

DOI: 10.13376/j.cbls/2023174

文章编号: 1004-0374(2023)12-1594-05



姜勇, 北京天坛医院国家神经系统疾病临床研究中心大数据中心主任, 流行病学与卫生统计学博士。中华预防医学会慢病分会、流行病学分会、健康管理分会、健康测量与评价专委会委员。科技部“十四五”“常见多发病防治研究”专项管理专家组成员。

## 国内外卒中队列研究进展

郭 藏<sup>1,2</sup>, 谷鸿秋<sup>1</sup>, 王拥军<sup>1,2</sup>, 姜 勇<sup>1,2\*</sup>

(1 昌平实验室, 北京 102206; 2 首都医科大学附属北京天坛医院, 国家神经系统疾病临床医学研究中心, 北京 100070)

**摘要:** 卒中是危害人类健康的主要疾病之一, 但卒中可防可控。队列研究在病因学研究中具有不可替代的地位和作用, 可为疾病的防控提供证据支撑。在卒中领域, 国内外已经有一些应用队列研究的实例, 并取得了一定的成果, 对于卒中防控具有重要的医学价值和社会价值。该文对经典的卒中社区队列和临床队列研究进展进行综述, 为我国未来卒中队列建设的方向提供思路。

**关键词:** 卒中; 社区队列; 临床队列

**中图分类号:** Q81; R743 **文献标志码:** A

## Research progress on stroke cohort domestically and internationally

GUO Cang<sup>1,2</sup>, GU Hong-Qiu<sup>1</sup>, WANG Yong-Jun<sup>1,2</sup>, JIANG Yong<sup>1,2\*</sup>

(1 Changping Laboratory, Beijing 102206, China; 2 China National Clinical Research Center for Neurological Diseases, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China)

**Abstract:** Stroke is one of the major diseases that significantly impact human health, yet it is preventable and manageable. Cohort studies hold an irreplaceable position in etiological research, providing evidence-based support for disease prevention and control. In the field of stroke, both domestically and internationally, there have been instances of cohort studies that have yielded certain achievements, holding significant medical and societal value for stroke prevention and control. This article provides an overview of the progress on classic stroke community and clinical cohort studies, aiming to offer insights into the direction of future stroke cohort development in China.

**Key words:** stroke; community cohort; clinical cohort

卒中是危害人类健康的主要疾病之一, 其高发病率、高复发率、高致残率和高致死率给患者家庭和社会造成了沉重的负担。队列研究是流行病学最基本的观察性研究设计之一, 可判定暴露对结局的作用, 研究结果外推性好, 在病因学研究中具有不

可替代的地位和作用。在卒中领域, 国内外不乏应用队列研究的实例, 既有社区队列, 也有临床队列,

收稿日期: 2023-07-31; 修回日期: 2023-11-15

\*通信作者: E-mail: jiangyong@ncrcnd.org.cn

这些队列已经取得了一定的成果, 对降低卒中的发病和复发风险具有重要的医学价值和社会价值。本文主要介绍卒中领域内代表性社区队列和临床队列研究的数据概况、研究成果和未来发展计划, 为今后我国卒中队列建设提供参考依据。

## 1 社区队列研究

### 1.1 国外持续多年的经典社区队列研究

在卒中领域, 国外有一些社区队列研究成果颇丰, 如弗莱明翰研究 (The Framingham Study, FHS)、社区动脉粥样硬化风险 (Atherosclerosis Risk In Communities, ARIC) 研究、北部曼哈顿研究 (The Northern Manhattan Study, NOMAS)、卒中地理和种族差异原因 (The Reasons for Geographic And Racial Differences in Stroke, REGARDS) 研究、牛津血管研究 (The Oxford Vascular Study, OXVASC)、鹿特丹研究 (The Rotterdam Study, RS) 等。ARIC 识别了卒中的危险因素, 包括生活方式、生化指标、影像学指标等, 在此基础上, 构建了卒中预测模型; 也报告了卒中发病率、患病率、死亡率等流行趋势; 随着实验室检查的改进, 未来还会探索卒中发生的基因组学、代谢组学、蛋白质组学和其他机制途径<sup>[1]</sup>。NOMAS 评估了卒中的种族和民族差异, 发现了不同种族和民族间卒中发病、亚型、危险因素和理想心血管健康状态存在差异, 为减少卒中差异的干预措施制定奠定了基础<sup>[2]</sup>。REGARDS 是一项基于人群的全国性队列研究, 主要目的是确定美国东南部和非洲裔美国人中风死亡率过高的原因, 以减小卒中的地理和种族差异<sup>[3]</sup>。OXVASC 前瞻性地评估了英国牛津郡人群心血管事件, 在卒中发生后, 收集了患者全面的数据, 他们收集了患者病前 mRS 评分和既往病史数据; 此外, 对患者进行详细的临床检查, 评估了卒中严重程度, 并收集脑部及血管成像、心电图及血液样本等信息<sup>[4]</sup>。特别地, FHS 历史最悠久, RS 更新方法和结果最及时频繁, 这里重点介绍这两项经典队列研究。

FHS 是美国持续时间最长的心血管队列研究, 对于发现卒中的危险因素有重要意义。随着时间的推移, FHS 收集的数据内容也不断丰富: 1948 年的第一轮调查主要收集了研究对象的病史、吸烟情况、人体测量数据以及生物样本 (用于检测血糖和胆固醇浓度); 1979 年就开始使用超声波检查、运动测试、精神状态检查 (MMSE) 和其他神经认知测试; 1980 年后陆续开始测量大量的生物标志物;

1990 年后推行霍尔特监测 (用于房颤监测); 20 世纪 90 年代末建立血管功能评估站, 可获得内皮功能、主动脉强度及外周动脉眼压; 2009 年后开始引入更先进的成像模式, 如磁共振成像; 近年来开展基因组研究, 成为第一批获得全基因单核苷酸多态性 (SNP) 数据的队列之一, 并且是基因型和表型数据库 (DBGAP) 最大的数据贡献者之一<sup>[5]</sup>。在 70 多年的时间里, FHS 影响了流行病学家对血压和房颤与卒中风险关系的理解, 并探索了遗传因素与卒中风险的关联。它在 1970 年证实了收缩压对于卒中风险的重要性不亚于舒张压或平均动脉压的观点; 在 1978 年发现房颤是卒中的危险因素, 房颤患者的卒中风险是非房颤者的 5.6 倍; 在 2010 年发表的文献指出, 父母有卒中病史的子女发生卒中风险增加了 3 倍<sup>[6]</sup>。21 世纪, FHS 引领潮流, 跨越传统范畴走向分子流行病学时代, 开始引入人类基因组绘图技术, 对中风的遗传决定因素进行广泛的基因组联合研究。FHS 走在遗传流行病学研究的前沿, 未来会进一步阐明遗传因素对卒中风险的作用, 特别是特定的卒中亚型, 以助于更好地预防个体卒中发生。

RS 是荷兰 1990 年开始的一项持续的前瞻性队列研究, 旨在阐明中老年慢性病的危险因素对于卒中的影响, 其目标是关注卒中的发生频率、病因和早期识别。RS 研究以其高频率的随访和广泛的研究成果为卒中研究提供了重要证据。RS 自 1990 年招募基线参与者以来, 随后每 3~6 年对其进行一次随访, 目前已随访至第七次; 自 2007 年以来几乎每隔两年都在欧洲流行病学杂志 (*European Journal of Epidemiology, Eur J Epidemiol*) 上更新研究目标、研究设计及主要研究成果。RS 参与者都在基线接受了详尽的检查, 在家中接受了 2 h 的面谈, 然后在指定地点接受 5 h 的身体检查, 重点放在收集影像和生物样本信息。为了确定卒中事件, RS 通过访谈评估基线卒中病史, 并在医院记录中证实, 通过研究数据库与医院的电子病历系统链接, 以确定和补充信息。RS 也不断补充脑成像数据, 20 世纪 90 年代对部分研究对象进行了脑磁共振成像, 2005 年引入 1.5 Tesla 扫描仪, 对所有无禁忌证的研究对象进行了脑成像<sup>[7]</sup>。近年来, RS 在卒中领域贡献或领导了 GWAS 以及其他基因组学工作, 以探索卒中的新遗传位点及其临床的表型, 并使用肌电图对周围神经评估<sup>[8-9]</sup>。除了研究了卒中的危险因素、疾病负担和长期预后之外, RS 正在积极地将卒中病因研究转化为疾病风险预测研究及公共卫生预

防研究,如发现外观正常的白质微结构的细微变化可以预测卒中发生、改进弗莱明翰卒中风险评估模型及鼓励对可改变的危险因素进行干预<sup>[10-12]</sup>。

## 1.2 国外大型社区队列研究

卒中发生是基因、环境、生活方式的复杂相互作用所致。纵然是历史悠久的FHS、RS,也无法解析卒中的病因链,仍需要大型队列研究来精细解析卒中的病因结构。目前,国外达到50万左右规模的人群队列有四项:欧洲10国的European Prospective Investigation into Cancer and Nutrition (EPIC)、美国的NIH-AARP Diet and Health Study (NIH-AARP)、英国的Million Women Study (MWS)和UK Biobank (UKB)。除此之外,法国的Nutrinet-Sante Study、瑞典的LifeGene仍在队列成员募集期,目标也是50万人,而美国All of Us研究旨在招募100万人以上。这里重点介绍UKB和All of Us<sup>[13]</sup>。

UKB是一个大型前瞻性队列研究及生物医学数据库,旨在通过构建大规模的人类信息资源库来探索基因、环境、生活方式和健康的联系。在2006—2010年期间,UKB完成了50万基线参与者的招募和数据采集,采集内容包括问卷、身体测量数据、生物样本等;特别地,采集了20万参与者的外显子数据及10万参与者的全身影像数据。UKB的随访依托英国国家健康服务系统强大的支持和保障,链接可用的常规国家数据集,包括初级保健、死亡登记、癌症筛查数据和特定疾病的登记数据、住院数据等。在卒中领域,UKB主要产出内容包括基因与卒中的关系、基因与环境的交互作用对卒中的影响、影像学研究以及人工智能分析等。UKB研究已经建立了多基因分级评分,并评估了该评分和健康生活方式(当前不吸烟、健康饮食、BMI<30 kg/m<sup>2</sup>以及每周2次及以上的适度身体活动)与卒中事件的相关性,结果发现无论有无遗传风险,不利的生活方式会增加卒中风险<sup>[14-15]</sup>;另外,UKB利用最新的外显子组测序数据发现导致脑小管病的NOTCH3、HTRA1和COL4A1/2单基因变异与卒中风险增加相关,同时发现控制心血管危险因素有助于降低卒中风险,以上结果从侧面反映出即使人体存在着卒中易感基因,也能通过健康生活方式达到预防卒中的目的<sup>[16]</sup>。根据UKB的计划和随访工作的推进,越来越多的基因、环境、生活习惯以及人体影像学数据会被采集、发布、挖掘,未来UKB研究也会提供更多关于卒中预防的宝贵经验。

All of Us是一项雄心勃勃的健康和遗传学研究

队列,旨在通过将百万人基因测序数据与其健康记录关联,加速健康研究和医学突破。All of Us已拥有大规模的生物样本信息库并初步形成大规模队列。最近报道显示,该研究已产生第一个全基因组数据集,包含近十万的全基因组序列,并已有47.4万多人参与,预期2026年底达到100万参与者的初始目标<sup>[17]</sup>。All of Us数据种类多,收集了包括问卷调查、体格检查、生物样本、电子健康记录、数字化健康数据(可穿戴设备)、基因组、医疗理赔数据及地理环境数据等多样性数据,并计划开展长期的随访监测。All of Us对复杂来源及不同类型的数据进行了规范管理,先对数据进行标准化处理,随后整合多源异构数据,并储存到标准数据库;另外,利用区块链技术等建立安全的云环境,对数据进行云存储和云共享。All of Us解决了传统生物医学研究中非种族少数群体样本量不足的问题,如计划包括45%以上少数人种和族群,已揭示了非种族少数群体的卒中负担<sup>[18]</sup>。

## 1.3 国内经典社区队列研究

与欧美发达国家相比,我国的队列研究起步较晚,20世纪八九十年代陆续开始开展,近10余年陆续开展大规模且有影响力的队列研究。泰州队列(Taizhou Longitudinal Study, TZL Study)和中国慢性病前瞻性队列(China Kadoorie Biobank, CKB)是我国规模最大的社区队列研究。泰州队列基线人群已达20万例,获得了每人1000多个表型、上百种暴露资料,并持续跟踪10余年,还收集了150万份生物样本,建成了人类遗传资源平台和健康大数据平台<sup>[19]</sup>。CKB最大的优势是存储了51万余人的基线调查和两次重复测量的样本,存储规模达到58.8万份,并实现了样本存储管理的全程电子化和所有样本全方位监管,这在世界上居于领先地位<sup>[13]</sup>。在卒中领域,国内已有一些社区队列产生了有价值且有影响力的科研成果,发表在医学四大刊(*Lancet*、*BMJ*、*JAMA*和*N Engl J Med*),如中国动脉粥样硬化性心血管疾病风险预测(prediction for ASCVD risk in China, China-PAR)研究和CKB。对于卒中,它们重点关注了肥胖<sup>[20]</sup>、环境污染物<sup>[21]</sup>、社会经济<sup>[22]</sup>对卒中的影响。值得一提的是,国内有一些社区队列启动了影像学检查,未来会为卒中的早期识别、准确诊断、早诊早治提供支持,如泰州脑影像子队列(The Taizhou Imaging Study, TIS)<sup>[23]</sup>、多血管评估认知损伤和血管事件研究(PolyvasculaR Evaluation for Cognitive Impairment and vaScular Events,

PRECISE)<sup>[24]</sup>等;另外,CKB和TZL study均也开展了基因组研究,未来会为研究遗传因素在卒中发生发展中的作用及相应的分子机制提供数据支持。

虽然国内的社区队列研究给卒中防控提供了高质量的本土化证据,但卒中队列的发展也存在着一系列问题和挑战:第一,上述有些队列的样本来自特定社区,研究证据可能无法适用于全中国人群,如TZL Study和PRECISE;第二,随着研究的推进,研究对象年龄增长,部分研究参与者可能会流失,难以持续长期追踪;第三,生物样本库、影像学库的运行和存储均需要长期维持,需要大量的资金人力投入,也是一项挑战。

未来我国社区队列研究会继续整合生活方式、遗传、环境等多维态数据,进一步探索卒中的复杂机制。基于多组学和影像学的深入分析,制定更精准的卒中预防和干预策略。加强数据管理和保护研究对象隐私,建立数据资源共享平台,推动数据共享开放,探索和验证中国人群的科研证据,支撑卒中防治决策和实践。

## 2 临床队列

临床队列是基于登记注册的真实世界研究,可用来探索疾病预后的危险因素,进而为临床试验研究提供设计和干预方案,从而提高疾病的诊疗质量。目前,国内外已有一些临床队列实例取得了宝贵的成果,如美国“跟着指南走”卒中登记(Get With The Guidelines-Stroke, GWTG-Stroke)研究、南伦敦卒中登记(South London Stroke Register, SLSR)研究、英格兰和威尔士的国家卒中审计计划(Sentinel Stroke National Audit Programme, SSNAP)、21个国家参与的TIA registry.org项目及中国国家卒中登记研究-III(The Third China National Stroke Registry, CNSR-III)。

缩小指南和实践之间的差距是医疗改善的目标,其中最成功的例子就是GWTG-Stroke。GWTG-Stroke利用患者管理工具(Outcome、Inc、Cambridge、Mass)收集数据,数据内容包括人口统计学、病史、基线脑部影像资料、住院治疗和事件、出院带药及出院目的地。它的研究结果强调了缺血性卒中治疗时间的关键性,鼓励缩短卒中患者接受溶栓治疗时间和血管内治疗时间,以改善预后<sup>[25-26]</sup>。

当前的卒中登记研究大多没有关注到卒中预后的长期结局,大多随访至出院或者发病后90天。目前,TIA registry.org和CNSR-III已随访至5年,

其中,TIA registry.org已在*N Engl J Med*上发表了一篇关于短暂性脑缺血发作或轻微缺血性卒中后5年卒中复发风险研究<sup>[27]</sup>。CNSR-III最大的特点是收集了患者完整的影像学信息、颅内动脉评估数据及多基因组学数据,建立了卒中多组学研究平台,形成了目前全球规模最大的缺血性卒中多组学数据库<sup>[28]</sup>。传统新药研发很难找到新的干预靶点,而多组学队列可为研发治疗卒中新药物提供新思路,基于高质量多组学队列,利用多组学联合分析可探索卒中的治疗新靶点,然后开展细胞实验、动物实验和临床研究,这将会加速药物研发进程,提高临床转化成功率。

## 3 总结与展望

目前,国内外已经有了一些队列研究产出为卒中防控做出了突出贡献。在卒中领域,早期大多数队列研究关注的暴露都是传统的危险因素。随着影像学、基因组学和人工智能技术的快速发展,有越来越多的队列探索卒中早期诊断的影像学标志物及研究卒中发生和发展的遗传学机理。与欧美发达国家相比,我国的社区队列研究起步较晚,获得的数据和取得的研究成果也更少,未来需要建立全国代表性卒中社区队列,应用智慧医疗、无创检测、可穿戴和物联网等先进数据采集技术收集更准确、精细的数据,完善影像和多组学研究。可喜的是,我国已经有卒中临床队列引入了先进的影像技术及人工智能技术,也建立了生物样本库及多组学研究平台,并且有国家政府支持,未来继续完善多组学数据采集、分析、管理和共享平台,有可能在卒中队列领域实现弯道超车。

### [参 考 文 献]

- [1] Wright JD, Folsom AR, Coresh J, et al. The ARIC (Atherosclerosis Risk In Communities) Study: JACC focus seminar 3/8. *J Am Coll Cardiol*, 2021, 77: 2939-59
- [2] Sacco RL. Stroke disparities: from observations to actions: Inaugural Edward J. Kenton Lecture 2020. *Stroke*, 2020, 51: 3392-405
- [3] Howard VJ, Cushman M, Pulley L, et al. The reasons for geographic and racial differences in stroke study: objectives and design. *Neuroepidemiology*, 2005, 25: 135-43
- [4] Downer MB, Li L, Carter S, et al. Associations of Multimorbidity With Stroke Severity, Subtype, Premorbid Disability, and Early Mortality: Oxford vascular study. *Neurology*, 2023, 101: e645-52
- [5] Andersson C, Johnson AD, Benjamin EJ, et al. 70-year

- legacy of the Framingham Heart Study. *Nat Rev Cardiol*, 2019, 16: 687-98
- [6] Romero JR, Wolf PA. Epidemiology of stroke: legacy of the Framingham Heart Study. *Glob Heart*, 2013, 8: 67-75
- [7] Hofman A, Breteler MM, van Duijn CM, et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol*, 2007, 22: 819-29
- [8] Ikram MA, Brusselle G, Murad SD, et al. The Rotterdam Study: 2018 update on objectives, design and main results. *Eur J Epidemiol*, 2017, 32: 807-50
- [9] Ikram MA, Brusselle G, Ghanbari M, et al. Objectives, design and main findings until 2020 from the Rotterdam Study. *Eur J Epidemiol*, 2020, 35: 483-517
- [10] Evans TE, O'Sullivan MJ, de Groot M, et al. White matter microstructure improves stroke risk prediction in the general population. *Stroke*, 2016, 47: 2756-62
- [11] Bos D, Ikram MA, Leening M, et al. The revised Framingham stroke risk profile in a primary prevention population: the Rotterdam Study. *Circulation*, 2017, 135: 2207-9
- [12] Bos MJ, Koudstaal PJ, Hofman A, et al. Modifiable etiological factors and the burden of stroke from the Rotterdam study: a population-based cohort study. *PLoS Med*, 2014, 11: e1001634
- [13] 李立明, 吕筠. 大型前瞻性人群队列研究进展. *中华流行病学杂志*, 2015, 36: 1187-9
- [14] Rutten-Jacobs LC, Larsson SC, Malik R, et al. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank participants. *BMJ*, 2018, 363: k4168
- [15] Elliott J, Bodinier B, Bond TA, et al. Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. *JAMA*, 2020, 323: 636-45
- [16] Cho B, Harshfield EL, Al-Thani M, et al. Association of vascular risk factors and genetic factors with penetrance of variants causing monogenic stroke. *JAMA Neurol*, 2022, 79: 1303-11
- [17] Sullivan F, McKinstry B, Vasishta S. The "All of Us" research program. *N Engl J Med*, 2019, 381: 1883-4
- [18] Leasure AC, Acosta JN, Both C, et al. Stroke disparities among nonracial minorities in the All of Us research program. *Stroke*, 2021, 52: e488-90
- [19] 陈兴栋, 蒋艳峰, 徐萍, 等. 大型人群队列遗传资源建设与利用. *遗传*, 2021, 43: 980-7
- [20] Chen Z, Iona A, Parish S, et al. Adiposity and risk of ischaemic and haemorrhagic stroke in 0.5 million Chinese men and women: a prospective cohort study. *Lancet Glob Health*, 2018, 6: e630-40
- [21] Huang K, Liang F, Yang X, et al. Long term exposure to ambient fine particulate matter and incidence of stroke: prospective cohort study from the China-PAR project. *BMJ*, 2019, 367: l6720
- [22] Levy M, Chen Y, Clarke R, et al. Socioeconomic differences in health-care use and outcomes for stroke and ischaemic heart disease in China during 2009-16: a prospective cohort study of 0.5 million adults. *Lancet Glob Health*, 2020, 8: e591-602
- [23] Jiang Y, Cui M, Tian W, et al. Lifestyle, multi-omics features, and preclinical dementia among Chinese: The Taizhou Imaging Study. *Alzheimers Dement*, 2021, 17: 18-28
- [24] Pan Y, Jing J, Cai X, et al. Polyvascular Evaluation for Cognitive Impairment and vascular Events (PRECISE)-a population-based prospective cohort study: rationale, design and baseline participant characteristics. *Stroke Vasc Neurol*, 2021, 6: 145-51
- [25] Man S, Xian Y, Holmes DN, et al. Association between thrombolytic door-to-needle time and 1-year mortality and readmission in patients with acute ischemic stroke. *JAMA*, 2020, 323: 2170-84
- [26] Jahan R, Saver JL, Schwamm LH, et al. Association between time to treatment with endovascular reperfusion therapy and outcomes in patients with acute ischemic stroke treated in clinical practice. *JAMA*, 2019, 322: 252-63
- [27] Amarenco P, Lavallée PC, Monteiro Tavares L, et al. Five-year risk of stroke after TIA or minor ischemic stroke. *N Engl J Med*, 2018, 378: 2182-90
- [28] Cheng S, Xu Z, Bian S, et al. The STROMICS genome study: deep whole-genome sequencing and analysis of 10K Chinese patients with ischemic stroke reveal complex genetic and phenotypic interplay. *Cell Discov*, 2023, 9: 75