

· 药物研究 ·

贝叶斯规整化神经网络模型

预测吲哚烷胺类化合物对 5-HT_{1B/1D} 受体亲和力*

温秋玲, 杨 博, 戴 康

(中南民族大学药学院, 武汉 430074)

【摘要】 目的 利用贝叶斯规整化神经网络模型研究 45 种吲哚烷胺对 5-羟色胺(5-HT)_{1B/1D} 受体激动活性的定量构效关系。方法 选取 115 种与结构相关的拓扑参数、几何参数和疏水性参数等, 通过主成分分析法进行参数减元, 建立基于 10 种主成分变量的活性预测贝叶斯规整化神经网络模型, 并利用去一法(LEAVE-ONE-OUT)对模型进行交叉验证。结果 应用残差绝对值的平均值(MAE)进行筛选, 得到隐含层神经元数目为 10 的模型为最佳模型。在该模型下, 吲哚烷胺对 5-HT_{1B} 受体和 5-HT_{1D} 受体亲和力的实验值和预测值一元相关系数平方(R^2)分别为 0.990 5 和 0.988 7。结论 模型显示吲哚烷胺 5-HT 受体激动作用与其结构有密切关系。贝叶斯规整化神经网络结合主成分分析方法有良好的预测能力, 且稳定可靠, 有望在 5-HT_{1B/1D} 受体激动药新药设计中得到广泛应用。

【关键词】 吲哚烷胺; 5-羟色胺_{1B/1D} 受体; 定量构效关系; 贝叶斯规整化神经网络

【中图分类号】 R965; R971 **【文献标识码】** A **【文章编号】** 1004-0781(2010)05-0555-04

The Bayesian Regularized Neural Network Prediction Model for Indolealkylamines Affinity to 5-HT_{1B/1D} Receptors

WEN Qiu-ling, YANG Bo, DAI Kang (College of Pharmacology, South-Central University for Nationalities, Wuhan 430074, China)

ABSTRACT Objective A quantitative structure-activity relationship (QSAR) model was built on the base of 45 indolealkylamines and by the method of Bayesian regularized neural network (BRNN). This model was used to predict activities (anti-migraine) of 5-HT_{1B/1D} receptor agonists. **Methods** One hundred and fifteen QSAR indices were used to elucidate the structural characters of indolealkylamines, which included topological, geometrical and hydrophobicity indices and so on. 115 variables were compressed by principle component analysis (PCA), and several PCA variables were chosen as input ones of model. BRNN model was trained and LEAVE-ONE-OUT (LOO) was used as a cross-validation method. **Results** In the present study, we obtained an optimized model with ten nerve centers in the hidden layer. The correlation coefficient for predicted vs experimental value was 0.990 5 for 5-HT_{1B} and 0.988 7 for 5-HT_{1D}. **Conclusions** The results indicate that the structures of indolealkylamines have good relationship with the affinity to the receptors. They also show that the BRNN model exhibits a good predictive ability and is robustness. We expect that this method may be widely applied in the area of new drug design for anti-migraine.

KEY WORDS Indolealkylamines; 5-HT_{1B/1D} receptors; QSAR; Bayesian regularized neural network

偏头痛属于自主神经系统疾病, 普通人群发病率高。过去人们认为神经和体液相互作用所产生的某些生化媒介物是偏头痛发病的重要因素。随着对偏头痛发病机制的深入研究, 人们认识到它们的有效作用来自于类似 5-羟色胺 1 (5-HT₁) 受体尤其是 5-HT_{1B/1D} 受

体的活性。1995 年, 舒马曲坦作为第一个专门抗偏头痛药物进入市场, 随后一系列曲坦类选择性 5-HT_{1B/1D} 受体激动药被逐步开发出来。目前, 国外药理学工作者对曲坦类药物研究很多, 主要通过药物分子设计-合成-药理学评价, 从吲哚类生物碱中开发抗偏头痛药物^[1,2]。然而, 对于偏头痛类药物的定量构效关系研究, 笔者尚未见国内有相关文献报道。笔者在本实验中通过定量构效关系(QSAR)研究吲哚环取代基对 5-HT_{1B/1D} 受体活性的影响, 为尽可能显示结构方面信息, 笔者尽可能多选取二维和三维 QSAR 参数, 并在此基础上建立活性预测模型。

曲坦类药物是吲哚类烷胺, 属 5-HT 化学衍生物,

【收稿日期】 2009-06-11 **【修回日期】** 2010-01-06

【基金项目】 * 国家自然科学基金资助项目(基金编号: 30471432)

【作者简介】 温秋玲(1985-), 女, 广西贵港人, 在读硕士, 研究方向: 药物合成、药物分子设计。电话: (0)13707181484, E-mail: wenquanjiufang@163.com。

【通讯作者】 戴 康, 男, 湖南长沙人, 副教授, 研究方向: 药物合成、药物分子设计。E-mail: kangdai688@163.com。

均含有一个咪唑环和一个基本氮原子,两者通过通常含有两个亚甲基的生物碱链相连^[3]。曲坦类药物属选择性 5-HT_{1B/1D}受体激动药,由于在咪唑环上 3-位和 5-位取代基的改变导致它们对受体的亲和力大小及生物膜透过性方面各不相同。药物在吸收、分布、代谢、排泄等药动学过程涉及到易化扩散、主动转运等非线性过程。为更好地模拟药物结构与受体亲和力的关系,与线性预测模型相比,神经网络模型具有更好的预测性能。反向神经网络是经常用于结构活性预测的神经网络,但存在易陷入局部最小化和过拟合等缺点。利用贝叶斯理论自动选择最优正则参数,在此基础上建立贝叶斯规整化神经网络(BRNN)具有传统神经网络快速收敛和充分利用数据的先验信息优点^[4,5],并具有稳健、拟合良好且概括能力强等优点,在毒理学、药效学及药物对肝药酶反应活性上得到成功应用。但在化合物分子对受体亲和力研究领域的应用笔者尚未见报道。笔者在本实验中应用 BRNN 方法建立咪唑烷胺对 5-HT_{1B/1D}受体的亲和力预测模型,现报道如下。

1 材料与方法

1.1 数据集与 QSAR 参数 笔者从文献[2,3]中选取 45 个具有 5-HT_{1B/1D}受体激动活性的化合物建立模型。为更全面地描述化合物结构,选取包括分子结构、拓扑和脂水分配系数、分子表面积、分子体积等 115 种二维和三维 QSAR 参数。115 种参数通过 Discovery Studio 程序计算获得。所有分子对 5-HT_{1B}受体和 5-HT_{1D}受体的亲和力 pIC₅₀数据均来自文献[6~9]。

1.2 主成分分析 首先,将所有 QSAR 参数通过式①进行标准化转换,使其数值控制在[0,1]范围内,①

$$P_{标准} = 2 \times \frac{P_{原} - P_{min}}{P_{max} - P_{min}}$$

然后对已完成标准化转换的参数进行 Principal component analysis 程序计算,得到化合物的 115 个主成分变量。最后根据各主成分变量方差占有率大小选取变量进入神经网络对化合物对受体亲和力大小的预测。该方法详见文献[10]。

1.3 模型的构建 笔者在本实验中根据贝叶斯原理,以训练数据集为基础,通过式②确定最似然的正规化参数的 BRNN 模型。

$$② P_{(\alpha, \beta/D, M)} = \frac{P_{(D/\alpha, \beta, M)} P_{(\alpha, \beta/M)}}{P_{(D/M)}}$$

其中 D 表示训练数据集, M 表示网络模型, P_(α,β/D,M) 为后验分布概率及似然函数。

将数据集随机分成两个子集,其中训练组(40 个化合物)用来训练神经网络,测试组(5 个化合物)用来检验模型。神经网络模型、主成分分析在 MATLAB 中调试通过,并用去一交叉验证法(LOO, LEAVE ONE

OUT)进行模型验证。

1.4 统计学方法 BRNN 预测完成后,得到咪唑类生物碱对 5-HT_{1B/1D}受体的亲和力 pIC₅₀预测值。将此预测值与实验值进行比较,并进行回归分析。应用相关系数平方(R²)和实验值-预测值一元回归系数(A)及残差绝对值平均值(MAE)进行检验。对所选数据进行多元线性回归分析,同时对预测值与文献数据进行比较,以验证模型的效果。这些统计分析工作在 Excel 和 SAS 中完成。

2 结果

2.1 主成分分析 笔者在本实验中选取化合物 QSAR 参数 115 个,采用主成分分析方法对参数进行减元。图 1 显示咪唑烷胺的主成分变量 1 和主成分变量 2 之间相关关系,该散点图显示各化合物主成分变量 1 和主成分变量 2 均匀分布,说明主成分变量具有代表性。为达到对参数减元的目的,通过计算各主成分的方差占总方差的比值,选取方差占有率最大的几个主成分变量组成数据集。笔者首先选取主成分变量 1 到主成分变量 4 组成数据集,进行贝叶斯神经网络预测化合物对受体的活性,所得结果显示训练组和测试组预测效果特别差且很不稳定(结果未列出)。为改善模型存在的问题,笔者逐步增加选取的主成分变量个数,发现模型预测效果和稳定性得到改善,当数据集中主成分变量增大到 10 个,模型预测效果和稳定性都很好。因此,笔者最终选取前 10 个主成分变量(数据集总方差占有率为 0.958 3)组成数据集,也就是使组成数据集的 n 个主成分变量的方差占有率之和 ≥ 0.95,结果主成分变量分别为 4, 9, 10, 11, 13 个,方差占有率分别为 0.856 7, 0.949 3, 0.958 3, 0.965 8, 0.976 9。

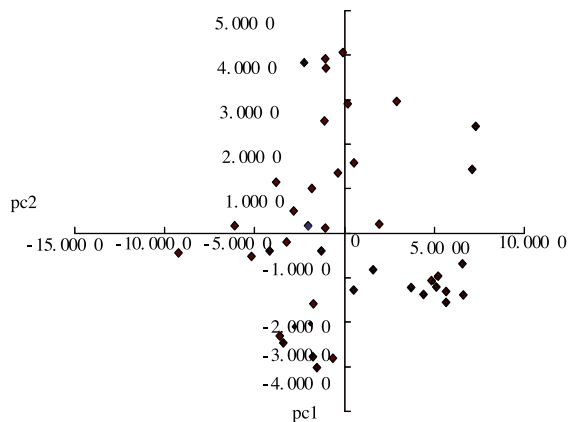


图 1 主成分变量之间的相关性

Fig. 1 Correlation analysis of the principal components

2.2 神经网络结构的确定 BRNN 采用三层神经网络设计,分别是输入层、隐含层和输出层,其中输入层神经元 10 个,等于变量个数,输出层 1 个,输出预测数

据。最终模型形成 10- N -1 模型,其中隐含层的神经元数目 N 是决定模型预测效果的关键因素。笔者通过 BRNN 模型预测吲哚烷胺对 5-HT_{1B}受体和 5-HT_{1D}受体活性预测,以残差绝对值的平均值(MAE)为指标,考察 N 在 6~20 范围内的神经网络模型(图 2),发现:①在对吲哚烷胺对 5-HT_{1B}受体的活性预测中, $N < 8$ 时和对 5-HT_{1D}受体活性预测中 $N < 9$ 时,神经网络模型预测效果较差,模型也不稳定,训练进度和效果受初始化的影响很大;②而在对吲哚烷胺对 5-HT_{1B}受体的活性预测中 $N > 13$ 和对 5-HT_{1D}受体活性预测中当 $N > 14$ 时,由于发生过拟合,神经网络模型的预测效果反而变差。故笔者最终选取 $N = 6, 8, 10$ 和 12 的神经网络作为对吲哚烷胺对 5-HT_{1B}受体的活性预测和 $N = 7, 10, 12$ 和 13 的神经网络作为对吲哚烷胺对 5-HT_{1D}受体的活性预测的优化模型。

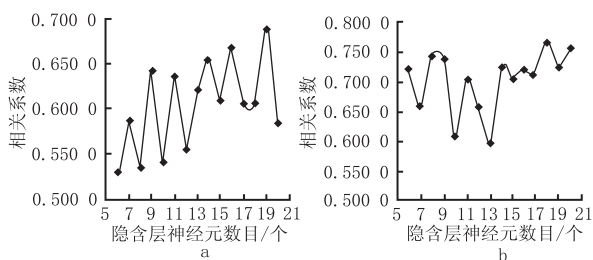


图 2 BRNN 神经网络性能与隐含层神经元数目关系

a. 5-HT_{1B}受体; b. 5-HT_{1D}受体

Fig. 2 Relationship between neural network performance and hidden neuron number

a. 5-HT_{1B} receptor; b. 5-HT_{1D} receptor

2.3 模型预测效果 根据神经网络优化结果,笔者选取 $N = 6, 8, 10, 12$ 的神经网络作为对吲哚烷胺对 5-HT_{1B}受体的活性预测,选取 $N = 7, 10, 12, 13$ 的神经网络作为对吲哚烷胺对 5-HT_{1D}受体的活性预测优化模型。表 1 显示经过优化的 BRNN 对训练组数据的预测结果,相关系数平方和实验值-预测值一元回归系数越接近 1,模型越稳定,预测效果越好。通过比较,笔者得到 10-10-1 结构神经网络是对吲哚烷胺对 5-HT_{1B}受体的活性预测更优的模型和 10-10-1 结构神经网络是对吲哚烷胺对 5-HT_{1D}受体活性预测的最优化模型,因为它们具有更好的拟合度。图 3 显示这两个最优化神经网络模型的预测值和实验值之间的线性拟合关系。虽然对吲哚烷胺对 5-HT_{1B}受体的活性预测模型 10-6-1 和对吲哚烷胺对 5-HT_{1B}受体的活性预测模型 10-7-1 的相关系数均是这么多个模型中最接近 1,但由于它们均存在不稳定性,所以笔者选取相关系数次之的模型作为对吲哚烷胺对 5-HT_{1B/1D}受体活性的最优化预测模型。通过比较,这两个模型相关系数均比多

元回归模型大且接近 1,可见这两个最优化的神经网络模型较之多元回归模型预测效果更好。

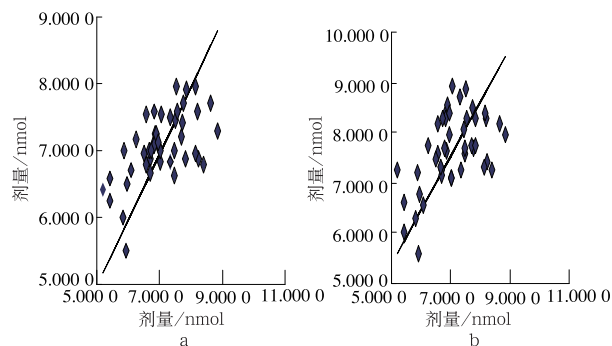


图 3 最优化神经网络模型预测值与实验值的拟合关系
a. 5-HT_{1B}受体拟合图; b. 5-HT_{1D}受体拟合图

Fig. 3 The fitting relationship between predictive value and experimental value of the optimized neural network model

a. the fitting figure of 5-HT_{1B} receptor; b. the fitting figure of 5-HT_{1D} receptor

表 1 不同模型相关系数与实验值-预测值的一元回归系数比较

Tab. 1 Comparison of contiguous coefficient and one-dimensional regression coefficient between predictive value and experimental value in different models

受体类型与模型	相关系数 R^2	实验值-预测值的 一元回归系数 A
5-HT _{1B} 受体		
10-6-1	0.991 1	0.989 1
10-8-1	0.990 3	1.002 4
10-10-1	0.990 5	0.992 7
10-12-1	0.989 1	0.995 3
多元回归	0.668 0	
5-HT _{1D} 受体		
10-7-1	0.989 5	1.068 1
10-10-1	0.988 7	1.076 4
10-12-1	0.986 8	1.066 7
10-13-1	0.988 5	1.089 2
多元回归	0.769 8	

3 讨论

笔者通过计算选取与结构密切相关的拓扑参数、几何参数以及疏水性参数等 115 个参数,通过主成分分析方法进行参数减元,采用 BRNN 预测 45 个吲哚烷胺对 5-HT_{1B/1D}受体的活性。通过 QSAR 研究发现,吲哚烷胺的吲哚母环上 3 位和 5 位上的取代基对化合物的活性影响很大。当 3 位取代基为 2 或 3 个亚甲基连接叔胺氮原子的生物碱链以及母环 5 位上为吸电子取代基(如磺酰胺基、砜基和三氮唑等)取代时,吲哚烷胺对 5-HT_{1B/1D}受体的活性较高,这点在已经上市的成

功用于治疗偏头痛的曲坦类药物的临床研究已得到证实。同时由于这些取代基不同, 吲哚烷胺在 5-HT_{1B}受体和 5-HT_{1D}受体间呈现出一定选择性。

本研究结果表明, BRNN 在吲哚类生物碱对 5-HT_{1B/1D}受体活性的 QSAR 研究方面可靠, 具有拟合良好、稳健和概括能力强等优点, 克服了反向传播神经网络易陷于局部最小化和过拟合等缺点。后续研究将通过此模型进行新的吲哚烷胺对 5-HT_{1B/1D}受体的活性预测, 为后续合成工作提供依据, 这将使得药物开发成本降低, 周期缩短。可见神经网络在药物的研发领域具有广阔应用前景。

[DOI] 10.3870/yydb.2010.05.001

[参考文献]

[1] KUNAL R, PARTHA P R. Comparative QSAR studies of CYP1A2 inhibitor flavonoids using 2D and 3D descriptors [J]. *Chem Biol Drug Des*, 2008, 72(5): 370 - 382.

[2] DU Q S, HUANG R B, WEI Y T, *et al.* Fragment-based quantitative structure-activity relationship (FB-QSAR) for fragment-based drug design[J]. *J Comput Chem*, 2009, 30(2): 295 - 304.

[3] YU A M. Indolealkylamines: biotransformations and potential drug-drug interaction [J]. *AAPS J*, 2008, 10(2): 242 - 253.

[4] NICULESCU S P. Artificial neural networks and genetic algorithms in QSAR[J]. *J Mole Stru Theochem*, 2003, 122

(22): 71 - 83.

[5] 戴康, 王晓琼. 神经网络在药理学评价中的应用[J]. *医药导报*, 2007, 26(9): 1048 - 1049.

[6] RAZZAQUE A, HEALD M A, PICKARD J D, *et al.* Vasoconstriction in human isolated middle meningeal arteries: determining the contribution of 5-HT_{1B}-receptor and 5-HT_{1F}-receptor activation [J]. *Br J Clin Pharmacol*, 1999, 47(1): 75 - 82.

[7] RAMADAN N M, SKLJAREVSKI V, PHEBUSL A. 5-HT_{1F} receptor agonists in acute migraine treatment: a hypothesis [J]. *Cephalalgia*, 2003, 23(8): 776 - 785.

[8] KARAMJIT S J, VIKKI B, MICHAEL B, *et al.* Discover of 4-[3-(trans-3-Dimethylaminocyclobutyl)-1H-indole-5-ylmethyl]-(4S)-oxazolidin-2-one (4991W93), a 5-HT receptor partial agonists and a potent inhibitor of electrically induced plasma extravasation [J]. *J Med Chem*, 2001, 44(5): 681 - 693.

[9] JOSE' L, CASTRO L J, STREET A R, *et al.* 3-[2-(Pyrrolidin-1-yl) ethyl] indoles and 3-[3-(piperidin-1-yl) propyl] indoles: agonists for the 5-HT receptor with high selectivity over the h5-HT Subtype [J]. *J Med Chem*, 1997, 40, 3497 - 3500.

[10] WANG Y H, LI Y, LI Y H, *et al.* Modeling km values using electrotopological state: substrates for cytochrome P₄₅₀ 3A4-mediated metabolism [J]. *Bioorg Med Chem Lett*, 2005, 15: 4076 - 4084.

枣味不同提取物镇静催眠作用研究*

王晓雯, 陆兔林, 毛春芹, 徐佳佳

(南京中医药大学, 210046)

[摘要] 目的 研究枣味(酸枣仁、五味子)不同提取物的镇静催眠作用。方法 分别制备枣味水提取物、醇提取物、醇水双提合并物, 将小鼠随机分为空白组、对照组和实验组。观察各提取物对小鼠自主活动次数及睡眠时间的影响。结果 与空白组相比, 枣味醇提取物高剂量组和醇水双提取物中剂量组小鼠的自主活动次数明显减少; 各组小鼠睡眠时间明显延长, 其中醇提取物和醇水双提取物中、高剂量组催眠作用最显著。结论 采用醇提或醇水双提方法所得的酸枣仁、五味子提取物具有明显的镇静催眠作用。

[关键词] 酸枣仁; 五味子; 镇静; 催眠

[中图分类号] R286; R965 [文献标识码] A [文章编号] 1004 - 0781(2010)05 - 0558 - 03

Sedative and Hypnotic Effects of Different Extracts from *Semen Ziziphus Spinosa* and *Fructus Schisandrae Chinensis*

WANG Xiao-wen, LU Tu-lin, MAO Chun-qin, XU Jia-jia (Nanjing University of Traditional Chinese Medicine, Nanjing 210046, China)

ABSTRACT Objective To study the sedative and hypnotic effects of different *Semen Ziziphus spinosa* and *Fructus Schisandrae Chinensis* extracts. **Methods** The water extract, alcohol extract and alcohol-water extract of *Semen Ziziphus spinosa* and *Fructus Schisandrae Chinensis* were prepared. The mice were randomly divided into blank group, contrast group and