

试论生成式人工智能的医疗应用 能力与风险边界*

王 硕^{1,2} 刘天语³ 汪 琛¹ 刘瑶瑶¹



摘要: 生成式人工智能医疗应用的伦理探讨需要对两个基本问题进行谨慎、清晰的划界。一方面,“能力”的划界旨在澄清生成式人工智能医疗应用的核心优势与技术局限,基于物种差异,明确生成式人工智能医疗应用相较于传统医疗人工智能的关键差异,避免陷入“ChatGPT 全能论”或“ChatGPT 无能论”的窠臼。另一方面,“风险”的划界旨在区分生成式人工智能医疗应用带来的问题类别及特征,厘清不同风险的独特性质与根源,包括传统人工智能的共有风险、生成式人工智能的特有风险以及生成式人工智能同医疗实践相结合的特性风险,推动风险的分类管理并推进伦理治理的不断完善。

关键词: 生成式人工智能, 医疗应用, ChatGPT, 能力划界, 风险划界

中图分类号: R-05 **文献标识码:** A **文章编号:** 1002-0772(2024)12-0001-05

DOI: 10.12014/j.issn.1002-0772.2024.12.01

Exploring the Capabilities and Risk Delineation of Generative AI in Medical Applications WANG Shuo^{1,2}, LIU Tianyu³, WANG Chen¹, LIU Yaoyao¹. 1. School of Social Sciences, Tsinghua University, Beijing 100084, China; 2. Kavli Institute for the Physics and Mathematics of the Universe, The University of Tokyo, Chiba 277-8583, Japan; 3. School of Humanities, Tsinghua University, Beijing 100084, China

Abstract: Ethical discussions on the application of generative artificial intelligence (AI) in healthcare require careful and clear delineation of two fundamental issues. On one hand, the delineation of "capabilities" aims to clarify the core advantages and technical limitations of generative AI in medical applications. Based on species differences, it is essential to delineate the key differences between generative AI medical applications and traditional medical AI, avoiding falling into the trap of "ChatGPT omnipotence" or "ChatGPT impotence". On the other hand, the delineation of "risks" aims to distinguish the categories and characteristics of problems brought about by generative AI medical applications, clarifying the unique nature and roots of different risks. This includes common risks of traditional AI, unique risks of generative AI, and characteristic risks associated with the integration of generative AI with medical practice. This promotes the classification and management of risks and advances the continuous improvement of ethical governance.

Key Words: generative artificial intelligence, medical application, ChatGPT, capability delineation, risk delineation

1 划界: 生成式人工智能的医疗应用伦理的逻辑起点

当前,生成式人工智能的技术进步正有力地驱动医疗领域的深刻变革与新质生产力的不断发展。生成式人工智能医疗应用依托机器学习技术和大规模医疗健康数据

集,有效整合了自然语言处理、图像识别、预测建模、内容生成等前沿技术,基于对海量医疗数据进行精确分析,能够处理医疗文本^[1]、辅助诊断流程^[2]并提供定制化、个性化的医疗建议^[3]。医疗与生成式人工智能的深度融合有助于减轻医生的工作负担^[4]和优化患者的就医体验^[5],有利于提高医疗行业的智能化水平,对促进社会医疗进步和增进公共健康福祉具有重要意义。

对于生成式人工智能医疗应用的理论与实践,学界已经进行了一些初步探讨。现有研究涉及生成式人工智能医疗应用的赋能场景^[6]、发展历程^[7]、伦理风险^[8]和治理机制^[9]等重要问题,对理解生成式人工智能在医疗领域的潜力和挑战具有重要意义。然而,现有关于生成式人工智能医疗应用的哲学与社会科学研究,尚未充分汲取生成式人工智能技术的条件特征,仍多是对既往“医疗人工智能研究”和“医疗机器人研究”的补充性延展,鲜有突出生

*基金项目: 2023 年中国博士后科学基金第 74 批面上资助项目(2023M741978); 2023 年国家资助博士后研究人员计划项目(GZC20231384)

1. 清华大学社会科学学院 北京 100084

2. 东京大学卡维里宇宙物理学与数学研究所 日本千叶 277-8583

3. 清华大学人文学院 北京 100084

作者简介: 王硕(1999-), 男, 博士研究生, 研究方向: 科技伦理。

通信作者: 刘瑶瑶(1993-), 女, 博士, 助理研究员, 研究方向: 医学人类学、科技与社会。

E-mail: sanfy611@mail.tsinghua.edu.cn

成式人工智能的新质特点。换言之,这些研究思路更多聚焦一般性伦理问题在生成式人工智能领域的解释套用,对生成式人工智能医疗应用的本质特征和异质潜能把握不充分,致使其限于传统人工智能哲学与社会科学研究的框架边界,未能直面生成式人工智能的颠覆性影响^[10]。

事实上,理解新兴技术的涌现与发展,确实需要借助过往相似属种的概念与理论作为知识基础。在这个意义上,不能脱离过往有关医疗人工智能的研究来独立探讨生成式人工智能医疗应用的潜力与风险。然而,更重要的是,生成式人工智能与既往人工智能虽然“同根同源”,但却有着本质的变化,哲学研究者的使命应当是对其进行准确的技术与社会意义上的“划界”,也就是区分其与传统技术的边界,并将其作为思辨性哲学讨论的起点。对于生成式人工智能医疗应用的伦理探讨,需要对两个基本问题进行谨慎、清晰的划界,并将其作为讨论其他问题的逻辑起点。

第一个基本问题是“能力”的划界。生成式人工智能在医疗领域究竟能做什么,不能做什么?这个问题的本质在于理解生成式人工智能的技术能力及其在医疗领域的应用范畴。如果没有清晰地讨论这个问题,我们可能会过高估计或低估生成式人工智能的能力,从而陷入“ChatGPT全能论”或者“ChatGPT无能论”的二极化“泥沼”。过分夸大其能力可能会导致人类医生的主体权威消解,而低估其能力则可能错失利用这一技术改善医疗服务与发展新质生产力的机会窗口。

第二个基本问题是“风险”的划界。生成式人工智能为医疗领域带来的诸多挑战,究竟哪些是新的风险,哪些是已经得到充分讨论的风险?这个问题的核心在于评估和平衡生成式人工智能应用的风险及其实际与潜在影响。如果未能清晰地讨论此问题,我们可能无法充分预见和应对各类风险问题。这既不利于增进公众对新技术的理解和信任,也不利于在各个治理环节中做好充分的预期和准备。

2 生成式人工智能医疗应用的能力划界

2.1 ChatGPT全能论与ChatGPT无能论

对于生成式人工智能医疗应用能力的认识,存在ChatGPT全能论和ChatGPT无能论的两种极化观点。全能论蕴含着浓厚的技术进步主义预设,主张生成式人工智能能够有效解决医疗领域的各种挑战,甚至超越人类医生。

全能论的主张包括:首先,生成式人工智能的通用性使其能够适应多种医疗场景,从解析复杂医疗记录到提供医学见解,特别是在常见疾病诊断、患者历史分析和医学研究等领域。其次,生成式人工智能的拟人交互性,尤其是情感性的形式表达,增强了医疗服务的人性化。最后,在疾病分类和医疗文献检索领域,生成式人工智能已经显示出与人类医生相当甚至更高的准确率,预示其在未来医疗领域的进一步应用和创新潜力^[11]。

无能论则对生成式人工智能在医疗领域的应用能力持有严格审视的态度,其主要观点包括:首先,生成式人工智能在医学知识深层次理解方面存在不足,难以应对复杂病例、制定治疗计划或理解患者的具体医疗需求。例如,由于缺乏专业的医学知识,生成式人工智能通常无法准确分析医学影像如X光片或磁共振成像^[12]。其次,生成式人工智能在把握医学数据的临床意义和解读复杂语境以及隐含信息方面存在局限,在应用中往往会默化并抹除个体独特性,难以深入了解每个患者的独特状况与既往病史,因此难以全面理解患者描述的症状并提供针对性的诊疗建议。最后,生成式人工智能无法在所有医疗领域都实现相同的性能。例如,在外科领域,它无法处理手术过程中的并发症或其他意外情况,也无法替代成熟外科医生的经验和专业知识^[13]。

诚然,作为自主能动性技术工具的生成式人工智能既不可能是万能的,更不可能是完全无能的。全能论和无能论这两种极端观点,虽然不完全贴合实际,但却为我们提供了一种哲学探讨的方法,那就是通过抽离条件的思想实验,更清醒地认识到生成式人工智能的重要潜力以及可能的效用“软肋”。这样的极化观点能够通过观点间的对垒与张力,推动我们深入思考和探索,而非简单接受或拒绝技术的可能性。但要实现对生成式人工智能实际的把握,需要准确理解和界定生成式人工智能在医疗领域可以做什么、不能做什么,以及在哪些方面能够提供真正的价值。

2.2 传统人工智能和生成式人工智能的属种差异

现在,我们有必要将这个问题回溯到对于传统人工智能和生成式人工智能之间差异的讨论。传统人工智能和生成式人工智能之间的区别并非只是技术细节上的,相反,二者代表了人工智能发展中的两个不同阶段,各自拥有独特的能力和应用领域。理解这两者之间的关系和差异,将帮助我们清晰地看到生成式人工智能在医疗领域中的潜力和限制,从而更有效地探讨和定义其能力边界。这种属种关系的探讨为我们提供了一个结构性框架,使我们更加清晰地认识到,生成式人工智能医疗应用与传统医疗人工智能存在很大差异。

人工智能可以被看作是一个更广泛的“属”,它包含了多个不同的子领域和二级技术,生成式人工智能则可以被视为人工智能领域内的一个特定“种集”,代表了人工智能的一个分支簇,专注新数据的产出与新内容的自主生成。

按照这样的划分,传统人工智能和生成式人工智能既是同“属”异“种”的关系,也是“一脉相通”的发展关系。传统人工智能通常关注数据的分析与处理,旨在依托特定数据集进行训练以执行特定任务,例如语言转译、信息识别、文本归类等。生成式人工智能则更进一步,在对数据进行理解与分析的基础上,还能够创造全新的内容输

出,比如生成全新的文本,并且能够完成模拟人类对话等更具挑战性的任务^[14]。

2.3 属种差异与能力增强

当我们将这种属种差异投射到医疗领域,便会清晰认识到,生成式人工智能医疗应用与传统医疗人工智能相比,呈现出以下关键差异。见图 1。

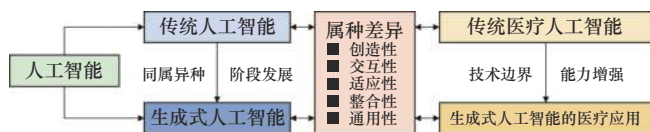


图 1 生成式人工智能医疗应用的能力划界

一是技术自主的创造性生成。传统医疗人工智能往往局限于对已有信息的分析和处理,在诊断辅助、病例分析和影像识别等方面具有比较广泛的应用。相比之下,生成式人工智能不仅能精准分析医疗数据,还有望具备循证迭代以生成各类医疗文本的能力,在自动起草病历报告、出院总结等医疗文本方面具有巨大潜力^[15]。

二是技术类人的交互能力。传统医疗人工智能在更多情况下被作为一种“后台工具”,只能对人类医生进行医疗决策起到辅助作用,在与患者的直接交流中可能显得十分有限。而生成式人工智能则具有更加自然和灵活的交互能力,能以更接近人类的方式表达,实现与患者的拟人对话,更加精准地考虑到患者情感和心理学需求。

三是技术独有的适应性学习。传统医疗人工智能需要在明确的数据集上进行训练,因此更专注于特定任务。而深度的知识积累及生成能力为生成式人工智能提供了更强大的适应性,使其具备对医学领域更为全面的认知,能够更好地理解和适应多样化的医疗场景和需求。

四是技术兼备的整合性效能。传统医疗人工智能在处理非标准医疗数据时可能会面临限制。而生成式人工智能能够处理更广泛的数据类型,特别是在处理非结构化数据时(如自由文本的医疗记录和患者询问),可以展现更高的灵活性和准确性。比如,ChatGPT 在医疗文本内容处理方面表现良好,能够在电子健康记录、临床笔记和医学研究论文等自由文本中提取有效信息^[16]。

五是技术场景的通用性适用。传统医疗人工智能主要适用于明确定义的任务,如疾病诊断或治疗计划的制定。生成式人工智能则在多个医疗领域内展现出更大的灵活性和广泛性,包括但不限于临床研究、康复教育和健康管理,从而为医疗服务提供更全面的支持。比如,ChatGPT 可以帮助抗寄生虫疫苗等前沿医疗科技的开发^[17],协助应对意外分娩等医疗紧急情况^[18]。

2.4 生成式人工智能的特性与技术边界

基于以上分析发现,生成式人工智能医疗应用的核心主要集中在创造性的“生成”。例如,生成式人工智能医疗应用可以为撰写病史和医疗报告、模拟患者对话情境以及编写临床文档等任务提供支持。然而,这也意味着在

那些对“生成”需求不大的领域和场景,其应用功能将会受到一定限制。也就是说,这在一定程度上反映出了生成式人工智能的技术边界,即它们并非万能。而这一点,往往被万能论或全能论的支持者所忽视。

生成式人工智能擅长的是根据已有数据和模式“生成”新的输出,更多是对文本语言及其语义内在关联性的形式化掌握^[19],而非实质性的理解。就后者而言,生成式人工智能同临床医生相比依然存在不少差距^[20]。比如,就需要复杂医学知识和深入专业判断的任务而言,由于文本语料库非完备性的客观现实,生成式人工智能有可能无法准确理解复杂的医学情境,导致得出错误的医学结论或给出完全不切实际的医疗建议。再比如,有别于西方生物学知识体系的中医理论包含了“阴阳”“五行”“气血”等丰富的抽象概念与类比想象,生成式人工智能可能会面临难以精确理解和解释这些表达的深层含义的挑战。

所以,在某些关键医疗领域,如精准医学、复杂疾病的多因素分析、个性化治疗决策、心理关怀咨询等,生成式人工智能可能并不是最佳选择。在这些领域,医学专家的经验、直觉判断以及涉及需要高度人类共情和个性化关怀的任务是不可替代的。

3 生成式人工智能医疗应用的风险划界

2024 年 1 月,联合国世界卫生组织发布《卫生领域人工智能的伦理与治理:多模态大模型指南》,明确提及卫生大模型的挑战和风险^[21]。生成式人工智能医疗应用的潜在风险成为本文研究的另一划界问题,其关键在于区分和识别新出现的风险与那些已经被广泛讨论的风险。生成式人工智能医疗应用的风险可以被划分为三类:一般人工智能的风险、生成式人工智能的一般风险以及生成式人工智能与医疗应用结合的特性风险。见图 2。这种划分方式有助于我们更清楚地理解每种风险的独特性质和根源,从而使我们能够更有效地分类应对和治理这些风险。

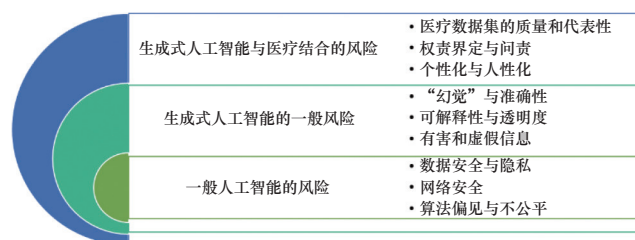


图 2 生成式人工智能医疗应用的风险划界

3.1 一般人工智能风险

应当认识到,一些风险并非独属于生成式人工智能,而是传统人工智能领域的共通问题。这些风险的深入讨论能够为理解和应对生成式人工智能医疗应用的挑战提供一个基准性的框架。

一是数据安全和隐私风险。生成式人工智能医疗应用处理大量个人健康信息,包括疾病历史、遗传数据和其他私人信息。这些敏感且有价值的信息可能成为网络攻

击目标,导致隐私泄露和个人信息滥用,进而降低患者对医疗机构的信任度,引发医患矛盾^[22]。

二是网络数据安全风险。比如数据投毒、对抗性攻击等网络安全威胁。数据投毒攻击通过篡改训练数据来误导模型学习错误的模式,而对抗性攻击则通过细微修改输入数据来诱导模型做出错误决策。这些攻击可能导致潜在的医疗事故,危及患者生命健康。

三是算法偏见和社会不公平^[23]。这些偏见往往源于训练数据的不平衡或是潜藏于数据采集等阶段的人类偏见^[24],可能导致模型在医疗决策过程中对特定群体产生歧视,如基于种族、性别或经济地位的不公平治疗,进而带来算法价值偏见矫治的难题^[25]。

3.2 大语言模型与生成式人工智能风险

还有一类风险是大语言模型与生成式人工智能的特性风险,是相较于传统人工智能发展所出现的新风险。这些风险也不仅限于医疗领域,而是生成式人工智能普遍面临的问题。

一是“幻觉”问题和真实性挑战。正如我们时常会感觉 ChatGPT 在“一本正经地胡说八道”,“幻觉”指的是听起来合理但实际上不准确甚至完全错误的内容,造成一种“事实欺诈”问题。有研究指出,ChatGPT 很难排除“幻觉”的医疗信息,Med-PaLM 生成的文本往往也会包含较多不正确的医疗内容^[26]。这一问题源于生成式人工智能的训练数据可能不够准确或完整,以及这些模型缺乏因果推断能力,它们只能学习文本之间的概率关联,而不能深入理解信息^[27]。

二是可解释性和透明度问题^[28]。生成式人工智能的理论基础和机制目前仍然是一个“黑箱”^[29],这使得其输出难以完全控制或预测。由于这些模型的复杂性和庞大的参数规模,加之生成式人工智能的自我与外源激发性特征^[30],追踪单一数据在模型中的处理过程变得极其困难,难以提供模型决策过程的有效解释。在实际的医疗诊断中,生成式人工智能难以清楚解释其诊断逻辑和病灶识别依据,而这对医生和患者理解其诊断结果非常重要。此外,模型的闭源性和验证难度也有可能增加了用户的疑虑和不信任。

三是传播有害和虚假信息风险^[31]。生成式人工智能或许可以为用户生成一系列看似专业和可信的医疗建议,而事实上,生成内容却包含着不准确的病理信息或不科学的治疗方法。正如社交媒体上广为流传的调侃“百度搜病,癌症起步”,生成式人工智能可能给用户提提供不准确、夸张和不科学的医疗信息。而高度拟人化的特征可能导致患者过度信赖自动生成的医学内容,从而接触到有害或误导性信息,增加医疗领域的不适当焦虑和自我诊断风险。

3.3 生成式人工智能与医疗结合的特性风险

需要注意到,生成式人工智能医疗应用还面临一些独

特的风险,这些风险是生成式人工智能技术特性与医疗实践特点相结合的产物。

一是医疗训练数据的质量和代表性问题。例如,使用公共数据集训练的生成式人工智能可能会忽略其他重要的医疗情况和条件,尤其是当这些数据集只限于特定地区或人群时。因此,生成式人工智能医疗应用需要进一步同大量且高度验证的医学知识资源更好结合^[32]。

二是权责界定和问责机制问题。主要有两种医疗情形下的责任备受关注。一方面,人工智能生成文本的医疗知识或信息产权归属依然存在争议。另一方面,生成式人工智能的参与可能引入了新的责任主体,医疗事故的责任承担亦需要更加清晰的界定。因此,在开发、实施和使用生成式人工智能过程中涉及的多方利益相关者之间的权责关系也需要明确和规范。

三是个性化与人性的伦理考量。当患者过度依赖生成式人工智能做出决策时,他们可能不再根据自己的个体经验和价值观来进行决断,这可能导致患者的自主选择权受到削弱,同时也降低了决策的个性化程度。此外,医生同患者的有效沟通是确保医生全面理解患者状况并提供个性化诊疗建议的关键。过度依赖生成式人工智能的建议可能减少了患者与医生之间的深入交流和共同决策。因此,生成式人工智能在医疗决策中的角色和在不同医疗任务中的应用需要进行深入的伦理考量。

4 讨论与启示

本文强调了对生成式人工智能医疗应用的能力和进行“划界”的重要性。“划界”旨在明确揭示生成式人工智能医疗应用的具体能力和局限,并充分阐明新出现的风险与已广泛讨论的风险的区别。这对于生成式人工智能医疗应用的伦理探讨、理论研究以及现实治理都具有重要价值和意义。

一方面,基于划界的伦理探讨前提。本文的分析表明,对生成式人工智能医疗应用的能力进行准确的评估至关重要。“能力划界”需要明确生成式人工智能与传统人工智能的关键区别,并在此基础上明确上述区别对于医疗实践的具体影响。我们应避免对生成式人工智能能力的过度估计或低估,不盲目崇拜也不应一概否定。在伦理探讨中,我们有必要基于对生成式人工智能真实能力的深刻洞察,避免由于对生成式人工智能能力和局限的错理解而做出过度的理论推断和得出片面的学术结论。

另一方面,基于划界的新兴风险治理。区分传统人工智能技术共有的风险、生成式人工智能独特的风险以及生成式人工智能同医疗实践相结合产生的特有风险,是对作为一项新兴科技的生成式人工智能进行有效管理的关键。其中,部分传统风险已有充分的学术讨论和政策响应。而生成式人工智能和医疗实践结合产生的新风险则亟需重视,对此,要加强研究并制定针对性策略。此外,生成式人工智能医疗应用的治理不仅需要遵循人工智能伦理原

则, 还需兼顾生命医学传统的伦理要求, 以确保这些技术的负责任和有效应用。

当然, 特别需要指明的是, “划界”并不意味着某种固定的严格界限。从历史的视野来看, 新兴科学技术的划界是一个不断发展变化的动态进程。在一定的技术条件下, 昨天的风险可能变成今天的机遇, 今天的伦理担忧也许会成为未来的“技术期待”。例如, 许多研究认为生成式人工智能在价值判断和同理心等方面存在局限, 然而, 也有研究发现 ChatGPT-4 已经能够在医学执照考试中涉及同理心、道德判断和专业精神的问题上取得优秀表现, 甚至在一定程度上超越了人类医师^[33]。这种技术演化的持续性告诉我们, 对待迅速迭代的新兴医疗技术, 要用动态和发展的观念来看待, 要以历史性的眼光把握科学技术的能力与风险的边界, 同时也要以前瞻式和预见性的警觉, 防范技术伦理的新兴风险。

参考文献

- [1] 梅景墨, 高艳红. ChatGPT 在检验医学中的应用挑战[J]. 标记免疫分析与临床, 2023, 30(4): 695-698.
- [2] DAVE T, ATHALURI S A, SINGH S. ChatGPT in medicine: An overview of its applications, advantages, limitations, future prospects, and ethical considerations[J]. Front Artif Intell, 2023, 6: 1169595.
- [3] KHAN R A, JAWAID M, KHAN A R, et al. ChatGPT: Reshaping medical education and clinical management[J]. Pak J Med Sci, 2023, 39(2): 605-607.
- [4] PATEL S B, LAM K. ChatGPT: The future of discharge summaries[J]. Lancet Digit Health, 2023, 5(3): e107-e108.
- [5] BETZLER B K, CHEN H, CHENG C Y, et al. Large language models and their impact in ophthalmology[J]. Lancet Digit Health, 2023, 5(12): e917-e924.
- [6] 康砚澜, 郭倩宇, 张文强, 等. 基于知识增强的医学语言模型: 现状、技术与应用[J]. 医学信息学杂志, 2023, 44(9): 12-22.
- [7] LI Y, LI Z, ZHANG K, et al. ChatDoctor: A medical chat model fine-tuned on a large language model meta-AI (LLaMA) using medical domain knowledge[J]. Cureus, 2023, 15(6): e40895.
- [8] LIEBRENZ M, SCHLEIFER R, BUADZE A, et al. Generating scholarly content with ChatGPT: Ethical challenges for medical publishing[J]. Lancet Digit Health, 2023, 5(3): e105-e106.
- [9] 肖仰华, 徐一丹. 大规模生成式语言模型在医疗领域的应用: 机遇与挑战[J]. 医学信息学杂志, 2023, 44(9): 1-11.
- [10] 李正风, 刘瑶瑶. 科技伦理治理要准确把握新科技革命及其伦理问题的新特点[J]. 科学通报, 2024, 69(13): 1677-1680.
- [11] DIGIORGIO A M, EHRENFELD J M. Artificial intelligence in medicine & ChatGPT: De-tether the physician[J]. J Med Syst, 2023, 47(1): 32.
- [12] CADAMURO J, CABITZA F, DEBELJAK Z, et al. Potentials and pitfalls of ChatGPT and natural-language artificial intelligence models for the understanding of laboratory medicine test results: An assessment by the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group on Artificial Intelligence(WG-AI)[J]. Clin Chem Lab Med, 2023, 61(7): 1158-1166.
- [13] CHENG K, SUN Z, HE Y, et al. The potential impact of ChatGPT/GPT-4 on surgery: Will it topple the profession of surgeons?[J]. Int J Surg, 2023, 109(5): 1545-1547.
- [14] MOOR M, BANERJEE O, SHAKERI Z, et al. Foundation models for generalist medical artificial intelligence[J]. Nature, 2023, 616(7956): 259-265.
- [15] ALI S R, DOBBS T D, HUTCHINGS H A, et al. Using ChatGPT to write patient clinic letters[J]. Lancet Digit Health, 2023, 5: e179-e181.
- [16] CASCELLA M, MONTOMOLI J, BELLINI V, et al. Evaluating the feasibility of ChatGPT in healthcare: An analysis of multiple clinical and research scenarios[J]. J Med Syst, 2023, 47(1): 33.
- [17] ŠLAPETA J. Are ChatGPT and other pretrained language models good parasitologists?[J]. Trends Parasitol, 2023, 39(5): 314-316.
- [18] SANTO D S E, JOVIANO-SANTOS J V. Exploring the use of ChatGPT for guidance during unexpected labour[J]. Eur J Obstet Gynecol Reprod Biol, 2023, 285: 208-209.
- [19] 陈小平. 大模型关联度预测的形式化和语义解释研究[J]. 智能系统学报, 2023, 18(4): 894-900.
- [20] SINGHAL K, AZIZI S, TU T. et al. Large language models encode clinical knowledge[J]. Nature, 2023, 620(7972): 172-180.
- [21] World Health Organization. Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models[EB/OL]. (2024-01-18)[2024-03-06]. <https://www.who.int/publications/i/item/9789240084759>.
- [22] 颜见智, 何雨鑫, 骆子焜, 等. 生成式大语言模型在医疗领域的潜在典型应用与面临的挑战[J]. 医学信息学杂志, 2023, 44(9): 23-31.
- [23] 王 硕, 文侃骁, 刘天语, 等. 医疗机器人的伦理风险与治理探析[J]. 医学与哲学, 2023, 44(23): 16-21.
- [24] 汪 琛. 医疗人工智能伦理治理的问题、困境与求解[J/OL]. 科学学研究, 1-18(2024-03-13)[2024-03-31]. <https://doi.org/10.16192/j.cnki.1003-2053.20240313.001>.
- [25] 刘天语, 王 硕, 刘鸿宇. “算法制”伦理: 数字社会的组织伦理新向度[J]. 自然辩证法研究, 2023, 39(6): 78-84.
- [26] HARRIS E. Large language models answer medical questions accurately, but can't match clinicians' knowledge[J]. JAMA, 2023, 330(9): 792-794.
- [27] RAO A, PANG M, KIM J, et al. Assessing the utility of ChatGPT throughout the entire clinical workflow: Development and usability study[J]. J Med Internet Res, 2023, 25: e45659.
- [28] BISWAS S. ChatGPT and the future of medical writing[J]. Radiology, 2023, 307(2): e223312.
- [29] THIRUNAVUKARASU A J, TING D S J, ELANGOVAN K, et al. Large language models in medicine[J]. Nat Med, 2023, 29(8): 1930-1940.
- [30] 陈小平. 大模型: 从基础研究到治理挑战[J]. 中国人工智能学会通讯, 2024, 14(1): 2-9.
- [31] FERRES J M L, WEEKS W B, CHU L C, et al. Beyond chatting: The opportunities and challenges of ChatGPT in medicine and radiology[J]. Diagn Interv Imaging, 2023, 104(6): 263-264.
- [32] KUNG T H, CHEATHAM M, MEDENILLA A, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models[J]. PLoS Digit Health, 2023, 2(2): e0000198.
- [33] BRIN D, SORIN V, VAID A, et al. Comparing ChatGPT and GPT-4 performance in USMLE soft skill assessments[J]. Sci Rep, 2023, 13(1): 16492.

收稿日期: 2024-01-05

修回日期: 2024-04-02

(本文编辑: 于 磊)