Neural networks to estimate the risk for preeclampsia occurrence

Costas K. Neocleous, Panagiotis Anastasopoulos,
Kypros H. Nikolaides, Christos N. Schizas, Kleanthis C. Neokleous

***Abstract*—A number of neural network schemes have been applied to a large data base of pregnant women, aiming at generating a predictor for the risk of preeclampsia occurrence at an early stage. The database was composed of 6838 cases of pregnant women in UK, provided by the Harris Birthright Research Centre for Fetal Medicine in London. For each subject, 24 parameters were measured or recorded. Out of these, 15 parameters were considered as the most influencing at characterizing the risk of preeclampsia occurrence. A number of feedforward neural structures, both standard multilayer and multi-slab, were tried for the prediction. The best results obtained were with a multi-slab neural structure. In the training set there was a correct classification of the 83.6% cases of preeclampsia and in the test set 93.8%. The preeclampsia cases prediction for the totally unknown verification test was 100%.**

I. Introduction

Preeclampsia is a syndrome that may appear during pregnancy and can cause perinatal and maternal morbidity and mortality. It affects approximately 2% of pregnancies [1]-[2]. It is characterized by hypertension (if isolated called Pregnancy-Induced Hypertension (PIH)) and by significant protein concentration in the urine (proteinuria). Such a high blood pressure may result in damage to the maternal endothelium, kidneys and liver [3]-[4].

The time that preeclampsia may occur is during the late 2nd or 3rd trimester. It has also been observed that it is more common to women on their first pregnancy.

The prevailing conditions that lead to preeclampsia are not well understood, hence its detection depends on signs or investigations. The likelihood of developing preeclampsia is increased by a number of factors in the maternal history, including Afro-Caribbean ethnicity, nulliparity, high body mass index (BMI), and previous or family history of preeclampsia [6]. However, screening by maternal history alone will detect only 30% of those who will develop the condition, with a false positive rate of 10%. Thus, the early diagnosis of preeclampsia is difficult, and the prognosis even more hard.

Attempts of preeclampsia prevention by using prophylactic interventions have been largely unsuccessful [7]-[8].

Costas Neocleous (Corresponding author) is with the Department of Mechanical Engineering, Cyprus University of Technology, Lemesos, CYPRUS. e-mail: costas.neocleous@cut.ac.cy.
Christos Schizas and Kleanthis Neokleous are with the Department of Computer Science, University of Cyprus, 75 Kallipoleos, 1678, POBox 20537, Nicosia, CYPRUS. e-mails: schizas@ucy.ac.cy; kleneokl@cs.ucy.ac.cy.
Panagiotis Anastasopoulos and Kypros Nikolaides are with the Harris Birthright Research Centre for Fetal Medicine, King's College Hospital Medical School, Denmark Hill, SE5 8RX, London, United Kingdom. e-mails: panasta@yahoo.com, fmf@fetalmedicine.com.

For this reasons, any tool that may improve the detection of preeclampsia, as for instance a reliable predictor or a method for the effective and early identification of the high-risk group, would be of great help to obstetricians and of course to pregnant women.

In recent years, neural networks and other computationally intelligent techniques have been used as medical diagnosis tools aiming at achieving effective medical decisions incorporated in appropriate medical support systems [9]-[11]. Neural networks in particular have proved to be quite effective and have also resulted in suitable patents [12]-[13].

II. Data

This study follows the result of a prospective screening study for preeclampsia in singleton pregnancies from the greater London area and South-East England. All of the women were attending the clinical center for routine assessment of risk for chromosomal abnormalities.

This is performed by measurement of fetal nuchal translucency thickness, maternal serum free human chorionic gonadotropin (fhCG) and serum pregnancy-associated plasma protein A (PAPP-A) at 11 to 13+6 weeks of gestation. Gestational age was derived from the fetal crown-rump length (CRL).

Written informed consent was obtained from the women agreeing to participate in the study, which was approved by King's College Hospital Ethics Committee.

Patients were asked to complete a questionnaire on maternal age, ethnic origin (White, Afro-Caribbean, Indian or Pakistani, Chinese or Japanese, or mixed), cigarette smoking during pregnancy, alcohol intake during pregnancy, drug abuse during pregnancy, medical history, medication, parity (parous or nulliparous if no delivery beyond 23 weeks), obstetric history (including previous pregnancy with preeclampsia), and family history of preeclampsia (sister, mother, or both). The maternal weight and height were measured, and the BMI was calculated in kilograms per meter squared.

The Blood Pressure (BP) was taken by automated devices, and the arm with the highest final MAP for the subsequent analysis of results was used [14]-[15].

For the measurement of uterine artery pulsatility index (UPI) a sagittal section of the uterus was obtained with each uterine artery along the side of the cervix at the level of the internal cervical os. Pulsed wave Doppler was used with the sampling gate set at 2 mm to cover the whole vessel with the angle of insonation being less than 50°. When three similar consecutive waveforms were obtained, the UPI was measured and the mean PI of the left and right arteries was calcu-

lated.

The database was composed of 6838 cases of pregnant women. These were provided by the Harris Birthright Research Centre for Fetal Medicine (FMF) in London. For each woman, 24 parameters were logged. Some of these parameters were socio-epidemiologic, others were records from a clinical examination and a third group from laboratory measurements.

The number of cases that preeclampsia occurred were only 116, which is a very small sub-database (1.7%) to be used for training of a neural system aiming at achieving good generalization. The definitions of pre-eclampsia used were those of the guidelines of the International Society for the Study of Hypertension in Pregnancy [16].

From the available data some parameters were excluded, based on recommendations from medical experts, thus only 15 parameters, out of the total of 24, were ultimately considered as the most influencing at characterizing the risk of preeclampsia occurrence, and were used in the built-up of the neural predictor. These parameters are shown in Table 1.

TABLE I
PARAMETERS THAT WERE USED
FOR PREECLAMPSIA PREDICTION

Mean arterial pressure (MAP)
Uterine pulsatility index (UPI)
Serum marker PAPP-A
Ethnicity
Weight
Height
Smoking? (Y/N)
Alcohol consumption? (Y/N)
Previous preeclampsia case?
Conception (spontaneous, ovulation drug or IVF)
Medical condition of pregnant woman
Drugs taken by the pregnant woman
Gestation age (in days) when the crown rump length (CRL) was measured
Crown rump length
Mother had preeclampsia? (Y/N)

The parameters were encoded in appropriate numerical scales that could make the neural processing to be most effective.

A test set of 36 cases was extracted and used to test the progress of training. This data set included 16 cases (44%) of women that exhibited preeclampsia.

Also, a verification data set having 9 cases of which 5 were with preeclampsia (56%), were extracted to be used as totally unknown to the neural network, and thus used for checking the prediction capabilities of each network.

III. Neural Predictor

A number of feedforward neural structures, both standard multilayer, of varying number of layers and neurons per layer, as well as multi-slab of different structures, sizes, and activation functions, were systematically tried for the prediction. This was done in a planned manner so that the best architecture would be obtained.

Considering the results obtained by such a systematic search, it was possible to ultimately select and use a multislab neural structure of four slabs connected as depicted in Figure 1.

All weights were initialized to 0.3, while the learning rate was the same for all connections, having value of 0.1. Similarly, the momentum rate was 0.2 for all links.

The test set was applied at the end of each epoch to test the progress of training. If the results of the testing at time t were better than those at time $t - 1$, the weights were saved as a better set.

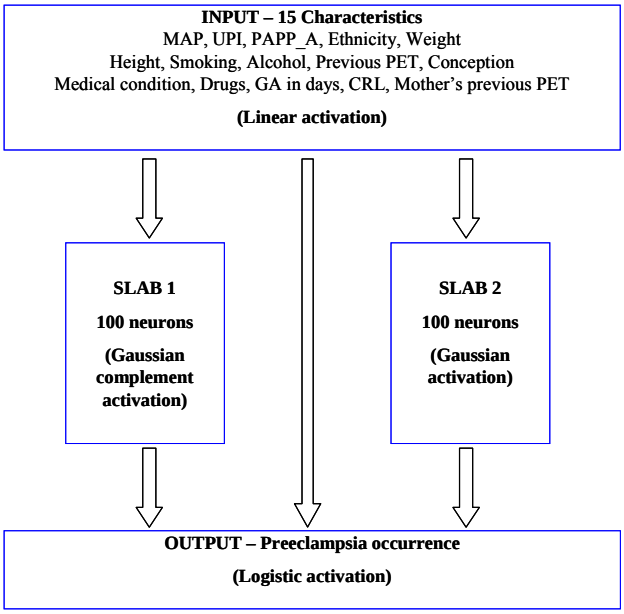

Fig. 1. The neural structure that was ultimately selected and used.

The training progress was monitored to observe whether there was improvement during the application of the training and test set data. For most network structures attempted, there was little generalization improvement after about 1200 epochs, as depicted in Figure 2.

Different sets of inputs were used to find an effective neural structure that would predict preeclampsia to an acceptable level. The inputs that were ultimately selected are those shown in Table 1.

IV. Results

In Table 2 an overall picture of the prediction results is presented.

The best results were obtained with a multi-slab neural structure of the type described in Figure 1. In the training set there was a correct classification of the 83.6% cases of preeclampsia and in the test set 93.8%. The preeclampsia cases

predicted correctly for the totally unknown verification test was 100%. In this set however, two subjects were predicted to exhibit preeclampsia, while they didn't.

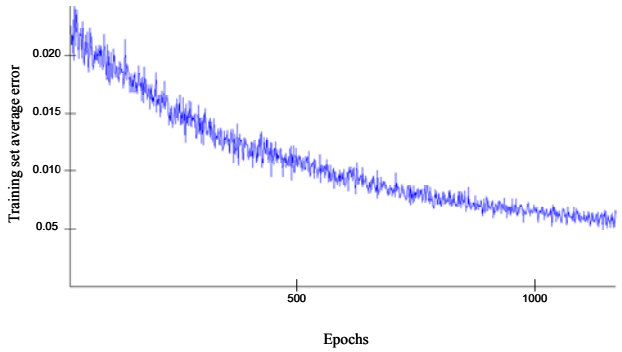

Fig. 2. Typical progress of training: Training set average error vs Epochs.

TABLE 2
PREECLAMPSIA PREDICTION RESULTS

	TRAINING SET	TEST SET	VERIFICATION SET
No of subjects in the database	6793	36	9
No of preeclampsia cases	116	16	5
Percentage of preeclampsia cases	1.7	44.4	55.6
Cases predicted correctly	3024	26	7
Percentage of cases predicted correctly	44.5	72.2	77.8
Preeclampsia cases predicted correctly	97	15	5
Percentage of Preeclampsia cases predicted	83.6	93.8	100

From a parameter contribution analysis it was observed that all 15 parameters contributed to the prediction. However, the mean arterial pressure and the uterine pulsatility index were the most influential, while the "Ethnicity" parameter was the least contributing.

V. Conclusion and Future Work

Attempts to predict preeclampsia using multivariate statistics have been reported in the past [17], but the use of neural network methodologies is very rare [18].

Furthermore, in other areas of fetal medicine prediction model methodologies have been equally necessary and clinically meaningful [19]-[20].

Based on the results obtained, it may be concluded that the neural structure has been shown to be an effective and reliable predictor for this set of data. Indeed, it has identified all five preeclampsia cases in the totally unknown verification set. It also identified correctly two more case as non-preeclamptic. However, it predicted preeclampsia to two unknown subjects, while these women did not manifest this. This means that the doctors should probably look more carefully to such cases, which in any case were on the safe side.

The association between Afro-Caribbean race and obesity with increased risk of pre-eclampsia is well documented [17]; [21], a finding not supported by our study. This could be due to a small number of such cases in the training data set. Thus, the network needs to be tested further on a completely new and more expanded database of preeclampsia cases, involving a larger sample of Afro-Caribbean cases.

The choice of 11+0 to 13+6 weeks as the gestational age for screening was made because this has been established as the first hospital visit of pregnant women at which combined sonographic and biochemical testing for chromosomal and other major defects is carried out [22].

Acknowledgments

The FMF foundation is a UK registered charity (No. 1037116).

We would like to kindly acknowledge Dr Leona C. Poon for her contribution to the initial organization of the parameters from the original clinical database.

References

[1] World Health Organization, (2005), Make Every Mother and Child Count, **World Health Report**, Geneva, Switzerland.

[2] Lewis G.(ed), (2004), Why Mothers Die 2000–2002: The Sixth Report of Confidential Enquiries Into Maternal Deaths in the United Kingdom. London, United Kingdom, RCOG Press.

[3] Drife J., Magowan B. (eds), (2004), **Clinical Obstetrics and Gynaecology**, Ch. 39, pp 367-370, Saunders.

[4] Douglas K, Redman C., (1994), Eclampsia in the United Kingdom. **Br Med J**, 309 (6966):1395-400.

[5] Moffett A., Hiby S., (2007), How does the maternal immune system contribute to the development of pre-eclampsia?, **Placenta**.

[6] Villar J., Abdel-Aleem H., Merialdi M., Mathai M., Ali M., Zavaleta N., Purwar M., Hofmeyr J., Nguyen T., Campo´donico L., Landoulsi S., Carroli G., Lindheimer M.(2006), World Health Organization Calcium Supplementation for the Prevention of Preeclampsia Trial Group, **Am J Obstet Gynecol**.,194:639–649.

[7] Rumbold A., Crowther C., Haslam R., Dekker G., Robinson J., (2006), ACTS Study Group. Vitamins C and E and the risks of preeclampsia and perinatal complications. **N Engl J Med**. 354:1796 –1806.

[8] Yu C., Smith G., Papageorghiou A., Cacho A., Nicolaides K., (2005), An integrated model for the prediction of pre-eclampsia using maternal factors and uterine artery Doppler velocimetry in unselected low-risk women. **Am J Obstet Gynecol**,193:429–436.

[9] Brause R., (2001), Medical Analysis and Diagnosis by Neural networks, Computer Science Department, Frankfurt a. M., Germany.

[10] Temurtas F., (2009), A comparative study on thyroid disease diagnosis using neural networks, **Expert Systems with Applications: An International Journal archive**, vol.36:1.

[11] Tourassi G., Floyd C., Lo J., (1999), A constraint satisfaction neural network for medical diagnosis, **Neural Networks**, vol.5.

[12] Computer-based neural network system and method for medical diagnosis and interpretation. **US Patent 5839438**.

[13] Local diagnostic and remote learning neural networks for medical diagnosis. **WIPO**, WO/2001/026026.

[14] National Heart Foundation of Australia, (2004), Hypertension management guide for doctors. At: http://www.heartfoundation.org.au. Accessed April 1, 2006.

[15] Khaw A., Kametas N., Turan O., Bamfo J., Nicolaides K., (2008). Maternal cardiac function and uterine artery Doppler at 11–14 weeks

in the prediction of pre-eclampsia in nulliparous women, **BJOG** 115:369–376.

[16] Davey D., MacGillivray I., (1988), The classification and definition of the hypertensive disorders of pregnancy. **Am J Obstet Gynecol**, 158: 892–898.

[17] Eskenazi B., Fenster L., Sidney S., (1991), A multivariate analysis of risk factors for preeclampsia. **JAMA**, 266:237–241.

[18] Mello G., Parretti E., Ognibene A., Mecacci F., Cioni R., Scarselli G., Messeri G., (2001), Prediction of the development of pregnancy-induced hypertensive disorders in high-risk pregnant women by artificial neural networks, **Clinical Chemistry and Laboratory Medicine**, 39(9):A25-A86.

[19] Heath V., Southall T., Souka A., Novakov A., Nicolaides K., (1998), Cervical length at 23 weeks of gestation: relation to demographic characteristics and previous obstetric history, **Ultrasound Obstet Gynecol**, 12:304–311.

[20] Iams J., Goldenberg R., Meis P., Merger B., Moawad A., Das A., Thom E., McNellis D., Copper R., Johnson F., Roberts J., (1996), The length of the cervix and the risk of spontaneous premature delivery, **The New England J of Med**, 334:567-72.

[21] Duckitt K., Harrington D., (2005), Risk factors for pre-eclampsia at antenatal booking: Systematic review of controlled studies, **Br Med J**, 330: 565–572.

[22] Nicolaides K., (2004), Nuchal translucency and other first-trimester sonographic markers of chromosomal abnormalities, **Am J Obstet Gynecol**, 191: 45–67.