该等证券的权益，还可能与这些公司具有其他投资银行相关业务联系。因此，投资者应注意本报告可能存在的客观性及利益冲突的情况，本公司将不会承担任何责任。本报告版权仅为本公司所有，任何机构或个人于未经本公司书面授权的情况下，不得以任何形式翻版、复制、转售、转发及或向特定读者以外的人士传阅，否则有可能触犯相关证券法规。

如需索取更多有关证券的信息，请与我们联系。

#### 对于接收此份报告的英国投资者

本报告仅提供给符合(I)不时修订之英国 2000 年金融服务及市场法令 2005 年(金融推广)令(“金融服务令”)第 19(5) 条之人士及(II) 属金融服务令第 49(2) (a) 至(d) 条(高净值公司或非公司社团等)之机构人士，未经招银国际环球市场书面授权不得提供给其他任何人。

#### 对于接收此份报告的美商投资者

招银国际环球市场不是在美国的注册经纪交易商。因此，招银国际环球市场不受美国就有关研究报告准备和研究分析员独立性的规则的约束。负责撰写本报告的全部或部分内容的分析员，未在美国金融业监管局(“FINRA”)注册或获得研究分析师的资格。分析员不受旨在确保分析师不受可能影响研究报告可靠性的潜在利益冲突的相关 FINRA 规则的限制。本报告仅提供给美国 1934 年证券交易法(经修订) 规则 15a-6 定义的“主要机构投资者”，不得提供给其他任何人。接收本报告之行为即表明同意接受协议不得将本报告分发或提供给任何其他人士。接收本报告的美国收件人如想根据本报告中提供的信息进行任何买卖证券交易，都应仅通过美国注册的经纪交易商来进行交易。

#### 对于在新加坡的收件人

本报告由 CMBI (Singapore) Pte. Limited (CMBISG) (公司注册号 201731928D) 在新加坡分发。CMBISG 是在《财务顾问法案》(新加坡法例第 110 章)下所界定，并由新加坡金融管理局监管的豁免财务顾问公司。CMBISG 可根据《财务顾问条例》第 32C 条下的安排分发其各自的外国实体，附属机构或其他外国研究机构编制的报告。如果报告在新加坡分发给非《证券与期货法案》(新加坡法例第 289 章)所定义的认可投资者，专家投资者或机构投资者，则 CMBISG 仅会在法律要求的范围内对这些人士就报告内容承担法律责任。新加坡的收件人应致电 (+65 6350 4400) 联系 CMBISG，以了解由本报告引起或与之相关的事宜。