

# 基于 ETL 技术的医院质控数据库系统设计与实现

崔金广, 窦一峰, 蒙文涛, 吴秀春

(天津市宝坻区人民医院网络信息中心, 天津 301800)

**摘要:**针对公立医院三级绩效考核数据采集分散、数据统计公式和标准不规范、内容反馈和呈现不直观等问题, 本文构建了基于数据仓库(ETL)技术的医院质控数据库系统, 主要包括时间维度与科室维度下的数据字典、指标分类、数据录入、报告展示、对比输出等内容。依托大数据分析, 通过 Kettle 分别由 HIS、LIS、PACS、EMR、卫生统计、OA 等信息系统提取数据, 将其加工、整理后存储到质控数据库, 并开发设计 B/S 架构的医院质控数据库系统, 对质控数据进行统一管理。最终实现及时准确的提取、加工、整理质控指标数据的目的, 旨在为院领导决策、职能处室和临床科室提供准确有效的数据支持。

**关键词:**医院质控数据库; 医疗质量; 质控指标; 设计

中图分类号: R197

文献标识码: B

DOI: 10.3969/j.issn.1006-1959.2021.04.005

文章编号: 1006-1959(2021)04-0016-06

## Design and Implementation of Hospital Quality Control Database System Based on ETL Technology

CUI Jin-guang, DOU Yi-feng, MENG Wen-tao, WU Xiu-chun

(Network Information Center, Tianjin Baodi Hospital, Tianjin 301800, China)

**Abstract:** In view of the scattered collection of three-level performance appraisal data in public hospitals, irregular data statistical formulas and standards, and unintuitive content feedback and presentation, etc. this article builds a hospital quality control database system based on data warehouse (ETL) technology, which mainly includes data dictionary, index classification, data entry, report display, comparison output, etc. under the time dimension and department dimension. Relying on big data analysis, Kettle extracts data from HIS, LIS, PACS, EMR, health statistics, OA and other information systems respectively, and stores them in the quality control database after processing and sorting. And develop and design a hospital quality control database system with B/S structure, and conduct unified management of quality control data. In the end, the purpose of timely and accurate extraction, processing, and sorting of quality control index data is realized, and it aims to provide accurate and effective data support for hospital leadership decision-making, functional departments and clinical departments.

**Key words:** Hospital quality control database; Medical quality; Quality control index; Design

医疗质量与安全是医疗工作的生命线和医院管理的核心内容, 探索以数据库为核心的医疗大数据项目建设, 运用信息技术构建监管平台, 建立医院质控指标智能监控和现场督察有机结合的管理模式, 通过建立系统化、信息化、精准化的大型公立医疗质量管理与控制指标监测数据库(MQMCIMD), 将有效提升我院的管理效率和综合竞争力, 对于医院长远发展具有十分重要的理论和实践意义<sup>[1]</sup>。目前三级公立医院均面临不同程度的问题, 如数据采集分散, 不能及时有效的给临床科室提供高效、精准的数据支持; 无质控数据库系统、考核方式较为单一、数据统计公式和标准不规范等<sup>[2]</sup>。相关研究指出<sup>[3,4]</sup>, 借鉴目标管理法、平衡记分卡和关键指标法的思想构建多维度动态绩效考核指标体系是必要的, 可完善绩效管理应用, 促进效率质量运行稳步提升。随着大数据、云计算等信息技术的迅猛发展, 借助信息技术服务平台初步建立绩效考核指标评价体系的信息化系统, 对于落实绩效考核结果运用的具体操作方法有重要的意义<sup>[5]</sup>。本文以数据挖掘中的数据提取、转换、加载为主, 在前期对我院基本情况进行调研的基础上, 开发了适合我院院情的质控数据库系统。该

基金项目: 天津市卫生健康科技项目(编号: ZC20095)

作者简介: 崔金广(1985.3-), 男, 天津人, 本科, 高级工程师, 主要从事医疗信息化软件开发工作

通讯作者: 吴秀春(1979.6-), 男, 天津人, 硕士, 高级工程师, 主要从事医院信息化工作

系统可以将医院各科室的医疗质量、收入、运营效率等内容进行统计分析, 分析科室投入与支持, 加强成本管理, 降低医疗资源的浪费, 优化资源配置, 同时通过对各考核指标的联动分析, 可以有效提升科室的工作效率, 从而寻找科室工作中的薄弱环节。从数据整合清洗到统计分析, 从数据可视化报告展示到数据的预测预警, 探析大数据时代的医疗大数据建设, 将能够有效提升我院的管理效率和综合竞争力<sup>[6-8]</sup>。现就系统建设及应用情况总结如下。

### 1 系统设计与实现

质控数据库系统开发完成后部署在可连接医院行政办公网的服务器上, 科室质控员通过浏览器访问质控数据库系统。质控数据库系统包括医疗指标、医技指标、重点疾病、重点手术、择期手术后并发症、急性心力衰竭、医院工作报表等 7 个主要功能页面, 具体数据库框架见图 1。

通过借鉴德国医院透明管理制度与标准委员会提出的集门诊、住院、护理、院感、药事、手术等六大板块为一体的覆盖全院质量与安全的信息监管与反馈平台结构<sup>[9]</sup>, 我院设计了质控数据库系统, 以患者和员工为导向, 以质量安全为保证, 依托门急诊和住院患者为基础, 开展手术、护理、院感和药事等相关指标的归口和制定工作, 以信息化为手段, 针对所有指标实行精准化管控。

**1.1 系统架构** 质控数据库系统采用 B/S 结构, 开发

工具选用 Eclipse, 采用 MVC 设计模式, 后端采用国内开源 Java 框架 JFinal, 前端采用 LayUI 框架+Enjoy 模板引擎+Echarts 图形库+Ajax 等。使用开源 ETL 工具 Kettle 进行数据提取、加工、整理<sup>[10]</sup>。质控

数据库系统分为 UI 层、Control 层、Service 层、Dao 层和实体层。在医院虚拟化资源池划出专用虚拟化服务器, 提供 SQL Server 2008 数据库服务和 Undertow 网站服务, 见图 2。

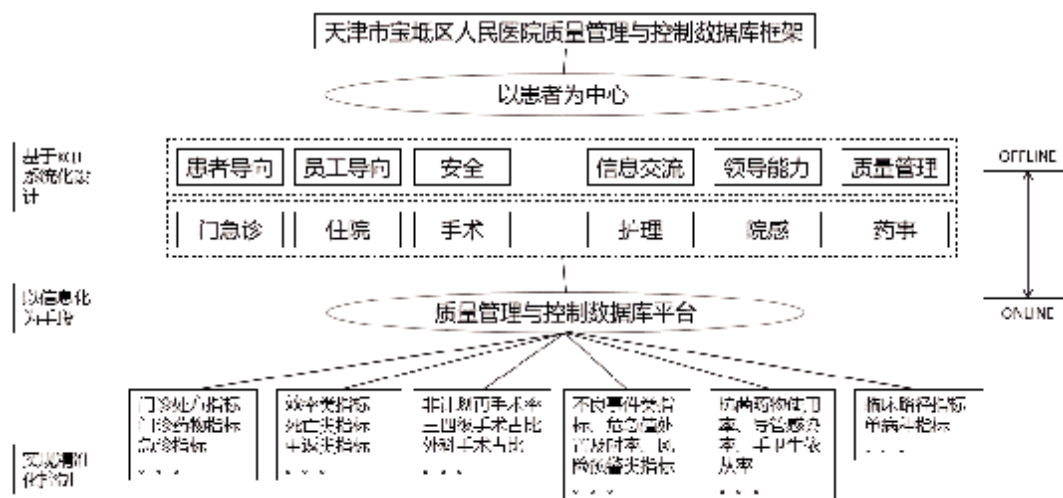


图 1 我院质控数据库框架

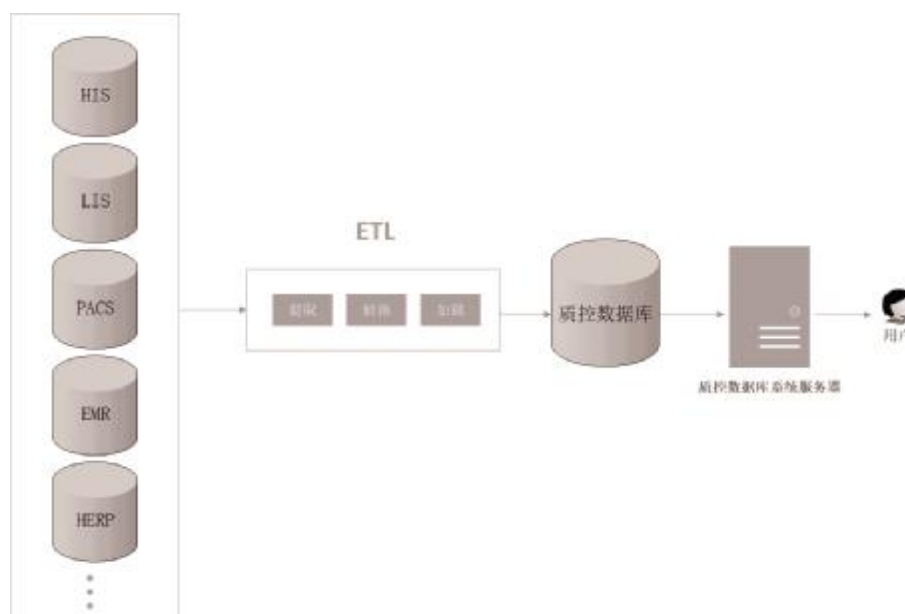


图 2 质控数据库系统技术架构示意图

## 1.2 关键技术

**1.2.1 ETL 技术** ETL 即抽取、转换、加载<sup>[11]</sup>。数据抽取即从源数据源抽取目的数据源需要的数据; 数据转换是将从源数据源获取的数据按照业务需求, 转换成目的数据源要求的形式, 并对错误、不一致的数据进行清洗和加工; 数据加载即将转换后的数据装载到目的数据源。医院质控数据库系统涉及几百项质控指标, 这些指标数据分布在 HIS、LIS、PACS、EMR、HERP、OA、卫生统计等信息系统中。而这些系统又分别采用 Oracle、SQL Server、MySQL 等主流数据库。应用 ETL 技术可以高效、稳定的由上述系统数据库中定时自动抽取、加工、整理医院质控数据库

系统所需的基础数据并存储到质控数据库中。

**1.2.2 ECharts 图形库+Ajax 技术** ECharts 能够在 PC 端和移动设备上流畅运行, 兼容当前绝大部分浏览器, 底层依赖轻量级的 Canvas 库 ZRender, ECharts 提供直观、生动、可交互、可高度个性化定制的数据可视化图表<sup>[12]</sup>。Ajax 是由 Javascript、xml、XMLHttpRequest 组合在一起、能实现异步提交的功能, 是一种创建交互式网页应用的网页开发技术<sup>[13]</sup>。通过 ECharts 和 Ajax 的配合使用, 可在不刷新页面的情况下, 异步请求后台数据, 动态刷新数据分析图。

## 2 系统功能

**2.1 医疗指标** 医疗指标包括院、科两级的医疗质控

指标数据,具体包括诊疗人次、出院人数、入出院诊断符合率、出院者平均住院日、病床使用率、治愈好转率、甲级病历率、平均病床工作日、手术例数、术前平均住院日、三、四级手术占手术总例数的比例、外科手术人次占外科出院人次比例、术前与术后诊断符合率、住院时间超过30天患者人数、急诊死亡人数、急诊抢救人数、急诊抢救成功人数、当天再入院率、2~14天再入院率、15~30天再入院率等37项重点医疗指标。医疗指标数据以月为单位,通过科室和年份两个下拉框选项,可统计查看每月院科两级全年数据明细,见图3。在全院状态下,还可查看每月科室明细。点击数据分析按钮后,弹出数据分析页面,通过科室和数据类型两个下拉框选项,可生成院科两级37项数据指标折线图或柱形图,见图4。

**2.2 重点疾病** 重点疾病包括国家规定的18种重点疾病的院科两级数据,通过年月选择可按时间范围统计查看18种重点疾病的全院指标数据,包括总例数、死亡例数、15d再入院率、30d再入院率、平均住院日、平均住院费用等6项指标。点击疾病名称,可统计查看科级明细数据,见图5。

**2.3 重点手术** 重点手术包括国家规定的18种重点手术和16种恶性肿瘤手术的院科两级数据。通过年月选择可按时间范围统计查看18种重点手术和16种恶性肿瘤手术的全院指标数据,包括总例数、死亡例数、术后非预期再手术、平均住院日、平均住院费用等5项指标。点击手术名称,可统计查看科级明细数据。

**2.4 择期手术后并发症** 择期手术后并发症可按时间范围统计查看肺栓塞、深静脉血栓、败血症、出血

或血肿、伤口裂开、猝死、呼吸衰竭、骨折、生理/代谢紊乱、肺部感染、人工气道意外脱出等11项手术并发症的全院汇总数据。点击月份可统计查看该月份发生手术并发症的病人明细,包括并发症名称、科室名称、住院号、住院次数、姓名、性别、年龄、主要诊断等7项数据指标。

**2.5 急性心力衰竭** 急性心力衰竭为单病种数量指标,可按时间范围统计急性心力衰竭病人数据明细,见图6,数据指标包括科室、住院号、住院次数、姓名、住院天数、住院费用、入院时间、首次心电图时间、心功能评估、住院后开始药物治疗时间、出院时药物治疗情况、健康宣教时间、针对性健康宣教是否完成、患者满意度等。

**2.6 医技指标** 医技指标包括我院病理科、放疗科、高压氧、检验科、康复科、麻醉科、内镜中心、输血科、疼痛科、心电图室、血透室、药剂科、影像科、营养科等14个医技科室的近300项指标。由于医技科室绝大多数指标均不能由医院各系统直接获取,本系统采取人工录入的方法由各科室质控员将指标数据手动录入到质控数据库中,见图7。科室质控员可通过数据分析功能,对其所录入指标数据,逐项生成折线图或是柱形图。

**2.7 医院工作报表** 医院工作报表的数据完全由天津市卫生统计系统中的医院工作报表数据中提取,包括门急诊人次、门诊人次、急诊人次、急诊抢救人数、急诊抢救成功人数、急诊死亡人次、观察室收容人数、健康检查人数、节假日人次、期末实有病床数、上期末留院人数、入院人数、他科转入、出院人数、住

医疗指标

科室	月份	诊疗人次	出院人数	出院者平均住院日(天)	平均病床工作日
全院	201901	101683	4101	5.6	37.7
全院	201902	82857	3272	5.1	29.7
全院	201903	102110	3918	5.6	36.8
全院	201904	97457	3840	5.6	36.4
全院	201905	104422	4066	5.4	36.7
全院	201906	90065	3706	5.7	34.9
全院	201907	101635	4234	5.3	37.1
全院	201908	103226	3850	5.4	35.2
全院	201909	96624	3725	5.5	34.1
全院	201910	90002	3830	5.6	36.4
全院	201911	95112	3823	5.6	35.5
全院	201912	107996	3946	5.8	36.8

图3 全年明细数据

院死亡人数、转往他科、期末留院人数、实际开放总床日数、平均开放病床数、实际占用总床日数、出院占用总床日数、手术总人次例数、平均病床工作日、病床使用率、平均病床周转次数、出院者平均住院日、住院危重病人抢救人次、住院危重病人抢救成功人次、病理检查人数等 29 项指标数据,见图 8。

我院医院工作报表数据由 2008 年至今,已经产生了近 12 年数据,通过科室、年份、单位(月、季、年)、数据类型四个维度的组合,可以生成院科两级 29 项指标折线图或柱形图。通过对历史数据的分析,不仅可以为院领导的决策提供数据支持,也可为职能处室制定政策提供数据支持。

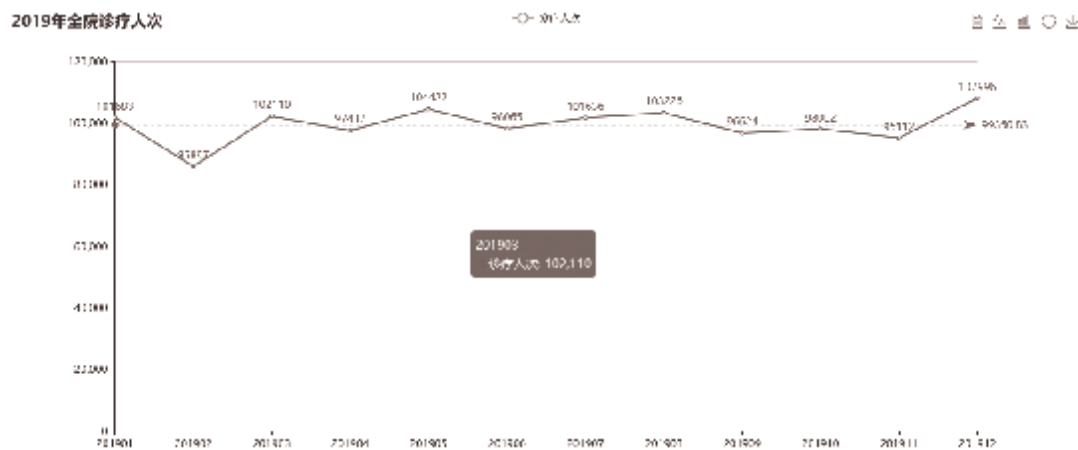


图 4 数据指标的可视化展示

2019-07 - 2019-08 住院新数据明细

科室名称	出院例数	死亡例数	15日再入院率(%)	30日再入院率(%)	平均住院日	平均住院费(元)
内分泌科	2	0	0	0	9	12579.16
普内科	1	0	0	0	3	3326.94
神经内科	378	1	0.151	0	6.46	9462.34
神经外科	2	0	0	0	26.5	79976.14
中医科	5	0	20	0	11.4	6099.62
重症医学科	6	1	0	0	3.33	14040.47

< 1 > 页码 1 页 共 6 条 12/5/2021

图 5 科级疾病明细数据

急性心力衰竭

科室	住院号	住院次数	姓名	住院人数	住院费用	入院时间
心血管科	866491	1		1	21246.13	2018-12-31 11:01:47
心血管科	508366	17		3	7980.01	2018-12-30 00:54:17
心血管科	867874	1		7	9488.47	2019-01-18 09:19:51
心血管科	845907	4		0	1	2019-02-09 14:44:07
心血管科	863010	2		1	3230.10	2019-02-10 01:01:00
心血管科	870063	1		8	34257.18	2019-02-14 11:24:46
心血管科	808885	4		1	1970.71	2019-02-21 23:32:44
心血管科	870766	1		4	4106.37	2019-02-22 22:01:47
心血管科	482610	3		13	50610.04	2019-02-15 23:40:56
心血管科	617968	3		11	11947.78	2019-02-17 09:23:31
心血管科	536642	5		5	10458.08	2019-03-02 18:55:14
心血管科	426301	5		6	6729.67	2019-03-04 10:30:34

< 1 2 3 ... 35 > 页码 1 页 共 412 条 12/5/2021

图 6 急性心力衰竭数据明细

图 7 医技指标数据录入

科室	月、季、年	门诊数据					
		门诊人次	门诊人次	急诊人次	急诊死亡人次	急诊抢救人次	急诊抢救成功人数
内科	201901	101000	90360	11315	29	609	600
内科	201902	100997	90211	10905	25	523	520
内科	201903	102110	90533	10777	27	527	525
内科	201904	97437	86455	10902	24	635	631
内科	201905	104422	94766	9556	25	617	616
内科	201906	98065	88118	10477	7	587	585
内科	201907	101636	92704	10537	21	576	575
内科	201908	100376	90254	10777	35	606	605
内科	201909	98674	87558	10035	23	677	675
内科	201910	98002	88445	10555	24	650	648
内科	201911	98112	88165	10477	24	641	638
内科	201912	107535	97288	10637	34	684	682

图 8 指标数据展示

### 3 讨论

近年来,随着医疗大数据的体量越来越大,越来越多的公立医院对于医疗大数据的价值挖掘越来越重视,而信息技术的飞速发展恰好为医疗大数据的有效利用和创新挖掘提供了积极的探索途径。医疗大数据形式种类多样,时效性高,本文通过运用基于提取、转换和加载的 ETL 技术对其进行分析处理,提取出日常质量控制需要的相关指标数据,并对结果进行对比展示和报告分析。本次提出的医疗质控数据库主要包含 7 大类共 198 项质控指标,可溯源性达 95%,包括质量、经营、绩效 3 个层次,涵盖基础质量、环节质量和终末质量指标,指标评估完成率

达 100%,同时评价指标具有特异性和专业性,数据 100%自动采集、整合,准确性达到 98%以上。

质控数据库系统的研发成功,为我院的医疗质量与安全管理提供了有效的数据支持,并极大的节约了经济成本,达到了项目预期效果。实践证明,系统化设计、信息化手段、精准化监管的三位一体模式能有效实现质量与安全的线上线下管理,从而确保医院质量与安全管理的效率与效能。

#### 参考文献:

- [1]徐祥贵,丁强,陈彦,等.“组团式”援疆促进医疗质量持续改进的实践与探索[J].现代医院管理,2018,16(3):13-16.
- [2]李泉,李雪婉,孙鉴.区域卫生信息化进程中大数据分析在公立医院绩效考核应用[J].医学信息,2014,27(4):1.

- [3]谢世堂,沈慧,曹桂,等.基于融合理念的公立医院内部绩效考核体系探索研究[J].中国医院管理,2017(10):22-24.
- [4]胡军,贾垒,袁晋华.现代化医院医疗质量管理的体系建设与实践[J].卫生软科学,2018,32(10):39-41.
- [5]管雅文.整合生物信息的临床数据中心建设方案[J].中国数字医学,2019,14(2):57-59.
- [6]魏小林,叶舟,拓宽前.基于 ETL 技术的医院绩效管理系统设计与实现[J].中国数字医学,2015,10(11):22-24.
- [7]Kushiro N,Matsuda S,Takahara K.Model Oriented System Design on Big-data[J].Procedia Computer ence,2014(35):961-968.
- [8]贾军华,徐洋阳.医疗大数据在医院绩效管理中的应用探讨[J].中国管理信息化,2019,22(24):48-49.
- [9]廖家智,陈安民,徐永健,等.精准化、信息化、系统化的医疗质量与安全体系构建 [J]. 中华医院管理杂志,2016 (32):123-125.
- [10]赵贵兵.Kettle 应用案例分析[J].电脑知识与技术,2019,15(9):95-96.
- [11]刘繁艳.基于数据仓库的 ETL 技术研究[J].软件导刊,2007(11):24-26.
- [12]周丹,尹安春,张秀杰.基于 Web of Science 的肾脏病学领域护理研究的可视化分析[J].中国实用护理杂志,2020,36(4):315-321.
- [13]朱鹏志.基于 Ajax 和 SSH 技术的办公自动化系统的实现[D].青岛大学,2020.
- 收稿日期:2020-09-30;修回日期:2020-10-09  
编辑/钱洪飞