

脑卒中大型医学数据库应用及研究进展

陈亦豪 常健博 魏俊吉 冯铭 王任直

【摘要】 随着全球人口老龄化的加剧,脑卒中已成为 60 岁以上人群中仅次于恶性肿瘤的第二大死亡原因。大型医学数据库通过系统集成医学数据,提供流行病学资料,从研究假设转化为指导临床实践的循证医学证据,广泛应用于临床医学研究。本文重点介绍世界范围内不同脑卒中数据库的特点,阐述其在脑卒中中的应用及研究进展,以及人工智能技术在数据库建设中的应用,以期促进脑卒中数据库的建设并开展高质量的脑卒中临床研究。

【关键词】 卒中; 数据库; 医学信息学应用; 人工智能; 综述

Update on application and research progress of large-scale medical database in stroke

CHEN Yi-hao, CHANG Jian-bo, WEI Jun-ji, FENG Ming, WANG Ren-zhi

Department of Neurosurgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

Corresponding author: WANG Ren-zhi (Email: wangrz@126.com)

【Abstract】 With rapid aging worldwide, stroke disease has become the second leading cause of death in people over 60 years old, following the cancer disease. The large-scale medical database provides epidemiological data by systematically integrating medical data and transforming research hypotheses into evidence-based medicine evidence to guide clinical practice, widely used in clinical medical research. This article describes different stroke databases to explain their application and research progress in stroke diseases, and describes the application of artificial intelligence (AI) in database construction, aiming to promote the construction of stroke databases and conduct high-quality stroke research.

【Key words】 Stroke; Database; Medical informatics applications; Artificial intelligence; Review

Conflicts of interest: none declared

随着全球人口老龄化的加剧,脑卒中已成为仅次于恶性肿瘤的第二大死亡原因,也是造成残疾的主要原因^[1]。中国目前有超过 1000 万例脑卒中患者,每年新发病例约 250 万例^[2]。脑卒中可造成永久性神经损伤,产生严重并发症甚至导致死亡,影响患者生活质量,并对患者家庭及社会带来较大经济负担。中国脑卒中发病率存在地区差异^[3],加之各地区人群存在文化习俗、饮食、生活习惯上的不同,使得脑卒中中的一级预防以及个体化和精准治疗的开展存在一定困难,亦无法为脑卒中的临床治疗提供循证医学支持,故亟待建立大型地方或国家级

医学数据库以提供和实现数据调研、存储、挖掘、共享。20 世纪 90 年代以来,西方发达国家陆续创建以疾病诊疗和质控为核心的大型医学数据库和脑卒中网络登记中心,并以此为基础开展大量的临床研究^[4-6]。2007 年,中国国家卒中登记(CNSR)数据库建立^[7],填补了国内大型脑卒中数据库的空白。随后,国内脑卒中数据库构建体系日益完善,对开展脑卒中的流行病学研究及反馈中国脑卒中诊疗质量发挥重要作用^[3,8]。然而,通过对比 CNSR 数据库与其他国家大型数据库中脑卒中诊疗质量发现,同美国等已具备成熟脑卒中诊疗水平的高收入国家相比,中国在脑卒中有效干预和严格遵循脑卒中治疗方针等方面仍存在较大差距^[9]。在人工智能(AI)技术快速发展的今天,大数据挖掘以及数据共享已成为促进科学研究进展必不可少的环节之一。如何实现高效的数据导入、高质量的数据管理和维

doi: 10.3969/j.issn.1672-6731.2021.03.004

作者单位: 100730 中国医学科学院 北京协和医学院 北京协和医院 神经外科

通讯作者: 王任直, Email: wangrz@126.com

护,以及跨模态数据库的进一步探索,是脑卒中精准防治的基础,也是未来脑卒中数据库亟待解决的核心问题及发展方向。

一、脑卒中数据库应用历史及现状

1. 国际脑卒中数据库的应用及发展 最早的脑卒中数据库可溯源至 1978 年,由美国国立卫生研究院(NIH)国家神经疾病与卒中研究所基于系统性数据收集方法和标准化诊断算法,最终获得一整套脑卒中患者电子病历资料并将其存储至数据中心,然后研究者根据这些数据开展一系列涉及脑卒中诊断、临床实践和预后的前瞻性研究^[10]。参与该数据库建立和应用的团队包括神经专科医护人员、流行病学家、统计学家和计算机工程师等。1979 年,世界卫生组织(WHO)牵头组建的心脑血管疾病登记中心监测心血管疾病的趋势和决定因素研究(MONICA)项目,旨在探索不同人群心脑血管疾病的发病趋势和危险因素,在不到 10 年的时间内迅速发展成为涵盖 41 个分中心共 118 个合作单位的国际合作研究中心^[11]。在脑卒中数据库建设的早期,上述数据库凭借多中心、大样本量和多学科协作的模式,成为脑卒中数据库建设之典范。不足之处在于,早期的数据库结构和数据类型较为简单。20 世纪 90 年代以后,随着电子计算机技术的进一步发展和普及,大量脑卒中数据库及网络登记中心开始创建和应用。1991-1996 年开展的前瞻性随机对照试验——国际卒中试验(IST),旨在研究早期服用阿司匹林和肝素对缺血性卒中预后的影响,该数据库拥有完整的基线及随访数据,现已对外开放并允许进行二次分析^[12]。IST 数据库的开放大大提高了数据的利用率,对脑卒中数据库资源共享具有重要意义。但后续利用该数据库进行的脑卒中研究普遍存在共享数据字段相对单一等实验设计的不足。1999-2003 年,由 86 所美国医院参与的网络国家急性卒中治疗和操作规范化计划(Ethos)中心投入使用,该中心通过收集、分析并反馈脑卒中治疗的有效率,而提出有针对性的脑卒中治疗质量改善措施。经研究发现,通过对脑卒中治疗质量控制的登记和及时反馈促进了脑卒中护理技术的进步^[13]。此后,多中心参与的国家级大型数据库成为脑卒中诊治质量控制的重要工具。传统医学数据库的出现,极大地保存了不同阶段脑卒中患者的人群结构、发病特点和治疗方式等临床资料,为同步和后期开展病例对照研究提供了丰富的数据资源。但

此类医学数据库仍以文本资料为主,影像学资料较少,在脑卒中研究方向上无明显创新性。此后,随着研究人员对脑卒中认知的不断加深,一类以前瞻性队列研究为载体,聚焦于特定研究方向的大型多中心脑卒中数据库应运而生,如 2008 年公布的 INTERACT 试验(Intensive Blood Pressure Reduction in Acute Cerebral Haemorrhage Trial)数据库^[14]以及重组活化因子 VII(rFVIIa)治疗出血性卒中随机对照试验数据库^[15]等。此类前瞻性设计的数据库多涵盖包括影像学资料在内的丰富的数据类型,虽旨在探讨降压和重组活化因子 VII 治疗出血性卒中的有效性和安全性,但在不同研究方向上,同样的数据仍具有较高的可应用性,是 21 世纪以来脑卒中数据库建设的主导趋势之一。

2. 国内脑卒中数据库的应用及发展 因国内脑卒中诊治水平参差不齐、医疗保险类别和覆盖范围不一、登记网络组建协调性差等,使得国内脑卒中网络登记数据库的创建晚于欧美等发达国家。2007 年,国家卫生部(现为国家卫生健康委员会)和科技部启动建立了中国 CNSR 数据库,截至 2018 年 3 月共纳入 15 166 例脑卒中患者,覆盖 27 个省(自治区)和 4 个直辖市,历经前瞻性收集患者社会人口学资料、探究病因、提高诊疗和护理质量 3 个阶段,填补了中国脑卒中相关数据库的空缺^[16]。2015 年,中国卒中协会牵头的中国卒中中心联盟(CSCA)成立,该联盟以国内大型三级医院为基础,组成多中心脑卒中数据监测平台和临床研究网络,通过开展随机对照试验对比分析脑卒中各种治疗方法的有效性,对于夯实循证医学证据、有效反馈和促进脑卒中诊疗和护理质量具有积极作用。截至 2017 年 7 月,共 1576 所医院计 433 264 例脑卒中患者的临床资料在该联盟中得以保存和维护^[17]。2017 年,北京大学第一医院牵头建立的中国急性缺血性卒中治疗结局登记(CASTOR)数据库通过分析社会人口学数据、治疗成本和效益、不良事件和并发症,以探寻中国最佳成本-效益的脑卒中治疗方式^[18]。虽然国内脑卒中医学数据库因多中心参与、数据量大的优势可以开展大规模数据调研,但是海量的数据处理工作将限制数据的可靠性,应严格把关数据采集、录入和检验。此外,国家级大型数据库并不能较好地体现不同地区和不同级别医院脑卒中诊疗水平的差异,亟待建设区域性脑卒中数据库以解决上述问题。基于此,上海脑卒中服务系统作为区域性脑卒

中登记数据库,由上海市政府于 2013 年筹建,通过自动提取结构化电子病历的核心数据,分析美国国立卫生研究院卒中量表(NIHSS)的应用率、静脉溶栓比例和血管再通成功率等关键绩效指标,对评价区域性脑卒中诊疗和护理质量、构建脑卒中护理质量改善模型具有不可或缺的作用^[19]。然而,由于经济水平的地区差距,国内现有的脑卒中网络模型并不完全适用于经济相对薄弱的中西部地区。随着国家医疗保险体系的不断完善,多学科诊疗模式(MDT)及大数据共享理念的普及,未来中国脑卒中数据库还将进一步覆盖中西部地区和城乡区域,并不断提高其数据质量。

3. 综合型数据库在脑卒中中的应用 除国内外特定的脑卒中相关数据库外,大型综合医学数据库的创建历史更长、使用范围更广、重复利用率更高。1987 年,英国通用实践研究数据库(GPRD)建立,截至 2009 年共注册 160 万例在英国综合性医疗机构就诊的患者,同步记录社会人口学资料、就诊过程、诊断信息和治疗方式,根据性别和年龄的分布基本上可代表英国国民总体,其中,1997-2006 年该数据库共注册 48 239 例脑卒中患者。Toschke 等^[6]分析 GPRD 数据库中脑卒中相关数据,证实积极降压治疗对改善脑卒中患者预后具有重要作用。2002 年,中国台湾地区全民健康保险研究数据库成立,旨在通过循证医学研究制定临床决策和医疗政策,共纳入超过 99% 的台湾地区人口。2015 年,Hsieh 等^[20]利用该数据库的原始数据进行脑卒中合并肾功能障碍患者的流行病学调查。即便如此,综合型数据库由于缺少脑卒中相关特异性数据标签,难以进行深层次的数据挖掘和提出前瞻性的临床研究假设。

二、基于特定研究目标的脑卒中数据库的建立与应用

相比纳入数据全面的脑卒中数据库,基于特定研究目标的脑卒中数据库可对脑卒中进一步深层次探索。例如,2008 年建立的新兴预测脑卒中后语言功能结局的脑卒中解剖与功能成像数据库,基于完备的临床随访数据以及中枢神经系统解剖和 fMRI 数据,实现预测脑卒中患者语言和认知功能康复的目的,并助力开展神经损伤恢复机制的进一步研究^[21]。出血性卒中是脑卒中的重要组成部分,病死率和病残率均较高。中国出血性卒中患者因脑出血产生的医疗费用接近缺血性卒中患者住院费用的 2 倍^[8]。目前,国内外脑卒中数据库均以缺血

性卒中数据库为主,少有聚焦脑出血的数据库建设和临床数据分析^[21-23]。2018 年,由新加坡国家神经科学研究所(NNI)公布的新加坡脑卒中数据库,涵盖 2009 年 1 月至 2013 年 11 月共 1338 例自发性脑出血患者的前瞻性数据,并基于该数据库的数据分析提出较传统脑出血评分量表更精准的脑出血预后预测模型,即脑出血分级量表^[22]。2019 年,中国医学科学院北京协和医院牵头的国内多中心颅内出血影像学数据库建设项目成立,旨在借助医疗大数据和人工智能技术,突出以头部影像学资料为核心的多维数据分析,以构建脑出血精准诊断、评估和预测体系^[24]。目前该数据库涵盖国内至少 32 所医疗机构共 5287 例脑出血患者计 14 609 张头部 CT 扫描图像。该数据库实施全临床文本资料存放和完整脑出血影像标签登记,为后期开展脑出血影像组学分析和构建预测模型发挥重要作用。未来脑出血影像学数据库将进一步扩大病例容量,开展更多的临床合作研究,致力构建中国出血性卒中更全面的数据共享平台。

三、脑卒中数据库的目标及要求

脑卒中数据库的目标是获得流行病学数据统筹医疗规划,收集医学资料促进科学研究发展,反馈和改善医疗和护理质量。高质量的数据资源是脑卒中数据库的核心评价指标,这要求提供原始数据信息的医疗机构配备成熟的电子病历系统和先进的影像设备,病历得以妥善保存并具有较高的内涵价值。此外,前期进行数据采集及审查人员培训、制定详细的数据库构建协议和高质量的病例报告表格,同步进行定期的数据审查等,是确保数据准确性及完整性的重要方式。近年来,电子数据采集系统(EDC)被广泛应用,通过自动检查每例患者所有数据元素的完整性、元素值范围及编码的正确性,挑选出重复条目和逻辑错误并及时反馈,便于数据质量的监测^[17]。Schwamm 等^[23]认为,构建和运行脑卒中登记数据库需具备以下因素:(1)制定标准化数据标签并定义。(2)资金充足。(3)应用电子数据采集系统。(4)保证病例随访的完整性以及开放、活泼的团队合作。(5)在数据收集、验证、清理、输入、分析、报告生成和反馈的整个过程中,应充分保护患者的个人隐私。这需要项目实施前进行培训教育、设立数据库开放权限、数据元素匿名化、数据编码和安全的数据传输方式,以最大程度避免患者个人信息泄露。

四、人工智能技术在数据库建设中的应用

近年来,随着人工智能技术在医学影像自动分割^[25-27]、辅助疾病诊断^[28-29]和预测疾病转归^[30-31]等成功应用案例的不断出现,人工智能技术也为数据库的建立提供了新的方向。例如,利用计算机技术对数据进行批量去隐私化处理,不仅保证患者信息安全,更有利于推动数据共享。应用基于人工智能的自然语言处理方法快速构建文本标签^[32],可极大节约人工创建和填写标签的时间。Lee等^[33]研发出一种自动提取放射学报告关键信息的人工智能模型,该模型能够自动生成患者有无骨折的关键标签内容。Wakamiya等^[34]利用网络信息挖掘技术和自然语言处理技术捕获并分析基于社交媒体(Twitter)的医学相关数据,验证这种信息处理技术的可行性。此外,影像资料的自动分割技术对于影像数据的分类和保存具有重要意义^[25-26]。北京协和医院神经外科建立了一种基于卷积神经网络(CNN)的颅内血肿自动分割模型,同手动分割方法具有良好的一致性^[35]。近年关于脑出血相关水肿的自动分割也见诸报道^[36-37]。在人工智能技术的辅助下,构建脑卒中数据库的各个环节会更加便捷,数据内容会更加精确,数据类型也会更加全面。同时,人工智能技术开始在数据库建设中发挥作用,而源于大型数据库的海量数据也是人工智能应用模型开发的学习资料保障^[38-39]。Misawa等^[40]利用大型结肠镜检查视频数据库中的数据验证了一种人工智能技术辅助的结肠息肉检测系统,提高了该系统检测结肠病变的敏感性和特异性。Inaguma等^[41]采用基于大型数据库的机器学习预测模型,发现尿蛋白含量增加是慢性肾病患者肾小球滤过率迅速下降的危险因素。Stadler等^[42]前瞻性设计一种带有详细注释的肿瘤成像数据库,以用于未指定任务的人工智能模型的开发。未来,人工智能技术和数据库建设还将继续融合协作,智能化的数据加工、存储和运输也将成为新型数据库建设的主流趋势。

综上所述,脑卒中和大型综合医学数据库的不断出现和优化,极大促进脑卒中研究及医疗和护理质量的提高。虽然越来越多的数据库实现从结构单一、数据质量低向高质量、多学科协作模式的转化,但数据缺失、共享和隐私保护的矛盾等问题仍然存在。此外,中国脑卒中病例数巨大,有限的医疗资源和参差不齐的地区诊治水平也使得高质量的数据生成和数据库建设面临巨大挑战,这要求从

根源上重视原始病历质量、精心维护数据的保存和运输、合理调动数据的共享并注重患者隐私保护。未来,随着信息化和人工智能技术的发展,以及对跨模态影像数据库的进一步探索,中国脑卒中数据库必将发挥重要的作用。

利益冲突 无

参 考 文 献

- [1] Johnston SC, Mendis S, Mathers CD. Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling[J]. *Lancet Neurol*, 2009, 8:345-354.
- [2] Wang W, Jiang B, Sun H, Ru X, Sun D, Wang L, Wang L, Jiang Y, Li Y, Wang Y, Chen Z, Wu S, Zhang Y, Wang D, Wang Y, Feigin VL; NESS - China Investigators. Prevalence, incidence, and mortality of stroke in China: results from a nationwide population - based survey of 480 687 adults [J]. *Circulation*, 2017, 135:759-771.
- [3] Liu L, Wang D, Wong KS, Wang Y. Stroke and stroke care in China: huge burden, significant workload, and a national priority [J]. *Stroke*, 2011, 42:3651-3654.
- [4] Reeves MJ, Mullard AJ, Wehner S. Inter-rater reliability of data elements from a prototype of the Paul Coverdell National Acute Stroke Registry[J]. *BMC Neurol*, 2008, 8:19.
- [5] Abboud H, Labreuche J, Arauz A, Bryer A, Lavados PG, Massaro A, Munoz Collazos M, Steg PG, Yamout BI, Vicaut E, Amarenco P; OPTIC Registry Investigators. Demographics, socio-economic characteristics, and risk factor prevalence in patients with non-cardioembolic ischaemic stroke in low - and middle - income countries: the OPTIC registry[J]. *Int J Stroke*, 2013, 8 Suppl A100: 4-13.
- [6] Toschke AM, Wolfe CD, Heuschmann PU, Rudd AG, Gulliford M. Antihypertensive treatment after stroke and all - cause mortality: an analysis of the General Practitioner Research Database (GPRD)[J]. *Cerebrovasc Dis*, 2009, 28:105-111.
- [7] Wang Y, Cui L, Ji X, Dong Q, Zeng J, Wang Y, Zhou Y, Zhao X, Wang C, Liu L, Nguyen-Huynh MN, Claiborne Johnston S, Wong L, Li H; China National Stroke Registry Investigators. The China National Stroke Registry for patients with acute cerebrovascular events: design, rationale, and baseline patient characteristics[J]. *Int J Stroke*, 2011, 6:355-361.
- [8] Wang Y, Li Z, Zhao X, Wang D, Li H, Xian Y, Liu L, Wang Y. Stroke care quality in China: substantial improvement, and a huge challenge and opportunity[J]. *Int J Stroke*, 2017, 12:229-235.
- [9] Li Z, Wang C, Zhao X, Liu L, Wang C, Li H, Shen H, Liang L, Bettger J, Yang Q, Wang D, Wang A, Pan Y, Jiang Y, Yang X, Zhang C, Fonarow GC, Schwamm LH, Hu B, Peterson ED, Xian Y, Wang Y, Wang Y; China National Stroke Registries. Substantial progress yet significant opportunity for improvement in stroke care in China[J]. *Stroke*, 2016, 47:2843-2849.
- [10] Bronstein K, Murray P, Licata-Gehr E, Banko M, Kelly-Hayes M, Fast S, Kunitz S. The Stroke Data Bank project: implications for nursing research[J]. *J Neurosci Nurs*, 1986, 18:132-134.
- [11] WHO MONICA Project Principal Investigators. The World Health Organization MONICA Project (monitoring trends and determinants in cardiovascular disease): a major international collaboration[J]. *J Clin Epidemiol*, 1988, 41:105-114.
- [12] International Stroke Trial Collaborative Group. The International Stroke Trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19 435 patients with acute

- ischaemic stroke[J]. *Lancet*, 1997, 349:1569-1581.
- [13] Hills NK, Johnston SC. Duration of hospital participation in a nationwide stroke registry is associated with improved quality of care[J]. *BMC Neurol*, 2006, 6:20.
- [14] Anderson CS, Huang Y, Wang JG, Arima H, Neal B, Peng B, Heeley E, Skulina C, Parsons MW, Kim JS, Tao QL, Li YC, Jiang JD, Tai LW, Zhang JL, Xu E, Cheng Y, Heritier S, Morgenstern LB, Chalmers J; INTERACT Investigators. Intensive blood pressure reduction in acute cerebral haemorrhage trial (INTERACT): a randomised pilot trial [J]. *Lancet Neurol*, 2008, 7:391-399.
- [15] Mayer SA, Brun NC, Begtrup K, Broderick J, Davis S, Diringer MN, Skolnick BE, Steiner T; FAST Trial Investigators. Efficacy and safety of recombinant activated factor VII for acute intracerebral hemorrhage [J]. *N Engl J Med*, 2008, 358:2127-2137.
- [16] Wang Y, Jing J, Meng X, Pan Y, Wang Y, Zhao X, Lin J, Li W, Jiang Y, Li Z, Zhang X, Yang X, Ji R, Wang C, Wang Z, Han X, Wu S, Jia Z, Chen Y, Li H. The Third China National Stroke Registry (CNSR-III) for patients with acute ischaemic stroke or transient ischaemic attack: design, rationale and baseline patient characteristics[J]. *Stroke Vasc Neurol*, 2019, 4:158-164.
- [17] Wang Y, Li Z, Wang Y, Zhao X, Liu L, Yang X, Wang C, Gu H, Zhang F, Wang C, Xian Y, Wang DZ, Dong Q, Xu A, Zhao J. Chinese Stroke Center Alliance: a national effort to improve healthcare quality for acute stroke and transient ischaemic attack: rationale, design and preliminary findings [J]. *Stroke Vasc Neurol*, 2018, 3:256-262.
- [18] Sun W, Ou Q, Zhang Z, Qu J, Huang Y. Chinese acute ischemic stroke treatment outcome registry (CASTOR): protocol for a prospective registry study on patterns of real-world treatment of acute ischemic stroke in China [J]. *BMC Complement Altern Med*, 2017, 17:357.
- [19] Dong Y, Fang K, Wang X, Chen S, Liu X, Zhao Y, Guan Y, Cai D, Li G, Liu J, Liu J, Zhuang J, Wang P, Chen X, Shen H, Wang DZ, Xian Y, Feng W, Campbell BC, Parsons M, Dong Q. The network of Shanghai Stroke Service System (4S): a public health-care web-based database using automatic extraction of electronic medical records[J]. *Int J Stroke*, 2018, 13:539-544.
- [20] Hsieh CY, Chen CH, Li CY, Lai ML. Validating the diagnosis of acute ischemic stroke in a National Health Insurance claims database[J]. *J Formos Med Assoc*, 2015, 114:254-259.
- [21] Seghier ML, Patel E, Prejawa S, Ramsden S, Selmer A, Lim L, Browne R, Rae J, Haigh Z, Ezekiel D, Hope TMH, Leff AP, Price CJ. The PLORAS database: a data repository for predicting language outcome and recovery after stroke [J]. *Neuroimage*, 2016, 124(Pt B):1208-1212.
- [22] Han JX, See AAQ, King NKK. Validation of prognostic models to predict early mortality in spontaneous intracerebral hemorrhage: a cross-sectional evaluation of a Singapore Stroke Database[J]. *World Neurosurg*, 2018, 109:e601-608.
- [23] Schwamm L, Reeves MJ, Frankel M. Designing a sustainable national registry for stroke quality improvement[J]. *Am J Prev Med*, 2006, 31(6 Suppl 2):S251-257.
- [24] Wang RZ, Chang JB, Feng M. Prospects for precise diagnosis, assessment, prediction and treatment of hemorrhagic stroke [J]. *Zhongguo Xian Dai Shen Jing Ji Bing Za Zhi*, 2019, 19:618-621. [王任直, 常健博, 冯铭. 出血性卒中精准诊断、评估、预测及治疗展望[J]. *中国现代神经疾病杂志*, 2019, 19:618-621.]
- [25] Renard F, Guedria S, Palma N, Vuillerme N. Variability and reproducibility in deep learning for medical image segmentation [J]. *Sci Rep*, 2020, 10:13724.
- [26] Ding Y, Acosta R, Enguix V, Suffren S, Ortmann J, Luck D, Dolz J, Lodygensky GA. Using deep convolutional neural networks for neonatal brain image segmentation [J]. *Front Neurosci*, 2020, 14:207.
- [27] Yang B, Chen W, Luo H, Tan Y, Liu M, Wang Y. Neuron image segmentation via learning deep features and enhancing weak neuronal structures [J]. *IEEE J Biomed Health Inform*, 2020.[Epub ahead of print]
- [28] Kong Y, Kong X, He C, Liu C, Wang L, Su L, Gao J, Guo Q, Cheng R. Constructing an automatic diagnosis and severity-classification model for acromegaly using facial photographs by deep learning[J]. *J Hematol Oncol*, 2020, 13:88.
- [29] Shin H, Oh S, Hong S, Kang M, Kang D, Ji YG, Choi BH, Kang KW, Jeong H, Park Y, Hong S, Kim HK, Choi Y. Early-stage lung cancer diagnosis by deep learning-based spectroscopic analysis of circulating exosomes [J]. *ACS Nano*, 2020, 14:5435-5444.
- [30] Skrede OJ, De Raedt S, Kleppe A, Hveem TS, Liestøl K, Maddison J, Askautrud HA, Pradhan M, Nesheim JA, Albrechtsen F, Farstad IN, Domingo E, Church DN, Nesbakken A, Shepherd NA, Tomlinson I, Kerr R, Novelli M, Kerr DJ, Danielsen HE. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study [J]. *Lancet*, 2020, 395:350-360.
- [31] Saha S, Pagnozzi A, Bourgeat P, George JM, Bradford D, Colditz PB, Boyd RN, Rose SE, Frupp J, Pannek K. Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model[J]. *Neuroimage*, 2020, 215:116807.
- [32] Juhn Y, Liu H. Artificial intelligence approaches using natural language processing to advance EHR-based clinical research [J]. *J Allergy Clin Immunol*, 2020, 145:463-469.
- [33] Lee C, Kim Y, Kim YS, Jang J. Automatic disease annotation from radiology reports using artificial intelligence implemented by a recurrent neural network[J]. *AJR Am J Roentgenol*, 2019, 212:734-740.
- [34] Wakamiya S, Morita M, Kano Y, Ohkuma T, Aramaki E. Tweet classification toward twitter-based disease surveillance: new data, methods, and evaluations[J]. *J Med Internet Res*, 2019, 21:e12783.
- [35] Chang JB, Jiang SZ, Chen XJ, Lok KH, Lee YL, Zhang QH, Wei JJ, Shi L, Feng M, Wang RZ. Consistency evaluation of an automatic segmentation for quantification of intracerebral hemorrhage using convolution neural network [J]. *Zhongguo Xian Dai Shen Jing Ji Bing Za Zhi*, 2020, 20:585-590. [常健博, 姜桑桐, 陈显金, 骆嘉希, 李沃霖, 张庆华, 魏俊吉, 石林, 冯铭, 王任直. 基于卷积神经网络的自发性脑出血血肿分割方法的一致性评价[J]. *中国现代神经疾病杂志*, 2020, 20:585-590.]
- [36] Dhar R, Falcone GJ, Chen Y, Hamzehloo A, Kirsch EP, Noche RB, Roth K, Acosta J, Ruiz A, Phuah CL, Woo D, Gill TM, Sheth KN, Lee JM. Deep learning for automated measurement of hemorrhage and perihematomal edema in supratentorial intracerebral hemorrhage[J]. *Stroke*, 2020, 51:648-651.
- [37] Zhao X, Chen K, Wu G, Zhang G, Zhou X, Lv C, Wu S, Chen Y, Xie G, Yao Z. Deep learning shows good reliability for automatic segmentation and volume measurement of brain hemorrhage, intraventricular extension, and peripheral edema [J]. *Eur Radiol*, 2021.[Epub ahead of print]
- [38] Stourac J, Dubrava J, Musil M, Horackova J, Damborsky J, Mazurenko S, Bednar D. FireProtDB: database of manually curated protein stability data[J]. *Nucleic Acids Res*, 2021, 49:D319-324.
- [39] Zhou D, Tian F, Tian X, Sun L, Huang X, Zhao F, Zhou N,

Chen Z, Zhang Q, Yang M, Yang Y, Guo X, Li Z, Liu J, Wang J, Wang J, Wang B, Zhang G, Sun B, Zhang W, Kong D, Chen K, Li X. Diagnostic evaluation of a deep learning model for optical diagnosis of colorectal cancer[J]. Nat Commun, 2020, 11:2961.

[40] Misawa M, Kudo SE, Mori Y, Hotta K, Ohtsuka K, Matsuda T, Saito S, Kudo T, Baba T, Ishida F, Itoh H, Oda M, Mori K. Development of a computer - aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video) [J]. Gastrointest Endosc, 2020. [Epub ahead of print]

[41] Inaguma D, Kitagawa A, Yanagiya R, Koseki A, Iwamori T,

Kudo M, Yuzawa Y. Increasing tendency of urine protein is a risk factor for rapid eGFR decline in patients with CKD: a machine learning - based prediction model by using a big database[J]. PLoS One, 2020, 15:e0239262.

[42] Stadler CB, Lindvall M, Lundström C, Bodén A, Lindman K, Rose J, Treanor D, Blomma J, Stacke K, Pinchaud N, Hedlund M, Landgren F, Woisetschläger M, Forsberg D. Proactive construction of an annotated imaging database for artificial intelligence training[J]. J Digit Imaging, 2021, 34:105-115.

(收稿日期:2021-03-05)
(本文编辑:袁云)

· 小词典 ·

中英文对照名词词汇(二)

感觉神经传导速度
sensory nerve conduction velocity(SNCV)

Karnofsky 功能状态评分
Karnofsky Performance Status(KPS)

国际标准化比值 international normalized ratio(INR)

国际头痛疾病分类第 2 版
International Classification of Headache Disorders Second Edition(ICHD- II)

国际卒中试验 International Stroke Trial(IST)

国家临床数据库 National Clinical Database(NCD)

国家神经科学研究所 National Neuroscience Institute(NNI)

核糖核苷酸还原酶 ribonucleotide reductase(RR)

红细胞沉降率 erythrocyte sedimentation rate(ESR)

琥珀酸脱氢酶 succinate dehydrogenase(SDH)

还原型辅酶 I 四氮唑还原酶
reduced nicotinamide adenine dinucleotide-tetrazolium reductase(NADH-TR)

环磷酸鸟苷 cyclic guanosine monophosphate(cGMP)

环磷酸腺苷 cyclic adenosine monophosphate(cAMP)

患者报告结局 patient-reported outcome(PRO)

Logistic 回归 Logistic regression(LR)

Glasgow 昏迷量表 Glasgow Coma Scale(GCS)

活化部分凝血活酶时间
activated partial thromboplastin time(APTT)

机器学习 machine learning(ML)

肌酸激酶 creatine kinase(CK)

基于诊断的分级预后评分
Diagnosis Specific-Graded Prognostic Assessment Score (ds-GPA)

O⁶-甲基鸟嘌呤-DNA 甲基转移酶
O⁶-methylguanine-DNA methyltransferase(MGMT)

DNA 甲基转移酶 DNA methyltransferase(DNMT)

甲氧基异丁基异腈 methoxyisobutylisonitril(MIBI)

甲状旁腺激素 parathyroid hormone(PTH)

甲状腺髓样癌 medullary thyroid carcinoma(MTC)

监测心血管疾病的趋势和决定因素研究
Monitoring of Trends and Determinants in Cardiovascular Diseases(MONICA)

健康相关生活质量 health-related quality of life(HRQoL)

降钙素基因相关肽 calcitonin gene-related peptide(CGRP)

胶质母细胞瘤 glioblastoma(GBM)

卷积神经网络 convolution neural network(CNN)

决策树 decision tree(DT)

抗癫痫药物 antiepileptic drugs(AEDs)

可逆性后部脑病综合征
reversible posterior encephalopathy syndrome(RPES)

可逆性脑血管收缩综合征
reversible cerebral vasoconstriction syndrome(RCVS)

库欣综合征 Cushing's syndrome(CS)

类固醇生成因子-1 steroidogenic factor-1(SF-1)

立体定向放射外科 stereotactic radiosurgery(SRS)

X 连锁 α 地中海贫血伴精神发育迟滞综合征蛋白
X-linked α-thalassaemia retardation syndrome protein (ATRX)

连线测验 Trail Making Test(TMT)

颅脑创伤 traumatic brain injury(TBI)

颅内静脉窦血栓形成
cerebral venous sinus thrombosis(CVST)

美国癌症研究学会
American Association for Cancer Research(AACR)

美国国家癌症研究所 National Cancer Institute(NCI)

美国国立癌症数据库 National Cancer Data Base(NCDB)

美国国立卫生研究院 National Institutes of Health(NIH)

美国国立卫生研究院卒中量表
National Institutes of Health Stroke Scale(NIHSS)

美国国立综合癌症网
National Comprehensive Cancer Network(NCCN)