

基于临床医疗科研信息共享系统 AECOPD 证症相关性分析研究*

胡金亮¹, 李素云^{1**}, 余海滨¹, 张润顺², 周雪忠³, 张兴红¹, 李 真¹

(1. 河南中医学院第一附属医院 郑州 450000; 2. 中国中医科学研究院广安门医院 北京 100053;
3. 北京交通大学计算机与信息技术学院 北京 100044)

摘 要:目的:利用复杂网络开展慢性阻塞性肺疾病急性加重期(AECOPD)证症相关性分析研究,为AECOPD 证候诊断标准的修订提供科学依据,并探索有关技术方法的推广应用。方法:利用当前我院临床医疗科研信息共享系统,开展大数据(2011年8月~2013年1月间3 000份格式化电子病历数据)导入数据库,通过ETL软件进行数据处理,运用复杂网络进行数据挖掘。获取AECOPD常见证候常见症状的相关度MISCR值(证候与症状的相关性),并与现有标准对比分析。结果:痰热壅肺、肺肾气虚、痰浊阻肺等各证候对应的症状群与现有标准对应分析发现临床所见证候多虚症、实证夹杂,但个别症状与对应的证候相矛盾。结论:基于临床医疗科研信息共享系统,开展AECOPD证症相关性分析研究,为中医证候诊断标准的验证和修订提供有力的证据和方法学研究。

关键词:临床医疗科研信息共享系统 证候分布 证症相关性分析 复杂网络

doi: 10.11842/wst.2013.07.019 中图分类号:R241 文献标识码:A

在面临实现中医药现代化的今天,对“证候”进行规范是中医学医、教、研活动和国际交流的需要。近年来,数据挖掘等多学科交叉研究用于中医证候规范化的研究工作取得了阶段性的成果和宝贵的经验,其中慢性阻塞性肺疾病急性加重期(Acute Exacerbation Chronic Obstructive Pulmonary Disease, AECOPD)证候诊断标准的研究有多项报道^[1-8]。但这些标准还需回到临床进一步验证,而验证方法学的研究尤为重要。

中医临床医疗科研信息共享系统的应用为中医临床研究提供客观、真实大数据、多种数据挖掘方法,弥补了传统临床医学科研方法与传统统计学方法的不足,充分挖掘并利用现有的医院临床信

息,通过构建格式化电子病历模板(该模板可根据各自的专业特点进行改进),利用ETL软件把格式化电子病历、HIS、LIS等数据导入构建数据库,用复杂网络、水晶报表、VEKA等数据挖掘方法进行信息挖掘,再把挖掘出的有用信息指导临床,体现出实践到理论再回到实践的过程,本研究正是基于以上特点把共享系统作为中医证候诊断标准的验证方法学研究平台。

1 数据的收集、预处理

1.1 数据收集

对临床研究多中心数据的数据抽取、转换和导入处理,并以临床术语库为基础进行数据规范整理,形成可分析利用的临床数据集和多维数据库。

收稿日期:2013-06-19

修回日期:2013-09-24

* 国家中医药管理局2012中医药行业科研专项(HY12079675):中医临床医疗科研信息共享系统的应用及推广研究,负责人:李真。

** 通讯作者:李素云,主任医师,博士生导师,主要研究方向:慢阻肺中医证候诊断标准及疗效分析。

1.2 数据预处理

源数据的预处理 ;ODS 数据的预处理 ;细节数据仓库的预处理 ;七大规则库的预处理形成等。

2 数据挖掘

基于复杂网络的 AECOPD 证候分布有 :痰热壅肺证、痰浊阻肺证、肺肾气虚证等 22 个证候,因篇幅有限,其中以痰热壅肺证、痰浊阻肺证、肺肾气虚证排在前 3 位的证候与现行版 AECOPD 证候诊断标准进行对比分析。

2.1 基于复杂网络痰热壅肺证相关性研究

痰热壅肺证相关性数据 :与现有 AECOPD 证候诊断标准对比分析除有胸闷、咳嗽、咳痰、喘息、痰粘稠、痰色黄等实症外,对应症状还有乏力、纳呆、眠差、气短、头晕、心慌、动则加重等虚症,虚实夹杂,还有痰少、痰色白、痰难咯等相矛盾症状。结果见表 1 和图 1。

2.2 基于复杂网络肺肾气虚证相关性研究

肺肾气虚证相关性数据 :与现有 AECOPD 证候诊断标准对比分析除有咳嗽、胸闷、纳呆、眠差、乏力、喘息、气短、动则加重、痰少、头晕、心慌、多汗等虚症外,还有痰粘稠、痰难咯、痰色白、大便干结、颈背部疼痛、痰中带血、口干等实症,还有发热、小便黄、痰色黄等相矛盾的症状。结果见表 2 和图 2。

2.3 基于复杂网络痰浊阻肺证相关性研究

痰浊阻肺证相关性数据 :与现有 AECOPD 证候诊断标准对比分析符合。结果见表 3 和图 3。

此外,以上结论看到痰热壅肺、肺肾气虚证都有虚实夹杂症,这在临床多见,而肺肾气虚证有发热、痰色黄、大便干结、小便黄等实症可能患者偶因受外邪入里化热;痰热壅肺证表现出痰少、痰色白、痰难咯等相矛盾症状因患者年老体弱或疾病后期邪实正虚也会有以上症状。

表 1 痰热壅肺证相关性

节点一	节点二	频度	MISCR
喘息	痰热壅肺	64	31.205 1
动则加重	痰热壅肺	63	30.385 8
乏力	痰热壅肺	65	32.649 0
咳嗽	痰热壅肺	65	31.811 2
咳痰	痰热壅肺	65	31.811 2
眠差	痰热壅肺	64	31.402 1
纳呆	痰热壅肺	64	31.200 8
难咯	痰热壅肺	63	30.785 6
气短	痰热壅肺	63	30.785 6
痰色白	痰热壅肺	63	30.785 6
痰色黄	痰热壅肺	63	31.209 7
痰少	痰热壅肺	63	30.785 6
头晕	痰热壅肺	63	30.795 1
心慌	痰热壅肺	63	30.795 1
胸闷	痰热壅肺	65	32.213 8
质粘稠	痰热壅肺	63	30.584 4

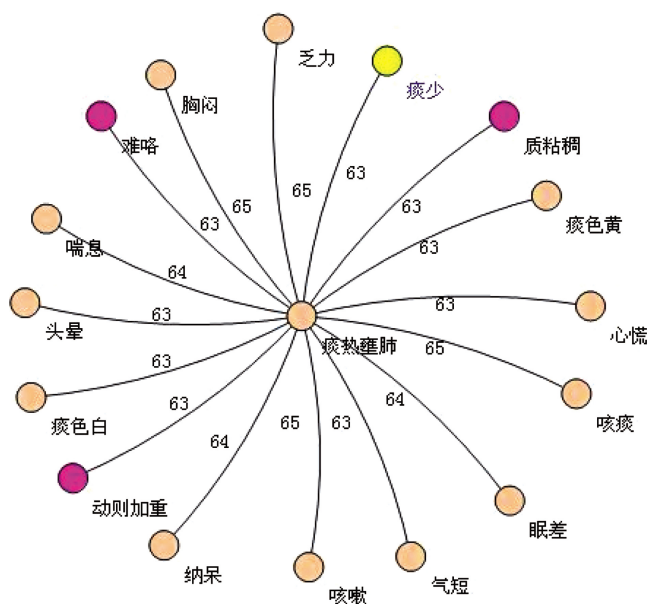


图 1 痰热壅肺证相关性图

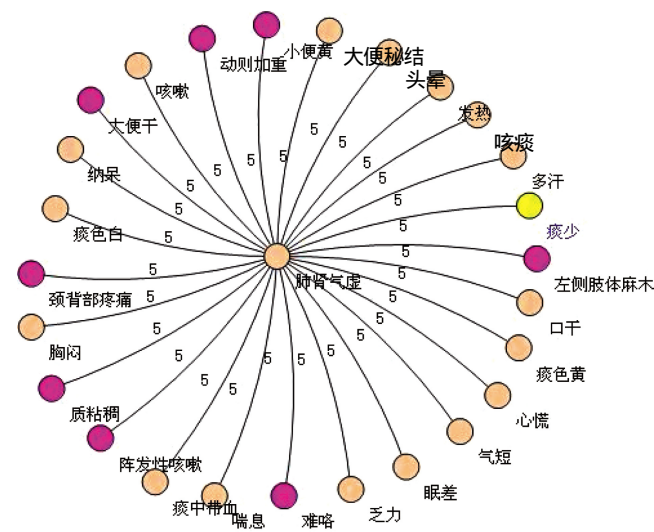


图 2 肺肾气虚证相关性

表 2 肺肾气虚证症相关性

节点一	节点二	频度	MISCR
喘息	肺肾气虚	5	0.666 8
大便干	肺肾气虚	5	0.951 4
大便干结	肺肾气虚	5	0.702 7
动则加重	肺肾气虚	5	0.675 7
多汗	肺肾气虚	5	0.707 9
发热	肺肾气虚	5	0.697 8
乏力	肺肾气虚	5	0.671 1
颈背部疼	肺肾气虚	5	1.136 8
咳嗽	肺肾气虚	5	0.654 6
咳痰	肺肾气虚	5	0.654 6
口干	肺肾气虚	5	0.743 5
眠差	肺肾气虚	5	0.671 4
纳呆	肺肾气虚	5	0.667 1
难咯	肺肾气虚	5	0.684 6
气短	肺肾气虚	5	0.684 6
痰色白	肺肾气虚	5	0.684 6
痰色黄	肺肾气虚	5	0.693 2
痰少	肺肾气虚	5	0.684 6
痰中带血	肺肾气虚	5	1.101 2
头晕	肺肾气虚	5	0.684 5
小便黄	肺肾气虚	5	0.754 1
心慌	肺肾气虚	5	0.684 5
胸闷	肺肾气虚	5	0.662 9
阵发性咳	肺肾气虚	5	1.291 5
质粘稠	肺肾气虚	5	0.680 2
左侧肢体麻木	肺肾气虚	5	1.136 8

表 3 痰浊阻肺证症相关性

节点一	节点二	频度	MISCR
喘息	痰浊阻肺	36	14.927 2
动则加重	痰浊阻肺	35	14.095 0
乏力	痰浊阻肺	35	14.002 4
反酸	痰浊阻肺	35	14.870 1
咳嗽	痰浊阻肺	36	14.647 1
咳痰	痰浊阻肺	36	14.647 1
眠差	痰浊阻肺	35	14.004 1
纳呆	痰浊阻肺	36	14.929 6
头晕	痰浊阻肺	35	14.280 6
胃脘胀痛	痰浊阻肺	35	14.870 1
心慌	痰浊阻肺	35	14.280 6
胸闷	痰浊阻肺	36	14.832 5
质粘稠	痰浊阻肺	35	14.187 2

3 讨论

通过以上数据分析提示 AECOPD 各证候对应的临床症状频数多为咳嗽、咳痰、胸闷、喘息、纳差、乏力、神疲、动则喘甚等,分析以上 3 个证候尽管不同症状出现的频次相同,但相对应的症状相关性却不同,每个证候对应的症状群均包括痰、瘀、虚,说明 AECOPD 期多病程长、多虚,久则痰瘀内生,外邪引发,印证了复杂网络具有自联想、自组织、自分类的特点,对中医复杂数据进行网络化建模,并用复杂网络的研究方法进行分析考察,揭示中医临床数据中的核心共性知识^[9-11],方法合理。

与现有标准中的相应证候对应的症状群对比分析,痰热壅肺证、肺肾气虚证都有虚实夹杂症。这些在临床较常见,我们在临床诊断时,一般通过分清寒热虚实主次来判证型。一方面以证候诊断标准作为临床诊断依据,同时考虑临床的复杂性,进一步验证现有标准。随着将来临床数据量的急剧增大及全国多中心平台的共享将为现有诊断标准的修订提供更好的方法学。

参考文献

1 李建生,王至婉,余学庆,等.慢性阻塞性肺疾病急性加重期证候诊断标准的建立.中华中医药杂志,2010,25(7):971~975.
2 李建生,王至婉,李素云,等.慢性阻塞性肺疾病稳定期证候诊断标准的建立.辽宁中医杂志,2012,39(7):1199~1202.
3 中华中医药学会内科分会肺系病专业委员会.慢性阻塞性肺疾

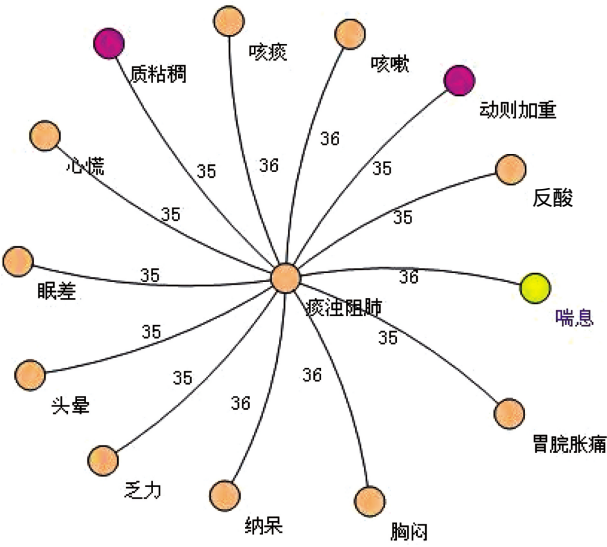


图 3 痰浊阻肺证症相关性图

- 病中医证候诊断标准(2011 版). 中医杂志, 2012, 53(2): 177~178.
- 4 Viktor Mayer-Schonberger, Kenneth Cukier 著. 盛杨燕, 周涛译. 大数据时代: 生活、工作与思维的大变革. 杭州: 浙江人民出版社, 2013: 1~247.
 - 5 Hand D, Manniala H, Smith P. Principles of Data Mining. MIT Press, Cambridge: MA, 2001.
 - 6 Berson A, Smith S, Thearling K. Building Data Mining Applications for CRM. McGraw-Hill, New York, 2000.
 - 7 Han J, Kamber M. Data Mining: Concepts and Techniques. Morgan Kaufmann, San Francisco, 2000.
 - 8 Inmon W H, Hackathorn R. Using the Data Warehouse. John Wiley & Sons, 1998: 50~150.
 - 9 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用(第一版). 北京: 清华大学出版社, 2005.
 - 10 刘保延, 周雪忠, 张润顺, 等. “医疗与临床科研信息共享系统”中医电子病历系统信息基本要求. 中国数字医学, 2012, 7(10): 58~60.
 - 11 胡镜清, 刘保延, 王永炎. 中医临床个体及诊疗信息特征与数据挖掘技术应用分析. 世界科学技术-中医药现代化, 2004, 6(1): 14~16.

Correlation Analysis of Syndromes and Symptoms of TCM AECOPD Based on Clinical Research Information Sharing System

Hu Jinliang¹, Li Suyun¹, Yu Haibin¹, Zhang Runshun², Zhou Xuezhong³, Zhang Xinghong¹, Li Zhen¹

(1. First Affiliated Hospital of Henan College of TCM, Zhengzhou 450000, China;

2. Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing 100053, China;

3. Beijing Jiaotong University, College of Computer and Information Technology, Beijing 100044, China)

Abstract: This study was aimed to analyze correlation between traditional Chinese medicine (TCM) syndromes and symptoms of acute exacerbation chronic obstructive pulmonary disease (AECOPD) based on complex network in order to provide a scientific basis for TCM syndrome amendments of diagnostic criteria, and explore relevant technologies to promote its application. By way of the current Clinical Research Information Sharing System, data and data processing were carried out (EHR 3000 between August 2011 and January 2013) into the database by ETL software. The MISCR (correlation between common symptoms and common syndromes) from both common symptoms and common syndromes of AECOPD was analyzed with existing diagnosis standards of AECOPD. The comparative analysis on TCM syndromes such as heat phlegm obstructing the lung, lung and kidney *qi* deficiency, phlegm turbid obstructing the lung and their correspondent existing diagnosis standards of AECOPD showed that the common TCM syndromes in the clinic are the mixture of deficiency and excess syndromes. However, some individual symptom is contradicted to its correspondent syndrome. It was concluded that based on the Clinical Research Information Sharing System, the research on correlation analysis between syndromes and symptoms of AECOPD can provide strong evidence and methodology for the validation and revision of TCM diagnostic criteria.

Keywords: Clinical Research Information Sharing System, syndrome distribution, correlation between common symptoms and common syndrome, complex network

(责任编辑 李沙沙 张志华, 责任译审 王 晶)

更正:

由于通讯作者笔误 现将 2013 年第 15 卷第 6 期 1480 页第一作者姓名“余海宾”更正为“余海滨”特此说明。