

欧盟医疗器械法规对人工智能医疗器械产品的监管分析

王雅文, 刘英慧*, 郭兆君, 梁宏, 盛恒松

国家药品监督管理局医疗器械技术审评中心 北京

【摘要】基于人工智能和机器学习 (Artificial Intelligence and Machine Learning, AI/ML) 医疗器械近年来发展迅速。然而, 我们对基于 AI/ML 的医疗器械是如何在欧洲获得批准的情况知之甚少。新的欧洲医疗器械法规 (European Medical Device Regulation, EU MDR) 的发布实施日期为 2021 年 5 月 26 日, 该法规提出了在人工智能软件产品应用于临床实践之前需要满足的相关要求。欧盟在该领域的发展尚处于起步阶段, 自 2015 年以来, 欧盟有 80 多种 AI 产品上市, 临床试验和同品种临床数据分析是主要上市途径, 本文从欧盟法规对人工智能医疗器械产品上市的相关要求进行阐述, 总结前沿监管思路与方法, 以期对我国人工智能产品的监督与管理提供参考。

【关键词】人工智能; 机器学习; 医疗器械

【基金项目】国家科技部科技创新 2030—“新一代人工智能”重大项目 (2020AAA0105002)

Regulatory analysis of EU medical device regulations on artificial intelligence medical device products

Yawen Wang, Yinghui Liu*, Zhaojun Guo, Hong Liang, Hengsong Sheng

Center For Medical Device Evaluation. NMPA, Beijing, China

【Abstract】 Medical devices based on artificial intelligence and machine learning (AI/ML) have developed rapidly in recent years. However, we know little about how AI/ML based medical devices are approved in Europe. The new European medical device regulation (EU MDR) is issued and implemented on May 26, 2021. The regulation puts forward the relevant requirements to be met before the application of artificial intelligence software products in clinical practice. The development of EU in this field is still in its infancy. Since 2015, more than 80 AI products have been launched in the EU. Clinical trials and clinical data analysis of the same variety are the main ways to launch AI products. This paper expounds the relevant requirements of EU regulations on the launch of AI medical devices, and summarizes the frontier regulatory ideas and methods, In order to provide reference for the supervision and management of artificial intelligence products in China.

【Keywords】 Artificial intelligence; Machine learning; Medical devices

引言

在医学领域中, 基于人工智能和机器学习 (Artificial Intelligence and Machine Learning, AI/ML) 技术推动着医疗向精准化方向发展, 从诊断、监护、治疗到康复, 各细分领域都将开启智能化时代^[1,2]。利用人工智能技术进行影像三维分割、病理图像分析处理、疾病早期筛查、药物设计、个性化精准医疗等方面发挥着重要作用^[3]。AI/ML 的发展很大程度上依赖于软件开发, 在具体应用中已被证明通过优化技术, 利用软件学习能力, 可以有效的提高图

像质量、降低辐射剂量、识别病理位置、创建和优化诊疗协议、准确分割病理区域和器官。医疗器械制造商也越来越多地应用该类技术更好的为医疗保健机构或医疗机构提供服务并改善患者的治疗效果^[4]。

自 2021 年 5 月 26 日起, 欧盟医疗器械法规 (European Medical Device Regulation, EU MDR) 取代欧盟医疗器械指令, 该法规对医疗器械审批上市提出了更为严格的要求, 其中包括人工智能医疗器械可用于临床使用之前需要满足的监管需求, 使欧

*通讯作者: 刘英慧

盟成为促进人工智能安全及可信的全球领导者^[5]。该法规涉及人工智能医疗器械风险类别的界定、支持措施、数据质量及上市后安全控制等多个方面的内容^[6]。本文旨在通过对欧盟人工智能医疗器械监管制度的分析, 总结监管思路与方法, 以期为我国行政管理部门和产品制造商对人工智能医疗器械产品的监督与管理提供参考。

1 欧盟人工智能医疗器械软件定义

医疗器械软件 (Medical device software, MDSW) 即在医疗器械预期使用过程中列出的一个或多个医疗目的提供信息处理的软件, 例如诊断、预防、监测、质量或缓解疾病, 人工智能医疗器械软件适用于可自行提供诊断或治疗的软件, 也适用于仅提供旨在告知医护人员做出最终诊断或治疗决定的信息的软件^[7]。

无论此类软件是可执行程序、交互式软件、网络服务还是对电子表格中数据的简单处理, 包括某些产品虽不符合医疗器械定义但驱动或影响医疗器械硬件使用的软件也受到 EU MDR 的管理。

除了 EU MDR 法规外, 同时也有以下适用于医疗器械软件的相关欧盟法规或指导文件。

2 欧盟 AI/ML 医疗器械的监管要求

符合医疗器械定义的软件同时必须符合 EU MDR 中列出的专用要求, 称为通用安全和性能要求 (General Safety and Performance Requirements, GSPRs)^[8]。该产品需在验证合格后通过 CE 认证投放到欧洲市场进行使用, 软件产品在特定条件下 (医疗机构判断市售产品无法满足目标患者群体的特定需求等) 不需要进行 CE 认证即可在医疗器械机构内部进行开发、使用和维护^[9]。法规中对高风险人工智能医疗器械制定了详细的评价原则, 首先训练人工智能的数据应符合要求, 对于个人隐私和数据保护需符合《通用数据保护条例》和《执法指令》中的要求。其次应确保技术鲁棒性, 在医疗器械整个生命周期中技术应可靠、结果应精确, 如给出诊断结果, 结果是可复制的^[10]。高风险产品需通过机构评估才具有上市的可能性, 法规中要求对公告机构进行评估和监督, 并对公告机构、评估机构的工作流程等进行了详细规定。图 1 列出了欧盟医疗器械软件监管路线。

2.1 人工智能医疗器械分类

在欧盟医疗器械法规体系中, I 类医疗器械的风险最低, 制造商可以通过自我声明符合法规的方式将其投放市场进行使用。如果使用过程中可能发生死亡或不可逆转的伤害, 将依据不同的风险等级分为 IIa 级、IIb 级, III 级。对于 IIa 类及更高级别, 需要通过指定机构共同审查产品的符合性和评估产品是否满足 EU MDR 要求的符合性, 并提供符合性证书和质量管理体系证书^[11]。基于对上市产品的考虑, 随着器械风险等级增加, 获得监管批准的途径和要求变得更加严格。医疗器械的预期用途, 对其是否作为医疗器械及进行风险类别管理等重要的问题起决定因素。判定预期用途所带来的风险包括两个方面: (1) 实际的医疗目的, 如是否可以对疾病或损伤进行诊断、治疗或监测。(2) 授权使用, 在特定患者人群和使用环境下需确认法规是否适用并, 将产品按照风险程度进行分类; 还应根据临床评价和器械技术文档确定最终用途。软件被视为有源设备, 因此也要考虑与有源设备相关的其他适用的规则。

2.2 上市产品情况

在带有具备 CE 标志的 AI/ML 医疗设备的风等级 I 与制造商的原产国 (以色列和英国) 以及风险等级和公司规模具有很强的关联性。在欧洲, 制造商可以自由选择他们的医疗器械在哪个欧洲国家/地区进行 CE 认证——即不必在制造商所在的国家进行认证。

在欧洲, 目前有 80 个 AI/ML 医疗器械产品获得 CE 认证, 其中产品种类主要集中在放射学领域, 通过对该产品上市后的特殊性进行分析, 可以发现通过收集放射学检查所形成的数据建立数据库难度较低, 这些数据库在 AI 学习中可以进行高效利用, 并给临床专家提供了非常有用的资源。目前, 在欧盟已上市的基于 AI/ML 的产品中归类为风险等级 III 的仅有两种。

2.3 临床评价

在欧盟, 医疗器械在上市前均需通过临床评价方式验证医疗器械在按照制造商所设定的预期用途使用时的安全性和有效性。

临床评价是通过医疗器械在使用过程中产生的临床数据进行分析, 临床评价的方式可以通过相关器械的临床使用数据的收集和分析, 或从科学

文献中对上市后产品临床使用数据进行分析, 特别是通过上市后临床病例的随访来进一步证明产品的安全性和有效性^[12]。在使用同品种器械数据时需先确定与同品种产品具有等同性, 包括非临床研究中对产品性能、生物学等进行详细比较, 同时也需要提供同品种器械相关技术文件^[13]。只有按照 EU MDR 的标准证明与受试器械等效时, 才能使用可比器械的数据, 该标准要求对临床、技术和生物学 (如适用) 特征进行详细比较。

临床数据要求具有足够的数量 (涵盖所有适应症) 和质量 (数据集的现状和适当的统计方法), 才能进行全面评估该器械是否安全并实现预期的临床用途。这也被称为充分的临床证据。一般临床评价过程见图 2。

首先是确定临床评价范围及与医疗目的相关的

临床预期用途, 临床评价计划中 (CEP) 包含了这一点, 它确定了支持临床需求、评估剩余风险和证明预期收益所需的数据类型和数量。制造商必须证明他们所提供的临床证据是充分的和合理的, 包括所适用标准和指导文件、以及与设备管理的医疗状况相关的信息、基准设备以及目标人群可用的医疗替代品。

当识别患者相关数据 (即, 在临床评估计划阶段告知足够的临床证据支持预期的临床目的) 时, 软件制造商必须考虑安全有效的三个方面内容: 与临床需求的关联性, 技术性能, 临床表现 (图 2)。器械对患者和用户构成的风险、临床预期用途、制造商的声明和现有临床数据的可用性 (三类器械及植入类器械需进行临床验证) 决定其是否需要通过临床试验研究作为临床性能验证的一部分。

表 1 适用于医疗器械软件的相关欧盟法规概述

法规 (EU) 2017/745	MDR: 2021 年 5 月 26 日实施
法规 (EU) 207/2012	医疗器械使用电子说明书规定
MEDDEV2.7/1 (修订版 4)	医疗器械临床评价指南
MDCG2019-1	医疗资格和分类指南
MDCG2019-16	医疗器械网络安全指南
数据保护条例 2016/679	通用数据保护条例 (GDPR)

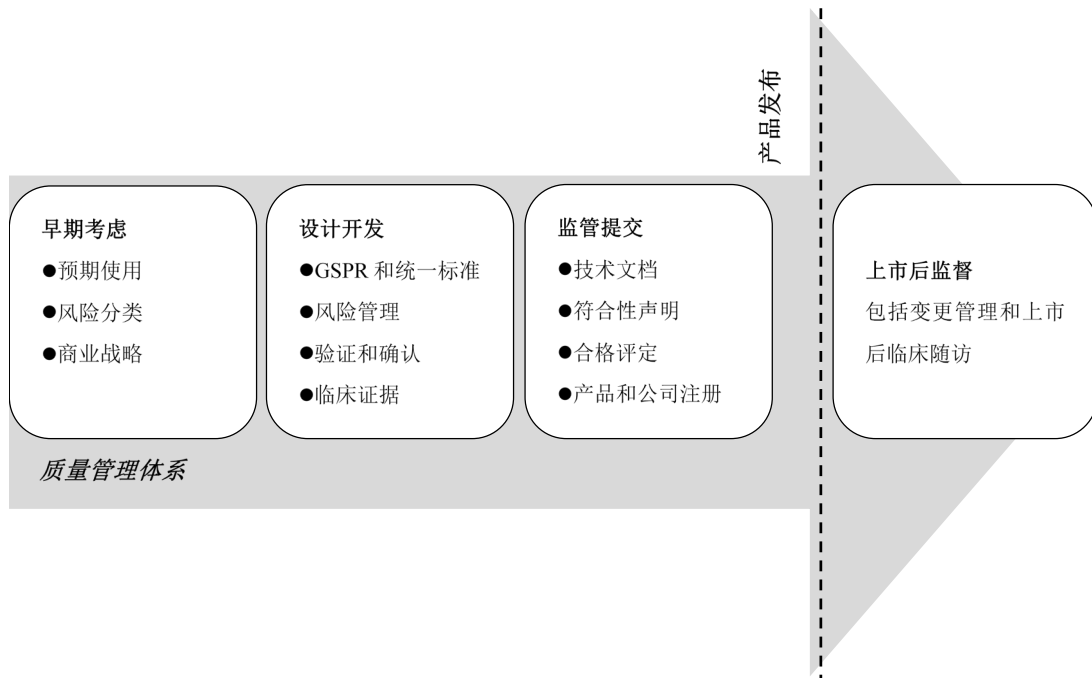
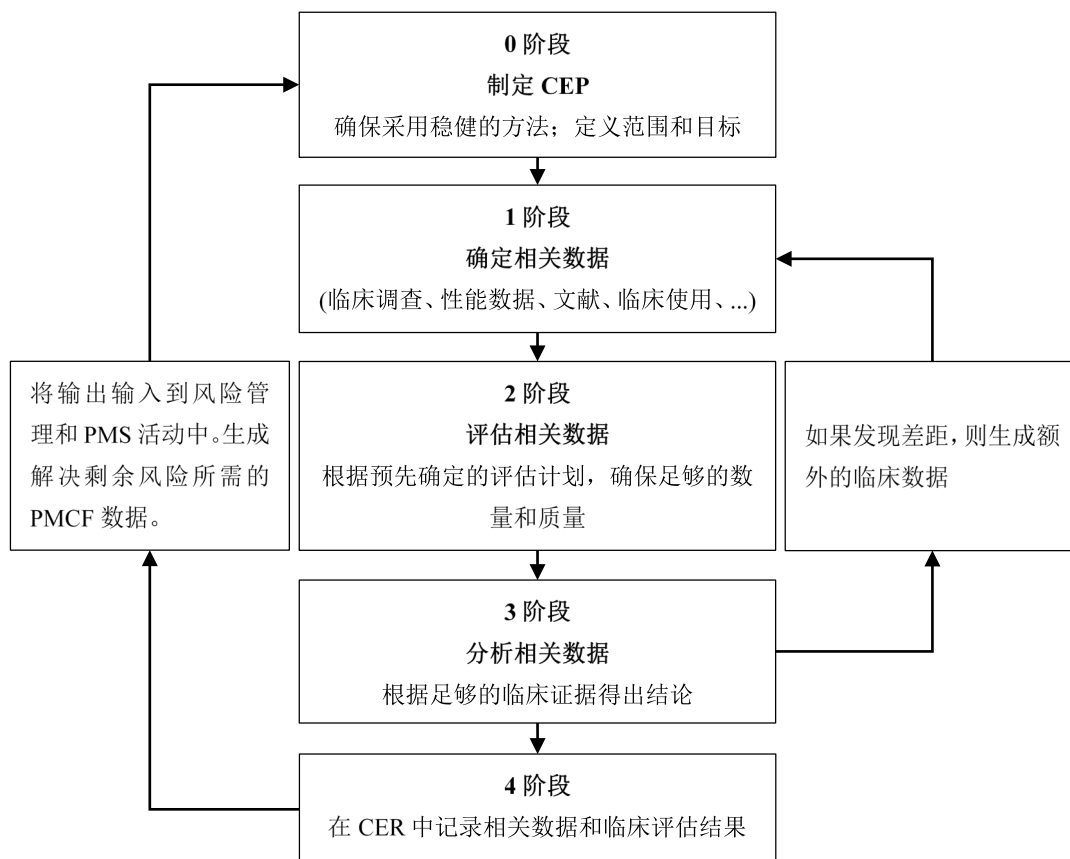


图 1 医疗设备软件的监管路线图

表 2 MDCG 2019-11 医疗器械(MDR)和体外诊断器械(IVDR)法规中的软件定性和分类指南

医疗保健状况或患者状况	欧盟医疗器械软件(MDSW)提供的信息对医疗保健决策的重要性		
	高	中	低
	治疗或诊断	推动临床管理	通知临床管理
危急情况或病人情况	III级	II b 级	II a 级
严重情况或病人情况	II b 级	II a 级	II a 级
不严重情况或病人情况	II a 级	II a 级	II a 级



有效的临床关联: 支持 MDSW 输出与目标疾病之间有效临床关联的证据;

技术性能验证: 证明 MDSW 能够准确、可靠和精确地生成预期输出，形成输入数据;

临床性能验证: MDSW 能够根据预期目的产生临床相关输出的证据

图 2 MDSW 临床证据的组成部分^[31]

临床评价中特别需要关注的是新技术对临床诊疗和患者治疗效果所收集的数据量是非常有限的，因此需要考虑在临床数据研究设计中的受试人群的权力、安全、尊严和福祉，以及生成的临床数据是否具有科学性、有效性和可靠性。依据 ISO 14155 《医疗器械临床研究良好临床实践标准》要求，临床研究过程必须接受科学和伦理审查。

AI/ML 的准确度往往依赖于大数据的质量。因此，在数据评估过程中，评估人员应当收集具有统

计意义的数据，以及按照相关标准和医疗管理系统中生成的数据（即足够数量和质量的数据）^[14]。在分析阶段，评估人员进行全面评估，以确定数据是否符合与设备相关的临床安全性和性能要求。临床评估报告（CER）记录了评估结果、所有输入和每个临床评估阶段，CER 作为欧洲医疗器械强制性技术文件提供给欧盟监管机构。

2.4 上市后监测

上市后监测（Post-Market Surveillance, PMS）

旨在通过观察和分析日常实际使用情况,对器械全生命周期的风险受益进行持续监测以便发现以前未知的风险,不断提高器械的质量。维护 AI/ML 的软件系统不仅是满足监管要求的必要条件,也是改善风险管理和提高设备质量的必要条件。EU MDR 要求制造商积极主动收集器械在真实诊疗环境下的使用数据,并在合理期限内保留部分记录和数据,共监管部门进行查询,记录包括:器械使用记录、如何选取训练数据与验证数据的记录,必要时还应保留原始记录。

同时,制造商必须准确记录所有不良事件。建立不良事件的收集、评估和反馈机制,并在必要时启动警戒程序。出现意外事件发生时及时通过 EUDAMED 数据库向主管当局报告事故,包括相关的纠正和预防措施。

由于人工智能具有不断学习的特性,制造商通常对其可采用性、临床实用性的验证有限,上市后可能会与上市前设定的临床预期用途产生一定偏差^[15]。因此,需要根据临床评价的结果,进行上市后临床随访 (PMCF),以确保具有充分的数据验证来支持产品实际使用的预期用途。从本质上讲, MDSW 收集了许多真实世界产生的数据,这些数据可以分析并用于多种目的,包括但不限于对 MDSW 的监测和纠正、检测系统性误用、了解用户交互、监控临床性能和提高软件效率过程中收集的相关数据。

当制造商变更经 CE 认证的医疗器械(例如,预期用途、产品声明或设计)或经批准的质量管理体系时,他们必须向通告机构提交更改计划。评估机构将对拟变更需求进行综合分析,并决定是否需要进行新的评定,或是否向欧盟提供相关评价资料来进行变更。

3 讨论和结论

人工智能技术与医疗的发展密不可分,在快速融合的过程中,对使用者的隐私及安全等方面的影响已经引起了社会各界的高度关注。各国都加快了对人工智能产品的立法约束。欧盟率先在人工智能医疗器械立法,为其他国家的立法探索提供了一定的参考。我国在 2019 年成立了人工智能医疗器械标准化技术委员会,之后陆续发布了《人工智能医疗器械注册审查指导原则》(征求意见稿)等系列参考性文件,在逐步发展的过程中不断完善适应于我

国人工智能发展的相关立法文件和标准。

医疗设备制造商越来越多地应用人工智能 (AI) 来创新其产品并改善患者治疗效果。尽管基于人工智能的算法为改善患者预后提供了良好的前景,但它们在临床实践中仍存在安全性和有效性的问题从而引入新的风险。为避免过重监管阻碍创新,欧盟引入了“领域”与“用途”相结合的判定标准对人工智能类医疗器械进行风险级别的划分,我国在此过程中也发布了《人工智能医用软件产品分类界定指导原则》,充分注重安全与发展的平衡。

人工智能产品因自身的特殊性,通过深度学习不断完善、更新和迭代。最直接的是训练数据集发生变化,数据集的质量是保证产品性能和安全的基础^[16]。欧盟要求在临床工作流程中引入商用人工智能工具时,不允许根据临床需要改变其预期用途。内部开发的人工智能软件监管时,要求开发人员对监管政策有透彻的了解,临床机构有条件将研究转化为临床实践。国际医疗器械监管机构论坛 (IMDRF) 中临床评价部分是我国参与该组织后完成的重要研究之一,我国应快速的在该领域中开辟新的临床评价工具,促进人工智能医疗器械质量的提升,引导产业健康发展。

总之,让研发人员和制造商了解欧盟医疗器械法规如何影响基于人工智能的医疗设备软件的开发,对人工智能的医疗设备软件的开发有着重要作用;同时对监管人员对此类产品的注册审批及上市后监管工作有着重要的指导意义。

参考文献

- [1] Muchlematter U J, Daniore P, Vokinger K N. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): a comparative analysis[J]. Lancet Digit Health, 2021, 3(3): e195-e203.
- [2] Zechendorf E, Vassen P, Zhang J, et al. Heparan Sulfate induces Necroptosis in Murine Cardiomyocytes: A Medical-In silico Approach Combining In vitro Experiments and Machine Learning[J]. Front Immunol, 2018, 9: 393.
- [3] Barragan-Montero A, Javaid U, Valdes G, et al. Artificial intelligence and machine learning for medical imaging: A technology review[J]. Phys Med, 2021, 83: 242-256.

- [4] Rahman R, Kodesh A, Levine S Z, et al. Identification of newborns at risk for autism using electronic medical records and machine learning[J]. *Eur Psychiatry*,2020,63(1):e22.
- [5] Pucchio A, Eisenhauer E A, Moraes F Y. Medical students need artificial intelligence and machine learning training[J]. *Nat Biotechnol*,2021,39(3):388-389.
- [6] Watson J, Hutyrá C A, Clancy S M, et al. Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: what can we learn from US academic medical centers?[J]. *JAMA Open*,2020,3(2):167-172.
- [7] Beckers R, Kwade Z, Zanca F. The EU medical device regulation: Implications for artificial intelligence-based medical device software in medical physics[J]. *Phys Med*,2021,83:1-8.
- [8] Gola M, Capasso L, Mondoni G, et al. From building regulations and local health rules to the new local building codes: a national survey in Italy on the prescriptive and performance requirements for a new performance approach[J]. *Ann Ig*,2020,32(5 Suppl 1):36-51.
- [9] Alwaysheh A, Wilcke J, Elvinger F, et al. Review of Medical Decision Support and Machine-Learning Methods [J]. *Vet Pathol*,2019,56(4):512-525.
- [10] Lo C M, Jack L Y. The use of multimedia medical data and machine learning for various diagnoses[J]. *Comput Methods Programs Biomed*,2018,165:A1.
- [11] Fischer I, Steiger H J. Toward automatic evaluation of medical abstracts: The current value of sentiment analysis and machine learning for classification of the importance of PubMed abstracts of randomized trials for stroke[J]. *J Stroke Cerebrovasc Dis*,2020,29(9):105042.
- [12] Png M E, Dritsaki M, Gray A, et al. Economic evaluation on plan of a randomised controlled trial of intra-nodular injection of anti-TNF and placebo among patients with early Dupuytren's disease: Repurposing Anti-TNF for Treating Dupuytren's Disease (RIDDD)[J]. *Wellcome Open Res*,2018,3:156.
- [13] Sheppard L, Dewey H, Bernhardt J, et al. Economic Evaluation Plan (EEP) for A Very Early Rehabilitation Trial (AVERT): An international trial to compare the costs and cost-effectiveness of commencing out of bed standing and walking training (very early mobilization) within 24 h of stroke onset with usual stroke unit care[J]. *Int J Stroke*,2016,11(4):492-494.
- [14] Ralbovsky N M, Lednev I K. Towards development of a novel universal medical diagnostic method: Raman spectroscopy and machine learning[J]. *Chem Soc Rev*,2020,49(20):7428-7453.
- [15] Borakati A. Evaluation of an international medical E-learning course with natural language processing and machine learning[J]. *BMC Med Educ*,2021,21(1):181.
- [16] Monlezun D J, Carr C, Niu T, et al. Meta-analysis and machine learning-augmented mixed effects cohort analysis of improved diets among 5,847 medical trainees, providers, and patients[J]. *Public Health Nutr*,2021:1-27.
- 收稿日期:** 2022年8月12日
出刊日期: 2022年10月11日
引用本文: 王雅文, 刘英慧, 郭兆君, 梁宏, 盛恒松, 欧盟医疗器械法规对人工智能医疗器械产品的监管分析[J], 国际医学与数据杂志 2022, 6(5): 9-14. DOI: 10.12208/j.ijmd.20220205
- 检索信息:** RCCSE 权威核心学术期刊数据库、中国知网 (CNKI Scholar)、万方数据 (WANFANG DATA)、Google Scholar 等数据库收录期刊
- 版权声明:** ©2022 作者与开放获取期刊研究中心 (OAJRC) 所有。本文章按照知识共享署名许可条款发表。 <https://creativecommons.org/licenses/by/4.0/>



OPEN ACCESS