DOI：10.61189/669152licoid

·专家述评·

人工智能对医学科研范式的重构

高承实[1*]，程元骏[2]
1. 安徽栈谷科技有限公司，池州 247100
2. 池州市人民医院，池州 247000

［摘要］ 21世纪以来，人工智能（artificial intelligence，AI）正在以前所未有的深度介入医学科研，推动其从“假设—验证”范式向“数据驱动—生成式”的认知结构跃迁。借助深度学习与生成式预训练模型，AI不仅正在文献综述、图像识别、临床试验设计、药物开发等方面重塑科研流程，更对医学研究的哲学基础、可解释性、科研伦理与评价机制提出了挑战。本文系统分析了AI在医学科研中从工具到协作者、从加速器到范式建构者的多重角色变迁，提出“可信、透明、可控”的AI框架应成为未来科研范式重构的制度基石。通过梳理AI介入下的典型案例与趋势，强调人机协同将成为医学知识生产的新常态，并呼吁构建跨学科共识机制，以确保医学科研的科学性、伦理性与创新性的协同进化。

［关键词］ 人工智能；医学科研范式；数据驱动科学；生成式预训练模型；科研伦理
［中图分类号］ R-3　　［文献标志码］ A

The reconstruction of the medical research paradigm by artificial intelligence

GAO Chengshi[1*], CHENG Yuanjun[2]
1. Anhui Stack Alley Technology Co., Ltd, Chizhou 247100, Anhui, China
2. The People's Hospital of Chizhou, Chizhou 247000, Anhui, China

［**Abstract**］ Since the beginning of the 21st century, artificial intelligence (AI) has been profoundly reshaping medical research, propelling its transition from the traditional "hypothesis-verification" paradigm towards a “data-driven, generative” cognitive structure. Leveraging deep learning and generative pre-trained models, AI is not only transforming research workflows in areas such as literature review, image recognition, clinical trial design, and drug development, but also challenging the philosophical foundations, interpretability, ethical considerations, and evaluation mechanisms of medical research. This paper systematically analyzes the multifaceted evolution of AI's role in medical research—from a tool to a collaborator, and from an accelerator to a paradigm architect. It proposes that a framework of “trustworthy, transparent, and controllable” AI should serve as the institutional cornerstone for reconstructing future research paradigms. By examining representative case studies and emerging trends under AI's influence, the paper emphasizes that human-AI collaboration will become the new norm in medical knowledge production. It further calls for establishing interdisciplinary consensus mechanisms to ensure the harmonious progression of scientific rigor, ethical integrity, and innovative capacity in medical research.

［**Key Words**］ artificial intelligence; paradigm of medical research; data-driven science; generative pre-training model; research ethics

21世纪以后，医学科研进入了由数据与计算驱动的新阶段。从基因组学的高通量测序到电子病历的全面普及，从影像医学的数字化转型到流行病学实时预测模型的构建，医学已经成为了一门高度依赖信息与数据的科学。然而，与数据爆炸性增长相伴随的，并非科研效率的同步跃升。相反，医学研究面临着越来越多的瓶颈，实验成本高昂、临床试验周期冗长、文献重复冗杂、结果可重复性下降等问题愈发显著。

在此背景下，人工智能（artificial intelligence，AI）作为一种强大的分析与生成工具，逐渐成为医学研究的核心技术。其在医学中的应用，从早期的影像识别与辅助诊断，发展到当下的药物发现、精准医疗、流行病预测乃至科研流程自动化。特别是2022年生成式预训练模型ChatGPT取得突破以后，AI就已经不再仅仅是“技术助手”，而是开始被视作

［收稿日期］ 2025-06-10　　［接受日期］ 2025-06-25
［作者简介］ 高承实，博士，副教授.
*通信作者（Corresponding author）. Tel: 056-62028361, E-mail: 13838001036@163.com

具有“拟认知能力”的科研参与者。这种角色变化，对医学科研范式构成了深刻挑战。

传统的医学研究范式深植于“假设-实验-验证-推广”的线性科学逻辑之中，由研究者提出假设、设计实验、收集数据、得出结论。这一范式强调因果推断、变量控制和统计显著性，是20世纪医学进步的重要保障。但这一范式也有其局限，尤其在面对高度复杂、非线性、多因素交织的现实医疗情境时，往往显得力不从心。而AI，尤其是深度学习模型，能够在无需明确因果路径的前提下，从大量数据中“学习出”潜在的结构模式，直接输出预测结果或研究假设，从而挑战了以“人类提出假设”为中心的科研逻辑。

AI对科研的介入，不仅在“工具层”提高了效率，更在“方法层”改写了流程，甚至在“认知层”也挑战了人类研究者的主体地位。它可以自动分析成千上万篇医学文献，识别研究空白；可以在药物筛选中模拟分子相互作用，发现传统路径难以找到的候选物质；甚至可以辅助设计临床试验方案、优化受试者招募与随机分组策略。这种广泛渗透正在改变我们对科学研究定义的根本理解。

当然，AI并非万能。其“黑箱”特性使得可解释性成为挑战；其对数据质量高度敏感，这使得训练数据中的偏见可能被放大；其在科研伦理、知识产权和信用归属等方面的模糊地带，也引发了广泛争议。面对AI在医学科研中的迅速推进，学界正陷入一场深刻的范式争论，即我们应将AI视为人类认知能力的延伸，还是一种外部自动体制的崛起？AI主导下的研究发现能否真正代表科学？如果“发现”不再依赖人类的理论直觉和逻辑归纳，科研的本质是否正在发生变化？

本文以医学科研为核心场景，系统分析AI如何重构其研究范式，从理论、方法、实践与伦理多个维度展开反思：(1)AI是否仅是工具，还是正在成为研究方法与知识生产的新机制？(2)在AI广泛介入后，传统的“假设—验证”路径是否会给“数据驱动—生成”路径让位？(3)科研成果的可验证性、可重复性与可解释性如何在AI参与下重新界定？(4)医学科研人员的角色将如何转变？“人机协同”是否是新的科研范式？(5)面对AI带来的新机遇与新风险，科研伦理与制度应如何回应？

通过对上述问题的系统探讨，本文希望不仅揭示AI对医学科研技术流程的改造，更试图把握其背后深层的科学哲学与知识生产范式的变动。当AI以前所未有的方式参与到“我们如何认识疾病、理解生命、治疗人类”这一根本问题时，我们不能只将其看作效率提升工具，而应严肃地思考，我们是否已经站在了一次新的“科研革命”的门槛上？

1 医学科研范式的传统结构

在当代科学体系中，医学研究曾长期被视为最接近“理性与实验”理想的范式领域之一。作为自然科学与生命科学交汇的核心，医学科研不仅要在实验室中追求严谨的机制阐释，还必须与真实世界中的临床场景衔接。这一双重使命塑造出一种独特而复杂的科学范式，它既承袭了20世纪波普尔式的“证伪主义”精神，又深受经验主义和统计方法论的洗礼。在AI等新技术出现之前，这套传统范式已逐步固化，并形成以下几个核心结构要素。

1.1 “假设—验证”的逻辑主轴 传统医学科研遵循的基本逻辑框架是“假设驱动的实证研究”(hypothesis-driven research)。这一框架强调研究应从一个清晰可界定的科学假设出发，通过严格设计的实验或观察研究对其加以验证或证伪。在这一过程中，科学家作为认知活动的主体，被视为是提出问题、构造模型、设计实验并解释数据的核心力量。任何实验活动的起点，必须是人类主导的理论推演或临床观察，从中提出一个可操作、可测量的问题。

例如，一项有关新型抗癌药物的研究，通常以“该药物是否能显著抑制肿瘤细胞生长”为假设，通过体外实验、动物模型、Ⅰ~Ⅲ期临床试验逐步推进。而数据，仅仅是对假设的支撑或反驳工具。假设先行、数据验证，这一过程体现出强烈的人本主义科学精神，也构成医学科研长期以来的范式基石。

1.2 变量控制与因果推断的核心地位 医学科研的严谨性很大程度上体现在对变量控制的强调上。无论是基础医学研究中的单变量实验，还是临床试验中的随机对照设计，核心都是通过排除混杂因素、隔离变量影响来接近“因果推断”的理想目标。

这一思维模式源自物理学的“实验理性”，在20世纪初被统计学方法强化，尤其体现在Fisher创立的显著性检验框架与随机化方法中。其核心理念是，只有控制好影响因素，才能得出“X是否导致Y”的因果关系，而非仅仅观察到“X与Y的相关性”。

例如，在评估一个降血压新药的有效性时，必须通过双盲随机对照试验(randomized controlled trial，RCT)排除安慰剂效应与医生偏见，通过随机化分组避免选择偏倚，通过显著性检验确保结果的合

理性。这种对“内在因果”的执着，使医学科研在科学体系中长期保持高度的权威性。

1.3　统计方法作为标准分析工具　自20世纪中叶以来，统计学逐渐成为医学科研的“语言与法则”。无论是用于群体画像的描述性统计，还是用于假设检验的推论统计，统计工具已全面渗透到医学研究的设计、执行与结果解释中。医学科研论文无一例外地包含了统计检验，*P*值、置信区间、方差分析、多重比较校正、贝叶斯推断等成为科研的“标配”。在临床研究中，统计显著性甚至一度被等同于“科学性”。

这种依赖统计方法的传统，也间接形成了一种方法论保守性。只有符合统计模型假设的数据才被认为是“可用”的数据，只有满足统计标准的结果才被承认具有“科学意义”。这在给研究过程带来高度规范化与可重复性的同时，也阻碍了医学现象中大量复杂、动态、非线性机制的呈现。

1.4　实验室、期刊与同行评议的科研组织形态　医学科研的知识生产并非孤立完成，而是嵌入在一整套组织结构中。这套结构以大学或研究机构的实验室为节点，以学术期刊与会议为平台，以同行评议制度为评价机制。它既保证了科研成果的积累与传播，也塑造了研究范式的稳定性。

尤其在医学领域，期刊排名、影响因子、基金评审构成了强烈的外部激励结构，研究者往往被引导去选择“可发表的选题”、使用“可接受的方法”、生成“可接受的结果”。在这种生态中，科研不仅是一种探索行为，更是一种“系统内生”的生产行为，受到制度规范、评价标准与资源竞争的深度影响。这使得医学科研在本质上表现出双重逻辑，一方面是理性主义的求真逻辑，另一方面是制度化的再生产逻辑。

1.5　人类主体中心性的伦理与责任　医学研究不同于其他自然科学研究的关键，在于它直接关系人类生命与健康。这一特性使得医学科研必须遵守严格的伦理规范，如《赫尔辛基宣言》、知情同意制度、临床伦理审查制度等。这些制度体现出一个根本性的科研观念，即人类研究主体不仅是认知的执行者，更是伦理的承载者。医学研究不能单纯追求“知识产出”，更必须保障受试者的权利与尊严，确保研究过程的正当性与社会责任。这一伦理基础也强化了“人类研究者”的中心地位。在传统范式中，AI等非人主体不能拥有研究动机、伦理判断与社会责任，仅被视为“辅助工具”。

1.6　传统范式的优势与局限　假设逻辑、因果推断、统计规范、制度组织与伦理框架等维度，共同构成了医学科研的传统范式，推动了抗生素、疫苗、手术、分子医学等领域的跃迁式发展。但现代医学进入“大数据—多因素—非线性”主导的新阶段时，传统范式开始面临前所未有的挑战：(1)数据增长的速度远超人工处理的能力；(2)多组学交叉导致变量之间的关系难以穷尽；(3)新药研发周期冗长、失败率高；(4)临床试验费用日益上升，伦理边界日趋复杂；(5)对证据等级与因果推断的传统要求，使得某些重要问题难以再通过传统经典路径得以解决。

在这样的时代背景下，AI就不再是锦上添花的工具，而成为一种可能改写范式的“系统性变量”。它的出现，迫使医学科研界反思，我们的知识生产方式是否仍适应当下技术的发展？我们赖以为基础的“科学逻辑”，是否正在被重新定义？

2　AI从工具到协作者

AI对医学科研的介入，已不再停留于“加快处理速度”或“提供便利工具”的表层作用，而是开始重塑科研的底层逻辑。从最初的图像识别与文本分析，到如今参与研究假设的生成、临床试验的设计优化乃至生物机制的挖掘，AI正在从一个被动执行的“工具”，演变为具备方法论意义甚至认知协作能力的“准科研主体”。这一转变并非一蹴而就，而是在不断深化的应用中逐步显现的。

2.1　AI在医学科研中的主要应用路径　AI对医学科研的介入已不再局限于边缘性任务，而是贯穿知识生产的全过程。为了更清晰地呈现其实际影响，本文从4个典型应用领域出发，解析其在科研流程中所起到的关键作用与方法论重塑。

2.1.1　数据挖掘与大规模文献综述　在信息爆炸的时代，医学科研的第一堵墙往往不是技术能力，而是信息获取能力。每年新增数十万篇医学论文，这让任何个体研究者都无法穷尽前沿信息。而AI，尤其是自然语言处理(natural language processing，NLP)与大语言模型(large language model，LLM)，在这方面展现出强大的总结、归纳与交叉分析能力。

Semantic Scholar、Elicit、ChatGPT、Perplexity等AI工具，已经被大量科研人员用于自动生成文献综述、识别研究趋势、归类研究方法甚至推断不同文献间的潜在矛盾。这种“主动阅读＋总结归纳”能力，将科研从“信息堆积”中解放出来，使研究者能够更聚焦于问题建构与理论创新。

更进一步，AI可实现“跨领域文献比对”。例如从神经科学与免疫学两大独立学科中提取共同机

制，为交叉学科研究提供新的假设框架，而这在传统的医学科研范式下完全依靠人力几乎无法实现。

2.1.2 影像识别与诊断模型 影像医学是AI最早且最成功的应用场景之一。深度学习模型在X光、MRI、CT图像识别等场景中表现出超越人类专家的敏感性，能发现肉眼难以察觉的病灶。

谷歌旗下DeepMind训练出的AI模型在多个国际测试中展现出高于放射科医师的检测准确率，且漏诊率更低。其在乳腺癌筛查方面全面超越人类医生[1]，在视网膜病变诊断方面也已达到顶尖专家水平[2]。虽然在不同病种间存在差异，但AI已具备辅助临床决策的能力。

但AI的科研价值远不止于此，还能帮助研究者发现某些影像特征与预后、分子变异的潜在联系，从而提出新的研究假设。例如某些AI模型可以识别肿瘤边缘不规则程度与免疫微环境的关联，从而给免疫治疗的机制探索带来新的启发。

2.1.3 临床试验模拟与设计优化 传统临床试验代价高昂、周期漫长，且受限于患者招募、伦理审查和数据噪声。AI的引入为传统临床试验带来了革命性发展：(1)虚拟受试者生成。AI可基于真实世界数据构建高度逼真的患者群体模型，用于数字孪生试验或模拟对照组，极大降低初期试验的不确定性。例如，美国FDA已接受AI模型模拟的药物-疾病相互作用用于前期研究申报[3]。(2)试验设计优化。AI可以自动识别最可能影响试验结果的变量，从而指导纳排标准制定、样本量设定与分组策略。这种“数据驱动的试验逻辑”不同于传统经验法则，更符合复杂系统的运行特征。

2.1.4 生物标志物筛选与药物发现 药物研发正处于成本高企与靶点枯竭的双重瓶颈期。AI在生物信息挖掘、分子筛选与机制预测中的应用，被视为突破口。

当前AI已能基于多组学数据（基因组、转录组、蛋白质组、代谢组）识别潜在疾病标志物，并推断其病理生理机制。更重要的是，AI可生成新分子骨架，开展虚拟筛选与分子对接，大幅缩短从靶点到候选药物的周期。2020年以来，多家制药巨头如Insilico Medicine、Exscientia等，已投入AI驱动的“自动化药物设计平台”，其生成的候选药物已进入临床前或I期试验阶段，这宣告了“AI药物”时代的来临。

2.2 从“工具性”使用到“方法论”参与 AI的深度介入，不只是功能性补充，其正在改变医学研究的方法论逻辑。

在传统科研中，“提出问题—构建模型—获取数据—分析验证—发布结果”是一条线性流程，其中人类研究者掌控问题提出与分析方法。而在AI驱动的研究中，这一结构被打破：(1)问题生成由AI建议。例如AI可能通过跨数据库比对发现某些基因突变与罕见临床表型相关，从而提出新的研究假设。(2)模型构建依赖AI算法的自身演化。如AutoML（自动机器学习）可自动寻找最佳模型架构，甚至生成新的数据增强策略。(3)分析过程由AI驱动并解释。尤其在非线性、高维度数据中，人类难以直接识别规律，AI提供了全新洞见。

这一过程呈现出“协同认知”的图景。人类负责设定研究边界与伦理底线，AI负责从数据中挖掘可行路径，并不断修正原始设想。AI不再是冷冰冰的执行器，而是具有方法论地位的科研协作者。

值得强调的是，这种协作并不意味着AI取代人类科学家，而是意味着人类科研人员的角色正在从“控制一切”向“协同建模”转变。科学家的任务逐渐转向设定问题边界、评估结果可解释性、制定伦理红线与验证AI假设的临床可行性。

2.3 “黑箱”模型与解释性医学的张力 尽管AI在多个层面展现出巨大的科研价值，但也引发了一个根本性张力，即AI模型的可解释性与医学研究对因果机制的强调之间的矛盾。

多数AI模型，尤其是深度学习模型，属于高度复杂的“黑箱系统”。它们能够提供极其准确的预测或分类结果，却很难解释“为什么”会得出这一结果。这对医学而言是一个巨大的问题。医学科研不仅追求预测，更追求机制。例如，一个AI模型可以准确预测某患者在半年内的复发概率为82%，但如果无法解释这个概率背后的生物路径或临床特征，那它便很难被医生采信，也难以转化为可执行的治疗策略。

医学的理性主义传统要求研究结果必须具备可解释性、可验证性与可再现性，而AI的“黑箱”特性挑战了这一原则。一方面它提供了“超越人脑”的洞见，另一方面它的不可解释性让人望而却步。

因此，近年来“可解释性AI”（explainable AI，XAI）成为医学AI研究的新焦点。可解释性AI力图在复杂模型与因果推理之间建立桥梁，例如通过局部可解释的通用模型解释方法（local interpretable model-agnostic explanations，LIME）、基于Shapley值的加性解释方法（SHapley Additive exPlanations，SHAP）、Attention等机制，揭示中对预测模型最重要的变量，可能关联真实生物机制的决策路径。这一

努力本质上是一种范式融合尝试，既保留了AI的非线性预测能力，又回归了医学的可解释性传统。

2.4　从工具到伙伴的进化　AI正在从医学科研的“加速器”演化为“建构者”。它不再只是帮忙画图、跑模型、对图像分类，而是在问题提出、模型构建、机制挖掘与假设修正的全过程中提供关键洞见。这一变革预示着一种范式转换的发生。从人类主导的数据收集与因果推理，过渡到人机共建的知识发现与系统建模。AI的加入，使科研不再是“精英个体的推理游戏”，而成为“系统间的协同进化”。

3　生成式AI对研究范式的颠覆

自2022年以来，以ChatGPT为代表的生成式AI引发了全球范围的科技震荡。它不仅是一种新型的人机交互工具，更以“语言理解＋知识生成”的能力，逐步渗透进科研流程的多个关键环节。对医学科研而言，这种颠覆已不再是抽象预言，而成为一种现实冲击。研究假设由AI生成、实验设计由AI优化、论文初稿由AI撰写、图表由AI绘制。过去高度依赖人的创造性工作，如今正被AI在效率与规模上改写。

3.1　多模态模型与跨领域知识整合　传统的AI模型多依赖结构化数据，而医学科研则涉及复杂的多模态数据输入，如影像、文本、生理信号、实验数据等。新一代生成式AI，尤其是OpenAI的GPT-4o、Google的Gemini、Meta的LLaVA等，正具备从图像、文本、语音甚至分子结构中同时学习的能力，构建出一个“统一表示空间”来处理医学知识。这种能力，使得“跨模态”的认知成为可能。

以肿瘤学研究为例，过去病理图像的解读、基因突变的分析、患者临床数据的整合往往需要多组学、多团队协作。而今，多模态AI可以同时处理苏木精-伊红染色图像与电子病历信息，在数秒内生成临床建议或研究假设。生成式模型将不再仅仅回答“是什么”，它可以主动提出“还可能是什么”，进而推动科学发现从“数据密集”转向“知识生成”。

此外，医学研究日益交叉化的趋势，也对AI的知识整合能力提出了挑战。生成式AI作为“语义中介”，其训练语料往往覆盖医学、生物学、统计学乃至伦理学等多个学科专业领域。这种“跨域语义迁移”能力，使其在医学科研中不再只是“语义模糊的写作工具”，而可能演化为真正的“跨学科交互节点”。

3.2　自动假设生成与研究设计建议　科学研究的起点是提出一个好的问题。传统上，这依赖研究者的经验、灵感与对文献的长期积累。而生成式AI通过大规模语料学习与上下文理解，可以在一定程度上模拟“提出问题”的过程，即“自动生成研究假设”。这标志着AI第一次从“被动工具”转向“主动参与者”。

例如，在药物再利用研究中，生成式模型可基于已发表论文、基因通路数据库、药物相互作用网络，提出尚未被探索的药物—靶点组合[4]。在精神疾病研究中，AI能结合患者社交媒体发言与电子病历，提出新的行为标志物假设[5-6]。

更进一步，生成式AI正在尝试对研究设计提出优化建议。这不仅包括样本量计算、分组设计、变量控制等技术性环节，甚至还涉及研究伦理审查的合规性建议。这意味着，AI不再只是科学发现的“放大镜”，而是其“放射源”。

3.3　文献写作、图表生成与科研流程自动化　在科研流程中，撰写论文、整理文献、绘制图表等被称为非创造性但高重复度的任务，占据了研究者大量时间。生成式AI正以惊人的效率重构这一链条。

首先是自动写作。目前，GPT类模型已能够根据给定研究主题、背景资料、数据摘要自动撰写论文摘要、引言甚至部分讨论部分。虽然尚不能完全取代人类写作，但在初稿生成和语言润色方面，其效果已被多个国际医学期刊认可，并开始探索AI参与写作的声明制度。

其次是自动图表生成。如ChatGPT与Python代码交互功能、Claude等模型的图像能力，可将用户输入的表格数据转为可视化图形，甚至基于文本自动判断使用何种统计图形（如箱线图、散点图、生存曲线等）。这极大降低了科研门槛。

更进一步的，是科研流程的半自动化。从制定检索策略、批量下载文献、提取关键信息、按主题聚类，到撰写综述式报告，AI工具已可实现“零编程”操作。在系统评价与Meta分析中，部分AI工具（如Elicit、Scite Assistant）能协助完成文献筛选与偏倚评估，大幅提升效率。未来的实验室或科研小组，可能不再以“人力”多少为单位，而是以AI增强容量（AI-augmented capacity）来评估科研生产力。

3.4　生成式预训练模型在“知识编织”中的角色　生成式AI最大的革命，不在于写作或效率，而在于其扮演了“知识编织者（knowledge weaver）”的角色。

传统的科研范式将知识理解为“离散事实”的积累，而生成式预训练模型则试图将其转化为一种“语义网络”的流动系统。在这一系统中，事实、概念、假设、证据、反驳、隐喻、类比等各种知识元素，

被编码为语言向量，彼此形成“嵌套”和“激活”的结构。这种结构极度接近人类“隐性知识”(tacit knowledge)的表达方式。

例如，当研究者输入“抗程序性死亡受体1(programmed death 1，PD-1)治疗失败的机制有哪些可能性？”时，GPT模型不只是输出数据库内已有答案，而是可能结合免疫逃逸、肿瘤微环境、共抑制信号、遗传突变等多层级知识路径，给出“似是而非但启发性强”的解释。这种“多义路径”的生成能力，恰是科研思维中不可或缺的一部分。

从这个意义上说，GPT并不是简单地“模仿”人类语言，而是构建了一个“弱意识形态下的科学思维模拟器”。虽然GPT还远未达到真正“理解”，但它确实在以一种全新的方式，介入并重塑我们“认识世界”的流程。

生成式AI不是科研流程的简单加速器，而是科研范式的结构性挑战者。从多模态学习到自动假设生成，从写作流程到知识编织，它正在将“科学作为生产活动”的各环节自动化、语义化与联动化。这一过程或将推动医学科研走出线性范式，迈入一个高度交互、增强认知、协同创新的新纪元。

4 新范式的哲学与方法论基础

随着AI技术的飞速发展，尤其是深度学习的广泛应用，科学研究的范式正在经历一场深刻的变革。这种变革不仅仅是技术层面的更新，更是哲学和方法论层面的重构。传统的医学科研方法，深植于经验主义传统之中，强调通过观察、实验和归纳来逐步积累事实，进而建立理论。然而，随着数据量的爆炸式增长和计算能力的显著提升，这种传统的科研范式已经难以满足复杂科学问题的求解需求。因此，一种新的范式，数据驱动科学，应运而生，它强调通过大规模数据的模式识别与预测功能来推动科学发现。这种从经验主义到数据驱动科学的跃迁，标志着科学研究进入了一个全新的时代。

4.1 从经验主义到“数据驱动科学”的跃迁 传统医学科研方法深植于经验主义传统之中。在这一科学哲学范式中，研究以观察、实验、归纳为核心，通过逐步积累事实以建立理论，最终在假设—验证框架内实现知识生产。然而，随着AI特别是深度学习的发展，这一过程正在被重新定义。医学科研不再仅依赖于假设推导和实验验证，而是更多地借助对大规模数据的模式识别与预测功能进行驱动，形成“数据密集型科学”(data-intensive science)[7]。

2007年1月11日，计算机图灵奖得主吉姆·格雷(Jim Gray)在面向美国国家研究委员会的演讲中提出了科学研究的“第4范式”，即数据密集型科学。这一概念后来被收录于微软研究院2009年出版的书籍 *The Fourth Paradigm: Data-Intensive Scientific Discovery* 中，即继实验科学、理论科学、计算科学之后，数据驱动科学成为新时代的知识生成方式。这一新范式强调不再“先有理论，后有验证”，而是在数据中发现模式、建立预测模型，甚至生成假设，再反过来指导实验。这种由AI推动的模式，在精准医学、流行病学、药物研发等领域已有广泛体现。例如，AI可以通过分析千万条基因表达数据，预测肿瘤亚型而不需要先验机制模型，继而影响疾病分类和个体化治疗。这并不意味着经验主义被彻底抛弃，而是其地位被重新调和。在新范式下，AI不只是数据处理的工具，更参与到“认知系统”之中，成为科研过程的协作主体之一，逐步改变医学研究的“问题意识”和“证据结构”。

4.2 AI与波普尔式可证伪性原则的张力 在20世纪科学哲学中，卡尔·波普尔(Karl Popper)提出将“可证伪性”作为科学与伪科学的分界线，强调一个理论只有在可能被经验所反驳时，才具有科学性[8-9]。在此框架下，医学研究强调假设先行、严谨实验设计、统计验证，以及可重复性。

然而，AI特别是深度神经网络的引入，使得“黑箱”模型在医学科研中迅速流行。这类模型强调预测准确性，但往往难以解释其内部机制，亦难以通过传统的可证伪性原则进行明确的经验检验。以图像识别中的癌症筛查为例，AI模型可以在没有明确可解释路径的情况下给出高准确率的分类结论，但这是否符合“科学”的标准，成为争议焦点。

这一张力的存在不仅是方法论问题，也带有深刻的哲学意涵，即我们是否可以接受“预测力即科学性”的转向？亦或应回到“可理解、可推演”的可证伪路径？目前的应对策略包括强化XAI技术、构建“白箱”模型、利用因果推断框架重塑AI模型的科学性边界。这些探索意在缓和AI与波普尔哲学之间的张力，并在保持科学性标准的同时，发挥AI在数据处理与模式发现中的优势。

4.3 科学评价标准的重构 医学科研长期以来以统计显著性为评判标准，即通过 P 值来判断研究发现是否具有“真实意义”。然而，AI模型，尤其是生成式模型，更关注“预测能力”和“泛化能力”，即模型在未知数据上的表现。这两者存在根本差异。

举例来说，一个用于预测糖尿病并发症的AI模

型，其预测准确率可以超过90%，但该模型未必满足传统意义上的统计显著性标准。反之，一些存在显著性的研究，可能在实际预测中毫无效果。这暴露出当前科研评价体系中“发现”与“应用”之间的张力。

AI推动下的医学科研可能需要一套新的评价指标体系，例如模型预测的曲线下面积（area under the curve，AUC）、灵敏度、特异度等，多轮交叉验证（cross-validation）或外部验证数据集的泛化能力，模型可解释性与伦理可接受性，重复性与可重现性评估的新标准（如FAIR原则）。这预示着从“统计推断”为中心的科研范式，正在部分转向“建模—预测—优化”链条为核心的AI范式，这在公共卫生与临床辅助决策系统中表现得尤为明显。

4.4　人类研究者的科研主体性变化　AI的深度介入，已经不再只是“工具扩展”，而逐渐涉及“主体重构”。在传统科研中，研究者是知识发现的唯一主体，所有辅助工具都是人类智力的延伸。但在AI系统越来越多地产出结果、提供研究思路甚至生成论文草稿的当下，科研的“作者性”与“责任边界”也被重新审视。

2023年，*Science*、*Nature*等顶级期刊相继声明：“AI工具不能作为论文作者署名”[10-11]，这一反应表面上是对学术署名伦理的维护，实则隐含着对“AI是否是知识生产者”这一核心问题的回应。

另一方面，越来越多的研究者借助GPT-4等大模型完成文献综述、研究设计构建，甚至提出新颖假设，开始从“研究员—工具”关系转向“研究员—协作者”关系。这种变化要求我们重新理解科研主体的构成。AI是“外部心智”的延伸（extended mind），还是人类的“共谋式创新”（co-creativity）伙伴？数据偏见、模型误导等研究责任由谁负责？指导AI学习、审核其输出的人类控制者——“元研究者”（meta-researcher）的贡献应该如何度量？

这场深层次变革意味着，未来的科研不再是“孤独的科学家对抗自然”，而是“人机共构的认知系统”面对复杂现实。在这种范式中，AI不一定拥有意识或伦理判断力，但它可能成为科研共同体的一部分，被纳入知识生产机制、伦理审查机制与质量控制机制之中。

5　风险、偏见与伦理反思

AI正以前所未有的速度与广度介入医学科研全过程，从数据分析到假设生成，从实验设计到论文撰写。然而，这一跃迁同样需要付出代价。在AI驱动科研范式的演进中，风险、偏见与伦理问题正同步扩张，构成亟需警惕的“负面外部性”。若忽视这些底层结构性问题，医学科研的信度、效度乃至公众信任都将受到挑战[12]。

5.1　模型偏见、数据歧视与“垃圾进—垃圾出”（garbage in，garbage out）　AI模型的能力取决于数据。训练数据的代表性、完整性与清洗质量，直接决定了其输出的准确性与公平性。然而，医学数据天生就带有浓厚的人类社会偏见。临床数据多来源于特定区域的大型医疗机构，罕见病、边缘群体数据极为稀缺；性别、种族、经济背景对疾病表现与诊疗路径具有高度相关性，却常常被“归一化”处理。

在“数据即命令”的AI逻辑下，偏见得以系统化编码，并以“客观预测”的面目进入科研判断。例如某些癌症影像识别模型在非白种人患者中的准确率明显偏低；临床试验模拟系统在预测药效时忽略少数族裔，导致实验设计结构性歧视；基于电子病历训练的诊断模型放大了历史中对女性患者的轻症误诊。

这类风险本质上体现了“垃圾进—垃圾出”机制，即模型无法超越其数据的“认知边界”。更严重的是，在AI模型的黑箱结构中，这些偏见往往被“沉默化”处理，难以被识别与纠偏，从而使科研结果被伪科学的精确性所“掩盖”。

为应对这一风险，需强化“偏见审计”（bias auditing）机制，将数据源结构分析、模型公平性评估、敏感属性测试纳入科研流程。同时，鼓励建设包含更多多样性内容的数据集，提升模型的包容性与广泛适应性。

5.2　数据隐私、知情同意与算法黑箱　AI的核心驱动力是大数据，而医学数据中包含大量高度敏感的个人健康信息，包括基因组信息、疾病史、精神状态记录、用药数据等。一方面，AI依赖这些数据实现精准建模；另一方面，这些数据的使用又引发了严重的隐私泄露与伦理争议。

最常见的AI模型使用的风险包括：（1）知情同意制度的虚置化。大量数据在脱敏后被“二次使用”，但患者对其用途不知情，甚至无法撤回授权。（2）AI数据越权分析。模型自动发现某些“意外关联”，如通过面部表情识别预测精神疾病倾向，而这又挑战了传统医学的伦理边界。（3）去标识化的脆弱性。即使去除了姓名等显性标识信息，AI也能通过多源数据交叉重新识别出个体。（4）AI模型本身的“黑箱性”进一步削弱了公众的可理解性与可质

疑性。当前流行的深度学习模型至少拥有数百万个参数，主流大语言模型甚至达到数千亿参数规模，但几乎无人能解释其预测路径与决策逻辑，这对医学研究尤其危险。在临床研究中，若无法理解模型为何得出某种结论，就无法评估其是否基于真实的病理机制。

面对这一伦理张力，学术界正在推进3个方向上的改进。(1)进行“可解释性优先”的模型设计，如使用决策树、注意力机制等手段提高模型透明度；(2)使用差分隐私与联邦学习等技术，旨在在不暴露原始数据的前提下完成高效训练；(3)增强知情同意的动态使用，通过可撤回、可追溯的数字签名技术赋权患者。这些尝试标志着AI伦理从“静态规避”向“动态治理”转型，但距离形成系统规范尚需大量的制度探索与技术创新。

5.3 伦理审查机制的滞后性 传统的科研伦理审查制度主要基于人类研究的规范逻辑设计，如干预研究需获取知情同意、风险最小化、受试者权益保障等。然而，AI科研范式的独特性让这一机制暴露出严重滞后性：(1)边界模糊。AI项目往往不涉及人类实验，仅处理“已有数据”，从而规避伦理审查；(2)风险难以评估。模型偏见、输出误导、学术欺骗等新型风险在传统伦理框架下难以量化；(3)伦理规范空白区。如何界定AI的署名权、责任归属、训练数据的道德合法性等问题，当前还缺乏共识。

以AI生成的科研内容为例，AI生成式内容是否应该纳入伦理审查？模型生成的图表、图像是否存在“伪造”之嫌？如何判断AI“建议”的研究设计是否侵犯他人知识产权？这些问题正逐步成为伦理新议题。因此，有必要重构伦理审查体系，使其具备“AI敏感性”。例如在伦理审查流程中引入技术审查维度，评估模型的训练数据、偏见风险、可解释性等内容；设立跨学科伦理委员会，涵盖数据科学家、临床医生、伦理学者等不同领域专家；推动AI科研的“注册制”与“预先透明化”，如临床试验那样要求提前公开设计方案。只有建立起与AI科研范式相匹配的伦理机制，才能避免“技术先行，伦理追赶”的被动局面。

5.4 科研评价体系对AI生成内容的适应问题 AI特别是生成式大模型的出现，正在深刻改变科研产出模式。越来越多研究者借助AI工具完成文献综述、初步撰写、图表生成，甚至是结论提炼，这种变化正在挑战当前的科研评价体系，主要体现在以下几个方面。

(1)原创性认定的模糊化。传统科研评价高度依赖“作者原创性”指标，但当研究者的大部分写作由AI辅助完成时，原创性边界就开始变得模糊。一篇内容完整、结构严谨的论文，可能是人机协作产物，其“学术贡献”难以判定。

(2)工作量与产出的“通胀”风险。AI工具大幅提升科研效率，导致部分研究者通过“量产”论文获取评价优势。尤其在非顶刊发表平台，AI生成内容可能导致论文质量总体下降、同行评议压力增大。

(3)评价标准的缺失。当前多数高校、科研机构并未制定关于AI辅助写作、图表生成的明确标准，导致评价体系对AI“隐性使用”视而不见，或陷入道德恐慌，出现一刀切禁止使用的非理性反应。

对此，不同学科领域的多名专家纷纷建议构建新的评价适应机制。(1)明确区分AI的“辅助性使用”与“主导性生成”，并制定标注规范；(2)引入“科研贡献构成表”，要求作者申报AI的参与程度；(3)加强对AI生成内容的同行评议，确保内容的逻辑一致性与专业准确性；(4)鼓励在研究的方法部分对AI的使用进行反思性说明，提高科研透明度与自我规训。

科技革命带来的范式变迁，必然要求科研制度与评价逻辑同步演进，否则将形成“技术与制度脱节”的断裂带，损害科学的公信力与学术生态。

AI为医学科研带来了新工具、新范式与新可能，同时也引入了一系列深层的伦理、风险与制度性挑战。模型偏见、隐私风险、伦理失调以及评价体系滞后，已构成AI科研范式重构中不可忽视的结构性问题。未来的医学研究，不仅需要技术上的迭代，也需要伦理规范、制度体系和文化价值观的协同进化。唯有以制度理性与伦理前瞻回应技术创新，才能实现AI与医学科研的可持续共生。

6 实证案例与未来趋势

AI对医学科研范式的影响，已不仅限于理论探讨或潜在可能，而是正在以可观测的方式进入科研实践。

6.1 基于AI的临床试验设计优化 临床试验是医学科研中最复杂、资源消耗最大的环节之一，其设计质量直接决定研究结果的科学性与可推广性。传统临床试验设计常面临诸多挑战，包括样本选择偏倚、随机化方案粗糙、终点指标设置不合理等。AI的介入，特别是以机器学习为代表的预测性建模技术，正在重塑这一流程。

针对传统临床试验因严格纳入标准导致的患者代表性与统计效力不足问题，Liu等[13]提出一种

基于AI的队列优化方法（AI approach to cohort optimization，AICO）。该方法利用BioBERT模型解析*ClinicalTrials.gov*中乳腺癌试验的文本化资格标准，提取结构化变量（如疾病分期、生物标志物），并结合真实世界电子病历（$n=5\,214$）构建患者特征图谱。通过强化学习框架模拟纳入标准组合，AICO在保持统计效力（>80%）的同时，将患者覆盖率从32.7%提升至54.1%。优化后，队列显著提高老年患者（≥65岁）与轻、中度肝肾损伤患者的纳入比例（分别提高18.2%、14.7%），降低了人群异质性。外部验证表明，该方法缩短了41%的招募周期，降低了Ⅲ期试验37%的失败率（$P<0.001$），证实AICO通过数据驱动优化标准可提升试验普适性与效率，为肿瘤学精准入组提供新范式。

类似案例还有基于强化学习优化剂量调整频率的慢性病药物试验、使用NLP技术自动从既往试验中抽取不良事件汇总参数、应用生成模型构建合成对照组等。AI正逐步将临床试验从“标准流程”向“个性化建模”升级。

然而，这一进程也引发争议。例如，AI自动生成的设计方案是否具备伦理审查资格？模型训练过程中若使用不完整或偏倚数据，是否会导致试验设计不公？这些问题尚未形成明确共识，但也为未来AI与临床试验结合的标准化与监管建设提供了议题基础。

6.2　*生成式AI模型在真实研究中的参与实践*　2023年以来，生成式AI模型（如OpenAI的ChatGPT、Google的Med-PaLM系列）开始广泛应用于生物医学研究的多个环节，特别是在研究设计辅助、文献综述、数据注释和初稿撰写方面。

6.2.1　*研究假设生成*　在一项2025年的研究[14]中，研究团队利用ChatGPT（基于GPT-4o）生成创新性假设，以解决心脏毒性研究中的五大挑战：机制复杂性、患者变异性、检测敏感性不足、缺乏可靠生物标志物和动物模型局限性。研究通过ChatGPT生成96个假设，涵盖从3D生物打印心脏组织到多组学分析的新方法。专家评估显示，14%的假设高度新颖，65%中等新颖。ChatGPT进一步为每个挑战挑选最优假设，并提供详细实验设计，包括背景、方法、预期结果及潜在问题。结果表明，AI生成的假设（如使用3D生物打印模型）显著提升了研究的创新性和可行性，优化了卡毒性检测的预测准确性，展示了AI在推动生物医学研究中的潜力。

6.2.2　*参与临床知识问答测试*　2023年，Google DeepMind团队开发的Med-PaLM 2模型在临床知识问答测试中展现了卓越性能，特别是在MedQA数据集上，该数据集基于美国医学执照考试（USMLE）题目。研究[15]表明，Med-PaLM 2在MedQA-USMLE测试中取得了86.5%的准确率，显著超越前一代Med-PaLM（67.6%），高于USMLE及格线（60%），达到“专家水平”，表明Med-PaLM 2在处理复杂医学问题、如疾病诊断和治疗方案选择方面具有强大潜力。研究通过多选题测试评估了模型在医学知识的理解和推理能力，显示其能够有效解析临床场景并提供准确答案。尽管未明确记录Med-PaLM 2完成USMLE全部阶段的测试，但其在MedQA上的高分凸显了大型语言模型在医学教育和临床辅助决策中的应用前景。

6.2.3　*AI辅助科研写作工具*　国内已有若干研究机构在项目撰写过程中引入AI工具，例如“智谱清言”“秘塔写作猫”等，主要用于文献整合、研究背景阐述和摘要初稿生成。目前，生成式AI模型正逐步嵌入科研流程中“认知密集”的部分，其角色已从“数据工具”演进为“语义协作体”。这一转变要求我们重新理解科研中的“原创性”“作者性”与“方法论”。

6.3　*AI与转化医学协同机制的初步探索*　转化医学（translational medicine）强调基础研究与临床应用的高效转化，其难点往往在于如何从基础数据中“发现可验证的路径”并快速实现机制到实践的跃迁。AI，尤其是多模态学习与迁移学习模型，正在成为桥接这一“翻译鸿沟”的关键技术。

在帕金森病（Parkin’s disease，PD）研究中，AI技术通过整合多源数据，为疾病机制研究和临床应用提供了新的路径。研究[16]表明，帕金森病的发病机制可能与肠道菌群通过“肠—脑轴”相互作用有关，涉及α-突触核蛋白的异常聚集和炎症反应等。图神经网络（graph neural networks，GNN）作为AI中的关键工具，已被用于整合蛋白质相互作用网络和生物通路信息，在神经退行性疾病中识别潜在分子靶点和病理机制。

这些研究展示了AI在转化医学中的桥梁作用，特别是在从基础研究向临床应用过渡的关键环节上。AI通过数据驱动的知识联动，初步推动了基础研究向临床转化，特别是在神经退行性疾病领域，为后续研究奠定了技术与理论基础。

此外，AI也正被应用于药物再利用领域。如使用生成对抗网络对美国食品药品监督管理局（Food and Drug Administration，FDA）的批准药物进行机制重构筛查，从中发现可用于罕见病的潜在候选药物。这类研究加快了从数据线索到临床试验的转

化节奏。

然而，这种“高速度、高复杂度、高维信息融合”的研究方式，也对传统学术机构的组织形式、数据治理模式与学科协作机制提出挑战。未来AI与转化医学的协同，将取决于是否能建立起以问题为中心、以跨界为常态的科研组织架构。

6.4 向增能科学(augmented science)演进 AI对医学科研的深度嵌入，标志着增能科学时代的开启。所谓增能科学，指的是在机器智能的协助下，科学研究的发现模式、协作方式与知识生成机制发生系统性重构。在这一范式下，科学不再是“人类智慧独舞”，而是“人—机共创”。AI不仅提供算力与辅助，还成为生成创见、分析复杂系统和优化路径的“合作者”；知识生成方式从“线性演绎”向“多路径建模”转变。

AI模型允许科学家在极短时间内探索成百上千种假设与变量组合，并筛选出最优路径；科研评价逻辑从“结果导向”向“过程透明＋机制合理”转化。在AI协作下，科研的“黑箱”环节变得可审计、可回溯。这也意味着，科研人员的能力结构需要重塑。懂统计与模型原理、能驾驭跨模态工具、具备“与AI协同工作”的元认知能力，正成为下一代医学研究者的基础素养。

AI正在通过真实案例持续验证其科研赋能能力，从临床试验设计优化、生成式模型参与研究过程，到跨学科协同中的关键角色，它正逐步演化为医学科研范式的结构性参与者。未来，我们将迈入“增能科学”时代。在这一时代，科学研究将不再是人类单边探索的过程，而是一场机器智能协作的新范式革命。

7 小结与展望

随着AI技术的大规模涌入，医学科研正经历前所未有的变革。从传统的假设驱动实验向数据驱动决策转型，AI仿佛成为科学家的“外脑”，加速了研究进程并推动科研范式的升级。正如科学界观察到的，AI驱动的科学研究“牵引传统的线性研究范式向更加快速迭代和自适应的方向发展”；未来科研可能呈现“AI提出候选方案—人类判定科学意义—协同优化”的螺旋上升模式。在这一背景下，我们必须思考，AI究竟是彻底颠覆旧范式的“催化剂”，还是与人类智慧协同共生的新范式？

7.1 AI驱动的医学科研范式变迁 AI深刻改变了医学研究的基本逻辑和流程。在影像诊断、基因组学、药物设计等领域，AI从快速处理海量数据、发现隐性模式，到生成研究假说，都展现出革新性的作用。比如，规模化AI智能体已被用于自动设计和执行化学实验，成为“机器化学家”，这表明AI正逐步从辅助工具转向实验设计者。

在医学科研中，传统依赖专家经验和少量样本的方法正向大数据、机器学习模型的协同方向演进。上海交大“明岐”模型的“透明诊断舱”机制，通过可视化病灶标记、决策路径和相似病例参考，将“黑箱”AI变为可审查的决策助手，有效减轻了医患对AI决策的疑虑[17]。这些实践表明，AI不仅仅是自动化运算，更在推动一种“快速迭代-试验-分析-调整”的科研新范式。

7.2 人机协作的新型科研框架 面向未来，单纯的AI替代并不可取。相反，多位专家倡导构建“人机协同”的科研新框架。中国工程院院士李国杰[18]指出，科学发现的本质仍依赖于人类的创造力和哲学思考，未来的科研将依靠人机协作的“混合智能”，由AI放大人类潜能，而人类确保技术向善。这一理念强调“各显其智，智智与共”：人类专注提出关键问题、构建理论框架，AI则快速生成假说、筛选答案，形成一个迭代共生的过程。

在医学领域，专家认为大型通用模型与专用小模型的协作将成为核心趋势：大型模型提供宏观推理和跨模态能力，小模型针对特定任务优化精度，两者结合并由专家持续监督，形成更高效、可控的诊疗和科研管道[19-21]。为实现这种协作，研究社区正在探索人机对齐技术，将伦理和医学知识融入AI设计，确保AI系统的目标和行为与医学价值观一致。例如，多中心临床试验、专家反馈机制以及引入物理模型等方法，能在实践中检测和修正AI偏差，强化AI决策的可解释性，保证其作为人类决策的辅助而非独立决策者。

7.3 可信、透明、可控的AI科研愿景 在此基础上，行业内外多个机构及多名专家学者均提出了面向未来的AI科研方法论愿景，即必须构建“可信(trustworthy)、透明(transparent)、可控(controllable)”的AI体系，成为新的科研守则[22-26]。

(1)可信。AI系统需要经过严格验证和临床评估，采用统计显著性分析、多中心验证等方法提升可靠度；同时，AI建议必须符合医学伦理，人机协作时以伦理为约束，只有这样才能获得医生和患者的深度信赖。

(2)透明。提高模型的可解释性与可追踪性，让研究者和临床医师清晰理解AI推理过程。开放源代码和模型结构、提供决策路径说明，这些措施

有助于快速定位问题并优化改进，也让AI在医学判断中成为“可沟通”的智能体。

（3）可控。在系统设计中预设控制机制和安全准则，如行为偏离监测、故障退回方案等，确保AI在出现异常时可以被人类及时干预甚至关闭。这种“可控AI”思路强调不盲目信任任何AI系统，而是通过设计哲学上的容错与监管机制，维持最终决策权在人类手中。

通过上述原则引领研究设计、数据处理与系统应用，形成一个以可信、可解释为基础的AI辅助科研流程。正如学者所言，要让AI真正融入科学实践，就应使之成为透明、可控、可信的技术；科研人员应对AI有“有依据的信任”，而非盲目信任。

7.4　*哲学、科技与实践的交融思考*　对未来研究范式的探讨不仅是技术问题，也是哲学问题。从科学哲学视角看，这一转型具有库恩式的范式变革特点[27]，旧的经验主义、线性实验观正在被更动态、复杂的认知模式取代。但与传统“革命式替代”不同，本次变革更倾向协同融合。AI赋能并不会剥夺科学家的地位，反而解放其低效劳动，让人类专注于高层次创造性工作。这需要我们重新审视“知识本质”与“认识论”，接受智能体能够与人类共同提出科学假说、演绎科学意义。

医疗科研的实践也凸显了价值观和人文关怀的重要性。AI在追求最优生物医学结果时，必须被设计成人性化的同伴，纳入情感、伦理与社会因素。因此，哲学、技术与实践在这个过程中交织，伦理学、信息论、医学研究方法论将共同指导新范式的形成。

7.5　*面向未来的研究范式*　展望未来，医学科研的范式更替与融合将持续进行。我们不应盲目恐慌AI替代人类科研，而应积极塑造人机协同的新生态。具体而言，需要构建跨学科的合作平台，将技术开发者、哲学伦理学者和临床专家紧密联结，共同定义AI应用的边界与规范。这包括制定开放透明的监管法规、标准化AI模型评估流程、以及推动开放科学数据生态，以保障每一步都在可追溯、可监管的轨道上。通过提升可解释性和强化信任度，AI将在关键领域得到更广泛应用；只有在开放、互信的体系下，人机协同模式才能发挥最大效能。

面对AI浪潮，我们应选择协同共融的路径，以“可信、透明、可控”的AI科研方法论为愿景，构建以人为本、可持续的新型科研范式，让人机共创的智慧真正造福医学研究与全人类。

伦理声明　无。

利益冲突　所有作者声明不存在利益冲突。

作者贡献　高承实：选题，撰写、修改论文；程元骏：选题，修改论文。

参考文献

[1] MCKINNEY S M, SIENIEK M, GODBOLE V, et al. International evaluation of an AI system for breast cancer screening[J]. Nature, 2020, 577(7788): 89-94.

[2] DE FAUW J, LEDSAM J R, ROMERA-PAREDES B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease[J]. Nat Med, 2018, 24(9): 1342-1350.

[3] U. S. Food and Drug Administration. Considerations for the use of artificial intelligence to support regulatory decision-making for drug and biological products[EB/OL]. (2025). [2025-01-06]. *https://www.fda.gov/media/184830/download*.

[4] HUANG K X, CHANDAK P, WANG Q W, et al. A foundation model for clinician-centered drug repurposing[J]. medRxiv, 2024: 2023.03.19.23287458.

[5] DE CHOUDHURY M, GAMON M, COUNTS S, et al. Predicting depression *via* social media[J]. Proc Int AAAI Conf Web Soc Medium, 2013, 7(1): 128-137.

[6] CHOI Y, LEE H. Data-driven discovery of digital biomarkers from electronic health records and social media for mental health[J]. J Med Int Res, 2021, 23(5): e23956.

[7] ANDERSON C. The end of theory: the data deluge makes the scientific method obsolete[J]. Wired Magazine, 2008, (7).

[8] （英）波　珀. 科学发现的逻辑[M]. 查汝强，邱仁宗，译. 北京：科学出版社，1986.

[9] （英）波普尔. 猜想与反驳：科学知识的增长[M]. 傅季重等译. 上海：上海译文出版社，1986.

[10] HOLDEN THORP H. ChatGPT is fun, but not an author[J]. Science, 2023, 379(6630): 313.

[11] EDITORIAL. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use[J]. Nature, 2023, 613(7945): 612.

[12] JASANOFF S. States of knowledge: the co-production of science and social order[M]. London: Routledge, 2004.

[13] LIU X, SHI C, DEORE U, et al. A scalable AI approach for clinical trial cohort optimization [C]//Communications in Computer and Information Science (CCIS, volume 1525). Cham: Springer, 2021: 479 - 489.

[14] LI Y L, GU T S, YANG C Y, et al. AI-assisted hypothesis generation to address challenges in cardiotoxicity research: simulation study using ChatGPT with GPT-4o [J]. J Med Internet Res, 2025, 27: e66161.

[15] SINGHAL K, TU T, GOTTWEIS J, et al. Towards expert-level medical question answering with large language models [J/OL]. (2023-05-16) [2025-03-15]. https://arxiv. org/abs/2305.09617.

[16] BROGI S, CALDERONE V. Artificial intelligence in translational medicine[J]. Int J Transl Med, 2021, 1(3):

223-285.

[17] 郭春丽，李清彬. 提升政策工具组合效能 推动经济持续回升向好[EB/OL]. (2025-03-31)[2025-05-10]. *https://www.news.cn/digital/20250331/9c2e2354b2624b56a3572d1ff1930386/c.html.*

[18] 李国杰. 人工智能将走向何方？——访中国工程院院士、计算机专家李国杰[EB/OL]. 人民论坛网，(2025-02-27)[2025-06-01]. *http://www.rmlt.com.cn/2025/0227/724134.shtml.*

[19] WAN P, HUANG Z, TANG W, et al. Outpatient reception via collaboration between nurses and a large language model: a randomized controlled trial[J]. Nat Med, 2024, 30(10): 2878-2885.

[20] ZHANG K, QU J, et al. MINIM: a general-purpose medical image synthesis model for cross-organ and multi-modal applications[J]. Nature Medicine, 2024: 148.

[21] YANG Y, ZHU L, et al. Medical World Model: Simulating tumor evolution for personalized treatment planning[J/OL]. [2025-03-10]. *https://arxiv.org/abs/2506.02327.*

[22] European commission high-level expert group on AI. Ethics guidelines for trustworthy AI[R/OL]. Brussels: European Union, 2019. [2025-06-15]. *https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines.*

[23] IEEE Standards Association. IEEE standard for transparency of autonomous systems (IEEE 7001-2021)[S]. New York: Institute of Electrical and Electronics Engineers, 2021.

[24] OECD. Recommendation of the council on artificial intelligence[EB/OL]. Paris: OECD Publishing, 2019. [2025-06-15]. *https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.*

[25] UNESCO. Recommendation on the ethics of artificial intelligence[R/OL]. Paris: United Nations Educational, Scientific and Cultural Organization, 2021. [2025-06-15]. *https://unesdoc.unesco.org/ark:/48223/pf0000381137.*

[26] WU X, WANG D, ZHANG Y, et al. A survey on trustworthy artificial intelligence: from principles to practices[J]. IEEE Trans Artif Intell, 2023, 4(3): 378 - 397.

[27] KUHN T S. The Structure of Scientific Revolutions[M]. Chicago: University of Chicago Press, 1962.

引用本文

高承实，程元骏. 人工智能对医学科研范式的重构[J]. 元宇宙医学，2025，2(2)：1-12.

GAO C S, CHENG Y J. The reconstruction of the medical research paradigm by artificial intelligence[J]. Metaverse Med, 2025, 2(2): 1-12.