

# AI制药：从降本增效到分子创新，数据生产构 筑长期壁垒

行业投资评级：强大于市|维持

中邮证券研究所 医药团队

盛丽华（首席分析师）

徐智敏（分析师）

SAC编号：S1340525060001

SAC编号：S1340525100003

➤ **前言：AI+制药行业的投资价值在于行业的现状分析和未来判断。**投资思路上来看，我们认为在于理解当下AI在制药中的作用地位、商业模式和成长空间、行业成长的关键要素和竞争壁垒。

➤ **AI在制药中的作用是什么？增效和创新。**

基于现有的技术和未来的发展状况，我们认为药物研发基于实验科学的本质不会因AI的迭代升级而发生颠覆。AI在制药端最成熟的落地应用在于提升了**临床前的降本增效**：AI虚拟筛选极大地降低了真实试验中所需要投入的化合物的数量，从而有效缩短了临床前研发周期、大幅减少研发成本。此外，**AI分子生成摆脱认知偏见具有创新价值**，目前的AI分子（如TNIK）进展进入临床2期阶段，有望逐步兑现其创新价值。

➤ **市场规模已超百亿美元，AI+CRO/AI+Biotech仍是企业造血盈利的大趋势**

市场空间来看，全球AI赋能药物研发费用市场规模由2023年的119亿美元有望增长至2032年的746亿美元，CAGR达22.6%。行业投融资热度整体高昂，国内有所收缩。从长远的商业化来看，我们认为AI+CRO/AI+Biotech仍是企业造血盈利的大趋势：一方面利于快速创收，另一方面利用合作项目可以完成企业自身算法模型迭代升级，构筑自身长期竞争力。

➤ **算法+数据深度绑定，高质量数据的生产能力是核心竞争力。**

算法+数据是行业核心技术迭代升级的关键生产要素。算法突破的关键因素在于研发人才团队，而数据的竞争在于高质量数据生产能力（而非传统试验数据的堆积），以实现数据积累和算法迭代之间的闭环和正向循环。大规模高质量的数据是行业稀缺资源，原因在于纳入训练集的数据要求苛刻+数据不共享（“数据孤岛”）。由于实验数据涉及核心利益，长期来看“数据孤岛”的行业特征不会改变，因而**高质量数据生产能力是长期的核心竞争力**。

➤ **相关标的**：行业容易形成马太效应，建议关注领先布局者：英矽智能、晶泰控股、泓博医药、成都先导。

➤ **风险提示**：研发失败风险、行业竞争加剧风险、地缘政治风险、政策变动风险

# 目录

- 一 **AI在制药中的作用是什么？增效和创新**
- 二 **市场规模超百亿美金，分子实体是商业化重点**
- 三 **高质量数据的生产能力是核心竞争力**
- 四 **相关标的**
- 五 **风险提示**

—

# AI在制药中的作用是什么？增效和创新

# 1.1 AI+药物研发：概述

- AI制药是指将NLP、深度神经网络，生成模型等AI技术与传统制药环节相结合，提升新药研发效率，拓展药物创新空间的技术应用。AI医疗大模型基于深度学习、自然语言处理等技术，整合海量生物医学数据(如蛋白质结构、基因序列、临床试验数据)，赋能药物研发全流程，包括新药发现(靶点发现、虚拟筛选、蛋白质结构预测)、新药开发(ADMET预测、晶型预测、合成路径)、临床试验优化(患者分层与剂量预测)。
- 具体应用：基于物理化学原理的AI算法提升**小分子药物**研发效率；在细胞与基因治疗(CGT)领域，AI也可赋能抗原识别、载体设计、放大优化等关键环节。

图表1：药物研发大模型

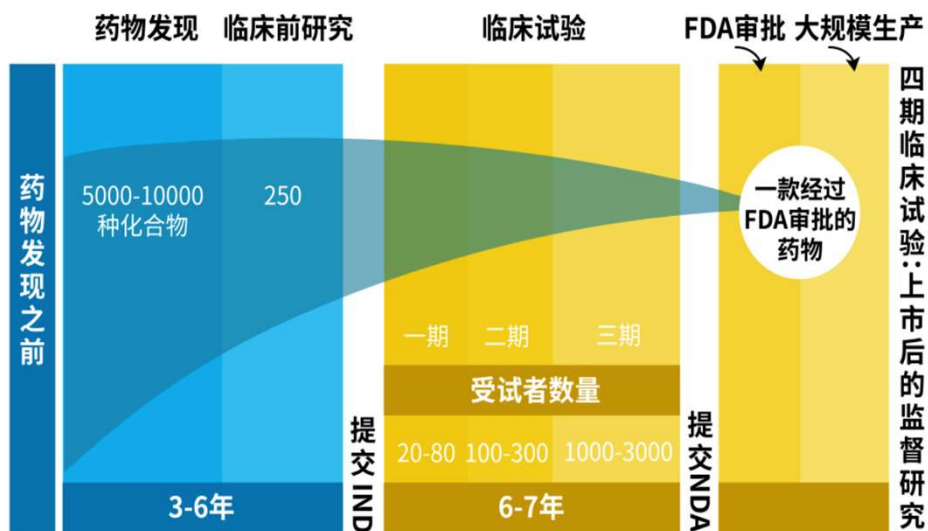


资料来源：智药局，中邮证券研究所

## 1.1 AI+ 药物研发：AI技术不会颠覆药物研发的逻辑

- AI制药在谈论AI+新药研发的融合之前，首先需做出判断：AI技术是否会颠覆制药行业？我们认为，不能。基于以下因素考虑：
  - 1、应用层面：AI计算的“虚拟”数据无法替代临床的“真实”数据。AI尚未能预测药物在人体系统中的反应，而临床试验中个体异质性还会使复杂程度几何倍增。
  - 2、监管层面：我们认为药物的安全性是监管考虑的首要因素，而临床试验是安全性证明的唯一来源。长期来看不存在替代品，药物开发的审批流程长期不会改变。
 因此，药物研发的基于实验科学的本质不会因AI的迭代升级而发生颠覆。

图表2：传统药物研发流程：临床试验难以被AI替代

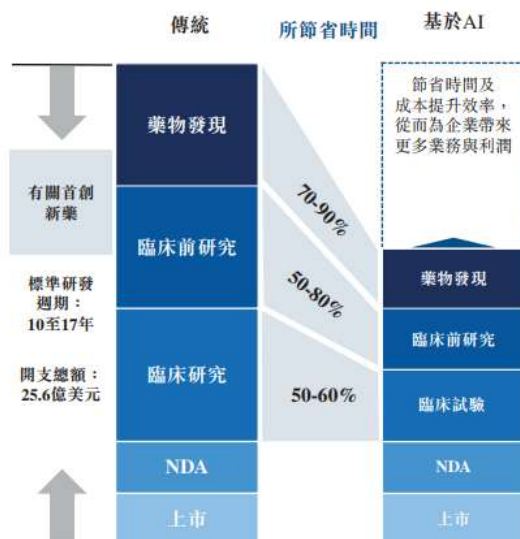


资料来源：美国药品研究和制造商协会，中邮证券研究所

# 1.1 AI+ 药物研发：应用重心在临床前研发阶段

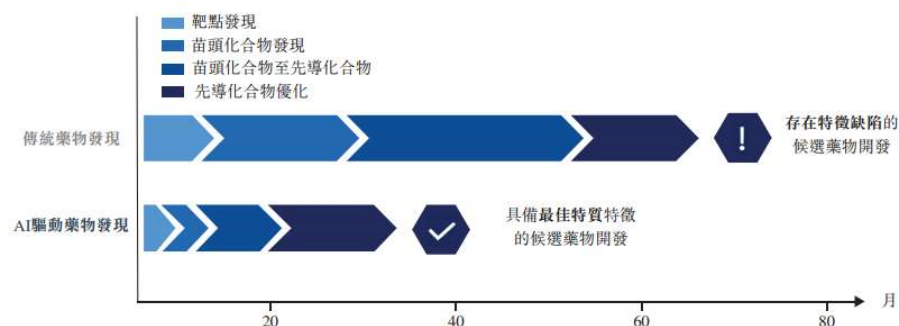
➤ **AI发挥降本增效的作用之处在于临床前的研发环节。**其主要原因在于药物发现环节可提供技术附加值多。其中虚拟筛选或是最为重要的AI应用，可快速从上百万化合物中通过计算的方式筛选出少量更具有潜在活性的化合物，极大地降低了真实试验中所需要投入的化合物的数量，进入临床前阶段后，通过ADMET预测等方式来协助研究优化分子降低失败风险，从而有效缩短了研发周期同时大大降低了研发成本。此外，AI主要的应用场景还包括靶点发现与验证（研发最早期，利用LLM等方式实现大范围的靶点搜寻）和分子生成（有望产生不存在的新分子）等。

图表3：AI与传统研发模式在创新药物研发对比



资料来源：英矽智能招股书，弗若斯特沙利文，中邮证券研究所

图表4：AI制药的应用重点在临床前阶段

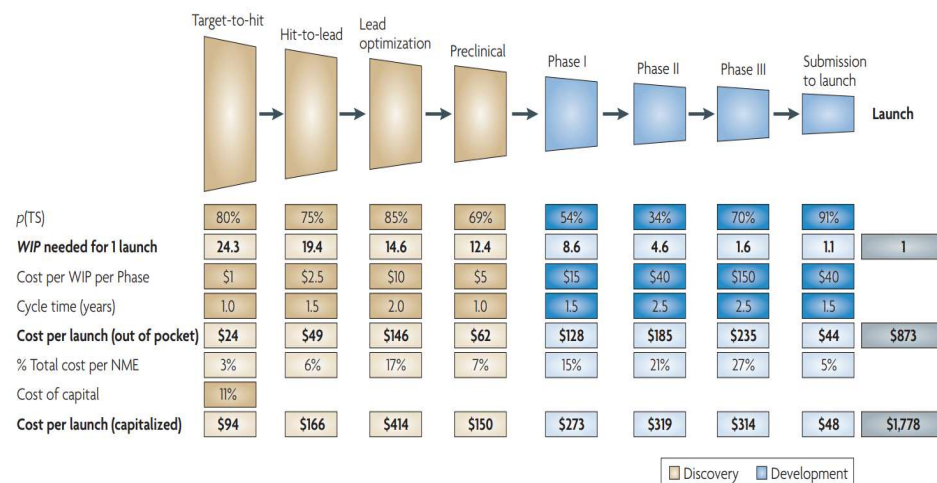


资料来源：英矽智能招股书，弗若斯特沙利文，中邮证券研究所

# 1.1 AI+ 药物研发：应用重心在临床前研发阶段

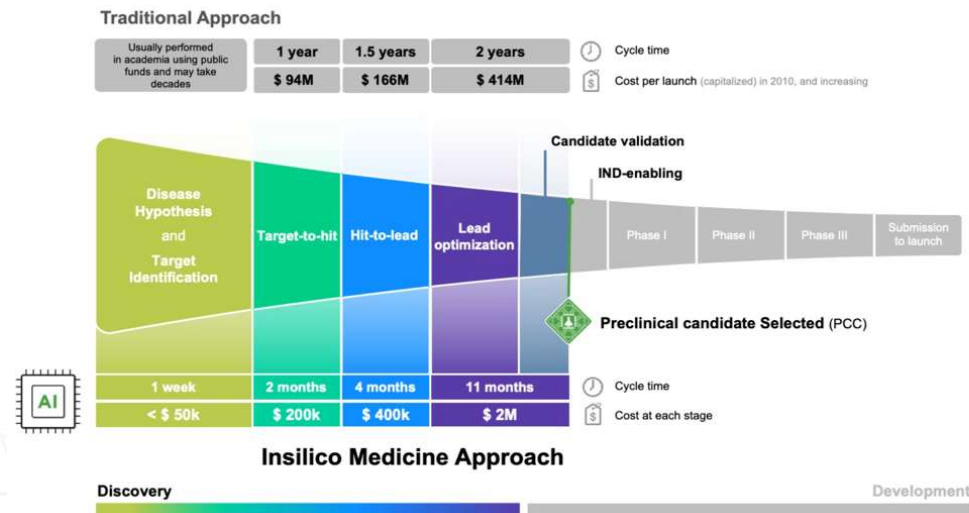
➤ **AI赋能下药物发现研发费用节约达90%+，周期大幅缩减。**从研发流程来看，依次为：靶点 (target) 发现及验证、苗头化合物 (hit) 筛选、先导化合物 (lead) 的发现、临床前候选药物 (PCC) 的生成。根据Nature Reviews 发表的统计数据，在由target到hit、由hit到lead和由lead优化 (至PCC) 的三个阶段传统方法的平均资本支出分别为94、166和414百万美元。根据Insilico Medicine公布的数据，AI能够将上述阶段花费分别缩减至0.2、0.4和2百万美元，缩减比例均为90%以上。同时，AI可将上述阶段的研发周期由传统方法的1年、1.5年和2年缩减至2个月、4个月和11个月，分别节省83%、78%和54%。（注：由于发布时间和统计范围等，两者的数据口径可能存在不一致，计算所得数据仅供参考。）

图表5：传统药物研发流程及对应的开发支出一览



资料来源：Nature Reviews，中邮证券研究所

图表6：AI在药物发现各环节实现大幅缩短周期和节省开支的作用



资料来源：英矽智能官网，中邮证券研究所

# 1.1 AI+ 药物研发：应用重心在临床前研发阶段

- 原理上来看，AI基于数据驱动，能够极大提升海量、多维数据的处理效率。靶点发现、hit筛选以海量的数据库为对象，lead和PCC环节需要进行多指标的评价，AI的契合有效提升了应用效率，详情可参见下方表格。我们认为，AI的效率提升的效果：1、数量上大幅减少了非必要化合物的合成与投入；2、质量上基于数据驱动的优化方法提高成功率。

图表7：AI在药物研发环节中的技术原理和赋能效果总结

研发环节	技术原理	赋能效果
靶点发现及验证	NLP技术、知识图谱技术，通过对海量、多来源、甚至异构的数据库进行交叉引用比对	可以覆盖目前研究尚不充分的疾病；搜索关联度及广度远超专家经验，避免遗漏
苗头化合物 (hit) 筛选	以用于预测各类化合物性质的深度学习模型为核心通过深度学习网络，基于相对明确的底层原理（如QSAR/QSPR等），快速以量子力学的精度对候选库中小分子化合物的特性（包括其自身的物理/化学特性、与靶点的结合模式及亲和力等）进行计算模拟，进而进行打分筛选。	减少湿实验与筛选投入，不受专家经验及化合物库体量限制，更易得到新颖分子
先导化合物 (lead) 发现	由于需要通过骨架跃迁、基团替换等方法修饰苗头化合物的化学结构，进而提升hit的生物活性、结构构效、毒理学性质等，会通过同时预测药物结构与多个指标间的关系，对修饰后的分子进行多指标评价	通过基于化学基团、化合物构效关系等精准预测ADMET等性质，可以减少对于化合物分子合成及测试的湿实验投入
生成临床前候选药物 (PCC)	以用于预测在动物/人体实验中ADMET性质的深度学习模型为核心	ADMET模型实用性提高，节约实验成本

资料来源：量子位，中邮证券研究所

## 1.2 AI+药物研发：Think out of box的创新价值

- AI赋能的核心还在于以分子生成为代表的创新价值，即业界所说的“Think out of box”。在传统研发路径下，化合物设计需要依据已有规则，多样性会受到编码规则、构建模块库等的限制。而基于生成对抗网络、强化学习等模型，AI实现由数据驱动，通过已有数据库学习完成**自主**的分子设计和生成。
- 我们认为，AI+制药的创新有三条主要落地路径：
  - 1、靶点的创新：AI所能触及的数据空间广度与维度突破了人类的框架限制，因而利于发掘出新适应症的靶点或者老适应症的新靶点，具备开发FIC的潜力。
  - 2、分子结构的创新：由AI生成的分子不受限于过往规则，因此在分子骨架等结构上更加具备创新意义。
  - 3、药物重定向：基于疾病/药物相似性网络等，对已有药物关联寻找新适应症。对已上市药物的重定向可以直接进入临床二期，开发效率提升显著。

## 1.2 AI+ 药物研发：Think out of box的创新价值

- **医疗大模型日益发展，AI创新药物在曲折中发展。** 上述路径具备理论上的可行性，并已有相关的进展。但需要强调的是目前尚未有AI自主设计的分子药物成功上市，同时首个进入临床的AI设计分子DSP-1181已停止研发；截止2026年，英矽智能的AI设计分子Rentosertib完成2a期临床，有望实现首个AI驱动发现的分子落地。
- 客观来看，AI设计的分子虽然具备新颖性，但脱离“box”同样可能会使产生的分子在现有的合成框架下难以触达。其限制来自于没有足够高质量的数据可提供。我们认为基于传统的数据，AI在生物路径、化学反应过程等生化知识体系的“学习”并未获得理想效果，因而造成了早期项目的进展并不顺利，但我们认为新技术的突破必然是曲折的，随着技术的迭代，后继者们仍在快速补位，推动AI创新分子的落地。

图表8：部分AI设计分子项目折戟

公司	成立年份	管线	适应症	靶点	进展	备注
Exscientia (现已合并到 Recursion)	2012	EXS-21546	癌症	A2A受体	1/2期研究不及预期终止	2023年10月宣布终止
		DSP-0038	强迫症	5-HT1A受体	1期研究不及预期终止	2020启动1期，全球首个由人工智能设计进入临床的药物，2022年停止开发
BenevolentAI	2013	BEN-2293	特应性皮炎	PanTrk	2期试验未达预期	2023年4月，BEN-2293的II期试验未达到预期治疗效果
Relay Therapeutics	2016	RLY-2139	乳腺癌	CDK2	未进入临床	2023Q3宣布暂停

资料来源：BioPlus, 药渡, 中邮证券研究所

图表9：仍在进展中的AI相关管线

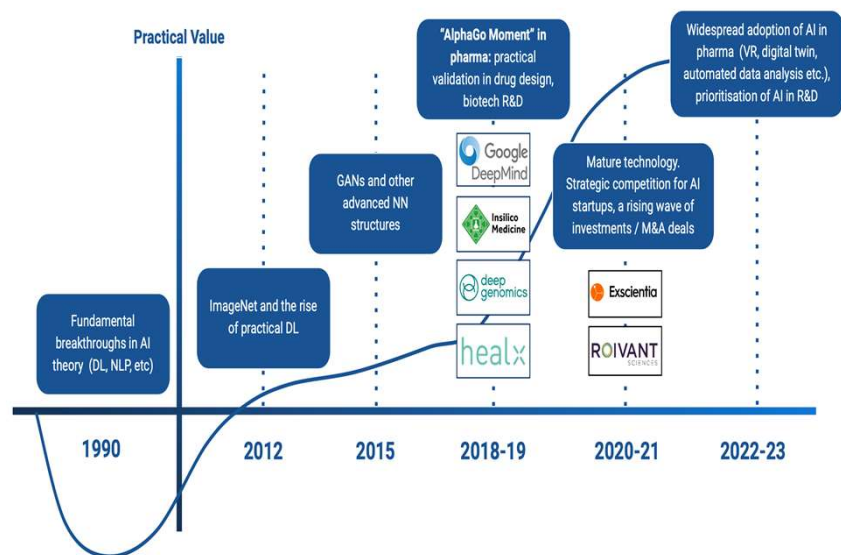
药物	公司	靶点	适应症	当前阶段
INS018_055 (PandaOmics)	英矽智能 (Insilico Medicine)	TNIK 抑制剂	特发性肺纤维化 (IPF)	Ila 期 (II/III 期无缝试验)
RLY-4008	Relay Therapeutics	FGFR2	胆管癌 (FGFR2 融合)	Phase I/II (ReFocus)
RLY-2608	Relay Therapeutics	PI3Kα (变构抑制)	PI3Kα 突变型 HR+/HER2- 转移性乳腺癌	Phase III (ReDiscover-2)
REC-994	Recursion	抗氧化剂 (Superoxide Scavenger)	脑海绵状血管畸形 (CCM)	Phase II (SYCAMORE)
BEN-8744	BenevolentAI	PDE10 (外周选择性)	溃疡性结肠炎 (UC)	Phase Ia (健康志愿者) 已完成

资料来源：各公司官网, 中邮证券研究所

## 1.3 底层技术决定AI效用，算法+数据是“限速步骤”

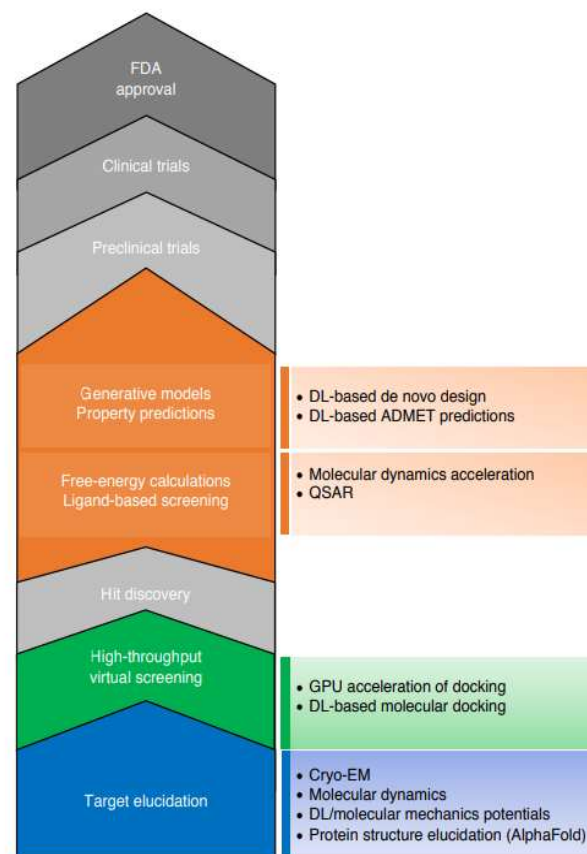
- **行业底层技术框架决定AI的价值效用**：计算机辅助设计（CADD）的出现和应用先于AI+ 新药研发（AIDD），是计算机参与新药研发的底层技术框架。AI的价值效用在于：利用AI在组合学和优化问题的计算优势，提升了CADD在具体应用层面的效果表现。具体来看，基于深度学习等AI模型加速了药物开发各环节的进展，详情可参考右图。

图表10：AI在药物研发的实用价值快速提升



资料来源：Deep Pharma Intelligence，中邮证券研究所

图表11：深度学习（DL）模型对CADD工作流程的提升



资料来源：Nature Machine Intelligence，中邮证券研究所

# 1.3 底层技术决定AI效用，算法+数据是“限速步骤”

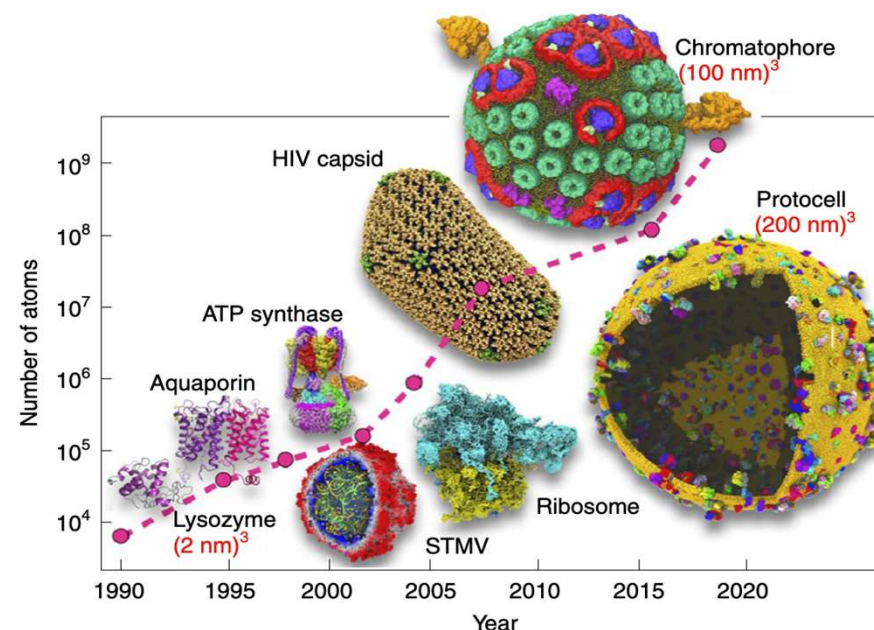
- 回顾发展历程来看，我们认为，AI增效实现“0-1”在于以下三点：
  - I. 虚拟筛选的算法成熟，底层原理（QSAR/QSPR）清晰，能够很好地与AI模型结合；
  - II. 训练数据集丰富、公开：配体库作为虚拟筛选的数据来源，拥有数百亿分子，同时实现了商业化便于获取；
  - III. AI算法的开发、GPU等硬件的快速升级使得AI大数据处理能力的飞速提升。

图表12：不同AI算法模型在药物研发中的应用

Learning category	Architecture/algorithm	Application
Generative models	Graph neural network (GNN)	Molecular representation <sup>11</sup>
		ADMET prediction <sup>11</sup>
		Energy prediction <sup>12</sup>
		Drug-target interaction <sup>18</sup>
		Molecular property prediction <sup>13</sup>
	Generative adversarial networks (GANs)	De novo drug design <sup>14</sup>
	Variational autoencoder (VAE)	De novo drug design <sup>15</sup>
	Transformer	Drug-target interaction <sup>20</sup>
		Molecular representation <sup>21</sup>
		De novo drug design <sup>22</sup>
Representation learning	SMILES2vec	Continuous molecular embedding <sup>23</sup>
	ProtVec	Continuous molecular embedding <sup>24</sup>
Transfer learning	DenseNet (CNN)	virtual screening <sup>27</sup>
	Multitask learning	Toxicity prediction <sup>18</sup>
Federated learning	Horizontal federated learning	Multiple pharmacokinetic parameter prediction <sup>19</sup>
		Binding mode, affinity and activity prediction <sup>20</sup>
		ADMET prediction <sup>21</sup>
		QSAR analysis <sup>21</sup>
Deep reinforcement learning	Stack-RNN + {reward-based RL}	Solubility prediction <sup>22</sup>
	Long short-term memory (LSTM) + {policy gradient, regularized policy gradient, advantage actor-critic network, proximal policy optimization}	De novo drug design <sup>23</sup>
	LSTM + {actor-critic network}	De novo drug design <sup>24</sup>
	RNN + {policy-based RL}	De novo drug design <sup>25</sup>
Interpretable artificial intelligence	Uncertainty estimation	Activity prediction <sup>26</sup>
		Reaction prediction <sup>27</sup>
		Molecular property prediction <sup>28</sup>
		ADMET <sup>29</sup>
		Chemical reactivity prediction <sup>30</sup>
	GNN	

[ ] indicates the policy iteration algorithm used for reinforcement learning (RL).

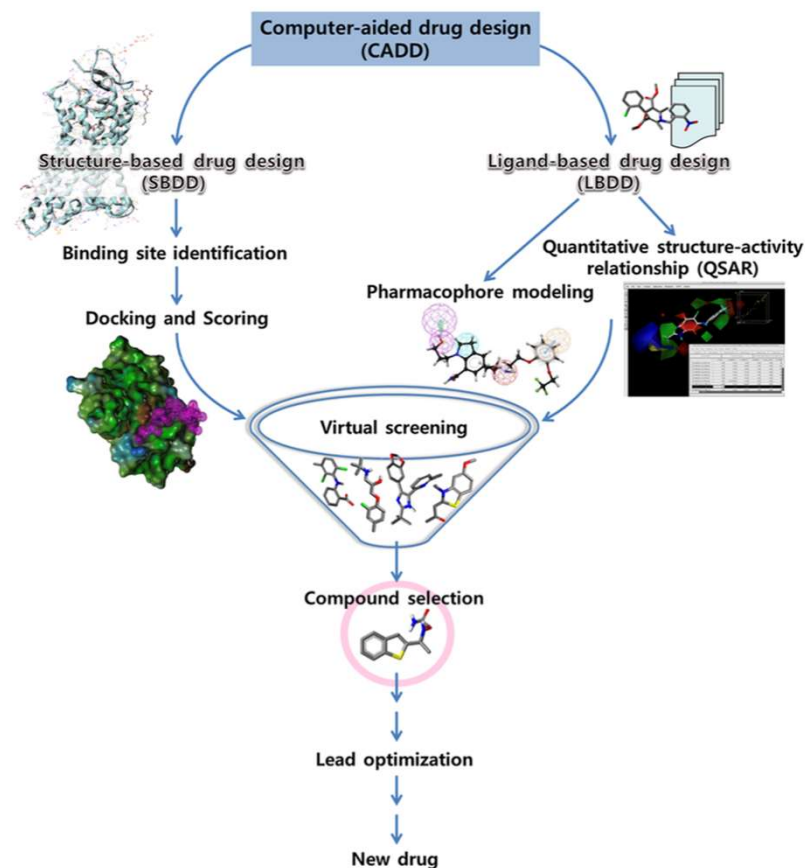
图表13：GPU硬件发展拓展了生物研究系统尺寸



## 1.3 底层技术决定AI效用，算法+数据是“限速步骤”

- 回顾发展沿用上述框架分析，分子生成（AI创新价值的代表）的限制因素为算法和数据：
  - **1、尚未有算法能够同时兼顾分子结构创新和合成性：**分子生成采用的是深度学习等模型，基于虚拟筛选形成的数据库进行“学习”，但是其学习框架依旧在CADD的基础之上：基于配体的药物设计（LBDD）/基于结构的药物设计（SBDD）。因此通过建模获得的分子符合基础的药化原理，但是往往会缺乏合成性原则导致无法在真实世界中合成。
  - **2、分子生成所需的数据更多维，且来源往往受到保护不会公开。**分子生成的数据需要考虑到多个维度：符合化学规则、生物途径、可合成等。同时，AI模型的训练需要兼顾成功与失败的数据，而往往公开的数据都是成功的案例。
- AI+制药公司的项目分为两类——仿制药与创新药。创新药的研发AI的表现如前文所述，尚未概念性验证，AI管线整体偏临床早期；而在仿制药的开发上，由于靶点、阳性参考物（仿制的创新药分子）结构等已确立，以CADD的传统方法即可完成开发，AI的边际效应可能不大。
- 综上，我们认为AI增效作用将会助力AI+药物研发在“1-10”阶段快速发展。同时，我们看好AI的创新价值，分子生成具备诞生BIC/FIC的潜力。实现这一价值，在于算法的突破和数据的获取或者分享，而两者的攻克都需长期的付出和投入。

图表14：CADD计算原理示意图



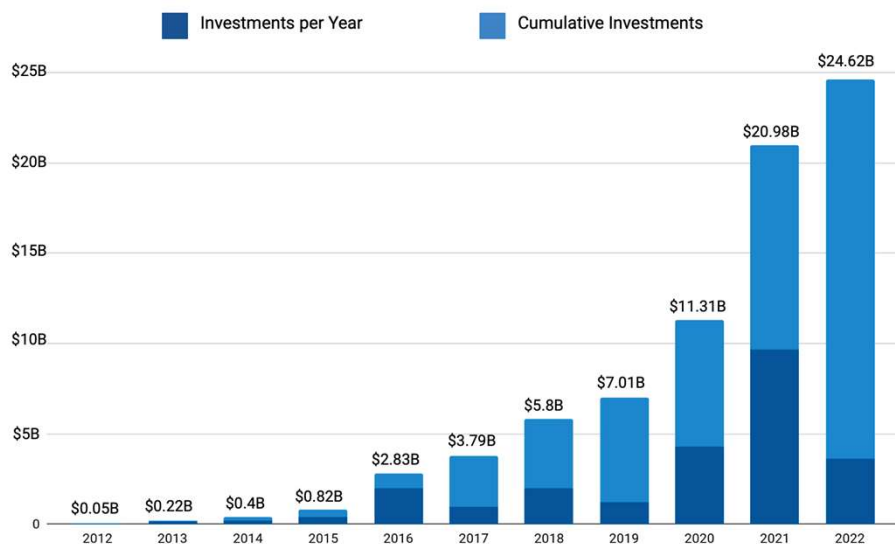
二

市场规模超百亿美金，分子实体是商业化重点所在

## 2.1 AI+制药融资热度有所起伏，产生结构性分化趋势

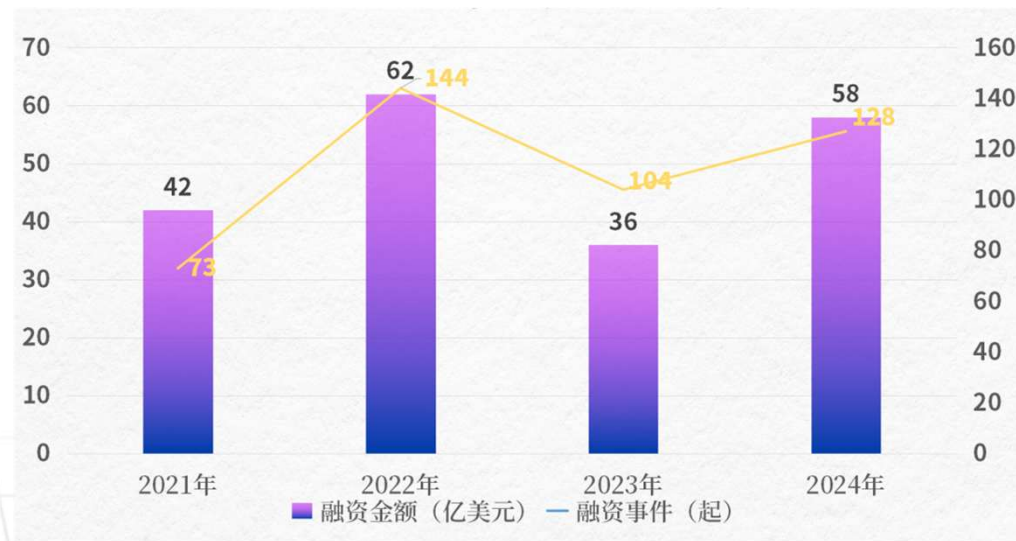
- 自2015年以来，全球AI+新药研发投融资快速增长，至2022年累计达246亿美元。国内融资数量及金额在2019年后开始大幅上涨，2021年达99.6亿元。2022受全球经济衰退影响，全球范围内的融资热度有所下降。
- **24年热度回暖，但产生结构性分化。**据智药局监测，2024年全球AI+药物研发融资总金额为58亿美元，融资笔数128起。其中美国AI药物研发融资事件64起，中国37起，其他国家和地区为27起，投融资活动仍然主要活跃在中国、美国和欧洲。从资金占比来看，美国仍然占据全球AI制药的融资高地，中国的AI制药资金占比则为10%。而刨除晶泰科技、BioAge Labs等公司IPO的影响，中国AI制药初创公司的融资额占比进一步下降，仅占全球融资额的8%。

图表15：2012-2022年AI+制药投融资情况



资料来源：Deep Pharma Intelligence，中邮证券研究所

图表16：2021-2024年全球AI制药融资统计



资料来源：智药局，中邮证券研究所

## 2.1 行业尚未形成统一的范式，企业间不尽相同

- 目前尚未有AI自主/辅助设计的药物上市。从技术角度来看，AIDD并未证明持续有效性且并未出现突破性的算法可以跳出CADD的框架。因此，行业内不存在统一的技术范式，企业间的算法和参与研发的环节不尽相同，详情可见下表。
- 国内外的企业参与研发环节有所差异，化合物筛选和先导化合物设计合成环节是大多企业聚焦的环节。我们认为背后的原因在于：1、两者的底层原理清晰，scaling up效果显著；2、随着研发环节逐渐向后，产生的数据量会逐渐减少，对于AI的依赖度降低，AI+制药公司切入难度大。
- 目前国内外龙头已表现出了全流程覆盖的趋势，如Schrodinger、Exscientia、英矽智能、Recursion等。但是同时也存在公司只参与部分环节。

图表17：国内外代表企业药物研发环节分布

药物研发环节	主要参与者
靶点发现	<ul style="list-style-type: none"> <li>Schrodinger、Relay Therapeutics、Recursion Pharmaceuticals、BenevolentAI、Deepmind</li> <li>英矽智能、英飞制药、科辉智药、燧坤智能、星药科技、亿药科技</li> </ul>
化合物筛选	<ul style="list-style-type: none"> <li>Schrodinger、Relay Therapeutics、Recursion Pharmaceuticals、Exscientia、Atomwise、BenevolentAI</li> <li>晶泰科技、英矽智能、百图生科、星亢原、冰洲石生物、望石智慧、宇道生物、英飞制药、科辉智药、剂泰医药、燧坤智能、康迈迪森、智药科技、星药科技、亿药科技</li> </ul>
先导化合物设计合成	<ul style="list-style-type: none"> <li>Schrodinger、Relay Therapeutic、Exscientia、Atomwise、BenevolentAI、Deepmind</li> <li>英矽智能、晶泰科技、百图生科、星亢原、冰洲石生物、望石智慧、深度智耀、宇道生物、英飞制药、科辉智药、剂泰医药、燧坤智能、康迈迪森、智药科技、智化科技、星药科技、亿药科技</li> </ul>
ADMET预测	<ul style="list-style-type: none"> <li>Schrodinger、Exscientia</li> <li>晶泰科技、望石智慧、英飞制药、剂泰医药、星药科技</li> </ul>
临床前结果预测	<ul style="list-style-type: none"> <li>英矽智能、冰洲石生物</li> </ul>
临床试验设计 患者招募，分组	<ul style="list-style-type: none"> <li>Lantern Pharma、LAM Therapeutics</li> <li>零氮科技</li> </ul>
药物重定位	<ul style="list-style-type: none"> <li>Lantern Pharma、Atomwise、BenevolentAI、Oncocross、LAM Therapeutics</li> <li>冰洲石生物、剂泰医药、燧坤智能、康迈迪森</li> </ul>

资料来源：量子位智库，中邮证券研究所

图表18：国内外代表企业药物研发环节分布

公司	AI平台			临床管线 <sup>1</sup>	實驗室	AI技術	主要資產1 期臨床試驗週期	重點領域	科研成果
	靶點發現	分子生成	臨床試驗						
InSilico Medicine	√ (商業可用)	√ (商業可用)	√ (商業可用)	2個I期 8個I期	智能實驗室 與濕實驗室	生成式AI/機 器學習/深度 學習	12個月	纖維化、腫瘤 學、免疫學	超過250篇同行評 審出版物
Relay Therapeutics	√ (內部)	√ (內部)	無公開可用資料	1個II期 1個I期 1個I期 1個I期	無公開可用 資料	機器學習/ 深度學習	33個月	精準腫瘤學與 遺傳疾病	在會議與研討會上 多次展示
Schrodinger	√ (商業可用)	√ (商業可用)	無公開可用資料	2個I期	無公開可用 資料	機器學習/ 深度學習	1期未完成	腫瘤學	數百篇關於生物製 劑及小分子藥物發 現的出版物
Recursion	√ (商業可用)	√ (商業可用)	√ (商業可用)	1個II期 3個I期 1個I期	濕實驗室與 乾實驗室	機器學習/ 深度學習/LLM	未披露	腫瘤學、炎症 與免疫學、罕 見疾病	多篇出版物與海報 以及開放數據集
AbCellera	無公開可用 資料	√ (商業可用)	無公開可用資料	2個I期	濕實驗室與 乾實驗室	機器學習	尚未啟動臨 床試驗	內分泌系統疾 病、代謝疾病 及免疫性炎症	多篇出版物與海報 以及開放數據集
XtalPi	√ (商業可用)	√ (商業可用)	無公開可用資料	2個I期 2個IND審批 2個IND申 請	濕實驗室與 乾實驗室	生成式AI/機 器學習/深度 學習	1期未完成	腫瘤學、神經 學、呼吸學、 炎症性炎症、 眼科	多篇出版物與海報

资料来源：英矽智能招股书，中邮证券研究所

## 2.1 行业尚未形成统一的范式，企业间不尽相同

图表19：全球AI制药领先企业的算法及应用

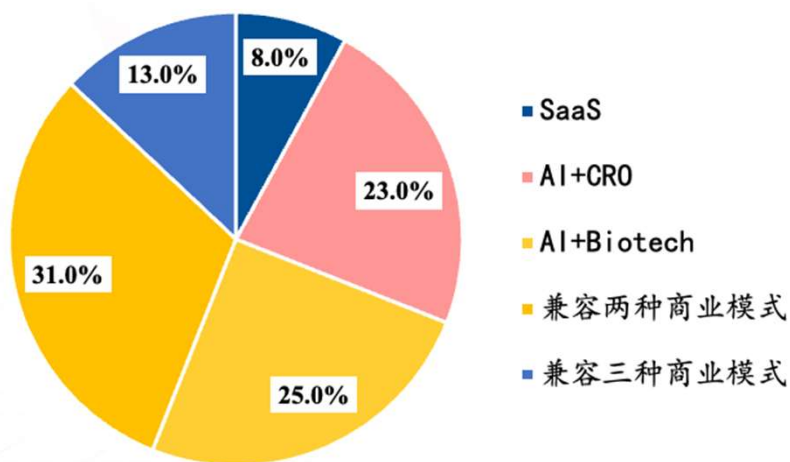
公司名称	使用的计算方法	简介
Ardigen	生物信息学,DL,ML,NLP	ardigen活跃于实验室信息管理系统、生物临床数据分析、大数据集成以及定制应用程序开发等领域
Atomwise	ML,DL(CNN),化学信息学	Atomnet是第一个运用深度CNN的药物发现算法，它已经在癌症、神经系统疾病、抗病毒药物、抗寄生虫药物及抗生素等方面进行探索。
BenchSci	NLP,ML,DL	试剂（如抗体）的解码开放和封闭数据访问，用可操作的见解展示已发布的数据
BenevolentAI	ML,DL,符号人工智能,NLP	从文本挖掘和语义链接演变为知识图来处理复杂多因素疾病，识别新靶点，小分子药物发现和患者分层。
BioSymetrics	NLP,ML,DL	处理原始表型、成像、药物和基因组数据集。允许研究人员将快速分析和机器学习功能集成到现有业务流程。
C4X Discovery	ML,DL,化学信息学,生物信息学	C4X创新的基于DNA的目标识别平台(Taxonomy3(R))人类基因数据集，以确定新的患者特异性靶点。
CytoReason	ML,DL,符号人工智能,化学信息学,生物信息学	CytoReason可以访问无与伦比的专有和公共数据，并结合尖端机器学习技术，创造其独特的生物疾病、组织和药物模型。
EURETOS	ML,DL,生物信息学	Euretos通过用户友好的应用程序提供基于云的研发平台的直接访问入口，还允许集成公司专有数据和安全环境中的公共数据。
Exscientia	ML, DL,生物信息学,化学信息学	该公司使用ML来预测ADME，新颖性，合成可达性，分子药理学。
Healx	ML,NLP,符号人工智能,化学信息学,生物信息学	Healx AI平台使用自然语言处理从已发表的资料、补充生物医学数据库和专有技术、策划数据中提取疾病知识。
IKTOS	ML,DL,化学信息学	Iktos已经发明并正在开发一种基于DL的配体新药设计，专注于多参数优化(MPO)
Insilico Medicine	DL,GANs,GANs+,强化学习,符号信息学,ML,化学信息学,生物信息学	全面的DL管线。生物学:信号通路，用于目标ID的DNNs和HTS分析。化学:GANs- RL用于新分子生成。
MindRank	小样本学习,ML,DL	MindRank旨在利用机器学习来设计和开发小分子，具有良好临床前疗效和安全性的药物。
NURITAS	DL,生物信息学	预测食源性生物活性肽的治疗潜力。允许研究人员:经济有效地开发针对特定疾病的天然食物来源的高度靶向治疗方法。
OWKIN	ML,联邦学习	Owkin结合了生物学和机器学习的专业知识来推进精准医学。
Peptilogics	生成式AI,DL,生物信息学,化学信息学,计算生物物理学,量子化学	Peptilogics的Nautilus平台结合了生成式人工智能、预测模型、计算生物物理学、量子化学和专门建造的超级计算以生产针对不同靶点和治疗的多参数优化引物区域。
STRUCTURA	ML(随机梯度下降和分支和边界最大似然优化)	cryoSPARC系统™能够实现高通量的蛋白质结构发现，在机器学习的帮助下，从低温电镜数据中提取分子复合物。
XtalPi	ML	XtalPi的ID4平台提供了精确的物理化学和化学预测用于药物设计的小分子候选药物的药性，固体形态选择，以及药物开发的其他关键方面。

资料来源：Deep Pharma Intelligence，中邮证券研究所，（注：参考资料有时限性，表格可能无法反映所有公司最新的技术迭代情况）

## 2.2 商业模式：Saas、AI+CRO/Biotech

- 目前行业的商业模式分为三种，其中CRO/Biotech的模式占主要：
  - ❑ SaaS：公司向药企或者药物研发CRO出售软件服务以获取收入。
  - ❑ AI+CRO：利用AI技术优势为药企/CRO等公司提供药物研发的外包服务。
  - ❑ AI+Biotech：以AI赋能，管线自研的Biotech路线，通过license out/合作/自主推进管线上市的方式创收。

图表20：AI制药商业模式分布情况



资料来源：智药局，蛋壳研究院，中邮证券研究所

图表21：AI制药商业模式从单一到混合



资料来源：蛋壳研究院，中邮证券研究所

## 2.2 商业模式：SaaS并非好的新入局模式

- SaaS模式的优劣：
  - 优点：对于数据的依赖度低、不承担对于药物管线研发失败的风险、资本投入低、现金回流快、利润率高；
  - 缺点：赋能新药研发环节有限、技术服务收入低、市场规模小，收益贡献度占比不高
- 我们认为SaaS的商业模式并不适合行业参与者，基于以下判断：1、SaaS为To B业务，客户群体数量有限、市场容量不高；2、头部公司已触达天花板，其余参与者避开非理性竞争。
- 头部公司Schrödinger的软件服务行业领先，经过30多年的技术积累，全球TOP20的制药巨头均为SCHRÖDINGER的客户。2024软件收入达1.8亿美元，大客户（超10万美元）留存率达95%。2020-2024软件增长率处于下滑趋势，25年Q1-3有所改善，但综合来看我们认为这预示着行业规模有限，公司或已接近SaaS天花板。而对于竞争者而言，需要证明自身软件产品的非劣/更优，意味着难度大投入高，回报率不理想。

图表22：薛定谔2019-2025Q3收入情况

	软件收入 (百万美元)	药物发现收入 (百万美元)	软件收入 yoy(%)	药物发现收入 yoy(%)	净利润 (百万美元)
2019	66.7	18.8			-24.6
2020	92.5	15.6	39%	-17%	-24.5
2021	113.2	24.7	22%	59%	-100.4
2022	135.6	45.4	20%	84%	-149.2
2023	159.1	57.5	17%	27%	40.7
2024	180.4	27.2	13%	-53%	-187.1
2025Q1-Q3	130.2	38.4	29%	107%	-135.8

资料来源：Wind，公司年报，中邮证券研究所

图表23：2022-2024薛定谔不同ACV客户留存情况

年份	ACV500万+ 美元客户数	ACV100万+ 美元客户数	ACV50万+ 美元客户留存率	ACV50万+ 美元客户留存率	ACV10万+ 美元客户留存率
2022	4	18	-	-	均大于92%
2023	4	27	54	98%	
2024	8	31	61	100%	95%

资料来源：Wind，公司年报，中邮证券研究所

## 2.2 商业模式：回归研发本质，分子实体仍是商业化重点

- AI+CRO：通过技术服务外包的形式与大量药企以及CRO公司达成合作，共同推进管线开发并获取服务收入，同时利用广泛合作沉淀多维度的数据以支持其算法模型不断进行优化和迭代。
- AI+Biotech：通过推进自有管线以更快地验证公司AI算法平台能力，同时以自主/合作推进管线上市或以授权交易阶段性管线成果给具有临床开发能力的药企的方式进行创收。
- **上述两者区别在于知识产权的归属，但从商业实现的目的来看均系回归研发的本质，即依赖分子实体商业化创收。从行业的发展角度来看，我们认为两者都可以通过合作获取多维实验数据，用于公司自身算法训练迭代优化，从而实现正向发展的良性循环。**

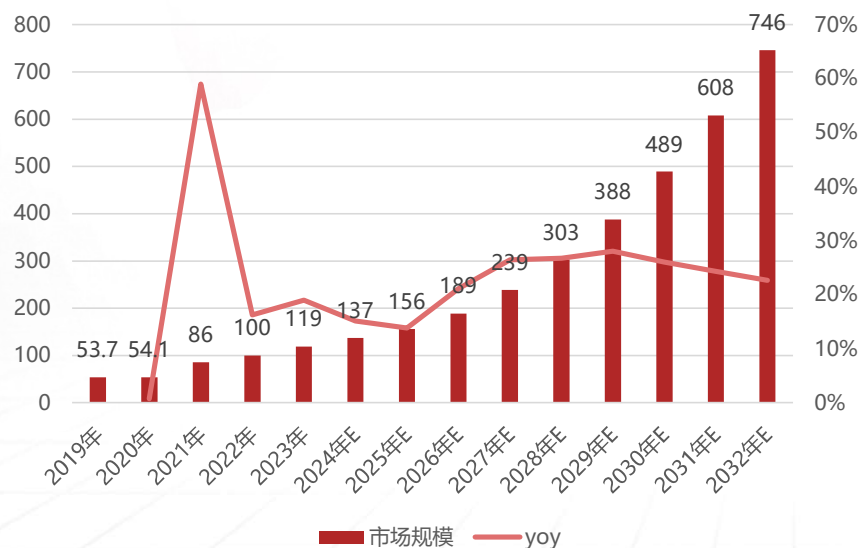
图表24：AI制药商业模式总览

商业模式	说明	优点	缺点
SaaS	提供软件平台服务，向药企、药物研发CRO出售软件	对数据依赖度低、不承担对于药物管线研发失败的风险、资本投入低、现金回流快、利润率高	赋能新药研发环节有限、技术服务收入低、市场规模小、收益贡献度占比不高
AI CRO	通过技术服务外包的形式与药企、CRO等公司合作，共同推进管线开发并获取服务收入，同时利用广泛合作提升自身技术平台	不用对药物研发最终结果负责，现金回流快、研发风险低； 利用广泛合作沉淀多维度的数据以支持其算法模型不断进行优化和迭代	需要大量高质量药物研发数据、同时需要清洗整合数据并建立干湿一体化，投入支出高； 需要高客户覆盖度来打开收入天花板
AI Biotech	开发内部管线为主，推进管线上市或license out	新药物上市利润空间巨大、市场规模不受限制	药物开发风险高、管线价值汇报周期长、资本投入大

## 2.3 全球AI赋能药物研发费用市场规模已超百亿美金

- 在上述讨论中，我们认为商业模式回归到以药物研发为导向的AI+CRO/Biotech。参考英矽智能招股书，2023年全球AI赋能药物研发费用市场已达119亿美金，2032有望增至746亿美金，CAGR为22.6%。

图表25：全球AI赋能药物研发费用市场规模（亿美元）

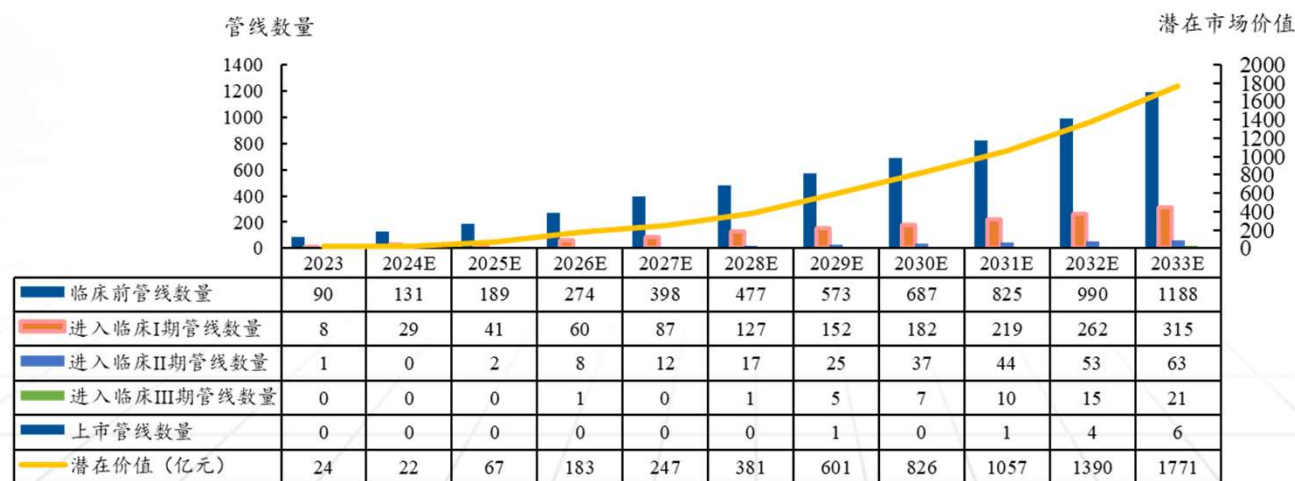


资料来源：英矽智能招股书，弗若斯特沙利文，中邮证券研究所

## 2.3 全球AI赋能药物研发费用市场规模已超百亿美金

- 此外，我们参考2023的公开资料以CRO的收入方式来测算国内管线的潜在市场空间，参考了Benevolent自研管线在不同临床阶段对外的付款价格。根据量子位数据来源，我们更新并梳理了国内管线研发的进度，临床成功率引用QLS和Frost & Sullivan的公开数据，据测算AI+制药的国内潜在市场价值有望2028年达380亿，2033年超1770亿，具体数据可参见下表。
- 关键假设：
  - 临床前管线数量的增速：参考NMPA公布数据，2018-2021化学药物IND数量CAGR为41%。我们假设2024-2028年，AI+药物的增长率略高于整体水平，而在2028-2033年达到稳定增长期。
  - 临床研发周期：AI目前的管线作用发挥在临床前，因此我们假设所有的临床前管线在1年内推进结束。临床后的周期参考传统药物开发经验，QLS数据来源：从I期临床到批准平均需要10.5年，这包括I期临床2.3年，II期临床3.6年，III期临床3.3年，审批阶段1.3年。

图表26：国内AI制药潜在市场价值测算



资料来源：NMPA，量子位，QLS，Frost&Sullivan，中邮证券研究所

三

行业竞争：高质量数据的生产能力是核心竞争力

## 什么是行业的竞争力？

- 我们认为算法（模型）+数据是行业核心技术迭代升级的关键生产要素，其生产的壁垒在于研发人才+成熟的实验平台。以下指标可作为公司竞争力的参考指标：
  - I. 研发人才：CADD/AIDD+药物研发团队兼备是算法突破的关键构件。
  - II. 高质量数据生产能力：具备成熟的实验平台能够实现数据积累和算法迭代之间的闭环和循环
  - III. 管线进度和数量：技术实力的背书
  - IV. 合作企业的数量和合作深度。
- 对于药物研发最重要的是获得一个有效的药物分子，并顺利通过三期临床上市。而有效的分子在AIDD理论上由算法+数据决定。算法与数据是深度绑定的。数据是算法开发的基础，也是算法迭代升级的必需品。**短期来看，算法是技术类型公司获得研发合作的关键因素，而更好的数据促使AI做出更好的决策，数据是行业内长期的关键竞争要素**

## 3.1 算法、数据重要性大于算力

- **我们认为算力并非关键要素**，基于两个判断：
  - I. 算力影响的是计算处理的时间，结果上表现为数天的延迟。然而对于整体药物研发周期而言，数天的影响可以忽略不计；
  - II. 从行业产业链来看，上游的互联网巨头可提供算力支持。同时以NVIDIA为代表的硬件公司前瞻性布局AI赛道，芯片系统满足AI+医疗、生命科学应用的需求。
- **什么是好的算法？** 算法需要完成对原有CADD框架的突破，做到分子创新性和可解释性兼顾。算法的突破是方法学的创新，属于科研范畴的产出结果。因此**研发人才是算法突破的关键因素**。我们认为具备领先的AI/CADD算法开发能力+具备丰富药物开发经验的人才团队是算法竞争力的重要指标。
- **好的数据的标准是什么？大规模+高质量**。大规模的数据是AI模型训练准确度提高的前提条件。高质量有AI和实验的双重考量：从AI角度来看，需满足以下特征：结构良好、完全注释、可重现和具有鲁棒性；从实验角度来看，需要多维度、多体系、兼顾成功与失败的训练集等要素。**大规模高质量的数据是行业稀缺资源，原因在于纳入训练集的数据要求苛刻+数据不共享（“数据孤岛”）。**
- 由于实验数据涉及核心利益，长期来看，“数据孤岛”的行业特征不会改变。因此，竞争力的判断关键在于**哪些公司拥有或者具备生产高质量数据的能力**。

## 3.2 高质量数据生产能力是核心竞争力

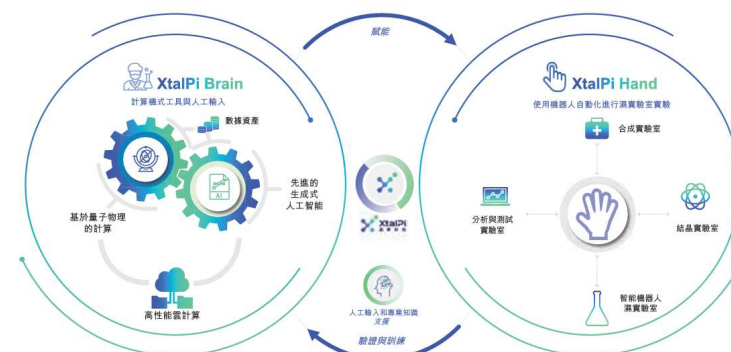
- 我们认为传统的制药企业和CRO巨头并不一定拥有高质量的数据。原因在于：
  - ❑ 传统实验方式难以达到AI训练的数据采集点的苛刻要求。如，代谢类数据对于药物反应的时间记录往往跨度较大，使得AI预测的准确度提升受阻。同时，AI需要足量的失败案例的数据集，与研发成功的目的相悖。
  - ❑ 实验数据需要经过清洗、信息化才能用于AI训练。清洗在于对数据进行人工的判别是否符合训练标准，需要药化专家的经验判断。信息化需要CADD/AIDD的专业人员的处理。传统药企和CRO可能没有相应的人才队伍/研发团队缺乏人工智能专业知识。
- 高质量实验数据生产能力在于实现数据积累和算法迭代之间的闭环和循环。对应到公司层面：具备成规模的实验平台+ AIDD/CADD计算方面具有丰富积累。基于AI训练需求可以设计针对性的实验方案，进而利用实验所生成的数据“喂养”AI模型，进入迭代循环。
- 结合上述，我们认为两类公司具备高质量的数据的生产能力：1、内部具有实验平台的AI技术型公司；2、内部具有CADD经验累积的CRO公司。即试验平台+AI团队+药物研发团队是自身壁垒构建的三要素

图表27：AI制药研发环节的主要挑战



资料来源：Deep Pharma Intelligence，中邮证券研究所

图表28：AI制药研发流程之间的循环关系

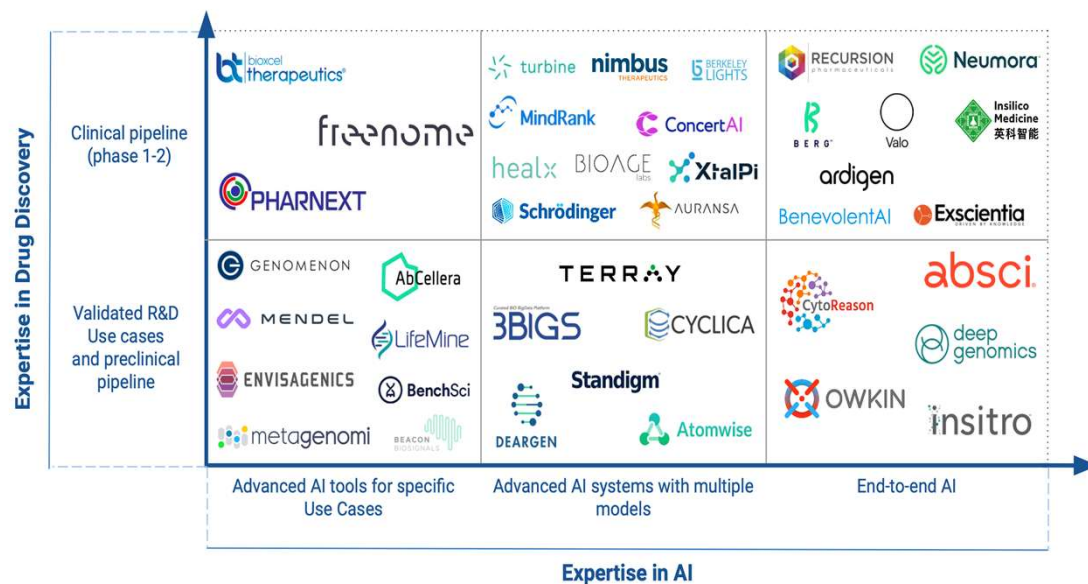


资料来源：晶泰科技招股书，中邮证券研究所

## 3.3 行业格局逐渐分化，马太效应利好行业领跑者

- **管线开发进度差距逐渐显现，马太效应利好领跑者。**早期创立的企业凭借先发优势完成了技术优势和项目经验积累，在行业内处于领先地位。管线顺利进入临床阶段是对AI+制药企业的研发实力的背书。我们认为，行业会形成马太效应的原因在于技术能够scale up：同一技术/模型在不同训练场景下能够完成迭代，数据越多，迭代越完善。而技术越完善的公司意味着更多的合作机会，通过合作创收的同时利用项目数据继续迭代自身算法。
- 在利用管线进度和AI技术的双维度考量下，全球Top40的AI+制药行业格局逐渐分化，详情可见下图。值得注意的是，国内的晶泰科技和英矽智能位列图中的高位。

图表29：国内外代表企业在AI与药物发现维度比较



资料来源：Deep Pharma Intelligence，中邮证券研究所

# 四

## 相关标的

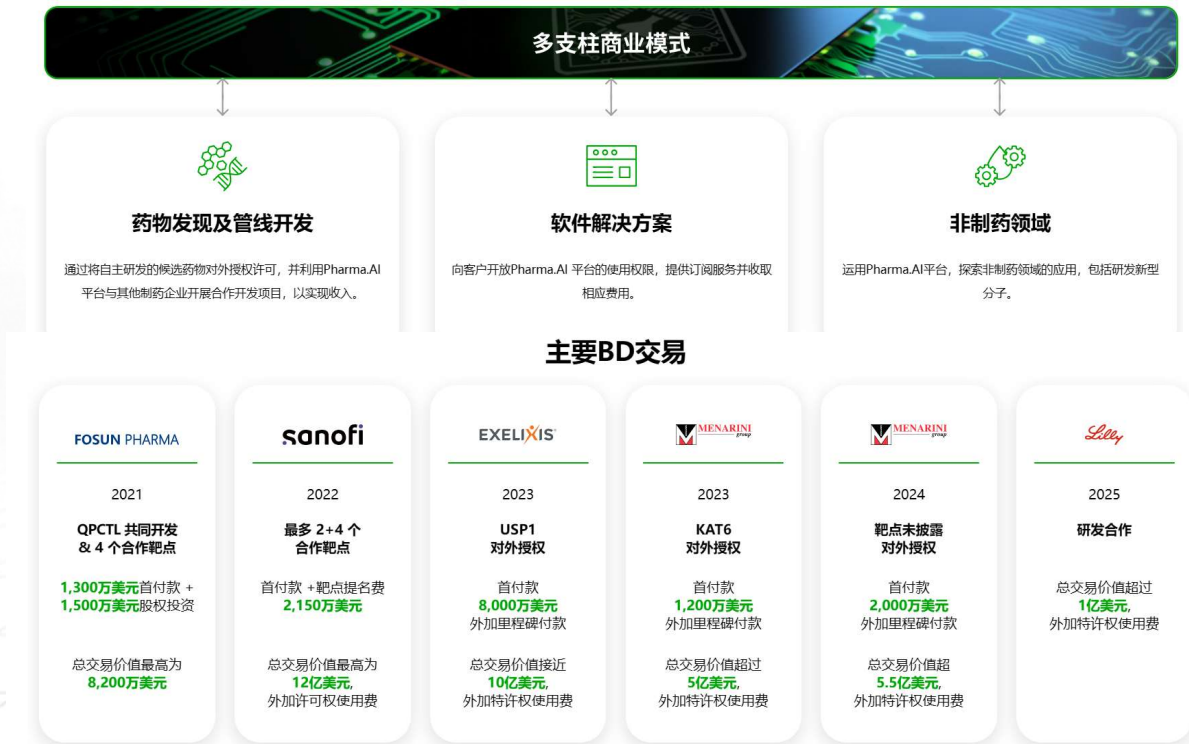
## 相关标的

- 建议关注领先布局者：英矽智能、晶泰控股、泓博医药、成都先导等。

# 英矽智能：Pharma.AI+自动化实验室，加速药物创新

- 英矽智能是一家由生成式人工智能驱动的全球先锋生物科技公司，利用其专有的 Pharma.AI 平台和先进的自动化实验室，加速药物发现并推动生命科学领域的创新。通过整合人工智能和自动化技术，英矽智能正在为纤维化、肿瘤学、免疫学、疼痛、肥胖和代谢紊乱等未满足的疾病领域提供创新药物解决方案。公司具备多支柱商业模式，同时已与海内外多个知名药企达成了BD交易合作。

图表30：英矽智能：多支柱商业模式，达成多个BD交易合作



资料来源：英矽智能官网，中邮证券研究所

# 英矽智能：Pharma.AI+自动化实验室，加速药物创新

➤ 英矽智能由人工智能驱动发现的产品管线中，已有10个分子获得临床试验许可，其中进展最为领先的Rentosertib（原名ISM001-055），已完成2a期临床研究并获得令人鼓舞的积极结果。

图表31：英矽智能管线概览

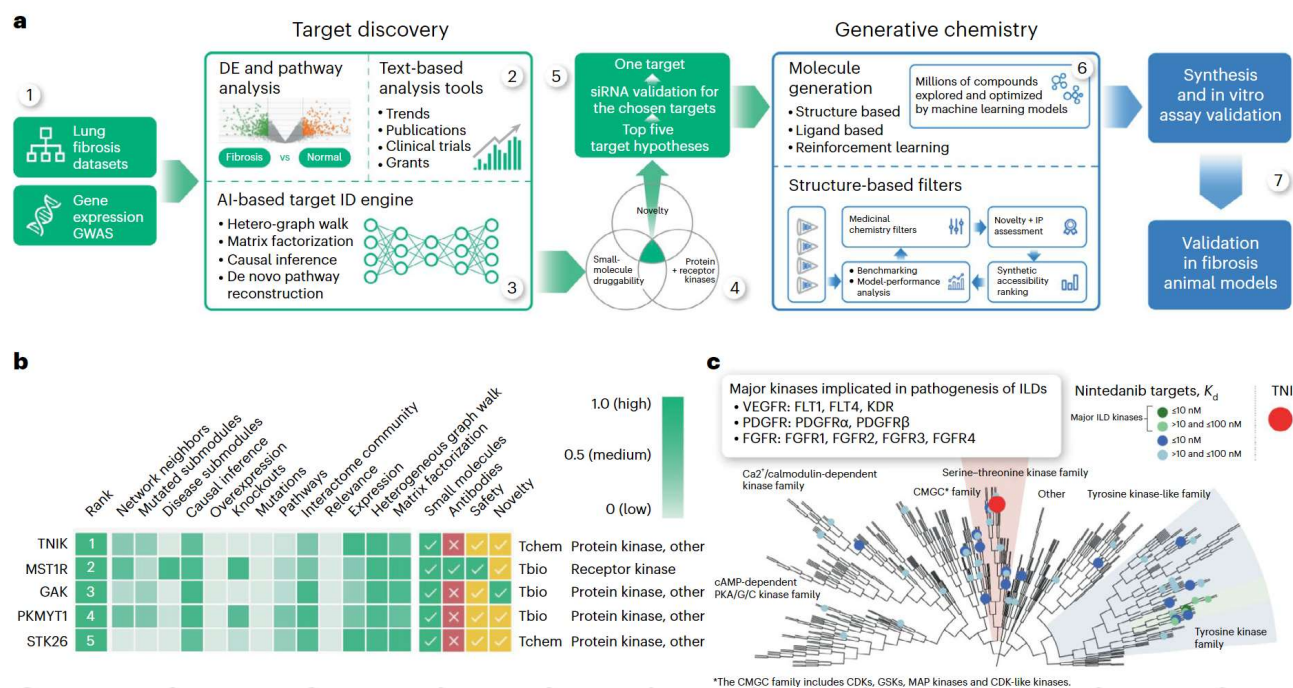


资料来源：英矽智能官网，中邮证券研究所

# 英矽智能：Pharma.AI+自动化实验室，加速药物创新

- **公司AI平台的创新分子正逐步兑现平台价值**：英矽智能利用其PandaOmics平台，通过深度学习和因果推断，从复杂的组学数据中识别出了TNIK (TRAF2- and NCK-interacting kinase) 作为特发性肺纤维化 (IPF) 的潜在抗纤维化靶点。通过**Chemistry42**平台进行分子生成和优化，最终筛选出了候选分子 INS018\_055。该分子在体内外实验中表现出强效的抗纤维化活性，并具有良好的药代动力学性质。

图表32：英矽智能TNIK分子开发流程图



# 晶泰控股：AI+机器人赋能，溢出效应覆盖多领域

- 晶泰是一家基于量子物理，以人工智能赋能和机器人驱动的创新平台型科技公司。利用量子物理、人工智能、云计算、机器人实验等前沿技术领域打造了面向未来产业的创新型研发平台。该平台凭借共性技术集成优势，满足微观分子世界探索的需求，在生物医药行业成效显著，大幅提升了研发效率和成功率，其溢出效应已覆盖至新材料、石油化工、新能源、农业和沙漠治理等领域。

图表33：晶泰科技技术部分核心亮点展示



# 晶泰控股：AI+机器人赋能，溢出效应覆盖多领域

- **公司平台在药物发现具备极高效率：**在无参考化合物的情况下，28天内利用XFEP和主动学习成功预测GPX4活性位点中4个关键药效团，实现特大化合物库的快速筛选。并结合自动化合成快速筛选出3个靶向GPX4活性分子( $IC_{50} < 10\mu M$ )。

图表34：晶泰XFEP平台效率大幅提升



# 五

## 风险提示

## ■ 风险提示:

- 研发失败风险：AI+制药本质仍属于新药研发的范畴，而新药研发的风险性高，从临床前到临床试验再到上市需要经过多环节的验证，存在研发失败风险。
- 行业竞争加剧风险：随着投融资等外部因素带动，行业竞争可能会趋于激烈。
- 地缘政治风险：BD合作模式部分涉及跨国跨地区，存在地缘政治因素扰动风险。
- 政策变动风险：行业易受监管、药物审批等政策扰动。

# 感谢您的信任与支持!

THANK YOU

## 分析师声明

撰写此报告的分析师（一人或多人）承诺本机构、本人以及财产利害关系人与所评价或推荐的证券无利害关系。

本报告所采用的数据均来自我们认为可靠的目前已公开的信息，并通过独立判断并得出结论，力求独立、客观、公平，报告结论不受本公司其他部门和人员以及证券发行人、上市公司、基金公司、证券资产管理公司、特定客户等利益相关方的干涉和影响，特此声明。

## 免责声明

中邮证券有限责任公司（以下简称“中邮证券”）具备经中国证监会批准的开展证券投资咨询业务的资格。

本报告信息均来源于公开资料或者我们认为可靠的资料，我们力求但不保证这些信息的准确性和完整性。报告内容仅供参考，报告中的信息或所表达观点不构成所涉证券买卖的出价或询价，中邮证券不对因使用本报告的内容而导致的损失承担任何责任。客户不应以本报告取代其独立判断或仅根据本报告做出决策。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，中邮证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。

中邮证券及其所属关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，也可能为这些公司提供或者计划提供投资银行、财务顾问或者其他金融产品等相关服务。

《证券期货投资者适当性管理办法》于2017年7月1日起正式实施，本报告仅供中邮证券签约客户使用，若您非中邮证券签约客户，为控制投资风险，请取消接收、订阅或使用本报告中的任何信息。本公司不会因接收人收到、阅读或关注本报告中的内容而视其为签约客户。

本报告版权归中邮证券所有，未经书面许可，任何机构或个人不得存在对本报告以任何形式进行翻版、修改、节选、复制、发布，或对本报告进行改编、汇编等侵犯知识产权的行为，亦不得存在其他有损中邮证券商业性权益的任何情形。如经中邮证券授权后引用发布，需注明出处为中邮证券研究所，且不得对本报告进行有悖原意的引用、删节或修改。

中邮证券对于本申明具有最终解释权。

# 免责声明

## 公司简介

中邮证券有限责任公司，2002年9月经中国证券监督管理委员会批准设立，是中国邮政集团有限公司绝对控股的证券类金融子公司。

公司经营范围包括：证券经纪，证券自营，证券投资咨询，证券资产管理，融资融券，证券投资基金销售，证券承销与保荐，代理销售金融产品，与证券交易、证券投资活动有关的财务顾问等。

公司目前已经在北京、陕西、深圳、山东、江苏、四川、江西、湖北、湖南、福建、辽宁、吉林、黑龙江、广东、浙江、贵州、新疆、河南、山西、上海、云南、内蒙古、重庆、天津、河北等地设有分支机构，全国多家分支机构正在建设中。

中邮证券紧紧依托中国邮政集团有限公司雄厚的实力，坚持诚信经营，践行普惠服务，为社会大众提供全方位专业化的证券投、融资服务，帮助客户实现价值增长，努力成为客户认同、社会尊重、股东满意、员工自豪的优秀企业。

## 投资评级说明

投资评级标准	类型	评级	说明
报告中投资建议的评级标准： 报告发布日后的6个月内的相对市场表现，即报告发布日后的6个月内的公司股价（或行业指数、可转债价格）的涨跌幅相对同期相关证券市场基准指数的涨跌幅。 市场基准指数的选取：A股市场以沪深300指数为基准；新三板市场以三板成指为基准；可转债市场以中信标普可转债指数为基准；香港市场以恒生指数为基准；美国市场以标普500或纳斯达克综合指数为基准。	股票评级	买入	预期个股相对同期基准指数涨幅在20%以上
		增持	预期个股相对同期基准指数涨幅在10%与20%之间
		中性	预期个股相对同期基准指数涨幅在-10%与10%之间
		回避	预期个股相对同期基准指数涨幅在-10%以下
	行业评级	强于大市	预期行业相对同期基准指数涨幅在10%以上
		中性	预期行业相对同期基准指数涨幅在-10%与10%之间
		弱于大市	预期行业相对同期基准指数涨幅在-10%以下
	可转债评级	推荐	预期可转债相对同期基准指数涨幅在10%以上
		谨慎推荐	预期可转债相对同期基准指数涨幅在5%与10%之间
		中性	预期可转债相对同期基准指数涨幅在-5%与5%之间
		回避	预期可转债相对同期基准指数涨幅在-5%以下

## 中邮证券研究所

### 北京

邮箱：yanjiusuo@cnpsec.com

地址：北京市东城区前门街道珠市口东大街17号

邮编：100050

### 上海

邮箱：yanjiusuo@cnpsec.com

地址：上海市虹口区东大名路1080号大厦3楼

邮编：200000

### 深圳

邮箱：yanjiusuo@cnpsec.com

地址：深圳市福田区滨河大道9023号国通大厦二楼

邮编：518048



**中邮证券**

CHINA POST SECURITIES