

· 论 著 ·

数据仓库在数字化医院的应用

宋 斌,陈海东,雷 勇,黄绍君,卓飞豹

[摘要] 数字化医院是我国现代医疗发展的新趋势,也是区域医疗信息化的一种重要实现方式。数据仓库和数据挖掘技术则是信息技术领域的新兴技术。本文将数据仓库和数据挖掘技术应用与数字化医院领域,建立基于医院信息系统(HIS)的数据仓库并使用数据挖掘技术,将大量源数据有效地转化为有用的知识信息,并服务于决策过程。

[关键词] 数据仓库;数字化医院;决策支持;区域医疗

中图分类号: R197. 32 文献标志码: A 文章编号: 1672-271X(2010)06-0519-04

Data warehouse application in digital hospital

SONG Bin, CHEN Hai-dong, LEI Yong, HUANG Shao-jun, ZHUO Fei-bao. The First Affiliated Hospital of Fuzhou General Hospital, Nanjing Military Area Command, Putian, Fujian 351100, China

[Abstract] Digital hospital is a new trend of the modern medical development, also an important implementation method of the regional medical informationization. Data warehouse and data mining are new techniques in information technology. In this paper, we apply this technology to digital hospital field, and build the hospital information system (HIS) based data warehouse and data mining technology. A lot of source data can be transformed to useful information so the information can be used in the decision process.

[Key words] data warehouse; digital hospital; decision support system; regional medical

随着数字化医院概念和计算机技术的发展,医院信息系统(HIS)已被各大医院所广泛使用。经过多年的积累,HIS 系统已收集到了海量的信息,同时其信息容量仍在呈几何级数增长。这些大量的信息成为了各个医院的宝贵财富。

然而,随着管理和临床数据的大量积累,HIS 原有相对简单的统计功能已不能满足人们日益增长的需求。人们需要随时获取患者、资金、物流、工作量等方面的数据、指标和报表,需要采用复杂的统计分析方法和数据挖掘技术反复处理海量历史数据以便从中总结出临床医学和医院管理方面的新知识。因此,如何整合大量的历史数据,通过信息的分类、整理、分析,从中发现医疗业务和经营的内在规律,从而为管理人员的决策提供支持,已成为现代 HIS 系统的发展目标^[1]。

本文在实现医院信息数字化的基础上,结合区域医疗信息平台,并采用数据仓库、联机分析处理和

数据挖掘技术等新兴信息技术相结合的方法建立决策支持系统以帮助医院制定有效的决策。

1 相关背景

1.1 数字化医院 数字化医院是我国现代医疗发展的新趋势,数字化医院系统是由医院业务软件、数字化医疗设备、网络平台所组成的三位一体的综合信息系统。数字化医院工程有助于医院实现资源整合、流程优化,降低运行成本,提高服务质量、工作效率和管理水平。数字化医院是对现代医院的一种特征性描述,是医院信息化发展过程中的一种必然产物。所谓信息化,其实是泛指重视信息利用的一种理念;而数字化则强调了信息内容的数字化,是专指数字化信息,并利用数字化的信息进行计算机处理。数字化是包含在信息化之内的一个更具体的概念,是在信息化基础上的一个发展。因此,数字化医院的实质是高度利用计算机、网络通讯等现代化技术实现信息化,但对数字技术应用的要求更高,对数据利用更全面,对信息的共享程度更高。建设数字化医院就是利用数字化设备和信息技术,实现医疗信息利用最大化、信息共享通用化,管理程序规范化、患者服务人性化、领导决策科学化的目标^[2]。

基金项目: 南京军区医学科技创新课题(06MA140; 07MA091)

作者简介: 宋 斌(1963-),男,江苏海门人,硕士,主任医师,从事卫生事业管理工作

作者单位: 351100 福建莆田,南京军区福州总医院第一附属医院

1.2 医院当前信息化现状 随着信息化的深入发展,当前医院存在如下问题急需解决:①各类信息系统大多数都是面向业务事务处理的系统,比如病历系统、门诊工作站等等,都是以具体的各项业务完成为基准,没有很好的统计分析对比功能;②信息系统已经运行了多年,积累了大量数据,部队医院军字一号系统已经运行了十多年的,积累的大量的历史数据,但是没有充分发挥应用;③“数据孤岛”,各个业务系统都有各自的环境,标准不统一,数据孤立,无法实现信息共享;④管理决策层对数据分析基础平台的需求日益强烈,管理层需要以统一、准确、有效的数据为基础,做出科学决策^[3]。

1.3 区域医疗信息化 实现区域医疗信息化,在信息共享基础上实现远程实时会诊、医学信息传输及手术示教等,对于我国医疗卫生体制改革具有重要的现实意义。随着我军卫勤保障体制改革不断推进,军队医院也向集团化转变;军队远程医疗会诊网络已建立,并不断完善;构建区域信息平台对于医疗

服务体系带来了巨大的后续效益^[4]。

区域医疗信息系统的关键之处就在于能够有效整合病人的信息,将同一个病人的所有诊疗信息、健康信息和基本资料信息进行归类汇总,整合到同一数据平台,不同的应用都将信息汇总到这个平台,实现远程实时会诊及远程康复指导,远程医学影像、检验数据、心电图信息的实时传输和区域内医疗单位检查结果互认;通过建立军人电子健康档案,便于及时调阅并贯穿于医疗保健全过程,最大限度满足各级人员的健康需求^[5]。这种归档不是简单的数据复制,而是以病人为中心的数据整理。通过区域医疗信息平台及含有通用数据标准、接口标准等,以网络信息化服务为手段,对区域有限的卫生资源进行整合,实现区域内各个部门信息采集、传送、共享、交换和协同服务。如图 1 所示,整个区域医疗信息平台的总体框架,其底层为数据交换和共享层、中间为数据中心和服务支撑层、上层为业务应用层。

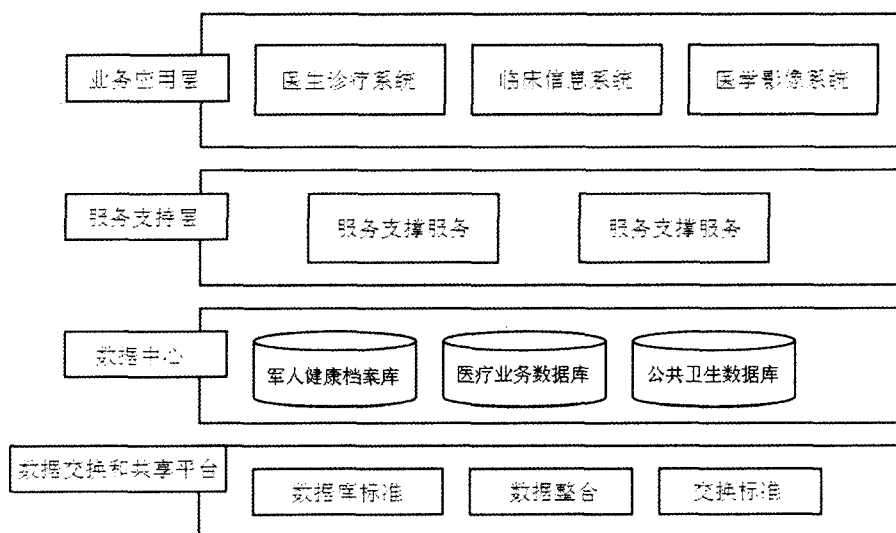


图 1 区域医疗信息平台

2 数据仓库

面对积累的大量业务数据,通过建设数据仓库对其进行整合并有效挖掘,是进行数据深层次利用的主要手段。在医院建设统一、规范的数据仓库,也是当前医院数字化、信息化发展的重要课题。因此,通过数据仓库技术来建立高效和谐的区域医疗信息系统,是区域医疗信息化发展的一种趋势。

2.1 数据仓库的概念 “数据仓库”一词没有一个统一的定义,其概念是由著名的数据仓库专家 W. H. Inmon 于 1992 年在其著作《Building the Data

Warehouse》中提出来的^[6]。数据仓库是面向主题的、集成的、随时间变化的、非易失性的数据集合,用于支持管理层的决策过程。数据仓库中的数据不是将业务处理系统中的数据简单集合,而是对各种源数据进行抽取、筛选、清理、综合而得到的数据集合。数据仓库中所存储的数据不经常进行更新处理,它主要用于查询和分析^[7]。

数据仓库的主要功能是企业级的多维信息查询、联机分析处理(OLAP)在线分析处理、数据挖掘和预测等功能。数据仓库从多个信息源中获取原始数据,经整理加工后,存储在数据仓库内部中,并通

过向用户提供访问工具,统一、协调和集成的信息环境,支持医疗机构全局的决策过程和对医疗机构经营管理的深入综合分析。同时,数据仓库应用还是一个不断深入、不断完善的过程,完成了一个主题、一个应用后,又会提出更多的应用需求。

以主题为导向的数据仓库是围绕着医院的基本实体设计的,如 HIS 中数据仓库的设计主题可以以医疗业务、病人结构、病人流动状况、床位占用率、资金流动等情况考虑,它们都是密切联系的,通过数据仓库的建立,可以有利于实现数据的关系化、规则化,并可提供动态的、多维的数据统计、分析和查询。

2.2 数据仓库的结构 数据仓库的总体层次结构由基本功能层、管理层和环境支持层三部分组成。①基本功能层:包括从数据源抽取数据,对所抽取的数据进行筛选、清理,将清理后的数据加载到数据仓库中,根据用户的需求完成数据仓库的复杂查询、决策分析和知识的挖掘等功能;②管理层:包含数据管理与元数据管理两部分,主要负责对数据仓库中数据的抽取、清理、加载、更新和刷新等操作进行管理。只要使这些操作正常完成,才能源源不断地为数据仓库提供新的数据源,为使用者正确利用数据仓库进行决策分析和知识挖掘;③环境支持层:主要包含数据传输和数据仓库基础两大部份。包括网络协议、网络操作系统、数据存贮系统等。

3 数据仓库技术在数字化医院中的应用

3.1 医院数据仓库的构建 当前医院数据仓库的构建步骤如下:①首先是设计数据仓库的主题并建立规范的数据标准,比如医生主索引、病人主索引、病种主索引等。建立数据仓库模型首先必须进行需求分析。数据仓库是按照决策分析的主题来组织数据的,因此设计数据仓库的主题至关重要。②其次对当前医院的各个系统进行统一抽取,并按照统一的数据标准进行清洗,转化成统一的、规范的数据。医院数据仓库的数据源非常复杂,既包括操作型数据库如 HIS、人事管理系统等,又包括非结构化的外部数据如各种文件。各种来源的数据在结构和组织方式上可能存在较大的差异,也可能包含噪声、空缺、不一致的数据,因此数据源的数据必须经过严格的抽取、清洗和转换,才能加载到数据仓库。③最后根据医院各项业务需求,通过商业智能工具分主题灵活展现各项信息。数据仓库建设流程如图 2 所示。

3.2 数据仓库在数字化医院中的优势

传统的统计查询与基于数据仓库的辅助决策支持系统^[8]对比见表 1。将数据仓库技术引入了医疗信息系统中,从根本上改变了原来的报表方式。由原来被动的、固定的报表模式,转变为各级领导可以

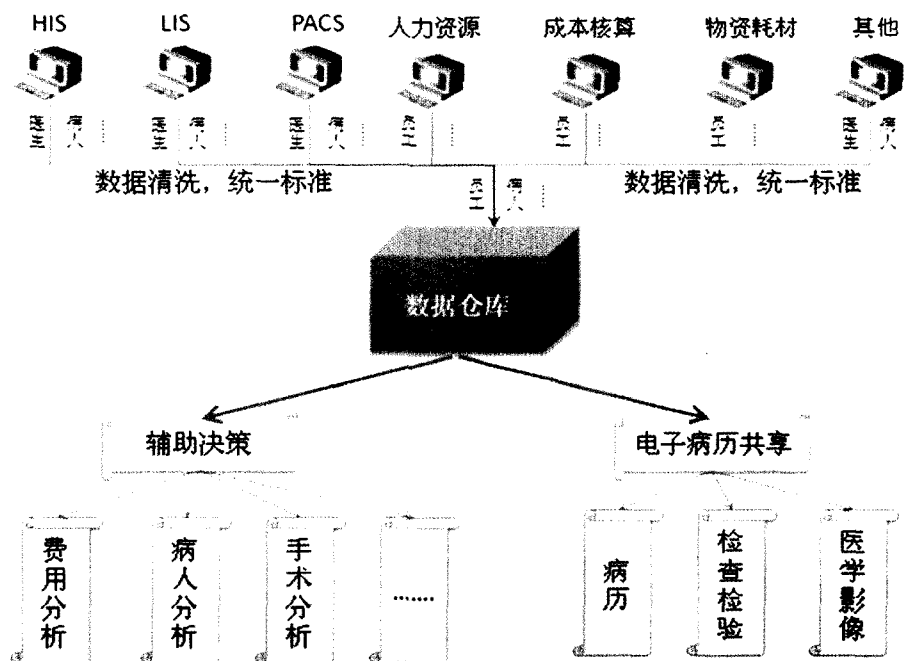


图2 数据仓库建设流程

表 1 查询统计系统与辅助决策支持系统对比

传统的查询统计系统	基于数据仓库的辅助决策支持系统
面向应用的	面向主题的
独立的	共享的
详细的数据	综合、提炼的数据
当前的数据	当前及历史的数据
为日常工作服务的	为管理层服务的
孤立的某个业务方面的查询	可覆盖医院各个层面的综合查询
简单的数据表查询	联机分析、数据挖掘

根据自己的想法,根据医院当前的发展目标,从现有的模型中主动挖掘自己想要的数 据,充分挖掘医院信息系统的信息资源,从而提高医院信息系统的应 用分析水平。

对医院决策者来讲,数据仓库可提供以下主要功能:①多层次、多角度的数据挖掘和分析。比如可以分析门住院病人疾病、地区、年龄、职业以及性别等分布,还可以根据 需要进一步找出哪段时间、哪些人、进行了哪些检查治疗、治疗效果如何等,为流 行病学调查和疾病预防提供依据。可以通过分析门诊病人的挂号、就 诊、检查、治疗以及取药等时间分布,为医院制定科学合理的门诊就医 流程提供依据。②辅助决策功能。例如,当医院发现药品比例过高时, 可以按时间段分析哪些医生在哪段时间内药品使用超过了合理比例, 从而为医院决策提供依据。当医院手术间很忙时,可以通过分析各手 术间进行手术的时间分布,看是否可以适当调整手术安排时间来提 高手术间的利用率或是需要增加手术间。③提高科研能力。医院科研 可以充分利用数据仓库来挖掘各种信息资料,加强这些信息资料后续 的综合分析和利用。可为医学科研、技术层面总结更多各种疾病的救 治经验,以及寻找和总结救治过程中一些规律性的东西。④提高医院 预测能力。预测准确的高低也是衡量决策水平高低的重要标志。数据 仓库是一个集成的系统,它除了可以采集来自医院本身的数据,还可 以采集医院外部的数据,如本地区的人群分布、职业分布、饮食习惯 和医疗资源配置等,通过多方的数据采集和科学的数学模型来预测未 来。

4 结 语

总之,在数字化医院的基础上建设数据仓库,构

建决策支持系统,能够利用数据仓库工具充分分析信息,从而为医院带来效益。

对于管理层和决策层而言,通过使用数据仓库,对门诊住院数据进行分析,对就医人员进行分析,为管理者提供决策依据,能够主动争取到病源,提高经济效益。对医务工作者而言,使用数据仓库对疾病的治疗进行分析,能够提高医疗质量。而对患者而言,则可以通过数据仓库提供给患者全方位的医疗服务以及在线的医疗帮助。

【参考文献】

[1] 陈金雄,刘雄飞,王庆森.数据仓库与医院管理决策[J].中国医院,2004,8(10):12-14.

[2] 范启勇,王杰宁,邬惊雷. 字化医院的内涵、特征和基本框架[J]. 中华医学管理杂志,2006,22(1):57-58.

[3] 吴志成. 军区医院数字化建设需把握的几个问题[J]. 东南国防医药,2009,11(3): I-II.

[4] 李 磊. 究区域医疗信息化建设[J]. 医学信息,2009,22(7):1134-1136.

[5] 凌小明,沈建坤,程钦安,等. 为部队网络医疗服务新模式的创建[J]. 东南国防医药,2009,11(6):561-562.

[6] Inmon WH. 数据仓库(Building the Data Warehouse)[M]. 王志海,林友芳,译. 北京:机械工业出版社,2005.

[7] 陈京民. 数据仓库与数据挖掘技术[M]. 北京:电子工业出版社,2002.

[8] 罗 晶,袁慧颖. 用数据仓库技术有效利用医院信息资源[J]. 海军总医院报,2009,22(2):111-112.

(收稿日期:2010-10-27)

(本文编辑:孙军红)