THE METHOD OF ARTIFICIAL NEURAL NETWORK APPLIED TO EXPLORE THE EFFECTING FACTORS OF HEPATIC CANCER RECURRENCE AFTER HEPATECTOMY *

HE, Jia HE,Xian-min
Department of health statistics, Second Military Medical University
Shanghai 200433, China

ZHANG, Zhi-jian
Eastern Hepatobiliary Surgery Hospital, Shanghai 200438, China

[ABSTRACT] **Objective:** To explore the affecting factors of liver cancer recurrence after hepatectomy. **Method:** The BP artificial neural network. **Result:** To select the affecting factors with statistical significance to liver cancer prognosis. The eighteen factors are selected by uni-factor analysis. The nine factors are selected by multi-factor analysis. **Conclusion**: The nine factors selected can be as important indexes to evaluate the recurrence of liver cancer after hepatectomy. The artificial neural network is a better method to analyze the clinical data. The study can provide the basis with scientific and objective data for evaluating prognosis of liver cancer.
[KEY WORDS] artificial neural network; liver cancer; affecting factors

Primary hepatic cancer (PHC) is one of the most malignant tumors, and is threatening the people's health. Now, hepatectomy is still the first choice of treatment for PHC, but the recurrence rate is higher, and life quality of patients is lower. The prognostic factors of PHC are complex, and the influence of them can not be measured by number. In this paper, artificial neural network (ANN) is used to analyze the data of 1457 cases from the Eastern Hepatobiliary Surgery Hospital and to study the influencing extent of factors on the recurrence of PHC. This study can provide the basis with scientific and objective data for evaluating prognosis of liver cancer patients.

1.Data

Data from the patients which had curative or relatively curative hepatectomy of PHC pathologically proved in our hospital between Jan. 1 1990 and Dec. 31 1995. In Jan. 1997, 1725 questionaires were sent out. With various methods of follow-up, 1457 ones were reclaimed, and the follow-up rate (84.5%) is higher relatively. Among them, 974 patients recurred, 483 censored (including lost follow-up, did not recur and died of other causes and so on). Fifty four possible influencing factor indexes were collected in this study, including patients' general state, symptoms and signs, laboratory examination, operation condition, tumor condition, postoperative situation, postoperative therapy, and so on.

2.Methods

ANN was used in this paper. ANN is a new, edge and crossed subject that includes computer science, informatics and medicine. The method of back propagation (BP) is used more universally. It derived from the biology neural network, and includes one input layer (corresponding to the explanatory variables in multiple analysis), one or more hidden layers (no corresponding object) and one output layer (corresponding to responding variables).

*** The study is supportrd by Natural Science Foundation in China (NSFC)**

Before survival analysis with BP network, the training must be done to adjust the coefficients of explanatory variables. The criterion is to reduce the error between the network's actual and expect output vectors of training data set. In the course of training, the error feeds back to the network, and the link weights between input and hidden layers, between hidden and output layers are adjusted, so this method calls BP neural network. With a non-traditional and non-linear way, BP network deals with the information. It does not require the normality and independence of the variables, so it can be used to analyze the data with outlier and multiple collinearity. On theory, BP network has superiority to traditional methods. The statistical method used in this paper is maximum likelihood estimate.

In this paper, the method of BP neural network-Cox regression model was used. First univariate analysis was used to analyze the 54 influencing factors, and then multivariate analysis was used to analyze the factors with statistical significance in univariate analysis, and the influencing extent of the factors was calculated. The software NoSA used in this paper was developed by XIA-Jielai, professor of department of health statistics of Fourth Military Medical University, and supported by the Natural Science Foundation in China(NSFC).

3.Results

(1)Univariate analysis

Each of the 54 factors was analyzed with BP neural network, 32 of them were of statistical significance, according to the clinical practice, they are combined into 18 factors (Tab 1). From tab 1, we can know the influencing extent of each factor on recurrence of PHC. The values of the third column are all positive, which indicate the greater the indexes measuring value is, the shorter the postoperative disease-free survival time is. For instance, the first factor: 'whether or not accepted TAE therapy before operation', the coefficient is positive, so for patients who did not accept TAE therapy before operation, the postoperative disease-free survival time was shorter. The fourth column is relative risk(RR), the greater the RR is, the more disadvantageous the prognosis is. If the influencing factor is multi-category data, the result is that the group compares with the group measured the smallest value. For example, the fourth factor is AFP, and the value was classified into seven groups according to magnitude of AFP. The first group: AFP<=20. The statistical results of other groups are to be compared with the first group. We can see, as AFP rises, the coefficient, RR values and Wald increase also, but the P value decreases. It indicates that the greater the AFP is, the shorter the postoperative disease-free survival time is, so AFP is an important index indicating the prognosis of PHC. We can see, the RR of 'AFP after operation' is the biggest (5.85), that is to say, 'AFP not decreased to normal' compared with 'decreased to normal', for disease–free survival time after hepatectomy , the risk of the former is 5.85 times to that of the later. We can see also, the RR of 'Portal tumor thrombus' is bigger, the risk of the patients with tumor thrombus is 4.64 times to that of who hadn't. According to the magnitude of RR (RR>3), the factors with serious effect on the prognosis are: Portal tumor thrombus, Distance from verge of resection to tumor (<1cm), TNM stage (IV stage) and Postoperative AFP value.

(2)Multivariate analysis

To recognize the factors influencing the postoperative recurrence for PHC after hepatectomy on the whole, 18 significant factors have been input the BP neural network-Cox regression model and been analyzed, results as tab 2. Nine significant factors have been found in multivariate analysis, and they are all risk factors, the same to say, the bigger the value is, the more disadvantageous the prognosis is. The order of risk factors according to their RR values is: Postoperative AFP> TNM stage (IV stage)> Portal tumor thrombus > Distance from verge of resection to tumor (<1cm)> Growing pattern of tumor> Postoperative immune herbal therapy> Postoperative TAE

therapy>TNM stage (III stage)> Preoperative TAE therapy> Other apparatus metastasis.

Tab 1 The results of univariate analysis with BP neural network

Affecting factor indexes	Measuring of the variable	Coefficient	Relative Risk	Wald	P value
Preoperative TAE therapy	Yes: 1, no: 2	0.5368	1.71	4.06	0.0000
Clinical stage	Sub clinical: 0, clinical: 1	0.3485	1.42	4.56	0.0000
Portal tumor thrombus	No: 0 , Yes: 1,	1.5345	4.64	8.73	0.0000
AFP	AFP<=20：0				
AFP(1)	20<AFP<=100：1	0.0029	1.00	0.03	0.9792
AFP(2)	100<AFP<=1000：2	0.0154	0.98	0.17	0.8657
AFP(3)	1000<AFP<=5000：3	0.1949	1.21	1.85	0.0641
AFP(4)	5000<AFP<=10000：4	0.3607	1.43	2.50	0.0124
AFP(5)	10000<AFP<=20000：5	0.4573	1.58	3.25	0.0012
AFP(6)	AFP>20000：6	0.5711	1.77	4.54	0.0000
Extent of resection	Curative : 0, Relatively curative: 1	0.9949	2.70	11.72	0.0000
Distance from verge of resection to tumor	s>=2cm: 0				
Distance (1)	1<s<2cm: 1	0.4741	1.61	3.65	0.0003
Distance (2)	S<=1cm: 2	1.2762	3.58	9.50	0.0000
Tumor number	Single: 0, multiple: 1	0.6674	1.95	8.81	0.0000
Tumor size	d<3cm：0				
Tumor size (1)	3<d<5cm：1	0.3206	1.38	2.81	0.0050
Tumor size (2)	5<d<8cm：2	0.3577	1.43	3.12	0.0018
Tumor size (3)	8<d<10cm：3	0.7560	2.13	5.81	0.0000
Tumor size (4)	d>10cm：4	1.0452	2.84	8.48	0.0000
Growing pattern of tumor	Expand: 0, invade: 1	0.8641	2.37	11.63	0.0000
Daughter nodule around tumor	no: 0, Yes: 1	0.7787	2.18	11.93	0.0000
Vascular tumor thrombus in tumor nest	no: 0, Yes: 1	0.5612	1.75	7.40	0.0000
TNM stage	I stage:0				
TNM stage (1)	II stage：1	0.1961	0.82	0.97	0.3303
TNM stage (2)	III stage：2	0.6606	1.94	3.32	0.0009
TNM stage (3)	IV stage：3	1.3113	3.71	6.44	0.0000
Jaundice after operation	no: 0, Yes: 1	0.4359	1.55	3.32	0.0009
Ascites after operation	no: 0, Yes: 1	0.2708	1.31	2.10	0.0354
AFP after operation	Decrease to normal: 0, not to normal: 1	1.7661	5.85	16.42	0.0000
Other apparatus metastasis	no: 0, Yes: 1	0.4490	1.57	4.00	0.0001
Postoperative TAE therapy	Yes: 0, no: 1	0.2372	1.27	2.45	0.0142
Postoperative Immune herbal therapy	Yes: 0, no: 1	0.6370	1.89	8.80	0.0000

Tab 2 Results of multivariate analysis with BP neural network

Affecting factor indexes	Coefficient	Relative Risk	Wald	P value
Preoperative TAE therapy	0.3581	1.43	2.26	0.0238
Portal tumor thrombus	0.7536	2.12	3.60	0.0003
Distance from verge of resection to tumor				
Distance (1)	0.1340	1.14	0.91	0.3617
Distance (2)	0.6431	1.90	4.05	0.0001
Growing pattern of tumor	0.5482	1.73	6.31	0.0000
TNM stage				
TNM stage (1)	0.2207	0.80	0.94	0.3481
TNM stage (2)	0.3860	1.47	1.64	0.0100
TNM stage (3)	0.7836	2.19	3.20	0.0014
Postoperative AFP	1.1108	3.04	8.14	0.0000
Other apparatus metastasis	0.2852	1.33	2.15	0.0313
Postoperative TAE therapy	0.3989	1.49	3.45	0.0006
Postoperative immune herbal therapy	0.4241	1.53	5.27	0.0000

4.Discussions

In medical field, the clinical data is complex: they are non-linear, and has inevitably multiple collinearity and outlier. The data are incomplete and always have undefined information. All make the analysis results contradictory or irrational. Results of analyzing clinical data with traditional statistical methods are not ideal. In the past, general regression model was always used to study such problems, but it is just the approximate description for the system, and the data should satisfy some conditions, so its applicability is confined. When the conditions can not be satisfied, the results of the model are always unexplainable.

The method of ANN can be used to study the causal relationship between the input and output system. Neural network model is a net structure composed of some basic units according to some rules. It bases on the simple computing, deals with information through the competing and cooperating among the basic units, and returns the macroscopical results, so it is a latent-computing method. The output values of the basic units and the biases themselves are not macroscopical, they are just the carrier to store and transfer information on the sense of microcosmic. Though the neural network model also describes the relationship between input and output, it has difference from conventional parameter models: 1. It does not depend on the form of the model, so has extensive applicability. 2. The error of some unit is restrained by the microcosmic structure of the network, so the model is robust. For the goodness on theory, this method may have satisfied results on analyzing complex problems in clinical study.

In this paper, BP neural network is used to analyze the prognostic factors of PHC. Eighteen influencing factors were picked with univariate analysis, and nine with multivariate analysis. The 9 factors can be considered as important indexes influencing the recurrence of PHC after hepatectomy. The results of this study accord with the clinical practice well, and provide quantified data for clinical study. This study indicates: the results are satisfied when ANN is used to study the prognosis of PHC.

Acknowledgements

THIS RESEARCH WAS SUPPORTED BY NSFC.

REFERENCES

1. ZHANG, ZJ. & WU, MC. (1999). Significance of TNM Classification in Prognostic Evaluation of Hepatocelluar Carcinoma Following Surgical Resection (in Chinese). Chinese J. Oncology, Vol. 21, No. 4, 293-295.
2. ZHANG, ZJ. & WU, MC. (1999). The Effect of Preoperative Transcatheter Hepatic Arterial Chemoembolization on Disease-free Survival after hepatectomy for Hepatocellular Carcinoma (in Chinese). Chinese J. Oncology, Vol. 21, No. 3, 214-216.
3. Sun, G. & Zhou, BS. (1999). Research of Lung Cancer Diagnosis Based on Artificial Neural Network (in Chinese). Chinese J. Health Statistics, Vol. 16, No. 3,142-144.
4. Zhou, LF. & Gao, ES. (1998). A Preliminary Inquiry of the Comparison of the Logistic Regression and Back Propagation Neural Network (in Chinese). Chinese J. Health Statistics, Vol. 15, No. 1,1-4.