**“****辽宁省中医智慧医疗大数据标准体系建设规范”**

**地方标准修订项目**

**标准编制说明**

**一、工作简介**

**（一）任务来源**

“辽宁省中医智慧医疗大数据标准体系建设规范”修订项目，辽宁省卫生计生委于2021年6月立项，计划编号：2021260。本标准由辽宁省卫生计生委提出，由辽宁省卫生计生委归口，由辽宁中医药大学附属医院进行项目修订。

本标准负责起草单位：辽宁中医药大学附属医院

本部分主要起草人：孔德昭、孙凤磊、郝秀炜。

本部分参加起草人：王科雯、罗文晔、李明慧、杨伟、李可大、张春杰。

**（二）制定标准的必要性和意义**

中医智慧医疗大数据标准化体系建设的作用，是有针对性地采集、存储海量医疗数据，并且进行标准化处理，让医疗数据在聚合、分析后，能够驱动临床医学、精准医学等实践应用。

目前，辽宁省尚没有中医智慧医疗大数据标准化体系建设的建设指南。因此，辽宁中医药大学附属医院组织专家编写《辽宁省中医智慧医疗大数据标准化体系建设规范》（简称《建设规范》），旨在为中医及中西医结合医院建设中医智慧医疗大数据标准化体系建设提供规范和指导，推进我省中医医疗大数据的应用发展。

**（三）起草单位**

辽宁中医药大学附属医院

**（四）协作单位**

辽宁中医药大学、中国中医科学院、杭州深睿博联科技有限公司、东软集团股份有限公司、中国医科大学附属医院、哈尔滨医科大学附属第一医院、北京中医药大学、大连理工大学。

**（五）主要起草人及其所做的工作**

**表 1 辽宁省中医智慧医疗大数据标准体系建设规范编写小组成员与分工表**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 姓名 | 性别 | 年龄 | 职称/职务 | 工 作 单 位 | 任 务 分 工 |
| 孔德昭 | 女 | 36 | 课题组总负责人 | 辽宁中医药大学附属医院 | 标准建设规范编写总负责人 |
| 孙凤磊 | 男 | 38 | 课题组研究员 | 杭州深睿博联科技有限公司 | 协助编写标准建设规范 |
| 郝秀炜 | 男 | 25 | 课题组研究员 | 辽宁中医药大学 | 协助编写标准建设规范，专家问卷发放与回收 |
| 王科雯 | 女 | 23 | 课题组研究员 | 辽宁中医药大学 | 标准建设前期准备工作，专家问卷发放与回收 |
| 罗文晔 | 女 | 23 | 课题组研究员 | 辽宁中医药大学 | 标准建设前期准备工作，专家问卷发放与回收 |
| 李明慧 | 女 | 25 | 课题组研究员 | 辽宁中医药大学 | 标准建设前期准备工作，专家问卷发放与回收 |

**（六）主要工作过程**

本标准的起草经历了如下几个阶段：

1.2021年11月至2022年5月，由辽宁中医药大学附属医院中医药循证医学研究中心牵头，成立标准研究小组，前期通过在知网、万方、维普等数据库以“数据标准”、“中医临床”、“体系建设”等检索词查阅相关文献，收集整理相关国家标准、行业标准。由研究小组结合查阅收集的相关资料，拟定《辽宁省中医智慧医疗大数据标准化体系建设规范》各维度及其内容条目的初稿。为了更好体现辽宁省的地域性特点，电话咨询辽宁省临床医学，信息管理，科研教学方面的权威专家，对各维度及其相应的条目进行优化。

2. 依据《辽宁省中医智慧医疗大数据标准化体系建设规范》初稿编写调查问卷。2022年12月向52位行业专家发放第一轮调查问卷。

3.回收整理问卷内容、专家意见，经小组成员研讨后，对初稿进行完善。2023年2月向52位行业专家发放第二轮调查问卷，并附上第一轮问卷修改内容，收集整理问卷内容、专家的意见；

4.将方案公开，倾听专家的意见，进行充分的学术交流，将方案存在的问题进行充分沟通，确定最终地方标准方案。

**二、标准编制原则和确定地方标准主要内容**

（一）标准编制原则

本文件主要针对中医智慧医疗大数据标准化体系建设中的数据安全、数据采集数据汇聚、数据标准统一、数据治理、数据质检等技术内容做出了详细规定。

本文件适用于各级中医及中西医结合医院建设中医智慧医疗大数据标准化体系。本文件与《“健康中国2030”规划纲要》（2016年中共中央、国务院印发）、《关于促进和规范健康医疗大数据应用发展的指导意见》（2016年国务院办公厅印发）、《关于印发促进大数据发展行动纲要的通知》（2015年国务院印发）、《辽宁省新一代人工智能发展规划》（辽政发〔2017〕57号）等相关的行业法律法规、发展规划纲要有较好的配套性，对中医智慧医疗大数据标准化体系建设进行标准化管理，极大促进了中医药行业发展。

（二）地方标准主要内容

本次地方标准确定了中医智慧医疗大数据标准化体系建设的数据安全、数据采集与汇聚、数据标准统一、数据治理、数据质检，为工作组工作内容的重中之重，因此，对于这部分问题，我们将其内容进行整理，具体内容如下：

**1数据安全：规定了中医智慧医疗大数据标准化体系的等级保护、部署模式、硬件安全、链路安全、数据访问和储存。**

1.1等级保护

中医智慧医疗大数据标准化体系建设应达到信息安全三级等保的要求。（具体要求按照物理环境、计算环境、区域边界、通信网络进行分等级安全防护建设，同时，在此基础上还需要建设集中的安全管理中心，对部署在物理环境、计算环境、区域边界、通信网络上的安全策略与安全机制实现集中管理）。

1.2部署模式

中医智慧医疗大数据标准化体系的数据部署，建议采用院内网+VPN模式，外网访问医院内部网络需使用VPN，对其中部分重要数据的访问建议实施内外网物理隔离，保障医院数据安全可控。

1.3硬件安全

为确保医疗大数据数据安全，所有相关的硬件设备应部署在医院中心机房， 统一纳入医院固定资产管理，包括硬件日常巡检、维护等。同时，医院应对硬件具有访问、维护等权限，所有安全设备的syslog 按照医院要求，输出并记录到医院内部日志服务器。

1.4链路安全

为保证链路的私密性和安全性，应布置防火墙、流量控制器、堡垒机、大数据审计设备等安全设备。同时，防火墙应部署在医院内网，保证VPN到医院内网之间的数据安全和访问控制。

1.5数据访问

数据访问应做到全面可控制，不同用户、不同业务系统的运营维修人员或系统管理员的权限不同，看到的业务系统的数据亦不同。

1.6数据储存

数据储存应采用加密存储，并做到先加密再存储，并对数据库的数据进行再备份，保证数据储存的安全。

**2.数据采集与汇聚：规定了数据采集的范围，采集的方式，数据采集后的脱敏与清洗，数据的验收标准和构建主索引系统（EMPI）。**

2.1数据采集范围

数据采集范围包括但不限于：HIS、EMR、LIS、PACS等系统的原始数据。

2.2数据拓展

中医智慧医疗大数据标准化体系应具备扩展对接院外数据的能力，院外数据包括但不限于：患者可穿戴设备数据、气象学数据、环境学数据、基因数据、诊疗指南数据、文献库、院外诊疗、随访数据、量表类数据等。

2.3数据采集方式

数据的采集可根据其实际情况选择一次性融合或二次提取融合，数据的采集方式包括但不限于：数据同步、数据备份、数据镜像、ETL工具和增量抽取等方式。需注意与院内原有系统的数据融合交互方式、安全性等。

2.4构建主索引系统（EMPI）

主索引系统的构建应包括但不限于：患者主索引算法配置、唯一标识的产生、匹配和交叉引用管理、标识及基本信息的更新通知等内容。

2.5数据脱敏

采集后的医疗数据应参考国内相关法律法规或者国际较为认可的美国 HIPAA 法案，对18类敏感数据进行脱敏。脱敏的内容包括患者姓名、身份证号码、电话号码等能够唯一标识某一个患者的数据信息。

2.6脱敏操作

所有脱敏的操作均在内部专用服务器上，由统一的工具完成，保证整个脱敏过程的安全。数据脱敏方法包括但不限于：替代、混洗、数值变换、加密、遮挡、空置插入或删除等。

2.7数据清洗

采集后的医疗数据应采用清洗信息化技术做格式、内容上的清洗处理。保障数据的准确性、一致性、可用性。

2.8数据验证

经过脱敏、清洗后的数据应对其进行数据验证。验证部分要明确数据治理方和数据应用方分别采取何种验证方式（推荐采用全量验证、逻辑验证等方法）。

**3.数据标准统一：对医疗核心数据的存在性、完整性、质量及归档做一系列标准规范，使业务人员、技术人员在提到同一个指标、名词、术语的时候有着一致的理解。**

3.1医学术语标准

3.1.1基础数据元变量如患者基本信息、就诊信息、医嘱信息等业务域中标识符号代码应符合GB/T 2261.4-2003、WS445-2014、WS363-2011、WS/T 303-2009的规定，并实时更新最新标准。

1. 3.1.2 中医疾病名称及证候应符合GB/T 15657-2021，GB/T 16751.1-2023，GB/T 16751.2-2021，还可参考《国际疾病分类第十一次修订本（ICD-11）》的规定，并实时更新最新标准。。

3.1.3 针刺腧穴定位、手法及编码应符合GB/T 12346-2021的规定，并实时更新最新标准。

3.1.4 中草药名称及编码应符合ISO 18668-1:2016、GB/T 31774-2015的规定，并实时更新最新标准。

3.1.5 中草药应用中所对应的四气五味、归经主治、毒性及副作用等数据元属性规范应遵循2020版《中华人民共和国药典》、中国中医药出版社出版的《中药学》教材、还可参考名家医案、临床用药或人用经验等规律挖掘或总结等。

3.1.6 常用医学术语参考标准包括但不限于：《常用临床医学名词（2019年版）》、《中医药临床术语集》、《中国中医药学主题词表》、《中医古典语言系统》、《中医药学语言系统》等。

3.1.7 对于除针灸、穴位敷贴、耳穴压豆、中药熏洗、刮痧、火罐之外，尚无参考标准的中医特色诊疗医学术语名称，应尽可能遵循国家已发布的相应医学术语标准，并注明该疗法无国家标准，仅供参考。

3.1.8 现代医学数据变量如西医诊断、手术操作及西药治疗等业务域中数据元取值标准，可参考相应国家标准、国际标准或行业标准及疾病领域指南。

3.2数据分类分级

依据中医健康医疗数据特点、数据的重要程度和风险等级制定数据分类分级标准，应符合GB/T 38327-2019、GB/T 39725-2020的规定。

3.3主数据管理

主数据管理应包括但不限于：医院内部主数据模型管理、数据管理、数据应用和数据服务等。

3.4元数据管理

元数据管理应包括但不限于：原始数据的元数据管理、通用数据模型的元数据管理以及主题数据模型的元数据管理等。

3.5数据模型

3.5.1 数据模型的设计应当遵循先设计物理模型，再设计概念模型，最后设计主题域模型的顺序。

3.5.2 通用数据模型的建立应参照真实世界中医医疗业务场景，中医诊疗临床思维逻辑将中医四诊、中医诊断、中草药处方、针灸治疗等属性进行具化设计，将表述相同类型相同内容的最小颗粒度字段进行具化表达。（最小颗粒度字段具体包括定义每个字段的数据格式、字段长度、值域、内容约束）。

3.5.3 基于辽宁省特色专科专病数据库，可建立对应的专科专病中医数据模型。

**4.数据治理：在数据挖掘的算法指导下，对文字表达不同但含义相对一致的字段信息进行归一，并将以自然语言的方式录入的医疗数据，根据医学语境转化可用于储存、查询、统计、分析和挖掘的数据结构。**

4.1数据归一

4.1.1 归一的处理首先以标准表作为基准表，将医院数据中实际出现的医疗词向基准表上做映射，并对其结果进行人工核查，对无法自动映射的名称进行人工标注说明。人工标注人员应进行统一的培训保证其专业性。

4.1.2 中医证候数据归一过程中，需要建立中医证候同义词图谱，对临床书写不规范的证候进行标准化处理。中医证候同义词图谱应实时更新纳入最新的中医证候研究。

4.2数据结构化

4.2.1 中医的结构化对象数据包括但不限于：一诉五史、中医特色治疗、中医四诊、中医诊断病-证、中药治疗等非结构化数据。可利用自然语言技术解析其中的关键信息，形成结构化数据，最后由数据质量专家人工进行进一步核对。

4.2.2 数据结构化应从多个独立维度对数据依据主题字段进行划分。

4.2.3 对于比较复杂的文本，可基于数据与知识共同驱动的路径，结合最新的自然语言处理的技术，搭建共性框架，并针对个性化需求，有针对性的开展识别。

4.2.4 结构化框架的构建可由专业医疗人员标记和数据挖掘技术两个方面结合来提供解决方案。

4.2.5 结构化框架的评估主要对其结果的准确率、召回率进行评估，可根据评估的结果和需求目标改进结构化算法中存在的不足。

**5.数据质检：验证经处理、生产后的数据与原业务系统数据的一致性、完整性、正确性。**

5.1数据质量验收

数据质量验收应从以下几个方面进行：

1.完整性：主要考察字段数据是否存在异常值、离群值、不符合数据范围值、NULL 等空值。

2.规范性：数据项的数据格式，数据长度，时间范围，数值范围是否正确。

3.标准性：数据项的数据是否与字典一致，如性别的取值应该是男，女，未知。

4.正确性：数据项内容是否正确，如入院记录中月经史不为空时，患者性别不应该为男。

5.一致性：数据项间内容是否一致，如考察病案首页上的患者、手术、诊断、费用等信息是否同 HIS 或其他系统中的信息一致。

6.时序性：流程时间分布合理性，如出院时间需要晚于入院时间。

7.逻辑性：数据质检应对表内、表间、不同访视点进行数值逻辑核查。

8.整合性：相关系统对应数据项目可对照或关联，如医疗费用明细表里的医嘱标识应该在医嘱标示在中西医处方表、药品类医嘱表、草药处方表或者非药品医嘱表里。

9.唯一性：记录的唯一性，患者唯一ID号是确保不同数据表的融合连接的关键字段，且必须记录好患者信息的时间，如对同一患者不同医嘱所对应的医嘱时间。

10.简洁性：数据变量名的表述是否简洁、易于理解。

5.2隐私数据检查

查看是否按照要求进行去隐私化处理，比如患者姓名、地址、身份证号等。还需对多个字段关联成主键导致隐私泄漏的情况进行检查。

5.3数据溯源

构建数据溯源核查模型，数据溯源模型应包含who、when、where、how、which、what、why七个部分的内容。支持展示规范数据、不规范数据和空数据的数据情况，以及异常数据涉及的患者列表和数据值域详情，原始异常值等。支持溯源到异常数据在实际病历文书中的位置，并将异常数据高亮显示。

**三、主要试验（或验证）的分析、综述报告，技术经济论证， 预期经济社会生态效益分析。**

根据辽宁省医疗数据特点，制定辽宁省中医智慧医疗大数据标准化体系建设规范，对中医药基础研究、临床治疗和科技创新发展具有很好的指导意义。形成中医智慧医疗大数据标准化体系建设标准并在行业推广应用，为大数据标准化体系建设提供思路和方法，规范建设中医智慧医疗大数据标准化体系，优化中医医疗资源配置和诊疗服务，为患者提供多样化的就医方式和智慧化的中医医疗服务。

此研究成果可促进中医药数据的多元化采集、主题化汇聚、集约化管理、知识化处理及社会化服务，填补本省内中医医疗数据标准化体系建设的空白，推动中医药数据标准化的学术进步。创新建立“纵向到底”“横向到边”“条块结合”“互联互通”、功能比较完备、标准规范统一、安全可靠的中医智慧医疗大数据体系、拓宽中医药服务领域、扩大中医院服务范围、创新促进中医药医疗数据大健康产业的发展，将中医药的独特优势惠及百姓。

辽宁省中医智慧医疗大数据标准化体系建设以智慧医疗人口健康信息化为基础，健康服务政策为支撑，带动省内中医药健康产业发展新形态，具有显著的社会效益和经济效益。

**四、与现行有关法律、法规和国家标准、行业标准、地方标准的关系**

（一）本标准规范按照 GB/T 1.1-2020《标准化工作导则 第 1 部分： 标准的结构和编写规则》的要求和规定编写标准内容。

（二）本文件与《“健康中国2030”规划纲要》（2016年中共中央、国务院印发）、《关于促进和规范健康医疗大数据应用发展的指导意见》（2016年国务院办公厅印发）、《关于印发促进大数据发展行动纲要的通知》（2015年国务院印发）、《辽宁省新一代人工智能发展规划》（辽政发〔2017〕57号） 等相关的行业法律法规、发展规划纲要有较好的配套性。

**五、征求意见和分歧处理情况。**

2022 年 12 月～2023 年 2 月，广泛征求了辽宁中医药大学、辽宁中医药大学附属医院、杭州深睿博联科技有限公司、东软集团股份有限公司、中国医科大学附属医院、哈尔滨医科大学附属第一医院、中国中医科学院、北京中医药大学、大连理工大学等 52 位专家的意见。依据上述内容编写德尔菲调查问卷，共发出两轮德尔菲专家问卷，第一轮德尔菲专家问卷发送 52 份，收回 52 份，专家提出相应条目修改意见 15 条，采纳条目 47条，新增条目 4 条。 第二轮德尔菲专家问卷发送 52 份，收回 42 份，专家提出相应条目修改意见 6 条，采纳条目 51条。本次德尔菲专家问卷各条目变异系数均＜0.25，Kendall′s W 协调性检验专家具有一致性（P＜0.01），专家意见具有一致性。

**六、推动标准实施的措施建议**

（一）实施建议

《辽宁省中医智慧医疗大数据标准化体系建设规范》以辽宁中医药大学附属医院信息化建设为模版，各级中医及中西医结合医院可根据当地医院实际情况，充分参考《辽宁省中医智慧医疗大数据标准化体系建设规范》，以保证中医医疗大数据标准化体系建设顺利进行，促进中医药医疗数据大健康产业的发展，建议作为推荐性标准实施。

（二）管理措施

中医智慧医疗大数据标准化体系的建设的核心目的是服务于最终用户，包括一线医生、护士、科研人员、管理人员。但由于医疗大数据标准化体系中汇集了大量包含患者敏感信息的医疗数据。如何保障科学、合理、安全的使用，需要医疗机构以及其数据管理部门制定相应的管理流程、管理制度，以便更好在推广实践中落实。

**七、拟作为强制性地方标准的须写明明确的法律法规依据**

无。

**八、其他应说明的事项**

无。