

哮喘处方相似度研究*

李园白¹ 崔蒙¹ 杨阳^{1#} 朱晓博¹ 于琦¹ 田野¹ 奚怀平¹ 祖娜² 王双玲³

(1 中国中医科学院中医药信息研究所 北京 100700; 2 北京中医药大学; 3 通县中医医院)

摘要: 目的 通过信息学的相似度比较方法,计算哮喘处方之间相似度,筛选特殊的哮喘处方。方法 通过三层比较法分析处方之间的相似程度,第一个层次是利用 Jaccard 集合相似度比较法分析处方中药组成相似程度,第二个层次是利用夹角余弦相似比较法分析处方中药组成剂量的相似程度,第三个层次是利用夹角余弦相似比较法分析处方功效特征相似度。结果 筛选出部分特殊哮喘处方。其中处方药物组成特殊的哮喘处方 557 个,组成剂量特殊的哮喘处方 16 个,特殊功效的哮喘处方 0 个。结论 大多数处方都是在前人基础上形成,独创性的处方比较少。

关键词: 方剂; 相似度; 集合; 夹角余弦; 哮喘

中图分类号: R2-03

Similarity of formulas for treating asthma*

LI Yuan-bai¹, CUI Meng¹, YANG yang¹, ZHU Xiao-bo¹, YU Qi¹, TIAN Ye¹, XI Huai-ping¹, ZU Na², WANG Shuang-ling³

(1 Institute of Information on Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700; 2 Beijing University of Chinese Medicine; 3 Tongxian Hospital of Chinese Medicine)

Abstract: Object Through similarity comparison methods of informatics, to calculate the similarity among these formulas and screen special formulas for treating asthma. **Methods** The similarity among the formulas for treating asthma was analyzed by using three-layer comparison method. The first layer was to analyze the similarity of Chinese medicinal composition in these formulas by using Jaccard method, the second layer was to analyze the similarity of dose of Chinese medicinals in these formulas by using angle cosine method, and the third layer was to analyze the similarity of effects of these formulas by using angle cosine method. **Results** There were some special formulas screened, and among them 557 with special medicinal composition, 16 with special dose and no one with special effect. **Conclusion** The most of formulas for treating asthma are designed on the base of previous foundation and there are a few formulas with originality.

Key words: formulas; similarity; set; angle cosine; asthma

医案在中医领域的地位举足轻重,读者阅读医案可以学习和借鉴前人的中医思想。通过分析以往关于医案的按语类资料,发现读者阅读医案主要有 2 个目的,一方面是学习其他医师的诊疗思想,另一方面是寻找特殊的治疗方法,开拓诊疗思路。故从

寻找特殊治法的阅读目的出发,以哮喘处方为例,探索获得特殊哮喘处方的方法。

1 资料

以公开发表于期刊中的哮喘医案处方作为评价对象,评价这类处方的特殊性,把特殊的处方筛选出

李园白,女,博士,副研究员

通信作者: 杨阳,女,博士,助理研究员,主要研究方向: 中医药信息学, E-mail: sanyangkt@126.com

* 中国中医科学院自主选题研究项目(No. ZZ0402055, ZZ040308)。

来。当评价“特殊”处方时,需要有“普通”特性的处方作为参照物,衡量评价对象的特殊程度,故选用了一些经典常见的方剂作为参照对象。

1.1 评价对象

选择期刊中公开发表、治疗成人哮喘有效的医案中的中草药口服处方作为评价对象。医案入选标准:诊断为哮喘且无其他合并症;治疗方法是中草药治疗;药物剂型为汤剂;给药方式为口服;处方有中草药名称、剂量;最后治疗效果为有效或痊愈。

全部处方来源于1347篇期刊文献,涉及期刊126种、1212册,包含2779个处方信息。期刊文献发表年限为20年(1988年—2007年)。所有医案检索来源为权威的期刊数据库^[1-3]

1.2 参照对象

选择流传较久、公认度较高、研究病例数较大的哮喘处方信息作为参照对象。其中流传较久的处方指古籍哮喘处方(涉及古籍240种)^[2],公认度较高的处方指现代诊疗理论书籍中的哮喘处方(涉及书籍44种)^[4],研究病例数较大处方指发表于期刊中的10例以上临床研究文献中的哮喘处方(涉及文献1293篇)。3部分处方共有2013个。

2 方法

由浅入深从3个层次分析处方之间的相似程度。

2.1 中药组成相似度比较

方剂相似度的比较仅比较中药组成的不同。选择Jaccard相似系数计算处方组成的相似度^[5]。

Jaccard相似系数计算方法如下:

$$\text{Sim}(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \quad \text{公式(1)}$$

例如:SIM(A方, B方) = SIM(半夏; 橘红; 白茯苓; 甘草) / (半夏; 天南星; 橘红; 枳实) = 2/6 = 0.33

2.2 中药组成剂量相似度比较

对中药组成和剂量组成的复杂信息进行比较,采用夹角余弦相似法^[6]。但因为很多中药的用量单位不同,有的是“克”、有的是“枚”,而且大多数中药的常用剂量范围不同,需要进行标准化处理,消除单位或剂量波动幅度不同造成的影响。

首先进行剂量的规范化,根据中药数据特点,自行设计了剂量规范化公式,如下:

$$X_i = \frac{X_i}{X_{\min} + X_{\max}} \quad \text{公式(2)}$$

例:大黄常用剂量(3~12g),实际应用剂量为10g,那么标准化剂量为0.67,所有剂量被规范化后,利用夹角余弦相似度计算组成剂量的相似度:

$$\text{Sim}(X, Y) = \frac{\sum XY}{\sqrt{\sum X^2 \times \sum Y^2}} \quad \text{公式(3)}$$

例如:

A方标准化剂量:大黄为0.60,厚朴为0.40。

B方标准化剂量:大黄为0.50,枳实为0.80。

$$\begin{aligned} \text{Sim}(A方, B方) &= \text{Sim}((\text{大黄}; \text{厚朴}) (\text{大黄}; \text{枳实})) \\ &= \frac{(0.6 \times 0.5 + 0.4 \times 0.8)}{\sqrt{(0.6^2 + 0.5^2) \times (0.4^2 + 0.8^2)}} \\ &= 0.67 \end{aligned}$$

2.3 处方功效特征相似度比较

对方剂之间的功效复杂信息进行比较,同样采用夹角余弦相似法。未将原处方中注释的功效做为比较依据,而是通过单味中药的功效集合形成整个处方功效集合,然后对功效集合之间进行相似程度比较。例如:某方功效处理结果:滋阴0.29;补血0.79;益精0.29;填髓0.29;活血2.17;调经0.93;止痛1.93;润肠1.17;通便1.17;平肝0.43;养血0.43

预处理后的处方功效集合即可利用夹角余弦相似法进行相似度比较,计算方法同组成剂量的相似度比较。

3 结果

共有2779个哮喘处方方剂信息被评价,每个处方均与2013个参照方剂相比较,每个评价处方均与参照方剂进行3种相似度比较(处方组成相似度、处方组成剂量相似、处方功效相似度),每个评价处方在1种相似度比较中均获得2013个相似度,选择其中最大值作为最终相似度。2779个处方中,每个处方均有3种相似度值。每种相似度中的2779个值均进行排序,即可获得同种类别相似度值低的哮喘处方,这些处方非常特殊,因为它们和2013个参照方剂比较后,均不相似。2779个哮喘处方被评价后的结果分布及最特殊的结果(前5名)见表1~表4。

表 1 处方相似度比较

Table 1 Results of comparison in similarity of formulas

相似度 Similarity	中药组成相似度 Similarity of compositions	中药组成剂量相似度 Similarity of doses	处方功效特征相似度 Similarity of effects
0 ~ 0.33	557	16	0
0.34 ~ 0.66	1 841	1 328	84
0.67 ~ 1.00	381	1 435	2 695

表 2 组成相似度最低处方(前 5 名)

Table 2 First five lowest similarity of formula compositions

项目 Item	方剂组成 Composition of formulas	来源 Origin	相似度 Similarity
自拟方	血竭 0.5 g	刘宇. 血竭 - 伤科良药用途多 [J]. 新中医, 2004, 36 (8): 53 - 54.	0.076 9
自拟方	猪肾 3 g	龚士澄. 豚颠断哮 [J]. 北京中医, 1988, (4): 1 - 10.	0.083 3
截喘汤加味	谷芽 12 g, 合欢皮 15 g, 开金锁 15 g	王佩芳, 姜光华. 病证相结合专方治专病 - 姜春华临证治病的独特经验 [J]. 上海中医药杂志, 1993, (3): 4 - 7.	0.100 0
独圣散	瓜蒂 3 g	瞿桂凤. 独圣散临床应用举隅 [J]. 四川中医, 1998, 16(4): 56.	0.111 1
过敏煎加味	过敏煎加味, 茜草 10 g, 丹参 15 g, 黛蛤散 3 g, 旋覆花 3 g, 牵牛子 3 g, 五灵脂 3 g, 香附 3 g	董振华, 秀元, 范爱平. 祝谌予治疗哮喘的经验 [J]. 浙江中医杂志, 1994, 29(1): 19 - 20.	0.142 9

表 3 组成剂量相似度最低处方(前 5 名)

Table 3 First five lowest similarity of formula doses

项目 Item	方剂组成 Composition of formulas	来源 Origin	相似度 Similarity
自拟方	萝卜汁 300 mL, 何首乌 3 g, 苦杏仁 3 g, 麻黄 3 g, 半夏 3 g, 槟榔 3 g, 枳壳 3 g, 佛手 3 g, 厚朴 3 g	陈金利. 哮喘急性发作时的治法遣药 [J]. 吉林中医药, 2000, 20(2): 1 - 2.	0.02
自拟方	鸡蛋 1.5 个, 清香油 500 g, 大黄 60 g	侯激流. 麻杏石甘汤的临床应用举隅 [J]. 中医药学刊, 2004, 22(10): 1945.	0.07
截哮蛋	鸡蛋 3 个	洪广祥. 哮证治疗之我见 [J]. 中医杂志, 1988, 29(3): 7 - 9.	0.07
自拟方	豚颠 3 个	龚士澄. 豚颠断哮 [J]. 北京中医, 1988, (4): 1 - 10.	0.08
自拟方	大枣 10 g, 黄芪 15 g, 素肉 74.4 g	王红影. 六君子汤治疗慢性哮喘证的临床运用 [J]. 湖南中医杂志, 1993, 9(4): 25 - 26.	0.12

表 4 功效相似度最低处方(前 5 名)

Table 4 First five lowest similarity of formula effects

项目 Item	方剂组成 Composition of formulas	来源 Origin	相似度 Similarity
自拟方	赭石 15 g	向楠, 裴迅. 重镇方药的临床应用 [J]. 湖北中医杂志, 2006, 28(1): 45 - 46.	0.48
截喘汤加味	谷芽 12 g, 合欢皮 15 g, 开金锁 15 g	王佩芳, 姜光华. 病证相结合专方治专病 - 姜春华临证治病的独特经验 [J]. 上海中医药杂志, 1993, (3): 4 - 7.	0.49
补肺汤	金匱肾气丸加味, 蛤蚧 3 g, 茜草 30 g, 地龙 15 g, 补骨脂 15 g, 仙鹤草 30 g	段海忠, 吴贵权, 唐明芝. 顽固性哮喘验案三则 [J]. 云南中医中药杂志, 1996, 17(4): 27 - 28.	0.49
自拟方	牡蛎 15 g, 龙骨 15 g, 皂矾 10 g, 鸡内金 100 g, 大黄 20 g, 西洋参 20 g	韩运才, 王莺歌. 哮喘病内外结合治验体会 [J]. 陕西中医, 2005, 26(8): 808 - 809.	0.49
补肺汤	金匱肾气丸加味, 地骨皮 20 g, 茜草 30 g, 地龙 15 g, 补骨脂 15 g, 仙鹤草 30 g	段海忠, 吴贵权, 唐明芝. 顽固性哮喘验案三则 [J]. 云南中医中药杂志, 1996, 17(4): 27 - 28.	0.50

病理特点,探讨建立评价女性绝经期尿路感染的动物模型。

实验结果表明,去势组小鼠在去势后5 d内全部17只均先后出现并维持间情期表现,去势小鼠阴道脱落细胞角化指数明显低于假手术组,术后第17天血清 E_2 水平明显低于假手术组($P < 0.01$),说明去势组小鼠符合女性绝经期阴道变化的特点。去势小鼠受细菌感染后的尿培养结果和膀胱病理切片结果均表明,去势感染组小鼠全部膀胱感染,同时,假手术感染组小鼠膀胱病理切片也存在黏膜血管扩张充血、炎细胞浸润等表现。实验结果初步表明,绝经期小鼠可以出现合并泌尿系感染的病理表现,但对于这一病理变化尚需进一步研究。

参考文献:

- [1] 张绍芬,主编. 绝经——内分泌与临床[M]. 北京: 人民卫生出版社, 2005.
- [2] 徐叔云. 药理实验方法学[M]. 2版. 北京: 人民卫生出版社, 1999: 1293.
- [3] 孙敬方,主编. 动物实验方法学[M]. 北京: 人民卫生出版社, 2001: 329.

(上接第590页)

4 讨论

从相似度比较结果分析,大多数处方都是在前人基础上形成,独创性的处方比较少。从哮喘处方相似度结果比较中可以看出,从第一层次组成角度分析,组成不相似的处方只占总体相似度的一小部分,而且从前几项分析,单味药偏多;从第二层次组成剂量角度分析,组成不相似的仅占总体相似度的极少部分,与组成角度相比,数量少多了,说明被比较的处方间,剂量大的药物比较相似;而第三层次功效角度分析,不相似的处方几乎没有,说明哮喘处方用药的功效类别基本相似。

在研究的整个过程中,在数据的处理(包括数据的规范、剂量的标准化、功效的叠加等)、相似度方法的选择、参照方剂的选择等多个环节,均是探索性的研究,每个环节都非常重要,每个环节的合理性越高,分析结果越准确。其中数据规范的合理全面可以保证相同的药物不被当成不同的药物处理,不同的药物不会被当做相同的药物处理;剂量的合理标准化,保证中药之间剂量换算合理规范;功效的合理叠加方法,可以保证整个处方功效比较的合理性;相似度方法选择的合理性,可以保证特殊处方被清晰地筛选出来;参照方剂具有普遍性,可以保证不会

出版社, 2001: 329.

- [4] 段恒, 陆华. 资冲颗粒灌肠和灌胃对去势小鼠雌激素水平的影响[J]. 北京中医药大学学报, 2007, 30(3): 185-187.
- [5] 孙建. 实验性肾盂肾炎动物模型的建立与发病机理及中药治疗的研究[J]. 中医药信息, 1997, 4(1): 45.
- [6] 尹艳艳, 明亮, 李前进, 等. 芪杞消炎颗粒对慢性尿路感染大鼠治疗作用的实验研究[J]. 中国中医药科技, 2004, 11(1): 25-27.
- [7] 王旭丹, 袁学勤, 周勇, 等. 柴苓通淋片防治小鼠实验性泌尿道感染的实验研究[J]. 中国中医基础医学杂志, 2006, 12(11): 818-820.
- [8] 吉勤, 杨蕊娇, 曹东. 尿淋清颗粒治疗反复性尿路感染的药效学研究[J]. 中国中西医结合肾病杂志, 2009, 10(1): 68-70.
- [9] 张胜华, 余兰香, 甄瑞贤, 等. 积雪草苷的抗菌作用及对小鼠实验性泌尿系统感染的治疗作用[J]. 中国新药杂志, 2006, 15(20): 1746-1749.
- [10] 杨丽娟, 刘如意, 任会勋, 等. 八正合剂药理作用的实验研究[J]. 河南中医学院学报, 2005, 11(6): 16-18.

(收稿日期: 2012-04-03)

干扰特殊处方的筛选……。

通过不同相似度计算方式确定了较为合理的相似度测度方法,在中医处方相似性研究方面,提供了有意义的方法学上的探索,在筛选特殊的处方研究方面进行了尝试。但是由于中药方剂相似性的方法学研究尚不成熟,上述每一个环节均可以深入研究。希望在不久的将来,该类处方信息处理技术运用得更真、更深、更实用。

参考文献:

- [1] 中国学术期刊(光盘版)电子杂志社. CHKD 期刊全文数据库[EB/OL]. [2011-12-03]. <http://www.cnki.net>.
- [2] 中国中医科学院中医药信息研究所. 中医药数据库检索系统[EB/OL]. [2011-12-03]. <http://cowork.cintcm.com/engine/windex.jsp>2008-03-07/2008-03-07.
- [3] 中国医学科学院医学信息研究所. 中国生物医学文献数据库[EB/OL]. [2011-12-03]. <http://share.cintcm.com/cbmbin/login.dll>2008-03-07/2008-03-07.
- [4] 薛清录. 全国中医图书联合目录[M]. 北京: 中国古籍出版社, 1991.
- [5] 管延勇, 王洪凯, 史开泉. 集合的粗相似度量[J]. 模糊系统与数学, 2006, 20(1): 134-139.
- [6] 张宇, 刘雨东, 计钊. 向量相似度测度方法[J]. 声学技术, 2009, 28(4): 532-536.

(收稿日期: 2011-12-23)