

DOI:10.3969/j.issn.1673-5501.2010.01.003

基于数据挖掘技术建立的儿童肺炎中医辨证规范量表的信度和效度研究

王雪峰¹ 董丹¹ 梁茂新² 张斌³ 吕玉霞⁴ 姜之炎⁵ 崔振泽⁶ 李燕宁⁷ 刘小凡⁸ 许允佳⁹
黄燕⁶ 张明卫³

摘要 目的 辨证论治是中医防治疾病最主要的方法和手段,完善现有优势病种的诊疗规范,使证候的诊断客观化、标准化和规范化,有利于中医学的推广与发展。本研究提出一种基于数据挖掘技术生成中医辨证动态量表的方法,将其应用于儿童肺炎的中医辨证规范。**方法** 选择6家中医和1家西医院作为儿童肺炎中医辨证规范研究的参研单位,将同时符合中医儿童肺炎喘嗽常证、西医儿童肺炎轻症诊断标准的患儿纳入研究组;辽宁中医药大学附属医院诊断的儿童非肺炎病例作为验证组。对所采集的儿童肺炎临床证候信息进行聚类挖掘和关联分析,生成儿童肺炎所属各证的基本构成和构成比、儿童肺炎各证所属症状的构成比,在证中证候组成模型和证间证候贡献模型的基础上构造儿童肺炎辨证量表。对量表的信度和效度进行评价。**结果** 研究组纳入1 063例,验证组纳入2 000例。通过对既往儿童肺炎中医研究文献的复习,汇总216项对儿童肺炎病证诊断有临床意义的证候变量,覆盖与儿童肺炎有关的30个方面。经过聚类处理后,有临床意义与统计学意义的指标有109项。痰热闭肺、风热闭肺、气阴两虚、阴虚肺热、痰浊未尽、痰热未清、风寒闭肺和毒热闭肺证量表的信度分别为0.742、0.583、0.666、0.258、0.370、0.559、0.797和0.861。内容效度系数较高的症状为:痰热闭肺证为小便(0.611)、咽喉症状(0.585)、口渴(0.575)、脉象/数(0.548)、厚苔(0.442)、痰/性质(0.42)、发热/性质(0.418)和咳嗽/程度(0.412);风热闭肺证为咽部症状(0.658)、咳嗽/程度(0.572)、汗/程度(0.535)和脉象/数(0.52);气阴两虚证为脉象/无力(0.876)、脉象/细(0.847)、咳嗽/声音(0.742)、腻苔(0.695)和痰/性质(0.551);阴虚肺热证为脉象/数(0.551)、痰-性质(0.535)、脉象/细(0.488)、剥脱苔(0.409);痰浊未尽证为脉象/数(0.534)、脉象/滑(0.533)、痰/量(0.41)和厚苔(0.407);痰热未清证为脉象/滑(0.611)、咳嗽/程度(0.598)、痰/性质(0.548)、痰/量(0.428)、痰/颜色(0.427)、脉象/数(0.425)和厚苔(0.406);风寒闭肺证为小便(0.927)、咽部症状(0.9)、精神神志/程度(0.879)、痰/性质(0.732)、腻苔(0.63)、哮鸣音(0.55)和脉象/紧(0.409);毒热闭肺证为咳嗽/程度(0.891)、口渴(0.824)、大便(0.813)、指纹性质(0.8)、发热/性质(0.771)、指纹程度(0.748)、腻苔(0.642)、咳嗽/声音(0.58)和发热/温度(0.481)。**结论** 数据挖掘技术支持儿童肺炎的病证结合辨证规范和辨证标准的建立,是可行的技术方法。研究所形成的儿童肺炎辨证量表是有效的辨证标准化的评估方法。

关键词 肺炎; 儿童; 数据挖掘; 辨证规范; 量表

Reliability and validity study on the scale of differentiation of symptoms and specification of traditional Chinese medicine in children pneumonia by data excavation technique

WANG Xue-feng¹, DONG Dan¹, LIANG Mao-xin², ZHANG Bin³, LV Yu-xia⁴, JIANG Zhi-yan⁵, CUI Zhen-ze⁶, LI Yan-ning⁷, LIU Xiao-fan⁸, XU You-jia⁹, HUANG Yan⁶, ZHANG Ming-wei³ (1 Affiliated Hospital of Liaoning University of TCM, Shenyang 110032, China; 2 Liaoning University of TCM, Shenyang 110032, China; 3 Northeast University, Shenyang 110032, China; 4 Affiliated Hospital of Heilongjiang University of TCM, Harbin 150040, China; 5 Longhua Affiliated Hospital of Shanghai University of TCM, Shanghai 200032, China; 6 Dalian Children's Hospital, Dalian 116012, China; 7 Affiliated Hospital of Shandong University of TCM, Jinan 250011, China; 8 Affiliated Hospital of Chengdu University of TCM, Chengdu 610072, China; 9 Guangdong Province Hospital of TCM, Guangzhou 510120, China)

基金项目 国家科技部“十五”攻关课题:2004BA721A05

作者单位 1 辽宁中医药大学附属医院 沈阳,110032;2 辽宁中医药大学 沈阳,110032;3 东北大学计算机中心 沈阳,110032;4 黑龙江中医药大学附属医院 哈尔滨,150040;5 上海中医药大学附属龙华医院 上海,200032;6 辽宁省大连市儿童医院 大连,116012;7 山东中医药大学附属医院 济南,250011;8 成都中医药大学附属医院 成都,610072;9 广东省中医院 广州,510120

通讯作者 王雪峰, E-mail:lnzywx@163.com



Corresponding Author: WANG Xue-feng, E-mail: lzzywxf@163.com

Abstract Objective The improvement of the present standardized diagnosis and treatment norms along with the unified symptom-pattern identification will objectify and standardize the diagnosis of symptoms, thus to promote and develop traditional Chinese medicine (TCM). This study presented a method that generated TCM dynamic scales based on the data excavation to be applied to the establishment of symptom differentiation standards on child pneumonia. **Methods** Seven hospitals were chosen as multicentre participating study unities, whose cases were all consistent with pneumonia diagnosis standard in western medicine and TCM. The non-pneumonia children from Affiliated Hospital of Liaoning University of TCM were taken as verification group. The collected information of clinical symptoms in children pneumonia was manipulated by clustering excavation and association analysis. The analysis of the clustering excavation results was performed to form the basic composition and constituent ratio of each symptom belonged to children pneumonia, symptomatic composition belonged to children pneumonia and differentiation of symptoms and specification scale of TCM in children pneumonia. Child pneumonia differentiation of symptoms and signs scales were produced based on the models composed by penthemeron in symptoms and contribution models composed by penthemeron between symptoms. And the reliability and validity of scales were evaluated. **Results** The study group and verification group contained 1 063 cases and 2 000 cases, respectively. Through literature review on children pneumonia of TCM, 216 symptoms and penthemeron variances were summarized, which were significant to the diagnosis of children pneumonia and covered 30 aspects related to the children pneumonia. 109 indexes with statistical and clinical significance were obtained after clustering management. The reliabilities of phlegm-heat stagnating in lung, wind-heat stagnating in lung, Qi and Yin deficiency, lung-heat due to Yin deficiency, stagnation of phlegm in the interior, accumulation of phlegm and heat, pathogenic wind-cold factors tightening the lung and pathogenic heat stagnating in lung scales were 0.742, 0.583, 0.666, 0.258, 0.370, 0.559, 0.797 and 0.861, respectively. The symptoms in high validity coefficient: phlegm-heat stagnating the lung symptoms were urination (0.611), throat symptoms (0.585), thirsty (0.575), pulse-frequent (0.548), thick tongue fur (0.442), sputum-quality (0.42), fever-quality (0.418), cough-degree (0.412); wind-heat stagnating the lung symptoms were pharyngeal portion symptoms (0.658), sweat-degree (0.535), cough-degree (0.572) and pulse-frequent (0.52); Qi and Yin deficiency the symptoms were pulse-weakness (0.876), pulse-parvus (0.847), cough-voice (0.742), slimy tongue fur (0.695), sputum-quality (0.551); lung-heat due to Yin deficiency symptoms were pulse-frequent (0.551), sputum-quality (0.535), pulse-parvus (0.488), exfoliative fur (0.409); stagnation of phlegm in the interior the symptoms were pulse-frequent (0.534), pulse-slippery (0.533), sputum-quantity (0.41), thick tongue fur (0.407); accumulation of phlegm and heat the symptoms were pulse-slippery (0.611), cough-degree (0.598), sputum-quality (0.548), sputum-quantity (0.428), sputum-color (0.427), pulse-frequent (0.425), thick tongue fur (0.406); pathogenic wind-cold factors tightening the lung symptoms were urination (0.927), throat symptoms (0.900), mind-degree (0.879), sputum-quality (0.732), slimy tongue fur (0.63), wheezing (0.55), pulse-tight (0.409); pathogenic heat stagnating the lung symptoms were cough-degree (0.891), thirsty (0.824), sedes (0.813), fingerprint quality (0.8), fever-quality (0.771), fingerprint degree (0.748), slimy tongue fur (0.642), cough-voice (0.58), fever-temperature (0.481). **Conclusions** The chiasm subjects, data excavation technique were used in the paper to offer a beneficial exploration of the differentiation of symptoms and specification of TCM. Studies showed that data excavation technique was an effective technical method to be used to support the establishment of the symptom differentiation norms and the standards for differentiation of symptoms and signs. The differentiation of symptoms scale assessment of children pneumonia was an effective standardized evaluation method.

Key words Pneumonia; Children; Data excavation; Differentiation of symptoms and specification; Scale

辨证论治是中医认识疾病和治疗疾病的基本原则,是中医学对疾病的一种特殊的研究和处理方法。从目前中医辨证论治整体状况来看,辨证种类繁多,证出多门,辨证过程存在主观性强和可重复性差等诸多问题^[1],制约了中医理论发展,如何形成统一规范、理清证型表述方式、完善现有疾病尤其是优势病种的诊疗规范,使证候的诊断客观化、标准化和规范化,将有利于中医学的发掘、继承和推广。

儿童肺炎中医称为肺炎喘嗽,多年的临床实践和基础

研究表明,中医药在改善儿童肺炎临床症状及防治儿童肺炎方面起到了重要的作用,体现了整体综合疗效的优势。本课题组在国家科技部“十五”攻关项目的资助下,先后分两期完成了有关儿童肺炎的中医系列研究,前期项目——儿童肺炎中医证治规律研究(2001AB701A16a),通过多中心、大样本和安慰剂对照的RCT研究,应用证候演变概率方法得出儿童肺炎的中医证候演变规律特点,形成了比较规范化的病证结合的综合治疗方案;后期项目——以儿童

肺炎为示范建立中医辨证规范及疗效评价方法指标体系的研究(2004 BA721A05),采用了前期项目已经证实有效的病证结合综合治疗方案作为后期项目研究的干预因素,引入了数据挖掘技术,对儿童肺炎中医辨证规范方法学的建立进行了探索性研究。

1 方法

1.1 研究路线 本研究选择性应用数据挖掘技术^[2],在密集采集儿童肺炎临床证候信息的基础上,实现基于粗糙集和信息熵的数据预处理,使用 T-HG 算法的数据聚类 and DFP-生长算法的数据关联等支持辨证规范的数据挖掘技术,在此基础上分别形成证中证候组成模型(是数据集中同一类中的数据及其类别特性分布的表示,由聚类挖掘结

果的分析生成)、证间证候贡献模型(是对证候在证间判别的贡献作用信息的表示,对证中证候组成进行熵值处理后得出证候对证的贡献)、量表辨证模型和反馈模型,共同构成本研究有机的整体。研究路线流程图如图 1 所示。

1.2 多中心参研单位资质 本研究选择国内 6 所中医药大学的附属医院,均具有较好的中西医诊疗和辨证水平;同时,出于课题的设计和将来研究成果推广的需要,还纳入一家西医院辽宁省大连市儿童医院,作为中医辨证规范化研究在西医院开展的试点。东北大学计算机中心作为本研究的数据处理单位。

1.3 样本量计算 东北大学计算机中心根据本研究儿童肺炎数据挖掘、建模和验证数据的需要,设定儿童肺炎样本量 $\geq 1\ 000$ 例,验证病例 $\geq 2\ 000$ 例。

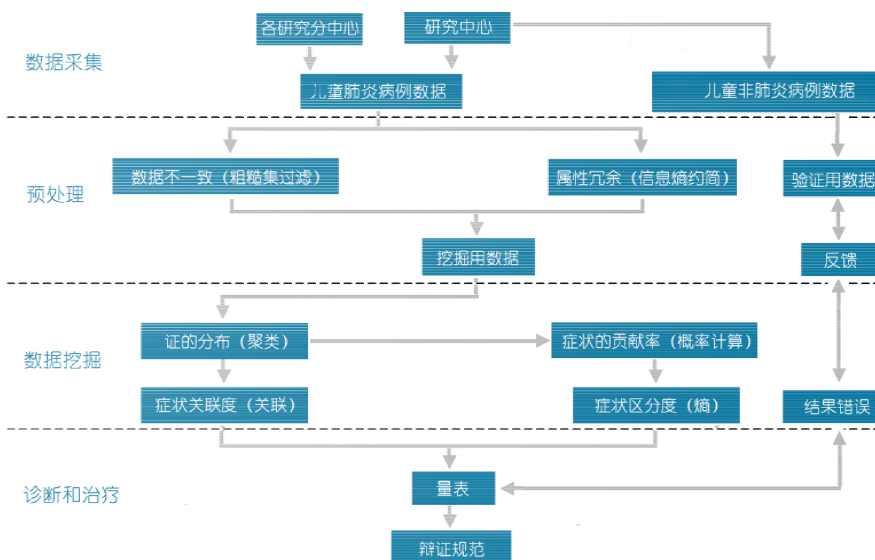


图 1 儿童肺炎中医辨证方法学建立的研究路线流程图

注 粗糙集过滤:发现和人工录入数据中的异常信息,对错误数据进行判定和过滤;信息熵约简:去除那些对挖掘工作没有意义的字段和对挖掘目标不具有决策支持作用的数据;聚类:将数据对象分组成为多个类或簇,在同一个簇中的对象之间具有较高的相似度,而不同簇中对象间的差别较大,将具有某种相同特征的数据聚在一起,根据病例数据的内部特征得到儿童肺炎客观存在的各证,并分析得出各证的分布及证的构成等信息;关联:挖掘某事件和他事件之间的依赖或关联性。如果两项或多项属性之间存在关联,那么其中一项的属性值就可以依据其他属性值进行预测,本研究用于探讨儿童肺炎各证候与各证以及各证候之间隐含的关联关系

1.4 诊断标准

1.4.1 儿童肺炎西医诊断标准^[3] 参照《小儿四病防治方案》和 1987 年全国儿童呼吸疾病会议制定的儿童肺炎(轻症)诊断标准,建立本研究儿童肺炎诊断标准如下:①急性发病;②有发热(发热程度可高可低,亦可无发热),咳嗽,喘憋,呼吸困难(如鼻翼煽动和三凹征等),婴幼儿、体弱儿及营养不良患儿可表现不明显),发绀;③肺部闻及中细湿啰音或呼吸音减弱;④胸部 X 线摄片可见斑片状阴影。

1.4.2 儿童肺炎中医辨证标准 参照《中医病证诊断疗效标准》^[4]和《中医儿科学》^[5]肺炎喘嗽常证,建立本研究中儿童肺炎喘嗽常证诊断标准,限于篇幅原因不一列出。

1.5 儿童肺炎病例的纳入标准 同时满足以下条件者被纳入:①符合西医儿童肺炎轻症诊断者;②符合中医肺炎喘嗽常证诊断者;③年龄 1~14 岁;④患儿或其监护人了解本研究的目的和实施过程,并签署书面知情同意书者。

1.6 儿童肺炎病例的排除标准 符合以下任一条者被排除:①对试验药物^[6]过敏者;②符合儿童肺炎诊断标准但合并有心、肝、肾和造血系统等严重原发性疾病的患儿;③病例入选后发现不符合纳入标准或符合排除标准者;④病例入选后自动退出或转院治疗者;⑤病例入选后因各种原因信息采集不全者。

1.7 儿童非肺炎病例的诊断和纳入标准 ①年龄 1~14



岁;②非呼吸系统疾病(非肺系病证)的住院患儿;③患儿或其监护人了解本研究目的,签署书面知情同意书者。

1.8 分组设计 研究组为7个参研单位纳入的儿童肺炎病例;验证组为儿童非肺炎病例。验证组病例用于验证数据集以检测辨证标准的准确性。为便于质量控制,减少偏倚,保证验证数据的可靠性,验证组病例由辽宁中医药大学附属医院儿科采集。

1.9 质量控制 本研究启动前,邀请国内10余位知名的中医学专家,对本课题研究路线、方法学进行了可能产生结果偏倚的论证,同时对儿童肺炎中医辨证的层次模型、数据挖掘后辨证属性判定的逻辑分析规则、图像处理技术和中医舌诊可视化、数字化分析等进行了充分地论证。本研究制定了详细的临床研究培训计划,培训内容包括:研究方案的统一认识,对舌诊、肺部听诊、胸部X线片采集和中药敷胸等操作规程进行现场培训观摩,东北大学计算机中心对病例报告表数据采集系统和数据库报送系统进行了上机操作培训。在各参研单位设立负责人和监督员,负责人负责遵循临床试验方案、把握总体质量和进度;监督员主要职责包括,确保试验记录数据准确、完整,试验过程医学伦理的监督。采用异地(网络和电话)和实地监查的方法。

1.10 治疗方法 采用本课题组前期RCT研究证实有效的综合治疗方案^[6],予西医基础治疗和中医辨证治疗。

1.11 观察指标及时间

1.11.1 观察指标 ①临床症状和体征:为尽可能穷尽儿童肺炎咳嗽临床所有症状与体征信息,收集1994至2006年国内医学期刊中儿童肺炎临床疗效观察的中医和中西医结合类文献,提取所有临床症状和体征;②客观化指标:选用统一的数码照相机采集舌象和将胸部X线片翻拍,用统一的美国3M公司生产的电子听诊器采集肺部啰音;③实验室和影像学指标:血、尿、粪常规和CRP;MP、ADV、RSV和IV血清特异性IgM抗体;摄胸部X线片。

1.11.2 观察时点 为采集到符合儿童肺炎中医证候演变规律的数据,依据儿童肺炎干预中的病程规律,每天观察症状、体征、舌诊和肺部听诊各1次。

1.12 数据收集和整理 各参研单位集中双人双份录入数据后,上传至东北大学计算机中心行盲态审核和一致性检验。本研究主要负责人、数据挖掘专家和监督员共同对数据库进行备份锁定。

1.13 证的级别划分 证的级别划分是专家结合聚类 and 关联挖掘结果进行逻辑分析后得出,逻辑分析规则为:①根据各症状及体征在证间的权重大小(即组间区别度)依次排序;②根据各症状及体征在组内的贡献率大小排出证内贡献率次序;③入选区别度排在前3位且贡献率≥0.1(10%)的症状及体征;④入选关联分析中只出现在每类前件的症状体征;⑤出现同类的2个症状体征仅有程度差别时,选择贡献率更高者;⑥出现2个或多个互相矛盾的症状体

征时,以贡献率高者入选;⑦舌象、脉象信息不参与组间权重比较,直接进入一级证中代表证的信息;⑧2个同类症状体征(例如2种脉象或舌象),在有组成复合症状体征的临床依据前提下,可酌情合并成复合症状;⑨本属正常的数据项,即便贡献率≥0.1,也不入选诊断标准之中。各证可分为:①特异证:带有疾病信息的证候,对疾病诊断具有绝对贡献作用;②一级证:携带证的信息,对证的诊断具有重要作用的证候;③二级证:参与类证鉴别的证候或疾病本脏证候;④三级证:对疾病诊断有一般支持作用的证候或疾病他脏证候。

1.14 统计学方法 所有数据在Oracle数据库中用Brio分析工具及数据挖掘工具进行分析。数据挖掘前进行基于粗糙集的数据预处理,即对症状属性进行约简,保留与挖掘有关的症状属性;采用T-HG算法的聚类挖掘用以得出儿童肺炎中医辨证规范中证的分布和症状贡献率;采用DFP-生长算法关联挖掘用以发现症状属性间的关联关系;采用克隆巴赫α法对量表信度进行考察;采用Pearson's相关系数对量表的内容效度进行考察。

2 结果

2.1 一般情况 研究组收集1072例儿童肺炎病例,其中辽宁中医药大学附属医院182例,辽宁省大连市儿童医院167例,广东省中医院150例,上海中医药大学附属龙华医院150例,成都中医药大学附属医院123例,山东中医药大学附属医院150例,黑龙江中医药大学附属医院150例。1072例中排除9例,实际纳入1063例(图2);其中男557例,女506例,年龄1~14岁。验证组纳入2000例儿童非肺炎病例,其中男1055例,女945例,年龄1~14岁。

2.2 儿童肺炎所属各证的基本构成 检索1994至2006年国内医学期刊中儿童肺炎咳嗽临床疗效观察的中医和中西医结合类文献4861篇,筛选出216项对儿童肺炎病证

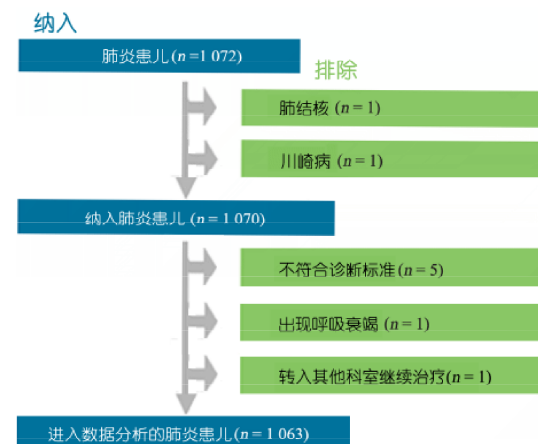


图2 儿童肺炎(咳嗽常证)纳入和排除流程图

诊断有临床意义的证候变量,覆盖与儿童肺炎喘嗽有关的30个方面,如气促、发热、咳嗽、痰、肺部听诊、汗、舌苔和脉象等。经过聚类处理后,筛选出有临床意义和统计学意义的指标109项。本研究通过聚类分析得出10个聚类类别,其中有2个类别进行人工逻辑分析后未能确定其证型归属名称,定名为未辨证1和未辨证2。由图3可以看出,发病初期,痰热闭肺证和风寒闭肺证所占的比例分别为51.5%和32.7%,随着病程的演变,在肺炎病程的3~6d呈现明

显的下降趋势;气阴两虚证、阴虚肺热证和痰热未清证在肺炎病程第3天开始升高,随着病情演变,至病程第8天,这3证所占比例分别由4.1%、1.7%和7.5%升至21.1%、7.7%和16.0%,并进入平台期;未辨证2在病程第5天呈上升趋势,至病程第10天达最高点;其余各证包括风寒闭肺证、毒热闭肺证、痰浊未尽证以及未辨证1由于临床证候信息量较少,演变趋势和规律不明显。儿童肺炎各证所属症状的构成见表1。

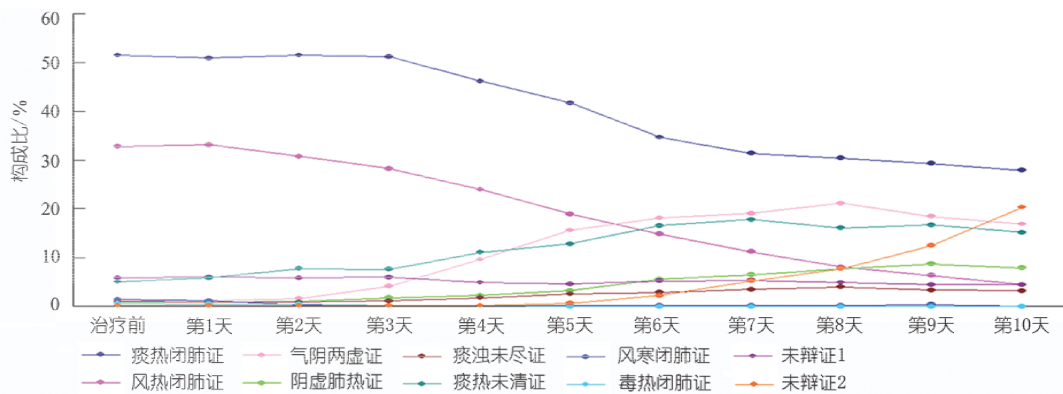


图3 儿童肺炎(喘嗽常证)所属各证的基本构成比

表1 儿童肺炎(喘嗽常证)各证所属症状的构成

证名	症状构成(症状类别,贡献率)
痰热闭肺证	中细湿啰音(0.3),中等量湿啰音(0.1),少量湿啰音(0.3);黄苔(0.5),阵咳(0.5),脉象-数(0.5),轻微咽赤(0.4),中等量痰(0.2),小便短赤(0.2),时常喉间痰鸣(0.2),厚苔(0.4),脉象/滑(0.3);痰黏稠(0.2),偶闻喉间痰鸣(0.2),渴喜冷饮(0.1);咽赤(0.3),大便干结(0.2),稍感乏力(0.2),夜卧不安(0.2),精神欠佳(0.2),盗汗(0.2),气促(0.2)
风热闭肺证	少量湿啰音(0.4),干鸣音(0.2),固定中小水泡音(0.1),中细湿啰音(0.2),中等量湿啰音(0.1);咳声重浊(0.6),食欲不振(0.5),食量减少1/3~2/3(0.5),汗出(0.2),咽部干(0.2),薄苔(0.5),数/脉象(0.5),浮/脉象(0.2),指纹风关(0.1),白苔(0.6);阵咳(0.5),有痰咯不出(0.5),自汗(0.4),轻微咽赤(0.4),咽赤(0.3),中等量痰(0.2),肿痛(0.1),频咳(0.1);痰量/少量(0.6),痰颜色/白(0.2),偶闻喉间痰鸣(0.2),口唇/红赤(0.2),体温37.3~38℃(0.1),高热(0.1)
气阴两虚证	固定中小水泡音(0.1),少量湿啰音(0.2),呼吸音粗(0.7);白苔(0.8),偶咳或声咳(0.7),微汗(0.6),薄苔(0.6),舌色/红(0.5),有自汗(0.5),咳声无力(0.4),痰颜色/白(0.3),脉象/细(0.3),盗汗(0.3),无力(0.3),腻苔(0.4)
阴虚肺热证	少苔(0.4),剥脱苔(0.2),脉象/细(0.2),脉象/数(0.4);偶咳或声咳(0.6),微汗(0.4),盗汗(0.2),咳声无力(0.1),痰颜色/黄(0.1);轻微咽赤(0.3)
痰浊未尽证	痰性质/清稀(0.1),指纹性质/淡紫隐隐(0.1),白苔(0.6),脉象/滑(0.4),舌色/红(0.4),少苔(0.1)痰量/少量(0.6),无汗(0.5),咳声重浊(0.5);偶咳或声咳(0.6),有痰咯不出(0.4)
痰热未清证	呼吸音粗(0.7);无汗(0.7),脉象/滑(0.4),痰颜色/黄(0.1),薄苔(0.5),黄苔(0.5),舌色/红(0.4);咽部干(0.1),痰性质/清稀(0.1);痰黏稠(0.2),中等量痰(0.1)
风寒闭肺证	脉象/浮(0.5),紧/脉象(0.5),流鼻涕或黄涕(0.4),稍感乏力(0.4),夜卧不安(0.3),痰黏稠(0.3),气促(0.3),体温37.3~38℃(0.3),面红(0.2),食少纳呆(0.1),腻苔(0.4);食量减少1/3~2/3(0.5),精神欠佳(0.4),口唇/红赤(0.3),干结(0.3),痰颜色/白(0.3),高热(0.3),活动减少(0.1);无汗(0.5),食欲不振(0.4),阵咳(0.4)
毒热闭肺证	呼吸音粗(0.8);痰量/少量(0.6),精神欠佳(0.5),有痰咯不出(0.5),腻苔(0.5),口唇/红赤(0.5),咽赤(0.5),频咳(0.4),干结(0.4),指纹性质/紫(0.4),渴喜冷饮(0.4),高热(0.3),偶闻喉间痰鸣(0.3),指纹性质/浮(0.2),风关(0.2),刺激性干咳(0.2),指纹性质/滞(0.2),潮红(0.2),肿痛(0.2),鼻塞(0.2),体温38.1~39℃(0.2),舌苔黄白相兼(0.1),活动减少(0.1),恶心(0.1);食欲不振(0.5),黄苔(0.5),厚苔(0.4),稍感乏力(0.4),夜卧不安(0.3),气促(0.2),体温37.3~38℃(0.2),面红(0.2);咳声重浊(0.5),食量减少1/3~2/3(0.4),微汗(0.4),自汗(0.3)

注:括号中数据为病程第1~10天的总体贡献率。表内红色字体为特异症;褐色字体为一级症;蓝色字体为二级症;黑色字体为三级症

由表 1 看出,在儿童肺炎实证阶段的证型中痰热闭肺证和风湿闭肺证的特异症信息较集中,即带有疾病信息的证候在各证型中占主导地位,毒热闭肺证带有较少的特异症信息,而风寒闭肺证无特异症信息。在肺炎实证阶段,带有肺系实热病性病机的特异症信息较集中,这与临床实际情况相符合;在儿童肺炎虚证阶段大部分证型无特异症信息或带有较少的特异症信息,如阴虚肺热证、痰浊未尽证、痰热未清证,而气阴两虚证的特异症信息相对较集中。在肺炎实证阶段,痰热闭肺证、风湿闭肺证、毒热闭肺证及风寒闭肺证中的一级症较集中,贡献率较高,而虚证中阴虚肺热证、痰浊未尽证和痰热未清证中的一级症较少,说明在肺炎虚证阶段,代表肺炎实热病性病机的证候信息逐渐减少。

2.3 儿童肺炎中医辨证量表条目说明 表 2 对儿童肺炎病程中出现的各项证候信息进行了定量或半定量划分,如肺炎主证中发热是定量划分方法,气促、咳嗽和痰等是半定量划分方法。表中证候信息及级别的划分是辨证诊断的前

提,记录了患儿在临床各个阶段症状的基本信息,同一个患儿的病历报告表根据不同证的量表可以计算出不同的分值,经过标准化处理,根据程度级别为医生提供辅助诊断。

2.4 儿童肺炎中医辨证量表证候权值 图 3 为经过聚类和相关挖掘后得出的轻度及中度证候对各证贡献率的权值,由图 3 可见,在儿童肺炎实证阶段的 4 个闭肺证型[毒热(图 3A)、风寒(图 3B)、风湿(图 3C)和痰热(图 3D)]中轻度证候分别有 16、15、14 和 17 项,中度证候分别有 4、7、5 和 4 项;而在肺炎虚证阶段的阴虚肺热证(图 3E)仅有 6 项轻度证候,没有中度级别的证候。说明肺炎实证阶段,患儿证候信息较集中,至肺炎虚证阶段则逐渐解离,只有一些轻度级别的证候,直至痊愈。毒热闭肺证、风寒闭肺证、风湿闭肺证、痰热闭肺证和阴虚肺热证正常症状权值均为 0;风寒闭肺证和痰热闭肺证重症证候口渴的权值分别为 5.175 和 3.918。痰热未清证、痰浊未尽证和气阴两虚证信息量较少,分析价值有限。

表 2 儿童肺炎(喘嗽常证)中医辨证量表

症状	正常	轻	中	重
气促/程度	呼吸平稳	气促	气急鼻翼煽动	喘憋不能平卧
呼吸次数/·min ⁻¹	18~20	~30	~40	>40
发热/温度	<37.3℃	~38℃	~39℃	>39℃
发热/性质	无热	高热	壮热	身热不扬
咳嗽/程度	偶咳或无咳	阵咳	频咳	剧咳
咳嗽/声音	咳声重浊	咳声无力	刺激性干咳	百日咳样痉咳
痰/量	无痰	少量	中等量	大量
痰/颜色	白	黄	黄白相间	带血丝
痰/性质	清稀	黏稠	有痰咯不出	有痰核或成块
痰/声音	无喉中痰鸣	偶闻喉间痰鸣	时常喉间痰鸣	痰鸣持续不断
肺部听诊/量	无	少量湿啰音	中等量湿啰音	密集湿啰音
精神神志/程度	精神佳	精神欠佳	烦躁不安	精神萎靡
汗/程度	无汗	微汗	汗出	多汗
口渴	无口渴	渴喜热饮	渴喜冷饮	渴不欲饮
面色	正常	面红	潮红	萎黄
口唇	红润	红赤	发绀或青紫	干燥
咽部现象	无咽赤	轻微咽赤	咽赤	明显咽赤
咽部症状	干	肿痛	痒	
鼻部症状	无	鼻塞	流清涕/黄涕	鼻孔干燥如烟煤
乏力	无	稍感乏力	活动减少	疲惫不堪
饮食/食欲	正常	食欲不振	食少纳呆	厌食
睡眠	正常	夜卧不安	少寐	夜不能寐
大便	正常	干结	溏薄	
小便	正常	短赤	清长	
舌色	淡红	红/淡白/紫暗		
舌苔/性质	白苔	黄苔		
舌苔/程度	薄苔	厚苔/少苔/剥脱苔/腻苔		
指纹/程度	不显于风关	风关	气关	命关
指纹性质	淡紫隐隐	红/紫/淡/滞/浮/沉		

注 为对四级分类量表的各个级别项的说明。临床根据患儿的临床证候可进行判断。表中仅列出能够划分程度级别的证候,尚有一些性质属性的证候如发热性质、咳嗽声音、痰颜色、痰声音、舌苔性质、指纹性质等证候列于二分类量表中参与辨证和疗效判定

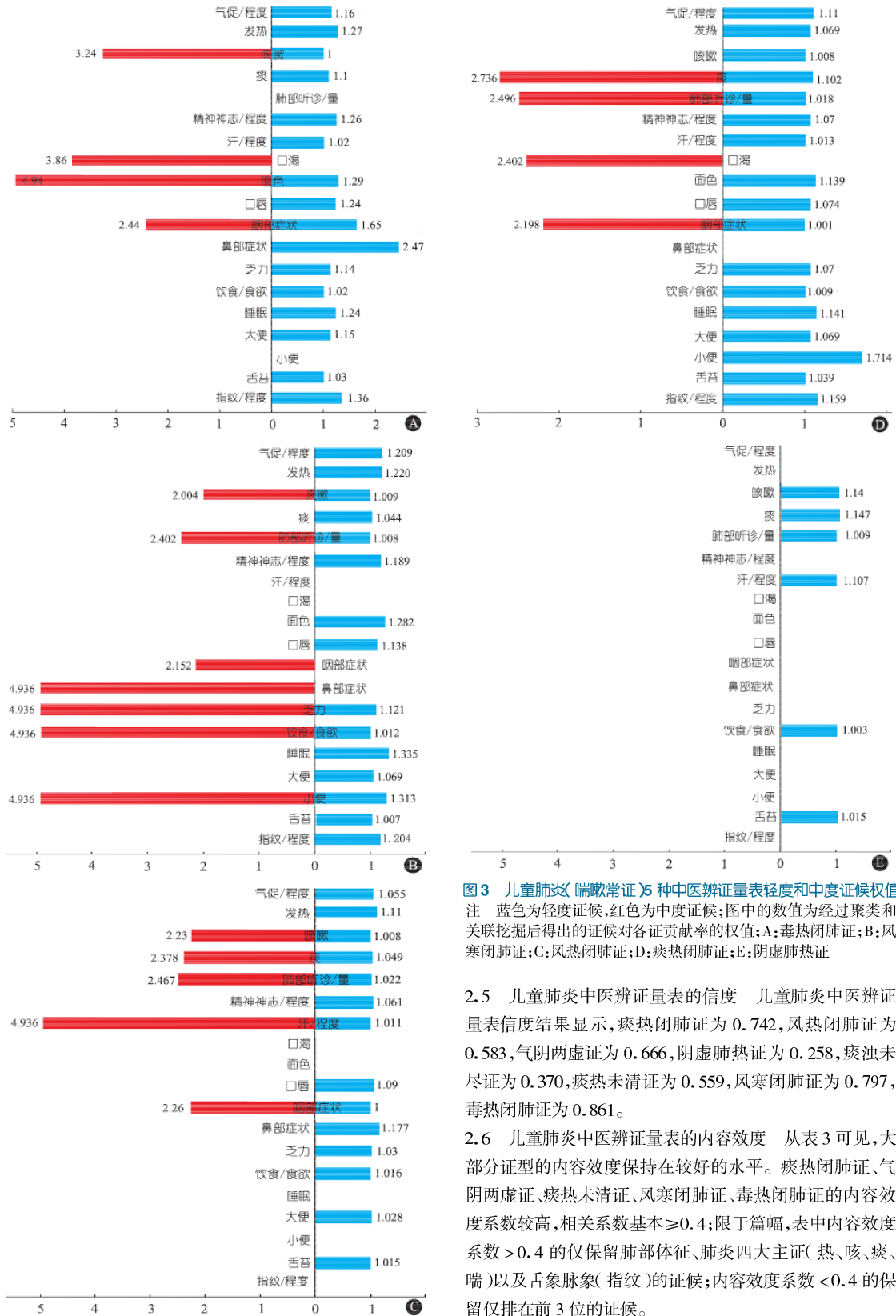


图3 儿童肺炎(喘嗽常证)5种中医辨证量表轻度及中度证候权重值
注:蓝色为轻度证候,红色为中度证候;图中的数值为经过聚类和关联挖掘后得出的证候对各证贡献率的权重;A:毒热闭肺证;B:痰热闭肺证;C:风热闭肺证;D:痰热闭肺证;E:阴虚肺热证

2.5 儿童肺炎中医辨证量表的信度 儿童肺炎中医辨证量表信度结果显示,痰热闭肺证为 0.742,风热闭肺证为 0.583,气阴两虚证为 0.666,阴虚肺热证为 0.258,痰浊未尽证为 0.370,痰热未清证为 0.559,风寒闭肺证为 0.797,毒热闭肺证为 0.861。

2.6 儿童肺炎中医辨证量表的内容效率 从表3可见,大部分证型的内容效率保持在较好的水平。痰热闭肺证、气阴两虚证、痰热未清证、风寒闭肺证、毒热闭肺证的内容效率系数较高,相关系数基本 ≥ 0.4 ;限于篇幅,表中内容效率系数 > 0.4 的仅保留肺部体征、肺炎四大主证(热、咳、痰、喘)以及舌象脉象(指纹)的证候;内容效率系数 < 0.4 的保留仅排在前3位的证候。

表3 儿童肺炎(咳嗽常证)中医辨证量表的内容效度(主要证型)

证型	效度系数 ≥ 0.4	效度系数 < 0.4
痰热闭肺证	0.611(小便);0.585(咽部症状);0.575(口渴);0.548(脉象/数);0.442(厚苔);0.42(痰/性质);0.418(发热/性质);0.412(咳嗽/程度)	0.357(肺部听诊/量);0.356(气促/程度);0.354(脉象/滑)
风热闭肺证	0.658(咽部症状);0.572(咳嗽/程度);0.535(汗/程度);0.52(脉象/数)	0.382(中细湿啰音);0.375(自汗);0.359(发热/性质)
气阴两虚证	0.876(脉象/无力);0.847(脉象/细);0.742(咳嗽/声音);0.695(腻苔);0.551(痰/性质)	0.393(厚苔);0.374(乏力);0.358(自汗)
阴虚肺热证	0.551(脉象/数);0.535(痰/性质);0.488(脉象/细);0.409(剥脱苔)	0.386(汗/程度);0.382(盗汗);0.378(咳嗽/程度)
痰浊未尽证	0.534(脉象/数);0.533(脉象/滑);0.41(痰/量);0.407(厚苔)	0.391(痰/性质);0.381(舌苔/性质);0.324(肺部听诊/量)
痰热未清证	0.611(脉象/滑);0.598(咳嗽/程度);0.548(痰/性质);0.428(痰/量);0.427(痰/颜色);0.425(脉象/数);0.406(厚苔)	0.333(痰/声音);0.33(舌苔/性质);0.322(汗/程度)
风寒闭肺证	0.927(小便);0.9(咽部症状);0.879(精神神志/程度);0.732(痰/性质);0.63(腻苔);0.55(喘鸣音);0.409(脉象/紧)	0.354(脉象/浮);0.339(中细湿啰音);0.333(发热/性质)
毒热闭肺证	0.891(咳嗽/程度);0.824(口渴);0.813(大便);0.8(指纹性质);0.771(发热/性质);0.748(指纹程度);0.642(腻苔);0.58(咳嗽/声音);0.481(发热/温度)	0.381(鼻部症状);0.291(舌苔/性质);0.283(汗/程度)

注 括号内为证所辖证候,其前面数字为内容效度分析得出的与该证型的相关系数

3 讨论

3.1 数据挖掘技术在形成中医辨证规范中应用的意义
 中医辨证规范的研究大致始于20世纪80年代中期,至今已经历了20多个年头。很多学者在该领域进行了积极的探索。对于有中医诊疗优势的病种,进行辨证规范研究,形成优势病种的中医辨证规范和标准,可以有效推动中医学实践的发展。中医临床辨证要实现客观化和规范化,进行辨证的量表评定是一种有效的辨证标准化评估方法。对既往中医量表研究的回顾表明,现有研究症状指标的确定存在主观性、量化权值标准单一及缺乏针对性等问题,即量表是静态存在的。由于疾病不同的证中包含的证候不同,同一证候在不同的证中的重要程度也不同,因而对于同一疾病的不同证型使用相同的量表,就会失去中医辨证论治的特色,也就无法实现中医辨证规范的客观化。数据挖掘技术致力于发现隐含在数据中的关于事物本质和事物发展趋势的知识或规律,并为专家的决策提供支持^[7]。将其应用于中医辨证中,不仅可以从临床数据中辨析证与证候间的复杂关系,为中医专家归纳辨证规律提供支持。其从数据出发的特性不仅能提高中医辨证和疗效评价的客观性和准确性,还有可能发现新的医学知识以丰富中医理论,将在实现中医辨证从定性描述到定量分析转变的中医定量诊断研究中发挥有效的支撑作用。

3.2 基于数据挖掘技术构成儿童肺炎证候模型及探索证候特点的创新
 对儿童肺炎所属各证构成及演变规律结果

分析表明,肺炎实证阶段各证型(包括痰热闭肺证、风热闭肺证、风寒闭肺证和毒热闭肺证)随着病程的进展呈现下降趋势,体现了明显的疾病向愈过程;肺炎虚证或恢复期各证型(包括气阴两虚证、阴虚肺热证、痰浊未尽证和痰热未清证)则随着病程进展呈现上升趋势,与临床实际情况相符。在聚类和关联挖掘后经过逻辑分析后尚有两个类别未能确定其证型归属,暂命名为未辨证1和未辨证2。对其演变趋势分析表明,未辨证1由于证候信息较少,在整个肺炎病程中曲线变化不大,无明显增高或降低的趋势;未辨证2在肺炎病程第5天开始上升,第10天达最高点,呈现一定的规律性。未辨证2中包含的证候信息是否是肺炎后期或恢复期患儿余邪未清或淤血痰浊留滞等,值得进一步研究和探讨。

中医证候轻重程度的量化是中医辨证规范的内容之一。儿童肺炎中医辨证量表条目按正常、轻、中及重程度的划分,对于受试病例的选择、证和症状疗效的判定等都具有非常重要的作用。本研究对儿童肺炎各证所属症状体征轻重程度的量化做了进一步客观化和精细化工作。研究中尽可能通过转换后,对症状程度用数字来表达。如饮食/食欲,从食量减少1/3~2/3、食量减少≥2/3等进行程度定量等。具体的量化方法:从症状出现的频率进行量化,如偶咳、阵咳、频咳和剧咳;从症状持续时间进行量化,如胸闷胀满短时间缓解、缓解时间较长及持续不解;从症状性质的程度进行量化,如精神佳、精神欠佳、烦躁不安及精神萎靡;体

征的量化,如肺部听诊的啰音量和程度等分别计量处理。这样得到的量化条目正常、轻、中和重4个级别的界定清晰,量化术语简捷明快,临床信息更趋于客观化。基于此建立起来的儿童肺炎证候模型可以对患儿的证进行重新划分,结合逻辑分析和数理统计的方法生成症状对证的贡献率表,在后续研究中可利用关联挖掘算法生成症状对疗效的伴随关系表,针对患儿病情的变化通过关联挖掘和统计分析等处理得到疗效挖掘结果,经过专家干预生成客观的疗效判定标准量表,以进行疗效判定。

在研究过程当中,本研究对数据挖掘得出的儿童肺炎辨证规范进行总结分析,同传统教材及行业标准比较,认为有以下不同:①与统编教材和行业标准相比,痰热闭肺证无发热,增加了咽赤、小便短赤、大便干结和乏力等项目。无发热的因素:一是住院患儿多口服退热药以维持体温在正常范围。二是辨证为痰热闭肺证的患儿大多不是发病即来就诊,往往在门诊或家中用药无效才入院治疗。本研究中痰热闭肺证患儿共492例,入院前平均病程8.25d,平均发热4.25d,也说明临床医生在观察纳入的病例中,采集到发热症状的概率就极大降低了。②对风寒闭肺证的辨证,挖掘结果中增加了流涕、乏力、大便情况和食欲情况等项目。这些证候均是在临床客观采集基础上,运用数学方法处理后得出的较客观结果,这说明临床医生在对儿童肺炎的实际辨证中更应关注这几方面改变的情况,为今后制定新的更符合临床实际的辨证标准,提供了客观依据。③本次挖掘结果得到的儿童肺炎中医辨证标准中证候的表述均带有程度等级,如偶咳、阵咳和频咳等。在挖掘结果分析中发现,阵咳总是出现于有闭肺属性的证中。当儿童肺炎已进入虚证阶段,即带有闭肺病机属性的证消失时,咳嗽的程度均由阵咳转为偶咳或声咳。因此,可将阵咳作为儿童肺炎中具有闭肺属性证的标志性症状。④挖掘形成的风热闭肺证中恶风信息量较少,分析其原因是由于患儿年龄小,在叙述症状时对恶风的感受表述不清,导致临床采集到的恶风症状频次较少。提示,儿科疾病辨证标准的制定应有别于成人,对某些症状如恶风、恶寒等的表述应针对儿童年龄的特点重新加以修订,以更符合临床客观实际。

3.3 基于数据挖掘技术产生儿童肺炎病证结合量表的应用价值 本研究将数据挖掘技术应用于儿童肺炎的中医辨证规范研究中,根据中医辨证的实际过程,建立了一个自下而上的中医辨证的层次结构,然后采用数据挖掘技术对这个层次结构进行了数据挖掘建模。同时,提出了一种基于数据挖掘思想的中医辨证动态量表的生成方法,根据医师在临床客观采集到的儿童肺炎中医证候信息,通过聚类处理,挖掘出证及证中证候的组成分布与贡献率,根据证候在不同证中的贡献率动态地确定各个证对应的量表的证候组成,得出证候量化程度的权值。因此,由挖掘结果设计得出的儿童肺炎中医证的动态量表,与以往的静态量表相比,症

状条目减少,症状指标的确定由数据之间内在性质决定而不是完全由人决定,症状量化权值更有针对性。限于篇幅所限,文中只列出了儿童肺炎中医辨证四级分类量表,尚有部分只能划分为有或无的二级分类证候量表如舌苔、脉象等,得出方法与四级分类证候相同,不再赘述。本研究儿童肺炎中医辨证量表中的症状有了量化分值,可以依据分值大小区分其对该证诊断的贡献率大小;同时,相同症状在不同证中由区别度大小来区分其在各类证中权重的不同。可根据临床实际采集到的数据,按照数据内部客观存在的规律初步形成了一个新的辨证标准,该辨证标准中的证候有了量化权值,初步实现了计量诊断。通过该量表对患儿病例数据的标准计算,进行标准化评估,为医生辨证诊断提供辅助信息。

3.4 儿童肺炎病证结合量表的信度和效度分析 儿童肺炎中医辨证量表生成后对其进行了信度与效度分析,用以考察量表的可靠性与有效性。信度分析采用克隆巴赫系数考察内部一致性。此系数越大,内部一致性越高,同质性越好。一般认为系数越高信度越好。本研究显示,8个证型中毒热闭肺证、风寒闭肺证、痰热闭肺证和气阴两虚证克隆巴赫系数 >0.6 ,在理想范围内;风热闭肺证和痰热未清证克隆巴赫系数接近 0.6 ,较为理想;而痰浊未尽证和阴虚肺热证克隆巴赫系数在 $0.2\sim 0.4$,由于恢复期证型证候信息量较少,故该系数较低。信度分析说明该量表有较好的内部一致性和同质性。内容效度是指所选的条目是否能够代表所要测量的内容或主题。一般认为,相关系数越高,则量表的效度越高。本研究内容效度分析结果表明,各个证的量表总体内容效度较好,各个量表的内容效度在 ≥ 0.4 的证候所占比例分别为:痰热闭肺证71%、风热闭肺证32%、气阴两虚证33%、阴虚肺热证21%、痰浊未尽证36%、痰热未清证67%、风寒闭肺证55%、毒热闭肺证69%,总体看来,内容效度保持在较好的水平,说明各个量表可以对这8个儿童肺炎证型进行评定。

3.5 本研究的局限性与前瞻性探索 规范化和标准化是科学研究的基础,也是一门学科成熟的标志。中医的辨证规范研究一直是中医界普遍关注的攻关课题之一。本研究借助于交叉学科数据挖掘技术进行了探索性的研究,力图提高中医辨证规范的客观性和准确性。研究表明,数据挖掘技术可以用来支持中医治疗优势病种的病证结合辨证规范和辨证标准的建立,是行之有效的技术方法。同时,基于辨证挖掘所形成的辨证量表,是根据症状中证的贡献率及动态变化、伴随关联关系得出的。因此,构建的辨证量表具有动态、客观和针对性强的特点,可进一步进行中医证或病证结合的疗效评价研究,得出的疗效评价结果也更符合中医证的动态变化特点,更能客观反映临床患儿病情的变化规律。

本研究在整个研究过程设定了临床密集采集时点、详

细制定各项标准操作规程,严格规定病例纳入标准等,目的是力求最大限度地保证采集到的临床信息的客观性,使研究结果趋于客观、真实、符合临床。但在研究过程中也认识到,中医许多证候信息的表达本身就具有不确定性和模糊性的特点,存在复杂的非线性关系,这些数据不同于其他领域数据的显著特征,因此加大了中医数据挖掘的难度,研究过程还不能彻底消除人工干预,比如在聚类分析后对证属性的划分、挖掘得出结果后对证候取舍规则的制定和逻辑分析等,都需要人为制定和人工干预。

本研究最终除了得到适于儿童肺炎中医辨证诊断量表之外,还得到了具有儿童肺炎动态证候变化特点的证候模型。如证中证候组成模型、证间证候贡献模型、量表辨证模型和反馈模型。由于临床上需要考察的对象是多系统或多病种的,要将本研究形成的技术和方法应用于其他疾病则需要有在现有研究基础上进行系统改进,即将现有适用于儿童肺炎的平台系统进一步开发,形成针对不同疾病共性信息的共性模块,再根据具体病证的不同信息开发个性模块,以构成一套适用于不同疾病的量表和方法系统。一个量表的优劣,一个评价系统的科学与否,还必须经过多病证、多中心的反复临床试验验证和改进,从而将本次的研究成果进一步推广,扩大其应用范围。同时,证候模型的建立是一个多次重复的过程,需要在不断地根据临床信息修正模型的过程中使其与临床客观情况趋于一致,形成更趋于成熟的模型。在不断提高证候模型准确性的同时,进一步客观和科学地评价已建立的模型,才能摸清中医证候海量信息

之间的复杂关系和内在规律。本研究探索重点在于发现新的问题,提出新的理论假说,以指导临床实践。

本刊说明 鉴于本文中中医证的英文翻译与读者阅读理解之间存在距离和困难,本文表和图没有采用本刊稿约要求的英文表达形式,属本刊特例,特此说明。

参考文献

- [1] Wang XR(王雪峰), Wang JL, Dong D, et al. 小儿肺炎临床疗效观察类文献质量分析. Chinese Journal of Practical Pediatrics(中国实用儿科杂志), 2007, 22(4): 309-312
- [2] Zhang XH(张晓红), Zhang B, Yin Y, et al. 基于数据挖掘的小儿肺炎中医辨证动态量表的生成方法. 世界科学技术-中医药现代化, 2008, 10(1): 27-34
- [3] Zhang XK(张梓荆). 全国小儿呼吸道疾病学术会议纪要. Chin J Pediatr(中华儿科杂志), 1998, 29(1): 40
- [4] 国家中医药管理局. 中医病证诊断疗效标准. 南京: 南京大学出版社, 1994
- [5] 汪受传, 主编. 中医儿科学. 北京: 中国中医药出版社, 2002
- [6] Wang XF(王雪峰), Dong D, Yu JE, et al. The randomized controlled trials of etiology analysis in 840 cases of pneumonia in children and the development regularity of TCM symptom. Chin J Evid Based Pediatr(中国循证儿科杂志), 2006, 1(3): 170-176
- [7] Jiawei Han, Micheline Kamber, 主编. 数据挖掘概念与技术. 北京: 机械工业出版社, 2001

(收稿日期: 2009-11-08 修回日期: 2009-12-30)
(本文编辑: 张崇凡)

上海交通大学医学院附属上海儿童医学中心 2010 年国家继续医学教育项目(第一批)

项目名称	项目编号	负责人	学分	举办日期
先心病围术期护理和质量管理	2010-14-04-010	傅丽娟	10	3月1~5日
小儿危重症呼吸治疗与管理	2010-06-01-027	王莹	10	4月8~12日
护理人力资源管理策略	2010-14-05-044	楼建华	10	4月20~23日
小儿神经病学基础与临床诊治进展	2010-06-01-029	王治平	10	5月18~25日
新生儿、小婴儿复杂先心病围术期管理	2010-06-02-013	史珍英	10	7月1~7日
重症监护病房质量评价体系	2010-14-04-011	杨江兰	10	7月19~23日
新生儿危重症新型医疗模式研讨	2010-06-04-017	孙建华	10	8月16~20日

全国儿科遗传代谢、内分泌疾病诊疗新进展学习班通知

随着疾病谱的改变,儿童遗传代谢和内分泌疾病日益引起重视。为加快我国儿科遗传代谢内分泌疾病专业的发展,提高对相关疾病的诊疗水平,促进学科间的交流与合作,经国家继续医学教育委员会批准,定于2010年第2季度在武汉举办全国“儿科遗传代谢、内分泌疾病诊疗新进展”学习班[2010-06-01-048(国)]。本学习班由华中科技大学附属同济医院儿科主办,学员将被授予国家级继续医学教育I类学分10分。有意参加者请与同济医院儿科联系,以便通知具体相关事宜。联系地址:武汉市解放大道1095号同济医院儿科(邮政编码430030)。联系人:魏虹、罗小平。电话/传真:027-83662393。E-mail: ekdx@tjh.tjmu.edu.cn 或 xpluo@tjh.tjmu.edu.cn