《基于人工智能的病历质控应用技术规范》框架

标准号：T/XXX [团体代号] XXXX—202X

# 范围

该标准适用于AI病历质控系统开发和应用机构（二级及以上医院、技术供应商）。

规范内容：技术架构、规则库、算法验证、安全伦理

**排除场景**：纸质病历、非结构化数据

# 规范性引用文件

WS 445-2014 电子病历基本数据集

GB/T 39725-2020《信息安全技术 健康医疗数据安全指南》

国卫医发〔2021〕7号《病案管理质量控制指标》

T/CHIA 22-2021 区域医疗质量监管系统功能规范

# 术语定义

**AI病历质控**：基于NLP/ML的自动缺陷检测。

**三层****规则库**：基础规则（《病案管理质量控制指标（2021年版）》）；专科规则（地方病种）；内涵规则（诊疗逻辑）

**召回率（Recall）：**是机器学习领域的评估指标，指模型能正确识别出的目标缺陷数量占实际总缺陷数量的比例。

**缺陷召回率测****试：** 即通过对比 AI 系统与人工（医师）的缺陷识别结果，计算 AI 对病历中特定缺陷的 “漏检率”（召回率越低，漏检越多），从而衡量其检测能力。

**误报率（False Positive Rate, FPR）** ：是机器学习中的评估指标，指模型将正常样本错误识别为 “缺陷样本” 的比例。

**误报率测试：**即通过统计 AI 把无缺陷病历误判为有缺陷的情况，衡量其 “误诊” 概率，与之前的 “召回率测试”（衡量漏检）形成互补。

**全量质控覆盖率**：指 AI 系统实际完成扫描和质控的病历数量，占全院同期产生的总病历数量的比例。

**覆盖率测试：**用于衡量 AI 系统在“全量病历监控”场景下的完整性，即是否能对全院所有病历进行无遗漏的质控筛查，避免出现“部分病历未被检测”的情况。

# 技术架构（分层设计）

|  |  |
| --- | --- |
| **模块** | **核心要求** |
| 数据接入层 | 支持HL7 FHIR/CDA，兼容医院现有EMR系统 |
| 规则引擎 | 可视化编辑+热更新（**广西实践：动态扩展60+专科规则**） |
| AI分析层 | 模型版本管理+迭代日志 |
| 应用层 | 红/黄/蓝三级实时预警（阻断提交→提示修正→通过） |

# 应用流程

图示

AI 生成的内容可能不正确。

# 规则库建设

## 规则库应采用模块化分层架构，支持动态扩展与热更新，确保质控规则与临床实践同步演进。

## 规则分类体系

1. 基础规则应基于《病案管理质量控制指标（2021年版）》制定。必须覆盖《病案管理质量控制指标（2021年版）》发布的27项核心质控指标。
2. 专科规则应覆盖三甲医院常见病种库质控需求。

——心脑血管疾病

——肿瘤疾病

——围手术期管理

1. 内涵规则的临床逻辑校验应满足：

——诊断与检查项目的关联性

——治疗方案与诊断的一致性

——手术禁忌症核查

1. 内涵规则的知识来源至少包含：

——最新临床指南

——医院专家共识（需标注共识编号）

——循证医学证据

## ​规则管理规范

1. 规则标识符格式宜采用「规则类型\_专科代码\_版本号」的结构。
2. 规则库应支持热更新功能，紧急规则可实时生效。规则库宜设置3个月更新周期，新增或修订内容需经医院审核。
3. 规则管理记录每条规则的创建时间、修订历史及责任医师。重大规则变更应附临床验证报告。
4. 新增规则需通过历史病历回溯测试。样本量宜≥500份，准确率宜≥85%方可上线。
5. 规则库管理周期如下：

图表

AI 生成的内容可能不正确。

# 核心算法性能检测要求

1. 缺陷召回率测试宜采用AI系统和≥3名单数主任医师的双盲测试。宜抽取≥1000份含标注意义缺陷的病历。缺陷召回率测试结果阈值宜≥90%。

注1：缺陷召回率＝（检出标注意义缺陷÷实际存在标注意义缺陷）×100%。

注2：AI检出缺陷召回率值与主任医师检出值均≥90%。当AI检出缺陷召回率比值接近主任医师检出比值，说明AI对缺陷的识别能力接近医师。

1. 误报率测试宜统计AI误判为缺陷的正常病历比例。误报率测试结果阈值宜≤8%。

注：误报率=（AI误判为缺陷的正常病历数÷正常病历总数）×100%。

1. 全量质控覆盖率测试宜统计系统实际扫描病历数和全院总病历数，通过比值计算结果。全量质控覆盖率测试结果阈值宜≥60%。

注：全量质控覆盖率=（实际扫描病历数÷全院总病历数）×100%。

1. 响应延迟宜采用压力测试，响应时间阈值宜≤5s。

注：并发值宜根据业务峰值计算确定，必要时宜结合系统资源等维度修正：

基于业务峰值的计算公式：并发值=单位时间峰值业务量×单业务处理耗时×缓冲系数

基于系统资源的修正公式：并发值=（服务器CPU核数×单核心并发能力）×资源利用率阈值

示例：基于医院规模的并发值参考阈值示例：

|  |  |  |  |
| --- | --- | --- | --- |
| 医院类型 | 日门诊量 | 建议基础并发值 | 峰值并发值（含缓冲） |
| 社区卫生服务中心 | ≤500人 | 5-10 | 15-20 |
| 二级综合医院 | 500-2000人 | 10-20 | 30-50 |
| 三甲专科医院 | 2000-5000人 | 20-40 | 50-80 |
| 三甲综合医院 | ≥5000人 | 40-80 | 100-150 |
| **说明：**  1.峰值并发值=基础并发值×2至3倍（如三甲医院高峰期可能同时有100名医生上传病历）。  2.教学医院、急诊占比高的医院需在此基础上上浮 30%-50%。 | | | |

1. ​系统宜每30分钟更新展示核心指标动态。召回率连续3天下降≥5%时应触发告警。
2. ​由医院质控科牵头，宜每个季度抽取1%病历人工复核。

# 安全技术规范

## 数据安全

1. 数据安全应符合GB/T 39725-2020《信息安全技术 健康医疗数据安全指南》相关要求。
2. 数据源仅允许从医院内网EMR系统获取。原始病历数据缓存应保留≤24小时，任务完成后应自动销毁。
3. 严禁原始病历数据离开医院私有云环境。

## 审计机制

1. 操作日志应按以下期限规则保留：

——规则触发详情保留应≥5年；

——医生修正记录应永久保留；

——系统误报分析保留应≥3年。

1. 每年由省级卫健部门指定机构开展安全评估。

# 权责要求

1. 质控结果需经执业医师签名确认后方可生效。系统界面应强制做出「本结果仅供参考，法律责任由签名医师承担」的提示。
2. 《病历使用授权书》中应列有「您的匿名病历数据将用于AI质控系统改进，有权随时通过医院公众号撤回授权」的条款。

参考文献

附录

**附录A（规范性）**：规则库与WS 445字段映射表

**附录B（资料性）**：广西地方病种质控规则示例