

# 基于隐结构辨证规则的9 323例中医古籍咳嗽病案的方-证分析

陈丽平<sup>1,2</sup>, 李建生<sup>1\*</sup>, 邢静宇<sup>2</sup>, 胡久略<sup>2</sup>, 卞华<sup>2</sup>, 庞立业<sup>2</sup>

(1. 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; 2. 南阳理工学院 河南省张仲景方药与免疫调节重点实验室, 河南 南阳 473004)

**[摘要]** 目的:探索基于隐结构辨证规则的中医古籍咳嗽医案以方测证规律。方法:提取古籍中医肺病数据库内咳嗽病案9 323例,先用Lantern5.0(孔明灯)软件构建频次>100的前120味中药隐结构模型,再对获得的隐变量进行综合聚类 and 隐类诠释,最后通过综合聚类的辨证规则分析咳嗽的证候分布规律。结果:对120味中药建模得到46个隐变量,114个隐类,按照贝叶斯信息测度(BIC)评分,模型分数为-161 242.92,模型图形成以Y0, Y5, Y7, Y13, Y35, Y38为核心,且Y36, Y38, Y24, Y19, Y17涉及4个隐类。根据隐结构流线图汇总出10个综合聚类模型,其中Z5外寒内饮证的分值8.4最高,提示Z5的辨证规则结果支持度较高,中药半夏的互信息0.31和信息覆盖度63%最高;Z2肺气阴两虚证的分值-2最低,五味子的分值-15最低,提示Z2的辨证规则支持度较低;Z6风热犯肺证和Z10痰瘀阻肺证中连翘、羌活的分值均为19.1最高,提示这2味中药对其贡献度较大。以方测证得出咳嗽的常见证候有凉燥袭肺、肾阴虚、肺肾阴虚、外寒内饮、风热犯肺、肝火犯肺、肺热炽盛、痰湿阻肺、痰瘀阻肺9种,新发现的咳嗽证型有外寒内饮证和痰瘀阻肺证。结论:基于隐结构辨证规则的以方测证完善了咳嗽的证候分型,为中医隐性知识的显性化提供思路,同时为其他疾病的证候完善也提供方法参考,并有利于开发有价值的新处方。

**[关键词]** 隐结构模型; 辨证规则; 咳嗽; 医案; 以方测证

**[中图分类号]** R22;R242;R2-031;R287 **[文献标识码]** A **[文章编号]** 1005-9903(2020)18-0179-09

**[doi]** 10.13422/j.cnki.syfjx.20201621

**[网络出版地址]** <https://kns.cnki.net/kcms/detail/11.3495.R.20200629.0949.004.html>

**[网络出版日期]** 2020-6-29 13:41

## Prescriptions and Syndromes of 9 323 Cough Cases in Ancient Chinese Medical Books Based on Rules of Latent Structure Differentiation

CHEN Li-ping<sup>1,2</sup>, LI Jian-sheng<sup>1\*</sup>, XING Jing-yu<sup>2</sup>, HU Jiu-lue<sup>2</sup>, BIAN Hua<sup>2</sup>, PANG Li-ye<sup>2</sup>

(1. Respiratory Disease Prevention and Control of Provincial Department of Traditional Chinese Medicine to Build Collaborative Innovation Center, Henan Key Laboratory of Chinese Medicine for Respiratory Disease, Henan University of Chinese Medicine, Zhengzhou 450046, China;

2. Henan Provincial Key Laboratory of Zhang Zhongjing Formulae and Herbs for Immuregulation, Nanyang Institute of Technology, Nanyang 473004, China)

**[Abstract]** **Objective:** To explore cough cases recorded in ancient traditional Chinese medicine (TCM) books based on the Rules of Latent Structure Differentiation. **Method:** The 9 323 cough cases in the database of ancient TCM books and pulmonary diseases were extracted. At first, Lantern 5.0 software was used to construct the top 120 TCM hidden structure model with the frequency>100. Then, the obtained hidden variables were comprehensively clustered and interpreted. **Result:** The 46 hidden variables and 114 hidden classes were

**[收稿日期]** 20200619(013)

**[基金项目]** 国家自然科学基金青年科学基金项目(81704200);国家中医药领军人才支持计划一岐黄学者项目(国中医药人教函[2018]284号);河南省中医药研究专项(2019ZY3033)

**[第一作者]** 陈丽平,博士,副教授,从事中医文献整理与挖掘利用研究,E-mail:942202160@qq.com

**[通信作者]** \*李建生,博士,教授,从事呼吸疾病中医药防治的临床与基础研究,Tel:0377-65676568,E-mail:li\_js8@163.com

obtained through the modeling of 120 traditional Chinese medicines. According to the bayesian information measure (BIC) score, the model score was -161 242.92. The model diagram was formed with Y0, Y5, Y7, Y13, Y35 and Y38 as the cores, and Y36, Y38, Y24, Y19 and Y17 involved 4 hidden classes. Ten comprehensive clustering models were summarized according to the hidden structure flow chart, among which the highest score of Z5 external cold and internal decoction syndrome was 8.4, indicating that the result of syndrome differentiation rules of Z5 had a high degree of support, and the highest mutual information degree and information coverage of *Pinellia sinensis* were 0.31 and 63%. Z2 had a low score for the syndrome of deficiency of lung qi and Yin, which was -2, and *Schisandra* had the lowest score of -15, indicating that the syndrome differentiation rules of Z2 was less supportive. *Forsythia* and *Notopterygium* both had the highest score of Z6 wind-heat attack lung syndrome and Z10 phlegm and blood-stasis pulmonary syndrome, which was 19.1, indicating that these two TCMs had a great contribution to it. According to the test, the common syndromes of cough were cold and dryness attacking the lung, kidney-Yin deficiency, kidney-Yin deficiency of the lung, cold and internal decoction, wind heat invading the lung, liver fire invading the lung, lung heat burning, phlegm-dampness blocking the lung, phlegm-stasis blocking the lung. The newly discovered cough syndromes were external cold and internal rheum syndrome and phlegm-stasis blocking the lung syndrome. **Conclusion:** Based on the syndrome differentiation rules of hidden structure, the improved classification of cough syndromes provides ideas for the demonstration of tacit knowledge of TCM and methodological reference for the improvement of syndromes of other diseases, and is conducive to the development of valuable new prescriptions.

**[Key words]** latent structure model; rules of syndrome differentiation; cough; basis; syndrome differentiation by effects of prescription

“法随证立,方从法出,方以药成”,方药与证候的关系是中医辨证论治的核心内容,方剂中的药物配伍特征与疾病表现的临床证候具有高度相关性或者针对性<sup>[1]</sup>。以方测证是根据方剂的功效或其组成药物的功效推测或反证其对应证候,是通过病案中的显性知识推测潜藏隐性知识的过程。以方测证现代多运用在动物模型研究中,郭春荣等<sup>[2]</sup>通过“以方测证”分析判定地塞米松灌胃结合卵蛋白诱发哮喘制造肾阳虚哮喘模型;曾梅艳等<sup>[3]</sup>基于以方测证理论构建可靠、稳定的十二指肠溃疡脾胃虚寒证候大鼠实验模型,并对模型进行综合评价;邓杨春<sup>[4]</sup>以方测证探《伤寒论》方在温病中的临床运用,发现温病的各种情况都可以在伤寒方的基础之上稍微变化得出相应治疗方案。大部分以方测证探究侧重动物模型的实验研究,以西医思维验证中医方证功效,虽从逻辑出发,但未能建立客观的疾病模型,并不符合中医的辨证论治思维。通过中医文献整理和数据挖掘,运用以方测证法,才是完善中医学方证理论体系相对可靠的方法。中医古籍医案蕴含着大量的显性知识和隐形知识,如医案中的症状、中药和方剂等为显性知识,疾病诊察的切入点、重点关注的症/证、辨证依据、治疗着眼点、立法组方思路、遣药心得等辨治规律为隐性知识<sup>[5]</sup>。本

文以9 323例中医古籍咳嗽医案为数据源,基于隐结构模型的辨证规则进行方-证规律探究,为咳嗽证候的完善提供参考,同时为中医隐性知识的显性化提供思路,并有利于开发有价值的新处方。

## 1 资料与方法

**1.1 数据来源及预处理** 提取《古籍中医肺病数据库》内咳嗽病案9 323例<sup>[6]</sup>,涉及825种著作,根据《中国药典》<sup>[7]</sup>对中药进行处理,拆分药组,去掉修饰性副词或药物的炮制方法及产地,同一中药的异名规范为标准词,个别无法规范的数据保留原格式。见表1。

表1 9 323例咳嗽医案中药规范示例

Table 1 Example of traditional Chinese medicine specifications for 9 323 cough cases

规范前中药	规范后中药
瓜蒌皮、瓜蒌仁、瓜蒌实	瓜蒌
橘皮、橘络、橘红衣、橘红、桔络、桔红	陈皮
生龙牡	龙骨、牡蛎
姜半夏、法半夏、清半夏	半夏
生意米、生薏苡仁、炒薏米、薏仁、炒薏苡仁	薏苡仁
生晒参、红参、野山参	人参

**1.2 数据挖掘方法** 先用Lantern5.0隐结构分析软件(香港科技大学研制)构建咳嗽医案中药隐结构

模型,展现显变量和隐变量的关系,线条粗细揭示显变量对隐节点的贡献度,线条越细显示贡献度越小<sup>[8]</sup>。隐结构模型采用贝叶斯网络的模式对数据进行多维划分,基于累积互信息(CMI)和累积信息覆盖度(CDC)对隐类进行客观量化,并利用隐类概率(P)对这些证候进行诠释。隐结构模型图直观显示隐变量(latent variable)和显变量(manifest variable)之间的关系,用互信息(MI)和信息覆盖度(DC)量化隐变量间、隐变量与显变量之间的关联程度,用隐类概率刻画隐变量的分布情况,条件概率(CP)刻画显变量的分布情况,按照累积互信息 $\geq 95\%$ 的原则,先选取显变量的子集,再对获取的隐结构模型中不同的隐变量和其涉及的隐类逐一诠释,最后归纳和提取该病的证候和所用中药主次地位的意义,并进一步对数据行综合聚类后再分析其辨证规则,分别用阈值(threshold value)和分值(score)表示。

辨证规则是用来量化症状或中药在证候判别时提供的支持证据,主要指标为对数优势比(log odds ratio)。针对证候Z,通过症状 $X_i$ 在证候Z中出现概率的对数优势比,来量化证候判别的支持证据,从而形成证候Z的辨证规则。分别用阈值(threshold value)和分值(score)表示,阈值是指一组数据的整体评分,分值为每个显变量对数优势比。阈值的公式

$$\log \frac{P(Z = \sim s)}{P(Z = s)} - \sum_{i=1}^n \log \frac{P(X_i = 0|Z = s)}{P(X_i = 0|Z = \sim s)}$$

分值(对数优势比 odds)的公式

$$\log \left( \frac{P(X_i|Z = s)}{P(X_i|Z = \sim s)} / \frac{P(X_i = 0|Z = s)}{P(X_i = 0|Z = \sim s)} \right)$$

其挖掘步骤如下,①收集咳嗽医案数据;②构建中药数据的隐结构模型;③对获得的隐变量进行逐一诠释;④依据中医知识进行隐变量组合;⑤根据组合进行综合聚类,再揭示咳嗽的辨证分型;⑥综合运用综合聚类模型的诠释指标建立辨证规则,以方测证揭示咳嗽证候分布特点。通过反复尝试,隐结构模型能相对客观地解决中医辨证的定量问题,其模型和中医的非线性思维拟合较好<sup>[9]</sup>。

## 2 结果

**2.1 医案中120味中药隐结构模型** 针对频次>100的前120味中药构建隐结构模型,见图1,并用LTM-EAST(爬山法)进行分析,累积信息覆盖率为95%,共得到46个隐变量,每个隐变量下涉及2个到

多个隐类,计114个隐类,其中涉及2个隐类的有29个隐变量,涉及3个隐类的有12个隐变量,用Y0, Y1……Y45表示,评价标准为贝叶斯信息测度(BIC)评分,模型分数-161 242.92。图1形成以Y0, Y5, Y7, Y13, Y35, Y38为核心, Y36, Y38, Y24, Y19, Y17涉及4个隐类,先根据隐结构的拓扑结构和线条粗细将不同隐节点进行综合聚类,共分为10组,用10种颜色显示,分别用Z1~Z10表示,再通过每组用药的主要功效推测出所针对的主要证候,最后用综合聚类的辨证规则判断推测是否成立,诠释见表2。

**2.2 医案中120味中药隐结构模型10组综合聚类诠释** 根据图1进行隐结构模型的初步诠释,图中线条粗细代表与隐节点关联程度的大小,线条越粗提示关联越紧密,互信息和信息贡献度也越大,诠释结果提示治疗咳嗽常用的中药药性偏温性和凉性,药味偏辛、甘、苦味,归经多为肺、脾、肝经,咳嗽常见的证型有凉燥袭肺、肺气阴两虚、肾阴虚、肺肾阴虚、外寒内饮、风热犯肺、肝火犯肺、肺热炽盛、痰湿阻肺、痰瘀阻肺10种,见表2。

**2.3 医案中综合聚类Z1权重** 图1和表2根据线条和中医基础知识初步诠释的结果,判定证型的具体值需要根据综合聚类权重做最终判定。与Z1关系密切的是苦杏仁和贝母,其与Z1的互信息是0.15,0.12,信息覆盖度是37%,30%,Z1聚类的患者中应用该药组的概率是65%,Z1总信息量0.41,证型分值4.2,见表3。

**2.4 医案中综合聚类Z2权重** 与Z2关系密切的是人参和五味子,其与Z2的互信息是0.15,0.12,信息覆盖度是39%,31%,Z2聚类的患者中应用该药组的概率是61%,Z2总信息量0.39,证型分值-2,此证型分值为负分,提示拟合结果较差,需要剔除,见表4。

**2.5 医案中综合聚类Z3权重** 与Z3关系密切的是山药、山茱萸和茯神,其与Z3的互信息均为0.06,信息覆盖度分别为34%,32%,31%,Z3聚类的患者中应用该药组的概率是92%,Z3总信息量0.18,证型分值5.3,见表5。

**2.6 医案中综合聚类Z4权重** 与Z4关系密切的是麦冬和沙参,其与Z4的互信息均为0.13,信息覆盖度是20%,19%,Z4聚类的患者中应用该药组的概率是69%,Z4总信息量为0.43,证型分值6,见表6。

**2.7 医案中综合聚类Z5权重** 与Z5关系最密切的是半夏,其与Z5的互信息是0.31,信息覆盖度为

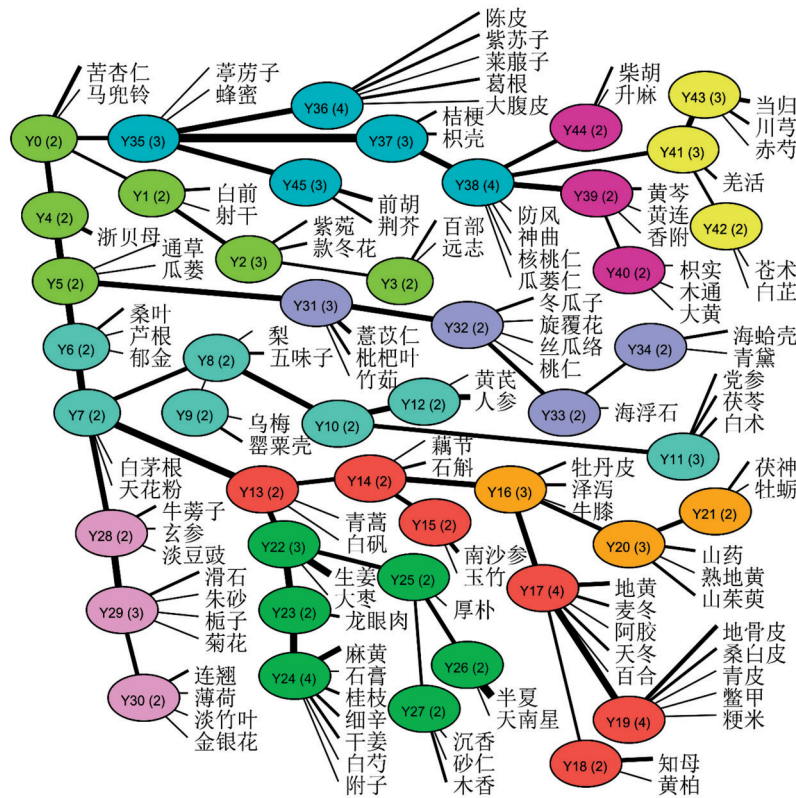


图1 9 323例咳嗽医案120味中药隐结构模型

Fig. 1 Hidden structure model of 120 traditional Chinese medicines for 9 323 cough cases

表2 9 323例咳嗽医案120味中药隐结构模型10组综合聚类诠释

Table 2 Preliminary interpretation of hidden structure model of 120 traditional Chinese medicines for 9 323 cough cases

隐节点	贡献度较大的显变量	性味归经	主要功效	证候
Y <sub>0</sub> -Y <sub>5</sub>	苦杏仁、浙贝母、白前、紫菀、款冬花、百部	偏温性、偏辛苦味, 归肺经	宣利肺气、疏风止咳	凉燥袭肺(Z1)
Y <sub>6</sub> -Y <sub>12</sub>	人参、党参、茯苓、白术、桑叶、五味子	温平凉并用、辛苦酸味, 归肺脾经	益气养阴、敛肺止咳	肺气阴两虚(Z2)
Y <sub>16</sub> , Y <sub>20</sub> , Y <sub>21</sub>	山药、山茱萸、茯神、牡丹皮、泽泻	温平并用、偏甘酸淡味、主归肾经	滋阴补肾	肾阴虚(Z3)
Y <sub>13</sub> -Y <sub>15</sub> , Y <sub>17</sub> -Y <sub>19</sub>	南沙参、石斛、地黄、麦冬、天冬、地骨皮、桑白皮	偏凉性、甘味、主归肺胃经	滋阴益胃、降逆下气	肺肾阴虚(Z4)
Y <sub>22</sub> -Y <sub>27</sub>	麻黄、半夏、生姜、龙眼肉、桂枝、细辛、干姜	偏温性、辛味、归肺经	解表散寒、温肺化饮	外寒内饮(Z5)
Y <sub>7</sub> , Y <sub>28</sub> -Y <sub>29</sub>	连翘、牛蒡子、滑石、金银花、淡竹叶、薄荷	偏凉性、辛苦味, 归肺经	辛凉透表、清热解毒	风热犯肺(Z6)
Y <sub>31</sub> -Y <sub>34</sub>	冬瓜子、枇杷叶、海蛤壳、青黛、薏苡仁、海浮石	偏凉性、偏苦味, 归肝肺经	清肝泻火	肝火犯肺(Z7)
Y <sub>38</sub> , Y <sub>44</sub> , Y <sub>39</sub> , Y <sub>40</sub>	黄芩、柴胡、升麻、黄连、大黄、木通	偏凉性、苦味, 主归肺经	清热解毒	肺热炽盛(Z8)
Y <sub>35</sub> -Y <sub>38</sub> , Y <sub>45</sub>	前胡、枳壳、桔梗、陈皮、紫苏子、葛根、荆芥	偏温性、辛甘味、归肺脾经	理气健脾、化痰止咳	痰湿阻肺(Z9)
Y <sub>38</sub> , Y <sub>41</sub> -Y <sub>43</sub>	羌活、当归、川芎、白芍、苍术、核桃仁	偏温性、辛甘酸味、归肝经	祛湿化痰、活血化痰	痰瘀阻肺(Z10)

63%, Z5聚类的患者中应用该药组的概率是69%, Z5总信息量0.49, 证型分值8.4, 见表7。

**2.8 医案中综合聚类Z6权重** 与Z6关系最密切的是连翘, 其与Z6的互信息是0.1, 信息覆盖度为50%, Z6聚类的患者中应用该药组的概率是88%, Z6总信息量0.21, 证型分值6.2, 见表8。

**2.9 医案中综合聚类Z7权重** 与Z7关系密切的是冬瓜子、枇杷叶和海蛤壳, 其与Z7的互信息分别

是0.08, 0.07和0.06, 用LTM-EAST(爬山法)进行分析, 信息覆盖度分别为39%, 32%, 26%, Z7聚类的患者中应用该药组的概率是89%, Z7总信息量0.22, 证型分值6.1, 见表9。

**2.10 医案中综合聚类Z8权重** 与Z8关系最密切的是黄芩, 其与Z8的互信息是0.13, 信息覆盖度为62%, Z8聚类的患者中应用该药组的概率是77%, Z8总信息量0.21, 证型分值3, 见表10。

表3 9 323例咳嗽医案综合聚类Z1权重

Table 3 Comprehensive cluster weight of Z1 in 9 323 cough cases

Z1(凉燥袭肺)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.65(s0,s1)		p=0.35(s0,s1)		分值/分
苦杏仁	0.15	37	0.15	37	0.83	0.17	0.27	0.73	4.3
浙贝母	0.12	30	0.24	58	0.90	0.10	0.42	0.58	3.5
紫菀	0.07	16	0.3	74	0.94	0.06	0.86	0.14	0.9
百部	0.05	12	0.34	83	0.99	0.01	1.00	0	-2.2
白前	0.04	10	0.36	89	1.00	0	0.96	0.04	8.4
瓜蒌	0.03	7	0.38	93	0.99	0.01	0.87	0.13	3.8
款冬花	0.02	6	0.39	95	0.95	0.05	0.90	0.10	0
马兜铃	0.02	5	0.40	98	0.99	0.01	0.89	0.11	4.8

表4 9 323例咳嗽医案综合聚类Z2权重

Table 4 Comprehensive cluster weight of Z2 in 9 323 cough cases

Z2(肺气阴两虚)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.61(s0,s1)		p=0.39(s0,s1)		分值/分
人参	0.15	39	0.15	39	0.97	0.03	0.76	0.24	-1
五味子	0.12	31	0.25	65	1.00	0	0.70	0.30	-15
茯苓	0.06	15	0.29	75	0.81	0.19	0.74	0.26	0
白术	0.05	14	0.31	81	0.98	0.02	0.91	0.09	-2
桑叶	0.05	12	0.34	88	0.84	0.16	0.99	0.01	6
罂粟壳	0.02	5	0.36	92	1.00	0	0.95	0.05	-9
芦根	0.02	5	0.37	94	0.94	0.06	1.00	0	16
黄芪	0.02	5	0.37	96	0.98	0.02	0.94	0.06	-3

表5 9 323例咳嗽综合聚类Z3医案权重

Table 5 Comprehensive cluster weight of Z3 in 9 323 cough cases

Z3(肾阴虚)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.92(s0,s1)		p=0.08(s0,s1)		分值/分
山药	0.06	34	0.06	34	0.98	0.02	0.79	0.21	4.7
山茱萸	0.06	32	0.09	49	0.99	0.01	0.96	0.04	3.9
茯神	0.06	31	0.13	75	0.99	0.01	0.64	0.36	6.3
泽泻	0.05	25	0.14	80	0.99	0.01	0.99	0.01	0.0
牡丹皮	0.04	22	0.15	83	0.97	0.03	0.85	0.15	2.5
熟地黄	0.04	20	0.16	90	0.99	0.01	0.78	0.22	4.6
牡蛎	0.03	15	0.18	100	1.00	0.00	0.80	0.20	5.9

2.11 医案中综合聚类Z9权重 与Z9关系密切的是前胡和枳壳,其与Z9的互信息为0.13和0.1,信息覆盖度是32%和25%,Z9聚类的患者中应用该药组的概率是76%,Z9总信息量0.4,证型分值4.8,见表11。

2.12 医案中综合聚类Z10权重 与Z10关系密切的是羌活和川芎,其与Z10的互信息均为0.06,信息覆盖度是47%和46%,Z10聚类的患者中应用该药组的概率是95%,Z10总信息量0.13,证型分值7.4,

见表12。

### 3 讨论

隐结构模型的辨证规则能相对客观地解决数据的模糊性和不确定性,基于无标签数据构建辨证分型模型和规则,根据中药应用特点对患者进行综合聚类分析,找到与证候对应的患者子类,利用患者子类的数据权重对证候进行量化以建立辨证规则。中药的隐结构模型含有多个隐变量,每个隐变量代表对患者的一种划分,综合聚类需要有灵活性

表6 9 323例咳嗽医案综合聚类Z4权重

Table 6 Comprehensive cluster weight of Z4 in 9 323 cough cases

Z4(肺肾阴虚)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.69(s0,s1)		p=0.31(s0,s1)		分值/分
麦冬	0.13	20	0.13	20	0.97	0.03	0.63	0.37	4.2
南沙参	0.13	19	0.26	38	0.99	0.01	0.41	0.59	7.3
地骨皮	0.06	16	0.30	53	1.00	0	0.93	0.07	8.6
桑白皮	0.04	15	0.33	64	0.90	0.10	0.92	0.08	-0.4
知母	0.03	10	0.34	70	0.97	0.03	0.97	0.03	0.3
天冬	0.03	9	0.35	76	0.99	0.01	0.94	0.06	3.4
地黄	0.02	9	0.36	80	0.97	0.03	0.90	0.10	1.8
阿胶	0.02	8	0.37	85	0.97	0.03	0.89	0.11	2.1
石斛	0.01	7	0.38	88	0.98	0.02	0.72	0.28	4.3
鳖甲	0.01	6	0.39	93	1.00	0	0.98	0.02	12.2
玉竹	0.01	5	0.39	94	1.00	0	0.81	0.19	6.1
百合	0.01	3	0.39	96	0.99	0.01	0.94	0.06	2.7

表7 9 323例咳嗽医案综合聚类Z5权重

Table 7 Comprehensive cluster weight of Z5 in 9 323 cough cases

Z5(外寒内饮)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.69(s0,s1)		p=0.31(s0,s1)		分值/分
半夏	0.31	63	0.31	63	0.97	0.03	0.03	0.97	11.2
干姜	0.05	11	0.36	72	0.99	0.01	0.93	0.07	3.3
龙眼肉	0.05	10	0.39	80	0.99	0.01	0.99	0.01	0.9
生姜	0.05	10	0.41	83	0.90	0.10	0.61	0.39	2.3
附子	0.04	7	0.44	88	1.00	0.00	1.00	0	2.0
细辛	0.04	7	0.46	92	0.99	0.01	0.91	0.09	2.8
桂枝	0.02	5	0.46	94	0.98	0.02	0.92	0.08	2.2
麻黄	0.02	5	0.48	98	0.94	0.06	0.81	0.19	1.8

表8 9 323例咳嗽医案综合聚类Z6权重

Table 8 Comprehensive cluster weight of Z6 in 9 323 cough cases

Z6(风热犯肺)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.88(s0,s1)		p=0.12(s0,s1)		分值/分
连翘	0.1	50	0.10	50	1.00	0.00	0.63	0.37	19.1
薄荷	0.03	17	0.11	54	0.96	0.04	0.71	0.29	3.3
栀子	0.03	16	0.14	67	0.98	0.02	0.77	0.23	3.8
牛蒡子	0.03	16	0.15	74	0.98	0.02	0.76	0.24	3.7
金银花	0.02	12	0.16	79	1.00	0.00	0.90	0.10	7.0
滑石	0.02	11	0.17	84	0.99	0.01	0.84	0.16	3.7
玄参	0.02	9	0.18	89	0.98	0.02	0.85	0.15	3.5
菊花	0.02	8	0.19	93	0.99	0.01	0.90	0.10	4.3
淡竹叶	0.02	8	0.20	95	0.99	0.01	0.90	0.10	4.3
天花粉	0.02	7	0.20	99	0.98	0.02	0.85	0.15	3.0

才能使患者类与证候的对应更贴切。以方测证是基于辨证论治的原则,在病证结合的前提下,根据

方剂的功效或其组成药物的功效推测或反证其对应证候的属性<sup>[10]</sup>。从10组综合聚类权重表3~12中

表9 9 323例咳嗽医案综合聚类Z7权重

Table 9 Comprehensive cluster weight of Z7 in 9 323 cough cases

Z7(肝火犯肺)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.89(s0,s1)		p=0.11(s0,s1)		分值/分
冬瓜子	0.08	39	0.08	39	0.99	0.01	0.61	0.39	6.5
枇杷叶	0.07	32	0.14	64	0.96	0.04	0.54	0.46	5
海蛤壳	0.06	26	0.17	78	0.98	0.02	0.66	0.34	3.9
薏苡仁	0.03	15	0.18	85	0.95	0.05	0.69	0.31	1.4
竹茹	0.03	14	0.19	89	0.98	0.02	0.77	0.23	4.6
丝瓜络	0.03	14	0.20	93	1.00	0.00	0.86	0.14	6.3
旋覆花	0.03	12	0.21	98	0.97	0.03	0.77	0.23	3.6

表10 9 323例咳嗽医案综合聚类Z8权重

Table 10 Comprehensive cluster weight of Z8 in 9 323 cough cases

Z8(肺热炽盛)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.77(s0,s1)		p=0.23(s0,s1)		分值/分
黄芩	0.13	62	0.13	62	0.99	0.01	0.61	0.39	10.4
柴胡	0.03	16	0.15	73	0.98	0.02	0.83	0.17	3.6
黄连	0.03	15	0.17	82	1.00	0.00	0.89	0.11	5.2
防风	0.02	8	0.18	87	0.99	0.01	0.90	0.10	3.1
大黄	0.01	7	0.19	90	0.99	0.01	0.93	0.07	3.9
枳实	0.01	7	0.20	94	0.99	0.01	0.93	0.07	3.9
升麻	0.01	5	0.20	96	0.99	0.01	0.94	0.06	3.5

表11 9 323例咳嗽医案综合聚类Z9权重

Table 11 Comprehensive cluster weight of Z9 in 9 323 cough cases

Z9(痰湿阻肺)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.76(s0,s1)		p=0.24(s0,s1)		分值/分
前胡	0.13	32	0.13	32	0.99	0.01	0.75	0.25	4.8
枳壳	0.10	25	0.2	51	1.00	0.00	0.78	0.22	5.6
桔梗	0.08	20	0.24	61	0.89	0.11	0.73	0.27	2.0
陈皮	0.07	17	0.29	74	0.81	0.19	0.35	0.65	3.1
紫苏子	0.04	10	0.32	81	0.97	0.03	0.72	0.28	3.6
葛根	0.03	8	0.34	86	0.99	0.01	0.99	0.01	-0.3
荆芥	0.03	8	0.36	92	0.98	0.02	0.98	0.02	-1.7
防风	0.02	6	0.37	94	0.99	0.01	0.96	0.04	1.7
莱菔子	0.02	4	0.38	96	0.99	0.01	0.91	0.09	3.0

分别遴选有代表性的互信息和信息覆盖度、隐类概率、辨证规则(阈值与分值)三类指标诠释其代表的辨证意义,以此来指导以方测证,从而完成方-证规律的挖掘。

互信息和信息覆盖度解决每味中药对其方剂的贡献度和纳入几味中药组成一个方剂能诠释完整的证候。表7中半夏和Z5的互信息0.31,信息覆盖度达63%>50%,提示半夏对外寒内饮证的贡献度最大,半夏善治脏腑之湿痰,其辛温而燥,是燥湿化

痰、温化寒痰之要药。《本经逢原》云:“半夏同苍术、茯苓治湿痰;同瓜蒌、黄芩治热痰;同南星、前胡治风痰;同芥子、生姜汁治寒痰。惟燥痰宜瓜蒌、贝母,非半夏所能治也”<sup>[11]</sup>。当纳入到桂枝、麻黄时其累积信息覆盖度已达98%>95%,纳入的这组中药已能诠释Z5,可以剔除其他冗余信息如厚朴、天南星等。

隐类概率完成对患者的分类,并为辨证规则的完成奠定基础。表6中Z4的患者被隐类概率分为

表12 9 323例咳嗽医案综合聚类Z10权重

Table 12 Comprehensive cluster weight of Z10 in 9 323 cough cases

Z10(痰瘀阻肺)	互信息	信息覆盖度/%	累积互信息	累积信息覆盖度/%	p=0.95(s0, s1)	p=0.05(s0, s1)	分值/分
羌活	0.06	47	0.06	47	1.00	0.00	0.63 0.37 19.1
川芎	0.06	46	0.10	75	0.98	0.02	0.45 0.55 5.8
防风	0.03	26	0.11	86	0.98	0.02	0.65 0.35 1.7
白芷	0.02	16	0.12	91	1.00	0.00	0.82 0.18 5.7
苍术	0.02	14	0.12	94	0.99	0.01	0.81 0.19 5.1
当归	0.00	3	0.13	101	0.92	0.08	0.79 0.21 1.6

两类,分别占总体的69%和31%,区分这两个类的重要中药是麦冬、南沙参等(累积互信息为0.26,累积信息覆盖度为38%),两个类的特点如类概率所示,基于这些特点,认为隐类Z4=s1是与肺肾阴虚对应的患者类,Z4有3个取值(假设两个概率较低的一类区别不重要,合二为一),把患者分成3类,通过分析,认为Z4=s0这类患者无肺肾阴虚,而Z4=s1和Z2这两类患者有肺肾阴虚,Z4=s1,地骨皮、鳖甲、玉竹、南沙参、天冬、百合的概率偏高;Z4=s2,鳖甲、知母的概率偏高,把s1和s2合并后,就得到一个与肺肾阴虚对应的患者类,其统计特征可以用来定量刻画肺肾阴虚。

基于综合聚类隐类概率的结果,得出判断肺肾阴虚的规则,阈值是6,总分≥6,判断为肺肾阴虚,总分<6,判断为非肺肾阴虚,准确度相对于基于综合聚类模型的辨证。分值是指对数优势比,麦冬在肺肾阴虚的人群中出现的概率0.97,不出现的概率是0.03,odds是0.97/0.03=32.33,麦冬在非肺肾阴虚的人群中出现的概率0.63,不出现的概率是0.37,odds是0.37/0.63=0.58,所以,麦冬是肺肾阴虚的正面证据,分值(score)是两个odds之比值的对数 $\log_2 \frac{32.33}{0.58} = 4.2$ ,其他中药的分值算法相同,其中分值≥0的是肺肾阴虚的正面辨证依据,分值<0的是负面依据,如桑白皮的分值是-0.4,根据辨证规则,若正面依据和负面依据的综合分数大于证候的阈值,即4.2+7.3+8.6-0.4+……+6.1+2.7>6,则判定为肺肾阴虚证。

表3~12资料显示Z1含止嗽散,Z2含补肺汤,Z3含六味地黄丸,Z4含沙参麦冬汤百合固金汤,Z5含小青龙汤,Z6含银翘散,Z7含旋覆花汤合黛蛤散,Z8含三黄枳实汤,Z9含三子养亲汤,Z10含川芎茶调散和当归芍药散的主要药组,同时发现了新的药组,如Z1苦杏仁、浙贝母、紫菀、百部、白前、瓜蒌、

款冬花、马兜铃。其中针对咳嗽肺气上逆的常用中药是苦杏仁,味苦降泄,肃降兼宣发肺气而止咳平喘,《本草便读》曰:“功专降气,气降则痰消嗽止”<sup>[12]</sup>。以方测证得出的证候有凉燥袭肺、肺气阴两虚、肾阴虚、肺肾阴虚、外寒内饮、风热犯肺、肝火犯肺、肺热炽盛、痰湿阻肺、痰瘀阻肺10种。其中Z2(-2)中人参(-1),五味子(-15),白术(-2),罂粟壳(-9),黄芪(-3)是负面辨证依据,根据辨证规则,-1-15+0-2+6-9+16-3=-8<-2,正面证据未超过负面证据,提示肺气阴两虚证的拟合较差,不符合辨证规则,故应剔除。

以方测证得出咳嗽的常见证候有凉燥袭肺、肾阴虚、肺肾阴虚、外寒内饮、风热犯肺、肝火犯肺、肺热炽盛、痰湿阻肺、痰瘀阻肺9种。与第2版《实用中医内科学》<sup>[13]</sup>、《中医药学名词·内科学·妇科学·儿科学》<sup>[14]</sup>不同的是外寒内饮、痰瘀阻肺证。隐结构模型的辨证规则除应用在以方测证挖掘方面,对以症测证、以病机测证、以药测方同样适用,为证候和方剂分布规律等隐性知识的挖掘提供参考。

【参考文献】

[1] 周荣荣,闫润红,王丽萍,等. 数据挖掘技术在中医方剂科学问题研究中的应用[J]. 中华中医药杂志, 2018, 33(9):4016-4020.

[2] 郭春荣,施旻昱,庄怡娴,等. 基于以方测证的肾阳虚哮喘大鼠模型建立与评价研究[J]. 中华中医药杂志, 2018, 33(10):4715-4718.

[3] 曾梅艳,宋厚盼,陈小娟,等. 基于以方测证理论探讨脾胃虚寒型十二指肠溃疡实验模型的构建及模型评价研究[J]. 中药药理与临床, 2019, 35(1): 183-188.

[4] 邓杨春. 以方测证探《伤寒论》方在温病中的临床运用[D]. 北京:河北京中医药大学, 2014.

[5] 杨光华,王萍,黄利兴,等. 名老中医学术传承研究方法探析[J]. 中医杂志, 2014, 55(21):1810-1812.

- [6] 陈丽平,李建生,蔡永敏,等. 中医肺病数据库构建及挖掘的思路与方法[J]. 中华中医药杂志, 2017, 32(12):5530-5534.
- [7] 国家药典委员会. 中华人民共和国药典. 一部[M]. 北京:中国医药科技出版社, 2015.
- [8] 张连文,袁世宏. 隐结构模型与中医辨证研究(I)——隐结构法的基本思想以及隐结构分析工具[J]. 北京中医药大学学报, 2006, 29(6):365-369.
- [9] 陈丽平,李建生,蔡永敏,等. 基于隐结构模型的名老中医辨治京津冀地区慢性支气管炎用药规律研究[J]. 北京中医药大学学报, 2018, 41(8):681-688.
- [10] 唐仕欢,杨洪军,黄璐琦. “以方测证”方法应用的反思[J]. 中国中西医结合杂志, 2007, 27(3):259-262.
- [11] 清·张璐. 本经逢原[M]. 顾漫,杨奕周,校注. 北京:中国中医药出版社, 2011:103-104.
- [12] 清·张秉成. 本草便读[M]. 张效霞,校注. 北京:学苑出版社, 2010:256-257.
- [13] 王永炎,严世芸. 实用中医内科学[M]. 上海:上海科学技术出版社, 2009:164-171.
- [14] 中医药学名词审定委员会. 中医药学名词·内科学·妇科学·儿科学[M]. 北京:科学出版社, 2011:30-32.
- [责任编辑 张丰丰]