

医院饮片信息数据库的设计初探

孟庆安

天津中医药大学第二附属医院药剂部, 天津 300150

摘要: 运用现代数据库、数据仓库的理论, 结合日益强大的计算机技术, 以关系型数据库为基础, 使用 Access 和模块化设计方式建立适合医院对饮片信息实际利用需求的饮片信息库。建立了数据库信息来源路径的标示方法; 确立了适合在医院范围内使用的中性饮片编码方法; 确立了对数据库中饮片信息“量子化”拆分的方法; 确立了由核心到外延, 分别构架, 采用接口连接的建库方法。通过初步应用, 数据库的设计既满足了医院对饮片信息利用的实际需求, 屏蔽各种庞杂的饮片信息, 又为饮片信息的科学研究利用保留了充裕的可扩展空间, 方便日后根据需要进行信息的补充和对接。

关键词: 数据库设计; 医院饮片; 饮片编码; 量子化; 置信级别

中图分类号: R288

文献标志码: A

文章编号: 0253-2670(2011)03-0610-03

Design of decoction pieces information database in hospital

MENG Qing-an

Department of Pharmacy, Second Affiliated Hospital, Tianjin University of Traditional Chinese Medicine, Tianjin 300150, China

Key words: database design; decoction pieces in hospital; coding of decoction pieces; quantization; confidence levels

目前, 中药饮片仍然是中医用于治疗疾病的主要手段。由于中药饮片在医院的应用涉及市场信息、采购、鉴别、养护、管理、炮制、调剂、制剂、方剂、药学服务、不良反应、经济效益等多个方面, 而每个方面又涉及许多专业性很强的其他领域。因此, 医院对中药饮片信息的需求是具体、简洁、方便扩展、易于对接。然而, 目前的中药信息化成果, 不是大而全^[1-2], 就是精而专^[3-4], 都是针对中药这个学科而构建的。能够按照医院对多种药品的同一类特定需求进行查询、筛选、聚类研究的数据库几乎没有; 能够按照特定医院的不同需求, 进行灵活变化, 适应性强的数据库更是难找。为解决这一现实问题, 笔者根据现代数据库的理论和技術, 结合日益发展的计算机软硬件技术, 设计构建一个面向医院中药饮片实际使用需求, 能够灵活调整的数据库, 使其既可以屏蔽各种庞杂的饮片信息, 又能为饮片信息的科学研究和利用保留可扩展的空间。

1 医院饮片数据库设计的关键问题分析

根据此数据库所要达到的目的和实现的功能, 数据库的设计首先要解决信息来源、编码方法、信息量子化、整体架构 4 个核心问题。

1.1 信息来源

数据是组成数据库的基本单元, 优质数据来自优质数据源^[5], 因此, 信息来源是判别数据库中所存数据有效性的保障。数据源不只是来自一种出版物, 而是集多种权威出版物之大成。由于各出版物在体例、用药名称、药物来源上不尽相同; 而医院中药饮片工作又牵涉极广, 所需饮片信息十分庞杂, 同一个药的同一类信息, 由于地区不同, 也不尽相同, 甚至可能会遇到在已有的可查询数据源上没有记录的药品。所以必须保证有源信息的可执行性, 以及一切收录信息的可溯查性。

1.2 编码

编码的可靠性、灵活性是保证数据库中所有饮片信息不出现重叠、歧义, 便于存储、调用及其独立性的关键。目前没有一个全国通行的编码方案, 现有的方案都是针对饮片本身而编制的, 如果不能作为国家标准, 就不具备与其他相关数据库接驳的能力。由于医院饮片信息数据库是分模块单独设计每个相对独立的部分, 最后靠接口完成连接, 因而要对大量的独立信息进行编码, 这些编码将担负各模块接口的重任。如果编码含有饮片过多的属性,

收稿日期: 2010-08-30

基金项目: 天津市卫生局立项课题 (2005014)

作者简介: 孟庆安 (1976—), 男, 天津市人, 主管药师, 硕士, 研究方向为医院中药学。Tel: (022)60335452 E-mail: wmx2k@hotmail.com

就会使数据库丧失灵活的按需扩展的能力,因此编码的设计必须保证具有绝对的独立性、不重复性、可靠性和相对的统一性。

1.3 信息的“量子化”

使用采集的大量中药饮片的数据,不论是数据仓库的挖掘算法,还是一般的数据库应用,最主要的不是建立方法,而是独立数据项的简单程度,越是简单的、不可分割的元数据^[6],越使同一属性的归纳简单易行,继而愈加方便计算机进行后续处理。因此这是源数据的预处理方法,是实现本数据库设计使用目的的关键方法。

1.4 数据库整体架构方案

整体架构方案是建设本数据库的总体工艺流程,好的方案可以使药品名称、别名、基本属性、分类、用法用量、图像调用、炮制、处方应付、成方处理、性状鉴别等方面根据需要,作为独立模块充实到数据库中,且变得有法可依,并可基于各个模块填充数据,进行实际应用开发。

2 关键问题的解决方案

2.1 采用设立置信级别的方法标示信息来源路径

在实际工作中,饮片所执行的标准各地并不完全相同,基于天津地区医院饮片主要以《中国药典》和《天津市饮片炮制规范》为法定标准的具体情况,确定以这两部标准上的信息为准。如果牵涉到这两部标准上没有收录的品种,再以国家标准→地方标准→现代典籍→古代典籍→专业期刊→约定→口口相授→网络→道听途说的顺序一级一级查找信息。这里的“约定”,用于标识来源于各地对于常用复方约定俗成的药品名称,如“斩毒剑”、“广东神曲”等。“口口相授”和“道听途说”用于标识那些尚无确切出版物记载,但已流传甚广和在局部使用的药品。为了使数据库的使用者可以判别和排序所查到的信息的可信度,设立置信级别信息,用以按需标识信息来源的可信度。

为此建立了两张关系表,一是 Authority,包括编号、名称、助记符、简称、置信级别等字段,收录所有用到的信息来源;一是 Criterion,存储信息来源的等级并设定等级的置信级别,包括编号、等级标准、标准级别,见图 1。

2.2 采用随机不重复编码方法

从编码学的角度来看,“便于记忆、简单实用、检索迅速、重码率低”应成为衡量药品编码系统好坏的基本准则^[7]。目前国内已有的医院药品编码,从形

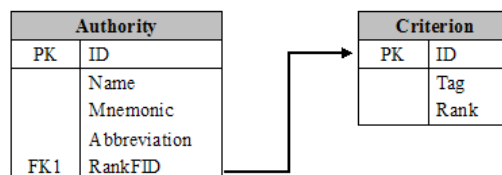


图 1 信息来源及置信级别数据结构图 (PK-主键, FK1-外键)

Fig. 1 Data structure diagram of information source and confidence levels (PK-primary key, FK1-foreign key)

式上看也是按照这 16 个字,分为了数字字母混合型、字母型、数字型 3 类^[8],但各有利弊。从整体来看,操作简便的编码方案检索迅速、重码率低,但不便记忆、实用性差;容易记忆的,检索速度慢、重码率高。所以这 3 类编码都不能达到上述 16 个字的标准。

基于上述分析,本研究以计算机自动生成的不重复的编号为药品的主键编码,以此杜绝人为差错,保证每个药品信息的唯一性。药品信息的调用和处理也以此编码实现。对于药品的录入、检索,使用助记符虽然重码率加大,但第一不会使药品信息混乱,第二符合人类记忆的自然规律。在使用时,如果出现重码,将以弹出窗口的形式将类似信息逐一列出以供选择,如此还可根据需要增减弹出窗口所含的信息量。经实际使用证实,在医院使用的范围内,这个方法适应于或快速或精确选取信息的需求。

2.3 应用“量子化”理念界定医院饮片实用信息

所谓中药信息的“量子化”是指通过合理的解析将原有的中药繁冗信息细化成由若干汉字组成的不可再细分的、具有相对独立内涵和排他作用的最小信息单位。这种最小信息单位称之为“量子”^[9]。

中药信息大致可以分成 2 种类型:一类是具有相对独立内涵和排他作用的词组型短小数据,如性味、归经、毒性、化学成分等;另一类是由多个汉字组合而成的大文本复杂数据,如药理、临床、各家论述等。为了使上述两类数据尤其是大文本复杂数据适宜于计算机处理,在中药数据库设计中引入了中药信息的“量子化”概念。

“量子”数据的采集非常繁杂,涉及专业领域广,又没有统一标准,已有的类似中药饮片数据仓库的建立,多偏向于特定数据挖掘的目的,故而他们设定的中药信息的“量子化”方案也偏向于特定目标,适应性不广,不适合临床。因此,以临床实用为指导思想,将中药信息分解为更适合临床使用的“量子”数据,以便应对今后各项需求。如将饮

片的来源分为类科、类种两个属性，饮片的性味归经分为性、味、归经 3 个属性等。

2.4 数据库构架方案

对于数据的“量子化”和数据库的构架，采取由内及外、分立组合的实施方案（图 2）。步骤如下：

①确定饮片的核心信息；②将信息“量子化”，即提炼药品的最小属性字段；③根据“量子化”后的结果，将能完整描述最小有用记录的属性组合成独立表，并留下对外接口；④根据需要，确定附着信息（如分类、组方、化学成分等），重复执行步骤②、③、④。

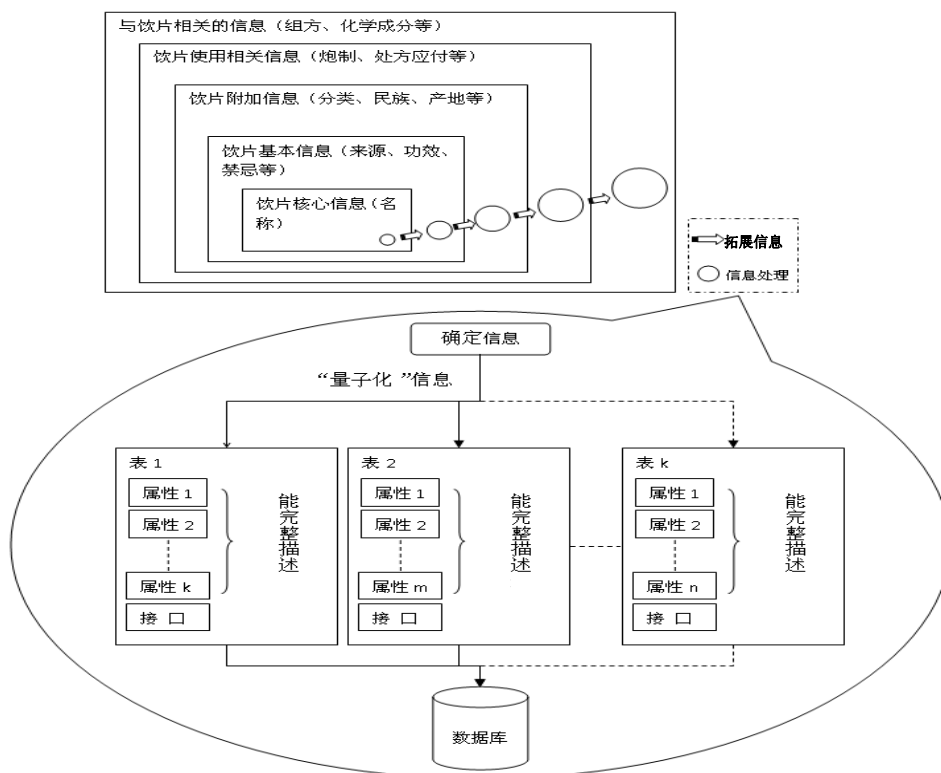


图 2 数据库构架方法及步骤示意图

Fig. 2 Frame of database and procedure diagram

3 结语

根据上述方法所构建的数据库经在本院实际使用证实，对于医院饮片在产地、化学成分、饮片科学鉴定、饮片调剂、商业、报告、单据、药学服务、仓储养护等方面的信息要求和工作内容，都完全可以按照本数据库的构架组织结构，共享已有结构和数据，实现了提取饮片信息的方便、灵活、简洁。且各组块的独立性很强，可以在今后进行细化研究时，进一步分解、组合，同时并不影响已有数据，并可快速调整应用程序，非常适合医院对饮片工作繁杂、多变的使用需求。总之该数据库实用性强且数据库构架具有很好的可扩展性，易于对接，提高了饮片管理的效率。

参考文献

[1] 张林. 数据库系统在中医药领域的应用与展望 [J].

中华中医药学刊, 2007, 25(1): 133-135.

- [2] 薛兴亚, 徐青, 章飞芳, 等. 中药信息数据系统构建 [J]. 世界科学技术—中医药现代化, 2006, 8(3): 91-94.
- [3] 王虹菲, 刘启贵, 刘辉, 等. 构建中药免疫数据库及对中药免疫规律的研究 [J]. 中草药, 2007, 38(2): 297-301.
- [4] 刘海波, 乔颖欣, 刘冰, 等. 中药化学数据库的数据仓库改造 [J]. 中国药学杂志, 2006, 41(9): 645-648.
- [5] 林端宜, 徐筱杰, 杨淑静, 等. 中药数据优化处理探讨 [J]. 中国中药杂志, 2008, 33(19): 2291-2293.
- [6] 王丹, 王文生. 元数据与数据元的内涵及其应用 [J]. 农业网络信息, 2005, (11): 27.
- [7] 叶显智, 施东辉, 程苏宁. 医院信息管理系统中药品编码初探 [J]. 西藏医药杂志, 2000, 21(4): 44-45.
- [8] 宋华. 医院中药饮片的编码研究 [J]. 中国药学杂志, 2004, 39(9): 712-714.
- [9] 吴园园, 章新友. 中药信息量子化研究 [J]. 江西中医学院学报, 2008, 20(2): 56-57.