J Med Syst (2011) 35:121–126
DOI 10.1007/s10916-009-9348-8

ORIGINAL PAPER

A Novel Method for Diagnosing Cirrhosis in Patients with Chronic Hepatitis B: Artificial Neural Network Approach

Mohammad Reza Raoufy · Parviz Vahdani · Seyed Moayed Alavian · Sahba Fekri · Parivash Eftekhari · Shahriar Gharibzadeh

Received: 19 May 2009 / Accepted: 9 July 2009 / Published online: 21 July 2009

Abstract We designed an artificial neural network (ANN) to diagnose cirrhosis in patients with chronic HBV infection. Routine laboratory data (PT, INR, platelet count, direct bilirubin, AST/ALT, AST/PLT) and age were collected from 144 patients. Cirrhosis in these patients was diagnosed by liver biopsy. The ANN's ability was assessed using receiver-operating characteristic (ROC) analysis and the results were compared with a logistic regression model. Our results indicate that the neural network analysis is likely to provide a non-invasive, accurate test for diagnosing cirrhosis in chronic HBV-infected patients, only based on routine laboratory data.

M. R. Raoufy · P. Eftekhari
Department of Physiology, School of Medical Sciences,
Tarbiat Modares University,
Tehran, Iran

P. Vahdani
Department of Infectious Diseases, Logman Hospital,
Tehran, Iran

S. M. Alavian
Baqiyatallah Research Center for Gastroenterology and Liver Disease, Baqiyatallah University of Medical Sciences,
Tehran, Iran

S. Fekri
Faculty of Medicine,
Shahid Beheshti University of Medical Sciences,
Tehran, Iran

S. Gharibzadeh (✉)
Neuromuscular Systems Laboratory, Faculty of Biomedical Engineering, Amirkabir University of Technology,
Tehran, Iran
e-mail: gharibzadeh@aut.ac.ir

Keywords Cirrhosis · Chronic hepatitis B · Artificial neural network · Logistic regression · Laboratory data

Introduction

Cirrhosis is one of the most important complications of chronic hepatitis B, which occurs in its final stage. The characteristic feature of cirrhosis is the destruction of liver parenchymal cells and their substitution by regenerative nodules which are surrounded by fibrotic tissue [1, 2]. Meantime, the risks of cirrhosis complications as hepatocellular carcinoma are increased [3].

About one fourth of mild chronic hepatitis B patients may be affected by cirrhosis in 1 to 13 years. Five year and 15 year survival of chronic hepatitis B with cirrhosis is 55% and 40%, respectively [1]. In chronic hepatitis B patients, rapid diagnosis of cirrhosis is really important and can improve the prognosis and reduce the complications [4].

Liver biopsy is the gold standard for evaluating different liver diseases as cirrhosis [1, 5]. However, it is invasive, difficult, expensive, and accompanied by significant morbidity and low mortality and is not permitted in all patients [1, 6, 7]. Therefore, presenting a non-invasive, accurate, simple, and inexpensive method for diagnosing the cirrhosis can be truly helpful.

An ANN is a structure of simulated neurons that are connected together somewhat in the same way as natural neurons are thought to be connected in the brain. The capability and advantages of ANNs are due to their special features including nonlinear, adaptive, and parallel processing. Each neuron of ANNs receives inputs either from a number of other neurons or from an external stimulus. The weighted sum of these inputs passes through a basis

function and the resulted argument is applied to an activation function that finally yields the output of the neurons. The manner in which connections are made between these neurons (viz. the topology) defines the flow of information in the network and is called the architecture of the network. Useful architectural configurations include single layer, multilayer, feed-forward, feedback and lateral connectivity. Indeed, it is relatively straightforward to build an ANN model from input–output data. Behavior of a network depends greatly on the interactions between these building blocks. There are three types of neuron layers: input, hidden, and output layers. The more the layers used, the greater the power the network possesses. On the other hand, an excessive number of layers often appear to be counter-productive. It may cause slower convergence in the back-propagation learning. Generally speaking, three-layer network could be adequate as a universal approximator of any nonlinear function [8].

In this research, we present a cross-sectional study to develop an ANN model, based entirely on routine laboratory data, in order to diagnose cirrhosis in patients with chronic hepatitis B.

Materials and methods

Patient selection

The study was conducted in Tehran Hepatitis Center, between 1997 and 2007, on patients with hepatitis B. Patients enrolled in this study included all of the chronic hepatitis B patients who have not cirrhotic manifestations. Total number of these patients was 144. The affection of these patients to cirrhosis was diagnosed later by liver biopsy. The patients were randomly divided into two groups: The first group included the data of 11 cirrhotic patients and 75 non-cirrhotics. This group was used to train the models (ANN and logistic regression). The second group, the remainder of cirrhotics (eight persons) and non-cirrhotics (50 persons) were used to test the models.

We used independent t test to compare the characteristics of the patients different across training and testing sets.

Diagnostic variables

In a preliminary statistical study, the different simple and routine laboratory tests of chronic hepatitis B patients (including liver function tests, cell blood count and coagulation tests) were evaluated in order to select those which are significantly different between cirrhotics and non-cirrhotics (Table 1). In these patients who had not clinical manifestations, these laboratory data were used as inputs of the models. These inputs were age (I_1), PT (I_2), INR (I_3), platelet count (I_4), direct bilirubin (I_5), AST/ALT (I_6), and AST/platelet count (I_7). The outputs of the models were supposed to be the presence ($U=1$) or absence ($U=0$) of the cirrhosis (Fig. 1).

Development of ANN model

In order to design the ANN, we used the Neural Network Toolbox of the MATLAB 7 software. The ANN used in this study was a standard backpropagation neural network with three layers: an input layer, a hidden layer and an output layer. The input, hidden, and output layers contained seven, 15 and one neurons, respectively (Fig. 1). The number of the network layers, hidden neurons and the stopping criteria were determined through trial-and-errors process because no commonly accepted theory exists for predetermining the optimal number of neurons in the hidden layer [9]. The default transfer function for hidden layers and output layer was tansig and purelin, respectively. The function newff was used to create the network object in training feed-forward network. The default back-propagation training algorithm is Levenberg–Marquardt (trainlm). We preprocessed the input and target values: map them into the interval [−1, 1]. This simplifies the problem for the network. It also ensures that targets fall into the range that your new feed-forward network can reproduce.

A major problem of neural networks is their tendency to over-fit the data, i.e. to generate networks that are too closely adapted to the training data. One way to deal with over-fitting is to divide randomly the data into a training set and a validation set as described above and to test the network response to validation set. In our study, the network performed proper response to validation set (second group of data). So it can be claimed that over-fitting has not occurred in the training phase.

The ANN was trained 5,000 times (epochs). The mean standard error was 1.27907e−011. The performance of the ANN was evaluated by calculating the area under the ROC curve (AUC).

Development of logistic regression model

A logistic regression model was developed using SPSS for Windows 15.0 (SPSS Inc). The training and testing datasets were the same as those used in ANN; thus there were an ANN and a logistic model for training and testing datasets. The logistic regression performance was assessed with ROC analysis using the same testing set.

Comparison of the models

While the testing set was entered to ANN and logistic regression model as inputs, the diagnosis was determined as

Table 1 The distribution of the dataset between cirrhotic and non-cirrhotic chronic HBV-infected patients

	Cirrhotic	Non-cirrhotic	*p* value
Age (year)	48.05 (9.80)	33.04 (12.86)	<0.0001
PT (s)	14.89 (2.57)	13.31 (0.78)	<0.0001
INR	1.45 (0.46)	1.07 (0.12)	<0.0001
Platelet count (1/μl)	117,947.37 (52,830.41)	206,147.83 (78,341.72)	<0.0001
WBC (cells/μl)	7,676.84 (8,190.18)	6,288.83 (1,633.05)	0.094
ALT (IU/Lit)	75.58 (101.71)	108.61 (127.63)	0.284
AST (IU/Lit)	76.89 (108.28)	63.48 (52.09)	0.382
ALP (IU/Lit)	218.13 (96.35)	243.20 (181.16)	0.600
Total bilirubin (mg/dl)	1.35 (0.74)	1.15 (1.16)	0.503
Direct bilirubin (mg/dl)	0.45 (0.22)	0.27 (0.14)	<0.0001
AST/ALT	1.41 (1.14)	0.73 (0.40)	<0.0001
AST/ platelet count	0.00068 (0.00077)	0.00033 (0.00030)	0.001

Data are mean (SD)

PT prothrombin time, *INR* international normalized ratio, *WBC* white blood cell, *ALT* alanine transaminase, *AST* aspartate transaminase, *ALP* alkaline phosphatase

the outputs of these models. Accuracy (the number of correct predictions divided by total predictions), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio positive (LR^+) and likelihood ratio negative (LR^-) of these models were calculated. The discriminating power of these diagnosis models can be measured by ROC curves. ROC analysis estimates a curve that describes the inherent trade-off between sensitivity and specificity of a prediction tool. Discriminatory power is measured by AUC, which is a particularly important metric for evaluating prediction tools because it is the average sensitivity over all possible specificities. AUC may range from 0 to 1, with area of 1.0 representing perfect discrimination and an area of 0.5 representing what is expected by chance alone [10].

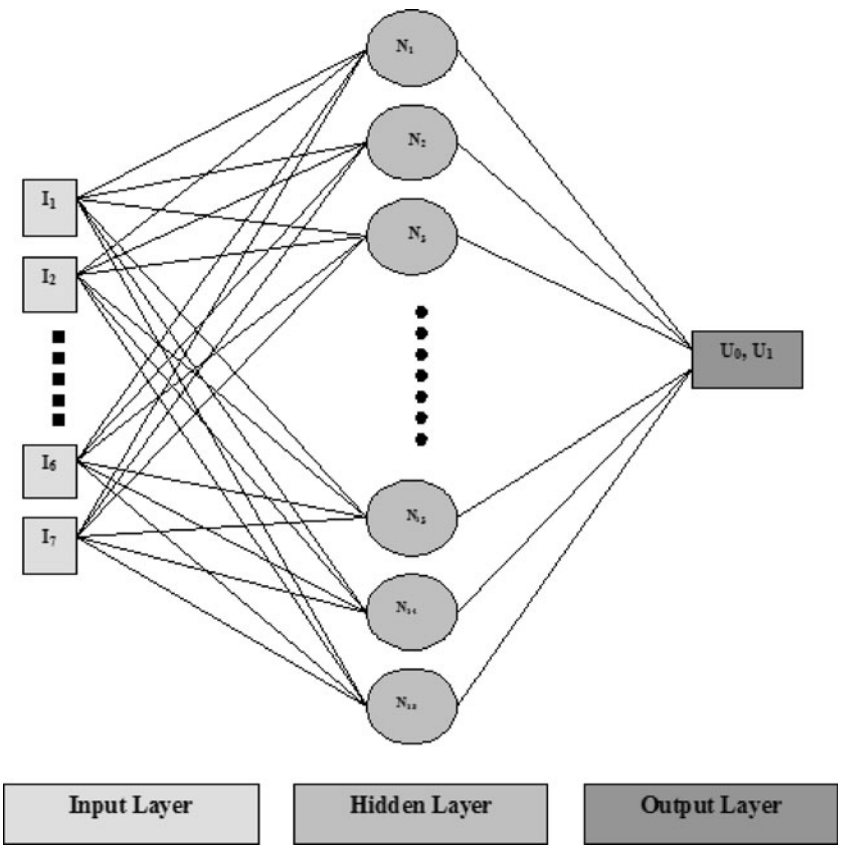

Fig. 1 The ANN (artificial neural network) designed with input, hidden, and output layers

Table 2 The characteristics of patients: comparison of training and testing sets

	Training set (n=86)	Testing set (n=58)
Cirrhosis	11 (13%)	8 (14%)
Age (year)	36.52 (14.83)	32.79 (10.90)
PT (s)	13.49 (1.32)	13.64 (1.29)
INR	1.11 (0.25)	1.14 (0.19)
Platelet count (1/μl)	195,081.40 (87,200.00)	188,293.10 (65,523.00)
Direct bilirubin (mg/dl)	0.29 (0.16)	0.30 (0.16)
AST/ALT	0.80 (0.53)	0.84 (0.68)
AST/PLT	0.00037 (0.00041)	0.00040 (0.00039)

Data are mean (SD) or numbers (percentage)

PT prothrombin time, *INR* international normalized ratio, *ALT* alanine transaminase, *AST* aspartate transaminase

Results

In this research, the studied population included 144 chronic HBV-infected patients who underwent liver biopsy, with 86 patients allocated at random to the training set and the remaining 58 to the testing set (Table 2). The characteristics of the patients were not significantly different across groups ($p>0.05$). Cirrhosis affected 13.2% of patients (19 persons), of which 11 cases were used in the training set and eight in the testing set.

As it is shown in Fig. 2, the ANN was able to diagnose seven out of eight cirrhotic patients and 46 out of 50 non-cirrhotic ones in testing set. These amounts were six out of eight and 44 out of 50 in logistic regression test, respectively. Table 3 shows the ANN and logistic regression performance in diagnosing the cirrhosis in chronic hepatitis B patients, compared to the gold standard liver biopsy.

Some predictive performance indices as sensitivity, specificity, PPV, NPV, LR^+, LR^-, and AUC are compared between ANN and logistic regression models in Table 4.

Figure 3 depicts the ROC curves for the ANN and logistic regression models. Results of the comparison of the ROC curves between two models are shown in Table 5. The ANN model significantly outperformed the logistic regression model ($p<0.0001$).

Discussion

Long-term follow-up of patients with chronic hepatitis B requires serial liver biopsies to reassess repeatedly the hepatic histology [11–13]. However, liver biopsy is associated with a finite risk of severe complications, patient discomfort, and expense [1, 6, 7]. Therefore, it is not suitable for regular monitoring of disease progression. For these reasons, chronic HBV-infected patients would greatly benefit from a reliable, non-invasive method of diagnosing cirrhosis. There are some previous studies which attempt to diagnose cirrhosis in chronic hepatitis infected patients by ANN method; however, their laboratory tests are expensive and less accessible [13, 14]. In some studies the research is done on hepatitis C [13, 15]. In this study, we developed an ANN model, based entirely on routine and inexpensive laboratory data, in order to diagnose cirrhosis in patients with chronic hepatitis B.

In designing the ANN, we used variables that although showed meaningful difference between cirrhotic and non-

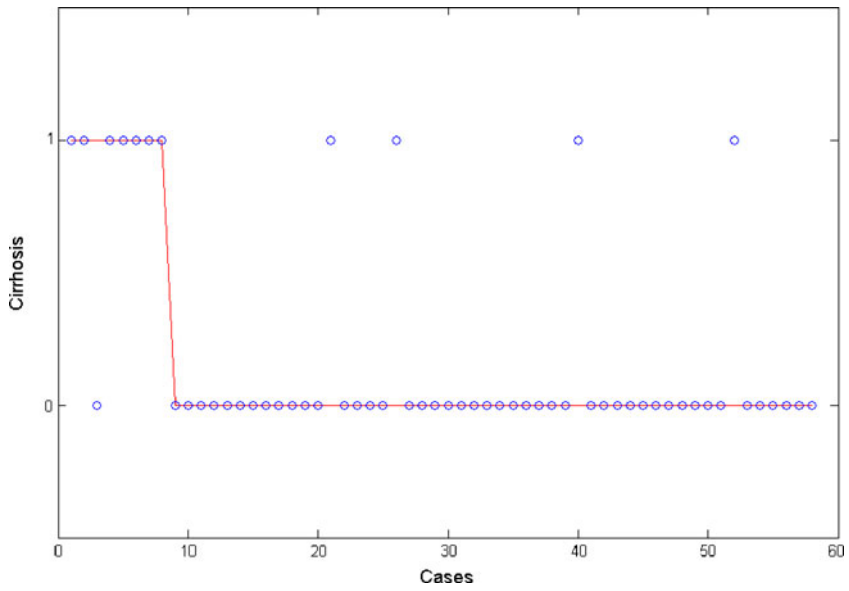

Fig. 2 ANN (artificial neural network) test results. The *continuous line* is the real output and the *circles* are the ANN output. "*1*" means presence and "*0*" means absence of the cirrhosis

Table 3 The ANN (artificial neural network) and logistic regression performance in diagnosing cirrhosis in chronic HBV-infected patients, compared to liver biopsy

		Liver biopsy report		Total
		Cirrhotic (N=8)	Non-cirrhotic (N=50)	
ANN	Positive	7	4	11
	Negative	1	46	47
Logistic regression	Positive	6	6	12
	Negative	2	44	46

cirrhotic patients in the preliminary statistical analysis, but were not directly diagnostic of cirrhosis. For validating the ANN, we used test data which were not utilized for training.

The low number of cirrhotic patients in this study may be one reason of relative decrease in sensitivity and positive predictive value. The ANN in this study had four false positive cases. This may be the consequence of false negative biopsy results, as it is claimed in some other studies [1, 6, 7]. Increasing the number of patients will certainly improve the ANN accuracy.

The AUC is a measure of a model's discriminatory power. According to the observation by Swets et al. [16], an AUC of ≥0.7 is diagnostically useful. In our study, both models discriminate well (AUC≥0.7). The ANN model significantly outperform the logistic regression model (AUC=0.898 vs. 0.815, p<0.0001) in the testing set. The ANN model also has better simultaneous sensitivity and specificity.

All prior studies using regression methodologies were based on the assumption that all required patient information would be available at the time of initial evaluation. Frequently, information is not available, either because patients can be poor historians or the practitioner forgets to ask pertinent questions, forgets to collect key observations, or is unable to retrieve all laboratory results in a timely fashion. Patients with missing data will be excluded in logistic regression, whereas the ANN will test all patients regardless of missing data and the ANN accuracy is not significantly compromised by this absence.

Table 4 Comparison of predictive performance of ANN (artificial neural network) and logistic regression

	ANN	Logistic regression
Accuracy (%)	91.38	86.21
Sensitivity (%)	87.5	75
Specificity (%)	92	88
PPV (%)	63.64	50
NPV (%)	97.87	95.65
LR^+	10.94	6.25
LR^-	0.14	0.28
AUC	0.898	0.815

PPV positive predictive value, *NPV* negative predictive value, LR^+ likelihood ratio positive, LR^- likelihood ratio negative

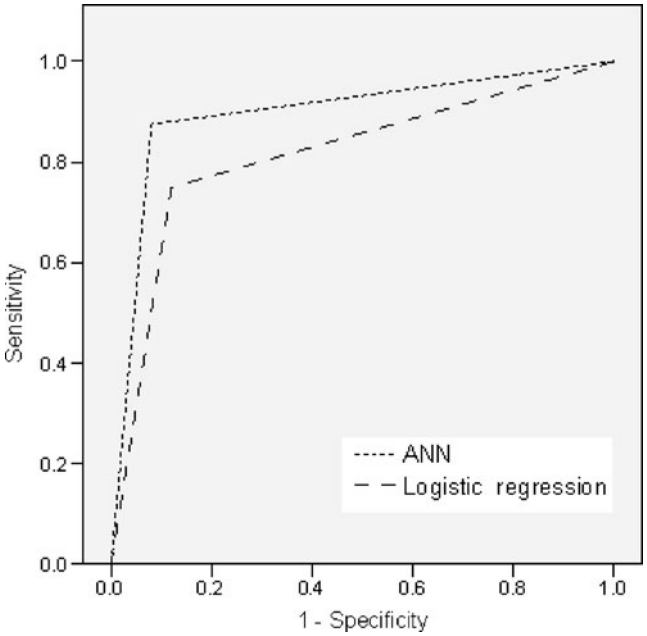

Fig. 3 Comparison of ROC (receiver operating characteristic) curves between ANN (artificial neural network) and logistic regression

A neural network approach is also preferable in that ANNs are model independent and flexible in being able to use mixes of categorical and continuous variables. ANNs also have the advantage that they can learn to predict arbitrarily complex nonlinear relationships between independent and dependent variables by including more processing elements in the hidden layer or more hidden layers in the ANN. These advantages make the ANN a more robust paradigm for application to a real-world setting [9, 17]. The real time use of the ANN is not difficult. The number of hospitals that have an electronic medical record is growing rapidly. Once trained, the ANN could reside in the background of the clinical information systems. The data used by our ANN is the standard information routinely collected in the patients with chronic hepatitis B. Once

Table 5 Comparison of ROC (receiver operating characteristic) curves

	Difference between AUC	Standard error	95% confidence interval	p value
ANN vs. logistic regression	0.083	0.098	0.126–0.039	<0.0001

ANN artificial neural network, *AUC* area under the ROC curve

entered into the electronic record, these data could then be used by the ANN to generate the probability of the predicted outcome. ANN accuracy could also be continuously improved over time because it can constantly be retrained as more patients are accumulated.

On the contrary, the "black box" interpretation is a major obstacle to the acceptance of ANNs as a tool for the medical decision support systems. However, an accurate second opinion is often helpful in medical decision making with or without a detailed understanding of how it works.

There are a number of limitations to this study that need to be addressed. First, the ANN was not tested in real time. It is not clear how physicians will respond if given ANN predicted outcome of cirrhosis or non-cirrhosis. Second, this study was carried out at a single institution. These findings must be corroborated on patients from multiple locations. Portability will be a critical factor to the future use of the ANN in this setting.

Overall, using the routine laboratory data, the designed model was able to diagnose the cirrhosis in chronic hepatitis B patients. This model is a novel high accuracy, non-invasive, inexpensive, and rapid method, which can be used clinically.

In future researches, Cohort studies using more samples could improve such new models and make them capable of predicting the stage of the disease.

References

1. Dennis, L. K., Eugene, B., Anthony, F., Stephen, H., Dan, L., and Larry, J., *Harrison'*, *s principles of internal medicine*, 16th edition. Mc Graw Hill, New York, 2005.
2. Heidelbaugh, J. J., and Bruderly, M., Cirrhosis and chronic liver failure: part I. Diagnosis and evaluation. *Am. Fam. Phys.* 74 (5) 756–762, 2006.
3. Leemans, W. F., Janssen, H. L., and de Man, R. A., Future prospectives for the management of chronic hepatitis B. *World J. Gastroenterol.* 13 (18)2554–2567, 2007.
4. Weickert, U., Buttmann, A., Jakobs, R., Schilling, D., Eickhoff, A., and Riemann, J. F., Diagnosis of liver cirrhosis: a comparison of modified ultrasound and laparoscopy in 100 consecutive patients. *J. Clin. Gastroenterol.* 39 (6)529–532, 2005.
5. Lefkowitch, J. H., Liver biopsy assessment in chronic hepatitis. *Arch. Med. Res.* 38 (6)634–643, 2007.
6. Sebastiani, G., and Alberti, A., Non invasive fibrosis biomarkers reduce but not substitute the need for liver biopsy. *World J. Gastroenterol.* 12 (23)3682–3694, 2006.
7. Munteanu, M., Non-invasive biomarkers FibroTest-ActiTest for replacing invasive liver biopsy: the need for change and action. *J. Gastrointestin. Liver Dis.* 16 (2)173–174, 2007.
8. Dariani, S., Keshavarz, M., Parviz, M., Raoufy, M. R., and Gharibzadeh, S., Modeling force-velocity relation in skeletal muscle isotonic contraction using an artificial neural network. *Biosystems*. 90 (2)529–534, 2007.
9. Tu, J. V., Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J. Clin. Epidemiol.* 49:1225–1231, 1996.
10. Hanley, J. A., and McNeil, B. J., The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 143:29–36, 1982.
11. Lok, A. S., and McMahon, B. J., Chronic hepatitis B: update of recommendations. *Hepatology.* 39:857–861, 2004.
12. Chinese Society of Hepatology, Chinese Medical AssociationChinese Society of Infectious Diseases, Chinese Medical Association, Guideline on prevention and treatment of chronic hepatitis B in China (2005). *Chin. Med. J. (Engl)*. 120 (24)2159–2173, 2007.
13. Haydon, G. H., Jalan, R., Ala-Korpela, M., Hiltunen, Y., Hanley, J., Jarvis, L. M., Ludlum, C. A., and Hayes, P. C., Prediction of cirrhosis in patients with chronic hepatitis C infection by artificial neural network analysis of virus and clinical factors. *J. Viral. Hepat.* 5 (4)255–264, 1998.
14. Poon, T. C., Hui, A. Y., Chan, H. L., Ang, I. L., Chow, S. M., Wong, N., and Sung, J. J., Prediction of liver fibrosis and cirrhosis in chronic hepatitis B infection by serum proteomic fingerprinting: a pilot study. *Clin. Chem.* 51 (2)328–335, 2005.
15. Piscaglia, F., Cucchetti, A., Benlloch, S., Vivarelli, M., Berenguer, J., Bolondi, L., et al., Prediction of significant fibrosis in hepatitis C virus infected liver transplant recipients by artificial neural network analysis of clinical factors. *Eur. J. Gastroenterol. Hepatol.* 18 (12)1255–1261, 2006.
16. Swets, J. A., Measuring the accuracy of diagnostic systems. *Science*. 240:1285–1293, 1988.
17. Lin, C. S., Li, Y. C., Mok, M. S., Wu, C. C., Chiu, H. W., and Lin, Y. H., Neural network modeling to predict the hypnotic effect of propofol bolus induction. *Proc. AMIA Symp.* 450–454, 2002.