

操作简便等特点,可作为早期诊断 MP 感染的重要指标,值得临床推广应用。

4.2 分离培养 痰、鼻和咽拭子、胸水、脑脊液、脓液等标本可用于分离培养 MP。但 MP 的分离培养阳性率不高,且需要时间长,不适于临床快速诊断,目前在临床上推广应用仍有困难。

4.3 快速诊断 聚合酶链反应(PCR)检查标本中 MP 核酸,简易、快速、特异性及敏感性较高,适用于年幼和免疫损害患儿 MP 感染的早期诊断。

5 治 疗

MP 是一种介于细菌和病毒之间的微生物,是目前能独立生活的最小的微生物,无细胞壁结构^[11]。对支原体最有抑制活性及常用于支原体感染治疗的抗菌药物是四环素类、大环内酯类及一些氟喹诺酮类。

5.1 大环内酯类是治疗支原体肺炎最主要的抗菌药物,虽然目前较常用的仍然是阿奇霉素,但在临床上已呈现疗效下降趋势,提示 MP 对阿奇霉素耐药性增强^[12]。

5.2 对红霉素耐药 MP 感染可改用四环素。其他药物如氯霉素、氯林可霉素以及氨基糖苷类抗生素体外试验虽有抑菌作用,但临床应用表明这些药物作用很弱或无效。近年有文献提出可短期使用糖皮质激素治疗支原体肺炎,可能会减少肺外并发症的发生^[13]。

5.3 支原体肺炎疗效长,一般都不少于 3 个疗程,疗程一定要足,否则容易复发,更易诱导 MP 产生耐药性。

MP 感染症状常持续或迁延反复数周或数月,因而 MP 感染也是慢性咳嗽和反复呼吸道感染重要病原之一^[14]。由于 MP 对一般治疗肺炎的药物有耐药性,因此及时进行检测有助于诊断和治疗。而 MP 感染的诊断常缺乏病原学标准,且临床表现无特异性,早期血清学抗体水平可能为阴性,而且临床上常用的 β -内酰胺类抗生素治疗无效。MP 不仅可以引起严重肺炎,还可引起严重的肺外并发症和持久的后遗症,故应充分认识其流行病学和临床特征,以利于临床早期诊断,及时治疗,改善预后。

参考文献:

[1] 游春萍. 抗体效价测定及诊断性治疗在肺炎支原体感染诊治中的应用[J]. 齐齐哈尔医学院学报, 2009, 30(14): 1671-1672.
[2] Morozumi M, Ito A, Murayama SY, et al. Assessment of realtime

PCR for diagnosis Mycoplasma pneumoniae pneumonia in pediatric patients[J]. Can J Microbiol, 2006, 52(2): 125-126.
[3] Vervloet LA, Marguet C, Camargos PA. Infection by Mycoplasma pneumoniae and its importance as an etiological agent in childhood community-acquired pneumonias[J]. raz J Infect Dis, 2007, 11(5): 507-514.
[4] Biscardi S, Lorrot M, Marc E, et al. Mycoplasma pneumoniae and asthma in children [J]. Clin Infect Dis, 2004, 38(10): 1341-1346.
[5] Hardy RD, Jafri HS, Olsen K, et al. Mycoplasma pneumonia induces chronic respiratory infection, airway hyper reactivity, and pulmonary inflammation; A murine model of infection-associated chronic reactive airway disease[J]. Infect Immunol, 2002, 70(2): 649-654.
[6] 黄汉菊. 医学微生物学[M]. 北京: 高等教育出版社, 2008: 156.
[7] Hassan J, Irwin F, Dooley S, et al. Mycoplasma pneumoniae infection in a pediatric population; analysis of soluble immune markers as risk factors for asthma[J]. Hun Immunol, 2008, 69(12): 851-855.
[8] Stemach I, Podsiadlowicz-Borzecka M, Grzolenski T, et al. Humoral and cellular immunity in children with Mycoplasma Pneumoniae infection; a 12 year prospective study[J]. Clin Diagn Lab Immunol, 2005, 12(10): 1246-1250.
[9] 陈伟红. 532 例小儿肺炎支原体感染的流行病学分析[J]. 检验医学与临床, 2010, 7(22): 2494-2495.
[10] Atkinson TP, Balish MF, Waites KB. Epidemiology, clinical manifestations, Pathogenesis and laboratory detection of Mycoplasma pneumoniae infection [J]. FEMS Microbiol Rev, 2008, 32: 956-973.
[11] 陈美元, 高翔, 竺康强. 吸入布地奈德混悬液辅治小儿支原体肺炎的疗效观察[J]. 临床肺科杂志, 2010, 15(11): 1584-1585.
[12] 陈晓凤. 阿奇霉素治疗小儿支原体肺炎疗效之变化[J]. 临床肺科杂志, 2010, 15(10): 1380-1381.
[13] Yoshinouchi T, Ohtsuki Y, Fujita J, et al. A study on intraalveolar exudates in acute mycoplasma pneumoniae infection[J]. Acta Med Okayama, 2002, 56(2): 111-116.
[14] 陈志敏. 儿童肺炎支原体感染诊治研究进展[J]. 临床儿科杂志, 2008, 26(7): 562-565.

(收稿日期: 2011-10-13)

• 综 述 •

数据挖掘的研究进展及在临床医学中的应用

周攀综述, 王开正[△]审校

(泸州医学院附属医院检验科, 四川 泸州 646000)

关键词: 数据库; 临床医学; 研究; 数据挖掘
DOI: 10. 3969/j. issn. 1673-4130. 2012. 10. 039

文献标识码: A **文章编号:** 1673-4130(2012)10-1230-03

随着信息化技术的快速发展,数据库的应用不断深化,各行各业数据呈爆炸式增长,医学领域同样如此。仅登录美国 GenBank 数据库中 DNA 序列总量于 2002 年就已超过 280 亿个碱基对^[1]。传统数据库管理系统和统计学分析方法已无法有效地分析这些数据,导致“数据爆炸、知识贫乏”的现象。在这种背景下数据挖掘(data mining, DM)应运而生。DM 是一

种将传统数据分析方法和处理大量数据的复杂算法相结合的技术,主要用于从大型数据库中提取有用信息^[2-3]。医院信息管理系统(HIS)和检验信息系统(LIS)每天都在不停地产生大量数据,这些数据日积月累,数量非常庞大。如何有效利用 DM 技术,从海量检验医学数据中发现有价值的信息是当前值得研究的课题。

[△] 通讯作者, Tel: (0830) 3165730; E-mail: wkz13808281350@163. com。

1 DM 概述

DM 是数据库知识发现(knowledge discovery in database, KDD)不可缺少的一部分,而 KDD 是将未加工的数据转换为有用信息的整个过程(图 1),包括一系列转换步骤,从数据的预处理到 DM 的后处理^[1]。其最早是在 1989 年举行的第 11 届美国人工智能协会(american association for artificial intelligence, AAAI)学术会议上提出的,是近年来随着人工智能和数据库技术的发展而出现的一门新兴技术^[4],其开发与研究应用是建立在先进的计算机技术、超大规模数据库的出现、对巨大量数据的快速访问、对这些数据应用精深的统计方法计算的能力这 4 个必要条件基础上的,以数据库、人工智能和数理统计三大技术为支柱^[5]。

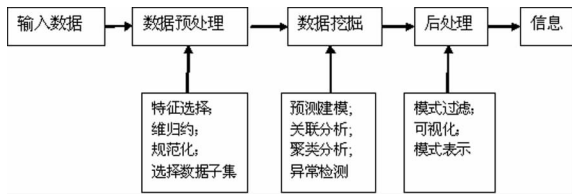


图 1 数据库 KDD 过程

2 DM 的基本模式及在临床医学中的应用

DM 的任务通常有两大类:预测任务和描述任务。预测任务主要是根据其他属性的值,预测特定属性的值,主要有分类(classification)和回归(regression)2 种模式。描述任务的目标是导出概括数据中潜在联系的模式(相关、趋势、聚类、轨迹和异常),主要有关联分析、聚类分析、异常检测 3 种模式^[1]。

2.1 预测建模(predictive modeling) 涉及以说明变量函数的方式为目标变量建立模型。有 2 种模式:分类和回归。分类是用于预测离散的目标变量。在临床医学中,疾病的诊断和鉴别诊断就是典型的分类过程。Melgani 和 Bazi^[6]以美国麻省理工学院的心律失常数据库的心电图为原始数据,采用不同分类模型,对心电图的 5 种异常波形和正常波形进行分类。回归是用于预测连续的目标变量。回归可广泛应用于医学研究中如医疗诊断与预后的判别、多因素疾病的病因研究等。Burke 等^[7]采用各种回归模式对影响乳腺癌患者预后的因素进行回归分析。

2.2 关联分析(association analysis) 用来描述数据中强关联特征的模式,用于发现隐藏在大型数据集中的令人感兴趣的联系。所发现的模式通常用蕴涵规则或特征子集的形式表示。关联分析主要应用于 DNA 序列间相似搜索与比较、识别同时出现的基因序列、在患者生理参数分析中的应用、疾病相关因素分析等^[5]。有学者对 37 000 例肾病患者进行了追踪观察,监测肾小球过滤率、尿蛋白水平和贫血状况,结果发现以上 3 种生理指标中的任何一项异常都伴随着心脏病发病率的上升,这种肾病与心脏病“关联”的现象可发生在肾病的早期阶段^[8]。

2.3 聚类分析(cluster analysis) 旨在发现紧密相关的观测值群组,使得与属于不同簇的观测值相比,属于同一簇的观测值相互之间尽可能类似。聚类分析在医学领域中主要用于 DNA 分析、医学影像数据自动分析以及多种生理参数监护数据分析、中医诊断和方剂研究、疾病危险因素等方面^[5]。罗礼溥和郭宪国^[9]利用聚类分析对云南省 25 县(市)现有的 112 种医学革螨的动物地理区划进行分析,发现云南省医学革螨的分布明显地受到自然地理区位和特定的自然景观所制约。

2.4 异常检测(anomaly detection) 用来识别其特征明显不同于其他数据的观测值。这样的观测值称为异常点(anomaly)

或离群点(outlier)。异常检测的目标是发现真正的异常点,避免错误地将正常对象标注为异常点。换言之,一个好的异常检测器必须具有高检测率和低误报率,其主要应用于检测欺诈、网络攻击、疾病的不寻常模式等^[2]。

3 DM 的方法及研究趋势

在 DM 算法的理论基础上,DM 常用方法:(1)生物学方法包括人工神经网络、遗传算法等;(2)信息论方法包括决策树等;(3)集合论方法包括粗糙集理论、近邻算法等;(4)统计学方法;(5)可视化技术等方法。DM 经过十几年的蓬勃发展,很多基本算法已较为成熟,在其基础上进行更加高效的改进和算法提高显得比较困难,如传统的频繁模式和关联规则挖掘在近几年的国际著名会议和期刊上已不再作为重要的研究主题^[10]。

近年来众多国内外知名学者相继探讨 DM 的最新方向。Yang 和 Wu^[11]汇总形成了 DM 领域十大挑战性问题报告;Agrawa 等^[12]探讨了 DM 的现状并展望了未来的发展方向;Piatetsky-shapiro 等^[13]讨论了 DM 新的挑战性问题,并主要探讨在生物信息学(bioinformatics)、多媒体挖掘(multimedia mining)、链接挖掘(link mining)、文本挖掘(text mining)和网络挖掘(web mining)等领域所遇到的挑战。与国外相比,DM 在国内的研究和应用始于 20 世纪 90 年代初,主要是对 DM 方法的介绍和推广,20 世纪 90 年代后期和 21 世纪初进入蓬勃发展阶段,当前 DM 已成为大型企业进行经营决策时所必须采用的方法,证券和金融部门已将 DM 作为今后重点应用的技术之一。有学者以 HIS 和 LIS 数据库信息为数据源,人工神经网络为工具,概率论为依据,对常规检验结果和质谱指纹图数据进行 DM 并应用于临床实践^[14-16]。

4 临床医学 DM 的特点

DM 作用于医学数据库跟挖掘其他类型的数据库相比较,具有其自己的特点。以电子病历、医学影像、病历参数、化验结果等临床数据为基础建立的医学数据库是一个复杂类型数据库,这些临床信息具有隐私性、多样性、不完整性、冗余性、异质性和缺乏数学性质等自身的特殊性和复杂性,使得医学 DM 与常规 DM 之间存在较大差异。医学 DM 方法包括统计方法、机器学习方法、神经网络方法和数据库方法等。将这些不同的挖掘方法应用到疾病的诊断、治疗和预后分析以及医疗管理等各个领域,从疾病的诊治、医疗质量管理、医院管理、卫生政策研究与医疗资源利用评价等方面去获取诸如概念、规律、模式等相关知识;用于对疾病进行分类、分级、筛选危险因素、决定治疗方案和开药数量等^[5]。

5 我国医学 DM 的现状与展望

生命科学的快速发展以及系统生物学(system biology)的出现和蓬勃发展为研究现代医学模式和中医药学提供了可能的新思路和新方法。通过基因组学、蛋白质组学等方法阐述复杂生命迫切需要 DM 等相关计算分析方法处理海量的基因、蛋白、染色质数据如基因调控网络的研究、蛋白质交互网络的挖掘等^[10]。在我国医学数据极为丰富,但运用 DM 技术分析和处理这些数据资源的研究尚处于起步阶段。有些大学(如第二军医大学、哈尔滨医科大学、泸州医学院等)已经面向医学本科生及研究生开设了相关课程,上海交通大学医学院也向医学专业研究生开设了《生物医学数据挖掘》的课程^[17],泸州医学院检验医学系开设了《检验医学信息学》课程,从检验医学信息的来源、综合、提炼和利用过程均进行了详细介绍^[18]。这些课程的开设旨在使学生及医学科学研究者了解这些知识,能理性地应用这些数学工具,并建立和其他学科领域研究人员合作的

基础。

医学 DM 是一门涉及面广、技术难度大的新兴交叉学科,是计算机技术、人工智能、统计等技术手段与现代医疗相结合的产物,需要从事计算机、统计学的科研人员与广大医务工作者之间的通力合作。随着理论研究的深入和进一步的实践摸索,医学 DM 必将在疾病的诊疗、医学科研与教学以及医院管理等方面发挥不可估量的巨大作用。

参考文献:

- [1] Benson DA, Karsch-Mizrachi I, Lipman DJ, et al. GenBank. Nucleic acids research [DB/OL]. Maryland: 2005, [2011-11-30]. <http://www.ncbi.nlm.nih.gov/Genbank/GenbankSearch.html>.
- [2] Pang-Ning T, Michael S, Vipin K. 数据挖掘导论[M]. 范明, 范宏建, 译. 北京: 人民邮电出版社, 2011: 2-3.
- [3] Han Jw, Kamber M. Data mining: concepts and techniques[M]. 2nd ed. San Francisco: Morgan Kaufmann, 2006.
- [4] 于长春, 贺佳, 范思昌, 等. 数据挖掘技术在医学领域的应用[J]. 第二军医大学学报, 2003, 24(11): 1250-1252.
- [5] 崔雷. 医学数据挖掘[M]. 北京: 高等教育出版社, 2006.
- [6] Melgani F, Bazi Y. Classification of electrocardiogram signals with support vector machines and particle swarm optimization [J]. IEEE Trans Inf Technol Biomed, 2008, 12(5): 667-677.
- [7] Burke HB, Rosen DB, Goodman PH. Comparing the prediction accuracy of artificial neural networks and other statistical models for breast cancer survival [M]// Tesouro G, Touretzky D, Leen T. Advances in neural information processing systems. Cambridge: MIT Press, 1995: 1063-1067.
- [8] McCullough P, Jurkovic C, Pergola P, et al. Independent components of chronic kidney disease as a cardiovascular risk state: resu-

hs from the Kidney Early Evaluation Program (KEEP) [J]. Arch Inter Med, 2007, 167(11): 1122-1129.

- [9] 罗礼溥, 郭宪国. 云南省医学革螨区系聚类分析[J]. 中国病原生物学杂志, 2006, 1(6): 420-423.
- [10] 王金龙. 数据挖掘研究进展[J]. 青岛理工大学学报, 2007, 28(4): 80-82.
- [11] Yang Q, Wu XD. 10 challenging problems in data mining research [J]. International Journal of Information Technology and Decision Making, 2006, 5: 597-604.
- [12] Agrawal R. Next frontier [C]// Proc. 12th ACM SIGKDD international conference on Knowledge discovery and data mining. Copenhagen: ACM Press, 2006.
- [13] Piatetsky-Shapiro G, Grossman R, Djeraba C, et al. What Are The Grand Challenges for Data Mining? KDD-2006 Panel Report [J]. SIGKDD Explorations, 2006, 8(2): 70-77.
- [14] 姜伟, 王开正, 白克镇, 等. 精神分裂症血清蛋白指纹图人工神经网络诊断模型研究[J]. 中国神经精神病学杂志, 2008, 34(1): 27-30.
- [15] 王专, 李小琼, 王开正, 等. 血清蛋白指纹图与大肠癌预后诊断[J]. 世界华人消化杂志, 2010, 18(35): 3745-3751.
- [16] 谭小林, 任兴军, 王开正, 等. 食管癌血清差异蛋白质组测定结果分析[J]. 重庆医学, 2011, 40(7): 641-646.
- [17] 龚著琳, 陈瑛, 苏懿, 等. 数据挖掘在生物医学数据分析中的应用[J]. 上海交通大学学报: 医学版, 2010, 30(11): 1420-1423.
- [18] 王开正, 陶华林. 检验医学信息学[M]. 成都: 四川科学技术出版社, 2010: 180-198.

(收稿日期: 2011-10-06)

• 综 述 •

Graves 眼病自身抗原研究进展

吴 青¹综述, 杨于力²审校

(1. 重庆医科大学附属儿童医院药剂科 400014; 2. 第三军医大学西南眼科医院, 重庆 400038)

关键词: Graves 眼病; 自身抗原; 受体, 甲状腺激素; 眼外肌抗原

DOI: 10. 3969/j. issn. 1673-4130. 2012. 10. 040

文献标识码: A

文章编号: 1673-4130(2012)10-1232-03

Graves 眼病是一种器官特异性自身免疫性疾病, 被认为是由针对自身抗原的抗体造成的自身免疫反应。目前比较公认的自身抗原为促甲状腺激素受体(TSHR), 另外有研究发现可能导致 Graves 眼病的自身抗原还包括甲状腺过氧化物酶(TPO)、甲状腺球蛋白(Tg)以及眼外肌抗原等^[1]。到目前为止, 对 Graves 眼病的各种自身抗原的本质及其相互间关系仍不清楚, 下面仅就各种 Graves 眼病的自身抗原作一简要介绍。

1 Tg

很早以前, Tg 就被认为是 Graves 眼病的自身抗原, 其是由甲状腺细胞产生的一种二聚化糖蛋白, 用于甲状腺激素的合成, 其相对分子质量为 660 000, 有 2 个相同的亚单位。Tg 共有 2 767 个氨基酸。其前蛋白为含有 19 个氨基酸的前导序列, 其后为 2 748 个氨基酸的成熟单体^[2]。有研究表明 Tg 的碳末端区域与乙酰胆碱酯酶有超过 64% 的同源性^[3]。Tao 等^[4]用人眼肌膜制备的抗原免疫小鼠获得抗 Tg 抗体, 所以他们认为人抗 Tg 抗体可以与眶内非 Tg 抗原发生交叉反应, 但

是这种交叉反应的本质并不清楚。Weetman 等^[5]进一步研究发现抗 Tg 抗体可以与眼肌乙酰胆碱酯酶发生交叉反应。他们同时提出一个假设, 在 Tg 和眼肌细胞乙酰胆碱酯酶之间的交叉反应可能在 Graves 眼病的发病机制中起重要作用。

2 TPO

TPO 参与甲状腺激素生成过程中的两个重要反应, Tg 酪氨酸的碘化以及碘化的酪氨酸的分子内交联过程, 最终形成三碘甲腺原氨酸(T3)和四碘甲腺原氨酸(T4)。微粒体蛋白位于甲状腺滤泡细胞膜表面, 被认为是 Graves 病的一种自身抗原, 导致了补体介导的细胞毒作用^[6]。虽然目前对其本质还不很清楚, 但有报道称 TPO 是微粒体蛋白的重要组成部分^[7]。Prummel 和 Wiersinga^[6]发现 TPO 抗体(TPOAb)水平与甲状腺内淋巴细胞浸润和促甲状腺素(TSH)水平均相关。朱烨^[7]等检测了 92 例甲状腺疾病患者及 66 例其他疾病患者血清中 TPOAb 及甲状腺微粒体抗体(TMAb)水平, 结果发现甲状腺疾病组两种抗体阳性率与非甲状腺疾病组比较, 差异有统计学