 DOI: https://doi.org/10.12349/iser.v7i4.10050

Application of Big Data Technology in Hospital Diagnosis and Treatment Data Management

Yang Yang

Heji Hospital Affiliated to Changzhi Medical College, Changzhi, Shanxi, 046000, China

Abstract

The healthcare industry has entered a data-driven development phase, with diagnostic and treatment data exhibiting characteristics of large scale, diverse types, rapid growth, and high value density. Big data technology, leveraging capabilities such as massive data processing, multimodal integration, deep mining analysis, and intelligent early warning decision-making, provides innovative solutions for the full lifecycle management of diagnostic and treatment data, demonstrating significant value in multiple aspects. This paper explores the application of big data technology in hospital diagnostic and treatment data management, systematically elucidates its practical value, outlines mainstream technology types and applicable scenarios, and proposes five key application strategies: data governance, clinical assistance, risk control, resource optimization, and research empowerment. These efforts aim to transform diagnostic and treatment data into a core element for enhancing medical service quality, safeguarding public health, and driving high-quality industry development.

Keywords

Big data technology; Diagnostic and treatment data; Data management; Smart hospital; Clinical decision-making

大数据技术在医院诊疗数据管理中的应用

杨洋

长治医学院附属和济医院，中国 · 山西 长治 046000

摘　要

医疗行业进入数据驱动发展阶段，诊疗数据呈现出规模大、类型多、增长快、价值密的特点。大数据技术凭借海量数据处理、多模态融合、深度挖掘分析、智能预警决策等能力，为诊疗数据全生命周期管理提供了全新解决方案，在许多方面展现出显著价值。本文围绕大数据技术在医院诊疗数据管理中的应用展开探讨，系统阐述应用价值，梳理主流技术类型与适用场景，提出数据治理、临床辅助、安全风控、资源优化、科研赋能五大应用策略，让诊疗数据真正成为提升医疗服务水平、保障群众健康、推动行业高质量发展的核心要素。

关键词

大数据技术；诊疗数据；数据管理；智慧医院；临床决策

1 引言

诊疗数据是医院运行的核心资产，也是医疗服务的重要依据。电子病历、检验检查、医学影像、手术麻醉、用药记录、生命体征等信息，共同构成了覆盖诊断、治疗、护理、康复、随访全流程的诊疗数据体系。随着信息化建设不断深入，医院数据量呈指数级增长，数据类型从结构化扩展到文本、影像、音频等非结构化，传统集中式存储与人工处理方式面临效率低、整合难、挖掘浅、风险高、共享弱等多重困境。大数据技术的成熟与普及，为破解这些难题提供了可行路径。通过数据采集、存储、清洗、分析、挖掘、可视化与安全管控，诊疗数据得以有序整合、高效流转、深度利用，既能辅助临床精准决策，也能支撑医院精细化管理，更能推动医学研究与公共卫生服务升级 [1]。

2 大数据技术在医院诊疗数据管理中的应用价值

2.1 提升诊疗质量与安全水平

诊疗数据的全面整合与智能分析，能够为临床提供更完整、更准确的患者信息，减少信息不对称带来的漏诊、误诊与重复检查。大数据系统可实时整合病史、过敏史、检验结果、影像报告、用药记录等信息，形成统一患者视图，辅助医生快速把握病情。智能审方、异常指标预警、并发症风险预测等功能，可提前识别用药冲突、检验危急值、手术高风险等隐患，将医疗干预由事后转向事前、事中，显著降低不良事件发生率，守护患者生命安全 [2]。

【作者简介】杨洋（1991—），中国山西长治人，本科，助理工程师，从事医疗大数据治理与应用研究。

2.2 优化医疗服务流程与效率

传统诊疗流程存在环节多、耗时长、信息孤岛等问题，患者排队久、检查繁、等待长，医护人员重复录入、跨系统查询工作量大。大数据技术打通门诊、住院、检验、检查、药房、收费等系统壁垒，实现数据自动流转、智能调度与流程优化。智能预约、分时就诊、检查自动排程、报告主动推送、床位智能分配等应用，可大幅缩短无效等待时间，提升就诊流畅度。数据驱动的流程再造让诊疗环节更顺畅、资源利用更充分，整体服务效率明显提升，医患双方负担均有效减轻[3]。

3 诊疗数据管理中主流大数据技术类型分析

3.1 多源数据采集与集成技术

医院诊疗数据分散在 HIS、LIS、PACS、EMR、手术麻醉、心电、监护等多个系统中，格式不一、标准不同。多源数据采集与集成技术通过 ETL 工具（如图 1）、接口适配、消息中间件等，实现结构化、半结构化、非结构化数据统一抽取、转换与加载，支持实时与批量采集，确保数据及时、完整、准确汇聚。该技术是数据管理的基础，解决数据分散、孤岛问题，为后续处理提供统一数据源。

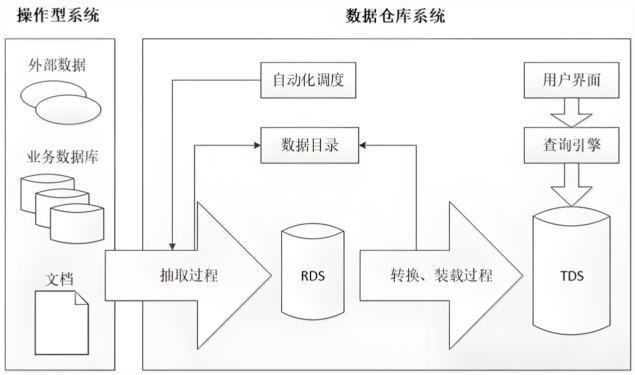


图 1 ETL 工具应用过程示意图

3.2 分布式存储与计算技术

诊疗数据体量巨大、增长迅速，传统数据库难以支撑存储与查询压力。分布式存储采用横向扩展架构，支持海量数据安全存储，保障高可用与高容错。分布式计算框架依托并行处理能力，快速完成大规模数据统计、分析与挖掘，满足秒级查询与实时计算需求。该技术解决数据“存不下、算不动”问题，为大数据应用提供底层支撑[4]。

3.3 数据治理与标准化技术

原始诊疗数据存在缺失、重复、错误、不一致等问题，直接影响应用效果。数据治理与标准化技术包括数据清洗、转换、脱敏、归一、编码统一等环节，按照国际与行业标准对诊断、手术、药品、耗材等进行规范映射，提升数据质量。通过元数据管理、数据血缘追踪、质量监控与异常告警，形成持续优化机制，确保数据真实、可靠、可用、可信。

4 大数据技术在诊疗数据管理中的应用策略与实践案例

4.1 全流程数据治理策略——夯实数据管理底座

全流程数据治理策略以“标准统一、质量可靠、安全可控、可管可用”为目标，覆盖诊疗数据从采集、存储、清洗、标准化、更新到运维的全生命周期（如图 2）。通过建立数据标准、质量规则、安全规范与责任体系，解决数据散乱、错误、不一致、不规范等问题，提升数据完整性、准确性、一致性、时效性与可用性[5]。

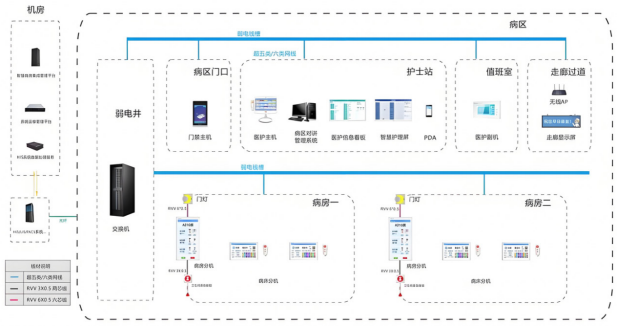


图 2 全流程数据治理平台实施过程

例如，医院上线大数据诊疗数据治理平台，覆盖门诊、住院、检验、检查、手术、药品、耗材全流程数据。平台对接院内 23 个业务系统，每日采集数据超 3 亿条，通过统一标准实现数据规范化治理。针对电子病历文本不规范、诊断编码混乱、检验结果缺失、重复录入等问题，系统自动执行清洗规则，补全缺失字段、修正错误逻辑、合并重复记录、统一 ICD 编码与手术操作编码，数据完整率从 78% 提升至 96%，一致率达 99% 以上。平台建立数据质量监控看板，实时显示异常数据类型、数量、来源科室与责任人，自动推送整改提醒，形成问题发现—整改—复核—闭环机制。针对隐私数据，自动进行脱敏处理，对姓名、证件号、手机号等信息加密或屏蔽，严格按权限开放访问，全程留痕审计。

4.2 临床智能辅助策略——提升精准诊疗能力

临床智能辅助策略以患者为中心，整合多维度诊疗数据，运用大数据分析与 AI 算法，为诊断、治疗、用药、护理、康复提供实时、个性化支持。通过构建患者统一视图、疾病风险预测、诊疗方案推荐、智能审方、异常预警等功能，将数据优势转化为临床能力优势，减轻医护负担、减少人为疏漏、优化决策质量，推动经验诊疗走向精准诊疗。

例如，医院上线大数据临床辅助决策系统，嵌入门诊与住院诊间流程。系统实时整合患者基本信息、病史、过敏史、检验检查、影像报告、用药记录、随访结果等全维度数据，形成完整个人视图。医生开具处方时，系统自动进行用药审核，对剂量异常、过敏禁忌、配伍冲突、重复用药等实时预警，拦截不合理用药。针对慢性病患者，系统基于历史数据生成

风险评估，提示并发症概率与筛查建议。在住院病区，系统对检验检查危急值自动预警，直达责任医生与护士，缩短响应时间。对手术患者，术前自动评估麻醉风险、出血风险、感染风险，生成个性化预案；术中实时采集生命体征数据，动态预警异常波动；术后跟踪康复指标，预测并发症并提前干预。

4.3 医疗安全风控策略——筑牢质量安全防线

医疗安全风控策略以“早发现、早预警、早干预、早改进”为目标，利用大数据实时监控诊疗全流程风险点，包括诊疗行为、用药安全、院感防控、手术安全、耗材使用、费用合规等。通过构建风险模型、设置预警阈值、自动识别异常、推送干预提醒、跟踪整改效果，形成主动式安全管理闭环，将风险控制在萌芽状态。

例如，医院构建大数据医疗安全风控平台，覆盖 16 个风险模块。平台实时采集诊疗行为数据，自动识别过度检查、不合理用药、违规耗材使用等行为并预警。院感模块实时监测住院患者感染相关指标，对高危科室、高危病例实时提醒，协助快速排查与处置。手术安全模块对术前评估不全、术中生命体征异常、术后感染风险升高等自动预警，降低手术不良事件。平台对高警示药品、抗菌药物使用全程监控，未按规范使用自动拦截并提示依据。针对重点人群与重点疾病，系统动态评估风险等级，分级干预。风控平台还具备质量分析功能，定期生成安全报表，定位薄弱环节，推动持续改进。

4.4 资源智能优化策略——提升医院运营效能

资源智能优化策略以“供需平衡、高效利用、降低浪费、提升效益”为目标，运用大数据分析门诊、住院、手术、设备、药品、耗材、人力等资源运行数据，结合历史规律与未来预测，实现智能调度与动态配置。通过智能预约、床位分配、手术排程、设备调度、库存预警、人力排班等功能，缓解资源紧张、减少闲置浪费、缩短患者等待、提升运行效率。

例如，医院上线大数据资源优化运营平台，构建全院运营驾驶舱。门诊模块基于历史流量与季节趋势，智能分配号源、分流时段、调整诊室开放，缩短候诊时间。住院模块自动评估床位使用率、周转率与病情等级，推荐最优分配方案，减少空床与等待。手术模块综合手术类型、时长、医生、设备、恢复室占用等因素，智能排程与冲突检测，提升手术室利用率。设备模块实时监控 CT、MRI 等使用率，智能预约与错峰安排，减少患者等待。药品与耗材模块实时监控消耗，设置安全库存，自动预警与补货，降低缺货与积压。人力模块基于工作量智能推荐医护排班，均衡负荷、保障休息。平台多端同步展示关键指标，异常自动预警。运行后，患者检查等待时间明显缩短，手术室利用率提升，床位周转加快，药品耗材浪费减少，人力成本优化。

4.5 科研数据赋能策略——加速医学研究创新

科研数据赋能策略以“高效提取、规范管理、安全共享、支撑研究”为目标，依托大数据平台构建专病数据库，实现病例快速筛选、数据自动提取、随访信息整合、研究项目管理与多中心协同。通过标准化、结构化、去标识化处理，在保护隐私前提下为临床研究提供高质量数据，降低科研人员数据收集整理负担。

例如，医院建设大数据科研专病库平台，覆盖心血管、肿瘤、神经、呼吸等重点专科。平台对接全院诊疗数据，按研究方案自动提取病例信息，构建标准化数据库。以心血管疾病为例，研究人员设定纳入排除条件，系统几分钟内完成筛选，自动生成基线特征、治疗方案、疗效指标、随访结果数据表。平台支持自定义变量、统计分析与可视化展示，方便直接用于论文与报告。通过隐私计算技术，实现与多家医院数据联合建模，扩大样本量、提升结论代表性。系统对研究项目全流程管理，数据使用留痕、合规可控。平台助力多项课题快速完成，多篇论文发表于高水平期刊，多项成果应用于临床路径优化与指南制定。科研人员从繁重数据工作中解放，专注研究设计与分析。科研赋能策略让数据成为研究核心生产力，推动临床与科研良性循环，为医院高质量发展注入持久动力。

5 结语

在医疗高质量发展与数字化转型背景下，医院应持续推进大数据技术深度应用，完善数据治理体系、强化安全合规、深化场景创新、推动数据要素价值释放。让诊疗数据真正服务于精准诊疗、高效管理、科学研究与群众健康，让技术更有温度、医疗更有品质、管理更有智慧。未来，随着技术不断进步与应用持续深化，大数据将在推动医疗模式变革、提升全民健康水平、助力健康中国建设中发挥更加重要而深远的作用。

参考文献

[1] 刘静. 医院统计工作中大数据技术应用效能分析[J].活力,2026,43(02):109-111.

[2] 王梓云. 基于大数据技术的医院人事管理创新[J].人力资源,2025,(12):143-145.

[3] 蒋京员. 大数据技术支持下的医院运营管理创新策略研究[J].中国管理信息化,2025,28(03):156-158.

[4] 郭亮. 基于大数据的智能辅助诊疗管理系统设计[J].电脑编程技巧与维护,2022,(02):84-85+143.

[5] 李锐. 大数据时代基于安徽省分级诊疗信息系统的探究[J].信息与电脑(理论版),2018,(18):122-123.