BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 6(6)

An ANN model for treatment prediction in HBV patients

Sajid Iqbal[1, 2]*, Khalid Masood[1], Osman Jafer[2]

[1]Bioinformatics lab, National Center of Excellence in Molecular Biology, Lahore, Pakistan; [2]Central veterinary research laboratory, Department of molecular biology and genetics, Dubai, United Arab Emirates; Sajid Iqbal - Email: sibhatti@mbg.ae; *Corresponding author

Received May 16, 2011; Accepted June 01, 2011; Published June 06, 2011

Abstract:
Two types of antiviral treatments, namely, interferon and nucleoside/nucleotide analogues are available for hepatitis infections. The selection of drug and dose determined using known pharmacokinetics and pharmacodynamics data is important. The lack of sufficient information for pharmacokinetics of a drug may not produce the desired results. Artificial neural network (ANN) provides a novel model-independent approach to pharmacokinetics and pharmacodynamics data. ANN model is created by supervised learning of 90 patients sample to predict the treatment strategy (lamivudine only and Lamivudine + Interferon) on the basis of viral load, liver function test, visit number, treatment duration, ethnic area, sex, and age. The model was trained with 68 (77.3%) samples and tested with 20 (22.7%) samples. The model produced 92% accuracy with 92.8% sensitivity and 83.3% specificity.

Keywords: ANN (Artificial neural networks), Hepatitis, Prediction, Treatment

Background:
Viral hepatitis is one of the important public health problems. There are 350 million patients, who are carrier of hepatitis B at present, worldwide **[1]**. The Majority of the Asian population is affected. Particularly the situation in Pakistan is alarming **[2]**. Pakistani population has a carrier rate of 3-4%. Some of the community based studies showed the presence of 31% hepatitis B core antibodies and 4.3% of hepatitis B surface antigen on tested samples **[3]**. There are two types of antiviral treatments, Interferon and nucleoside/nucleotide analogues. The expenses for antiviral treatment are considerable in developing and developed countries **[1]**. Hence the appropriate treatment selection has a major role in the cost and cure of the disease. The dose of the drug and selection of the drug can be determined by the knowledge of drug's pharmacokinetics and pharmacodynamics. The lack of sufficient information for pharmacokinetics of a drug may not produce the desired results. Artificial Neural Network (ANN) provides a novel model-independent approach to pharmacokinetics and pharmacodynamics data **[4]**. This is why ANN can predict the drug treatment accurately based on the factors or symptoms of the patients. ANN has been used as a prediction tool for classification and decision support in many fields like medicine **[5]**. ANN is based on artificial intelligence (AI) established in the area of computer sciences, which is capable of producing software which can do sophisticated, intelligent, and can perform computations similar to those of human brain **[4]**. Neural Networks collect the information by detecting relationship patterns in data as the brain does **[4]**.

Computational artificial intelligence is mainly concerned with the construction of AI programs that perform diagnosis and make therapy recommendations based on purely statistical and probabilistic methods. **[6]** ANN provides a more advanced model of univariate and multivariate analysis, which imitates the nervous system in its function. It consists of a group of input and output neurons, simulating receptors and effectors, and a group of hidden neurons layer which simulates inter-neurons. **[5]** Neural Network is basically an interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal neuron. The processing ability of network is stored in the inter-unit connection strengths, or weights, obtained by process of adaptation to, or learning from, a set of training patterns **[6]**. The therapy recommendation is too important for the curing of hepatitis B especially with time duration. Hence the concept was developed to create an ANN system through supervised learning of Hepatitis B patients samples to predict the treatment strategy i.e. lamivudine only or lamivudine + Interferon on the basis of viral load, visit number, treatment duration, ethnic area, sex, and age. Following were the aims of this study. (a) To set up an artificial neural network system. (b) To develop a computational model to recommend the treatment plan of the newly registering patients based on ANN prediction. (c) To develop a database of hepatitis B patients. (d) To develop a Management Information System for handling present and upcoming HBV patients.

Methodology:
Sera samples were collected from HBV infected patients, DNA was extracted and then PCR was performed before the start of the treatment, at 3 months, at 6 months and in follow up periods. Two types of treatment were given to the patients i.e. Lamivodine only and Lamivodine + Interferon. The complete demographic, diagnostic and treatment history of patients were being recorded. The front end of system was developed in Dot net technologies i.e. Visual Studio 2005, at backend a normalized database has been developed in MYSQL.

Statistical Analysis and testing:
Neural Network modeling was applied to get the required results by sigmoid activation function. ANN has been developed by creating single-layered Perceptron by Feed-forward Networks and back propagation methods through supervised learning of a sample of 90 patients to predict the treatment strategy. The sigmoid function was calculated through logistic regression. The system was built in two phases, in phase one the system was given training to learn and in second phase it was test against training.

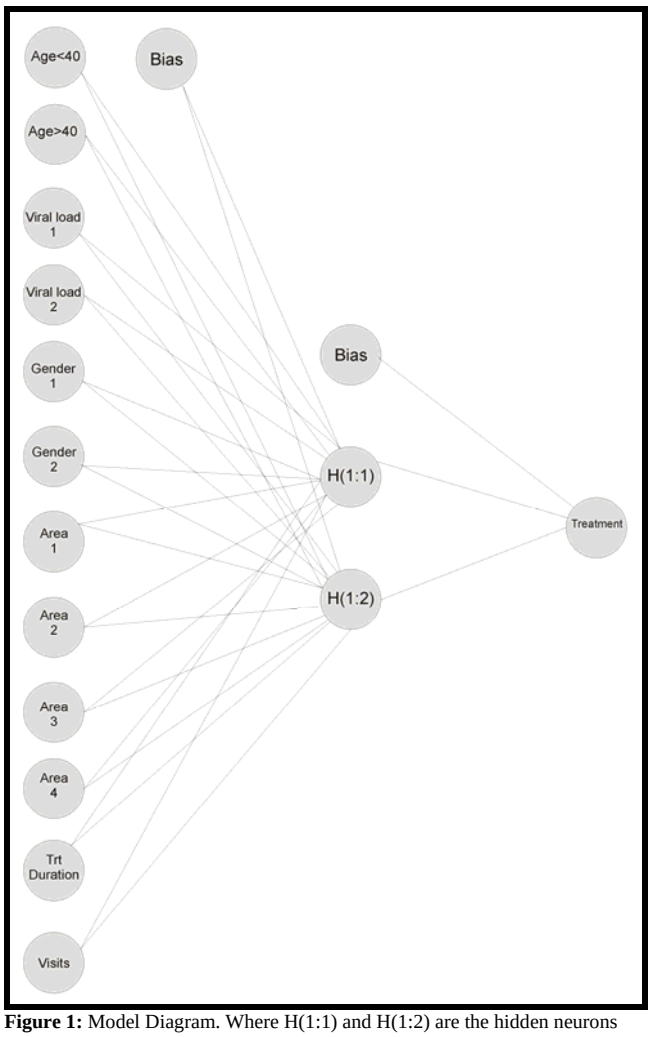

Figure 1: Model Diagram. Where H(1:1) and H(1:2) are the hidden neurons

Results and Discussion:

Hepatitis B virus infections are well known to be the major cause of mortality and morbidity worldwide, especially among developing countries **[7]**. Hepatitis B is endemic in Africa, Eastern Europe, Asia and America **[2]**. Our study is primarily focused on Pakistani population. The Pakistan HBV infection rate is known to be increasing day by day owing to the lack of proper health facilities, poor economic status, etc. **[7]**. The drug selection for HBV infection is an important factor not only for the treatment of the disease, but also has an impact on the economic status of the country. The cost of anti-viral therapy is always considerably high, globally. In addition, the drug may have many side effects like fatigue, flu like syndrome etc. Most importantly, only a few patients may achieve a response after therapy **[1]**. Hence, prediction or suggestion of a treatment is crucial. Therefore, a system was developed which predicts accurate treatment and treatment procedure. In addition, a database information system was developed, to save the patient data. The latest technique used for accurate prediction is artificial neural networks, which is actually a novel independent approach to analyze the data for prediction **[4]**. For a linear relationship, certain assumptions must be met within the data. In case of data deviation from assumptions is met, assumptions-free modeling approach such as Artificial Neural Network is utilized.

An ANN was developed by creating single layer Perceptron, Feed-forward Networks and back propagation method. The system was created through supervised learning of a sample of 90 patients, to predict the antiviral treatment strategy, i.e. Lamivodine only and Lamivodine+Interferon on the basis of viral load, visit number, treatment duration, ethnic area, sex, and age. In previous study it has been proved that the factors like age, genotype, viral load, Alt levels, and liver function tests are the strong predictors for antiviral therapy **[1]**. The same concept was carried out in another study where ANN was developed using SNP's **[9]**. The age was categorized into two groups above 40 and below 40 yrs. The viral load has two categories, below 10,000 and above 10,000. Ethnic area was categorized into NWFP, Punjab, Sindh and Baluchistan. There were 73 (81.1%) males and 17(18.9 %) females. Hence, the prevalence is high among males. 72 (80%) patients were below age 40 and 18 (20%) were above 40, hence it is more common among young men. 38 (42.2%) patients belong to NWFP, 25 (27.8%) belong to Punjab, 16 (17.8%) from Sindh and 11(12.2%) from Baluchistan. The set of all known samples were broken into two orthogonal (independent) sets, training set and testing set to test the performance of the neural network. The system was provided with 68 (77.3%) samples to train the network, and 20 (22.7%) were tested against training. A total of six variables were used as input against which a hidden layer of two neurons were developed. The weights were assigned between the input and the hidden neurons. The weights of hidden neuron and output were calculated by using the Sigmoid Activation function **(Table 1 see Supplementary material)**. In input layer, H (1:1) is the weight between the input and one hidden neuron in the hidden layer. H (1:2) is with neuron 2. While in hidden layer 1 the weights are H (1:1) and H (1:2) for the hidden neurons and the output. The diagram of the neural networks is shown in **Figure 1**. A significant reduction in the error was found with training and testing the system showing high precision level. The overall standard error (SE) of the system was 1.214. To validate the prediction model, we calculated sensitivity (SEN), specificity (SPE), positive prediction value (PPV) and negative prediction value (NPV). The overall accuracy by the ANN was 92% which is found to be almost same in another study of FFNN using SNP's **[9]**. The value of SEN, Sep, PPV, and NPV were 0.928, 0.8333, 0.93 and 0.83 respectively.

Conclusion:

We have developed a prediction model to predict the appropriate drug through an advanced methodology based on phenotypic characteristics. The results have shown that a trained ANN model is a promising method for providing the inference from measure of phenotypic traits. Our model achieves a higher successful rate of prediction and allows patients and doctors to make more informed decision. Based on the experience of appropriate therapy for HBV infection with this system we can predict 92% satisfactory therapy for the infection with low error rate. The error can be further reduced by increasing the sample size. The choice of the drug is not only helping to cure the disease but also it reduces the side effects of the drug and the cost of the therapy especially in the developing world.

References:

[1] Mao QG *et al*. *World J Gastroenterol*. 2010 **16**: 3465 [PMID: 20632453]
[2] http://eprints.hec.gov.pk/2104/1/2022.htm
[3] Jafri W *et al*. *BMC Infect Dis*. 2006 **6**: 101 [PMID: 16792819]
[4] Agatonovic-Kustrin S & Beresford R. *J Pharm Biomed Anal*. 2000 **22**: 717 [PMID: 10815714]
[5] Naumovic R *et al*. *Biomed Pharmacother*. 2010 **64**: 633 [PMID: 20888177]
[6] Jiang J *et al*. *Comput Med Imaging Graph*. 2010 **34**: 617 [PMID: 20713305]
[7] Alam MM *et al*. *BMC Infect Dis*. 2007 **7**: 115 [PMID: 17922910]
[8] Biglarian A *et al*. *Asian Pac J Cancer Prev*. 2010 **11**: 533 [PMID: 20843146]
[9] Lin E *et al*. *Pharmacogenomics* 2006 **7**: 1017 [PMID: 17054412]

Edited by P Kangueane
Citation: Iqbal *et al*. Bioinformation 6(6): 237-239 (2011)

BIOINFORMATION

open access

Supplementary material:

Table 1: Likelihood parameter estimates for synaptic weights

Parameter Estimates				
Predictor		Predicted		
		H(1:1)	H(1:2)	Output Layer treatment
Input Layer	(Bias)	-0.36	0.43	
	[age_cat=.00]	-0.26	-0.28	
	[age_cat=1.00]	0.28	0.57	
	[viral_load=.00]	-1.77	-0.51	
	[viral_load=1.00]	1.64	0.97	
	[gender=0]	0.38	0.04	
	[gender=1]	-0.58	-0.44	
	[Area=1]	0.10	0.65	
	[Area=2]	0.20	-0.76	
	[Area=3]	-0.88	-1.05	
	[Area=4]	1.60	1.18	
	Treatment duration	3.74	0.93	
	Visit	-1.65	-1.32	
Hidden Layer 1	(Bias)			0.884305
	H(1:1)			-4.37528
	H(1:2)			-2.01677