

《元宇宙医学》第一届编委会

荣誉主编

樊 嘉

杨雄里

钟南山

主 编

白春学

顾建英

副主编

陈荣昌

高承实

贾泽军

刘 雷

连 勇

潘 毅

宋振举

王伟林

杨达伟

Niels H. Chavannes

Charles A. Powell

Christoph Thüemmler

编 委（按拼音排序）

白 莉

白春学

曹 鑫

陈 颖

陈大明

陈荣昌

陈智鸿

方金武

甘舜轩

高承实

葛海燕

顾建英

何 萍

胡 洁

胡伟国

胡 征

黄翔宇

贾泽军

姜 杰

金 晶

冷 非

李立亮

连 勇

刘 雷

陆俊羽

马礼兵

潘 毅

漆 远

沈 涛

盛 斌

施国明

宋海涛

宋元林

宋振举

苏士成

孙 湛

王 健

王 凯

王 洵

王伟林

王悦虹

韦 球

魏益平

辛弘毅

薛新颖

杨 冬

杨 丽

杨达伟

杨志刚

于广军

张纪阳

张宇鸣

张玉侠

张子强

赵列宾

赵明明

赵郑拓

周建军

朱智勤

邹 最

Niels H. Chavannes

Charles A. Powell

Christoph
Thüemmler

Lian Wu

Emeli Zhang

Shangming Zhou



Metaverse in Medicine 元宇宙医学

ISSN 3006-4236



元宇宙医学

二〇二五年九月 第二卷第三期



2025 年 9 月 第 2 卷第 3 期

“AI+ 肺癌防治中心” 启动会

为贯彻落实“健康中国 2030”战略，推动人工智能赋能肺癌防治工作，促进医学新质生产力在临床实践中的落地，经研究决定召开 AI+ 肺癌防治中心启动会。现将有关事项通知如下：

一、会议宗旨

2020 年，全球癌症新发病例约 1930 万例，癌症死亡近 1000 万例。其中肺癌新发病例数约为 221 万例，占癌症新发病例总数的 11.4%。肺癌导致的死亡约 180 万例，为癌症死亡首位。2022 年中国肺癌新发病例数已经超过 106 万，肺癌死亡例数超过 73 万。就肺癌患者的 5 年存活率而言，2019 年中国为 19.7%，还远远低于在 2010 至 2014 年期间诊断的亚洲国家中的日本（33%）和韩国（25%）肺癌患者。

如何解决上述问题，国际元宇宙医学联盟，中国肺癌防治联盟特提出“AI+ 肺癌防治分中心行动”，即赋能有一定基础的医院，应用物联网和元宇宙医学技术，践行救治“AI+ 肺癌防治分中心”的会诊培训和示范工作。得到赋能的分中心将全力配合中国肺癌防治联盟 A 级专家团队会诊并学习相关技术。对于行动效果优秀的医院，且考核达标后，可被认证为“AI+ 肺癌防治分中心及示范基地”，利于为其联袂的肺结节诊治分中心或医联体医院提供各期肺癌的会诊和示范业务。旨在助力推广“智能救治 AI+ 肺癌防治分中心行动”，为健康中国 2030 规划做出应有的贡献。

二、会议时间

2025 年 10 月 27 日（星期一）下午 13:30

三、会议地点

复旦大学附属中山医院 18 号楼 4 楼绍青厅

四、会议内容

- （一）宣布“AI+ 肺癌防治中心”正式成立；
- （二）发布 BAIMGPT 社会面开放计划（首批 30 万条 AI 医疗咨询服务）；
- （三）部署国家“四大慢病重大专项”成员单位应用推广；
- （四）介绍国际合作与未来规划。

五、会议议程

时间	内容	主讲 / 主持人
13:30-13:35	开场介绍与嘉宾致辞	宋元林主任
13:35-13:45	院领导致辞	宋振举副院长
13:45-13:55	国际专家致辞	Charles Powell 教授
13:55-14:15	启动讲话：《AI+ 肺癌防治的战略意义与全球首创地位》	白春学教授
14:15-14:25	中心建设汇报：《AI 肺癌防治中心总体规划》	宋元林主任
14:25-14:30	启动仪式与合影留念	全体嘉宾

六、出席人员

宋振举副院长、Charles Powell 教授、白春学教授、宋元林主任、杨达伟副主任医师及相关多学科专家共约 30 人。

复旦大学附属中山医院
上海呼吸物联网医学工程技术研究中心
中国肺癌防治联盟（CAALC）
国际元宇宙医学协会（IAMM）
2025 年 10 月 9 日

元宇宙医学

YUAN YU ZHOU YIXUE



季刊

2024 年创刊

第 2 卷第 3 期 (总第 7 期)

2025 年 9 月 30 日出版

主 办

复旦大学附属中山医院

编 辑

《元宇宙医学》编辑委员会

200032, 上海市枫林路 179 号

电话: (021) 51217570

网址: www.zentimepublishing.com

com

E-mail: metamedsh@vip.163.com

主 编

白春学

顾建英

副主编 (按拼音排序)

Nels H. Chavannes

陈荣昌

高承实

贾泽军

刘 雷

连 勇

Charles A. Powell

潘 毅

宋振举

Christoph Thüemmler

王伟林

杨达伟

编辑部主任

贾泽军

联合出版

《元宇宙医学》编辑部;

Zentime Publishing Corporation

Limited

ISSN 3006-4236

2025 年版权归复旦大学附属中山医院《元宇宙医学》编辑部所有

本刊刊出的文章除特别声明外, 不代表主办单位和本刊编委会的观点。本刊如有印装质量问题, 请向本刊编辑部调换。

目 次

专家述评

DID 在元宇宙医学中的应用探索..... 高承实(1)

构建阿尔茨海默病的未来: 从早期诊断到个性化干预的人工智能元

宇宙..... 林继先, 王 华, 汤罗嘉(11)

综 述

虚拟现实在心肺复苏培训中的应用与进展: 挑战与展望

..... 甘舜轩(19)

方法学研究

大语言模型和专病 GPT 赋能 OSA 咨询和管理的临床应用效果对比

研究方案

陆俊羽, 蔡沁怡, 白春学(25)

大语言和专病 GPT 赋能咨询和管理肺结节意义的对比研究方案

..... 杨达伟, 王 源, 白春学(31)

产学研融合

基于大语言模型的中文电子病历命名实体识别

..... 程 洁, 刘独玉, 陈思旭, 等(39)

伦理与法规

人工智能工具医学使用的科学伦理与规范重构..... 白春学,贾泽军,杨达伟,等(48)

指南与共识

人工智能照护中国专家共识

.... 中国肺癌防治联盟,上海呼吸物联网医学工程技术研究中心,国际元宇宙医学协会赋能基层医院
肺癌筛查管理专家组(56)

消 息

中山医院将建设国家人工智能应用中试基地 (24)

本期执行主编 高承实
本期责任编辑 臧天茹,孙梦瑶
本期责任校对 殷 悦,姬静芳,王 迪,翟铖铖

METaverse IN MEDICINE

Established in 2024; Volume 2, Number 3, September 30, 2025

Sponsor

Zhongshan Hospital, Fudan University

Editor

Editorial Committee of *Metaverse in Medicine*
179 Fenglin Road, Shanghai 200032, China
Tel: (86)21-51217570
http: www.zentimepublishing.com
E-mail: metamedsh@vip.163.com

Editor in Chief

BAI Chunxue
GU Jianying

Associate Editor

Nels H. Chavannes
CHEN Rongchang
GAO Chengshi
JIA Zejun
LIU Lei
LIAN Yong
Charles A. Powell
PAN Yi
SONG Zhenju
Christoph Thüemmler
WANG Weilin
YANG Dawei

Editorial Director

JIA Zejun

Co-publishing

Editorial Office of *Metaverse in Medicine* & Zentime Publishing Corporation Limited

ISSN 3006-4236

Copyright © 2025, Zhongshan Hospital, Fudan University

CONTENTS IN BRIEF

Commentary

- Exploration of the application of did in metaverse medicine
..... GAO Chengshi (1)
- Building the future of Alzheimer's disease: an AI-driven metaverse from early diagnosis to personalized intervention
..... LIN Jixian, WANG Hua, TANG LuoJia (11)

Review

- Application and progress of virtual reality in cardiopulmonary resuscitation training: challenges and prospects GAN Shunxuan (19)

Methodology

- A comparative study plan on the clinical application effects of large language models and specialized GPTs in OSA consultation and management
..... LU Junyu, CAI Qinyi, BAI Chunxue (25)
- A comparative study of the significance of GPT-enabled counseling and management of pulmonary nodules
..... YANG Dawei, WANG Yuan, BAI Chunxue (31)

Integration of IUR

- Named entity recognition in chinese electronic medical records based on large language models CHENG Jie, LIU Duyu, CHEN Sixu, et al (39)

Ethics and law

Reconstruction of scientific ethics and norms for the medical use of artificial intelligence tools
..... BAI Chunxue, JIA Zejun, YANG Dawei, et al(48)

Guidelines and consensus

Expert consensus on artificial intelligence care in China
..... Chinese Alliance Against Lung Cancer; Shanghai Engineer & Technology Research Center of Internet of Things for
Respiratory Medicine; The Expert Group for Empowering Lung Cancer Screening Management in Primary Hospitals of the
International Association for Metaverse Medicine(56)

News

Zhongshan Hospital Fudan University has been approved for the construction of a national artificial intelligence
application pilot base (24)

Acquisitions editor GAO Chengshi
Executive editors ZANG Tianru, SUN Mengyao
Executive proofreaders YIN Yue, JI Jingfang, WANG Di, ZHAI Chengcheng

DOI:10.61189/304718qrovcm

· 专家述评 ·

DID 在元宇宙医学中的应用探索

高承实*

安徽栈谷科技有限公司, 池州 247100



[摘要] 随着元宇宙医学的发展,患者身份管理、数据隐私保护及跨机构数据共享成为亟需解决的核心问题。去中心化身份(decentralized identity, DID)作为一种基于区块链的身份管理方法,为元宇宙医学提供了可信、可验证且可自主可控的身份体系。本文系统分析了 DID 的技术原理、架构模式及在医疗场景中的应用实践,梳理了国际、国内的探索案例,并对技术、法律、伦理和产业层面的风险与挑战进行了深入讨论。在此基础上,提出了未来发展方向,包括 DID 与隐私计算、数字孪生和 AI 医疗助手的融合,全球虚拟医院身份互认体系的构建,多方协作治理模式以及学科交叉研究的必要性。研究表明,DID 在解决“身份可信性、隐私可控性与数据共享性”三重矛盾方面具有显著优势,为元宇宙医学的安全、合规与高效发展提供了关键支撑。

[关键词] 去中心化身份(DID);隐私保护;数据共享;跨机构互操作性;数字孪生;多方协作治理

[中图分类号] R-0 **[文献标志码]** A

Exploration of the application of did in metaverse medicine

GAO Chengshi*

Anhui Stack Alley Technology Co., Ltd, Chizhou 247100, Anhui, China

[Abstract] The emergence of metaverse medicine has introduced critical challenges in patient identity management, data privacy, and cross-institutional data sharing. Decentralized identity (DID), a blockchain-based identity framework, provides a verifiable, trustworthy, and user-controlled solution for these challenges. This paper presents a systematic analysis of DID's technical principles, architectural designs, and applications in healthcare, reviewing both international and domestic implementations. It further examines associated risks and challenges across technical, legal, ethical, and industrial dimensions. Future directions are discussed, including the integration of DID with privacy-preserving computing, digital twins, and AI healthcare assistants, the development of globally interoperable virtual hospital identity systems, multi-party governance models, and interdisciplinary research initiatives. The findings indicate that DID effectively addresses the triad of identity trustworthiness, privacy control, and data shareability, offering foundational support for the secure, compliant, and efficient advancement of metaverse medicine.

[Key Words] decentralized identity (DID); privacy protection; data sharing; cross-institution interoperability; digital twin; multi-party governance

随着虚拟现实(VR)、增强现实(AR)、区块链以及人工智能等技术的迅速发展,医学正快速迈入元宇宙医学时代。元宇宙医学不仅包括沉浸式的医学教育、虚拟手术模拟和远程康复训练,还涵盖了跨地域协作、数字孪生人体、虚拟医院等新型医疗形态。然而,在这一过程中,身份管理与隐私保护成为绕不开的核心难题。

首先,医学数据具有高度敏感性。患者的病历、基因数据、影像资料往往涉及个人隐私乃至家族风险,一旦泄露可能造成严重的社会与法律后果。因此,元宇宙医学中的任何交互场景,无论是

医生资质验证、患者远程诊疗,还是跨机构病历共享,都必须建立在可信的身份识别与隐私保护基础之上。

其次,现有的医疗身份体系存在显著局限。一方面,大多数医院和医疗机构依赖于本地或区域性的电子病历(EMR/EHR)系统,导致数据难以跨机构、跨区域互操作,患者就医过程中往往需要重复检查和信息录入。另一方面,传统的身份验证机制高度中心化,例如依赖医院内部的账户体系、政府颁发的身份证号或医保卡号。这种模式在虚拟空间中难以直接套用,且容易产生身份伪造与数据滥

[收稿日期] 2025-09-12

[接受日期] 2025-09-25

[作者简介] 高承实,博士,副教授。

*通信作者(Corresponding author).Tel: 13838001036. E-mail: 13838001036@163.com

用的风险。

此外,远程医疗与虚拟医院的兴起,对身份信任体系提出了更高要求。在跨国界、跨区域的虚拟医疗服务中,医生如何证明其真实身份与执业资质?患者如何在不暴露全部隐私信息的情况下获得诊疗服务?医疗设备、虚拟人体(数字孪生)又该如何建立与现实世界的一一对应关系?这些问题都凸显了传统身份管理体系的不足。

在法律与合规层面,全球主要国家和地区均在加强医疗数据的隐私保护。例如,欧盟的《通用数据保护条例》(GDPR)^[1]、美国的《健康保险可携性与责任法案》(HIPAA)^[2]、中国的《个人信息保护法》^[3]均对医疗数据的收集、存储、处理和跨境传输提出了严格要求。这些法规不仅强调了“最小化收集”和“知情同意”的原则,也对身份识别和授权机制提出了要求。在元宇宙医学环境中,身份认证与数据访问更加复杂,如何在合规的同时实现高效的医疗服务,是亟待解决的难题。

1 去中心化身份(DID)的基础与原理

在此背景下,去中心化身份(decentralized identifier, DID)逐渐进入研究者与产业界的视野。与传统身份体系不同,DID 倡导“用户自主管理身份”,通过密码学与分布式账本技术实现身份标识的可验证与不可篡改。更重要的是,DID 能够结合可验证凭证(verifiable credentials, VC)等机制,在不暴露完整信息的前提下,向验证方证明“我是谁”或“我具备某种资质”^[4]。例如,患者可以仅证明自己“已年满 18 岁”而无需泄露出生日期,医生可以证明自己“拥有某项手术执业资质”而不必公开全部执业信息。这种特性为元宇宙医学中的身份与隐私管理提供了新的解决思路。

1.1 DID 的概念与发展背景 去中心化身份(DID)是指一种无需依赖中心化身份提供者,由用户自主管理的身份体系。与传统身份体系(如政府身份证、医院账户或第三方认证系统)不同,DID 的设计理念是“用户主权”,即用户能够自主管理自己的身份信息,并选择性地向验证方披露数据^[4]。

DID 概念源于分布式账本技术的发展。随着区块链在金融、供应链、物联网等领域的应用,研究者逐渐认识到,分布式账本天然具备数据不可篡改、可追溯、去中心化信任的特性,可以用来构建跨域、无需中心化机构管理的身份体系。W3C 在 2022 年发布了 DID 1.0 标准,定义了 DID 的核心结构、解析方法及与可验证凭证的交互规范^[4]。

在元宇宙医学背景下,DID 的价值尤为突出。它不仅保护患者隐私,实现数据可控共享,还可以用于验证医务人员资质、保障医疗设备的可信性,为数字孪生与虚拟医院提供可信身份基础。

1.2 DID 的核心要素 根据 W3C DID 标准及相关研究,DID 主要由以下三个核心要素组成^[4-5]。

一是 DID 标识符(decentralized identifier)。DID 是一种全局唯一的标识符,用于标识一个身份主体(个体、组织、设备、数字孪生等)。DID 的格式通常为“did:方法名:标识符”,例如“did:example:123456789abcdefghi”。“方法名”指定了 DID 的解析和注册规则,而“标识符”唯一标识了身份主体。

二是 DID 文档(DID document)。DID 文档是存储在区块链上或分布式存储系统中的 JSON 或 JSON-LD 数据结构。DID 文档包含公钥、服务端点、验证方法等信息,用于支持身份验证和加密通信。通过 DID 文档,验证方可以获得与 DID 关联的公钥,从而完成身份认证与消息加密。

三是去中心化验证机制。DID 的可信性依赖于分布式账本或去中心化网络,而非单一机构。用户持有私钥,通过签名证明自己对 DID 的控制权。验证方使用公钥验证签名,从而确认身份的真实性。

通过上述三个要素,DID 实现了身份的自主管理、可验证以及不可篡改。

1.3 可验证凭证 DID 本身只提供身份标识,但医疗场景经常需要证明某种属性或资质,例如“医生具备心脏手术资质”或“患者已完成疫苗接种”。此时,可验证凭证(VC)成为关键机制^[4]。

VC 的核心思想是,发行方颁发凭证,持有者保存凭证,验证方验证凭证。在元宇宙医学中,VC 的应用示例,首先是对医务人员进行资质验证,医院或专业机构作为发行方,颁发医生执业凭证(VC),医生将凭证存储在个人 DID 钱包中,参与远程手术或虚拟诊疗时,可向虚拟医院验证资质而无需泄露完整信息。之后对患者健康状态进行证明。疫苗接种、慢病管理、检验结果等证明可作为 VC 颁发给患者,患者可选择性披露某些信息,例如仅证明“已完成疫苗接种”,而无需公开具体接种日期或个人敏感信息。

通过 VC,DID 可以实现“数据可用不可见”,这是元宇宙医学中身份与隐私管理的核心优势。

1.4 DID 与传统身份体系的比较 传统 EMR/EHR 与医保电子凭证以“机构—平台”为中心,患者账号受限于院端或国家库,跨院需重复注册、调阅依赖接口开放,存在数据孤岛与隐私泄漏风险。DID 将身份与

凭证拆分为患者自持的加密钥匙与可验证凭证,跨院、跨省无需新建账户,通过链上信任注册表即可秒级验证,真正实现“我的数据我做主”。在落地层面,DID 可复用现有 EMR/EHR 的数据出口与医保结算通道,医院继续按 HL7/FHIR 保存完整病历,仅将摘

要、授权与状态写入 VC,既保留集中系统的大容量、高性能,又获得去中心化网络的互操作性和隐私保护,二者互补形成“链上身份+链下数据”的新型医疗数字底座。DID 与传统中心化身份的比较及其在元宇宙医学中的意义如表 1 所示。

表 1 DID 与传统中心化身份的比较及其在元宇宙医学中的意义

特性	传统中心化身份	DID(去中心化身份)	元宇宙医学意义
管理主体	政府、医院、第三方平台	用户自主管理	患者/医生可自主控制信息披露
数据存储	中心化数据库	分布式账本 / 联盟链	跨机构互操作,防篡改、可追溯
信任模式	信任中心化机构	密码学 + 分布式共识	虚拟医院、远程诊疗环境中无需单点信任
隐私保护	受限,中心化平台可访问	可选择性披露,结合零知识证明等	保护患者敏感信息,实现合规数据共享
可扩展性	跨机构复杂,互操作差	方法可扩展,可支持多种医疗场景	支持数字孪生、虚拟医院、远程诊疗等新场景

从表 1 可以看出,DID 在患者数据主权、医务人员资质验证、跨机构互操作性和隐私保护方面具有显著优势,这也是本文后续分析其在元宇宙医学中应用的理论基础。

1.5 DID 的技术优势与医学应用价值 DID 在元宇宙医学中具有多重优势和价值。首先是身份自治,患者和医务人员可完全控制自己的身份信息,不依赖中心化平台。其次是数据可控共享,通过 VC 或零知识证明,允许有限信息披露,实现跨机构医疗数据共享。三是防篡改与可追溯,利用分布式账本技术,所有身份操作均可记录、验证和追溯。四是跨平台互操作性,DID 体系基于开放标准,可与不同医院系统、虚拟医疗平台和国际标准(如 HL7 FHIR)对接。最后还支持数字孪生与虚拟设备绑定,每个数字孪生或 IoT 设备均可通过 DID 绑定身份,实现可信的虚拟一现实对应关系。

综合来看,DID 不仅是一种技术机制,更是元宇宙医学中身份治理与隐私保护的基础设施,为后续的应用场景分析和技术实现提供了理论支撑。

2 医学应用场景分析

DID 在元宇宙医学中的典型应用场景,涵盖了患者健康身份管理、医务人员资质认证、虚拟医院与远程诊疗,以及医疗设备与数字孪生体身份。上述场景中的应用,揭示了 DID 的实际价值与潜在应用路径。

2.1 患者健康身份管理 在元宇宙医学环境中,患者健康信息的跨机构共享和隐私保护是核心问题。传统医疗体系下,患者的病历信息分散在不同医院、实验室和健康管理平台,跨机构访问往往依赖纸质病历或中心化系统授权,效率低且存在隐私泄

露风险。通过 DID 与可验证凭证(VC)的结合,患者可以实现自主管理健康身份,同时支持可控共享。具体应用包括以下几个方面。

一是个人健康档案(PHR)管理。患者通过个人 DID 钱包管理电子健康档案;医疗机构可通过验证患者 DID 与 VC 来确认身份及健康状态;支持选择性信息披露,例如仅提供过敏史、免疫接种记录或慢病管理状态,而不泄露全部病历信息。

二是跨医院病历整合。利用 DID 进行患者身份统一识别,确保不同机构的病历信息可溯、可验证。结合零知识证明,可在不透露敏感数据的前提下,实现数据可信共享。例如患者在虚拟医院接受远程诊疗时,医生只需验证关键健康信息即可制定治疗方案,而无需访问完整历史病历。

三是动态授权与隐私控制。患者可根据不同场景动态授权数据访问,例如一次性授权、限时授权或按功能授权。权限变化通过区块链记录,实现操作可追溯,防止滥用。

通过以上应用,DID 为患者提供了“自主可控、可验证、跨机构互操作”的健康身份管理,实现了元宇宙医学中身份与隐私的双重保护。

2.2 医务人员资质与执业认证 在虚拟医疗环境中,医生和护理人员的身份验证与资质确认尤为关键。传统认证模式依赖中心化机构颁发的纸质证书或电子认证,这种证书或认证存在以下局限:跨境远程手术或远程会诊难以实时验证资质;纸质或单一数字证书易被伪造或篡改;不利于虚拟医院、远程诊疗等新型医疗模式的扩展^[6]。

DID 与 VC 提供了一种新的解决方案。在资质颁发与存证环节,医疗监管机构或医院作为 VC 发行方,将医生执业资质、手术权限等信息颁发给其

DID;VC 数据通过区块链或分布式账本存储,保证不可篡改。在动态验证与使用场景环节,医生在虚拟手术、远程诊疗或学术交流中,可通过 VC 向医院或患者验证资质,而无需公开完整信息,并且支持动态更新,例如手术许可到期、培训更新等。在跨境与跨机构互操作环节,DID 体系支持全球标准化解析,使医生资质在不同国家或地区的虚拟医院互认,减少跨境医疗服务的身份信任壁垒,提高远程医疗效率。

通过这种方式,DID 体系不仅确保医务人员身份的真实性,也为虚拟医疗场景提供了安全可信的资质管理机制。

2.3 虚拟医院与远程诊疗 随着 VR/AR 技术的发展,虚拟医院和远程诊疗逐渐成为元宇宙医学的重要组成部分。在此场景下,身份验证与隐私保护面临更高要求。一是双向的身份验证,患者 DID 与医生 DID 的双向验证可以确保双方均为真实、授权的主体。结合零知识证明技术,患者可在不泄露全部个人信息的前提下完成身份确认。二是虚拟诊疗流程设计。患者通过 DID 钱包提交必要健康信息;医生通过 VC 验证执业资格;虚拟医院系统通过区块链或联盟链验证双方身份及权限,完成安全诊疗。三是数据隐私与合规,所有身份与权限操作均可链上记录,实现可追溯性。结合加密存储、访问控制和隐私计算,满足 GDPR、HIPAA 等法规要求。四是跨平台互操作性。支持多种虚拟医疗平台和远程诊疗系统的身份互认,便于实现患者健康数据和诊疗记录的跨平台共享。

通过上述机制,DID 为虚拟医院与远程诊疗提供了身份可信、数据安全、隐私可控、跨平台互操作的技术基础。

2.4 医疗设备与数字孪生体身份 在元宇宙医学中,医疗设备及数字孪生体的身份管理同样关键。医疗 IoT 设备的 DID 化,使得每个智能医疗设备(如可穿戴设备、监测仪器)均可注册 DID,这样既确保了数据来源可信,也防止了数据伪造或篡改。

一方面是数字孪生身份绑定。患者的数字孪生体可通过 DID 与真实个体绑定,实现虚拟与现实的可信对应。结合区块链记录,确保数字孪生体操作和数据可溯源。另一方面是跨设备协作与数据整合。多种医疗设备数据通过 DID 与患者 DID 关联,实现统一身份管理。支持智能诊疗系统进行数据分析、AI 辅助诊断,而不暴露敏感信息。通过对设备与数字孪生体的身份治理,DID 不仅保障了数据可信性,也为虚拟医疗环境中复杂实体的管理提

供了标准化方法。

综上所述,DID 在元宇宙医学中呈现出广泛的应用价值。在患者健康身份管理方面,实现了跨机构、跨场景的数据可信共享与隐私保护;在医务人员资质管理方面,保障远程医疗与虚拟手术的身份真实性与资质可验证性;在虚拟医院与远程诊疗方面,支持双向身份验证、隐私计算与合规管理;在医疗设备与数字孪生方面,确保数据来源可信、操作可追溯、虚拟实体与真实个体一致性。这一系列应用场景为元宇宙医学中的身份与隐私治理提供了可行路径,也为后续技术实现和案例分析提供了具体实践依据。

3 技术实现与系统架构

DID 在元宇宙医学中的技术实现路径和系统架构设计,涵盖了 DID 注册与解析、可验证凭证(VC)管理、身份验证流程、隐私保护机制以及与现有医疗信息系统的集成。

3.1 DID 注册与解析 DID 的技术实现首先依赖于身份的注册与解析机制,这一过程确保了每个身份的全局唯一性和可验证性^[7]。

DID 注册。用户或设备在注册时生成一对公私钥,通过公有链或联盟链发布 DID 文档。DID 文档包含公钥、服务端点、验证方法等信息,用于后续身份认证和数据交互。在医学场景中,患者、医务人员及医疗设备均可注册 DID,实现统一身份标识。

DID 解析。验证方在需要确认身份时,通过 DID 解析机制获取 DID 文档。解析方法依赖于具体的 DID 方法(method),如 did:example、did:ethr、did:indy 等,不同方法对应不同的底层账本和解析规则。在元宇宙医学中,解析过程必须保证高可用性和低延迟,以满足虚拟诊疗和实时交互需求。

链上与链下存储结合。DID 文档可存储在公链、联盟链或分布式存储系统中(如 IPFS、Arweave),保证数据不可篡改与可溯源。对于敏感信息(如患者病历、医疗设备参数),可采用链下加密存储,并在 DID 文档中存放索引或加密指针,兼顾安全性与性能。

3.2 可验证凭证(VC)管理 在 DID 架构下,可验证凭证(VC)用于证明身份属性和资质,是元宇宙医学中隐私保护与信任建立的核心机制。

VC 的颁发。医疗监管机构、医院或实验室作为 VC 发行方,为患者、医务人员或设备颁发凭证,凭证内容可以是执业资质、疫苗接种记录、病历摘要、设备认证信息等。VC 使用发行方私钥签名,确

保证书的真实性与完整性。

VC 的存储与管理。用户通过 DID 钱包或客户端存储 VC,并可在不同虚拟医疗平台携带和使用。支持多重凭证管理,如按场景、按时间、按角色进行分类存储。

VC 的验证与隐私保护。验证方使用发行方公钥验证 VC 的真实性。结合零知识证明(ZKP)技术,可以在不泄露全部属性的前提下验证特定资质。例如,患者可证明“已接种疫苗”而不泄露具体时间或品牌。这种机制在跨医院、跨地区、跨平台的远程诊疗中尤为关键。

3.3 身份验证与访问控制流程 在元宇宙医学中,身份验证流程需同时满足安全性、效率性与隐私性。在身份认证流程中,患者或医务人员通过 DID 钱包发起认证请求;系统解析 DID 文档,获取公钥及验证方法;使用签名或加密挑战—响应机制完成身份验证。

在授权与访问控制环节,用户基于 VC 授权访问特定资源(如病历、影像、设备数据),授权信息通过智能合约或链上策略存储,保证不可篡改和可追溯。支持细粒度权限控制,包括按角色、按时间、按功能授权。

在多方协作场景中,比如在虚拟手术或远程诊疗中,患者、医生、设备均可通过 DID 完成身份互认证,系统根据 VC 和访问策略动态控制数据流向,保证隐私保护和合规性。

3.4 隐私保护与安全机制 DID 在元宇宙医学中的隐私保护依赖多种技术手段。一是加密存储与通信。患者数据、VC、设备信息均采用对称或非对称加密存储;通信过程使用 TLS 或端到端加密,防止数据在传输中泄露。二是零知识证明(ZKP)与选择性披露。ZKP 技术允许用户证明某一属性的真实性,而无需公开具体信息。例如,患者可证明“血糖水平在安全范围”而不泄露具体数值。支持 VC 的选择性披露,实现最小化信息共享原则。三是链上审计与可追溯性。所有身份操作、凭证颁发和访问授权均记录在区块链公有链或联盟链上;审计日志可用于合规检查和事件追溯,防止滥用或身份冒用。四是防篡改与防重放攻击。利用区块链不可篡改特性保证身份数据及 VC 的完整性,签名和时间戳机制防止重放攻击,确保身份验证安全可靠。

3.5 系统架构设计 结合前述技术实现,元宇宙医学中 DID 系统的典型架构可分为五层,如图 1 所示。

身份层(DID 层)管理患者、医生、设备及数字孪

生体的 DID,提供身份注册、解析和公钥管理服务;凭证层(VC 层)颁发、存储、验证 VC,支持选择性披露与零知识证明,提供跨平台、跨机构的资质与健康状态认证;访问控制与策略层基于智能合约或链上策略进行权限管理,支持动态授权、时间/角色/功能控制和多方协作;应用服务层支撑虚拟医院、远程诊疗、数字孪生交互、医疗 IoT 数据管理等具体应用,提供身份验证接口、VC 验证接口及安全通信服务;数据存储与隐私保护层用于链上存储 DID 文档和策略信息,链下加密存储敏感数据,实现数据可溯源、不可篡改,同时保证隐私保护。

通过该架构设计,DID 系统可以在元宇宙医学中实现身份可信、数据安全、隐私可控、跨机构互操作的技术目标。

4 案例与实验探索

本节将以实际案例和实验为基础,分析 DID 在元宇宙医学中的落地尝试,包括国际探索和国内探索。通过对比不同实践经验,可以为未来在国内外医疗场景中推广 DID 提供参考。

4.1 国际探索

4.1.1 欧盟 EBSI 在医疗身份管理上的尝试 欧洲区块链服务基础设施(European Blockchain Services Infrastructure, EBSI)是欧盟委员会与成员国联合建设的跨境区块链基础设施,明确将 DID 与 W3C 可验证凭证(VC)作为其“数字信任基建”核心组件之一。EBSI 提供了专门的 DID 方法规范(did:ebsi/针对自然人与法人不同的 DID 方法)、DID 注册/解析(DID Registry API)、以及面向 VC 的数据模型与工具链(钱包、解析库、issuer registry 等),并在教育、职业资格证书与若干公共服务场景开展跨境试点。

就健康医疗方向,EBSI 的治理与技术栈支持将凭证颁发方(issuers)列入“可信发行者名单”,凭证以加密/选择性披露方式在持有者钱包中保存,验证时通过 DID 文档和信任注册表完成验证与状态检查,从而在合规(GDPR)约束下实现跨国凭证互认。若干学术/行动研究也用 EBSI 做了跨境凭证验证的实证(教育/证书领域的试点经验对医疗凭证设计具有直接借鉴意义)。EBSI 的医疗身份管理探索主要包括以下几个方面^[8]。

一是去中心化身份(DID)与医疗凭证。EBSI 推出了基于 DID 的可验证凭证(VC)框架,支持医疗从业者资质、患者健康信息和疫苗接种记录的数字化。在试点项目中,患者可通过 DID 钱包存储健康凭证,并在不同欧盟成员国的医疗机构之间共享,

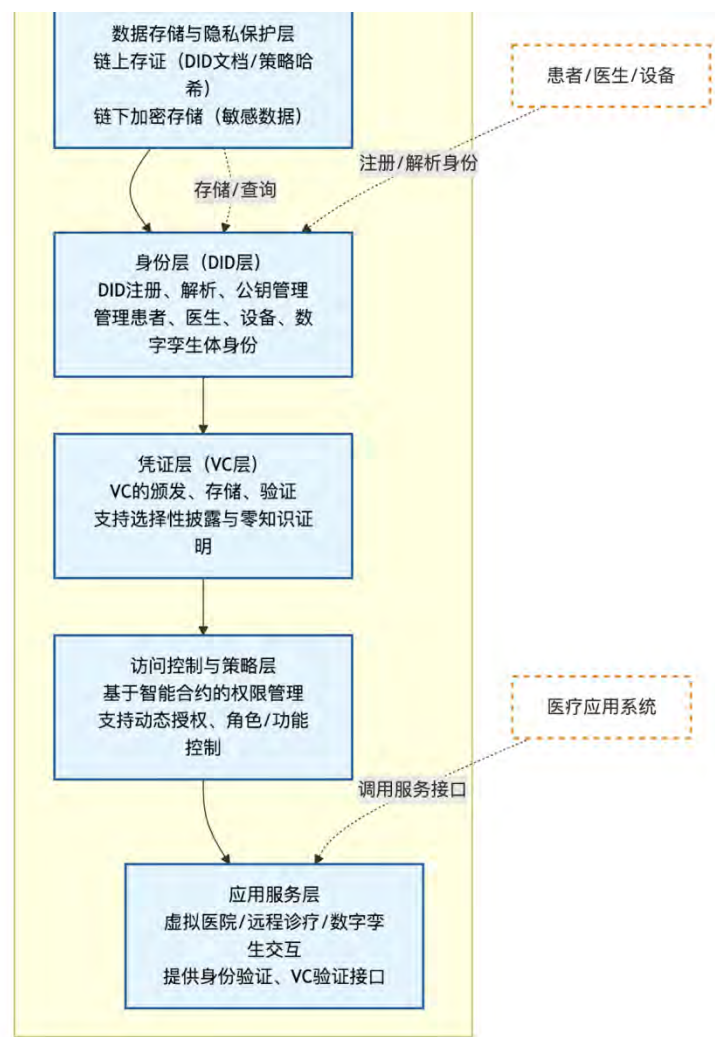


图1 元宇宙医学中 DID 系统的典型五层架构

实现跨境互认。

二是跨机构互操作性。EBSI 利用统一的 DID 方法和链上服务端点,实现医疗机构之间的身份互认与数据共享。支持 GDPR 合规,凭证数据采用加密存储和选择性披露,确保患者隐私。

三是试点成果与挑战。初步试点显示,使用 DID 与 VC 可有效减少跨机构病历验证时间。该试点尚未大规模进入生产环境,仍属 Pre-Production 阶段,在试点场景下能够显著提高医疗服务效率。面临的主要挑战包括,不同成员国法规差异、现有医疗信息系统兼容性、用户教育和数字素养。

欧盟 EBSI 统一 DID 方法当前版本为 did:ebasi v2.0,支持在 Polygon 与 Hyperledger Fabric 双链部署,而非单一公链。上述工作表明,EBSI 已把 DID/VC 从概念推进到可操作的跨境基础设施与试点工具,技术细节(DID 方法、DID Registry、VC data models, issuer registry、resolver libraries 等)均在官方 hub 文档中公开^[9]。

4.1.2 美国 MediLinker 基于 Hyperledger Indy/Aries 的医疗 DID/VC 原型与可用性评估

MediLinker 是由美国大学研究团队(以德州奥斯汀 Dell Medical School 等机构为主)开发的患者中心化身份管理原型,采用 Hyperledger Indy + Hyperledger Aries 实现去中心化标识与凭证交换,前端为用户钱包(可持有多类 VC:健康 ID、保险、药物清单、疫苗/COVID-19 状态、研究同意等)。该项目包括技术设计论文(阐述 DID/VC 模式、链上保存 DID 与凭证 schema,链下/离线数据用 FHIR 标准互操作)和以模拟就诊场景的可用性/功能性测试。

公开论文报告,在 30 人的模拟测试中 29 人完成全部活动,系统在七类临床场景(签到、共享凭证、撤销授权等)中均能实现预期功能;后续的可用性调查显示大多数受访者愿意采用该类患者自主管理的凭证模型(论文中披露了具体的完成率与用户反馈)。这些论文清晰记录了实现栈(Indy/Aries、钱包、凭证类型、与 FHIR 的接口说明)与测试结果,

表明 Hyperledger Indy/Aries 在“医疗场景的原型/POC”中已具备可操作性,但仍处于实验/小规模试点与用户测试阶段,而非大规模生产部署^[10]。

4.1.3 美国 MITRE 与 Mayo Clinic 分布式身份实验 美国 MITRE 和 Mayo Clinic 开展了针对分布式身份在医疗领域的实验,重点探索患者和医疗从业者的可信身份管理。

MITRE 基于 Hyperledger Indy 和 Aries 框架,构建了实验性的医疗 DID 系统。支持患者对病历、实验室结果和疫苗记录的自主控制,并可选择性分享给医生或保险机构^[11]。

Mayo Clinic 将 DID 与远程诊疗平台结合,医生和患者通过 DID 钱包完成身份验证。实验显示,使用 DID 与 VC 可以显著提升远程诊疗的安全性和用户信任度,同时在试点场景下减少重复验证步骤,降低中心化身份管理成本^[12]。

美国相关实验给我们的关键启示是,分布式身份在医疗场景中可增强信任、保障隐私,但对系统性能、用户操作便捷性和合规性提出了高要求。

4.2 国内探索

4.2.1 医保电子凭证与健康码的潜在对接 中国的医保电子凭证和健康码体系已广泛应用于日常就医和公共健康管理,其与 DID 的对接具有潜在价值。在医保电子凭证方面,电子凭证已实现患者身份验证和医疗费用结算数字化,但仍依赖中心化机构管理。如果引入 DID,患者可自主控制凭证授权,增强跨机构、跨省就医的便利性。在健康码与疫情防控数据方面,健康码记录的个人健康状态和疫情相关信息,可通过 DID 与 VC 的方式实现可验证共享。患者可选择性披露健康状态,兼顾隐私保护和防疫要求。

4.2.2 医疗联盟链案例 国内部分医疗联盟链也在探索去中心化身份的应用,杭州、深圳等地具备联盟链基础设施,学者已提出引入 DID 的框架^[13],但尚未公开落地试点,技术路径可参考欧盟 EBSI 与国际 MITRE 经验。相应的应用实践可以包括医院间病历信息共享、医务人员资质验证、远程会诊身份确认均可通过 DID 实现。初步研究显示,DID 有助于降低身份管理复杂度,提高数据共享效率,同时兼顾隐私保护。

通过国际、国内的案例可以看出,DID 在元宇宙医学中具有以下特点。一是身份可信、跨机构互操作。DID 可统一患者、医生和设备身份,实现跨平台和跨机构的数据共享。二是隐私可控、选择性披露。结合 VC 与零知识证明技术,患者可自主管理

数据共享范围。三是实践成熟度差异明显。欧盟和美国在实验性应用上更为成熟,而国内在联盟链与电子凭证层面逐步探索。这些案例为后续技术标准制定、系统部署和政策规划提供了经验和借鉴,同时也揭示了跨国法规、技术兼容性和用户接受度等未来挑战。

5 风险与挑战

随着去中心化身份(DID)在元宇宙医学中的应用逐步推进,尽管其在身份可信、隐私可控和跨机构互操作性方面展现出显著优势,但在技术、法律、伦理及产业层面仍面临诸多风险与挑战。

5.1 技术层面的挑战 可扩展性问题。元宇宙医学涉及海量患者、医疗机构及设备数据。DID 系统需支持高并发身份注册、凭证颁发和验证。因此,首先面临的挑战就是链上性能限制。大部分 DID 方法依赖区块链或联盟链存储 DID 文档或索引,链上吞吐量直接影响注册和验证效率。当前公链如 Ethereum 每秒可处理交易数有限,可能无法满足大规模医疗场景需求。其次是链下存储与计算压力。为保障隐私,敏感信息通常在链下采用加密存储,链上仅存索引或哈希。大规模医疗数据的加密、解密及访问控制可能导致系统计算压力增加,影响响应速度。

跨链互操作性。元宇宙医学必然涉及不同区块链网络、联盟链和医疗信息系统,不同 DID 方法的底层账本不一致,导致解析、验证和凭证共享复杂化。Cosmos IBC 与 Polkadot DID 模型为医疗凭证跨链提供了现成的底座,IBC 轻客户端可让医院联盟链在分钟级内同步患者 DID 文档与 VC 状态,无需新建网关即可与区域健康链互通;Polkadot 平行链架构则允许医保、药监、科研三条异构链各自运行专用 DID 方法,通过中继链共享信任根,实现“一键跨域”授权。借助跨链 DID,患者在国内就诊后,可将疫苗 VC 通过 IBC 实时传递至海外旅行保险链,自动触发保障生效,真正将“链上健康护照”融入全球医疗生态。但跨链互操作目前仍缺乏统一标准,需要协议层协调、智能合约桥接和中继机制支持。

用户友好性。DID 钱包和 VC 管理对非技术用户仍存在一定门槛。医疗场景中,患者和医务人员操作复杂度过高可能影响接受度和使用体验。需要设计简单、易操作的身份管理客户端,同时保证安全性。

5.2 法律与合规层的挑战 隐私法规冲突。欧盟

GDPR 强调数据主体控制权和数据最小化原则,而中国《个人信息保护法》在数据存储和跨境传输上有不同要求。跨境医疗数据交换过程中,DID 系统需兼顾不同法律框架,确保合法合规。

责任认定与审计。去中心化身份系统去除了单一中心化机构,这导致在发生数据泄露或身份滥用时,责任归属复杂。必须结合链上审计和可追溯机制,明确医疗机构、平台和用户各方责任。

5.3 伦理层的挑战 数据自主权与公共利益平衡。DID 提供患者对数据的完全控制权,但在公共卫生、疫情防控等场景中,过度自主可能影响公共利益。如何在保护隐私的同时,确保必要的公共健康数据共享,是伦理层面亟需解决的问题。

信息不对称与数字鸿沟。不同患者在数字素养、设备获取能力上存在差异。高技术门槛可能导致部分群体难以获得虚拟医疗服务,增加社会不平等。

信任与透明度。虽然区块链和 DID 技术本质上增强透明性,但复杂的加密和零知识证明机制可能令用户难以理解验证过程。如何提升用户信任,同时保证技术复杂度不影响可用性,是伦理设计的重要考虑。

5.4 产业层的挑战 基础设施建设。医疗 DID 系统需要高可靠的链上节点、存储系统和安全通信通道。对医院、远程诊疗平台和虚拟医疗企业而言,部署和维护成本较高。

商业模式可持续性。虚拟医疗和元宇宙医疗场景中,DID 系统的盈利模式尚不明确。如何在保证数据隐私与安全的前提下,实现系统的商业可持续发展,是产业化推广的关键问题^[7]。

标准化与生态建设。缺乏统一标准和规范可能导致各平台 DID 和 VC 无法互操作。医疗元宇宙生态中,标准化工作、跨机构合作和行业联盟建设亟需推进。

技术、法律、伦理与产业层面的挑战相互交织。技术层面,可扩展性、跨链互操作性和用户友好性仍需优化。法律与合规层面,跨境法规差异和责任认定问题影响系统推广。伦理层面,需平衡数据自主权与公共利益,同时关注信息不对称和信任问题。产业层面,基础设施投入、商业模式和标准化建设是推广关键。这些挑战表明,元宇宙医学中 DID 的广泛落地仍需技术优化、政策引导、伦理审视及产业协同,以实现真正安全、可信、可持续的数字身份管理。

6 未来展望

随着去中心化身份(DID)在元宇宙医学中的探索不断深入,技术成熟度、产业应用和政策环境均呈现加速发展趋势。未来,DID 将不仅是身份管理工具,更可能成为元宇宙医学可信体系的核心支撑。

6.1 技术融合 未来元宇宙医学将呈现多技术融合的趋势,DID 将与隐私计算、数字孪生及 AI 医疗助手深度结合。

首先是 DID + 隐私计算。隐私计算技术(如同态加密、联邦学习、多方安全计算)可在保证数据不泄露的前提下进行跨机构分析。与 DID 结合后,患者数据可以安全参与算法训练或远程诊疗,同时保持身份不可伪造和隐私可控。

其次是 DID + 数字孪生。每位患者的数字孪生体可以与 DID 绑定,实现虚拟医疗环境中的个性化互动。数字孪生体不仅承载病历和健康状态,还能在虚拟诊疗、手术模拟、康复训练中充当可信代理。

再次是 DID + AI 医疗助手。AI 医疗助手依托 DID 认证和 VC 验证,能在虚拟诊疗场景中为患者提供个性化建议,同时确保交互数据的安全性。结合智能合约,AI 助手的诊疗行为和决策可被链上审计,提高可追溯性与合规性。

技术融合将进一步提升元宇宙医学中身份管理的可信度、安全性和效率,同时为远程诊疗、虚拟医院及健康数据分析提供基础保障。

6.2 产业走向 未来元宇宙医学的产业发展将趋向全球化和平台互操作化。全球虚拟医院身份互认体系将满足跨境医疗、远程诊疗及国际科研合作的需求,将推动全球虚拟医院建立统一的身份互认体系。基于 DID 和 VC,患者身份、医疗从业者资质和设备认证可以实现跨国、跨平台互认,从而打破地域限制。

产业链延伸与商业模式创新。DID 作为基础设施,可延伸至医疗保险结算、药品供应链、健康管理服务等环节。企业可基于 DID 提供增值服务,如虚拟健康档案管理、个性化健康推荐、数字孪生手术训练平台。

标准化与生态建设。全球虚拟医疗生态需要统一身份标准、协议规范和数据互操作标准,以推动产业健康发展。标准化可降低跨平台应用成本,提升系统可靠性和用户接受度。

6.3 治理模式 DID 在元宇宙医学中的应用将促使治理模式由单一机构管控向多方协作转变。

首先是多方协作治理。政府、医疗机构、技术平台和患者共同参与 DID 系统管理。智能合约和链上策略可保证治理过程透明可审计,确保各方权责明确。

其次是动态授权与策略管理。随着医疗场景的多样化,访问权限和数据使用策略将动态调整。DID 系统结合 VC,可实现按角色、按时间、按功能的精细化权限控制,提高治理灵活性和合规性。

最后是合规与标准监管。去中心化身份的推广必须在法律法规框架内开展,结合区块链审计机制实现监管与隐私保护的平衡。跨国应用场景下,需要建立国际协作的治理机制,协调不同国家隐私法规的差异。

6.4 学科交叉 未来元宇宙医学中的 DID 研究将成为多学科交叉的前沿领域。医学提供临床需求和应用场景,定义患者、医生及医疗设备身份管理的功能和规范。计算机科学负责 DID、VC、零知识证明、隐私计算、数字孪生和智能合约等技术的研发与优化。法律研究跨境数据流通、患者隐私保护和责任认定等法律问题,为 DID 落地提供合规依据。伦理学探讨数据自主权、公共利益、数字鸿沟与信任等问题,为系统设计提供伦理指引。

通过学科交叉合作,可形成技术、法律和伦理的有机统一,推动元宇宙医学身份管理系统在安全、合规和可持续方向发展。

未来展望显示,DID 在元宇宙医学中的应用前景广阔。在技术层面,DID 与隐私计算、数字孪生和 AI 医疗助手的融合将显著提升身份可信度和医疗服务效率。在产业层面,全球虚拟医院身份互认体系和标准化生态建设将推动跨境虚拟医疗发展。在治理层面,多方协作治理模式、动态权限管理和合规监管将成为新趋势。在学科交叉方面,医学、计算机、法律与伦理的深度融合将为技术落地提供全面支撑。

DID 将不仅是身份管理工具,更是元宇宙医学可信生态的核心支撑,为未来虚拟医疗服务的安全、合规与高效提供坚实基础。

7 结 论

本研究系统梳理了去中心化身份(DID)在元宇宙医学中的理论基础、技术架构、应用场景及实际案例,并对风险、挑战和未来发展进行了深入分析。通过前文的探讨,可以得出以下主要结论。

7.1 DID 在元宇宙医学中的核心价值 DID 通过提供自主、可信、可验证的身份标识,为元宇宙医学中

的各类参与主体建立了统一的身份基础。其核心价值主要体现在三个方面。一是身份可信性。DID 采用分布式账本和加密算法,使患者、医生及医疗设备的身份可被安全验证,防止伪造和冒用。通过可验证凭证(VC),医疗从业者资质、患者健康信息及数字孪生均可在虚拟环境中得到可信确认。二是隐私可控性。患者数据通过选择性披露、零知识证明等技术实现自主管理。数据使用、访问范围和共享对象均可由数据主体自主控制,有效保护隐私,同时满足医疗和科研的合规要求。三是跨机构共享性。DID 提供统一身份体系,支持不同医院、虚拟诊疗平台及跨境医疗机构的数据互操作。医疗信息可以在保持隐私的前提下进行安全交换,为远程诊疗、全球虚拟医院及医疗科研提供基础支撑。

7.2 解决“身份+隐私+共享”三重矛盾 元宇宙医学面临的核心矛盾,一是身份可信性与去中心化。传统中心化身份管理存在单点失效风险,而 DID 在分布式账本上实现身份验证与可追溯。二是隐私保护与数据共享。患者隐私与医疗数据共享需求存在天然冲突,DID 与 VC、隐私计算结合可实现选择性共享与最小化数据暴露。三是跨机构互操作性与安全性。多机构、多平台协作要求身份和数据互操作,同时保证安全合规,DID 提供可验证的身份锚点,为安全数据交换提供支撑。

因此,DID 在元宇宙医学中实现了身份可信、隐私可控与跨机构数据共享的平衡,为解决这一“三重矛盾”提供了可行路径。

7.3 未来研究与政策建议 基于本文分析,未来研究与政策可从以下几方面推进。

在技术优化方面,提升 DID 系统的可扩展性和跨链互操作性,优化链上链下协同存储与计算机制;改进用户界面和交互体验,降低非技术用户操作门槛。在产业生态建设方面,推动全球虚拟医院身份互认体系和标准化协议建设,形成跨境、跨平台的医疗身份生态;鼓励初创企业和科研机构探索创新商业模式,实现技术落地与可持续发展。在法律与合规指导方面,协调不同国家和地区的隐私法规差异,建立跨境数据流通合规框架;明确去中心化身份的责任归属与审计机制,为监管提供依据^[5]。在伦理与社会研究方面,平衡数据自主权与公共利益,关注数字鸿沟和信息不对称问题;探索医学、计算机、法律与伦理交叉的多学科研究,为系统设计提供指导。

通过技术、政策、伦理及产业的协同推进,DID 有望在元宇宙医学中发挥核心作用,促进可信、隐

私友好和高效的医疗服务发展。

总的来看,DID 在元宇宙医学中的应用不仅是身份管理手段,更是虚拟医疗生态可信体系的基石。它为患者自主控制数据、保障隐私安全、实现跨机构数据共享提供了技术支撑;同时,也对法律合规、伦理规范和产业模式提出了新要求。未来,随着技术优化、标准化推进和多学科协同,DID 将在全球虚拟医疗体系中发挥越来越重要的作用,为元宇宙医学的可持续发展提供坚实保障。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 高承实:选题,撰写、修改论文。

参考文献

- [1] 欧洲议会与理事会. 欧盟通用数据保护条例(General Data Protection Regulation, GDPR)[S]. (2016-04-27) [2025-09-01]. <https://gdpr-info.eu/>
- [2] U. S. CONGRESS. Health Insurance Portability and Accountability Act of 1996 (HIPAA) [S]. Washington, D. C. : U. S. Government Printing Office, 1996.
- [3] 全国人民代表大会. 中华人民共和国个人信息保护法[S]. (2021-08-20) [2025-09-01]. http://www.npc.gov.cn/npc/c2/c30834/202108/t20210820_313088.html
- [4] SPORNY M, LONGLEY D, CHADWICK D, 等. 去中心化标识符(DID)v1.0: 核心架构、数据模型与表示法[EB/OL]. (2022-07-19) [2025-09-01]. <https://www.w3.org/TR/did-core/>.
- [5] ALLEN C. The Path to Self-Sovereign Identity [EB/OL]. (2016-04-26) [2025-09-02]. <https://www.lifewithalacrity.com/article/the-path-to-self-sovereign-identity/>
- [6] ZHUANG Y, ZHANG W, NI J, 等. 基于区块链的医疗数据共享去中心化身份管理研究[J]. 工业信息化学报, 2022, 18(12): 9292-9301.
- [7] 葛 锐, 廖方宇, 姜 文, 等. 分布式数字身份及其在科学数据共享中的应用前景[J]. 数据与计算发展前沿, 2023, 5(5): 46-59.
- [8] European Commission.. Application and Exploration of the European Blockchain Services Infrastructure (EBSI) in Healthcare Identity Management [EB/OL]. (2025-07-09) [2025-09-02]. <https://digital-strategy.ec.europa.eu/en/policies/european-blockchain-services-infrastructure>
- [9] European Commission. What is EBSI? [EB/OL]. Digital Building Blocks, (2025-07-01) [2025-09-25]. <https://ec.europa.eu/digital-building-blocks/sites/spaces/EBSI/pages/590447955/What%2Bis%2BEBSI>
- [10] KHURSHID A, HOLAN C, COWLEY C, et al. Designing and testing a blockchain application for patient identity management in healthcare[J]. JAMIA Open, 2021, 4(3): oaaa073.
- [11] M P V, QURASHI S N, SOBIA F, et al. Decentralized identity management using blockchain for healthcare systems[C]//2024 IEEE Silchar Subsection Conference (SILCON 2024). November 15-17, 2024, Agartala, India. IEEE, 2024: 1-6.
- [12] Mayo Clinic. Distributed identity for remote telehealth: Joint POC with MITRE [EB/OL]. (2021-06-30) [2025-09-25]. <https://newsnetwork.mayoclinic.org/discussion/mayo-clinic-platform-and-mitre-explore-distributed-identity-for-telehealth/>.
- [13] 王 伟, 李 娜. 医疗联盟链中去中心化身份(DID)的应用研究[J]. 中国医疗信息化, 2022, 28(3): 45-49.

引用本文

高承实. DID 在元宇宙医学中的应用探索[J]. 元宇宙医学, 2025, 2(3): 1-10.

GAO C S. Exploration of the application of did in metaverse medicine[J]. Metaverse Med, 2025, 2(3): 1-10.

DOI:10.61189/061708adttmb

· 专家述评 ·

构建阿尔茨海默病的未来：从早期诊断到个性化干预的人工智能元宇宙

林继先¹, 王 华², 汤罗嘉^{3,4*}

1. 上海市闵行区中心医院神经内科, 上海 201199

2. 上海市闵行区中心医院信息科, 上海 201199

3. 复旦大学附属中山医院急诊科, 上海 200032

4. 上海市闵行区中心医院院长办公室, 上海 201199

[摘要] 阿尔茨海默病(Alzheimer's disease, AD)作为全球性公共卫生危机,面临早期诊断困难、治疗手段有限与照护压力沉重等多重挑战。人工智能(artificial intelligence, AI)与元宇宙的深度融合为AD全病程管理提供了新的解决思路。沉浸式交互环境结合AI分析可实现早期筛查与风险分层;基于数字孪生与自适应算法的个体化数字干预有望延缓疾病进展;同时,沉浸式模拟培训为照护者与临床医护人员提供高效支持,提升照护质量与决策能力。AI元宇宙将重塑AD的诊断、治疗与照护体系,为突破当前困境开辟新路径。

[关键词] 阿尔茨海默病;元宇宙;人工智能;数字疗法;数字孪生

[中图分类号] R 749.1 **[文献标志码]** A

Building the future of Alzheimer's disease: an AI-driven metaverse from early diagnosis to personalized intervention

LIN Jixian¹, WANG Hua², TANG LuoJia^{3,4*}

1. Department of Neurology, Central Hospital of Minhang District, Shanghai, Shanghai 201199

2. Department of Information, Central Hospital of Minhang District, Shanghai, Shanghai 201199

3. Department of Emergency Medicine, Zhongshan Hospital, Fudan University, Shanghai 200032

4. President's Office, Central Hospital of Minhang District, Shanghai, Shanghai 201199

[Abstract] Alzheimer's disease (AD), as a major global public health crisis, faces multiple challenges including difficulties in early diagnosis, limited therapeutic options, and heavy caregiving burdens. The deep integration of artificial intelligence (AI) and metaverse technologies offers innovative solutions for comprehensive AD management. Immersive environments combined with AI analytics enable early screening and risk stratification; digital twins and adaptive algorithms facilitate personalized digital interventions that may slow disease progression; while immersive simulation training provides efficient support for caregivers and healthcare professionals, enhancing care quality and decision-making capacity. The AI-driven metaverse will reshape AD diagnosis, treatment, and caregiving systems, opening new pathways to overcome current obstacles.

[Key Words] Alzheimer's disease; metaverse; artificial intelligence; digital therapy; digital twin

阿尔茨海默病(Alzheimer's disease, AD)作为痴呆症最常见的形式,已成为一个全球性的公共卫生问题,是全球第7大致死原因,也是导致老年人残疾和依赖的主要因素^[1]。全球痴呆症患者2019年约为5700万,每20年翻倍增长,到2050年将接近1.53亿,给全球卫生系统和社会经济带来巨大压力^[2]。尽管各国已投入大量资源应对这一全球性难

题,但AD的临床管理仍然面临3大核心挑战。

第一,诊断困境严峻。许多患者的早期症状,如记忆问题,常被误认为是正常衰老现象。目前缺乏单一决定性的AD诊断方法,往往依赖于排除其他病因后的临床复杂评估,不但确诊滞后而且漏诊率高^[3-4]。一项纳入30257例患者的系统综述与荟萃分析^[5]显示,从家属或患者首次识别症状到明确

[收稿日期] 2025-09-09

[接受日期] 2025-09-25

[基金项目] 闵行区医防协同医防融合项目(认知健康综合评估诊疗和研究医防融合平台建设)。Supported by Minhang District medical prevention coordination and integration project (Construction of an Integrated Cognitive Health Assessment, Diagnosis, Treatment, and Research Platform for Medical-Preventive Fusion)

[作者简介] 林继先,博士,副主任医师。E-mail: linjixian@fudan.edu.cn

*通信作者(Corresponding author). Tel: 021-64041990, E-mail: tang.luojia@zs-hospital.sh.cn

诊断的平均间隔为3.5年;即便在高收入国家,也仅有50%~65%的痴呆病例得到临床诊断,年轻起病与额颞叶痴呆等亚型的确诊时间则更长。初诊路径分散、影像、体液标志物与功能评估之间的割裂,是造成诊断延误的重要原因^[6]。

第二,治疗手段有限。现阶段针对AD的治疗手段均以改善症状为主,达不到治愈效果^[1]。近年来,以新药仑卡奈单抗与多奈单抗为代表的抗淀粉样蛋白单克隆抗体可以延缓轻度AD患者的临床进展;仅有极少数患者会因为抗体治疗的副作用死亡,比如引发淀粉样蛋白相关影像学异常(amyloid-related imaging abnormalities, ARIA)。即使大多数ARIA患者并无明显临床症状体征,但为了最大程度保障患者安全,仍需要在治疗前后多次进行MRI监测和评估,大大增加疾病的长期管理难度^[7]。

第三,照护压力沉重。痴呆症的照护成本几乎等于癌症、心脏疾病与中风的总和^[8-9]。据2015年AD国际组织(ADI)报导,每位照护者每天的平均工作时长可达6 h,其家庭照护压力体现在多个方面:(1)全球痴呆症约50%的经济费用由家庭成员承担;(2)痴呆症的日常生活活动,如饮食、穿衣、梳洗等需要高强度的照料,包括出行、就医、清洁护理都需要花费大量的时间与精力,照护者的精神负担沉重。

面对多重挑战,只有技术创新才能破局AD的临床管理困境。医学元宇宙是由沉浸式拓展现实(extended reality, XR)[虚拟现实(virtual reality, VR)、增强现实(augmented reality, AR)、混合现实(mixed reality, MR)],实时网络与边缘、云计算,互操作标准与可信数据治理共同支撑、与现实世界并行的持久交互空间。在医学元宇宙中,临床行为、环境线索与个体状态被情境化测量与闭环利用^[10]。

人工智能(artificial intelligence, AI)在该框架内承担“智能层”职能:对多模态数据,包括影像、体液标志物、行为学与语音、步态、眼动等数字指标进行特征提取、情境理解、自适应决策与个体化预测建模(包括数字孪生与反事实模拟)^[11]。从技术视角出发,AI被视为元宇宙的核心关键技术之一^[12-13];也有学者^[11]强调AI在塑造交互与体验中的驱动力,认为二者在本质上是相互融合的技术体系。重要的是,AI与元宇宙的协同为医学研究与临床应用提供了一种新的范式。如果说AI是驱动精准医疗的智能大脑,那么元宇宙就是承载和实现这一智能的、可交互的数字身体。

1 AI与元宇宙协同在AD早期诊断与评估中的应用

在明确了AD的临床难点后,技术创新便显得尤为关键。AI与元宇宙的结合不仅仅是概念上的前沿探索,更已逐步渗透到疾病的早期诊断和评估实践中。

1.1 超越传统量表:元宇宙中的认知评估 传统纸笔量表在文化依赖性、学习效应与灵敏度方面存在局限,难以完整记录AD发病的临床前阶段与轻度阶段的细微变化^[3]。沉浸式XR评估(immersive VR, iVR)将测量嵌入日常化情境,如导航任务、虚拟超市购物、虚拟厨房备餐等,提供更高的生态效度与可量化指标^[4],使得认知评估管理更加生动立体与全面。iVR被认为可用于AD评估的工程前提保障^[4]。

在技术层面,低时延交互(motion-to-photon latency, MTP)是沉浸式评估AD的前提保障。若延迟过高,会导致晕动并污染反应时、操作轨迹等认知指标。结合边缘计算与5G、6G的分布式渲染,可将时延进一步压缩至临床可接受水平。通过系统无关相机法可高精度量化MTP,为临床应用建立质控阈值^[14-15]。采用低时延联合高频追踪技术可以稳定地获取AD筛查人群的眼动、手势、身体姿态等数字行为标志,减少设备方差,提高早期差异的检出性^[16-17]。

内嗅皮层—海马网络的病变可导致AD患者早期的空间导航与定向障碍,被认为是AD早期最敏感的一个脑部功能域。fMRI研究^[18]显示,即使在认知正常的ApoE ε4携带者中,也可观察到类网格细胞表征的减弱,并伴随可测量的导航偏差。沉浸式导航任务可通过操控地标可用性、路径整合与自体中心—客体中心转换,逐层测试空间策略。与传统记忆任务检测比较,沉浸式导航任务能在更早期区分AD高危个体^[19-20]。

工具性日常生活活动能力量表(instrumental activities of daily living, IADL)的VR评估与情景或执行—内侧颞叶—额顶网络失衡机制存在密切关联,在老年与痴呆高危人群中显示出可行性和有效性。在路径上,把关键指标纳入AD风险分层与纵向随访,与既有量表交叉校准,与一般量表比较对疾病的诊断敏感性更高^[21]。沉浸式情境将IADL拆解为可量化步骤,如目标—搜索—选择—序列—结账上菜,输出时长、错误类型次数、序列偏差、路径效率等数字指标,较纸笔量表更具生态效度^[22]。经验证的工具包括:虚拟现实日常评估实验室

(validation of the virtual reality everyday assessment lab, VR-EAL)、虚拟超市范式(VSP/VST、MEMOSHOP等)区分轻度认知障碍(mild cognitive impairment, MCI)与健康对照并评估老年人情景记忆^[22-23]。VR-EAL是一种用于评估日常注意、情景或前瞻记忆及执行功能的工具,已被验证具有良好效度与可用性,且几乎不引发眩晕^[24]。综上所述,iVR用于认知评估不仅提高了生态效度,还能通过细粒度指标捕捉到传统量表难以识别的早期损伤,应用前景广泛。

1.2 AI赋能的隐蔽标志物检测 元宇宙环境能够捕获海量数字生物标志物的多模态数据流^[25]。用户在虚拟环境中交互时,传感器通过捕捉眼动轨迹、步态动力学、语音语调乃至键盘敲击模式的微小变化^[26],通过特征融合与自适应建模,将这些原始信号转化为AD早期风险分层的数字生物标志物。

1.2.1 眼动追踪 在沉浸式场景中可采集扫视、反扫视、注视停留与视觉配对等指标,对额顶叶—海马网络功能异常敏感^[27]。研究^[28]表明,基于眼动的范式对MCI具有区分潜力。在大样本数字筛查中,眼动+绘画的多模态模型的AUC为0.958,提示其在社区早筛中的可行性与可扩展性^[29]。VR-眼动一体化评估能够跨不同教育背景预测蒙特利尔认知评估量表(MoCA)评分,并判别认知障碍^[30]。

1.2.2 瞳孔测量 提供与认知负荷-去甲肾上腺素系统相关的补充信号。定量瞳孔反射在探索性研究中显示对AD的诊断潜力与预后价值,任务诱发的瞳孔反应亦与遗传风险与MCI状态相关,为AD的节律与视网膜通路改变提供无创窗口^[31]。

1.2.3 步态与双任务行走 是反映“认知—运动接口”的重要指标。在沉浸式走廊或结合IMU压力步道的环境中,可连续采集步速、步长、步时变异性、步态对称性及转向动力学等数字步态标志。若同时叠加口头减法、词语流畅或空间更新等并行任务,即可形成双任务范式。通过计算双任务代价,能够量化认知负荷对步态控制的影响,从而刻画认知与运动之间的耦合^[32]。步态变异性与转向动力学被认为与灰质体积、认知表现相关,提示其作为早期进行性下降标志的潜力^[33]。前瞻性队列研究(Gait and Brain)^[34]表明,双任务步态可预测未来MCI人群痴呆转化的风险级别;沉浸式走廊与IMU压力步道结合能输出步速、变异性与转向动力学,适合作为风险分层与纵向随访的指标,有助于区分主观认知下降与MCI^[35]。

1.2.4 手部追踪与自然交互技术 利用头显内置相机实时估计手或指关节姿态,避免了控制器依赖^[36]。研究^[37-38]表明,常见头显在指尖位置的平均误差约1.1 cm,关节角误差约9.6°,已可用于时钟绘制、药盒分拣或厨房任务的功能量化,捕捉早期运动执行困难。

1.2.5 语音与语言分析 在沉浸式对话或导航叙述任务中可被动采集,特征包括语速、停顿分布、语义连贯度与句法复杂度。语音与语言分析是AD数字生物标志物核心通道之一^[39]。纵向研究^[40]结果显示,通过认知测评过程中的语音,可预测患者6年内的AD转化,准确率超过78%。

1.2.6 多模态融合建模 多模态融合建模能弥补单一信号的脆弱性。将元宇宙场景下获得的眼动轨迹、绘画或路径学、语音特征与基本人口学信息融合,可提升单模态易受情境或任务影响的脆弱性,多模态模型总体AUC优于单模态^[41]。在社区研究中,眼动+绘画模型的AUC显著高于单模态^[29];语音+眼动的小样本研究亦显示分类增益^[42]。

综上所述,沉浸式导航与IADL任务提供高生态效度的数据采集界面;AI负责将眼动、步态、语音、路径学等数字标志物与影像、体液证据进行多模态融合,实现高效的AD早期识别与风险分层。

2 AI与元宇宙在AD治疗与干预中的应用

如果说早期诊断与风险分层是管理AD的起点,那么干预和治疗则是改变疾病进程的核心环节。AI与元宇宙技术的引入,使得个体化认知训练与动态干预成为可能。本节将重点讨论个性化数字疗法与数字孪生在AD干预中的前景。

2.1 个性化的数字认知疗法 元宇宙为认知干预提供了理想的平台。荟萃分析^[43]表明,基于VR的康复训练能够显著改善MCI与痴呆症患者的整体认知,涉及记忆、执行功能、注意与视觉空间加工等多个领域。与传统纸笔或电脑化训练相比,沉浸式任务的生态效度与迁移潜力更高。

2.1.1 处方学逻辑 典型训练方案包括虚拟超市购物(VSP、VST、MemoShop)、虚拟厨房烹饪与导航任务等,处方剂量多为每周2~3次、单次30~60 min、持续6~12周。研究^[24,44]表明,完成率与训练收益呈正相关,提示依从性是疗效关键决定因素。因此,应在设计时预设最小有效剂量与依从性阈值,并通过游戏化机制、即时反馈与难度自适应维持患者参与度^[45]。

2.1.2 AI自适应调控 沉浸式干预的优势在于可

实时采集轨迹、反应时、错误模式等细粒度指标。AI 模型可基于这些特征动态调整任务难度(例如购物任务中物品数量与干扰物复杂度,导航任务中的地标可见性),维持“挑战-能力”平衡。这一闭环自适应框架有望提高疗效并降低训练倦怠^[9]。

2.1.3 迁移效应 情境化 IADL 训练不仅改善认知测验得分,还能促进日常功能改善。随机对照试验证明,虚拟超市与厨房训练在 MCI 患者中改善了购物能力、备餐速度与多任务处理能力,部分研究^[21,46]显示可延缓向痴呆的临床转化。

综上,AI 驱动的沉浸式认知疗法实现了从“统一标准化训练”向“个体化处方与实时调控”的跃迁。未来研究需通过多中心 RCT 和长期随访验证其对延缓疾病进展的实际效果。

2.2 疾病管理的数字孪生设想 数字孪生是指面向单个个体的动态、可交互虚拟副本,持续整合其多模态信息(临床资料、影像学与体液生物标志物、遗传组学或多组学以及在沉浸式情境下采集的数字行为学数据),并以此支撑预测与决策^[47-48]。在 AD 领域,其被广泛视为精准医疗与预测性干预的重要方向^[26]。

2.2.1 疾病进展建模 在 AD 情境中,数据驱动的疾病进展模型,为数字孪生提供“时序-亚型”的结构化骨架。SuStaIn 算法通过同时建模“表型异质性+时间顺序”,在 AD、FTD 等数据集上识别出不同的进展亚型与阶段,已成为刻画“事件-阶段-亚型”关系的代表性方法;其时间尺度扩展进一步面向纵向或半纵向数据改进了时序估计的稳定性^[49]。事件驱动建模(event-based model, EBM)将疾病视为一系列“生物标志物异常事件”的累积序列,可从横断面数据估计事件顺序与不确定性,并在家族性 AD、亨廷顿病等场景得到验证与扩展(含后续的稀疏或并行事件形式)^[50-51]。将上述进展模型嵌入数字孪生框架,可对个体进行“实时病程定位”与亚型归类,并与 AT(N)标志物或行为数据(沉浸式导航或 IADL 的轨迹、错误谱系、眼动指标等)联合用于风险分层与预后预测^[26,47]。

2.2.2 非侵入数据推断体液标志物 走向“早期数字诊断”:新近研究探索以非侵入记录+数字孪生预测 CSF A β 、p-Tau 等体液标志物,期望降低腰穿依赖并推进早期诊断。初步结果显示,以多模态非侵入数据联合孪生体进行学习,可在独立数据集上取得有意义的判别或回归性能,为临床转化提供了可行性线索^[52]。

2.2.3 反事实模拟与个体化处方 数字孪生不仅用

于表征当前状态,更可用于反事实模拟。例如,在“虚拟个体”上比较标准照护、加入情境化 VR 训练和新药物策略等路径方案的长期结局差异,为临床医生对真实患者开具个体化处方提供量化证据^[53]。方法学上,可结合因果机器学习或与外部或历史对照的数据融合来估计个体化治疗效应,估计结果需配套外部验证与校准,避免样本选择与混杂带来的偏倚累积^[54-55]。

2.2.4 生成式 AI 与合成数据 在真实临床数据有限时,生成式模型可创造合成患者数据,用于训练分类器或扩展对照组,从而提升模型的稳健性与推广性^[56]。不过,合成数据无法预测未在源数据中出现的新型不良反应,因此需要建立风险缓解策略,以维护数据完整性、可追溯性和问责制^[57]。

面向 AD 的数字孪生在 3 类任务上最具前景:(1)个体化预测,包括病程定位、转归概率;(2)干预模拟,量化多路径干预 AD 的风险与收益,在数字孪生上预演 AD 不同疗法的长期效果,为现实世界的患者选择最优化、最个性化的治疗方案。;(3)研究加速,基于外部或合成对照支撑的 AD 小样本推断与试验设计。这样就可以在数字孪生上“预演”不同疗法对 AD 的长期效果,为现实世界的 AD 患者选择最优化、最个性化的治疗方案。

3 赋能照护者与医疗专业人员

AD 的管理不仅聚焦于患者本身,还深刻依赖于照护者和医疗专业人员的支持。如何提升照护者的应对能力和医护人员的诊疗水平,是维持长期照护质量的重要前提。

3.1 针对照护者的沉浸式培训与支持 AD 长期照护常依赖家庭成员或专业背景较低人员,往往缺乏系统培训,应对 AD 患者的激越、沟通障碍和夜间游走等情境经常会感到无力。元宇宙技术可将这些高风险情境转化为低风险、可复用的沉浸式脚本,供照护者反复演练。

沉浸式体验与共情提升。已有 VR 项目通过第一人称视角模拟患者感受,例如“虚拟迷路”或“信息混乱”情境,让照护者直接体验 AD 患者的困惑和焦虑。这类沉浸式培训被证实能够增强照护者的同理心和耐心,改善沟通与行为管理^[58-60]。

去激化训练。研究表明,照护者在 VR 环境中反复练习处理激越、抵抗护理或拒绝进食等情境,可以显著提高他们的知识、信心和处置一致性。VR-SIM Carers 项目以及圣克拉拉大学的混合现实试点已将“日落综合征”“拒药”“夜间游走”等典型

场景脚本化,照护者可通过虚拟患者的即时反馈了解自身反应的有效性,形成练习—反馈—再练习的闭环。

成效与指标。系统性证据^[61]显示,接受沉浸式培训的照护者在照护知识测验、同理心量表和自我效能感评分上显著提升;在 Zarit 照护负担量表上也观察到下降趋势。同时,沉浸式平台还能自动记录语言使用、反应延迟和操作轨迹,用于客观化评估。

因此,元宇宙平台不仅为 AD 照护者提供了低风险的培训环境,还可能在真实世界中降低应激水平和流失率,对可持续照护体系具有积极意义。

3.2 加速医疗专业人员的技能提升 随着 AD 新疗法的出现,临床医生需要学习新的诊疗流程。元宇宙为此提供了几近完美的培训工具。医生可以在超逼真的虚拟诊室中,对 AI 驱动的虚拟病人进行诊疗全流程的模拟。比传统方法更高效、成本更低,且能确保医生在接触真实患者前已具备充分的技能。

新型抗淀粉样药物的临床应用对医生提出了全新的临床诊疗流程的规范化要求,包括患者遴选、MRI 监测节点设置、ARIA 识别与处置^[7]。元宇宙在 AD 管理的应用提供了可交互、可重复的沉浸式演练环境,让医护在短时间内能较快地掌握复杂流程,包括 AD 患者基线与随访 MRI 要求 ARIA 分级与暂停、减量、继续观察地不同处置策略等,获得即时反馈^[7,62]。

模拟案例与反应训练。虚拟病人可出现“输注反应”“轻度 ARIA-E”“症状性 ARIA-H”等场景,医护需判断影像表现、选择管理策略。系统记录操作路径与决策准确率,便于形成个性化反馈与改进。

成效与限制。现有研究^[63]显示,沉浸式培训能提高知识掌握率、诊疗流程完整率和决策一致性,同时提升沟通技巧。不过,对于资深临床医生,知识增益有限;因此建议在培训对象上分层设计:面向新入职医护作为核心培训手段,面向资深医生则作为复杂情境应急演练。

综上,沉浸式培训在照护者与医护人员两方面均展现出情境化、可量化、可推广的优势,有望成为 AD 全病程管理体系的重要补充。

4 挑战、伦理考量与展望

尽管 AI 与元宇宙在诊断、干预和照护等方面展现出巨大潜力,但从实验室走向临床应用仍然存在诸多挑战。这些挑战不仅涉及技术实现的可行性,也包括伦理、隐私与监管等层面的深远考量。

4.1 技术与现实的鸿沟 虽然医学元宇宙与 AI 在实验环境中表现出前景,但在临床转化过程中仍存在多重障碍。技术层面,高性能 XR 头显价格较高,且对老年群体不够友好;边缘计算和高速网络尚未普及,导致交互延迟问题影响数据质量与使用体验^[37]。临床层面,如何将这些复杂的数字系统无缝嵌入到繁忙的临床 workflow 中,以及如何通过严格的临床试验证明其有效性并获得监管机构的批准,是必须跨越的障碍^[64]。此外,一个关键限制是,目前大多数研究为小样本或单中心试验,缺乏跨中心的外部验证;此外,数字孪生与合成数据难以预测前所未有的不良反应,因此需建立阈值预注册、模型不确定性报告和跨中心验证等方法学保障^[65]。

4.2 伦理的边界 AI 和元宇宙在在认知能力受损的弱势群体中应用,需重点关注伦理的问题。数据隐私与安全是首要挑战。沉浸式场景会采集眼动、语音、轨迹等高维数据,均属于《通用数据保护条例》(GDPR)第 9 条所界定的敏感健康数据。为合规应用,应坚持数据最小化(仅保存必要特征而非原始视频)、目的限定与可审计访问控制,并结合联邦学习与安全聚合以减少数据泄露风险^[66]。

算法公平性是另一个关键议题。已有商业健康算法因训练数据偏差而低估某些族群的健康风险。在 AD 筛查与预测中,性别、教育背景和文化差异都可能导致模型性能不均衡。研究应预注册公平性指标(分群灵敏度或特异度)、引入人工复核环节,并邀请受影响群体代表参与治理过程^[67]。

同意能力波动在认知障碍人群中尤为突出。推荐采用“过程同意”模式:不仅在首次签署时获得同意,还需在每次沉浸式训练或评估前再次确认理解与意愿,并在会话过程中持续监测不适信号,允许随时退出。如参与者被判定缺乏完全能力,则由法律授权代理人代行,但仍需最大化尊重患者即时意愿并避免不必要的负担^[68-69]。

另外,人格与尊严保护同样重要。沉浸式场景可能引发“去人性化”体验,如过度控制或欺骗性设计。为避免负面影响,应限制单次会话时长、设置安全退出机制,并避免引发持久心理创伤的情境。

4.3 未来研究路线图 为推动该领域的稳健发展,未来研究应重点聚焦于以下几个方面。(1)临床证据积累:开展大规模、多中心、长期随访的随机对照试验,以形成高等级证据并检验其真实世界效用。(2)机制与数据融合:推动机制导向的生物学模型与数据驱动的计算模型相结合,并引入因果推断方法,以揭示 AD 进展的关键机制并指导靶向干预。

(3)平台与数据共享:建立开放协作的研究平台和标准化数据资源,促进跨学科和跨地域的科研合作。(4)伦理与监管框架:由技术专家、临床医生、伦理学者和患者代表共同制定专门的伦理与监管指南,确保技术发展的公正性与包容性。

5 小 结

AI与元宇宙的融合正在为AD管理开辟新路径:(1)沉浸式任务可捕捉多模态行为数据并经AI融合为数字生物标志物,用于早期识别与风险分层;(2)基于评估任务改造的VR训练与数字孪生则支持个体化干预、病程定位与反事实模拟;(3)沉浸式脚本化培训为照护与医护提供复用的场景演练和决策支持。未来,我们还需确保高质量证据、合规与伦理保障,以及跨设备和跨中心的互操作性与可推广性,以实现AD管理元宇宙的常规化应用,。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 林继先:查阅文献,撰写、修改论文;王华:查阅文献、修改论文;汤罗嘉:选题、策划、修改论文。

参考文献

- [1] SANTOS V D, COSTA A C, JUNIOR N C, et al. Virtual reality interventions and their effects on the cognition of individuals with Alzheimer's disease: a systematic review and meta-analysis[J]. *J Alzheimers Dis*, 2025, 103(1): 68-80.
- [2] KRUSE C S, SEN K Y, ARMENTA V, et al. Leveraging mHealth and virtual reality to improve cognition for Alzheimer's patients: a systematic review[J]. *Healthcare (Basel)*, 2022, 10(10): 1845.
- [3] YI Y L, HU Y Y, CUI M X, et al. Effect of virtual reality exercise on interventions for patients with Alzheimer's disease: a systematic review[J]. *Front Psychiatry*, 2022, 13: 1062162.
- [4] CLAY F, HOWETT D, FITZGERALD J, et al. Use of immersive virtual reality in the assessment and treatment of Alzheimer's disease: a systematic review[J]. *J Alzheimers Dis*, 2020, 75(1): 23-43.
- [5] KUSORO O, ROCHE M, DEL-PINO-CASADO R, et al. Time to diagnosis in dementia: a systematic review with meta-analysis[J]. *Int J Geriatr Psychiatry*, 2025, 40(7): e70129.
- [6] GIEBEL C, SILVA-RIBEIRO W, WATSON J, et al. A systematic review on the evidence of misdiagnosis in dementia and its impact on accessing dementia care[J]. *Int J Geriatr Psychiatry*, 2024, 39(10): e6158.
- [7] CUMMINGS J, APOSTOLOVA L, RABINOVICI G D, et al. Lecanemab: appropriate use recommendations[J]. *J Prev Alzheimers Dis*, 2023, 10(3): 362-377.
- [8] WIMO A, PRINCE M, GAUTHIER S. Global estimates of informal care[J]. 2018.
- [9] YANG Q, ZHANG L X, CHANG F Y, et al. Virtual reality interventions for older adults with mild cognitive impairment: systematic review and meta-analysis of randomized controlled trials[J]. *J Med Internet Res*, 2025, 27: e59195.
- [10] WANG Y T, SU Z, ZHANG N, et al. A survey on metaverse: fundamentals, security, and privacy[J]. *IEEE Commun Surv Tutor*, 25(1): 319-352.
- [11] SOLIMAN M M, AHMED E, DARWISH A, et al. Artificial intelligence powered Metaverse: analysis, challenges and future perspectives[J]. *Artif Intell Rev*, 2024, 57(2): 36.
- [12] HUYNH-THE T, PHAM Q V, PHAM X Q, et al. Artificial intelligence for the metaverse: a survey[J]. *Eng Appl Artif Intell*, 2023, 117: 105581.
- [13] GHAZINOORY S, PARVIN F, SAGHAFI F, et al. Metaverse technology tree: a holistic view[J]. *Front Artif Intell*, 2025, 8: 1545144.
- [14] STAUFFERT J P, NIEBLING F, LATOSCHIK M E. Latency and cybersickness: impact, causes, and measures. A review[J]. *Front Virtual Real*, 2020, 1: 582204.
- [15] WARBURTON M, MON-WILLIAMS M, MUSHTAQ F, et al. Measuring motion-to-photon latency for sensorimotor experiments with virtual reality systems[J]. *Behav Res Methods*, 2023, 55(7): 3658-3678.
- [16] ERGEN M, SAOUD B, SHAYEA I, et al. Edge computing in future wireless networks: a comprehensive evaluation and vision for 6G and beyond[J]. *ICT Express*, 2024, 10(5): 1151-1173.
- [17] LAKSHMINARAYANAN V, RAVIKUMAR A, SRIRAMAN H, et al. Health care equity through intelligent edge computing and augmented reality/virtual reality: a systematic review[J]. *J Multidiscip Healthc*, 2023, 16: 2839-2859.
- [18] KUNZ L, SCHRÖDER T N, LEE H, et al. Reduced grid-cell-like representations in adults at genetic risk for Alzheimer's disease[J]. *Science*, 2015, 350(6259): 430-433.
- [19] COUGHLAN G, LACZÓ J, HORT J, et al. Spatial navigation deficits—overlooked cognitive marker for preclinical Alzheimer disease?[J]. *Nat Rev Neurol*, 2018, 14: 496-506.
- [20] ALLISON S L, FAGAN A M, MORRIS J C, et al. Spatial navigation in preclinical Alzheimer's disease[J]. *J Alzheimers Dis*, 2016, 52(1): 77-90.
- [21] BUELE J, VARELA-ALDÁS J L, PALACIOS-NAVARRO G. Virtual reality applications based on instrumental activities of daily living (iADLs) for cognitive intervention in older adults: a systematic review[J]. *J Neuroeng Rehabil*, 2023, 20(1): 168.
- [22] SAGASPE P, AMIEVA H, DARTIGUES J F, et al. Validity and diagnostic performance of a virtual reality-based supermarket application "MemoShop" for assessing episodic memory in normal and pathological aging[J]. *Digit Health*, 2023, 9: 20552076231218808.
- [23] YAN M L, YIN H R, MENG Q Y, et al. A virtual supermarket program for the screening of mild cognitive impairment in older adults: diagnostic accuracy study[J]. *JMIR Serious Games*,

- 2021, 9(4): e30919.
- [24] KOURTESIS P, COLLINA S, DOUMAS L A A, et al. Validation of the virtual reality everyday assessment lab (VR-EAL): an immersive virtual reality neuropsychological battery with enhanced ecological validity [J]. *J Int Neuropsychol Soc*, 2021, 27(2): 181–196.
 - [25] NTT. The digital future: battling dementia with brain bio-digital twins [EB/OL]. [2025-07-28]. https://group.ntt/en/magazine/blog/brain_bio-Digital_twins/.
 - [26] WICKRAMASINGHE N, ULAPANE N, ANDARGOLI A, et al. Digital twins to enable better precision and personalized dementia care[J]. *JAMIA Open*, 2022, 5(3): ooac072.
 - [27] WOLF A, TRIPANITAK K, UMEDA S, et al. Eye-tracking paradigms for the assessment of mild cognitive impairment: a systematic review[J]. *Front Psychol*, 2023, 14: 1197567.
 - [28] NOROUZI M, KAFIEH R, CHAZOT P, et al. Insights from the eyes: a systematic review and meta-analysis of the intersection between eye-tracking and artificial intelligence in dementia[J]. *Aging Ment Health*, 2025, 29(8): 1367–1375.
 - [29] ZHANG X N, ZHANG F F, HOU S J, et al. Effectiveness of digital screening tools in detecting cognitive impairment among community-dwelling elderly in Northern China: a large cohort study[J]. *J Prev Alzheimers Dis*, 2025, 12(3): 100080.
 - [30] XU Y, ZHANG C, PAN B B, et al. A portable and efficient dementia screening tool using eye tracking machine learning and virtual reality[J]. *NPJ Digit Med*, 2024, 7(1): 219.
 - [31] GRAMKOW M H, CLEMMENSEN F K, LINDBERG U, et al. Prognostic value of light reflex pupillometry in Alzheimer's disease—a longitudinal cohort study[J]. *Alzheimers Res Ther*, 2025, 17(1): 152.
 - [32] ZHONG Y, HUANG S Q, ZOU M X, et al. Gait analysis in older adults with mild cognitive impairment: a bibliometric analysis of global trends, hotspots, and emerging frontiers[J]. *Front Aging*, 2025, 6: 1592464.
 - [33] MUKLI P, DETWILER S, OWENS C D, et al. Gait variability predicts cognitive impairment in older adults with subclinical cerebral small vessel disease[J]. *Front Aging Neurosci*, 2022, 14: 1052451.
 - [34] MONTERO-ODASSO M M, SARQUIS-ADAMSON Y, SPEECHLEY M, et al. Association of dual-task gait with incident dementia in mild cognitive impairment: results from the gait and brain study [J]. *JAMA Neurol*, 2017, 74 (7) : 857–865.
 - [35] SALZMAN T, LAURIN E, THIBAUT C, et al. A systematic review and meta-analysis of dual-task outcomes in subjective cognitive decline [J]. *Alzheimers Dement (Amst)*, 2025, 17 (1): e70054.
 - [36] SCHAUMBURG M, IMTIAZ A, ZHOU R J, et al. Immersive virtual reality for older adults: challenges and solutions in basic research and clinical applications [J]. *Ageing Res Rev*, 2025, 109: 102771.
 - [37] ABDLKARIM D, DI LUCA M, AVES P, et al. A methodological framework to assess the accuracy of virtual reality hand-tracking systems: a case study with the Meta Quest 2[J]. *Behav Res Methods*, 2024, 56(2): 1052–1063.
 - [38] BORZELLI D, BOARINI V, CASILE A. A quantitative assessment of the hand kinematic features estimated by the oculus Quest 2[J]. *Sci Rep*, 2025, 15: 8842.
 - [39] QI W H, ZHU X H, WANG B, et al. Alzheimer's disease digital biomarkers multidimensional landscape and AI model scoping review[J]. *NPJ Digit Med*, 2025, 8(1): 366.
 - [40] AMINI S, HAO B R, YANG J M, et al. Prediction of Alzheimer's disease progression within 6 years using speech: a novel approach leveraging language models [J]. *Alzheimers Dement*, 2024, 20(8): 5262–5270.
 - [41] ZHANG R, SHENG J H, ZHANG Q, et al. A review of multimodal fusion-based deep learning for Alzheimer's disease [J]. *Neuroscience*, 2025, 576: 80–95.
 - [42] CHEN J, HAO C, ZHANG X, et al. An intelligent screener for mild cognitive impairment *via* integrated eye-tracking and the digital clock drawing test [J]. *J Alzheimers Dis*, 2025: 13872877251350101.
 - [43] JABIN M S R, YAROSON E V, ILODIBE A, et al. Ethical and quality of care-related challenges of digital health twins in older care settings: protocol for a scoping review [J]. *JMIR Res Protoc*, 2024, 13: e51153.
 - [44] CHEN I C, WU C Y, HU Y L, et al. Development and validation of virtual reality cognitive training for older adults with mild cognitive impairment: protocol for a mixed-methods program evaluation study [J]. *Clin Interv Aging*, 2024, 19: 1855–1865.
 - [45] YANG J G, THAPA N, PARK H J, et al. Virtual reality and exercise training enhance brain, cognitive, and physical health in older adults with mild cognitive impairment[J]. *Int J Environ Res Public Health*, 2022, 19(20): 13300.
 - [46] DRUMMOND D, GONSARD A. Definitions and characteristics of patient digital twins being developed for clinical use: scoping review[J]. *J Med Internet Res*, 2024, 26: e58504.
 - [47] KATSOULAKIS E, WANG Q, WU H M, et al. Digital twins for health: a scoping review [J]. *NPJ Digit Med*, 2024, 7 (1): 77.
 - [48] YOUNG A L, MARINESCU R V, OXTOBY N P, et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference [J]. *Nat Commun*, 2018, 9(1): 4273.
 - [49] CS P, NP O, AL Y, et al. Parsimonious EBM: Generalising the event-based model of disease progression for simultaneous events[J]. *Neuroimage*, 2025, 311: 121162.
 - [50] FONTEIJN H M, MODAT M, CLARKSON M J, et al. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease [J]. *Neuroimage*, 2012, 60(3): 1880–1889.
 - [51] AMATO L G, LASSI M, VERGANI A A, et al. Digital twins and non-invasive recordings enable early diagnosis of Alzheimer's disease [J]. *Alzheimers Res Ther*, 2025, 17 (1): 125.

- [52] WU C Y, CHEN L, DICKSON J R, et al. Synthetic control methods for n-of-1 and parallel-group trials in Alzheimer's disease: a proof-of-concept study using the I-CONNECT [J]. *Alzheimers Dement*, 2025, 21(7): e70460.
- [53] RIAHI V, DIOUF I, KHANNA S, et al. Digital twins for clinical and operational decision-making: scoping review [J]. *J Med Internet Res*, 2025, 27: e55015. [PubMed]
- [54] SHEN D, AGARWAL A, MISRA V, et al. Obtaining personalized predictions from a randomized controlled trial on Alzheimer's disease [J]. *Sci Rep*, 2025, 15(1): 1671.
- [55] TAN K, BRYAN J, SEGAL B, et al. Emulating control arms for cancer clinical trials using external cohorts created from electronic health record-derived real-world data [J]. *Clin Pharmacol Ther*, 2022, 111(1): 168–178.
- [56] GIUFFRÈ M, SHUNG D L. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy [J]. *NPJ Digit Med*, 2023, 6(1): 186.
- [57] INOUE K, IIZUKA H, KAMIO H, et al. A pilot study on the use of virtual reality to educate community members about dementia [J]. *Phys Occup Ther Geriatr*, 2024, 42 (2) : 170–185.
- [58] WANG J Y, LI Q, CUI J, et al. Effectiveness of virtual reality on the caregiving competence and empathy of caregivers for elderly with chronic diseases: a systematic review and meta-analysis [J]. *J Nurs Manag*, 2023, 2023: 5449955.
- [59] GUEORGUEVA I, CHOW K, CHUA L, et al. Donanemab exposure-response in early symptomatic Alzheimer's disease [J]. *Alzheimers Dement*, 2025, 21(7): e70491.
- [60] STAFFORD A, BENDER S, PARSONS K, et al. Evaluating a virtual reality dementia training experience using psychophysiological methods: a randomised controlled study [J]. *Australas J Ageing*, 2024, 43(3): 523–532.
- [61] HUNG L, MANN J, WALLSWORTH C, et al. Facilitators and barriers to using virtual reality and its impact on social engagement in aged care settings: a scoping review [J]. *Gerontol Geriatr Med*, 2023, 9: 23337214231166355.
- [62] DE DOMENICO M, ALLEGRI L, CALDARELLI G, et al. Challenges and opportunities for digital twins in precision medicine from a complex systems perspective [J]. *NPJ Digit Med*, 2025, 8(1): 37.
- [63] SEL K, HAWKINS-DAARUD A, CHAUDHURI A, et al. Survey and perspective on verification, validation, and uncertainty quantification of digital twins for precision medicine [J]. *NPJ Digit Med*, 2025, 8(1): 40.
- [64] HEMDAN E E, SAYED A. Smart and secure healthcare with digital twins: a deep dive into blockchain, federated learning, and future innovations [J]. *Algorithms*, 2025, 18(7): 401.
- [65] UMAIR M, KHAN M S, HANIF M, et al. Privacy-preserving dementia classification from EEG via hybrid-fusion EEGNetv4 and federated learning [J]. *Front Comput Neurosci*, 2025, 19: 1617883.
- [66] OBERMEYER Z, POWERS B, VOGELI C, et al. Dissecting racial bias in an algorithm used to manage the health of populations [J]. *Science*, 2019, 366(6464): 447–453.
- [67] FORD E, MILNE R, CURLEWIS K. Ethical issues when using digital biomarkers and artificial intelligence for the early detection of dementia [J]. *Wiley Interdiscip Rev Data Min Knowl Discov*, 2023, 13(3): e1492.
- [68] PYER M, WARD A. Developing a dementia friendly approach to consent in dementia research [J]. *Aging Ment Health*, 2024, 28(2): 294–301.
- [69] EVANS C J, YORGANCI E, LEWIS P, et al. Processes of consent in research for adults with impaired mental capacity nearing the end of life: systematic review and transparent expert consultation (MORECare_Capacity statement) [J]. *BMC Med*, 2020, 18(1): 221.

引用本文

林继先,王 华,汤罗嘉. 构建阿尔茨海默病的未来:从早期诊断到个性化干预的人工智能元宇宙[J]. 元宇宙医学,2025,2(3):11–18.

LIN J X, WANG H, TANG L J. Building the future of Alzheimer's disease: an AI-driven metaverse from early diagnosis to personalized intervention [J]. *Metaverse Med*, 2025, 2(3): 11–18.

DOI: 10.61189/524688gdrjug

· 综述 ·

虚拟现实在心肺复苏培训中的应用与进展:挑战与展望



甘舜轩*

南京柏拓视觉科技有限公司, 南京 210019

[摘要] 心肺复苏(cardiopulmonary resuscitation, CPR)是挽救心搏骤停患者生命的核心技能,其普及程度直接决定了院外抢救的成功率。然而,无论在公众教育还是专业医护培训中,传统CPR教学方式仍存在很多局限。近年来,随着人工智能、虚拟现实(virtual reality, VR)、增强现实以及元宇宙相关技术的发展,沉浸式交互训练逐渐被用于医学教育领域,成为CPR培训模式创新的重要方向。本文综述了CPR VR系统的研究进展和挑战,为未来技术开发提供思路。

[关键词] 心肺复苏;虚拟现实;元宇宙;医学教育;沉浸式培训;人工智能;技能迁移

[中图分类号] R-1 **[文献标志码]** A

Application and progress of virtual reality in cardiopulmonary resuscitation training: challenges and prospects

GAN Shunxuan*

Nanjing Baituo Visual Technology Co., Ltd, Nanjing 210019, Jiangsu, China

[Abstract] Cardiopulmonary resuscitation (CPR) is a critical life-saving skill, and its widespread adoption is closely associated with the survival rate of out-of-hospital cardiac arrest. However, traditional CPR training remains several limitations. With the rapid development of artificial intelligence, virtual reality (VR), augmented reality, and metaverse-related technologies, immersive training has emerged as a promising innovation in medical education, particularly in CPR instruction. This article reviews the research progress and challenges of VR-based CPR system, and provides ideas for future technology development.

[Key Words] cardiopulmonary resuscitation; virtual reality; metaverse; medical education; immersive training; artificial intelligence; skill transfer

近年来,随着虚拟现实(virtual reality, VR)、增强现实(augmented reality, AR)及混合现实(mixed reality, MR)技术的不断成熟,基于沉浸式交互的医学教育模式逐渐受到重视,其中心肺复苏(cardiopulmonary resuscitation, CPR)培训成为研究与应用的重点方向之一。

1 CPR VR 相关研究综述

1.1 全球CPR VR培训发展概况 欧美国家的CPR VR教育起步较早。早在21世纪10年代,美国心脏协会(American Heart Association, AHA)和欧洲复苏委员会(European Resuscitation Council, ERC)就提出应探索数字化和模拟技术在急救培训中的应用,以弥补传统假人培训的不足^[1]。VR与AR在CPR教学中的应用不仅提升了学习者的操作兴趣和沉浸感,还提高了技能掌握率,改善了记忆保持效果。

欧美的一些高校和急救机构已经将CPR VR模块纳入医学教育课程与继续教育项目,并尝试用于CPR科普。

在亚洲,尤其是中国、日本和韩国,CPR VR的研究与应用正处于快速发展阶段。例如,日本研究者开发了一套用于日本院前创伤急救评估(Japan Advanced Trauma Evaluation and Care, JPTEC)课程中初评环节的VR系统,用来替代传统的30 min面对面培训。该VR系统简单易用,即使没有专业技能,也可仅用2 d就完成VR内容制作。学员可在该系统内反复练习CPR,仅需15 min,就可以显著提升学员的实操测试通过率^[2]。在中国,一些高校和医院也已经开始探索低成本的CPR VR方案,以期解决师资短缺与场地限制的问题。然而,亚洲地区的研究多集中于应用探索和效果评估,尚缺乏大规模的随机对照试验和长期随访数据。

[收稿日期] 2025-09-08

[接受日期] 2025-09-24

[作者简介] 甘舜轩.

*通信作者(Corresponding author). Tel: 025-58715810, E-mail: 33971873@qq.com

1.2 经典研究与典型案例

1.2.1 相关研究 已有多项研究验证了CPR VR培训对提高操作准确性和学习主动性的积极效果。Sun等^[3]对15项研究进行荟萃分析,发现VR或AR培训效果在按压深度、频率及整体掌握度方面与传统培训相当;部分研究显示,VR组在某些指标上可能优于传统培训组。Nikiema等^[4]的系统综述指出,基于VR的严肃游戏能够显著提升学习者的CPR技能获取、自我效能感及知识保持。Artero等^[5]的随机对照试验结果显示,VR训练可以有效提升学员的按压正确率与自信心,尤其适用于混合教学模式。

1.2.2 产品与系统 CPR VR已经逐渐从“概念验证”走向“临床教育与公众普及”的双重应用场景。目前,市场上已出现多种CPR VR产品。(1)CPR VR App:部分基于Oculus、HTC Vive平台,能够提供虚拟场景与语音或视觉反馈。(2)Haptic-enabled systems:结合数字人和VR场景的系统,如Yigitbas等^[6]提出的AR或VR BLS培训平台,能够提供触觉反馈和自动化评估。(3)移动端:已有团队开始尝试开发轻量化的手机+简易VR头显CPR培训工具,以降低成本并支持大规模推广。

1.3 教学效果评估与比较 CPR VR培训已展现出多种优势。(1)提升操作准确性:研究^[7]表明,VR培训组的按压频率、深度及位置的准确性显著高于对照组。(2)提高学习效率:VR培训能够在更短时间内帮助学员掌握基本技能,尤其适合时间有限的非专业群体。(3)促进记忆保持:研究显示,VR训练在技能保持方面优于传统培训。一周或数周后的追踪测试中,VR组正确率下降幅度更小。(4)优化学习体验:VR训练提供更强的沉浸感和参与度,受训者普遍反馈“更有趣、更真实”,从而激发学习者的自主性。

然而,CPR VR也存在多处不足。(1)迁移效度问题:目前缺乏充分证据证明VR训练技能在真实急救环境中完全迁移。(2)个体差异:部分受训者出现VR晕动症或不适感,影响培训效果。(3)成本限制:高端VR硬件的价格较高,阻碍了相关VR产品在资源有限环境中的普及^[8]。

现有研究证明了VR或AR在CPR培训中的应用价值,它能够提高操作准确率、学习效率和主观能动性,为传统培训提供有力补充。然而,CPR VR在大规模推广、真实环境迁移、数据安全和成本控制等方面仍存在挑战。未来,需要进行更多高质量的随机对照试验和长期随访研究,同时结合人工智能(artificial intelligence, AI)与多模态交互技

术,以推动元宇宙医学教育体系的发展。

2 实现CPR VR培训的技术手段

基于VR的CPR培训系统能够提供沉浸式学习体验,其沉浸式效果很大程度上取决于所采用的技术。目前,相关技术研究与应用主要集中在硬件设备、软件平台、交互方式以及数据驱动的反馈机制4个层面。

2.1 硬件设备

2.1.1 头戴式显示器(head-mounted display, HMD) HMD是VR系统的核心组件,目前主流设备包括Oculus Quest、HTC Vive、Pico Neo等。这些设备能够提供沉浸式三维场景,并具备空间定位和运动追踪功能,使学习者在虚拟环境中能够完成CPR操作的空间定向。高端HMD支持6自由度(6DoF)追踪,可准确捕捉学习者头部与上肢的动作,沉浸式进行胸外按压和人工呼吸操作。

2.1.2 动作捕捉与传感器 手势识别与动作捕捉是CPR VR系统的关键环节。目前主要采用红外光学传感器(如leap motion)、惯性测量单元(IMU)以及摄像头结合的多模态追踪技术。这些传感器可实时捕捉学习者的双手位置、按压动作与角度,从而评估学习者的按压位置是否正确。部分系统还配备力敏传感器,以测量按压深度和力度。但是,光学传感器的成本比较高昂,并且对于设备的性能要求比较高,目前在实际开发过程中主要会采用摄像头进行用户行为的捕捉。无论使用哪种动作捕捉方式,由于目前HMD的普及程度远低于手机、平板等消费电子产品,其手势操作方案多样且缺乏统一标准,从而显著增加了用户的学习成本。因此,在现阶段,动作捕捉技术的应用应着眼于夯实基础功能的用户体验,而非追求功能的多样性,这是降低技术门槛的关键。

2.1.3 力反馈装置与触觉交互 传统CPR假人的优势在于能够提供真实阻力和反馈。因此,一些CPR VR系统尝试结合力反馈装置或混合模拟假人。例如,由浙江大学医学院附属邵逸夫医院设计开发的CPR VR训练平台(图1),采用VR头显与传感器集成的模拟假人作为核心交互设备,并创新地简化CPR流程和教学模式,集中优化按压力度、频率与时间等关键参数的实时反馈,在维持虚拟场景的沉浸感的同时,提供了触觉反馈,提升了操作真实性。

2.2 软件平台

2.2.1 虚拟患者生成 相关软件通常基于Unity或Unreal Engine开发,能够模拟院外心脏骤停等急救

场景。场景包含虚拟患者、环境噪音、围观人群等要素,以增加紧迫感 and 情境真实性。部分系统还支持

多种难度等级,从单一心脏骤停患者到复杂环境中的多任务急救,便于分层教学。



图1 CPR VR 产品介绍

2.2.2 实时反馈模块 实时反馈是CPR VR系统的核心优势之一。系统可通过颜色提示、语音指导或图形化界面显示学员的按压深度、频率和手位偏差(图2)。例如,当按压过浅时,系统会给出“压得不够深”的提示;当频率过快时,则提醒“请降低节奏”。这种即时反馈有助于纠正错误并增强学习记忆。

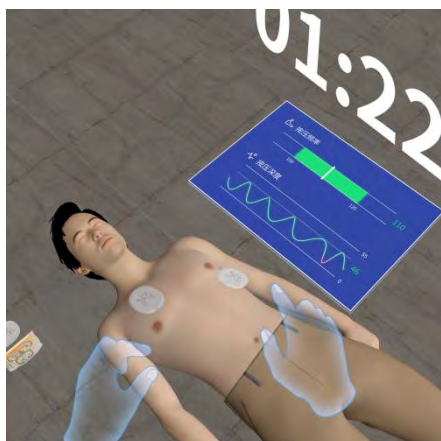


图2 CPR VR 产品图形化界面

2.2.3 AI教学与个性化学习 近年来,AI被逐渐应用至CPR VR培训。AI模块能够根据学习者的操作数据生成综合评分,并提供个性化改进建议^[9]。一

些研究提出可以利用机器学习预测学员的技能掌握曲线,从而调整训练强度和内容,定制个体化学习方案。

2.3 交互方式

2.3.1 手势与语音交互 手势识别允许学员在VR场景中完成按压动作,而无需外接复杂控制器。同时,语音交互可模拟急救指令场景,例如学员对虚拟旁人发出“拨打120”的指令,以提升团队协作与临场感。

2.3.2 多用户协作模式 部分VR平台支持多人协同训练。学员可在同一虚拟场景中分工合作,例如一人负责胸外按压,另一人负责人工呼吸或操作自动体外除颤器。这种模式不仅增强互动性,还能训练团队的急救配合能力。

2.3.3 混合现实与增强现实 相比单一VR技术,增强现实(augmented reality, AR)与混合现实(mixed reality, MR)技术能够在现实环境中叠加虚拟元素。例如,通过AR眼镜在假人模型上叠加实时按压反馈参数,实现虚实结合的训练模式。这种教学方式即保留了假人的物理阻力,又增加了数据化反馈,更容易推广。但由于AR眼镜本身的性能以及产品显示技术的限制,导致AR眼镜易受周围复杂光学环境的干扰,沉浸感不强,叠加的信息无法清晰显

示,结构配重不合理导致无法长时间佩戴,未来仍需要进一步完善。

2.4 数据驱动与反馈机制 数据采集与分析是 CPR VR 系统的另一大优势。采集的数据主要包括动作数据(按压深度、频率、角度、位置偏差),时间数据(反应时间、操作流程完成时间),生理数据(心率、呼吸频率)。这些数据不仅用于学员个人反馈,也可上传至云端,支持教师或机构的批量分析与远程监督。未来,结合大数据与 AI 算法,CPR VR 系统可为学员建立个人技能档案,实现纵向跟踪和动态改进。

当前 CPR VR 培训的技术手段呈现出“硬件—软件—交互—数据”的综合特征。在硬件方面,HMD、动作捕捉与力反馈假人共同构建沉浸体验;灶软件方面,虚拟病人场景与实时反馈模块提高学习效果;在交互方面,手势、语音及多人协作增强了临场感与真实性;在数据层面,系统能够实现量化评估与个性化学习路径。

尽管技术发展迅速,但在成本控制、硬件普及度、交互精度和用户体验稳定性方面仍有挑战。这些问题的解决,将推进 CPR VR 系统在元宇宙医学教育框架下实现大规模应用。

3 挑战与不足

尽管基于 VR 技术的 CPR 培训取得了积极进展,相关研究也不断涌现,但从现有证据和应用情况来看,CPR VR 系统在教育推广和实际应用中仍然面临诸多挑战与不足。这些问题主要集中在技能迁移效率、硬件与成本限制、用户体验、技术精度、数据与伦理风险等方面。

3.1 技能迁移与有效性不足 虚拟环境能否真实转化为临床与公众急救情境中的有效技能,是 CPR VR 培训面临的核心问题之一。部分随机对照试验表明,VR 训练可以显著提升学员在模拟测试中的表现,但对于真实院外急救事件中的表现尚缺乏直接证据。目前,多数研究的随访期较短,难以评估长期记忆保持和行为迁移效果^[10]。此外,VR 训练在心理压力模拟方面仍有不足,无法完全再现真实急救情境中的紧张氛围,这可能影响其在真实场景中的应用价值。

3.2 硬件成本与推广局限 高端 VR 头显、动作捕捉传感器和力反馈设备的采购与维护成本较高,限制了其在学校、社区和基层医疗机构中的大规模普及。尽管已有研究尝试开发轻量化、低成本的移动端解决方案,但这些系统在交互精度、反馈真实性

方面往往有所妥协。同时,设备更新迭代快,培训机构在投资时可能面临“技术折旧”的风险。对于资源有限的发展中国家而言,硬件成本高是 CPR VR 难以大规模推广的主要原因之一。

3.3 用户体验与接受度问题 VR 系统的使用体验直接影响学习效果。部分受训者在使用 VR 设备时出现头晕、恶心等症状,导致培训中断或效果下降。此外,不同年龄群体和技术熟悉程度的学员在接受度上存在差异:年轻人普遍适应较快,而年长的学员可能因系统操作复杂而降低其学习兴趣。VR 培训在提升学习兴趣方面虽具优势,但若体验不佳,可能反而影响参与度和普及性^[10]。

3.4 技术精度与标准化不足 动作捕捉和力度检测的精度仍是限制 CPR VR 培训效果的关键因素。目前多数 VR 系统能够提供按压频率与深度的反馈,但在手位偏差、胸廓回弹等关键操作的检测上仍不够精确。各系统评估指标与算法差异明显,缺乏统一规范,导致培训效果存在差异^[3],亟需行业标准化建设。

3.5 数据安全与伦理挑战 CPR VR 系统往往需要采集学员的动作数据、操作表现,甚至包括心率等生理参数。这些数据一旦上传至云端,可能面临隐私泄露与滥用的风险。此外,随着 AI 教练和大数据分析的应用,如何确保数据处理的透明性与合规性,成为亟待解决的问题。医学教育类期刊近年逐渐强调在虚拟教学研究中应包含伦理审批与数据保护说明,这提示未来 CPR VR 研究必须在设计阶段就重视隐私与伦理规范。

CPR VR 培训在技术创新和教育效果上展现出巨大潜力,但当前仍处于探索与过渡阶段。其推广与应用受到技能迁移证据不足、硬件成本高、用户体验不稳定、技术精度有限以及数据安全问题等多方面挑战的制约。未来发展需要更多高质量的随机对照试验与长期随访研究,同时推动行业标准化建设与成本优化,并在数据安全与伦理规范方面建立健全的框架,才能真正实现 CPR VR 的大规模应用和价值落地。

4 展望

随着 VR、AR、AI 和元宇宙概念的不断融合,CPR 培训在未来的发展方向逐步清晰。虽然现有研究已证实 VR 在提升学习兴趣、操作准确率和促进技能保持方面的潜力,但要真正实现 VR 在医学教育与公共卫生中的大规模应用,仍需从技术升级、个性化教育、跨学科融合、推广策略与政策标准

化等方面持续推进。

4.1 数字孪生与个性化学习路径 未来 CPR VR 系统有望结合数字孪生技术,为学习者建立个体化虚拟模型,动态跟踪学习进展并模拟真实操作表现。通过采集多模态数据(动作、语音、生理指标),系统可生成个性化学习曲线,并结合 AI 教练提供定制化训练方案。例如,对于按压深度不足的学员,系统可自动增加相关模块的训练频率,并提供针对性的语音或视觉提示。这种基于个体化数据驱动的学习方式,有助于解决当前“一刀切”培训效果存在个体差异的问题。

4.2 AI 教学与智能反馈 AI 技术的引入将使 CPR VR 培训迈向智能化。未来,系统不仅能实时评估学员操作,还能借助机器学习模型预测学员的潜在错误,并提前干预。AI 教练可以在虚拟场景中扮演导师,提供自然语言交流、即时纠错与心理激励,从而增强学习者的沉浸感和学习黏性。同时,结合大语言模型与多模态交互,系统可实现语音、手势、眼动等多通道输入,进一步提升训练的自然性与交互性。

4.3 多模态交互与沉浸式体验升级 未来 CPR VR 培训不应仅局限于手部动作识别和力反馈,而应扩展至多模态交互,例如以下几方面。(1)语音交互:允许学员在模拟场景中向虚拟旁人发出急救指令;(2)眼动追踪:用于评估学员在紧急情境下的注意力分布,帮助改进教学策略;(3)全身动作捕捉:训练团队协作与整体应急流程,而不仅是胸外按压环节。此外,随着触觉反馈技术与可穿戴设备的发展,学员可获得更真实的力反馈与身体沉浸感,从而提高训练的真实性与学习迁移效率。

4.4 公共健康普及与大规模推广 在应用层面,未来 CPR VR 系统有望在学校、社区、企业或事业单位中广泛部署,实现公众 CPR 教育常态化。轻量化、低成本的移动端 VR 方案将成为大规模推广的关键。例如,利用智能手机和便携式 VR 头显开展 CPR 基础培训,辅以 AI 评分系统和云端平台进行远程监督。这种模式不仅降低了成本,还能突破地域与师资限制,使 CPR 教育得到普及。

4.5 跨学科融合与元宇宙医学生态 CPR VR 的发展不仅是技术升级,更是跨学科融合的结果。构建完整的元宇宙医学生态,亟需医学、教育学、计算机科学、人机交互、心理学与公共政策等学科的协同共建,以形成坚实的支撑体系。在该生态中,CPR VR 将作为重要应用场景之一,与其他虚拟医疗教

育、手术模拟、慢病康复平台共同发展,形成覆盖教学、临床和公共健康的闭环。

4.6 标准化与政策支持 目前,CPR VR 培训尚缺乏统一的行业标准,不同系统之间的评估指标差异显著,影响了研究结果的可比性。未来应由国际组织(如 AHA、ERC)牵头,制定基于 VR 的 CPR 培训技术规范 and 认证体系,将其纳入急救教育的标准流程。此外,政策支持至关重要。各国政府和医疗监管机构若能在经费补助、认证体系建设和推广政策上给予支持,将极大促进 CPR VR 技术的落地与普及。

未来 CPR VR 培训的发展应聚焦于数字孪生与个性化学习、AI 教练与智能反馈、多模态交互与沉浸升级、大规模公共普及、跨学科融合与元宇宙医学生态建设。在政策与标准的引导下,CPR VR 技术有望从教育创新工具演变为公共健康战略资源,成为提升社会整体急救能力与医疗教育水平的重要组成部分。

5 小 结

CPR 作为最关键的院外急救技能,其培训质量与普及率直接关系到患者生存率。近年来,VR、AR 及元宇宙相关技术为 CPR 教育带来了全新的发展机遇。CPR VR 培训在提高操作准确率、提升学习效率、促进记忆保持以及增加学习主观能动性方面均表现出一定优势,其可扩展性和可及性为大规模推广提供了新的可能。

然而,当前 CPR VR 的应用仍面临诸多挑战,包括虚拟训练到现实场景的技能迁移效率不足、硬件成本较高、用户体验差异明显、动作捕捉与反馈精度有限,以及数据隐私安全挑战与伦理问题等。这些因素在一定程度上限制了 CPR VR 的广泛推广和标准化应用。

我们认为,CPR VR 技术的下一个突破点在于深度智能化。其中,AI 教练的集成尤为关键。它不仅需提供实时纠错,更应能基于学习者历史数据(如错误模式、学习速度),构建个性化数字孪生模型,动态生成针对性的强化训练模块。要实现这一目标,首要解决的是高精度、低延迟的多模态数据融合问题(如动作、眼动、语音、生理信号),其次需要开发更强大的行为预测算法。相比之下,元宇宙概念的融合虽具想象空间,但在 CPR 培训的具体落地路径和实际教育价值上,目前仍需更多的可行性论证。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 甘舜轩:选题、撰写、修改。

参考文献

- [1] American Heart Association. 2020 American Heart Association guidelines for CPR & emergency cardiovascular care [J]. Circulation, 2020, 142(16 suppl 2).
- [2] KIYOZUMI T, ISHIGAMI N, TATSUSHIMA D, et al. Development of virtual reality content for learning Japan prehospital trauma evaluation and care initial assessment procedures[J]. Acute Med Surg, 2022, 9: e755.
- [3] SUN R, WANG Y X, WU Q Y, et al. Effectiveness of virtual and augmented reality for cardiopulmonary resuscitation training: a systematic review and meta-analysis [J]. BMC Med Educ, 2024, 24(1): 730.
- [4] NIKIEMA S L W, BARRO S G, KANTAGBA Y M K, et al. Advancing CPR training with virtual reality: a review [M]// Digital health and informatics innovations for sustainable health care systems. Amsterdam: IOS Press, 2024.
- [5] ALCÁZAR ARTERO P M, PARDO RIOS M, GREIF R, et al. Efficiency of virtual reality for cardiopulmonary resuscitation training of adult laypersons: a systematic review [J]. Medicine, 2023, 102(4): e32736.
- [6] YIGITBAS E, KROIS S, RENZELMANN T, et al. Comparative evaluation of AR-based, VR-based, and traditional basic life support training [C]. 2022 IEEE 10th International Conference on Serious Games and Applications for Health (SeGAH), Sydney, Australia; IEEE, 2022: 1-8
- [7] 王思瑶. 虚拟现实技术在医学生心肺复苏培训中的应用 [D]. 南方医科大学, 2019.
- [8] 李娜. 基于虚拟现实智能救护培训平台的培训需求调查及心肺复苏培训效果分析 [D]. 南昌大学, 2021.
- [9] American CPR Care Association. How AI revolutionizes CPR and AED training: key benefits [EB/OL]. [2025-08-12]. <https://cprcare.com/blog/use-of-ai-in-cpr-training/>.
- [10] 陈震. 基于虚拟现实技术的护理人员心肺复苏培训方案构建研究 [D]. 电子科技大学, 2024.

引用本文

甘舜轩. 虚拟现实在心肺复苏培训中的应用与进展:挑战与展望[J]. 元宇宙医学, 2025, 2(3): 19-24.

GAN S X. Application and progress of virtual reality in cardiopulmonary resuscitation training: challenges and prospects [J]. Metaverse Med, 2025, 2(3): 19-24.

·消息·

中山医院将建设国家人工智能应用中试基地

近日,国家发改委正式批复,由中山医院承担建设“国家人工智能应用中试基地(医疗领域临床医学科研方向)”的任务。该基地以“聚焦临床科研、破解行业痛点、构建生态标杆”为核心定位,致力于打造覆盖“研发—测试—验证—应用”全链条的创新支撑平台。

基地将建设国产化算力支撑平台,通过自建微调/推理算力集群,严格保障隐私数据“不出域”,为安全高效的数据处理与模型训练提供坚实基础。基地将打造全链条模型服务体系,以国产自主可控的基础大模型和AI框架为底座,构建涵盖模型工程、Agent(智能体)工程、AI资产管理在内的全流程服务能力。开发开放工具链,打造模块化应用开发组件,提升医疗多智能体系统构建效率,并向行业开放共享。

在具体应用层面,基地将面向医学科研、智能诊疗能力提升,智能医疗器械开发,智能医用机器人研发等方向,落地不少于49个特色应用场景,并发布超过10项配套标准规范。这些创新将显著提升诊断精准度和效率,使医务人员能够更专注于复杂病例的攻关和治疗方案的优化。同时,患者也将受益于更高效的诊断流程、个性化的治疗方案,以及从诊前咨询到术后随访的全流程优化服务。

来源:解放日报

记者:顾泳

2025 年 8 月 7 日

DOI: 10.61189/246300fnss

· 方法学研究 ·

大语言模型和专病 GPT 赋能 OSA 咨询和管理的临床应用效果对比研究方案

陆俊羽¹, 蔡沁怡², 白春学^{2*}

1. 重庆大学附属仁济医院(重庆市第五人民医院)呼吸与危重症医学科, 重庆 400062

2. 复旦大学附属中山医院呼吸与危重症医学科, 上海呼吸物联网医学工程技术研究中心, 上海市呼吸病研究所, 上海 200032

[摘要] 本研究方案旨在对比大语言模型(如 DeepSeek GPT)与专病 GPT(如 BAIMGPT)在阻塞性睡眠呼吸暂停(OSA)咨询和管理中的临床应用效果。采用多中心真实世界研究设计, 纳入 1000 例 OSA 患者或高风险人群, 通过用户横向评价和第三方专家审核, 评估两种模型在便捷性、亲切感、安全感、问题理解准确性、回答准确性、语音交互、视觉赋能及需求程度等方面的表现。研究重点关注两种模型在 OSA 筛查、诊断准确性和个性化防治中的作用, 并探讨其在提升患者教育、医生培训和基层医疗覆盖方面的潜力。研究结果将为人工智能技术在 OSA 诊疗中的优化应用提供实证依据, 推动精准医疗和健康管理的发展。

[关键词] 阻塞性睡眠呼吸暂停; BAIMGPT; DeepSeek GPT**[中图分类号]** R 563.8 **[文献标志码]** A

A comparative study plan on the clinical application effects of large language models and specialized GPTs in OSA consultation and management

LU Junyu¹, CAI Qinyi², BAI Chunxue^{2*}

1. The Fifth People's Hospital of Chongqing, Chongqing 400062, China

2. Department of Pulmonary and Critical Care Medicine, Zhongshan Hospital, Fudan University, Shanghai Engineer & Technology Research Center of Internet of Things for Respiratory Medicine, Shanghai Respiratory Research Institution, Shanghai 200032, China

[Abstract] This research plan aims to compare the clinical application effects of large language models (such as DeepSeek GPT) and specialized disease GPT (such as BAIMGPT) in the consultation and management of obstructive sleep apnea (OSA). A multi-center real-world research design is adopted, involving 1000 OSA patients or high-risk individuals. Through user cross-sectional evaluations and third-party expert reviews, the performance of the two models in aspects such as convenience, friendliness, security, accuracy of problem understanding, accuracy of answers, voice interaction, visual empowerment, and the degree of patient needs is assessed. The research focuses on the roles of the two models in OSA screening, diagnostic accuracy, and personalized prevention and treatment, and explores their potential in enhancing patient education, doctor training, and coverage of primary medical care. The research results will provide empirical evidence for the optimized application of artificial intelligence technology in OSA diagnosis and treatment, and promote the development of precision medicine and health management.

[Key Words] obstructive sleep apnea; BAIMGPT; DeepSeek GPT

咨询和管理在 OSA 的防治中扮演着至关重要的角色。首先, OSA 筛查可以帮助患者早诊早治, 了解病情, 理解治疗方案, 使他们愿意参与“云+端”互动筛查、诊断和管理, 从而更好地配合治疗^[1-3]。此外, 专家分身可以应用丰富的肢体语言和三维道

具以及图文知识, 有效地教会端医生或初诊医生掌握 OSA 的一二级预防策略。自我管理则是患者在日常生活中的实践, 可以通过使用 OSapp 5A 这样的工具来提高二级预防的效率, 特别是在提升患者的自我管理能力方面。自我管理不仅有助于患者更

[收稿日期] 2025-06-23**[接受日期]** 2025-09-15

[基金项目] 四大慢病重大专项(2024ZD0529300), 中国民族卫生协会“卫生健康技术重点推广项目”(XM20250388). Supported by Noncommunicable Chronic Diseases National Science and Technology Major Project (2024ZD0529300) and Key Promotion Project of Health and Wellness Technology of the Chinese Ethnic Health Association (XM20250388).

[作者简介] 陆俊羽, 博士. E-mail: junyulu@aliyun.com

*通信作者(Corresponding author). Tel: 021-64041990, E-mail: bai.chunxue@zs-hospital.sh.cn

好地控制症状,还能提高生活质量,减少并发症和共患病影响患者的风险。总之,咨询和管理对于提高 OSA 患者对疾病的认识、促进健康生活方式的改变、优化治疗效果以及降低并发症的发生率都具有重要的意义。因此,在 OSA 的防治中,加强咨询和管理是不可或缺的一部分。

目前,OSA 的咨询和管理不足导致 OSA 的防治也存在一些限制。首先,尽管很多学科已经开展了 OSA 的诊治工作,并且制定了不同的诊治共识指南,但依旧难以满足广大患者的需求。这可能是因为 OSA 的症状,如夜间打鼾、白天嗜睡等,可能被患者忽视,导致未能得到及时的诊断和治疗。此外,OSA 常与一些慢性疾病相关联,如高血压、中风、糖尿病等,而这些疾病的患者往往需要跨学科的综合治疗,这在实际医疗工作中可能会遇到协调和资源分配上的挑战。其次,在基层医疗机构中,由于缺乏专业的睡眠医学知识和设备,医生可能无法准确地对 OSA 进行筛查和评估。这可能导致一些高危个体未被及时发现,从而错失了早期干预的机会^[4-6]。在自我管理方面,患者对 OSA 的认识不足,可能是阻碍有效自我管理的一个主要障碍。许多患者可能不了解如何通过改变生活方式来缓解症状,或者不知道如何在家中正确使用无创机械通气等设备。此外,持续的教育和支持对于帮助患者坚持治疗计划至关重要,但在实践中可能难以实现。最后,隐私和数据安全问题在应用物联网技术进行 OSA 诊疗的过程中尤为重要,需要确保所有通过网络传输的数据都得到妥善的保护,避免泄露给未经授权的第三方。

简言之,目前在 OSA 防治工作中,咨询和管理的痛点和难点主要包括患者意识的提高、跨学科综合治疗的实施、基层医疗能力的提升、患者自我管理能力的增强以及隐私和数据安全保障挑战。解决这些问题需要多方面的努力,包括公众教育、政策引导、技术创新以及医疗服务模式的转变。

1 应用专病 GPT 解决目前 OSA 咨询和管理的问题

(1)提高筛查率:通过宣传和科普提高知晓率,使大众知道 OSA 对健康的危害,早诊早治的意义,推动筛查;(2)提高早诊率:应用从简单的 IV 级设备到复杂的 I 级设备监测,提高 OSA 的早期诊断率,同时节省医疗资源和费用;(3)提高诊断准确性:通过应用专病 GPT,可以对大量的医学文献和临床数据进行深度分析和处理,从而获得丰富的医学知识和推理能力。这些知识可用于指导医生进行更准确

的诊断,减少漏诊和误诊的发生。(4)个性化防治方案:专病 GPT 可以根据患者的具体情况,提供个性化的治疗方案和建议。这有助于确保每位患者都能获得最适合自己的防治方案,提高治疗效果和患者满意度。(5)增强患者自我管理意识:通过专病 GPT,患者可以获取有关 OSA 的详细知识和自我管理技巧。这有助于患者更好地了解自己的病情,掌握有效的自我管理方法,提高生活质量。(6)提高医生专业素养:专病 GPT 还可以用于医学教育和培训,帮助相关医生提高专业素养和服务能力。这将有助于医生更好地满足患者的咨询和管理需求,提升整体医疗服务水平。总之,应用专病 GPT 可以有效解决目前 OSA 咨询和管理存在的问题,提高诊断准确性、治疗个性化和患者自我管理能力^[7-9],进一步提升 OSA 综合诊疗水平。

但是,大语言模型 GPT 是否能精准地赋能诊疗和管理患者,仍不得而知。基于小规模的研究结果,大语言模型 GPT 技术,难以精准赋能咨询和管理 OSA。而专病 GPT,即针对特定疾病开发的预训练转换器模型,有潜力解决 OSA 咨询和管理中存在的痛点和难点。为此,本研究将通过横向对比 BAIMGPT 和 DeepSeek GPT 在 OSA 咨询和管理的临床应用效果,以便通过以下应用方式造福用户或患者。

2 研究方法

2.1 研究目的 研究旨在探讨大语言模型和专病 GPT 赋能 OSA 咨询和管理的临床应用效果。本研究基于这一研究目的设计了在真实世界中,随机自身横向对比研究 BAIMGPT 和 DeepSeek 等 GPT 技术,在咨询和管理 OSA 中的便捷性、亲切感、安全感、问题理解度、回答准确性,语音交互、视觉赋能和需求程度的差别方面,探讨大语言模型和专病 GPT 赋能咨询和管理 OSA 的临床应用效果。

研究期望使用户或患者获得如下收益。(1)提高 OSA 教育获益率、知晓率、筛查率和咨询解答获益率:通过数字人 GPT,用户或患者可以获取更多关于 OSA 的知识,提高教育的获益率。同时,GPT 可以帮助用户或患者了解 OSA 的知晓率和筛查率,确保用户或患者能够及时发现并处理 OSA。此外,GPT 还可以提供咨询解答服务,帮助用户或患者解决关于 OSA 的各种问题,提高咨询解答的获益率。(2)改善培训获益率:数字人 GPT 可以改善用户或患者教育的培训获益率。通过 GPT,用户或患者可以获得更加个性化和高效的医疗服务,提高培训的

效果和获益率。(3)降低延误诊断率:通过数字人GPT,用户或患者可以获得更准确的医疗信息和咨询服务,减少因信息不足而导致的延误诊断率。这有助于提高医疗服务的效率和准确性,确保用户或患者能够及时得到正确的诊断和治疗。(4)改善防治效果:数字人GPT可以帮助用户或患者提高健康意识和管理能力,改善防治效果。通过GPT生成的医学知识和健康建议,用户或患者可以更好地了解如何预防和管理OSA,提高防治效果。(5)实现健康中国规划的强基层广覆盖愿景:数字人GPT的应用可以将医疗服务普及到更广泛的基层地区,实现健康中国规划的强基层广覆盖愿景。通过GPT的远程医疗教育和科普宣传功能,可以将医疗服务延伸到用户或患者的家庭和社区,提高医疗服务的可及性和质量。

2.2 研究设计 对大语言模型GPT和专病GPT用于咨询和管理OSA的临床意义的相关研究进行综述,了解当前大语言模型GPT和专病GPT咨询和管理OSA的研究进展^[10-12]和挑战。

在真实世界中随机自身横向对比研究BAIMGPT和DeepSeek等GPT技术,在咨询和管理OSA中的便捷性、亲切感、安全感、问题理解度、回答准确性、语音交互性、视觉赋能和需求程度的差别方面,探讨大语言模型和专病GPT赋能咨询和管理OSA的临床应用效果。

2.3 研究方案 (1)项目启动与文献调研:明确研究目标和任务,组建研究团队,进行文献调研和模型构建的前期准备工作。基于大规模医学文本数据、临床数据和OSA专病知识库,构建和优化专病GPT赋能咨询和管理OSA的BAIMGPT。(2)模型构建与优化:基于研究结果和医学数据,构建和优化医学数字人BAIMGPT模型,对模型优化的要求为对咨询和管理OSA的便捷性、亲切感、安全感、问题理解度、回答准确性、语音交互、视觉赋能和需求程度。

2.3.1 用户横向比较评价2种GPT咨询和管理OSA的应用效果 横向对比研究专病BAIMGPT和大语言模型DeepSeek咨询和管理OSA的便捷性、亲切感、安全感、问题理解度、回答准确性、语音交互性、视觉赋能和需求程度。

具体方法:向参与试验的用户介绍2种GPT使用方法,提出规定问题(包括什么是OSA,如何诊断OSA等),然后给用户提出3~6个自选问题的机会,由2种GPT分别自主回答。每个问题回答后,均由用户自主选择A、B、C三个评价之一作出评价(A:非

常满意;B:满意;C:不满意)。

评价内容包括以下8方面。

便捷性:BAIMGPT和DeepSeek GPT使用中,不受时间和地域的限制,满足用户的便捷、高效服务需求的满意程度。

亲切感:BAIMGPT和DeepSeek GPT使用中,能与用户建立更加亲切关系,提升诊疗体验和管理信心的满意程度。

安全感:BAIMGPT和DeepSeek GPT使用中,能与用户建立更加信任关系,提升诊疗安全感的满意程度。

问题理解度:BAIMGPT和DeepSeek GPT使用中,各个GPT理解用户问题的满意程度。

回答准确性:BAIMGPT和DeepSeek GPT使用中,各个GPT回答用户问题的准确程度。

语音交互性:BAIMGPT和DeepSeek GPT使用中,是否可以语音交互问答,更方便用户。

视觉赋能:BAIMGPT和DeepSeek GPT使用中,是否可以通过GPT界面利用视觉赋能用户理解,更方便用户应用。

需求程度:对BAIMGPT和DeepSeek GPT用于咨询和医疗服务的需求前景作出评价。

2.3.2 第三方评价理解问题及回答问题的准确性 在本研究中,理解问题及回答问题准确性最为重要,是直接影响到各种GPT的临床应用效果的两个主要因素,所以采用双重评价方法,即用户和第三方(与这一研究利益无关的非研究者担任)的双重评价方法。后者为由第三方直接调用在真实世界问答时,BAIMGPT和DeepSeek等2种GPT对用户理解问题及回答问题的真实世界电子记录,由第三方专家与共识指南、专著或专家经验的符合程度对比评价。并由第三方专家核对理解问题及回答问题准确性(A:不准确;B:基本准确;C:完全准确)。

2.4 研究工具 研究使用BAIMGPT和DeepSeek GPT。

3 临床研究计划

3.1 研究中心和研究人群

复旦大学附属中山医院、重庆大学附属仁济医院(重庆市第五人民医院)、广西壮族自治区人民医院、浙江大学附属第一医院、上海市老年医学中心、吉林大学第二医院、大连大学附属中山医院、复旦大学附属浦东医院、上海百汇医院、南宁市第一人民医院等分中心在门诊或体检发现的OSA用户或患者。

3.2 研究实施

3.2.1 通过教育孵育和筛选入组用户或患者 对于各医院门诊或体检中心体检发现的 OSA 用户或患者(经多导睡眠监测确诊),首先应用数字人“十六字防治 OSA”进行用户或患者教育,筛选入组病例,审核受试者入选与排除标准,筛选合格者入选。

对于通过教育孵育和筛选为合格入组的用户或患者,由研究者在公开场合向受试者或其家属解释,邀请其签署知情同意书,明确试验目的、方法、可能的风险和益处以及他们的权利和责任。在知情同意书中应明确告知受试者本试验目的为 GPT 赋能 OSA 咨询和管理的效果评估试验,不能代表真实世界医生的诊疗意见。如果用户或患者需要就诊真实世界资深专家,本研究专家(包括分中心专家)可以在咨询后,在真实世界临床工作的门诊接诊。介绍研究项目签好知情同意书之后,即可向用户或患者提供的基于 GPT 的咨询。

3.3 研究步骤

由研究者为用户或患者授权使用 GPT,提出规定问题(包括什么是 OSA,如何诊断 OSA)后,给其提出 3~6 个自选问题的机会,由 2 个 GPT 分别回答。每个问题回答后,从前述 8 个方面(便捷性、亲切感、安全感、问题理解度、回答准确性、语音交互性、视觉赋能、需求程度)作出评价(A:非常满意;B:满意;C:不满意)。

3.4 临床资料的收集

本研究的临床资料采用纸质版和电子版《病例报告表》收集信息,资料的收集与保存符合 GCP 要求。纸质版《病例报告表》由复旦大学附属中山医院保存。录入后数据资料由统计学负责人独立处理。

3.5 研究数据分析

3.5.1 用户或患者横向比较评价两种 GPT 咨询和管理 OSA 的意义 横向对比评价专病 BAIMGPT 和大语言模型 DeepSeek 和两种 GPT 咨询和管理 OSA 的便捷性、亲切感、安全感、问题理解、回答准确,语音交互、视觉赋能和需求程度。

3.5.2 第三方专家评价 2 种 GPT 理解问题及回答问题的准确性 基于这一研究中,理解问题及回答问题准确性最为重要,会影响到各种 GPT 的临床应用效果,所以采用双重评价方法:即用户和第三方(与这一研究利益无关的非研究者担任)的双重评价方法。后者为由第三方直接调用在真实世界问答时, BAIMGPT 和 DeepSeek 等 2 种 GPT 对用户理解问题及回答问题的真实世界电子记录,与公认的权威共

识指南、专著或专家经验的符合程度对比评价,由第三方核对理解问题及回答问题准确性。

3.6 终点指标

(1)主要终点:2 种 GPT 问题理解和回答问题的准确性。(注解:回答问题准确性,指回答问题时,答案与权威共识指南的符合程度。)

(2)次要终点:便捷性、亲切感、安全感、需求程度。同时,讨论 2 种 GPT 在 OSA 筛查和管理中的优势、局限性和未来发展方向。

4 统计分析

样本含量约 1 000 例。根据研究设计,选择相应的研究数据的统计与分析方法。

对 BAIMGPT 和 DeepSeek 2 种 GPT,横向比较的试验结果进行统计分析,对比研究大语言和专病 GPT 赋能咨询和管理 OSA 的临床效果,讨论各种 GPT 在 OSA 筛查和 OSA 咨询管理中的优势、局限性和未来发展方向。其中主要包括:(1)分析医学数字人专大融合 GPT 在 OSA 筛查和 OSA 咨询管理中的应用潜力,包括其在亲切感、安全感和信任度、对提高筛查知晓率、积极性及配合管理、辅助筛查、决策等方面的能力。(2)探讨医学数字人专大融合 GPT 对问题理解和回答准确性在管理方面的应用价值,如通过模拟医生的专业知识,提供个性化的健康建议和管理方案。(3)评估用户在使用医学数字人专大融合 GPT 时的体验,包括语音交互、视觉赋能和需求程度等方面,以指导产品的持续改进和优化^[13-15]。(4)确定医学数字人专大融合 GPT 在教育和培训中的作用,特别是在提升医护人员对 OSA 筛查和管理的认识和技能方面。

后续将撰写研究报告和论文,总结研究成果,并在学术会议和期刊上发布。

5 研究相关伦理学

5.1 伦理委员会审核 本研究是全球首个 GPT 赋能筛查 OSA 的需求和体验研究,研究者将每年(如果适用)向伦理委员会提交研究年度报告。在研究中止和/或完成时,研究者也将书面通知伦理委员会;研究者也将及时向伦理委员会报告所有研究工作中发生的变化(如方案和/或知情同意书修订)。

5.2 书面知情同意 本项目获伦理批准后,在使用受试者数据或样本前,将通过电话随访的形式联系受试者或受试者家属,为其解释、提供研究目的相关的信息和可能发生的后果,并尽可能邀请其回院签署知情同意书。知情同意书应作为临床试验的

重要文档保留备查。如联系到受试者,但受试者不愿回院签署书面知情同意书,则在进行电话知情,获得口头同意后,使用其样本或数据;三次及以上联系不到受试者,方可使用其样本或数据以上过程将全程保存记录,对于联系不上的人群依据卫计委《涉及人的生物医学研究伦理审查办法》第三十八条和三十九条,应用可识别身份信息的人体材料或者数据进行研究,已无法找到该受试者,且研究项目不涉及个人隐私和商业利益的,可以免除签署同意书。

5.3 保密措施 本研究会将用户或患者的影像数据上传到珠海横琴圣澳云智科技有限公司推出的“OSA 智能辅助分析系统”中进行 AI 分析。上传之前,我们会将用户或患者影像信息经过匿名化处理。为最大程度保护用户或患者隐私,我们将在分析计算完成后,将会立即删除服务器中用户或患者的影像资料。

通过本项目研究的结果可能会在医学杂志上发表,但是参与研究受试者的个人信息将受到严格保密,必要时除了医院伦理委员会和相关研究人员可以查阅用户或患者的资料外,其他人不会获得这些信息。受试者将通过他们的姓名字母缩写和在研究入选时提供的编号来鉴定。

受试者将了解到这些数据的匿名性和他们保护自己隐私的权利。但是也必须了解到这些数据

将被提交给课题负责单位或者政府管理机构的事实,而且还可能提交给卫生部等机构,进行检查和评价。参与医师将保留 1 份受试者个人资料清单(受试者资料和相应的受试者姓名),以便确认记录。

6 附 录

6.1 BAIMGPT BAIMGPT 为上海白春学人工智能科技工作室研发,且具有独立知识产权和注册商标的 BAIMGPT。其核定使用服务项目国际分类为 9、10、35、38、41、42、44(图 1),包括虚拟现实装置、医疗记录和档案的计算机化管理、互联网广播服务、教育、培训、医疗培训和教学、通过互联网提供医疗信息服务和医疗诊所服务,以及医疗转诊的行政服务。该工作室与国际元宇宙医学协会和中国肺癌防治联盟合作,研发了专病与大语言融合知识库。2024 年 4 月 10 日,在上海市闵行区医学会元宇宙医学专业委员会举行成立大会上,全球领先的元宇宙医学数字人 GPT——BAIMGPT1.0 正式亮相。基于该知识库专家牵头制定的首部及持续更新的亚太和中国肺结节评估管理共识指南和 OSA 诊治管理的专家共识,12 部专著,以及 100 余名国际知名或国内知名专家团队,在元宇宙医学和医学数字人 GPT 方面,具有顶层设计、学术引领、科技创新和智能惠众等方面的优势。



图 1 BAIMGPT 注册商标

6.2 DeepSeek DeepSeek, 全称杭州深度求索人工智能基础技术研究有限公司,成立于 2023 年 7 月 17 日,使用数据蒸馏技术,得到更为精炼、有用的数据。由知名私募巨头幻方量化孕育而生,专注于开发先进的大语言模型(LLM)和相关技术。2025 年 1 月 31 日,英伟达宣布 DeepSeek-R1 模型登陆 NVIDIA NIM。同一时段内,亚马逊和微软也接入 DeepSeek-R1 模型。英伟达称,DeepSeek-R1 是最先进的大语言模型。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 陆俊羽、蔡沁怡:撰写论文;白春学:修改、审核论文。

参考文献

- [1] PEPPARD P E, YOUNG T, BARNET J H, et al. Increased prevalence of sleep-disordered breathing in adults [J]. Am J Epidemiol, 2013, 177(9): 1006-1014.
- [2] PUNJABI N M. The epidemiology of adult obstructive sleep apnea[J]. Proc Am Thorac Soc, 2008, 5(2): 136-143.

- [3] PATEL S R, WHITE D P, MALHOTRA A, et al. Continuous positive airway pressure therapy for obstructive sleep apnea[J]. *Ann Intern Med*, 2003, 139(6): 447–455.
- [4] KAPUR V K, AUCKLEY D H, CHOWDHURI S, et al. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American academy of sleep medicine clinical practice guideline[J]. *J Clin Sleep Med*, 2017, 13(3): 479–504.
- [5] BONSIGNORE MR, BARBÉ F, DE GAETANO A, et al. Sleep apnea management in primary care. *Eur Respir J*. 2021, 57(2).
- [6] ARNALDO PR, JOVANOVIĆ D, SMITH DL. Challenges in managing obstructive sleep apnea in resource-limited settings[J]. *Sleep Health*, 2020, 6(1): 11–18.
- [7] TOPOL E J. High-performance medicine: the convergence of human and artificial intelligence[J]. *Nat Med*, 2019, 25(1): 44–56.
- [8] ESTEVA A, ROBICQUET A, RAMSUNDAR B, et al. A guide to deep learning in healthcare[J]. *Nat Med*, 2019, 25(1): 24–29.
- [9] WANG F, CASALINO L P, KHULLAR D. Deep learning in medicine—promise, progress, and challenges[J]. *JAMA Intern Med*, 2019, 179(3): 293–294.
- [10] JOHNSON K W, TORRES SOTO J, GLICKSBERG B S, et al. Artificial intelligence in cardiology[J]. *J Am Coll Cardiol*, 2018, 71(23): 2668–2679.
- [11] SHEN J Y, ZHANG C J P, JIANG B S, et al. Artificial intelligence versus clinicians in disease diagnosis: systematic review[J]. *JMIR Med Inform*, 2019, 7(3): e10010.
- [12] CHEN J H, ASCH S M. Machine learning and prediction in medicine – beyond the peak of inflated expectations[J]. *N Engl J Med*, 2017, 376(26): 2507–2509.
- [13] TOPOL E J. Deep medicine: how artificial intelligence can make healthcare human again[M]. Basic Books, 2019.
- [14] LUXTON D D. Artificial intelligence in psychological practice: current and future applications and implications[J]. *Prof Psychol Res Pract*, 2014, 45(5): 332–339.
- [15] MIOTTO R, WANG F, WANG S, et al. Deep learning for healthcare: review, opportunities and challenges[J]. *Brief Bioinform*, 2018, 19(6): 1236–1246.

引用本文

陆俊羽, 蔡沁怡, 白春学. 大语言模型和专病 GPT 赋能 OSA 咨询和管理的临床应用效果对比研究方案[J]. 元宇宙医学, 2025, 2(3): 25–30.

LU J Y, CAI Q Y, BAI C X. A comparative study plan on the clinical application effects of large language models and specialized GPTs in OSA consultation and management[J]. *Metaverse Med*, 2025, 2(3): 25–30.

DOI: 10.61189/672599zeyzsy

· 方法学研究 ·

大语言和专病 GPT 赋能咨询和管理肺结节意义的对比研究方案

杨达伟^{1,2,3}, 王 源¹, 白春学^{1,2,3*}

1. 复旦大学附属中山医院呼吸与危重症医学科, 上海 200032

2. 复旦大学附属中山医院(厦门)呼吸与危重症医学科, 厦门 361000

3. 上海呼吸物联网医学工程技术研究中心, 上海市呼吸病研究所, 中国肺癌防治联盟, 上海 200032

[摘要] 本研究旨在对比通用大语言模型(DeepSeek GPT)与专病优化模型(BAIMGPT)在肺结节咨询与管理中的临床价值。通过多中心真实世界研究, 招募 12 家医院的 1000 例肺结节患者, 采用随机自身对照设计, 评估两种模型在便捷性、亲切感、安全感、问题理解准确性、回答专业性、语音交互、视觉赋能及需求程度 8 个维度的表现。核心方法包括: 用户双重评价, 三方盲审, 终点指标。研究预期验证专病 BAIMGPT 在提升筛查知晓率、个性化管理及医患信任度的优势, 并为 AI 赋能肺癌早筛提供循证依据。方案通过伦理审查, 数据匿名化处理, 兼顾创新性与安全性。

[关键词] 肺结节; BAIMGPT; DeepSeek GPT**[中图分类号]** R 563 **[文献标志码]** A

A comparative study of the significance of GPT-enabled counseling and management of pulmonary nodules

YANG Dawei^{1,2,3}, WANG Yuan¹, BAI Chunxue^{1,2,3*}

1. Department of Respiratory and Critical Care Medicine, Zhongshan Hospital, Fudan University, Shanghai 200032, China

2. Department of Respiratory and Critical Care Medicine, Zhongshan Hospital, Fudan University, Xiamen 361000, China

3. Shanghai Respiratory Internet of Things Medical Engineering Technology Research Center, Shanghai Institute of Respiratory Diseases, China Lung Cancer Prevention and Treatment Alliance, Shanghai 200032, China

[Abstract] This study aims to compare the clinical value of a general large language model (DeepSeek GPT) and a disease-specific optimized model (BAIMGPT) in pulmonary nodule consultation and management. Through a multicenter real-world study, 1,000 patients with pulmonary nodules from 12 hospitals will be recruited for a randomized self-controlled trial evaluating the performance of the two models across eight dimensions: convenience, friendliness, sense of security, accuracy in question comprehension, response professionalism, voice interaction, visual empowerment, and demand level. The core methodologies include dual user evaluation, third-party blinded review, and endpoint assessment. The study was expected to validate BAIMGPT's advantages in improving screening awareness, personalized management, and physician-patient trust, providing evidence-based support for AI-powered early lung cancer screening. The protocol has passed ethical review, with anonymized data processing ensuring both innovation and safety.

[Key Words] plmonary nodule; BAIMGPT; DeepSeek GPT

1 The importance of pulmonary nodule consultation and management in promoting early diagnosis and treatment of lung cancer

Pulmonary nodule consultation and management

play a crucial role in improving the early diagnosis and treatment of lung cancer^[1]. Through professional medical consultation, patients can receive accurate information about the nature of pulmonary nodules, potential cancer risks, and necessary medical

[收稿日期] 2025-06-23**[接受日期]** 2025-09-15

[基金项目] 四大慢病重大专项(2024ZD0529300), 上海市科学技术委员会项目基金(20DZ2254400, 21DZ2200600). Supported by Fund of Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0529300) and Science and Technology Commission of Shanghai Municipality (20DZ2254400, 21DZ2200600).

[作者简介] 杨达伟, 博士, 副主任医师, 副研究员. E-mail: yang.dawei@zs-hospital.sh.cn

*通信作者 (Corresponding author). Tel: 021-64041990, E-mail: bai.chunxue@zs-hospital.sh.cn

measures, which is essential for the timely detection and treatment of lung cancer. Management focuses on encouraging patients to take responsibility for their health under medical guidance, including adhering to medical advice for regular check-ups, making lifestyle adjustments, and continuous health monitoring. In terms of demand, the promotion of low-dose computed tomography (LDCT) screening programs^[2] has led to the early detection of an increasing number of pulmonary nodules. Consequently, there is a growing need for medical professionals who can accurately interpret these screening results and provide appropriate advice. At the same time, patients need to understand how to monitor their health in daily life and when to seek medical assistance.

Clinically, pulmonary nodule consultation and management help raise public awareness about the early diagnosis and treatment of lung cancer. This not only improves treatment outcomes and extends patient survival but also reduces the high medical costs associated with late-stage disease detection. Moreover, effective management can alleviate the burden on healthcare systems and optimize the allocation of medical resources. In short, pulmonary nodule consultation and management are integral to raising public health awareness and promoting early detection and treatment of lung cancer, making them an indispensable part of modern healthcare services.

2 Current Challenges and Pain Points in Pulmonary Nodule Consultation and Management

Currently, pulmonary nodule consultation and management face several challenges and pain points. (1) Lack of widely adopted consensus guidelines: Differences between regions and medical institutions, as well as insufficient familiarity and experience with consensus guidelines among physicians, often lead to suboptimal consultation quality, delaying timely diagnosis and treatment of early-stage lung cancer. (2) Lack of professional knowledge: Many patients and their families have limited understanding of pulmonary nodules, making it difficult to make informed decisions. Some non-specialist doctors may also provide misguided advice due to unfamiliarity with

consensus guidelines, potentially leading to inappropriate patient behaviors. (3) Uneven distribution of medical resources: In some areas, quality medical resources are scarce, limiting access to professional and timely consultations. In other regions where resources are abundant, patients may face high medical costs. (4) Challenges in follow-up management: Long-term follow-up is essential for managing pulmonary nodules, requiring strong patient adherence and management capabilities. In practice, many patients fail to attend follow-up appointments for various reasons, affecting disease management and outcomes. (5) Inadequate psychological and social support: Patients with pulmonary nodules often experience anxiety or depression, and existing support systems may not meet their needs. Furthermore, families and society often lack sufficient awareness and understanding of pulmonary nodules, leading to insufficient support for patients^[3].

To address these issues, it is necessary to leverage new forms of productivity to strengthen public education on pulmonary nodules, enhance awareness, promote the widespread adoption of diagnostic standards and consensus guidelines, optimize the allocation of medical resources, improve the service capacity of primary care institutions, enhance patient follow-up management, and provide more comprehensive psychological and social support services.

3 How GPT can address current challenges in pulmonary nodule consultation and management

In recent years, the development of artificial intelligence has attracted widespread attention, with natural language processing (NLP) emerging as a key area of application. As a leading model in NLP, GPT (Generative Pre-training Transformer) has achieved significant success in text generation, classification, and is gradually being applied in medicine and healthcare^[4], GPT demonstrates significant potential in pulmonary nodule consultation and management, contributing to medical innovation and improving the efficiency and accuracy of healthcare services, thereby better meeting patient needs.

From a socio-economic perspective, such

technologies are expected to accelerate the digital transformation of the healthcare industry^[5], increase the level of intelligent healthcare services, reduce costs, and improve accessibility.

However, whether large language models like GPT can accurately empower diagnosis and management remains uncertain. Small-scale studies suggest that general-purpose GPT models struggle to provide accurate consultation and management for pulmonary nodules. In contrast, specialized GPT models, i.e., disease-specific pretrained transformer models, show promise in addressing the pain points in pulmonary nodule consultation and management.

This study aims to compare the performance of BAIMGPT and DeepSeek GPT in the consultation and management of pulmonary nodules to evaluate their clinical relevance and potential to benefit users or patients through the following applications. (1) Patient Consultation: Using natural language processing, GPT can communicate directly with patients, understand their questions and concerns, and provide preliminary diagnostic suggestions. This reduces the burden on physicians and improves the efficiency and quality of patient consultations. (2) Disease Prediction and Diagnosis: By analyzing electronic medical records, GPT can help predict potential health issues, including the nature and progression of pulmonary nodules^[6]. This assists doctors in making faster diagnoses and offering personalized management plans. (3) Medical Knowledge Graph Construction: GPT can aid in constructing comprehensive medical knowledge graphs, helping physicians understand the relationships between different diseases and enhance the overall knowledge base for diagnosis and treatment^[7]. (4) Medical Text Summarization: GPT is capable of processing large volumes of medical literature and extracting key information, enabling healthcare professionals to quickly grasp the latest medical advancements.

In summary, the application of specialized GPT models is expected to improve various aspects of pulmonary nodule management and diagnosis — from enhancing diagnostic accuracy to optimizing the patient consultation experience — thus holding significant practical value.

4 Research methodology

4.1 Research objectives

This study aims to investigate the clinical significance of large language models (LLMs) and disease-specific GPTs in pulmonary nodule consultation and management. To achieve this objective, we designed a real-world, randomized self-controlled comparative study evaluating BAIMGPT, DeepSeek GPT, and related technologies across multiple dimensions: convenience, approachability, perceived safety, question comprehension, answer accuracy, voice interaction, visual assistance, and demand level. The study seeks to elucidate the clinical value of both general-purpose LLMs and specialized GPTs in pulmonary nodule care. Anticipated benefits for users/patients include the following aspects. (1) Enhanced educational outcomes, awareness, screening rates, and consultation efficacy: Through digital human GPT technology, users/patients can acquire comprehensive knowledge about pulmonary nodules, thereby improving educational benefits^[8]. The GPT systems facilitate increased awareness and screening participation, ensuring timely detection and management of pulmonary nodules. Additionally, the GPT-powered consultation services effectively address various patient inquiries, significantly improving consultation outcomes. (2) Improved training effectiveness: The digital human GPT optimizes training benefits for user/patient education. By delivering personalized and efficient medical services, the technology enhances both the quality and effectiveness of educational interventions. (3) Reduced delayed diagnosis rates: Access to accurate medical information and consultation services through digital human GPT minimizes diagnostic delays caused by information gaps. This improvement enhances healthcare service efficiency and accuracy, ensuring timely and appropriate diagnosis and treatment. (4) Enhanced prevention and management outcomes: The digital human GPT strengthens health awareness and self-management capabilities, significantly improving prevention and treatment efficacy^[9]. By providing AI-generated medical knowledge and health recommendations, users/patients gain better

understanding of pulmonary nodule prevention and management strategies. (5) Supporting China's Healthy China initiative for grassroots healthcare coverage: The application of digital human GPT extends medical services to broader primary care settings, realizing the Healthy China vision of comprehensive grassroots coverage^[10]. Through remote medical education and science popularization functions, the technology expands healthcare accessibility to patients' homes and communities, improving both the availability and quality of medical services.

4.2 Study design

This real-world randomized self-controlled cross-comparison study evaluates BAIMGPT, DeepSeek GPT, and other technologies in terms of convenience, approachability, perceived safety, question comprehension, answer accuracy, voice interaction, visual assistance, and demand level for pulmonary nodule consultation and management. The study aims to explore the clinical significance of general LLMs versus specialized disease-focused GPTs in pulmonary nodule care.

4.3 Literature review

A comprehensive review of existing research on GPT-based LLMs and specialized disease GPTs for pulmonary nodule consultation and management will be conducted to understand current advancements and challenges in this field.

4.4 Study protocol design

The study protocol includes the following aspects. (1) Project Initiation & Literature Review: Define research objectives and tasks, assemble the research team, conduct preliminary literature reviews and model development preparations^[11]. Based on large-scale medical text data, clinical data, and a specialized pulmonary nodule knowledge base, construct and optimize BAIMGPT, a disease-specific GPT for pulmonary nodule consultation and management. (2) Model Development & Optimization: Build and refine the Medical Digital Human BAIMGPT model based on research findings and medical data^[12]. The model optimization requirements shall focus on: convenience, approachability, sense of security, question

comprehension, answer accuracy, voice interaction, visual empowerment, and demand level for pulmonary nodule consultation and management^[13].

4.5 Cross-user comparative evaluation of three gpts in pulmonary nodule consultation and management

A horizontal comparison study was conducted between the specialized BAIMGPT and the general large language model DeepSeek, focusing on eight key dimensions in pulmonary nodule consultation and management: convenience, approachability, sense of security, question comprehension, answer accuracy, voice interaction, visual empowerment, and demand level. The specific methodology involved: Introducing trial participants to the three GPT systems and their usage methods; Presenting standardized questions (including "What are pulmonary nodules?" and "How are pulmonary nodules diagnosed?"); Allowing users to pose 3-6 self-selected questions, with each GPT autonomously responding to all questions. Later, users independently selected one of three evaluation grades: (A: Very Satisfied; B: Satisfied; C: Dissatisfied). The evaluation covered the following eight aspects:

Convenience: Ability to provide time/location-independent and efficient service (All GPTs can achieve this).

Approachability: Ability to establish a personalized, reassuring interaction (Only Digital Human GPT can achieve A).

Perceived Safety: Ability to build trust and confidence (Only Digital Human GPT can achieve A).

Question Comprehension: How well each GPT understands user queries.

Answer Accuracy: Alignment of responses with medical accuracy.

Voice Interaction: Usability of voice-based Q&A (Models without voice capability cannot score A).

Visual Assistance: Effectiveness of visual aids in enhancing understanding (Models without visual interfaces cannot score A).

Demand Level: User-rated need for GPT in healthcare: Urgent Need; Need; No Need.

4.6 Third-party evaluation of question comprehension and answer accuracy

In this study, question comprehension and

answer accuracy constitute the two most critical factors directly influencing the clinical applicability of various GPT models. Therefore, a dual-evaluation approach was adopted involving: User self-assessment; Independent third-party evaluation (conducted by non-researchers without conflicts of interest). The third-party evaluation methodology included: Direct analysis of real-world interaction logs from BAIMGPT, DeepSeek GPT, and other GPT models^[14], comparative assessment against consensus guidelines, authoritative medical literature and expert knowledge from Grade-A specialists of the China Lung Cancer Prevention Alliance was conducted. Accuracy is evaluated by third-party experts (A: Inaccurate; B: Partially accurate; C: Fully accurate).

4.7 Study design selection

An open-label, multicenter real-world study was conducted to horizontally compare BAIMGPT and DeepSeek GPT across eight dimensions: convenience, approachability, security perception, question comprehension, answer accuracy, voice interaction, visual empowerment, and demand level. This design evaluated the efficiency of GPT-enabled pulmonary nodule consultation and management^[15].

4.8 Research tool selection The study utilized BAIMGPT and DeepSeek GPT systems (see Supplementary for technical specifications).

5 Clinical research plan

5.1 Research centers and study population

The study will be conducted across multiple centers, including: Zhongshan Hospital, Fudan University, Henan Provincial People's Hospital, The First Affiliated Hospital of Zhejiang University, The First Affiliated Hospital of Chongqing Medical University, Shanghai Geriatric Medical Center, Xinqiao Hospital, Army Medical University, The Second Hospital of Jilin University, Zhongshan Hospital Affiliated to Dalian University, Pudong Hospital, Fudan University, Xinhua Hospital, Shanghai Jiao Tong University, Shanghai Parkway Health, The First People's Hospital of Nanning. The study population consists of pulmonary nodule patients or users identified during outpatient visits or health check-ups at these centers.

5.2 Participant education and screening

Patients diagnosed with pulmonary nodules via thin-slice CT scans will first undergo education using the Digital Human's "16-Word Lung Cancer Prevention and Treatment" program. Eligible participants will be screened based on inclusion and exclusion criteria.

5.3 Informed consent process

Qualified participants (or their families) will receive a detailed explanation of the study's purpose, methods, potential risks, benefits, and their rights and responsibilities before signing the informed consent form (ICF). The ICF will clarify that: The study involves GPT-assisted consultation and management and does not replace real-world medical diagnosis. If participants require real-world expert consultation, study-affiliated specialists can provide clinical follow-up in outpatient settings. After consent is obtained, participants will receive GPT-based consultations.

5.4 Study procedures

Participants will be granted access to three GPT models (BAIMGPT and DeepSeek GPT variants) to: Answer standardized questions (e. g., "What are pulmonary nodules" and "How are they diagnosed"), and pose 3-6 self-selected questions.

Each GPT's responses will be evaluated on:

Convenience. Satisfaction with accessibility and efficiency (time/region-independent).

Approachability. Ability to foster a personalized, reassuring interaction.

Perceived safety. Degree of trust and confidence in the GPT's responses.

Question comprehension. Accuracy in understanding user queries.

Answer accuracy. Alignment with evidence-based medical knowledge.

Voice interaction. Usability of voice-based Q&A.

Visual assistance. Effectiveness of visual aids in enhancing understanding.

Perceived need. User-rated demand for GPT in healthcare (urgent need / need / no need).

5.5 Data collection

Clinical data will be recorded via paper-based and electronic Case Report Forms (CRFs), compliant

with GCP standards. Paper CRFs will be archived at Zhongshan Hospital, Fudan University; electronic data will be processed independently by the lead statistician.

6 Data analysis

6.1 User evaluation of GPT models

A cross-comparison of BAIMGPT (specialized disease model) and DeepSeek GPT (general LLM) will assess: convenience, approachability, safety, comprehension, accuracy, voice/visual features, and demand.

6.2 Third-party expert validation

An independent, conflict-free third party will evaluate: real-world response accuracy by comparing GPT outputs to consensus guidelines, authoritative texts, and A-tier expert opinions from the China Lung Cancer Prevention Alliance.

6.3 Endpoints

Primary endpoint: Accuracy of question comprehension and responses (measured against guidelines^[16]). Secondary endpoints: Convenience, approachability, safety, and demand. Exploratory Analysis: Strengths, limitations, and future directions of GPT in lung cancer screening and nodule management.

7 Summary and discussion

The Sample Size is estimated to be around 1,000 cases. Appropriate statistical and analytical methods will be selected based on the study design.

A statistical analysis will be conducted on the comparative experimental results of GPT and three DeepSeek GPT models, focusing on the contrast between general large language models and specialized disease-empowered GPTs in pulmonary nodule consultation and management. The discussion will explore the advantages, limitations, and future directions of various GPT models in lung cancer screening and pulmonary nodule management, including the following aspects. (1) Analyzing the potential of Medical Digital Human-integrated GPT in lung cancer screening and pulmonary nodule consultation management^[17], particularly its ability to enhance screening awareness, motivation, and

compliance through approachability, perceived safety, and trustworthiness in aiding screening decision-making. (2) Investigating the application value of Medical Digital Human-integrated GPT in problem comprehension and response accuracy for management purposes, such as simulating physician expertise to provide personalized health recommendations and management plans. (3) Evaluating user experience with Medical Digital Human-integrated GPT^[18], including voice interaction, visual empowerment, and demand assessment, to guide continuous product refinement and optimization^[19]. (4) Assessing the role of Medical Digital Human-integrated GPT in education and training, particularly in improving healthcare professionals' knowledge and skills in lung cancer screening and management^[20].

Research findings will be compiled into reports and academic papers for dissemination at conferences and in peer-reviewed journals.

8 Research ethics

8.1 Ethics committee review

This study is the first research on GPT-empowered lung cancer screening needs and patient experience. Upon approval, it will help enhance the influence of Shenzhen-Zhongshan Hospital and serve as a critical reference for future transformations in medical practice. Therefore, we specially request an expedited ethics committee review. The researchers will submit annual progress reports (if applicable) to the ethics committee. Upon study termination and/or completion, the researchers will provide written notification to the ethics committee. Any changes in the research (e.g., protocol and/or informed consent revisions) will also be promptly reported.

8.2 Written informed consent

After obtaining ethical approval, before using participant data or samples, researchers will contact participants or their families via follow-up phone calls to explain the study's purpose, potential implications, and invite them to return to the hospital to sign the informed consent form (ICF). The ICF will be retained as a key document for the clinical trial. If a participant is unwilling to return to sign the ICF, verbal consent may be obtained via phone before using their data/

samples. If a participant cannot be contacted after three or more attempts, their data/samples may still be used in compliance with ethical regulations. This entire process will be fully documented. For individuals who cannot be reached, in accordance with Articles 38 and 39 of the National Health Commission's Ethical Review Measures for Biomedical Research Involving Human Subjects, if the study does not involve personal privacy or commercial interests, the requirement for signed consent may be waived when using identifiable human materials or data from untraceable participants.

8.3 Confidentiality measures

The research team will upload users' or patients' imaging data to the "Pulmonary Nodule Intelligent Auxiliary Analysis System" developed by Zhuhai Hengqin Shengao Yunzhi Technology Co., Ltd. for AI analysis. Before uploading, we will anonymize the imaging information of users or patients. To maximize privacy protection, all uploaded imaging data will be promptly deleted from the server upon completion of the analysis.

The results obtained from this research may be published in medical journals, but the personal information of study participants will remain strictly confidential. Only the hospital ethics committee and authorized researchers may access the data when necessary. Participants will be identified only by their initials and assigned study numbers upon enrollment.

Participants will be informed about the anonymity of their data and their right to privacy protection. However, they must also acknowledge that the data

may be submitted to the responsible research institution, government regulatory agencies (such as the Ministry of Health), for inspection and evaluation. The attending physicians will maintain a confidential list of participants' personal information (study records and corresponding names) for verification purposes.

This initiative collaborates with the International Metaverse Medical Association and the China Lung Cancer Prevention and Treatment Alliance to develop a specialized disease and large language model-integrated knowledge base. On April 10, 2024, at the inaugural meeting of the Metaverse Medical Professional Committee of the Shanghai Minhang Medical Association, the world-leading Metaverse Medical Digital Human GPT—BAIMGPT 1.0—was officially unveiled. Based on this knowledge base, experts have led the formulation of the first and continuously updated Asia-Pacific and Chinese consensus guidelines for pulmonary nodule evaluation and management, as well as expert consensus on OSA diagnosis and treatment management, along with 12 specialized publications. Supported by a team of over 100 internationally and domestically renowned experts, this project holds advantages in top-level design, academic leadership, technological innovation, and intelligent public benefit in the fields of metaverse medicine and medical digital human GPT. **Supplementary** BAIMGPT is developed by Shanghai Baichunxue Artificial Intelligence Technology Studio and holds independent intellectual property rights and a registered trademark (Figure 1).



Figure 1 BAIMGPT Registered Trademark

These encompass virtual reality devices, computerized management of medical records and files, internet broadcasting services, education, training, medical training and teaching, provision of medical information services and clinical services via

the Internet, and administrative services for medical referrals.

On April 10, 2024, at the inaugural meeting of the Metaverse Medicine Committee of the Minhang District Medical Association in Shanghai, the world's

leading metaverse medical digital human GPT—BAIMGPT 1.0—was officially unveiled.

Based on this knowledge base and led by experts, the first and continuously updated Asia-Pacific and China consensus guidelines on pulmonary nodule assessment and management, as well as expert consensus on OSA diagnosis and treatment management, have been developed. Supported by 12 monographs and a team of over 100 internationally or nationally renowned experts, BAIMGPT demonstrates advantages in top-level design, academic leadership, technological innovation, and intelligent public service in the fields of metaverse medicine and medical digital human GPT.

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 杨达伟、王源:撰写论文;白春学:修改、审核论文。

参考文献

- [1] SHARMA A, KAPOOR R. Current Challenges in Pulmonary Nodule Management: A Review [J]. *Clinical Respiratory Medicine*, 2021, 29(4): 1020–1030.
- [2] ZHANG M, LI X. A Survey of AI in Pulmonary Disease Diagnosis [J]. *Journal of Pulmonology and AI*, 2022, 28(6): 112–119.
- [3] ZHANG Y, et al. Development of Disease-Specific GPT Models for Healthcare [J]. *AI Research Journal*, 2021, 16(4): 75–82.
- [4] 刘 磊, 张 明. 预训练语言模型GPT在医学中的实践应用 [J]. *医学人工智能杂志*, 2021, 11(2): 88–101.
- [5] ANDERSON L, ZHANG H. AI and Healthcare Efficiency: A Systematic Review [J]. *Healthcare Technology Reviews*, 2022, 25(3): 115–123.
- [6] HE J, et al. Improving Diagnostic Accuracy with AI in Lung Cancer [J]. *Journal of Diagnostic AI*, 2022, 7(2): 99–106.
- [7] LIU X, TANG F, REN M. Ethical Challenges in AI-Powered Medical Technologies [J]. *Bioethics and AI*, 2023, 18(1): 80–91.
- [8] CHOU F, YANG K. Enhancing Patient Education Through AI Models [J]. *Patient Education and AI*, 2022, 5(4): 220–228.
- [9] HUANG Y, LIU C. The Potential of AI in Pulmonary Disease Screening and Treatment [J]. *Journal of Respiratory Health*, 2021, 18(3): 140–149.
- [10] XIE W, ZHANG L. Machine Learning in Lung Cancer Management [J]. *Journal of Machine Learning in Medicine*, 2021, 14(6): 88–97.
- [11] PATEL A, SINGH R. Early Detection of Pulmonary Diseases Using AI [J]. *Pulmonary Medicine Advances*, 2022, 22(3): 33–44.
- [12] CHEN Q, WANG T, Zhao Y. GPT Models for Predictive Analytics in Healthcare [J]. *Journal of Digital Health*, 2023, 19(3): 145–153.
- [13] LEE T, HUANG Y. Changes in Pulmonary Nodule Screening Protocols Due to AI [J]. *AI in Cancer Detection*, 2021, 14(5): 98–108.
- [14] TAN J, et al. The Role of DeepSeek GPT in Medical Consultations [J]. *Artificial Intelligence in Medicine*, 2022, 9(2): 55–62.
- [15] WANG Y, ZHAO H. Comparative Effectiveness of Deep Learning Models in Radiology [J]. *AI in Healthcare Journal*, 2022, 5(1): 67–74.
- [16] BROWN D, LIU S. Comparative Evaluation of GPT Models in Clinical Settings [J]. *AI and Health Systems*, 2023, 11(1): 45–53.
- [17] WANG Y, ZHANG H. The Future of AI in Pulmonary Nodule Diagnosis and Management [J]. *Clinical Medicine Insights*, 2021, 10(4): 212–221.
- [18] KIM H, LEE J. Patient Trust and Acceptance in AI Medical Systems [J]. *Journal of Medical AI Ethics*, 2023, 5(2): 45–52.
- [19] PATEL S, WANG J, Huang L. Impact of AI in Real-World Pulmonary Nodule Screening [J]. *Journal of Lung Cancer Screening*, 2023, 30(1): 45–54.
- [20] SMITH J, LIU Y, ZHAO H. Applications of AI in Early Diagnosis of Lung Cancer [J]. *Journal of Medical AI*, 2023, 15(4): 234–245.

引用本文

杨达伟,王源,白春学. 大语言和专病GPT赋能咨询和管理肺结节意义的对比研究方案[J]. 元宇宙医学, 2025, 2(3): 31–38.

YANG D W, WANG Y, BAI C X. A comparative study of the significance of GPT-enabled counseling and management of pulmonary nodules [J]. *Metaverse Med*, 2025, 2(3): 31–38.

DOI:10.61189/502047ilpumd

· 产学研融合 ·

基于大语言模型的中文电子病历命名实体识别



程洁, 刘独玉*, 陈思旭, 钱姝妤

西南民族大学电气工程学院, 四川 成都 610041

[摘要] 命名实体识别(NER)作为自然语言处理(NLP)的核心任务,在电子病历(EMRs)中识别疾病、症状等医学实体,对临床辅助决策和医学知识库构建意义重大。但传统方法依赖大量标注数据与复杂模型,训练及推理成本较高。本文提出一种融合语义检索与提示学习的大语言模型生成式医学NER方法。首先,构建句子级向量数据库,对电子病历进行语义编码以实现可检索表示;然后,基于输入语句进行语义相似度检索,将相似示例动态注入提示模板,引导模型完成实体抽取;最后,通过结构化特殊标记生成实体类型标注结果,实现直接解码输出。实验表明,该方法在自建电子病历数据集和瑞金医院糖尿病数据集上均表现良好,尤其在低资源场景下具备较强的鲁棒性与迁移能力。

[关键词] 命名实体识别;电子病历;大语言模型

[中图分类号] TP 391 **[文献标志码]** A

Named entity recognition in chinese electronic medical records based on large language models

CHENG Jie, LIU Duyu*, CHEN Sixu, QIAN Shuyu

Southwest Minzu University, College of Electrical Engineering, Chengdu 610041, Sichuan, China

[Abstract] Named Entity Recognition, as a core task in Natural Language Processing, plays a crucial role in identifying medical entities such as diseases and symptoms in Electronic Medical Records, which is of great significance for clinical decision support and the construction of medical knowledge bases. However, traditional methods rely heavily on large amounts of annotated data and complex models, resulting in high training and inference costs. This paper proposes a generative medical NER method that integrates semantic retrieval and prompt learning with large language models. First, a sentence-level vector database is constructed to semantically encode EMRs for retrievable representations. Then, based on the input sentence, semantic similarity retrieval is performed, and similar examples are dynamically injected into a prompt template to guide the model in entity extraction. Finally, entity type annotation results are generated through structured special markers, enabling direct decoding output. Experimental results demonstrate that the proposed method performs well on both a self-constructed EMR dataset and the Ruijin Hospital diabetes dataset, and exhibits strong robustness and transferability, especially in low-resource scenarios.

[Key Words] named entity recognition; electronic medical records; large language models

NER作为NLP领域的一项基础任务,其主要目标是识别出文本中有特定含义的实体,并对这些实体进行标注^[1]。它也是自然语言处理领域中许多研究问题的基础,如关系抽取^[2]、事件抽取^[3]、知识图谱^[4]、机器翻译^[5]、问答系统^[6-7]等。电子病历作为医疗信息的核心载体,包含大量非结构化数据,蕴含着疾病、症状、检查、治疗等关键医学实体信息。从电子病历中高效、准确地抽取医学实体,不仅是深入分析医疗文本的基础,也是提升临床决策与治疗水平、实现智能医疗辅助的关键步骤。

当前,主流的医学NER方法多基于序列标注框架^[8],如BiLSTM-CRF、BERT-CRF等。这类方法通

常依赖复杂的模型结构和大规模高质量标注语料,在处理医学文本时面临诸多挑战,如语言表达复杂、实体边界模糊、标签类别重叠等,使得模型泛化能力不足,迁移困难,且训练与部署成本较高。

近年来,大语言模型(large language models, LLMs)以其卓越的语言理解与生成能力,为医学NER任务提供了新的解决思路。通过合理设计提示(prompt),可以在无需微调参数的前提下引导模型完成实体识别,有效缓解标注资源依赖问题。然而,LLMs的性能高度依赖于提示的上下文质量。传统手工设计的提示模板往往通用性强而针对性弱,难以有效捕捉特定医学语境下的实体语义,从而影

[收稿日期] 2025-08-14

[接受日期] 2025-09-16

[作者简介] 程洁,硕士研究生. E-mail: 19822909065@163.com

*通信作者(Corresponding author). Tel:15828397145, E-mail: liuduyu10000@163.com

响识别效果的稳定性与准确性。

为此,本文提出一种结合语义检索与提示学习的上下文实体标记生成方法,旨在提升提示内容的相关性和大模型识别能力。该方法通过构建以句子级为单位的向量数据库,对已标注电子病历进行语义编码,实现可检索的语义表示;随后,利用语义检索机制从中动态选取与输入文本最相似的句子作为示例,并将其注入提示模板中,引导大语言模型进行实体抽取;最后,采用结构化标记格式输出实体及其类型,实现统一建模与直接解码。该方法将传统 NER 任务由“序列标注”转化为“文本生成”,不仅简化了建模流程,也更契合大语言模型的生成式能力。

为验证方法的有效性,本文在两个中文医学实体识别数据集上进行了系统实验,即自建数据集(包括糖尿病和小儿肺炎两个病种的电子病历)和瑞金医院糖尿病数据集。实验结果表明,本文方法在两个数据集上均取得了优异的性能,尤其在低资源条件下表现出显著优势,充分验证了结合语义检索与提示机制的大语言模型生成式方法在中文电子病历信息抽取中的可行性和潜力。

1 研究背景

命名实体概念提出之初,研究多基于规则与字典方法^[9],这类方法在特定场景中有效,但人工依赖高、泛化性差,难适应语言多样性,且在开放领域易出现漏检与误判。因此,引入了机器学习方法,如支持向量机(SVM)、隐马尔可夫模型(HMM)以及条件随机场(CRF)等。Zhou 等^[10]融合 SVM 与 HMM,在 GENETAG 数据集上达到了 0.83 的 F1 值。Jonnalagadda 等^[11]与 Jiang 等^[12]基于 CRF 在 I2B2 2010 数据集上分别取得了 0.82 和 0.84 的 F1 值。此类方法通过特征建模提升了识别效率,但仍高度依赖人工特征设计,难以捕捉深层语义。

为解决这些问题,深度学习方法应运而生,通过模型自动完成特征提取、训练和预测,常用模型包括卷积神经网络(CNN)、循环神经网络(RNN)、长短期记忆网络(LSTM)和 Transformer 等。Collobert 等^[13]提出基于 CNN 的多层神经网络结构,适用于包括 NER 在内的多种 NLP 任务。RNN 适用于序列建模,但易出现梯度消失或爆炸。为此,LSTM 基于梯度问题进行了改进。Huang 等^[14]结合 BiLSTM 和 CRF,利用双向状态获取上下文信息。Transformer^[15]基于多头自注意力机制,能更好地捕捉长距离依赖关系,是 NLP 研究中的重要突破,但

其效果仍受限于训练语料的规模与质量。

基于迁移学习的思想,预训练语言模型(PLMs)的出现为这一问题提供了解决思路。Peters 等^[16]提出了 ELMo,该模型利用 BiLSTM 获取上下文表示。谷歌提出的 BERT^[17],通过堆叠 Transformer 与 MLM 训练构建双向语言模型,被广泛应用并衍生出了许多变体模型。例如,HU 等^[18]提出 BERT-Attention-BiLSTM 用于中医短文本匹配。郭等^[19]提出 RoBERTa-wwm-ext-BiGRU-EGP 模型,能够有效提升医学实体解析与文本分析能力。

随着 PLMs 规模的不断扩大,微调成本急剧上升。提示学习(prompt learning)^[20]通过调整下游任务的结构与输入,使其更契合预训练语言模型,从而提升模型参数的利用率与任务性能。在此背景下,GPT-3、ChatGPT、ChatGLM 以及 Qwen 等具备强大语言理解与生成能力的大语言模型(LLMs)相继涌现。基于 LLMs 的提示学习范式,逐步成为应对下游任务特别是 NER 任务的一种新趋势。Schick 等^[21]提出 PET,将任务转化为完形填空模板并利用软标签提升低资源条件下的模型性能;GPT-3^[22]通过通过任务描述实现少样本高性能;Cui 等^[23]利用 BART 构建模板排序促进知识迁移;Ma 等^[24]提出 Template-free Prompt 直接预测类别词;Lee 等^[25]则通过 Demonstration-based Learning 提升低资源下的模型表现;Shen 等^[26]统一实体定位与分类并采用动态图匹配提升标签质量。

基于生成式 LLMs 的 NER 方法将传统序列标注任务转化为结构化文本生成任务,即以输入文本与提示为条件,生成带有实体标记或结构化信息的输出。该方法在少样本、嵌套实体和跨领域迁移中具备良好的泛化与语义理解能力,尤其适用于中文医学电子病历的复杂实体与语义歧义处理。

2 研究方法

本文针对医学命名实体识别任务,提出了一种结合语义检索与提示学习的大语言模型生成式识别方法。该方法整体流程主要包括以下三个部分。

①实体向量数据库构建:将电子病历中的句子作为基本单元,利用预训练模型进行语义编码,构建可供检索的文本向量集合;②基于语义检索的提示生成:根据输入句子从向量数据库中检索相似样本,并将其作为示例动态注入提示模板中,辅助模型完成实体抽取;③实体标记生成与解码:大语言模型生成嵌入结构化特殊标记的输出句子,直接输出带实体类别的识别结果,便于解析与评估。整体

架构如图 1 所示。

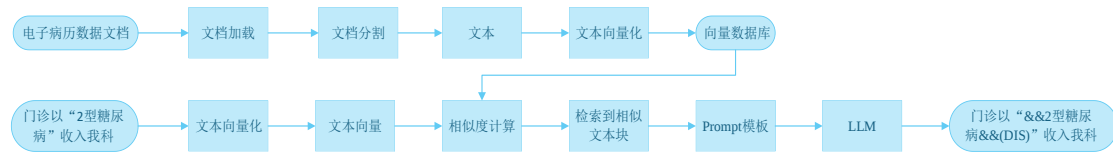


图1 基于LLM的NER流程图

2.1 实体向量数据库构建 为了为提示增强的生成式医学命名实体识别模型提供可检索的、结构化的医学领域知识支撑,本文设计并构建了一个面向医学电子病历的向量数据库。该向量数据库以电子病历中的句子级文本为基础,通过预训练医学嵌入模型获取其稠密语义表示,为后续提示生成阶段提供高质量、可追溯的示例支撑。向量数据库的构建过程包括以下关键步骤。

2.1.1 文档加载与分割 首先加载经过预处理和人工标注后的中文电子病历数据集。原始数据以“content”字段存储完整病历文本。为提高检索粒度与表示一致性,本文采用基于中文句号“。”的切分策略,将每条“content”文本划分为若干语义完整、粒度一致的句子级片段,并以“text”字段形式保存。例如:“text”:“糖尿病史10余年,子宫癌术后8年,心脏支架术后,无传染病史。”;“text”:“临床表现为反

复发作的喘息、胸闷、咳嗽等症状。”该操作确保了每条文本具有明确边界、语义集中,适用于句子级向量建模与相似度检索。

2.1.2 语义向量编码 bge-large-zh-v1.5^[27]是由BAAI(Beijing Academy of Artificial Intelligence)开发的一款基于Transformer架构的高性能中文文本嵌入模型,如图2。该模型采用1024维的嵌入空间,能够对句子级文本进行高质量的语义编码,适用于相似度计算与语义检索等任务。其核心技术包括指令微调(instruction fine-tuning)和硬负样本挖掘(hard negative mining),显著提升了向量分布的判别能力与检索效果。同时,bge-large-zh-v1.5支持最长8192个token的长文本输入,兼顾语义精度与处理效率。在C-MTEB等中文基准评测中,该模型表现优异,展现出在多语言任务中的广泛适应性与应用潜力。

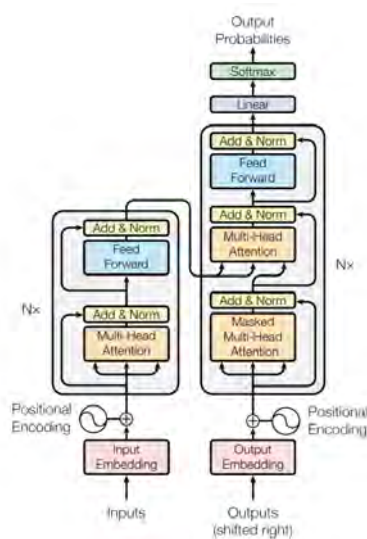


图2 Transformer架构图^[15]

bge-large-zh-v1.5的文本表示可形式化包含输入嵌入与位置编码、Transformer编码和向量归一化三个过程。①输入嵌入与位置编码(embedding + positional encoding):假设语料库中总共包含 N 个句子,对于其中任一句子 $x_j(j = 1, 2, \dots, N)$ 可表示为

$x_j = w_1^j w_2^j \dots w_n^j$,首先通过嵌入层获得词级输入:

$$E^j = [e_1^j; e_2^j; \dots; e_n^j] + [p_1^j; p_2^j; \dots; p_n^j] \quad (2-1)$$

其中, $e_i^j \in \mathbb{R}^d$ 为 w_i^j 的嵌入向量, $p_i^j \in \mathbb{R}^d$ 为 w_i^j 的位置编码, $d = 1024$ 表示嵌入维度, n 为输入序列的token总数, E^j 为 w_i^j 的输入嵌入矩阵,维度为 $\mathbb{R}^{n \times d}$ 。

②Transformer 编码:经过多层 Transformer 编码器进行上下文建模。

$$H_{(0)}^i = E^j \quad (2-2)$$

其中, $H_{(0)}^i$ 为初始隐状态, 维度为 $\mathbb{R}^{n \times d}$ 。

$$H_{(l)}^i = \text{TransformerLayer}^{(l)}(H_{(l-1)}^i) \quad (2-3)$$

其中, $l = 1, 2, \dots, L$, L 为 Transformer 编码器的总层数。

每层包含多头注意力(multi-head attention)和前馈网络(FFN), 输出最后一层的隐状态表示为:

$$H_{(l)}^i = [h_{[\text{CLS}]}^i; h_2^i; \dots; h_n^i], h_i^i \in \mathbb{R}^d \quad (2-4)$$

其中, $h_{[\text{CLS}]}^i$ 表示 Transformer 编码器输出的 [CLS] token 对应的向量表示。

③向量归一化(CLS token + L2 Normalization):

$$\vec{x}_j = \frac{h_{[\text{CLS}]}^{(j)}}{\|h_{[\text{CLS}]}^{(j)}\|} \quad (2-5)$$

其中, $\|\cdot\|$ 表示向量的 L2 范数。

(3) 向量集合构建 将语料库中的所有句子完成编码并归一化后, 构建向量集合(即向量数据库) $\mathcal{X}$:

$$\mathcal{X} = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N\} \quad (2-6)$$

该集合仅包含句子级文本向量, 不包含实体标注信息, 保证了数据库结构的简洁性与通用性。同时, 通过句子索引的方式, 可在提示生成阶段快速映射与追溯原始样本。

2.2 基于语义检索的提示生成 在完成向量数据库的基础上, 本文进一步设计了基于语义检索的提示生成机制, 旨在充分利用向量数据库中蕴含的医学知识示例, 动态增强大语言模型的生成式命名实体识别能力。该模块的核心流程如下。

(1) 输入句子编码与相似度计算: 本文采用 Top_K 相似度检索策略, 旨在为每一个输入句子构造高质量的语义提示。具体而言, 使用嵌入模型 bge-large-zh-v1.5(同 2.1) 将待识别的输入句子 q 转化为稠密向量表示 $\vec{q}$, 并与向量集合 $\mathcal{X}$ 中的所有向量进行余弦相似度计算, 以检索与其最接近的 Top_K 个向量所对应句子作为提示示例。

余弦相似度计算:

$$\text{sim}(\vec{q}, \vec{x}_j) = \frac{\vec{q} \cdot \vec{x}_j}{\|\vec{q}\| \cdot \|\vec{x}_j\|} \quad (2-7)$$

其中, $\cdot$ 表示向量点积。显然, $\text{sim}(\vec{q}, \vec{x}_j)$ 的取值范围为 $[-1, 1]$, 值越接近 1 表示两个句子的语义越接近。

$$\text{Top_K}(q) = \arg \max_{x_j \in \mathcal{X}} \text{topK} \text{sim}(\vec{q}, \vec{x}_j) \quad (2-8)$$

其中, $\arg \text{topK}$ 表示从候选向量集合中, 返回与 $\vec{q}$ 相似度最高的前 K 个向量对应的句子, 即 Top_K 个句

子, 作为最相关的提示示例。

(2) 示例标注重构: 检索后, 将原始“text”中出现的实体按照特殊标记格式“&&实体&&(实体类型)”替换, 从而得到符合提示规范的“输入-输出”示例。所有 Top_K 示例会拼接并注入到提示模板中的 {RAG_text} 占位符处, 可为大语言模型提供具备相似语义背景和准确标注结构的示例。

(3) 提示模板动态填充: 设计提示模板如表 1 所示。该提示模板涵盖了实体抽取规则、实体类型定义、输出格式约束等内容, 并支持动态插入检索示例和当前待识别句子 {query}, 以构建完整的模型输入。

(4) 提示引导的生成推理: 经过填充后的完整提示将作为输入送入大语言模型, 引导其识别当前句子中的医学实体。由于提示中嵌入了结构化规则和真实示例, 模型可直接参考标注风格进行仿写, 进而提高抽取结果的准确率、实体边界判断的清晰性及类型分辨的一致性。

本节所设计的提示生成机制融合了向量检索与提示模板的优点, 具备可扩展、可控、可追溯的特点, 为生成式实体识别任务提供了更可靠的输入结构。

2.3 实体标记生成与解码 本文采用显式可视的特殊标记方法, 即将识别出的实体及其对应类型以“&&实体&&(实体类型)”的格式嵌入原句中, 作为模型的直接输出结果, 如图 3 所示。相较于传统 BIO(begin-inside-outside)序列标注方式, 该标记方法具有明显的优势: BIO 标注需要对每个分词位置进行标签预测, 解码时再通过后处理算法组合实体, 容易导致边界模糊、重叠错误或标签不一致等问题; 而本文中的特殊标记方法则能在原文中直接呈现实体及其实体类型, 可立即校验, 无需复杂的后处理, 特殊标记不仅机器可解析, 如正则表达式即可抽取, 同时也方便人工快速核对, 有助于医学专家对生成结果进行验证, 提高整体的可追溯性与可解释性。

在解码阶段, 大语言模型会根据提示注入的信息以及语义检索返回的相似示例, 逐步推理并输出结构化标记的句子。系统后续可通过简单的正则匹配或规则算法, 快速解析出实体及其类别。

3 研究实施

本实验的环境配置如表 2 所示, 大语言模型框架是 LangChain, BGE(BAAI general embedding)模型是 bge-large-zh-v1.5, Transformers 版本为 4.52.3。

表 1 提示模板设计

你需要对用户输入的句子进行医疗实体抽取。实体抽取的规则有：	
1. 实体必须出现在句子中；	
2. 实体必须与实体类型相关；	
3. 实体和实体之间不能重叠。	
抽取的实体类型有：	
1. 疾病(DIS):患者所患的具体疾病或诊断名称。	
2. 症状(SYM):感受到了或表现出的一种令人不快的感觉或表现出的现象,可能是某种潜在疾病或病状的症状。	
3. 体征(SIG):包括体格检查和可观察到的临床异常。	
.....	
9. 时间(TIM):时间以及持续时间。包括症状、疾病的具体时间以及持续时间。	
10. 治疗手段(PAH):为了解决疾病或者缓解症状而施加给患者的治疗程序,干预措施。包括药物治疗、手术治疗、医学营养治疗等。	
11. 身体部位(BOP):身体部位或器官。	
根据用户输入的句子,识别对应的实体和实体类型。在句子中出现的实体,用"&&实体&&(实体类型)"表示,如:&&支气管肺炎&&(DIS)。	
以下是输出示例,不能输出其他对话以及对实体识别任务过程的解释：	
{RAG_text}	
用户输入：	
{query}	



图 3 输出表示

Figure 3 Output representation

表 2 本研究环境配置表

类别	型号
操作系统	Ubuntu 22.04.5 LTS
CPU	Intel(R) Xeon(R) Gold 6330 CPU @ 2.00GHz
显卡	NVIDIA GeForce RTX 4090
CUDA	12.4
显存	24GB
内存	134GB
Python	3.9.18

3.1 数据集 本文所使用的医学领域数据集包含自建电子病历数据集和瑞金医院糖尿病数据集。自建电子病历数据集包含四川省某市多家三甲医院真实的 600 份糖尿病电子病历数据和 600 份小儿肺炎电子病历数据。在医学专家的指导下,本文作者所在实验室的两位研究生 A 和 B 采用双盲多轮标注方法标注并构建了自建语料库。

根据病历特点,定义了 11 类医疗实体,实体类

型及其含义如表 3 所示。将自建数据集拆分为训练集和测试集,其中训练集有 748 420 个字符,实体数量有 107 624 个,测试集有 93 672 个字符,实体数量有 13 554 个。训练集用于构建实体向量数据库,以

支持语义检索与示例选取;测试集则作为实验评估集,用于检验所提出方法在命名实体识别任务中的性能表现。

表 3 自建数据集(糖尿病+小儿肺炎)

实体类型	含义	示例
疾病(DIS)	患者所患的具体疾病或诊断名称	糖尿病、小儿肺炎
症状(SYM)	患者自我感受到或主诉的主观不适	头晕、恶心
体征(SIG)	医生客观检查或仪器检测到的体征	心音低钝、呼吸音粗
身体部位(BOP)	涉及人体结构的名称	双下肢、腹部
时间(TIM)	具体时间及持续时间	1 年、3 天前
程度(DEG)	病情或症状的严重程度	剧烈、轻度
治疗手段(PAH)	施加给患者的治疗程序、干预措施	手术、二级护理
药物(DRU)	具体的药品名称	布洛芬、二甲双胍
药物修饰词(MDM)	用药的频率、剂量以及投药方式	一日两次、口服
医学检查(AUE)	实验室检查和辅助检查	CT、血常规
医学检查结果(RES)	通过医学检查得到的数值或结果	阳性、8.2mmol/L

瑞金医院糖尿病数据集来源于中文糖尿病研究领域的权威杂志,共包含 363 篇文档,约 250 万字,医疗实体划分为 15 类,实体类型和实体含义如表 4。将数据集拆分为训练集和测试集,其中训练

集有 662 106 个字符,实体数量有 48 714 个,测试集有 257 258 个字符,实体数量有 19 069 个。该数据集的训练集与测试集的作用同自建数据集一样。

表 4 瑞金医院糖尿病数据集

实体类型	含义	示例
Disease(DIS)	患者所患的具体疾病或诊断名称	糖尿病、高血压
Reason(REA)	引发疾病或症状的原因、诱因	营养状态差
Symptom(SYM)	临床表现,包括症状和体征	嗜睡、关节痛
Test(TES)	各类医学检查、实验室检查	空腹血糖、TG
Test_Value(TEV)	检查项目对应的数值、结果	23.7kg/m ² 、升高
Drug(DRU)	明确的药物名称	双膦酸盐
Frequency(FRE)	药物使用频率	每周 1 次、餐前
Amount(AMO)	用药剂量	2 ~ 5 g、20mg/d
Method(MET)	用药方式	口服、静脉
Treatment(TRA)	非药物的治疗方式	降糖治疗、PEI
Operation(OPE)	手术、介入行为	肠道手术
SideEff(SIE)	用药引起的不良反应或副作用	低血糖风险
Anatomy(ANA)	身体或器官部位名称	骨髓干细胞
Level(LEV)	描述病情、症状或指标程度的词语	急性、4 期
Duration(DUR)	症状、疾病或用药的持续时间	4 周、<10 年

3.2 评价指标 采用精确率 P 、召回率 R 以及 $F1$ 值这 3 个指标评估本文方法的性能。

精确率 P :模型预测为正例(即正确识别出的实体)中,实际真正正确的比例。

其中 TP 为真正例,表示被正确识别的实体数量, FP 为假正例,表示错误识别为实体的数量。

召回率 R :数据中所有实际的正例(即真实实体)中,被模型正确识别出的比例。

$$P = \frac{TP}{TP + FP} \tag{3-1}$$

$$R = \frac{TP}{TP + FN} \tag{3-2}$$

其中 FN 为假负例,表示模型未能识别出的真实实体数量。

$F1$ 值:精确率 P 和召回率 R 的调和平均。

$$F1 = \frac{2 \times P \times R}{P + R} \tag{3-3}$$

3.3 实验结果与分析 为验证本文所提出的结合语义检索与提示学习的大语言模型生成式命名实体识别方法的有效性,在自建小儿肺炎与糖尿病混

合病种的中文电子病历数据集、瑞金医院糖尿病数据集这两个数据集上进行了实验。比较模型涵盖传统序列标注模型 BERT+BiLSTM+CRF、采用 API 调用的大语言模型 GLM-4-Flash(表 5 简写为 GLM-4)和本地部署的开源模型 Qwen2-7B-Instruct(表 5 简写为 Qwen2)的多种提示设置(0-shot, few-shot),以及本文提出的基于语义增强的提示生成方法(RAG+few-shot)。实验结果见表 5。

表 5 实验结果

模型	自建数据集			瑞金医院数据集		
	Precision	Recall	F1	Precision	Recall	F1
BERT+BiLSTM+CRF	93.00	92.00	93.00	46.00	42.00	44.00
GLM-4+0-shot	16.03	6.35	9.10	14.94	9.12	11.33
GLM-4+5-shot	44.22	51.04	47.39	49.30	40.02	44.17
GLM-4+10-shot	48.79	55.63	51.98	51.73	40.99	45.74
RAG+GLM-4+1-shot	81.98	80.90	81.44	48.22	42.07	44.93
RAG+GLM-4+3-shot	88.27	89.38	88.82	52.62	49.59	51.06
RAG+GLM-4+5-shot	89.86	91.00	90.42	55.51	52.43	53.93
RAG+Qwen2 +1-shot	84.71	60.25	70.42	28.89	28.58	28.74
RAG+Qwen2 +3-shot	94.61	92.44	93.51	42.70	40.62	41.63
RAG+Qwen2 +5-shot	94.65	94.59	94.62	47.37	40.91	43.90

从实验结果可见,在自建电子病历数据集上,传统结构化模型 BERT+BiLSTM+CRF 实现了 93.00% 的 $F1$ 值,表现出较强的基线性能。而基于生成式大语言模型的方案在引入语义检索与 few-shot 提示后进一步提升了识别效果,尤其是 RAG+Qwen2-7B-Instruct+5-shot 模型达到了 94.62% 的 $F1$ 值,全面超越传统方法。这一结果表明:语义检索机制能够有效选取与目标文本语义相关的提示示例,在提升实体识别能力的同时,提高了大模型的泛化性和稳健性。进一步从 few-shot 配置来看,生成式模型对提示示例极为敏感。例如, GLM-4 模型在 0-shot 下的 $F1$ 值仅为 9.10%,而随着示例数量增加至 10-shot, $F1$ 值提升至 51.98%。在引入语义检索机制后,性能提升更加显著: RAG+GLM-4 在 5-shot 配置下即达 90.42%, Qwen2 系列即使仅使用 3 个示例也能实现 93.51% 的 $F1$ 值,在 5-shot 时达到峰值 94.62%。

在瑞金医院糖尿病数据集上,任务复杂度更高。该数据集中包含大量医学英文缩写、嵌套结构等复杂实体类型,导致实体识别任务面临更大挑战。传统结构化模型 BERT+BiLSTM+CRF 在该数据集上取得了 44.00% 的 $F1$ 值。相比之下, GLM-4 在

无语义增强的 5-shot 设置下仅达 44.17%,而引入语义增强策略后, RAG+GLM-4+5-shot 明显提升至 53.93%,性能显著提升,验证了语义检索机制在复杂医疗语境中的适用性和有效性。 Qwen2 系列在该数据集上的表现也有一定改善, RAG+Qwen2-7B-Instruct+5-shot 模型实现了 43.90% 的 $F1$ 值,达到了和传统模型差不多的效果。

综上所述,实验结果充分验证了本文方法在不同类型医学文本中的广泛适应性与有效性。尤其值得强调的是,自建混合病种数据集在实体类型覆盖全面、语义表达丰富且标注规范等方面具有显著优势,为模型提供了高质量的语义支持和训练基础。正是基于这一高质量标注,生成式大语言模型在该数据集上得以充分学习,在 5-shot 设置下达到了 94.62% 的 $F1$ 值,优于传统结构化方法,表现出更强的识别能力和泛化性能。相比之下,瑞金医院糖尿病数据集由于实体嵌套复杂、缩写频繁等特性,整体识别难度更高, $F1$ 值普遍偏低。然而,随着提示示例数量增加,模型表现持续上升,说明其具备良好的学习能力。而在引入语义检索机制后,性能提升尤为明显。这一结果从侧面进一步印证了高质量的标注语料对于医学 NER 任务至关重要,而自

建数据集的标注质量正是支撑模型取得优异性能的核心保障。

同时,在 few-shot 低资源配置下,模型也展现了出色的少样本学习能力,即便在示例数量极少的情下,也能依托语义提示完成精准的实体识别,充分体现了语义增强策略的有效性。因此,本文提出的方法不仅适用于结构复杂、表达多样的多病种电子病历场景,也为后续更大规模的医疗信息抽取任务提供了可扩展、可迁移的解决思路。

4 结 论

本文面向中文电子病历命名实体识别任务,提出了一种结合语义检索机制与提示学习的生成式大语言模型识别方法,以提升模型在低资源、多病种、真实医疗文本环境下的实体识别能力。本文方法通过构建实体向量数据库,从训练集中检索与输入语义相似的句子作为 few-shot 提示,引导大语言模型在无需参数微调的前提下完成实体识别任务。相比传统依赖大规模标注数据的序列标注方法,本文方法不仅有效缓解了标注资源匮乏的问题,同时充分发挥了生成式语言模型在语义建模与上下文理解方面的优势,适用于实体类型复杂的医学文本。

尽管如此,本研究仍存在一定局限性。一方面,生成式识别方法缺乏位置约束机制,在实体密集或边界模糊的场景中易产生幻觉式输出;另一方面,大语言模型推理资源开销较大,尚难以直接部署于资源受限的实际临床环境。此外,当前模型主要聚焦于实体层级,对于更高阶的结构化信息如事件、因果关系的抽取能力仍有待提升。

未来工作可从以下几个方面展开探索:(1)引入结构控制机制或后验验证模块,以提升输出实体的准确性与一致性;(2)借助参数高效微调方法,推动模型的轻量化与边缘部署;(3)拓展至医学事件、时间表达、因果关系等复杂信息的识别任务,进一步增强模型在医疗信息抽取中的应用广度。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 程洁:论文选题、撰稿;刘独玉:论文修改;陈思旭:论文修改;钱姝好:共同标注数据。

参考文献

[1] 祁鹏年,廖雨伦,覃 颢. 基于深度学习的中文命名实体识别研究综述[J]. 小型微型计算机系统, 2023, 44(9): 1857-

1868.
[2] YIN L B, MENG X, LI J X, et al. Relation extraction for massive news texts[J]. Comput Mater Continua, 2019, 60(1): 275-285.
[3] MA X, LU Y, LU Y, et al. Biomedical event extraction using a new error detection learning approach based on neural network [J]. Computers, Materials & Continua, 2020, 63 (2) : 923-941.
[4] ZHOU H, SHEN T, LIU X, et al. Survey of knowledge graph approaches and applications [J]. Journal on Artificial Intelligence, 2020, 2(2): 89-101.
[5] QIU J, LIU Y, CHAI Y H, et al. Dependency-based local attention approach to neural machine translation [J]. Comput Mater Continua, 2019, 59(2): 547-562.
[6] SHARMA Y, GUPTA S. Deep learning approaches for question answering system [J]. Procedia Comput Sci, 2018, 132: 785-794.
[7] DOU Z Y, WANG X, SHI S M, et al. Exploiting deep representations for natural language processing [J]. Neurocomputing, 2020, 386: 1-7.
[8] LEE J, YOON W, KIM S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining[J]. Bioinformatics, 2020, 36(4): 1234-1240.
[9] 隋 臣. 基于深度学习的中文命名实体识别研究[D]. 杭州: 浙江大学, 2018.
[10] ZHOU G D, SHEN D, ZHANG J, et al. Recognition of protein/gene names from text using an ensemble of classifiers[J]. BMC Bioinformatics, 2005, 6(Suppl 1): S7.
[11] JONNALAGADDA S, COHEN T, WU S, et al. Enhancing clinical concept extraction with distributional semantics [J]. J Biomed Inform, 2012, 45(1): 129-140.
[12] JIANG M, CHEN Y K, LIU M, et al. A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries [J]. J Am Med Inform Assoc, 2011, 18(5): 601-606.
[13] COLLOBERT R, WESTON J, BOTTOU L, et al. Natural language processing (almost) from scratch [J]. J. Mach. Learn. Res., 2011, 12: 2493-2537.
[14] HUANG Z, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging[J]. Arxiv, 2015.
[15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[A]. arXiv, 2023.
[16] PETERS M, NEUMANN M, IYYER M, et al. Deep contextualized word representations [C]//Proceedings of the 2018 Conference of the North American Chapter Of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). New Orleans, Louisiana. Stroudsburg, PA, USAACL, 2018: 2227-2237.
[17] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [A]. arXiv, 2019.
[18] HU X S, ZHANG H J, SUN Y J. Chinese medical short text

- matching model based on fine-tuning BERT-attention-BiLSTM [C]//2023 IEEE/ACIS 23rd International Conference on Computer and Information Science (ICIS). June 23–25, 2023, Wuxi, China. IEEE, 2023: 91–96.
- [19] 郭振华, 宋 波. 基于RBEGP的中文电子病历命名实体识别研究[J]. 电脑知识与技术, 2024, 20(16): 6–10.
- [20] LIU P F, YUAN W Z, FU J L, et al. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing [J]. ACM Comput Surv, 2023, 55 (9) : 1–35.
- [21] SCHICK T, SCHÜTZE H. Exploiting cloze-questions for few-shot text classification and natural language inference [C]// Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Online. Stroudsburg, PA, USAACL, 2021: 255–269.
- [22] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners [C] //Proceedings of the 34th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc. , 2020: 1877–1901.
- [23] CUI L Y, WU Y, LIU J, et al. Template-based named entity recognition using BART [C]//Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Online. Stroudsburg, PA, USAACL, 2021: 1835–1845.
- [24] MA R T, ZHOU X, GUI T, et al. Template-free prompt tuning for few-shot NER [C]//Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Seattle, USA. Stroudsburg, PA, USAACL, 2022: 5721–5732.
- [25] LEE D H, KADAKIA A, TAN K M, et al. Good examples make A faster learner: simple demonstration-based learning for low-resource NER [C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Dublin, Ireland. Stroudsburg, PA, USAACL, 2022: 2687–2700.
- [26] SHEN Y L, TAN Z Q, WU S H, et al. PromptNER: prompt locating and typing for named entity recognition [C]// Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Toronto, Canada. Stroudsburg, PA, USAACL, 2023: 12492–12507.
- [27] XIAO S T, LIU Z, ZHANG P T, et al. C-pack: packed resources for general Chinese embeddings [C]//Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. Washington DC USA. ACM, 2024: 641–649.

引用本文

程 洁, 刘独玉, 陈思旭, 等. 基于大语言模型的中文电子病历命名实体识别[J]. 元宇宙医学, 2025, 2(3): 39–47.

CHENG J, LIU D Y, CHEN S X, et al. Named entity recognition in chinese electronic medical records based on large language models[J]. Metaverse Med, 2025, 2(3): 39–47.

人工智能工具医学使用的科学伦理与规范重构



白春学¹, 贾泽军¹, 杨达伟¹, 高承实^{2*}

1. 复旦大学附属中山医院, 上海 200032

2. 安徽栈谷科技有限公司, 池州 247100

[摘要] 随着人工智能技术的快速发展, AI 工具在医学研究与学术写作中展现出显著作用, 从数据处理、文献检索到论文写作、图表可视化和跨学科合作均实现效率提升。AI 不仅是辅助工具, 更逐渐成为科研伙伴, 在假说生成、实验设计和多模态数据分析中发挥创造性, 推动科研范式向“人机共研”转型。当前已出现多种应用场景, 如医学论文初稿生成、科研诚信检测、医学影像报告和科研图表的自动撰写。与此同时, AI 的广泛应用也对署名权、数据可追溯性、内容可信度及隐私保护提出挑战。国际指南如 ICMJE《医学期刊学术出版推荐规范》以及 Science、Nature 的公开声明均强调, AI 工具不得列为作者, 其使用需透明披露且责任归属人类研究者。为此, 亟需建立以透明披露、责任归属、质量可验证和合规开放为核心的规范体系, 以确保 AI 在医学科研中实现增效与创新的双重价值, 并为数字时代可持续科研生态奠定基础。

[关键词] 人工智能; 医学科研范式; 数据驱动科学; 生成式模型; 科研伦理

[中图分类号] R-3 **[文献标志码]** A

Reconstruction of scientific ethics and norms for the medical use of artificial intelligence tools

BAI Chunxue¹, JIA Zejun¹, YANG Dawei¹, GAO Chengshi^{2*}

1. Zhongshan Hospital, Fudan University, Shanghai 200032, China

2. Anhui Stack Alley Technology Co., Ltd, Chizhou 247100, Anhui, China

[Abstract] With the rapid development of artificial intelligence, AI tools have demonstrated significant value in medical research and academic writing, enhancing efficiency in data processing, literature retrieval, manuscript drafting, visualization, and interdisciplinary collaboration. Beyond serving as auxiliary tools, AI is increasingly becoming a research partner, contributing to hypothesis generation, experimental design, and multimodal data analysis, thereby fostering a paradigm shift toward “human - AI co-research”. Typical applications include rapid drafting of medical manuscripts, research integrity checks, and automated generation of imaging reports and scientific figures. However, the widespread adoption of AI also raises challenges concerning authorship, data traceability, content reliability, and privacy protection. International guidelines such as the ICMJE Recommendations and public statements from Science and Nature explicitly emphasize that AI tools cannot be listed as authors, that their use must be transparently disclosed, and that ultimate responsibility lies with human researchers. Therefore, it is urgent to establish a normative framework centered on transparency, accountability, verifiability, and compliant openness, ensuring that AI delivers both efficiency and innovation while laying the foundation for a sustainable research ecosystem in the digital era.

[Key Words] artificial intelligence; medical research paradigm; data-driven science; generative models; research ethics

近年来, 人工智能(artificial intelligence, AI)技术的迅速发展, 尤其是大语言模型、深度学习与多模态算法的广泛应用, 正在深刻改变医学研究与学术写作的方式。从医学影像智能诊断到临床数据挖掘, 从自动化文献综述生成到学术论文初稿撰写, AI 作为工具的能力不断突破传统科研的效率与边界。尤其是在 ChatGPT 等生成式语言模型的推动下, 科研人员在知识发现、论文写作与学术交流等环节中, 获得了前所未有的支持与便利, 甚至被

成为医生第二大脑。面对这一新兴力量, 医学研究和写作正站在一个新的历史关口, 即如何在充分发挥 AI 工具优势的同时, 确保学术的严谨性与科研的规范性?

传统意义上的学术规范与科研伦理, 是在以人工为中心的研究范式中逐步形成的。这一体系强调研究者对数据收集、分析与论文写作负有直接、不可替代的责任。然而, 在 AI 工具已经能够承担部分甚至大部分科研劳动的今天, 原有规范的适用

[收稿日期] 2025-06-23

[接受日期] 2025-09-15

[作者简介] 白春学, 博士, 教授, 主任医师. E-mail: bai.chunxue@zs-hospital.sh.cn

* 通信作者 (Corresponding author). Tel: 13838001036. E-mail: 13838001036@163.com

性正受到挑战。例如,研究人员是否可以在未披露的情况下使用 AI 生成文献综述? 如果论文的部分文本由 AI 起草,是否仍然符合署名与原创性的学术伦理? 面对这些问题,学术共同体在短期内往往采取“禁止”或“限制”策略,但这种基于旧有研究方式所延续的规范,很可能会阻碍新工具的合理使用,从而错失推动科学进步的重要契机。

医学研究与写作的核心目标是促进科学发展、推动其临床应用、改善人类健康。在这一根本目标的指引下,学术规范也应当与时俱进,主动吸纳能够提升科研效率与质量的新工具,而不是成为阻碍创新的桎梏。AI 工具不仅可以减轻研究者的重复性劳动,更能够在跨学科数据整合、临床证据快速生成、科研假说的形成与验证等方面发挥独特作用。如果能够在制度层面合理引导和规范 AI 的使用,它们将成为科研共同体的“增能工具”(augmentation tools),而不是学术伦理的威胁。

与此同时,我们必须正视 AI 工具可能带来的风险与挑战。一方面,生成模型仍然存在“幻觉”问题,其输出内容可能缺乏可验证性;另一方面,过度依赖 AI 可能引发学术诚信危机,模糊研究者与工具的责任边界。此外,医学研究具有高度的伦理敏感性,涉及患者隐私、临床数据安全与社会责任,AI 工具的应用必须严格嵌入医学伦理框架,避免技术滥用。因此,制定新的学术规范,不仅是技术问题,更是价值与责任的再平衡。

本文旨在全面审视 AI 工具在医学研究与论文写作中的应用现状,探讨其对科研方法论、成果呈现与学术规范的冲击,并提出新的伦理与规范框架,以实现效率、质量与责任的平衡。我们认为,学术共同体不应当回避 AI 工具的应用,而应以积极开放的姿态,推动规范重构,让人工智能成为医学科学的新型基础设施。这不仅是医学研究在数字时代的必然选择,也是实现科学可持续发展的关键路径。

1 AI 工具在医学研究与写作中的应用现状

AI 在医学中的应用已从最初的辅助性工具逐步发展为知识生产与科研范式转型的重要工具。医学是一门高度依赖数据与证据的学科,从影像学、临床记录到基因组学与科研文献,每一个环节都蕴含着潜在的规律。然而,长期以来,受制于计算能力与分析工具的不足,这些数据未能被充分挖掘。随着深度学习、自然语言处理(natural language processing, NLP)以及大语言模型(large language

model, LLMs)的突破,AI 已经能够对海量、异质化的医学数据进行系统化处理与知识抽取,由此推动科研流程在数据处理、成果写作以及协作教育等方面的深刻变革。

1.1 数据处理与知识发现 医学科研的核心在于通过数据揭示规律、发现新知。AI 技术的兴起,使数据处理与知识发现成为医学 AI 应用的前沿议题。

(1)医学影像诊断辅助。医学影像是 AI 技术率先落地的重要领域。早期研究集中在计算机辅助检测(computer aided detection, CAD),如乳腺 X 线片中标注可疑区域^[1]。随着卷积神经网络(convolutional neural networks, CNN)在 2012 年 ImageNet 挑战中的突破^[2],深度学习成为主流方法,并在肺结节检测、乳腺癌筛查、糖尿病视网膜病变识别等任务中展现出接近甚至超越专科医生的水平^[3]。

近年来,多模态影像(CT、MRI、超声)的结合联袂纵向数据建模使 AI 能够实现疾病动态演化的追踪。同时,研究重点从“替代医生诊断”转向“增强医生决策”。例如,Grad-CAM、注意力机制等可解释性方法揭示了模型的决策依据^[4],为医生提供额外参考。更重要的是,大规模影像数据的自动标注与分割逐渐成为临床科研的重要工具,推动疾病亚型探索与表型分层研究^[5]。

(2)大规模文献综述与知识图谱。医学研究高度依赖文献积累。传统的系统综述与 Meta 分析往往耗时数月至数年,难以应对文献爆炸式增长。早期的信息检索工具(如 PubMed 的关键词匹配)只能有限缓解^[6]。随着 NLP 与信息抽取技术的发展,大规模文献的自动筛选、分类与主题聚合成为可能。

近年来,LLMs 显著提升了文献处理能力。一方面,它们能基于数万篇文章生成研究趋势报告,捕捉跨领域联系^[7];另一方面,知识图谱逐渐成熟,能够将疾病、基因、药物关系结构化呈现,并支持药物再定位与新发疾病机制研究^[8]。由此,AI 已不再只是“加速综述工具”,而是可能成为科研发现的催化剂。

(3)临床数据挖掘与模式识别。电子病历(electronic health records, EHR)、可穿戴设备与多组学数据的积累,使临床研究进入大数据时代。早期数据挖掘依赖统计模型与浅层机器学习,用于风险预测或简单分类,但难以处理高维、非结构化与多模态数据。深度学习突破了这一限制,能够发现潜在模式与风险因子^[9]。

已有研究显示,基于 EHR 的预测模型在心衰、

败血症等风险预警中具有显著优势^[10]。同时,多组学数据的整合推动了精准医学的进展。例如,AI可通过模式识别优化临床试验入组,提升统计效力与临床相关性。更重要的是,AI模型还能通过聚类或潜变量建模发现新疾病亚群或机制性联系,从而推动疾病分类体系更新。

总体来看,AI在医学影像、文献综述与临床数据挖掘中,不仅提升了效率,更在生成新知识方面展现出潜力。它正逐步从“辅助工具”演化为“知识引擎”,为医学研究范式的转型奠定基础。

1.2 科研写作与成果呈现 科研写作是知识转化为学术成果的重要环节。随着AI技术的发展,学术写作的多个步骤开始被重塑。

(1)自动化摘要与文献综述生成。摘要与综述是科研写作中最耗时的部分。AI工具能够根据研究主题或初步数据自动生成研究背景与初步文献回顾,为研究者提供写作框架^[11]。这类工具并不替代学术判断,而是通过生成草稿减轻重复劳动,使研究者集中于学术价值与逻辑论证。已有研究表明,AI生成的综述在结构完整性与覆盖率上可与人工撰写相近,但仍需专家审校^[12]。

(2)学术语言优化与跨语种写作。英语是国际医学期刊的主要发表语言,长期以来对非母语研究者形成障碍。基于LLM的学术工具能够进行语法校正、风格优化甚至跨语种翻译。研究表明,这类工具能显著提高稿件被国际期刊接受的概率^[13]。在知识全球化的背景下,AI在语言优化中的作用不仅是“润色”,更是推动学术成果跨语种传播的重要途径。

(3)图表生成与可视化。数据可视化是科研表达的重要组成部分。传统图表工具在效率与复杂度方面存在局限,而AI能够根据数据自动生成统计图表,甚至实现复杂医学过程的三维可视化^[14]。近年来,AI驱动交互式可视化平台逐渐兴起,在临床研究报告和跨学科合作中显著提升了科研成果的可读性与影响力^[15]。

AI在科研写作环节的应用,核心价值并不在于替代研究者,而在于重塑学术表达的过程。它通过生成草稿、优化语言与可视化表达,降低了写作壁垒,推动了学术成果的全球传播与跨学科共享。

1.3 科研协作与教育 除数据与写作外,AI还在科研协作与教育领域展现潜力。医学科研往往依赖跨学科、多地域合作,而教育与培训则决定了科研共同体的可持续发展。

(1)智能助理与科研团队协作。科研团队日益

成为知识生产的主要单元^[16]。然而,跨地域、跨学科合作往往面临沟通与协调成本。AI工具作为“智能助理”,能够辅助会议安排、任务跟踪与实时文献检索。例如,自动生成会议纪要与行动清单的NLP系统已被用于科研团队管理^[17]。此外,AI驱动的知识图谱与推荐系统能揭示潜在跨学科联系,推动创新议题的产生^[18]。在新冠疫情等全球公共卫生危机中,这类工具加速了国际科研协作与知识共享^[19]。

(2)教育、培训与知识普及。医学教育正在经历由AI驱动的变革。自适应学习与智能辅导系统(intelligent tutoring system, ITS)能够根据学生表现动态调整教学进度与难度^[20]。在医学教育中,这种方法已用于病例推理、科研写作训练等场景^[21]。

在科研培训中,AI工具能模拟科研写作流程、自动批改研究设计,帮助青年学者快速掌握学术规范^[22]。在社会层面,AI驱动的开放科普平台能将复杂科研成果转化为通俗内容^[23],缩短科研与公众的知识鸿沟。

AI在科研协作与教育中的应用,推动了科研体系从“个体驱动”走向“智能协同”。它不仅提升了团队效率和教育质量,还扩展了知识传播的社会影响力。

综上所述,AI在医学科研的三个核心环节——数据处理与知识发现、科研写作与成果呈现、科研协作与教育——均展现出显著的学术价值和应用前景。在数据层面,AI实现了从影像识别到临床模式发现的知识提炼;在写作层面,它降低了科研表达与传播的门槛;在协作与教育层面,它推动了科研共同体的组织方式与人才培养模式的转型。简言之,AI正在成为医学知识生产的重要基础设施,为精准医学与转化医学的发展奠定坚实基础。

2 人工智能对科研规范与学术伦理的冲击

人工智能工具在医学研究与写作中展现出强大潜力,但与此同时,现有的科研规范和学术伦理体系正面临前所未有的挑战。传统的学术规则主要基于“人”在知识生产中的核心地位,而AI工具的广泛应用正在模糊研究者与工具的边界,从而引发了制度与实践的紧张关系。

2.1 现有学术规范的局限性 在深入分析AI带来的具体伦理挑战之前,有必要首先反思现有学术规范的生成逻辑及其内在局限。学术伦理体系并非一蹴而就,而是伴随着科研实践逐渐积累、形成的社会契约。其核心假设是,知识生产依赖于具有人

格与责任意识的研究者。然而,当 AI 工具在研究全过程中被广泛应用时,这一假设的适用性正在逐渐减弱。

(1)基于人工写作与传统研究方式形成的伦理体系。现有的学术规范与伦理框架,主要是在人工主导的研究与写作环境下逐渐确立的。署名权的界定、数据的管理规范,以及论文的原创性要求,均以“研究者是唯一知识创造主体”为前提。例如,《温哥华作者署名标准》明确要求作者必须对研究设计、数据分析和论文撰写承担实质性责任^[24]。然而,AI 工具的出现突破了这一假设,它们不仅能在写作中生成学术性文本,还能在研究过程中进行模式识别与数据推理,事实上参与了知识生产。由此引发的疑问包括,AI 生成的内容能否被认定为“原创”?若 AI 在研究设计中提出了重要假设,是否应被视为研究贡献?这些问题均超出了传统伦理框架的边界。

(2)对 AI 工具的使用缺乏明确规定。近年来,部分顶级期刊(如 *Science*、*Nature*)已发布声明,禁止将 AI 工具列为论文作者^[25]。然而,对其在研究与写作过程中的合理使用范围,学界尚未达成统一共识。一些学术机构采取保守排斥的态度,强调 AI 输出不可直接作为研究成果引用;另一些机构则主张有限容纳,允许在语言润色、翻译等环节使用 AI,但要求披露其作用。正如 Lund 等人^[26]所指出的,这种不一致性使得研究人员在使用 AI 工具时处于灰色地带,既担心违反规范,又难以放弃其高效性。换言之,制度供给滞后于实践需求,导致学术共同体在面对 AI 时缺乏明确的行为指南。

2.2 潜在风险 在理清学术规范的局限后,更为迫切的问题是识别 AI 在科研中的具体风险。这些风险不仅威胁到学术诚信,还可能动摇学术共同体赖以维系的信任机制。

(1)虚假数据和幻觉内容带来的学术诚信问题。生成式 AI 工具的“幻觉”现象,已成为学术界最为担忧的风险之一^[27]。所谓“幻觉”,是指 AI 在缺乏事实支撑的情况下生成虚假或错误内容,但表现形式却极为逼真。在医学研究中,这种现象可能导致引用不存在的文献、生成虚构的数据,甚至得出未经验证的结论。一旦缺乏严格的人工复核,研究成果的真实性与可重复性将受到严重威胁。与传统的学术不端(如抄袭、篡改数据)不同,AI 生成的虚假信息往往难以通过常规审查手段识别,从而对学术诚信构成系统性冲击。

(2)作者身份与署名伦理。AI 工具是否应被视

为论文的“作者”,是学术伦理面临的核心争议之一。根据现行的国际规范,作者必须具备承担法律与学术责任的能力,而 AI 显然不具备这一条件。因此,多数学术机构明确否认 AI 的作者资格。然而,完全忽视 AI 的作用也可能损害研究的透明度与可追溯性。学界正在探索折中方案,例如在论文的“方法”或“致谢”部分披露 AI 的具体作用^[28]。这表明,署名伦理的挑战并非在于是否承认 AI,而在于如何合理界定其“工具”属性与“贡献”性质。

(3)数据隐私与医学伦理冲突。医学研究高度依赖患者数据,而隐私保护始终是核心伦理议题。AI 工具在处理临床数据时,如果缺乏严格的数据脱敏与合规机制,可能导致敏感信息泄露。此外,AI 的“黑箱”特性意味着其决策过程难以解释,这与医学研究所强调的透明性和责任追溯相冲突^[29]。因此,如何在保障数据安全与患者隐私的同时,合理发挥 AI 的分析能力,成为医学伦理面临的又一重大难题。

2.3 制度滞后与实践冲突 在识别规范局限与潜在风险之后,必须正视一个现实问题,即 AI 在科研实践中的普及速度,已经远远超出了学术制度的响应速度。这种不协调不仅导致研究者的行为困惑,也削弱了学术共同体的制度权威。

(1)科研人员广泛使用与期刊、机构限制。大量调查^[30]显示,科研人员已在不同程度上依赖 AI 工具进行文献检索、语言润色和初步数据分析。然而,许多期刊和科研机构对此仍持保守态度,对 AI 使用缺乏系统性规范,甚至采取一刀切的禁止政策。这种制度与实践的错位,会导致研究者在实际操作中不得不“隐性使用”AI,从而形成事实上的规范失效与制度空转。

(2)制度更新滞后的风险。更为严重的是,学术规范的更新速度远远落后于 AI 技术的发展节奏。如果不能及时建立与之匹配的伦理与监管框架,既可能抑制 AI 的合理应用,也可能纵容无序使用带来的风险^[31]。更重要的是,制度滞后可能使学术共同体在国际规则制定中失去主动权,从而在全球学术竞争中处于不利地位。因此,制度与实践的错位不仅是技术应用的短期矛盾,更是学术治理的战略性挑战。

总体而言,人工智能工具对科研规范与学术伦理的冲击具有深远性和系统性。一方面,现有规范基于传统人工写作与研究方式建立,无法充分应对 AI 带来的原创性、署名与数据合规等问题;另一方面,生成式 AI 的幻觉风险、黑箱特性与隐私挑战,

进一步加剧了学术共同体的信任危机。在制度层面,科研人员的普遍使用与期刊、机构的保守限制之间存在显著张力,而规范更新的滞后则可能导致学术秩序的失衡。由此可见,医学研究这一伦理高度敏感的领域亟需建立一套既能保障科学严谨性,又能包容技术创新的全新学术规范与伦理框架。

3 新的科研工具使用伦理与规范框架

人工智能工具的广泛应用已成为医学研究与写作的重要趋势。面对新技术对学术规范与伦理体系的挑战,学术共同体不能仅停留在“禁止”与“限制”的被动应对层面,而应主动构建一套新的科研工具使用框架,使其既能保障科研质量与学术诚信,又能充分发挥 AI 的增效作用。本文提出四项核心原则,分别是工具透明化原则、责任归属原则、质量与可验证性原则、合规与开放原则。

3.1 工具透明化原则 在新型科研生态中,透明性被认为是重建信任的首要前提。缺乏披露不仅可能引发对学术成果真实性的质疑,还会加深学术共同体对 AI “隐性使用”的不安。因此,工具透明化原则应作为框架的基础性要求。

(1)明确披露使用范围与功能。科研人员在研究与写作过程中,应如实披露 AI 工具的使用情况,包括具体应用环节(如文献综述撰写、语言润色、数据模式识别)、所使用的工具类型(如 ChatGPT、Claude、Copilot 等),以及工具在成果形成中所扮演的角色。这一做法与学术研究对方法透明性的基本要求相一致,有助于提升成果的可追溯性^[32]。

(2)制度化披露机制。透明不仅依赖于个人自律,更需要制度保障。期刊与科研机构应建立统一的披露机制。例如,在论文投稿环节设立“AI 使用声明”模块,要求作者说明是否以及如何使用 AI 工具。通过制度化披露,可以有效避免因模糊使用引发的争议,同时为后续评估 AI 工具的实际学术影响提供系统性数据支撑^[33]。

3.2 责任归属原则 在强调透明的同时,更为关键的问题是如何界定责任。AI 工具虽能辅助研究,但其本质仍是工具,缺乏法律人格与责任意识。因此,责任归属原则应当成为学术伦理框架的核心支点。

论文作者要对最终内容负责任。无论 AI 工具在研究与写作中发挥何种作用,论文作者始终是科研成果的唯一责任主体。研究结论的科学性、数据的真实性、文字表述的准确性,都必须由作者亲自审查与确认。这一责任不可由 AI 工具分担或

替代。

AI 不具备署名资格。学界对 AI 是否应被列为作者的讨论已趋于共识,即 AI 工具不具备法律主体地位,无法承担学术不端或失实结论的责任,因此不应列入作者署名。然而,完全忽略 AI 的作用可能削弱成果透明度。因此,在方法学部分或致谢部分合理标注 AI 工具的辅助角色,既确保署名伦理的清晰性,又维护学术交流的开放性。

3.3 质量与可验证性原则 即便有了透明与责任机制,如果缺乏质量与可验证性的保障,科研成果仍将失去其学术价值。鉴于生成式 AI 存在“幻觉”和不稳定性,学术共同体必须构建可验证性机制。

(1)保证 AI 生成结果的可追溯性。医学研究对数据与结论的真实性有极高要求。科研人员在使用 AI 工具生成内容时,应确保结果均可追溯至可信来源。例如,AI 生成的文献综述必须逐一核实引用是否真实存在,并确认观点与原始研究相符。

(2)建立人工复核机制。生成式 AI 的输出不可直接等同于可靠证据。科研人员应将其作为“初稿”或“参考”,而非“最终版本”,并建立人工复核机制。在制度层面,期刊编辑与同行评审也应加强把关,确保论文不因未经验证的 AI 输出而损害学术严谨性。

(3)推动可验证的技术规范。从长远来看,开发者与学术界可合作推动 AI 工具引入可验证性设计。例如,自动生成参考来源、标记生成与编辑过程的“数字水印”。这些技术与规范结合的措施,有助于缓解“黑箱输出”带来的风险。

3.4 合规与开放原则 在强调透明、责任与质量之后,还必须正视更为宏观的制度问题,即如何在合规性与开放性之间寻求平衡。医学研究不仅涉及学术规范,还深受法律、隐私与伦理制度的制约。

(1)兼容医学伦理与数据隐私保护。医学研究高度依赖患者数据,因此 AI 工具的应用必须严格遵循隐私与伦理审查要求。临床数据在输入 AI 模型前应充分脱敏,且不得上传至未经审查的公共平台。同时,研究团队应优先选择本地化、安全可控的 AI 使用环境,以确保数据合规流转。

(2)鼓励国际化标准与共享规范。AI 工具的应用是全球性现象,规范的制定若仅局限于单一国家或地区,将导致制度割裂。国际学术机构如 ICMJE 与 WHO,可以在推动 AI 使用伦理与规范共识方面发挥关键作用。这种跨国框架不仅能增强学术共同体的信任,还能提升研究的可比性与协同力。

(3)平衡开放性与约束性

规范的制定既不能过于宽松,以致放任滥用;也不能过于严格,导致科研人员无法正当使用 AI。理想的做法是在开放与约束之间建立动态平衡,既承认 AI 工具的学术价值,鼓励其合理应用;又通过制度与伦理约束,防止其偏离科学研究的核心价值。

新的科研工具使用伦理与规范框架,不应仅是“补丁式”的修正,而应是一种系统性重构。通过确立工具透明化、责任归属、质量与可验证性、合规与开放四大原则,可以在保障学术诚信与科研质量的同时,最大限度发挥人工智能的增效作用。这一框架既回应了医学研究的现实需求,也为全球学术共同体在数字时代重塑学术规范与科学伦理提供了参考蓝图。

4 AI 赋能下的医学研究与写作新范式

人工智能工具的快速演化正在推动医学研究与写作进入全新的发展阶段。从最初的效率提升,到科研方法论的重构,再到学术共同体与出版体系的制度创新,AI 的介入不仅改变了研究的形式,更深刻影响了知识生产的逻辑。展望未来,AI 有望推动医学研究与写作迈向新的学术范式。

4.1 从效率工具到科研伙伴 当前,AI 工具多被视作“辅助性”工具,用于数据处理、语言润色或图表生成。但随着技术的进步,AI 将逐渐从效率工具演变为科研伙伴,能够在假说生成、研究设计与跨学科整合中发挥创造性作用。例如,AI 可以基于大规模医学数据库,快速提出潜在的因果关系假设,辅助研究者选择最优研究路径。

更进一步,AI 的推理能力将突破传统统计模型的局限,能够在多模态数据(影像、基因组、临床记录等)中发现复杂的非线性关系,为精准医学提供更为细致的研究视角。在这一过程中,研究者的角色将从单一的知识生产者,转向“问题提出者与判断者”,而 AI 则承担“探索与发现者”的功能,二者形成动态互动的人机共研模式。

4.2 可编程科学与智能实验设计的科研方法论再定义 AI 的嵌入将推动科研方法论发生根本性变革。传统科研强调人工制定实验设计与数据分析方案,而在 AI 赋能下,科研将走向“可编程科学”。研究人员可以通过自然语言或代码指令,调用 AI 模型自动化完成实验流程,从假设生成到数据收集、从建模分析到结果解释,形成高度智能化的“端到端科研链”。

这一变革带来的最大优势在于动态迭代与实

时优化。不同于传统实验设计的固定性,AI 驱动的实验可以在过程中不断修正假设和参数配置,实现“边研究边优化”。例如,在临床试验中,AI 能够根据实时数据反馈动态调整受试者分组策略,从而缩短实验周期,提高试验效率与结论可靠性。这种“智能实验”模式,不仅提升了科研效率,还可能重塑科学的基本范式,即从“人类驱动”走向“人机协同驱动”的知识生产逻辑。

4.3 期刊与学术共同体的规范创新 随着 AI 的深度应用,学术出版与共同体运作也将迎来结构性革新。

(1)出版流程智能化。未来,医学期刊的出版流程将因 AI 的介入而大幅革新。AI 可以在稿件初审中自动完成语言校对、格式核查、相似度检测,甚至进行初步的逻辑一致性评估。这不仅缩短了出版周期,也能将人力资源释放到更高价值的内容评估环节,从而提升整体出版效率。

(2)审稿机制革新。AI 也将改变同行评审机制。未来可能出现“人机联合审稿”模式。AI 工具先对论文的事实准确性、数据可验证性进行快速审查,随后由专家学者对研究价值与创新性进行深度评议。这种双重机制不仅能够减轻审稿压力,还能提升评审的客观性与一致性。更重要的是,通过大规模积累 AI 辅助审稿数据,还能推动形成客观化的“科研质量指标”,减少传统同行评审中的主观偏差。

4.4 元宇宙医学语境下的 AI 学术生态 在进一步展望中,AI 的学术功能不应局限于当下的研究与出版体系,更应该将其延伸至更具沉浸感和协作性的未来科研空间中。随着元宇宙概念与医学研究的交融,AI 将作为底层驱动力,支撑起跨空间、跨学科、跨机构的科研互动与知识共建。

(1)跨空间与沉浸式科研协作。在元宇宙医学的发展框架下,AI 工具的作用将进一步扩展至跨空间、跨学科的知识交互与协作中。借助沉浸式环境与虚拟科研空间,研究人员可以在元宇宙中共享实时实验数据,利用 AI 进行智能化模拟实验与可视化呈现,实现“虚拟临床试验”与“数字孪生患者”的学术探索。

(2)智能化虚拟科研基础设施。AI 将成为元宇宙科研生态的“智能中枢”。例如,在虚拟实验室中,AI 可以自动化协调资源调度、实验流程管理与数据分析;在全球科研网络中,AI 可作为翻译、对话与推理的中介,使跨语言、跨文化的科学交流无缝进行。这一趋势意味着,AI 不仅是研究过程中的

“助手”,更是未来学术生态的制度性与技术性基础设施。

总体来看,人工智能正在推动医学研究与写作实现三重转型,即从效率驱动到伙伴协作,从传统方法论到智能化科学,从静态规范到动态共建,其影响不仅限于提高科研生产力,更将深刻地体现在对学术规范、知识逻辑与学术生态的重构上。在元宇宙医学的语境下,AI 不仅是技术手段,更是学术共同体的新型基础设施。

未来,如何在促进科研创新的同时确保学术诚信、伦理合规与社会责任,将成为学术界与产业界共同面临的重大课题。只有通过科学合理的制度设计与跨学科的共同治理,人机共研的学术生态才能真正加速人类对生命健康的认知与实践。

5 结 语

人工智能工具的迅速发展普及,正以前所未有的深度与广度重塑医学研究与学术写作的生态。从数据处理、文献检索到知识发现,从科研写作、图表可视化到学术交流与合作,AI 的介入不仅显著提高了效率,更在研究设计、假说生成及跨学科整合中发挥了潜在的创造性作用。可以说,AI 不仅是研究者的“智能助手”,更正在成为医学知识生产的“伙伴”,甚至成为医生的第二大脑,推动科研从传统的人力密集模式向“人机共研”的协作化、智能化新范式转型。

然而,这一技术变革同时也带来了学术伦理、数据安全与制度规范的深层挑战。现有学术规范体系过度依赖人工研究与写作假设,未能充分覆盖 AI 的使用场景,导致署名权、数据可追溯性、内容可信度及患者隐私保护等方面出现潜在冲突。制度滞后与实践高速发展的不平衡,既可能引发学术诚信风险,也可能阻碍科研创新的合理开展。

因此,学术共同体必须主动建构一套新的科研伦理与规范框架。核心原则应涵盖透明披露 AI 使用、明确责任归属、保证结果质量与可验证性、遵循合规与开放标准。在此基础上,在确保科研严谨性与社会责任的前提下,AI 工具的应用才能够进一步实现科学发现的加速、跨学科合作的拓展和医学研究模式的智能化升级。长远来看,AI 将成为医学科研与知识生产的基础设施,与大数据、云计算和其他数字技术共同构建数字时代的可持续科研生态。

伦理声明 无。

利益冲突 所有作者声明不存在利益冲突。

作者贡献 白春学:选题、修改论文;贾泽军、杨达伟:参与讨论;高承实:选题、撰写、修改论文。

参考文献

- [1] DOI K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential [J]. Comput Med Imaging Graph, 2007, 31(4/5): 198–211.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]// Advances in Neural Information Processing Systems 25. 2012, 1097–1105.
- [3] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks [J]. Nature, 2017, 542(7639): 115–118.
- [4] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization [C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22–29, 2017, Venice, Italy. IEEE, 2017: 618–626.
- [5] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis [J]. Med Image Anal, 2017, 42: 60–88.
- [6] MOHER D, LIBERATI A, TETZLAFF J, et al. 系统评价与荟萃分析优先报告条目 (PRISMA) 声明 [J]. 中国循证医学杂志, 2009, 9(9): 1085–1091.
- [7] LEE B, LEE S, KIM D, et al. Large language models for systematic review automation: evidence synthesis at scale [J]. Journal of Medical Internet Research, 2023, 25: e48625.
- [8] PERCHA B, ALTMAN R B. A global network of biomedical relationships derived from text [J]. Bioinformatics, 2018, 34(15): 2614–2624.
- [9] MIOTTO R, LI L, KIDD B A, et al. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records [J]. Sci Rep, 2016, 6: 26094.
- [10] SHICKEL B, TIGHE P J, BIHORAC A, et al. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis [J]. IEEE J Biomed Health Inform, 2018, 22(5): 1589–1604.
- [11] ELHADAD N. AI-assisted drafting of literature reviews: benefits and limitations [J]. Journal of the American Medical Informatics Association, 2019, 26(8–9): 785–787.
- [12] HOPE T, CLARK J, BENDER E M, et al. Can large language models accurately summarize systematic reviews? A human evaluation [J]. Journal of the American Medical Informatics Association, 2022, 29(11): 1923–1932.
- [13] KIM Y, LEE M, KIM D, et al. (2023). Towards Explainable AI Writing Assistants for Non-native English Speakers. arXiv.
- [14] RAUBER P E, FADEL S G, FALCÃO A X, et al. Visualizing the hidden activity of artificial neural networks [J]. IEEE Trans Vis Comput Graph, 2017, 23(1): 101–110.
- [15] ZHANG X, WANG Y, LIU C. AI-driven interactive visualization platforms in clinical research reporting: a

- systematic review[J]. *Journal of Biomedical Informatics*, 2023, 140: 104339.
- [16] WUCHTY S, JONES B F, UZZI B. The increasing dominance of teams in production of knowledge [J]. *Science*, 2007, 316 (5827): 1036–1039.
- [17] SEE A, LIU P J, MANNING C D. Get to the point: summarization with pointer-generator networks [C]//*Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vancouver, Canada. Stroudsburg, PA, USA: ACL, 2017.
- [18] PERONI S, SHOTTON D. OpenCitations, an infrastructure organization for open scholarship [J]. *Quant Sci Stud*, 2020, 1 (1): 428–444.
- [19] NIAZI M K K, SIDDIQUI S A. COVID-19 research collaboration networks: a 50-country analysis [J]. *Scientometrics*, 2020, 125(3): 2557–2576.
- [20] KOEDINGER K R, CORBETT A T, PERFETTI C. The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning [J]. *Cogn Sci*, 2012, 36(5): 757–798.
- [21] CHAN L, GOH P S, WONG M L. Adaptive learning and intelligent tutoring systems in medical education: a systematic review [J]. *Medical Teacher*, 2022, 44(4): 392–401.
- [22] ZHANG Y, WANG X, LIU C. AI-driven scientific writing simulators for early-career researchers: development and evaluation [J]. *Journal of Medical Internet Research*, 2023, 25: e42851.
- [23] MINH T N, HOANG T Q. AI-powered open science communication platforms: bridging the gap between research and public understanding during COVID-19 [J]. *Health Communication*, 2021, 36(14): 2145–2154.
- [24] International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals (ICMJE Recommendations 2023) [S/OL]. (2023–12–18) [2025–09–02]. <https://www.icmje.org/recommendations/>.
- [25] Nature. AI tools such as ChatGPT must not be listed as authors on research papers [N]. *Nature News*, (2023–01–27) [2025–09–02]. <https://www.nature.com/articles/d41586-023-00107-4>.
- [26] LUND B D, WANG T, MANNURU N R, et al. ChatGPT and a new academic reality: AI-assisted research and the erosion of authorial integrity [J]. *Journal of Academic Librarianship*, 2023, 49(4): 102589.
- [27] JI Z W, LEE N, FRIESKE R, et al. Survey of hallucination in natural language generation [J]. *ACM Comput Surv*, 2023, 55 (12): 1–38.
- [28] HOLDEN T H. ChatGPT is fun, but not an author [J]. *Science*, 2023, 379(6630): 313.
- [29] PRICE W N, COHEN I G. Privacy in the age of medical big data [J]. *Nat Med*, 2019, 25: 37–43.
- [30] VAN DIS E A M, BOLLEN J, ZUIDEMA W, et al. ChatGPT: five priorities for research [J]. *Nature*, 2023, 614 (7947) : 224–226.
- [31] FLORIDI L. The governance of AI: ethical and policy challenges [J]. *Nature Machine Intelligence*, 2023, 5 (8) : 827–829.
- [32] ANDERSON K, CHEN L, ZHANG Y, et al. Transparent reporting of AI assistance in scientific writing: a framework for disclosure [J]. *Research Integrity and Peer Review*, 2023 (8): 12.
- [33] Nature Editorial. AI and the future of writing: transparency is key [J]. *Nature*, 2023, 618(7962): 214–215.

引用本文

白春学, 贾泽军, 杨达伟, 等. 人工智能工具医学使用的科学伦理与规范重构 [J]. 元宇宙医学, 2025, 2(3): 48–55.

BAI C X, JIA Z J, YANG D W, et al. Reconstruction of scientific ethics and norms for the medical use of artificial intelligence tools [J]. *Metaverse Med*, 2025, 2(3): 48–55.

DOI: 10.61189/574152vaftbf

·指南与共识·

人工智能照护中国专家共识



中国肺癌防治联盟,上海呼吸物联网医学工程技术研究中心,国际元宇宙医学
协会赋能基层医院肺癌筛查管理专家组

编者按 《人工智能照护中国专家共识》(以下简称《共识》)的发布,不仅是对AI医疗技术应用的规范指引,更勾勒出“科技-医疗-人文”多方协同的照护新图景——以AI为技术纽带,让医生、护士等照护主体打破领域壁垒,形成1+1>2的服务合力,最终实现“服务病人、造福社会”的核心目标。

从实践维度看,AI的价值在于为医疗照护“补位增效”而非“替代主导”。对医生而言,AI通过影像识别辅助肿瘤早期筛查,用多模态数据优化诊断决策,将医生从繁琐的数据分析中解放,专注于临床判断与医患沟通;对护士来说,智能排班系统、康复机器人与用药提醒工具,让护理工作从“经验驱动”转向“精准适配”,既减轻人力负担,又提升术后康复、老年照护的个性化水平。这种“AI+医护”的协作模式,本质是让技术成为医护人员的“延伸手脚”,而非独立主体,既保留了医疗服务的人文温度,又放大了专业能力的辐射范围。

更为关键的是,《共识》并未局限于技术应用,而是通过“安全性优先”“隐私保护”“公平可及”等原则,为多方协同划定了伦理与法律边界。例如,明确AI开发者、医疗机构、医护人员的责任划分,既保障技术研发的创新空间,又避免“责任真空”损害患者权益;强调数据采集的规范与透明,让患者从“被动数据提供者”变为“知情参与者”,夯实信任基础。这种“技术创新+规则护航”的双轮驱动,让AI照护不仅有“效率提升”的速度,更有“安全可靠”的厚度。

从社会价值层面看,这份《共识》更是推动优质医疗资源下沉的重要抓手。通过AI技术向基层医院、偏远地区延伸,县级医院的医生可借助智能辅助系统提升诊断能力,社区护士能通过可穿戴设备实现慢性病患者的远程监测——这正是“强基层、广覆盖”的生动实践,让不同地域、不同收入群体的患者,都能共享技术进步的红利。

《共识》的意义,远超一份技术规范文件。它搭建了一个“科技赋能医疗、医疗引导科技、伦理守护底线”的协同平台,让医生的专业、护士的细致拧成一股绳,最终指向“以患者为中心”的核心目标。这种多方联动的模式,不仅是AI照护的发展方向,更为未来智慧医疗的可持续发展提供了“中国方案”——唯有让技术、人才、规则形成合力,才能真正让AI照护成为守护全民健康、助力实现“健康中国2030”的重要力量。

【摘要】 随着人口老龄化加速与医疗资源日益紧张,人工智能(AI)技术在医疗照护领域显现出巨大潜力。融合机器学习、自然语言处理、计算机视觉等先进技术,AI在疾病监测、康复辅助、老年照护及心理健康等多个方面开创了创新解决方案。AI照护的精髓在于应用现代信息技术和AI算法,为疾病预防、监测、康复和日常健康管理提供智能支持。其目标是通过减少人力成本、提高医疗资源利用效率,实现患者全天候监测与个性化管理。在此过程中,安全、人性化、隐私保护、公平与可及性,以及透明与可解释性,构成了AI照护的五大基本原则,引领技术向正确方向发展。实际应用中,AI通过可穿戴设备实时监控生理指标,预警健康风险;智能机器人与VR技术为康复患者提供定制化指导;AI聊天机器人则成为心理慰藉的新途径。同时,AI还在优化护理流程、提升行政效率方面发挥重要作用。然而,AI照护的推进也会遇到技术障碍、社会接受度及政策法规等多重挑战,亟需技术创新、伦理指导、政策调整及多方协作的共同助力。在推广应用时,必须建立明确的数据采集与分析标准、人工复核机制、设备安全保障及AI身份清晰标识等规范基础上。在伦理与法律层面,AI照护需划清技术责任与医疗责任界限,确保知情同意的透明度,并为弱势群体提供特殊保护。此外,数据质量、持续学习及多学科合作也是推动AI照护技术进步的关键因素。面对技术局限、社会接受度不足及政策法规滞后等难题,未来AI照护的发展应聚焦于技术创新、伦理教育、政策完善及国际合作等方向。

【关键词】 人工智能;机器学习;自然语言处理;健康;照护;伦理

【中图分类号】 R 473 **【文献标志码】** A

【收稿日期】 2025-09-01

【接受日期】 2025-09-16

【基金项目】 四大慢病重大专项(2024ZD0529300). Supported by Noncommunicable Chronic Diseases National Science and Technology Major Project (2024ZD0529300).

*通信作者(Corresponding author). 白春学,博士,教授,主任医师. Tel: 021-64041990, E-mail: bai.chunxue@zs-hospital.sh.cn

Expert consensus on artificial intelligence care in China

Chinese Alliance Against Lung Cancer; Shanghai Engineer & Technology Research Center of Internet of Things for Respiratory Medicine; The Expert Group for Empowering Lung Cancer Screening Management in Primary Hospitals of the International Association for Metaverse Medicine

[Abstract] With the acceleration of population aging and the increasing shortage of medical resources, artificial intelligence (AI) technology has shown great potential in the field of medical care. Integrating advanced technologies such as machine learning, natural language processing, and computer vision, AI has created innovative solutions in disease monitoring, rehabilitation assistance, elderly care, and mental health. The essence of AI care lies in the application of modern information technology and AI algorithms to provide intelligent support for disease prevention, monitoring, rehabilitation and daily health management. Its goal is to achieve round-the-clock monitoring and personalized management of patients by reducing labor costs and improving the efficiency of medical resource utilization. In this process, security, humanization, privacy protection, fairness and accessibility, as well as transparency and explainability, constitute the five basic principles of AI care, leading the development of technology in the right direction. In practical applications, AI monitors physiological indicators in real time through wearable devices to warn of health risks. Intelligent robots and VR technology provide customized guidance for rehabilitation patients; AI chatbots have become a new way of psychological comfort. At the same time, AI also plays an important role in optimizing care processes and improving administrative efficiency. However, the promotion of AI care will encounter multiple challenges such as technical obstacles, social acceptance, and policies and regulations, and it urgently needs the joint help of technological innovation, ethical guidance, policy adjustment, and multi-party collaboration. When promoting the application, it is necessary to establish clear data collection and analysis standards, manual review mechanisms, equipment security guarantees, and clear identification of AI identities. At the ethical and legal level, AI care needs to draw a clear line between technical and medical responsibility, ensure transparency of informed consent, and provide special protections for vulnerable groups. In addition, data quality, continuous learning, and multidisciplinary collaboration are also key factors driving the advancement of AI care technology. In the face of technical limitations, insufficient social acceptance, and lagging policies and regulations, the future development of AI care should focus on technological innovation, ethics education, policy improvement, and international cooperation.

[Key Words] artificial intelligence; machine learning; natural language processing; healthy; care; ethics

在数字化技术加速变革的时代背景下,《人工智能(AI)照护专家共识》的制定恰逢其时,既是回应医疗健康领域创新发展需求的重要举措,更是应对AI技术在健康照护中应用带来的机遇与挑战的必然选择^[1]。这一由医疗、护理、科技、伦理等多学科领域代表性专家共同参与制定的指导性文件,具有重要的现实意义和深远的历史价值。

当前,AI技术的快速发展正在系统重构医疗服务的方式和内涵。AI算法在疾病预测、辅助诊断、个性化治疗方案制定等方面展现出巨大潜力,成为提升医疗质量与效率的重要推动力。然而,AI技术的复杂性与医疗行为的专业性决定了其应用必须建立在安全性、有效性和伦理性基础之上^[2]。制定《AI照护专家共识》正是为了填补这一领域的发展空白,为AI技术在健康照护中的规范应用提供方向指引。医疗健康领域的AI应用具有鲜明的技术密集型特征,同时又直接关系到人民群众的生命安全与健康权益。医疗、护理、科技、伦理等多学科领域的协同合作是解决这一复杂问题的关键。通过多学科专家的深度对话与共识凝聚,可以建立起涵盖

技术标准、临床应用规范、伦理审查机制等在内的完整体系^[3],确保AI医疗产品的开发和使用既符合技术创新规律,又满足医疗服务的本质要求。

《AI照护专家共识》对助力“健康中国2030”规划具有特殊意义。文件通过明确AI技术在医疗机构中的应用场景和发展路径,推动优质医疗资源下沉和均衡配置。这不仅有助于提升县级医院、社区卫生服务中心等基层机构的服务能力,更是实现“强基层、广覆盖、同质化、高质量”医疗照护目标的重要抓手。制定《AI照护专家共识》是一项具有里程碑意义的开创性工作,它不仅是对现有AI技术发展成果的系统总结,更为未来技术创新和应用拓展指明了方向。通过构建多方价值共识和行动指南,必将推动我国智慧医疗驶入规范、可持续的发展轨道,为实现全体人民享有更高水平的健康保障作出重要贡献。

1 AI照护的定义与范围

随着信息技术的飞速发展和人口老龄化的加剧,传统医疗资源的供需矛盾日益凸显,人们对智

能化、个性化的健康管理需求显著增加。在此背景下, AI技术逐渐成为现代照护领域的重要推手。通过整合机器学习(machine learning, ML)、自然语言处理(natural language processing, NLP)、计算机视觉(computer vision, CV)等先进算法, AI在疾病监测、康复辅助、老年护理和心理健康支持等方面已经展现出巨大潜力^[4]。为促进这一领域的健康发展, 确保技术与临床实践的有效结合, 需要明确AI照护的定义与适用范围。

1.1 AI照护的定义 AI照护是指利用现代信息技术, 结合AI算法, 在疾病预防、监测、康复以及日常健康管理等场景中提供智能化支持的一系列服务和解决方案。其核心技术包括但不限于ML、NLP、CV和知识图谱等。通过这些技术帮助医护人员完成复杂的数据分析和决策辅助工作, 从而提升照护质量和效率。具体而言, AI照护的核心目标在于: (1)为传统照护工作提供智能化支持, 降低人力成本; (2)提高医疗资源的利用效率, 缩小城乡之间、不同收入群体之间的健康差距; (3)实现对患者的全天候监测与个性化管理, 改善患者的治疗体验和康复效果。

1.2 AI照护的适用范围 AI照护的应用场景十分广泛, 涵盖多个领域。(1)疾病监测: 基于穿戴设备(如智能手环、血压计)采集生命体征数据, 利用ML模型进行实时分析, 帮助慢性病患者(如慢阻肺、糖尿病和高血压患者)实现病情的动态监测与预警^[5]。在肺癌等重大疾病的早期筛查中, AI技术可辅助医生快速识别影像中的可疑病变, 提升诊断准确性和效率。(2)康复辅助: 针对术后康复人群或运动损伤患者, 通过智能化康复方案生成系统, 提供个性化的锻炼计划和进度跟踪服务。利用增强现实(augmented reality, AR)技术, 为患者提供沉浸式康复训练体验, 增强其参与度和治疗效果^[6]。(3)老年护理: 针对老年人群体, AI照护可实现跌倒风险监测、用药提醒以及心理健康评估等功能, 帮助他们安全地独立生活^[7]。在认知障碍(如阿尔茨海默病)患者的日常管理中, AI技术能够通过语音或视觉交互提供定向提示和情感支持, 延缓病情进展。(4)心理支持: 针对抑郁症、焦虑症患者, 开发智能化的心理健康评估工具, 结合NLP技术进行情绪分析, 为患者提供适时的心理干预建议^[8]。通过虚拟助手(Chatbot)等形式, 为需要长期心理支持的群体提供全时空的情感陪伴和咨询。(5)健康管理: AI照护还可应用于健康风险评估、饮食与运动规划等领域, 帮助普通人群实现科学化的生活管理。例如, 通过

个人健康数据的分析, 生成个性化的健康计划, 并根据实时反馈进行动态调整^[9]。

1.3 适用对象 AI照护的主要目标群体包括^[10]以下群体。(1)老年人: 随着年龄增长, 老年人的身体机能逐渐衰退, 容易罹患慢性病或出现意外(如跌倒), AI技术可以帮助他们实现实时监测和风险管理, 提高生活质量。(2)慢性病患者: 对于需要长期服药、定期复诊的慢性病群体, AI照护能够提供用药提醒、病情跟踪与健康建议, 减少因管理不当导致的并发症风险。(3)术后康复人群: 通过智能化康复方案和实时反馈机制, 帮助患者加速恢复进程并降低复发概率。(4)残障人士: 在日常生活中的自理能力支持(如起居、饮食)和心理健康维护方面, AI技术能够提供显著的帮助。

AI照护作为一种创新性技术手段, 正在逐步改变传统的医疗照护模式。它不仅能够提升医疗服务的可及性和精准度, 还能为患者提供更优质的生活体验。未来, 随着技术的持续进步和社会接受度的提高, AI照护有望在更多领域得到广泛应用, 成为现代医疗体系中的重要组成部分。

2 AI照护的核心原则

AI技术在医疗照护领域的广泛应用为患者带来了诸多便利, 但同时也伴随着伦理、隐私和技术安全等多重挑战。为了确保AI照护的健康发展, 使其真正成为提升医疗质量、改善患者体验的重要工具, 必须明确并遵循一系列核心原则。这些原则不仅指导技术研发与应用实践, 还能有效规避潜在风险, 为AI照护的可持续发展奠定基础^[11-12]。

2.1 安全性优先 安全性优先内涵: AI系统的安全性是其落地应用的前提条件。AI技术在医疗领域的错误可能导致严重后果, 因此必须确保系统在设计、测试和使用过程中具备高度的安全性和可靠性。安全性优先的原则要求: (1)AI系统需经过严格的临床验证和风险评估, 确保算法的稳定性和准确性。(2)在实际应用中, 应建立多重保障机制(如人工审核和实时监控), 避免因技术故障或误判导致患者受到伤害。(3)安全性问题应在产品设计阶段就被充分考虑, 并贯穿整个生命周期的开发与运营^[13]。

2.2 以人为本 以人为本内涵: AI技术虽然高效, 但其本质上是工具而非替代品。在医疗照护中, 人始终是核心主体, AI的存在是为了增强而非取代人类的专业能力。以人为本的原则要求^[14]: (1)AI系统应作为辅助工具, 为医护人员提供决策支持或效

率提升服务,而不是完全替代人类的判断与情感关怀。(2)在实际应用中,需特别关注患者的情感需求,包括尊重患者的知情权、隐私权以及个性化偏好。(3)严格禁止 AI 技术在不具备充分伦理依据的情况下独立执行关键性医疗操作(如诊断和治疗决策)。

2.3 隐私保护 隐私保护内涵:医疗数据的敏感性和重要性决定了隐私保护是 AI 照护不可忽视的核心原则之一^[15]。通过技术和制度手段,确保患者信息的安全与合规使用。隐私保护的原则要求^[16-17]:(1)在数据采集、存储和传输过程中,严格遵守相关法律法规(如《通用数据保护条例》GDPR、《健康保险流通与责任法案》HIPAA 等)。(2)采用加密技术对患者数据进行处理,并确保数据匿名化存储,避免因技术手段或管理疏忽导致信息泄露。(3)建立透明的数据使用机制,明确告知患者数据的用途和范围,同时为其提供拒绝或撤销授权的权利。

2.4 公平性与可及性 公平性与可及性内涵:AI 照护的目标是提升整体医疗水平,因此必须避免技术鸿沟带来的不平等现象,确保不同群体能够公平受益。公平性与可及性的原则要求:(1)在技术研发和应用中,注意识别并消除潜在的算法偏见,避免因数据样本失衡或模型设计问题导致某些特定群体(如少数族裔、经济弱势群体)受到不公平对待^[18]。(2)推动 AI 技术在医疗资源匮乏地区(如偏远农村或发展中国家)的应用,帮助缩小城乡之间、国与国之间的健康差距^[19]。(3)提供多样化的产品形态和使用方式,降低技术门槛,确保不同年龄、教育背景的患者都能便捷地享受到 AI 带来的便利^[20]。

2.5 透明与可解释性 透明与可解释性内涵:AI 系统的决策过程必须透明化,避免“黑箱”操作对医疗安全和信任关系造成的负面影响。用户(包括患者和医护人员)有权了解 AI 结论背后的逻辑和依据。透明与可解释性的原则要求^[21]:(1)在设计 AI 系统时,优先选择具有较高可解释性的算法,并为用户提供清晰的决策依据说明。(2)对于复杂的 AI 模型,需提供可视化工具或其他辅助手段,帮助使用者理解其工作原理和输出结果。(3)建立明确的反馈渠道,允许用户对 AI 系统的决策提出疑问或建议,并在必要时采取人工干预措施。

AI 照护的核心原则是技术与人文的交汇点,体现了一个“以患者为中心、AI 为抓手、专家为重点”的多方协同发展的医疗生态体系。从安全性优先到公平性与可及性,这些原则不仅是技术研发和应用的基础,更是社会信任的基石。未来,随着技术

的进步和社会认知的提升,我们将能够更好地平衡效率与伦理,在 AI 照护领域实现更大的突破。

3 AI 照护应用场景与规范

AI 技术在医疗照护领域的广泛应用,为疾病监测、康复辅助、心理支持和行政管理等场景带来了创新性解决方案。然而,为了确保这些技术的安全性、有效性及可持续性,必须建立明确的应用场景规范,并对技术的落地实施进行严格把控。为此,需要探讨 AI 在不同照护场景中的应用潜力,并提出相应的规范化目标,以推动 AI 技术与医疗照护的深度融合。

3.1 疾病监测与预警 疾病监测与预警应用场景:AI 可通过分析可穿戴设备(如智能手表、健康手环)采集的数据,实时监测患者的生理指标(如心率、血压、活动量等),并结合 ML 算法预测潜在风险^[22]。例如,在老年照护中,AI 可用于预警跌倒风险或心率异常,从而帮助护理人员及时采取干预措施^[23]。疾病监测与预警的规范化目标:(1)明确数据采集与分析的阈值标准,确保预警准确性;(2)建立人工复核机制,避免因算法误差导致误报或漏报;(3)保障数据隐私安全,防止患者信息泄露^[24];(4)提供可解释性报告,帮助医护人员快速理解 AI 判断依据。

3.2 辅助康复 康复辅助应用场景:AI 在肢体康复训练中展现出广阔前景。例如,智能机器人可通过实时监测患者的运动轨迹和力度,提供个性化指导与反馈^[25]。此外,结合虚拟现实(virtual reality, VR)技术的 AI 系统还可模拟日常生活场景,帮助患者逐步恢复生活能力^[26]。康复辅助规范化目标:(1)根据患者个体情况调整康复方案,避免过度依赖 AI 系统;(2)确保设备安全性和稳定性,防止因技术故障导致意外伤害;(3)建立动态评估机制,定期更新训练计划以适应患者进展;(4)提供多模态交互界面(如语音、视觉反馈),提升用户体验。

3.3 心理与情感支持 心理与情感支持应用场景:AI 聊天机器人可为患者提供心理咨询和情感陪伴服务。例如,在长期住院或孤独症患者的照护中,AI 可通过 NLP 技术模拟人际对话,帮助缓解焦虑、抑郁情绪。心理与情感支持的规范化目标^[8, 27]:(1)在交互界面明确标注“AI 身份”,避免误导用户产生过度依赖;(2)严格限制敏感话题(如自杀倾向)的讨论范围,必要时及时转介专业心理咨询师;(3)建立伦理审查机制,确保算法不传播偏见或歧视性内容;(4)提供多种退出机制,帮助用户随时终止对话并获得人工支持。

3.4 优化行政管理 行政管理优化应用场景:AI技术可显著提升护理流程中的效率问题。例如,在护理排班管理中,AI可通过分析工作量和人员配置,生成最优排班方案;在药品管理中,AI可用于实时监测库存,并智能提醒补货^[28]。行政管理优化的规范化目标^[29]:(1)配备人工审核环节,避免因算法误判导致管理失误;(2)确保数据的安全性和隐私性,防止敏感信息外泄;(3)提供操作日志记录功能,便于追溯和审计;(4)建立风险预警机制,及时发现潜在系统异常或漏洞。

AI在医疗照护领域的应用前景广阔,但其落地实施必须建立在规范化、安全性和伦理性的基础之上。从疾病监测到行政管理优化,AI技术正在改变传统的照护模式,并为患者和医护人员带来更高效、精准的服务体验^[30]。未来,我们需要进一步加强技术研发与规范体系建设的协同创新,推动AI与医疗行业的深度融合,最终实现高质量、可信赖的强基层广覆盖AI照护服务。

4 人工智能照护的伦理与法律要求

AI技术在医疗照护领域的应用带来了诸多便利,但也引发了关于责任划分、知情同意以及弱势群体保护等伦理与法律问题^[31]。为确保AI技术的合理使用,并最大限度地保障患者权益,本共识就相关伦理与法律要求达成如下共识。

4.1 责任划分 责任划分核心问题:AI系统的开发和应用涉及多方主体,包括开发者、医疗机构和照护人员等。若AI系统在运行中出现错误或偏差(如算法误诊、操作失误),如何界定各方的责任成为亟待解决的问题^[32]。责任划分的规范要求^[11]如下。(1)明确技术责任与医疗责任的界限:AI开发者的责任主要集中在技术设计与测试,确保系统算法的安全性及准确性;医疗机构和照护人员则负责合理使用AI工具,并在必要时进行人工干预。(2)建立协同机制:若因AI系统错误导致患者受到损害,应由开发者、医疗机构以及相关责任人共同承担连带责任。同时,需通过合同约定等方式明确各方的权责划分。(3)完善风险防控体系:医疗机构应当配备专门人员对AI系统的运行进行监督,并建立应急预案,以应对可能出现的技术故障或伦理争议。

4.2 知情同意 AI技术的使用往往依赖于患者数据的收集和处理,但许多患者可能并不清楚AI在其治疗过程中的具体角色。如何全面告知患者相关信息,并获得其明确同意,是维护患者自主权的关键^[33]。知情同意的规范要求有以下几方面。(1)

清晰告知AI的参与程度:患者需被告知哪些医疗决策或操作是由AI系统完成的,以及这些决策可能带来的风险和局限性。医疗机构应提供通俗易懂的说明材料,并通过多种方式(如书面、口头)进行解释。(2)数据使用透明化:医疗机构应当向患者详细说明其个人信息及健康数据的具体用途、存储方式以及共享范围,尤其是在涉及第三方服务提供商时,需获得患者的额外授权。(3)建立动态知情同意机制:随着AI技术的不断迭代,医疗机构应定期更新相关信息,并在重要变更时重新获取患者的同意,以确保患者始终处于信息充分知情的状态。

4.3 弱势群体保护 弱势群体保护核心问题:在AI照护系统中,认知障碍患者(如痴呆症患者)由于其特殊性,可能无法完全理解AI的使用方式及其潜在风险^[34]。这种情况下,如何保护他们的合法权益尤为重要。弱势群体保护的规范要求。(1)设置个性化保护措施:对于认知障碍患者,应特别设计易于理解的知情同意文件,并由家属或法定监护人协助完成签署。同时,在AI系统的使用过程中,需定期评估患者的认知状态,并根据其变化调整保护措施。(2)限制数据收集与处理:在保障隐私的前提下,针对认知障碍患者的医疗决策应尽量减少AI的自主性,以避免因算法偏差导致患者权益受损。若有必要使用AI进行辅助决策时,需由专业医护人员进行双重审核。(3)建立投诉与申诉机制:认知障碍患者及其家属若认为其权益受到侵害,应当有明确的途径向相关部门提起投诉,并获得及时的调查和处理。

5 AI照护技术标准与评估

AI技术在医疗照护领域的应用迅速发展,为提升医疗服务质量和效率提供了新的可能性。然而,随之而来的技术要求和评估机制也亟需明确。基于现有研究与实践,需要就AI照护的技术标准与评估原则达成以下共识。

5.1 数据质量 数据质量核心问题:AI系统的性能很大程度上依赖于其训练数据的质量。若数据来源单一或缺乏代表性,可能导致模型在实际应用中出现偏差,进而影响决策的准确性和公平性。数据质量的技术要求^[35]如下。(1)覆盖多样化临床场景:训练数据应涵盖不同年龄段、疾病类型、种族和生活环境等多维度的患者信息,以增强模型的泛化能力。例如,在开发AI辅助诊断系统时,需包括典型病例与非典型病例,以及多种族人群的生理特征数据。(2)高质量标注与清洗:数据标注应由具有医学

背景的专业人员完成,并通过多重检验确保准确性。同时,需对异常值、重复数据及缺失信息进行处理,以减少噪声对模型训练的影响。(3)动态更新机制:随着时间推移,医疗领域的新知识和新技术不断涌现,AI系统应基于最新的临床数据进行迭代优化,避免因数据“老化”导致的决策偏差。

5.2 持续学习 持续学习核心问题:医疗领域的复杂性和动态变化要求 AI 系统具备持续学习的能力。若模型无法及时更新或调整,可能在面对新证据或个体化需求时表现不佳。持续学习的技术要求^[36]有以下几方面。

(1)定期更新机制:AI系统应预设自动化的数据采集、清洗和重组流程,并结合临床实践中的最新研究成果(如随机对照试验、Meta分析等)进行模型再训练。这种动态更新机制不仅提升了系统的准确性,也增强了其在不同临床环境下的适应性。(2)个体化学习模块:在临床应用中,每位患者的生理特征和病程进展均具有独特性。为此,AI系统需具备基于患者特定数据的自适应学习能力,例如根据用药反应、病情变化等因素动态调整诊疗方案建议。这需要结合强化学习等技术,在保护患者隐私的前提下实现个性化优化。(3)可解释性与透明度:在持续学习过程中,模型的决策逻辑应保持清晰可追溯,以便医疗人员理解和验证系统输出的安全性与合理性。例如,通过可视化工具展示模型更新前后的性能变化及潜在风险。

5.3 多学科协作 多学科协作的核心问题:AI照护系统的开发涉及计算机科学、医学、护理、伦理学等多个学科,单一领域的视角难以满足技术标准的全面性与评估的有效性^[37]。因此,组建多学科团队是确保系统安全性和有效性的必要条件。多学科协作的技术要求。

(1)明确职责分工:在项目启动阶段,需明确医生、护士、工程师和伦理学家等不同角色的任务与责任,例如数据标注由医护人员主导,模型开发由技术人员负责,伦理评估则贯穿整个开发流程。这种明确的分工有助于避免因沟通不畅导致的技术偏差或伦理疏漏。(2)跨领域知识整合:在实际应用中,医疗需求往往伴随着复杂的临床背景和伦理考量。因此,技术团队需通过定期研讨会、联合项目等方式,实现医学专业知识与AI算法开发的有效结合。例如,在设计患者风险预测系统时,需结合临床医生的经验与算法工程师的技术优势,确保模型既准确又具备临床适用性。(3)动态评估与反馈机制:多学科团队应对AI系统的性能进行持续跟踪

和评估,并根据实际应用中的反馈结果进行优化调整。这需要建立覆盖开发、测试、部署全生命周期的质量控制体系,并定期邀请患者代表参与意见征集,确保技术成果真正服务于临床需求。

AI在医疗照护领域的应用潜力巨大,但其成功依赖于高质量的技术标准和科学的评估机制。通过优化数据质量、构建持续学习能力以及推动多学科协作,可以有效提升AI系统的安全性、准确性和适应性^[3]。未来,研究者和从业者需共同努力,制定更完善的行业规范,并在实践中不断验证与改进,以实现技术进步与临床需求的最佳结合。

6 AI照护的挑战与未来方向

AI在医疗照护领域的应用正逐步改变传统的医疗服务模式。然而,这一过程中也伴随着诸多挑战与机遇,也需要详细探讨当前AI应用于医疗照护所面临的主要技术瓶颈、社会接受度问题及政策支持需求,并展望未来的发展方向。

6.1 技术局限性

在医疗照护中,患者不仅仅需要专业化的医疗服务,心理和情感支持同样是不可忽视的一部分。AI系统在处理复杂的临床决策时已表现出一定的能力,但在情感识别方面仍存在显著的技术限制。例如,在心理咨询或精神健康服务中,人类咨询师能够通过细腻的语言、面部表情和语调变化来判断患者的情绪状态,并提供相应的支持与建议。然而,现有AI系统在捕捉这些微妙的情感线索方面还不够精准^[38]。尽管NLP技术已取得进展,但在真正理解患者情感需求并给予个性化回应方面仍显不足。此外,复杂决策是另一个重要的技术瓶颈^[39]。医疗照护常常涉及多个变量的综合考量,如患者的病史、当前症状、实验室数据等。AI需要在短时间内分析这些信息并提供有效的治疗建议。然而,在处理高度不确定性或动态变化的临床情境时,现有的算法表现仍然有限。

6.2 社会接受度

尽管AI技术在提高医疗服务效率方面展现了巨大的潜力,但其推广和应用却面临着来自社会层面的抵触情绪。部分人群对AI替代传统人类医疗照护存在疑虑,尤其是在情感支持等需要较高人文关怀的领域^[40]。他们担心AI缺乏同理心和人性化,可能导致患者感到被忽视或不被理解。另一个重要问题是隐私担忧。医疗数据具有高度敏感性,确保这些信息在存储、处理和传输过程中的安全性至关重要。然而,过去几年频发的数据泄露事件使得

公众对AI系统持有较高的警惕态度,担心个人健康信息可能被滥用或不当使用^[41]。这种信任缺失可能会限制患者接受由AI提供的服务。

6.3 政策与法规支持不足

当前,AI医疗照护技术的快速发展超过了相关政策和法规的更新速度。缺乏统一的行业标准导致市场产品良莠不齐,服务质量参差不齐,难以保障患者利益^[42]。此外,大部分国家的医疗保险体系尚未将基于AI的医疗服务纳入报销范围,均限制了这些技术的普及应用。与此同时,现有法律框架在数据隐私、责任归属以及AI决策透明度等方面还存在明显的滞后性。例如,在AI系统出现误诊或错误建议的情况下,如何明确法律责任往往难以界定。这些问题亟需通过政策制定和法律法规完善来解决。

6.4 未来方向

尽管面临上述诸多挑战,但是AI照护系统依然具有广阔的发展前景,亟需克服挑战,推动发展。

(1)技术创新突破。在技术层面,未来需要加大对情感识别与复杂决策能力的研究投入。研究机构可以通过开发更先进的算法模型,结合深度学习与强化学习等技术来提升AI系统对人类情绪的理解和表达能力^[43]。例如,综合使用面部表情分析、语音语调识别以及NLP技术,能够为患者提供更有人情味的支持服务。同时,在复杂决策领域,可以探索融合多模态数据(如患者生理指标、医疗影像、基因数据等)的AI系统,有助于提高诊断和治疗建议的准确性和可靠性。此外,与传统临床知识结合的AI方法,可以让AI在提供个性化医疗解决方案方面更具效率和信心。

(2)伦理指导与公众教育。从社会层面来看,加强伦理指导和公众教育是推动AI照护技术广泛应用的关键^[44]。通过普及AI的基本原理、应用场景及其优势,可以消除人们对这项技术的误解以及信任危机。例如,开展面向大众的科普讲座或制作通俗易懂的宣传资料,能够帮助人们更好地理解AI在医疗领域的应用价值。此外,伦理委员会和相关监管机构需要制定明确的行为准则,确保AI系统的开发和使用遵循道德标准,并尊重患者隐私与权益。只有在技术发展的同时兼顾伦理考虑,才能真正实现科技服务于人的初衷。

(3)政策框架的完善。为了促进AI照护技术的健康发展,各国政府需尽快建立健全相关政策和法规体系^[45]。这包括制定统一的数据处理标准、提升数据安全性保障以及优化医疗服务回馈机制等具

体措施。例如,可以建立严格的安全评估和认证流程,确保所有用于医疗的数据处理平台符合相关隐私保护要求。同时,相关部门应积极推动医疗AI产品的临床试验与审批程序,加快新技术的商业化进程,让更多的患者能够享受到技术进步带来的便利。

(4)推动多领域协作。实现上述目标需要多方力量的协同合作。例如,医疗机构可以与技术企业建立战略合作关系,共同开发适用于具体诊疗场景的AI解决方案;学术界则应加强基础研究,推动算法和技术持续创新^[46];政策制定者需保持开放态度,及时调整法律法规以适应技术创新的需求。此外,在国际合作方面,各国可以通过共享经验、设立联合研发项目等措施,加速技术发展与传播^[47]。在全球化的背景下,共同面临的医疗问题需要全球范围内的解决方案,而AI医疗照护无疑是一个极具潜力的突破口。

人工智能照护中国专家共识

专家组名单(按姓氏拼音排序)

编写组组长

白春学 复旦大学附属中山医院呼吸与危重症医学科,上海市呼吸病研究所,上海呼吸物联网医学工程技术研究中心

张玉侠 复旦大学附属中山医院护理部

编写组成员

董春玲 吉林大学第二医院呼吸与危重症医学科

葛海燕 华东医院呼吸与危重症医学科

兰美娟 浙江大学医学院附属第二医院护理部

李可铖 嘉兴佰乐丰生物科技有限公司上海分公司

梁琼 南宁市第一人民医院

刘泽松 上海宇惯科技有限公司

卢叶婷 上海小度律师事务所

宋元林 复旦大学附属中山医院呼吸与危重症医学科

隋岩峰 维舍尼亚商贸(上海)有限公司

王洵 江南大学附属中心医院

吴学玲 上海交大医学院附属仁济医院

辛弘毅 上海交通大学溥渊未来技术学院

闫亚敏 复旦大学附属中山医院护理部

杨达伟 复旦大学附属中山医院呼吸与危重症医学科

张晓菊 河南省人民医院呼吸与危重症医学科
 赵春娇 南宁市第一人民医院
 赵 飞 长沙市中医医院(长沙市第八医院)呼吸
 内科/肺病科与危重症医学科
 朱唯一 上海交通大学医学院附属瑞金医院护
 理部
 朱文思 复旦大学附属中山医院呼吸与危重症医
 学科

执笔人

白春学 复旦大学附属中山医院呼吸与危重症医
 学科,上海市呼吸病研究所,上海呼吸物联网医学
 工程技术研究中心
 卢叶婷 海小度律师事务所
 宋元林 复旦大学附属中山医院呼吸与危重症医
 学科
 杨达伟 复旦大学附属中山医院呼吸与危重症医
 学科
 张玉侠 复旦大学附属中山医院护理部
 朱文思 复旦大学附属中山医院呼吸与危重症医
 学科

参考文献

- [1] WHO. Global Strategy on Digital Health 2020 - 2025 [R] : WHO Press, 2021.
- [2] ROJAS-GUALDRÓN D. Artificial intelligence in health care: the hope, the hype, the promise, the peril. national academy of medicine. . CES Med, 2022
- [3] LEKADIR K, FRANGI A F, PORRAS A R, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare[J]. BMJ, 2025, 388: e081554.
- [4] FENG G, WENG F L, LU W, et al. Artificial intelligence in chronic disease management for aging populations: a systematic review of machine learning and NLP applications[J]. Int J Gen Med, 2025, 18: 3105-3115.
- [5] RAJESH S. AI-enabled real-time health monitoring for elderly care: a smart solutions approach[J]. Int J Sci Res Eng Trends, 2025, 11(2): 2011-2015.
- [6] LINDROTH H, NALAIE K, RAGHU R, et al. Applied artificial intelligence in healthcare: a review of computer vision technology application in hospital settings [J]. J Imaging, 2024, 10(4): 81.
- [7] IMRAN R, KHAN S S. A systematic review on the efficacy of artificial intelligence in geriatric healthcare: a critical analysis of current literature[J]. BMC Geriatr, 2025, 25(1): 248.
- [8] GRAHAM S, DEPP C, LEE E E, et al. Artificial intelligence for mental health and mental illnesses: an overview[J]. Curr Psychiatry Rep, 2019, 21(11): 116.
- [9] BEKBOLATOVA M, MAYER J, ONG C W, et al. Transformative potential of AI in healthcare: definitions, applications, and navigating the ethical landscape and public perspectives[J]. Healthcare (Basel), 2024, 12(2): 125.
- [10] BORNA S, MANIACI M J, HAIDER C R, et al. Artificial intelligence support for informal patient caregivers: a systematic review[J]. Bioengineering (Basel), 2024, 11(5): 483.
- [11] 中华人民共和国科学技术部. 新一代人工智能伦理规范 [R]. 中华人民共和国科学技术部, 2021.
- [12] ALVAREZ M, WINFIELD A, HAVENS J. AUTONOMOUS INTELLIGENT S. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems [R]. IEEE Standards Association, 2016.
- [13] NING Y L, LIU M X, LIU N. Advancing ethical AI in healthcare through interpretability[J]. Patterns (N Y), 2025, 6(6): 101290.
- [14] TANG L, LI J X, FANTUS S. Medical artificial intelligence ethics: a systematic review of empirical studies [J]. Digit Health, 2023, 9: 20552076231186064.
- [15] MEHL G, TUNÇALP Ö, RATANAPRAYUL N, et al. WHO SMART guidelines: optimising country-level use of guideline recommendations in the digital age [J]. Lancet Digit Health, 2021, 3(4): e213-e216.
- [16] WILLIAMSON S M, PRYBUTOK V. Balancing privacy and progress: a review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare[J]. Appl Sci, 2024, 14(2): 675.
- [17] ESMAEILZADEH P. Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: a perspective for healthcare organizations [J]. Artif Intell Med, 2024, 151: 102861.
- [18] CHINTA S V, WANG Z C, PALIKHE A, et al. AI-driven healthcare: fairness in AI healthcare: a survey[J]. PLoS Digit Health, 2025, 4(5): e0000864.
- [19] CIECIERSKI-HOLMES T, SINGH R, AXT M, et al. Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review[J]. NPJ Digit Med, 2022, 5(1): 162.
- [20] WEINREB R N, LEE A Y, BAXTER S L, et al. Application of artificial intelligence to deliver healthcare from the eye [J]. JAMA Ophthalmol, 2025, 143(6): 529-535.
- [21] LOH H W, OOI C P, SEONI S, et al. Application of explainable artificial intelligence for healthcare: a systematic review of the last decade (2011-2022) [J]. Comput Methods Programs Biomed, 2022, 226: 107161.
- [22] SABRY F, ELTARAS T, LABDA W, et al. Machine learning for healthcare wearable devices: the big picture[J]. J Healthc Eng, 2022, 2022: 4653923.
- [23] PARAMASIVAM A, FERLIN DEVA SHAHILA D, JENATH M, et al. Development of artificial intelligence edge computing based wearable device for fall detection and prevention of elderly people[J]. Heliyon, 2024, 10(8): e28688.
- [24] RADANLIEV P. Privacy, ethics, transparency, and

- accountability in AI systems for wearable devices [J]. *Front Digit Health*, 2025, 7: 1431246.
- [25] NICORA G, PE S, SANTANGELO G, et al. Systematic review of AI/ML applications in multi-domain robotic rehabilitation: trends, gaps, and future directions [J]. *J Neuroeng Rehabil*, 2025, 22(1): 79.
- [26] ZANATTA F, GIARDINI A, PIEROBON A, et al. A systematic review on the usability of robotic and virtual reality devices in neuromotor rehabilitation: patients' and healthcare professionals' perspective [J]. *BMC Health Serv Res*, 2022, 22(1): 523.
- [27] LEE E E, TOROUS J, DE CHOUDHURY M, et al. Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom [J]. *Biol Psychiatry Cogn Neurosci Neuroimaging*, 2021, 6(9): 856–864.
- [28] WEI Q Y, PAN S C, LIU X Y, et al. The integration of AI in nursing: addressing current applications, challenges, and future directions [J]. *Front Med (Lausanne)*, 2025, 12: 1545420.
- [29] 国家药品监督管理局医疗器械技术审评中心. 人工智能医用软件产品分类界定指导原则 [R]. 国家药品监督管理局, 2021.
- [30] KARRAR H R, NOUH M I, KUFYAH M T, et al. The role of artificial intelligence in advancing hospital management information systems: review article [J]. *Saudi J Clin Pharm*, 2025, 4(1): 15–23.
- [31] NASIR M, SIDDIQUI K, AHMED S. Ethical–legal implications of AI-powered healthcare in critical perspective [J]. *Front Artif Intell*, 2025, 8: 1619463.
- [32] SOLAIMAN B. Legal and ethical considerations of artificial intelligence for residents in post-acute and long-term care [J]. *J Am Med Dir Assoc*, 2024, 25(9): 105105.
- [33] PARK H J. Patient perspectives on informed consent for medical AI: a web-based experiment [J]. *Digit Health*, 2024, 10: 20552076241247938.
- [34] WANGMO T, LIPPS M, KRESSIG R W, et al. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders [J]. *BMC Med Ethics*, 2019, 20(1): 98.
- [35] DE HOND A A H, LEEUWENBERG A M, HOOFT L, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review [J]. *NPJ Digit Med*, 2022, 5(1): 2.
- [36] VOKINGER K N, FEUERRIEGEL S, KESSELHEIM A S. Continual learning in medical devices: FDA's action plan and beyond [J]. *Lancet Digit Health*, 2021, 3(6): e337–e338.
- [37] GAZQUEZ-GARCIA J, SÁNCHEZ-BOCANEGRA C L, SEVILLANO J L. AI in the health sector: systematic review of key skills for future health professionals [J]. *JMIR Med Educ*, 2025, 11: e58161.
- [38] LIMA E S, PERICO C P, NICHIO B T L, et al. Emotional recognition technologies applied to health: Review and challenges [J]. *Braz J Psychiatry*, 2025.
- [39] CHENG S W, CHANG C W, CHANG W J, et al. The now and future of ChatGPT and GPT in psychiatry [J]. *Psychiatry Clin Neurosci*, 2023, 77(11): 592–596.
- [40] WITKOWSKI K, DOUGHERTY R B, NEELY S R. Public perceptions of artificial intelligence in healthcare: ethical concerns and opportunities for patient-centered care [J]. *BMC Med Ethics*, 2024, 25(1): 74.
- [41] KAUTTONEN J, ROUSI R, ALAMÄKI A. Trust and acceptance challenges in the adoption of AI applications in health care: quantitative survey analysis [J]. *J Med Internet Res*, 2025, 27: e65567.
- [42] PHAM T. Ethical and legal considerations in healthcare AI: innovation and policy for safe and fair use [J]. *R Soc Open Sci*, 2025, 12(5): 241873.
- [43] LIAN H L, LU C, LI S N, et al. A survey of deep learning-based multimodal emotion recognition: speech, text, and face [J]. *Entropy (Basel)*, 2023, 25(10): 1440.
- [44] GOKTAS P, GRZYBOWSKI A. Shaping the future of healthcare: ethical clinical challenges and pathways to trustworthy AI [J]. *J Clin Med*, 2025, 14(5): 1605.
- [45] 胡梦瑶, 袁 飞. 人工智能临床应用的法律规制 [J]. *中华医学杂志*, 2023, 103(18): 1363–6.
- [46] GISSELBAEK M, BERGER-ESTILITA J, DEVOS A, et al. Bridging the gap between scientists and clinicians: addressing collaboration challenges in clinical AI integration [J]. *BMC Anesthesiol*, 2025, 25(1): 269.
- [47] KLINGELHÖFER D, BRAUN M, DRÖGE J, et al. Research on artificial intelligence, machine and deep learning in medicine: global characteristics, readiness, and equity [J]. *Global Health*, 2025, 21(1): 36.

引用本文

中国肺癌防治联盟, 上海呼吸物联网医学工程技术研究中心, 国际元宇宙医学协会赋能基层医院肺癌筛查管理专家组. 人工智能照护中国专家共识 [J]. *元宇宙医学*, 2025, 2(3): 56–64.

Chinese Alliance Against Lung Cancer; Shanghai Engineer & Technology Research Center of Internet of Things for Respiratory Medicine; The Expert Group for Empowering Lung Cancer Screening Management in Primary Hospitals of the International Association for Metaverse Medicine. Expert consensus on artificial intelligence care in China [J]. *Metaverse Med*, 2025, 2(3): 56–64.

复旦大学附属中山医院简介

复旦大学附属中山医院（以下简称“中山医院”）始建于1937年，为纪念中国民主革命的先驱孙中山先生而命名，是中国人创建和管理的最早的大型综合性医院之一。中山医院是国家卫生健康委预算管理单位，国家发展改革委首批综合类国家医学中心建设单位，国家卫生健康委公立医院高质量发展试点单位。中山医院是复旦大学附属综合性教学医院，是教育部、国家卫生健康委与上海市人民政府共建共管医院，也是上海市第一批三级甲等医院。在全国三级公立医院绩效考核中，中山医院始终名列前茅，连续多年获得最高评级“A++”。中山医院连续多年在上海申康医院发展中心市级医院院长绩效考核中，位列上海市综合类医院第一名。

使命

以病人为中心，致力于提供优质、安全、便捷的医疗服务；
通过医疗、教育、科研和管理创新，促进医学事业的发展，提升民众的健康福祉。

愿景

世界一流的创新型、智慧型现代化医院。

核心价值观

严谨、求实、团结、奉献、创新、关爱。

