

基于 TOPSIS 法的前馈控制在丹参醇提液均化中的应用

陈兆昱, 金伟锋, 万海同, 何 昱*

浙江中医药大学药学院, 浙江 杭州 311402

摘 要: 目的 基于理想解排序 (technique for order preference by similarity to an ideal solution, TOPSIS) 法的前馈控制实现不同批次丹参醇提液的均化。方法 首先采用积分球漫反射收集各批次丹参药材的近红外光谱数据; 然后在不同批次丹参 8 种有效成分 (丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A、迷迭香酸) 提取工艺单因素和含量测定实验的基础上, 以提取时间、提取温度、乙醇体积分数为数值因素、批次为分类因素, 8 种有效成分含量经 TOPSIS 法综合评价后的得分为响应值, 进行 I-最优实验设计; 最后建立以近红外数据, 工艺参数为输入变量, 提取液有效成分含量的综合得分为输出变量的遗传神经网络模型, 以各批次丹参的近红外数据调整相应的醇提工艺参数, 使得不同批次丹参醇提液中 8 种有效成分的含量差异减小。结果 通过建立的模型调整相应的提取工艺使得 4 个批次丹参醇提液中各成分含量 RSD 值的平均值 24.9% 降低至 13.5%。结论 基于 TOPSIS 法的前馈控制能提升丹参醇提液质量的一致性, 为中药提取液的均化研究提供了参考。

关键词: 前馈控制; TOPSIS 法; 遗传神经网络; 丹参; 近红外光谱; 中药均化; 丹参素; 原儿茶醛; 丹酚酸 B; 丹酚酸 A; 丹参酮 I; 隐丹参酮; 丹参酮 II_A; 迷迭香酸

中图分类号: R283.6 文献标志码: A 文章编号: 0253-2670(2022)08-2302-10

DOI: 10.7501/j.issn.0253-2670.2022.08.006

Application of feedforward control based TOPSIS method to homogenize alcohol extraction process of *Salvia miltiorrhiza*

CHEN Zhao-yu, JIN Wei-feng, WAN Hai-tong, HE Yu

College of Pharmaceutical Sciences, Zhejiang Chinese Medical University, Hangzhou 311402, China

Abstract: Objective The feedforward control based on the technique for order preference by similarity to an ideal solution (TOPSIS) method is used to homogenize alcohol extracts of different batches of Danshen (*Salvia miltiorrhiza*). **Methods** Firstly, the near-infrared diffuse reflectance spectra of *S. miltiorrhiza* were obtained. Then a high-performance liquid chromatography method had been developed for the simultaneous quantification of eight active components (danshensu, protocatechualdehyde, salvianolic acid B, salvianolic acid A, tanshinone I, cryptotanshinone, tanshinone II_A, rosmarinic acid) contained in alcohol extracts of *S. miltiorrhiza*. On the basis of single factor experiment and content determination experiment, extraction time, extraction temperature, ethanol concentration were taken as numerical factors, the batch was taken as categorical factors, eight control targets in the alcohol extract which were comprehensively evaluated by TOPSIS method was taken as response of I-optimal experimental design for the experiment. Finally, a genetic neural network model which reflected the relationship between near-infrared data, process parameters and the effective components was established. Based on the established model, the extraction process was adjusted according to the near-infrared data to decrease the difference in eight active components caused by batch change. **Results** Based on the established models, extraction conditions were optimized according to near-infrared data. The RSD value of each component content in alcohol extract of *S. miltiorrhiza* was reduced from 24.9% to 13.5%. **Conclusion** The feedforward control based on the TOPSIS method has ability to improve the quality and consistency of the alcohol extract of *S. miltiorrhiza*. It provides a reference for research on the homogenization of Chinese medicine extraction process.

收稿日期: 2021-11-04

基金项目: 国家自然科学基金资助项目 (81873226); 国家科技重大专项重大新药创制 (2019ZX09301101); 浙江省“万人计划”科技创新领军人才项目 (2019)

作者简介: 陈兆昱, 硕士研究生, 主要从事中药有效成分提取分离方面研究。E-mail: 13626535279@163.com

*通信作者: 何 昱, 教授, 博士生导师。Tel: (0571)61768145 E-mail: heyu0923@sina.com

Key words: feedforward control; TOPSIS method; *Salvia miltiorrhiza* Bge.; genetic neural network; near infrared spectroscopy; homogenization of traditional Chinese medicine; danshensu; protocatechualdehyde; salvianolic acid B; salvianolic acid A; tanshinone I; cryptotanshinone; tanshinone II_A; rosmarinic acid

在满足一定质量要求的前提下,减少中药材批次变化带来的质量差异是中药提取液均化的主要目标。前馈控制是一种根据扰动量进行补偿的开环控制系统,能够明显减少由于物料变化带来的产品质量波动^[1-3],因此前馈控制在中药提取液均化方面的应用有较高的研究价值。目前已有将近红外光谱(near-infrared spectroscopy, NIRS)引入前馈控制中,减少产地变化带来提取液质量波动的报道^[4],但并未针对中药提取液成分复杂、有效成分多的特点进行研究。

理想解排序法(technique for order preference by similarity to an ideal solution, TOPSIS)是一种依据评价对象与理想解接近程度进行评分的综合评价法,在中药质量评价方面已有广泛的使用^[5-8]。本研究基于 TOPSIS 法构建前馈控制模型,针对提取液与理想解的接近程度进行控制,通过优化关键工艺的方式提升提取液与理想解的接近程度,从而达到对成分复杂的中药提取液进行控制的目的。

丹参为唇形科鼠尾草属植物丹参 *Salvia miltiorrhiza* Bge.的干燥根和根茎,有活血祛瘀、通经止痛、凉血消痛的功效^[9]。其主要有效成分可分为水溶性酚酸类和脂溶性丹参酮类 2 大类,是常见的活血化瘀药,常用于治疗心血管疾病^[10-11]。

在实验过程中,首先收集各批次丹参的 NIRS 数据,作为前馈控制模型的协变量,然后建立了丹参中水溶性有效成分(丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、迷迭香酸)和脂溶性有效成分(丹参酮 I、隐丹参酮、丹参酮 II_A)的定量方法,并对各个批次的丹参进行含量测定和提取工艺单因素实验。最后在此基础上进行 I-最优实验设计,建立了以近红外数据为协变量,提取工艺为可控变量,丹参醇提液中 8 种有效成分含量经 TOPSIS 法计算后的综合得分为因变量的遗传神经网络模型。根据各批次丹参近红外数据对提取工艺进行相应的优化,以减少不同批次丹参醇提液中 8 种有效成分的含量差异,使丹参醇提液质量均化。

1 材料

1.1 仪器

Agilent 1260 infinity 型高效液相色谱仪,配备 DAD 检测器,美国 Agilent 公司;FA1104N 十万分

之一电子分析天平,上海精密科学仪器有限公司;超纯水仪,美国 Millipore 公司;Antaris II 型傅里叶变换近红外光谱仪,美国 Thermo Fisher 公司。

1.2 药材及试剂

不同批次的丹参药材购自各地药房以及药材市场,详见表 1,并经浙江中医药大学药学院张水利教授鉴定为唇形科植物丹参 *S. miltiorrhiza* Bge.的干燥根和根茎。

对照品丹参素、原儿茶醛、丹酚酸 A、丹参酮 I、隐丹参酮、迷迭香酸、丹酚酸 B、丹参酮 II_A均购自成都德思特生物技术有限公司,批号分别为 DSTDD001501、DST201109-008、DST201022-013、DST200511-011、DSTDM002701、DSTDY008001、DSTDD000902、DSTDD001102,质量分数均大于 98%;甲醇、乙腈为色谱纯,美国 Tedia 公司;甲酸、

表 1 不同批次丹参的来源、产地、批号

Table 1 Information of medicinal materials of *S. miltiorrhiza*

编号	批号	产地	来源
S1	2020456	山东	安徽省亳州市中药材市场
S2	2013201	山东	安徽省亳州市中药材市场
S3	20101601	山东	浙江景岳堂药业有限公司
S4	20102501	山东	浙江景岳堂药业有限公司
S5	200714	山东	浙江震元医药连锁有限公司杭
S6	200918	山东	杭州方回春堂国药馆有限公司
S7	200802	山东	浙江中医药大学饮片厂
S8	2011250	山东	爱珍药品零售有限公司
S9	20210103	山东	河北康典药业有限公司
S10	201101471	山东	成都康美药业生产有限公司
S11	201120	山东	安徽广和中药股份有限公司
S12	200801	山东	普宁市泽群中药饮片有限公司
S13	221405	山西	安徽省亳州市中药材市场
S14	223021	山西	安徽省亳州市中药材市场
S15	200815	山西	杭州凝和堂中医门诊部有限公司
S16	2201026	山西	杭州太和堂中医门诊部有限公司
S17	2201029	山西	杭州民泰中药饮片公司
S18	2005021	四川	四川一片叶药业
S19	2009023	四川	通化市同福参茸制品有限公司
S20	203010	安徽	安徽省亳州市中药材市场
S21	203111	安徽	安徽省亳州市中药材市场
S22	209041	河南	安徽省亳州市中药材市场

95%乙醇均为分析纯。

2 方法与结果

2.1 各批次丹参药材 NIRS 的采集

取上述 22 批丹参药材粉碎, 过 80 目筛, 40 °C 下干燥 24 h 后, 均匀填充于近红外光谱仪的样品杯中, 采用积分球漫反射测定。以空气为参比, 分辨率为 8 cm⁻¹, 扫描次数 64 次, 扫描范围 4000~10 000 cm⁻¹, 每个样品测定 3 次, 求平均光谱。

为了避免 NIRS 中存在的冗余信息对模型的干扰, 在应用 NIRS 数据前多对其进行预处理及波长筛选^[12-13]。本研究采用 TQ Analyst 软件 (版本 9.5.0.76, 美国 Thermo Fisher 公司) 中自带的标准正则变化 (standard normal variate, SNV) 和多元散射校正 (multiplicative scatter correction, MSC) 2 种光程矫正方法组合一阶导数 (first derivative, 1st) 和二阶导数 (second derivative, 2nd) 2 种光谱格式, 共产生 4 种不同的预处理方式。此外还需将处理后的光谱数据采用 S-G 滤波平滑 (Svaitzky-Golay smoothing) 以消除导数引入的光谱噪音。

波长筛选方面, 连续投影法 (successive projections algorithm, SPA) 能充分寻找含有最低限度的冗余信息的变量组, 使得变量之间的共线性达到最小。同时能减少建模所用变量的个数, 提高建模的速度和效率, 故本研究采用连续投影法对预处理后的 NIRS 筛选波长。经过筛选后共将预处理和波长筛选后的 NIRS 数据进行主成分分析 (principal component analysis, PCA), 以主成分得分来反映药材的质量信息^[14-15], 用于遗传神经网络模型的构建。

为了对 4 种不同的预处理方法进行筛选, 本研究采用遗传神经网络模型中训练集和验证集的最大决定误差 (maximum absolute error, MAE)、决定系数 (determination coefficient, R^2) 及相关系数 (correlation coefficient, r) 值确定预处理方法, 以训练集和验证集 MAE 趋近于 0、 R^2 、 r 趋近于 1 为

优, 其中神经元数量选为 6, 主成分数量选为 2, 具体数据见表 2。根据表 2, 本研究决定采用 SNV+1st 作为 NIRS 数据的预处理方法, 其通过连续投影法筛选出的波长为 4 072.92、4 420.05、4 481.76、4 844.31、4 936.87、5 075.72、5 260.86、5 268.57、7 062.04、9 881.46 cm⁻¹。除预处理方式外, 主成分数量也是一个需要考量的因素。主成分数量越多, 累计方差贡献率越高, 其能解释的光谱信息也越多。但随着主成分数量增多, 模型需输入的变量也随之增多, 使得工艺寻优的难度增大。

表 2 预处理方法对训练集和验证集 MAE、 R^2 、 r 的影响
Table 2 Effect of pretreatment method on MAE, R^2 and r of training data and testing data

预处理方法	R^2_{train}	MAE _{train}	r_{train}	R^2_{test}	MAE _{test}	r_{test}
无预处理	0.741 4	0.029 6	0.862 8	0.834 2	0.018 5	0.925 7
SNV+1st	0.995 6	0.001 9	0.997 8	0.968 3	0.012 9	0.991 8
SNV+1st	0.996 5	0.001 8	0.998 3	0.940 3	0.019 8	0.970 6
SMC+2nd	0.994 7	0.003 1	0.997 3	0.929 2	0.021 2	0.964 2
SMC+2nd	0.964 4	0.007 8	0.982 0	0.966 8	0.012 7	0.992 1

因此, 本研究在选取主成分数量时, 采用不同主成分数量所建立模型的 MAE、 R^2 、 r 3 个指标和累计方差贡献率选择主成分数目, 结果见表 3。由表 3 可知, 当主成分数量 2 个及以上时模型有较好拟合度, 但考虑到主成分数量对后续计算难度的影响, 本研究最终选择主成分数量为 2 来反映药材的质量信息。此时 PCA 的累积方差贡献率为 72%, 能反映大部分的 NIRS 数据信息。

2.2 不同批次丹参中 8 种成分含量测定的方法学考察

2.2.1 混合对照品溶液的制备 分别精密称取丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A、迷迭香酸对照品适量, 加入 80% 甲醇制备成质量浓度分别为 0.340、0.110、3.100、0.340、0.300、0.560、0.400、0.260 mg/mL 的对照品储备液。精确吸取各对照品溶液适量, 配

表 3 主成分数量对训练集和验证集 MAE、 R^2 、 r 的影响

Table 3 Effect of number of principal components on MAE, R^2 and r of training data and testing data

主成分数量	累积方差贡献率/%	R^2_{train}	MAE _{train}	r_{train}	R^2_{test}	MAE _{test}	r_{test}
1	43	0.938 9	0.010 6	0.968 9	0.900 4	0.024 3	0.949 0
2	72	0.995 6	0.001 9	0.997 8	0.968 3	0.012 9	0.991 8
3	86	0.995 7	0.002 8	0.997 8	0.968 0	0.013 1	0.996 5
4	93	0.996 5	0.001 8	0.998 3	0.962 3	0.015 9	0.984 4

制成 8 种成分质量浓度分别为 0.257、0.004、0.906、0.019、0.014、0.063、0.060、0.059 mg/mL 的混合对照品溶液。

2.2.2 供试品溶液的制备 精密称取干燥后的丹参粉末 5 g (过五号筛), 置于圆底烧瓶中, 以 1:20 的液料比精密加入 70%乙醇, 浸泡 30 min 后, 在 80 °C 下加热回流 90 min, 放冷, 滤过, 续滤液移至 250 mL 量瓶中加 70%乙醇定容, 摇匀。为避免溶剂效应, 进样前取供试品 0.5 mL 与 0.5 mL 纯水混合, 过 0.22 μm 微孔滤膜。

2.2.3 色谱条件 色谱条件为 Agilent Zorbax SB-C₁₈ 色谱柱 (250 mm×4.6 mm, 5 μm) 及 Agilent Zorbax SB-C₁₈ 保护柱 (12.5 mm×4.6 mm, 5 μm), 流动相为 0.4%甲酸水溶液(A)-乙腈(B)-甲醇(C), 梯度洗脱: 0~8 min, 93%~80% A, 7%~20% B; 8~22 min, 80%~76% A, 20%~24% B; 22~28 min, 76%~70% A, 24%~30% B; 28~33 min, 70%~30% A, 30%~40% B, 0~30% C; 33~51 min, 30% A, 40% B, 30% C; 51~60 min, 30%~20% A, 40%~50% B, 30% C; 60~65 min, 20%~93% A, 50%~7% B, 30%~0 C; 柱温 30 °C; 体积流量 1.0 mL/min; 进样量 20 μL; 检测波长 280 nm (丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A) 及 330 nm (迷迭香酸)。色谱图见图 1。

2.2.4 线性关系考察 取“2.2.1”项下混合对照品溶液, 用 80%甲醇逐级稀释, 得到梯度质量浓度的混合对照品溶液, HPLC 进样检测, 以峰面积 (Y) 对各对照品溶液质量浓度 (X) 绘制标准曲线, 建立回归方程, 结果分别为丹参素 $Y=8.812 X+5.601$, $r=0.999 1$, 线性范围 0.513~25.660 μg/mL; 原儿茶醛 $Y=52.946 X+3.290$, $r=0.999 3$, 线性范围 0.083~4.151 μg/mL; 丹酚酸 B $Y=18.784 X+287.420$, $r=0.999 6$, 线性范围 18.113~906.660 μg/mL; 丹酚酸 A $Y=34.024 X+2.571$, $r=0.999 4$, 线性范围 0.385~19.245 μg/mL; 丹参酮 I $Y=35.350 X+6.594$, $r=0.999 2$, 线性范围 0.287~14.340 μg/mL; 隐丹参酮 $Y=57.383 X+47.884$, $r=0.999 7$, 线性范围 1.268~63.396 μg/mL; 丹参酮 II_A $Y=66.310 X+48.884$, $r=0.999 6$, 线性范围 1.207~60.337 μg/mL; 迷迭香酸 $Y=40.714 X+50.197$, $r=0.999 5$, 线性范围 1.177~58.870 μg/mL。

2.2.5 精密度试验 取同一混合对照品溶液, 按

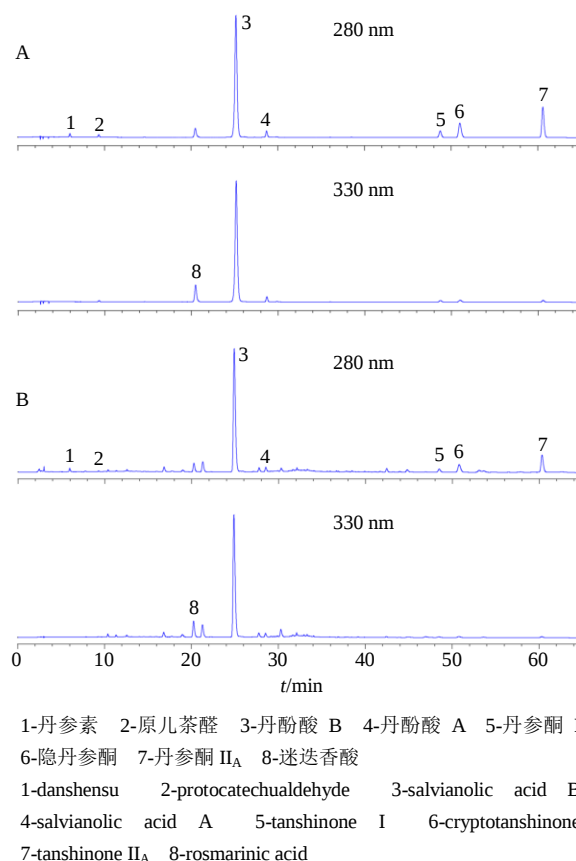


图 1 混合对照品溶液 (A) 和供试品溶液 (B) 的 HPLC 图
Fig. 1 HPLC chromatograms of mixed standards solution (A) and *S. miltiorrhiza* sample solution (B)

“2.2.3”项色谱条件检测, 连续进样 6 次, 记录峰面积, 结果丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A、迷迭香酸 8 个组分峰面积的 RSD 分别为 2.17%、1.82%、2.00%、2.51%、1.28%、2.01%、1.92%、1.79%, 表明仪器精密度良好。

2.2.6 稳定性试验 取同一供试品溶液 (S20) 分别于制备 0、2、4、8、12、24 h 后进样, 按“2.2.3”项色谱条件检测, 记录峰面积。结果丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A、迷迭香酸 8 个组分峰面积的 RSD 值分别为 2.80%、1.61%、1.73%、1.98%、2.43%、1.29%、1.50%、1.62%, 表明供试品溶液在 24 h 内稳定。

2.2.7 重复性试验 取相同批次的丹参粉末按“2.2.2”项方法平行制备 6 组供试品溶液, 按“2.2.3”项色谱条件检测, 结果丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A、迷迭香酸 8 个成分的质量分数分别是 0.730、0.016、32.013、0.306、0.470、0.527、1.043、2.038 mg/g,

其 RSD 分别为 2.73%、2.21%、2.02%、2.43%、2.36%、1.63%、2.55%、1.67%，说明该方法重复性良好。

2.2.8 加样回收率试验 取已测定 8 个成分质量分数的 S20 的丹参粉末 0.5 g，精密加入与粉末中各成分等量的对照品溶液，按“2.2.2”项方法平行制备 6 份，按“2.2.3”项色谱条件检测，结果丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A、迷迭香酸 8 个成分的平均加样回收率分别为 98.10%、97.30%、99.30%、97.50%、96.70%、100.80%、103.50%、99.10%，RSD 分别为 2.98%、1.77%、1.81%、2.04%、2.77%、1.73%、2.01%、2.63%。

2.3 单因素实验

选取丹参药材中的丹参素、原儿茶醛、丹酚酸 B、丹酚酸 A、丹参酮 I、隐丹参酮、丹参酮 II_A、迷迭香酸共 8 种有效成分作为丹参醇提液的控制对象。通过预试验发现 8 种成分随提取条件变化的趋势不尽相同，因此本研究通过熵权法赋权求综合得分的方式来综合评价各项提取工艺参数在单因素实验中对这 8 种成分的影响^[16]。熵权法是一种通过计算各指标信息熵反映其变异程度的赋权法，熵值越小，指标变异程度越大，其权重也就越大。通过熵权法求得的综合得分能反映出 8 种成分含量随单因素水平变化的趋势，进而对单因素水平进行筛选，其计算步骤如下。

第 1 步，将各项指标的数据标准化，标准化公式如下，其中 x_{ij} 表示第 i 个样品在第 j 个指标下的数据 ($i=1, 2, \dots, n; j=1, 2, \dots, m$)。

$$x'_{ij} = (x_{ij} - \min\{x_{1j}, \dots, x_{nj}\}) / (\max\{x_{1j}, \dots, x_{nj}\} - \min\{x_{1j}, \dots, x_{nj}\}) \quad (1)$$

第 2 步，计算各样品在第 j 各指标下的比重 p_{ij} 。

$$p_{ij} = x'_{ij} / \sum_{i=1}^n x'_{ij} \quad (2)$$

第 3 步，计算各项指标的熵值 e_j 。

$$e_j = -k \sum_{i=1}^n p_{ij} \ln p_{ij}, \text{ 其中 } k=1/\ln n \quad (3)$$

第 4 步，计算第 j 项指标的信息冗余度 d_j 。

$$d_j = 1 - e_j \quad (4)$$

第 5 步，计算第 j 项指标的权重。

$$w_j = d_j / \sum_{i=1}^m d_j \quad (5)$$

第 6 步，计算各单因素实验中的样品的总得分。

$$S_j = \sum_{i=1}^m d_j w_j x'_{ij} \quad (6)$$

精密称取丹参粉末进行提取，保持其他提取条件不变，分别考察提取时间、提取温度、乙醇体积分数、料液比、浸泡时间 5 个因素对丹参中 8 种有效成分的影响，经熵权法计算后的综合得分见表 4。由表 4 可知，各提取工艺参数的最优值分别为提取温度 90 °C，提取时间 90 min，浸泡时间 30 min，

表 4 各因素不同水平的综合得分

Table 4 Comprehensive scores of different levels in single factor experiment

提取时间/min	总得分	提取温度/°C	总得分	乙醇体积分数/%	总得分	料液比	总得分	浸泡时间/min	总得分
30	0.267	60	0.130	50	0.363	1 : 10	0.038	0	0.000
60	0.279	70	0.180	60	0.759	1 : 20	0.388	30	0.910
90	0.783	80	0.684	70	0.886	1 : 30	0.755	45	0.761
120	0.700	90	0.814	80	0.566	1 : 40	0.739	60	0.833
150	0.432	100	0.308	90	0.365	1 : 50	0.687	75	0.860

乙醇体积分数 70%，料液比 1 : 30。

2.4 不同批次丹参各成分含量的测定

选取单因素实验中各工艺参数的最优点为提取条件，测定不同批次丹参提取液在同一提取条件下各成分的含量。为了直观地反映不同批次丹参之间有效成分的含量差异，为后续实验设计选择批次提供依据，将丹参中水溶性成分和脂溶性成分两大类分别进行熵权法赋权，计算总得分，见图 2。

图 2 横坐标代表丹参中水溶性成分经熵权法赋

权后计算的总得分（丹参素、迷迭香酸、原儿茶醛、丹酚酸 B、丹酚酸 A），纵坐标代表丹参中脂溶性成分（丹参酮 II_A、隐丹参酮、丹参酮 I）经熵权法赋权后计算的总得分。

图 2 中不难看出 S9（水溶性成分含量低而脂溶性成分含量极高），S22、S13（水溶性成分高而脂溶性成分极低），S11（水溶性成分极高）4 个样本偏离中心区域较远。由于前馈控制是通过改变提取条件对提取液质量进行影响的，在提取液质量满足

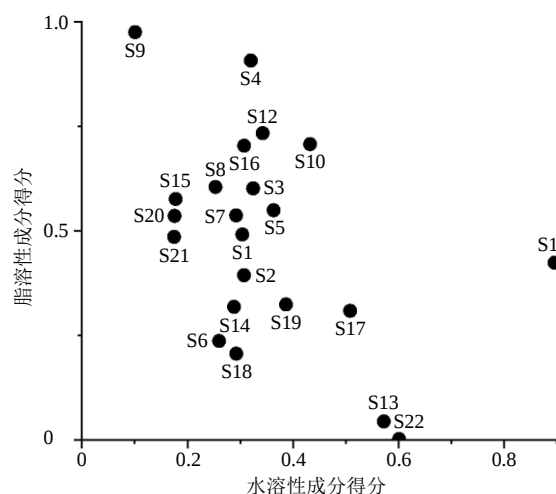


图 2 各组丹参水溶性成分和脂溶性成分得分分布图

Fig. 2 Score distribution of water-soluble components and fat-soluble components in each group of *S. miltiorrhiza*

一定的条件下, 前馈控制起到的效果是有限的, 远不足以弥补这 4 个批次药材本身质量与其他批次的差异, 因此这 4 个批次将不纳入代表批次的选择范围。选择分布较为分散但未与其他批次差异过大的 4 个批次 S8、S10、S19、S21 作为 I-最优实验设计的分类变量进行后续实验设计。4 个批次各成分含量见表 5。

2.5 I-最优实验设计

最优实验设计是一种灵活的实验设计方法, 可以用于不规则设计、含分类因素的设计和自定义模型设计^[17]。本实验选用 I-最优实验设计, 依据单因素实验的结果选取提取温度、提取时间、乙醇浓度为变量, 选择 S8、S10、S19、S21 4 个批次作为分类变量, 以 TOPSIS 分析的结果 *Q* 为响应值进行 I-最优实验设计, 结果见表 6。

表 5 4 个批次丹参的 8 种成分含量

Table 5 Contents of eight components in four batches of *S. Miltiorrhiza*

批次	质量分数/(mg·mL ⁻¹)							
	丹参素	原儿茶醛	丹酚酸 B	丹酚酸 A	丹参酮 I	隐丹参酮	丹参酮 II _A	迷迭香酸
S8	1.103	0.020	34.649	0.514	0.526	0.886	0.989	1.537
S10	1.396	0.040	35.956	0.953	0.330	0.823	1.655	1.662
S19	1.420	0.029	31.520	0.632	0.360	0.426	0.995	2.216
S21	1.052	0.027	29.048	0.514	0.231	0.724	1.236	1.156
RSD/%	15.44	27.76	9.50	31.81	33.92	28.47	25.65	26.72

表 6 丹参提取过程的 I-最优实验设计

Table 6 I-optimal design and results of extraction process of *S. miltiorrhiza*

试验号	提取温度/℃	提取时间/min	乙醇/%	批次	Q 值	试验号	提取温度/℃	提取时间/min	乙醇/%	批次	Q 值
1	100	120	80	S8	0.675	17	80	120	80	S21	0.584
2	80	120	70	S19	0.661	18	100	60	60	S21	0.515
3	80	120	60	S10	0.684	19	80	120	70	S8	0.606
4	100	90	80	S21	0.616	20	90	120	70	S10	0.521
5	90	90	70	S21	0.614	21	100	90	60	S19	0.583
6	90	60	70	S10	0.551	22	100	120	80	S19	0.593
7	90	90	80	S21	0.577	23	80	60	70	S21	0.501
8	90	90	60	S21	0.490	24	100	120	60	S21	0.538
9	90	90	60	S21	0.550	25	100	120	60	S8	0.858
10	80	90	60	S19	0.575	26	100	90	70	S8	0.719
11	90	60	80	S8	0.527	27	80	60	60	S8	0.578
12	100	120	70	S21	0.599	28	90	60	70	S10	0.731
13	80	120	70	S8	0.595	29	90	90	70	S19	0.651
14	100	90	60	S10	0.604	30	80	90	80	S10	0.665
15	90	60	80	S19	0.612	31	100	90	70	S8	0.693
16	100	60	80	S10	0.604	32	90	120	70	S10	0.501

2.6 前馈控制

前馈控制作为一种开环的控制系统，其应用于中药领域首先需表征物料信息，将物料与物料之间的差异数值化，使得物料变化对关键质量属性的影响可以在建立的数学模型中表达出来^[18]。然后在多个批次中收集工艺参数对最终产品关键质量属性影响的数据，建立生产参数、物料信息 2 部分数据为自变量，关键质量属性为因变量的回归模型。最终需要对某一批次的键质量属性进行控制时，依据相应的物料信息，调整可控的生产参数，使得关键质量属性达到控制目标^[19]。

中药中成分复杂，一旦需要关注的成分增多则构建相应设计空间的难度也随之增加。因此本研究采用 TOPSIS 法评价丹参提取液中 8 种成分含量，计算后的得分能综合反映提取液中各成分含量距离理想解的接近程度，该综合得分越大则该组提取液中各成分含量越接近于理想解。依据“2.1”项 NIRS 数据和“2.5”项 I-最优实验设计数据建立以近红外数据，提取工艺为自变量，丹参提取液中各成分含量经 TOPSIS 法计算的综合得分为因变量的人工神经网络模型。将建立的模型用于前馈控制中，根据各批次丹参的近红外数据调整提取工艺，使得各批次的综合得分均有最大值，此过程中各批次丹参提取液成分的含量均向理想解靠近，进而减少各批次提取液之间的质量差异且满足一定的质量要求。本实验应用的 TOPSIS 法具体步骤如下^[20]。

第 1 步，设定各成分的理想解 (x_{best})，本研究 x_{best} 设定为所选择 4 个批次在“2.4”项下测得数据的平均值，根据控制目标将原始数据正向化，公式如下。

$$x'_i = 1 - |x_i - x_{\text{best}}| / M, M = \max\{|x_i - x_{\text{best}}|\} \quad (7)$$

第 2 步，将正向化的数据标准化，标准化公式如下。

$$x''_i = (x'_i - x'_{\text{max}}) / (x'_{\text{max}} - x'_{\text{min}}) \quad (8)$$

设共有 m 个评价指标， n 个评价对象，则标准化后的数据矩阵为

$$Z = \begin{bmatrix} x''_{11} & \cdots & x''_{1m} \\ \cdots & \cdots & \cdots \\ x''_{n1} & \cdots & x''_{nm} \end{bmatrix} \quad (9)$$

第 3 步，确定最佳理想解 Z^+ 和最差理想解 Z^- ，最佳理想解 Z^+ 由每列指标中的最大值组成，最差理想解 Z^- 由每列指标中的最小值组成。

$$Z^+ = (\max\{x''_{11} \cdots x''_{n1}\}, \cdots, \max\{x''_{1m} \cdots x''_{nm}\}) = (Z_1^+ \cdots Z_m^+) \quad (10)$$

$$Z^- = (\min\{x''_{11} \cdots x''_{n1}\}, \cdots, \min\{x''_{1m} \cdots x''_{nm}\}) = (Z_1^- \cdots Z_m^-) \quad (11)$$

第 4 步，计算各评价对象最佳理想解 Z^+ 和最差理想解 Z^- 的距离，其中 D_i^+ 和 D_i^- 分别表示评价指标距离最佳理想解 Z^+ 和最差理想解 Z^- 的距离。

$$D_i^+ = [\sum_{j=1}^m (Z_i^+ - Z_{ij})^2]^{1/2}, D_i^- = [\sum_{j=1}^m (Z_i^- - Z_{ij})^2]^{1/2} \quad (12)$$

第 5 步，计算接近程度 (Q)， Q 为各评价对象距离最佳方案的贴近程度， Q 值越贴近 1 表示评价对象距离理想解越近， Q 值越贴近 0 表示评价对象距离理想解远。

$$Q = D_i^- / (D_i^+ + D_i^-) \quad (13)$$

2.6.1 遗传神经网络模型的建立 遗传神经网络是利用遗传算法找出神经网络的权重和阈值的数学模型，同时具备神经网络非线性拟合能力强和遗传算法收敛速度快和全局性好的优点，在中药提取工艺优化方面已有较多的应用^[21-22]。

本研究遗传神经网络模型由 MATLAB (版本 2016a, 美国 MathWorks 公司) 完成，采用提取工艺 (乙醇体积分数、提取时间、提取温度) 和物料信息 (近红外主成分得分) 为输入变量，最优实验设计中各组经 TOPSIS 法计算后的结果为输出变量，随机抽取 I-最优实验设计中的 29 组为训练集，剩余 3 组为测试集训练模型。

隐藏层神经元数目是决定模型优劣的一个重要因素，隐藏层数目过少则拟合效果差，数目多则模型容易出现“过拟合”现象，因此采用与“2.1”项中筛选近红外数据预处理方法相同的方法即模型的 MAE、 R^2 、 r 值对神经元数目进行筛选，具体数据见表 7，可知神经元数目为 6 个时最优。模型预测值与实测值误差分析如图 3 所示，由图可知模型的预测值和实测值基本一致，模型拟合效果良好。

表 7 神经元数量对的训练集和验证集 MAE、 R^2 、 r 的影响
Table 7 Effect of number of hidden neurons on MAE, R^2 and r of training data and testing data

神经元数目	R^2_{train}	MAE _{train}	r_{train}	R^2_{test}	MAE _{test}	r_{test}
3	0.715 5	0.026 3	0.845 9	0.924 1	0.019 8	0.990 9
4	0.990 0	0.006 5	0.995 0	0.899 7	0.023 3	0.951 5
5	0.995 6	0.001 9	0.997 8	0.957 7	0.016 7	0.983 5
6	0.995 6	0.001 9	0.997 8	0.968 3	0.012 9	0.991 8
7	0.995 6	0.001 9	0.997 8	0.955 4	0.017 2	0.980 1
8	0.995 6	0.001 9	0.997 8	0.958 8	0.016 5	0.982 6

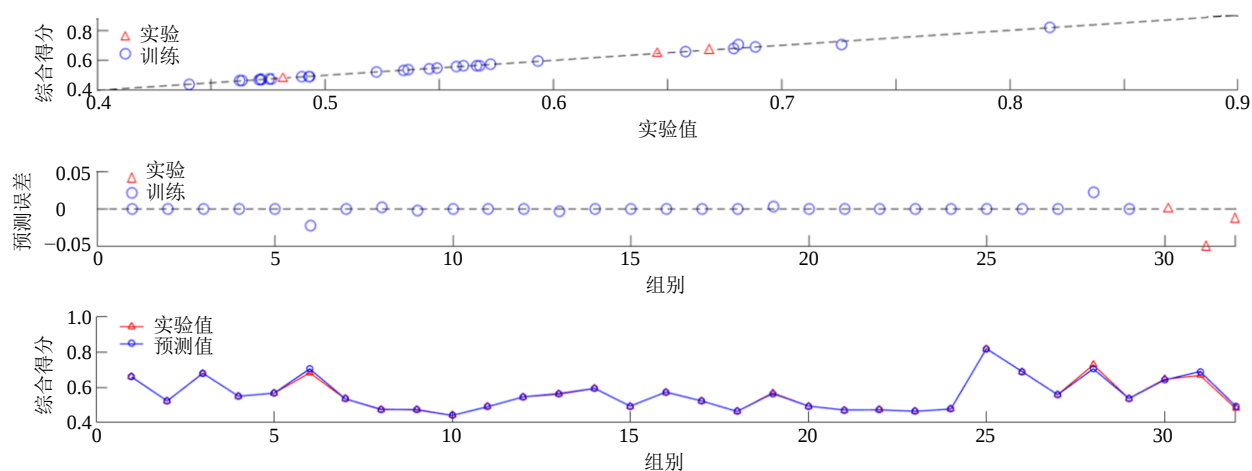


图 3 模型预测值与实测值误差分析

Fig. 3 Error analysis results of experimental data and model prediction data

2.6.2 前馈控制的实施及验证 基于所建立的遗传神经网络模型，采用遗传算法对设定的 4 个批次丹参的提取工艺参数进行优化^[23-24]，算法流程见图 4。设定种群大小为 1000，迭代次数为 200，使得相应

的醇提液的 TOPSIS 综合得分 Q 值最大，寻优结果为 S8：提取温度 98 °C，提取时间 115 min，乙醇体积分数 73%；S10：提取温度 81 °C，提取时间 120 min，乙醇体积分数 60%；S19：提取温度 80 °C，提取时间 120 min，乙醇体积分数 79%；S21：提取温度 99 °C，提取时间 94 min，乙醇体积分数 79%，模型的预测值及实际测定值见表 8。

为了直观地表示参与前馈控制前后批次有效成分含量的变化，仍将丹参分水溶性成分和脂溶性成分 2 大类进行按“2.4”项下权重赋权，计算各自总得分，以散点图的形式呈现，见图 5。由图 5 不难看出 4 个批次丹参药材的醇提液经前馈控制后与未进行前馈控制的相比，水溶性成分和脂溶性成分得分分布明显更加集中，且向目标值靠拢。从具体数据方面分析，比较“2.4”项表 5 与“2.6.2”项表 8 可以看出，4 个批次中 8 种有效成分的 RSD 值均有下降，RSD 的平均值从 24.9% 降到了 13.5%，表明前馈控制效果良好，提取液质量一致性得到了提升。若是按照理想解中各成分含量 $\pm 10\%$ 作为可接受范围，则 4 个批次丹参中 8 种有效成分含量不达标百分比由“2.4”项中占总数的 72% 降到了 38%，表明前馈控制对提取液中各成分含量起到了控制效果。

S8、S10、S21 丹参提取液的最优 Q 值均达到 0.7 以上，表明已经与理想解较为接近，而 S19 的最优 Q 值最大仅为 0.552，仍与理想解有一定距离。从图 5 可以看出 S19 前馈控制前后与理想解均有一定的差距，且主要差距是脂溶性成分方面。具体到成分来看，无论是前馈控制前的表 6 中还是前馈控制后的表 8 中 S19 批次提取液中隐丹参酮含量与理

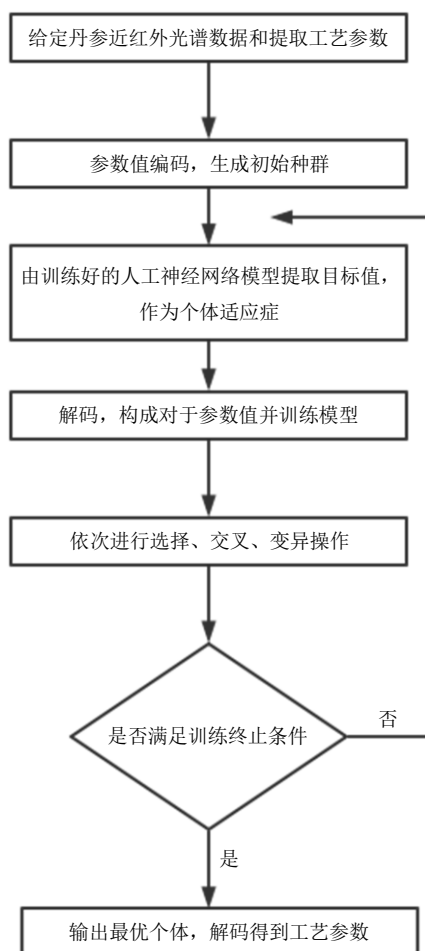


图 4 遗传算法优化工艺流程图

Fig. 4 Genetic algorithm optimization process flow chart

表 8 验证性实验结果

Table 8 Results of experiments for verification

批次	质量浓度/(mg·mL ⁻¹)								Q	
	丹参素	原儿茶醛	丹酚酸 B	丹酚酸	丹参酮	隐丹参酮	丹参酮	迷迭香酸	实际值	预测值
S8	1.287	0.031	30.792	0.657	0.438	0.754	1.197	1.387	0.770	0.783
S10	1.131	0.032	32.221	0.597	0.238	0.691	1.299	1.486	0.755	0.723
S19	1.363	0.030	28.958	0.594	0.420	0.403	1.019	2.054	0.552	0.567
S21	1.205	0.035	30.499	0.633	0.441	0.754	1.360	1.375	0.716	0.681
RSD/%	8.06	6.75	4.37	4.73	25.49	25.77	12.24	20.68		
x_{best} (1±10%)	1.243±0.124	0.029±0.003	32.793±3.278	0.653±0.063	0.362±0.036	0.715±0.720	1.219±0.122	1.643±0.164		

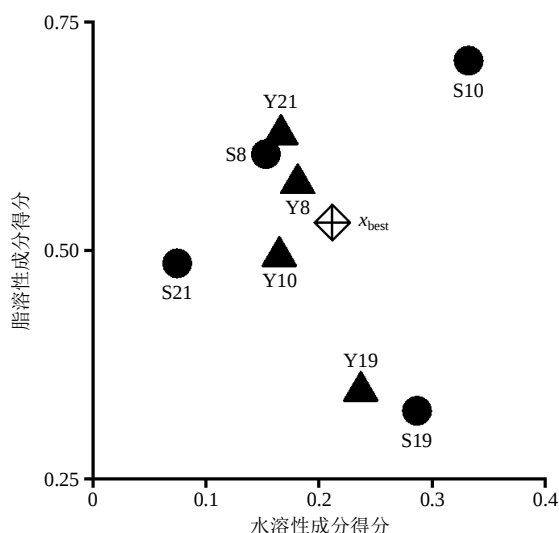


图 5 前馈控制前后对比图

Fig. 5 Comparison scatter plot before and after feedforward control

想解均有较大差距，可能是该批次药材本身隐丹参酮含量较低所致。

3 讨论

中药有多成分多靶点的特点，将前馈控制应用于中药提取过程中需要综合考量各有效成分，以达到减少批次变化带来的提取液质量浮动为目的。本研究使用 TOPSIS 法综合考量丹参提取液中 8 种有效成分的含量，建立了提取工艺和 NIRS 数据为自变量，提取液有效成分含量为因变量的人工神经网络模型，最终根据各批次丹参的近红外数据调整相应的提取工艺。结果表明此模式不仅减少了不同批次药材提取液之间有效成分的含量差异，且对它们的具体含量进行了一定的控制，实现了以中药提取液多种有效成分含量为目标的均化，为中药均化方面研究提供参考。

药材中有效物质的提取是中成药加工生产过程

中的关键工序，提取物的质量优劣与最终产品质量休戚相关。但中药材质量往往难以保持一致，使得其相应的提取物质量也参差不齐。针对提取过程建立前馈控制模型，则可对中药材质量波动产生的提取物质量差异进行补偿，减少提取液之间的质量差异。前馈控制模型一旦建立，只需依据模型改变提取工艺，便可提升中药产品的质量一致性。但是需要注意的是前馈控制不能完全消除原料带来的产品质量差异，其仅可作为中药质量控制的一部分。在对原料进行一定筛选的前提下^[25]，通过前馈控制对提取物进行控制，才能得到具有较好质量一致性的中药产品。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 黄赞扬, 王亚静, 祁东利, 等. 基于质量源于设计理念构建多糖组分干法制粒前馈控制模型 [J]. 中国实验方剂学杂志, 2022, 28(4): 130-137.
- [2] 严斌俊, 郭正泰, 瞿海斌, 等. 前馈控制方法及其在丹红注射液醇沉工艺品质提升中的应用 [J]. 中国中药杂志, 2013, 38(11): 1667-1671.
- [3] 李瑶瑶, 张凯旋, 熊皓舒, 等. 基于质量源于设计的复方丹参滴丸混合工艺前馈控制方法研究 [J]. 中国中药杂志, 2021, 46(11): 2816-2823.
- [4] 王晓宇, 李文龙, 瞿海斌. 基于药材光谱数据的前馈控制在三七醇提过程中的应用研究 [J]. 中国中药杂志, 2018, 43(15): 3127-3134.
- [5] Sun Y F, Kmonickova E, Han R L, et al. Comprehensive evaluation of wild *Cordyceps cicadae* from different geographical origins by TOPSIS method based on the macroscopic infrared spectroscopy (IR) fingerprint [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2019, 214: 252-260.
- [6] 周欣, 陈丽燕, 陈晓东, 等. 不同酸剂诱导人工沉香的

- 分析与评价 [J]. 中国中药杂志, 2017, 42(16): 3159-3166.
- [7] 李运, 张霁, 徐福荣, 等. 多指标决策分析 TOPSIS 对三七的质量评价研究 [J]. 中草药, 2017, 48(22): 4764-4771.
- [8] 张志鹏, 邓李红, 李美洲, 等. 基于化学模式识别和熵权 TOPSIS 法分析槐角不同部位的差异 [J]. 中药材, 2021, 43(6): 1413-1419.
- [9] 中国药典 [S]. 一部. 2020: 77.
- [10] Yang N Z, Ju A C, Li X, *et al.* Salvianolate injection in the treatment of acute cerebral infarction: A systematic review and a meta-analysis [J]. *Medicine (Baltimore)*, 2018, 97(47): e12374.
- [11] Du H X, He Y, Pan Y J, *et al.* Danhong Injection attenuates cerebral ischemia-reperfusion injury in rats through the suppression of the neuroinflammation [J]. *Front Pharmacol*, 2021, 12: 561237.
- [12] 宋侨. 近红外光谱技术在马应龙麝香痔疮膏质量控制中应用研究 [D]. 武汉: 湖北中医药大学, 2020.
- [13] Ni L J, Zhang F F, Han M Y, *et al.* Qualitative analysis of the roots of *Salvia miltiorrhiza* and *Salvia yunnanensis* based on NIR, UHPLC and LC-MS-MS [J]. *J Pharm Biomed Anal*, 2019, 170: 295-304.
- [14] Henriksen H C, Næs T, Rødbotten R, *et al.* Simultaneous modelling of process variables and raw material properties as measured by NIR. A case study from cellulose production [J]. *Chemom Intell Lab Syst*, 2005, 77(1/2): 238-246.
- [15] 李文龙, 刘绍勇, 薛东升, 等. 一种基于近红外光谱技术的熊胆粉粗提物洗涤工艺优化方法 [J]. 中国中药杂志, 2014, 39(1): 75-79.
- [16] 陈光玮, 金伟锋, 许守超, 等. 球面对称设计结合遗传神经网络优选黄芪-红花药对水提工艺 [J]. 中草药, 2021, 52(8): 2257-2266.
- [17] Smucker B, Krzywinski M, Altman N. Optimal experimental design [J]. *Nat Methods*, 2018, 15(8): 559-560.
- [18] 王晓宇, 李文龙, 瞿海斌. 前馈控制技术研究进展及其在中药质控中的应用展望 [J]. 中草药, 2017, 48(21): 4560-4564.
- [19] Yan B J, Li Y, Guo Z T, *et al.* Quality by design for herbal drugs: A feedforward control strategy and an approach to define the acceptable ranges of critical quality attributes [J]. *Phytochem Anal*, 2014, 25(1): 59-65.
- [20] 刘书斌, 张宏武, 冯晓莉, 等. 基于 AHP 法优化的熵权 TOPSIS 模型对不同产地黄花菜药材质量的综合评价 [J]. 中国实验方剂学杂志, 2019, 25(6): 168-174.
- [21] 黄鹏程, 金伟锋, 万海同, 等. 遗传神经网络与遗传算法优选黄芪皂苷微波提取工艺条件 [J]. 中草药, 2019, 50(16): 3815-3823.
- [22] 吴思佳, 金伟锋, 王航, 等. 基于遗传神经网络的双目标甘草提取工艺优化 [J]. 中华中医药杂志, 2019, 34(4): 1719-1721.
- [23] 石磊, 高卫红, 吕莉莉, 等. 基于 BP 人工神经网络和遗传算法的葛根总黄酮提取工艺优化研究 [J]. 中国中医急症, 2018, 27(2): 198-201.
- [24] 尹斐, 杨洁红, 方雨晨, 等. 基于 2 种分析方法的补阳还五汤中有效成分提取工艺优化研究 [J]. 中草药, 2018, 49(1): 135-141.
- [25] 钟文, 陈莎, 章军, 等. 重点是原料, 还是工艺?——以葛根苓连汤为例探讨中成药质量一致性控制方法 [J]. 中国中药杂志, 2016, 41(6): 1027-1032.

[责任编辑 郑礼胜]