

## 生物医学领域人工智能应用的伦理问题

郭旭芳, 刘 辉\*

(中国医学科学院 北京协和医学院 人文和社会科学学院, 北京 100730)

**摘要:**人工智能在生物医学领域的应用正在加速发展,在公共卫生、临床诊疗、健康管理方面展露出广阔的发展前景,同时也引起了许多潜在和现存的伦理困境。目前,讨论比较集中的伦理议题有:1)人工智能的道德地位;2)人工智能应用引起的各种公平性问题;3)个人权利与公共利益之间的张力问题。本文试图对上述议题进行初步的探讨,从平衡的观点出发,提出若干原则供学术界共同讨论和批评。

**关键词:**道德地位;公平性;公共利益;隐私;平衡

中图分类号:R52 文献标志码:A

## Ethical issues in artificial intelligence application in biomedicine

GUO Xu-fang, LIU Hui\*

(School of Humanities and Social Sciences, CAMS & PUMC, Beijing 100730, China)

**Abstract:** The application of artificial intelligence in the field of biomedicine is accelerating, showing broad prospects in public health, clinical diagnosis and health management. However, it will also cause many potential and existing ethical dilemmas. At present, the ethical issues that are attractive are: 1) the moral status of artificial intelligence; 2) fairness issues caused by artificial intelligence applications; 3) the contradiction between individual rights and public interests. This paper a discussed the above issues. From a balanced point of view, several principles are proposed for discussion and challenges from the academic community.

**Key words:** moral status; fairness; public interest; privacy; balance

英国学者托马斯·霍布斯(Thomas Hobbes, 1588—1679)曾在《利维坦》一书中提到可以制造一种“人工动物”,人工智能(artificial intelligence, AI)的雏形诞生<sup>[1]</sup>。1956年,美国达特茅斯会议上正式确立了“人工智能”这一术语,自此打开了人工智能领域的大门。人工智能早期的定义是:一种大部分工作集中在模拟或复制人类智能的工具<sup>[2]</sup>;人工智能是一种获取并表达知识,并且运用知识的学科。从医学和生物学的角度来看,人工智能也被称为

“一般智能”,指具有应对普遍环境的能力和智力<sup>[3]</sup>。如何定义人工智能尚无定论,但是可以总结出以下3个特点:1)根据人类活动研制出来的数字化系统;2)模拟人类的某些行为;3)执行人类能够进行的智力活动。

人工智能在生物化学、计算机技术、医疗卫生、经济金融等诸多领域发展迅速,给人类的生活带来极大便利,逐步成为国际竞争的新焦点,但同时人工智能发展的不确定性也带来新挑战,这些挑战不仅

收稿日期:2019-11-12 修回日期:2019-12-23

基金项目:北京协和医学院“中央高校基本科研业务费”(3332018134)

\*通信作者(corresponding author):liuhui@pumc.edu.cn

局限于技术层面,更多的是伦理方面的。生物医学领域是人工智能技术首先应用的领域之一,医学研究以及医疗保健对于全人类有着至关重要的作用。然而,有研究证明,人工智能设备在设计之时往往没有任何明确的伦理考虑<sup>[4]</sup>。因此,人工智能在医学中的伦理挑战就成为了不可忽视的问题。本文试图通过对人工智能现存矛盾以及潜在问题的梳理,探讨如何平衡人工智能带来的利益以及伦理风险。

## 1 人工智能道德地位和责任问题

多年来,医疗机器人已经被用于外科手术以及医疗保健中。外科手术中机器人的运用对于头颈部,心脏和泌尿外科等科室都是有益的<sup>[5]</sup>。这说明医疗领域人工智能的合理应用使得医疗专业人员的工作更有效率,但是医疗机器人在扩大运用范围,提高医院效率的同时对医生的道德主体地位提出了挑战,此外人工智能的使用也增加临床诊疗中的不确定性风险。

在手术中,如果医疗机器人引发了法律追溯责任问题以及伦理责任问题,那么谁应该因医疗机器人造成事故承担责任?例如,最近的神经技术可以使得大脑信号变成输出设备执行所需动作的命令<sup>[6]</sup>。在这种框架下,当不确定性风险发生时,责任主体应该如何判定?一个主要或仅由一个设备产生的行为的责任是否应该最终归于人类?在伦理条款,现行的“医疗道德守则”规定:专业的责任取决于患者的医生,护理等任意一方涉及医疗责任,导致对患者造成任何损害,主管医生应该按照医疗程序承担主要责任<sup>[7]</sup>。但人工智能与一般医疗设备有所区别,常常看上去是一种“自主行为”,医疗机器人能否成为责任主体,它能否拥有道德主体地位?

国际学术界针对如何判定机器人的道德地位这一问题已经讨论了10年,14个欧洲国家的156名人工智能专家曾联名表明态度——授予机器人法人资格“不恰当的”。大多数学者认为现阶段的人工智能机器人尚且只能属于工具范畴,因为机器人并不具备自主意识,缺乏理性思考。然而,图灵测试的结果意味着复杂的人工智能可能具有意识<sup>[8]</sup>。如果人工智能科技进一步发展,发展后的人工智能系

统极有可能具备自主制定行动策略、自主做出道德选择的能力。这一局面出现时,人类与智能机器人的道德地位又应该如何界定?

笔者认为对于人工智能自主性的发展应该有一定的限制。人工智能可以分为强人工智能(Strong Artificial Intelligence)和弱人工智能(Weak Artificial Intelligence),对于弱人工智能来说,不应具有道德地位;至于强人工智能,即具有自主意识的设备,本身就属于争议,强人工智能不仅在医学行为中道德地位难以确定,这一技术对人类的地位本身可能造成威胁,从平衡的观点出发,在完全了解强人工智能的利弊之前,应该谨慎对待此类技术在医学领域的应用。技术的发展应该控制好“人类”与“人工智能工具”之间的平衡这一问题,在医学领域的应用要控制系统行为的不可预知性和不确定性。

基于以上观点,人工智能在生物医学领域的应用仍然是工具属性,医疗机器人不具备道德地位,即不能独立承担事故的责任,也没有资格代替医生在医疗行为中的角色,而强人工智能及其应用应予以慎重对待。如,在发生医疗事故致使患者利益受到损害时,是应该由研发者还是由操作者来承担后果呢?笔者认为,应该针对事故情况具体分析,若是因为智能机器的程序故障导致医疗伤害,则主要责任应由研发者承担,操作者承担次要责任;若是因为操作不当导致医疗事故,损害患者权益,则应该有操作者承担责任。

## 2 医疗人工智能引发的公平性问题

人工智能有着惊人的潜力,可以加速生物学和医学的科学发现,并改善医疗保健结构,增进人群以及个人的健康,但是人工智能的应用也可能会加剧人群中现有的医疗保健差距。人工智能医疗的成本较为昂贵,因而其主要受益群众也仅限于少数,此种前提下,医疗资源分配不公平现象将会更为突出。除此之外,人工智能算法所基于的数据本身可能包含有意想不到算法偏见<sup>[9-10]</sup>,不可控的产生年龄、性别、种族方面的歧视,加强现有的社会不平等形式。例如:大脑植入是一种神经科学方面治疗的新手段,可是只有有能力支付的人才能被提供这种治疗方式<sup>[11-12]</sup>。在此情况下,应该如何平衡健康公平性的问题?

笔者认为,可以从以下3个方面进行考虑:第一,要实现人工智能健康效益和成本之间的平衡。在临床领域应用的技术应该是适宜的技术,对人工智能在生物医学领域的大规模应用需要充分论证,技术已成熟、有利于改善人的健康,且成本应具有可接受性,才能进行推广。第二,实现个人利益与社会利益之间的平衡。坚持以人民健康福祉为中心,将应用的人工智能技术产品分类,对人的健康具有重大意义的基本医疗手段国家社会应该予以承认,纳入医疗保障系统,以尽可能低的价格满足居民健康需要;另一方面,对于促进人类健康范围较窄且成本较高的人工智能医疗手段可规定由个人承担。第三,坚持“不伤害”的原则。对于可能引发健康伤害风险的人工智能需要控制其使用的范围,维护好生命健康权这一基本权益,特别是对社会弱势群体应该予以尊重和保护。

### 3 个人权利与公共利益之间的张力问题

#### 3.1 自由意志和公共利益之间的张力

当今网络技术时代,人工智能设备不可避免运用于数据采集、整合、储存、分析,而人工智能也需要使用个人健康信息来完成机器学习,提高诊断,治疗和辅助研究的能力。在医疗领域,医学研究者可以通过流行病学分析,鉴定某一疾病与基因、环境的关系,从而达到对于疾病更好的预防,结果可以使得全社会受益。但是人工智能所需的个人健康信息是典型的医疗大数据,因此在利用医疗大数据时,不论是基于对于数据提供者的人格尊重,还是基于保护其利益,都应取得充分的知情同意,但是大数据采集在充分的知情同意程序前却面临困境。例如,欧洲现行的《一般数据保护条例》(General Data Protection Regulation,简称GDPR)提出要求:对个人数据(personal data)的每一次加工或分析(processing or analysis),都要获取数据对象—专门的、知情的、明确的同意。这个条例下,研究人员会面临着两大压力:数据对象难于寻求以及严重的经济负担,导致医学研究受限、整个产业体系发展难以推进。

#### 3.2 个人隐私与公众利益之间的张力

生物医学大数据是人工智能在医学领域运用的基础,患者健康信息包括个人社会信息、病例数据甚至影像、磁共振结果都可以作为人工智能(机器学习)

习)的必要资料,若没有充足的病例数据,人工智能的成熟不可想象,由此产生的健康效益难以保证。有研究表明机器诊断能力不输于放射学专家,甚至能超越后者<sup>[14]</sup>。基于对个人隐私的保护,一般的解决思路便是去识别化(de-indentification),即把个人身份信息从数据中消除。但是,近年来,随着人工智能系统和大数据识别技术的发展,又出现了再识别化(re-indentification)这一技术,通过大数据和人工的手段把分散在各处的信息整合并重新生成个人身份信息,从而使去识别化的手段失效,换言之,即便去识别化手段是完全有效的,也没有就去识别化的具体标准作出统一规定。因此可以说隐私保护和人工智能发展之间、社会健康福利之间的张力问题一直存在,而且可能更加尖锐。

为生物医学信息库大数据下平衡个人权利以及公众福利的张力问题提供以下几个解决思路:1)权利的集中与责任的集中相平衡。个人的数据集中形成大数据、大数据推动人工智能的发展,实际上是个人权利向中心节点的集中(让渡),这个节点可以是政府、高校,也可以是医院,甚至是公司,机构节点集中保管和使用大规模个人数据就必须承担相应的数据安全、隐私保护、规范使用等责任。2)政府部门承担监管责任。因为大数据设计人数众多,应该有政府有关部门制定大数据管理制度,防止个人数据滥用现象的发生。3)个人应提供医疗数据或者其他健康数据,并且应具有从相关人工智能服务中获益的资格。人工智能因个人数据而获得提升,理应以服务于公众健康为目的。利用了个人医疗或者生物数据应该以并且对于所使用数据的个人予以一定程度上的回馈,比如,提供健康咨询等。

尽管人工智能在生物医学领域的应用上面临着很多技术和伦理的挑战,但是我们应该肯定一点——人工智能的合理运用可对现代医学发展起到推波助澜的作用。现阶段对于人类最大的挑战在于:如何定义人工智能设备的地位,并保障其安全性、有利性、保密性、公平性等问题。针对文中所提及的现存以及潜在的伦理困境,提出一些人工智能运用的原则:

1)以“不伤害”原则作为一切人工智能应用的底线,寻求人工智能合理使用的平衡点。人工智能

设备设计的初衷是“以人类为中心”,更好地服务于人类社会,无论是在身体健康层面上,还是在精神心理层面上,如若人工智能对人类造成了伤害,则此类设备应该重新进行风险评估并且严格监管。

2)明确人工智能的工具属性,肯定人类的道德主体地位。人工智能无权作为一个独立的道德体存在,更不能作为责任的承担主体。医学领域的人工的目的是为人类健康服务的,这是人工智能存在的价值,不可以为了经济利益而背离这一价值。

3)保持适度的谨慎,人工智能潜在风险难以预测,对于人工智能的运用应该“取长避短”,不应该畏惧它,但也不能肆无忌惮的滥用,反对技术潮热,伦理的研究始终应该走在技术的前端。

4)政府对于人工智能系统应尽监测责任,机构应尽保护的责任。对于生物信息数据库,政府部门应该制定一个严格的监管机制,以防止医疗数据的滥用;通过制定相关政策来保障个人权利和公众健康利益之间的平衡。

## 参考文献:

- [1] 周程,和鸿鹏. 人工智能带来的伦理与社会挑战[J]. 人民论坛, 2018,2:26-28.
- [2] Shortliffe EH. Artificial intelligence in medicine: weighing the accomplishments, hype, and promise[J]. Yearb Med Inform, 2019,28:257-262.
- [3] Park WJ, Park JB. History and application of artificial neural networks in dentistry[J]. Eur J Dent, 2018,12: 594-601.
- [4] Ienca M, Wangmo T, Jotterand F, *et al.* Ethical design of intelligent assistive technologies for dementia: a descriptive review[J]. Sci Eng Ethics, 2018,24:1035-1055.
- [5] Sepehrpour AH, Garas G, Athanasiou T, *et al.* Robotics in cardiac surgery[J]. Ann R Coll Surg Engl, 2018,100: 22-33.
- [6] Keskinbora KH. Medical ethics considerations on artificial intelligence[J]. J Clin Neurosci, 2019,64:277-282.
- [7] Ishihara K, Fukushi T. Introduction: roboethics as an emerging field of ethics of technology[J]. Account Res, 2010,17:273-277.
- [8] Warwick K, Shah H. Passing the turing test does not mean the end of humanity[J]. Cognit Comput, 2016,8: 409-419.
- [9] Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy[J]. J Med Internet Res, 2019,21: e13216. doi: 10.2196/13216.
- [10] Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges[J]. PLoS Med, 2018,15:e1002689.doi: 10.1371/journal.pmed.1002689.
- [11] Keskinbora KH, Keskinbora K. Ethical considerations on novel neuronal interfaces[J]. Neurol Sci, 2018,39: 607-613.
- [12] Cheung EH. A new ethics of psychiatry: neuroethics, neuroscience, and technology[J]. J Psychiatr Pract, 2009,15:391-401.
- [13] 李欣菱,郭芳芳,周振,等.基于深度学习的人工智能胸部CT肺结节检测效能评估[J]. 中国肺癌杂志, 2019, 22:336-340.