13th Workshop on ADC Modelling and Testing
Sep. 22-24, 2008, Florence, Italy

SOM Approach in Monitoring and Diagnosis of Obesity-Hypertension

Octavian Postolache[1], Monica Ferreira[2], Gabriela Postolache[3], Pedro Silva Girão[1]

[1]*Instituto de Telecomunicações/Instituto Superior Tecnico, Lisboa,Portugal*
[2]*Hospital Santa-Maria, Lisboa,*
[3]*Escola Superior de Saude, Universidade Atlântica, Oeiras*
Emails: poctav@alfa.ist.utl.pt, gabrielap@uatla.pt, psgirao@ist.utl.pt

Abstract - Obesity-hypertension is an emerging concept in pathophysiology. Obesity and hypertension have been turned into an epidemic afflicting all the word, being among the mainly factors that have been burning the health costs. This paper reports a study whose purpose was to develop an objective method to better diagnose and manage this pathophysiology. A data management and data mapping system was developed. Kohonen topological mapping was used in the classification of obesity-hypertension considering clinical characteristics and laboratory results. Thus, the n-dimensional space of physiopathological parameters was converted into a 2D space of the following obesity classes: healthy subject, overweight, obesity class I, obesity class I-hypertension, obesity class II, obesity class II-hypertension, obesity class III, and obesity class III-hypertension. Transient changes in the individual state could also be analyzed using the proposed self-organizing map based model. Characteristics of the designed maps, such as topology and quantification errors, were studied.

keywords: neural network, data-mapping, obesity-hypertension syndrome

I. Introduction

Obesity is rapidly turning into an "epidemic" afflicting much of the industrialized world. A study carried by the Institute of European Food Studies (IEFS) in 1997 in the 15 European Member States showed that the prevalence of obesity is higher in United Kingdom (12%), followed by Spain (11%), while it was lower in Italy, France and Sweden (7%) [1]. Medical expense for obese people is at least 25% higher than for normal weight people [2]. Obesity accounts for 2-6% of total health care costs in several developed countries; some estimates put the figure as high as 7% [3]. The true costs are undoubtedly much higher as not all obesity-related conditions are included in the calculations.

There is a lot of knowledge on obesity, but thoroughly view of the phenomenon remains to be done. Despite a sustained preventive work against increase obesity and hypertension, the efforts to manage obesity and hypertension have been soundly defeated. A new perspective is needed for better diagnosis and management of obesity-hypertension. Few works on stratification of risk factors and associated clinical conditions with obesity have been done [4]. To deal with these issues, much research is needed to develop improved statistical methods. In the last years, bio-informatics approach to detect complex pattern and dynamics has been developed. Many researchers are exploring variations and modifications of logistic regression, and automated detection of informative combined effects (e.g. [5]). Additional explorations are being conducted in data mining and machine learning research. Data reduction involves a collapsing or mapping of the data to a lower dimensional space. Pattern recognition, on the other hand, involves extracting patterns from the data to discriminate between groups by using the full dimensionality of the data. Examples of pattern recognition methods include cluster analysis [6], cellular automata [7], support vector machines [8], self-organizing maps [9] and neural networks [10]. We used a self-organizing map (SOM) algorithm because of its ability to visualize multidimensional data in a two-dimensional format, and to make data reduction and abstraction by generating prototype vectors from measurement data. The SOM is especially suitable for exploratory data analysis of large data sets [11,12,13].

We focused on potential cumulative risk of hypertension and obesity. Organ damage and associated clinical condition in obese people increase with the extent of risk factor clustering. Furthermore, according to the ESH-ESC guidelines, hypertension induces high-added risk for target organ damage, diabetes, or associated clinical conditions [14]. Therefore it is worth to diagnose obesity-hypertension relation.

The problem with the diagnosis and management of obesity is that the relationship between different items (e.g. laboratory results and/or symptoms) is not always well established, and that there exists a myriad of exceptions for every rule. Worldwide, 45% of all physicians reported never measuring waist

circumference, and 52% overestimated the waist girth that puts their patient at risk. More than half (59%) of at risk patients had not been informed by their doctors about the link between abdominal obesity and heart disease [15].
Since today, data analysis related with obese population using SOMs was little explored. A study in Finland has shown that SOM is a usefully tool in describing and modelling data for extracting cluster related to insulin resistance syndrome and cardiovascular disease from a large prospective population-based study in middle aged men [16]. Also generational trends in obesity in the United States were better analysed using wavelet based SOM analysis that emphasized that older people had a lower prevalence of obesity than younger adults [4]. These results challenged an assumption that all age groups were equally exposed to increasing sedentary behaviours, unhealthy foods, and over consumption of food calories.
The aim of this study was to describe the usefulness of the SOM algorithm in modelling and presentation of multivariate clinical related with ethiopathogeny of obesity-hypertension syndrome.

II. Methods

The data acquisition, data management and data mapping system software block diagram is presented in Fig.1. Thus, the input data associated with the patients obtained from instruments are uploaded in the database that is connected to the self-organizing map (SOM). The input, output structure and training algorithm, the most important items in the designing of a clinical parameters distribution model, are discussed next.

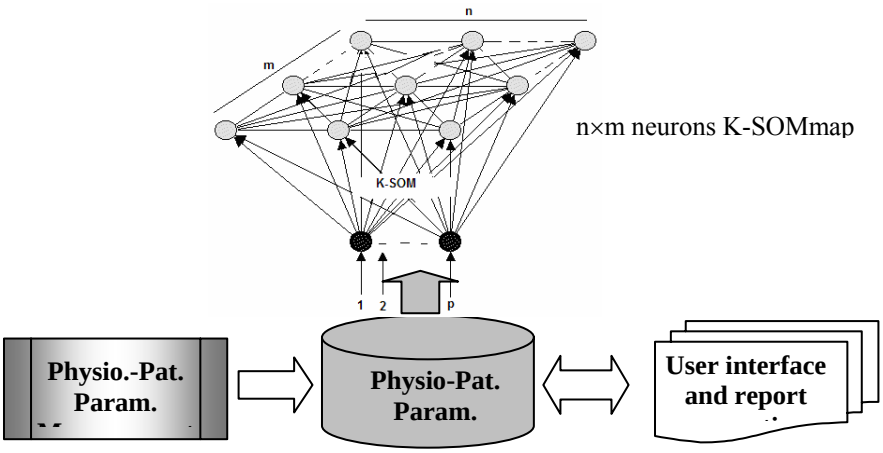

Fig.1. The data management and data mapping system block diagram

A. The input of the model

There is a lack of studies that may give a thorough view on the main clinical indicators [see 19], which may better describe pathophysiological changes to obesity and obesity-hypertension. Since there were no well-characterized real datasets available that fit all obesity classes described in our study, a simulation study was needed. The values for different obesity classes were adjusted according to published data, considering a Gaussian distribution of the values. The maximum values for the morbid obese group were built taking into account clinical cases of obesity described in the 19th century – Daniel Lambert. At the time of the Lambert's death, in 1809 in Stamford, England, he was 39 years of age and weighed 336 kg. The SOM was designed using the simulated data expressed by 1600x19 matrix (SOM_{19}) of individuals under test. Using 19 parameters for each individual, an accurate characterization of obesity-hypertension or tendency is carried out. However, in practice a small number of clinical data are used on diagnosis of obesity hypertension. In this case a SOM approach using 1600X10 matrix (SOM_{10}) was used to design a low complexity SOM. In Table 1 are represented the data range for variables include in SOM_{19} and SOM_{10} characteristics to the normal and obesity-hypertension class III. The model input data include essential clinical information from an obesity-hypertension associations that are expressed by the values of the following parameters: body mass index (BMI), waist girth, waist-to hip ratio, blood pressure, insulin, triglyceride, plasma free fatty

acids, leptin, protein C, glucose, fibrinogen, angiotensin II, atrial natriuretic peptide, HDL cholesterol [17], total cholesterol [18,19], LDL cholesterol [20,21], heart rate as an index of sympathetic autonomic control activation [21] and circulating ghrelin levels [22].

Table 1. Variables range used to describe obesity-hypertension syndrome class III

Variables	Normal	Obesity class III+Hypertension
Body mass index [a] (kg/m^2)	18.5-24.9	40-85
Waist girth [a] (cm)	58-88	150-170
Waist-to-hip ratio (cm/cm)	0.6-0.9	1.4-1.6
Systolic blood pressure [a] (mmHg)	90-125	140-190
Diastolic blood pressure [a] (mmHg)	55-84	90-120
Heart rate [a] (bpm)	55-95	90-140
Serum triglycerides [a] (mg/dL)	30-173	200-350
Total cholesterol [a] (μmol/L)	250-680	200-350
Serum HDL cholesterol [a] (mg/dL)	50-60	35-50
Serum LDL cholesterol [a] (mg/dL)	50-130	30-40
Plasma free fatty acids (μmol/L)	250-680	700-1100
Leptin (ng/mL)	4-15	9-35
Insulin (pmol/L)	36-140	500-1600
Ghrelin (ng/mL)	0.3-0.9	0.1-0.2
Angiotensin II (pg/mL)	5-16	10-22
Atrial natriuretic peptide (fmol/L)	15-27	9-12
Glucose (mg/dL) [a]	70-110	90-220
Fibinogen (g/L)	2-4	10-12
Protein C reactive (mg/L)	5-15	20-30

a – variable included in SOM_{19} and SOM_{10} algorithm

The data from 30 adult patients (12 men and 18 women) with voluntary participation in the study were used to test sensibility of SOM algorithm to classify patients with obesity-hypertension syndrome. The main characteristics of the subjects included on the study are represented in the Table 2.

Table 2. The main characteristics of the patients included on the study

	Median	Average
Age (y)	58.50(33-87)	60.37
BMI(kg/m^2)	24.37 (18.57-46.88)	25.21
Waist Girth (cm)	101.00 (67-131)	100.90
SAP (mmHg)	158.10 (112-189)	161.00
DAP (mmHg)	81.50 (56-101)	78.90

The simulated data were not defined using sex-specific observation points. However, a future study considering the data distribution versus sex and age will be considered. Although the parameter settings are not exhaustive of the physiopathologically plausible situations, the outlined conditions are reasonable, mainly designed to differentiate obesity and hypertension features.
Actually diagnosis of obesity is made mainly according to the BMI index:

$$BMI \left[kg / m^2\right] = \frac{m}{h^2} \quad (1)$$

Normal people have a BMI in the range 18.5-24.9. Obesity is defined when BMI>30; morbid obesity corresponds to BMI>35 [23,24,25]. A continuous relationship between gradation of BMI and health risk and between waist circumference and health risk exists. Individuals with central obesity (android or visceral obesity type) are considered to have higher cardiovascular risk. Moreover, to better characterize central obesity, waist-to-hip ratio [17,20] and visceral adipose tissue assessed by multidetector computerized tomography [26] should be used.

B. Model architecture and training

A Kohonen Self-Organizing Map (K-SOM) [15,16] was designed and implemented for our study taking into account a data compression that preserves the most important topological and metric relationship of the primary data. The internal parameters of K-SOM (weights connection) are obtained based on an unsupervised learning process. Thus, for a given number of map cells (neurons) the K-SOM prototype vector, m_{PPi}, associated with the cells is randomly initialised and updated during the training according to the following learning rule:

$$m_{PPi}(t+1) = m_{PPi}(t) + h_{ci} \cdot \left(pp(t) - m_{PPi}(t)\right) \quad (2)$$

where *pp(t)* represents an input vector randomly drawn from the input data set at time t, *pp(t)*=[pp_1 pp_2 ….pp_q], and h_{ci} is called neighbourhood kernel around the winner cell c and defined by:

$$h_{ci} = \alpha(t) \times \exp\left(-\frac{\|r_i - r_c\|^2}{2 \times \sigma^2(t)}\right) \quad (3)$$

$0<\alpha(t)<1$ is the learning rate at time t, $\|r_i - r_c\|$ is the distance between cells c and I within the output space (map), and $\sigma(t)$ corresponds to the width of the neighbourhood function [15,16]. A practical approach concerning the evolution of topology error and quantification errors for different number of neurons is considered on the map.

C. The output of the model

We developed a model that simulates and optimises an obesity diagnosis. The model's output is expressed by the designed map including m×n neurons distributed in different clusters (8 clusters in the present application) associated with the obesity classes. Based on the best matched unit (BMU), an algorithm for new incoming data permits the visualization on the designed K-SOM map of individual diagnosis, thus providing a fast obesity-hypertension diagnosis. The accuracy of the match for the input vectors was studied.

III. Results and discussion

Using a set of 19 physiopathological parameters considered for 1600 individuals distributed in 8 classes, an 8×25 K-SOM_{19} was designed. Additionally, a reduced number of considered parameters for data mapping were considered. Thus, for 10 parameters, the 8 classes are expressed by 9 X 22 K-SOM_{10}. The cluster separation for both cases is presented in Fig. 2. The K-SOMs performance was expressed by quantification errors qe_{10}=1.35 and qe_{19}=2.16 while the topologic errors were: te_{10}=0,04 and te_{19}=0.04. In order to better express the distribution of the input data between clusters, two representations were used on the designed map: hit histograms and bar-plane distribution. For the above considered case, the bar-plane distribution is presented in Fig. 3.
The complexity of the etiopathogenesis of the obesity-hypertension syndrome is visualized in the SOM hit histogram. Thus, overweight tendency is expressed by distribution of some physiological parameters interference, expressed by superposition of parts of hit histogram. A normal (red color histogram) is superposed on overweight (blue colour histogram). A practical approach concerning the introducing of a new parameter for a better separation between the clusters, as well as the optimal K-SOM design, was considered. The utilization of the histogram can be used to better express the relative proportion of each component (e.g. leptin) on the sum of all components in the map unit. The incoming vector, including the considered parameters, is expressed on the map.

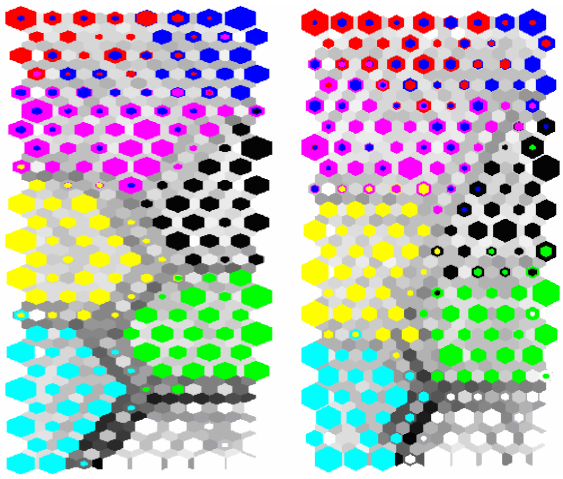

Fig.2. Obesity-hypertension diagnosis based on K-SOM results: SOM_{19} and SOM_{10} clusters of data distributions for 19 and 10 variables

In the hit-histogram, better clustering of data in SOM_{19} than in SOM_{10} can be observed (normal individuals are represented in red, the overweight individuals in blue, obesity class I in black, obesity class II in green and the obesity class III in white, mangenta – obesity class I+hypertension, yellow-obesity class II + hypertension, light-blue – obesity class III+hypertension,). The identification of normal, overweight and obesity class I clusters is not straightforward in SOM_{10}, and needs further clarification by using additional variables.

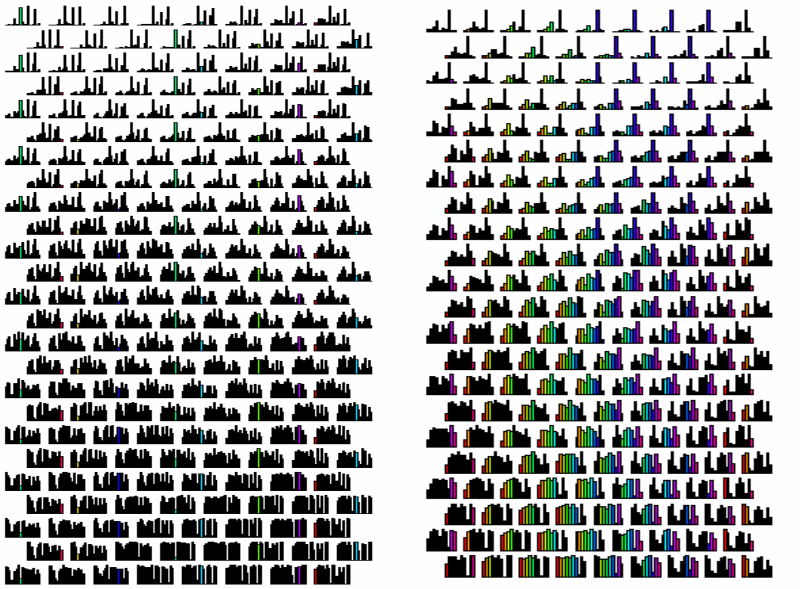

Fig. 3 Obesity-hypertension diagnosis based on K-SOM results: K-SOM bar-plane distribution for different obesity classes and different input parameters: 19 parameters on the left and 10 parameters on the right

The histogram expresses the distribution of the considered parameters for individuals for each K-SOM_{19} and K-SOM_{10} cells. From left to right the included parameters were defined in the methods paragraph (Table 1).
After SOM_{10} designing process, testing data obtained from a group of the study mentioned above was used to characterize the data mapping capabilities of the designed architecture. Thus, the representation accuracy associated with defined clusters of obesity classes was checked out. For a small dimension testing set (30X10) the designed SOM proves to be a good classifier: there was a 70% success of representation of the experimental input data on the right cluster, 10% indecision and 20% wrong classification.
An implemented SOM software module permits also to represent the patient's obesity state according to transient on obesity characteristics parameters measured on different time intervals (weeks, months).

IV. Conclusion

Obesity and overweight, as part of the metabolic syndrome, are important risk factors for the development of diabetes, hypertension, coronary heart disease, hyperlipidemia, stroke, sleep apnea syndrome, osteoarthritis and certain forms of cancer. In our developed model, we provide an instrument for primary care physicians to better diagnose and monitoring individuals with risk factors associated with the clinical obesity-hypertension syndrome. The presented classification of biomedical data gave satisfactory results on a preliminary screening examination stage. The self-organized map proved to be an important tool to better characterize obesity and obesity-hypertension association. The developed algorithm permits diagnosis and management of obesity-hypertension physiopathology.

References

[1] Varo JJ, Martinez-Gonzalez MA, Martinez JA, "Obesity prevalence in Europe", *An Sist Sanit Navar*, vol. 25, Suppl 1, pp.103-8, 2002.

[2] Lewis KK, Man LH, "Overweight and obesity in Massachusetts: epidemic, hype or policy opportunity?" *Issue Brief (Mass Health Policy Forum),* vol. 23;(30), pp.1-29, 2007.

[3] WHO, http://www.who.int/dietphysicalactivity/publications/facts/obesity/en.

[4] Garavaglia SB, Synthelabo S, "Generational trends in obesity in the United States: analysis with a wavelet coefficient self-organizing map", *Proceedings 2004 IEEE International Joint Conference on Neural Networks,* vol. 1, pp. 769- 774, 2004.

[5] Tahri-Daizadeh N, Tregouet D, Nicaud V, Manuel N, Cambien F, Tiret L, "Automated detection of informative combined effects in genetic association studies of complex traits", *Genome Res*, vol.13, pp. 1952-1960, 2003.

[6] Kaufman L, Rousseeuw PJ, *Finding Groups in Data : an Introduction to Cluster Anaysis*. Hoboken : Wiley, 2005.

[7] Wolfram S, *Cellular Automata and Complexity :* Collected Papers. Reading, MA : Addison-Wesley, 1994.

[8] Cristianini N, Shawe-Taylor J, *An introduction to support vector machines*, Cambridge University Press, 2000.

[9] Hastie T, Tibshirani R, Friedman JH, *The Elements of Statistical Learning. Data Mining, Interference, and Prediction,* New York : Springer, 2001.

[10] Ripley BD, *Pattern Recognition and Neuronal Networks*, Cambridge : Cambridge University Press, 1996.

[11] Kohonen T. *Self Organizing Maps*, 3rd ed. New York, Spinger-Verlag, 1997.

[12] Postolache OA, Silva Girão PMB, Pereira JMD, Ramos HMG, "Self-organizing maps application in a remote water quality monitoring system", *IEEE Transactions on Instrumentation and Measurement,* vol. 54(1), pp.322-329, 2005.

[13] Oja E, Valkealahti K, "Local independent component analysis by the self- organizing map", *Proceedings of International Conference on Artificial Neural Networks (ICANN'97),* pp. 553- 558, 1997.

[14] Guidelines Committee. 2003 European Society of Hypertension - European Society of Cardiology guidelines for the management of arterial hypertension", *J Hypertens*, vol. 21, pp. 1011-1053, 2003, http://www.eshonline.org/ documents/2003_guidelines.pdf.2003.

[15] Smith SC Jr, Haslam D. Abdominal obesity, waist circumference and cardiometabolic risk: awareness among primary care physicians, the general population and patients at risk-the Shape of the Nations survey. *Curr Med Res Opin*, vol. 23(1), pp. 29-47, 2007.

[16] Valkonen V-P, Kolehmainen M, Lakka H-M, Salonen J, "Insulin resistance syndrome revisited: application of self-organizing maps", *International Journal of Epidemiology* , vol. 31(4), pp. 864-871, 2002.

[17] Aneja A, El-Atat F, McFarlane SI, Sowers JR, "Hypertension and obesity", *Recent Progress in Hormone Research*, vol. 59, pp. 169-205, 2004.

[18] Ai M, Tanaka A, Ogita K, Sekine M, Numano F, Numano F, Reaven GM, "Relationship between hyperinsulinemia and remnant lipoprotein concentrations in patients with impaired glucose tolerance", *J Clin Endocrinol Metab*, vol. 85(10), pp. 3557-60, 2000.

[19] Aguilera CM, Gil-Campos M, Cañete R, Gil A, "Alterations in plasma and tissue lipids associated with obesity and metabolic syndrome", *Clin Sci (Lond*), vol. 114(3), pp. 183-93, 2008.

[20] Gupta R, Rastogi P, Sarna M, Gupta VP, Sharma SK, Kothari K, "Body-mass index, waist-size, waist-hip ratio and cardiovascular risk factors in urban subjects", *J Assoc Physicians India,* vol. 55, pp. 621-7, 2007.

[21] Kim JA, Park Y-G, Cho K-H, Hong M-H, Han H-C, Choi Y-S,Yoon D, "Heart rate variability and obesity indices: emphasis on the response to noise and standing", *J Am Board Fam Pract,* vol.18, pp. 97-103, 2005.

[22] Oner-Iyidoğan Y, Koçak H, Gürdöl F, Oner P, Issever H, Esin D, "Circulating ghrelin levels in obese women: a possible association with hypertension", *Scand J Clin Lab Invest,* vol. 67(5), pp. 568-76, 2007.

[23] Sowers JR, Epstein M, Frohlich ED, "Diabetes, hypertension, and cardiovascular disease: an update", *Hypertension,* vol. 37(4), pp. 1053-9, 2001.

[24] World Health Organization, *Obesity: Preventing and Managing the Global Epidemic. WHO Obesity Technical Report Series No. 894* World Health Organization Geneva, Switzerland, 2000.

[25] Health Canada, *Canadian guidelines for body weight classification in adults*, www.hc-sc.gc.ca/ fn-an/ alt_formats/ hpfb-dgpsa/ pdf/ nutrition/ weight_book-livres_des_poids, 2003.

[26] Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu CY, Vasan RS, Murabito JM, Meigs JB, Cupples LA, D'Agostino RB Sr, O'Donnell CJ, "Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study", *Circulation,* vol. 3; 116(1), pp. 39-48, 2007.