

# 中枢神经系统炎性脱髓鞘专病临床科研数据库设计与构建\*

武雷<sup>1</sup> 王冰<sup>2</sup> 余倩<sup>2</sup> 孙慧<sup>1</sup> 赵赫<sup>1</sup> 高赛<sup>1</sup> 郭荷娜<sup>3</sup> 黄燕宁<sup>1</sup>  
孟召友<sup>4</sup> 肖连臣<sup>5</sup> 许海珍<sup>6</sup> 黄德晖<sup>1</sup>

(<sup>1</sup> 中国人民解放军总医院第一医学中心神经内科医学部 北京 100853

<sup>2</sup> 浙江省数字医学诊断技术重点实验室 杭州 310030 <sup>3</sup> 陕西省人民医院神经内二科 西安 710068

<sup>4</sup> 陆军军医大学第二附属医院神经内科 重庆 400038 <sup>5</sup> 湖北文理学院附属医院神经内科 襄阳 441000

<sup>6</sup> 中国人民解放军联勤保障部队第 900 医院神经内科 福州 350005)

**[摘要]** 目的/意义 构建中枢神经系统 (central nervous system, CNS) 炎性脱髓鞘专病数据库, 为临床科研服务, 提高基层医疗水平。方法/过程 利用互联网收集病历数据, 经过处理和分析后, 构建 CNS 炎性脱髓鞘专病数据库。运用统计分析、自然语言处理、人工智能影像识别和数据可视化等技术, 整合分析数据库信息。结果/结论 形成标准化的 CNS 炎性脱髓鞘专病数据库大数据信息池, 将临床科研数据可视化, 兼顾患者宣教、专科医生培训、多中心远程会诊功能, 促进医疗科研成果转化, 为未来真实世界的临床研究提供参考依据, 优化诊疗路径, 赋能基层医院。

**[关键词]** 炎性脱髓鞘疾病; 数据库; 大数据平台

**[中图分类号]** R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2024.05.014

## Design and Construction of a Specialized Clinical Research Database for Inflammatory Demyelinating Diseases of the Central Nervous System

WU Lei<sup>1</sup>, WANG Bing<sup>2</sup>, YU Qian<sup>2</sup>, SUN Hui<sup>1</sup>, ZHAO He<sup>1</sup>, GAO Sai<sup>1</sup>, GUO Henna<sup>3</sup>, HUANG Yanning<sup>1</sup>, MENG Zhaoyou<sup>4</sup>, XIAO Lianchen<sup>5</sup>, XU Haizhen<sup>6</sup>, HUANG Dehui<sup>1</sup>

<sup>1</sup> Department of Neurology, First Medical Center of Chinese PLA General Hospital, Beijing 100853, China; <sup>2</sup> Key Laboratory of Digital Technology in Medical Diagnostics of Zhejiang Province, Hangzhou 310030, China; <sup>3</sup> Department of Neurology, Shaanxi Provincial People's Hospital, Xi'an 710068, China; <sup>4</sup> Department of Neurology, Second Affiliated Hospital of Army Medical University, Chongqing 400038, China; <sup>5</sup> Department of Neurology, Affiliated Hospital of Hubei University of Arts and Science, Xiangyang 441000, China; <sup>6</sup> Neurology Department, 900th Hospital of Joint Logistics Support Force of the Chinese People's Liberation Army, Fuzhou 350005, China

**[Abstract]** **Purpose/Significance** To construct a specialized database for inflammatory demyelinating disease of the central nervous system (CNS), so as to contribute to clinical research and improve the diagnostic and treatment capabilities of primary healthcare institutions. **Method/Process** Using the internet to collect medical data, after processing and analysis, the CNS inflammatory demyelinating

**[修回日期]** 2023-10-20

**[作者简介]** 武雷, 博士, 副主任医师, 发表论文 25 篇; 通信作者: 黄德晖。

**[基金项目]** 中国研究型医院学会科研课题 (项目编号: Y2023FH-SJKX09)。

disease database is constructed. Using statistical analysis, natural language processing (NLP), artificial intelligence (AI) image recognition and data visualization and other technologies, the database information is integrated and analyzed. **Result/Conclusion** A standardized big database for CNS inflammatory demyelinating diseases is constructed, which enables visualization of clinical research data, provides patient education and specialist training, and facilitates multi-center teleconsultations. The establishment of a specialized database for the CNS inflammatory demyelinating disease can promote the transformation of medical research achievements, provide references for future real-world clinical research, optimize the process of diagnosis and treatment, and improve the clinical capability of primary healthcare institutions.

[**Keywords**] inflammatory demyelinating disease; database; big data platform

## 1 引言

随着精准医疗、智慧医院理念的兴起,以及信息技术的进步,医院信息化的重点更聚焦于专科专病医疗大数据的研究。目前,中枢神经系统(central nervous system, CNS)炎性脱髓鞘疾病的临床研究相对较少,未形成规范化、个性化诊疗体系,不同层级医院差距较大,基层医生可能对其认知不足,导致诊断延迟,甚至误诊,治疗不充分、预防复发治疗缺失或不精准比较常见。因此,搭建 CNS 炎性脱髓鞘疾病的专病数据库,对整合规范相关临床数据、加强循证医学实践、服务临床科研具有重要意义。患者宣教、基层医院神经免疫亚专科医生教育培训、多中心远程会诊平台等功能也将有助于提高公众和基层医生对 CNS 炎性脱髓鞘疾病的认识,提升此类疾病的整体诊疗水平。

## 2 CNS 炎性脱髓鞘专病数据库构建

### 2.1 专病数据库构建背景

CNS 炎性脱髓鞘疾病是一类主要累及视神经、脑和脊髓,造成神经纤维髓鞘脱失而轴索相对保留的自身免疫性炎症性疾病<sup>[1]</sup>。常见的 CNS 炎性脱髓鞘疾病包括多发性硬化、视神经脊髓炎谱系疾病、髓鞘少突胶质细胞糖蛋白抗体相关疾病、急性播散性脑脊髓炎等<sup>[2-3]</sup>。该类疾病被列入国家第一批罕见病目录,其病因复杂、可累及多部位、临床表现多样,增加了鉴别与诊断的难度,需要结合详细完整的病史资料、神经系统检查、脑脊液检查、神经

影像以及神经电生理检查诊断。近年来 CNS 炎性脱髓鞘发病率逐渐增加并累及青壮年人群,其高复发性、高致残性使患者有一定概率遗留明显的神经功能障碍<sup>[4-6]</sup>,患者需要长期甚至终身服药以预防和减少疾病复发,延缓疾病残障进展。但由于高效的随访系统和集成平台的缺乏、患者主客观因素等影响,患者主动复诊、复查率低,失访率高,带来巨大的复发风险。为此,患者病历资料数据的收集、管理和分析显得尤为重要;同时,海量医疗数据可以为复杂疾病的研究提供新方向,为临床实践提供证据及建议。但目前真实世界治疗场景中产生的数据并未通过严格的数据质量监测、治疗流程控制,造成数据质量参差不齐<sup>[7]</sup>。此外,尽管医疗诊断的电子化进程在加快,但诸如随访记录、影像报告等数据未完全结构化,存在特定关键指标数据少或缺失等问题。临床科研中结构化与非结构化数据共存,信息系统开发面临多模态信息融合方面的挑战。

目前国内在 CNS 炎性脱髓鞘疾病患者队列的建立和临床科研数据的收集管理挖掘等方面还存在诸多问题,特别是缺乏广泛被认可的专病数据库,严重制约了患者管理水平和临床科研产出。因此本文介绍的 CNS 炎性脱髓鞘疾病专病数据库及多中心数据应用平台依托中国人民解放军总医院第一医学中心在神经内科疾病诊治方面的学科优势与专业特色,围绕中国人民解放军总医院第一医学中心及各临床中心、基层医院、患者群体三方构建临床科研数据库及多中心数据应用平台,整合临床研究资源,挖掘临床多模态数据中的潜在价值,对提高 CNS 炎性脱髓鞘疾病认知、优化诊疗路径、推进精准医学临床研究、探索具有中国特色的健康数据管

理方案具有重要意义<sup>[8]</sup>。

## 2.2 专病数据库系统架构设计

系统架构自底向上由基础数据层、数据交换

层、数据存储层、业务服务层和前端门户层 5 部分组成，见图 1。由此构建面向多层面的 CNS 炎性脱髓鞘专病数据库，为该病多变量相互作用的分析提供有效支撑。

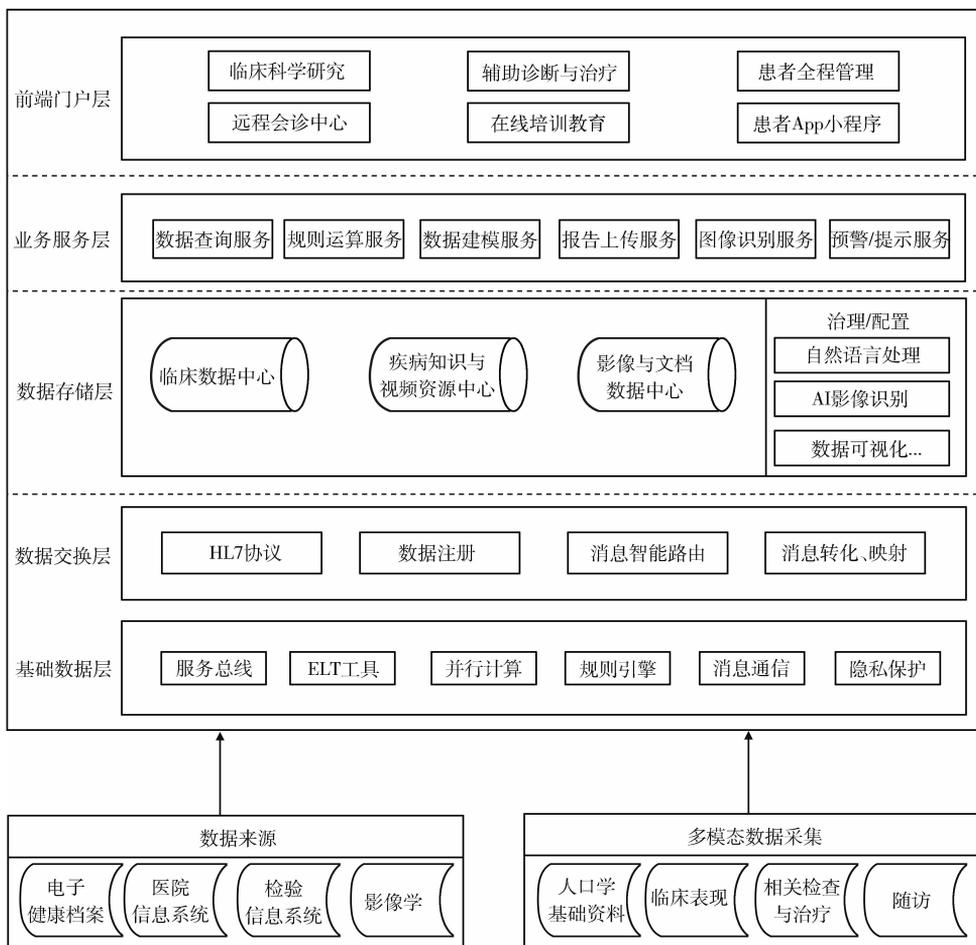


图 1 CNS 炎性脱髓鞘专病数据库系统架构

## 2.3 专病数据库数据结构模型建立

2.3.1 数据来源 纳入 CNS 炎性脱髓鞘疾病，特别是多发性硬化、视神经脊髓炎谱系疾病、髓鞘少突胶质细胞糖蛋白抗体相关疾病这 3 个主要疾病亚型，由神经内科医生收集患者的人口学资料，包括性别、年龄、发病年龄、生活地域、环境，以及疾病史、临床表现、实验室检验（脑脊液常规、生化、抗体、寡克隆区带、病原学检测、细胞因子）、脑电图、磁共振成像（magnetic resonance imaging, MRI）影像学报告、治疗情况（方案、效果）、复查抗体滴度和随访等信息，通过数据库应用平台定

期收集患者的疾病相关资料，安排专人整理汇总。数据采集依托各基层医院、科研单位等系统，通过鉴权认证接口上传临床研究数据，支持多源异构数据由分散的独立系统向平台内科研专病数据库汇集；数据库对包括 Excel 表单、临床数据、电子病历、影像学数据在内的多模态数据进行汇总，丰富 CNS 炎性脱髓鞘疾病的临床科研数据。

2.3.2 数据建模 基于 openEHR 国际标准体系中临床诊疗过程中的客观观察、医疗评估、诊疗计划、执行过程（observation、evaluation、instruction、action, OEIA cycle）等信息，建立 CNS 炎性脱髓鞘专病数据库的标准化信息模型。根据专病诊

疗路径设计数据集，共涉及 4 大类 13 个数据模块基本数据，见表 1。最终以事件时间为基准轴，将数

据拆分成以时间为维度的集合，便于临床医生对患者进行全程管理。

表 1 CNS 炎性脱髓鞘专病数据库部分标准数据集

类型	部分模块名称	部分子模块名称	部分字段名称
患者基础数据	患者	基本信息、体格检查	年龄、性别、民族、身高、体重、腰围、臀围、血压
	家族史	描述、家庭成员情况	与本人关系、是否健在、疾病或诊断
	既往史		描述、发病时间
	疾病史		疾病名称、发病时间
	现病史		描述、发病时间
	用药史		用药情况、过敏史、不良反应
	亚系列疾病		视神经脊髓炎谱系疾病、多发性硬化、髓鞘少突胶质细胞糖蛋白免疫球蛋白 G 抗体相关疾病
	病因	基本病因、诱发因素	感染、疫苗接种、情绪心理、合并其他情况
临床表现	临床表现		肢体乏力、感觉异常、共济失调、眼部症状、精神症状、认知功能下降，其他（失眠、尿频、尿失禁、便秘）
相关检查与治疗	相关检查	实验室检验、影像学检查、生化指标	脑脊液、生物标志物、眼科检查、脑部磁共振成像、脊髓磁共振成像、视神经磁共振成像、血生化、尿生化
	鉴别诊断	临床特征、影像学、抗体	脑部磁共振成像、脊髓磁共振成像、脑脊液寡克隆区带、水通道蛋白 4 抗体
	治疗情况	用药方案、康复治疗、生活指导	剂量、方法、不良反应
随访	随访		随访时间、随访记录

2.3.3 数据治理 医疗数据呈现海量、多模态、多源异构等特点，而病历、影像等非结构化数据在临床科研中逐渐发挥关键性作用，临床科研数据治理及标准的规范化成为构建临床专病库的重点内容。因此，在数据库等平台构建过程中，采用抽取、转换、加载（extraction - transformation - loading, ETL）工具和自然语言处理（natural language processing, NLP）、人工智能（artificial intelligence, AI）影像识别等操作，分层处理临床科研数据。其中，ELT 工具可完成 CNS 炎性脱髓鞘疾病定义及特征、相关药品医嘱等结构化数据的字段内容识别、标准数据字段转化等工作，使部分映射关系分明的数据直接写入临床专病库，通过数据质控将其输出为患者基本信息、就诊记录、用药情况、治疗时间、检查项目名称等多个字段；NLP 算法可处理电子健康档案及医院信息系统中诊疗记录、疾病文献、临床指南、健康档案等长文本，抽取患者手术史、家族史、临床症状、病情进展等碎片化信息，

进行词性标注、句法分析、语义识别等操作并完成结构化处理，再通过数据分析技术挖掘诊断过程、药物影响等方面信息，增强临床科研系统对患者病情的预测能力；由于同一患者很可能存在多份病历及 MRI 影像学报告，平台要从各基层医院数据库中跨系统调取此类非结构化数据，进行多种形式的关联及复杂逻辑计算，AI 影像识别技术可将非结构化报告信息与 MRI 影像报告识别为标准结构化数据，为数据分析及可视化提供便利条件。后结构化处理通用核心流程主要包括明确后结构化处理的文本或文件、机器分词、机器标注、变量抽取、标准映射、投入应用。

2.3.4 数据存储与隐私安全 数据存储平台由临床数据中心、疾病知识与视频资源中心、影像与文档数据中心组成。按模态对不同类型数据分区存储，有利于平台数据的规范化管理。为了使医疗机构间完成高效顺畅的信息传输，专病数据库与基层医院信息系统采用基于 HL7 协议来规范数据交

互<sup>[9-10]</sup>。HL7 在医疗健康数据交换领域已成为标准化卫生信息传输协议，汇集了不同应用软件间的接口标准格式，允许各医疗机构在异构系统间数据交互。

为保障数据安全，数据库采用数据去隐私化的保护措施，划分和管理账号权限<sup>[11]</sup>。数据库管理员可为每位医生建立独立的账号体系并授权和管理。同时，三重数据加密算法/高级加密标准（triple data encryption algorithm/advanced encryption standard, 3DES/AES）等加密技术组成网络传输安全机制，对通信中隐私数据脱敏，保障数据库应用平台组网安全以及医生与患者在平台互动的信息安全<sup>[12-14]</sup>。另外，部分结构化信息可通过简单映射提取关键字段，其他非结构化的病历数据、影像信息等则采取“NLP + AI 影像识别 + 人工校验”的半自动化方式，按照国际疾病分类（第十一次修订版）（International Classification of Diseases, 11th Revision, ICD - 11），使用罕见病编码规则进行术语映射，使数据模型更加标准、规范、安全<sup>[15]</sup>。

### 3 专病数据库平台功能及应用效果

#### 3.1 专病数据库平台功能

CNS 炎性脱髓鞘专病数据库平台提供了一系列创新性服务，便于临床科研高效开展，详细功能模块，见图 2。



科研中心		访视中心		数据应用	
初筛管理	中心管理	访视任务	统计描述	假设检验	
纳排管理	数据探索	访视日历	科研驾驶舱	数据洞察	
患者档案	数据导出	消息管理	数据洞察	智能检索	
研究对象	研究项目	调查问卷			
eCRF设计	数据稽查	访视计划			

图 2 CNS 炎性脱髓鞘专病数据库主要功能模块

医生 App 端为医生提供患者档案管理、纳排管理、日常随访、临床扩展致残量表评分（expanded disability status scale, EDSS）等功能，帮助医生更

方便地管理患者的治疗进展和评估疗效。医生 PC 端提供多中心研究项目创建能力，且拥有强大的电子病例报告表（electronic case report form, eCRF）设计功能，为患者数据的统一管理和分析提供可靠保障。此外，患者 App 端的问卷和随访日历能够帮助患者和医生及时沟通与监测，调整治疗方案，达到更好的治疗效果。

#### 3.2 应用效果实例

CNS 炎性脱髓鞘专病数据库自 2022 年 10 月正式上线运行后，在专病数据库建设方面初步形成应用示范。截至 2023 年 4 月，已有 1 100 多例患者数据集成，全部转化为可供查询分析的结构化数据，数据库建立了常态的增量入库管理机制，便于临床医生持续跟踪患者病情，医生不必在多个系统调阅病历及结果，通过关键词检索功能可更快速地定位到符合研究要求的患者，将患者历次就诊信息以结构化数据呈现，支持明细导出，便于对数据进一步综合分析和挖掘，有助于临床科研资源的统筹协调。

### 4 结语

本文通过建立 CNS 炎性脱髓鞘专病数据库，整合分散在全国各基层医院和不同地域的患者临床数据，通过数据治理、NLP 和 AI 影像识别等关键技术分层处理多模态复杂类型数据，使规范化的标准数据应用于分析及挖掘过程，提高临床科研数据可视化效果，促进医疗科研成果转化；同时，平台搭建多中心会诊板块，为基层医院疑难罕见病的转诊提供“绿色通道”，便于临床医生持续跟踪患者病情，帮助破除真实世界诊疗过程中可能出现的主客观困难，为患者早期预防、及时诊断提供便捷途径，也为患者诊疗提供个性化指导方案；平台推出在线教育板块，一方面为基层临床医生提供 CNS 炎性脱髓鞘疾病规范化诊治规范、经典病例、精品课程、前沿进展等资料，提高对 CNS 炎性脱髓鞘疾病的诊治能力；另一方面为患者推送 CNS 炎性脱髓鞘疾病科普知识，增强患者的疾病认识和随访的依从性。

人工智能在临床的广泛应用将促进临床医疗水

平的不断提高, 缩小不同地区、各级医院间的医疗水平差异, 提高同质化医疗水平, 推动我国医疗事业快速发展<sup>[16]</sup>。因此, 平台后续计划依据我国 CNS 炎性脱髓鞘疾病的流行病学特点, 通过政府、科研单位、医疗机构、患者群体等多中心推广, 建立具有中国区域特色的疾病队列, 强化精准医疗环境下对罕见病患者群体的服务能力, 推动我国医疗卫生事业临床科研持续健康发展。

**利益声明:** 所有作者均声明不存在利益冲突。

## 参考文献

- 1 戚晓昆. 应重视中枢神经系统炎性脱髓鞘病的相关问题 [J]. 中华医学杂志, 2012, 92 (43): 3.
- 2 PACHNER A R, DISANO K, ROYCE D B, et al. Clinical utility of a molecular signature in inflammatory demyelinating disease [J]. *Neurol neuroimmunol neuroinflamm*, 2018, 96 (1): e520.
- 3 HFTBERGER R, LASSMANN H. Inflammatory demyelinating diseases of the central nervous system [J]. *Handbook of clinical neurology*, 2018, 145 (1): 263–283.
- 4 MCGINLEY M P, GOLDSCHMIDT C H, RAE-GRANT A D. Diagnosis and treatment of multiple sclerosis: a review [J]. *Journal of the American medical association*, 2021, 325 (21): 765.
- 5 WEI W, MA D, LI L, et al. Progress in the application of drugs for the treatment of multiple sclerosis [J]. *Front pharmacol*, 2021, 13 (12): 724718.
- 6 BRUSCOLINI A, SACCHETTI M, CAVA M L, et al. Diagnosis and management of neuromyelitis optica spectrum disorders: an update [J]. *Autoimmunity reviews*, 2018, 17 (3): 195–200.
- 7 夏锋, 韦邦福. 精准医疗的理念及其技术体系 [J]. *医学与哲学 (临床决策论坛版)*, 2010, 31 (11): 1–3.
- 8 高景宏, 李明原, 王琳, 等. 健康医疗大数据在精准医疗领域的应用与挑战 [J]. *医学信息学杂志*, 2022, 43 (5): 15–20.
- 9 HL7 International. Health level seven international [EB/OL]. [2023-04-10]. <https://www.hl7.org/about/index.cfm?ref=nav>.
- 10 周芯宇, 陈伟, 吴国全, 等. REST API 设计分析及实证研究 [J]. *软件学报*, 2022, 33 (9): 3271–3296.
- 11 NABIZADEH F, MASROURI S, RAMEZANNEZHAD E, et al. Artificial intelligence in the diagnosis of multiple sclerosis: a systematic review [J]. *Multiple sclerosis and related disorders*, 2022, 59 (2): 103673.
- 12 王煜. 3DES 加密技术的新探讨与改进 [J]. *网络空间安全*, 2016, 7 (8): 30–32.
- 13 XU H Q. Improvement S-Box of AES algorithm based on FPGA [C]. *Sanya: 2017 International Conference on Computer, Electronics and Communication Engineering*, 2017.
- 14 柴绍杰, 张彩珍. AES 加密算法的改进及 FPGA 实现 [J]. *兰州交通大学学报*, 2020, 39 (3): 47–53.
- 15 孙宇航, 刘汉强, 王萍. ICD11 中罕见疾病编码可视化对我国的启示 [J]. *医学与社会*, 2019, 32 (12): 58–60.
- 16 HUANG J, XIN B, WANG X, et al. Multi-parametric MRI phenotype with trustworthy machine learning for differentiating CNS demyelinating diseases [J]. *Journal of translational medicine*, 2021, 19 (1): 377.

## 《医学信息学杂志》开通微信公众号

《医学信息学杂志》微信公众号现已开通, 作者可通过该平台查阅稿件状态; 读者可浏览当期最新内容、过刊等; 同时提供国内外最新医学信息研究动态、发展前沿等, 搭建编者、作者、读者之间沟通、交流的平台。可在微信添加中找到公众号, 输入“医学信息学杂志”进行确认, 也可扫描右侧二维码添加, 敬请关注!



《医学信息学杂志》编辑部