ICS 点击此处添加 ICS 号
CCS 点击此处添加 CCS 号

T

团 体 标 准

T/CSBME XXXX—XXXX

医疗数据可靠性管理技术要求

Technical Requirements for the Reliability Management of Medical Data

征求意见稿

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX－XX－XX 发布 XXXX－XX－XX 实施

中国生物医学工程学会 发 布

目　　次

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》的有关规定进行起草。

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。

本文件由中国生物医学工程学会提出

本文件由中国生物医学工程学会知识产权与标准化工作委员会归口。

本文件主要起草单位：华南理工大学、北方工业大学、广东工业大学、广州金域医学检验中心有限公司、科大讯飞华南有限公司、广州医科大学附属肿瘤医院、佛山市三水区人民医院、佛山市高明区人民医院、广州市天河区林和街道社区卫生服务中心。

本标准主要起草人：蔡宏民、刘晨、陈佳洲、钟熹、廖文雄、庞柳、林垂旭、刘斯、龙迁艳、周国华、邹文毕、张宇、傅俊方、唐丽、匡志明。

引　言

随着医疗信息化进程的加速推进，医疗数据作为支撑临床决策、科研创新和公共卫生管理的基础性战略资源，其可靠性管理已成为保障医疗质量安全的核心环节。然而，医疗数据的收集、存储、传输和使用过程中面临着诸多挑战。数据的错误、丢失、篡改或未经授权的访问都可能导致严重的后果。因此，建立一套完善的医疗数据可靠性管理技术要求，对于确保医疗数据的质量、安全性和可用性具有重要意义。

本文件聚焦医疗数据采集、存储、处理、共享、销毁等环节的技术要求，重点规范数据的完整性、规范性、一致性、安全性等多个核心维度，旨在实现医疗数据“来源可溯、过程可控、质量可信”的管理目标。

本文件的制定参考了国内外相关标准和最佳实践，结合了医疗行业的实际需求和特点，旨在为医疗机构、医疗信息化企业以及相关监管部门提供一套实用、可操作的医疗数据可靠性管理技术要求。通过本标准的实施，期望能够提升医疗数据的管理水平，促进医疗行业的数字化转型，为患者提供更加安全、高效和优质的医疗服务。

医疗数据可靠性管理技术要求

1 范围

本文件规定了医疗数据可靠性管理的术语和定义、全生命周期管理、技术要求、评价方法以及应用场景等内容。

本文件适用于医疗健康行业各类机构对医疗数据的保护与安全管理，包括但不限于医院、基层医疗卫生机构、健康管理机构、医疗信息化企业以及相关监管部门等在医疗数据的采集、存储、传输、处理、共享和销毁等全生命周期过程中的数据可靠性保障与管理。

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。

GB/T 39725—2020 信息安全技术 健康医疗数据安全指南

3 术语和定义

GB/T 39725—2020界定的以及下列术语和定义适用于本文件。

3.1

医疗数据 Medical Data

指在医疗健康领域中，通过各种健康医疗信息系统收集、存储、传输、处理和产生的，涵盖医疗服务、公共卫生、健康管理、医学科研、医疗保障等活动所产生的各类结构化、半结构化和非结构化数据，包括但不限于患者基本信息、病历资料、检验检查结果、医学影像数据、医疗费用信息、健康档案等。

3.2

数据可靠性 Data Reliability

数据可靠性是指数据在全生命周期内（采集、存储、处理、传输、共享和销毁）保持真实、准确、可信且未被篡改的特性。它反映了数据的质量和可信度，确保数据能够真实反映医疗活动的实际状态，为医疗决策提供准确依据。

3.3

数据规范性 Data Standardization

数据规范性是指数据在收集、存储、处理、传输和使用过程中，符合既定的规则、标准、法规和最佳实践要求的程度。

3.4

数据完整性 Data Integrity

数据完整性是指数据在内容和结构上保持完整、准确且未遗漏的特性。它要求数据在全生命周期内保持其原始状态，不丢失、不被篡改，并且符合预定义的格式和规则。

3.5

数据一致性 Data Consistency

数据一致性是指数据在不同系统、不同时间点之间保持逻辑上的一致性和协调性。它要求数据在跨系统共享、同步或更新时，始终保持相同的格式、内容和语义，避免数据冲突或矛盾。

3.6

数据安全性 Data Security

数据安全性是指保护数据免受未经授权的访问、使用、披露、篡改、破坏或丢失的能力，确保数据的保密性、完整性和可用性。

3.7

健康医疗信息系统 Health Information System

以计算机可处理的形式采集、存储、处理、传输、访问、销毁健康医疗数据的系统。

[来源：GB/T 39725—2020，3.6]

3.8

完全公开共享 Completely Public Sharing

数据一旦发布,很难召回,一般通过互联网直接公开发布

[来源：GB/T 39725—2020，3.11]

3.9

受控公开共享 Controlled Public Sharing

通过数据使用协议对数据的使用进行约束

[来源：GB/T 39725—2020，3.12]

3.10

领地公开共享 Enclave Public Sharing

在物理或者虚拟的领地范围内共享,数据不能流出到领地范围外

[来源：GB/T 39725—2020，3.13]

4 总则

4.1 目的与原则

本文件旨在确保医疗数据在全生命周期管理过程中的可靠性，具体包括数据的规范性、完整性、一致性和安全性，从而保障医疗数据的质量，为医疗服务、临床决策、医疗研究和公共卫生管理提供可信依据。医疗数据管理应遵循以下原则：

（1）合法性原则：数据的采集、存储、处理、传输、共享和销毁应符合国家法律法规及相关行业标准。

（2）最小化原则：数据的使用和共享应限制在实现医疗目的所必需的最小范围内。

（3）患者隐私保护原则：在数据管理过程中，应严格保护患者的个人隐私，确保数据不被泄露或滥用。

（4）可追溯性原则：数据的来源、处理过程和使用情况应可追溯，便于数据质量的监控和问题的溯源。

（5）持续改进原则：数据管理应根据技术发展和实际需求，持续优化和改进管理流程和技术措施。

4.2 基本要求

4.2.1 数据质量要求：医疗数据应具备规范性、完整性、一致性和安全性，确保数据能够真实反映医疗活动的实际状态。

4.2.2 数据安全要求：数据在采集、存储、处理、传输和共享过程中，应采取必要的安全措施，防止数据泄露、篡改或丢失。

4.2.3 数据管理要求：应建立完善的医疗数据管理体系，明确数据管理的责任主体、管理流程和技术措施，确保数据管理的规范化和标准化。

4.2.4 技术要求：数据管理应采用先进的技术手段，如数据加密、备份恢复、数据清洗、数据脱敏等，以提升数据的可靠性和安全性。

4.3 责任与分工

4.3.1 医疗机构：负责本机构医疗数据的采集、存储、处理和使用，确保数据的可靠性、完整性和一致性；同时，应建立健全数据管理制度，明确数据管理责任部门和人员。

4.3.2 医疗信息化企业：负责提供符合本标准要求的健康医疗信息系统，确保系统的数据管理功能符合技术要求；同时，应协助医疗机构开展数据管理的技术支持和培训工作。

4.3.3 监管部门：负责对医疗机构和医疗信息化企业的数据管理情况进行监督、检查和指导，确保本标准的有效实施。

4.4 持续改进与监督

4.4.1 医疗机构和医疗信息化企业应定期对数据管理流程和技术措施进行评估和优化，持续提升数据管理能力。

4.4.2 监管部门应建立监督机制，定期对数据管理情况进行检查，对不符合本标准要求的行为进行纠正和处罚。

5 数据全生命周期管理

5.1 数据采集

5.1.1 制定数据采集标准

应制定统一的数据采集标准，包括数据格式、命名规范、编码标准等，确保数据的一致性和准确性。数据所属机构的信息中心应统一下发数据采集标准，医疗机构宜根据标准创建相关数据的镜像库，以便采集相关数据。

5.1.2 优化技术基础设施

应采用多样化的数据采集工具和技术，如ETL工具，支持多源数据整合、数据清洗与预处理，确保数据的完整性和一致性。宜利用现代数据平台，如基于HL7 FHIR标准的平台，简化数据集成，减少错误。应选择可靠的数据库技术，根据数据量和查询需求优化数据库架构。

5.1.3 确认采集人员身份

应对采集数据的终端设备和操作人员进行身份验证，确保数据来源的合法性和真实性。宜采用多因素认证等强认证方式，防止非法设备和人员接入数据采集系统。

5.1.4 加强数据质量控制

在数据采集过程中应进行实时监控和日志记录，及时发现并处理问题。应对采集的数据进行清洗、去重、格式化等预处理，剔除错误和冗余信息。宜建立数据质量评估体系，定期检查和清洗数据。

5.1.5 确保数据安全与合规

应遵循数据最小化原则，只采集必要的数据。应采用数据加密、访问控制、身份鉴别等技术手段，保障数据在采集、传输等环节的安全。应遵守相关法规和伦理要求，确保数据的合法合规采集和使用。

5.1.6 强化人员培训与管理

应对医护人员和数据采集人员进行培训，确保其按照统一规范操作，减少人为误差。应建立专业的数据管理团队，负责数据的采集、整合、清洗等工作。

5.1.7 建立数据溯源与审计机制

应确保数据采集过程的可追溯性，对数据的来源、采集时间等进行监控。宜定期进行数据安全审计，强化防泄漏措施。设立数据随访和定期更新机制，建立系统性地对医疗数据进行长期和持续的跟踪、采集、分析和更新的规范化流程。

5.2 数据存储

5.2.1 数据分类分级

应按照数据类型或使用场景，分析大型医院和基层社区医院等不同医疗机构的数据量增长速度和业务特性，根据数据的敏感性和重要性进行分类分级管理，对不同级别的数据采取相应的存储和安全措施。

分类分级后，数据存储应依级别选不同存储介质与架构。常规数据应用普通硬盘阵列，较敏感数据应采用分布式的加密分片存储，高度敏感数据应考虑专用加密存储设备，并结合安全合规的云存储。对最高级别敏感数据还应采用物理隔离，将存储设备置于专门安全区域，使用独立网络与存储系统。

5.2.2 数据存储设备选择

应选择具备高可靠性、高耐用性和高性能的存储设备，以满足医疗影像等大容量数据的存储需求。同时，确保存储设备能够适应医院的复杂环境，如磁场干扰、辐射等。

5.2.3 数据备份与恢复

应定期进行全量备份和增量备份，将备份数据存储在安全的位置，以防止自然灾害或系统故障导致的数据丢失。应制定数据恢复计划，确保在数据遭受丢失、损坏或篡改等情况下能够及时恢复数据，保障业务的连续性。备份数据应与原始数据具有相同的安全级别，并进行定期验证和恢复演练。应采用一致性算法和校验机制，确保数据在存储和恢复过程中的完整性。应采用硬件冗余（如RAID技术）和分布式存储架构，确保单点故障不会导致数据丢失。

5.2.4 建立数据安全管理职责

应建立数据安全管理的责任部门和人员，明确各工作人员的职责。应定期对医护人员和数据管理人员进行数据安全意识和技能培训，提高其对数据保护的重视程度。

5.2.5 应急响应与灾难恢复

应建立详细的数据泄露或系统故障的应急响应机制，确保在安全事件发生时能够迅速采取措施，减少损失。

5.3 数据传输

5.3.1 建立网络隔离传输机制

根据医疗数据的分级情况，应对不同级别数据的传输网络进行隔离，防止低级别网络中的安全威胁蔓延到高级别网络。应采用物理隔离或逻辑隔离技术，如虚拟专用网络（VPN）、网闸等，保障数据传输的安全性。

5.3.2 采用安全传输协议

应使用安全的传输协议，如 HTTPS、SFTP 等，替代不安全的协议，如 HTTP、FTP 等，确保数据在传输过程中的安全性。协议应经过严格的安全测试和验证，防止协议漏洞被利用导致数据泄露。

5.3.3 数据加密传输

在数据通过网络进行传输时，应采用加密技术对数据进行保护，防止数据被窃取和篡改。对于敏感性较高的数据，如患者隐私数据（如病历、基因信息）等，宜采用AES-256（对称加密）结合RSA-4096算法实现端到端静态与传输加密，并借助同态加密技术支持隐私计算。对于敏感性较弱的数据，如一般医疗业务数据（如设备日志、排班信息），宜采用AES-128算法提升加密效率，并通过传输层安全性协议TLS 1.3保障传输安全，兼顾安全性与系统性能。

5.3.4 错误检测与校正机制

在医疗数据传输过程中，应分层采用不同的错误检测与校正机制以确保数据完整性。 物理层应通过CRC（循环冗余校验）或LDPC（低密度奇偶校验码）检测位级错误。传输层应采用TCP协议的ACK重传机制纠正丢包或乱序问题，并结合前向纠错算法在实时流媒体场景（如远程手术视频）中实现低延迟纠错。应用层通过哈希校验（如SHA-256）验证数据完整性，并在DICOM等医学影像传输协议中嵌入像素级校验码并修复图像数据损坏。同时，建议基于区块链的分布式存储记录数据哈希指纹，确保传输日志不可篡改，满足HIPAA对医疗数据完整性的合规要求。

5.3.5 监控传输过程

应部署入侵检测系统（IDS）和入侵防御系统（IPS），实时监控网络数据传输活动，及时发现并响应异常行为。应通过自动化跟踪系统和全面审计追踪，确保数据传输过程的合规性。

5.4 数据处理

5.4.1 建立安全数据处理环境

应为数据处理提供安全可靠的计算环境，包括服务器、操作系统、数据库管理系统、应用程序等的安全防护。应定期进行安全漏洞扫描和修复，安装杀毒软件和防火墙等安全防护工具，防止恶意软件和网络攻击对数据处理过程造成破坏。

5.4.2 数据脱敏处理

在进行数据处理时，对于敏感数据和不必要暴露的数据，应进行脱敏处理，以降低数据泄露风险。脱敏处理应遵循不可逆、可验证等原则，确保脱敏后的数据无法还原出原始敏感信息，同时保留数据的业务特征和分析价值。

5.4.3 授权审批

对涉及医疗数据的处理操作，如数据查询、统计分析、数据挖掘等，应建立严格的授权审批机制。确保只有经过授权的人员和应用程序才能进行相应的数据处理操作，并记录操作日志，便于事后审计和追溯。

5.5 数据共享

5.5.1 建立数据共享管理制度

确定数据的共享范围（完全公开共享、受控公开共享或领地公开共享），制定明确的数据共享管理细则，包括数据的使用范围、使用目的、使用方的资质审查等。实施共享操作审计，记录数据共享过程中的所有操作，确保数据使用的合规性。

5.5.2 共享授权要求

应对数据共享的合作方进行背景资质审查，签署数据使用协议，明确数据提供方和数据接收方的权利和义务，以及数据共享的范围、方式和用途等。在数据共享前，应对数据接收方进行身份验证和授权审批，确保数据仅在授权范围内被共享和使用。

5.5.3 数据脱敏与匿名化

根据数据共享的目的和接收方的需求，应选择合适的脱敏和匿名化方法，对敏感数据进行脱敏或匿名化处理，确保数据在共享后无法识别出特定个人或敏感信息主体，同时满足数据接收方的业务需求。

5.5.4 安全传输与存储

应采用安全的传输方式和存储介质，保障数据在共享过程中的安全性。在数据传输过程中，进行加密和完整性保护。在数据接收方存储数据时，按照数据提供方的安全要求进行存储和管理，防止数据泄露和滥用。

5.5.5 审计与监控

应部署监控系统，实时监控数据共享操作行为，记录日志并进行安全审计，及时发现和响应异常行为。

5.6 数据销毁

5.6.1 选择合适的销毁方式

应根据医疗数据的存储介质和数据类型，选择合适的销毁方式，如物理销毁（如硬盘消磁、粉碎等）、数据擦除（如使用专业的数据擦除软件对存储介质进行多次覆盖写入）等，确保数据无法被恢复。在销毁过程中，重点关注数据残留风险及备份数据的同步销毁，确保所有副本均被彻底清除。

5.6.2 建立销毁审批与记录流程

应建立数据销毁审批流程，对需要销毁的数据进行审核和批准，确保销毁的数据符合相关法规和政策要求。同时，对数据销毁过程进行记录，包括销毁的数据内容、销毁时间、销毁方式、执行人员等信息，便于事后审计和追溯。

5.6.3 建立销毁监察机制

应针对不同存储介质和设备，建立不可逆的销毁技术及流程，设立销毁监察机制，防止数据销毁阶段出现数据泄露。

5.6.4 专业人员与培训

应确保数据销毁由经过培训的专业人员操作，避免误操作或越权删除。对相关人员进行数据销毁流程和安全意识培训，定期开展销毁演练，确保销毁过程的高效性和安全性。

6 评价方法

6.1 评价目标和范围

明确评价的目标是评估医疗数据在全生命周期内的可靠性，包括数据的完整性、规范性、一致性和安全性。确定评价范围，涵盖数据采集、存储、处理、传输、共享和销毁等环节。

6.2 评价指标体系

6.2.1 数据完整性：应确保数据完整、准确，不存在缺失或错误。应通过数据的完整性指标（数据集中非空值的比例）和准确性指标（与已知的正确数据进行对比，计算错误数据的比例）进行评价验证。
6.2.2 数据规范性：数据在收集、存储、处理、传输和使用过程中，应符合既定的规则、标准、法规和最佳实践要求的程度。应通过数据规范性指标（不符合指定格式的数据数量占总数据数量的比例）进行评价验证。
6.2.3 数据一致性：应确保数据在不同系统、不同时间点之间保持逻辑上的一致性和协调性。应检查数据在不同表或不同记录之间的一致性，并通过数据一致性指标（不一致数据数量占总数据数量的比例）进行评价验证。
6.2.4 数据安全性：应保护数据免受未经授权的访问、使用、披露、篡改、破坏或丢失，应确保数据的保密性、完整性和可用性。

6.3 评价方法

6.3.1 资料查验：查阅数据管理相关文件，如数据安全管理制度、操作规范、技术文档等，评估其是否符合标准要求。
6.3.2 人工核验：通过系统演示、日志检查等方式，验证数据操作记录、安全措施的实施情况。
6.3.3 工具测试：使用数据安全评估工具，对数据加密、脱敏、备份恢复等功能进行测试。
6.3.4 专家评审：邀请医疗数据管理专家，根据评价指标对数据可靠性进行打分，可采用模糊综合评价法或 TOPSIS 方法。

6.4 评价过程

6.4.1 选取数据样本：从医疗数据中选取具有代表性的数据样本，应涵盖不同业务领域、数据类型和敏感程度的数据。
6.4.2 建立分级评价标准：依据本标准规定的分级原则和方法，组织专家对选取的数据样本进行分级评价，确定每个数据样本的预期分级结果。

6.4.3 验证对比：将实际的数据分级情况与专家评估的预期分级结果进行对比，验证数据分级的准确性和合理性。如果存在差异，分析原因并进行调整和优化，确保数据分级符合标准要求。

6.4.4 结果分析：对收集到的数据进行统计分析，使用故障树分析、失效模式分析等工具，识别数据管理中的风险点。根据分析结果，评估医疗数据的可靠性水平，并与相关标准进行对比，判断是否符合要求。

6.4.5 发现问题与改进措施：根据评价结果，识别数据管理中的薄弱环节，提出改进建议，如优化数据存储架构、加强安全措施、提升人员培训等。制定改进计划，并跟踪改进效果。

6.4.6 文件记录与管理：将评价过程和结果进行详细记录，建立可追溯的档案，包括评价报告、改进方案、验证数据等。这些记录应符合医疗数据管理的合规性要求，并便于后续复查。

附 录 A
（资料性）
数据可靠性应用场景

A.1 转诊场景

本标准提出医疗数据的可靠性管理技术是支撑检查结果跨机构互认的基础保障，在转诊协同场景中尤为重要。基于医疗数据全生命周期可靠性管理机制，通过统一的数据采集规范、质控技术标准和互认评估体系，确保检查流程的数字化可溯源、影像数据的完整性可校验。当患者在基层医院完成影像学检查后，上级医院可直接基于加密传输的标准化影像数据进行可靠性评估，结合诊断医生同质化技术水平的双重认证，有效避免重复检查并保障诊断结论的延续性，降低每年不必要的医疗支出。

A.2 多中心医疗科研协同

在多中心医疗科研协同场景中，医疗数据可靠性管理有望为跨机构联合研究提供了重要的技术支撑。传统模式下，临床研究常依赖物理硬盘拷贝传输多中心数据，面临存储介质损坏、病毒侵入风险以及PB级数据处理效率低下等问题。本标准提出的医疗数据可靠性管理技术，在类似平台有望通过严格的数据质量评估、持续的错误监测与修正以及全流程的数据溯源机制，确保各医院参与协同的元数据真实准确、完整一致，从而有力支撑多中心协同机制下基于这些数据构建的模型的可靠性与稳定性，为医疗决策提供坚实可靠的数据基础。

A.3 大批量样本突发处理场景

在疫情监测、集中体检或门诊高峰等特殊时段，检验科常面临短时间内大量样本集中涌入的挑战。本标准提出的医疗数据可靠性管理技术，可有效支撑高负荷下的样本高效处理与数据稳定运行。在数百样本并发处理时，通过并行数据采集、多线程报告生成和压力优化机制，保障数据不丢失、不错配、不延迟。同时，系统应具备良好的设备接口调度与负载均衡能力，支持多台检验设备数据同时上传，结合缓存与异步处理机制，避免通道阻塞，提升突发场景下的响应效率。在样本量激增的情况下，系统仍需自动执行质控规则，对异常结果实时标记和预警，确保检验数据质量不下降。通过全流程的可靠性控制，实现大规模检测的高效、准确与可追溯，为应急响应和群体筛查提供有力支撑。

参 考 文 献

[1] GB/T 39725—2020 信息安全技术-健康医疗数据安全指南。
[2] YY/T 1833.2—2022 人工智能医疗器械 质量要求和评价 第 2 部分：数据集通用要求。
[3] DB 4401/T 304.1—2024 卫生健康数据采集规范 第 1 部分：采集与交换。
[4] DB 4403/T 497—2024 全民健康信息平台医疗服务数据采集规范。
[5] T/JSIA 0004—2024 健康医疗大数据数据质量管理规范。
[6] T/SHIA 014.1—2024 影像检查质量控制及共享互认标准第 1 部分：数据采集。
[7] T/GDWJ 013—2022 广东省医疗健康数据安全分类分级管理技术规范。
[8] 国家卫生健康委员会，《国家健康医疗大数据标准、安全和服务管理办法（试行）》，2018年4月2日。
[9] 中国通信标准化协会，《基于人工智能的多中心医疗数据协同分析平台参考架构》，2022年4月24日。
[10] 中国国际科技促进会，《医疗健康大数据安全要求》，2024年2月1日。
[11] 中国电子技术标准化研究院，《信息安全技术 健康医疗数据隐私保护安全要求》，2023年9月11日。