信息生态系统视角下生成式人工智能辅助临床决策的人机协同研究进展及展望

杨 娇* 郭 蕊
首都医科大学公共卫生学院 北京 100069

【摘 要】生成式人工智能(Generative Artificial Intelligence,GAI)正加速融入医疗服务体系,并正在重塑医疗信息生态系统。本文基于信息生态理论,从信息环境、信息流与信息人三个层面系统梳理GAI辅助医生临床决策的人机协同研究进展。分析发现,在信息环境层面,GAI技术能力快速演化,但医疗准入政策相对滞后;在信息流层面,现有评估范式难以反映真实临床场景中诊断推理的动态性与逻辑性;在信息人层面,人机协作中医生的信息权衡等关键机制难以明晰,协同效果存在较高不确定性。基于此,本文提出三方面研究展望:一是构建国家级评估与监管框架,并持续开展场景化测评;二是建立诊断全流程动态评估与产品全生命周期的质量监测体系;三是深入开展医生与GAI的人机协同互动与认知过程研究。

【关键词】生成式人工智能; 大语言模型; 临床决策支持
中图分类号:R197 文献标识码:A doi:10.3969/j.issn.1674-2982.2025.12.006

Generative Artificial Intelligence for clinical decision-making: Progress and prospects of Human-AI collaboration from an information ecology perspective

YANG Jiao, GUO Rui
School of Public Health, Capital Medical University, Beijing 100069, China

【**Abstract**】Generative Artificial Intelligence (GAI) is rapidly integrating into healthcare service systems and is reshaping the medical information ecosystem. Based on information ecology theory, this study systematically reviews the research on human-AI collaboration in GAI-assisted clinical decision-making across three dimensions: the information environment, information flows, and information actors. Our analysis shows that, at the level of the information environment, GAI's capabilities are advancing rapidly, while regulatory and admission policies remain lagging. At the level of information flows, existing evaluation paradigms struggle to capture the dynamic and logical nature of diagnostic reasoning in real clinical settings. At the level of information actors, key mechanisms such as physicians' information weighing in human-AI collaboration remain unclear, leading to considerable uncertainty in collaborative performance. Based on these findings, the study proposes three research directions: (1) establishing a national evaluation and regulatory framework with ongoing vignette-based assessments; (2) developing a dynamic evaluation system covering the entire diagnostic process and quality monitoring across the GAI product lifecycle; and (3) conducting in-depth studies on the human-AI interaction and cognitive processes between physicians and GAI.

【**Key words**】Generative Artificial Intelligence; Large language model; Clinical decision support

1 引言

以生成式人工智能(Generative Artificial Intelligence,GAI)为代表的新一轮智能技术革命,正在深刻改变知识生产与决策支持的模式。人工智能正从以识别和判断为主的"感知智能"向具备复杂推

* 基金项目:国家自然科学基金面上项目(72574152);首都卫生管理与政策研究基地开放性课题(2025JD01)
作者简介:杨娇(2001年—),女,硕士研究生,主要研究方向为数智健康。E-mail:Yangjiao@mail.ccmu.edu.cn
通讯作者:郭蕊。E-mail:guorui@ccmu.edu.cn

理能力的“认知智能”纵深发展，并在医疗等知识密集型领域加速重塑信息流动结构与决策逻辑。

2025 年初，DeepSeek 的发布标志着 GAI 进入规模化应用阶段，国内多家医院相继部署并尝试其本地化接入。截至 2025 年 3 月，已有 300 余家医院完成部署[1]，医疗场景中的生成式模型实践显著加速，推动了医疗场景中“通用智能”到“专业智能”的跃迁。然而，技术突破并不意味着临床实践的质量保障。2024 年世界卫生组织发布的《生成式人工智能医疗应用伦理指南》指出，GAI 可能生成虚假、不准确或带有偏见的医学信息，其在临床决策中的使用潜藏算法安全、数据偏倚等技术风险与人机协作风险。

相比于传统的判别式人工智能，GAI 拥有自主生成、语义理解、推理链条构造与交互式对话等能力，使其从“辅助工具”转变为直接介入临床知识生产与认知活动的智能参与者。因此，GAI 融入医疗系统不再是技术模块的单点输入，而是对医疗信息生成、流动与解释的全面介入。依赖单一性能评估或个别应用验证，难以解释其在真实世界运行过程中的复杂性。

基于此，有必要从更系统的角度审视 GAI 在医疗领域的使用现状、风险与互动机制，尤其是其与医生之间正在形成的新的协同关系。本文引入信息生态理论作为理论分析框架，旨在从信息环境、信息流质量与信息人互动三个方面，梳理并分析目前相关文献的研究进展，以明晰 GAI 融入医疗实践所面临的挑战和风险，为其安全与可持续应用提供理论基础。

2 资料与方法

2.1 文献检索策略

本研究在 PubMed、Web of Science、中国知网和万方数据库进行文献检索，检索表达式详见表 1。英文文献检索截至 2025 年 4 月 14 日，中文文献检索截至 2025 年 12 月 4 日，共检索文献 5 152 条。

表 1　相关检索表达式

数据库	检索表达式
中国知网	SU=(“大语言模型”+“生成式人工智能”)*(“医疗”+“诊断”+“诊疗”+“临床推理”+“医疗决策”+“临床决策支持”)
万方数据库	主题:(“生成式人工智能” OR“大语言模型”)AND 主题:(“医疗” OR“诊断”OR“诊疗”OR“临床推理”OR“医疗决策”OR“临床决策支持”)
Web of Science	TS= "generative artificial intelligence" OR "large language model" OR "Artificial Intelligence Generated Content" OR "Artificial general intelligence" AND TS="medical" OR "diagnosis" OR "clinical decision making" OR "clinical decision reasoning" OR "clinical decision support"
PubMed	"generative artificial intelligence[Title/Abstract] " OR "large language model[Title/Abstract]" OR "Artificial Intelligence Generated Content[Title/Abstract]" OR "Artificial general intelligence[Title/Abstract] " AND "medical[Title/Abstract]" OR "diagnosis[Title/Abstract]" OR "clinical decision making[Title/Abstract]" OR "clinical decision reasoning[Title/Abstract]" OR "clinical decision support[Title/Abstract]"

2.2 文献纳入排除标准

文献的纳入标准为：(1)研究主题涉及 GAI 在医疗诊断、临床决策或临床推理中的应用；(2)研究对象包含医生或临床实践情境；(3)对于评论或综述类文献，若基于原始数据进行了新的分类、统计或分析，可纳入作为补充性数据来源。排除标准为：(1)文献类型为纯评论、观点文章、新闻、书评、会议摘要、信件等非研究性文章，且未进行二次分析；(2)无法获取全文或全文信息不足以支持分析；(3)重复文献。

2.3 筛选与分析方法

对检索到的文献关键词、标题和摘要进行阅读，排除明显不符合主题的文献，对于不能确定是否符合主题的文献，进一步阅读全文，最后筛选文献 25 篇。在此基础上，依据信息生态理论的三个分析层面(信息环境、信息流、信息人)，选取能够反映相关关键机制或典型问题的代表性的关键文献进行分析，梳理医疗实践中医生与 GAI 人机协作的问题与挑战。

3 分析框架：生态信息理论

信息生态理论源于信息科学与系统理论的交叉融合，强调在动态环境中，信息人、信息、信息技术与信息环境之间的持续互动与反馈。[2]与传统以技术性能为核心的 AI 评估范式相比，信息生态理论提供了一个更具系统性与动态性的分析视角，可揭示 GAI 融入医疗实践后引发的整体性变化。

该理论框架主要包括三个核心层面(图1):

(1)信息环境层面:指医疗信息系统运行的宏观环境,包括政策导向、技术基础与组织条件。GAI的引入不仅依赖于算力、算法与数据资源等技术要素,也受到宏观政策与医院内部需求的共同影响。技术创新、政策约束与医院需求共同塑造了GAI应用的场景与发展路径。

(2)信息流层面:指信息的生成、流通与解释过程。GAI的生成式特征重塑了医疗知识的生产与传播逻辑,使诊断推理过程从以医生经验为中心转向人机交互的信息流模式。然而,这一过程也带来了信息流的可靠性、可溯源性与准确性的新型挑战。

(3)信息人层面:指医生与GAI在诊疗活动中的角色定位与协同方式。GAI不再只是被动工具,医生与AI之间的互动逐步从工具性使用走向认知层协作。

综上,信息生态理论为分析"医疗+GAI"的结构性问题提供了一个系统化框架。它能够同时解释GAI落地过程中受到政策、组织与技术条件的限制(信息环境),揭示医疗场景中信息质量风险的形成机制(信息流),并刻画医生(信息人)与GAI在协同决策中的互动行为。基于此,本文将从信息环境、信息流与信息人三个方面展开分析,为理解GAI医疗应用的挑战及其治理路径提供理论支撑。

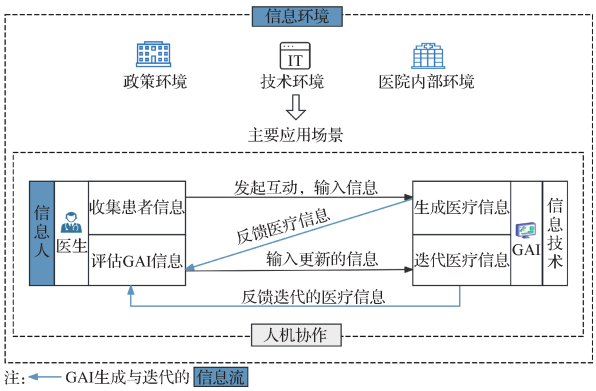

图1 医生与GAI协作的医疗信息生态系统

4 信息环境视角下GAI医疗应用的现状与挑战

信息环境是GAI与医生行为发生的一切场景和条件的要素集合,包括技术环境、政策环境和医院内部环境,三重驱动共同影响GAI在医疗实践中的落地。

在技术环境层面,技术环境决定了可实现的功能边界。"生成式"意味着GAI可以学习和模拟事物内在规律,并通过上下文学习,根据用户需求自主创造出新的内容,而不仅仅是模式识别或分析。实现了从"狭义人工智能"转向"通用人工智能"[3],能够适应更多新任务[4],承担更复杂的临床推理任务,满足更多样化的临床需求。因此,相比于传统人工智能,更适合部署在结构化强、需要具备一定推理能力的环节。然而,"生成式"本身具有不确定性,其输出的稳定性、逻辑一致性以及与临床语境适应性仍存在限制,这也构成其在医疗场景使用时的关键技术风险。

在政策环境层面,政策环境提供了试点空间并设定风险边界。一方面,国家明确将人工智能作为推动医疗服务模式创新的关键技术,积极引导其在临床领域的应用。鼓励GAI在分诊、预问诊、辅助诊疗等领域开展应用试点。另一方面,将安全性、可解释性与责任边界等作为应用的基本要求,强调完善算法风险治理、伦理安全等机制。在"推动应用"与"强化治理"并行的政策逻辑下,不断引领推动GAI在医疗领域应用场景落地。但是,目前仍缺乏针对GAI作为医疗产品的准入标准,使得制度对模型能力的约束存在一定模糊空间。

在医院内部环境层面,医院内部需求决定了GAI应用场景落地的优先次序。医院主要基于自身业务需求和流程特征,推动GAI系统的本地化部署与定制开发。面对诊疗效率与诊疗质量的双重压力,医院倾向优先部署在保证诊疗质量的前提下,能够显著提升诊疗效率、减轻医生负担的功能模块。

技术进步提供了能力边界,政策塑造了试点与合规框架,医院需求决定了优先落地的功能点。三者交互催生了当前医院中最常见的三类GAI典型应用场景:预问诊、分诊与辅助诊疗。首先,GAI的上下文理解与生成能力能够支撑病史采集、病情初筛与诊疗决策等高信息密度任务,从技术上具备可行性;其次,国家政策明确鼓励在预问诊、分诊与辅助诊疗等领域开展试点,为其提供制度空间;再次,医院在提升诊疗效率与质量方面具有迫切需求,使这些场景成为最优先部署的环节。因此,这些场景集中体现了技术环境、政策环境与医院内部环境三重因素的共同作用。这三类应用既是GAI展现其强大信息处理与交互能力的主要场景,也是其技术风险最突出的节点(图2)。

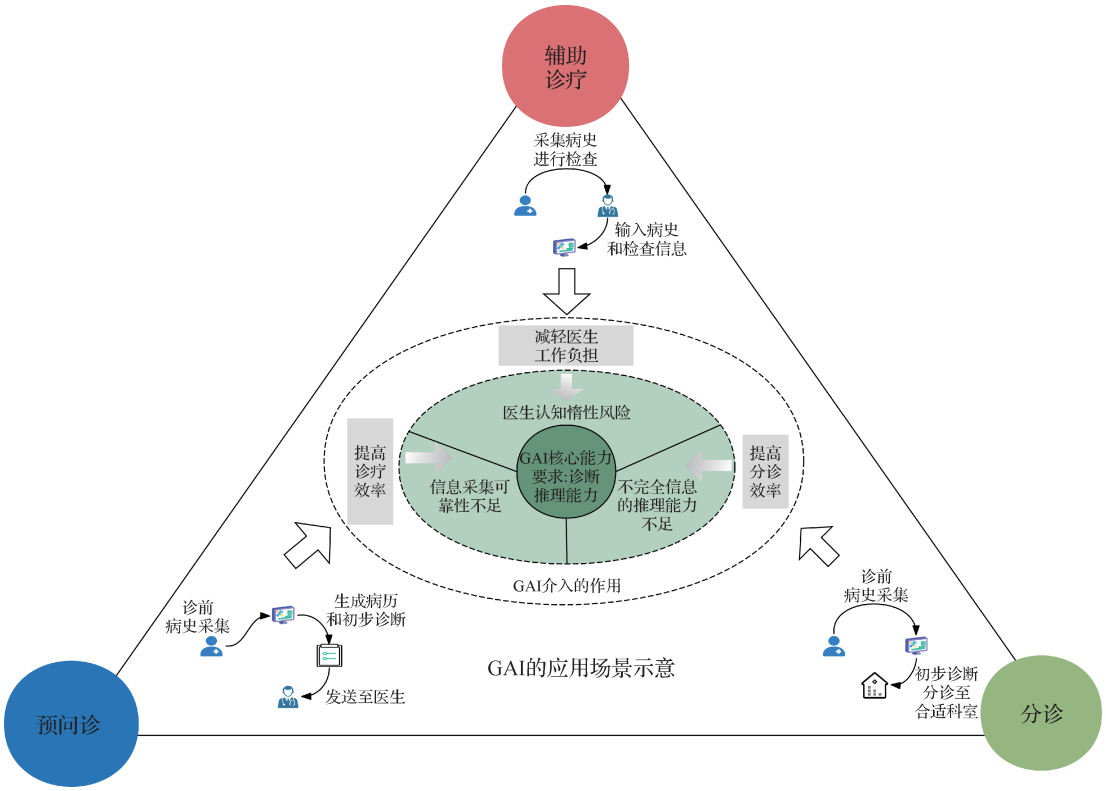

图 2　医生与 GAI 人机协作三大主要应用场景

4.1 预问诊：输入噪声与“同理心”导致信息采集可靠性不足

病史采集能力高度依赖患者的信息输入质量，患者作为非专业人士，往往难以准确、全面地描述自身症状，这导致 GAI 的输入源存在天然的“噪声”。这也直接影响后续医生与 GAI 协同的基础信息质量。同时，不同科室之间病史采集的侧重点可能存在一定差异。而医生不仅根据患者的生理状况和症状进行判断，还考虑心理、情感和社会因素对健康的影响。这些因素在医学问诊中至关重要，但对于 AI 技术来说相对复杂和难以量化。医生需要花费一定时间与对 GAI 采集的病史进行筛选与纠正，从而影响协同效率。

4.2 分诊：训练数据不足，基于不完全信息的决策存在不稳定性

分诊是前瞻性的、基于不完全信息的开放任务，需要预测病情动态发展的可能性，这对模型的推理能力提出了更高要求。[5]首先，其训练数据多来源于已确诊的病例报告或学术文献，缺乏大量关于“初筛—鉴别—最终分诊”的动态、连续的真实世界训练数据，导致其“学习”不足。最后，分诊存在风险权衡的复杂性。分诊是医疗安全（避免分诊不足）与运营效率（避免过度分诊）的权衡，GAI 难以内化并理解这种涉及医疗伦理、资源分配和潜在法律后果的复杂决策逻辑。因此，医生在使用 GAI 分诊建议时，需要在自身专业判断与 GAI 建议之间不断验证，从而增加医生的认知负荷。

4.3 辅助诊疗：医生认知惰性风险

在辅助诊疗方面，GAI 则能够综合考虑患者病史、家族史、检查信息等，生成诊断列表与后续检查建议，帮助医生快速进行疾病筛查与诊断。[6]同时，还可基于疾病诊疗指南，提供治疗方案与建议[7]，减轻医生的工作负担。然而，GAI 的高效辅助是一把“双刃剑”，极易诱发医生的认知惰性与自动化偏见——即不假思索地接受 GAI 的输出结果，缺乏批判性运用，可能引发严重的临床决策错误。

上述应用场景中暴露的风险均指向同一核心问题，即 GAI 的诊断推理能力，GAI 所生成的信息流质量，在可靠性和稳定性方面仍存在一定问题。要进一步理解其风险结构，需要系统审视现有研究如何评估 GAI 的诊断推理能力。

5 信息流的质量危机：现有GAI评估范式存在显著局限，难以推动GAI的发展

临床诊断是一个复杂的动态过程，需要具备较强的诊断推理能力，即需要综合分析患者的主诉、体格检查、实验室及影像学信息，并做出最终诊断。其强调诊断过程的整体性、合理性以及诊断的正确性，被视为诊断支持研究的核心标准。但在对GAI的诊断推理能力评估研究中，现有方法存在明显局限，未能反映真实的临床实践。目前仍未发展出GAI质量评估的“金标准”，使得其信息流质量存疑，实际应用存在难以预估的风险。

5.1 静态评估难以体现真实临床环境，GAI的信息整合能力存疑

相比于传统AI，GAI的优势之一在于其互动特性，能够根据医生反馈的信息进行实时信息更新与反馈。但是当前GAI的评估逻辑仍延续了传统AI的“知识应答”模式，将临床诊断这一动态、迭代的认知过程，简化为了静态、单向的知识问答。且简单地采用多项选择题和较为简单的开放式题目进行测量[8-10]，GAI虽能够表现出较好的应试能力，但这更反映其知识储备，而非信息整合的诊断推理能力。即使在情境化测试中，研究者一次性输入完整病例信息，仍无法重现医生在真实诊疗中与信息的交互关系，从而高估了模型的诊断表现。[11-12]

5.2 GAI的诊断推理过程存在明显黑箱性质，推理过程评估缺失

相比于传统AI，GAI的算法黑箱性质更为明显，因果推断机制薄弱。而诊断是一个环环相扣的完整链条，而现有评估多聚焦于“鉴别诊断—检查—诊断—治疗”中某一单一环节的表现（表2），且评价指标呈现结果导向，进一步放大了这一“黑箱效应”。在这种结果导向体系下，推理过程的因果逻辑被压缩，诊断质量难以被验证。科学地评估其推理能力不仅有助于揭示潜在风险，也为解决模型可解释性不足提供了重要路径。已有的少量推理过程性评估研究表明[13-14]，GAI具备一定的类人推理潜力，但其仍存在一定局限性。这意味着，GAI虽能生成表面合理的诊断结论，却未必具备稳定的推理加工机制。换言之，信息流的“形式正确”并不等于“逻辑真实”，而这正是现有评估范式未能解决的核心困境。

总之，现有评估范式揭示了GAI的诊断推理能力在动态、复杂的真实临床环境中远未成熟，其生成的信息流质量既缺乏真实性验证，也缺乏稳定性的保障。而GAI已从技术验证阶段逐渐发展至临床应用探索阶段，这意味着医生在协作过程中必须面对一个关键的不确定性：当GAI的信息流质量尚未定性时，医生作为核心的“信息人”将如何理解、利用或质疑这类信息？二者之间的协同决策效果究竟如何？

表2 GAI诊断推理能力评价有关文献

维度	评价指标	评价对象	科室	结论
鉴别诊断	准确性[15]	GPT-4.0	—	ChatGPT-4.0在74.6%的案例中准确地创建了一个鉴别诊断名单。诊断中正确的案例中，70.2%在首位就提供了正确的诊断，23.4%在第二位提供了正确的诊断，6.4%在第三位提供了正确的诊断。
检查	检查结果解读的准确性[16]	Meditron、Clinical Camel、Llama 2 Chat、OASST、WizardLM	内科	GAI表现相当糟糕，正确率低。
	完整性[16]	Meditron、Clinical Camel、Llama 2 Chat、OASST、WizardLM	内科	GAI偶尔不要求进行影像学检查。
	医学图像描述准确性[17]	Qwen2-VL、Qwen2-VL(FT)、Qwen2-VL(FT) + VisRAG	胸外科	未经微调的Qwen2-VL在图像描述中主要为基础性表述，缺乏对关键病灶的明确诊断提示。而Qwen2-VL(FT)在经过LoRA微调后能更准确地使用医学术语并识别异常表现。基于VisRAG引入外部诊断报告后，Qwen2-VL(FT)能够生成更为全面和精准的描述。
诊断	完整性[18]	GPT-3.5和GPT-4.0	内科、急诊科	GPT-4.0和GPT-3.5，在完整性方面都获得了高分，分别为4.3和3.7。
	一致性[19]	GPT-4.0	口腔科	对60道问题答案的一致性较高（85.4%），但仍有改进空间。
	相关性[20]	ChatGPT	骨科	相关性得分（4.43分）高于平均水平（3分）。
	准确性[21]	Med-MLLM、GPT-2、GPT-3、ChatGPT、GPT-4	感染科	Med-MLLM（97%）表现出了与GPT-4.0（98%）相当的准确率，并高于GPT-2.0（87%）、GPT-3.0（91%）和ChatGPT（93%）。

表2　GAI诊断推理能力评价有关文献（续）

维度	评价指标	评价对象	科室	结论
治疗	治疗方案的合理性[22]	ChatGPT	全科（9种常见疾病，包括传染病和非传染病）	85.2%（23/27）出现了不必要或有害的药物建议。即使在正确诊断的情况下，ChatGPT 推荐此类药物的概率为 59.3%（16/27）。
	治疗方案的正确性[23]	LangChain-LLM、GPT-4.0	—	LC、GPT-4.0 的治疗方案在召回率和 F1 分数上均优于临床医生（P<0.05），且 LC 和 GPT-4.0 之间差异不大。
诊断推理全流程	结构化反思工具得分百分比[24]	GPT-4.0	内科、全科和急诊科	与使用搜索引擎的医生相比，医生使用 GPT4.0 并不能提高对具有挑战性的临床病例的诊断推理能力。
	R-IDEA 总分[13]	GPT-4.0	内科	使用由 R-IDEA 分数衡量的临床推理方面，GPT4.0 的表现优于医生，GPT4.0 的错误临床推理案例多于住院医师。

6 信息人的互动困境：医生-GAI协同决策的未知性

在生成式人工智能融入诊疗过程的背景下，医生与GAI之间的互动逐渐成为临床决策的重要组成部分。GAI不再是传统意义上的工具型系统，而是能够生成诊断相关信息、提出鉴别诊断并构建推理路径的“认知参与者”。这种身份变化使得医生的决策行为不再是对工具的单向调用，而是与GAI形成信息交换和认知协作的动态过程。在信息流质量不稳定背景下，医生与GAI的互动成为关键问题。然而，与技术层面的快速演化相比，关于医生如何处理、整合并利用 GAI 所生成信息的研究明显滞后。使协同决策的效能机制、风险路径及其对临床质量的影响仍缺乏系统性认识，进而影响了 GAI 医疗应用的规范发展。

6.1 缺乏医生与GAI的人机协作过程性分析

目前，国内研究仍停留在GAI信息流质量的探究层面，缺乏医生与GAI协作的实证研究。相比之下，国际研究已进入人机协作效果的实证探索阶段，以医生与GAI协作诊断的正确性作为协同效能的关键标尺，但得出的结论呈现高度异质性。[14, 25-26]除了可能受到GAI技术的影响[27-30]，从信息生态学的视角来看，还可能受到医生与GAI交互行为的影响。这种结果导向的研究范式忽视了协同决策的动态演变过程，使得医生与GAI之间的信息互动机制如同“黑箱”一般难以窥见全貌。在真实的临床生态中，医生作为信息处理的主体，其决策过程是一个高度动态且复杂的认知博弈过程。医生首先会基于自身经验形成初步诊断假设，随后在与GAI的交互中不断进行信息校验、质疑与修正。这一微观层面的信息处理流程，即信息如何在人类认知与算法逻辑之间被权衡、整合与迭代，是理解协同效能的关键所在。然而，现有研究过于聚焦群体层面的诊断正确性对比，忽视了个体层面决策行为的动态演变及与GAI之间复杂的信息博弈过程，无法合理解释为何在同一GAI辅助下，不同医生的决策效能存在显著差异。这使得我们既难以优化协同策略，也无法有效防范因盲目信任而导致的错误信息在生态中扩散的风险。

因此，还应关注医生在GAI辅助前后，对其决策行为的影响。已有研究显示，GAI 介入确实能够提升医生的决策表现。[26]例如，一项针对 GAI 分诊辅助的研究发现，医生往往愿意根据 GAI 建议调整初始判断，最终分诊正确率显著提升。但该研究并未分析“GAI是否正确”这一关键变量如何在协同过程中影响医生的决策路径。相比之下，针对传统AI的研究已尝试将传统AI的正确性纳入分析，并对医生的人机协作行为进行分类（表3），这一类研究数量虽有限，但也为理解医生与GAI的多样化互动方式提供了分析框架。

表3　AI辅助诊断对医生决策行为改变影响情况的有关文献

作者（年份）	初始诊断		AI诊断		最终诊断	
	正确	错误	正确	错误	正确	错误
WANG D Y（2023）[28]		√	√		√	
	√			√		√
WIES C（2024）[31]	√		√		√	
	√			√	√	
		√		√	√	
	√		√			√
ROSENBACKE R（2024）[32]		√	√			√
		√		√		√

6.2 人机认知差异是协同决策的底层挑战

医生的决策行为往往取决其自身知识经验与GAI信息的推理博弈过程，而诊断推理本质上是人类的认知能力在医学诊断决策场景下的一种应用。但

由于GAI与医生的认知存在本质上的差异，医生的决策依赖感知线索和环境上下文，AI则依赖数据与相关性分析，这种认知差异可能是导致协同决策中产生误解与风险的主要原因。二者在信息加工方式上的不一致，使得医生在采纳AI建议时需要对其可信度进行动态校准，而这种信任过程本身会受到任务复杂度、信息冲突与时间压力等多重因素的影响。已有研究提出了坚持自身判断、依赖AI建议与并行推理下的迭代整合等典型互动模式，为理解人机协作的认知特征提供了初步框架。[33-34]但目前研究仍主要集中在信任变量，缺乏能够整合认知行为、情境因素与信息特征的综合性解释框架。

7 研究展望

7.1 信息环境：构建国家级评估与监管框架

当前的信息环境面临评估标准缺失等核心挑战，亟需构建一个系统性的国家行动框架。首要任务是推动国家层面出台专项政策，并构建以高质量、标准化评测数据集为核心支撑的新型评估体系。政策应聚焦三个方面：一是应明确GAI在医疗场景中的合规使用，明确其作为医疗器械的监管属性和准入标准。制定GAI在医疗中的风险分级策略、产品界定规则和注册审评标准。二是启动国家级基准数据集建设工程，形成符合真实诊疗场景的高质量评测数据、任务集与指标体系。该体系应超越传统上仅关注最终输出结果的单一模式，转向一个涵盖包括推理过程准确性、诊断结论正确性等维度在内的过程性评估与结局性评估有机统一的综合评估模式。三是持续开展场景化测评，既检验模型的诊断结论，也量化其推理链条的合理性，使评估真正贴近实际临床的诊断过程。

7.2 信息流：建立诊断全流程动态评估与产品全生命周期的质量监测体系

为确保GAI在真实临床情景下所产生信息流的可靠性与安全性，首要任务是革新当前静态化的评估范式，动态评估GAI的推理能力。需从诊断全流程出发开发测评工具，量化GAI随着医生不断更新输入信息时的推理迭代能力与持续性表现。此外，信息生态并非静态不变，而是处于动态演化之中。GAI的持续学习和进化，以及新的医疗技术的不断涌现，都将不断重塑医疗信息生态。因此，我们需要建立一套完善的监测和评估体系，加强国家人工智能应用基地建设，实现覆盖产品全生命周期的连续性测评，及时发现和解决信息生态中出现的问题，确保GAI能够更好地服务于医疗事业。

7.3 信息人：厘清医生与GAI的人机协同互动与认知处理过程

从信息生态的视角看，人机协同的可靠性不仅取决于GAI的技术性能，也取决于医生的使用策略、协同模式设计以及临床组织环境的规范约束。理解并优化医生这一关键“信息人”与GAI间的互动模式，是实现稳定、高效共生的基础。未来研究必须揭开协同决策的“微观黑箱”，从宏观结果比较转向微观过程机制阐释。具体而言，应分析医生在GAI辅助下的使用和认知过程，包括与GAI的信息互动情况、GAI生成信息的理解与评估过程、结合自身经验与GAI信息的推理方式以及最终决策行为等。通过梳理这些环节，厘清GAI辅助决策背景下医生与GAI的协同决策模式及医生的决策信息处理方式，为弥合GAI技术开发与临床使用之间的差距提供依据。

作者贡献：杨娇负责文章撰写与修改，郭蕊负责文章构思与论文审定。

作者声明本文无实际或潜在的利益冲突。

参 考 文 献

[1] ZENG D, QIN Y, SHENG B, et al. DeepSeek's "Low-Cost" Adoption Across China's Hospital Systems: Too Fast, Too Soon?[J]. JAMA, 2025, 333(21): 1866-1869.

[2] 张迪，张力伟. 数智信息生态系统：内涵、构成与机制[J]. 现代情报，2024，44(4)：11-21.

[3] KRISHNAN R, RAJPURKAR P, TOPOL E J. Self-supervised learning in medicine and healthcare[J]. Nature Biomedical Engineering, 2022, 6(12): 1346-1352.

[4] 张熙，杨小汕，徐常胜. ChatGPT及生成式人工智能现状及未来发展方向[J]. 中国科学基金，2023，37(5)：743-750.

[5] LEVINE D M, TUWANI R, KOMPA B, et al. The diagnostic and triage accuracy of the GPT-3 artificial intelligence model: an observational study[J]. The Lancet Digital Health, 2024, 6(8): e555-e561.

[6] CARUCCIO L, CIRILLO S, POLESE G, et al. Can ChatGPT provide intelligent diagnoses? A comparative study between predictive models and ChatGPT to define a new medical diagnostic bot[J]. Expert Systems with Applications, 2024, 235: 121186.

[7] AYOUB M, BALLOUT A A, ZAYEK R A, et al. Mind + Machine: ChatGPT as a Basic Clinical Decisions Support

Tool[J]. Cureus, 2023, 15(8): e43690.
[8] WANG H, WU W, DOU Z, et al. Performance and exploration of ChatGPT in medical examination, records and education in Chinese: Pave the way for medical AI[J]. International Journal of Medical Informatics, 2023, 177: 105173.
[9] YANEVA V, BALDWIN P, JURICH D P, et al. Examining ChatGPT Performance on USMLE Sample Items and Implications for Assessment[J]. Academic Medicine, 2024, 99(2): 192-197.
[10] ARMITAGE R C. Performance of Generative Pre-trained Transformer-4 (GPT-4) in Membership of the Royal College of General Practitioners (MRCGP) -style examination questions[J]. Postgraduate Medical Journal, 2024, 100(1182): 274-275.
[11] HIROSAWA T, KAWAMURA R, HARADA Y, et al. ChatGPT-Generated Differential Diagnosis Lists for Complex Case-Derived Clinical Vignettes: Diagnostic Accuracy Evaluation[J]. JMIR Medical Informatics, 2023, 11: e48808.
[12] SURAPANENI K M. Assessing the Performance of ChatGPT in Medical Biochemistry Using Clinical Case Vignettes: Observational Study[J]. JMIR Medical Education, 2023, 9: e47191.
[13] CABRAL S, RESTREPO D, KANJEE Z, et al. Clinical Reasoning of a Generative Artificial Intelligence Model Compared With Physicians[J]. JAMA Internal Medicine, 2024, 184(5): 581-583.
[14] GOH E, GALLO R, HOM J, et al. Large Language Model Influence on Diagnostic Reasoning: A Randomized Clinical Trial[J]. JAMA Network Open, 2024, 7(10): e2440969.
[15] SHIEH A, TRAN B, HE G, et al. Assessing ChatGPT 4.0's test performance and clinical diagnostic accuracy on USMLE STEP 2 CK and clinical case reports[J]. Scientific Reports, 2024, 14(1): 9330.
[16] HAGER P, JUNGMANN F, HOLLAND R, et al. Evaluation and mitigation of the limitations of large language models in clinical decision-making[J]. Nature Medicine, 2024, 30(9): 2613-2622.
[17] 韩普, 李雄, 陈文祺, 等. 基于大语言模型的医疗健康领域知识服务模式研究[J]. 现代情报, 2025, 44(1): 1-18.
[18] LAHAT A, SHARIF K, ZOABI N, et al. Assessing Generative Pretrained Transformers (GPT) in Clinical Decision-Making: Comparative Analysis of GPT-3.5 and GPT-4[J]. Journal of Medical Internet Research, 2024, 26: e54571.
[19] SUáREZ A, DíAZ-FLORES GARCíA V, ALGAR J, et al. Unveiling the ChatGPT phenomenon: Evaluating the consistency and accuracy of endodontic question answers[J]. International Endodontic Journal, 2024, 57(1): 108-113.
[20] MAGRUDER M L, RODRIGUEZ A N, WONG J C J, et al. Assessing Ability for ChatGPT to Answer Total Knee Arthroplasty-Related Questions[J]. The Journal of Arthroplasty, 2024, 39(8): 2022-2027.
[21] LIU F, ZHU T, WU X, et al. A medical multimodal large language model for future pandemics[J]. NPJ Digital Medicine, 2023, 6(1): 226.
[22] SI Y, YANG Y, WANG X, et al. Quality and Accountability of ChatGPT in Health Care in Low- and Middle-Income Countries: Simulated Patient Study[J]. Journal of Medical Internet Research, 2024, 26: e56121.
[23] 马泽冰, 刘怡兵, 范华雨, 等. 基于大语言模型的肌肉减少症诊治应用: 与临床医生诊断决策的对比研究[J]. 中国循证医学杂志, 2025, 25(7): 775-782.
[24] GOH E, GALLO R, HOM J, et al. Influence of a Large Language Model on Diagnostic Reasoning: A Randomized Clinical Vignette Study[J]. medRxiv, 2024 7(10): e2440969.
[25] WAN P, HUANG Z, TANG W, et al. Outpatient reception via collaboration between nurses and a large language model: a randomized controlled trial[J]. Nature Medicine, 2024, 30(10): 2878-2885.
[26] GOH E, BUNNING B, KHOONG E C, et al. Physician clinical decision modification and bias assessment in a randomized controlled trial of AI assistance[J]. Communications Medicine, 2025, 5(1): 59.
[27] 韩春光. 人工智能临床决策支持系统在Ⅰ-Ⅲ期乳腺癌术后辅助治疗中的应用性研究[D]. 合肥: 安徽医科大学, 2022.
[28] WANG D Y, DING J, SUN A L, et al. Artificial intelligence suppression as a strategy to mitigate artificial intelligence automation bias[J]. Journal of the America Medical Informatics Association, 2023, 30(10): 1684-1692.
[29] CHANDA T, HAUSER K, HOBELSBERGER S, et al. Dermatologist-like explainable AI enhances trust and confidence in diagnosing melanoma[J]. Nature Communicatons, 2024, 15(1): 524.
[30] JABBOUR S, FOUHEY D, SHEPARD S, et al. Measuring the Impact of AI in the Diagnosis of Hospitalized Patients: A Randomized Clinical Vignette Survey Study[J]. JAMA, 2023, 330(23): 2275-2284.
[31] WIES C, HAUSER K, BRINKER T J. Reply to: False conflict and false confirmation errors are crucial components of AI accuracy in medical decision making[J]. Nature Communications, 2024, 15(1): 6897.

[32] ROSENBACKE R, MELHUS Å, STUCKLER D. False conflict and false confirmation errors are crucial components of AI accuracy in medical decision making[J]. Nature Communications, 2024, 15(1): 6896.
[33] REVERBERI C, RIGON T, SOLARI A, et al. Experimental evidence of effective human-AI collaboration in medical decision-making[J]. Scientific Reports, 2022, 12(1): 14952.
[34] JUSSUPOW E, SPOHRER K, HEINZL A, et al. Augmenting Medical Diagnosis Decisions? An Investigation into Physicians' Decision-Making Process with Artificial Intelligence[J]. Information Systems Research, 2021, 32(3): 713-735.

[收稿日期:2025-11-27 修回日期:2025-12-12]
(编辑 赵晓娟)

世卫组织警告:医疗人工智能快速扩张,法律与伦理保障严重滞后

在人工智能正迅速改变医疗面貌之际,世卫组织欧洲区域办事处周二发布最新报告《健康领域的人工智能:欧洲区域准备现状》,警告在医疗领域大规模采用人工智能的同时,相关法律和伦理保障却明显滞后,已难以有效保护患者和医务人员。

这份报告基于欧洲区域53个成员国中50国的反馈,首次系统呈现了人工智能在医疗体系中的应用、监管与风险现状。报告指出,人工智能已在诊断、行政工作优化和提升医患沟通方面发挥重要作用,但随着应用加速扩散,责任划分模糊、数据安全不足和法律框架缺失等问题日益突出。

准备程度差异显著,监管滞后成最大隐患

尽管几乎所有国家都意识到人工智能在诊疗、疾病监测和个性化医疗中的潜力,但整体准备程度则"成碎片化且不均衡"。仅有4个国家制定了医疗人工智能国家战略,另有7国正在制定之中。

部分国家已迈出前进步伐:爱沙尼亚整合电子健康记录与人口数据库以支持人工智能工具,芬兰投资开展面向医务人员的人工智能培训,西班牙则在基层医疗试点人工智能辅助早期疾病识别。

然而,全区域86%的国家认为法律不确定性是人工智能应用的首要障碍,78%将资金不足视为主要问题。仅8%的国家制定了医疗人工智能责任标准,明确人工智能造成错误或伤害时的责任归属。

应用广泛但投入不足

目前,32个国家已采用人工智能辅助诊断,半数国家引入聊天机器人用于患者支持,超过一半国家明确了医疗领域人工智能的优先应用方向,但仅四分之一为此提供专项资金。各国推动医疗人工智能的主要动机包括提升患者护理质量、缓解医疗系统人力压力以及提高效率和生产力。

患者安全、隐私与公平成为核心关切

报告指出,人工智能在医疗体系的扩张与普通民众息息相关,涉及三大关键问题:患者安全、公平获得医疗服务以及数字隐私。人工智能依赖数据进行学习,一旦数据存在偏见或缺失,可能导致漏诊、误诊或资源分配不公。

报告呼吁各国制定符合公共卫生目标的人工智能国家战略,投资发展人工智能能力建设,加强法律与伦理保障机制,以透明方式吸引公众参与,并完善跨境数据治理。

(摘编自:联合国新闻网)