

DOI:10.13288/j.11-2166/r.2014.07.020

中医药防治传染病专栏

中医药治疗流行性感冒文本挖掘结果与
诊疗方案的比较分析赵 静¹, 郭洪涛², 韩经丹¹, 李 立¹, 寇 爽¹, 高 原¹,
姜 帆¹, 吕 诚¹, 王燕平¹, 吕爱平[✉]

(1. 中国中医科学院中医临床基础医学研究所, 北京市东直门内南小街 16 号, 100700; 2. 河南中医学院附属第一医院)

基金项目: 国家中医药管理局中医药行业科研专项资助项目 (200907001); 国家自然科学基金资助项目 (81102872); 中国中医科学院基本科研业务费自主选题资助项目 (ZZ060814, Z0256)

摘要 目的 对中医药治疗流行性感冒 (简称流感) 文本挖掘结果及诊疗方案进行比较分析, 为流感防控方案的制定提供文献学支持。方法 对中国生物医学文献数据库中收集的流感相关文献进行文本挖掘, 得到流感的常见中医证候和方药, 并与中华人民共和国卫生部颁布的 7 个不同版本流感诊疗方案中的中医证候和方药使用情况进行对比, 分析其异同。结果 流感常见中医证候为风热犯卫、湿热证、少阳病; 常用方剂为银翘散、麻杏石甘汤与黄连解毒汤; 常用中药依次为: 板蓝根、金银花、黄连、黄芩、连翘, 且金银花常与连翘、黄芩、板蓝根联用。在流感诊疗方案中, 出现频率较高的中医证候为风热犯卫、热毒袭肺、气营两燔; 常用方剂为麻杏石甘汤、银翘散; 常用中药依次为: 生石膏、炙甘草/生甘草、炙麻黄、杏仁、黄芩。结论 流感证候以风热犯卫的流感轻症多见, 治疗方药以银翘散与麻杏石甘汤为主, 文本挖掘与诊疗方案中证候与方药基本一致, 但文本挖掘结果以流感轻症为主, 诊疗方案则兼顾流感重症。

关键词 流行性感冒; 证候; 中药; 诊疗方案; 文本挖掘

文本数据挖掘是指从文本数据中抽取有价值的信息和知识的计算机处理技术, 是数据挖掘中的一个重要分支, 近年来被广泛地用于中医药研究中^[1-2]。在新发突发传染病中医药方案制定过程中, 文本挖掘技术发挥了积极作用。在方案制定前, 全面检索相关文献, 为专家制定方案提供有据可循的支持与参考, 弥补了中医药传染病临床研究不足的缺陷; 在方案制定后, 对疾病规律与方案使用情况进行总结, 与临床和基础研究结果一并为方案的修订、更新提供依据。

随着 2009 年甲型 H1N1 流感、2013 年人感染 H7N9 禽流感等新型流感的不断爆发, 国内报道的流感诊疗相关文献急剧增加。我们借助文本挖掘技术, 对近年来中医药治疗流行性感冒 (简称流感) 的文献进行系统挖掘, 并结合流感诊疗方案对文本挖掘结果进行分析, 以期对流感的防控提供更多支

持与参考。

1 资料和方法

1.1 文本数据收集

检索中国生物医学文献数据库 (CBM), 以“(流行性感冒 or 流感) AND (中医 or 中药)”为检索策略, 检索日期截止至 2013 年 7 月 17 日。获得每篇文献的流水号、标题、摘要、主题词等信息。

1.2 文本数据处理

将收集的数据按照下载顺序整合, 以 ANSI 编码格式保存, 并利用自主研发的文本提取软件 (软著登字第 0261882 号, 登记号 2010SR073409) 对“文本数据收集”中下载的非结构化文本数据进行机标关键词 (简称关键词) 的信息提取, 保存成格式化的、便于数据库 (Access) 和大型数据库 (SQL) 处理的格式。

1.3 数据一次清洗

将文本数据处理产生的数据表导入 SQL 中,

* 通讯作者: lap64067611@126.com, (010) 64014411 转 2343

根据表格内容对表格进行命名, 针对“序号”和“机标关键词”进行处理。根据课题组既往构建的分层切片算法^[3], 通过数据一次处理、逆向规则降噪清洗等步骤^[4-6], 最后得到的结果应用 Cytoscape 软件成图。

2 流感文本挖掘结果

本研究共检索到文献 1351 篇, 通过文本挖掘得到流感的证候、治疗方药等相关结果。对于证候与中药的使用情况, 采用一维频次 (即频数) 来显示; 对于不同证候及中药间的关系, 采用二维频次 (即两个证候或中药同时出现) 所形成的网络图来表示。

2.1 常见中医证候

表 1 示, 近年流感证候以热证、湿证居多, 其中风热犯卫、湿热证与少阳病居前 3 位; 其次为风热证、寒热错杂、表寒里热、风寒束表证。病位多在表, 即卫分、少阳及肺。根据不同辨证方法证候诊断有所不同, 其中卫气营血辨证中多表现为卫分证和气分证, 六经辨证中多表现为少阳病、太阳病, 八纲辨证多表现为湿热证、风热证、寒热错杂、表寒里热、风寒束表、风热表证、风热犯肺、肺热壅盛等证。

表 1 流行性感冒常见中医证候频次 (频次 ≥ 3)

证候	频次	证候	频次
风热犯卫	13	气营两燔	3
湿热证	11	热毒闭肺	3
少阳病	10	太阳病	3
风热证	8	太阳伤寒	3
寒热错杂	7	痰热闭肺	3
表寒里热	6	卫分证	3
风寒束表	6	心阳虚	3
风热表证	5	风寒湿痹	3
风热犯肺	5	风热疫毒	3
肺热壅盛	4	风热闭肺	3
气分证	4		

图 1 为流感常见中医证候二维网络图, 每个节点代表一个证候, 证候间的连线表示两个证候在文献中同时出现, 风热证、风热犯表和风热犯卫等证位于网络的中心, 与其他证候连线较多, 为流感的核心证候。

2.2 常用中药

表 2 示, 治疗流行性感冒常用中药前 10 位中以清热解毒药、清热燥湿药为主 (7 味), 其次为和解少阳 (1 味)、调和药性 (1 味)、滋阴润肺 (1 味) 及消食健脾 (1 味)。

表 2 流行性感冒常用中药频次表 (频次 ≥ 20)

中药	频次	中药	频次
板蓝根	104	黄芪	26
金银花	99	牛蒡子	25
黄连	94	鱼腥草	25
黄芩	83	生石膏	24
连翘	82	薄荷	24
柴胡	72	莪术	22
穿心莲	51	桔梗	22
甘草	49	藿香	22
天冬	43	葛根	21
菊花	42	牛黄	20
麦芽	42	生地黄	20
麻黄	33	栀子	20
大青叶	29		

2.3 常用药对

表 3 示, 通过对常用药对分析, 使用频率最多的药对为金银花与连翘, 其次为黄芩与金银花、板蓝根与金银花。图 2 示, 流感治疗过程中使用频率较高的药物组合为银翘散包含的主要药物 (以菱形表示): 金银花、连翘、竹叶、淡竹叶、荆芥、牛蒡子、淡豆豉、薄荷、桔梗和芦根; 麻杏石甘汤包含的药物 (以八角形表示): 麻黄、杏仁、生石膏和甘草; 以及黄连解毒汤 (以灰色圆形表示) 中的黄连、黄芩、栀子。

表 3 流行性感冒常用中药药对频次表 (频次 ≥ 10)

药对	频次	药对	频次
金银花 连翘	54	甘草 桔梗	14
黄芩 金银花	29	黄连 连翘	14
板蓝根 金银花	26	桔梗 连翘	13
黄芩 连翘	24	金银花 桔梗	13
板蓝根 连翘	22	板蓝根 大青叶	13
甘草 金银花	17	金银花 牛蒡子	13
甘草 连翘	17	薄荷 甘草	11
连翘 牛蒡子	16	黄连 金银花	11
薄荷 连翘	16	金银花 荆芥	11
黄连 黄芩	16	甘草 牛蒡子	11
薄荷 金银花	15	连翘 竹叶	10

3 流感诊疗方案中中医证候和中药使用情况

对 2009 年后中华人民共和国卫生部颁布的 4 版甲型 H1N1 流感方案^[7-10]、1 版流感方案^[11] 和 2 版人感染 H7N9 禽流感方案^[12-13] 的中医证候一维频次统计可以看出, 出现频率较高的为风热犯卫、热毒袭肺、气营两燔、热毒壅肺/毒热壅肺, 详见表 4。

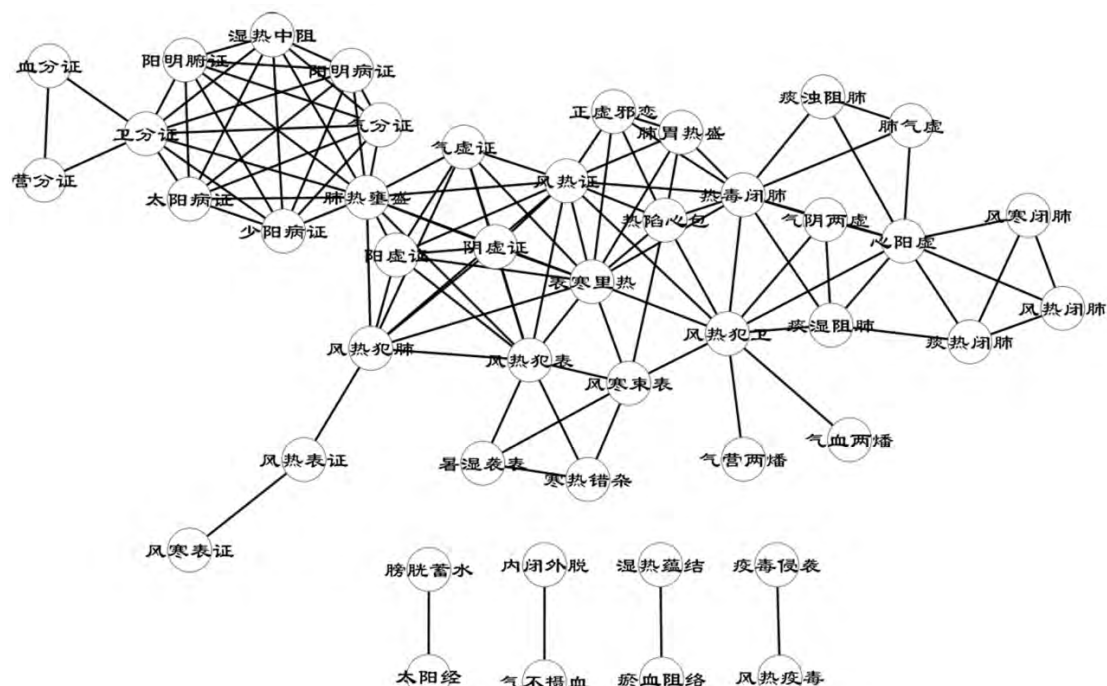


图 1 流行性感冒常见中医证候二维网络图

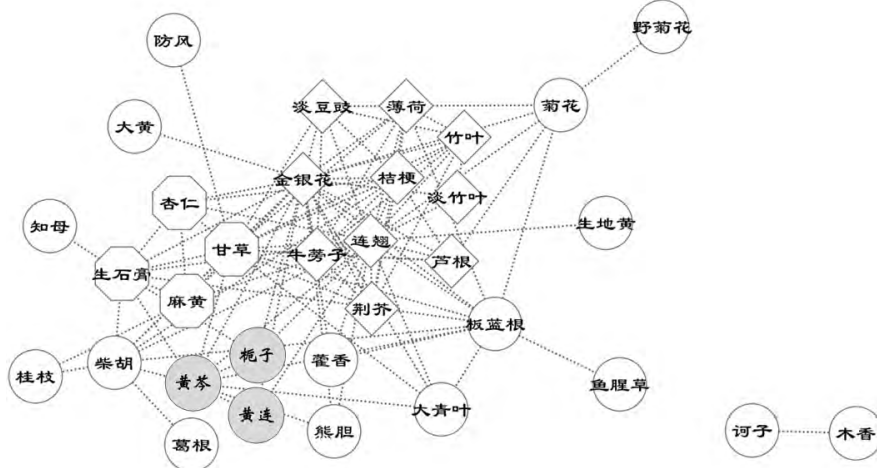


图2 流行性感冒常用中药组合网络图

表 4 流感诊疗方案常见中医证候频次 (频次 ≥ 1)

证候名称	频次
风热犯卫	4
热毒袭肺	4
气营两燔	4
热毒壅肺/毒热壅肺	4
疫毒犯肺, 肺失宣降	2
疫毒壅肺, 内闭外脱	2
毒壅气营	2
毒热闭肺	1
毒热内陷, 内闭外脱	1

表5 流感诊疗方案常用中药频次 (频次 ≥ 6)

中药名称	频次
生石膏	17
炙甘草/生甘草	16
杏仁	16
炙麻黄	12
黄芩	12
金银花	10
浙贝母	9
连翘	8
知母	8
大黄	7
青蒿	7
牛蒡子	6
桑叶	6
桔梗	6
竹叶	6
芦根	6

表5示,流感诊疗方案的常用中药,多为麻杏石甘汤、银翘散的主要药物,与流感的文本挖掘中高频的中药基本吻合。

4 文本挖掘结果与流感诊疗方案的比较分析

在应对甲型 H1N1 流感和人感染 H7N9 禽流感过程中, 中华人民共和国卫生部颁布的 7 版诊疗方案^[7-13]中都包含中医诊疗内容, 为全国中医药临床救治工作早期介入、及时开展提供了有效保障。

文本挖掘结果显示, 流感常见证候以热证、湿证居多, 病位多在表, 卫分、少阳及肺, 最常见的证候是风热犯卫, 与 2009 年颁布的 4 版甲型 H1N1 流感指南^[7-10]及 2011 年流感指南中常见证候^[11]一致。文本挖掘流感证候二维网络图显示, 流感的核心证候除风热证、风热犯表、风热犯卫、风热犯肺等证候之外, 还有肺热壅盛、表寒里热、热毒闭肺、热陷心包等证候, 这些证候不仅符合近年流感病位在肺、以热为主的证候特点, 也显示了流感发病急骤、传遍迅速易陷心包的特征, 这也是 2013 年 H7N9 两版方案^[12-13]中的常见证候。

流感常用方药对比分析显示, 文本挖掘流感治疗药物以清热解毒及清热燥湿药物为主。对药物组合统计结果显示, 这些药物正是临床常用治疗流感的银翘散、麻杏石甘汤以及黄连解毒汤, 这也是甲型 H1N1 流感及人感染 H7N9 禽流感治疗中常见的基础方。其中, 由银翘散和麻杏石甘汤化裁而得的金花清感方为应对甲型 H1N1 流感过程中形成, 研究证实该方对甲型 H1N1 流感具有治疗作用, 可作为奥司他韦的替代药物用于甲型 H1N1 流感的治疗^[14]。

综合比较文本挖掘与诊疗方案结果可以看出, 两者在流感常见证候与治疗方药方面基本一致, 即流感以风热犯卫证多见, 常用治疗方药以银翘散与麻杏石甘汤为主。但是文本挖掘结果中临床使用最多的方剂为银翘散、其次是麻杏石甘汤; 而诊疗方案中出现最多的为麻杏石甘汤、其次为银翘散。由此可见, 临床报道 (文本挖掘结果) 主要以流感轻症 (风热犯卫) 居多, 而诊疗方案中则不仅包含轻症, 还兼顾流感重症, 即热毒袭肺与气营两燔等证, 这些证候的主要治疗方药为麻杏石甘汤^[7-13]。因此, 两者在常用方剂的频次与顺序上存在差别。提示在制定流感重症诊疗方案时应单独针对流感重症制定检索策略进行检索与挖掘, 从而避免临床报道中轻症居多的偏倚。

5 文本挖掘技术在流感诊疗方案制定中的应用

尽管中医学对流感的认识历史悠久, 在与流感等传染病长期斗争的历史过程中积累了丰富的临床

经验, 但如何将已有的临床经验进行挖掘与转化, 从而在新型流感的应对过程中发挥重要的作用是中医药诊疗方案制定过程中需要解决的问题。文本挖掘技术的出现使快速整理大量文献成为可能, 不仅被运用到流感疾病规律总结及知识发现过程中, 也运用到流感的中医药诊疗方案制定中。

循证医学的发展及临床实践指南制定方法的完善为当今医疗实践提供了重要保障, 流感中医药方案的形成即是在循证医学原则指导下, 在结合中医学自身特色的基础上, 参照指南制定方法而形成的。考虑到中医药临床研究起步较晚, 可作为高级别证据的临床研究较少, 而相关诊疗经验存在于大量的文献中, 因此我们提出将文本挖掘技术运用到流感等方案的制定与修订过程中, 为方案的形成提供文献学的循证支持, 以弥补流感中医药临床研究证据的不足。2009 年甲型 H1N1 流感和 2013 年人感染 H7N9 禽流感方案的制定即借助了文本挖掘技术, 为新型流感等的快速、有效防控提供了支持。尽管文本挖掘技术在运用过程中也存在一定不足^[15], 但在新发突发传染病不断爆发的今天, 该技术不仅可以快速总结以往临床实践中的疾病规律与用药情况, 为临床积极防控提供借鉴和参考, 也使有效整合传染性疾病的循证证据提供了可能。

参考文献

- [1] 郭洪涛, 郑光, 张弛, 等. 利用数据挖掘技术探索类风湿关节炎与糖尿病“同证”的科学基础 [J]. 世界科学技术—中医药现代化, 2010, 12(5): 818-822.
- [2] 谭勇, 郭洪涛, 郑光, 等. 利用文本挖掘技术探索中药治疗疾病的用药规律 [J]. 世界科学技术—中医药现代化, 2010, 12(5): 823-827.
- [3] Zheng G, Jiang M, He X, et al. Discrete derivative: a data slicing algorithm for exploration of sharing biological networks between rheumatoid arthritis and coronary heart disease [J]. BioData Min, 2011, 4:18.
- [4] Guo HT, Zheng G, Li ZX, et al. Searching Rules of Acupuncture on Periarthritis of Shoulder by Data Slicing Algorithm [C/OL]. (2011-08-18). http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6006348&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D6006348
- [5] 李立, 周琦, 郑光, 等. 基于文本挖掘技术分析中成药、西药对慢性胃炎的治疗规律 [J]. 中国实验方剂学杂志, 2011, 17(24): 228-231.
- [6] 梁非, 展俊平, 李立, 等. 基于文本挖掘方法探索寒性热性中药的病证方药相应规律 [J]. 中国实验方剂学杂志

- 志, 2013, 19(15): 333-337.
- 〔7〕 钟南山, 李兰娟, 王辰, 晁恩祥等代表卫生部组织编写. 甲型 H1N1 流感诊疗方案(2009 年第一版) 〔J〕. 中华医学杂志, 2009, 89(22): 1519-1520.
- 〔8〕 钟南山, 李兰娟, 王辰, 晁恩祥等代表卫生部甲型 H1N1 流感临床专家组. 甲型 H1N1 流感诊疗方案(2009 年第 2 版) 〔J〕. 中华医学杂志, 2009, 89(36): 2526-2528.
- 〔9〕 钟南山, 李兰娟, 王辰, 晁恩祥等代表卫生部甲型 H1N1 流感临床专家组. 甲型 H1N1 流感诊疗方案(2009 年第 3 版) 〔J〕. 中华医学杂志, 2009, 89(45): 3173-3175.
- 〔10〕 中华人民共和国卫生部. 甲型 H1N1 流感诊疗方案(2010 年版) 〔J〕. 国际呼吸杂志, 2011, 31(2): 81-84.
- 〔11〕 中华人民共和国国家卫生和计划生育委员会. 流行性感冒诊断与治疗指南(2011 年版) 〔EB/OL〕. (2011-02-17). <http://www.moh.gov.cn/zhuzhan/wsbmgz/201304/7fd812ef8dc54291af5060957eea304c.shtml>
- 〔12〕 中华人民共和国国家卫生和计划生育委员会. 人感染 H7N9 禽流感诊疗方案(2013 年第 1 版) 〔EB/OL〕. (2013-04-02). <http://www.moh.gov.cn/zhuzhan/wsbmgz/201304/839aecf249f94b5e90f03a3c37471dd9.shtml>
- 〔13〕 中华人民共和国国家卫生和计划生育委员会. 人感染 H7N9 禽流感诊疗方案(2013 年第 2 版) 〔EB/OL〕. (2013-04-10). <http://www.moh.gov.cn/yjb/bmdt/201304/3d890ae7adfe4566b82fe675e6711ed7.shtml>
- 〔14〕 Wang C, Cao B, Liu QQ, et al. Oseltamivir compared with the Chinese traditional therapy maxingshigan - yinqiaosan in the treatment of H1N1 influenza: a randomized trial 〔J〕. Ann Intern Med, 2011, 155(4): 217-225.
- 〔15〕 郭洪涛, 郑光, 赵静, 等. 基于文本挖掘分析甲型 H1N1 流感的中医药治疗特色 〔J〕. 世界科学技术—中医药现代化, 2011, 13(5): 772-776.

Comparative Analysis of Text Mining Results and TCM Treatment Programs for Influenza

ZHAO Jing¹, GUO Hongtao², HAN Jingdan¹, LI Li¹, KOU Shuang¹, GAO Yuan¹, JIANG Fan¹, LÜ Cheng¹, WANG Yanping¹, LÜ Aiping¹

(1. China Academy of Chinese Medical Sciences, Beijing 100700; 2. The First Affiliated Hospital of Henan University of TCM)

ABSTRACT Objective To compare and analyze the text mining results and TCM treatment programs for influenza, and to literature support for prevention and control programme of influenza. **Methods** The articles related to influenza in Chinese Biomedical Literature Database were searched and collected. The TCM patterns and formulas were obtained by text mining and compared with the TCM patterns and formulas in seven different versions of influenza treatment programs issued by National Health and Family Planning Commission of the People's Republic of China.

Results The commonly seen TCM patterns were wind-heat invading defense pattern, damp-heat pattern and lesser yang disease. The commonly used formulas were *Yinqiao Powder* (银翘散), *Maxing Shigan Decoction* (麻杏石甘汤) and *Huanglian Jiedu Decoction* (黄连解毒汤). The commonly used herbs were Banlangen (板蓝根, *Radix Isatidis*), Jinyinhua (金银花, *Flos Lonicerae Japonicae*), Huanglian (黄连, *Rhizoma Coptidis*), Huangqin (黄芪, *Radix Scutellariae*) and Lianqiao (连翘, *Fructus Forsythiae*). Jinyinhua was commonly combined with Lianqiao, Huangqin and Banlangen. In the influenza treatment programs the commonly seen TCM patterns were wind-heat invading defense pattern, heat toxin assailing the lung pattern and dual blaze of qi and construction. The commonly used formulas were *Maxing Shigan Decoction* (麻杏石甘汤) and *Yinqiao Powder* (银翘散). The commonly used herbs were Shengshigao (生石膏, *Gypsum Fibrosum*), Zhigancao (炙甘草, *Radix et Rhizoma Glycyrrhizae Praeparata cum Melle*) / Shenggancao (生甘草, *Radix et Rhizoma Glycyrrhizae*), Zhimahuang (炙麻黄, *Ephedrae Herba cum Liquido Fricta*), Xingren (杏仁, *Armeniaca Semen*) and Huangqin. **Conclusion** The commonly seen TCM pattern of influenza is wind-heat invading defense pattern. The commonly used formulas are *Maxing Shigan Decoction* (麻杏石甘汤) and *Yinqiao Powder* (银翘散). The patterns and formulas in the influenza treatment programs are basically the same as the text mining results. However, the text mining results are mainly mild influenza and both mild and severe influenza can be seen in the treatment programs.

Keywords influenza; syndrome; Chinese medicinals; treatment program; text mining

(收稿日期: 2013-12-30; 修回日期: 2014-01-27)

编辑: 焦爽