

团 体 标 准

T/ISC XXX—XXXX

医疗行业专病高质量数据集建设指南 第 1 部分：卒中专病

High-quality Datasets for Specialized Disease Construction Guidelines Part 1:
Stroke

（征求意见稿）

XXXX - XX - XX 发布

XXXX - XX - XX 实施

中 国 互 联 网 协 会 发 布

目 次

前 言 II

引 言 III

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

 3.1 个人健康医疗数据 1

 3.2 健康医疗数据 1

 3.3 健康医疗数据控制者 2

 3.4 数据架构 2

 3.5 数据元 2

 3.6 数据标注 2

 3.7 标识符 2

4 符号和缩略语 2

5 总体要求 3

6 数据需求 3

 6.1 应用场景 3

 6.2 数据范围 3

 6.3 数据可使用性 3

 6.4 数据质量需求 3

 6.5 数据安全需求 3

7 数据规划 4

 7.1 数据架构 4

 7.2 实施计划 5

8 数据采集 5

 8.1 数据采集方式 5

 8.2 核心数据元与采集要求 5

9 数据预处理 7

 9.1 数据清洗 8

 9.2 术语与编码标准化 8

 9.3 数据脱敏 8

10 数据标注 8

 10.1 文本问答对构建 8

 10.2 图像类数据标注 11

 10.3 波形数据标注 12

参 考 文 献 13

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是《医疗行业专病高质量数据集建设指南》系列标准的第1部分，《医疗行业专病高质量数据集建设指南》系列标准包含以下部分：

——第1部分：卒中专病

——第2部分：心力衰竭

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国互联网协会提出并归口。

本文件起草单位：首都医科大学、西昌市人民医院、中国信息通信研究院、……

本文件主要起草人：严治、范从华、韩林月、桑新雨、武雅文、冯天宜、闵栋、……

引 言

为规范医疗行业卒中专病高质量数据集建设，包括数据需求、数据规划、数据采集、数据预处理、数据标注等维度的基本要求，制定本文件。

本文件的编制旨在规范卒中专病临床全生命周期的数据治理体系，构建科学、规范、可用的高质量数据集。通过统一数据标准，提升卒中专病临床数据支撑人工智能训练的标准化水平，提高卒中人工智能辅助诊疗、预后预测等智能应用的研发效率与质量。

本文件的编制同步参考《医院卒中中心建设与管理指导原则（试行）（2016版）》、《医院卒中中心建设指引（2024版）》等核心指导文件的业务与信息化要求。

医疗行业专病高质量数据集建设指南 第1部分：卒中专病

1 范围

本文件规定了医疗行业卒中专病高质量数据集的建设指南，涵盖卒中专病数据集建设的数据需求、数据规划、数据采集、数据预处理、数据标注等内容。

本文件适用于指导各级医疗机构、科研院所、信息化企业及相关组织开展卒中专病高质量数据集建设与应用，支持训练和评估人工智能大模型，以赋能区县级医院及以下基层医疗机构提升急性卒中诊疗、卒中预防和康复的能力。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

ISO 8601 日期和时间的表示方法

GB/T 18391.1—2009 信息技术 元数据注册系统（MDR） 第1部分：框架

GB/T 39725—2020 信息安全技术 健康医疗数据安全指南

GB/T 42755—2023 人工智能 面向机器学习的数据标注规程

GB/T 37964—2019 信息安全技术 个人信息去标识化指南

《医院卒中中心建设与管理指导原则（试行）（2016版）》

《医院卒中中心建设指引（2024版）》

《中华人民共和国数据安全法》

《中华人民共和国个人信息保护法》

《人工智能科技伦理审查与服务办法（试行）》

《涉及人的生命科学和医学研究伦理审查办法》

3 术语和定义

下列术语和定义适用于本文件。

3.1

个人健康医疗数据

单独或者与其他信息结合后能够识别特定自然人或者反映特定自然人生理或心理健康的相关电子数据。

[来源：GB/T 39725—2020，3.1]

3.2

健康医疗数据

个人健康医疗数据以及由个人健康医疗数据加工处理之后得到的健康医疗相关电子数据。

[来源：GB/T 39725—2020，3.2]

3.3

健康医疗数据控制者

能够决定健康医疗数据处理目的、方式及范围等的组织或个人，包括提供健康医疗服务的组织、医保机构、政府机构、健康医疗科学研究机构、个体诊所等。

[来源：GB/T 39725—2020，3.5]

3.4

数据架构

对企业业务主要数据类型和来源、逻辑数据资产、物理数据资产和数据管理资源结构和交互的描述。

[来源：ISO TR 21965:2019, 3.2.6]

3.5

数据元

由一组属性规定其定义、标识、表示和允许值的数据单元。

[来源：GB/T 18391.1—2009，3.3.8]

3.6

数据标注

给数据样本指定目标变量和赋值的过程。

[来源：GB/T 42755-2023，3.1]

3.7 标识符

微数据中的一个或多个属性，可以实现对个人信息主体的唯一识别。

[来源：GB/T 37964-2019，3.6]

4 符号和缩略语

下列符号和缩略语适用于本文件。

ASPECTS: Alberta卒中项目早期CT评分 (Alberta Stroke Program Early CT Score)

CT: 计算机断层扫描 (Computed Tomography)

CTP: 计算机断层扫描灌注成像 (Computed Tomography Perfusion)

CTA: 计算机断层扫描血管成像 (Computed Tomography Angiography)

DSA: 数字减影血管造影 (Digital Subtraction Angiography)

DICOM: 医学数字成像和通信 (Digital Imaging and Communications in Medicine)

DNT: 入院至静脉溶栓开始时间 (Door-to-Needle Time)

DPT: 入院至动脉穿刺时间 (Door-to-Puncture Time)

DRT: 入院至血管再通时间 (Door-to-Reperfusion Time)

EMRS: 电子病历系统 (Electronic Medical Record System)

GCS: 格拉斯哥昏迷量表 (Glasgow Coma Scale)

LIS: 实验室信息系统 (Laboratory Information System)

MRI: 磁共振成像 (Magnetic Resonance Imaging)

mRS: 改良Rankin量表 (modified Rankin Scale)

mTICI: 改良脑梗死溶栓分级 (modified Thrombolysis in Cerebral Infarction)

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

PACS: 影像归档和通信系统 (Picture Archiving and Communication System)

TOAST: 急性缺血性卒中病因分型 (Trial of Org 10172 in Acute Stroke Treatment)

5 总体要求

卒中专病高质量数据集建设应按照生命周期有序展开, 包括数据需求、数据规划、数据采集、数据预处理、数据标注等阶段。各阶段主要按以上顺序逐步开展, 同时会对其他阶段进行反馈, 或者会在其他阶段反馈下进行迭代优化。

6 数据需求

6.1 应用场景

卒中专病高质量数据集建设应用场景宜覆盖卒中识别与院前急救、静脉溶栓、机械取栓、二级预防、卒中后认知障碍检查及管理、康复和远期随访等核心场景。

6.2 数据范围

卒中专病高质量数据集建设数据范围宜覆盖患者基础信息、入院与转运信息、既往病史与危险因素、既往用药史、临床症状与发病时间、生命体征、神经功能体征与神经功能量表、检验检查、影像及辅助检查、病因分型、再灌注治疗、急性期治疗和二级预防药物、并发症与不良事件、出院情况、康复与随访等内容。

6.3 数据可使用性

健康医疗数据控制者在构建卒中专病高质量数据集时, 宜确认用于人工智能训练的卒中专病数据是否可通过医疗业务系统或设备获取, 并是否已履行必要的隐私脱敏与伦理审查程序, 确保数据可合法合规使用。

6.4 数据质量需求

健康医疗数据控制者在构建卒中专病高质量数据集时, 宜明确数据完整性、准确性、一致性等质量特性:

a) 完整性: 关键必填项缺失率达到一定比例以上的数据, 宜自动标记为待校验数据, 经核查补全或确认后再纳入数据集;

b) 准确性: 数据内容应真实反映患者临床诊疗实际情况, 与原始诊疗记录保持一致。例如: 数值型指标需严格限制在合理值域内, 疾病与手术操作编码应准确对应标准字典 (如ICD编码), 基于文本抽取的关键信息 (如卒中病因分型) 及影像标注结果 (如核心梗死区勾画), 宜通过规则校验或人工抽检, 确保与原始病历记录及影像客观表现相符;

c) 逻辑一致性: 数据质量规则宜覆盖医学逻辑一致性校验, 例如“静脉溶栓”为“是”时, 溶栓药物名称、溶栓开始时间和给药剂量等数据元不应为空, 且溶栓开始时间应晚于发病时间与到院时间。

6.5 数据安全需求

健康医疗数据控制者在构建卒中专病高质量数据集时, 涉及的伦理审查与知情同意程序宜符合《人工智能科技伦理审查与服务办法 (试行)》、《涉及人的生命科学和医学研究伦理审查办法》相关要求, 并参照GB/T39725-2020的规定进行数据全生命周期安全管理。

7 数据规划

7.1 数据架构

健康医疗数据控制者在构建卒中专病高质量数据集时，宜以核心数据元为基本管理单元，对每项核心数据元同步标识数据产生时间、数据加工层级和数据置信度等级等标签属性，以形成可追溯、可分层、可评价的数据集架构。核心数据元及其标签属性标识示例见表 1。

表 1 卒中专病数据集架构示例

核心数据元	数据值示例	数据产生时间	数据加工层级	数据置信度等级
发病时间	2026-06-22T08:30	2026-06-22T09:00	二次加工数据	C 级
首次血压	168/96 mmHg	2026-06-22T09:10	原始数据	A 级
血糖值	7.8 mmol/L	2026-06-22T09:12	原始数据	A 级
NIHSS 评分	12 分	2026-06-22T09:28	二次加工数据	B 级
ASPECTS 评分	8 分	2026-06-22T10:05	二次加工数据	A 级
TOAST 分型	大动脉粥样硬化型	2026-06-23T10:30	三次加工数据	B 级
脑出血血肿体积	30 cm³	2026-06-22T10:30	二次加工数据	A 级
蛛网膜下腔出血 Fisher 分级	3 级	2026-06-22T11:00	二次加工数据	A 级

7.1.1 数据产生时间

健康医疗数据控制者在构建卒中专病高质量数据集时，应记录每项核心数据元产生时间精确值，时间格式应参照ISO 8601，记录为YYYY-MM-DD THH:MM或YYYY-MM-DD。

7.1.2 数据加工层级

健康医疗数据控制者在构建卒中专病高质量数据集时，应明确标识每项核心数据元的数据加工层级，以数据溯源链路、是否引入前置主观判断为核心区分依据，具体包括以下三级：

d) 原始数据：未经人工加工、解读或计算的原生客观数据，包括原始DICOM影像、监护仪原始生命体征数值、床旁快速检测原始结果、LIS检验原始结果、体格检查原始录像、问诊原始录音；

e) 二次加工数据：以单一原始数据为直接信息来源，由专业人员基于经验判断进行归纳、解读、整理形成的描述性内容，或由系统依据单一原始指标自动公式计算所得数据，包括病程记录中的基于查体和问诊的症状体征的描述、影像报告中对影像征象的描述和结论、单指标衍生计算值、查体现场基于原始观测的NIHSS/GCS/mRS量表评分；

f) 三次加工数据：综合多源数据，整合多维度临床信息，经人工、AI模型或人机协同综合分析、逻辑推演后形成的、具备诊疗指导意义的结论型数据，包括疾病病因分型、溶栓/取栓适配性综合结论、跨维度临床诊断、基于他人病程记录、查体记录等已加工信息间接评定的NIHSS/GCS/mRS量表评分。

7.1.3 数据置信度等级

健康医疗数据控制者在构建卒中专病高质量数据集时，应综合采集方式、加工层级、时间戳，明确标识每项核心数据元的数据置信度等级。具体包括以下四级：

a) A级（高可信可溯源数据）：时间戳完整，精确至时分秒；由设备直接采集、系统调取的原始客观数据或由具备资质的专业人员按照标准化流程评估并留存可追溯依据的数据，如具备相应资质医师评估的NIHSS评分、具备相应资质影像医师确认的ASPECTS评分；

b) B级（规范临床数据）：时间戳完整，允许仅精确到日的非关键病史数据；通过人工实时录入或诊疗行为发生后24小时内完成的人工补录；记录人员身份明确，记录内容符合卒中临床诊疗逻辑，具备常规的可追溯性；

c) C级（主观/近似数据）：时间戳精度受限，仅精确至日；依赖患者/家属回忆口述录入的数据（如发病时间、口述既往史、估算体重）；存在一定主观偏差或近似性的数据；

d) D级（待校验/存疑数据）：关键时间戳缺失；数据采集方式不明或超24小时不合规事后补录；数据存在明显医学逻辑冲突且无法提供原始依据进行校验的数据。针对判定为D级的数据，应实施以下管理要求：

a) 隔离存储：D级数据应与A/B/C级数据实施分区隔离；

b) 建立人工复核校验机制：由具备神经专科资质人员溯源原始病历、影像资料等开展多方校验。经人工溯源校验补齐缺失信息或排除逻辑冲突的数据，视为复核通过，可重新判定置信度等级；否则，维持D级；

c) 使用限制规则：未完成有效复核或者复核后维持D级的数据不应纳入高质量数据集且禁止对外共享，仅允许内部质控复盘、探索性研究使用。

7.2 实施计划

卒中专病高质量数据集构建实施计划应覆盖数据采集、数据预处理和数据标注等阶段，明确实施步骤、责任分工、工具平台、时间安排和交付物，涉及多机构数据协作时，宜明确数据提供、数据处理、数据审核、数据使用和结果反馈等协作机制。

8 数据采集

8.1 数据采集方式

健康医疗数据控制者在采集数据时，应明确标识每项核心数据元的具体采集方式。具体包括以下五种：

a) 院前移动卒中设备直接采集：由救护车车载监护仪、车载心电图设备、便携式血糖仪等设备，在院前急救和转运过程中采集的数据；

b) 系统接口调取：由EMRS、LIS、PACS或集成大数据平台通过接口调取的数据；

c) 人工实时录入：由医护人员在诊疗过程中同步录入的问诊、查体、评分、禁忌证核查等数据；

d) 事后人工补录：由医护人员在诊疗结束后24小时内根据病历、纸质记录、回忆或其他资料补充录入的数据；

e) 可穿戴设备采集：由患者佩戴的物联网腕带等可穿戴设备持续采集的生命体征、定位及活动状态等数据。

8.2 核心数据元与采集要求

8.2.1 患者基础信息

a) 核心数据元：性别、年龄、体重等；

b) 采集要求：健康医疗数据控制者从EMRS或集成大数据平台采集患者基础信息时，宜截取其文本段，分离性别、年龄等关键信息。

8.2.2 入院与转运信息

a) 核心数据元：就诊时间、到达本院途径、首次就诊科室、入院时间、入院时初步诊断、住院科室等；

b) 采集要求：健康医疗数据控制者从EMRS、院前急救系统或集成大数据平台采集入院与转运信息时，宜截取其文本段，分离就诊时间、就诊科室等关键信息。

8.2.3 既往病史与危险因素

a) 核心数据元：既往糖尿病史、高血压病史、房颤病史、冠心病史、卒中病史、颅内出血/肿瘤/巨大动脉瘤史、主动脉弓夹层史、家族史、血液病病史、风湿免疫病史、中毒病史、结构性心脏病史、肺动静脉瘘病史、药物滥用史、大型外科手术史、严重头颅外伤史、吸烟史、饮酒史等；

b) 采集要求：健康医疗数据控制者从实时问诊资料或EMRS或集成大数据平台中采集病史及危险因素信息时，宜截取其文本段，分离既往病史、个人生活习惯等关键信息。

8.2.4 既往用药史

a) 核心数据元：激素使用史（雌激素、孕激素）、近期抗凝/抗血小板药物使用史、降糖药/降压药使用史等；

b) 采集要求：健康医疗数据控制者从实时问诊资料或EMRS或集成大数据平台中采集既往用药史信息时，宜截取其文本段，分离药物名称、药物类型、给药剂量等关键信息。

8.2.5 临床症状与发病时间

a) 核心数据元：主要症状、明确发病时间、醒后卒中/末次正常时间及发现症状时间（发病时间不明）、症状完全缓解时间等；

b) 采集要求：健康医疗数据控制者从实时问诊资料或集成大数据平台中采集问诊信息时，宜截取其文本段，分离主要症状、发病时间等关键信息。

8.2.6 生命体征

a) 核心数据元：血压、心率、呼吸、体温、血氧饱和度等；

b) 采集要求：健康医疗数据控制者从监护仪设备或集成大数据平台中采集生命体征信息时，宜截取其文本段，分离血压、心率等关键信息。

8.2.7 神经功能体征与神经功能量表

a) 核心数据元：意识水平、定向力（时间/地点/人物）、指令配合度、失语情况、构音障碍、凝视、视野、面瘫、上下肢运动、感觉障碍、共济失调、吞咽功能筛查（洼田饮水试验等）等，NIHSS评分、mRS评分、GCS评分等；

b) 采集要求：健康医疗数据控制者从实时查体资料或集成大数据平台中采集体格检查信息时，宜截取其文本段，分离意识水平、意识水平提问/指令等关键神经功能体征信息及NIHSS评分、mRS评分等关键量表信息，或根据神经功能体征直接计算神经功能量表信息。

8.2.8 检验检查

a) 核心数据元：血小板计数、白细胞计数、红细胞计数、血红蛋白、肝肾功能、D-二聚体、同型半胱氨酸、INR、PT、APTT、血糖、血脂、心肌酶谱、电解质、炎症指标（如CRP、PCT）、血气分析、其他危急值结果、红细胞沉降率、抗“O”、ANA、ENA、ANCA、ACL、蛋白C、蛋白S等；

b) 采集要求：健康医疗数据控制者从EMRS或集成大数据平台采集患者检验信息时，宜截取其文本段，分离血小板计数、血糖等关键信息。

8.2.9 影像及辅助检查

a) 核心数据元：颅脑 CT/CTP/CTA、颅脑 MRI、颅脑 DSA、心电图、心脏超声、脑彩超、颈动脉超声等检查结果/报告、ASPECT 评分等；

b) 采集要求：健康医疗数据控制者从 EMRS 或集成大数据平台采集患者影像信息时，宜截取其文本段，分离颅脑 CT/CTP/CTA、颅脑 MRI 等关键信息。

8.2.10 病因分型

a) 核心数据元：缺血性卒中 TOAST 病因分型及其改良版、CISS 分型、OCSP 分型、出血性卒中病因分型、责任血管判定等；

b) 采集要求：健康医疗数据控制者从 EMRS、集成大数据平台或经审核确认的辅助分析结果中采集患者诊断信息时，宜截取其文本段，分离病因分型等关键信息。

8.2.11 再灌注治疗

a) 核心数据元：静脉溶栓治疗，包括但不限于静脉溶栓药物类型、给药剂量、DNT；血管内治疗，包括但不限于血管内治疗类型、DPT、DRT、取栓次数、mTICI 分级、取栓后补救措施等；

b) 采集要求：健康医疗数据控制者从 EMRS 或集成大数据平台采集患者治疗信息时，宜截取其文本段，分离静脉溶栓药物、给药剂量等关键信息。

8.2.12 急性期治疗和二级预防药物

a) 核心数据元：抗血小板药物、抗凝药物、调血脂药物、降糖药物、降压药物等药物类型及剂量；

b) 采集要求：健康医疗数据控制者从 EMRS 或集成大数据平台采集急性期治疗和二级预防药物信息时，宜截取其文本段，分离关键信息。

8.2.13 并发症与不良事件

a) 核心数据元：卒中相关性肺炎、脑水肿与颅内压增高、颅内出血转化、系统性栓塞、心梗/心衰、肺栓塞、压疮、深静脉血栓、癫痫、药物不良反应、介入手术相关并发症等；

b) 采集要求：健康医疗数据控制者从 EMRS 或集成大数据平台采集患者诊疗记录时，宜截取其文本段，分离卒中相关性肺炎、脑水肿等关键信息。

8.2.14 出院情况

a) 核心数据元：出院时间、实际住院天数、出院方式、出院去向、主要诊断名称、出院时神经功能评分、出院时开具的药物种类等；

b) 采集要求：健康医疗数据控制者从 EMRS 或集成大数据平台采集出院情况信息时，宜截取其文本段，分离关键信息。

8.2.15 康复与随访

a) 核心数据元：康复治疗方案、终点/不良事件、死亡原因和日期、卒中复发预警症状及就医行为、随访 mRS 评分、二级预防药物依从性等；

b) 采集要求：随访数据应按 3 个月、6 个月、1 年等固定周期记录，支持长期队列研究，且随访数据需同步采集复发 TIA/卒中、颅内出血、血管再狭窄终点事件，所有随访记录绑定患者唯一标识与随访医师工号。

9 数据预处理

9.1 数据清洗

健康医疗数据控制者应对采集的原始数据进行清洗，包括：

- a) 缺失值处理：针对核心数据元缺失的记录，宜溯源补全，无法补全的记录，不宜纳入高质量数据集；
- b) 异常值处理：针对超出临床合理范围或违背临床时序逻辑的异常值，应溯源核查原始记录，无法核实的异常值标记为低置信度数据，或按预设校验规则予以剔除；
- c) 重复数据处理：识别并清理因跨系统对接重传等原因产生的重复记录（如重复的医嘱、检验明细）；
- d) 冗余数据处理：剔除与卒中专病分析及AI建模无关的冗余数据。

9.2 术语与编码标准化

健康医疗数据控制者应对采集的原始数据进行术语与编码标准化，包括：

- a) 诊断术语：参照国际疾病分类编码ICD-11，统一临床疾病诊断名称；
- b) 量表术语：参照国内外卒中领域通用神经功能评估量表规范及标准缩写，统一神经功能评估量表的名称、缩写等，如美国国立卫生研究院卒中量表（NIHSS）、改良Rankin量表（mRS）、改良脑梗死溶栓血流分级（mTICI）、阿斯伯格样早期CT评分（ASPECTS）等；
- c) 药物与操作术语：参照国家药典通用名规范、手术操作分类编码ICD-9-CM-3，统一卒中相关药物、诊疗操作的标准名称。

9.3 数据脱敏

健康医疗数据控制者应对采集的原始数据进行脱敏处理，包括：

- a) 直接标识符：参照GB/T 37964—2019，对姓名、身份证件号、联系方式等直接标识符进行删除或加密替换处理等；
- b) 准标识符：参照GB/T 37964—2019，对出生日期、地址、住院号等准标识符采取泛化编码、屏蔽处理等；
- c) 影像数据：参照DICOM PS 3.15: Appendix E以及GB/T 39725-2020对DICOM图像及DICOM头文件中的敏感标签及隐私文本信息进行删除或替换处理等。

10 数据标注

10.1 文本问答对构建

健康医疗数据控制者在构建大模型训练所需问答对或指令微调数据时，应针对不同决策场景，提取上述所需核心数据元，生成面向不同场景的问答对。核心决策场景及问答所需数据元、问答输出要求详见表2。

表 2 卒中专病多场景问答对构建

决策场景	问答所需核心数据元	问答输出
------	-----------	------

决策场景	问答所需核心数据元	问答输出
卒中识别与院前急救	1. 生命体征：血压、心率、呼吸、体温、血氧饱和度等； 2. 主要症状； 3. 核心时间：卒中明确发病时间、醒后卒中/末次正常时间及发现症状时间（发病时间不明）等； 4. 检验结果：院前指尖血糖等检查结果； 5. 既往病史：高血压史、糖尿病史、冠心病史等病史及相关疾病控制情况。	1. 卒中可能性预警； 2. 院前急救处置建议； 3. 转运决策建议； 4. 预挂号/绿色通道启动建议； 5. 循证依据：上述判定对应的临床指南依据。
静脉溶栓	1. 既往病史与用药：既往颅内出血史、口服抗凝药等； 2. 核心时间：卒中明确发病时间、醒后卒中/末次正常时间及发现症状时间（发病时间不明）等； 3. 神经功能：NIHSS 评分、发病前 mRS 评分等； 4. 检验结果：血糖等检查结果； 5. 影像结果：颅脑 CT，明确有无颅内出血、大面积低密度梗死灶；颅脑 MRI，判定 DWI/T2-Flair 是否匹配等。	1. 溶栓适配结论（四级）： ① 推荐常规溶栓 ② 相对适用，需谨慎溶栓 ③ 获益风险不明，需进一步评估 ④ 存在禁忌，应禁止溶栓 2. 获益/风险量化分析：对比分析静脉溶栓与常规药物治疗的血管再通概率、神经功能改善预估及风险评估； 3. 循证依据：上述判定对应的临床指南依据； 4. 进阶临床决策建议：明确是否需紧急大血管评估、是否符合桥接取栓指征及溶栓前后用药管控建议。
机械取栓	1. 既往病史与用药：既往颅内出血史、房颤史、抗凝药服用史、溶栓药物、降压药服用情况等； 2. 核心时间：卒中明确发病时间、醒后卒中/末次正常时间及发现症状时间（发病时间不明）等； 3. 神经功能：NIHSS 评分、发病前 mRS 评分等； 4. 检验结果：凝血功能、血小板计数、肌酐清除率等检查结果； 5. 影像结果：CTA/MRA/DSA 大血管成像，明确闭塞血管部位；ASPECT 评分/CTP/MRP 灌注成像、CTP Mismatch 比值，明确梗死核心体积、缺血半暗带范围。	1. 取栓适配结论（四级）： ① 推荐常规取栓 ② 相对适用，需谨慎取栓 ③ 获益风险不明，需暂缓评估 ④ 存在禁忌，应禁止取栓 2. 获益/风险量化分析：对比分析机械取栓与最佳药物治疗的获益/风险评估； 3. 循证依据：上述判定对应的临床指南依据； 4. 进阶临床决策建议：明确取栓方案、围手术期管理方案等。

决策场景	问答所需核心数据元	问答输出
卒中后认知障碍检查及管理	1. 基础信息：患者年龄、身体基础状况等； 2. 核心时间：卒中明确发病时间、醒后卒中/末次正常时间及发现症状时间（发病时间不明）等； 3. 认知功能评估：MMSE、MoCA 等量表评分； 4. 伴随情况：情绪障碍、睡眠障碍、日常生活能力； 5. 检验结果：血糖等检查结果； 6. 影像结果：颅脑 CT 等结果。	1. 认知损害分层判定（无下降/轻度认知障碍/卒中后痴呆）； 2. 后续筛查与检查方案建议； 3. 药物、认知训练等个体化干预方案； 4. 照料风险提示、长期随访计划； 5. 循证依据：上述判定对应的临床指南依据。
早期康复方案	1. 基础信息：患者年龄、身体基础状况、活动耐力。 2. 既往病史：高血压史、糖尿病史、冠心病史等病史及相关疾病控制情况； 3. 核心时间：卒中发病时间、当前病程（溶栓/取栓后时长）； 4. 生命体征：血压、心率、血氧等； 5. 神经功能：发病前 mRS 评分、发病时及当前 NIHSS 评分、具体意识状态、神经功能缺损情况； 6. 影像结果：明确有无出血转化、梗死部位及范围；	1. 康复启动适配结论（四级）： ① 推荐立即启动早期康复 ② 建议延迟 1-3 天启动康复 ③ 生命体征/病情不稳，暂缓评估 ④ 存在禁忌，不建议启动康复 2. 康复方案：制定标准化康复方案，针对肢体、言语、吞咽障碍的个体化康复计划、明确康复强度及调整、康复全程监护重点等。
病房健康宣教	1. 基础信息：患者及家属文化程度、卒中相关健康认知水平； 2. 既往病史、用药及危险因素：高血压史、糖尿病史、高血脂病史及心房颤动等病史及相关疾病控制情况，抗栓、降压、降糖等药物服用情况，患者饮食、作息习惯等。 3. 患者当前病情：溶栓/取栓后、有无并发症； 4. 神经功能。	1. 宣教定位：现阶段的宣教目的及目标家属； 2. 认知与行为干预：明确需纠正的认知误区、需规范的生活及就医行为，并明确对应指南依据； 3. 宣教实施建议：明确宣教重点内容，包括用药指导、饮食护理、并发症识别、体位护理等；明确宣教方式，包括口头讲解、宣传手册、视频科普等；明确宣教频次及家属同步宣教建议。
预防指导	1. 疾病特征：卒中类型（缺血性/出血性）、病因分型； 2. 既往病史、用药及危险因素：既往卒中/短暂性脑缺血发作（TIA）病史、高血压史、糖尿病史、高血脂病史、心房颤动史，抗栓、降压、降糖等药物服用情况，吸烟、	1. 预防目标：明确针对该患者病因分型的核心控制目标； 2. 干预策略：制定个体化干预方案，评估干预预期收益，并明确对应指南依据； 3. 实施建议：明确抗栓、降压、降脂等药物调整方案，制定饮食、运动、戒烟、限酒

决策场景	问答所需核心数据元	问答输出
	饮酒、饮食、运动、作息习惯等； 3. 神经功能：发病前 mRS 评分、神经功能恢复情况； 4. 检验结果：肝肾功能、凝血功能等检查结果。	等生活方式干预措施，明确危险因素定期监测频率、高危人群筛查建议。
远期随访	1. 既往病史与用药：高血压史、糖尿病史、高血脂病史、心房颤动史，抗栓、降压、降糖等药物服用情况等； 2. 核心时间：卒中明确发病时间、醒后卒中/末次正常时间及发现症状时间（发病时间不明）等； 3. 神经功能：出院时及当前 mRS 评分、NIHSS 评分，神经功能恢复情况； 4. 影像结果：近期颅脑 CT/MRI、血管检查等影像学复查结果； 5. 治疗方式：溶栓/取栓/保守治疗； 6. 并发症与不良事件：出血转化、感染、深静脉血栓等。	1. 随访计划：明确随访时间及核心目标； 2. 随访内容：明确神经功能评估、危险因素监测、药物调整等随访内容； 3. 延续性管理建议：明确康复延续建议、卒中复发预警信号告知。

10.2 图像类数据标注

健康医疗数据控制者宜基于医学知识(如临床指南、专家共识、医学教科书),对颅脑 CT/CTP/CTA、颅脑 MRI、颅脑 DSA 等影像数据开展标注,明确标注形式、规则、属性规范等核心要求,标注规范和要求如下。

10.2.1 标注形式

- 健康医疗数据控制者开展图像类数据标注时,标注形式主要分为分割标注和标框标注。
- a) 分割标注:健康医疗数据控制者宜采用轮廓勾勒方式对“脑梗死核心区”、“颅内出血血肿区域”或“脑水肿累及范围”等目标病灶或特定解剖结构边缘进行精准的像素水平勾画,以获取精确的形态、面积或体积信息;
- b) 标框标注:健康医疗数据控制者宜使用矩形边界框对所有可疑动脉瘤、大血管狭窄/闭塞部位、微出血灶等目标病变区域进行标注,若影像中存在多个独立病灶(多发病灶),应按病灶分布逐一进行标注。

10.2.2 标注规则

- 健康医疗数据控制者开展图像类数据标注时,标注规则包括但不限于:
- a) 覆盖全面:病灶类标注应覆盖所有受累区域,包括散在的腔隙性脑梗死或微小病灶;
- b) 解剖精确:结构类标注应按脑血管供血区解剖分级(如 ASPECT 评分对应的脑区划分)精准勾画,确保无遗漏、无界限模糊与过度标注;
- c) 若病灶内包含亚结构(如脑出血血肿内的“岛征”及脑梗死区域内的“出血转化”),应在主病灶

轮廓内进行嵌套式独立勾画。

d) 针对微小病灶（如<3mm 的腔隙灶与血管周围间隙的鉴别）等主观性较强的标注场景，宜建立多级审核与疑难病灶仲裁机制。当标注医师与审核医师产生分歧时，交由高级别专家进行仲裁，或作排除处理。

10.2.3 标注属性规范

健康医疗数据控制者开展图像类数据标注时，标注属性规范包括但不限于：

a) 计量单位：长度单位为厘米（cm）或毫米（mm），面积单位为平方厘米（cm²）或平方毫米（mm²），体积单位为立方厘米（cm³）或立方毫米（mm³），血管狭窄程度使用百分比（%）；

b) 数值记录：应对标注的病灶或结构属性记录客观数值，若属性关联特定的临床量表或诊断分级，宜按照卒中相关临床指南、专家共识等完成等级判定与结果界定。

10.3 波形数据标注

健康医疗数据控制者宜基于医学知识（如临床指南、专家共识、医学教科书），对卒中监护相关的波形数据（如心电图波形、连续血压波形、脑电波形）开展标注，明确标注形式、规则、属性规范等核心要求，标注规范和要求如下。

10.3.1 标注形式

健康医疗数据控制者开展波形类数据标注时，标注形式包括但不限于特征点标注、片段标注、参数计算标注、周期标注、趋势标注、对比标注。

10.3.2 标注规则

健康医疗数据控制者开展波形类数据标注时，标注规则包括但不限于：

a) 多通道协同规则：对于多导联或多通道数据（如 12 导联心电图、多通道脑电图），应明确标注是针对单一通道还是全局通道；在进行段级标注时，各通道的时间轴对齐误差应控制在设备采样率允许的最小精度范围内；

b) 质控要求：不合格或受干扰的异常波形，必须明确标注其干扰类型或异常原因。

10.3.3 标注属性规范

健康医疗数据控制者开展波形类数据标注时，标注属性规范包括但不限于：

a) 计量单位：血压及颅内压使用毫米汞柱（mmHg），心电/脑电波形振幅使用毫伏（mV）或微伏（μV），心率使用次/分（bpm），血氧饱和度（SpO₂）及其他百分比类指标使用百分号（%）；

b) 数值记录：应对标注的波形特征或生理参数记录客观数值，若属性关联特定的诊断分级，宜按照卒中相关临床指南、专家共识等完成等级判定与结果界定。

参 考 文 献

- [1] 《人工智能驱动的卒中研究及管理临床核心通用数据元标准专家共识》
 - [2] DICOM PS3.15:2023 Annex E—Attribute Confidentiality Profiles
-