

您所在的位置： 首页 >> 新闻中心 >> 行业动态

医疗人工智能伦理若干问题探讨

作者：李明 等 发布时间：2020-01-15

人工智能在医疗行业领域的应用与实践带来了新的伦理问题的挑战。新的人工智能技术可能对患者产生潜在的风险，这些因素主要来自算法偏见、隐私安全、医生接受度、贵权和技术滥用等，当前医疗领域也面临着缺乏医疗人工智能伦理规范和审核机制。医疗人工智能未来对整个医疗健康体系势必产生重大的影响，有助于医疗健康体系的重塑，促进公民的健康民主。期待行业组织或者政府制定医疗人工智能伦理规范。

1.引言

人工智能（Artificial Intelligence，AI）已经在医疗健康领域进行了广泛性的探索性应用，部分已经形成商业化的使用。人工智能将产生广泛的影响，显著地影响医学实践，改变患者体验和医生的日常生活。与此同时，AI先进的技术也带来一些列的新的伦理挑战。医疗伦理自古以来就贯彻整个医学发展的各个阶段，人工智能作为创新性的技术，对医疗健康体系势必产生重要的影响。本文就医疗人工智能伦理若干问题进行探讨。

2.当前医疗人工智能伦理的主要问题

2.1. 算法偏见对患者安全潜在风险因素

涉及患者安全的风险主要表现为 AI 提供的辅助诊断、治疗建议以及操作可能误导医生进行错误的决策，直接或者间接对患者产生不同程度的机体伤害。这种危害从技术上主要来自系统的算法偏见（Algorithmic Bias）与漏洞。算法偏见指的AI系统中产生不公平结果的系统性和可重复性错误。偏差可能是由于许多因素造成的，包括但不限于算法本身的设计、数据编程、收集、选择或用于训练算法的方式相关的意外或非预期的使用或决策。计算机产生的偏见几乎无处不在 偏见的原因往往不在准则中，而是在更广泛的社会过程中。到其结果中的无意识和制度偏差。算法偏见主要分为“数据驱动的偏见”、“人为造成的偏见”与“机器自我学习造成的偏见”三类别。

“数据驱动造成的偏见”指由于原始训练数据存在偏见性，导致算法执行时将歧视带入决策过程。鉴于算法本身不会质疑其所接收到的数据，只是单纯地寻找、挖掘数据背后隐含的结构和模式，如果人类输入给算法的数据一开始就存在某种偏见或喜好，那么算法会获得的输出结果也会与人类偏见相同。

例如图像识别在进行机器学习和建立学习算法的过程中，存在潜在的偏差。训练图像往往来源于具体某个合作医疗机构的数据，其图像标识过程中带有鲜明的该医疗机构的诊断特色和习惯。

此外，算法概念化的框架本身包含了工作组的主观假设，这种基于“主观”上的数据结果成为所谓“客观”的数据输入。尽管国际影像厂商设备的图像质量趋同同质性，但更多的不同厂商影像也存在不同的设备型号、数字化成像质量的程度上的差异，同时模型可能造成或加剧按年龄、性别、种族、地区或其他受保护特征划分的医疗保健人群不公平现象。这些因素导致应用的范围存在偏差。

人为造成的偏见是指算法设计者为获得了某些利益，或者为了表达自己的一些主观观点而设计存在歧视性的算法。这里包括了主观和客观的偏见，一方面由于人和技术本身问题造成的，另外一方面也有人刻意而为之。斯坦福大学有学者认为，用于医疗保健应用的人工智能开发者的价值观可能并不总是与临床医生的价值观一致。例如可能会有诱惑，引导系统采取临床行动，提高质量指标，但不一定是为患者护理。这些算法可能会在被潜在的医院监管机构审查时，扭曲为公众评估提供的数据。

而也有人为的偏见是来自医学本身特点，尤其在治疗方案上有着区域、学派的分歧，治疗方案选择上存在人为的偏见，尽管询证医学已经成为主导，但具有主观色彩的经验医学在临床实践上仍有效。很多时候，没有治疗唯一的“金标准”，这也就是为什么经常出现等级性质的方案推荐，有着专家共识、临床指南、临床规范的不同的标准。此外人为的偏见，也有可能来自系统开发机构工程师伦理知识的匮乏，缺乏行业专业领域知识。

机器自我学习造成的偏见，随着算法复杂程度的日益提高，通过机器学习过程形成的决策越来越难以被解释人工智能内部的代码、算法存在“黑箱”，导致无法控制和预测算法的结果，而在应用中产生某种不公平倾向。

2.2.患者隐私安全性泄露风险

进入信息化时代后，患者个人健康信息数据（PHI，Personal Health Information）一直被关注，亦有相应的行业规范和法规政策进行管理。随着数据搜集、机器学习、人工智能等技术的使用，个人信息泄露的情况频繁发生。个人隐私保护、个人健康信心重要性日益凸显。当前对于医疗 AI 应用涉及患者的隐私问题、自我保护策略和可接受性尚还没有清晰的策略。数据更多的泄露来源是在为 AI 进行训练的患者数据。其中人体基因组数据的泄露引起潜在的生物危害和商业利益更为突出。有报告显示，公众对个人隐私泄露问题的顾虑占到调查人群的59.72%。

2.3.医生对医疗 AI的接受性

从技术角度而言，医生不了解人工智能算法的局限性或范围，这种不透明的“黑箱”效应从而产生了盲目信仰或怀疑的可能性，不信任的占据多数。由于医疗 AI 当前还处于相对早期的临床应用阶段，无论医生还是民众对医疗 AI 的信任程度未达到期望。很多医疗 AI 系统没有很好的在整合医疗工作流程中，其医疗 AI 解决方案并没有贯穿整个医疗全流程，部分操作反倒增加医生现有工作负担，其带来的价值没有很好体现。此外，医生对医疗 AI 新技术的各种顾虑，不自觉地表现出排斥和质疑的心理，某种程度延缓了新技术的迅速推广实施。对于新技术的接受程度上，具有海外进修培训经历的高学历的中青年医生对 AI技术应用更持有理性的态度，年轻医生好奇心居多，同时也有部分老医生对新技术持有质疑和排斥的心理。此外，在医疗不同场景下的应用，也导致对 AI 的评价与接受度不同，有学者研究在医学培训教育上医生对 AI 的作用认可程度最高。

2.4.医疗贵权问题

医生对 AI 系统临床应用另一方面的顾虑是谁对医疗 AI 提供的决策最终负责，如果出现错误，谁来承担后果。当前阶段，医疗 AI 的最后结果仍需要人工校验审核，医生承担由此对患者诊疗结局的责任。未来临床广泛引入医疗 AI 后如何实现问责，现在仍不明确。

当人工智能应用结果导致临床纠纷、人类伦理或法律的冲突问题时，是否能够从技术层面对人工智能技术开发人员或设计部门问责。并在人工智能应用层面建立合理的责任和赔偿体系，保障医疗 AI 在临床发挥其应有的价值，而不至于陷入尴尬困境。

2.5.医疗 AI滥用导致过度医疗与潜在风险

医疗AI在特定领域特别是影像识别上对已经高于人类平均水平，而且就有较高的效率，作为工具在临床上也有滥用的风险，从而导致过度医疗。根据某城市三级医院某院网站公开报道，自医院引入肺结节影像识别 AI 技术后，肺结节筛查人次三年中增长73%，筛查直接带来的收益增长 81%，加上手术治疗，医院总收益增长100%。随着效率的提升，医院业务水平得到极大的提高。从公开的数字看，AI 给医院带来丰富的收益回报。但从所在地区的卫生统计年鉴来看，所在地区的非传染性疾病的疾病谱分布情况未发生明显改变，而该院肺部结节患者和手术量短期大幅增加可以质疑利用 AI工具进行了过度医疗。这种过度医疗无疑对患者身体上和经济学上都带来伤害。在NLST（National Lung Screening Trial）研究中，CT筛查组中96.4%的阳性结节为良性，我国农村肺癌早诊早治项目的假阳性率也较高。过高的假阳性可能导致过度诊断、过度治疗、医疗资源的浪费及增加受检者焦虑心理。

2019 年 7 月 17 日美国神经网络公司Neuralink在旧金山举行新闻发布会，宣布脑机接口系统获得新突破。此公司研发“像缝纫机一样”的机器人，这个机器人可以将超纤细的线植入大脑深处。目的是将人工智能直接植入人类大脑皮层以提高人类智能水平（即脑机接口技术），Neuralink的初衷在于帮助人类解决脑部疾病。目前研究人员已经在猴子身上进行实验，让猴子能用大脑来控制电脑。根据Neuralink官方公布的实验信息，他们把老鼠头部连接到USB-C端口，通过电线将老鼠的想法传到计算机。当软件分析其大脑时，可以通过扬声器听到老鼠神经元发射的信号。计划明年用于人体试验。从理论上，这种技术可以实现双向控制，也就意味着计算机可以控制大脑以至于控制人类的行为，这无疑会产生很大的风险。

2.6.医疗伦理审查缺乏

国家药品监督管理局出台了对于医疗 AI 的相关指导规范，并已启动了认证流程，从范围、风险、临床实验上进行了规定。医院伦理委员会主要负责本医疗机构中药物和医疗器械临床试验、器官移植、涉及人的医学研究、相关技术应用、动物实验以及医疗管理等活动的医学伦理学审查。目前在医院内实际的应用还未有对 AI 相应的伦理审查机制，有医院参照药物临床实验引入的伦理审查机制对医疗 AI 引入进行伦理审查，很难做到完全适用。

在实践中，有的伦理审查委员对于“不伤害”狭义地理解为不伤害身体，对风险的认识仍停留在仅关注生命健康、安全方面，忽视个人数据被不正当处理、个人信息泄露、侵害隐私权所带来的社会心理风险。更多医院对医疗 AI 引入并未经过伦理审查。也有医院在对患者应用医疗 AI 系统时尝试用患者知情同意书来规避一些潜在风险。

3.医疗人工智能伦理原则

3.1.医学伦理原则

医学伦理基本规范沿袭《希波克拉底宣言》直至现代的《日内瓦宣言》，其基本原则包括对患者的不伤害原则、有利原则、公正原则和尊重原则。任何医疗新技术在医疗中应用都基于医疗伦理道德，在伦理范围内规范医疗实践的准则和秩序。其新技术最终目的是为了人类健康福祉服务，而非对人类的伤害。

3.2.人工智能的伦理原则

众多产业组织在人工智能伦理上制定了规范原则。目前，两个影响较为广泛的人工智能伦理共识：《阿西洛马人工智能原则(Asilomar AI Principles)》和国际电气电子工程师协会(IEEE, Institute of Electrical and Electronics Engineers)组织倡议的《人工智能伦理标准》，鼓励技术人员优先考虑道德规范自主和智能技术创造中的考虑因素。很多国际巨头如：IBM、Google、Microsoft、Intel 等都制定了AI 的产品原则和伦理规定。这些原则其大部分都涉及：安全、透明、保护隐私、防止滥用等，核心都是充分维护人类根本利益。可以看出都是以人为核心理念。

3.3.医疗人工智能伦理规范

目前美国医学会（AMA，American Medical Association）对医疗 AI 明确提出了规范，特别强调促进精心设计、高质量、经临床验证的医疗保健人工智能的发展，即：a.根据以用户为中心的，特别是针对医师和其他医务成员，进行最佳实践的设计和评估；b.透明；c.符合引线标准的再现性；d.识别并采取措施解决偏见，避免引入或加剧医疗保健差异，包括对弱势群体测试或部署新的人工智能工具时；以及e.保护患者和其他个人的隐私利益，并保存信息。

一些中国医疗行业协会也成立了人工智能分会，对医疗 AI 伦理规范上也做了探索。但仅有中国医师协会超声医师分会发布了《中国超声医学人工智能行为准则：北京宣言》，从“制定规范、科学管理；实现医工结合、促进转化；以临床为中心，使患者利益最大化”3个方面 13条细则阐述了准则。

目前在跨学科领域的医疗人工智能还少有对伦理上的研究与探索，推测其原因在于医疗 AI 还尚未在临床得到广泛的应用，临床的迫切需求不够强烈。综合来看，医疗人工智能的伦理规范，仍是以人为核心，强化患者安全，保护患者隐私，AI 技术要透明，防止滥用，最终目的是促进疾病恢复，维护人类健康，实现健康民主。

4.对医疗健康体系的影响

尽管医疗 AI 不可能取代医生，但是医生相当多的重复性、规律性的诊疗工作可以被AI 所代替。长远来看，人工智能将产生广泛的影响，彻底改变医学实践，改变患者体验和医生的日常生活医务人员职业形态必将发生改变。这就对医生的临床技能提出了更高的要求，更加注重对患者的人文关怀。而相当一部分医生面临职业的严峻挑战。

当医疗 AI 应用逐渐向消费者端渗透。主要表现为个人、组织更加便捷地获取各专业领域及生活方式最佳适宜决策的建议，带来开放、可及、民主式的医疗，给公众和医生更多机遇，也弥补之前医患之间“信息不对称”的鸿沟，患者可以更加积极主动地参与自身的诊疗过程。实现了“健康民主”从而对整个的医疗健康体系产生重大的变革。

5.策略与期望

为医疗人工智能的伦理创造“法律”或“规则”是具有挑战性的，因为伦理道德界限很难教授软件系统或者机器，而且医学和相关影响因素本身的复杂性也决定了医疗AI伦理规范定义的艰巨性。更核心本质是人类的道德与伦理的规范性实现问题。

医疗人工智能系统发现、理解并指出人类在决策过程中的不一致性时，它们还可能揭示出人类自身偏见、狭隘和认知偏见的方式，反过来促进人类采取更公正或平等的观点。形成共同价值观的过程中，人类可能比人工智能进步得更多。

利用技术工具在人工智能中建立道德规范，部署人工智能的人面临的问题之一是深度学习和神经网络的黑盒子性质。这使得建立透明度和检查偏见变得困难。作为技术的提供方，越来越多的公司正在部署技术，搭建平台来帮助解决这个问题。例如IBM的AI OpenScale、开放源码工具，以及来自人工智能初创企业的解决方案，可以提供更大的透明度，并检查偏见。在实现上，要确保创建道德的治理结构并确保人工智能系统的责任，明确的角色和结构，将人工智能道德责任分配给关键人员和团队，并赋予他们权力。利用调整现有的治理结构，在特定的团队中构建问责制来实现。例如，组织中现有的道德主管(例如首席道德官)也可以被委托负责研究人工智能中的道德问题。建立内部和外部委员会，负责合乎道德地部署人工智能，这些委员会是独立的，因此不会受到急于部署人工智能的人的压力。

此外，对于使用的机构，医院伦理委员会需要对医疗 AI 伦理的制度完善，全方位的对产品和方案进行评估，既要保证安全不伤害也要顾及技术进步性的意义。

2019年7月，中央深化改革委员会召开第九次会议，审议通过《国家科技伦理委员会组建方案》。会议指出，科技伦理是科技活动必须遵守的价值准则。组建国家科技伦理委员会，目的就是加强统筹推进规范和指导协调，推动构建覆盖全面、导向明确、规范有序、协调一致的科技伦理治理体系。要抓紧完善制度规范，健全治理机制，强化伦理监管，细化相关法律法规和伦理审查规则，规范各类科学研究活动。这将是国家最高层级的组织机构，指导所有科技类的伦理规范。

对未来展望，一方面期望中华医学会、中国医院协会在国家级学会层面上制定医疗 AI 伦理规范，以进行统一的指导。另外一方面对于医疗 AI 生产者与使用者有必要进行教育和培训了解医疗 AI 的伦理挑战，同时确保患者充分了解如何使用这些工具来决定其诊疗。更好的利用医疗 AI 工具为人类提供健康服务。

本文节选自《医学与哲学》杂志2019年11月刊

CHIMA课题《医务人员对医疗人工智能（AI）接受度与行为反应研究》

作者：李明、李昱熙、戴康、李小花

摘要、关键词、参考文献略



医疗信息化相关书籍上架，请点击此处查看

上一篇： 2019年美国HIT从业人员的5大热点岗位

下一篇： 医疗信息化新闻一周回眸