

利用文本挖掘技术探索复方丹参注射液应用规律

田景平^{1,2}, 郑光^{2,3}, 郭洪涛⁴, 蔡峰², 吕诚², 姜森², 张弛², 吕爱平^{2*}

(1. 广州中医药大学研究生院, 广州 510006; 2. 中国中医科学院中医临床基础医学研究所, 北京 100700;
3. 兰州大学信息学院, 兰州 730000; 4. 河南中医学院第一附属医院, 郑州 450000)

[摘要] 目的: 利用文本挖掘技术探索复方丹参注射液临床应用规律。方法: 在中国生物医学文献服务系统中收集复方丹参注射液文献数据, 采用基于敏感关键词频数统计的数据分层算法, 挖掘复方丹参注射液相关中医证候、疾病、症状、中医治法、针灸穴位、西药等规律。这些规律通过一维频次表及二维网络图进行展示。结果: 中医证候方面, 以瘀、虚因素为主; 脑梗死、心绞痛、糖尿病并发症、心力衰竭、肝硬化、病毒性心肌炎等疾病为其主要治疗疾病; 胸闷、头晕、头痛、溃疡、腹痛为其主要治疗症状; 中医治法以活血化瘀为主; 相关针灸穴位主要有球后、足三里、风池、肾俞等; 常与低分子右旋糖酐、胞二磷胆碱、阿司匹林、维生素 B 簇等西药联用。结论: 文本挖掘能够比较客观的总结中西医用药规律, 为临床应用提供有益的探索和参考。

[关键词] 文本挖掘; 数据分层算法; 复方丹参注射液; 应用规律

[中图分类号] R287.6 **[文献标识码]** A **[文章编号]** 1005-9903(2013)04-0330-05

Application Regularity of Compound Salvia Miltiorrhiza Injection with Text Mining Technique

TIAN Jing-ping^{1,2}, ZHENG Guang^{2,3}, GUO Hong-tao⁴, CAI Feng²,
LV Cheng², JIANG Miao², ZHANG Chi², LV Ai-ping^{2*}

(1. Guangzhou University of Traditional Chinese Medicine (TCM), Guangzhou 510006, China; 2. Institute of Basic Research In Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China;
3. School of Information, Lanzhou University, Lanzhou 730000, China;
4. The First Affiliated Hospital of Henan University of TCM, Zhengzhou 450000, China)

[收稿日期] 20120705(013)

[基金项目] 国家自然科学基金杰出青年项目(30825047); 国家自然科学基金面上项目(30973975, 81072982); 青年基金项目(30902003); 中国博士后基金面上项目(20110940553); 中国中医科学院自主选题项目(Z0172)

[通讯作者] * 吕爱平, 博士生导师, 研究员, 从事临床评价研究, Tel: 010-64014411-3301, E-mail: lap64067611@126.com

- [10] 杨霓芝, 徐大基, 刘旭生, 等. 尿毒康治疗脾肾气(阳)虚兼瘀浊内阻慢性肾功能衰竭的临床研究[J]. 中医杂志, 1999, 40(3): 161.
- [11] 张小鹿, 陶明龙, 黄芳. 益肾泄浊化瘀汤治疗慢性肾衰竭临床观察[J]. 上海中医药杂志, 2009, 43(8): 37.
- [12] 董小革. 健脾益肾活血降浊法治疗慢性肾衰竭 30 例临床观察[J]. 河北中医, 2010, 32(8): 1134.
- [13] 章关根. 内服抗衰汤配合药浴治疗慢性肾功能衰竭 50 例[J]. 中国中医药科技, 2008, 15(3): 229.
- [14] 张秋林, 洪钦国, 汤水福. 尿毒清胶囊治疗慢性肾功能衰竭疗效观察[J]. 湖北中医杂志, 2007, 29(1): 8.
- [15] 王小娟, 郭建生, 雷颖, 等. 肾衰颗粒对慢性肾功能不全患者血液流变学的影响及疗效观察[J]. 中国中医药信息杂志, 2007, 14(2): 12.
- [16] 王小娟, 雷颖, 郭建生, 等. 肾衰颗粒对慢性肾衰竭患者血浆内皮素、一氧化氮的影响[J]. 中国实验方剂学杂志, 2007, 13(1): 64.
- [17] 席淑清, 路现平. 益肾安合剂治疗慢性肾衰竭的临床研究[J]. 河北中医, 2008, 30(10): 1023.
- [18] 刘家生. 益肾调中泄浊治疗慢性肾功能衰竭 43 例临床观察[J]. 安徽中医临床杂志, 2001, 13(2): 99.
- [19] 周琴, 常杰, 戴海萍. 中西医结合治疗慢性肾功能衰竭 60 例[J]. 中国中医药信息杂志, 2009, 16(12): 57.

[责任编辑 邹晓翠]

[Abstract] Objective: To explore the application regularity of compound *Salvia Miltiorrhiza* (SM) injection. **Method:** The data set on compound SM injection was downloaded from Chinese BioMedical literature database. The regularities of syndrome, diseases, symptoms, traditional Chinese medicine (TCM) treatment method, acupuncture point and western medicines on compound SM injection were mined out by data slicing algorithm. The results are demonstrated with tables and networks. **Result:** TCM syndrome of compound SM injection was mainly correlated with blood stasis and deficiency. The diseases such as cerebral infarction, angina pectoris, diabetic complications, heart failure, liver cirrhosis and viral myocarditis were often treated with compound SM injection. The symptoms such as chest distress, dizziness, headache, ulcer and abdominal pain were frequently reported. Activating blood and resolving stasis were main Chinese medicinal therapy of compound SM injection. Acupuncture points such as retrobulbar, Zusanli, Fengchi and Shenshu were commonly related with compound SM injection. For western medicines, it was usually used together with low molecular dextran, citicoline, aspirin and vitamin B group. **Conclusion:** Text mining approach provides a novel method in the summary of application regularity in both TCM and western medicine.

[Key words] text mining technique; data slicing algorithm; compound *Salvia Miltiorrhiza* injection; application regularity

复方丹参注射液是由丹参、降香两种中药制备而成的复方制剂,每毫升相当于丹参、降香各 1 g。丹参为唇形科鼠尾属植物丹参的根,其性味苦、微寒,归心、心包、肝经,具有活血祛瘀、调经止痛、除烦安神、凉血消痈之功效^[1];降香为豆科黄檀属植物降香檀、印度黄檀的树干或根部心材,其性味辛、温,归肝、脾、心经,具有活血散瘀、止血定痛、降气、辟秽之功效^[1]。现代药理研究证实^[1]丹参、降香具有降低血液黏度、降低血脂、抑制血小板聚集、改善微循环等作用。近些年来,我国对复方丹参注射液的研究显示:复方丹参注射液对脂多糖诱导的兔弥漫性血管内凝血有良好的拮抗作用^[2];复方丹参注射液能够改善挫伤后兔视网膜的微循环、提高局部的营养代谢,延缓和抑制细胞凋亡的发生^[3];复方丹参注射液也能明显提高大鼠血清中超氧化物歧化酶(SOD)活性,降低肌酸激酶(CK)、乳酸盐脱氢酶(LDH)、丙二醛(MDA)含量,从而保护心肌缺血性损伤^[4]。基于丹参、降香的广泛药理作用,复方丹参注射液在临床上得以广泛应用,研究其临床应用规律有着重要意义。

复方丹参注射液研究所报道的相关文献大量存在于现有数据库中。本文借助课题组不断成熟的文本挖掘技术^[5-6],结合原文献回溯,人工阅读分析等方法,对现有复方丹参注射液的中文文献进行挖掘,分析复方丹参注射液的临床应用规律。

1 材料与方法

1.1 文本数据收集 在中国生物医学文献数据库(Chinese BioMedical Literature Database, CBM,

<http://sinomed.cintcm.ac.cn/index.jsp>)中,“精确匹配”状态下以“复方丹参注射液”为检索词进行检索,共得到文献 4 534 篇(检索日期:2012 年 6 月 19 日),下载后作为文本数据挖掘的原始数据。

1.2 文本数据处理 将收集来的数据,按照下载的先顺序,整合到一个平面文件(后缀 txt)里面,以 ANSI 编码格式保存。然后,利用专有的文本提取工具(软件著作权,软著登字第 0261882 号,登记号 2010SR073409),对下载的非结构化的 txt 文本数据进行信息提取,保存成格式化的、便于大型关系型数据库(Microsoft SQL Server,以下简称 SQL)处理的格式,然后导入 SQL 中进行下一步的挖掘分析。假设每一篇文章的贡献度是相同的,一篇文章中重复出现的关键词,只需要计算一次,据此构建算法进行数据清洗工作^[5]。清洗完毕后的数据,既可以提取挖掘对象的一维频次,也可以得到挖掘对象的二维关系,进行可视化呈现。抽出不同频次的关键词对,用 Cytoscape 2.8 软件进行可视化处理,形成可视化的网络图,然后结合专业知识进行解析,一旦发现不合理的结果,即回溯原文献数据集,如果是噪音,仍按算法进行噪音清洗,直至噪音降到满意为止。最后的结果可视化成图,结合专业知识进行解析。

2 复方丹参注射液文本挖掘结果的评价和分析

2.1 中医证候文本挖掘结果 文本挖掘共提取到中医证候 40 项,按文献频次高低选择前 6 个列出(括号内为文献篇数,以下同):气虚血瘀(18)、气滞血瘀(12)、血瘀(11)、肾阴虚(6)、痰湿阴肺(5)、气阴两虚(5)。其中以气虚血瘀文献频次最高。

共提取到 77 组证候组合,把提取到的证候组合构建网络图(图 1)。图 1 为文献频次 ≥ 2 的相关证候的网络分布图,网络中证候的显示度和结点的大小正相关;连线表示相关性。

图中可见以气虚血瘀证为主,涉及到气滞血瘀、阴虚阳亢等瘀、虚证候。综合图、表,复方丹参注射液的证候以瘀、虚因素为主。

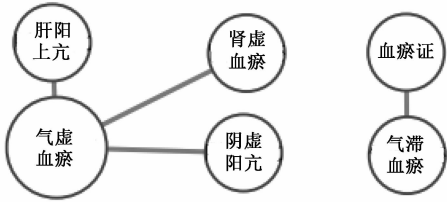


图 1 复方丹参注射液相关证候网络(频次 ≥ 2)

2.2 疾病文本挖掘结果 疾病共提取到 76 项,按疾病在文献中出现的频次高低选取前 9 项列出:脑梗死(870)、心绞痛(286)、糖尿病(212)、新生儿缺氧缺血性脑病(180),心力衰竭(120)、肝硬化(84)、心肌炎(59)、急性胰腺炎(46)、过敏性紫癜(43)。其中“糖尿病”多为“糖尿病周围神经病变”、“糖尿病肾病”、“糖尿病足”等糖尿病并发症;心肌炎多为“病毒性心肌炎”。挖掘到的疾病以心血管疾病居多,其他疾病如肝硬化、急性胰腺炎、过敏性紫癜等也为复方丹参注射液所常见治疗疾病。回溯原文献数据集发现复方丹参注射液多为治疗过敏性紫癜或过敏性紫癜肾炎^[1]的报道,但也有致过敏性紫癜的报道^[8]。

共提取到疾病组合 56 组,把提取到的疾病组合结果构建网络图(图 2)。图 2 为相关疾病文献频次 ≥ 6 的网络分布图,网络中疾病的显示度和结点的大小正相关;连线表示相关性。

脑梗死,心绞痛、糖尿病等疾病一维频次较高,但很少参与二维频次的构建。网络图中以心力衰

竭、心肌梗死等心血管疾病为主。综合图、表,复方丹参注射液在脑梗死心绞痛、糖尿病并发症、肝硬化等疾病方面应用较多,其中以心血管疾病为多。

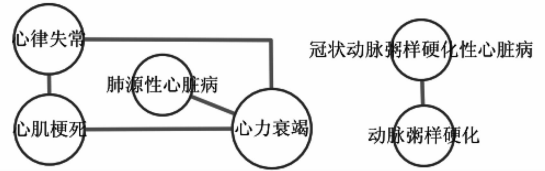


图 2 复方丹参注射液相关疾病网络(频次 ≥ 6)

2.3 症状文本挖掘结果 症状共提取到 100 项,按文献频次高低选取前 10 项作表(表 1)。回溯原文献数据集发现:“溃疡”多为“消化性溃疡”、“皮肤溃疡”等;“功能障碍”多为“关节功能障碍”、“神经功能障碍”、“胃肠功能障碍”等。由表 1 可见溃疡、头痛、功能障碍、耳聋等为其常见治疗症状。

表 1 复方丹参注射液症状文献频次比较

No.	症状	文献频次	No.	症状	文献频次
1	溃疡	39	6	咳嗽	27
2	头痛	36	7	发热	24
3	功能障碍	28	8	腹痛	23
4	耳聋	28	9	胸闷	21
5	变形	27	10	头晕	20

共提取到症状组合 381 组,把提取到的症状组合结果构建网络图(图 3)。图 3 为相关症状频次 ≥ 3 的网络分布图,网络中症状的显示度和结点的大小正相关;连线表示相关性。

回溯原文献数据集发现:恶心、呕吐多为复方丹参注射液的不良反应;图中以胸闷、头晕、腹痛等为核心,涉及到心悸、麻木、头痛、关节痛等心脑血管、消化、运动等系统疾病症状。综上,复方丹参注射液治疗症状以胸闷、头晕、头痛、溃疡、腹痛为主,涉及到心、脑血管等系统疾病症状。

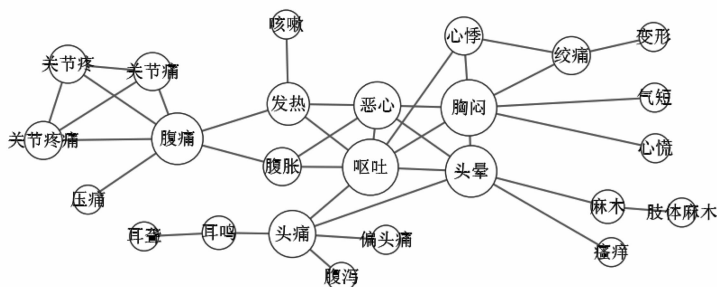


图 3 复方丹参注射液相关症状网络(频次 ≥ 3)

2.4 中医治法文本挖掘结果 中医治法共提取到

17 项,按文献频次高低选择前 6 个列出:活血化瘀

参注射液文献中相关西药文献频次 ≥ 8 的网络分布图,网络中西药的显示度和结点的大小正相关;连线表示相关性。

图 6 可见,以低分子右旋糖酐、阿司匹林等为核心,分别与溶栓、抗凝、改善微循环等药物联用;中间小图可见以甲钴胺为核心,与维生素 B 簇等营养神经药物联用。综合表、图可见低分子右旋糖酐、胞二磷胆碱、阿司匹林、维生素 B 簇等西药为复方丹参注射液较常联用西药,这些药物以主治心、脑血管缺血性病变为主。

3 讨论

文本挖掘技术是以计算语言学、统计数理分析为理论基础,服务于生物、医药、文献研究等学科的新兴的交叉学科^[9]。应用于中医药领域,文本挖掘能从海量的中医药文献中发现知识以促进中医临床研究和中药复方研发等多个方面。根据中医理论和专业知识,利用数据挖掘技术对中医药文献库与生物医学信息进行处理,为中西医结合研究提供新的思路和途径,并且结果更加客观,可重复性强^[10]。

本研究对 4 534 篇“复方丹参注射液”文献进行文本挖掘,得到复方丹参注射液相关中医证候、疾病、症状、治法、针灸、西药等挖掘结果,是对复方丹参注射液临床应用规律的较全面总结。中医证候方面,以瘀、虚因素为主;主治疾病以脑梗死、心绞痛等心、脑血管疾病为主;症状方面,以胸闷、头晕等为主,涉及到心、脑血管等系统疾病症状;中医治法以活血化瘀为主;针灸穴位方面,主要有球后、足三里、风池、肾俞等,针灸穴位配伍选用以主治脑梗死等缺血性病变为主;西药方面主要与低分子右旋糖酐等主治心、脑血管缺血性疾病为主的西药联合应用。由此可见,复方丹参注射液的应用研究比较集中于心、脑血管疾病,基于其广泛的药理作用及良好的临床效果,可拓展其在消化、呼吸系统等方面疾病的应用研究。

文本挖掘的结果,通过回溯原文献,人工阅读评价,使结果更为客观,如图 3,恶心、呕吐多为复方丹参注射液的不良反应;选取高频次部分参与网络关

系图的构建,可以集中反映临床研究情况,如图 5,选取穴位相关频次 ≥ 5 作网络图,结果分为两个小网络图,这两个小网络图分别反应了临床配伍应用主治不同的疾病,如中间小图以治疗糖尿病性神经病变为主。因此,通过这种文本挖掘组合,能够比较客观的总结中西医用药规律,为临床应用提供有益的探索和参考。

[参考文献]

- [1] 南京中医药大学. 中药大辞典[M]. 上海:上海科学技术出版社,2006:643,2077.
- [2] 林熙,祈洁珍,邱鹏新,等. 复方丹参注射液抗脂多糖诱导的兔弥漫性血管内凝血[J]. 中国病理生理杂志,2011,27(3):464.
- [3] 王志玉,史爱云,刘轶. 复方丹参注射液治疗兔视网膜挫伤的实验研究[J]. 国际眼科杂志,2011,11(12):2079.
- [4] 王秀丽,李耿,赵保胜,等. 复方丹参冻干粉针的制备及其对急性心肌缺血大鼠的保护作用[J]. 中国实验方剂学杂志,2010,16(7):171.
- [5] Zheng G, Jiang M, He X J, et al. Discrete derivative: a data slicing algorithm for exploration of sharing biological networks between rheumatoid arthritis and coronary heart disease[J]. Bio Data Min,2011,4:18.
- [6] 谭勇,杨静,赵宁,等. 利用文本挖掘技术探索中西药治疗慢性乙型肝炎的用药规律[J]. 中国实验方剂学杂志,2011,17(24):311.
- [7] 陈昌陆. 低分子肝素联合复方丹参注射液治疗儿童过敏性紫癜性肾炎 33 例疗效观察[J]. 医学信息,2011,24(9):4886.
- [8] 张学文. 静脉滴注复方丹参注射液致过敏性紫癜[J]. 药物不良反应杂志,2011,13(2):130.
- [9] 薛为民,陆玉昌. 文本挖掘技术研究[J]. 北京联合大学学报:自然科学版,2005,19(4):59.
- [10] Li S, Zhang Z Q, Wu L J, et al. Understanding ZHENG in traditional Chinese medicine in the context of neuro-endocrine-immune network[J]. IET Syst Biol, 2007, 1(1):51.

[责任编辑 邹晓翠]