

3 种关键工艺参数辨识方法的比较研究

刘爽悦, 沈金晶, 李文龙, 瞿海斌, 龚行楚*

浙江大学药学院药物信息学研究所, 浙江, 杭州 310058

摘要: 目的 研究多元线性回归法、逐步回归法和标准偏回归系数法 3 种关键工艺参数辨识方法的特点。方法 以黄芩水提液酸沉碱溶工艺关键参数辨识为例比较 3 种方法。结果 3 种方法所得结果基本一致, 即关键参数为药液质量分数、酸沉 pH 值、碱溶 pH 值和静置温度。结论 3 种方法的阈值选择都有主观性; 多元线性回归法思路最简单, 但辨识关键工艺参数不如逐步回归法灵敏; 标准偏回归系数法的优势在于能综合参数对多个指标的影响, 并且易于加权计算。

关键词: 黄芩; 酸沉; 碱溶; 关键工艺参数; 纯化

中图分类号: R283.3 文献标志码: A 文章编号: 0253-2670(2016)18-3193-06

DOI: 10.7501/j.issn.0253-2670.2016.18.009

Comparison on three methods for identification of critical process parameters

LIU Shuang-yue, SHEN Jin-jing, LI Wen-long, QU Hai-bin, GONG Xing-chu

Pharmaceutical Informatics Institute, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China

Abstract: Objective To investigate the characteristics of multiple linear regression method, stepwise regression method, and standardized partial regression coefficient method for the identification of critical process parameters. **Methods** The acid precipitation and alkaline dissolution process of *Scutellariae Radix* (SR) water extract was investigated as an example. **Results** Similar results were obtained using three different methods. The concentration of SR water extract, pH value of acid precipitation, pH value of alkali dissolution, and refrigeration temperature were considered as critical process parameters. **Conclusion** The selection of threshold values of the three methods all are subjective. Multiple linear regression method is simple, but is not as sensitive as stepwise regression method. The effects of a parameter on several process indices can be considered simultaneously and easily weighted using standardized partial regression coefficient method.

Key words: *Scutellariae Radix*; acid precipitation; alkaline dissolution; critical process parameter; purification

质量源于设计 (Quality by design, QbD) 是当前国际制药领域的先进理念, 强调通过了解制药过程保障药品质量^[1-2]。这必然要求明确各制药工艺对药品质量的可能影响, 具体内容包括: 确定关键生产工艺, 辨识关键工艺参数, 获得关键工艺参数和药品或中间体质量的定量关系等。其中, 辨识关键工艺参数能明确药品制造过程中重点控制的目标参数, 对于实施 QbD 理念至关重要^[3-4]。

常用的关键工艺参数辨识方法包括: 多元线性回归法、标准偏回归系数法^[5]、逐步回归法、敏感性分析法^[6]和失效模式和效果分析法^[7]等。其中, 多元线性回归法需要建立工艺参数和工艺评价指标

间的多项式模型, 比较模型中各项系数的统计检验 P 值和预设阈值的相对大小确定关键参数。该法已用于辨识醇沉^[8]和水沉^[9]等工艺的关键参数。标准偏回归系数法需计算各参数对应各工艺评价指标的标准偏回归系数绝对值并求和, 和值较大的参数更可能是关键参数。目前该法已用于确定醇沉工艺关键参数^[10]。逐步回归法通过设定移入和移出项的阈值, 在建立多元线性模型时直接将关键工艺参数对应的项保留于模型中。该法已用于辨识中药浓缩液关键质控指标^[11-12]。

黄芩 *Scutellariae Radix* 是唇形科黄芩属植物黄芩 *Scutellaria baicalensis* Georgi 的干燥根, 具有清

收稿日期: 2016-04-22

基金项目: 浙江省中医药科学研究基金项目 (2015ZB023)

作者简介: 刘爽悦, 硕士研究生。E-mail: mirandalusy@163.com

*通信作者 龚行楚, 副教授。Tel/Fax: (0571)88208426 E-mail: gongxingchu@zju.edu.cn

热燥湿、泻火解毒、止血安胎等功效^[13-14]。黄酮类化合物为黄芩主要有效成分^[15-16]，包括野黄芩苷、黄芩苷和汉黄芩苷等。黄芩提取物收录于《中国药典》2015 年版，用于制备银黄口服液、痰热清注射液、茵栀黄软胶囊等多种中成药。其制法一般采用酸沉碱溶工艺：水提黄芩后调酸沉淀，然后取沉淀并用碱液复溶。目前对酸沉碱溶工艺的研究很少，其关键工艺参数仍然不明。

本实验将以黄芩水提液酸沉碱溶工艺的关键参数辨识为例，比较多元线性回归法、逐步回归法和标准偏回归系数法，讨论各方法特点，为其他中药生产工艺关键参数辨识提供参考。

1 仪器与材料

DHG-9123A 电热恒温鼓风干燥箱，上海精宏实验设备有限公司；S40 pH 计、AB204-N 精密电子天平，德国 Mettler-Toledo 公司；J-26XPI 高速离心机，美国 Beckman Coulter 公司；THD-1008W 低温恒温槽，宁波天恒仪器厂；ZCY-15B 恒温槽，宁波天恒仪器厂；GM-1.0A 隔膜真空泵，天津津腾实验设备有限公司；Agilent 1100 高效液相色谱仪，Agilent 科技有限公司，配四元梯度泵、自动进样器、柱温箱、紫外检测器、ChemStation 工作站。

对照品野黄芩苷（批号 130617，质量分数 99.43%）、黄芩苷（批号 130609，质量分数 99.23%）和汉黄芩苷（批号 130603，质量分数 99.39%）均购自上海融禾医药科技有限公司；乙腈，色谱纯，默克股份两合公司；纯净水，浙江娃哈哈实业股份有限公司；纯化水，由 Milli-Q Synthesis 水纯化系统（德国 Millipore 公司）制得；甲醇，色谱纯，默克股份两合公司；磷酸，色谱纯，Fluka Analytical；氢氧化钠、无水乙醇，分析纯，上海凌峰化学试剂有限公司；盐酸，优级纯，国药集团化学试剂有限公司。黄芩药材来自上海凯宝药业股份有限公司，经浙江大学药学院李文龙博士鉴定为唇形科黄芩属植物黄芩 *Scutellaria baicalensis* Georgi 的干燥根。

2 方法与结果

2.1 黄芩水提液的制备

黄芩药材置于烘箱内以 55 ℃ 干燥至恒定质量后，称取 290 g 置于 5 L 圆底烧瓶中，然后加入 1.6 L 沸腾的纯净水开始提取。加热回流提取 1.0 h 后滤过过滤液。再用 1.2 L 沸腾的纯净水第 2 次提取，加热回流提取 0.5 h 后滤过过滤液。合并 2 次滤液后离心，称质量后测定其中野黄芩苷、黄芩苷和汉黄

芩苷的量。

2.2 酸沉碱溶精制

将黄芩水提液稀释至实验所需质量分数，加盐酸溶液调低 pH 值后，在预设温度下保温一定时间。保温结束后取出样品并置于低温恒温槽中静置一定时间后，抽滤得到沉淀物。加沉淀物 8 倍量的水混悬，用 10% NaOH 溶液调 pH 值，再加等量乙醇，搅拌溶解。收集滤液称质量后测定其中野黄芩苷、黄芩苷和汉黄芩苷的量。

2.3 分析方法

采用 HPLC 法测定黄芩水提液和酸沉碱溶滤液中野黄芩苷、黄芩苷和汉黄芩苷的量^[17]。分析条件：色谱柱为 Waters Atlantis-T₃ 柱（250 mm×4.6 mm，5 μm）；以 0.03% 磷酸水溶液为流动相 A，0.03% 磷酸乙腈溶液为流动相 B，进行梯度洗脱：0~15 min，5%~20% B；15~25 min，20% B；25~45 min，20%~45% B；45~46 min，45%~90% B；检测波长 335 nm（野黄芩苷）、280 nm（黄芩苷和汉黄芩苷）；体积流量 1 mL/min；进样量 5 μL；柱温 40 ℃。

2.4 部分因子实验设计及评价指标的确定

本研究采用 7 因素 2 水平的部分因子实验设计，考察药液质量分数（A，以生药质量分数表示）、酸沉 pH 值（B）、保温温度（C）、保温时间（D）、静置温度（E）、静置时间（F）和碱溶 pH 值（G）7 个参数对黄芩水提液酸沉碱溶工艺精制效果的影响。表 1 中列出的部分因子实验设计由 Design-Expert 软件（Version 8.0，美国 Stat-Ease 公司）产生。各因素的高低水平结合工业生产经验设定，也列于表 1 中。工艺评价指标为野黄芩苷、黄芩苷和汉黄芩苷保留率。

保留率 = 滤液体积 × 滤液中指标成分质量浓度 / (水提液体积 × 水提液中指标成分质量浓度)

从表 1 可知，野黄芩苷保留率在 9.1%~55.8%，黄芩苷保留率在 18.6%~92.0%，汉黄芩苷保留率在 6.3%~81.3%，说明这 3 种活性成分在酸沉碱溶工艺中存在明显损失。有文献报道野黄芩苷和黄芩苷在强酸或强碱条件均易降解^[18-19]。

2.5 3 种关键工艺参数辨识方法比较

辨识关键工艺参数时采用 3 种不同方法。

2.5.1 多元线性回归法 对于多元线性回归法，采用公式 (1) 建立工艺评价指标和工艺参数的定量模型。关键工艺参数对应的偏回归系数检验 *P* 值应小于或等于预设 *P* 值阈值。

$$Y = a_0 + \sum_{i=1}^7 a_i X_i \quad (1)$$

其中 Y 为工艺评价指标实际值, a_0 为常数项, a_i 为偏回归系数, X_i 为工艺参数的编码值

采用公式 (1) 建立工艺参数与工艺评价指标间

的模型, 模型决定系数 (R^2) 及各项 P 值列于表 2。各模型 R^2 在 0.83~0.87, 说明数据中大部分变异均能被当前模型解释。 P 值较小意味着对应工艺参数为关键参数的可能性较大。对于不同工艺评价指标, 相应的关键参数有可能不同。改变 P 值的阈值为

表 1 部分因子实验条件及结果

Table 1 Fractional factorial designed experiments and results

运行序	工艺参数设置 (括号内为编码值)							活性成分保留率/%		
	A/(g·g ⁻¹)	B	C/°C	D/min	E/°C	F/h	G	野黄芩苷	黄芩苷	汉黄芩苷
1	0.120 (+1)	3 (+1)	50 (-1)	40 (-1)	5.0 (-1)	20 (+1)	8.5 (+1)	21.7	39.6	44.1
2	0.052 (-1)	1 (-1)	50 (-1)	80 (+1)	5.0 (-1)	20 (+1)	8.5 (+1)	42.6	75.3	69.7
3	0.120 (+1)	1 (-1)	50 (-1)	40 (-1)	30.0 (+1)	12 (-1)	8.5 (+1)	49.3	75.0	79.4
4	0.120 (+1)	1 (-1)	50 (-1)	80 (+1)	30.0 (+1)	20 (+1)	6.5 (-1)	53.0	86.0	80.5
5	0.052 (-1)	3 (+1)	90 (+1)	80 (+1)	5.0 (-1)	20 (+1)	6.5 (-1)	30.2	63.2	31.4
6	0.086 (0)	2 (0)	70 (0)	60 (0)	17.5 (0)	16 (0)	7.5 (0)	47.5	81.8	70.9
7	0.120 (+1)	1 (-1)	90 (+1)	40 (-1)	5.0 (-1)	20 (+1)	6.5 (-1)	54.6	88.4	80.0
8	0.120 (+1)	1 (-1)	90 (+1)	80 (+1)	5.0 (-1)	12 (-1)	8.5 (+1)	55.8	91.1	79.6
9	0.052 (-1)	1 (-1)	50 (-1)	40 (-1)	5.0 (-1)	12 (-1)	6.5 (-1)	47.7	89.3	81.3
10	0.086 (0)	2 (0)	70 (0)	60 (0)	17.5 (0)	16 (0)	7.5 (0)	47.7	84.3	72.6
11	0.120 (+1)	3 (+1)	90 (+1)	80 (+1)	30.0 (+1)	20 (+1)	8.5 (+1)	23.2	44.5	16.7
12	0.120 (+1)	3 (+1)	50 (-1)	80 (+1)	5.0 (-1)	12 (-1)	6.5 (-1)	29.8	57.9	38.5
13	0.052 (-1)	1 (-1)	90 (+1)	80 (+1)	30.0 (+1)	12 (-1)	6.5 (-1)	48.7	92.0	69.8
14	0.052 (-1)	3 (+1)	50 (-1)	40 (-1)	30.0 (+1)	20 (+1)	6.5 (-1)	20.9	37.6	27.2
15	0.052 (-1)	1 (-1)	90 (+1)	40 (-1)	30.0 (+1)	20 (+1)	8.5 (+1)	32.9	65.3	64.2
16	0.086 (0)	2 (0)	70 (0)	60 (0)	17.5 (0)	16 (0)	7.5 (0)	48.7	89.0	76.4
17	0.120 (+1)	3 (+1)	90 (+1)	40 (-1)	30.0 (+1)	12 (-1)	6.5 (-1)	29.4	53.1	17.4
18	0.052 (-1)	3 (+1)	50 (-1)	80 (+1)	30.0 (+1)	12 (-1)	8.5 (+1)	9.1	18.6	6.3
19	0.052 (-1)	3 (+1)	90 (+1)	40 (-1)	5.0 (-1)	12 (-1)	8.5 (+1)	17.8	34.6	19.4

表 2 多元线性回归模型中的偏回归系数、 P 值与 R^2 值

Table 2 Partial regression coefficients, P values, and R^2 values of multivariate linear regression models

指标	野黄芩苷保留率		黄芩苷保留率		汉黄芩苷保留率	
	偏回归系数	P 值	偏回归系数	P 值	偏回归系数	P 值
常量	0.370		0.670		0.540	
A	0.042	0.033*	0.037	0.235	0.042	0.207
B	-0.130	<0.001*	-0.200	<0.001*	-0.250	<0.001*
C	0.012	0.513	0.033	0.290	-0.030	0.351
D	0.011	0.522	0.029	0.357	-0.013	0.689
E	-0.021	0.244	-0.042	0.185	-0.052	0.126
F	-5.3×10^{-3}	0.762	-7.3×10^{-3}	0.810	0.014	0.666
G	-0.039	0.045*	-0.077	0.025*	-0.029	0.369
R^2	0.861		0.836		0.868	

* $P < 0.05$

0.05、0.10、0.15，所得关键参数列于表 3。当 P 阈值为 0.05 和 0.10 时，关键参数为药液质量分数、酸沉 pH 值和碱溶 pH 值。当 P 阈值为 0.15 时，静置温度也被判断为关键参数。

2.5.2 逐步回归法 对于逐步回归法，在公式 (1) 的基础上将移入或移出特定项的 P 值分别设定为 0.05、0.10、0.15。最终模型中各项对应的参数为关键工艺参数。在不同移入和移出项 P 值时所得模型系数、 R^2 及相应偏回归系数值见表 4。各模型 R^2 在 0.75~0.83，说明模型中保留的项能解释数据大部分

变异。对于野黄芩苷保留率和黄芩苷保留率，当 P 值为 0.05、0.10、0.15 时，所得结果完全相同。对于汉黄芩苷保留率，当 P 值提高到 0.15 时，模型中的项数以及模型 R^2 均增加。不同 P 值时所得关键工艺参数见表 3。当移入和移出项 P 值为 0.05 和 0.10 时，关键参数为药液质量分数、酸沉 pH 值和碱溶 pH 值。当移入和移出项 P 值为 0.15 时，关键参数包括静置温度。

2.5.3 标准偏回归系数法 对于标准偏回归系数法。首先对各个关键工艺评价指标采用公式 (2) 进

表 3 不同筛选方法所得关键工艺参数

Table 3 Critical process parameters selected using different methods

判断指标	多元线性回归法 P 值			逐步回归法 P 值			标准偏回归系数法绝对值之和		
	0.05	0.10	0.15	0.05	0.10	0.15	0.61~0.72	0.52~0.61	0.33~0.52
A	✓	✓	✓	✓	✓	✓		✓	✓
B	✓	✓	✓	✓	✓	✓	✓	✓	✓
C									
D									
E			✓			✓			✓
F									
G	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓ 关键工艺参数
✓ critical process parameter

表 4 逐步回归模型中的偏回归系数、 R^2 和 P 值

Table 4 Partial regression coefficients, R^2 values, and P values of stepwise regression models

移入和移出 项的 P 值	野黄芩苷保留率			黄芩苷保留率			汉黄芩苷保留率		
	模型中的项	偏回归系数	P 值	模型中的项	偏回归系数	P 值	模型中的项	偏回归系数	P 值
0.05	常量	0.370		常量	0.670		常量	0.540	
	A	0.042	0.021 2	B	-0.200	<0.000 1	B	-0.250	<0.000 1
	B	-0.130	<0.000 1	G	-0.077	0.022 1			
	G	-0.039	0.030 9						
	R^2	0.830		R^2	0.749		R^2	0.787	
0.10	常量	0.370		常量	0.670		常量	0.540	
	A	0.042	0.021 2	B	-0.200	<0.000 1	B	-0.250	<0.000 1
	B	-0.130	<0.000 1	G	-0.077	0.022 1			
	G	-0.039	0.030 9						
	R^2	0.830		R^2	0.749		R^2	0.787	
0.15	常量	0.370		常量	0.670		常量	0.540	
	A	0.042	0.021 2	B	-0.200	<0.000 1	B	-0.250	<0.000 1
	B	-0.130	<0.000 1	G	-0.077	0.022 1	E	-0.052	0.106 5
	G	-0.039	0.030 9						
	R^2	0.830		R^2	0.749		R^2	0.820	

行标准化处理, 然后通过多元线性回归建立工艺参数和指标的定量模型, 多元线性回归模型见公式(3), 然后将各偏回归系数的绝对值相加, 加和值大于阈值的项对应参数为关键参数。多元线性回归和逐步回归计算均由 Design-Expert 8.0 软件完成。

$$Y' = (Y - \bar{Y}) / s \quad (2)$$

其中 Y 为工艺评价指标标准化后的值, \bar{Y} 为工艺评价指标平均值, s 为工艺评价指标标准差

$$\bar{Y} = b_0 + \sum_{i=1}^7 b_i X_i \quad (3)$$

其中 b_0 为常数项, b_i 为偏回归系数

对表 1 中各个关键工艺评价指标以公式(2)进行标准化, 用多元线性回归方程进行拟合, 获得模型的确定系数及偏回归系数结果见表 5。偏回归系数绝对值大小反映了工艺参数对各指标的影响程度。将其加和后可以获得某工艺参数对多个指标的综合影响。选择不同阈值可能获得不同关键工艺参数, 表 3 中列出了一些可能的阈值及所得关键参数。由表 3 可知, 阈值越低, 所得的关键工艺参数越多。

表 5 标准偏回归系数模型 R^2 、偏回归系数及其绝对值之和
Table 5 Partial regression coefficients, sums of their absolute values, and model R^2 values obtained using Equation 3

项目	保留率偏回归系数			偏回归系数绝对值之和
	野黄芩苷	黄芩苷	汉黄芩苷	
A	0.290	0.160	0.160	0.61
B	-0.880	-0.850	-0.940	2.67
C	0.080	0.140	-0.110	0.33
D	0.079	0.120	-0.048	0.25
E	-0.150	-0.180	-0.190	0.52
F	-0.037	-0.032	0.052	0.12
G	-0.270	-0.340	-0.110	0.72
R^2	0.861	0.836	0.868	—

3 讨论

3 种方法的共同点是都使用了线性回归, 建模时没有平方项和交叉项, 所以不能体现出工艺参数的非线性作用和交互作用。3 种方法的阈值选择对最终得到的关键工艺参数都可能有较大影响。前 2 种方法中 P 值阈值设得越小, 或者第 3 种方法中标准偏回归系数绝对值之和阈值越大, 得到的关键工艺参数数量越少。3 种方法的阈值选择都有一定主观性。当实验数据量较多时, 分析结果的可靠性较

高, 可以考虑设定更严格的阈值。本实验 3 种方法所得结果基本一致, 可以认为药液质量分数、酸沉 pH 值、碱溶 pH 值和静置温度是黄芩酸沉碱溶工艺的关键参数。

比较表 2 和表 4 可知, 逐步回归法所得模型中各项的 P 值均小于多元线性回归法所得模型中相应项的 P 值。其原因在于逐步回归法建模时会删去部分不重要的项, 这能使保留在模型中的项更为显著。所以逐步回归法比多元线性回归法更敏感, 有可能在 P 值阈值相同的情况下辨识出更多关键参数。多元线性回归法和逐步回归法均考察某个参数对某个指标的影响是否显著, 而标准偏回归系数法的优点在于能体现出某个参数对所有评价指标的综合影响。如果各个指标之间的权重不同, 那么在计算标准偏回归系数绝对值之和时也很容易相应进行加权。当工艺评价指标数量较多时, 建议采用标准偏回归系数法。

参考文献

- [1] 许之麟. 基于 QbD 理念的党参水提醇沉工艺研究 [D]. 杭州: 浙江大学, 2016.
- [2] 张小飞, 邢传峰, 果秋婷. 基于质量源于设计 (QbD) 理念优化穿心莲内酯固体脂质纳米粒 [J]. 中草药, 2015, 46(2): 194-200.
- [3] 薛晶, 朱克旭, 崇小萌, 等. 水分对阿莫西林克拉维酸钾颗粒稳定性的影响 [J]. 中国药学杂志, 2016, 51(3): 224-229.
- [4] 王永香, 米慧娟, 张传力, 等. Box-Behnken 响应面法优化热毒宁注射液金银花和青蒿 (金青) 的醇沉工艺研究 [J]. 中草药, 2015, 46(5): 671-678.
- [5] 严斌俊, 郭正泰, 瞿海斌, 等. 丹红注射液醇沉关键工艺参数筛选方法 [J]. 中国中药杂志, 2013, 38(11): 1672-1675.
- [6] Degerman M, Westerberg K, Nilsson B. Determining critical process parameters and process robustness in preparative chromatography-A model-based approach [J]. *Chem Eng Technol*, 2009, 32(6): 903-911.
- [7] Gong X C, Zhang Y, Pan J Y, et al. Optimization of the ethanol recycling reflux extraction process for saponins using a design space approach [J]. *PLoS One*, 2014, 9(12): e0114300.
- [8] Zhang L, Gong X C, Qu H B. Optimizing the alcohol precipitation of danshen by response surface methodology [J]. *Separ Sci Technol*, 2013, 48(6): 977-983.
- [9] Gong X C, Chen H L, Chen T, et al. Unit operation optimization for the manufacturing of botanical injections

- using a design space approach: A case study of water precipitation [J]. *PLoS One*, 2014, 9(8): e104493.
- [10] 许之麟, 黄文华, 龚行楚, 等. 设计空间法优化党参一次醇沉工艺 [J]. 中国中药杂志, 2015, 40(22): 4411-4416.
- [11] 闫安忆, 龚行楚, 瞿海斌. 一种中药醇沉前浓缩液关键质量控制指标的辨析方法 [J]. 中国中药杂志, 2012, 37(11): 1558-1563.
- [12] 张 寒, 闫安忆, 龚行楚, 等. 丹参注射液生产中一次醇沉上清液浓缩工艺质控指标研究 [J]. 中国中药杂志, 2011, 36(11): 1436-1440.
- [13] 陈佩东, 徐丹洋, 孔祥鹏, 等. 黄芩中黄酮类成分的分 离鉴定及其体外对凝血系统的影响 [J]. 中草药, 2012, 43(12): 2333-2336.
- [14] 中国药典 [S]. 一部. 2015.
- [15] 刘志辉, 李俊生, 郑 啸, 等. 透皮促渗剂对黄芩总黄酮凝胶膏剂成分体外经皮渗透的影响 [J]. 中草药, 2015, 46(18): 2703-2711.
- [16] 李俊生, 吕佳佳, 王兴慧, 等. 黄芩总黄酮及其单体的溶解性及体外经皮渗透性能研究 [J]. 中草药, 2014, 45(2): 200-207.
- [17] 刘爽悦. 黄芩水提物纯化过程质量控制方法研究 [D]. 杭州: 浙江大学, 2015.
- [18] 钟海军, 邓英杰, 徐春莲, 等. 灯盏花素的理化性质研究 [J]. 中国药房, 2013, 24(7): 608-610.
- [19] 郑辛甜, 瞿海斌. 紫外光谱法结合多元曲线分辨研究黄芩苷的转化规律 [J]. 高等学校化学学报, 2012, 33(7): 1450-1455.