

因子分析在中医方药量效关系研究中的应用

熊旺平, 周 娴, 杜建强, 刘玉晖, 晏 丽, 徐国良*

江西中医药大学, 江西 南昌 330004

摘 要: 目的 以经典方葛根芩连汤为例, 对中医方药临床剂量与效应的变化规律进行研究。方法 葛根芩连汤给药根据剂量分为9个组(1.65~28.05 g/kg), 每组同样的剂量给药6次, 提取血清蛋白、血糖、胰岛素、糖化血红蛋白和胰岛素抵抗指数5项指标的因子。结果 提取2个公共因子, 方差贡献率分别为55.601%和24.157%。第1因子能代表血糖、糖化血红蛋白和胰岛素抵抗指数, 而第2因子则适合代表血清蛋白和胰岛素。结论 在分析葛根芩连汤治疗糖尿病的实验数据中, 因子分析法取得了较好的效果, 可以消除人为因素的干扰, 是研究中药临床量效关系的有效方法。

关键词: 因子分析; 量效关系; 方差贡献率; 葛根芩连汤; 糖尿病

中图分类号: R285.5 文献标志码: A 文章编号: 0253-2670(2014)19-2820-04

DOI: 10.7501/j.issn.0253-2670.2014.19.017

Application of Factor Analysis in study on dose-effect relationship in traditional Chinese medicine prescription

XIONG Wang-ping, ZHOU Xian, DU Jian-qiang, LIU Yu-hui, YAN Li, XU Guo-liang

Jiangxi University of Traditional Chinese Medicine, Nanchang 330004, China

Abstract: Objective By taking classical prescription of Gegen Qinlian Decoction (GQD) as the example, this research aims at studying the law of change between the clinical dosage of Chinese medicine and its effect. **Methods** The test groups were divided into nine groups according to nine different dosages of GQD (1.65—28.05 g/kg). Each group was given the same dosage for six times. Then the factors of serum protein, blood sugar, insulin, glycated hemoglobins, and insulin resistance index were extracted. **Results** Two common factors were extracted and the variance contribution rates were 55.601% and 24.157%, respectively. The first factor could represent blood sugar, glycated hemoglobins, and insulin resistance index, while the second factor could represent serum protein and insulin. **Conclusion** Through analyzing the experimental data of GQD in the treatment of diabetes mellitus, it is found that Factor Analysis can achieve better results in eliminating the interference from human factors. Therefore, Factor Analysis is an effective and suitable method for studying the dose-effect relationship of Chinese material medica in clinical therapy.

Key words: Factor Analysis; dose-effect relationship; variance contribution rate; Gegen Qinlian Decoction; pharmacodynamics; diabetes mellitus

“中医不传之秘在于药量”, 中医药学是我国传统医学的瑰宝, 其优势在于临床疗效, 而其进一步发展所面临的突出问题仍然是疗效问题^[1]。影响中医疗效的关键因素除辨证论治、方剂配伍、中药药性及药材质量以外, 与方药的用量有着密切的关系。许多先进的技术手段也逐渐引入量效关系的研究中, 研究人员从临床研究、文献研究、数据挖掘、数理统计研究、药物化学及药理研究等各个角度分别对量效关系进行阐释, 在数据挖掘方面很多研究

人员最初用多元线性回归来描述中药方药剂量同疗效间的关系^[2]。但是由于方药成分非常复杂, 成分之间往往是非线性的关系, 中药的量效关系与化学药常见的“S”型曲线有着显著的区别^[3]。中药是一个高度复杂的化学物质体系, 其复杂性不仅表现在组成方剂的化学组分的复杂性及各组分相互关系的复杂性, 也体现在方剂与人体相互关系的复杂性。因此, 方剂的量效关系研究不能照搬化学药量效关系的研究方法, 需要同时考虑多种影响因素和配伍

收稿日期: 2014-04-11

基金项目: 国家重点基础研究发展计划(“973”计划)基金资助项目(2010CB530603); 国家自然科学基金资助项目(61363042)

作者简介: 熊旺平(1982—), 男, 讲师, 硕士, 研究方向为医药数据挖掘及医药信息化。Tel: 13576916505 E-mail: xiaoxiongxp@126.com

*通信作者 徐国良, 男, 博士, 教授, 主要从事中药药理、代谢组学及中药药理学研究。E-mail: 726531715@qq.com

情况,研究难度高。有研究者将人工神经网络、主成分分析法、模糊统计及灰色识别等方法应用到方药用量的预测中来^[4-6],但是由于常用的神经网络、主成分分析法等存在的缺点以及容易产生过拟合的现象,阻碍了其在方药量效关系模型上的应用。

因子分析法是将具有错综复杂关系的变量或者样品综合为数量较少的几个因子,以再现原始变量与因子之间的相互关系,同时根据指标相关性的把变量分组,在尽量减少信息丢失的前提下,在众多指标中提取出少量的不相关指标,然后根据方差贡献率确定权重,进而计算出综合得分的一种方法^[7],构建的中药量效关系的数据模型具有较好的拟合效果、预测能力以及稳健性。

因子分析法是假定总体是一个 p 维变量:

$$\bar{x} = (X_1, X_2, \dots, X_p)$$

它的均值向量 $E(\bar{x}) = \bar{\mu}$, 协方差矩阵 $V = (\sigma_{ij})_{p \times p}$ 都存在。

从总体中提取的综合变量: $F_1, F_2, \dots, F_m$ ($m < p$) 称为 (总体的) 公共因子。一般来说,公共因子不可能包含总体的所有信息,每个变量 X_i 除了可以由公共因子解释的那部分外,还有一些公共因子解释不了的部分,称这部分为变量 X_i 的特殊因子,记为 ε_i , a 为各公共因子的回归系数变量。 X_i 的信息 = 公共因子可以表达部分 + 公共因子不可表达部分,这就是所谓因子模型。目前,公共因子可以表达的部分由公共因子的线性组合表示,即上面的因子模型可以写成以下的形式:

$$X_i - \mu_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{im}F_m + \varepsilon_i \quad (i=1, 2, \dots, p)$$

1 数据处理过程

葛根芩连汤出自张仲景《伤寒论》,由葛根、

黄芩、黄连、甘草 4 味药组成,为太阳病误下后形成表邪未解、邪热内陷,出现下利、喘、脉促的太阳阳明合病的表里双解剂,具有止泻、抗缺氧、解热、抑菌和抗心律失常、降糖等药理作用。临床上用于治疗急性肠炎、细菌性痢疾、肠伤寒、胃肠型感冒、慢性前列腺炎急性发作、糖尿病等疾病。

2 型糖尿病 (中医消渴证) 由遗传因素、内分泌功能紊乱、微生物感染及其毒素、自由基毒素、精神及神经因素等各种致病因子作用于机体导致胰岛功能减退或缺失、胰岛素抵抗而引发的糖、蛋白质、脂肪、水和电解质等一系列代谢紊乱综合征^[8],临床上以高血糖为主要特点及与之相关的一系列并发症的常见病。因此本研究通过血清蛋白、血糖、胰岛素、糖化血红蛋白和胰岛素抵抗指数等指标的变化来进行分析。

通过葛根芩连汤 9 个给药组 (剂量 1.65~28.05 g/kg), 每组同样的剂量给药 6 次,实验对象是 2 型糖尿病大鼠模型^[9],观察血清蛋白、血糖、胰岛素、糖化血红蛋白和胰岛素抵抗指数 5 项指标。首先对数据进行聚类分析,进行去噪处理,删除对方差值影响大的数值,得到 9 组不同剂量相对准确的 5 项指标的值。见表 1。

在处理过程中,在血清蛋白、血糖、胰岛素、糖化血红蛋白和胰岛素抵抗指数 5 项指标中提取 2 个公共因子,进行因子分析。表 2 中初始特征值中的合计便是每个公共因子的方差贡献率,系统计算出全部 5 个因子的方差贡献值,并按降序排列。方差的百分比是每个因子的方差贡献占总方差的比率,即方差贡献率。其后的提取平方和载入表示在未经旋转时,被提取的 2 个公共因子 (表 2 中为第 1、2 因子) 各

表 1 不同剂量对应的指标值

Table 1 Corresponding index values of different doses

组号	剂量 / (g·kg ⁻¹)	血清蛋白 / (mmol·L ⁻¹)	血糖 / (mmol·L ⁻¹)	胰岛素 / (U·mL ⁻¹)	糖化血红蛋白 / %	胰岛素抵抗指数 / %
1	1.65	1.77	22.85	6.90	9.25	7.01
2	4.95	2.54	21.13	6.97	9.10	6.53
3	8.25	1.80	20.15	7.47	9.20	6.68
4	11.55	1.98	18.00	7.57	9.27	6.06
5	14.85	1.89	17.78	7.33	8.55	5.86
6	18.15	1.98	16.43	6.66	8.57	4.98
7	21.45	1.59	21.38	7.18	9.40	6.84
8	24.75	1.98	23.70	7.15	9.35	7.52
9	28.05	2.13	22.05	7.23	9.45	7.07

表2 方差解释表

Table 2 Total variance explanation

因子	初始特征值			提取平方和载入			旋转平方和载入		
	合计	方差贡献率 / %	累积方差贡献率 / %	合计	方差贡献率 / %	累积方差贡献率 / %	合计	方差贡献率 / %	累积方差贡献率 / %
1	2.780	55.601	55.601	2.780	55.601	55.601	2.721	54.424	54.424
2	1.208	24.157	79.758	1.208	24.157	79.758	1.267	25.334	79.758
3	0.816	16.330	96.088						
4	0.195	3.898	99.986						
5	0.001	0.014	100.000						

自方差贡献值以及方差贡献率。从中可以看到,在未经旋转时,提取的第1公共因子的方差值为2.780,方差贡献率为55.601%;第2公共因子的方差值为1.208,方差贡献率为24.157%。同时,2个公共因子可以解释总方差的79.758%,即血清蛋白等5个指标的近80%的信息可以由这2个公共因子来解释。

可以看到,经过旋转后,得到的新公共因子的方差贡献值、方差贡献率和累积方差贡献率与未经旋转相比,每个因子的方差贡献值有变化,但累积方差贡献率不变。表3、4为因子载荷矩阵表,可以明显看到,旋转后的载荷矩阵比未旋转时更容易解释因子意义。现以旋转后的因子载荷矩阵为例说明。由于因子载荷是变量与公共因子的相关系数,因此对一个变量来说,载荷绝对值较大,因子与其关系更密切,也更能代表这个变量。按照这一观点,第1因子(F_1)更能代表血糖、糖化血红蛋白和胰岛素抵抗指数,而第2因子(F_2)则更适合代表血清蛋白和胰岛素。

表5为因子得分系数矩阵表,其给出因子得分公式:

$$F_1 = 0.079 \times \text{血清蛋白} + 0.375 \times \text{血糖} - 0.011 \times \text{胰岛素} + 0.318 \times \text{糖化血红蛋白} + 0.359 \times \text{胰岛素抵抗指数}$$

$$F_2 = -0.605 \times \text{血清蛋白} - 0.187 \times \text{血糖} + 0.631 \times \text{胰岛素} + 0.108 \times \text{糖化血红蛋白} + 0.004 \times \text{胰岛素抵抗指数}$$

对2个因子得分加权求和,权数取方差值或方差百分比。本研究旋转后第1因子方差值为2.721,第2因子方差值为1.267。可以用以下公式求综合得分(F_z):

$$F_z = 2.721 F_1 + 1.267 F_2$$

葛根芩连汤9组给药组的疗效因子得分、综合得分及排名情况见表6。

2 分析与总结

(1) 因子和综合得分为负数并不代表该给药组的量会产生副作用,而是说明该给药组的疗效效果

表3 未经旋转的因子载荷矩阵表

Table 3 Unrotated component matrix

指标	因子	
	1	2
血清蛋白	-0.111	0.735
血糖	0.923	0.309
胰岛素	0.309	-0.751
糖化血红蛋白	0.926	-0.053
胰岛素抵抗指数	0.981	0.079

表4 旋转后的因子载荷矩阵表

Table 4 Rotated component matrix

指标	因子	
	1	2
血清蛋白	0.033	-0.742
血糖	0.966	-0.124
胰岛素	0.158	0.796
糖化血红蛋白	0.898	0.232
胰岛素抵抗指数	0.978	0.112

表5 因子得分系数矩阵表

Table 5 Coefficient matrix of factor score

指标	因子	
	1	2
血清蛋白	0.079	-0.605
血糖	0.375	-0.187
胰岛素	-0.011	0.631
糖化血红蛋白	0.318	0.108
胰岛素抵抗指数	0.359	0.004

低于样本平均水平。对比空白组,葛根芩连汤用药效果总体效果还是比较好的,综合竞争力大于0的给药组有6组,占样本数的66.7%。其中,葛根芩连汤第8给药组(24.75 g/kg)以2.93分居榜首,最低得分第6给药组(18.15 g/kg)为-6.26,相差

表 6 葛根芩连汤给药效果各因子得分、综合得分及排名情况表

Table 6 Score, comprehensive score, and rank of each factor after administration of GQD

组号	剂量 / (g·kg ⁻¹)	第 1 因子		第 2 因子		综合评价	
		得分	排名	得分	排名	得分	排名
1	1.65	0.675 42	3	-0.280 53	7	1.48	5
2	4.95	0.276 87	5	-1.800 90	9	-1.53	7
3	8.25	0.057 37	6	1.077 20	2	1.52	4
4	11.55	-0.443 93	7	1.081 37	1	0.16	6
5	14.85	-1.267 53	8	0.553 76	4	-2.75	8
6	18.15	-1.811 70	9	-1.048 56	8	-6.26	9
7	21.45	0.454 61	4	0.906 95	3	2.39	2
8	24.75	1.190 26	1	-0.244 78	6	2.93	1
9	28.05	0.868 63	2	-0.244 51	5	2.05	3

9.19 分, 样本标准差为 0.968 3, 说明葛根芩连汤因剂量的不同, 在疗效上的差距有明显的分化现象。且在药量 1.65 g/kg 基础上持续加量, 相应的综合得分并没有呈现递增或递减的趋势, 所以相应的量效关系不是简单的线性关系。

(2) 不同的给药组对应的每个因子的排名与最后的综合排名并不完全一致, 而且有的给药组的单项排名与其综合排名还有很大差距, 这主要是因为计算总分时各单项因子所占的权重不同。其中第 1 因子的权重为 54.424%, 第 2 个因子的权重为 25.334%。

(3) 从各个因子得分情况来看, 第 1 因子代表的是血糖、糖化血红蛋白和胰岛素抵抗指数, 第 1 因子得分大于 0 的给药组有 6 组, 其中, 第 8 给药组 (24.75 g/kg) 以 1.190 26 的得分名列第 1 因子得分榜首, 最低得分第 6 给药组 (18.15 g/kg) 为 -1.811 70, 相差为 3.001 96 分, 第 1 因子的权重最大, 为 54.424%, 因此在第 1 因子上得分较高的给药组往往综合排名较前, 表明葛根芩连汤对 2 型糖尿病模型动物的血糖、糖化血红蛋白和胰岛素抵抗指数等指标影响较大。对 2 型糖尿病患者, 经胰岛素控制血糖的基础上, 以葛根芩连汤为主, 多方加味治疗可以取得较好的临床疗效。

此外, 本研究的样本容量显然偏小, 因此必须做检验验证因子分析是否有效。KMO 检验值为 0.646, 小于 0.7, 但大于 0.5。效果差一些, 但可供参考。而 Bartlett 检验值 Sig=0, 变量间的相关性不显著, 故本次因子分析是符合数据分析条件的。由于因子之间的线性关系不显著, 原有变量重组出

来的因子之间的线性关系较弱, 因子参与数据建模能够有效地解决变量多重共线性等给分析方法应用带来的诸多问题, 大大缩减了模型的内部结构, 提高了模型的解释能力。本研究对科学地阐释方药量效关系及其影响因素, 系统地研究、总结和提炼方药剂量理论, 对于提高中医方药的临床疗效、指导临床合理选择剂量, 安全有效地用药具有重要的意义。

参考文献

- [1] 于 淼, 焦拥政, 仝小林, 等. 中药复方实验研究量效关系的进展与思考 [J]. 中医杂志, 2011, 52(17): 1517-1520.
- [2] 宋姚屏, 李 崑, 吴孟旭, 等. 方剂量效关联的现代研究方法概述 [J]. 辽宁中医杂志, 2006, 33(10): 1256-1257.
- [3] 孙塑伦, 翁维良, 杨龙会, 等. 中医临床研究实施设计与优化 [M]. 北京: 中国中医药出版社, 2008.
- [4] 陈 峰. 医用多元统计分析方法 [M]. 北京: 中国统计出版社, 2006.
- [5] 张建逵, 康廷国, 窦德强. 林下山参与园参无机元素的聚类分析和主成分分析 [J]. 中草药, 2012, 43(9): 1835-1840.
- [6] 尚尔鑫, 范欣生, 段金廛, 等. 方剂性味归经配伍规律的图形化知识挖掘研究 [J]. 世界科学技术—中医药现代化, 2008, 10(1): 39-44.
- [7] 李 林, 李晓冬, 宋剑南, 等. 应用匹配矩阵、因子分析和聚类分析法研究冠心病痰瘀证候特征 [J]. 中西医结合学报, 2006, 4(4): 343-347.
- [8] 安丽萍, 王英平, 刘晓梅, 等. 五味子油对链脲佐菌素诱导的 2 型糖尿病大鼠血糖的影响 [J]. 中草药, 2012, 43(3): 552-556.
- [9] 刘玉晖, 王跃生, 王颖怡, 等. 葛根芩连汤整方剂量变化治疗糖尿病大鼠的量效关系研究 [J]. 中国实验方剂学杂志, 2013, 19(12): 258-261.

• 药材与资源 •

福建金线莲 DALP 遗传多样性分析

杨春勇¹, 李 戈¹, 王艳芳¹, 高微微², 李荣英¹, 唐 玲^{1*}

1. 中国医学科学院 北京协和医学院药用植物研究所云南分所 西双版纳州傣药南药重点实验室, 云南 景洪 666100

2. 中国医学科学院 北京协和医学院药用植物研究所, 北京 100193

摘要: **目的** 通过分子标记技术对不同产地野生福建金线莲及近似种的遗传多样性进行分析, 了解不同产福建地金线莲及近似种的遗传差异。 **方法** 采用直接扩增片段长度多态性 (DALP) 分子标记技术, 对产多个产地的野生福建金线莲及近似种的 18 个居群进行多样性检测。 **结果** 筛选出 6 个引物组合, 总共扩增得到 355 个多态性位点, 遗传多态性为 100%, 平均每组引物扩增所得多态位点为 59.2 个。其中 14 个福建金线莲居群多态百分率为 95.77%, 其平均观测等位基因数 (N_a) 为 1.273 4, 有效等位基因数 (N_e) 为 1.074 4, Nei's 基因多样性指数 (H) 为 0.056 2, Shannon 信息指数 (I) 为 0.096 9, 遗传分化系数 (G_{st}) 为 0.423 0, 基因流 (N_m) 为 0.681 9。 **结论** 不同福建金线莲居群之间存在较高的遗传分化、基因交流较小。地理隔离和资源锐减可能是造成福建金线莲居群间基因流动受到限制的原因。

关键词: 福建金线莲; 遗传多样性; 种质资源; 遗传分化系数; 基因流

中图分类号: R282.7 **文献标志码:** A **文章编号:** 0253-2670(2014)19-2824-05

DOI: 10.7501/j.issn.0253-2670.2014.19.018

Genetic diversity of *Anoectochilus roxburghii* by DALPYANG Chun-yong¹, LI Ge¹, WANG Yan-fang¹, GAO Wei-wei², LI Rong-ying¹, TANG Ling¹

1. Key Laboratory for Dai and Southern Medicine of Xishuangbanna Dai Autonomous Prefecture, Yunnan Branch Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, and Peking Union Medical College, Jinghong 666100, China

2. Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, and Peking Union Medical College, Beijing 100193, China

Abstract: Objective To investigate the genetic polymorphism of wild *Anoectochilus roxburghii* and its allied species by molecular marker and find out the genetic differences between them. **Methods** Direct amplification of length polymorphism (DALP) was applied to evaluate the genetic variation of 18 populations of *A. roxburghii* and its allied species from different places. **Results** Six pairs of primers were selected to detect 355 polymorphic loci in 18 populations, the percentage of polymorphic bands (PPB) was 100%, and the amplified average polymorphic locus by each pair of primers was 59.2. The percentage of PPB of 14 populations of *A. roxburghii* was 95.77%, its observed number of alleles (N_a) was 1.2734, effective number of alleles (N_e) was 1.074 4, Nei's gene diversity (H) was 0.056 2, Shannon's information index (I) was 0.096 9, coefficient of gene differentiation (G_{st}) was 0.423 0, and estimated gene flow (N_m) was 0.681 9. **Conclusion** The results indicate that *A. roxburghii* has larger genetic differentiation and lower gene flow among populations. Geographic isolation and wild resource loss may be the major causes of the limited gene flow.

Key words: *Anoectochilus roxburghii* (Wall.) Lindl; genetic diversity; germplasm resources; gene differentiation; gene flow

金线莲别名金线兰、金丝草, 在民间以全草作为药用, 为珍稀名贵中药材, 其主要基原植物为福建金线莲 *Anoectochilus roxburghii* (Wall.) Lindl. 和台湾金线莲 *A. formosanus* Hayata^[1-2]。其中, 福建金线莲具备清热凉血、祛风利湿、除湿解毒等功效, 具有较高的药用价值^[3]。同时由于生长习性独特、

野生资源分布较少, 再加上人为的过度采挖, 造成其野生资源的日益匮乏和 market 价格的居高不下。另外, 由于形态特征近似, 一些同属以及斑叶兰属和血叶兰属的其他种在市场上也称为金线莲, 造成了金线莲种源混杂, 影响了金线莲的药材品质。

金线莲规范化栽培是保护野生资源、缓解药材

收稿日期: 2014-04-19

基金项目: 中央级公益性科研院所基本科研业务费专项基金资助项目: 云南省金线莲的种质资源收集及遗传多样性研究 (YZYN-12-04); 中国医学科学院药用植物研究所创新团队研究计划: 热带名贵药用植物种质创新与繁育 (121306)

作者简介: 杨春勇 (1981—), 男, 助理研究员, 主要从事药用植物栽培研究。Tel: (0691)2134865

***通信作者** 唐 玲 (1984—), 女, 助理研究员。E-mail: tl129@126.com