

DOI:10.3724/SP.J.1008.2010.00691

应用在线联机分析处理探讨腰椎后路减压植骨内固定术治疗腰神经根损伤效果的影响因素

Application of on-line analytical processing in evaluating factors influencing effect of posterior lumbar internal fixation in treatment of lumbar spinal nerve root damage

史国栋, 欧阳跃平*, 史建刚, 贾连顺, 袁文

第二军医大学长征医院骨科, 上海 200003

[摘要] **目的** 应用在线联机分析处理(online analytical process, OLAP)探讨腰椎后路减压植骨内固定术治疗腰神经根损伤效果的影响因素。**方法** 选取2007年1月至2008年6月第二军医大学长征医院61例行腰椎后路减压植骨内固定术的腰神经根损伤患者作为研究对象, 汇总患者的临床资料, 包括: 基本的住院号、性别、职业、年龄、病史、症状持续时间等, 建立数据仓库, 在此平台上应用OLAP技术进行数据的多维分析处理, 验证手术治疗效果的影响因素。**结果** 经过分析处理发现在各项相关因素的综合作用中, 神经根受压迫时间较为重要, 直接影响术后治疗效果, 符合临床实际。**结论** OLAP技术能有效处理大宗腰神经根损伤患者的临床资料, 有利于指导临床实践。

[关键词] 腰椎后路减压植骨内固定术; 神经根病; 数据挖掘; 联机分析处理; 临床研究

[中图分类号] R 681.533 **[文献标志码]** B **[文章编号]** 0258-879X(2010)06-0691-03

腰椎后路减压植骨内固定术治疗腰神经根损伤的效果影响因素较多, 包括患者年龄、性别、主要症状、症状持续时间等, 相关临床数据具有数据量大、数据维数多、数据类型复杂等特点^[1-4], 无法被传统的数据处理方法很好地处理, 严重制约了相关临床研究的发展。数据联机分析处理(online analytical process, OLAP)给相关研究的进步带来了希望。OLAP是一种用于组织大型商务数据库和支持商务智能的技术, 是目前较先进的数据分析处理方法, 具有灵活的分析功能、直观的数据操作和分析结果可视化表示等优点, 从而使基于大量复杂数据的分析变得轻松而高效, 以利于迅速作出正确判断^[5-7]。近年来, 其被广泛应用于信息产业界等多个领域, 也逐步被应用于医学实践, 取得初步成效^[8-10]。

本研究对2007年1月至2008年6月第二军医大学长征医院收治的临床及影像学诊断明确、病例资料完整的61例行腰椎后路减压植骨内固定术的腰神经根损伤患者的临床资料进行汇总, 尝试采用OLAP技术对汇总资料进行处理, 以指导临床实践。

1 资料和方法

1.1 数据来源及获取 选取第二军医大学长征医院2007年1月~2008年6月临床及影像学诊断明确、病例资料完整的61例行腰椎后路减压植骨内固定术的腰神经根损伤患者作为研究对象, 汇总所有临床资料。61例患者中男38例、女23例, 年龄23~75岁, 平均(49±1)岁; 病程最长为24个月, 最短为1h。汇总患者的相关信息, 包括住院号、性别、居

住地、职业、年龄、病史、过敏药物、主要症状、症状持续时间、手术方式、治疗方案、用药方案、术后症状的改善等。

1.2 数据联机分析处理

1.2.1 临床数据的单维汇总 获得的临床数据必须进行有效分类存储后形成数据仓库才能进行高效处理。本研究将不同的临床数据存储在不同的表格, 避免了数据的冗余, 使得查询响应速度加快, 操作更加简便。通过表的相互连接来进行各种属性上的查询。根据实际的需求和选择的技术, 本研究选取的患者基本情况表包括: Symp_history表、Segment表、Age表、Recover_situation表, 并对数据属性进行了映射, 保留患者ID识别号码、发病持续时间、年龄、腰椎病变累及节段、手术后的恢复程度, 形成患者具体数据表。

1.2.2 临床数据的多维分析 为分析各维数据间的交互连接, 分析各自间的相关性, 需要采用OLAP对上述数据进行多维分析处理。具体步骤: 根据表1可以统计在不同范围的属性中患者数量, 以此来得到一系列相关属性间的关系。选取数值型的数据来进行分析, 从Age、Recover_situation、Symp_history等3个维上观察数据, 将年龄和症状持续时间进行分段。

数据的预处理过程主要是涉及数据变换、数据规约和概念分层。数据变换采用了最小-最大规范化对原始数据进行线性变换。假定 \min_A 和 \max_A 分别为属性A的最小和最大值。数据变换的基本公式:

$$v' = \frac{v - \min_A}{\max_A - \min_A} (\text{new_max}_A - \text{new_min}_A) + \text{new_min}_A$$

针对恢复程度的数据处理上, 将评价数据映射到[0, 10]

[收稿日期] 2009-12-20 **[接受日期]** 2010-04-30

[基金项目] 国家自然科学基金(30300359), Supported by National Natural Science Foundation of China(30300359).

[作者简介] 史国栋, 博士, 副主任医师。E-mail: shgd008@163.com

* 通讯作者(Corresponding author). Tel: 021-81885632, E-mail: oyaopin@yahoo.com.cn

的区间之内,如将2号患者的恢复程度评价是6.2,根据计算方式 $\frac{6.2-2.2}{9.1-2.2} \times 10 = 5.80$ 。数据划分成 $[0, 2], (2, 4], (4, 6], (6, 8], (8, 10]$ 5个分段,并且将数据分别规约为2、4、6、8、10(图1)。

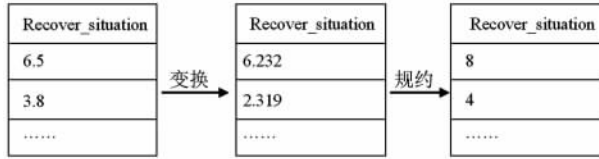


图1 数据预处理示意图

概念分层允许挖掘多个抽象层上的数据,是数据挖掘一种强有力的工具,通过将属性域划分为区间,离散化技术可以用来减少给定连续属性值的个数,减少属性值的数量(图2)。因此在这里病症的持续时间以月、半年、1年等多个属性

值上进行度量,并且在此基础上进行联机分析处理。

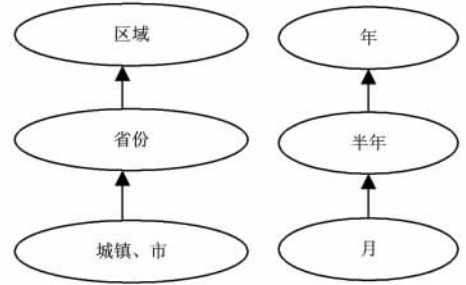


图2 一个基于不同值个数的模式概念分层的自动产生

1.2.3 主题数据的展示 多维数据模型围绕中心主题(如,数量)组织,该主题用事实表表示。事实是数值度量的。把它们看作数量,是因为想根据它们分析不同维之间的关系。在表1中可以从这3维中观察数据,表示的度量是该属性上的对象个数,也可以以3维数据方的形式表示这些数据(图3)。

表1 从年龄、症状持续时间、术后恢复程度上的统计数据

| Symp_history t/month | Age="0-19" | | | | | Age="20-29" | | | | | Age="30-39" | | | | | Age="40-49" | | | | |
|-------------------------|------------|---|---|---|----|-------------|---|---|---|----|-------------|---|---|---|----|-------------|---|---|---|----|
| | 2 | 4 | 6 | 8 | 10 | 2 | 4 | 6 | 8 | 10 | 2 | 4 | 6 | 8 | 10 | 2 | 4 | 6 | 8 | 10 |
| 0-6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 2 | 1 |
| 7-12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 1 | 0 | 0 |
| 13-18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 19-24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

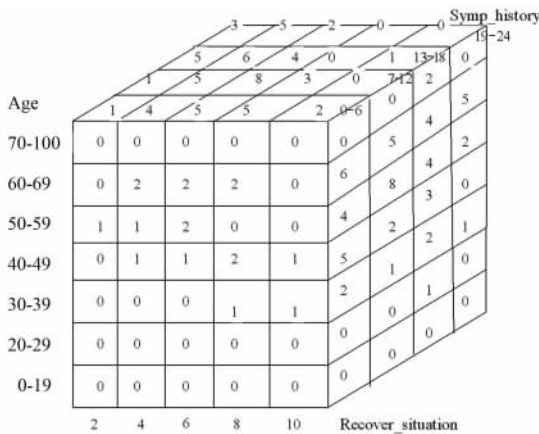


图3 三维数据立方体描述各个属性特征的病例数量

2 结果

2.1 年龄因素 结果(图4A)表明:70岁之前,随着患者年龄的增大,腰椎神经根损害的发病率呈递增趋势,70岁之后,发病数锐减。分析其原因可能是:本组临床资料是对患腰椎神经根损害后行腰椎后路减压内固定术患者的回顾性分析,未行手术治疗者没有纳入该组资料,70岁以上患者行手术治疗风险成倍增加而大多选择保守治疗。

2.2 病因分析 结果(图4B)表明:20~29岁年龄组神经根损害患者术后恢复效果明显优于其他年龄组,随着年龄的增加,术后恢复效果递减。分析其原因可能为:年轻患者腰椎神经根损害的原因多为腰椎间盘突出,症状明显,就诊较早,早期接受受压神经根减压,取得较好疗效;年龄较大患者多为椎管狭窄症等腰椎退变性疾病,缓慢发病,就诊较晚,或者初次就诊多行保守治疗,神经根受压迫得不到及时的减压。

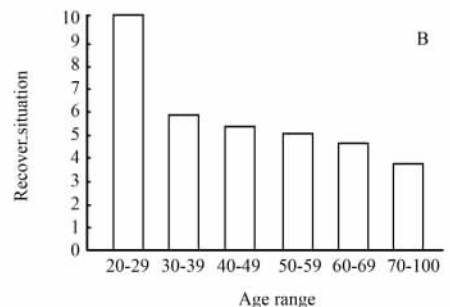
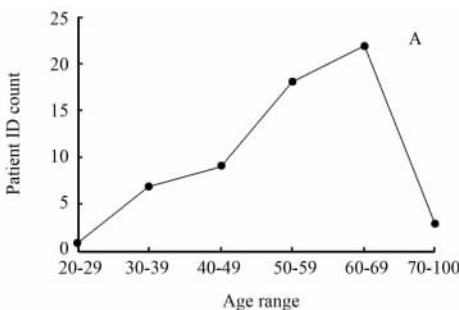


图4 年龄与就诊数量(A)、恢复程度(B)间的相关性

2.3 病程及压迫节段 结果(图 5A)表明:病程与术后疗效有着较密切的关系,发病时间短而获得减压的患者术后恢复满意,发病时间较长者术后疗效明显弱于早期手术者。神经

根损害时压迫节段也是影响预后的重要因素,其他因素相同的条件下压迫节段多者术后疗效差(图 5B)。

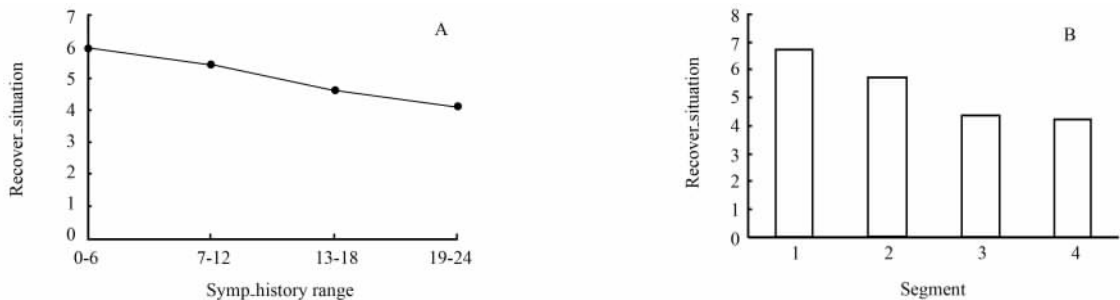


图 5 患病时间(A)、受损节段(B)与恢复程度的关系

3 讨论

本研究采用 OLAP 技术进行数据处理,从多维多个粒度上观察数据,同时建立数据模型后在该模型上进行一系列的操作,通过数据转换、数据规约、概念划分等基本技术,建立 3 维、4 维数据立方体后再进行联机分析处理操作,减少了计算量,并且可以作进一步的数据分析。经过综合分析后发现在各项相关因素的综合作用中,神经根受压迫时间较为重要,直接影响术后疗效,与临床研究结果基本类似。

在数据库中,数据存储在不同的数据表中,若采用传统的数据库技术,依据数据表的连接来进行计算时计算量比较大,并且得到结果后数据的中间计算过程不具有再利用价值。而根据 OLAP 技术,数据立方体就可以很好处理此类问题,数据立方体建立之后,可以在立方体上进行一系列的操作,得到不同的结果,使得各类查询分析的计算量减少,大大提升了工作效率,并且可以为进一步的分析研究做好铺垫。应用此种方法,可以将大宗腰椎神经根损伤患者的临床资料进行数据处理,综合分析,指导临床手术方案的制定、指导治疗、准确判断预后,对提高该类疾病的预防治疗水平有重要意义。

[参考文献]

[1] Breslow L A, Ratwani R M, Trafton J G. Cognitive models of the influence of color scale on data visualization tasks[J]. Hum Factors, 2009, 51: 321-338.

[2] Kitab S A, Miele V J, Lavelle W F, Benzel E C. Pathoanatomic basis for stretch-induced lumbar nerve root injury with a review of the literature[J]. Neurosurgery, 2009, 65: 161-167.

[3] Hey H W, Hee H T. Lumbar degenerative spinal deformity: surgical options of PLIF, TLIF and MI-TLIF[J]. Indian J Orthop, 2010, 44: 159-162.

[4] Nottmeier E W, Seemer W, Young P M. Placement of thoracolumbar pedicle screws using three-dimensional image guidance: experience in a large patient cohort[J]. J Neurosurg Spine, 2009, 10: 33-39.

[5] Bernier E, Gosselin P, Badard T, Bédard Y. Easier surveillance of climate-related health vulnerabilities through a Web-based spatial OLAP application[J]. Int J Health Geogr, 2009, 8: 18.

[6] Bédard Y, Henriques W D. Modern information technologies in environmental health surveillance. An overview and analysis [J]. Can J Public Health, 2002, 93 Suppl 1: S29-S33.

[7] Campins-Falcó P, Herráez-Hernández R, Verdú-Andrés J, Cháfer-Pericás C. On-line determination of aliphatic amines in water using in-tube solid-phase microextraction-assisted derivatization in in-valve mode for processing large sample volumes in LC[J]. Anal Bioanal Chem, 2009, 394: 557-565.

[8] Ordonez C, Chen Z. Evaluating statistical tests on OLAP cubes to compare degree of disease[J]. IEEE Trans Inf Technol Biomed, 2009, 13: 756-765.

[9] Zhou X, Chen S, Liu B, Zhang R, Wang Y, Li P, et al. Development of traditional Chinese medicine clinical data warehouse for medical knowledge discovery and decision support[J]. Artif Intell Med, 2010, 48(2-3): 139-152.

[10] Scotch M, Parmanto B, Monaco V. Evaluation of SOVAT: an OLAP-GIS decision support system for community health assessment data analysis[J]. BMC Med Inform Decis Mak, 2008, 8: 22.

[本文编辑] 贾泽军