Study on the Balance Activity of Mice Based on BP-ANN

Lilan Zhu¹, Shaohui Du², Xiaoli Liang³

Department of Physical Education, Guangdong Industry Polytechnic, Guangzhou, Guangdong, China

*Corresponding author’s e-mail: zhulilan_8983@sina.com, dushaohui66@163.com, liangxiaoli66@yeah.net

Abstract. In order to study the quantitative structure-activity relationship (QSAR) of the balance activity (pH) for 30 benzodiazepinoxazole derivatives to male mice, the molecular electronegativity distance vector (MD) of above compounds was calculated by program according to molecular topological environment in this paper. The five-variable (M₆, M₄₆, M₂₅, M₆₃, M₇₀) QSAR model of pH for above compounds was constructed by using leaps-and-bounds regression method. The result demonstrates that the model is robustness and good prediction ability by using Rcv², F tests. The M₆, M₄₆, M₂₅ and M₇₀ were used as the input neurons of artificial neural network (ANN), and a 5:3:1 network architecture was employed. A satisfying BP-pH model could be constructed with the back-propagation algorithm, with the correlation coefficient (R²) and the standard error(SD) being 0.928 and 0.117, respectively, showing that the relationship between pH and these structural parameters has a good nonlinear correlation. Form the three parameters of the model, it is known that the dominant influence factors of increased balance activity are the microscopic fragments: –CH₃, –CH₂–, >C<, –NH–, –N< and –X in the molecules.

1. Introduction

Balance ability refers to the ability to maintain body posture and control body center of gravity, which is the basic ability of all static and dynamic activities. It can be said that almost all human activities are carried out under the condition of maintaining physical balance, especially athletes, who need good balance ability to be competent. The development of balance ability can generally be carried out through static and dynamic balance activities, such as standing on one foot, standing long jump, etc., but none involves drug control of human balance ability [1, 2]. Yoshimoto M. et al [3] give the effects of 30 benzoxazole derivatives on the balance ability (pH) of male mice, realize the control of drugs on the balance ability of mice, and develop a bio-mathematical model of balance ability. Yoshimoto M. et al. established a five-element equation for the balance ability of mice, and its determination coefficient (R2) was only 0.645, with low correlation and poor prediction ability. In order to further improve the quality of the model and increase the correlation, this paper studies the mathematical rules of the balance ability of mice [3] based on the quantitative structure-activity relationship (QSAR) method [4-10], reveals the microstructure affecting the balance activity (pH) of mice, provides scientific reference for the balance ability training, and provides a new attempt for the research of sports science.
2. Materials and methods

2.1. Experimental subject and balance activity data
Experimental animals: Male mice weighing 20-25g.
Experimental drug: The basic structures of 30 benzodiazepine derivative molecules [3] are shown in Fig.1, Wherein R₁, R₂, R₃, R₄, R₅ and R₆ are substituent groups [3].

![Figure 1. Basic structure of benzodiazepinoxazole derivatives](image)

Balance ability experiment: The above compounds were dissolved in 0.85% saline containing 0.5% tragacanth gum as a suspension. Immediately after oral administration to mice, the mice's front paws are suspended under a smooth glass rod (diameter 1.3 cm). Examining the dose that causes 50% of them to fall down within one minute, i.e. measuring the concentration required for 50% of them to lose balance, has been expressed by "ED50", and the unit is mol·dm⁻³. According to the equilibrium principle of physical chemistry, there is a logarithmic relationship between the free energy change (ΔGr) of chemical reaction and concentration. Therefore, the negative logarithm of ED₅₀ (pH=−logED₅₀) should be taken for modeling, and the specific data of pH are shown in Table 1[3].

2.2. Molecular electronegativity distance vector (MD)
One of the keys to QSAR research is to numerically characterize abstract molecular structures, to establish variables describing molecular structures. Commonly used methods are topological index, quantitative parameters, etc. According to the limitations of several famous topologies, Liu Shushen et al. [11-13] proposed an electrical distance vector (M₆) that can fully display topology, geometry and electrical characteristics. The calculation process is detailed in document [11-13]. Using the electrical distance vectors of 30 benzodiazepine drug molecules in the program, the MD entering the model is shown in Table 1.
Table 1. The molecular electronegativity distance vector ($M_D$) and balance ability ($p_H$) of benzodiazepinoxazole derivatives to male mice

| No. | $M_6$ | $M_{25}$ | $M_{46}$ | $M_{63}$ | $M_{70}$ | $p_{H,exp}$ | $p_{H,cal.1}$ | $p_{H,cal.2}$ |
|-----|-------|----------|----------|----------|----------|-------------|---------------|---------------|
| 1   | tr    | 0        | 4.6504   | -0.2796  | 0        | 0.0467      | 3.84          | 4.02          |
| 2   | ts    | 0        | 4.6844   | -0.2734  | 0        | 0.0486      | 4.12          | 4.15          |
| 3   | va    | 0        | 2.9278   | -0.1716  | 0        | 0.0300      | 4.03          | 4.34          |
| 4   | tr    | 0        | 2.9579   | -0.1683  | 0        | 0.0313      | 4.57          | 4.42          |
| 5   | tr    | 0        | 8.5289   | -0.7973  | 0        | 0.1523      | 4.32          | 4.18          |
| 6   | tr    | 0.0938   | 4.6051   | -0.2753  | 0.0985   | 0.0764      | 4.40          | 4.28          |
| 7   | ts    | 0        | 0        | 0        | 0        | 0           | 4.76          | 4.70          |
| 8   | va    | 0        | 4.6394   | -0.2692  | 0.0986   | 0.0776      | 4.73          | 4.61          |
| 9   | tr    | 0.0948   | 2.9107   | -0.1696  | 0.0617   | 0.0480      | 4.37          | 4.37          |
| 10  | tr    | 0        | 2.9411   | -0.1663  | 0.0620   | 0.0489      | 4.76          | 4.67          |
| 11  | tr    | 0.0772   | 0        | 0        | 0        | 0           | 4.25          | 4.51          |
| 12  | ts    | 0        | 0        | 0        | 0        | 0           | 4.49          | 4.70          |
| 13  | va    | 0        | 6.8689   | -0.6861  | 0.1375   | 0.1375      | 4.30          | 4.59          |
| 14  | tr    | 0        | 0        | 0        | 0        | 0           | 4.89          | 4.70          |
| 15  | tr    | 0.0925   | 12.3960  | -1.4008  | 0.2457   | 0.3252      | 4.08          | 4.24          |
| 16  | tr    | 0        | 8.6253   | -0.7794  | 0        | 0.1567      | 4.71          | 4.50          |
| 17  | ts    | 0.1378   | 4.5927   | -0.2753  | 0.0985   | 0.0812      | 4.61          | 4.38          |
| 18  | va    | 0.3741   | 8.5184   | -0.7855  | 0.1729   | 0.2171      | 4.48          | 4.69          |
| 19  | tr    | 0.1394   | 2.9030   | -0.1696  | 0.0617   | 0.0510      | 4.59          | 4.39          |
| 20  | tr    | 0        | 12.5280  | -1.3665  | 0.2453   | 0.3297      | 4.74          | 4.93          |
| 21  | tr    | 0.0949   | 8.4674   | -0.7855  | 0.1716   | 0.2044      | 4.71          | 4.84          |
| 22  | ts    | 0.3782   | 6.8780   | -0.6766  | 0.1362   | 0.1882      | 4.87          | 4.77          |
| 23  | va    | 0        | 6.9615   | -0.6711  | 0        | 0.1413      | 4.94          | 4.87          |
| 24  | tr    | 0        | 8.5646   | -0.7677  | 0.1717   | 0.2076      | 5.07          | 5.34          |
| 25  | tr    | 0.0959   | 6.8351   | -0.6766  | 0.1352   | 0.1773      | 5.38          | 5.00          |
| 26  | tr    | 0        | 10.9320  | -1.2517  | 0.2105   | 0.3054      | 5.39          | 5.25          |
| 27  | ts    | 0        | 6.9285   | -0.6616  | 0.1355   | 0.1801      | 5.66          | 5.47          |
| 28  | va    | 0.1360   | 12.3450  | -1.4008  | 0.2457   | 0.3466      | 5.43          | 5.09          |
| 29  | tr    | 0.1394   | 8.4364   | -0.7855  | 0.1716   | 0.2171      | 4.91          | 5.30          |
| 30  | tr    | 0.1409   | 6.8088   | -0.6766  | 0.1352   | 0.1882      | 5.27          | 5.38          | 5.26          |

2.3. Statistical regression analysis

The electrical distance vector ($M_D$) of 30 benzodiazepine drug molecules is the independent variable, and the equilibrium activity $p_H$ [3] of mice is the dependent variable. The QSAR model is established by regression analysis using the leaps-and-bounds regression (LBR) program. The quality of regression models is usually tested by Fisher statistics. Its critical value is $F_{(a,b-1),b}$, where $\alpha$ is the significance level, $b$ is the number of independent variables entering the model, and $f$ is the number of compounds contained in the model. Cross-validation correlation coefficient ($R_{cv}^2$) is generally used to evaluate the prediction ability of the model. According to "Jackknife" method, a model is established to predict the rejected compounds. The cross-validation coefficient ($R_{cv}^2$) is obtained by correlating the predicted values with the experimental values. It is generally believed that $R_{cv}^2$ is greater than 0.5, and the established model has good prediction ability [14].

At present, the most widely used ANN is a three-layer neural network based on back-Propagation (BP) algorithm [15-17]: input layer, hidden layer and output layer. Among them, the input layer accepts external data input, the hidden layer processes and converts the input data, and the output layer generates output (i.e. prediction) results. Each layer in the network contains several neurons, of which the number of neurons in the input layer and the output layer is determined by the
number of variables in the model, and the number of neurons in the hidden layer can be determined by trial calculation or rules.

3. Results and discussion

3.1. Multivariate linear QSAR model of pH

YOSHIMOTO M et al. [3] established a five-element equation for the equilibrium activity pH of 30 benzoxazole derivatives on mice by using structural parameters, with the determination coefficient R²=0.645 (R=0.803) and the estimated standard error S₀ = 0.293. In this paper, the best five-element QSAR model of pH [3] and MD is established by the leaps-and-bounds regression method as follows:

\[ pH = 4.703 - 2.507M_{6} + 7.909M_{46} - 0.118M_{25} - 8.891M_{63} + 44.582M_{70} \]

\[ f = 30, R^2=0.781(R=0.884), R_{cv}^2 = 0.656, F = 17.104, S_0 = 0.230 \]

3.2. Quality inspection of QSAR model for pH

Under the significance level \( \alpha=0.05 \), the Fisher critical value \( F_{0.05 (5, 24)} = 2.62 \) of model (1) was obtained by looking up the table. Fisher statistic (17.104) of model (1) is much larger than the critical value. It shows that 95% of the regression relationships expressed by this model are very significant and have good fitting degree, i.e. dependent variables are closely related to independent variables.

\( R_{cv}^2 \) of model (1) is 0.656, which is greater than 0.5, reflecting the good prediction ability of the model. The predicted value (pH.cal.1) and corresponding experimental value (pH.exp.) of the balance ability of mice given in model (1) are shown in Table 2. As can be seen from Table 1, they basically coincide. In addition, \( R^2=0.781(R=0.884) \) in this paper is obviously better than Yoshimoto M. et al's \( R^2=0.645(R=0.803) \), and its \( S_0 \) is only 69.3% of Yoshimoto Met al's [3], which proves the success of model (1).

3.3. Artificial neural network model

Although the quality of model (1) is obviously better than Yoshimoto's five-element regression equation, its correlation is slightly poor, and its R² is less than 0.8. The quality of the regression model is also determined by the determination coefficient (R², also known as error reduction ratio). As a rule of thumb, the equation with regression coefficient R² >0.80 is considered reasonable [18]. Therefore, on the basis of the linear model (1), the artificial neural network method is adopted to establish the nonlinear equation to improve the fitting level.

In order to avoid over-fitting and over-training of the established neural network model, the unit number (U) of the best hidden layer is found according to the suggested rules of Xu et al. [19], and the definition formula is:

\[ \rho(=N/M) \geq 1 \]

(2)

Where \( N \) and \( M \) are respectively the number of samples and the total weight of the network, respectively. \( M \) is defined as:

\[ M=(I+1)U+(U+1)Q \]

(3)

In formula (3): I, U and Q are the number of cells in the input layer, the hidden layer and the output layer, respectively. In this paper, \( I=5(M_6, M_{46}, M_{25}, M_{63}, \text{and} \ M_{70}) \), \( Q=1(\text{pH}) \) and \( N=30 \), and \( U=3.42 \) can be obtained by taking \( \rho=1.2 \). Therefore, a 5:3:1 network structure is adopted to build the model. In order to further avoid over-fitting and over-training, the data set is divided into three groups: training set, verification set and test set (tr, va and te in Table 2 in turn). The number of compounds in each set is 18, 6 and 6 in turn. The purpose of the verification set is to monitor the training process, that is, to stop training automatically when the error of the verification set starts to rise, to prevent over-training.
and over-fitting of the network, and to reduce the training time. The established model is $R^2 = 0.930$ ($R = 0.964$) for the training set and $R^2 = 0.928 (R = 0.963)$ for the whole. The two models are very close, which shows that the established model is stable and there is no over-training or over-fitting. The predicted values given by the model (see "pHcal. 2" in Table 1) are in good agreement with the experimental values. The estimated standard error SD is 0.117, which is only 39.9% of Yoshimoto M. et al. [3]. The BP-ANN model is obviously better than the research results of Yoshimoto et al. The weights and biases of BP-ANN model are given in Table 2.

Table 2. Weights and Biases of the BP-ANN Model

| Interlayer variation                  | Weights  | Biases  |
|--------------------------------------|----------|---------|
| From the input layer to the hidden layer | 47.800  | -16.258 | 87.274 | -14.823 | -106.420 |
| From the hidden layer to the output layer | -0.954 | -1.030 | -0.279 | -10.385 | -2.843  |
|                                      | 1.396    | 1.405   | 0.2098 | 14.207  | 3.517   |
|                                      | -7.106   | 81.919  | 80.738 | 14.207  | -6.139  |

3.4. QSAR Equation Structure Analysis

In the model (1), $M_6$ reflects the interaction of –CH$_3$ and –NH–, $M_{25}$ reflects the interaction of –CH$_2$– and –X, $M_{46}$ reflects the interaction of >C< and –X, $M_{63}$ reflects the interaction of –NH– and –X, $M_{70}$ reflects the interaction of –N< and –X. Therefore, the microstructure affecting the balance ability of benzodiazepine drugs to mice pH is mainly: –CH$_3$, –CH$_2$–, >C<, –NH–, –N<, –X. The first three are non-polar groups and the last three are highly electronegative polar groups. The former can have hydrophobic interaction with the target protein in mice, while the latter can form hydrogen bonds or coordination bonds with the target protein.

4. Conclusion

(1) Based on the electrical distance vector and the regression method of the best subset of variables, the best five-element QSAR model of benzoazxole drugs on the equilibrium pH of mice was established. Through statistical indexes $R^{cv}_2$ and $F$ test, the model has strong correlation and prediction ability. Its $R^2$ and $S_D$ are obviously superior to the research results of Yoshimoto M.

(2) The correlation ($R^2 = 0.928$, $S_D = 0.117$) of the BP-pH model established with $M_6$, $M_{46}$, $M_{25}$, $M_{63}$ and $M_{70}$ as input variables is significantly better than that of the five-variable regression equation ($R^2 = 0.781$, $S_D = 0.230$). The model shows that five variables have a good nonlinear relationship with pH, rather than a simple linear relationship.

(3) The microstructure that affects the balance ability (pH) of diazepine drugs to mice is –CH$_3$, –CH$_2$–, >C<, –NH–, –N<, –X. Therefore, the interaction between benzodiazepine drug molecules and target enzymes in mice is mainly hydrogen bond and hydrophobic interaction.

In this study, a bio-mathematical model of the balance ability of mice is proposed, which reveals the microstructure groups that affect the balance ability of mice and the mechanism of action with target enzymes in mice. It provides a theoretical reference for drug control for the training of human balance ability, and has certain enlightenment for scientific research of sports science.

Acknowledgments

This work was financially supported by Natural Science Foundation of Guangdong Industry Polytechnic (KJ2019-032) fund.

References

[1] Yuan K. J.; Wu J. H. Effects of square dance and taijiquan on the dynamic balance ability of middle-aged and elderly women. Journal of Beijing Sport University, 41 (2018), 82 - 86.

[2] Zhang Y., You Y. H., Liu B. l., Zhang M., Cui G.F., Zhang Q. X. Research of square dance and fitness walking exercise on static balance ability in middle-aged and older women. Journal of Beijing Sport University, 40 (2017), 46 - 50.
[3] Yoshimoto M., Kamioka T., Miyadera T., Kobayashi S., Takagi H., Tachikawa R. QSAR in minor tranquilizers benzodiazepinooxazole derivatives. Chem. Pharm. Bull., 25 (1977), 1378 - 1390.

[4] Zhu L. L. A Free-Wilson model of antioxidative activities for flavonoids compounds. Acta Nutrimenta Sinica, 34 (2012), 392 - 394.

[5] Zhu L. L. QSAR model for prediction of the aroma intensities of volatile compounds in chinese rice wine. Food Science, 34 (2013), 138 - 141.

[6] Liao L. M., Huang X., Lei G. D. Structural characterization and octanol/water partition coefficients (LogP) prediction for oxygen-containing organic compounds. Chin. J. Struct. Chem., 36 (2017), 1243 - 1250.

[7] Zhu L. L. Study on the relative activity for thiazide diuretics in doping based on artificial neural network. Journal of Guangdong Industry Polytechnic, 18 (2019), 7 - 10.

[8] Yu M., Huang J. J., Ma M., Fu R. Y., Yan Y., Zhang F. S., Yin J. F., Xie N. N. Zinc chelating activity and quantitative structure-activity relationship of tripeptides. Chem. J. Chin. Univ., 39 (2018), 234 - 240.

[9] Tong J. B., Qin S. S., Lei S., Wang Y. A 3D-QSAR study of HIV-1 integrase inhibitors using RASMS and Topomer CoMFA. Chin. J. Struct. Chem., 38 (2019), 867 - 881.

[10] Wang C., Feng C. J. QSAR studies on the inhibitory activity of levofloxacin-thiadiazole HDACi conjugates to histone deacetylases. Chin. J. Struct. Chem., 37: 1679 - 1688.

[11] Liu S. S., Yin C. S., Li Z. L., Cai S. X. (2001)QSAR study of steroid benchmark and dipeptides based on MEDV-13. J. Chem. Inf. Comput. Sci., 41 (2018), 321 - 329.

[12] Liu S. S., Li Z. L., Yin C. S., Wang L. S. VSMP: a novel variable selection and modeling method based on the prediction. J Chem Inf Comput Sci, 43(2003), 964 - 969.

[13] Jing G. H., Li X. L., Zhou Z. M. Quantitative structure-activity relationship (QSAR) study of toxicity of substituted aromatic compounds to photobacterium phosphoreum. Chin. J. Struct. Chem., 29(2010), 1189-1196.

[14] Xia S., Feng Y., Cheng J.G., Luo H.B., Li Z., Li Z. M. QAAR exploration on pesticides with high solubility: An investigation on sulfonyleurea herbicide dimers formed through p-p stacking interactions. Chin. Chem. Let., 25(2014), 973 – 977.

[15] Strahinja Z. K., Sanja O. P-K., Lidija R. J., Pavle T. J., Evgenija A. D., Jovana J. A. Comprehensive QSRR modeling as a starting point in characterization and further development of anticancer drugs based on 17α-picolyl and 17 (E)-picolinylidene androstane structures. Euro. J. Pharm. Sci., 93 (2016), 1 - 10.

[16] Nanda G., Basudeb A., Tapan K. G. A QSAR study of antiplatelet agents using artificial neural network - correlation with micelle-water partition coefficient, Bioorganic & Medic. Chem. Letters, 7 (1997), 877 - 880.

[17] Hadi N., Abbas F., Mehrab N. (2012) Quantitative structure-retention relationship analysis of nanoparticle compounds by GA-PLS, GA-KPLS and L-M ANN. Chin. J.Struct. Chem., 31: 569 - 573.

[18] Xu X., Li X. G. QSAR for predicting biodegradation rates of polycyclic aromatic hydrocarbons in aqueous systems. Chin. J. Struct. Chem., 31 (2012), 1212 – 1221.

[19] Xu L., Shao X. G. Methods of chemometrics. Beijing: Science Press, 2004, pp: 279 - 298.