

医院信息数据质控策略：确保医疗数据的准确性和可靠性

吴玉清

身份证号：445224198612116923

DOI:10.61369/ME.2025040014

摘 要： 阐述医疗数据质量的重要性及多维度概念，介绍医院数据治理体系的架构、实施路径，包括元数据管理等。还提及自然语言处理等技术在数据质控中的作用，以及多种质控模型和机制，强调基于治理框架等策略的成效，展望未来发展。

关 键 词： 医疗数据质量；数据治理；质控策略

Hospital Information Data Quality Control Strategy: to Ensure the Accuracy and Reliability of Medical Data

Wu Yuqing

ID: 445224198612116923

Abstract： This paper highlights the importance and multidimensional nature of medical data quality, outlines the architecture and implementation paths of hospital data governance systems, including metadata management. It also discusses the role of technologies such as natural language processing in data quality control, as well as various quality control models and mechanisms. The paper emphasizes the effectiveness of strategies based on governance frameworks and looks ahead to future developments.

Keywords： medical data quality; data governance; quality control strategies

引言

医疗数据质量至关重要，涵盖准确性、完整性、一致性和时效性等多个维度，是数据治理理论框架的重要组成部分，符合 ISO/TS 22220 等国际医疗标准体系要求。随着医疗行业的不断发展，相关政策法规日益完善，如 2018 年实施的《医疗质量管理办法》强调了医疗质量的重要性，包括医疗数据质量。从诊疗决策到法规合规，医疗数据质量影响深远。不准确的数据可能引发误诊，对治疗效果产生不良影响，同时法规也对患者隐私保护和数据安全提出了明确要求。因此，构建科学有效的医院数据治理体系，采用先进技术和方法保障数据质量成为关键。

一、医疗数据质量的理论基础

（一）数据质量基本概念

数据质量是指数据满足特定需求和应用场景的程度。它涵盖多个维度，其中准确性指数据能正确反映所描述的对象或事件的真实情况，例如医疗记录中的诊断结果应与实际病情相符^[1]。完整性强调数据不应有缺失值或关键信息遗漏，如患者的基本信息和完整的诊疗过程记录都应完整存在。一致性要求在不同数据源或数据集中，相关数据应保持逻辑上的一致，避免出现矛盾信息。时效性则关注数据是否在合适的时间内被记录和更新，以保证其对决策和医疗过程的有效性。这些概念是理解和评估医疗数据质量的基础，同时也是数据治理理论框架的重要组成部分，在

国际医疗标准体系如 ISO/TS 22220 中也有明确的定位和要求。

（二）医疗数据特殊性

医疗数据具有特殊性，其质量至关重要。从诊疗决策依赖性来看，临床数据、影像数据和检验数据是医生进行诊断和制定治疗方案的重要依据^[2]。这些数据的准确性和完整性直接影响诊疗决策的正确性，进而影响患者的健康与安全。例如，不准确的检验数据可能导致误诊，从而对治疗效果产生不利影响。从法规合规性角度，HIPAA、GDPR 等法规对数据治理提出了严格要求。这些法规旨在保护患者的隐私和数据安全，确保医疗数据的合法使用和管理。医疗机构必须遵守相关法规，建立完善的数据质量控制体系，以保障医疗数据的质量。

二、医院数据治理体系构建

（一）治理组织架构设计

医院数据治理体系的组织架构设计至关重要。构建包含领导小组、技术专家组和科室质控员的三层次治理框架，形成全面覆盖的数据治理层级。领导小组负责整体决策与战略规划，为数据治理提供方向指引^[3]。技术专家组凭借专业知识和技能，对数据治理过程中的技术难题进行攻克和指导。科室质控员则深入各个科室，负责本科室数据的质量把控，确保数据从源头的准确性。同时，明确数据治理委员会的职责，使其在协调各方资源、制定标准规范等方面发挥核心作用。建立跨部门协作机制，打破部门壁垒，促进信息流通和协同工作，保障数据治理工作在全院范围内高效、有序开展。

（二）全流程实施路径

医院数据治理体系构建的全流程实施路径需涵盖多方面。从元数据管理入手，明确数据的定义、来源等关键信息，为数据的准确理解和使用奠定基础^[4]。主数据标准化是重要环节，确保如患者基本信息等核心数据在不同系统中的一致性和准确性。同时，要对数据全生命周期进行管控，包括数据的产生、存储、使用、共享和销毁等各个阶段。结合 PDCA 循环，进行需求分析，了解医院各部门对数据的需求；制定合理的方案，明确治理目标和具体措施；在执行过程中严格监控，及时发现和解决问题，形成一个不断优化的闭环管理流程，以提升医院数据治理的效果和质量。

三、智能分析技术应用

（一）数据质量控制技术

1. 智能数据清洗

自然语言处理技术在非结构化电子病历归一化处理中具有重要作用。通过对病历文本的分析，可识别其中的关键信息，并将其转化为统一的格式。这有助于提高数据的一致性和可用性，为后续的数据分析和应用提供良好的基础。同时，基于规则引擎的异常值检测算法能够有效识别数据中的异常情况。该算法依据预先设定的规则，对数据进行扫描和判断，及时发现不符合正常范围或逻辑的数据点。这种智能数据清洗方法能够显著提升医院信息数据的质量，确保医疗数据的准确性和可靠性，为医疗决策提供有力支持^[5]。

2. 动态校验机制

开发实时逻辑校验模块对于医院信息数据质控至关重要。该模块可在数据录入或更新时即时检查逻辑错误，例如检查患者的年龄与所患疾病的常见发病年龄是否相符，检验数据是否符合医学常理等。同时，集成医学知识图谱能进一步提升校验的准确性和全面性。医学知识图谱包含了丰富的医学概念、关系以及临床路径知识^[6]。通过将实时录入的数据与知识图谱中的临床路径进行比对，可以验证临床操作是否合规。例如，在一个手术治疗过程中，可依据知识图谱检查术前准备、手术步骤以及术后护理等

环节是否符合标准流程，从而确保医疗数据在各个环节的准确性和可靠性，为医疗决策提供有力支持。

（二）质量评估模型构建

1. 统计学监控模型

在构建统计学监控模型时，可运用六西格玛方法建立关键质量指标（CQIs）统计过程控制图^[7]。此方法通过对数据的收集、整理和分析，确定关键质量指标，进而构建控制图。控制图能够直观地展示数据的波动情况，及时发现数据异常点。通过对数据的长期监控，可以了解数据质量的稳定性和趋势。同时，六西格玛方法强调过程的改进，通过对异常数据的分析，找出导致数据质量问题的根本原因，采取相应的改进措施，从而不断提高数据质量，确保医疗数据的准确性和可靠性，为医院的决策提供有力支持。

2. 机器学习评估

采用随机森林算法构建多维度质量评估模型，用于智能诊断数据质量缺陷。该算法具有处理高维数据和非线性关系的优势，能有效挖掘数据特征与质量缺陷之间的复杂关联。通过对大量标记数据的学习，模型可以准确识别数据中的异常模式和潜在问题，为质量控制提供决策支持。同时，利用交叉验证等技术评估模型的性能，确保其具有良好的泛化能力和稳定性。通过不断优化模型参数和调整算法结构，提高模型对不同类型数据质量问题的识别准确率，从而更好地服务于医院信息数据的质控工作，确保医疗数据的准确性和可靠性^[8]。

四、全链路质控策略实施

（一）标准化管理策略

1. 术语体系标准化

基于 SNOMED CT 和 LOINC 建立临床术语映射机制是实现术语体系标准化的重要举措。SNOMED CT 作为一种全面的临床术语系统，涵盖了广泛的医学概念和关系，能够准确描述临床情况^[9]。LOINC 则专注于实验室检验和临床观测指标的标准化编码。通过建立两者之间的映射机制，可以使不同系统中的数据在语义层面实现互通。这有助于解决医疗信息系统中因术语不一致导致的数据理解和整合困难问题，提高数据的准确性和可靠性，为全链路质控策略的实施提供坚实的基础，确保医疗数据在不同环节和系统之间能够准确传递和有效利用。

2. 接口标准化

按照 HL7 FHIR 规范构建医疗信息系统互操作接口对于确保数据采集规范性至关重要。HL7 FHIR 提供了一种标准化的方式来表示和交换医疗信息，使得不同的医疗信息系统之间能够更好地进行互操作。通过遵循 HL7 FHIR 规范，可以确保接口能够正确地采集和传输各种医疗数据，包括患者的基本信息、诊断结果、治疗计划等。这有助于减少数据错误和不一致性，提高数据的准确性和可靠性。同时，HL7 FHIR 规范还支持数据的验证和完整性检查，进一步确保了数据的质量。因此，在医疗信息系统的接口标准化过程中，应积极采用 HL7 FHIR 规范，以提升全链路质控

水平，确保医疗数据的准确性和可靠性^[10]。

（二）多层次质控体系

1. 院级质控平台

院级质控平台是多层级质控体系的关键环节。该平台可开发集成数据质量监控仪表盘，其具备强大功能。通过对医院各类信息数据的实时监测，能敏锐捕捉到异常数据。一旦发现异常，系统会自动启动追踪机制，精准定位异常数据的产生源头及相关影响因素。同时，会及时发出预警信号，以便相关工作人员能够迅速做出反应。这不仅提高了数据质控的效率，还能最大程度减少因数据错误而可能导致的医疗风险，确保医疗数据在准确性和可靠性方面达到较高标准，为医院的整体医疗服务质量提供有力保障。

2. 科室闭环管理

建立多层次质控体系下的科室闭环管理对于医院信息数据质控至关重要。首先需建立专科数据质量改进小组，小组成员应包括科室主任、医疗骨干及信息专员等。该小组负责本科室数据质量的全面把控，依据相关标准和规范对数据进行审核与评估。同时，实施基于 DRGs 的病案首页专项质控，针对病案首页中的关键信息，如诊断信息、手术操作信息等进行严格核查。通过与 DRGs 分组标准进行比对，及时发现并纠正可能影响分组准确性的数据错误。科室内部应形成数据收集、审核、反馈及整改的闭环流程，确保每一个数据问题都能得到及时有效的解决，从而提升医院整体信息数据的准确性和可靠性。

（三）安全保障机制

1. 权限管控策略

实施基于 RBAC 模型的细粒度访问控制，即根据用户的角色分配权限。不同角色的医护人员、管理人员等被赋予相应的操作权限，如医生可查看和修改患者病历，而财务人员只能访问费

用相关信息。通过这种方式，严格限制数据访问范围，确保数据安全。同时集成双因素认证技术，除了常规的用户名和密码登录外，增加如手机验证码、指纹识别或面部识别等第二因素认证。这大大增强了登录的安全性，防止因用户名和密码泄露导致的非法访问，有效保障了医院信息数据在权限管控方面的安全，为数据的准确性和可靠性提供基础保障。

2. 审计追踪机制

构建可追溯的数据血缘图谱以及实现全操作链路的区块链存证是关键。数据血缘图谱能清晰呈现数据的来源、处理过程及去向，为审计追踪提供直观依据。通过记录数据在各个环节的变化情况，包括数据的生成、修改、传输等，确保任何数据操作都有迹可循。区块链存证则进一步增强了数据的安全性和不可篡改性。它以分布式账本的形式记录数据操作信息，多个节点共同维护，保证了数据的真实性和完整性。一旦数据出现问题，可以快速定位到具体环节和责任人，从而有效保障医疗数据的准确性和可靠性，为医院的信息管理和医疗决策提供坚实支撑。

五、总结

通过系统总结可知，基于治理框架、智能分析和全链路管控的数据质控策略取得了一定成效。这些策略在确保医疗数据准确性和可靠性方面发挥了重要作用。同时展望未来，人工智能与区块链融合技术在医疗数据质量追溯方面具有广阔发展前景，其有望进一步提升数据质量的管控水平。提出建立区域医疗数据质量联盟的跨机构协作机制也至关重要。这种机制可以促进不同机构之间的合作与交流，整合各方资源，共同致力于提高医疗数据质量，为医疗行业的发展提供更有力的数据支撑，以更好地服务于患者和医疗工作者。

参考文献

- [1] 郑益力. 公立医院医疗大数据治理问题研究——以温州市 E 医院为例 [D]. 上海师范大学, 2021.
- [2] 彭狄狄. 医疗数据治理与协同的演化博弈研究 [D]. 华南理工大学, 2023.
- [3] 邓婷婷. 我国政府数据治理风险的防范策略研究 [D]. 中国矿业大学 (江苏), 2022.
- [4] 张行. 健康医疗智能数据治理系统设计与实现 [D]. 山东大学, 2021.
- [5] 曹洁. 基于时序逻辑的围术期医疗数据治理方法研究及应用 [D]. 中国科学院大学, 2022.
- [6] 邓军增. 医院健康医疗数据治理探讨 [J]. 医学信息学杂志, 2021, 42(8): 14-17.
- [7] 张永芳, 吴永明, 王胜男, 等. 广东省脑死亡质控数据回顾分析及策略探讨 [J]. 中国卫生质量管理, 2022, 29(3): 15-17, 21.
- [8] 龙思哲. 基于数据中台的医院信息系统数据治理方案的探讨 [J]. 当代医学, 2021, 27(29): 193-194.
- [9] 吴龙, 严晓明, 陈秀娟, 等. 医疗数据湖建设及医疗数据治理探索 [J]. 医学信息学杂志, 2022, 43(6): 8-13.
- [10] 刘鸿. 谈对试油数据质控管理的认识 [J]. 信息系统工程, 2021(6): 71-73.