

· 数据挖掘 ·

基于关联规则和复杂系统熵聚类的 颜正华教授治疗胃脘痛用药规律研究

吴嘉瑞, 张冰*, 杨冰
(北京中医药大学, 北京 100102)

[摘要] **目的:** 基于中医传承辅助系统软件, 分析颜正华教授治疗胃脘痛的用药经验。**方法:** 收集颜正华教授治疗胃脘痛的病案, 采用关联规则 apriori 算法、复杂系统熵聚类 etc 无监督数据挖掘方法, 确定处方中各种药物的使用频次及药物之间的关联规则, 分析颜正华教授治疗胃脘痛的用药经验。**结果:** 对筛选出的 150 个处方进行分析, 确定处方中药物的使用频次, 药物之间的关联规则, 挖掘出 49 个核心组合和 29 个新处方。**结论:** 颜正华教授治疗胃脘痛经验丰富, 多用疏肝理气、和胃止痛之品。中医传承辅助系统对于挖掘名老中医临床经验具有重要实用价值。

[关键词] 颜正华; 胃脘痛; 关联规则; 聚类算法

[中图分类号] R287.6 **[文献标识码]** A **[文章编号]** 1005-9903(2012)20-0001-05

Analysis on Patterns of YAN Zheng-hua's Prescriptions for Treating Stomachache by Apriori and Clustering Algorithm

WU Jia-rui, ZHANG Bing*, YANG Bing
(Beijing University of Chinese Medicine, Beijing 100102, China)

[Abstract] **Objective:** To analyze the composing experience of professor YAN Zheng-hua's prescriptions for treating stomachache by using traditional Chinese medicine (TCM) inheritance support system. **Method:** The prescriptions for stomachache composed by YAN Zheng-hua were collected, then the data were input in the TCM inheritance support system and the composing principles were analyzed by using data-mining methods, such as revised mutual information, complex system entropy clustering and unsupervised hierarchical clustering. **Result:** Based on the analysis of 150 prescriptions, the frequency of each herb and the association rules among the herbs were used, 49 core combinations and 29 new prescriptions are mined from the database. **Conclusion:** Professor YAN Zheng-hua is good at soothing the liver and regulating qi to relieve stomachache. TCM inheritance support system can be used to analyze clinical experiences of senior TCM doctors.

[Key words] YAN Zheng-hua; stomachache; association rules; clustering algorithm

颜正华教授是国医大师, 首都国医名师, 治验甚众, 尤擅长消化系统病证等的诊疗。本研究首先收集、整理颜正华教授治疗胃脘痛处方, 进而基于中医传承辅助系统软件, 采用关联规则 apriori 算法、复

杂系统熵聚类 etc 无监督数据挖掘方法, 分析处方中药物的使用频次及药物之间的关联规则、处方规律, 探讨颜正华教授治疗胃脘痛的用药经验, 以期阐明颜正华教授胃脘痛用药思想提供参考。

[收稿日期] 20120607(391)

[基金项目] 国家科技支撑计划(2007BAH0B01); 北京市自然科学基金(7112075); 北京市中医药科技发展基金(JJ-2010-70); 北京中医药大学科研创新团队项目(2011-CXTD-14); 北京中医药大学自主课题(JYB22-JS020)

[第一作者] 吴嘉瑞, 副教授, 博士, 硕士生导师, 从事临床中药学研究, Tel: 010-84738662, E-mail: exogamy@163.com

[通讯作者] * 张冰, 教授, 博士, 博士生导师, 从事临床中药学研究, Tel: 010-84738606, E-mail: zhangbing@263.net

1 资料与方法

1.1 处方来源与筛选 本研究以 2005 - 2010 年颜正华教授在北京中医药大学国医堂出诊处方为来源进行筛选,以《中医内科学》教材胃脘痛主要症状为评判标准,共筛选胃脘痛处方 150 首。

1.2 分析软件 “中医传承辅助系统(V1.1)”软件,由中国中医科学院中药研究所提供。

1.3 处方的录入和核对 将上述筛选后的处方录入“中医传承辅助系(V1.1)”。录入完成后,由双人负责数据的审核,以确保数据的准确性。通过“中医传承辅助系统(V1.1)”软件中“数据分析”模块中的“方剂分析”功能,进行用药规律挖掘^[1-3]。

1.4 数据分析

1.4.1 提取数据源 在“中医疾病”项中输入“胃脘痛”,提取出治疗胃脘痛的全部方剂。

1.4.2 频次统计分析 将胃脘痛方剂中每味药的出现频次从大到小排序,并将“频次统计”结果导出。

1.4.3 组方规律分析 “支持度个数”(表示在所有药物中同时出现的次数)分别设为 30,“置信度”设为 0.9,按药物组合出现频次从大到小的顺序进行排序;“规则分析”分析所得的规则。

1.4.4 新方分析 首先进行聚类分析(核心算法包括改进的互信息法、复杂系统熵聚类),在聚类分析前,先选择合适的相关度和惩罚度,然后点击“提取组合”按钮,发现新组方(基本算法是无监督的熵层次的聚类),并可以实现网络可视化展示^[4-6]。

2 结果

2.1 用药频次 对颜正华教授 150 首胃脘痛处方中的药物频次进行统计,使用频次高于 30 的有 20 味药,使用频次前 3 位分别是陈皮、佛手、香附。见表 1。

表 1 方剂中使用频次 30 以上的药物情况

序号	中药名称	频次	序号	中药名称	频次
1	陈皮	113	11	茯苓	47
2	佛手	96	12	枳壳	46
3	香附	95	13	柴胡	46
4	白芍	94	14	旋覆花	43
5	煨瓦楞子	80	15	炒枣仁	43
6	赤芍	80	16	绿萼梅	37
7	当归	68	17	黄连	36
8	丹参	62	18	吴茱萸	32
9	砂仁	57	19	青皮	32
10	紫苏梗	56	20	夜交藤	31

2.2 基于关联规则分析的组方规律分析 按照药物组合出现频次由高到低排序,前三位分别是“佛手、陈皮”,“陈皮、香附”,“佛手、香附”。见表 2。分析所得药对的用药规则,结果见表 3。关联规则

网络见图 1。

表 2 胃脘痛处方中支持度为 30 条件下药物组合频次

序号	药物模式	频次	序号	药物模式	频次
1	佛手,陈皮	81	11	白芍,佛手	64
2	陈皮,香附	81	12	香附,煨瓦楞子	63
3	佛手,香附	73	13	佛手,陈皮,香附	62
4	白芍,香附	73	14	白芍,陈皮,香附	61
5	白芍,陈皮	68	15	赤芍,佛手	60
6	白芍,赤芍	67	16	佛手,香附,煨瓦楞子	59
7	佛手,煨瓦楞子	66	17	白芍,煨瓦楞子	58
8	赤芍,陈皮	66	18	白芍,赤芍,香附	57
9	赤芍,香附	64	19	白芍,佛手,香附	56
10	陈皮,煨瓦楞子	64	20	佛手,陈皮,煨瓦楞子	55

表 3 胃脘痛处方中药物组合关联规则(置信度 > 0.94)

序号	规则	置信度
1	白芍,紫苏梗 - > 香附	1
2	白芍,陈皮,紫苏梗 - > 香附	1
3	佛手,旋覆花 - > 煨瓦楞子	0.971 429
4	吴茱萸 - > 黄连	0.968 75
5	青皮 - > 陈皮	0.968 75
6	吴茱萸,黄连 - > 煨瓦楞子	0.967 742
7	砂仁,紫苏梗 - > 陈皮	0.967 742
8	赤芍,紫苏梗 - > 香附	0.967 742
9	紫苏梗 - > 陈皮	0.967 742
10	香附,紫苏梗 - > 陈皮	0.964 286
11	佛手,紫苏梗 - > 陈皮	0.961 538
12	佛手,砂仁 - > 陈皮	0.956 522
13	赤芍,佛手,煨瓦楞子 - > 香附	0.954 545
14	佛手,香附,紫苏梗 - > 陈皮	0.954 545
15	白芍,赤芍,佛手,煨瓦楞子 - > 香附	0.95 2381
16	砂仁,香附 - > 陈皮	0.948 718
17	砂仁 - > 陈皮	0.947 368
18	赤芍,佛手,陈皮,煨瓦楞子 - > 香附	0.947 368
19	白芍,紫苏梗 - > 陈皮,香附	0.945 946
20	白芍,紫苏梗 - > 陈皮	0.945 946
21	白芍,香附,紫苏梗 - > 陈皮	0.945 946
22	紫苏梗,煨瓦楞子 - > 香附	0.945 946
23	紫苏梗,煨瓦楞子 - > 佛手	0.944 444
24	紫苏梗,煨瓦楞子 - > 陈皮	0.944 444
25	赤芍,柴胡 - > 白芍	0.944 444
26	白芍,赤芍,陈皮,煨瓦楞子 - > 香附	0.944 444
27	香附,紫苏梗,煨瓦楞子 - > 佛手	0.942 857
28	香附,紫苏梗,煨瓦楞子 - > 陈皮	0.941 176
29	佛手,紫苏梗,煨瓦楞子 - > 香附	0.941 176
30	佛手,紫苏梗,煨瓦楞子 - > 陈皮	0.941 176
31	佛手,砂仁,香附 - > 陈皮	0.941 176
32	陈皮,紫苏梗,煨瓦楞子 - > 香附	0.941 176
34	陈皮,紫苏梗,煨瓦楞子 - > 佛手	0.941 176
35	砂仁,煨瓦楞子 - > 陈皮	0.941 176

表 5 基于复杂系统熵聚类的治疗胃脘痛的核心组合

序号	核心组合	序号	核心组合
1	旋覆花-甘草-炒神曲	26	砂仁-炒神曲-薏仁
2	竹茹-全瓜蒌-元明粉	27	延胡索-炒川楝子-青皮
3	木香-生薏苡仁-炒谷芽	28	延胡索-青皮-柴胡
4	木香-生薏苡仁-炒麦芽	29	玉竹-煨瓦楞子-香附
5	木香-炒谷芽-大腹皮	30	生龙牡-夜交藤-炒枣仁
6	木香-炒麦芽-大腹皮	31	生龙牡-夜交藤-珍珠母
7	续断-土茯苓-怀牛膝	32	生龙牡-夜交藤-酸枣仁
8	续断-鱼腥草-怀牛膝	33	生龙牡-酸枣仁-白菊花
9	枳壳-青皮-柴胡	34	佛手-煨瓦楞子-香附
10	枳壳-柴胡-郁金	35	佛手-煨瓦楞子-白术
11	神曲-砂仁-炒神曲	36	佛手-紫苏梗-香附
12	旋覆花-佛手-煨瓦楞子	37	夜交藤-炒枣仁-生龙骨
13	大枣-炙甘草-半枝莲	38	夜交藤-炒枣仁-桑寄生
14	生薏苡仁-炒谷芽-泽泻	39	夜交藤-生龙骨-远志
15	生薏苡仁-炒麦芽-泽泻	40	柴胡-赤芍-郁金
16	生薏苡仁-泽泻-绿萼梅	41	决明子-火麻仁-蒲公英
17	黄芩-砂仁-郁金	42	决明子-火麻仁-生首乌
18	麦冬-煨瓦楞子-香附	43	炒白芍-炙甘草-桂枝
19	麦冬-紫苏梗-香附	44	火麻仁-生首乌-熟地黄
20	当归-炒谷芽-泽泻	45	火麻仁-生首乌-旱莲草
21	当归-丹参-泽泻	46	薏仁-绿萼梅-桑寄生
22	当归-炒麦芽-泽泻	47	白芍-青皮-柴胡-赤芍
23	当归-泽泻-绿萼梅	48	白芍-青皮-炒白芍-赤芍
24	当归-泽泻-炒白术	49	白芍-柴胡-赤芍-香附
25	当归-绿萼梅-茯苓		

表 6 基于熵层次聚类的治疗胃脘痛新处方

序号	候选新处方
1	旋覆花, 谷芽, 麦芽, 佛手, 煨瓦楞子
2	白茅根, 贝母, 土茯苓, 鱼腥草
3	全瓜蒌, 决明子, 生首乌, 升麻, 瓜蒌
4	炒黄柏, 何首乌, 沉香木, 黄花地丁, 石韦, 瞿麦, 野菊花, 车前草, 紫花地丁
5	红花, 黄芪, 桃仁, 天冬
6	谷芽, 柏子仁, 麦芽, 紫苏梗
7	茵陈, 虎杖, 鸡内金, 炒山栀
8	贝母, 苦参, 百合, 牡蛎, 芡实
9	生薏苡仁, 绿萼梅, 茯苓, 炒谷芽, 炒麦芽
10	麦冬, 生地黄, 枸杞子, 沙参, 玉竹, 杏仁
11	当归, 绿萼梅, 茯苓, 炒谷芽, 炒麦芽
12	白豆蔻, 制香附, 炒神曲, 高良姜, 薏仁

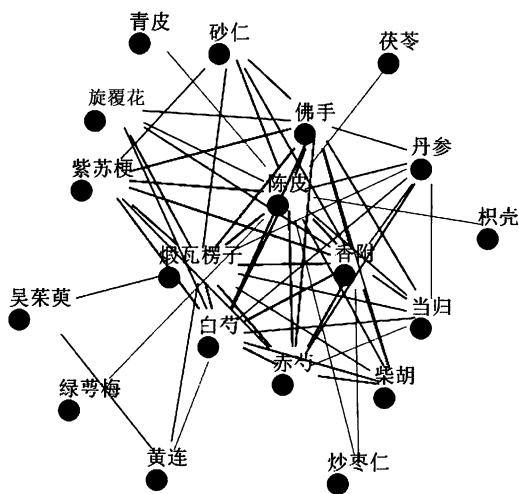


图 1 关联规则网络展示 (支持度 30, 置信度 0.9)

2.3 基于熵聚类的方剂组方规律分析

2.3.1 基于改进的互信息法的药物间关联度分析

依据方剂数量, 结合经验判断和不同参数提取数据的预读, 设置相关度为 8, 惩罚度为 4, 进行聚类分析, 得到方剂中两两药物间的关联度, 将关联系数 > 0.03 的药对列表。见表 4。

表 4 基于改进的互信息法的药物间关联度分析

药对	关联系数	药对	关联系数
煨瓦楞子-白术	0.042 37	白芍-黄连	0.031 525
煨瓦楞子-炒枳壳	0.038 104	沙参-煨瓦楞子	0.031 44
麦冬-石斛	0.037 45	煨瓦楞子-怀山药	0.031 44
煨瓦楞子-生地黄	0.036 876	柴胡-煨瓦楞子	0.030 962
生白术-大腹皮	0.034 025	干荷叶-大腹皮子	0.030 807
麦冬-紫苏梗	0.032 562	冬瓜仁-大腹皮子	0.030 807
白芍-旋覆花	0.032 311	麦冬-怀山药	0.030 575
玉竹-枸杞子	0.031 823	白芍-砂仁	0.030 458
柴胡-紫苏梗	0.031 744		

2.3.2 基于复杂系统熵聚类的药物核心组合分析 以改进的互信息法的药物间关联度分析结果为基础, 按照相关度与惩罚度约束, 基于复杂系统熵聚类, 演化出 3~4 味药核心组合。见表 5。

2.3.3 基于无监督熵层次聚类的新处方分析 在以上核心组合提取的基础上, 运用无监督熵层次聚类算法, 得到 29 个新处方, 见表 6。新方核心组合药物网络和新方药物网络见图 2~3。

续表 6

序号	候选新处方
13	生白芍, 佩兰, 莲子肉, 莲子心, 藿香梗
14	砂仁, 紫苏梗, 陈皮, 沙参, 麦冬, 玉竹, 生地黄
15	延胡索, 香橼皮, 合欢皮, 青皮, 柴胡
16	牡蛎, 芡实, 怀山药, 山萸肉, 鸡内金
17	白花蛇舌草, 半枝莲, 灵芝, 姜竹茹, 无柄灵芝
18	秦艽, 紫苏叶, 槟榔, 高良姜, 蔻仁
19	石见穿, 降香, 薤白, 石决明, 葛根, 石斛, 钩藤
20	生龙牡, 夜交藤, 炒枣仁, 生牡蛎, 生龙骨
21	茯苓皮, 金钱草, 赤小豆, 秦艽, 桑枝, 丝瓜络
22	枳实, 郁李仁, 火麻仁, 银花藤, 蒲公英, 生首乌
23	山茱萸, 五味子, 制首乌, 黄柏, 知母, 川续断
24	石决明, 葛根, 天花粉, 密蒙花, 谷精草, 山楂, 青箱子
25	石决明, 天麻, 怀牛膝, 竹茹, 生白芍, 莲子肉
26	炒白芍, 炙甘草, 桂枝, 大枣, 白花蛇舌草, 半枝莲
27	焦麦芽, 焦神曲, 炒扁豆, 大腹皮子
28	荆芥, 连翘, 防风, 桔梗, 蝉衣, 制僵蚕
29	柏子仁, 生龙牡, 珍珠母, 酸枣仁, 白菊花

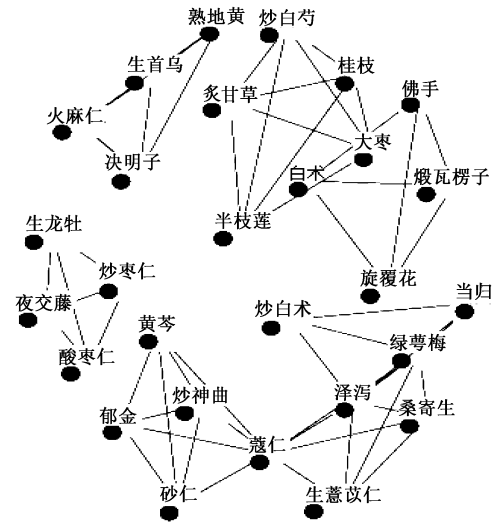


图 3 治疗胃脘痛新方药物网络图

陈皮等。经过聚类算法分析,常用药对包括:煨瓦楞子-白术、煨瓦楞子-炒枳壳、麦冬-石斛、煨瓦楞子-生地黄、生白术-大腹皮、麦冬-紫苏、白芍-旋覆花、玉竹-枸杞子、柴胡-紫苏梗等。基于复杂系统熵聚类的治疗胃脘痛的核心组合主要有:旋覆花-甘草-炒神曲、竹茹-全瓜蒌-元明粉、木香-生薏苡仁-炒谷芽、木香-生薏苡仁-炒麦芽、木香-炒谷芽-大腹皮、木香-炒麦芽-大腹皮等。基于熵层次聚类的治疗胃脘痛新处方主要有:①旋覆花、谷芽、麦芽、佛手、煨瓦楞子;②白茅根、贝母、土茯苓、鱼腥草;③全瓜蒌、决明子、生首乌、升麻、瓜蒌;④炒黄柏、何首乌、沉香木、黄花地丁、石韦、瞿麦、野菊花、车前草、紫花地丁;⑤红花、黄芪、桃仁、天冬等。

以上研究结果较好地验证了颜正华教授诊疗胃脘痛的治疗经验。颜教授认为,胃脘痛之病机虽变化多端,却总以虚实为纲,治疗不外补泻两途;补泻之中兼参寒热缓急。寒者散寒,停食者消食,气滞者理气,热郁者泄热,血瘀者化瘀,阴虚者益胃养阴,阳弱者温运脾阳。既往医案研究表明,颜教授胃脘痛治验思想全面,常从肝气犯胃、胃络瘀阻、寒邪伤胃、饮食失节、湿热阻胃、脾胃虚寒、阴伤胃痛等方面综合考量,辨证论治^[7-10]。纳入本研究的病案以肝气犯胃或脾胃气虚兼有气滞者居多,故颜老处方所用药物以疏肝理气、和胃止痛为主。如单味药出现频次最高者为陈皮,其味辛、苦,性温,归脾、肺经,功能理气健脾、燥湿化痰,善治中焦寒湿气滞、脘腹胀痛,亦可用于食积气滞之脘腹胀痛。佛手亦为出现频次较高的药物,其味辛、苦,性温,归肝、脾、胃、肺经,功能疏肝解郁、理气和中、燥湿化痰,用于脾胃气滞之脘腹胀痛。又如煨瓦楞子,味咸,性平,归肺、胃、肝

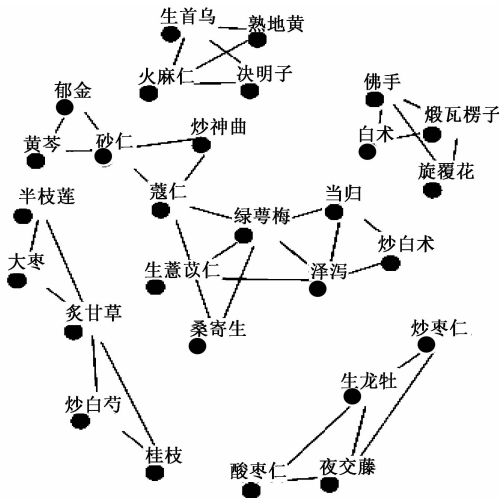


图 2 新方的核心组合药物网络图

3 讨论

本研究应用中医传承辅助系统软件,运用关联规则和聚类算法分析颜正华教授治疗胃脘痛的用药经验。经过关联算法分析,提炼出颜教授治疗胃脘痛常用的药物有陈皮、佛手、香附、白芍、煨瓦楞子、赤芍、当归、丹参、砂仁、紫苏梗、茯苓、枳壳、柴胡、旋覆花、炒枣仁、绿萼梅、黄连、吴茱萸、青皮等,这些药多数具有理气、止痛、健脾、和中、活血等功效。常用药物组合有:①当归、陈皮;②白芍、当归;③当归、佛手;④当归、香附;⑤当归、煨瓦楞子;⑥丹参、当归;⑦当归、赤芍;⑧陈皮、茯苓;⑨陈皮、砂仁;⑩白芍、

经。功能消痰软坚、化痰散结、制酸止痛。煨用后可制酸止痛,为颜教授治疗肝胃不和、胃痛吐酸最常用药。

颜教授认为,肝为刚脏,喜条达而主疏泄,如肝失于疏泄,横逆犯胃,则气机阻滞,胃脘胀痛。颜教授治疗胃脘痛证属肝气犯胃者常以疏肝理气、和胃解郁立法^[11-13]。本研究显示,出现频次和置信度较高的药物配伍均具有疏肝理气之功。如香附、白芍出现频次仅次于陈皮、佛手。李时珍称香附为“气病之总司,妇科之主帅。”香附味辛、微苦、微甘,性平,归肝、脾、三焦经,具疏肝解郁、调经止痛之功,常用于脾胃气滞、脘腹胀痛。白芍味苦、酸,微寒,归肝、脾经,能养血敛阴、柔肝止痛、平抑肝阳,治疗血虚肝郁、胁肋脘腹疼痛效佳。再如,颜教授方剂中,当归和白芍常同时出现。二者相伍能养血理血,治疗筋脉挛急、胃脘疼痛。其中,当归味甘、辛,性温,补血活血,调经止痛,用于血虚兼有瘀滞的心腹疼痛、月经不调等。白芍苦酸微寒,功善养血敛阴,柔肝止痛,善用于血虚偏热的患者,且柔肝兼能平肝,对于血虚肝旺的疼痛效果尤佳。上述两药一温一寒,一动一静,相制相合,相辅相成。

本研究基于中医传承辅助系统平台开展颜正华教授用药规律数据挖掘研究,获得了既往传统医案整理和统计学研究未获得的新知识、新信息,为颜正华教授胃脘痛治疗的深入挖掘和传承提供了参考。当然,应用关联规则和聚类研究得到的组方规则、核心组合和新处方等还需进一步临床验证,数据挖掘方法亦有其局限性,需结合中医药理论与实践综合分析、评价。同时,本研究过程亦表明,中医传承辅助系统为深入分析、挖掘名老中医经验提供了良好的平台,值得进一步推广和应用。

[参考文献]

[1] 杨洪军,赵亚丽,唐仕欢,等.基于熵方法分析中风病

方剂中药物之间关联度[J].中国中医基础医学杂志,2005,11(9):706.

[2] 唐仕欢,陈建新,杨洪军,等.基于复杂熵聚类方法的中药新药处方发现研究思路[J].世界科学技术——中医药现代化,2009,11(2):225.

[3] 李健,卢朋,唐仕欢,等.基于中医传承辅助系统的治疗肺癆方剂用药规律分析[J].中国实验方剂学杂志,2012,18(2):254.

[4] 李健,卢朋,张瑞贤,等.中医方剂大辞典中治疗肺癆方剂用药规律分析[J].中国实验方剂学杂志,2012,18(10):1.

[5] 卢笑晖,单琦玮.基于中医传承辅助系统分析的卢尚岭教授治疗头风病用药经验[J].中国实验方剂学杂志,2012,18(9):5.

[6] 孙敬昌,王燕平.基于中医传承辅助系统的治疗水肿方剂用药规律分析[J].中国实验方剂学杂志,2012,18(10):11.

[7] 吴嘉瑞,张冰.国医大师颜正华临床经验实录[M].北京:中国医药科技出版社,2011:149.

[8] 吴嘉瑞,张冰.颜正华胃脘痛治验举隅[J].中华中医药杂志,2009,24(12):1594.

[9] 张冰,吴嘉瑞.国医大师颜正华临床经验与用药思想探赜[M].中华中医药杂志,2009,24(6):742.

[10] 常章富.颜正华教授治疗慢性胃炎的经验[J].北京中医药大学学报,1996,19(3):24.

[11] 吴嘉瑞,张冰,叶恩培.国医大师颜正华诊疗消化系统常见病用药规律分析[J].国际中医中药杂志,2012,34(6):532.

[12] 吴嘉瑞,张冰,叶恩培.基于数据库分析的颜正华消化系统用药总体规律研究[J].中国实验方剂学杂志,2012,18(2):242.

[13] 张冰,吴嘉瑞.颜正华治学思想探析[J].中医杂志,2012,53(7):550.

[责任编辑 何伟]