ESANN 2010 proceedings, European Symposium on Artificial Neural Networks - Computational Intelligence and Machine Learning. Bruges (Belgium), 28-30 April 2010, d-side publi., ISBN 2-930307-10-2.

Neural models for the analysis of kidney disease patients

Emilio Soria[1], José David Martín[1], Mónica Climente[2], Amparo Soldevila[3], Antonio Serrano[1] *

1- Intelligent Data Analysis Laboratory (IDAL) - Electronic Engineering Department
ETSE, University of Valencia, C/ Dr. Moliner 50, 46100 Burjassot, Valencia, Spain
2- Pharmacy Service. University Hospital, Dr. Peset, Valencia, Spain.
3- Nephrology Service, Hospital La Fe, Valencia, Spain.

Abstract. This work uses Machine Learning techniques and other classical approaches to analyze both physiological variables and treatment characteristics in patients undergoing chronic renal failure. Firstly, the use of Self-Organizing Maps is proposed in order to extract qualitative knowledge. Secondly, the Hemoglobin concentration is predicted one-month ahead by models based on the Multilayer Perceptron; the prediction uses information from two months (the current month and the previous one). Achieved results support the usefulness of these tools in daily clinical practice.

1 Introduction

Secondary anemia due to end-stage renal disease (ESRD) is a clinical situation that is present in more than 90% of patients undergoing periodic hemodialysis. In the framework of Chronic Kidney Disease (CKD), anemia is mostly due to the lack of Erythropoietin (EPO) production. As renal function performs worse in the ESRD, there is a decrease in the EPO production. It finally leads to the anemia. The symptoms due to anemia are quite variable, and they usually appear when Hb concentration is lower than 11 g/dl. The patients' quality-of-life (QoL) decreases, and changes in the cardiac function may also involve an increase in the ratio of morbidity and mortality.

In these circumstances the administration of Erythropoiesis Stimulating Agents (ESAs) like recombinant human erythropoietin or darbepoetin is the treatment of choice trying to maintain the Hb levels between 11 and 12 g/dl. The clinical investigation performed in recent years regarding the administration of different rHu-ESAs (recombinant Human-ESAs) has allowed the optimization of treatment and, consequently, the improvement in the QoL in patients undergoing chronic dialysis. However, ESAs are an expensive treatment that increases the high costs of the renal replacement therapy through dialysis and transplantation, and it is not without risks. The main secondary effects of ESAs are arterial hypertension, and thromboembolic complications including thrombosis of vascular access [1], [2]. Consequently, dosage optimization is critical to ensure adequate pharmacotherapy as well as patient's safety.

* This work has been partially supported by the Spanish Inter-Ministry Committee of Science and Technology under projects TIN2007-61006 and CSD2007-00018.

Although the guidelines for the administration of ESAs are followed in daily clinical routine, the response is not always the expected one since many complex and non-linear relationships tend to appear in the assimilation of the drug by the patient, and usual protocols take into account the average population response to the treatment [3]. Therefore, in spite of the general success of ESA therapy for correcting anemia and maintaining Hb levels in patients with CKD, the management of renal anemia with these agents often is labor-intensive and time-consuming. The response to the treatment with ESAs is highly dependent on the patient [4], [5]. The same dosages may have very different responses in different patients, most notably the so-called ESA-resistant patients, who do not respond to these agents, even after receiving high dosages.

This work proposes a neural network approach as a medical decision aid, being the decision in this case the administration of ESAs. This work is a continuation of previous experiences developed by the authors in this field [6]. The models proposed in this work pursuit two main goals.

First, a qualitative knowledge extraction by using Self-Organizing Maps (SOMs), which is the most extended visual data mining model in high dimensionality problems. This way, the profiles existing in the data are shown up, and therefore, it is possible to visualize the different physiological behaviors present in the database. This knowledge is crucial to know the relationships among variables, showing for example that the effect of EPO administrations is patient-dependent.

Second, a predictor model of the Hb level for next month using physiological variables and the ESAs administrations in the current month and also in the past. By using this model, it is possible to check whether or not the Hb evolution matches the desired one.

2 Neural models

2.1 Self-Organizing Maps

The Self-Organizing Map consists of a set of neurons which are usually arranged in a one- or two-dimensional grid: an input layer, formed by n neurons (one neuron for each input variable) and an output layer in which the information is processed (the two-dimensional map) [7]. Although higher dimensional grids are also possible, they are hardly ever used because of their problematic visualization. Every neuron has a fixed position in the grid and is represented by an n-dimensional weight vector where n is the dimensionality of the input space. A pattern x is randomly chosen from the data set in each training step. Then, the neuron whose weight vector is the most similar to the user pattern is found; this neuron is called Best Matching Unit (BMU). The weight vectors of the BMU and its neighborhood are updated in order to bring the n-dimensional vectors of the map closer to the pattern x. That update is carried out by using the difference between the input vector and the weight vector. The neighborhood kernel determines which neurons around the BMU are updated, and how this update affects each neuron [8]. Once the map training is finished, the visualization of the two-dimensional map provides qualitative information about how the input variables are related to each other.

2.2 Multilayer perceptron

This model consists in a layered arrangement of individual computation units known as *artificial neurons*. The neurons of a given layer feed with their outputs the neurons of the next layer. The inputs x_i to a neuron are multiplied by adaptive coefficients w_i, called synaptic weights, which represent the connectivity between neurons; then a non-linear function is applied to the result of the linear combination. Neurons from a specific network are grouped together in layers that form a fully connected network. The first layer contains the input nodes, which are usually fully connected to hidden neurons and these are, in turn, connected to the output layer. In our case, only one output neuron is necessary, since only one variable is predicted at each time. The objective is to determine the best network parameters to model the relationship between the input and output variables. Although there are many other possible choices, its well-known behaviour as universal function model justifies its use [7].

3 Results

Data were prospectively collected throughout 2008 for patients with secondary anemia due to CKD in periodic hemodialysis. All were chronic hemodialysis stable patients, treated in hemodialysis centers affiliated to the University Hospital Dr. Peset (Valencia, Spain). ESAs were administered to patients through intravenous route, with a frequency determined by Hb levels (current and previous) and dose and type of ESA (there are two different kinds: EPO beta and darbepoetin). Patients were monitored monthly, and the following variables were collected: plasmatic concentration of Hb (g/dl), hematocrit (Ht) concentration (%), ferritin (ng/ml), administration of intra-venous iron (IV Fe) (mg/month), number of administrations of IV Fe, weekly dosage of EPO-beta (international units, IU), weekly dosage of darbepoetin-alpha (μg), number of administrations of ESAs and patient's age and gender.

After data preprocessing, a qualitative analysis of data was carried out. We used SOM for this task. The goal is to find qualitative relationships between Hb and the rest of used features (weight, gender, age, previous Hb values and EPO and Fe doses in the two previous months). Since two different kinds of EPO were used, a different neural model was developed for each one of them.

The neural model was tuned in terms of: a) the two possible types of learning (on-line and batch); b) the most common neighborhood functions (Gaussian and Mexican hat) c) the learning constants; d) the neurons coefficients' initialization (50 different initializations for each possibility) and e) the possible architectures. The SOMs were analyzed in order to determine the quantization error (adjustment of the neurons' vectors to the input data) and the topographic error (a measure of how well the neighborhood relationship is maintained between the original and the final bidimensional space), finally choosing the map that presented a lower product of the quantization error times the topographic one. In the case of EPO beta, the final map had an architecture formed by 16x10 neurons, while in the case of darbepoietin the architecture was 13x9. Only results with EPO beta will be shown for the sake of the length of the paper. Figure 1 (right) shows the different groups provided by the

already trained SOM; those groups provide an intuitive representation of the different patients' behaviors present in the data set. Moreover, Figure 1 (left) shows the number of patterns assigned to each neuron (the more filled the neuron, the higher the number of patterns assigned to it).

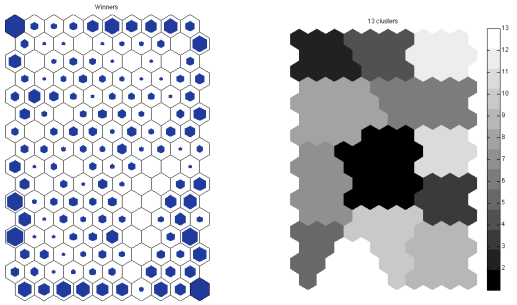

Fig. 1: EPO beta. Relative number of patterns assigned to each neuron (left) and the different groups present in the trained SOM (right).

Using the information about the different groups (it is obtained analyzing distances between prototypes using U-matrix [8]) and the areas of the map corresponding with the highest number of patterns, the next step is the analysis of the component maps, where the variables appear altogether (Figure 2).

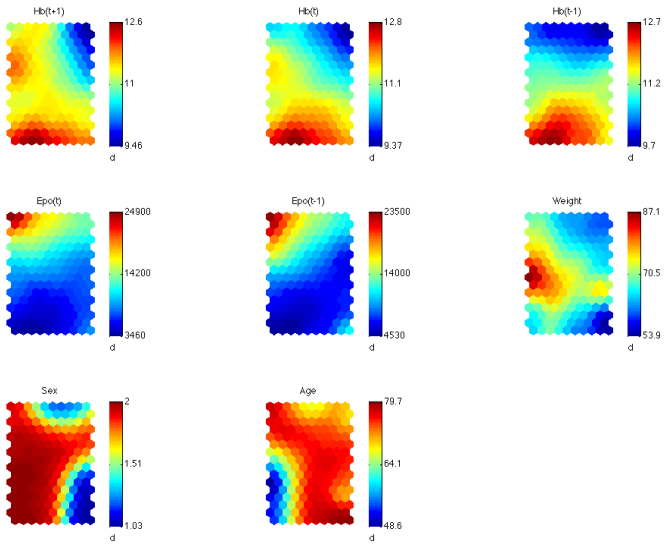

Fig. 2: EPO beta: Component maps showing the weights of the neurons.

Figure 2 shows the relationship between EPO doses and the increase of Hb levels, and an apparent two-month delay in the EPO effects. In fact, the bottom-left

corner of the map corresponds with low doses of EPO and good levels of Hb; the winner map (Figure 1, left) shows up that there are many patterns following this behavior; analyzing the remainder of variables, it can be seen that it corresponds with male (Gender map, value 2) and relatively young (Age map) patients. The lower-right corner of the map also represents many patterns (Figure 1, left); this corner corresponds with stable patients (Hb increment is quite low and and Hb levels within the target range); therefore, low values of EPO doses are administered; this corner represents basically female patients with low weight and relatively elderly. EPO-resistant patients are mapped onto the upper-left corner. These patients show low levels of Hb and no Hb increases after receiving EPO.

While the SOM approach provides a qualitative analysis of data, MLP is focused on a quantitative analysis instead. In particular, we focus on predicting next value of Hb (one-month ahead) using anthropometric features (weight, gender and age) as well as the current and previous-month values of Hb concentration and also the EPO doses administered in those months. Therefore, the tackled problem consists of seven input variables and only one outcome. The data set was split into two data sets: a training set (67% of the data) and a test set (remainder 33%). The splitting criterion was checked by testing that basic statistics of both data sets matched each other. The Levenberg-Marquardt learning algorithm was used due to its appropriate trade-off between accuracy and convergence speed [7]. The learning algorithm was run on-line since it is more appropriate for time series prediction than an off-line approach. It was early stopped by cross-validation. In order to avoid falling into local minima, each architecture (we used two layers and varied the number of neurons between 2 and 20) was initialized in 100 different ways.

Neural models were compared with linear models using the same input variables. Linear models were fitted using a robust least-squares method to avoid problems with outliers. Table 1 shows the achieved results for the two kinds of EPO using both linear and neural models. Instead of using global indices which do not give information about the model behavior at the patient level, Table 1 shows the quartiles (25, 50 and 75%) of the distribution of performance indices (RMSE: Root Mean Square Error; MAE: Mean Absolute Error; ME: Mean Error). Those indices are calculated for each individual patient.

	RMSE (g/dl)	MAE (g/dl)	ME (g/dl)
Linear	0.55/0.76/1.04	0.45/0.66/0.85	-0.20/-0.02/0.22
Neural	0.56/0.74/1.03	0.48/0.64/0.85	-0.18/0.00/0.22
Beta (1034 patterns)			
Linear	0.50/0.68/0.87	0.42/0.54/0.75	-0.19/0.07/0.19
Neural	0.48/0.61/0.83	0.38/0.49/0.69	-0.24/-0.05/0.06
Darbepoietin (632 patterns)			

Table 1**.** Achieved results for linear and neural models in the Hb prediction. The quartiles of the distribution of performance indices are shown.

Given the similarity in the prediction performance, linear and neural models were compared using a non-parametric statistical test (Man-Whitney, U test). These tests showed that there were no significant differences between the outputs of one model and the outputs of the other in the case of EPO beta. However, significant

differences did appear for darbepoietin ($p<0.05$). Before applying the tests, the non-normality of the models' outputs was also assessed by the Kolmogorov-Smirnov test ($p<0.05$).

4 Conclusions

This work has proposed the use of neural models as a clinical decision aid in EPO dose individualization for anemic patients with CKD undergoing hemodialysis. A qualitative approach where the goal has been to find different behaviors in the data set has been firstly proposed. It can be seen as the first step of a behavior-based model. This way, different therapies can be applied to different behaviors. It has been shown the ability of the model to explain different data profiles qualitatively.
Secondly, a quantitative approach has been proposed for the accurate prediction of Hb levels for next month using the variables measured currently and in the previous month. Linear models show a performance similar to neural models but they are statistically different each other in the case of darbepoietin ($p<0.05$) (not in the case of EPO beta). Prediction accuracy is within an acceptable range from a clinical point of view [9]
The achieved results may have been conditioned by the multi-center approach. The exclusion of outlier patients derived from comorbidities and special situations who need a more individualized treatment, would increase the predictive value of the model. Therefore, this is one of our next future tasks to do.

References

[1] L. Peterson. FDA Oncologic Drugs Advisory Committee (ODAC) Meeting on the Safety of Erythropoietin in Oncology. *Trends in medicine*, 1–4, 2004.

[2] A. Phrommintikul, S.J. Haas, M. Elsik, H. Krum. Mortality And Target Haemoglobin Concentration In Anaemic Patients With Chronic Disease Treated With Erythropoietin: A Meta-Analysis. *Lancet* 369 (9559), 381-8, 2007.

[3] A. Gaweda, M.K. Muezzinoglu, G.R. Aronoff, A.A. Jacobs, J.M. Zurada, M. Brier. Individualization of pharmacological anemia management using reinforcement learning, *Neural Networks* 18: 826–834, Elsevier, 2005.

[4] R.L. Pisoni, J.L. Braga-Gresham, E.W. Young, T. Akizawa, Y. Asano, F. Locatelli. Anemia management and outcomes from 12 countries in the Dialysis Outcomes and Practice Patterns Study (DOPPS). 1: 94-111, *Am J Kidney Dis*, 2004.

[5] E. Lacson E, N. Ofsthun, M. Lazarus. Effect of Variability in Anemia Management on Hemoglobin Outcomes in ESRD. *Am J Kidney Dis*, 41: 111-124, 2003.

[6] J.D. Martín, F. Gomez, E. Soria, J. Schmidhuber, M. Climente, N.V. Jiménez. A reinforcement learning approach for individualizing erythropoietin dosages in hemodialysis patients. *Expert Systems with Applications,* Elsevier, 36: 9737-9742, 2009.

[7] S. Haykin, *Neural Networks and Learning Machines*. Prentice Hall, New Jersey, 2009.

[8] Kohonen, T. *Self Organizing Maps*, Springer Verlag, 2000.

[9] National Kidney Foundation, K.D.O.Q.I.: *Guidelines for anemia of chronic kidney disease*. NKF K/DOQI Guidelines (2007), http://www.kidney.org.