

• 药剂与工艺 •

基于裂区的 D-最优实验设计法优化丹参川芎嗪注射液水提工艺

赵 芳¹, 陈泽麒¹, 李文竹¹, 苟 维², 刘 飞², 张程益², 瞿海斌^{1*}

1. 浙江大学药物信息学研究所, 浙江 杭州 310058

2. 贵州拜特制药有限公司, 贵州 贵阳 550008

摘 要: 目的 在质量源于设计 (quality by design, QbD) 理念的指导下优化丹参川芎嗪注射液 (*Salviae Miltiorrhizae* and *Ligustrazine Hydrochloride Injection*, SMLHI) 的水提工艺。方法 首先通过筛选实验设计确定加热温度、保温时间、药材规格和溶剂倍量为丹参水提工艺的关键工艺参数。随后为了提高实验效率, 采用基于裂区的 D-最优实验设计优化 SMLHI 的水提工艺。为了尽可能多的提取实验数据潜在信息, 采用了响应曲面法和人工神经网络 2 种方法建立关键工艺参数和关键质量属性之间的回归模型。最后, 选择较优的模型进一步采用基于满意度函数的多指标优化算法综合考察提取液各关键质量属性, 确定水提工艺的最佳操作条件。**结果** 基于裂区的 D-最优实验设计可以有效提高实验效率, 同时发现人工神经网络模型较响应曲面法模型具有更好的拟合能力和预测能力。以人工神经网络模型为基础进行多指标优化, 最终确定 SMLHI 水提工艺的最佳操作条件为加热温度 150 °C, 保温时间 84 min, 药材规格 11 cm, 溶剂用量为 10 倍。**结论** 利用基于裂区的 D-最优实验设计法优化了 SMLHI 水提工艺, 为传统的中药制药工艺研究过程中面临的工艺参数改动受限、更改成本高昂的情况提供合理的实验设计方案, 提高研究效率, 降低研究成本, 对不同试验规模下中药制药工艺的过程研究具有较大参考价值, 为中药的工艺二次开发提供了可供参考的新方法。

关键词: 质量源于设计; 裂区设计; 响应曲面法; 人工神经网络; 丹参川芎嗪注射液; 水提工艺; 多指标优化算法; 关键工艺参数; D-最优实验设计法; 关键质量属性; 满意度函数

中图分类号: R284.1 文献标志码: A 文章编号: 0253-2670(2021)01-0045-10

DOI: 10.7501/j.issn.0253-2670.2021.01.007

Optimization of water extraction process of *Salviae Miltiorrhizae* and *Ligustrazine Hydrochloride Injection* via split-plot based D-optimal experimental design

ZHAO Fang¹, CHEN Ze-qi¹, LI Wen-zhu¹, GOU Wei², LIU Fei², ZHANG Cheng-yi², QU Hai-bin¹

1. Pharmaceutical Informatics Institute, Zhejiang University, Hangzhou 310058, China

2. Guizhou Baite Pharmaceutical Co., Ltd., Guiyang 550008, China

Abstract: Objective Quality by design (QbD) concept was applied to guide the efficient optimization of the water extraction process for *Salviae Miltiorrhizae* and *Ligustrazine Hydrochloride Injection* (SMLHI). **Methods** Firstly, Plackett-Burman experimental design method was used to screen critical process parameters (CPPs). Secondly, in order to improve the study efficiency, split-plot based D-optimal methodology was applied to implement experimental design. To obtain more process information, response surface method (RSM) and artificial neural network (ANN) were used to establish the regression model between CPPs and critical quality attributes (CQAs) for the extraction. Finally, after comparison of the two models, the optimum operating conditions were determined with the multi-index optimization algorithm by a better regression model. **Results** The study efficiency was improved by the split-plot method to conduct the experimental arrangement. Then, the ANN method showed better modeling performance than the RSM in this study, and

收稿日期: 2020-09-07

基金项目: 国家“重大新药创制”科技重大专项 (2018ZX09201011-002)

作者简介: 赵 芳, 女, 博士研究生, 研究方向为中药制药工艺研究。E-mail: z_fang@zju.edu.cn

*通信作者: 瞿海斌, 博士生导师, 教授, 从事中药制药过程质量控制研究。Tel: (0571)88208428 E-mail: quhb@zju.edu.cn

then it was taken as the final optimum model for the multi-index optimization. The optimal operating conditions of the extraction process for SMLHI were follows: The heating temperature was 150 °C, extraction time was 84 min, specification of medicinal material was 11 cm, and the ratio of material to water was 10. **Conclusion** This study optimized the water extraction process for SMLHI by introducing the split-plot design into the process optimization of traditional Chinese medicine. Split-plot design methodology can provide reasonable experimental design solutions to improve research efficiency and reduce research costs for pharmaceutical process research in traditional Chinese medicine, which faces limited changes in process parameters and high costs. And it has a greater reference value for the process study of Chinese medicine pharmaceutical process under different test scale, and provides a new method for the secondary development of Chinese medicine pharmaceutical process for reference.

Key words: quality by design; split-plot design; response surface methodology; artificial neural network; *Salviae Miltiorrhizae* and *Ligustrazine Hydrochloride Injection*; water extraction process; multi-index optimization; critical process parameter; D-optimal design; critical material attribute; satisfactory function

在中药大品种二次开发中,一般需要采用实验设计方法优化制备工艺参数^[1-4]。但在实验过程中,可能面临某些难以改变的工艺参数、满足不了设计的完全随机化要求的困难。为了提高实验效率,需要引入裂区实验设计(split-plot experiment design)方法^[5-6]。与其他实验设计方法相比,裂区实验设计通过主区将一个或多个作为副区的完全随机设计、随机区组设计、拉丁方设计等结合起来,可以适用于实验中有部分因子较难改变的情况,在医药领域常用于新型药物制剂的配方设计与优化^[7-12]。在中药制药工艺研究中,在质量源于设计(quality by design, QbD)理念的指导下合理应用裂区实验设计,有助于避免在研究过程中因为改变 HTCF 而导致后续研究中断或对后续研究造成干扰,同时帮助提高实验效率,减少不必要的资源浪费。

响应曲面法(response surface methodology, RSM)是一种随机优化过程的统计学方法^[13],目前在中药生产过程工艺优化的研究中已经有了广泛的应用^[14-15],主要用于考察工艺参数间相互作用以及工艺参数的二次效应对工艺评价指标的影响。人工神经网络(artificial neural network, ANN)在过程建模、优化方面具有强大的信息提取能力,较 RSM 可以建立更加复杂关系的模型,尤其是非线性关系,是一种多项式回归建模的替代方法^[13]。近年来,ANN 在过程预测及工艺优化方面的应用越来越广泛,并常常表现出比 RSM 更好的预测能力^[16-17]。

丹参川芎嗪注射液(*Salviae Miltiorrhizae* and *Ligustrazine Hydrochloride Injection*, SMLHI)是由丹参提取液和盐酸川芎嗪组成的复方制剂,临床主要用于治疗闭塞性脑血管疾病及其他缺血性心血管疾病^[18-20]。本研究以 SMLHI 生产的首要环节丹参水提工艺为研究对象,通过风险分析与筛选设计

确定丹参水提的关键工艺参数(critical process parameter, CPP)。在此基础上,考虑到加热效率的调整难以随机进行,采用基于裂区的 D-最优实验设计。同时,采用 RSM 和 ANN 两种建模方法处理实验设计数据,建立丹参水提工艺 CPP 和关键质量属性(critical quality attribute, CQA)之间的定量关系。最后,选择较优的模型用于指导实际生产,并通过多指标优化算法考察丹参提取液的综合质量,确定最佳操作条件。

1 仪器与试剂

1.1 仪器

Cary60 型紫外-可见分光光度计及 Agilent1100 型高效液相色谱仪,配四元梯度泵、自动进样器、柱温箱、紫外检测器、ChemStation 工作站,安捷伦科技有限公司;XS105、AB204-N 型电子分析天平、SevenMult 型 pH 综合测试仪,梅特勒-托利多上海有限公司;SE6001F 电子天平,奥豪斯仪器有限公司;Milli-Q 超纯水机,美国 Millipore 公司;DHG-9123A 型电热恒温鼓风干燥箱,上海精宏实验设备有限公司;ZCY-15B 型数控超级恒温槽,宁波天恒仪器厂。

1.2 试剂

丹参药材由贵州拜特制药有限公司提供,经贵州拜特制药有限公司苟维工程师鉴定为正品丹参(植物基原为唇形科鼠尾草属植物丹参 *Salvia miltiorrhiza* Bunge 的干燥根和根茎),药材中丹酚酸 B 含量由课题组根据《中国药典》2015 年版方法^[21]测定而得;对照品丹参素钠(批号 171027,质量分数≥98%)、原儿茶醛(批号 171126,质量分数≥98%)、紫草酸(批号 170825,质量分数≥98%)、迷迭香酸(批号 171009,质量分数≥98%)、丹酚酸 B(批号 171102,质量分数≥98%)、丹酚酸 A

(批号 170920, 质量分数 $\geq 98\%$)均购自上海融禾医药科技发展有限公司; 苯酚(批号 20141121)购自美国阿拉丁工业公司; 硫酸(批号 20171201)购自国药集团化学试剂有限公司; 乙腈、甲醇、甲酸、乙酸为色谱纯, 购自德国 Merck 公司; 超纯水由超纯水系统(Milli-Q® Synthesis, 美国 Millipore 公司)制得。

2 方法与结果

2.1 丹参水提工艺操作流程

以筛选实验中心点实验条件为例, 取事先拣选好的 5 cm 长的特定丹参药材 145 g, 加 1160 mL 水(pH 7, pH 值用盐酸和氢氧化钠调节)浸泡 12 h, 期间开启数控恒温槽, 将油预热至 140 °C。浸泡完成后, 油浴加热升温至沸腾, 保温 60 min, 倾出一提液。随后加入 1160 mL 水溶液(pH 7)进行二次提取, 同样升温至沸腾后继续保温 60 min, 倾出二提液, 与一提液均匀混合, 作为后续测试所需样本溶液。

2.2 工艺评价指标选择及分析方法

丹参素、原儿茶醛、迷迭香酸、紫草酸、丹酚酸 B 以及丹酚酸 A 是目前丹参相关中药制剂中主要的可追溯质量标志物^[22-23], 本研究选择这 6 个成分的提出率作为提取工艺效率的指征, 作为工艺研究的关键质量属性(critical quality attribute, CQA)。此外, 提取液中杂质成分的存在不利于提取液的后续纯化及精制。预实验结果表明, 糖类成分是 SMLHI 原料丹参提取物中的主要杂质成分, 占提取物总固含量的 60%以上, 其存在会显著影响产品黏度、渗透压等^[14]。综上, 本实验选择丹参素、原儿茶醛、迷迭香酸、紫草酸、丹酚酸 B、丹酚酸 A、总糖及总固体的提出率为评价指标。

单个指标提出率计算方法如公式(1)。

提出率=提取液中含量(浓度)×提取液质量(体积)/药材质量 (1)

在本研究中, 采用高效液相色谱与紫外检测器联用(HPLC-UV)法^[18]测量样品中丹参素、原儿茶醛、迷迭香酸、紫草酸、丹酚酸 B 以及丹酚酸 A 6 种成分的含量; 根据《中国药典》^[21]规定的苯酚-硫酸法测定提取液的总糖含量, 药典规定的恒重法测定提取液的总固含量。

2.3 数据处理

2.3.1 筛选实验数据处理 为满足对中药注射液的安全性及有效性的双重保证, 采用加权多元线性

回归系数法筛选对结果影响较大的参数^[24]。该方法已在多项中药制药过程工艺研究中进行了实践, 且被证明具有较好的应用价值^[15-16, 25-26]。在本实验中, 考虑到酚酸类成分为丹参提取物中的主要活性成分, 因此在筛选 CPP 时重点考虑各工艺参数对酚酸类成分的影响, 以各酚酸类成分在标准提取液(本研究中, 以实验设计矩阵的中心点作为标准提取液)中的含量作为权重, 其与回归模型相关系数的乘积作为该指标调整后的相关系数(adjusted correlation coefficient, ACC)。以同一工艺参数在各指标模型中 ACC 的绝对值之和综合评价该工艺参数对生产工艺影响的大小, 进而筛选出对水提工艺影响较大的工艺参数, 进行进一步的优化。

2.3.2 优化实验数据处理 本研究为了更好地建立 CPP 与 CQA 之间的关系, 同时采用 RSM 和 ANN 两种建模方法, 并选择较优的模型用于指导 SMLHI 的生产。

(1) 基于 RSM 建立 CPP 和 CQA 之间定量模型: 基于裂区的实验设计 RSM 模型以公式(2)表示, 由常数项、一次项、二次项、交叉项、区组效应、副区误差和主区误差组成。

$$Y = a_0 + \sum_{i=1}^4 a_i X_i + \sum_{i=1}^4 a_{ii} X_i^2 + \sum_{i=1}^3 \sum_{j=i+1}^4 a_{ij} X_i X_j + \beta_k + \varepsilon_k + \varepsilon \quad (2)$$

a_0 是常数, a_i 、 a_{ii} 与 a_{ij} 为回归系数, Y 是水提工艺关键质量属性, X_i 是 CPP, β_k 是区组效应, ε_k 是副区误差, ε 是主区误差, 模型方程中移入或移除特定项的 P 值设定为 0.10

(2) 基于径向基函数网络建立 CPP 和 CQA 之间定量模型: 径向基函数网络(radial basis function network, RBF)作为 ANN 的一种, 是一种结构简单、收敛速度快、能够逼近任意非线性函数的网络, 较多层感知器相比较, 达到相同模型精度需要的参数更少, 对小样本数据有较好的拟合效果^[27-28]。本研究以各组的工艺条件作为输入向量, 输入节点个数等于 CPP 的个数; 隐含层中心点采用随机选取的方式进行, 基函数选用高斯函数, 如公式(3)所示; 输出层为线性单元, 与隐含层全连接, 因此实际输出可以用公式(4)表示。径向基函数网络结构如图 1 所示。

$$\Phi_i = G(\|X - u_i\|) = \exp(-\|X - u_i\|^2 / 2\sigma^2) \quad (3)$$

$$Y_n = \sum_{i=1}^n w_{in} \Phi_i \quad (4)$$

X 为输入向量(工艺参数), u_i 为基函数的中心, σ 为基函数的标准差, w_{in} 为权重系数

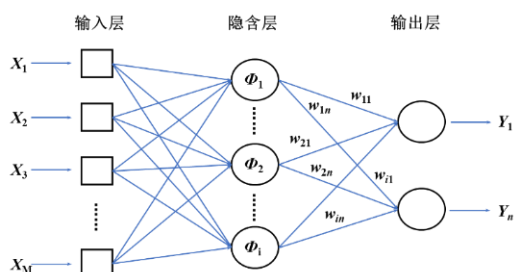


图 1 RBF 结构

Fig. 1 Structure of RBF

为了综合比较 2 种模型的性能,选择 R^2 和预测平均相对偏差 (mean relative error, MRE) 对 2 种建模方法进行比较分析。

$$\text{MRE} = 1/n \sum_{i=1}^n |(Y_{\text{pre}} - Y_{\text{real}})/Y_{\text{real}}| \quad (5)$$

Y_{real} 为关键质量属性测量值, Y_{pre} 为关键质量属性预测值

2.3.3 基于满意度函数确定最佳操作条件 在实际生产中,有效成分与杂质的控制往往较为矛盾,2 类物质的提出规律总是呈现出较为一致的趋势^[11,29]。为了在提取过程中尽可能多的提取出酚酸类成分,并尽可能减少杂质提出,从而降低产品后处理中分离与纯化的难度,提高产品的质量,本研究综合考量各 CQA,进一步利用多指标算法确定最佳的操作条件。

满意度函数法是一种常用的综合多个优化目标的方法,它将各响应指标的预测值分别通过各自的满意度函数转化成该指标的满意度,再将多个满意度合并成一个总体满意度,以便对得到的总体满意度函数进行单目标优化^[3,30]。总体满意度 (d) 的计算如公式 (6) 所示。

$$d = (\prod d_{v_n}^{w_n})^{1/\sum w_n} \quad (6)$$

d_{v_n} 为评价指标满意度, w_n 为权重

所有计算由 Microsoft Office Excel 2016 (美国 Microsoft 公司)、Design-expert 8.0.7 (美国 Stat-Ease 公司) 和 Matlab R2019b (美国 MathWorks 公司) 软件完成。

2.4 CPP 筛选

2.4.1 基于 Plackett-Burman (PB) 设计的 CPP 筛选实验 根据实际生产经验和文献调研^[11,29],挑选出 6 个待筛选的 CPP: 加热温度 (x_1)、浸泡时间 (x_2)、保温时间 (x_3)、饮片质量 (x_4 , 以药材中丹酚酸 B 含量表示)、药材规格 (x_5 , 药材切碎程度) 和溶剂 pH 值 (x_6)。并进一步采用 PB 实验设计考察上述 6 个待筛选的 CPP, 根据 SMLHI 生产过程的实际工

艺确定各参数的操作范围。各个参数的设定水平及筛选实验的具体工艺条件参数见表 1。此外,溶剂倍量会显著影响提取效率,在相关研究中均被视作 CPP。

表 1 SMLHI 水提工艺 PB 筛选实验的参数水平和实验条件
Table 1 Parameters, levels, and experimental conditions of PB screening design for water extraction process of SMLHI

组别	$x_1/^\circ\text{C}$	x_2/h	x_3/min	$x_4/\%$	x_5/cm	x_6
1	115 (-1)*	0 (-1)	45 (-1)	3 (-1)	1 (-1)	5.5 (-1)
2	165 (+1)	24 (+1)	75 (+1)	3	1	5.5
3	165	0	45	10 (+1)	1	8.5 (+1)
4	115	24	75	10	1	8.5
5	115	24	45	3	10 (+1)	8.5
6	165	0	75	3	10	8.5
7	165	24	45	10	10	5.5
8	115	0	75	10	10	5.5
9	140 (0)	12 (0)	60 (0)	7 (0)	5 (0)	7.0 (0)
10	140	12	60	7	5	7.0
11	140	12	60	7	5	7.0

*括号中数字为各参数的相应水平

*numbers in parentheses were setting levels of parameters, respectively

2.4.2 CPP 筛选实验结果 筛选实验样品溶液中各组分的含量测定结果见表 2, 各工艺指标相应的回归系数及权重系数见表 3。由表 3 可知, 6 个酚酸模型的决定系数 (R^2) 均大于 0.605, 总固体与总糖模型的 R^2 大于 0.663, 各个指标的回归模型均能描述工艺参数对指标的大部分影响。

水提工艺各影响酚酸提取效率的潜在 CPP ($x_1 \sim x_6$) 的原始及调整后回归系数绝对值之和分别为 0.262、0.065、0.170、2.404、0.000、0.000 与 0.071、0.018、0.047、7.539、0.000、0.000。说明药材中丹酚酸 B 含量、加热温度和保温时间 3 个参数调整后的回归系数绝对值之和大于其他工艺参数, 即这 3 个参数对水提工艺中各酚酸的提出率变化有较显著的影响。此外, 据表 3 可知, 药材规格可以显著影响总固体和总糖的提出率。由于药材中丹酚酸 B 的含量在实际生产中采购过程会予以控制, 因此本研究不进一步进行考虑。综上, 最终选择加热温度 (x_1)、保温时间 (x_3)、药材规格 (x_5) 和溶剂倍量 (x_7) 为需要优化指标。

2.5 CPP 优化

2.5.1 基于裂区的 D-最优实验设计 在丹参提取工艺的实际操作中, 无论是小试、中试还是商业规

表 2 SMLHI 水提工艺的 PB 筛选实验结果

Table 2 Results of PB design for screening design for water extraction process of SMLHI

序号	提出率/%							总固体
	丹参素	原儿茶醛	迷迭香酸	紫草酸	丹酚酸 B	丹酚酸 A	总糖	
1	0.198 7	0.015 6	0.299 2	0.213 4	4.785 7	0.016 1	12.59	36.04
2	0.148 0	0.025 0	0.074 1	0.089 9	1.351 8	0.069 3	22.55	46.85
3	0.186 5	0.012 0	0.185 3	0.188 9	3.974 0	0.257 5	14.73	33.45
4	0.221 4	0.026 6	0.103 0	0.150 3	2.674 5	0.197 0	20.57	38.36
5	0.119 5	0.012 3	0.102 6	0.101 3	1.788 9	0.005 0	12.75	29.65
6	0.195 1	0.033 0	0.118 1	0.104 7	1.713 4	0.112 0	18.22	37.10
7	0.339 4	0.038 2	0.324 7	0.252 3	4.802 9	0.570 6	11.83	31.97
8	0.223 4	0.026 2	0.100 0	0.151 6	2.684 4	0.200 8	19.87	36.74
9	0.208 5	0.023 5	0.084 4	0.140 9	2.527 4	0.174 1	21.64	39.18
10	0.349 5	0.035 8	0.252 7	0.256 5	4.951 5	0.630 9	17.61	41.25
11	0.160 9	0.025 3	0.099 0	0.093 8	1.454 2	0.080 0	25.47	44.88

表 3 SMLHI 水提工艺的 PB 实验各评价指标多元线性回归模型的 R^2 、回归系数及指标的权重系数Table 3 R^2 values, regression coefficients, and weight coefficient of indexes for regression models of PB design for water extraction process of SMLHI

参数	丹参素	原儿茶醛	迷迭香酸	紫草酸	丹酚酸 B	丹酚酸 A	总固体	总糖
常数	0.310	0.036	0.230	0.230	4.312	0.305	54.765	26.079
x_1	0.064	-0.011	/	-0.018	/	-0.169	-1.493	/
x_2	/	/	/	/	/	0.065	2.610	/
x_3	0.035	0.004	/	/	/	0.131	/	/
x_4	0.073	/	0.104	0.084	1.949	0.194	-2.554	-3.423
x_5	/	/	/	/	/	/	-5.992	-3.510
x_6	/	/	/	/	/	/	-2.124	-2.173
R^2	0.976	0.790	0.605	0.920	0.911	0.956	0.983	0.663
权重系数	0.316	0.037	0.139	0.214	3.812	0.276	0	0

“/”代表该工艺对模型影响不显著

“/” the slash represents the corresponding item with no significance

模下的试验,对提取罐温度进行实时调整都会影响实验效率与温控精度:首先,加热器的实时加热效率及温度本身较难调整;同时,一旦对加热器的加热效率进行调整,如不经过一段较长的等待时间,后续实验的升温曲线也会随之发生变化。此外,这样的变更也会极大增加时间成本与经费成本。因此,本实验采用基于裂区的 D-最优实验设计,将加热温度设为困难因子(hard-to-change factor, HTCF),保温时间、药材规格和溶剂用量(以倍量表示)设为易于改变因子(easy-to-change factor, ETCF),各工艺参数的水平设定与实验设计见表 4。

2.5.2 优化实验的 CQA 选择 本研究基于 SMLHI 实际生产情况和《中国药典》要求,在优化实验中对 CQA 进行了适当的调整。药典规定, SMLHI 中

的主要药效成分为盐酸川芎嗪与丹参素。丹参药材中成分主要为丹酚酸 B、丹参素等酚酸主要是由丹酚酸 B 经过一系列降解反应后产生^[31-32]。因此,基于对丹酚酸 B 降解路径的分析,本实验将紫草酸、丹酚酸 A 及丹参素折算成相应丹酚酸 B 物质的量,以四者物质的量的总和视作酚酸提出量,以该指标来综合评价酚酸类成分的提取规律。此外,考虑到杂质成分含量差异会影响提取液的物性参数以及后续纯化工艺,因此,选取总酚酸纯度(总酚酸纯度=酚酸提出量/总固体提出量)评价提取液质量。基于以上综合考虑,选择酚酸提出率、总酚酸纯度、迷迭香酸提出率、总糖提出率和总固体提出率作为优化工艺的 CQA。

2.5.3 优化实验结果 优化实验结果见表 4。

(1) RSM 模型: 根据表 4, 采用 Design-Expert 软件建立多指标优化模型, 经显著性分析, 得到模型的回归系数、 R^2 及 P 值见表 5。从表 5 可知, 各 CQA 模型 R^2 均大于 0.75, 调整 R^2 均大于 0.68, 表

明模型能解释绝大部分变异。其中, 加热温度对各个指标影响都不显著, 对比筛选实验结果, 加热温度的不同主要影响的是酚酸类成分组成。可能原因是加热温度不同对丹酚酸 B 降解存在一定程度的影

表 4 水提工艺裂区 D-优化实验参数水平设定及结果

Table 4 Parameters, levels, and results of D-optimal design for water extraction process of SMLHI

组别	序号	工艺条件				提出率/%				总酚酸 纯度/%
		$x_1/^\circ\text{C}$	x_3/min	x_5/cm	x_7	酚酸	迷迭香酸	总糖	总固体	
1	1	156.25	77.4	1	6.88	3.080 7	0.124 6	28.58	41.11	7.49
1	2	156.25	120	5	8.58	3.305 7	0.139 0	29.14	44.78	7.38
1	3	156.25	60	5	8.74	2.892 4	0.111 4	27.24	37.28	7.76
1	4	156.25	60	11	10	2.980 1	0.125 1	21.17	30.77	9.68
1	5	156.25	111	1	10	3.407 0	0.145 2	31.59	44.75	7.61
1	6	156.25	120	11	6	2.422 5	0.096 1	24.76	33.51	7.23
2	7	115 (-1)*	74.1	1 (-1)	9.3	3.113 8	0.128 1	30.89	41.72	7.46
2	8	115	120	11	6.7	2.634 5	0.119 1	24.99	34.87	7.55
2	9	115	60	5	10	2.806 2	0.104 1	24.54	33.87	8.29
2	10	115	60	11	7.3	2.418 9	0.095 2	18.65	28.90	8.37
2	11	115	60	5	6	2.859 4	0.116 9	25.78	34.75	8.23
2	12	115	102.9	1	6	3.446 5	0.139 4	38.37	48.64	7.09
3	13	140 (0)*	90 (0)	5 (0)	8 (0)	3.064 9	0.138 4	27.69	36.74	8.34
3	14	140	90	5	8	3.235 9	0.124 0	34.30	43.98	7.36
4	15	165	102.6	1	8.9	3.238 1	0.136 5	35.68	44.66	7.25
4	16	165 (+1)*	120 (+1)	5	10 (+1)	3.298 5	0.151 9	37.65	47.18	6.99
4	17	165	120	5	6	2.755 2	0.120 0	33.47	40.46	6.81
4	18	165	60	5	6	2.744 4	0.116 8	28.41	36.88	7.44
4	19	165	60 (-1)	11 (+1)	6 (-1)	2.321 9	0.092 9	19.22	27.45	8.46
5	20	140	90	1	8	3.145 4	0.125 9	32.48	44.61	7.05
5	21	140	90	1	8	3.113 1	0.124 4	29.82	42.33	7.35
6	22	165	104.1	11	6.9	2.757 3	0.131 5	21.57	35.53	7.76
6	23	165	120	11	10	2.930 5	0.144 7	22.99	36.30	8.07
6	24	165	81.9	5	10	2.996 6	0.129 8	26.98	41.72	7.18
6	25	165	60	1	6	2.860 1	0.110 2	25.11	36.76	7.78
6	26	165	120	1	6	3.193 5	0.127 8	29.03	44.20	7.23
6	27	165	60	1	10	3.206 5	0.123 6	23.84	39.79	8.06
7	28	140	90	11	8	2.720 0	0.112 8	21.15	34.30	7.93
7	29	140	90	11	8	2.490 7	0.109 7	16.64	30.67	8.12
8	30	115	120	5	6	3.071 4	0.134 9	30.62	43.05	7.13
8	31	115	120	5	10	3.271 9	0.142 5	31.18	46.73	7.00
8	32	115	120	1	7.1	3.097 3	0.126 3	31.03	44.25	7.00
8	33	115	100.2	11	10	2.945 6	0.160 0	21.42	34.27	8.60
8	34	115	78	11	6	2.686 2	0.098 2	23.22	30.79	8.73
8	35	115	60	1	6	2.903 9	0.114 5	24.28	37.42	7.76

*参数的相应水平

*numbers in parentheses were setting levels of parameters, respectively

表 5 水提工艺优化实验设计各评价指标 RSM 模型的回归系数、 R^2 和调整后 R^2

Table 5 Regression coefficients, R^2 values and adjusted R^2 values of RSM models for water extraction process of SMLHI

指标	酚酸提出率	迷迭香酸提出率	总糖提出率	总固提出率	总酚酸纯度
常量	2.940	0.120	29.63	39.37	7.59
x_1	-0.005 (0.843 6) ^{&}				
x_3	0.120 (0.000 4)	0.011	3.08	3.72	-0.46
x_5	-0.240	-0.005 (0.018 5)	-4.22	-5.16	0.44
x_7	0.140	0.010		1.03 (0.020 8)	0.18 (0.021 0)
x_3x_7		0.004 (0.061 8)			
x_1x_7	0.067 (0.037 4)				
x_5x_7	0.067 (0.071 7)	0.007 (0.003 5)			0.17 (0.081 5)
x_3^2			-2.06 (0.070 3)		
x_5^2			-2.58 (0.012 4)	-1.72 (0.027 1)	0.26 (0.047 4)
R^2	0.82	0.75	0.82	0.89	0.75
R^2_{adj}	0.77	0.68	0.79	0.87	0.68

[&]相应项的 P 值

[&]numbers in parentheses were P levels of parameters, respectively

响。提取时间和溶剂倍量对各成分提出率的相关系数都为正值，说明随着提取时间和溶剂用量增加各成分提出率都能增大，但延长提取时间会降低总酚酸纯度。药材规格对各成分提出率的相关系数均为负值，说明药材粉碎程度越大，各成分越容易提取出。实验结果与提取传质规律相吻合，溶剂倍量的增大，可以提高药材内外浓度差，增加传质推动力；药材规格越大，即药材比表面积越小，传质阻力就

越大^[33]。

(2) RBF 模型：BRF 模型隐含层节点个数太少模型会欠拟合，而节点个数过多会造成过拟合，因此对隐含层节点数进行了优化。随机选取优化实验设计中 28 组作为校正集，7 组作为验证集，各指标模型的 R^2 和交叉验证均方根误差 (root mean square error of cross validation, RMSECV) 随节点个数的变化趋势如图 2 所示。

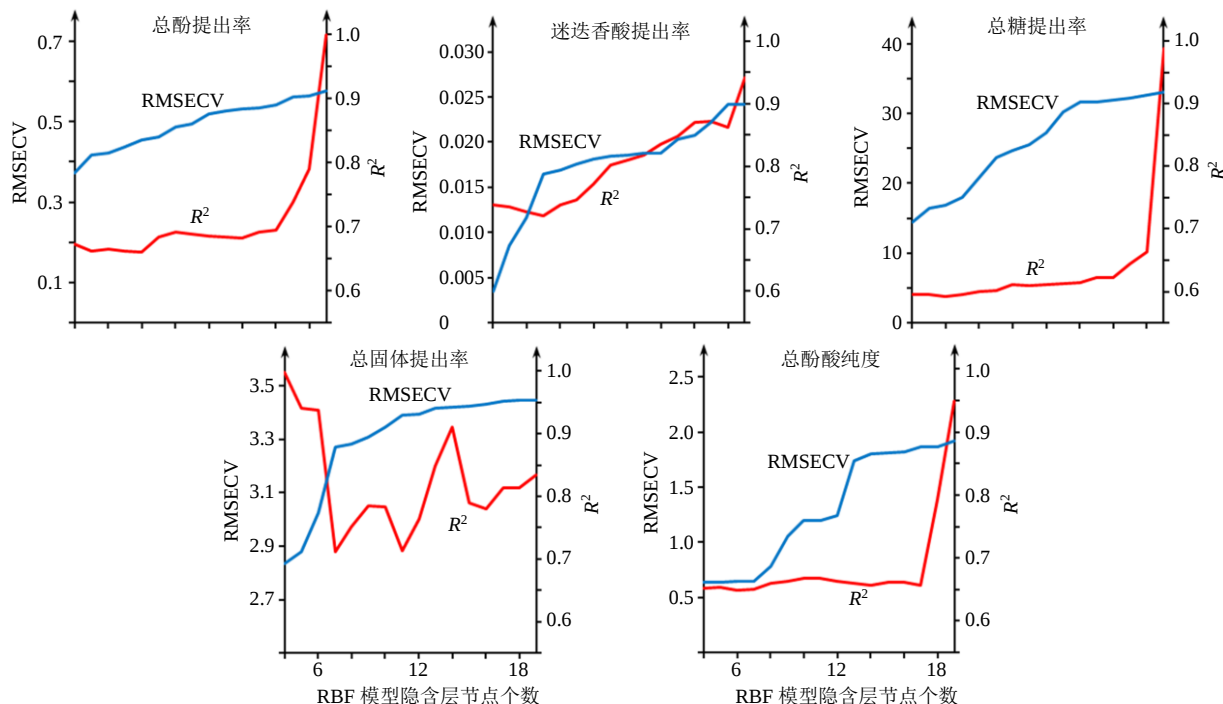


图 2 各 CQA 的 RBF 模型 R^2 和 RMSECV 随隐含层节点个数变化趋势

Fig. 2 Variation trend of R^2 values and RMSECV of RBF models for seven CQAs vary with number of hidden layer nodes

依据图 2 结果, 最终选择酚酸提出率、迷迭香酸提出率、总糖提出率、总固提出率和总酚酸纯度 RBF 模型的隐含层节点个数依次为 12、11、6、10、14, 并用 7 组验证集实验对训练好的 RBF 模型进行外部验证, 结果如表 6 所示。

从表 6 中可以发现, 单克药材中的酚酸提出率、

迷迭香酸提出率、总糖提出率、总固体提出率 RBF 模型外部验证均方根误差 (root mean square error of prediction, RMSEP) 依次为 0.260 6%、0.013 1%、4.859 4%、2.861 5%, 总酚酸纯度外部验证均方根误差为 0.419 3%, 结果均较好, 表明 RBF 可以较好的应用于丹参水提过程的设计建模分析, 且具有良好的预测性能。

2.5.4 RSM 和 RBF 模型比较 本研究以 R^2 和 MRE 为模型质量评价指标对 2 种模型进行比较分析。28 组校正集的 RSM 和 RBF 模型预测值和实测值关系见图 3, 各指标模型的 R^2 和预测 MRE 见表 7。容易发现, 除了总糖提出率模型 R^2 较 RSM 模型较差以外, RBF 模型性能均较优于 RSM 模型。因此, 本研究在所建 RBF 模型的基础上进行后续多指标优化。

表 6 RBF 模型性能评估结果
Table 6 Performance of RBF models

指标	隐含层个数	R^2	RMSECV	RMSEP
酚酸提出率	12	0.88	0.216 3	0.260 6
迷迭香酸提出率	11	0.82	0.017 5	0.013 1
总糖提出率	6	0.74	3.742 1	4.859 4
总固体提出率	10	0.91	3.048 8	2.861 5
总酚酸纯度	14	0.87	0.607 2	0.419 3

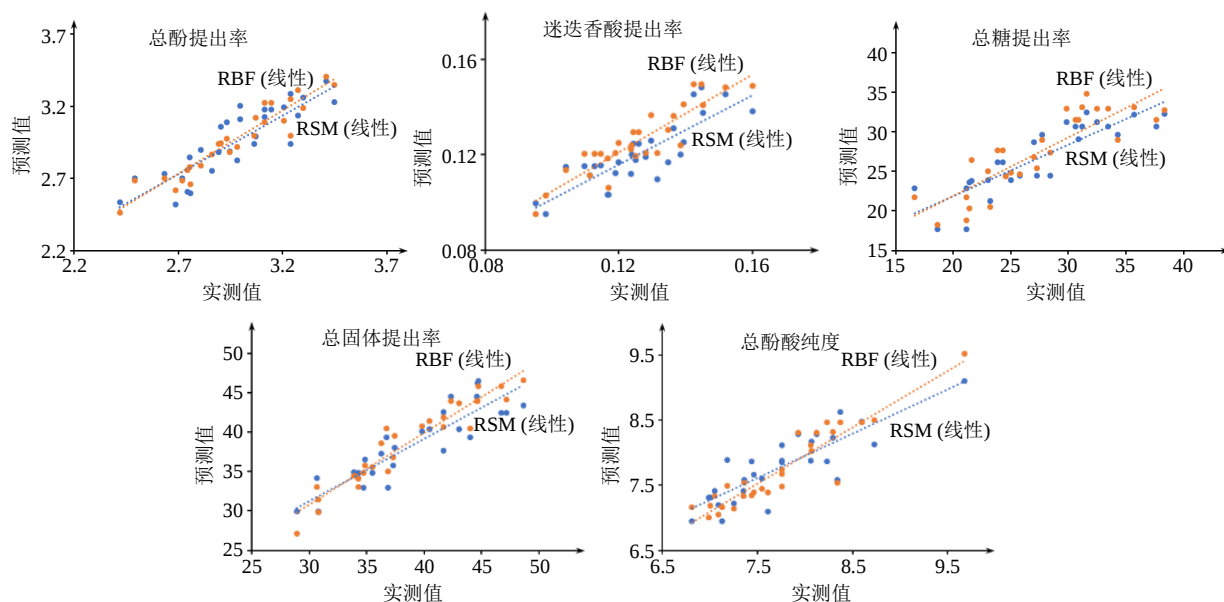


图 3 RSM 和 RBF 模型实测值和预测值关系

Fig. 3 Relationship between measured and predicted values of RSM and RBF models

表 7 水提工艺的 RSM 和 RBF 模型的 R^2 和 MRE
Table 7 R^2 values and MRE values of RSM and RBF models for water extraction of SMLHI

指标	RSM		RBF	
	R^2	MRE/%	R^2	MRE/%
酚酸提出率	0.82	2.878 5	0.88	1.896 8*
迷迭香酸提出率	0.75	4.961 0	0.82	3.407 9*
总糖提出率	0.82	7.117 6	0.74	6.716 4*
总固体提出率	0.89	3.898 7	0.91	2.720 9*
总酚酸纯度	0.75	2.843 9	0.87	1.753 1*

“*” 处相应指标代表模型具有较好的预测性能

Indexes with “*” in the table represented better model performances

2.6 基于满意度函数的最佳操作条件确定

根据生产情况和经验, 设定酚酸提出率、迷迭香酸提出率、总酚酸纯度需要满足的最小值以及总糖提出率、总固提出率需要满足的最大值, 并计算在操作范围内各指标能达到的最大值或最小值。在本研究中, 为了保证产品药效活性, 赋予有效成分较大权重。各指标设定的最值、权重、满意度计算公式见表 8, 在此基础上, 计算 RBF 模型的 SMLHI 水提工艺最佳操作条件。具有最佳满意度的操作参数为加热温度 150 $^{\circ}\text{C}$, 保温时间 84 min, 药材规格 11 cm, 溶剂倍量为 10; 满意度为 0.595 9。

表 8 水提工艺 CQA 范围及权重

Table 8 Limits and weights of CQA for water extraction process

评价指标	最小值/%	最大值/%	权重	评价指标满意度计算公式
酚酸提出率	2.500 0	/	3	$d_{Y_n} = (Y - Y_{\min}) / (Y_{\max} - Y_{\min})$
迷迭香酸提出率	0.100 0	/	2	
总酚酸纯度	7.00	/	2	
总糖提出率	/	30.00	1	$d_{Y_n} = (Y_{\max} - Y) / (Y_{\max} - Y_{\min})$
总固体提出率	/	40.00	1	

$Y_{\max}$ 、 $Y_{\min}$ 分别为计算所得操作范围内指标能达到的最大值、最小值

$Y_{\max}$, $Y_{\min}$ means the maximum and minimum value that an index could achieve

3 讨论

本研究在 QbD 理念的指导下,系统地考察了 SMLHI 生产过程的丹参水提工艺。首先,通过 PB 实验设计及加权回归系数法筛选出了影响丹参水提的 CPP。接着,为了提高实验效率与操作精度,采用基于裂区 D-最优实验设计法进行了优化实验设计,并通过 RSM、RBF 2 种建模方法建立了 CPP 和 CQA 间的定量关系,2 种模型的考察结果均较好。与 RSM 模型相比,训练所得的 BRF 模型整体性能较优,表明即使在数据量较小、实验运行次数有限的情况下,ANN 也能表现出较好的拟合和预测能力。但是,RSM 模型较 ANN 模型能够更加直观的表达工艺参数对指标的影响。最后,本研究使用多指标优化算法综合考察了丹参提取液质量,并基于较优的 ANN 模型确定了 SMLHI 水提工艺最佳的生产条件。

传统的中药制药过程工艺优化研究过程中,当工艺参数的改动受限或更改成本高昂时,为充分发掘因子间的交互作用,需要引入特定的研究方法。本研究在实验设计过程中利用了基于裂区设计的优化方法,有效避免研究工作中工艺参数的改动受限,通过该设计方法得到合理的实验方案,有助于在降低研究成本的同时仍能对过程实现精准优化。此外,本研究将实验设计与 ANN、多指标优化算法相结合,更好的建立了 SMLHI 提取工艺 CPP 和 CQA 之间的关系,并确定了丹参水提液质量最优的生产条件。本研究能够帮助提高研究效率并降低研究成本,对不同试验规模下中药制药工艺的过程研究具有较大参考价值,为中药制药工艺的研究与二次开发提供了新的参考方法。

利益冲突 所有作者均声明不存在利益冲突

参考文献

[1] Yu L X, Amidon G, Khan M A, et al. Understanding

pharmaceutical quality by design [J]. *Aaps J*, 2014, 16(4): 771-783.

- [2] 龚行楚, 陈滕, 瞿海斌. 基于质量源于设计理念的中成药二次开发研究进展 [J]. *中国中药杂志*, 2017, 42(6): 1031-1036.
- [3] 陈泽麒, 赵芳, 袁玮, 等. 基于质量源于设计(QbD)理念研究参麦注射液醇提水沉工艺 [J]. *中草药*, 2020, 51(17): 4433-4441.
- [4] 熊皓舒, 田埂, 刘朋, 等. 中药生产过程质量控制关键技术研究进展 [J]. *中草药*, 2020, 51(16): 4331-4337.
- [5] Draper N R, John J A. Response surface designs where levels of some factors are difficult to change [J]. *Aust NZ J Stat*, 1998, 40(4): 487-495.
- [6] Jones B, Nachtsheim C J. Split-plot designs: What, why, and how [J]. *J Qual Technol*, 2009, 41(4): 340-361.
- [7] Bingham D R, Sitter R R. Design issues in fractional factorial split-plot experiments [J]. *J Qual Technol*, 2001, 33(1): 2-15.
- [8] Kowalski S M, Cornell J A, Vining G G. Split-plot designs and estimation methods for mixture experiments with process variables [J]. *Technometrics*, 2002, 44(1): 72-79.
- [9] Måge I, Næs T. Split-plot design for mixture experiments with process variables: A comparison of design strategies [J]. *Chemom Intell Lab Syst*, 2005, 78(1/2):81-95.
- [10] Jones B, Goos P. A candidate-set-free algorithm for generating D-optimal split-plot designs [J]. *J R Stat Soc Ser C Appl Stat*, 2007, 56(3): 347-364.
- [11] Baria B, Upadhyay N, Singh A K, et al. Optimization of 'green' extraction of carotenoids from mango pulp using split plot design and its characterization [J]. *LWT*, 2019, 104:186-194.
- [12] Drosou K, Koukouvinos C. Supersaturated split-plot designs for robust parameter experiments [J]. *J Qual Technol*, 2020, 52(3): 249-264.
- [13] Jensen W A. Response surface methodology: Process and product optimization using designed experiments 4th

- edition [J]. *J Qual Technol*, 2017, 49(2): 186-188.
- [14] 李文竹, 张禄权, 李白玲, 等. 基于质量源于设计理念的丹参浓缩膏石硫工艺优化研究 [J]. 中草药, 2019, 50(10): 2302-2311.
- [15] 赵芳, 苟维, 刘飞, 等. 基于中心复合序贯设计法丹参川芎嗪注射液水沉工艺优化 [J]. 中草药, 2020, 51(1): 51-58.
- [16] Ameer K, Bae S W, Jo Y, et al. Optimization of microwave-assisted extraction of total extract, stevioside and rebaudioside-A from *Stevia rebaudiana* (Bertoni) leaves, using response surface methodology (RSM) and artificial neural network (ANN) modelling [J]. *Food Chem*, 2017, 229: 198-207.
- [17] Desai K M, Survase S A, Saudagar P S, et al. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan [J]. *Biochem Eng J*, 2008, 41(3): 266-273.
- [18] 陈丽冰, 何容, 李辉, 等. 恤彤注射液中间体丹参提取液的指纹图谱分析及多指标含量测定 [J]. 中国现代应用药学, 2020, 37(9): 1085-1089.
- [19] 李琳. 丹参川芎嗪注射液在脑血栓临床治疗中的作用 [J]. 中国医药指南, 2020, 18(20): 61-62.
- [20] 许珍晶. 丹参川芎嗪注射液治疗后循环缺血性眩晕的临床评价 [J]. 药物评价研究, 2017, 40(9): 1334-1337.
- [21] 中国药典 [S]. 一部. 2015: 76-77.
- [22] Li W, Polachi N, Wang X, et al. A quality marker study on salvianolic acids for injection [J]. *Phytomedicine*, 2018, 44: 138-147.
- [23] Liu C, Guo D A, Liu L. Quality transitivity and traceability system of herbal medicine products based on quality markers [J]. *Phytomedicine*, 2018, 44: 247-257.
- [24] 刘爽悦, 沈金晶, 李文龙, 等. 3 种关键工艺参数辨识方法的比较研究 [J]. 中草药, 2016, 47(18): 3193-3198.
- [25] 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.
- [26] 沈金晶, 龚行楚, 潘坚扬, 等. 基于质量源于设计理念的金银花水提液石灰乳沉淀工艺优化研究 [J]. 中国中药杂志, 2017, 42(6): 1074-1082.
- [27] 陈明. MATLAB 神经网络原理与实例精解 [M]. 北京: 清华大学出版社, 2013: 191-199.
- [28] Elfighi F M. A hybrid statistical approach for modeling and optimization of RON: A comparative study and combined application of response surface methodology (RSM) and artificial neural network (ANN) based on design of experiment (DOE) [J]. *Chem Eng Res Des*, 2016, 113: 264-272.
- [29] 王萍, 王宇鹤, 贺新怀, 等. PB 试验结合 Box-Behnken 响应面法优化丹参水提液中丹酚酸 B 提取工艺的研究 [J]. 中医药导报, 2016, 22(23): 55-59.
- [30] Costa N R, Lourenço J, Pereira Z L. Desirability function approach: A review and performance evaluation in adverse conditions [J]. *Chemom Intell Lab Syst*, 2011, 107(2): 234-244.
- [31] Pan J, Gong X, Qu H. Quantitative ^1H NMR method for hydrolytic kinetic investigation of salvianolic acid B [J]. *J Pharm Biomed Anal*, 2013, 85: 28-32.
- [32] 潘坚扬. 丹酚酸类物质降解规律研究 [D]. 杭州: 浙江大学, 2019.
- [33] 张晋. 五种根茎类药材黄酮类成分提取动力学模型的研究 [D]. 北京: 北京中医药大学, 2014.

[责任编辑 郑礼胜]