

# 基于 RBF 神经网络的加味生化汤药效模拟研究

陈 超, 沈志滨\*

(广东药学院中药学院, 广东 广州 510006)

**摘要:**目的 以加味生化汤的药效模拟为例, 建立中药复方的 RBF 神经网络药效学预测模型。方法 基于正交设计和 RBF 神经网络, 提出了中药复方药效模型, 以加味生化汤的实验数据对模型精度进行了验证和评价。结果 自检验和留一法检验的线性相关系数分别为 0.944、0.846。结论 所建模型可以对复方不同组合进行药效模拟, 可为进一步实验验证提供参考, 在寻找中药复方有效组分, 并探讨以有效组分为处方进行中药复方二次开发模式方面具有良好的应用前景。

**关键词:** 加味生化汤; RBF 神经网络; 药效; 模拟

中图分类号: R284.02

文献标识码: B

文章编号: 0253-2670(2009)08-1239-03

中药复方是在中医药理论指导下, 根据“君、臣、佐、使”的组方原则, 由数种单味药配伍构成的统一整体。作为一个复杂的系统, 中药复方具有各种系统普遍具有的非加合性的特征, 使得在揭示其配伍的科学内涵及作用机制上存在很多困难。中药拆方研究, 对于确定发挥作用的药效物质研究来说, 是一项关键的基础性研究工作。但是, 由于拆方研究的正交实验设计需要大量的、反复的工作, 限于实验条件和人力物力, 所研究的中药复方多为组分少的单方, 缺乏临床实践的代表性<sup>[1]</sup>。计算机模拟可提供有用的结果, 可为进一步实验验证提供参考。本课题组对香丹注射液中抗心肌缺血有效成分进行神经网络分析, 取得较好的效果。因此本实验基于正交设计和 RBF 神经网络, 提出了中药复方药效模型, 对加味生化汤的实验药效进行了模拟预测和验证评价。

## 1 方法

1.1 实验设计: 由加味生化汤 6 个组分 ABCDEF, 按郑青山等<sup>[2]</sup>提出的  $L_{12}(2^6)$  正交表设计(表 1)。每组 5 只雌性小鼠, 按 25 mL/kg ig 7 d, 前 3 d 同时 sc 己烯雌酚 1.8 mg/(kg·d), 7 d 后取子宫烘干称定质量, 记录 100 g 体质量的子宫干质量(mg), 以对雌激素的子宫增重作用, 数据为作用指数, 数值大表示作用强。

1.2 RBF 神经网络: 人工神经网络(artificial neural network, ANN)是人类在对大脑神经网络认识理解的基础上人工构造的、由多层神经元连接而成

表 1  $L_{12}(2^6)$  正交设计表

Table 1 Orthogonal design of  $L_{12}(2^6)$

组别	数值法	下标法	字母法	药效值
1	111111	$A_1 B_1 C_1 D_1 E_1 F_1$	ABCDEF	7, 8, 8, 8, 9
2	111222	$A_1 B_1 C_1 D_2 E_2 F_2$	ABC	7, 7, 7, 8, 8
3	112212	$A_1 B_1 C_2 D_2 E_1 F_2$	ABE	8, 8, 8, 8, 9
4	122121	$A_1 B_2 C_2 D_1 E_2 F_1$	ADF	7, 7, 7, 7, 8
5	122211	$A_1 B_2 C_2 D_2 E_1 F_1$	AEF	7, 7, 7, 8, 8
6	121122	$A_1 B_2 C_1 D_1 E_2 F_2$	ACD	6, 6, 6, 7, 7
7	211221	$A_2 B_1 C_1 D_2 E_2 F_1$	BCF	7, 7, 7, 8, 8
8	212121	$A_2 B_1 C_2 D_1 E_2 F_1$	BDF	6, 7, 7, 8, 8
9	212112	$A_2 B_1 C_2 D_1 E_1 F_2$	BDE	7, 7, 8, 8, 8
10	221112	$A_2 B_2 C_1 D_1 E_1 F_2$	CDE	6, 6, 7, 7, 7
11	221211	$A_2 B_2 C_1 D_2 E_1 F_1$	CEF	7, 7, 8, 8, 8
12	222222	$A_2 B_2 C_2 D_2 E_2 F_2$	—	6, 6, 7, 7, 7

的复杂网络, 能够进行复杂的逻辑操作和模拟复杂的非线性系统, 实现函数的非线性映射, 在语音识别、模式分类、自动控制等领域取得了广泛的应用, 近年来 ANN 亦被成功地推广应用到中药研究领域<sup>[3-5]</sup>。

ANN 主要分为 3 种类型, 即反馈网络模型、前向网络模型和自组织网络模型, 其中径向基函数(radial basis function, RBF)网络是以函数逼近理论为基础而构造的一类前向网络, 在函数逼近能力、分类能力和学习速度等方面均有出色表现<sup>[6]</sup>。RBF 神经网络(RBFNN)由 3 层组成, 输入层传递输入向量到隐含层, 径向基函数构成隐含层节点, 将低维的输入向量变换到高维空间, 使得在低维空间内的线

\* 收稿日期: 2008-12-05

基金项目: 国家科技支撑计划资助项目(2006BA108B03-06); 广东省科技计划资助项目(2008 B030301035); 广东省自然科学基金资助项目(8451022401001607); 广东药学院科研启动项目(2007ZYX05)

作者简介: 陈 超(1981-), 男, 江苏涟水人, 博士, 讲师, 主要从事化学生物信息学和药物信息学研究。Tel: (020)39352181

\* 通讯作者 沈志滨 E-mail: szb8113@yahoo.com.cn

性不可分问题在高维空间内线性可分, 输出层节点是简单的线性函数。高斯函数是最常用的径向基函数:

$$R_i(x) = \exp\left[-\frac{\|x - c_i\|^2}{2\sigma_i^2}\right], \quad i = 1, 2, \dots, m$$

$x$  是  $n$  维输入向量;  $c_i$  是第  $i$  个基函数的中心, 与  $x$  具有相同维数的向量;  $\sigma_i$  是第  $i$  个感知的变量, 它决定了该基函数中心点的宽度;  $m$  是感知单元的个数。

输入层实现  $x \rightarrow R_i(x)$  的非线性映射, 输出层实现  $R_i(x) \rightarrow y_k$  的线性映射, 即:

$$y_k = \sum_{i=1}^m w_{ki} R_i(x), \quad k = 1, 2, \dots, r$$

其中,  $r$  是输出节点数,  $w_{ki}$  是权重值。

## 2 结果与讨论

2.1 RBFNN 模型: 表 1 中各配伍组以其数值法编码, 如第 1 组即编码为“111111”, 作为 RBFNN 的输入向量, 而该组的平均药效作为 RBFNN 的预期输出。需要指出的是, 平均药效在输入前应进行归一化处理:  $x(i) = (x_0(i) - \text{Min}(x_0)) / (\text{Max}(x_0) - \text{Min}(x_0))$ ,  $x_0(i)$  表示第  $i$  组的原有平均药效,  $\text{Min}/\text{Max}$  分别表示取最小/最大值。实验使用 MATLAB 软件(R2007a 版)附带的神经网络工具箱进行 RBFNN 计算, 其中网络设计函数 newrb 有 2 个主要参数需要优化确定, 即平方和误差  $eg$  和展开常量  $sc$ , 其他参数采用默认值, 即隐含层数为 1, 输入层、隐含层和输出层的节点数分别为 6、6、1。

为了将预测药效与实验值进行比较, 本实验采用预测误差均方根(RMSEP)作为评价指标, 其定义为  $RMSEP = \sqrt{\sum (x^{\text{Predicted}} - x^{\text{Observed}})^2 / N}$ , 式中  $x^{\text{Predicted}}$  和  $x^{\text{Observed}}$  分别表示预测药效和实验值, 而  $N$  表示训练集样本数。

一般来说, 评价预测模型的优劣可以进行自检验、他检验或留一法检验<sup>[7]</sup>。自检验是将训练集直接代入 RBFNN 模型进行验算, 所以精度一般偏高, 在一定程度上可说明输入向量具有可分性。他检验则是对与训练集完全独立的样本集进行验算, 故精度偏低, 但却是一种较为严格的评价方法, 反映了输入向量和预测方法的泛化性, 能够客观估计输入向量的聚类特性。三者中, 以留一法检验最为严格和客观, 它是从数据集中依次将每个样本单独取出, 用剩下的样本集进行训练, 对取出的单个样本进行测试; 重复这样的过程, 直至数据集中每个样本都被单独测试过。

本实验采用网格法搜索参数区间 ( $2^{-10}$ ,  $2^{-9}$ , ...,

$2^{10}$ ), 确定  $eg = 2^{-3}$ ,  $sc = 2^4$  为最优参数, 此时留一法检验的 RMSEP 为最小, 仅为 0.1565。该参数组合下, 模型的自检验和留一法检验的结果见图 1。可以看到, 模型的预测值与实验值均吻合较好, 线性相关系数分别为 0.944、0.846。

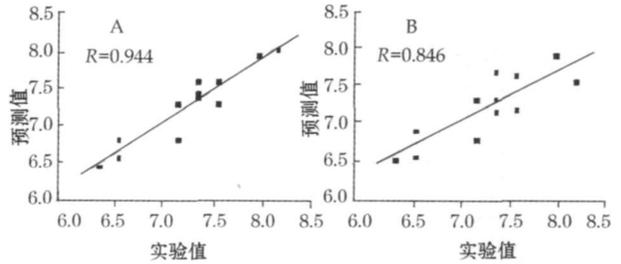


图 1 RBFNN 模型的自检验(A)和留一法(B)检验结果  
Fig 1 Detection of RBFNN model by self-consistency test (A) and jackknife test (B)

2.2 复方药效模拟: 利用所建立的 RBFNN 模型, 可以模拟本复方不同组合的药效。表 2 列出的是留一法检验所建模型的预测药效, 因留一法检验需运算  $N$  次(训练集样本数), 故表 2 中预测值实为  $N$  组预测值的平均结果, 其中最大药效组合为 ABEF, 最小药效组合为 CD, 药效区间为 6.269~8.327。与实验值相比, 12 个配伍组的平均绝对误差仅为 0.119, 小于 SVM (support vector machine, 支持向量机) 模型的预测误差; 尽管文献报道<sup>[2]</sup> 的预测误差更小, 但这是因为它的样本数少, 而且仅是自检验的结果, 本质上与留一法结果没有可比性。综上, 本实验所建模型具有良好的准确性, 可为进一步实验验证提供参考。

各组分对复方药效的贡献可以定义为配伍组中该组分 1 水平的平均药效和 2 水平的平均药效之差, 即  $M_i = \bar{x}_{1i} - \bar{x}_{2i}$ ,  $i = A, B, \dots, F$ 。将 RBFNN 的预测药效代入, 即可求出各组分的药效贡献值分别是: 0.205、0.648、-0.148、-0.181、0.584 和 0.292, 故各组分的重要程度排序为 B、E、F、A、C、D, 其中 B、E、F、A 为正性作用组分, 而 C、D 为负性作用组分。

## 3 结论

对中药复方有效部位中主要药效物质基础或者有效药效物质群的研究, 是中药复方化学研究的重点。本实验基于正交设计和 RBF 神经网络, 提出了中药复方药效的预测模型, 继而可选择性地对预测药效较好的组合或部位进行药效学实验, 快速、准确确定中药复方的有效部位, 并分析其质和量的变化与药效的关系, 建立复方量效模型。

表 2 不同组合的药效预测

Table 2 Prediction of responses in different components combination

配伍组	实验值	文献值 <sup>[2]</sup>		SVM <sup>a)</sup>		RBF NN		配伍组	实验值	文献值 <sup>[2]</sup>		SVM <sup>a)</sup>		RBF NN	
		预测值	误差	预测值	误差	预测值	误差			预测值	误差	预测值	误差	预测值	误差
ABCDEF	8 000	8 000	0	7 816	-0.184	7.997	-0.003	BC	/	7.133	/	7.252	/	7.099	/
ABC	7 400	7 400	0	7 442	0.042	7.304	-0.096	BCD	/	6.867	/	6.970	/	6.917	/
ABE	8 200	n a <sup>b)</sup>	/	7 990	-0.210	8.037	-0.163	BCDE	/	7.400	/	7.363	/	7.501	/
ADF	7 200	6 933	-0.267	7 013	-0.187	6.914	-0.286	BCDEF	/	7.733	/	7.643	/	7.792	/
AEF	7 400	n a <sup>b)</sup>	/	7 598	0.198	7.679	0.279	CDEF	/	7.067	/	7.124	/	7.146	/
ACD	6 400	n a <sup>b)</sup>	/	6 594	0.194	6.472	0.072	DE	/	6.867	/	6.918	/	7.001	/
BCF	7 400	n a <sup>b)</sup>	/	7 523	0.123	7.391	-0.009	DEF	/	7.200	/	7.196	/	7.293	/
BDF	7 200	7.333	0.133	7.360	0.160	7.357	0.157	EF	/	7.467	/	7.439	/	7.473	/
BDE	7 600	n a <sup>b)</sup>	/	7 493	-0.107	7.649	0.049	ABD	/	7.267	/	7.322	/	7.270	/
CDE	6 600	6.733	0.133	6.801	0.201	6.563	-0.037	ABDE	/	7.800	/	7.716	/	7.855	/
CEF	7 600	n a <sup>b)</sup>	/	7.398	-0.202	7.326	-0.274	ABDEF	/	8.133	/	7.924	/	8.145	/
—	6 600	n a <sup>b)</sup>	/	6.795	0.195	6.598	-0.002	ACE	/	7.267	/	7.269	/	7.240	/
ABEF	/	8.400	/	8.165	/	8.327	/	ACEF	/	7.600	/	7.533	/	7.531	/
CD	/	6.200	/	6.424	/	6.269	/	BCE	/	7.667	/	7.667	/	7.682	/
AB	/	7.533	/	7.574	/	7.452	/	BCEF	/	8.000	/	7.919	/	7.973	/
ABCD	/	7.133	/	7.164	/	7.121	/	CDF	/	6.533	/	6.766	/	6.563	/
ABCDE	/	7.667	/	7.562	/	7.706	/	平均绝对误差 <sup>c)</sup>		0.107		0.167		0.119	

a) 主要参数: LibSVM 2.81 工具包, RBF 核函数,  $\gamma=2^{-4}$ ,  $C=2^3$ , 留一法检验; b) 未报道; c) 平均绝对误差 =  $\sum |error| / N$ , 其中  $N$  为 5 或 12

a) main parameters: LibSVM 2.81 package, RBF kernel,  $\gamma=2^{-4}$ ,  $C=2^3$ , jackknife test; b) not available; c) mean absolute error =  $\sum |error| / N$ , where  $N$  equals to 5 or 12

参考文献:

[1] 严永清. 中药现代研究的思路与方法 [M]. 北京: 化学工业出版社, 2006

[2] 郑青山, 何迎春, 杨娟, 等. 药物相互作用分析与复方药效模拟: 一个方法学研究 [J]. 中国药理学通报, 2007, 23(8): 1106-1112

[3] 姚美村, 乔延江, 袁月梅, 等. 基于人工神经网络方法的中药功效归类研究 [J]. 中国中药杂志, 2003, 28(7): 689-691

[4] 范晓辉, 程翼宇. 中药复杂组效关系的变结构神经网络辨识方法 [J]. 高等学校化学学报, 2004, 25(11): 2004-2009

[5] 王泽, 顾勤兰, 王志祥. 基于遗传神经网络的中药葛根及同属植物的质量评价研究 [J]. 中草药, 2005, 36(10): 1580-1582

[6] 张玉瑞, 陈剑波. 基于 RBF 神经网络的时间序列预测 [J]. 计算机工程与应用, 2005, 41(11): 74-76

[7] Chen C, Tian Y X, Zou X Y, et al. Prediction of protein secondary structure content using support vector machine [J]. Talanta, 2007, 71(5): 2069-2073

## 五福化毒丸的质量标准研究

郑瑾, 张丽娟\*, 宋新波, 杨振容, 刘黄刚  
(天津中医药大学, 天津 300193)

摘要: 目的 建立中成药五福化毒丸的质量标准。方法 采用显微和薄层鉴别法进行定性鉴别, 采用高效液相色谱法进行测定。结果 显微法鉴别出成药中 11 味药; 采用薄层色谱法定性鉴别了方中的黄连和赤芍, 色谱斑点清晰, 分离度良好, 阴性无干扰; 以 HPLC 法测定该丸剂中甘草酸和连翘苷, 方法精密度、稳定性、重复性良好, 回收率分别为 95.75% 和 98.49%, RSD 分别为 0.55% 和 1.12% ( $n=6$ )。结论 本实验所确定的质量分析方法稳定可靠, 可以用于五福化毒丸的质量标准研究。

关键词: 五福化毒丸; 显微鉴别; 薄层色谱; 高效液相色谱; 甘草酸; 连翘苷

中图分类号: R286.02 文献标识码: B 文章编号: 0253-2670(2009)08-1241-05

五福化毒丸是《中国药典》2005 年版一部收载的传统中药成方制剂, 全方由水牛角浓缩粉、连翘、

青黛、黄连、牛蒡子(炒)、玄参、地黄、桔梗、芒硝、赤芍、甘草 11 味中药组成的传统蜜丸。其具有清热解

\* 收稿日期: 2008-10-27

作者简介: 郑瑾(1983-), 女, 天津人, 在读硕士, 主要从事中药提取纯化工艺和质量控制的研究。

Tel: (022)23111486 E-mail: nicolezhengjin@163.com

\* 通讯作者 张丽娟