

基于数据挖掘预防新型冠状病毒肺炎中医方药初探

赵 新^{1,2}, 张林旭³, 路雪婧^{1,2*}

1. 成都中医药大学, 四川 成都 610075

2. 中医药眼病防治与视功能保护四川省重点实验室, 四川 成都 610075

3. 河南中医药大学, 河南 郑州 450000

摘 要: 运用中医传承辅助平台 (V2.5) 挖掘分析各省市新型冠状病毒肺炎 (COVID-19) 预防方药, 总结探讨其潜在新方。收集检索国家及各省、自治区和直辖市卫健委、中医药管理局官方网站关于 COVID-19 中医药预防方案, 得到黑龙江、北京、天津、河北、河南、江西、四川、湖北、湖南等 17 个省市的 COVID-19 中医药预防方案, 运用频次统计、关联规则、基于熵聚类方法规律分析等方法进行药物规律及特点的挖掘。64 首处方中共包含 82 味中药, 使用频次较高的中药为黄芪、金银花等, 补虚药使用最多, 多用甘、温之品, 归肺经药物最多; 通过关联规则分析, 得到 8 个常用药物组合; 基于熵聚类方法规律分析得到 7 个潜在新方。各地多使用补虚扶正药物预防 COVID-19, 多从肺、脾胃着手, 虽具体细节不同, 但都体现了中医学的预防思想。

关键词: 新型冠状病毒肺炎; 中医药预防; 数据挖掘; 中医传承辅助平台; 药物规律

中图分类号: R183.3

文献标志码: A

文章编号: 0253-2670(2020)05-1147-06

DOI: 10.7501/j.issn.0253-2670.2020.05.008

Preliminary study of traditional Chinese medicine for preventing COVID-19 based on data mining

ZHAO Xin^{1,2}, ZHANG Lin-xu³, LU Xue-jing^{1,2}

1. Chengdu University of Traditional Chinese Medicine, Chengdu 610075, China

2. Sichuan Key Laboratory of Ophthalmopathy Prevention & Cure and Visual Function Protection with TCM, Chengdu 610075, China

3. Henan University of Traditional Chinese Medicine, Zhengzhou 450000, China

Abstract: To dig out and analyze the drug rule of COVID-19 prevention prescriptions from provinces and cities by using Traditional Chinese Medicine Inheritance Support System, summarize and explore its potential new prescription. The Chinese medicine prevention programs for COVID-19 were collected and searched from the official website of the Health Commission and State Administration Medicine of Traditional Chinese Medicine of the country and provinces, autonomous regions, and municipalities. TCM prevention programs in 17 provinces including Heilongjiang, Beijing, Tianjin, Hebei, Henan, Jiangxi, Sichuan, Hubei, and Hunan, etc were received. A total of 82 herbs were included in 64 prescriptions, the most frequently used Chinese herbs were *Astragalus membranaceus*, *Lonicera japonica*, etc. Tonifying deficiency drugs with sweet and warm natures were used the most, the Chinese herbs distributed in the lung channel was the most in channel tropism drugs. Analysis by association rule, eight combinations of commonly used drugs were obtained. Based on entropy clustering method rule analysis, seven potential new prescriptions were obtained. Tonifying deficiency drugs are often used in various places to prevent COVID-19, focusing on the lung, spleen and stomach. Although the specific details are different, they all reflect the preventive thinking of traditional Chinese medicine.

Key words: COVID-19; Chinese medicine prevention; data mining; Traditional Chinese Medicine Inheritance Support System; drug law

收稿日期: 2020-02-16

基金项目: 第三届国医大师传承工作室建设 (003109011001); 川派中医眼科名家-廖品正 (2015CP017); 基于“道术结合”思路与多元融合方法的名老中医经验传承创新研究 (2018YFC1704100); 西部地区名老中医学学术观点、特色诊疗方法和重大疾病防治经验研究 (2018YFC1704104)

作者简介: 赵 新, 在读研究生, 研究方向为青光眼、眼底病及视神经保护。Tel: 18382284466 E-mail: 979212981@qq.com

***通信作者** 路雪婧, 博士, 教授, 博士生导师, 研究方向为青光眼、眼底病及视神经保护。Tel: 18908178510 E-mail: xg1214@126.com

2019 年 12 月,我国武汉市出现不明原因病毒感染肺炎病例,通过对肺炎患者的样本进行测序,发现了前所未有的 β 冠状病毒,2020 年 2 月 21 日国家卫生健康委员会将“新型冠状病毒肺炎”英文名称修订为“COVID-19”^[1-2]。COVID-19 主要传播途径是经呼吸道飞沫传播,亦可通过接触传播^[3]。截止 2020 年 2 月 26 日 24 时,我国累计报告 COVID-19 确诊病例 78 497 例,现有疑似病例 2 358 例^[4]。“上工不治已病治未病”,我国及各省市、自治区的卫生健康委员会、中医药管理局根据疾病特点、当地气候等情况提出多个防治 COVID-19 的方案。本文尽可能全面收集全国各省市发布的 COVID-19 预防方案,利用数据挖掘,提炼核心的药物组合,分析总结一些潜在的用药规律,为开发具有预防 COVID-19 作用的方药提供参考。

1 资料与方法

1.1 资料来源

收集检索国家及各省、自治区和直辖市卫健委、中医药管理局官方网站关于 COVID-19 中医药预防方案,得到黑龙江、北京、天津、河北、河南、江西、四川、湖北、湖南等 17 个省市的中医药预防方案^[5-21],共 64 首口服中药处方。

1.2 分析软件

中医传承辅助平台(V2.5)是由中国中医科学院中药研究所联合中国科学院自动化研究所共同研究开发,能将临床数据与半自助模式、个性化需求有效地结合起来,便于总结疾病用药规律,探讨疾病的中医治则治法,发现中药新方^[22-24]。

1.3 数据录入与核对

运用中医传承辅助平台中的“临床采集”将各省市的预防方药逐一录入。为确保数据的正确性,录入完成后,由双人单独核对录入数据。

1.4 数据挖掘

运用中医传承辅助平台中的“数据分析”以及“统计报表”对收集的 64 首处方进行用药规律的挖掘。

2 结果

2.1 频次统计分析

频次指变量值中代表某种特征的数(标志值)出现的次数。收集的 64 首处方中共包括 82 味中药,共计 439 味次。在 82 味中药中,使用频次在 10 以上的药物有 13 味,依次为黄芪、金银花、防风、甘草等,见表 1。手动将 82 味中药按照教材《中药学》进行功效分类,将药物分类按照出现的频次从高到

低进行排序,见表 2。药物四气频次最高为温性,其次为寒性,具体结果见图 1。药物五味频次最高为甘味,其次为苦味,具体结果见图 2。药物归经频次最高为肺经,其次为脾经、胃经,具体结果见图 3。

2.2 关联规则分析

关联规则是发现大量数据中项集之间频繁模式、相关关系以及因果结构的一种算法,其可探寻处方中的高频药组,挖掘常用药物组合,总结配伍

表 1 药材频次分布

Table 1 Frequency distribution of CMM

序号	药材	频次	序号	药材	频次
1	黄芪	35	8	贯众	18
2	金银花	31	9	麦冬	15
3	防风	24	10	苍术	14
4	甘草	22	11	陈皮	13
5	藿香	20	12	芦根	12
6	炒白术	19	13	连翘	12
7	桔梗	19			

表 2 药物分类频次分布

Table 2 Frequency distribution of CMM classification

序号	药物分类	频次	序号	药物分类	频次
1	补虚药	144	9	活血化痰药	3
2	清热药	101	10	驱虫药	2
3	解表药	72	11	平肝熄风药	2
4	化湿药	48	12	止血药	1
5	化痰止咳平喘药	25	13	消食药	1
6	理气药	17	14	泻下药	1
7	利水渗湿药	13	15	温里药	1
8	收涩药	7	16	开窍药	1

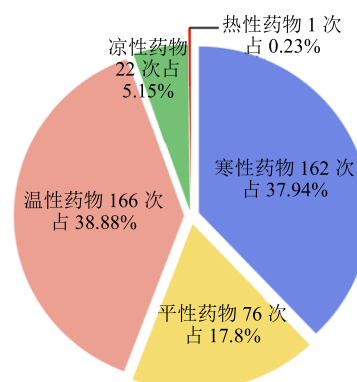


图 1 药物四气频次分布

Fig. 1 Frequency distribution of CMM four natures

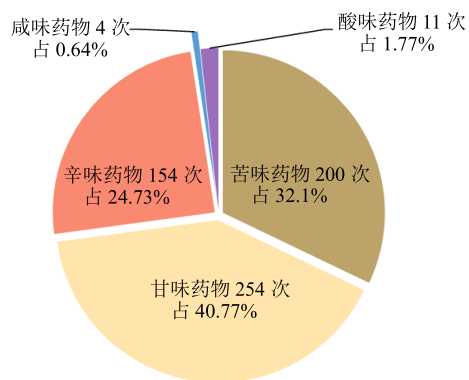


图 2 药物五味频次分布

Fig. 2 Frequency distribution of CMM five flavors

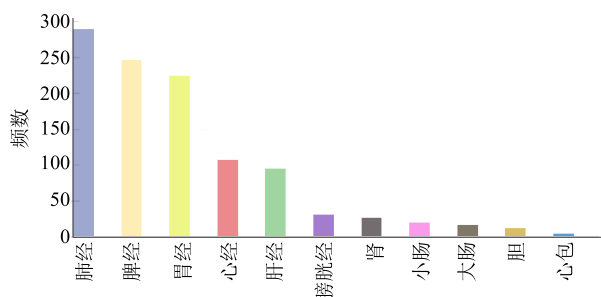


图 3 药物归经频次统计

Fig. 3 Statistics of CMM channel tropism

规律，为创制新方提供依据^[25]。设置支持度个数为 12（支持度个数表示药物组合在所选处方中出现的频次，支持度个数=方剂数量×支持度，平台默认支持度为 20%，如本次收集 64 首处方，系统默认支持度个数为 12），置信度为 0.6（置信度表示药物 A 出现时，药物 B 出现的概率，该值越接近于 1，概率越大，A 在左边，B 在右边，如防风→黄芪 0.958，表示防风出现时，黄芪出现的概率为 95.8%）。获取 8 条常用药物组合（共涉及 6 味中药），药物组合出现频次最高的是黄芪、防风，具体见表 3；在此基础上点击“规则分析”，共获取 12 条药物组合关联规则，并按置信度的大小排列，具体见表 4；6 味中药相互关系的网络展示见图 4。

表 3 常用药物组合

Table 3 Common CMM combination

序号	药物模式	频次	序号	药物模式	频次
1	黄芪、防风	23	5	金银花、黄芪	16
2	黄芪、炒白术	17	6	甘草、桔梗	13
3	炒白术、防风	17	7	黄芪、甘草	13
4	黄芪、炒白术、防风	16	8	黄芪、桔梗	12

表 4 药物组合关联规则

Table 4 Association rule of CMM combination

序号	规则	置信度	序号	规则	置信度
1	防风→黄芪	0.958	7	防风→炒白术	0.708
2	炒白术，防风→黄芪	0.941	8	黄芪，防风→炒白术	0.696
3	黄芪，炒白术→防风	0.941	9	桔梗→甘草	0.684
4	炒白术→黄芪	0.895	10	防风→黄芪，炒白术	0.667
5	炒白术→防风	0.895	11	黄芪→防风	0.657
6	炒白术→黄芪，防风	0.842	12	桔梗→黄芪	0.632

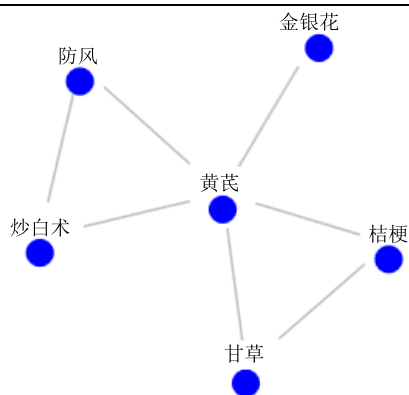


图 4 药物关系网络

Fig. 4 Network of CMM relationship

2.3 基于熵聚类方法规律分析

熵理论非常适合用于挖掘具有混合型、离散型和连续型等特点的中医药数据，可以定量、定性挖掘药物之间的相关性，也可以挖掘名医经验的核心组合以及隐藏的用药规律^[25]。为减少不必要的计算，同时防止数据的丢失，平台默认相关度为 8，惩罚度为 2。相关度是熵聚类算法中的一个重要参数，处方是由若干药物组成，一个药物与另一个药物必然存在一定的关联性，根据改进的互信息法的数学算法，药物与药物之间的关联可以进行定量描述，相关度为 8，表示仅取排序在前面的 1~7 的药物与药物之间的关联度，共 8 味药物，然后进行聚

类分析。惩罚度是为了减少负面数据信息干扰的一个参数,药物与药物之间的关联有正相关也有负相关,为了避免负相关对数据分析的影响,特设立惩罚度加以排除。点击新方分析,运用改进的互信息法和熵聚类的分析方法得到 14 条核心药物组合,具体见表 5,部分核心药物组合的网络展示见图 5;在核心组合基础上,使用无监督的熵层次的聚类算法,挖掘得到 7 个潜在新方,具体见表 6,网络展示如图 6 所示。

3 讨论

此次 COVID-19 属于中医“疫病”范畴^[3],发病急骤,传染性强,易于流行,预防工作迫在眉睫。

表 5 核心药物组合

Table 5 Core CMM combination

序号	核心组合	序号	核心组合
1	连翘_板蓝根_荆芥	8	连翘_板蓝根_白术
2	连翘_金银花_太子参	9	羌活_金银花_淡豆豉
3	桑叶_菊花_枸杞子	10	桑叶_菊花_青果
4	白芍_茯苓_枸杞子	11	白芍_茯苓_牛膝
5	白芍_菊花_枸杞子	12	白芍_菊花_牛膝
6	苍术_炒白术_贯众	13	防风_炒白术_玄参
7	藿香_薄荷_牛蒡子	14	藿香_陈皮_白术

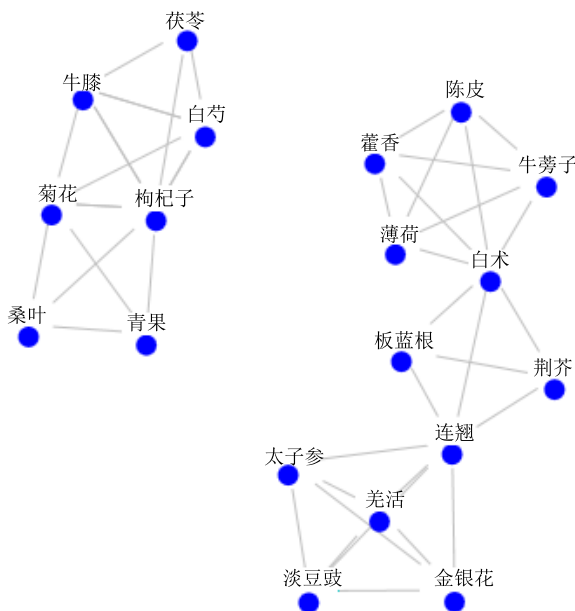


图 6 潜在新方网络

Fig. 6 Network of potential new prescription

早在几千年前,祖国医学就提出了预防疾病的重要性,《黄帝内经》曰:“圣人不治已病治未病,不治已乱治未乱……”,中医的“治未病”思想体现了“未

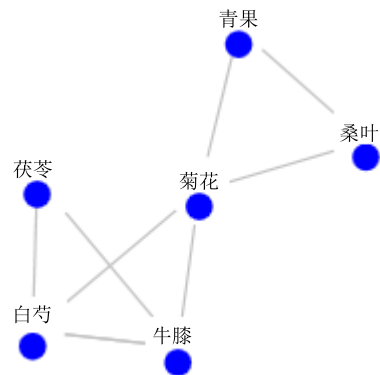


图 5 核心药物组合网络

Fig. 5 Network of core CMM combination

表 6 潜在新方

Table 6 Potential new prescription

序号	新方组合
1	连翘_板蓝根_荆芥_白术
2	连翘_金银花_太子参_羌活_淡豆豉
3	桑叶_菊花_枸杞子_青果
4	白芍_茯苓_枸杞子_牛膝
5	白芍_菊花_枸杞子_牛膝
6	苍术_炒白术_贯众_防风_玄参
7	藿香_薄荷_牛蒡子_陈皮_白术

病先防”的预防观念。正气不足是疾病发生的主导因素,邪气是发病的重要条件,疾病的发生主要与“正气”和“邪气”两方面有关。“正气存内,邪不

可干”，增强人体正气，抵御致病因素的侵害，邪气无隙可乘是“未病先防”的关键所在。正气概念源于《内经》，是一身之气抵抗外感邪气入侵时的称谓，又以“精气”“真气”称之。人体的“气”主要由肾中“先天之气”、肺呼吸的自然界“清气”、脾胃运化的“水谷精气”3 部分构成。肺为生气之主，脾胃为生气之源，肾为生气之根，肺和脾胃与后天之气的生成关系密切，故各地预防疾病，调充正气，多从肺、脾胃着手，补虚扶正药物和补益和中的甘味药物使用最多。

“邪之所凑，其气必虚”，各地预防方多以“玉屏风散”为加减。选用黄芪甘温之品，内补脾肺之气，外固表实卫；白术甘苦性温，归脾、胃经，可益气健脾、燥湿止泻、培土生金；防风为治风通用药，可泻肺实，除上焦风邪，与黄芪相配则祛邪不伤正，固表不留邪。有专家指出^[26-30]，本次疫病以“湿邪”为主，疫疠之邪从口鼻而入，直入肺脾。肺体清虚娇嫩，不耐寒热，易受邪袭。湿邪重浊黏腻，易蕴积脾胃，阻滞气机，出现脘痞、纳差、欲呕、便溏等胃肠不适。此次 COVID-19 首次发病于己亥年终，土运不及，人与天地相参，与日月相应，体内则可表现为脾土不足。脾主运化，为气血生化之源，脾虚失运，水谷精微无法运化，气、血、津液化生不足，无法营养四肢百骸，则出现乏力、四肢困重；气血亏虚，则正气不足，易受邪气侵袭，感受湿邪会进一步加重脾虚。故各地预防方药中多以补气益肺、健脾化湿之品为主，如黄芪、白术、党参、太子参、苍术、藿香、佩兰等。

COVID-19，病位在肺，基本病机特点为“湿、热、毒、瘀”^[31]。本病首发于 2019 年 12 月，今年冬天当寒反温，湿邪易夹燥，预防宜选用养阴润肺、生津润燥、开宣肺气的药物，如桔梗、芦根、麦冬、百合、北沙参，佐以配伍清热解毒药、化湿药、健脾渗湿药、活血化瘀药等。“六气伤人必随人身之气而变”，现代人们多食肥甘燥辣之品，生活节奏快，五志过极，木脏疏泄失常，郁而化热，预防方药中可酌加清热药。此外，有研究提出清热解毒类中药可抗病毒^[32]。

各地卫生部门，因地制宜，审证求因，酌情化裁，制定合适的预防方药。本文利用频次统计、关联规则和熵聚类方法挖掘 COVID-19 预防方药中显性、隐形的药物规律和潜在新方，熵聚类主要是挖掘低频次高相关药物的隐形关系，与频次统计和关

联规则挖掘的显性药物关系不同，7 个潜在新方组合的临床价值还需进一步结合实践验证。希望本文所挖掘的数据能够为疫病的防控提供一定的参考借鉴，降低疫情伤害。

参考文献

- [1] Zhu N, Zhang D, Wang W, *et al.* A Novel coronavirus from patients with pneumonia in China, 2019 [J]. *N Engl J Med*, 2020, 382: 727-733.
- [2] 国家卫生健康委员会. 国家卫生健康委关于修订新型冠状病毒肺炎英文命名事宜的通知 [EB/OL]. [2020-02-22]. <http://www.nhc.gov.cn/xcs/zhengcwj/202002/6ed7614bc35244cab117d5a03c2b4861.shtml>.
- [3] 国家中医药管理局办公室国家卫生健康委办公厅. 关于印发新型冠状病毒感染的肺炎诊疗方案（试行第五版）的通知 [EB/OL]. [2020-02-05]. <http://www.nhc.gov.cn/yzygj/s7653p/202002/3b09b894ac9b4204a79db5b8912d4440.shtml>.
- [4] 国家卫生健康委员会. 截至 2 月 26 日 24 时新型冠状病毒肺炎疫情最新情况 [EB/OL]. [2020-02-27]. <http://www.nhc.gov.cn/xcs/yqfkdt/202002/369fb309eded45abaf10645a9c8ceb40.shtml>.
- [5] 甘肃省卫生健康委员会. 关于印发甘肃省新型冠状病毒感染的肺炎中医药防治方案（试行第二版）的通知 [EB/OL]. [2020-02-01]. <http://wsjk.gansu.gov.cn/file.jsp?contentId=83488>.
- [6] 北京市中医管理局. 北京市中医管理局关于印发《北京市新型冠状病毒感染的肺炎防治方案（第二版）》的通知 [EB/OL]. [2020-01-30]. http://zyj.beijing.gov.cn/zyyxxw/202001/t20200130_1621630.html.
- [7] 黑龙江省卫生健康委员会. 关于印发《黑龙江省新型冠状病毒感染的肺炎中医药防治方案（试行）》的通知 [EB/OL]. [2020-02-04]. <http://wsjkw.hl.gov.cn/index.php/Home/Ztzt/show/newsid/722>.
- [8] 贵州省中医药管理局. 贵州省中医药管理局公布预防“新型冠状病毒肺炎”中医处方、民族医处方 [EB/OL]. [2020-01-22]. http://www.gz.xinhuanet.com/2020-01/22/c_1125494617.htm.
- [9] 河北中医. 河北省新型冠状病毒肺炎中医防治方案（试行第四版）发布 [EB/OL]. [2020-02-12]. <https://mp.weixin.qq.com/s/9QTqGDW6vkQyWX-VrjulqQ>.
- [10] 河南省中医管理局. 河南省发布新型冠状病毒感染的肺炎中医药预防方案 [EB/OL]. [2020-01-27]. <http://www.tcm.gov.cn/hydt/1700.htm>.
- [11] 湖北省人民政府. 抗击新型肺炎 湖北省中医院制定 8 个中医药防治协定方 [EB/OL]. [2020-01-25]. http://www.hubei.gov.cn/zhuanti/2020/gzxxgzbd/qfqk/202001/t20200125_2014934.shtml.

- [12] 陕西省卫生健康委员会. 陕西省卫生健康委关于印发新型冠状病毒感染的肺炎中医药防治方案的通知 [EB/OL]. [2020-01-23]. http://sxwjw.shaanxi.gov.cn/art/2020/1/23/art_10_67378.html.
- [13] 安徽省中医药管理局. 省中医药学会发布中医药预防呼吸道传染病处方 [EB/OL]. [2020-01-22]. <http://wjw.ah.gov.cn/ahtcm/NewsDetail.aspx?id=940>.
- [14] 天津市卫生健康委员会. 市卫生健康委关于印发天津市新型冠状病毒感染的肺炎中医药防治方案 (试行)的通知 [EB/OL]. [2020-01-27]. http://wsjk.tj.gov.cn/art/2020/1/27/art_14_70212.html.
- [15] 天津市卫生健康委员会. 市卫生健康委关于印发天津市新型冠状病毒感染的肺炎中医药防治方案 (试行第二版) 的通知 [EB/OL]. [2020-02-03]. http://wsjk.tj.gov.cn/art/2020/2/3/art_14_70468.html.
- [16] 天津市卫生健康委员会. 市卫生健康委关于印发天津市新型冠状病毒感染的肺炎中医药预防方案的通知 [EB/OL]. [2020-01-28]. http://wsjk.tj.gov.cn/art/2020/1/28/art_14_70285.html.
- [17] 四川省中医药管理局. 四川省中医药管理局关于印发《四川省新型冠状病毒感染的肺炎中医药防控技术指南》的通知 [EB/OL]. [2020-02-25]. <http://sctcm.sc.gov.cn/ctcm/cdzyyfkdt/2020/2/5/f936288f96344837b66ef952e4382946.shtml>.
- [18] 云南省卫生健康委员会. 云南省卫生健康委关于印发新型冠状病毒感染的肺炎中医药防治方案 (试行第二版) 的通知 [EB/OL]. [2020-02-11]. <http://ynswsjkw.yn.gov.cn/wjwWebsite/web/doc/UU158140615448559052>.
- [19] 湖北省人民政府. 中医专家的抗疫防护法 守正、防寒、祛湿、辟疫 [EB/OL]. [2020-02-08]. http://www.hubei.gov.cn/zhuanti/2020/gzxxgzbd/fkbp/202002/t20200208_2021480.shtml.
- [20] 中国新闻网. 海南省新型冠状病毒感染的肺炎中医药防治方案 (公众版试行第二版) [EB/OL]. [2020-02-03]. <http://www.hi.chinanews.com.cn/hnnew/2020-02-03/513709.html>.
- [21] 湖南省中医药管理局. 关于印发《湖南省新型冠状病毒感染的肺炎中医药诊疗方案 (试行第二版)》的通知 湘中医药函 (2020) 11 号 [EB/OL]. [2020-01-26]. http://tcm.hunan.gov.cn/tcm/xxgk/tzgg/202001/t20200126_11164462.html.
- [22] 吴嘉瑞, 唐仕欢, 郭位先, 等. 基于数据挖掘的名老中医经验传承研究述评 [J]. 中国中药杂志, 2014, 39(4): 614-617.
- [23] 卢朋, 李健, 唐仕欢, 等. 中医传承辅助系统软件开发与应用 [J]. 中国实验方剂学杂志, 2012, 18(9): 1-4.
- [24] 唐仕欢, 杨洪军. 中医组方用药规律研究进展述评 [J]. 中国实验方剂学杂志, 2013, 19(5): 359-363.
- [25] 杨洪军, 等. 中医传承辅助平台的开发与应用 [M]. 福州: 福建科学技术出版社, 2013.
- [26] 苗青, 丛晓东, 王冰, 等. 新型冠状病毒感染的肺炎的中医认识与思考 [J/OL]. 中医杂志, [2020-02-26]. <http://kns.cnki.net/kcms/detail/11.2166.R.20200205.1606.002.html>.
- [27] 李宝乐, 李小叶, 任顺平, 等. 结合 3~5 版《新型冠状病毒感染的肺炎诊疗方案》和古代文献初探新型冠状病毒肺炎的中医药辨治 [J]. 中草药, 2020, 51(4): 873-877.
- [28] 马家驹, 陈明, 王玉光. 新型冠状病毒 (2019-nCoV) 综合征中医证治述要 [J/OL]. 北京中医药, [2020-02-26]. <http://kns.cnki.net/kcms/detail/11.5635.R.20200207.1616.002.html>.
- [29] 王玉光, 齐文升, 马家驹, 等. 新型冠状病毒 (2019-nCoV) 肺炎中医临床特征与辨证治疗初探 [J/OL]. 中医杂志, [2020-02-26]. <http://kns.cnki.net/kcms/detail/11.2166.R.20200129.1258.002.html>.
- [30] 徐旭, 张莹, 李新, 等. 各地区中医药预防新型冠状病毒肺炎 (COVID-19) 方案分析 [J]. 中草药, 2020, 51(4): 866-872.
- [31] 国家中医药管理局办公室国家卫生健康委办公厅. 关于印发新型冠状病毒感染的肺炎诊疗方案 (试行第三版) 的通知 [EB/OL]. [2020-01-23]. <http://www.nhc.gov.cn/zyygj/s7653p/202001/f492c9153ea9437bb587ce2ffcbce1fa.shtml>.
- [32] 刘菊, 崔瑛, 白明学, 等. 基于中医药预防治疗新型冠状病毒肺炎的用药探析 [J/OL]. 中草药, [2020-02-26]. <http://kns.cnki.net/kcms/detail/12.1108.R.20200212.1133.002.html>.