

• 药理 •

基于 LARS 算法的川芎成分组合与血管活性关系分析

李振坤^{1,3}, 陈建新^{2,4}, 杨洪军^{3*}, 刘玫琦³, 黄璐琦³, 易建强⁴

(1. 江西中医学院, 江西 南昌 330004; 2. 北京中医药大学, 北京 100029;
3. 中国中医科学院中药研究所, 北京 100700; 4. 中国科学院自动化研究所, 北京 100190)

[摘要] 目的: 探索适合中药“组效关系”的分析方法。方法与结果: 以川芎为范例, 基于 LARS 算法, 按照均匀试验设计-药效试验-数学建模(模型验证)程序, 得到阿魏酸、丁基苯酞、藁本内酯 3 种成分不同组合与血管活性的关系, 实现了成分组合的活性优化。结论: LARS 算法是中药“组效关系”研究的有效方法之一。

[关键词] LARS 算法; 组效关系; 川芎

[中图分类号] 285.5 [文献标识码] B [文章编号] 1005-9903(2009)03-0024-04

Association between Rhizoma Chuanxiong Constituents Combinations and Vasodilatation Activity Based on Least Angle Regression Algorithm

LI Zhen-kun^{1,3}, CHEN Jian-xin^{2,4}, YANG Hong-jun^{3*}, LIU Mei-qi³, HUANG Lu-qi³, YI Jian-qiang⁴

(1. Jiangxi University of Traditional Chinese Medicine, Nanchang 330004, China;
2. Beijing University of Chinese Medicine, Beijing 100029, China;
3. Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China;
4 Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

[Abstract] **Objective:** To investigate the method of constituent-activity relationship suitable for Chinese herbs. **Methods and Results:** Rhizoma chuanxiong is set as a paradigm, based on the least angle regression algorithm, we find the association between the vasodilatation activity and the different concentration combination of ferulic acid, butylphthalide and ligustilide, according to the scheme of Uniform Experimental Design-Efficacy Experiment — mathematical modeling (model validation). We also optimize the different concentration combination of the three constituents. **Conclusion:** Least angle regression algorithm is an effective method for research of constituent-activity relationship of Chinese herbs.

[Key words] least angle regression; constituent-activity relationship; Rhizoma Chuanxiong

中药物质基础可以划分为 4 个层次: 有效药材(饮片)-有效部位-有效组分(群)-有效成分(群)。在不同层次上, 定性、定量描述中药组合与药理效应之

间的关联规律, 即“组效关系”, 是中药组方设计优化的关键科学问题之一。为了探索“组效关系”的研究方法, 本文以川芎的主要药效成分组合为范例, 基于 LARS(Least Angle Regression Algorithm) 算法, 按照均匀试验设计-药效试验-数学建模(模型验证)程序, 分析阿魏酸、丁基苯酞、藁本内酯 3 种成分组合与血管活性的关系, 以优化配比。

1 材料

1.1 动物 清洁级雄性 SD 大鼠, 体重(220±20)g,

[收稿日期] 2008-10-13

[基金项目] 国家自然科学基金(30600820); 国家重点基础研究发展计划(2006CB504700)

[通讯作者] * 杨洪军, Tel: (010) 64014411-2948; E-mail: hongjun0420@vip.sina.com

购自北京大学医学部实验动物科学部, 合格证号: SCFK(京)2006-0008。

1.2 样品及试剂 阿魏酸, 纯度 98%, 南京青泽医药科技开发有限公司, 批号 QZM007; 丁基苯酞, 纯度 94%, 由中国医学科学院药物研究所石建功研究员提供; 藜本内酯, 纯度 92%, 由中国中医科学院中药研究所李慧副研究员提供; 苯肾上腺素(PE), Sigma 公司产品; 乙酰胆碱(ACh), Sigma 公司产品; Krebs-Henseleit(K-H)液配方: NaCl 118.96 mmol·L⁻¹; KCl 4.73 mmol·L⁻¹; KH₂PO₄ 1.17 mmol·L⁻¹; MgSO₄ 1.17 mmol·L⁻¹; NaHCO₃ 25.0 mmol·L⁻¹; CaCl₂ 2.54 mmol·L⁻¹; 葡萄糖 11.1 mmol·L⁻¹。

1.3 仪器 ALC-M 型离体组织器官实验系统, 上海奥尔科特生物科技有限公司; BP110S 型电子天平, 德国塞多利斯公司。

2 方法及结果

2.1 主动脉环制备及血管活性检测方法 取雄性 SD 大鼠, 颈部脱臼, 剪开胸腔, 迅速取出胸主动脉条, 放置于 4 °C K-H 液中, 小心剥去附在胸主动脉的脂肪及结缔组织, 横切成 2~3 mm 长的血管环。将血管环悬挂于预置 5 mL K-H 液的浴槽中, 温度 (37.0 ± 0.2) °C 恒定, 持续通入 95% O₂ 和 5% CO₂ 的混合气体。标本的一端固定, 另一端经张力换能器连接 ALC-M 离体组织器官实验系统, 记录实验过程中张力的变化。稳定过程先以 0 g 张力开始, 维持 30 min 后调节其基础张力至 1 g, 平衡 1 h, 期间每隔 15 min 更换 1 次 K-H 液。主动脉环稳定后, 用 KCl (60 mmol·L⁻¹) 刺激, 以激发最大收缩, 待收缩幅度稳定后用 K-H 液清洗。待血管环重新稳定后, 用 PE (10⁻⁶ mol·L⁻¹) 收缩血管环达到峰值, 加入累计浓度的 ACh (10⁻⁹ mol·L⁻¹ ~ 10⁻⁵ mol·L⁻¹), 以血管舒张的程度检验血管内皮的完整性, 舒张幅度达 70% 以上者可认为内皮完整。

血管环用 K-H 液清洗数次, 使其保持基础张力。以 KCl (60 mmol·L⁻¹) 诱发最大收缩幅度为 100%, 以加入受试样品后的血管张力变化幅度与 KCl 诱发最大收缩幅度之间的比率作血管舒张反应的量-效曲线。

2.2 3 种成分的血管舒张活性 采用 GraphPad Prism 4.0 软件, 用于拟合阿魏酸、藜本内酯、丁基苯酞 3 种成分的量效曲线, 并计算出最大舒张效应和半最大效应浓度, 如表 1 所示。3 种成分均表现出

舒张血管活性。

表 1 3 种成分舒张血管作用 ($\bar{x} \pm s, n = 4$)

受试样品	起效浓度 ($\mu\text{g}\cdot\text{mL}^{-1}$)	最大舒张效应 (%)	半最大效应浓度 ($\mu\text{g}\cdot\text{mL}^{-1}$)
阿魏酸	16.25	92.15 ± 2.73	470.6 ± 43.1
藜本内酯	13.13	104.32 ± 14.63	1 652 ± 47.2
丁基苯酞	18.75	109.48 ± 3.30	173.7 ± 13.2

2.3 均匀试验设计及成分组合的血管活性检测 依据阿魏酸、藜本内酯、丁基苯酞 3 种成分单独作用的血管活性试验数据, 按照均匀设计试验表, 进行 3 因素 9 水平的均匀试验设计, 从 16~144 $\mu\text{g}\cdot\text{mL}^{-1}$, 平均分成 9 个水平, 具体组合见表 2。

表 2 3 因素 9 水平均匀试验设计表

试验 序号	因素		
	丁基苯酞 ($\mu\text{g}\cdot\text{mL}^{-1}$)	阿魏酸 ($\mu\text{g}\cdot\text{mL}^{-1}$)	藜本内酯 ($\mu\text{g}\cdot\text{mL}^{-1}$)
1	48	112	144
2	96	64	128
3	144	16	112
4	32	128	96
5	80	80	80
6	128	32	64
7	16	144	48
8	64	128	32
9	112	48	16

按照 2.1 的离体血管试验方法, 对均匀试验设计所确定的 9 个组合, 进行了血管活性检测, 结果如表 3 所示。

表 3 均匀设计试验组合的血管活性 ($\bar{x} \pm s, n = 4$)

试验 序号	丁基苯酞 ($\mu\text{g}\cdot\text{mL}^{-1}$)	阿魏酸 ($\mu\text{g}\cdot\text{mL}^{-1}$)	藜本内酯 ($\mu\text{g}\cdot\text{mL}^{-1}$)	合计 ($\mu\text{g}\cdot\text{mL}^{-1}$)	平均舒 张率(%)
1	49	106	139	294	55 ± 11.6
2	94	57	118	269	61 ± 9.4
3	140	16	112	268	51 ± 4.6
4	34	125	92	251	44 ± 4.9
5	82	77	78	237	61 ± 6.6
6	125	26	69	220	56 ± 11.0
7	17	140	50	207	31 ± 2.5
8	65	121	38	224	49 ± 2.5
9	112	46	22	180	36 ± 3.0

说明: 丁基苯酞、阿魏酸、藜本内酯的剂量为依据均匀设计的实际称量给药剂量(由于称量的原因实际试验与设计有些误差)。

由表 3 可知: 川芎主要成分丁基苯酚、阿魏酸、藜本内酯组合后, 3 种成分间的比例不同, 其血管舒张活性也不同。

2.4 数学建模及活性优化 针对“非线性、离散性、小样本”数据特点, 采用基于 LARS 算法的非线性多元回归方法, 对均匀设计得到的 9 个试验组合进行数学建模。

2.4.1 LARS 算法的定义 LARS 是 least angle regression 最小角度回归的缩写, 是 2004 年由美国学者提出 LARS 算法是一种回归算法^[1], 是求解下面的方程:

$$\min_{\beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2$$

$$\text{s. t. : } \sum_{j=1}^p |\beta_j| \leq c$$

$$\beta_k$$

LARS 算法是一个残差拟合过程, 当选择了一个变量以后, 从结果中减去该变量的影响, 得到的便是残差; 然后再在该残差的基础上选择新的特征。因此拟合过程也就是残差不断减小的过程。

2.4.2 LARS 算法的步骤

(1) 初始化, 把所有的系数 β_j 设置为 0。

(2) 先寻找与目标 y 最相关的一个变量, 记作 x_j , 把 β_j 朝着和 y 相关的符号方向增大, 把残差 $r = y - \bar{y}$ 沿着这个方向计算, 直到另外一个变量 x_k 被选中以使得 r 和 x_j 最相关。

(3) 朝着 β_j 和 β_k 联合的最小均方方向增大, 该方向与 β_j 和 β_k 的残差有相等的相关系数, 直到另外一个变量 x_m 被选中以满足 r 和 x_m 最相关。

(4) 直到所有的变量都出现, 算法收敛。

我们采用 LARS 算法, 得到的舒张血管活性 (Activity) 与丁基苯酚 (X), 阿魏酸 (Y) 和藜本内酯 (Z) 之间的数学模型为: Activity = 0.221 7X + 0.110 3 Y + 0.194 2Z + 0.01* 0.047 4* X* Z + 0* X* Y + 0* Y* Z + 0.000 1* X* Y* Z - 0.01* 0.166* Z² + 0* X² + 0* Y² + 1.952 4。

根据药效实验结果, 3 种成分之和的量越大, 其活性必然越大, 不存在活性优化的区间。因此, 为了更好的反映经预测组合后的活性较优于 3 种成分简单配比组合的活性, 我们将 3 种成分之和确定在一个较小的量也即 280 $\mu\text{g}\cdot\text{mL}^{-1}$ 。依据数学模型, 在 3 个成分之和为 280 $\mu\text{g}\cdot\text{mL}^{-1}$ 的前提下, 利用电脑程序

遍历所有的取值, 最终获得血管活性超过 80% 的组合, 如表 4 所示。

表 4 活性预测数据表

组合	丁基苯酚 ($\mu\text{g}\cdot\text{mL}^{-1}$)	阿魏酸 ($\mu\text{g}\cdot\text{mL}^{-1}$)	藜本内酯 ($\mu\text{g}\cdot\text{mL}^{-1}$)	活性区间 (%)
1	100	90	90	72.67~ 88.82
2	100	100	80	73.60~ 89.96
3	100	110	70	73.38~ 89.69
4	100	120	60	72.01~ 88.01
5	110	80	90	73.29~ 89.57
6	110	90	80	74.60~ 91.18
7	110	100	70	74.68~ 91.28
8	110	110	60	73.52~ 89.86
9	120	70	90	73.14~ 89.39
10	120	80	80	74.92~ 91.57
11	120	90	70	75.39~ 92.14
12	120	100	60	74.53~ 91.09
13	120	110	50	72.34~ 88.42
14	130	60	90	72.22~ 88.27
15	130	70	80	74.56~ 91.13
16	130	80	70	75.49~ 92.27
17	130	90	60	75.02~ 91.69
18	130	100	50	73.13~ 89.38
19	140	60	80	73.51~ 89.85
20	140	70	70	75.00~ 91.67
21	140	80	60	75.00~ 91.66
22	140	90	50	73.50~ 89.83
23	150	60	70	73.91~ 90.34
24	150	70	60	74.46~ 91.01
25	150	80	50	73.43~ 89.75

对预测结果进行了试验验证, 结果如表 5 所示, 实测活性在预测范围之内。

表 5 组合预测实际结果 ($\bar{x} \pm s, n = 4$)

组合	丁基苯酚 ($\mu\text{g}\cdot\text{mL}^{-1}$)	阿魏酸 ($\mu\text{g}\cdot\text{mL}^{-1}$)	藜本内酯 ($\mu\text{g}\cdot\text{mL}^{-1}$)	预测舒张率 范围 (%)	平均舒张率 (%)
3	100	110	71	74.096~ 80.57	78.596 \pm 3.03

3 讨论

“组效关系”是相对于“构效关系”而言, 在中药物质基础的不同层次上, 定性、定量研究中药组合 (包括化学部位组合、组分组合、成分组合等) 与药理效应之间的关联规律。“组效关系”体现了中药与西药的本质区别, 是中药组方设计优化的关键科学问

题之一,必将成为中药药理学新的研究领域。

由于中药物质基础的复杂性,在定量描述中药组合与药理效应之间关系时,必须借助复杂系统建模的方法。其中,基于 LARS 算法的非线性多元回归方法是在 LARS 线性方法^[1]基础上的改良方法,适合“非线性、离散性、小样本”数据分析。主要具有以下 3 个优点:(1)对数据一致性不作要求,对于偏差较大的数据依然能够给出良好的拟合能力,吻合中药药理实验数据高度离散的特点;(2)本文中改进的 LARS 回归方法不但能拟合已有的实验结果,而且能预测出成分之间的最优配比。同时,最优活性预测在一个可信区间内,而不是具体的某个值,这样便于进一步的实验验证,更符合中药药理数据的实际;(3)在刻画药物组分与药效关系过程中,方法能够自组织地选择组分之间的组合项,并自组织地剔除一些

组合项,比如从上述药效与 3 个组分之间的方程可以看出,药效与 X 和 Y , Y 和 Z 的组合项关系不大,其回归系数均为 0,这种特征选择功能为从组分或成分层次上,探讨“君臣佐使”和“七情和合”等配伍理论奠定了基础。

阿魏酸、丁基苯酚、藁本内酯是川芎中的 3 种主要成分,基于 LARS 回归算法,按照均匀试验设计-药效试验-数学建模(模型验证)程序,分析 3 种成分组合与血管活性的关系,并经过实验验证,证实了方法的可靠性,实现了成分组合的活性优化。

[参考文献]

- [1] Bradley Efron, Trevor Hastie, Iain Johnstone, *et al.* Least angle regression[J]. *The Annals of Statistics*, 2004, 32(2): 407-451.