

# 基于关联规则结合贝叶斯网络的症状间关系及其对证候诊断贡献度的研究： 以新型冠状病毒肺炎湿热蕴肺证为例

李建生<sup>1,2</sup> 春柳<sup>1</sup> 冯贞贞<sup>1</sup> 赵虎雷<sup>1</sup> 谢洋<sup>2</sup> 孙博倩<sup>1</sup> 刘文瑞<sup>1</sup>

<sup>1</sup>河南中医药大学,呼吸疾病中医药防治省部共建协同创新中心,河南省中医药防治呼吸病重点实验室,郑州 450046; <sup>2</sup>河南中医药大学第一附属医院呼吸科,郑州 450000

通信作者:李建生, Email: li\_js8@163.com

**【摘要】** 目的 以新型冠状病毒肺炎(简称新冠肺炎)湿热蕴肺证为例,探讨证候的症状间关系及其对证候诊断的贡献度,为建立证候诊断依据提供方法学支持。方法 基于654份新冠肺炎患者的临床调查数据,以湿热蕴肺证为例,采用SPSS Modeler 14.1软件,结合关联规则与贝叶斯网络数据挖掘技术,探讨症状间关系并明确症状(群)对证候诊断的贡献度。结果 654份新冠肺炎临床资料中涉及湿热蕴肺证患者121例,其中出现频率>40%的症状有发热(53.72%)、咳嗽(47.93%)、舌质红(45.45%)、脉数(43.80%)、苔腻(42.15%)、苔黄(41.32%)、乏力(40.50%)和纳呆(40.50%)。关联规则分析显示,二项关联关系较强的症状群包括发热、口渴、胸闷、气促、咳嗽、痰黄等;三项关联关系较强的症状群包括咳嗽、痰黄、痰黏稠,纳呆、呕恶、头身困重,发热、口渴、乏力等。以湿热蕴肺证(是=1,否=0)为目标变量,以出现频率>15%的症状为输入变量,建立贝叶斯网络模型,得出湿热蕴肺证症状(群)概率分布表,其中发热的父节点(每个输入变量的上级节点)只有1个(湿热蕴肺证),条件概率是0.54;咳嗽的父节点有痰黄、湿热蕴肺证,表示在湿热蕴肺证中咳嗽与痰黄存在直接因果关系,且在有痰黄条件下,咳嗽的条件概率为0.99。常见症状(群)及其对湿热蕴肺证诊断的贡献度为:发热、口渴(0.47),咳嗽、痰黄(0.49),胸闷、气促(0.46),纳呆、头身困重(0.61),苔黄腻、脉滑数(0.95)。结论 关联规则结合贝叶斯网络在阐释症状间关系及其对证候诊断的贡献度具有一定的可行性和客观性,为建立证候诊断依据提供了方法学支持。

**【关键词】** 证候; 湿热蕴肺证; 症状; 关联规则; 贝叶斯网络; 新型冠状病毒肺炎

**基金项目:** 河南省新型冠状病毒防控应急攻关项目(201100310400, 201100310500); 中原学者科学家工作室项目(豫财行[2018]204号); 河南省高校科技创新人才支持计划(19HASTIT008)

DOI: 10.3760/cma.j.cn121430-20200619-00923

## Correlation between symptoms and their contribution to syndrome based on association rule combined with Bayesian network: syndrome of lung damp-heat accumulation in coronavirus disease 2019

Li Jiansheng<sup>1,2</sup>, Chun Liu<sup>1</sup>, Feng Zhenzhen<sup>1</sup>, Zhao Hulei<sup>1</sup>, Xie Yang<sup>2</sup>, Sun Boqian<sup>1</sup>, Liu Wenrui<sup>1</sup>

<sup>1</sup>Co-construction Collaborative Innovation Center for Chinese Medicine and Respiratory Diseases by Henan and Education Ministry of China, Henan Key Laboratory of Chinese Medicine for Respiratory Disease, Zhengzhou 450046, Henan, China;

<sup>2</sup>Department of Respiratory, the First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou 450000, Henan, China

Corresponding author: Li Jiansheng, Email: li\_js8@163.com

**【Abstract】 Objective** To explore the correlation between symptoms and their contribution to syndrome based on syndrome of lung damp-heat accumulation in coronavirus disease 2019 (COVID-19), thus to provide methodological basis for the syndrome diagnosis. **Methods** Based on 654 clinical investigation questionnaires data of COVID-19 patients, a model based on syndrome of lung damp-heat accumulation was set. Using SPSS Modeler 14.1 software, association rules and Bayesian network were applied to explore the correlation between symptoms and their contribution to syndrome. **Results** There were 121 questionnaires referring to syndrome of lung damp-heat accumulation in total 654 questionnaires. The symptoms with frequency > 40% were fever (53.72%), cough (47.93%), red tongue (45.45%), rapid pulse (43.80%), greasy fur (42.15%), yellow tongue (41.32%), fatigue (40.50%) and anorexia (40.50%). Association rule analysis showed that the symptom groups with strong binomial correlation included fever, thirst, chest tightness, shortness of breath, cough, yellow phlegm, etc. The symptom groups with strong trinomial correlation included cough, yellow phlegm, phlegm sticky, anorexia, vomiting, heavy head and body, fever, thirst, fatigue, etc. Based on SPSS Modeler 14.1 software, with syndrome of lung damp-heat accumulation (yes = 1, no = 0) as target variable, and the selected symptoms with frequency > 15.0% as input variables, the Bayesian network model was established to obtain the probability distribution table of symptoms (groups), in which there was only one parent node (the upper node of each input variable) of fever, and the conditional probability was 0.54. The parent node of cough had yellow phlegm and syndrome

of lung damp-heat accumulation, indicating that there was a direct causal relationship between cough and yellow phlegm in syndrome of lung damp-heat accumulation, and the conditional probability of cough was 0.99 under the condition of yellow phlegm. The common symptom groups and their contribution to syndrome were as follows: fever and thirsty (0.47), cough and yellow phlegm (0.49), chest tightness and polypnea (0.46), anorexia and heavy cumbersome head and body (0.61), yellow greasy fur and slippery rapid pulse (0.95). **Conclusions** It is feasible and objective to analyze the correlation between symptoms and their contribution to syndromes by association rules combined with Bayesian network. It could provide methodological basis for the syndrome diagnosis.

**【Key words】** Syndrome; Syndrome of lung damp-heat accumulation; Symptom; Association rule; Bayesian network; Coronavirus disease 2019

**Fund program:** Henan Provincial Emergency Project for Prevention and Control of Coronavirus disease 2019 of China (201100310400, 201100310500); Central Plains Scholar Scientist Studio Project of China (2018-204); Program for Science Technology Innovation Talents in Universities and Colleges of Henan Province of China (19HASTIT008)

DOI: 10.3760/cma.j.cn121430-20200619-00923

证候是由具有内在联系的症状组成的症状群总称,是对病因、病位、病性、正邪关系及病势等作出的高度概括<sup>[1]</sup>。在证候诊断过程中,除考虑单一症状对证候诊断的贡献度以外,还应考虑到有些症状组合出现时可能会明显提高其诊断的贡献度<sup>[2]</sup>。现代科技的发展和计算机的应用将为证候诊断标准的研究提供强有力的技术支持<sup>[3]</sup>。因此,在明确常见证候主要症状的前提下,进一步探讨症状间的关系及其对证候诊断的贡献度,对证候诊断依据的确立有重要意义。本课题组基于597例新型冠状病毒肺炎(简称新冠肺炎)患者的654份临床调查数据,分析新冠肺炎的10个常见证候及症状特征。本研究以新冠肺炎中常见的湿热蕴肺证为例,结合关联规则与贝叶斯网络技术,建立能确定证候的症状(群)间关系及其对证候诊断贡献度的方法,以期为实现明确证候诊断依据提供方法学支持。

## 1 资料与方法

**1.1 病例来源:**收集2020年2月1日至26日河南省救治定点医院收治的597例新冠肺炎患者654份临床资料,其中湿热蕴肺证121例。

**1.2 伦理学:**本研究符合医学伦理学标准,研究方案经河南中医药大学第一附属医院伦理委员会审核通过(审批号:2020HL-015)。

### 1.3 病例诊断、纳入和排除标准

**1.3.1 西医诊断及分型标准:**参照《新型冠状病毒感染的肺炎诊疗方案(试行第四版)》<sup>[4]</sup>和《新型冠状病毒感染的肺炎诊疗方案(试行第五版)》<sup>[5]</sup>中的有关诊断及临床分型标准。

**1.3.2 中医辨证标准:**参照相关文献<sup>[6-7]</sup>中的辨证治疗要点,制定湿热蕴肺证辨证标准:发热、汗出不解,咳嗽,痰黄,气促,胸闷,口渴,口苦,或面唇紫暗,头身困重,脘痞,腹胀,纳呆或呕恶,便溏或便秘,倦怠乏力,舌质红或紫暗,苔黄腻,脉濡数或滑数。

**1.3.3 纳入标准:**符合新冠肺炎诊断标准的湿热蕴肺证患者。

**1.3.4 排除标准:**合并严重的精神疾病、意识障碍或者处于镇静状态而无法配合完成四诊资料采集的患者。

### 1.4 数据库的建立及数据预处理

**1.4.1 数据库建立:**参考相关诊疗方案,并进行多次线上专家研讨,制定“新型冠状病毒肺炎患者中医证候临床调查表”。调查表主要包括患者基本信息、临床症状、舌苔和脉象、诊断及分型等。以项目组研发的基于移动互联网技术的临床科研信息系统平台为主,辅以纸质版调查表进行多中心调查。调查人员均为从事临床一线工作,且参与新冠肺炎救治的高年资临床医师。每个中心单独采集资料,最后统一至项目组进行数据审核录入,建立新冠肺炎证候临床调查研究数据库。数据采集过程中具体质量控制参考已发表文献<sup>[6]</sup>。

**1.4.2 数据预处理:**基于新冠肺炎证候临床调查研究数据库,建立湿热蕴肺证新冠肺炎数据库。去除无关数据和容易识别的噪声数据,删去空白数据,对空缺数据和不一致数据按病例真实情况重新输入。将各症状出现情况分为两类并赋值(症状出现=1,症状未出现=0),且所有数据标准都为标志型。

**1.5 统计学方法:**应用SPSS 22.0软件收集患者的临床资料。计量资料呈正态分布以均数±标准差( $\bar{x} \pm s$ )表示,呈偏态分布以中位数(四分位数) $[M(Q_L, Q_U)]$ 表示;计数资料以例数(%)表示。

### 1.6 数据挖掘及处理方法

**1.6.1 关联规则:**使用SPSS Modeler 14.1软件构建“数据源-类型-Apriori-网络”数据流。筛选湿热蕴肺证的症状(频率>10%)输入关联规则模型,设置最小支持度为10%,最小置信度为80%,提升度>1,指定前项中包含最大项目数为1或2。分别进行二

项和三项关联分析,以挖掘关联较强的症状组合,并分析症状间的链接程度,颜色越深代表关联性越强。

### 1.6.2 贝叶斯网络

**1.6.2.1 模型建立:**使用SPSS Modeler 14.1软件构建贝叶斯网络数据挖掘模型,类型选用“标志”,过滤后导入贝叶斯网络。结构类型选用“TAN模型”,参数学习方法选用“对小单元格计数的贝叶斯调整”,模式选用“专家”,选用“包括特征选择预处理步骤”,缺失值选用“仅使用完整记录”,显著性水平为0.01,最大条件集为5。

**1.6.2.2 贝叶斯网络:**贝叶斯网络是基于概率论和图论构建的一种不确定性知识表达推理模型<sup>[8-9]</sup>。贝叶斯网络是一个有向无环图,其中节点代表研究对象中的变量,有向弧代表变量间的因果关系,变量之间的关系度量用节点间的条件概率来表示。

贝叶斯公式: $P(A|B) = P(A)P(B|A)/P(B)$ 。式中,A表示诊断为湿热蕴肺证这一事件;B表示湿热蕴肺证中出现的某一症状; $P(A)$ 是诊断为湿热蕴肺证的先验概率; $P(B|A)$ 是条件概率,表示在已诊断为湿热蕴肺证的条件下症状B出现的概率; $P(A|B)$ 是后验概率,即当症状B出现时对诊断为湿热蕴肺证的先验概率加以修正后的概率。

当有3个或3个以上变量时,贝叶斯定理也是成立的<sup>[10-11]</sup>。公式: $P(A|B,C) = [P(C|A,B)P(A,B)/P(B)P(C|B)] = [P(A)P(B|A)P(C|A,B)] / [P(B)P(C|B)]$ 。式中, $P(A|B,C)$ 即表示在症状B、症状C同时出现时诊断为湿热蕴肺证的条件概率,在本研究中将其作为诊断贡献度。

## 2 结果

**2.1 一般资料:**121例新冠肺炎湿热蕴肺证患者均纳入分析,男性68例,女性53例;年龄19~87岁,中位年龄48(35,56)岁;发病时间(9.97±5.90)d;轻型1例(0.8%),普通型80例(66.1%),重型38例(31.4%),危重型2例(1.7%)。有武汉旅居史59例(48.8%),有确诊患者密切接触史72例(59.5%);出现频率>40%的症状依次为发热、咳嗽、舌质红、脉数、苔腻、苔黄、乏力和纳呆(表1)。

**2.2 湿热蕴肺证症状间关联规则分析:**二项关联关系较强的症状群包括发热、口渴、胸闷、气促、咳嗽、痰黄等(表2);三项关联关系较强的症状群包括咳嗽、痰黄、痰黏稠,纳呆、呕恶、头身困重,发热、口渴、乏力等(表3)。高频症状间关联规则网状图显示(图1),胸闷与气促症状的关联性最强。

表1 121例新型冠状病毒肺炎湿热蕴肺证患者出现频率>10%的症状

症状	频次(次)	频率(%)	症状	频次(次)	频率(%)
发热	65	53.72	痰黄	23	19.01
咳嗽	58	47.93	口渴	23	19.01
舌质红	55	45.45	气促	22	18.18
脉数	53	43.80	脉滑数	22	18.18
苔腻	51	42.15	气短	19	15.70
苔黄	50	41.32	便溏	19	15.70
乏力	49	40.50	咽干	18	14.88
纳呆	49	40.50	呕恶	18	14.88
痰少	39	32.23	大便秘结	18	14.88
脉滑	34	28.10	舌质紫暗	18	14.88
痰黏稠	29	23.97	口唇紫暗	18	14.88
苔厚	29	23.97	咽痛	17	14.05
胸闷	28	23.14	口苦	17	14.05
舌质淡	27	22.31	腕痞	17	14.05
痰白	25	20.66	苔白	17	14.05
干咳	25	20.66	脉濡	17	14.05
头身困重	24	19.83	神疲	14	11.57
苔黄腻	24	19.83	低热	13	10.74

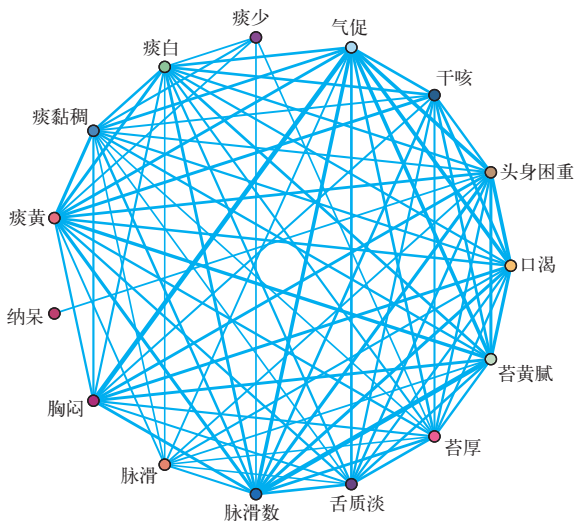
表2 新型冠状病毒肺炎湿热蕴肺证高频症状间二项关联分析

后项	前项	支持度(%)	置信度(%)	提升度
咳嗽=1.0	痰黄=1.0	19.01	100.00	2.09
胸闷=1.0	气促=1.0	18.18	100.00	4.32
舌质红=1.0	舌质紫暗=1.0	14.88	100.00	2.20
发热=1.0	口渴=1.0	19.01	95.65	1.78
纳呆=1.0	呕恶=1.0	14.88	94.44	2.33
脉数=1.0	脉濡=1.0	14.05	94.12	2.15
纳呆=1.0	头身困重=1.0	19.83	91.67	2.26
纳呆=1.0	大便秘结=1.0	14.88	88.89	2.20
咳嗽=1.0	痰黏稠=1.0	23.97	86.21	1.80
乏力=1.0	口唇紫暗=1.0	14.88	83.33	2.06
苔黄腻=1.0	脉滑数=1.0	18.18	81.82	4.13

表3 新型冠状病毒肺炎湿热蕴肺证高频症状间置信度>90%的三项关联分析

后项	前项	支持度(%)	置信度(%)	提升度
咳嗽=1.0	痰黄=1.0和痰黏稠=1.0	14.88	100.00	2.09
纳呆=1.0	呕恶=1.0和头身困重=1.0	11.57	100.00	2.47
发热=1.0	口渴=1.0和乏力=1.0	11.57	100.00	1.86
发热=1.0	口渴=1.0和纳呆=1.0	11.57	100.00	1.86
发热=1.0	口渴=1.0和舌质红=1.0	11.57	100.00	1.86
咳嗽=1.0	痰黄=1.0和痰少=1.0	10.74	100.00	2.09
胸闷=1.0	气促=1.0和乏力=1.0	10.74	100.00	4.32
胸闷=1.0	气促=1.0和纳呆=1.0	13.22	100.00	4.32
胸闷=1.0	气促=1.0和发热=1.0	10.74	100.00	4.32
纳呆=1.0	头身困重=1.0和发热=1.0	14.88	94.44	2.33
气促=1.0	胸闷=1.0和纳呆=1.0	14.05	94.12	5.18
咳嗽=1.0	痰黏稠=1.0和苔黄=1.0	12.40	93.33	1.95
发热=1.0	口渴=1.0和苔腻=1.0	11.57	92.86	1.73
发热=1.0	口渴=1.0和脉数=1.0	11.57	92.86	1.73
脉数=1.0	脉濡=1.0和脉滑=1.0	10.74	92.31	2.11
纳呆=1.0	呕恶=1.0和发热=1.0	10.74	92.31	2.28
痰黏稠=1.0	痰黄=1.0和痰少=1.0	10.74	92.31	3.85





注：节点（症状）之间形成无数个复杂的关系，将网状图设置为强链接较重，通过强弱不同的连接线表示症状间的关联程度，线型越粗代表症状之间的关联性越大

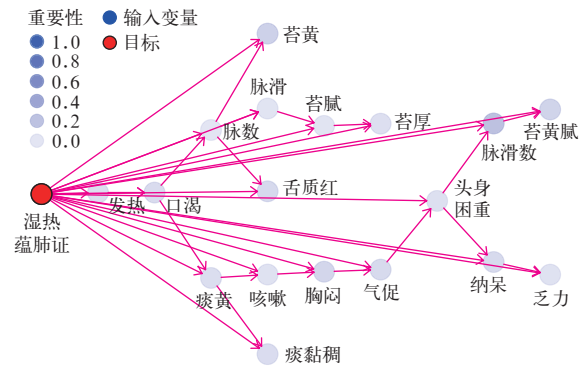
图1 新型冠状病毒肺炎湿热蕴肺证高频症状间关联规则网状图

2.3 贝叶斯网络判别结果(表4;图2):以湿热蕴肺证(是=1,否=0)为目标变量,以出现频率>15%的症状为输入变量,建立贝叶斯网络模型,得出湿热蕴肺证症状(群)概率分布表。其中,发热的父节点(每个输入变量的上级节点)只有1个(湿热蕴肺证),条件概率是0.54;咳嗽的父节点有痰黄、湿热蕴肺证,表示湿热蕴肺证中咳嗽与痰黄存在直接因果关系,且在有痰黄条件下,咳嗽的条件概率为0.99。

表4 新型冠状病毒肺炎湿热蕴肺证症状(群)贝叶斯网络条件概率

子节点	父节点	条件概率
咳嗽	痰黄、湿热蕴肺证	0.99
苔黄腻	脉滑数、湿热蕴肺证	0.81
痰黏稠	痰黄、湿热蕴肺证	0.78
气促	胸闷、湿热蕴肺证	0.78
纳呆	气促、湿热蕴肺证	0.72
脉数	口渴、湿热蕴肺证	0.61
舌质红	脉数、湿热蕴肺证	0.57
苔腻	脉滑、湿热蕴肺证	0.56
头身困重	纳呆、湿热蕴肺证	0.55
发热	湿热蕴肺证	0.54
苔黄	脉数、湿热蕴肺证	0.51
口渴	发热、湿热蕴肺证	0.50
乏力	头身困重、湿热蕴肺证	0.46
脉滑	脉数、湿热蕴肺证	0.38
脉滑数	纳呆、湿热蕴肺证	0.31
苔厚	苔腻、湿热蕴肺证	0.28
胸闷	咳嗽、湿热蕴肺证	0.19
痰黄	口渴、湿热蕴肺证	0.01

2.4 常见症状(群)的确定及对证候诊断的贡献度:将同时满足以下条件的症状组合作为新冠肺炎湿热



注：红色节点代表目标变量,是其余节点的父节点(每个输入变量的上级节点);蓝色节点代表输入变量;每个节点通过条件概率与其父节点相关;如2个节点间有箭头连接,说明两者间有因果联系

图2 新型冠状病毒肺炎湿热蕴肺证症状的贝叶斯网络图

蕴肺证常见症状(群):①关联规则中支持度>10%、置信度>80%及提升度>1的二项关联组合;②贝叶斯网络条件概率≥0.5的症状组合。最终得出5组常见症状(群):发热、口渴,咳嗽、痰黄,胸闷、气促,纳呆、头身困重,苔黄腻、脉滑数。经贝叶斯公式推导得出每组症状(群)出现时发生湿热蕴肺证的条件概率,作为对湿热蕴肺证诊断的贡献度,其中苔黄腻和脉滑数的贡献度最高,为0.95;胸闷和气促的贡献度最低,为0.46(表5)。

表5 新型冠状病毒肺炎湿热蕴肺证症状(群)对证候诊断的贡献度

症状(群)	证候	贡献度
苔黄腻、脉滑数	湿热蕴肺证	0.95
纳呆、头身困重	湿热蕴肺证	0.61
咳嗽、痰黄	湿热蕴肺证	0.49
发热、口渴	湿热蕴肺证	0.47
胸闷、气促	湿热蕴肺证	0.46

### 3 讨论

中医证候规范化研究对发展中医理论和提高临床诊治水平有极其重要的意义。每个证候都有其特异性的症状,可以是一个或几个症状或者几个有关症状形成的症状群。以往研究多针对单一症状或症状间的关系,缺乏症状或症状(群)对证候诊断贡献度的进一步研究<sup>[12-14]</sup>。相关症状间、症状群之间的关系及其对诊断的贡献度如何确定、各证候症状群形成的依据及其对证候诊断的贡献大小如何等尚需建立适宜方法解决<sup>[15]</sup>。新冠肺炎已成为国际公认的严重危害人类健康的重大公共卫生事件<sup>[16-17]</sup>。中医药在治疗新冠肺炎方面存在一定优势<sup>[18-20]</sup>。本研究以新冠肺炎湿热蕴肺证为例,探讨湿热蕴肺证症状的关系及各个症状(群)对该证候诊断的贡

献度,为证候诊断依据的建立提供适宜方法。

数据挖掘技术重视症状间的相互作用并充分考虑到中医证候复杂性的特点,在中医证候研究中扮演着越来越重要的角色。近年来,关联规则<sup>[21-25]</sup>和贝叶斯网络<sup>[26-29]</sup>在中医证候研究中得到了一定程度的应用。关联规则是处理相关关系中最为常用和成熟的方法之一<sup>[13]</sup>,通常使用支持度、置信度和提升度来衡量关联规则<sup>[30-31]</sup>,若支持度太低,表明症状群出现机会很少,缺乏特异性;置信度太低,表明症状群的可信度差;提升度则是指前项对后项的影响程度。贝叶斯网络包括有向无环网络图和条件概率表。网络中的每个节点表示1个变量,即1个症状,各症状之间的弧表示二者存在直接因果关系,条件概率表可以定量表示这些因果关系的强度。贝叶斯网络具有双向推理的特点<sup>[32]</sup>,对于2个以上的变量,可通过贝叶斯公式推导出它们同时存在时证候发生的条件概率。关联规则仅用于探讨症状间的关联程度,无法判别其对证候诊断的贡献度;贝叶斯网络仅考虑两两属性间的关联性,忽略了其他关联性。单一数据挖掘技术的运用尚不能全面解决症状间关系及其对证候诊断贡献度的问题。因此,本研究结合上述两种数据挖掘技术,依据关联规则强弱程度及贝叶斯网络的因果关系,筛选出新冠肺炎湿热蕴肺证常见症状(群),并通过条件概率得出症状(群)对证候诊断的贡献度。

本研究通过在关联规则中设定最小规则支持度为10%,最小规则置信度为80%,提升度>1,得出新冠肺炎湿热蕴肺证二项关联关系较强的症状(群)包括发热、口渴、胸闷、气促、咳嗽、痰黄等。在贝叶斯网络中通过节点之间的弧确定症状之间的联系,将同时符合关联规则强关联组合和贝叶斯网络直接因果关系(条件概率 $\geq 0.5$ )的症状组合设定为湿热蕴肺证的常见症状(群),如咳嗽和痰黄支持度为19.01%、置信度为100%、提升度为2.09,且在贝叶斯网络中二者存在直接因果关系(条件概率为0.99)。通过贝叶斯公式推导得出咳嗽和痰黄对湿热蕴肺证诊断的贡献度为0.49。据此,初步形成新冠肺炎湿热蕴肺证诊断依据:发热、口渴、咳嗽、痰黄、胸闷、气促、纳呆、头身困重、苔黄腻、脉滑数,诊断贡献度依次为0.47、0.49、0.46、0.61、0.95。该初步诊断依据有待专家形成共识并进行临床验证。

本研究尚存在一定不足:贝叶斯网络可显示两两属性间的关联性,但不能显示属性之间可能存在

2个以上症状的关联性,如湿热蕴肺证中咳嗽和痰黄、纳呆和头身困重为常见症状(群),关联规则中咳嗽、痰黄和痰黏稠,纳呆、呕恶和头身困重两组症状为强关联组合,但无法计算出这两个症状组合对该证候诊断的贡献度。因此,是否将“咳嗽、痰黄、痰黏稠”以及“纳呆、头身困重、呕恶”作为新冠肺炎湿热蕴肺证的诊断依据更为合理,尚有待研究。

利益冲突 所有作者均声明不存在利益冲突

## 参考文献

- [1] 朱春梅,顾巍杰,杨德才,等.基于症状/体征群的慢性胃炎常见证候特征选择[J].中国中医药信息杂志,2019,26(7):16-21. DOI: 10.3969/j.issn.1005-5304.2019.07.005.
- [2] Zhu CM, Gu WJ, Yang DC, et al. Feature selection of common syndromes of chronic gastritis based on feature groups [J]. Chin J Int Tradit Chin Med, 2019, 26 (7): 16-21. DOI: 10.3969/j.issn.1005-5304.2019.07.005.
- [3] 赵勇,邹小娟,白赞,等.证候量化诊断标准的探讨与研究[J].世界中西医结合杂志,2010,5(4):354-355. DOI: 10.3969/j.issn.1673-6613.2010.04.035.
- [4] Zhao Y, Zou XJ, Bai Y, et al. Discussion and research on quantitative diagnostic criteria of syndromes [J]. World J Integr Tradit West Med, 2010, 5 (4): 354-355. DOI: 10.3969/j.issn.1673-6613.2010.04.035.
- [5] 胡金亮,李建生,李素云.基于动态模糊神经网络的中医证候诊断标准建立方法应用研究[J].中医学报,2012,27(9):1136-1138. DOI: 10.16368/j.issn.1674-8999.2012.09.022.
- [6] Hu JL, Li JS, Li SY. Traditional Chinese medicine syndrome diagnostic standard building method applied research based on the dynamic fuzzy neural network [J]. J Chin Med, 2012, 27 (9): 1136-1138. DOI: 10.16368/j.issn.1674-8999.2012.09.022.
- [7] 国家卫生健康委员会.新型冠状病毒肺炎诊疗方案(试行第四版)[EB/OL].(2020-01-27)[2020-04-05].http://www.nhc.gov.cn/yzygj/s7653p/202001/4294563ed35b43209b31739bd0785e67.shtml.
- [8] National Health Commission of the People's Republic of China. Diagnosis and treatment of pneumonia caused by novel coronavirus (trial version 4) [EB/OL]. (2020-01-27) [2020-04-05]. http://www.nhc.gov.cn/yzygj/s7653p/202001/4294563ed35b43209b31739bd0785e67.shtml.
- [9] 国家卫生健康委员会.新型冠状病毒肺炎诊疗方案(试行第五版)[EB/OL].(2020-02-05)[2020-04-05].http://www.nhc.gov.cn/yzygj/s7653p/202002/3b09b894ac9b4204a79db5b8912d4440.shtml.
- [10] National Health Commission of the People's Republic of China. Diagnosis and treatment of pneumonia caused by novel coronavirus (trial version 5) [EB/OL]. (2020-02-05) [2020-04-05]. http://www.nhc.gov.cn/yzygj/s7653p/202002/3b09b894ac9b4204a79db5b8912d4440.shtml.
- [11] 李建生,李素云,谢洋,等.河南省524例新型冠状病毒肺炎患者临床特征及中医证候分布[J/OL].中医杂志,2020:1-6[2020-03-27].[优先发表].
- [12] Li JS, Li SY, Xie Y, et al. Clinical characteristics and distribution of TCM syndromes of 524 patients with novel coronavirus pneumonia in Henan Province [J/OL]. J Tradit Chin Med, 2020: 1-6 [2020-03-27]. [published online ahead of print March 12, 2020].
- [13] 李建生,李素云,谢洋.河南省新型冠状病毒肺炎中医辨证治疗思路与方法[J].中医学报,2020,35(3):453-457. DOI: 10.16368/j.issn.1674-8999.2020.03.101.
- [14] Li JS, Li SY, Xie Y. Ideas and methods of TCM syndrome differentiation and treatment for COVID-19 in Henan Province [J]. J Chin Med, 2020, 35 (3): 453-457. DOI: 10.16368/j.issn.1674-8999.2020.03.101.
- [15] Verron S, Tiplica T, Kobi A. Fault detection and identification with a new feature selection based on mutual information [J]. J Process Control, 2008, 18 (5): 479-490. DOI: 10.1016/j.jprocont.2007.08.003.
- [16] Straub D, Der KA. Bayesian network enhanced with structural reliability methods: methodology [J]. J Eng Mech, 2010, 136 (10): 1248-1258. DOI: 10.1061/(ASCE)EM.1943-7889.0000173.
- [17] 钱勤.基于贝叶斯智能分类的改进垃圾短信拦截平台构建[D].



- 上海:上海师范大学,2013.
- Qian Q. Bayes-based intelligent classification spam SMS interception platform build [D]. Shanghai: Shanghai Normal University, 2013.
- [11] 徐静. 北京一日游参团游客满意度测量及相关问题研究 [D]. 北京:北京交通大学,2018.
- Xu J. Measurement and related research on daycation tourist satisfaction [D]. Beijing: Beijing Jiaotong University, 2018.
- [12] 王伟杰,唐晓颇,王新昌,等. 基于临床辨证的类风湿关节炎常见中医证候 Logistic 回归分析 [J]. 中华中医药杂志, 2019, 34 (2): 807-810.
- Wang WJ, Tang XP, Wang XC, et al. Logistic regression analysis on TCM syndromes of rheumatoid arthritis based on syndrome differentiation [J]. CJTCMP, 2019, 34 (2): 807-810.
- [13] 王至婉,闫蔷薇,李建生,等. 慢性阻塞性肺疾病急性加重期各证候症状间关联模式研究 [J]. 中医杂志, 2018, 59 (24): 2120-2125. DOI: 10.13288/j.11-2166/r.2018.24.012.
- Wang ZW, Yan QW, Li JS, et al. Study on association pattern among syndromes in acute exacerbation stage of chronic obstructive pulmonary disease [J]. J Tradit Chin Med, 2018, 59 (24): 2120-2125. DOI: 10.13288/j.11-2166/r.2018.24.012.
- [14] 郝闻致,龚栋,薛飞飞,等. 基于扎根理论与内容分析法的肝郁证候定性定量研究方法探析 [J]. 中华中医药杂志, 2020, 35 (2): 607-610.
- Hao WZ, Gong L, Xue FF, et al. Analysis of the qualitative and quantitative research methods of liver stagnation syndrome based on grounded theory and content analysis [J]. CJTCMP, 2020, 35 (2): 607-610.
- [15] 李建生,余学庆,王至婉. 病证结合模式下证候诊断标准建立的关键环节 [J]. 中医杂志, 2013, 54 (15): 1261-1264. DOI: 10.13288/j.11-2166/r.2013.15.003.
- Li JS, Yu XQ, Wang ZW. Key links of establishing syndrome diagnosis standard based on the combination of disease and syndrome [J]. J Tradit Chin Med, 2013, 54 (15): 1261-1264. DOI: 10.13288/j.11-2166/r.2013.15.003.
- [16] 张嘉皓,李娟,苏连久,等. 新型冠状病毒肺炎相关性急性肾损伤的临床特点及危险因素分析 [J]. 中华危重病急救医学, 2020, 32 (4): 407-411. DOI: 10.3760/cma.j.cn121430-20200302-00198.
- Zhang JH, Li J, Su LJ, et al. Clinical characteristics and risk factors of acute kidney injury in coronavirus disease 2019 [J]. Chin Crit Care Med, 2020, 32 (4): 407-411. DOI: 10.3760/cma.j.cn121430-20200302-00198.
- [17] 刘昶权,邓西龙,潘越峻,等. 不同临床分型新型冠状病毒肺炎患者临床特征及肺部 CT 影像学特征分析 [J]. 中华危重病急救医学, 2020, 32 (5): 548-553. DOI: 10.3760/cma.j.cn121430-20200323-00203.
- Liu CQ, Deng XL, Pan YJ, et al. Clinical characteristics and CT imaging features of patients with different clinical types of coronavirus disease 2019 [J]. Chin Crit Care Med, 2020, 32 (5): 548-553. DOI: 10.3760/cma.j.cn121430-20200323-00203.
- [18] 李春艳,倪晋,李国妮,等. 延安市 4 例新型冠状病毒肺炎患者中西医结合治疗分析 [J]. 中国中西医结合急救杂志, 2020, 27 (2): 206-209. DOI: 10.3969/j.issn.1008-9691.2020.02.019.
- Li CY, Ni J, Li GN, et al. Analysis on 4 cases of corona virus disease 2019 with integrated Chinese and Western medicine treatment in Yan'an City [J]. Chin J TCM WM Crit Care, 2020, 27 (2): 206-209. DOI: 10.3969/j.issn.1008-9691.2020.02.019.
- [19] 刘小发,郭登洲,杨倩,等. 基于浊毒理论防治新型冠状病毒肺炎的辨证思路 [J]. 中国中西医结合急救杂志, 2020, 27 (1): 13-14. DOI: 10.3969/j.issn.1008-9691.2020.01.004.
- Liu XF, Guo DZ, Yang Q, et al. Dialectical thinking of preventing novel coronavirus pneumonia based on turbid toxin theory [J]. Chin J TCM WM Crit Care, 2020, 27 (1): 13-14. DOI: 10.3969/j.issn.1008-9691.2020.01.004.
- [20] 仕丽,王檀,方雅堃,等. 中医药治疗新型冠状病毒肺炎病例分析及思考 [J]. 吉林中医药, 2020, 40 (7): 841-846. DOI: 10.13463/j.cnki.jlzyy.2020.07.001.
- Shi L, Wang T, Fang YK, et al. Analysis and thinking of COVID-19 treated by traditional Chinese medicine [J]. Jilin J Tradit Chin Med, 2020, 40 (7): 841-846. DOI: 10.13463/j.cnki.jlzyy.2020.07.001.
- [21] 孙媛媛,王明春,刘恩顺,等. ARDS 证候特征的关联规则研究 [J]. 天津职业技术师范大学学报, 2013, 23 (3): 29-31, 35. DOI: 10.3969/j.issn.2095-0926.2013.03.008.
- Sun YY, Wang MC, Liu ES, et al. TCM syndrome research of ARDS based on association rules [J]. J Tianjin Univ Technol Educ, 2013, 23 (3): 29-31, 35. DOI: 10.3969/j.issn.2095-0926.2013.03.008.
- [22] 赵艳青,李青松,黄丽,等. 干眼中医证型分布规律及症状相关性研究 [J]. 中国中医眼科杂志, 2019, 29 (1): 19-24. DOI: 10.13444/j.cnki.zgzyygz.2019.01.005.
- Zhao YQ, Li QS, Huang L, et al. Study on distribution pattern of TCM syndromes of xerophthalmia and correlation between symptoms [J]. Chin J Chin Ophthalmol, 2019, 29 (1): 19-24. DOI: 10.13444/j.cnki.zgzyygz.2019.01.005.
- [23] 艾军,汪受传,戴铭,等. 基于症状关联规则的小儿肺炎中医证候病机学研究 [J]. 北京中医药大学学报, 2016, 39 (2): 140-143. DOI: 10.3969/j.issn.1006-2157.2016.02.011.
- Ai J, Wang SC, Dai M, et al. Study of TCM patterns and mechanism for pediatric pneumonia based on symptom correlation [J]. J Beijing Univ Tradit Chin Med, 2016, 39 (2): 140-143. DOI: 10.3969/j.issn.1006-2157.2016.02.011.
- [24] 张世君,齐冬梅,李运伦,等. 中医证候规范化研究中无监督方法探讨 [J]. 长春中医药大学学报, 2016, 32 (1): 1-5. DOI: 10.13463/j.cnki.cczyy.2016.01.001.
- Zhang SJ, Qi DM, Li YL, et al. Unsupervised methods in TCM syndrome standardization study [J]. J Changchun Univ Tradit Chin Med, 2016, 32 (1): 1-5. DOI: 10.13463/j.cnki.cczyy.2016.01.001.
- [25] 龚燕冰,倪青,王永炎. 中医证候研究的现代方法学述评 (一)——中医证候数据挖掘技术 [J]. 北京中医药大学学报, 2006, 29 (12): 797-801. DOI: 10.3321/j.issn.1006-2157.2006.12.001.
- Gong YB, Ni Q, Wang YY. Modern methodology of TCM syndrome study (I): data mining technology of TCM syndrome [J]. J Beijing Univ Tradit Chin Med, 2006, 29 (12): 797-801. DOI: 10.3321/j.issn.1006-2157.2006.12.001.
- [26] 龚燕冰,倪青,高思华,等. 2 型糖尿病不同并发症中医证候与血糖相关性的贝叶斯网络分析 [J]. 北京中医药大学学报, 2009, 32 (12): 815-818. DOI: 10.3321/j.issn.1006-2157.2009.12.006.
- Gong YB, Ni Q, Gao SH, et al. Bayesian network analysis on correlation between TCM syndromes and plasma glucose changes in different complications of type 2 diabetes [J]. J Beijing Univ Tradit Chin Med, 2009, 32 (12): 815-818. DOI: 10.3321/j.issn.1006-2157.2009.12.006.
- [27] 吴荣,聂晓燕,王阶,等. 基于贝叶斯网络的老中医治疗冠心病辨证规律研究 [J]. 中国中医药信息杂志, 2010, 17 (5): 98-99. DOI: 10.3969/j.issn.1005-5304.2010.05.053.
- Wu R, Nie XY, Wang J, et al. Study on the law of syndrome differentiation in the treatment of coronary heart disease based on Bayesian network [J]. Chin J Inf Tradit Chin Med, 2010, 17 (5): 98-99. DOI: 10.3969/j.issn.1005-5304.2010.05.053.
- [28] 唐启盛,曲森,包祖晓,等. 抑郁症中医证候的贝叶斯网络研究 [J]. 中医杂志, 2008, 49 (11): 1013-1015. DOI: 10.3321/j.issn.1001-1668.2008.11.027.
- Tang QS, Qu M, Bao ZX, et al. Bayesian network research on TCM syndromes of depression [J]. J Tradit Chin Med, 2008, 49 (11): 1013-1015. DOI: 10.3321/j.issn.1001-1668.2008.11.027.
- [29] 王至婉,李羚,李建生,等. 基于贝叶斯网络的慢性阻塞性肺疾病急性加重期证候与症状间的关联模式 [J]. 中华中医药杂志, 2019, 34 (9): 3964-3969.
- Wang ZW, Li L, Li JS, et al. Relationship between symptoms and syndromes in acute exacerbation of chronic obstructive pulmonary disease based on bayesian network [J]. CJTCMP, 2019, 34 (9): 3964-3969.
- [30] 薛薇. 基于 SPSS Modeler 的数据挖掘 [M]. 北京:中国人民大学出版社,2014.
- Xue W. Data mining based on SPSS Modeler [M]. Beijing: China Renmin University Press, 2014.
- [31] 方晶,乔飞,车军勇,等. 基于中医传承辅助系统的 40 例新型冠状病毒肺炎中医证候特点分析 [J]. 江苏中医药, 2020, 52 (4): 70-73. DOI: 10.19844/j.cnki.1672-397X.2020.04.011.
- Fang J, Qiao F, Che JY, et al. Analysis of traditional Chinese medicine syndrome characteristics of 40 cases of coronavirus disease 2019 based on traditional Chinese medicine inheritance auxiliary system [J]. Jiangsu J Tradit Chin Med, 2020, 52 (4): 70-73. DOI: 10.19844/j.cnki.1672-397X.2020.04.011.
- [32] 侯安农. 基于 T-S 动态故障树的高散时间贝叶斯网络分析方法 [D]. 秦皇岛:燕山大学,2019.
- Hou AN. Discrete-time Bayesian network analysis method based on T-S dynamic fault tree [D]. Qinhuangdao: Yanshan University, 2019.