30th Annual International IEEE EMBS Conference
Vancouver, British Columbia, Canada, August 20-24, 2008

Analysis of significant factors for dengue infection prognosis using the self organizing map

Tarig Faisal[a], Fatimah Ibrahim[a], *Member, IEEE*, Mohd Nasir Taib[b], *Senior Member, IEEE*

[a] Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia

[b] Faculty of Electrical Engineering, Universiti Teknologi Mara, Shah Alam, Selangor, 40450, Malaysia

Abstract—**This study presents a new approach to determine the significant prognosis factors in dengue patients utilizing the self-organizing map (SOM). SOM was used to visualize and determine the significant factors that can differentiate between the dengue patients and the healthy subjects. Bioimpedance analysis (BIA) parameters and symptoms/signs obtained from the 210 dengue patients during their hospitalization were used in this study. Database comprised of 329 sample (210 dengue patients and 119 healthy subjects) were used in the study. Accordingly, two maps were constructed. A total of 35 predictors (17 BIA parameters, 18 symptoms/signs) were investigated on the day of defervescence of fever. The first map was constructed based on BIA parameters while the second map utilized the symptoms and signs.**

The visualized results indicated that, the significant BIA prognosis factors for differentiating the dengue patients from the healthy subjects are reactance, intracellular water, ratio of the extracellular water and intracellular water, and ratio of the extracellular mass and body cell mass.

Keywords—**Self Organizing Map, Bioimpedance analysis, Symptoms and Signs, Dengue patients, Healthy subjects.**

I. Introduction

Dengue disease is considered as worldwide disease since it's widespread in many parts of tropical and subtropical world [1, 2]. According to world health organization (WHO), 2.5 billion people are facing the risk of dengue disease in these areas [3].

The WHO classified the dengue infection as dengue fever (DF) and dengue haemorrhagic fever (DHF)/dengue shock syndrome (DSS) [3]. DF begins with a sudden temperature increase accompanied by headache, myalgia, macular rash, loss of appetite nausea, vomiting, abdominal pain, metallic taste of food, change in psychological state and moderate thrombocytopenia [4]. Many cases of DF patients might progress to dengue haemorrhagic fever (DHF) due to the increase in the vascular permeability. The evidence of the haemorrhagic is the main sign of the progression from DF to DHF. The first sign of DHF is Fever or history of acute fever lasting 2–7 days [3]. Second sign is the hemorrhagic tendencies evidenced by at least one of the following: positive tourniquet test (TT), petechiae, purpura, ecchymoses; bleeding from mucosa, gastrointestinal tract, injection sites or other location; haematemesis or melena. Third sign is Thrombocytopenia (100 000 cells per mm^3or less). Lastly, Haemoconcentration (20% or more rise in the haematocrit (Hct) value relative to baseline average for the same age, sex and population) or sings of plasma leakage such as pleural effusion, ascites and hypoproteinaemia. DHF patients were classified into four categories DHFI, DHFII, DHFIII, and DHFIV. Both DHFIII and DHFIV are considered as dengue shock syndrome (DSS) which is fatality stage.

Laboratory and blood testing are required in order to monitor the progress in dengue patients. On the other hand, these tests might be risky for dengue patients since frequent blood taking from dengue patients will cause further injury to the subcutaneous tissue [5]. Few studies were conducted to overcome the above mentioned complexity [5, 6, and 7]. One of these studies was utilized the bioelectrical impedance analysis (BIA) technique for monitoring and classifying the daily risk in dengue patients [5]. The study was proposed the reactance as an indicator for classify and prognosis the risk in dengue patients. Another study utilized the symptoms, signs in dengue patients to predict the day of defervescence of fever [6]. The study achieved 90% prediction accuracy for predicting the day of defervescence of fever. Even though, those studies achieved significant results, they used statistical analysis in finding their predictors.

Therefore, this study utilized the SOM to hunt the significant prognosis factors in order to differentiate between the dengue patients and the healthy subjects.

Self-organizing map (SOM) is considered as one of the most powerful technique for visualizing, understanding, and exploring the complexity of the high-dimensional data. It assists the user for hunting the significant correlations among the data [8]. It has been used for several biomedical applications such as medical diagnosis, prosthesis control and gene expression [9, 10, and 11].

35 predictors comprised of 17 BIA parameters, 18 symptoms/signs, were used in this study. The results showed that the reactance, intracellular water, ratio of the extracellular water and intracellular water, ratio of the extracellular mass and body cell mass are the main BIA predictors which can use to differentiate between the dengue patients and the healthy subjects. Moreover, Anorexia, Abdominal Epigastic pain, Petechiea Rash and bleeding tendency are the only symptoms/signs presented with dengue patients in the day of defervescence of fever.

II. Self Organizing Map (SOM)

The self-organizing map (SOM) is the dimension reduction technique which maps high-dimensional data into a simple low-dimensional display. It used for simplify the complexity of the data. Accordingly, three stages need to be performed in order simplify the data. The stages are competition, cooperation, and adaptation [12, 13, 14, 15, and 16]. The general procedures for performing these stages start with providing the inputs data to the neurons in the network. Secondly, all the neurons will competent each other to find the winning neuron or the best matching unit (BMU). The BMU will be determined base on minimum Euclidean distance between the inputs data and the neurons. After finding the BMU, the cooperation among neighbouring neurons will occurs in order to find the spatial location of a topological neighbourhood. Finally, the synaptic weights of the winning neuron and its excited neighbourhood will be adjusted with respect to the input samples. After the network was trained, the SOM will display in two dimensions. This map consists of different clusters. Each cluster comprises with the similar data.

III. Methodology

The implemented technique for determining the significant prognosis factors in dengue patients is shown in Fig. 1.

Fig. 1 The block diagram for the implemented procedures

Database of 329 data set (210 dengue patients and 119 healthy subjects) obtained from [17] on the day of defervescence of fever were used. Defervescence of fever is defined as the day when a patient has no fever [18]. The present study was conducted using the SOM toolbox [19].

The BIA data were normalized to 0 and 1 in order to be treated by SOM in the same way. On the other hand, the symptoms and signs data didn't normalize since it were recorded as 1 for presenting the symptom or the sign in the patients and 0 for the absent of the symptom or the sign. Then, all the vectors of the data samples were labeled and fed it to SOM. The healthy subjects were labeled as (H) while the dengue patients were labeled as (D).

Separated trainings were conducted to define the significant BIA, symptoms, and signs predictors. The first training utilized the BIA data while the second training utilized the symptoms and signs data. The training parameters were initialized as follow: the learning rate was 0.5 in the ordering phase and 0.05 in the convergence phase. The neighborhood was chosen as Gaussian while the network topology was chosen as hexagonal. Several training was conducted to find the best size of the SOM for the two sets of the data. After the training was conducted, the best maps size was visualized and the significant predictors were defined.

A. Symptoms and signs

18 symptoms and signs were investigated at the day of defervescence of fever. The data comprised of Fever, Headache, Dizziness and fainting (dizz/fain) , Weakness lower limb (wllimb), Arthralgia , Myalgia, Body ache , Nausea, Vomit, Anorexia, Abdominal Epigastic pain (gastric), Chill and rigor (chillnr), Petechiea Rash (p.rash), Flush face (flushf), Bleeding tendency (bt), Hepatomegaly (hepa), Conjunctivitis (con'vitis), Macular.

B. BIA data

17 BIA parameters were used. The BIA parameters comprised of, Gender (GED), Resistance (RES), Reactance (REACT), Phase Angle (PA), Body Capacitance (BC), Body Cell Mass (BCM), Extracellular Mass (ECM), Fat Mass (FM), ECM/BCM Body Mass Index (BMI), Basal Metabolic Rate (BMR), Total Body Water (TBW), Intracellular Water (ICW), Extracellular Water (ECW), lean body mass (LBM), (ERB) which is (ECM/BCM), (ERI) which is (ECW/ICW), TRT which is (TBW/W), and the weight (W).

IV. Results

Several trainings were conducted to identify the best map size for the two sets of the data. Accordingly, the best map size, the quantization error and topographical error are as shown in Table I.

TABLE I
SOM'S SIZE

Data set	Map units	Quantization error	Topographic error
BIA	100	1.313	0.033
Symptoms/signs	100	2.030	0.012

A. BIA data analysis

The u-matrix and labels map which is shown in Fig. 2 were obtained after the training was conducted utilizing the BIA data. All the data sample was comprised with 18 variables. Accordingly, the component planes which are shown in Figure 3 reflect the contribution of these variables to the u-matrix.

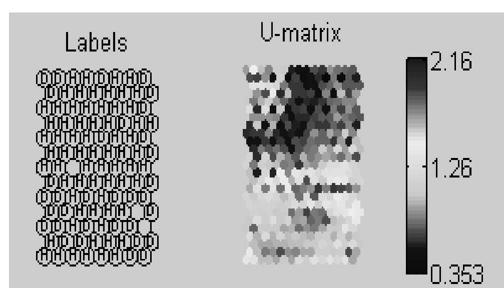

Fig. 2. Visualization of the u-matrix and the labels map for the BIA data

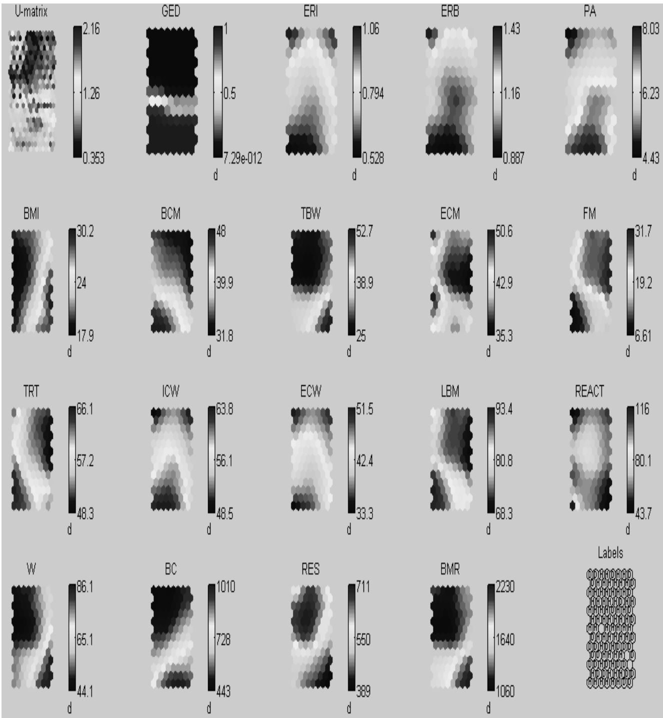

Fig. 3. Visualization of the components planes map for the BIA data

There are two main clusters presented in the u-matrix as shown in Fig. 2. The first cluster is appearing in the upper part of the map while the second cluster is in the lower part of the map. Based on gender component plane map shown in Figure 3, the upper cluster is related to the female data while the lower is related to the male data. The labels map shows that the dengue patients were distributed in the upper left and right corner, lower right corner and middle left side of the u-matrix.

Hunting the correlation between the BIA parameters and the dengue patients were based on the investigation of the similarity between the clusters in u-matrix and the component planes. The component planes showed that there are correlations exist among the u-matrix, ERI, REACT, ICW and ERB. The REACT and ICW component planes showed that the values of REACT and ICW in dengue patients are significantly low than the healthy patients. Moreover, the ICW is lower in the dengue female patients as compared with the dengue male patients. The ERI and ERB component planes showed that the values of the ERI and ERB are higher in the dengue patients as compared with the healthy patients. Furthermore, ERI, ERB, are higher in the dengue female patients as compared with the dengue male patients. The results of the BIA parameters were summarized in Table II.

TABLE II
SIGNIFICANT BIA PARAMETERS

	Dengue patients		Healthy subjects	
BIA parameters	Male	Female	Male	Female
REACT	Low	Low	Normal	Normal
ERI	High	Higher	Normal	Normal
ERB	High	Higher	Normal	Normal
ICW	Lower	Low	Normal	Normal

Hence, the significant BIA prognosis factors are; reactance, intracellular water, ratio of the extracellular water and entracellular weater, ratio of the extracellular mass and body cell mass.

B. *Symptoms and signs analysis*

Figures 4 and 5 show the SOM for the symptoms and signs in the dengue patients during the day of defervescence of fever.

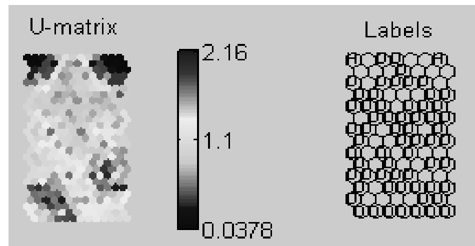

Fig. 4. Visualization of the u-matrix and the labels map for the symptoms and signs data

The u-matrix in Fig. 4 shows that there are three clusters; first cluster is in the upper right side of the map, second cluster in the upper left side of the map whilst the last cluster is the remaining part of the map. The labeling map indicate that the first and the second clusters are related to the healthy subjects since there are labeled as H. The last cluster represents the dengue patients.

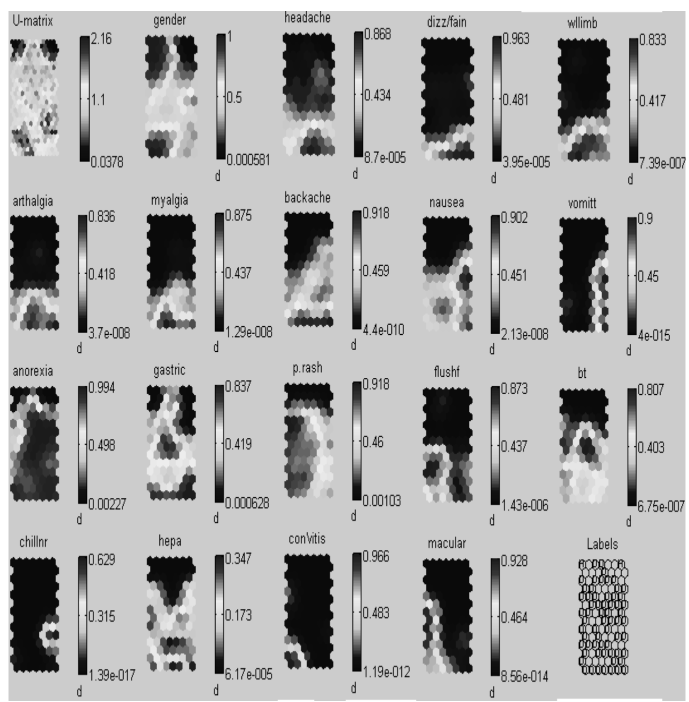

Fig. 5. Visualization of the components planes map for the symptoms and signs data

Figure 5 shows the component planes which reflect the contribution of symptoms and the signs in the u-matrix. The maps showed that the correlated component planes are

Anorexia (anorexia) and Abdominal Epigastic pain (gastric). Moreover, Petechiea Rash (p.rash) and bleeding tendency (bt) are semi-correlated since they represent in the majority of the dengue patients. The other component planes didn't show any significant for differentiating the dengue patients from the healthy subjects. Thus, those signs/symptoms didn't present in dengue patients in the day of defervescence of fever

Thus, the results indicated that only the Anorexia, Abdominal Epigastic pain, Petechiea Rash and bleeding tendency, are presented with dengue patients in the day of defervescence of fever. Therefore, these signs were considered as significant prognosis factors for dengue patients.

VI. Conclusion

This paper examined 35 factors, including 17 BIA parameters and 18 clinical symptoms/signs in dengue patients prognosis via self organizing map. The best map size was found as 100 units with 1.313 quantization error and 0.033 topographic error when the BIA parameters were used. On the other hand, the best map size was obtained as 100 units with 2.03 quantization error and 0.012 topographic error, when the symptoms and signs were used. Hunting the correlation by SOM was conducted based on the u-matrix and the component planes map. The results showed that the reactance, intracellular water, ratio of the extracellular water and intracellular water, ratio of the extracellular mass and body cell mass, are the significant BIA factors. On the other hand, Anorexia, Abdominal Epigastic pain, Petechiea Rash, and bleeding tendency are the main signs and symptoms that presented with dengue patients on the day of defervescence of fever.

Acknowledgment

This research is supported by E-Science fund grant Project Title: Development of an intelligent Prognostic system for dengue infections (11-02-03-1014).

References

[1] TP. Monath, "Dengue: the risk to developed and developing countries," *Proc Natl Acad Sci USA 91*, pp. 2395-2400, 1994.

[2] DJ. Gubler, "Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century," *Trends Microbiol*, vol 10, pp. 100-103, 2002.

[3] World Health Organization, *Dengue Haemorrhagic Fever: Diagnosis, Treatment, Prevention and Control,* 2^{nd} ed., Geneva, 1997.

[4] DJ. Gubler, "Dengue and dengue hemorrhagic fever," *Clin Microbiol Rev*, vol. 11, pp. 480-496, 1998.

[5] F. Ibrahim et al. "A Novel approach to classify Risk in Dengue Hemorrhagic Fever (DHF) using Bioelctrical Impedance," *IEEE Transaction on instrumentation and measurement*, vol. 54, no. 1, 2005.

[6] F. Ibrahim et al. "A Novel Dengue fever (DF) Dengue and Hemorrhagic fever (DHF) Analysis using artificial neural network," *Computer methods and programs in biomedicine*, vol. 79, pp. 273-281, 2005.

[7] F. lbrahim, et al. "Modeling of hemoglobin in dengue fever and dengue hemorrhagic fever using bioelectrical impedance," *Physiol. Meas.*, vol. 25, pp. 607-616, 2004.

[8] J. Vesanto and J. Ahola, "Hunting for Correlations in Data using the Self- Organizing Map," *In Proc. CIMA'99 Computational Intelligence Methods and Applications*, pp. 279-285, 1999.

[9] I. Christodoulos, S. Constantinos Pattichis "Medical Diagnostic Systems using Ensembles of Neural SOFM Classifiers," *IEEE*, 1999

[10] F. Sebelius et al. "Classification of motor commands using a modified self-organising feature map," *Medical Engineering & Physics 27*, pp. 403-413, 2005.

[11] Janne Nikkila et al. "Analysis and visualization of gene expression data using Self-Organizing Maps," *Neural Networks*, vol. 15, pp. 953-966, 2002.

[12] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, New Jersey, 2^{nd} Edition, 1999.

[13] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, pp. 1464-1480, September 1990.

[14] Arsuaga Uriarte, and F. Díaz Martín, "Topology Preservation in SOM," *international journal of applied mathematics and computer sciences*, vol. 1, no. 1, ISSN 1307-6906, 2004.

[15] T. Kohonen, *Self-Organizing Maps*, 2^{nd} ed., Springer-Verlag, Berlin, 1997.

[16] Jing Li, "Information Visualization with Self-Organizing Maps," Available: at http: //www14.in.tum. de/ konferenzen/ Jass05/courses/6/ Papers/09.pdf.

[17] F. Ibrahim, "Prognosis of Dengue Fever and Dengue Hemorrhagic Fever Using Bioelectrical Impedance," Ph.D thesis, University of Malaya, Malaysia, 2005.

[18] Oxford University Press, Concise medical dictionary, 3^{rd} ed., Oxford, 1990

[19] J. Vesanto, E. Alhoniemi, J. Himberg, K. Kiviluoto, J. Parviainen "Self-organizing map for data mining in MATLAB: the SOM toolbox," *Simulation News Europe*, pp. 25–54, 1999.