

基于决明子提取物谱效关系 GA-BP 神经网络模型的谱效评分及质量评价

鄢海燕¹, 王恒^{1,2}, 邹纯才¹

¹皖南医学院药学院, 安徽 芜湖 241002; ²安庆一一六医院, 安徽 安庆 246003

摘要:目的 构建决明子提取物谱效关系的 GA-BP 神经网络模型, 探索谱效评分用于中药质量控制新方法。方法 建立决明子提取物(0.1、0.2、0.4 g/mL)指纹图谱; 采用 5-Fu 制备小鼠肝损伤模型, 测定给予不同剂量(0.4、0.8、1.6 g/kg)决明子提取物的小鼠体质量、肝脏指数及血清 ALT、AST、肝组织 MPO、SOD、T-AOC 等药理学指标并利用 AHP-EWM 计算综合药效; 构建指纹图谱与综合药效的 GA-BP 神经网络模型, 获取相应预测综合药效。利用灰色关联度法建立指纹图谱与实测综合药效、预测综合药效的谱效关系并进行 Gaussian 拟合分析。利用指纹图谱相对峰面积和谱效关联度计算谱效评分, Z 比分数法检验数据可靠性, 确定谱效评分限度范围并对验证样品进行质量评价。结果 GA-BP 神经网络模型的综合药效预测值与实测值均非常接近, 误差小于 0.2。实测综合药效与预测综合药效的谱效关系数据经 Gaussian 拟合表明, SEE 和 RMSE 值均接近于 0, R-square 和 Adjusted R-square 值均大于 0.95, 实测综合药效与预测综合药效的谱效关系高度拟合。对谱效评分进行计算及 Z 比分数法检验, 确定谱效评分限度为 6.16~7.30。验证用各组的预测结果与实验结果相符, 且在谱效评分限度范围内, Z 比分数法检验结果表明数据可靠。结论 GA-BP 神经网络模型能较好地预测综合药效, 建立的谱效评分方法可用于样品的质量评价。

关键词: GA-BP 神经网络; 决明子提取物; 谱效关系

A GA-BP neural network model based on spectrum-effect relationship for assessing spectrum-effect score and quality evaluation of Cassia seeds extract

YAN Haiyan¹, WANG Heng^{1,2}, ZOU Chuncai¹

¹School of Pharmacy, Wannan Medical College, Wuhu 241002, China; ²Anqing 116 Hospital, Anqing 246003, China

Abstract: Objective To construct a GA-BP neural network model based on the spectrum-effect relationship of Cassia seeds extract and test its performance for quality control of Cassia seeds using spectrum-effect score. **Methods** The HPLC fingerprints of Cassia seeds extract (0.1, 0.2, and 0.4 g/mL) were established. In a mouse model of 5-Fu-induced liver injury treated with 0.4, 0.8, and 1.6 g/kg of Cassia seeds extract, the pharmacodynamics parameters were measured to calculate the comprehensive efficacy using AHP-EWM. A GA-BP neural network model between the fingerprints and comprehensive efficacy was constructed, and the corresponding predicted comprehensive efficacy was obtained. The spectrum-effect relationship between the fingerprints and the measured and predicted comprehensive efficacy was established using grey correlation method followed by Gaussian fitting analysis. The spectral efficiency score was calculated using the relative peak area of the fingerprints and the correlation degree of the spectral efficiency. The reliability of the data was tested using the Z-ratio score method. The limit range of the spectral efficiency score was determined and the quality of the verification samples was evaluated. **Results** The error between the predicted value using the GA-BP neural network model and the measured value of the comprehensive efficacy was less than 0.2. Gaussian fitting analysis showed good fitting between the spectrum-effect relationship data of the measured and predicted comprehensive efficacy. The limit of the spectral efficiency score was 6.16-7.30. The prediction results for each verification group were consistent with the experimental results and within the limit of spectral efficiency score, and the results of Z-ratio score analysis demonstrated good data reliability. **Conclusion** The GA-BP neural network model can effectively predict the comprehensive efficacy of Cassia seeds extract, and the established spectrum-effect scoring method can be used for quality evaluation of samples.

Keywords: GA-BP neural network; Cassia seeds extract; spectrum-effect relationship

中药指纹图谱技术作为现代中药质量控制的重要手段,推动了中药产业的技术革新与中药质量标准的进步^[1]。通过统计学分析^[2,3]、模式识别^[4,5]和智能算法^[6,7]等关联技术建立中药指纹图谱与药效指标间的数学模型,可实现谱效关联^[8,9]。当前,谱效关联虽构建了中药化学成分与药效指标的对应关系,但未能实

现“图谱解析→药效预测→质控评价”联动。若能在现有技术框架中引入动态建模机制,即通过数据分析运用机器学习算法^[10,11]解析指纹图谱与药效指标的深层关联,构建持续迭代的“图谱解析→药效预测→质控评价”循环机制,则有望实现通过实时解析指纹图谱信息预测药效趋势,建立以谱效融合为导向的新型质量控制方式,使中药质量监管从被动成分检测升级为谱效融合导向的管理模式,这是推动传统中药质量监管向科学化、智能化转变的重要方向和关键路径。为此,本研究采用层次分析-熵权法(AHP-EWM)组合

收稿日期:2025-04-28

基金项目:安徽高校省级自然科学研究重大项目(KJ2016SD60)

作者简介:鄢海燕,教授,E-mail: yhy0801@126.com

通信作者:邹纯才,教授,E-mail: zouchc@163.com

赋权计算综合药效^[12, 13]。基于决明子提取物指纹图谱和综合药效,采用遗传算法-反向传播(GA-BP)神经网络^[12, 13]结合 Gaussian 拟合进行模型可行性分析、模型构建和验证,预测综合药效,建立谱效关系并制定评价指标,以期对中药指纹图谱在药品质量控制中的深入运用提供参考。

1 材料和方法

1.1 实验动物

昆明种小鼠 72 只,体质量 18~22 g,购于济南朋悦实验动物繁育有限公司,许可证号:SCXK(鲁)2019-0003。实验经皖南医学院实验动物福利与伦理委员会批准(伦理批号:LLSC-2022-036)^[8, 9]。实验前适应性喂养 3 d。

1.2 试药

决明子提取物(皖南医学院药物制剂)。联苯双酯滴丸(万邦德制药集团有限公司);氟尿嘧啶(山东西亚化学股份有限公司,纯度≥99%),溶解于磷酸盐缓冲液备用;谷丙转氨酶(ALT)、谷草转氨酶(AST)、蛋白定量(TP)、超氧化物歧化酶(SOD)、髓过氧化物酶(MPO)和总抗氧化能力(T-AOC)试剂盒(南京建成生物)^[8, 9]。

1.3 仪器

Multiskan Go 全波长酶标仪(赛默飞世尔科技有限公司);JW-3021HR 高速冷冻离心机(安徽嘉文仪器装备有限公司);FA2004B 电子天平(上海越平科学仪器有限公司)等^[8, 9]。

1.4 溶液的制备

决明子提取物溶液的制备、混合对照品的制备见文献^[8, 9]。

1.5 决明子提取物溶液指纹图谱的构建

色谱条件、决明子 HPLC 指纹图谱的建立见文献^[8, 9]。

1.6 决明子提取物抗 5-Fu 致肝损伤作用

动物造模及给药,体质量、肝脏指数、血清 ALT 与 AST 检测、肝组织氧化指标检测见文献^[8, 9]。

1.7 AHP-EWM 综合赋权

1.7.1 AHP 主观赋权

1.7.1.1 构建 AHP 判断矩阵 参照文献^[12],根据 AHP 理论构建判断矩阵标准。根据 ALT、AST、MPO、T-AOC、SOD、体质量及肝脏指数 7 个药效指标在肝损伤评价中的内在信息,对 7 个评价指标重要程度排序为 ALT≈AST>MPO≈T-AOC≈SOD>体质量≈肝脏指数,对相对重要性进行打分,将 ALT 对 AST、MPO、T-AOC、SOD、体质量及肝脏指数的标度分别定为 1、3、3、3、5、5;其它依此类推,构建评价指标 h 与评价指标 l 成对比较的判断矩阵 $X_{h \times l}$ ($h=1, 2, \dots, n; l=1, 2, \dots, n$)。矩阵 X 的形

式见式(1),具体数据见表 1。

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} \quad (式 1)$$

式(1)中, x_{hl} 为评价指标 h 相对于评价指标 l 的比较结果。

表 1 指标成对比较的判断矩阵

Tab.1 Judgement matrix for paired comparison of the parameters

Evaluation	ALT	AST	MPO	T-AOC	SOD	Weight (g)	Liver index
ALT	1	1	3	3	3	5	5
AST	1	1	3	3	3	5	5
MPO	1/3	1/3	1	1	1	2	2
T-AOC	1/3	1/3	1	1	1	2	2
SOD	1/3	1/3	1	1	1	2	2
Weight (g)	1/5	1/5	1/2	1/2	1/2	1	1
Liver index	1/5	1/5	1/2	1/2	1/2	1	1

1.7.1.2 计算 AHP 权重系数 (w_j^s) 按式(2)对矩阵 X 进行几何平均处理,计算初始权重系数 $r(j)$ ($j=1, 2, \dots, n$);按式(3)对 $r(j)$ 进行归一化处理,获得 AHP 权重系数 w_j^s 。

$$r(j) = (x_{j1}x_{j2}\cdots x_{jn})^{1/n} \quad (式 2)$$

$$w_j^s = \frac{r(j)}{\sum_{j=1}^n r(j)} \quad (式 3)$$

1.7.1.3 一致性检验 按式(4)、式(5)、式(6)分别计算最大特征值 λ_{\max} 、一致性指标(CI)及一致性比例(CR)。

$$\lambda_{\max} = \left[\sum_{j=1}^n \left(\frac{\sum_{j=1}^n x_{jn} r(j)}{r(j)} \right) \right] / n \quad (式 4)$$

$$CI = (\lambda_{\max} - n) / (n - 1) \quad (式 5)$$

$$CR = CI / (n - 1) \quad (式 6)$$

计算结果表明,本文 CI 为 0.0014, CR 为 0.00023,均小于 0.1,表明矩阵 X 均有一致性。

1.7.2 EWM 客观赋权

1.7.2.1 药效指标数据矩阵的建立 实验动物数为 m , 药效指标数为 n , 构建药效指标原始数据矩阵 $Y_{m \times n}$, 见式(7)。

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \quad (式 7)$$

式(7)中, y_{mn} 为第 m 只小鼠的第 n 个药效指标数据。本文中 $m=1, 2, 3, \dots, p; n=1, 2, 3, \dots, q$ 。GA-BP 神经网络模型可行性分析及建立时, $p=24$; GA-BP 神经网络验证时, $p=6, q=7$, 即有 7 个药效指标。

1.7.2.2 矩阵转换 将药效指标原始数据矩阵 Y 转换为概率矩阵 $P_{m \times n}$, 见式(8)。

$$P = \frac{y_{mn}}{\sum_{i=1}^m y_{in}} \quad (式 8)$$

$P_{m \times n}$ 表示第 m 个药效指标下的第 n 个药效指标测定结果的概率。

1.7.2.3 评价指标信息熵(H_j)和EWM权重系数(w_j^o)的计算 药效指标信息熵(H_j)和EWM权重系数(w_j^o)的计算, 见式(9)、式(10)。

$$H_j = -\frac{1}{\ln m} \sum_{i=1}^m P_{mi} \ln P_{mi} \quad (式 9)$$

$$w_j^o = \frac{(1 - H_j)}{\sum_{j=1}^n (1 - H_j)} \quad (式 10)$$

$j=1, 2, 3, \dots, q$ 。本文 $q=7$, 即有 7 个药效指标。

1.7.3 组合权重系数(w_j)的确定 按“1.7.1.2”项下 AHP 计算方法得到 w_j^s , 按“1.7.2.3”项下 EWM 计算方法得到 w_j^o , 按式(11)计算各个评价指标的 w_j 。

$$w_j = \frac{w_j^s w_j^o}{\sum_{j=1}^n w_j^s w_j^o} \quad (式 11)$$

1.8 数据的标准化处理及综合药效的计算

1.8.1 指纹图谱共有峰峰面积标准化处理 参照文献[14], 采用均值化变换法对指纹图谱共有峰数据进行标准化处理。

1.8.2 药效数据标准化处理 参照文献[14], 采用最小值变换法对药效数据进行标准化处理。体质量、T-AOC、SOD 中各数据与该药效指标中数据最大值相除, 肝脏指数、ALT、AST 和 MPO 中药效指标中最小值与各数据相除, 得出各药效指标标准化值, 即 $a_{n(m)} = \frac{x_{n(m)}}{x_{n,max}}$ (n 为体质量、T-AOC、SOD, m 为实验动物数, x 为药效数值, x_{max} 的选择需包括正常组和模型组数据);

$b_{n(m)} = \frac{y_{n,min}}{y_{n(m)}}$ (n 分别为肝脏指数、ALT、AST 和 MPO, m 为实验动物数, y 为药效数值, y_{min} 的选择需包括正常组和模型组数据) 综合药效=各药效指标数据标准化结果乘以各自权重、求和。

1.9 决明子提取物指纹图谱与综合药效的谱效关系分析

以 Matlab 2014a 语言自编均值化邓氏关联度计算法程序, 建立决明子提取物指纹图谱与综合药效的谱效关系^[15, 16]。

1.10 GA-BP神经网络模型建立的可行性分析

基于决明子提取物指纹图谱与其对应的综合药效(ED)数据, 通过 Matlab 2014a 软件建立 GA-BP 神经网络预测模型。设定 GA 参数: 最大遗传代数 60, 变量的二进制位数 10, 代沟 0.95, 交叉概率 0.7, 变异概率 0.01, 随机产生的染色体个体数目 40, 压差为 2。BP 网络训练

循环次数 5000, 学习率为 0.05, 收敛误差 0.0000001, 其余为默认值。

将最优权值和阈值赋予 GA-BP 神经网络模型, 以表 1“建模用”中的 9 组数据(高、中、低浓度各取 3 组)进行学习训练, 3 组数据(高、中、低浓度各取 1 组)进行测试, 建立 GA-BP 神经网络模型(MF)。基于 GA-BP 神经网络模型(MF)获得“建模用”24 组决明子提取物指纹图谱对应的预测综合药效(MF)。按“1.9”项下方法, 建立决明子提取物指纹图谱-预测综合药效(MF)的谱效关系并对谱效关系(MF)进行分析。

1.11 GA-BP神经网络模型的构建与验证

在“1.10”项下 GA-BP 神经网络模型(MF)可行的基础上, 以表 1 中“建模用”的全部 24 组指纹图谱-综合药效(“建模用”ED)数据进行学习训练, 以表 1 中“验证用”的 6 组指纹图谱数据(高、中、低浓度各取 2 组)预测综合药效(MV)。按“1.9”项下方法, 建立决明子提取物指纹图谱-预测综合药效(MV)的谱效关系并对谱效关系(MV)进行分析。

1.12 谱效评分

1.12.1 谱效评分计算方法 谱效评分^[17, 18]: 计算决明子提取物指纹图谱各共有峰的相对峰面积并与其对应的谱效关联度相乘, 求和。

1.12.2 谱效评分限度的确定方法 对图 2 中建模用决明子 HPLC 指纹图谱进行处理, 以 8 号峰橙黄决明素为参照峰, 计算 31 个共有峰的相对峰面积并与其共有峰对应的关联度相乘求和, 计算建模用样品的谱效评分并以 Z 比分数法^[19, 20]进行数据检验, 确定谱效评分限度范围。

1.12.3 待检样品的谱效评分 将图 3 中验证用决明子指纹图谱共有峰数据及对应综合药效指标与图 2 中建模用决明子指纹图谱共有峰数据及对应综合药效指标分别按低、中、高浓度组组合, 按“1.9”项下方法计算关联度, 按“1.12.1”和“1.12.2”项下方法处理, 计算谱效评分并以 Z 比分数法进行数据检验。若样品的谱效评分在模型的谱效评分限度范围内, 样品此项检验合格, 相关数据纳入模型库中; 否则判定此项检验不合格, 相关数据不纳入模型库中。

1.13 统计学分析

通过 SPSS 18.0 软件进行统计学处理, 实验数据以均数±标准差表示, 组间比较采用独立样本 t 检验, $P < 0.05$ 认为差异有统计学意义。

2 结果

2.1 决明子提取物 HPLC 指纹图谱的建立

2.1.1 指标成分的标定 决明子提取物及混合对照品的

HPLC图显示(图1),各目标成分峰峰形较好,通过色谱峰行为及保留时间的一致性,确定决明子提取物指纹图

谱中8、13、23、29、30号峰分别为橙黄决明素、大黄酸、大黄素、大黄酚和大黄素甲醚。

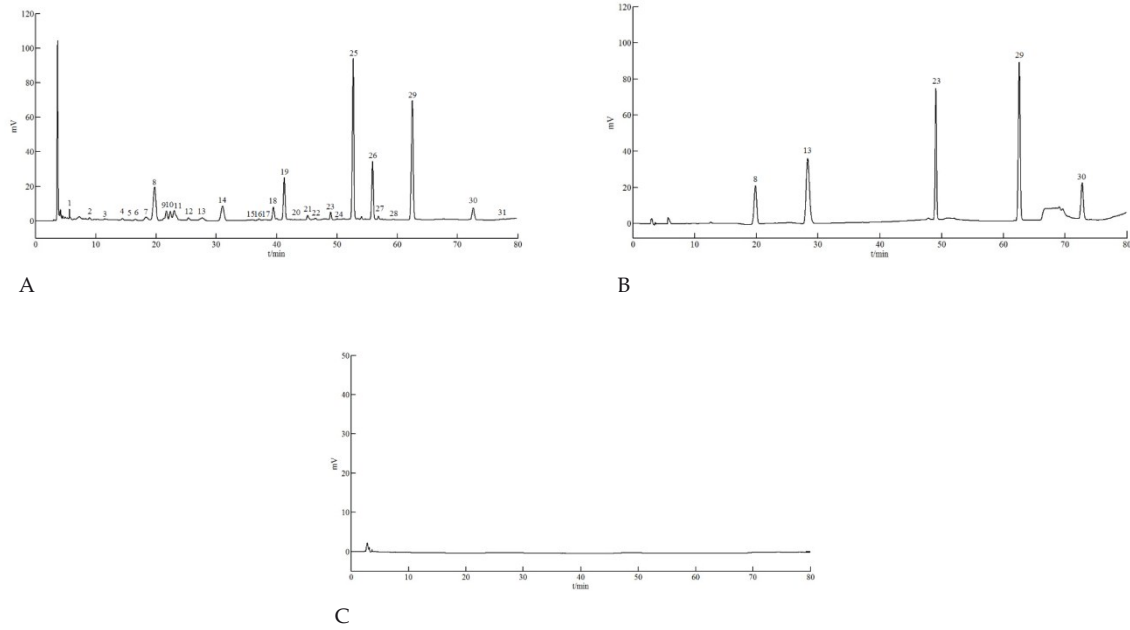


图1 决明子提取物(A)、混合对照品(B)及空白溶液(C)HPLC图

Fig.1 High-performance liquid chromatography (HPLC) of Cassia seeds extract (A), mixed reference substance (B) and blank solution (C).

2.1.2 方法学验证 精密度、稳定性、重复性试验,橙黄决明素、大黄酸、大黄素、大黄酚、大黄素甲醚保留时间的RSD值和峰面积的RSD值均小于5%,结果表明方法可靠。

决明子提取物指纹图谱导入《中药色谱指纹图谱相似度评价系统》(国家药典委员会,2012.130723 版本)软件,以S1为参照图谱,通过Mark峰匹配得到决明子提取物31个共有峰,生成共有模式及匹配数据(共有峰峰面积)。对共有峰峰面积进行标准化处理(图2)。

2.1.3 建模用决明子提取物指纹图谱的建立 将24批

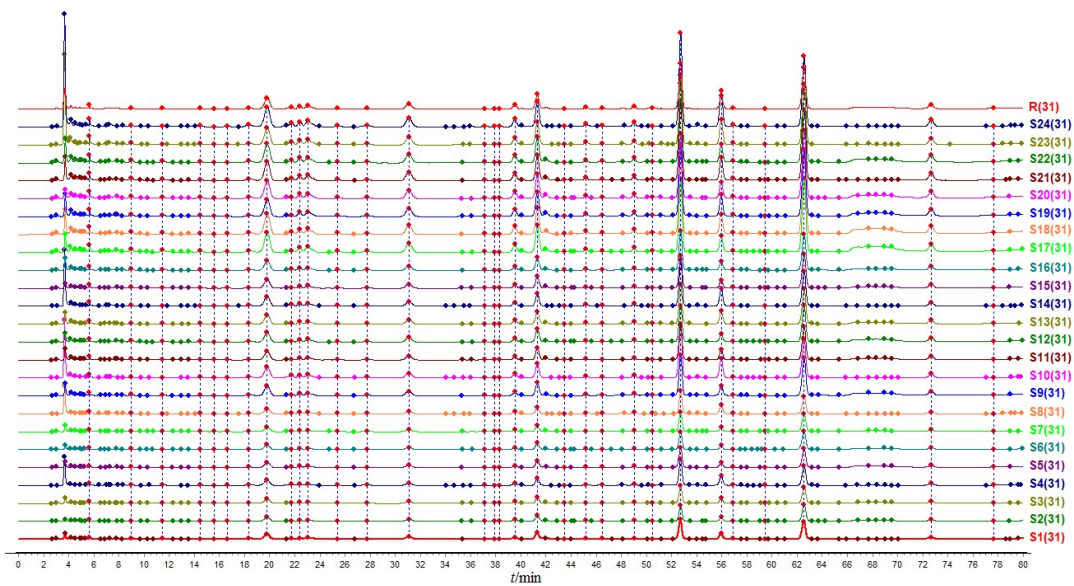


图2 建模用决明子HPLC指纹图谱

Fig.2 HPLC fingerprints of Cassia seeds used for modeling. S1-S8 is 0.01 g·mL⁻¹, S9-S16 is 0.02 g·mL⁻¹, and S17-S24 is 0.04 g·mL⁻¹; R is the common mode.

2.1.4 验证用决明子提取物指纹图谱的建立 以建模用决明子提取物指纹图谱的共有模式为参照,建立6批验证

用决明子提取物指纹图谱,生成匹配数据。对共有峰峰面积进行标准化处理(图3)。

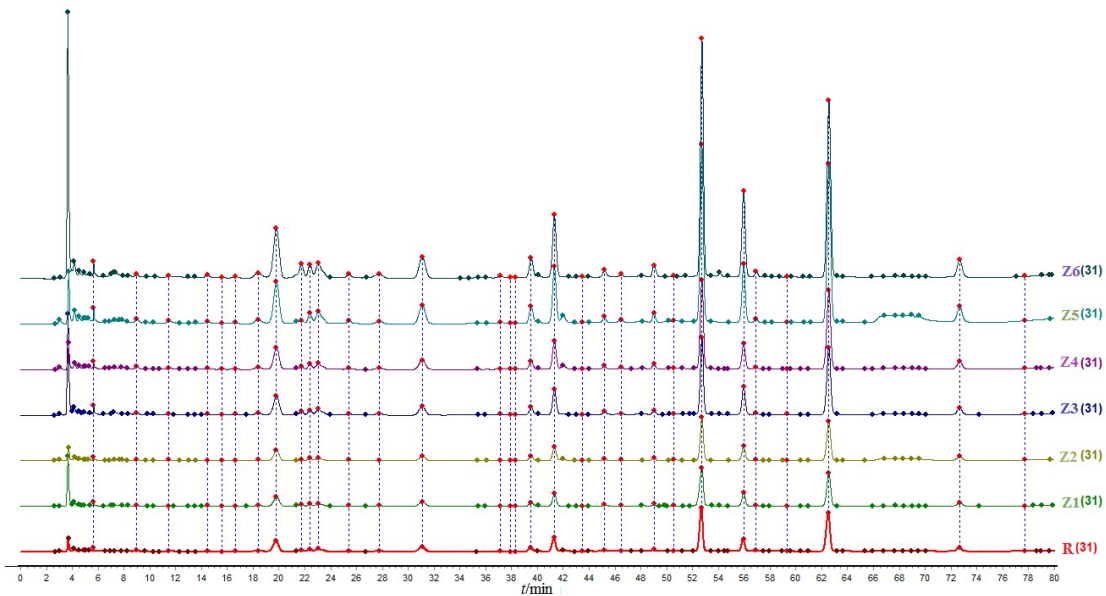


图3 验证用决明子HPLC指纹图谱

Fig.3 HPLC fingerprints of cassia seeds used for verification. R: Common mode; Z1, Z2: 0.01 g·mL⁻¹; Z3, Z4: 0.02 g·mL⁻¹; Z5, Z6: 0.04 g·mL⁻¹.

2.2 决明子提取物抗5-Fu致小鼠肝损伤的药效学结果

正常组、模型组、联苯双酯阳性药组的药效学指标数据($n=10$)分别为:小鼠体质量为 31.72 ± 2.02 、 19.90 ± 1.42 、 20.86 ± 1.32 g;肝脏指数为 54.86 ± 5.02 、 64.58 ± 2.59 、 57.90 ± 3.24 mg/g;ALT为 7.02 ± 1.90 、 18.43 ± 7.97 、 11.88 ± 4.81 U/L;AST为 26.37 ± 2.04 、 43.10 ± 11.36 、 34.33 ± 5.66 U/L;MPO为 0.372 ± 0.05 、 1.068 ± 0.36 、 0.449 ± 0.09 U·gprot⁻¹;T-AOC为 0.116 ± 0.03 、 0.065 ± 0.01 、 0.084 ± 0.01 mmol·gprot⁻¹;SOD为 423.0 ± 24.65 、 330.6 ± 32.30 、 393.8 ± 33.05 U·mgprot⁻¹。

与正常组相比,模型组小鼠体质量第3天出现下降($P<0.01$),肝脏指数降低($P<0.01$);血清中ALT和AST含量均升高($P<0.01$),MPO含量升高($P<0.01$),SOD与T-AOC活性降低($P<0.01$)。

与模型组相比,决明子提取物组减缓了小鼠体质量的下降,缓解肝脏指数降低($P<0.05$),降低了ALT、AST、MPO的含量($P<0.05$),增加了SOD与T-AOC活性($P<0.01$),且均呈剂量依赖关系(表2)。

体质量、肝脏指数、ALT、AST、MPO、T-AOC、SOD的组合权重分别为0.0707、0.0104、0.2829、0.3446、0.1862、0.0740、0.0312(建模用);0.0381、0.0310、0.3314、0.3236、0.1450、0.0830、0.0479(验证用)。综合药效: x_{\max} 或 y_{\min} 的取值包括正常组和模型组数据在内,体质量、肝脏指数、ALT、AST、MPO、T-AOC、SOD对应的 x_{\max} 或 y_{\min} 为35.01、44.94、4.53、23.45、0.26、0.19、463.68(表2)。

2.3 指纹图谱-综合药效[实验测定(ED)]的谱效关系结果

建立决明子提取物指纹图谱与综合药效(“建模用”ED)的谱效关系及决明子提取物指纹图谱与综合药效(“验证用”ED)的谱效关系(图4A)。

2.4 GA-BP神经网络模型的可行性分析结果

2.4.1 GA-BP神经网络模型[模型可行性(MF)]的建立 目标寻优的进化过程(图5A),将最优权值和阈值赋予新建的GA-BP神经网络,以表2中的9组数据(组别:S1-L1、S2-L2、S3-L3;S9-W1、S10-W2、S11-W3;S17-H1、S18-H2、S19-H3)进行学习训练(图5B);3组数据(组别:S5-L5、S13-W5、S22-H6)进行测试,预测结果分别为0.5008(实测结果为0.5282,绝对误差为-0.0274)、0.6045(实测结果0.5948,绝对误差为0.0097)、0.6768(实测结果为0.6635,绝对误差为0.0133),预测结果与实测结果较为接近。网络训练的均方误差曲线(图5C),相关系数 $r=1$,误差小于0.2(图5D)。GA-BP神经网络建立模型预测的综合药效(MF)(表2)。

2.4.2 谱效关系(MF)的数据分析 建立决明子提取物指纹图谱与综合药效(MF)的谱效关系(图4B)。对指纹图谱-综合药效(“建模用”ED)(图4A)和指纹图谱-综合药效(MF)(图4B)进行谱效关系数据拟合(图6,表3)。

由拟合数据可知(表3),SEE和RMSE接近于0,而R-square和Adjusted R-square均大于0.95,说明谱效关系(“建模用”ED-MF)与谱效关系(MF)拟合度非常好,

表2 决明子提取物抗5-Fu致肝损伤作用的药效学结果

Tab.2 Pharmacodynamic results of Cassia seeds extract against 5-Fu-induced liver injury in mice

Category	Group	Pharmacodynamic indicators and their standardized values (BZ)							Comprehensive pharmacological effects				
		Weight (g) /BZ	Liver index (mg·g ⁻¹)/BZ	ALT(U·L ⁻¹) /BZ	AST(U·L ⁻¹) /BZ	MPO (U·gprot ⁻¹) /BZ	T-AOC (mmol·gprot ⁻¹)/BZ	SOD/U·mgprot ⁻¹ /BZ	ED	MF	MB	MV	
Used for modeling	S1-L1	20.38/0.5821	62.49/0.7192	11.49/0.3941	28.81/0.8140	0.475/0.5474	0.074/0.3936	340.2/0.7336	0.5946	0.5803	0.5480		
	S2-L2	18.91/0.5401	65.47/0.6864	8.53/0.5309	30.10/0.7790	1.206/0.2156	0.090/0.4787	364.3/0.7856	0.5641	0.5494	0.5419		
	S3-L3	20.33/0.5807	63.13/0.7118	13.56/0.3340	38.75/0.6052	0.513/0.5068	0.083/0.4415	398.9/0.8603	0.5054	0.5008	0.4951		
	S4-L4	22.54/0.6438	56.33/0.7978	8.10/0.5589	29.77/0.7878	0.525/0.4952	0.066/0.3510	388.6/0.8380	0.6277	0.5009	0.5928		
	S5-L5	23.10/0.6598	54.72/0.8212	12.88/0.3516	39.96/0.5869	0.413/0.6295	0.072/0.3830	382.6/0.8251	0.5282	0.5008	0.5107		
	S6-L6	21.10/0.6027	55.08/0.8159	13.71/0.3303	46.38/0.5056	0.504/0.5159	0.088/0.4681	403.3/0.8698	0.4766	0.6478	0.4723		
	S7-L7	21.11/0.6030	57.69/0.7790	9.26/0.4890	27.59/0.8502	0.888/0.2928	0.081/0.4308	424.9/0.9164	0.5970	0.6208	0.5458		
	S8-L8	22.67/0.6475	60.67/0.7407	10.57/0.4286	25.57/0.9170	0.534/0.4869	0.069/0.3670	422.5/0.9111	0.6370	0.5008	0.6126		
	S9-W1	20.28/0.5793	56.17/0.8000	8.63/0.5244	35.35/0.6633	0.351/0.7407	0.098/0.5212	296.3/0.6389	0.6227	0.5583	0.5432		
	S10-W2	24.04/0.6867	60.49/0.7429	9.24/0.4901	30.83/0.7607	0.366/0.7104	0.092/0.4893	405.2/0.8739	0.6528	0.6334	0.5801		
	S11-W3	20.66/0.5901	55.83/0.8049	8.96/0.5056	31.67/0.7405	0.805/0.3230	0.095/0.5053	408.8/0.8815	0.5733	0.5193	0.5257		
	S12-W4	22.07/0.6304	58.01/0.7747	12.03/0.3764	38.04/0.6165	0.360/0.7222	0.094/0.5000	415.3/0.8958	0.5710	0.6762	0.5025		
	S13-W5	19.14/0.5467	62.94/0.7140	9.81/0.4614	32.59/0.7197	0.463/0.5616	0.097/0.5159	407.6/0.8791	0.5948	0.5884	0.5525		
	S14-W6	19.86/0.5673	57.65/0.7795	6.32/0.7167	32.90/0.7128	0.481/0.5405	0.087/0.4627	424.9/0.9164	0.6601	0.5852	0.6391		
	S15-W7	21.43/0.6121	54.58/0.8234	19.73/0.2295	32.08/0.7130	0.440/0.5909	0.084/0.4468	405.5/0.8744	0.5391	0.5336	0.5086		
	S16-W8	21.63/0.6178	60.53/0.7424	11.24/0.4027	26.82/0.8744	0.369/0.7046	0.089/0.4734	388.4/0.8377	0.6590	0.6456	0.5770		
	S17-H1	20.56/0.5873	63.46/0.7081	11.42/0.3966	27.64/0.8484	0.327/0.7951	0.099/0.5266	426.9/0.9207	0.6692	0.6773	0.6855		
	S18-H2	19.28/0.5507	55.78/0.8056	6.55/0.6916	34.28/0.6842	0.357/0.7283	0.101/0.5372	347.7/0.7498	0.6775	0.6769	0.5481		
	S19-H3	19.25/0.5498	52.71/0.8526	15.01/0.3017	35.42/0.6620	0.413/0.6295	0.112/0.5957	437.0/0.9425	0.5519	0.5149	0.4823		
	S20-H4	18.52/0.5290	51.59/0.8711	16.04/0.2822	28.38/0.8264	0.563/0.4618	0.094/0.5000	396.7/0.8556	0.5608	0.6472	0.5288		
	S21-H5	20.71/0.5915	64.65/0.6951	8.99/0.5039	30.88/0.7595	0.501/0.5190	0.104/0.5532	444.1/0.9578	0.6208	0.6744	0.6463		
	S22-H6	20.55/0.5870	60.00/0.7490	7.32/0.6187	30.95/0.7578	0.428/0.6075	0.098/0.5212	392.0/0.8455	0.6635	0.6693	0.6947		
	S23-H7	21.24/0.6067	50.14/0.8963	19.03/0.2379	34.25/0.6848	0.696/0.3736	0.109/0.5797	406.0/0.8755	0.4953	0.6640	0.4724		
	S24-H8	21.54/0.6153	60.09/0.7479	6.81/0.6647	24.73/0.9482	0.398/0.6533	0.096/0.5106	412.7/0.8901	0.7533	0.6772	0.7284		
Used for verification	Z1-ZL1	19.17/0.5476	66.50/0.6758	22.26/0.2034	33.31/0.7041	1.162/0.2238	0.076/0.4042	393.8/0.8492	0.4437			0.5664	
	Z2-ZL2	21.54/0.6153	63.41/0.7087	12.04/0.3762	32.40/0.7237	0.584/0.4452	0.071/0.3776	332.6/0.7174	0.5345			0.4729	
	Z3-ZW1	20.93/0.5978	59.52/0.7550	15.33/0.2953	27.40/0.8558	0.608/0.4276	0.079/0.4202	414.9/0.8949	0.5607			0.4919	
	Z4-ZW2	22.38/0.6392	59.90/0.7502	15.39/0.2943	30.12/0.7786	0.375/0.6933	0.088/0.4681	409.9/0.8839	0.5788			0.5526	
	Z5-ZH1	21.11/0.6030	63.81/0.7043	8.046/0.5628	38.50/0.6091	0.496/0.5242	0.095/0.5053	407.7/0.8792	0.5885			0.4762	
	Z6-ZH2	19.72/0.5633	52.36/0.8583	6.509/0.6956	28.24/0.8304	0.260/1.000	0.089/0.4734	381.0/0.8216	0.7710			0.6579	

S1-S24 are the fingerprints of cassia seeds extract used for modeling (Figure 2); L, W and H: Low, medium and high concentration groups of mouse models, respectively; Z1-Z6: Fingerprints of Cassia seeds extract used for verification (Figure 3); ZL, ZW, and ZH: Low, medium and high concentration groups for verification respectively; ED, MF, MB and MV: Experimental Determination, Model Feasibility, Model Building and Model Verification respectively.

GA-BP神经网络模型预测的综合药效具有可行。

2.5 GA-BP神经网络模型的建立与验证结果

2.5.1 GA-BP神经网络模型的建立 在“2.4”GA-BP神经网络模型(MF)可行的基础上,以表2中“建模用”的24组数据进行学习训练并建模,除GA参数最大遗传代数50外,其它参数同“2.4”。

目标寻优的进化过程、训练过程(图7A、B),网络训

练的均方误差曲线(图7C),相关系数 $r = 1$,表明GA-BP神经网络预测值与实测值非常接近,误差小于0.2(图7D),说明GA-BP神经网络模型可较好拟合指纹图谱与综合药效[模型建立(MB)]间的映射关系。

2.5.2 验证 应用“2.5.1”项下建立的GA-BP神经网络模型,对表1中“验证用”的6组数据进行分析,预测6组数据的综合药效[模型验证(MV)](表2),并建立指纹图

谱-综合药效(MV)的谱效关系(图8B)。

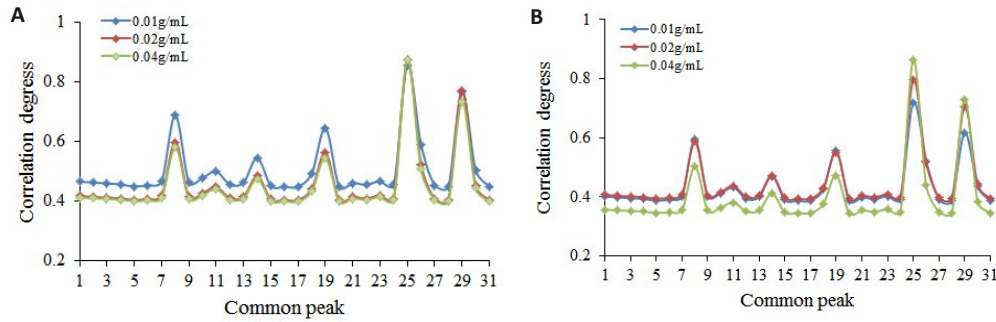


图4 指纹图谱-综合药效("建模用"ED)(A)及指纹图谱-综合药效(MF)(B)的谱效关系
Fig.4 Spectrum-effect relationship between fingerprint-comprehensive efficacy (ED for "modeling") (A) and fingerprint-comprehensive efficacy (MF) (B).

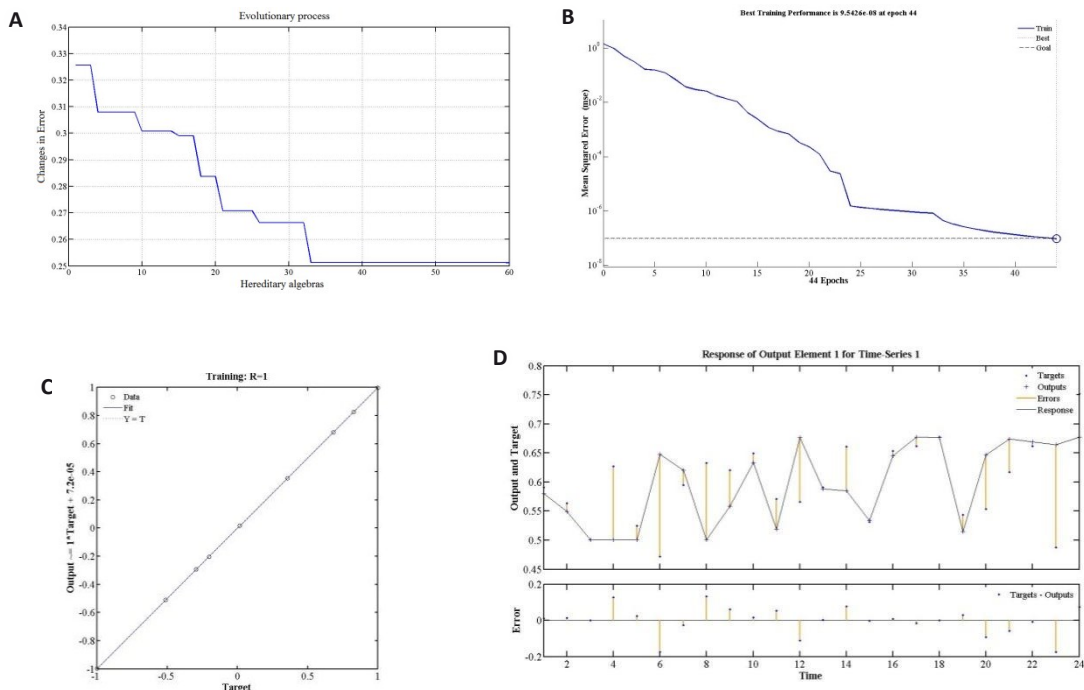


图5 GA-BP神经网络模型(MF)的目标寻优进化过程(A)、训练过程(B)、均方误差(C)、误差图(D)
Fig.5 Target optimization evolution process (A), training process (B), mean square error (C), and error diagram (D) of GA-BP neural network model (MF).

对指纹图谱-综合药效(“验证用”ED)(图8A)和的指纹图谱-综合药效(MV)(图8B)进行谱效关系数据拟合(图9、表4)。

由拟合数据可知(表4),SEE和RMSE接近于0,而R-square和Adjusted R-square均大于0.95,说明谱效关系(“验证用”ED、MV)拟合度高,通过GA-BP神经网络模型可以较好地预测综合药效并准确表达谱效关系。

2.6 谱效评分结果与分析

2.6.1 谱效评分限度范围的确定 对表2中建模用样品S1-S24进行处理,以8号峰橙黄决明素为参照峰,计算31个共有峰的相对峰面积并与各共有峰对应的关联度(图4A)相乘求和,计算谱效评分(表5)。

对表5中的谱效评分数据进行Z比分数法处理,获得主要参数:中位数=6.62, Q1=6.42, Q3=6.91, IQR=0.49, NIQR=0.36, 极小值=6.16, 极大值=7.30。按Z比分数法的计算公式计算Z值,绘制谱效评分Z比分数柱状图(图10)。

Z比分数中,各组的|Z|均小于2,结果满意(图10)。故本实验以建模用样品S1~S24谱效评分的最低值和最高值设定谱效评分限度范围为:6.16~7.30。

2.6.2 验证用样品谱效评分结果 对表2中验证用样品Z1-Z6进行处理,以8号峰橙黄决明素为参照峰,计算31个共有峰的相对峰面积并与各共有峰对应的关联度(图8A、B)相乘求和,计算谱效评分(实验测定)和谱效

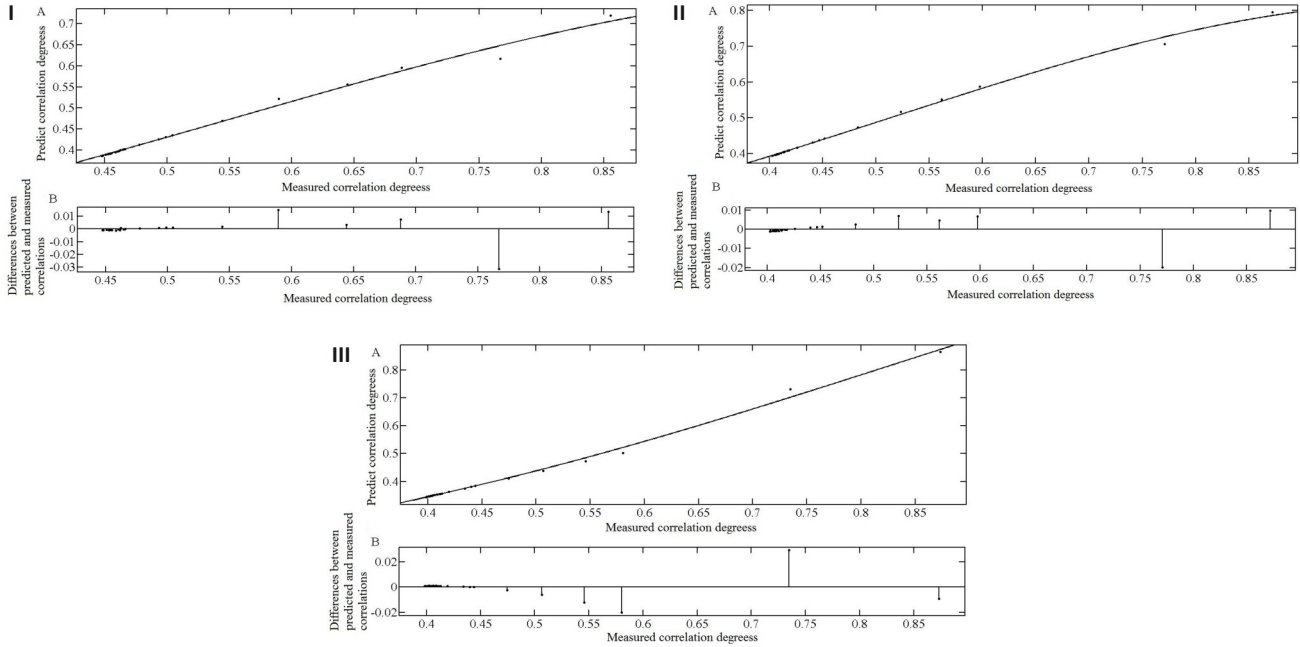


图6 低(I)、中(II)、高(III)剂量组谱效关系("建模用"ED-MF)Gaussian拟合曲线(A)和残差图(B)

Fig.6 Gaussian fitting curve (A) and residual diagram (B) of the spectrum-effect relationship (ED-MF for "modeling") for low (I), medium (II) and high (III) dose groups.

表3 低(I)、中(II)、高(III)剂量组谱效关系("建模用"ED-MF)Gaussian拟合曲线数据

Tab.3 Gaussian fitting curve data of spectrum-effect relationship (ED-MF for "modeling") for low (I), medium (II) and high (III) dose groups

Group	General model gauss	Coefficients (with 95% confidence bounds)			Goodness of fit			
		a	b	c	SSE	R-square	Adjusted R-square	RMSE
I		0.7850	1.1150	0.7935	0.0015	0.9927	0.9921	0.0073
II	$f(x) = a * \exp(-((x-b)/c)^2)$	0.8244	1.0330	0.7339	0.0006	0.9977	0.9975	0.0047
III		1.3500	1.4890	0.9330	0.0015	0.9961	0.9959	0.0074

评分(模型预测)(表6),验证用样品各组的实验测定和模型预测谱效评分均在谱效评分限度范围6.16~7.30。

2.6.3 验证用样品谱效评分的评价与分析 对表5、6中的谱效评分数据进行Z比分数法处理,获得主要参数:中位数=6.68, Q1=6.45, Q3=6.90, IQR=0.45, NIQR=0.33,极小值=6.16,极大值=7.30。按Z比分数法的计算公式计算Z值,绘制谱效评分Z比分数柱状图(图11)。验证用样品各组实验测定和模型预测谱效评分结果的|Z|均小于2,结果满意。

基于现有GA-BP神经网络模型,验证用样品各组的预测结果与实验结果相符,且在模型设定的谱效评分限度范围内,Z比分数法处理结果表明数据可靠,其数据可纳入模型数据库进行数据整合。

3 讨论

《中国药典》自2010年版开始收载中药指纹(特征)图谱技术,采用随行对照、特征峰分析、相似度3种模式

对中药质量进行评价^[21]。然而,这种以“化学等同性”为导向的静态评价模式存在以下不足:一是指纹图谱与药理学响应脱节。虽然现有谱效关系的研究报道较多^[22, 23],但中药指纹图谱与药效指标相关性尚未建立动态量化关系,导致质量控制指标与药理学评价之间缺乏直接联系;现有质控标准多基于历史数据设定成分含量阈值,无法动态响应药材种植、炮制工艺变化、中药提取物的制备及制剂制备过程等引发的药效波动;现有质量控制方法是多效一法,也就是对于中药可能有几种药效或功能主治,但质量控制方法只有一种^[24]。二是谱效关系的功能缺位。传统关联分析方法(如PLS^[25, 26]、灰色关联度^[27, 28])虽能建立中药指纹图谱-药理学间的关联,但多采用离线建模方式,仅能实现“已知图谱→已知药效”间的关联,缺乏谱药融合为导向的中药质量控制能力。

数据拟合的方式有多种,如 gustom equation、

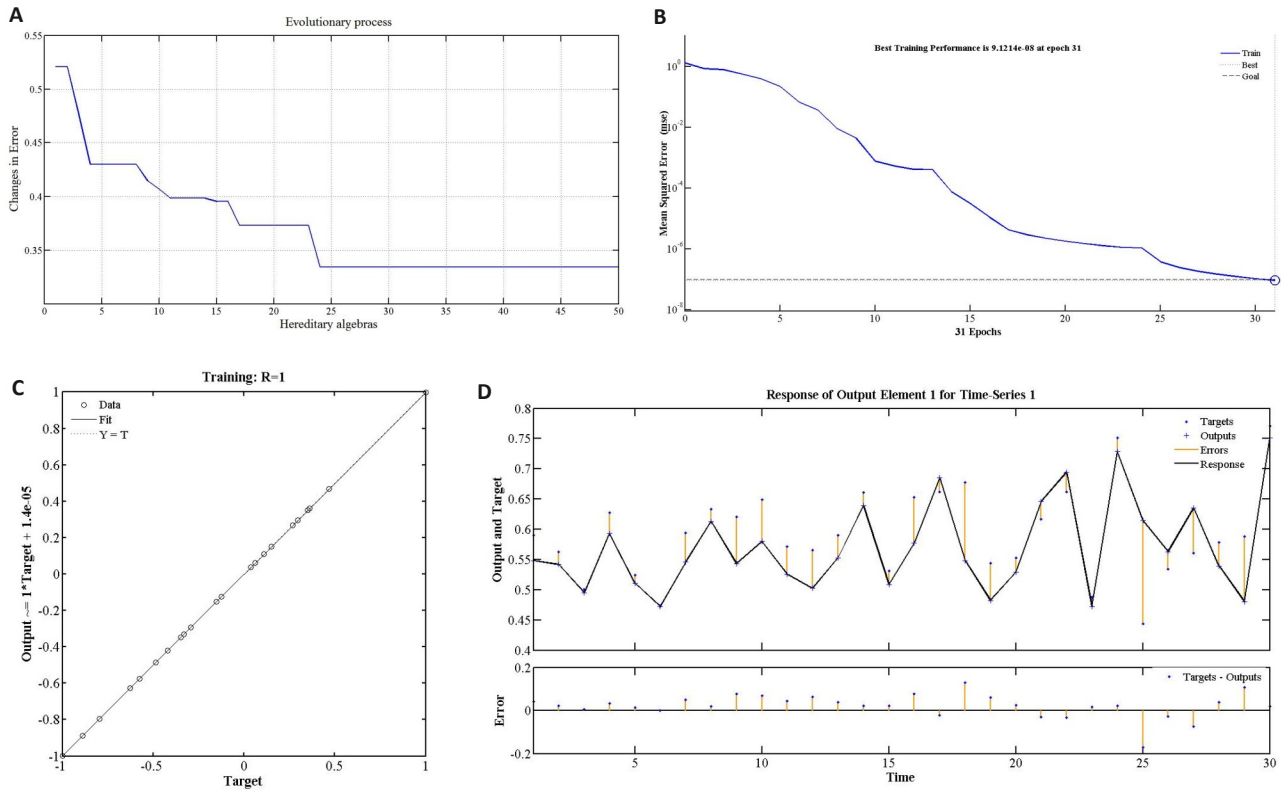


图7 GA-BP神经网络模型的目标寻优进化过程(A)、训练过程(B)、均方误差(C)、误差图(D)

Fig.7 Target optimization evolution process (A), training process (B), mean square error (C), error diagram (D) of GA-BP neural network model (MF).

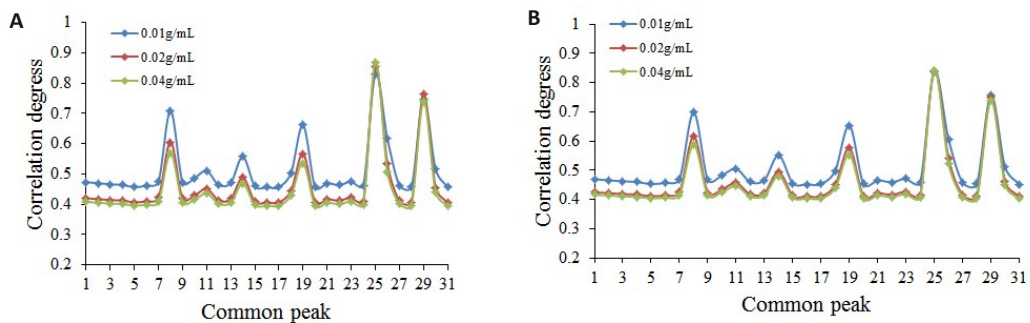


图8 指纹图谱-综合药效("验证用"ED)的谱效关系(A)和指纹图谱-综合药效(MV)(B)的谱效关系

Fig.8 Spectrum-effect relationship (A) of fingerprint-comprehensive efficacy (ED for "verification") and the spectrum-effect relationship (B) of fingerprint-comprehensive efficacy (MV).

exponential、polynomial、exponential、fourier、gaussian、power、interpolant、power、rational、sum of sine 等,经比较分析,gaussian 方式^[29]适合本实验。SEE是模型预测值与真实值之间所有单个误差的平方之和。RMSE是SEE均值的平方根。SSE、RMSE越接近于0,说明模型选择和拟合更好,数据预测也越成功^[30]。R-square(决定系数)是量化模型对因变量变化的解释能力。Adjusted R-square (校正决定系数)为R-square平方,在R-square基础上,根据样本量和特征数量调整后的指标,避免因增加无关变量而虚高R-square。R-square、Adjusted R-square越接近1,表明模型能有效

捕捉数据中的规律。

本文对决明子提取物谱效关系 GA-BP 神经网络模型的构建与应用进行了初步探讨,整个过程包括建模、数据拟合、谱效关系的建立、谱效评分的计算、Z 比分数法检验数据及样品质量的判断等可采用相应的计算机语言完成程序编写,为建立以谱效融合为导向的新型质量控制方式提供了方法参考。后期需在大样本量的基础上进一步优化谱效关联方法,提升 GA-BP 神经网络模型的精确性和稳定性,完善谱效评分的确定方法及谱效评分限度的合理设置,实现通过实时解析指纹图谱信息预测药效趋势,构建持续迭代的"图谱解析→药效预

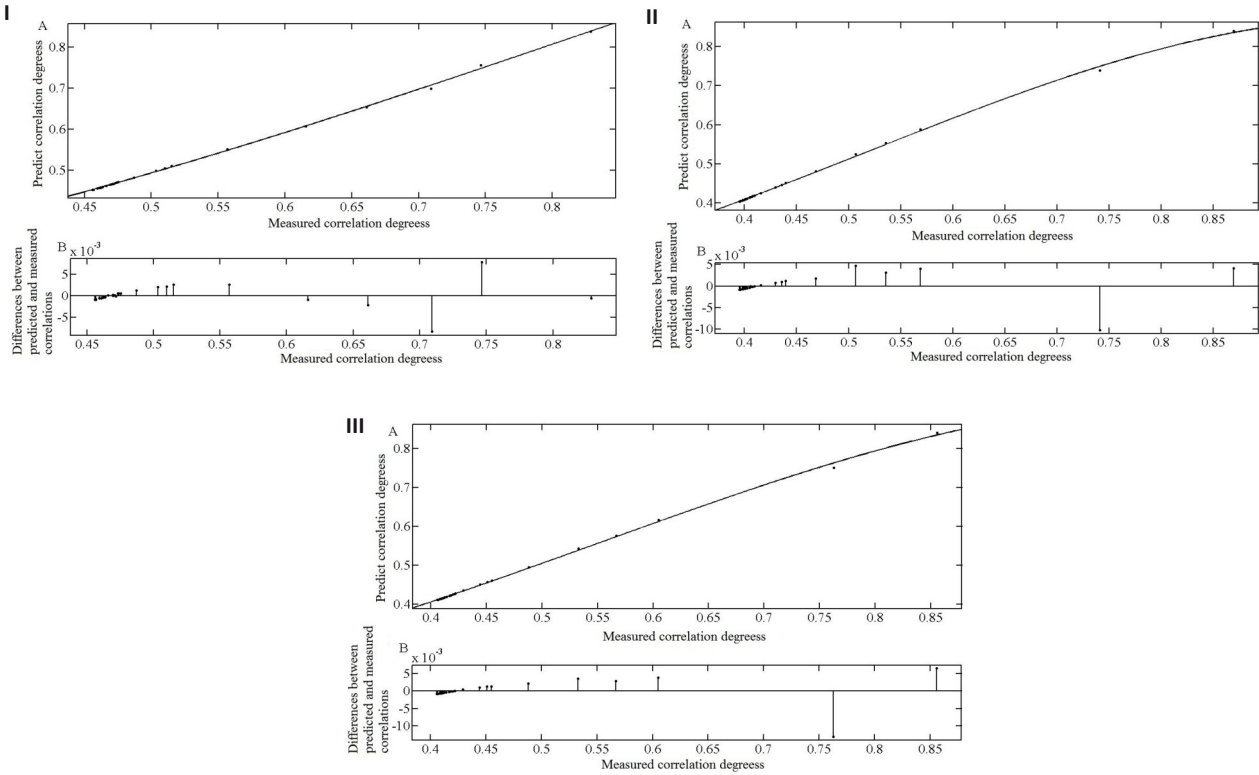


图9 低(I)、中(II)、高(III)剂量组谱效关系("验证用"ED、MV)Gaussian拟合曲线(A)和残差图(B)

Fig.9 Gaussian fitting curve (A) and residual diagram (B) of spectrum-effect relationship (ED, MV for "verification") for low (I), medium (II) and high (III) dose groups.

表4 低(I)、中(II)、高(III)剂量组谱效关系("验证用"ED、MV)Gaussian拟合曲线数据

Tab.4 Gaussian fitting curve data of spectrum-effect relationship (ED, MV for "verification") for low (I), medium (II) and high (III) dose groups

Group	General model gauss	Coefficients (with 95% confidence bounds)			Goodness of fit			
		a	b	c	SSE	R-square	Adjusted R-square	RMSE
I		1.3650	1.5680	1.0590	0.0002	0.9994	0.9994	0.0024
II	$f(x) = a \cdot \exp(-((x-b)/c)^2)$	0.9196	1.0930	0.7667	0.0003	0.9992	0.9991	0.0031
III		0.8751	1.0210	0.7122	0.0002	0.9994	0.9994	0.0026

表5 建模用样品谱效评分结果

Tab.5 Results of spectral efficiency score of samples used for modeling

Group	Spectral efficiency score (serial number)	Group	Spectral efficiency score (serial number)	Group	Spectral efficiency score (serial number)
S1-L1	7.10(a22)(b27)(c27)*	S9-W1	6.50(a8)(b9)(c9)	S17-H1	6.22(a2)(b2)(c2)
S2-L2	6.64(a13)(b14)(c14)	S10-W2	7.04(a21)(b26)(c26)	S18-H2	6.34(a5)(b5)(c5)
S3-L3	6.90(a18)(b23)(c23)	S11-W3	6.69(a14)(b16)(c16)	S19-H3	6.53(a9)(b10)(c10)
S4-L4	7.30(a24)(b30)(c30)	S12-W4	6.38(a10)(b7)(c7)	S20-H4	6.16(a1)(b1)(c1)
S5-L5	6.82(a16)(b20)(c20)	S13-W5	6.87(a17)(b21)(c21)	S21-H5	6.60(a12)(b13)(c13)
S6-L6	6.60(a11)(b12)(c12)	S14-W6	6.93(a20)(b25)(c25)	S22-H6	6.43(a7)(b8)(c7)
S7-L7	6.92(a19)(b24)(c24)	S15-W7	6.32(a4)(b4)(c4)	S23-H7	6.23(a3)(b3)(c3)
S8-L8	7.27(a23)(b29)(c26)	S16-W8	6.59(a10)(b11)(c11)	S24-H8	6.77(a15)(b19)(c19)

* is the sorting sequence number of spectral efficiency scores, where a represents the data in Table 5;b represents the data in Table 5+Table 6 (experimental determination);c represents the data in Table 5+Table 6 (model prediction).

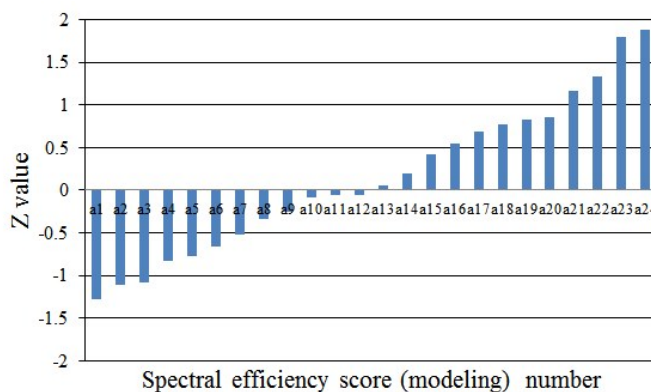


图10 谱效评分(建模)Z比分数柱状图
Fig.10 Spectral efficiency score (modeling) Z-ratio score histogram.

表6 验证用样品谱效评分结果

Tab.6 Results of spectral efficiency score of samples used for verification

Group	Spectral efficiency score		Absolute error	Relative error
	Experimental determination	Model prediction		
Z1-ZL1	6.69 (b17)*	6.70 (c17)*	0.01	0.15
Z2-ZL2	7.16 (b28)	7.17 (c28)	0.01	0.14
Z3-ZW1	6.36 (b6)	6.34 (c6)	-0.02	-0.31
Z4-ZW2	6.73 (b18)	6.71 (c18)	-0.02	-0.30
Z5-ZH1	6.67 (b15)	6.67 (c15)	0.00	0.00
Z6-ZH2	6.88 (b22)	6.88 (c22)	0.00	0.00

* is the sorting sequence number of spectral efficiency scores, where b represents the data in Table 5+Table 6 (experimental determination);c represents the data in Table 5+Table 6 (model prediction).

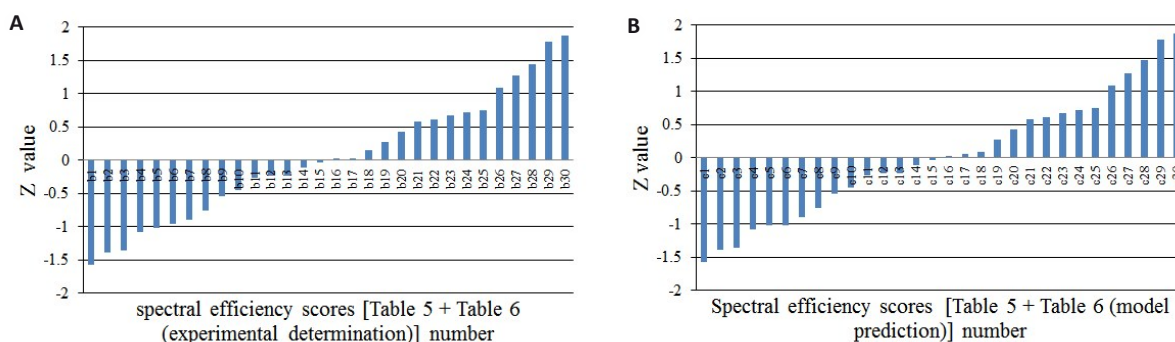


图11 谱效评分[表5+表6(实验测定)](A)和[表5+表6(模型预测)](B)Z比分数柱状图

Fig.11 Z-ratio score histogram of spectral efficiency scores [Table 5 + Table 6 (experimental determination)](A) and [Table 5 + Table 6 (model prediction)](B).

测→质控评价"质量控制模式,为中药质量监管从静态成分检测向动态谱效融合监控转型提供方法学参考和技术路径。

Declaration of interests: The authors declare no competing interests.

参考文献:

[1] 鄢海燕, 邹纯才. 基于谱效整合分子对接技术研究瓜蒌子及其壳、仁抗凝血的物质基础及作用机制[J]. 中国中药杂志, 2022, 47(5):

1370-82.
[2] 李柯翱, 米尔扎提·麦麦提, 张明惠, 等. 基于统计学分析的寒喘祖帕颗粒HPLC指纹图谱研究[J]. 中国现代中药, 2022, 24(8): 1555-60.
[3] 李宗民, 朱浩东, 于丹, 等. 基于多元统计学分析葛根指纹图谱研究[J]. 辽宁中医药大学学报, 2020, 22(11): 50-3.
[4] 吴秀平, 徐明珠, 孙宜春, 等. 基于UPLC多指标成分定量和指纹图谱结合化学模式识别法评价不同基原十大功劳木的质量[J]. 中草药, 2025, 56(8): 2914-24.
[5] 姜春风, 孔雪, 周星宇, 等. 基于指纹图谱、多成分含量测定及化学模式识别的柴胡郁金颗粒质量评价方法研究[J]. 药物分析杂志,

- 2025, 45(2): 201-8.
- [6] 董佳乐, 吴晓康, 杨宛颐, 等. 基于神经网络和多元统计分析的动态预测建模分析[J]. 数字技术与应用, 2024, 42(4): 205-7.
- [7] 王靖霖, 王东升, 苑宏英, 等. 指纹图谱技术与人工智能在污染物溯源解析中的应用研究[J]. 环境保护科学, 2022, 48(6): 130-7.
- [8] 王 恒, 李梦奇, 李燊星, 等. 决明子总蒽醌提取物抗氟尿嘧啶致小鼠肝损伤的谱效关系[J]. 南方医科大学学报, 2023, 43(5): 825-31.
- [9] 王 恒. 决明子对氟尿嘧啶致小鼠肝损伤的保护作用及相关机制研究[D]. 芜湖: 皖南医学院, 2022.
- [10] 韩愈琪, 刘雪莱, 上官文斌. 基于本构神经网络橡胶隔振器动态特性建模[J]. 噪声与振动控制, 2023, 43(3): 265-70.
- [11] 陈梅妹, 王 洋, 雷黄伟, 等. 基于多种机器学习算法和语音情绪特征的阈下抑郁辨识模型构建[J]. 南方医科大学学报, 2025, 45(4): 711-7.
- [12] 俞炎或, 邹纯才, 鄢海燕, 等. 基于谱效关联及AHP-EWM综合赋权的Box-Behnken响应面法结合GA-BP多指标优化白芍-甘草药对提取工艺[J]. 中国现代应用药学, 2025, 42(3): 410-23.
- [13] 汤瑜晨, 胡云飞, 郭抒博, 等. Box-Behnken响应面法结合基准关联度和AHP-EWM优化经典方清金化痰汤的提取工艺[J]. 分析测试学报, 2024, 43(11): 1794-802, 1812.
- [14] 邹纯才, 洪国君, 鄢海燕. 不同数据标准化方法对瓜蒌滴丸抗凝血作用谱效关系的影响[J]. 中国中药杂志, 2018, 43(9): 1864-70.
- [15] 史精干, 李燊星, 王 恒, 等. 决明子分散片对大鼠急性肝损伤的保护作用及其谱效关系研究[J]. 中国药学杂志, 2023, 58(14): 1295-304.
- [16] 黄 莉, 程索婷, 刘 竹, 等. 决明子脐贴对慢传输型便秘大鼠的治疗作用及谱效关系[J]. 南方医科大学学报, 2024, 44(4): 720-6.
- [17] 孟宪生, 包永睿, 王 帅, 等. 复方中药质量标志物的发现与量效色卡可视化技术[J]. 药学学报, 2019, 54(2): 222-7.
- [18] 孟宪生, 罗 曦, 贾梦楠, 等. 中药质量评价研究现状及“质-量”双标评价方法探讨[J]. 中草药, 2023, 54(22): 7281-6.
- [19] 鄢海燕, 魏美玲, 吴 雪, 等. 药学本科毕业设计“双导师制”协同培养模式的构建与教学评价[J]. 通化师范学院学报, 2023, 44(8): 123-32.
- [20] 李雪转. Z比分数法在实验室能力验证中运用的探讨[J]. 计量与测试技术, 2023, 50(7): 86-8.
- [21] 鄢海燕, 邹纯才. 《中国药典》(2010年版~2020年版)中药指纹(特征)图谱应用进展与展望[J]. 南方医科大学学报, 2022, 42(1): 150-5.
- [22] Li H, Zhao H, Chen L, et al. Spectrum-effect relationship between HPLC fingerprints and antioxidant activity of Qi-Fu-Yin based on multiple statistical correlation analysis[J]. Phytochem Anal, 2024, 35(7): 1565-76.
- [23] Liu G, Zhou L, Mu K, et al. Based on spectrum-effect relationship, network pharmacology, and molecular docking, the material basis and mechanism of antioxidant effect of Miao medicine Indigofera stachyoides lindl were studied[J]. ACS Omega, 2024, 9(41): 42199-211.
- [24] 鄢海燕, 邹纯才. 基于功能主治及其物质基础的中药多元化质量控制模式探讨[J]. 国际药学研究杂志, 2019, 46(4): 266-9.
- [25] 舒宇航, 简祈盼, 郑宇锟, 等. 基于灰色关联度与正交偏最小二乘法分析的黄连抗菌谱效关系研究[J]. 时珍国医国药, 2023, 34(9): 2298-302.
- [26] 王春艳, 侯益民, 李汉伟, 等. 基于偏最小二乘回归法的黑果蒺藜乙酸乙酯部位抗肿瘤谱效关系[J]. 中成药, 2021, 43(11): 3227-31.
- [27] 王志萍, 李林杰, 王昱涵, 等. 基于灰色关联度的金母颗粒抗炎作用的谱效关系研究[J]. 中药新药与临床药理, 2023, 34(5): 675-83.
- [28] 张国玉, 熊继东, 魏家艳, 等. 灰色关联度分析清半夏体外抗氧化作用的谱效关系[J]. 华西药学杂志, 2024, 39(4): 382-6.
- [29] 裴庆庆, 刘慧慧. 基于高斯函数拟合的多维数据三维可视化仿真[J]. 计算机仿真, 2024, 41(1): 206-10.
- [30] 王彩玲, 张国浩, 闫晶晶. 基于人工神经网络的透射水体亚硝酸盐含量模拟估计[J]. 中国无机分析化学, 2024, 14(7): 857-65.

(编辑: 林 萍)