

· 科学技术与社会 ·

医疗人工智能家长主义伦理困境及其治理

Ethical Dilemmas and Governance of Paternalism in Medical Artificial Intelligence

吕宇静 /LV Yujing 杜严勇 /DU Yanyong

(同济大学人文学院, 上海, 200092)
(School of Humanities, Tongji University, Shanghai, 200092)

摘要: 随着人类社会的深度智能化, 医疗人工智能的应用已经成为一种趋势, 并逐渐在实践中显现出家长主义的倾向。尽管这类新的参与者可赋予医患双方角色职责的转变, 且在数据管理和决策辅助方面具备潜在正当性, 但是不能忽略其中隐含的伦理风险。医患自主性的发挥、医疗有利原则的实现、医疗决策的可靠性及公正性都受到医疗人工智能家长主义的影响, 相关困境的纾解已变得日渐紧迫。其中, 患者赋权维度旨在激发患者参与决策的自主性, 有限的医生权威可适当弥补机器决策的局限, 医疗人工智能的技性改良是发挥其助益功能的关键, 还需要倡导医生、患者、医疗人工智能三元协同的民主关系, 以缓释医疗人工智能家长主义引发的伦理困境。

关键词: 医疗人工智能 家长主义 决策 伦理

Abstract: With the intelligent development of society in depth, the application of medical artificial intelligence (AI) has become a trend, and gradually gives rise to paternalistic tendencies. Although these participants, the AI systems, may reshape the roles and responsibilities of physicians and patients and they are often justified in domains such as data management and decision support, the hidden ethical risks should not be overlooked. This paternalism of medical AI can constrain both patient and physician autonomy, undermine the realization of the principle of medical beneficence, compromise the reliability and fairness of medical decision-making, and addressing these challenges has therefore become increasingly urgent. In response, patient empowerment promotes meaningful participation of patients in decision-making. At the same time, a moderated form of physician authority remains necessary to compensate for the limitations of machine decision-making and continued technical refinement of medical AI is the key to play its beneficial function. In addition, a more balanced and collaborative relationship among physicians, patients, and medical AI systems is needed to mitigate the ethical dilemmas associated with the paternalism in medical AI.

Key Words: Medical artificial intelligence; Paternalism; Decision-making; Ethics

中图分类号: TP18; R-052 DOI: 10.15994/j.1000-0763.2026.06.010 CSTR: 32281.14.jdn.2026.06.010

基金项目: 国家社会科学基金重大项目“人工智能伦理风险防范研究”(项目编号: 20&ZD041)。

收稿日期: 2025年9月27日

作者简介: 吕宇静(1990-)女, 河南郑州人, 同济大学人文学院博士研究生, 研究方向为科技伦理。Email: 2311585@tongji.edu.cn

杜严勇(1976-)男, 四川南充人, 同济大学人文学院教授, 研究方向为科技伦理。Email: yanyongdu@tongji.edu.cn

人工智能(AI)在医疗领域的快速发展和应用,既带来了前所未有的机遇,也伴随着巨大的挑战。医疗AI可借助海量数据进行高效检测并给出关联性的诊断建议,亦能通过智能方案优化患者管理以适当降低医疗成本,^[1]如今已被视为增强专业医疗人员水平的有效方法。^[2]但是,在融入具体实践的同时,医疗AI也引发了其决策风格的伦理隐忧。其中最具争议的话题即包括以算法建议为表现形式的新型医疗家长主义(Paternalism),又称机器家长主义,主要是指AI在没有或很少医患(医生与患者的简称)参与的情况下做出独立决策的情况。当AI在临床决策中发挥权威作用时,它可能优先考虑算法建议而不是来自患者、代理人或临床医生在内的人类判断。尽管医疗AI的决策对行为能力不足的患者具备一定的适用性,但是若存在特殊病况,AI可能无法保障结果的可靠性与准确性,而且不能排除医疗AI决策中裹挟着隐私泄露、偏见风险及归责问题。为了更加明确地把握医疗AI家长主义这一问题,文章从医疗AI家长主义的新特征出发,揭示医疗AI家长主义引发的伦理风险,并进一步呈现针对性的治理策略。

一、医疗AI家长主义的特征与积极意义

家长主义,源于拉丁语“pater”(父亲)。传统医疗家长主义模式中,患者的角色是同意医生的决定。拥有医学知识和技能的医生根据自己的专业判断来决定何为患者的最佳利益,达成一种医生视角的患者健康福祉。^[3]当代医学伦理学将其视为对患者自主性的不合理干涉。^[4]医疗AI的应用衍生了新型的家长主义:一方面AI自主地采取行动;另一方面根据它们的编码规则,机器可以成为“绝对”的影响力,且不具备被推翻的条件或可能性。^[5]尽管医疗AI能够提供可能性方案,但是这种模式让患者重新回到家长主义的关系中。

1. 医疗AI扩展决策力量

与传统的医疗家长主义不同,AI家长主义在决策过程中引入了额外的行为体。相应地,

医疗决策主体模式也发生改变。传统的医疗决策以医生-患者为主,医疗AI的加入扩展了这种既定风格,呈现为人工智能-医生、人工智能-患者与人工智能-医生-患者模式。

在人工智能-医生模式中,医疗AI作为医生的决策助手,逐渐地嵌入临床诊断流程,可以承担风险识别、数据管理与建议支持功能。虽然目前仍表现为形式上的医生决策主导,但是鉴于医疗AI的技术优势和算法效力,医生可能会倾向于接受AI提供的建议。在人工智能-患者模式中,医疗AI直接为患者提供初步的诊断建议或者推荐治疗路径。随着远程医疗与健康管理设备的普及,这种模式在实践中愈来愈常见。患者可随时使用AI系统进行医疗问答或个性化健康咨询,达成一种“去中介化”的互动方式,这显著提升了患者获取医疗信息的便捷性。但是,该模式中的算法预设或技术依赖构成一种新型的决策权威,呈现一种表象化的自主交互状态。在人工智能-医生-患者模式中,医疗AI作为决策力量,为专业医生和患者提供诊断建议和治疗帮辅。这时AI不只是信息载体,亦是影响医生与患者的干预力量。不仅医生将AI的输出作为诊断参考,患者也可能在缺乏理解状态下被动地接受AI建议,此时患者可能陷入医生和医疗AI叠加形成的双重权威。

在新的医疗家长主义中,专家性判断依据转移至医疗AI,医疗决策由“医生最了解”逐渐被“机器最了解”模式取代。^[6]医疗AI可以实例化为“专家iDoctor”,成为医疗团队的新成员,^[7]并执行临床医生的医疗判断功能。基于大数据与计算最优化的技术逻辑,医疗AI形成的治疗建议呈现出表象化的中立和客观形式,且会靶向推荐给医生和患者。这时AI系统推荐并排名治疗方案时,以一种对患者“居高临下”的态度呈现,是一种显性的家长主义作风。^[8]逐渐地,医疗AI的建议在临床实践中从辅助性地位跃迁,转向一种默认替代选项。尽管在现实决策操作中并未直接宣布剥夺医生与患者的选择权,但是面对医疗AI的算法推荐和建模方案时,人们可能会固化地接受系统提供的“最佳路径”。

2. 医患角色职责的转变

医疗AI的加入将医生的职责扩展至对技术理解与审查的综合把握。具体来讲,医生的角色由传统医疗实践中的主要决策者转变为医疗AI与患者之间的纽带。^[9]通过对医疗AI决策过程进行信息再表达,医生可将复杂算法转化为患者易理解的临床语言,并行阐释系统输出所固有的不确定性及适用边界。以知情同意为支撑的实践尝试,更有利于巩固医疗实践的信任基础。同时,医生需承担对医疗AI的审查职责。医疗AI本身携带系统性偏差,其输出的有效性严格依赖于临床情境。医生可对AI建议的准确性与稳健性等风险进行必要检视,从而防止技术性错误向临床风险的转化。当然,医生所承载的新角色与能力责任紧密相关,AI素养的不足将制约其在实践中的专业功能发挥。

患者的责任扩展至结合医疗AI进行自我管理。尤其是慢性疾病治疗中可穿戴设备的应用,传感设备可为患者提供持续的反馈,帮助患者进行自我健康管理。据统计,美国有120万糖尿病患者可能从AI技术中获益,在智能医疗设备的辅助下,患者可自觉改善生活方式并减少用药,([10], pp.225-226)这种即时的智能健康守护促进了患者福祉的实现。有学者指出,自我监测设备似乎对患者提出额外的要求,患者需对这些测量结果承担责任。^[11]此前,这种责任主要是由医生和护士负责,医疗AI的加入扩展了责任承担的范围,这也是医疗AI赋予患者角色责任转变的突出表现。在这个过程中,医疗AI家长主义的程度强弱与患者的主体性特质直接相关。如果患者对医疗AI和医学实践有基本的认知,那么患者对自主性发挥的弹性则相对较大。反之,医疗AI家长主义可能对患者造成潜在的技术权威压制。

3. 医疗AI家长主义的潜在正当性

在特定医疗条件下,AI家长主义具备一定的正当性与实践价值。其一,医疗AI拥有强大的数据基础和整合能力,可以在有限的时间内快速提供医疗决策建议。尤其是医生在面对复杂或者罕见的医疗场景时,可适当参考医疗AI提供的判例模型。此时,AI不是强家长

主义干预,而是作为一种技术补偿,弥补紧急情况下医生决策储备不及的困境,或能发挥医疗建议的辅助之效。其二,患者并非总是具备充分的知情条件和理性决策能力,在面对未知的医疗场景时,可能需要AI提供必要的引导服务和建议功能。前提是这种家长主义协助具备可选择性或可拒绝性,是建立在有限条件基础上的家长主义,若超乎患者自主性发挥的合理边界,技术建议则演变为家长式强制,影响医患主体性的发挥与个体权益的保障。

二、医疗AI家长主义的伦理困境

尽管上述讨论指出医疗AI家长主义具备部分现实性和正当性,但是医疗AI的介入也存在不可忽视的伦理风险。明确其中的具体表现与内在逻辑,才能对症下药。

1. 家长式AI决策影响自主性发挥

医学AI家长主义引发的伦理挑战之一即是对患者自主性的影响。自主权表明患者能够做出符合自身健康价值观的决策,同时免受不当操纵或胁迫,它包含基于信息进行推理与权衡并做出合乎逻辑的决策的能力。作为传统生命医学领域的核心原则,自主性的发挥受到意向性、理解及控制性因素的影响。([12], pp.109-112)

首先,患者的意向性是自主性发挥的关键因素。医疗AI家长主义建议通常以技术嵌入和单向推荐的方式呈现在患者面前,将患者设定为建议的默认接收者,即使该方案不符合患者医疗计划的预期。严格来讲,在医学背景下,任何医疗措施都不得违背具备行为能力的人的明确意愿,但此间,医疗AI的权威和自主性的不相容直接体现为患者意向性选择受限。其次,理解受阻是AI家长主义影响患者自主性的另一体现。患者对AI的作用及其输出可能无法完全理解,这会妨碍患者做出明智的决定,降低患者自主性的发挥和医疗决策的参与度。如果患者对自身医疗信息掌控不足,会直接违背知情同意原则并妨碍其做出理性决策。此外,医疗AI家长主义是自主性发挥的外在控制性因素。

鉴于医疗AI日渐成为“技术专家”并形成新的知识权威,患者对医疗AI建议的依赖性与日俱增,这成为操纵患者自主性的关键环节,限制其独立思考和解决问题的能力。([10], p.165)面对医疗AI应用的广泛性,可能存在这样的情况,即使患者不愿意,他们也可能迫于压力使用AI决策辅助工具。[13]

对医生而言,医疗AI引发的自主性干扰也不可忽略。毕竟人们天生就有技术依赖的倾向,尤其是个体认为技术比自己更加智能时,更易发生逃避思考和不愿承担责任的情形。([10], p.164)如今,嵌入式筛查辅助AI已在多家医疗机构部署,并逐渐展现出临床决策支持的潜力。其中,以乳腺X射线模型为代表的應用,其诊断准确性已被证实可与放射科医师相当甚至更佳的性能。[14]但若类似的医疗AI在筛查中逐渐占据主导,医生或会过多依赖医疗辅助AI而淡化自身的技能培养与专业自主性。[15]彼时,医疗AI家长主义作为新的决策力量,成为相较于传统医生权威之上的更高级的存在。

2. 计算偏好模糊真实医疗有利

患者的偏好和价值观是医学的三大支柱之一。当前,医疗AI可以根据技术逻辑对患者进行偏好预测,匹配患者在治疗时的决策意愿。某些情况下,AI甚至可以提供相对于替补决策者更为准确的预测,这可能是部分人群更倾向于选择AI的原因。[16]

不过,算法具备的排序优势可以直接决定用户所见的内容,以及呈现机器认为的重要方面。但问题在于,第三方难以对他者的效用作出精准判断,因为个体偏好排序往往取决于多重因素的互动。尤其对于无行为能力的患者来说,尚不清楚患者自身的真实偏好是什么,医疗AI可能无法做出利害攸关场景下符合实际且尊重患者的选择。[17]当处于紧张的临床情境时,患者面对少之又少的备选项,无形中被推入一种形式化的伪知情程序或“强制同意”。([18], p.31)且在自动化偏见(automation bias)的影响下,人们会倾向于将机器内容视为可替代的判断。[19]这种倾向可能导致临床人员在面对模糊信息或既有实践框架时,过度依

赖AI输出而忽视对其合理性的独立审视。譬如,商业化脓毒症预测模型(Epic Sepsis Model, ESM)被美国数百家医院部署,但是根据学者对27679名患者住院记录的研究,该系统不仅会发出误导性或重复性警告,而且还存在67%漏报率。[20]综合来看,这似乎并未为临床诊疗提供理想化的辅助。

另外,医疗AI输出的治疗偏好或许无法对复杂场景中的价值惯性,甚至会限制个人的选择自由,将自主权降低为广义模式而非个性化伦理需求。[21]一般而言,有利原则不仅强调行为主体增进患者福祉,还要求在具体实践中对潜在利益、风险及成本进行审慎权衡,以达到整体效用的最优化。([12], p.235)如果医疗AI的决策建议只是满足患者的偏好,则是一种机器理解的医学上的好,算法上仍缺乏仔细深入的现实性考虑。或言,决策的制定也不能脱离具体的价值情境,例如,在精神病学神经技术的应用中,由于无法理解社会、心理、生物和精神层面的复杂性,AI技术可能会简化实际风险, [22]由此形成的技术决策将妨碍医学进步的实现。

3. 机器不透明遮蔽医疗决策可靠性

AI重蹈医疗家长主义覆辙的重要原因,在于其内部工作原理的不透明性引发人们对算法知识的理解不足。[23]因此,当AI应用于医疗决策时,医护人员或许无法领会也不能解释它为什么会给出特定的诊断及建议。如前所述的脓毒症治疗系统,机器学习模型可从医院电子健康记录或持续监测中收集大量的临床数据,在症状出现前数小时预测患者患脓毒症的风险。但是这类医用AI假阳性率高且缺乏外部验证,加之有限的信息公开度,其中的安全性和有效性仍待商榷。[24]尽管也存在某些情况,算法能做出匹敌医生的可靠决策。但是综合来看,医疗AI家长主义决策的接受性堪忧,囿于机器的不透明性,这导致有些患者并不信任这项技术,或者更愿依赖医生的专业知识来做出治疗决策。

假设在推荐治疗方案中,人们可以适当地相信算法决策,但是这种输出的稳定性和安全

性尚不清楚。即使医疗AI决策符合患者的偏好，但若无确证的决策依据与明晰的判断链条，人们将无法充分评估和验证其输出的合理性。结果可能是隐形的医疗AI家长主义游走于技术黑箱，持续影响医疗决策的可信度与可靠性。

4. 自动化输出隐含医疗再歧视

医疗场景的诸多公正实践，往往呈现类似“抽签式”的不确定性。（[12]，p.288）医疗AI系统基于广泛的数据建模，容易复制和再现结构性偏见，并以自动化建议方式呈于医患面前。这种风险不只缘于算法本身的技术属性，也在于其决策结果往往被赋予“价值中立”的伪饰。其中衍生的依赖引发持续性的医疗偏见，如不能受到及时审视，可能会跟随未来决策基数的增加以至无限扩大。同样地，医疗数据来源的局限性也会引发AI决策中的再歧视。例如，海德堡研究人员设计了一个诊断黑色素瘤的算法，其训练集中使用了超10万张预标记的白人皮肤病变图像。尽管该算法某种程度上显示出相比专业医生的有效性，但是受限于样本的多样性，算法对其他肤色人群的诊断决策可能会存在偏见。（[10]，p.168）若推广应用，这种偏差经由医疗AI自动化输出，并施加于个体患者，无形之中扩大了歧视的广度和深度。

另外，囿于算法技术的多链整合，家长主义AI决策中的不公正难以具体归责至个体，呈现医疗决策偏见质询中的责任空档。尽管AI在医疗决策中的作用和责任发生新的变化，但若医疗AI作为决策主体，并不能对其具备风险的家长式决策后果负责，这也是当前医疗AI辅助诊断中的争议话题。而且，此间的正义侵蚀关联到深层次的决策结构，不同群体看似同等接受医疗AI的建议，受到家长式辅助，但是这种建议的匹配性和医疗有利原则的契合度大相径庭。还要指出的是，源于个体的知识基础和技术素养的差异，对医疗AI建议的理解和反馈机制存在偏倚，不能有效发挥机器建议的辅助作用，或者扩大不当家长主义建议的影响，构成一种新型的医疗实践不正义形式。

三、医疗AI家长主义的治理策略

鉴于医疗AI家长主义倾向引发的伦理挑战日益凸显，传统的技术治理与伦理约束难以应对AI家长主义带来的复杂问题。因此，需要重新思考应对医疗AI家长主义的核心路径。可以尝试患者赋权、重构医生的限定性权威、医疗AI技术的伦理升级并引入医疗民主，以一种多维协同的模式迎接已来的医疗AI时代。

1. 患者赋权：从默认接受到积极参与

赋权在医疗领域意义重大，是调和家长主义模式的可能路径。未来的患者应该更加主动，尽管在一些必要的情境下仍然离不开医生的诊断，（[18]，p.305）然而，公众对医生传统权威地位的绝对依附已不再如以往那般稳固。例如，在医疗AI技术的应用中，授权患者查看设备上的数据是其参与决策的重要体现。在技术引导的基础上，患者逐渐反思医疗专家或者医疗AI建议的可靠性并开始为自己行动，这种赋权即是个体拒绝家长主义的直接转变。这不仅意味着患者减少对专家的依赖，还展现患者积极主动寻求有利原则的期望，^[25]并开始积极管理自己的疾病与健康状况。一方面，应确保知情同意原则的真正实现，为患者提供自主性发挥的要件，带动患者共同加入决策活动中。可以采取确保在提供决策辅助之前始终获得患者的明确同意。包括对相关医务人员进行知情同意工作执行规范培训，告知患者知情同意的细节和参与的自愿性。另一方面，患者应具备一定的自主性与控制力，并且有充分的条件行使选择退出医疗AI诊断和治疗程序的自由。从专业化理论的角度来看，维护患者的自主性意味着让患者能够在既有条件中进行选择。这种建立在可及性和自主性基础上的赋权，才有利于患者更加主动地参与医疗决策。

2. 医生“权威”：家长式AI介入中的职责再转换

首先，医疗AI介入中的医生专家角色“限定”回归。尽管医疗AI家长主义现象促使医生角色发生变化，但是囿于AI技术发展中的潜在风险，需要重构医生在决策中的有限专家地位。算法建议应该支持医生的决策，而不应成

为最终决定项,将医疗AI的决策作为辅助性参考更具适当性,临床医生需保持他们既有的认识论权威,并且要兼顾对医疗AI负责任解释及推荐决策的考察。例如,定向实时早期预警系统(Targeted Real-time Early Warning System, TREWS)不仅被证实具备改善脓毒症患者临床指标的潜力,而且该系统采用去中心化的设置,不会通过弹窗和电话中断临床流程,医护人员可以自主决定是否及如何响应警报。^[26]毕竟,AI的本质在于模仿人类智能,可能无法完全具备专业医生独有的灵活性和创造性。在面对复杂或者罕见的医疗案例时,医疗AI家长式决策并非总能契合医学上的有利原则,医生应避免对算法诊疗的过分信任而产生自动化偏见,保持理性认知并提升AI素养,重塑作为临床专家的权威形象,适时做出有利患者的医疗判断。而且从现实出发,人们似乎对传统医生更加信任。一项关于患者和公众对医疗AI接受度的研究发现,虽总体上人们对AI持积极态度,但是参与者强烈倾向于将AI仅用作辅助工具,只有少数人认为这项技术应该取代医生。^[27]

其次,医生需要优化专家调解者角色及职责。医疗AI的决策受限于透明性问题,影响患者对其决策参与的认知度和接受性,部分人群对AI的使用存在相当大的抵触情绪,也可称之为“算法厌恶”。这种状态下,不论人们是否发现或知晓算法存在的不足,都可能出现直觉性的排斥心理。^[28]鉴于医疗AI与临床实践的结合已是大势所趋,如果AI的临床任务表现优异,医生有义务参考系统提供的建议或调整自身的判断,并负责任地向患者解释相关决策,^[29]促进患者对医疗AI的理解力和自主判断力,以发挥医疗AI的决策辅助功能。具体来讲,可以强化医生对AI系统的算法机制和医学推理逻辑的基本把握,尤其是明确医疗AI应用与诊断治疗系统结合时的相关功能要素。这些知识不仅关乎医生甄别AI建议的合理性边界,更影响医生对患者传达必要信息的解释力。如此一来,医生既可有效地履行己身义务,亦能精准地将关联内容告知患者并促进医疗福祉的最大化。

3. 技性升级:医疗伦理嵌入设计与修正评估

其一,优化医疗AI技术的伦理嵌入。在系统设计初期开展价值敏感分析,精准识别不同患者常见的伦理偏好及分歧,提升算法推荐决策与医学有利原则的多维对接,避免空泛的算法偏好引发的决策失误。有研究提出了“P4-LW”工具,一种在AI辅助下运行但受患者知情同意所规范的动态互动机制。该机制允许行为个体受控地暴露于假设的医疗场景来主动训练AI系统了解他们的偏好,以确保该工具直接从个人表达的价值观念中获得启发,而不是依赖外部的或可能具有误导性的数据源进行训练(相较于既有的P4模型而论)。^[30]此外,建立可调节的医疗AI模型,尤其注重凸显决策退出机制的设置,留给患者可自主选择的决策参与余地,弱化AI家长主义机制的程序设定。其中,应用程序需要为其建议的决策提供详尽说明,并允许患者具备推翻所有决策的选择自由。这不仅是医疗AI技术的伦理优化,更是保障患者自主性发挥的要件。

其二,注重医疗算法决策的修正性评估。医疗AI算法的局限性可能引发输出的不均衡,引入修正性评估对结果进行再校验与再校正,可适当提升其临床场景下的公正性。具体而言,修正性评估不仅要求对模型结果进行多维度分析,还需要结合外部验证机制与专家判断,对潜在的不合理偏差进行及时纠正。^[31]同时,修正性评估应当是动态且持续的过程,可随着临床环境的变化和医疗实践的需求而迭代。这种动态修正有益于减少医疗AI决策中隐含的偏倚,提升对病患群体的健康和权益保护。而且,将修正性评估纳入医疗AI的常态化治理,不仅有助于提高算法决策的透明度和可解释性,也能为实现医疗AI在公正基础上的发展奠定基础。

4. 医疗民主:协同合作的医患关系

在医疗AI的介入下,打破家长主义困局的方法之一即是从传统的医患二元关系模式中寻求突破,构建一种人工智能-医生-患者三元协同的民主决策关系。医疗民主的真谛在于,强化患者主动参与及医生协助指导,并通过技术手段加以调试,从而打破一般家长主义的笼

罩，促使医患双方转向平等合作的伙伴关系。（[18]，p.XII）在未来的医疗AI决策中，不仅医疗AI的绝对家长式作风不复，医生的“独断专行”也将弱化。智能医疗的协作进程促使患者从服从性和依赖性的权威中解放出来，并拥有相对自由和理性的选择权。至此，医生和患者的自主性与协作性将得到最大程度地发挥。

鉴于医疗AI家长主义作风从根本上与共享决策相悖，而且在完全自动化的环境中，很难确保自主权和共同决策能充分发挥作用。因此，为了避免新家长主义的专断风险，人们应该将自动化流程仅仅作为一种助益，而非成为自动化决策下的失控者。面对医疗AI应用的必然趋势，为了消解医疗家长主义决策中的潜在风险及危害，医疗实践需要走向一种“有限的家长主义”，这似乎是较为适合临床环境的家长主义选择。如此，有限的家长主义既可满足医疗AI的扩大化应用现实，也能最大程度地保护人类主体的决策自由。这样一来，AI的技术补偿和人类自主性发挥得到平衡，医生和患者都有充分表达意见的空间，是促进医疗民主的重要程序。

不过，医疗的价值标准是不断变化的，医疗民主并无一成不变的操作模式，理想的共同决策应是医生敏锐地感知患者的需求并按时间和环境的变化而及时调整方案。患者的声音可以被充分聆听和考虑，医生的决策远离武断式家长主义。或言，这是一种动态的医患相处模式，是有限的家长权威、自主性发挥和医疗AI技术设计的即时适配过程。循此逻辑，才可能构建一种合伦理的医疗民主模式，维护医生和患者的最大利益并提升整体医疗质量。

[参考文献]

- [1] Sunarti, S., Rahman, F. F., Naufal, M., et al. 'Artificial Intelligence in Healthcare: Opportunities and Risk for Future'[J]. *Gaceta Sanitaria*, 2021, 35: S67-S70.
- [2] Guan, J. 'Artificial Intelligence in Healthcare and Medicine: Promises, Ethical Challenges and Governance'[J]. *Chinese Medical Sciences Journal*, 2019, 34(2): 76-83.
- [3] Rubeis, G. 'Ethical Foundations: Medical Ethics and Data Ethics'[A], Rubeis, G. (Ed.) *Ethics of Medical AI*[C], Cham: Springer Nature Switzerland AG, 2024, 55-87.
- [4] Sirgiovanni, E. 'Should Doctor Robot Possess Moral Empathy?'[J]. *Bioethics*, 2025, 39(1): 98-107.
- [5] Spiekermann, S., Pallas, F. 'Technology Paternalism-Wider Implications of Ubiquitous Computing'[J]. *Poiesis Prax*, 2006, (4): 6-18.
- [6] Mc Dougall, R. J. 'Computer Knows Best? The Need for Value-flexibility in Medical AI'[J]. *Journal of Medical Ethics*, 2019, 45(3): 156-160.
- [7] Arnold, M. H. 'Teasing out Artificial Intelligence in Medicine: An Ethical Critique of Artificial Intelligence and Machine Learning in Medicine'[J]. *Bioethical Inquiry*, 2021, 18: 121-139.
- [8] Heyen, N. B., Salloch, S. 'The Ethics of Machine Learning-Based Clinical Decision Support: An Analysis Through the Lens of Professionalisation Theory'[J]. *BMC Medical Ethics*, 2021, 22(1): 112.
- [9] Triberti, S., Durosini, I., Pravettoni, G. 'A "Third Wheel" Effect in Health Decision Making Involving Artificial Entities: A Psychological Perspective'[J]. *Frontiers in Public Health*, 2020, 8: 117.
- [10] 贾格·辛. 未来医疗[M]. 希瑞译, 北京: 中信出版社, 2024.
- [11] Lucivero, F., Jongsma, K. R. 'A Mobile Revolution for Healthcare? Setting the Agenda for Bioethics'[J]. *Journal of Medical Ethics*, 2018, 44(10): 685-689.
- [12] 汤姆·比彻姆·詹姆士·邱卓思. 生命医学伦理原则[M]. 刘星等译, 北京: 科学出版社, 2022.
- [13] Sass, R. 'Equity, Autonomy, and the Ethical Risks and Opportunities of Generalist Medical AI'[J]. *AI and Ethics*, 2025, 5(1): 567-577.
- [14] 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.
- [15] Botha, N. N., Segbedzi, C. E., Dumahasi, V. K. et al. 'Artificial Intelligence in Healthcare: A Scoping Review of Perceived Threats to Patient Rights and Safety'[J]. *Archives of Public Health*, 2024, 82(1): 188.
- [16] Segers, S. 'What You Believe You Want, May Not Be What the Algorithm Knows'[J]. *Journal of Medical Ethics*, 2023, 49(3): 177-178.
- [17] Annoni, M. 'It Is Not About Autonomy: Realigning the Ethical Debate on Substitute Judgement and AI Preference Predictors in Healthcare'[J]. *Journal of Medical Ethics*, 2025, 51(7): 450-455.

- [18] 埃里克·托普. 未来医疗 [M]. 郑杰译, 杭州: 浙江人民出版社, 2016.
- [19] Mosier, K. L., Skitka, L. J., Burdick, M. D., et al. 'Automation Bias, Accountability, and Verification Behaviors'[A], Human Factors and Ergonomics Society. (Ed.) *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*[C], Los Angeles, CA: SAGE Publications, 1996, 40(4): 204-208.
- [20] Wong, A., Otles, E., Donnelly, J. P., et al. 'External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients'[J]. *JAMA Internal Medicine*, 2021, 181(8): 1065-1070.
- [21] Refolo, P., Sacchini, D., Raimondi, C., et al. 'Should Artificial Intelligence-Based Patient Preference Predictors Be Used for Incapacitated Patients? A Scoping Review of Reasons to Facilitate Medico-Legal Considerations'[J]. *Healthcare*, 2025, 13(6): 590.
- [22] Sauerbrei, A., Kerasidou, A., Lucivero, F., et al. 'The Impact of Artificial Intelligence on the Person-Centred, Doctor-patient Relationship: Some Problems and Solutions'[J]. *BMC Medical Informatics and Decision Making*, 2023, 23(1): 73.
- [23] Beisbart, C. 'Opacity Thought Through: On the Intransparency of Computer Simulations'[J]. *Synthese*, 2021, 199(3): 11643-11666.
- [24] Bignami, E. G., Berdini, M., Panizzi, M., et al. 'Artificial Intelligence in Sepsis Management: An Overview for Clinicians'[J]. *Journal of Clinical Medicine*, 2025, 14(1): 286.
- [25] Segers, S., Mertes, H. 'The Curious Case of "Trust" in the Light of Changing Doctor-patient Relationships'[J]. *Bioethics*, 2022, 36(8): 849-857.
- [26] Adams, R., Henry, K. E., Sridharan, A., et al. 'Prospective, Multi-site Study of Patient Outcomes After Implementation of the TREWS Machine Learning-based Early Warning System for Sepsis'[J]. *Nature Medicine*, 2022, 28(7): 1455-1460.
- [27] Young, A. T., Amara, D., Bhattacharya, A., et al. 'Patient and General Public Attitudes Towards Clinical Artificial Intelligence: A Mixed Methods Systematic Review'[J]. *The Lancet Digital Health*, 2021, 3(9): e599-e611.
- [28] 杜严勇. 厌恶算法还是欣赏算法? ——人工智能时代的算法认知差异与算法信任建构 [J]. *哲学分析*, 2022, 13(3): 151-165; 199.
- [29] Amann, J., Vayena, E., Ormond, K. E., et al. 'Expectations and Attitudes Towards Medical Artificial Intelligence: A Qualitative Study in the Field of Stroke'[J]. *PLoS One*, 2023, 18(1): e0279088.
- [30] Milian, R. D., Bhattacharyya, A. 'Beyond Presumed Autonomy: AI-assisted Patient Preference Predictors and the Personalised Living Will'[J]. *Journal of Medical Ethics*, 2025, 51(7): 1-2.
- [31] Ueda, D., Kakinuma, T., Fujita, S., et al. 'Fairness of Artificial Intelligence in Healthcare: Review and Recommendations'[J]. *Japanese Journal of Radiology*, 2024, 42(1): 3-15.

[责任编辑 李斌]