ARTIFICIAL NEURAL NETWORKS TO INVESTIGATE THE SIGNIFICANCE OF PAPP-A AND B-HCG FOR THE PREDICTION OF CHROMOSOMAL ABNORMALITIES

Costas Neocleous
Department of Mechanical Engineering
Cyprus University of Technology, Lemesos, Cyprus
costas.neocleous@cut.ac.cy

Andreas C. Neocleous
Department of Computer Science
University of Cyprus, Nicosia, CYPRUS
andrrrrr@hotmail.com

Kypros H. Nicolaides
Harris Birthright Research Centre for Fetal Medicine
King's College Hospital Medical School, London, United Kingdom
fmf@fetalmedicine.com

Christos N. Schizas
Department of Computer Science
University of Cyprus, Nicosia, CYPRUS
schizas@ucy.ac.cy

Kleanthis C. Neokleous
Department of Computer Science
University of Cyprus, Nicosia, CYPRUS
kleneokl@cs.ucy.ac.cy

Abstract

A systematic approach has been done, **to investigate different neural network structures for the appraisal of the significance** of the **free b-human chorionic gonadotrophin (b-hCG)** and the **pregnancy associated plasma protein-A (PAPP-A)** as important parameters for the prediction of the existence of chromosomal abnormalities in fetuses.

The **database** that has been used was **highly unbalanced**. It was composed of 35,687 cases of pregnant women. In the vast majority of cases (35,058) there had not been any chromosomal abnormalities, while in the remaining 629 (1.76%) some kind of chromosomal defect had been confirmed. 8,181 cases were kept as **a totally unknown database that was used only for the verification of the predictability of each network, and for evaluating the importance of PAPP-A and b-hCG as significant predicting factors.** In this unknown data set, there were 76 cases of chromosomal defects.

The system was trained by using 8 input parameters that were considered to be the most influential at characterizing the risk of occurrence of these types of chromosomal anomalies. Then, the PAPP-A and the b-hCG were removed from the in-puts in order to ascertain their contributory effects.

The best results were obtained when using a multilayer neural structure having an input, an output and two hidden layers. **It was found that both of PAPP-A and b-hCG are needed in order to achieve high correct classifications and high sensitivity of 88.2% in the totally unknown verification data set.** When both the b-hCG and PAPP-A were excluded from the training, the diagnostic yield dropped down to 65%.

INTRODUCTION

An effective first-trimester screening for fetal chromosomal abnormalities may be achieved by exploiting **a combination of certain maternal as well as fetoplacental parameters**.

It is well known that the risk for aneuploidies increases with maternal age, it is higher in women that had previously affected pregnancies, and it increases with the fetal nuchal translucency (NT) thickness. It is also higher in the cases where there is an absence of the fetal nasal bone and also when there is abnormal flow through the ductus venosus and across the tricuspid valve. Furthermore, it is also related to the maternal serum concentration of the placental products free ß-human chorionic gonadotrophin (b-hCG) and the pregnancy associated plasma protein-A (PAPP-A) (Nicolaides, 2004a).

The traditional approach to screening for each aneuploidy is to use some form of statistical techniques to estimate the patient-specific risk for each aneuploidy. These are usually implemented in suitable proprietary packages.

In our study, **artificial neural networks (ANN) had been constructed, trained and verified with a large unknown data set.** The diagnostic results for each network were compared with the rest, so that ultimately a good diagnostic tool will be obtained.

A large number of different neural network structures have been constructed and trained to the large data base of pregnant women, aiming at producing a neural classifier/predictor for the risk of presence of chromosomal abnormalities in fetuses. The neural structures that were attempted were mainly of the feedforward type, both of standard multi-layer, as well as of multi-slab topologies.

DATA

The database that was used had **35,687 cases of pregnant women**. The vast majority of these cases (35,058) were normal as far as the chromosomal abnormalities are concerned. The remaining 629 cases (1.76%) were confirmed as having some kind of chromosomal defect (Trisomy 21 (Down syn-drome), Trisomy 18 (Edwards syndrome), Trisomy 13 (Patau syndrome), Turner syndrome, Triploidy). That is, there is a prevalence of 1.76%.

This is **a highly unbalanced data set**, that makes it difficult to exercise exhaustive validation techniques other than the “split-sample method” (also known as “hold-out method”).

The data were provided by the **Fetal Medicine Foundation** of London. They were obtained from the greater London area and South-East England for pregnant women attending routine clinical and ultrasound assessment for the risk of chromosomal abnormalities.

For the present study, for each pregnant woman, a number of relevant parameters were

collected, encoded/converted into appropriate numbers, and suitably used for the training of the neural networks. These parameters are shown in Table 1.

TABLE I

INPUT PARAMETERS

1	**MA**, Maternal age
2	Information on previous occurrence of trisomy 21
3	Information on previous occurrence of trisomy 18
4	Information on previous occurrence of trisomy 13
5	**CRL**, **C**rown **R**ump **L**ength (mm)
6	**NT**, **N**uchal **T**ranslucency (mm)
7	Serum marker **PAPP-A**
8	**β-hCG**, Human Chorionic Gonadotropin

Out of the total of 35,687 cases, 629 cases (1.76%) were confirmed as having some form of chromosomal anomaly of T21, T18, T13, Triploidy or of the Turner Syndrome. The remaining 35,058 cases (98.2%) did not show any chromosomal abnormality. Thus, the data set is highly unbalanced as far as the distribution of the various classes. This makes the potential for building an effective neural network predictor to be a difficult task.

A subset of 8,181 cases (23%) were isolated and kept aside to be used as a totally unknown database in order to check the predictability of each attempted neural network, and later for the evaluation of the importance of PAPP-A and b-hCG as important parameters that contribute to an accurate prediction of the risk of occurrence of the genetic abnormality of interest. In this unknown data set, there were 76 cases (0.93%) of chromosomal defects. Namely, there were 20 cases of Down syndrome (0.06%), 28 cases of the Edwards syndrome (0.08%), 10 of Patau syndrome (0.03%), 5 of Triploidy (0.01%), and the remaining 13 of the Turner syndrome (0.04%).

It is emphasized that these cases were never used during the learning procedures of training of the neural networks, and thus, they were a reliable way for ascertaining the predictability of each network. Because the number of anomalous cases is very small, and such cases cannot be artificially generated, the anomalous cases in the verification set were confined to only a small, but substantial, percentage.

THE NEURAL PREDICTOR

A number of feedforward neural structures of standard multilayer type, having different number of layers and activations, as well as different neurons per layer were systematically built, trained and tested. Also, multi-slab topologies of different structures, sizes, and activation functions, were systematically built, trained and verified, in order to find the best performing structure to be used for the prediction of the totally unknown verification data set. This was done in a planned and systematic manner so that the best performing architecture would be obtained and finally used.

Table 2 shows a summary of the attempted structures and of the performances achieved for the totally unknown verification data set of the 8,181 cases.

TABLE 2

ATTEMPTED NETWORK TOPOLOGIES AND SUMMARY OF RESULTS

		CLASSIFICATIONS (%)				
INPUTS	**TOPOLOGY**	**ALL**	**NORMAL**	**ABNORMAL**	**SENSITIVITY**	**SPECIFICITY**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-20log-1log	**98.6**	**98.8**	**82.9**	**82.9**	**98.8**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-30log-1log	**98.5**	**98.7**	**84.2**	**84.2**	**98.7**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-40log-1log	**98.7**	**98.9**	**77.6**	**77.6**	**98.9**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-50log-1log	**99.0**	**99.2**	**78.9**	**78.9**	**99.2**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-20tanh-1log	**99.1**	**99.3**	**77.6**	**77.6**	**99.3**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-30tanh-1log	**99.1**	**99.3**	**77.6**	**77.6**	**99.3**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-40tanh-1log	**99.1**	**99.3**	**80.3**	**80.3**	**99.3**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-50tanh-1log	**98.9**	**99.0**	**85.5**	**85.5**	**99.0**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-15log--15log-1log	**97.8**	**97.9**	**88.2**	**88.2**	**97.8**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-25log--25log-1log	**98.6**	**98.8**	**78.9**	**78.9**	**98.8**
8:MA,Prev,CRL,NT,PAPP-A,b_hCG	FF: 8lin-50log--50log-1log	**98.4**	**98.5**	**81.6**	**81.6**	**98.5**
7:MA,Prev,CRL,NT,b_hCG	FF: 7lin-15log--15log-1log	**98.7**	**99.0**	**65.8**	**65.8**	**99.0**
7:MA,Prev,CRL,NT,b_hCG	FF: 7lin-50tanh-1log	**98.1**	**98.4**	**65.8**		
7:MA,Prev,CRL,NT,b_hCG	FF: 7lin-30log-1log	**98.1**	**98.4**	**65.8**		
7:MA,Prev,CRL,NT,PAPP-A	FF: 7lin-15log--15log-1log	**98.8**	**99.0**	**75.0**		
7:MA,Prev,CRL,NT,PAPP-A	FF: 7lin-50tanh-1log	**98.6**	**98.8**	**81.6**		
7:MA,Prev,CRL,NT,PAPP-A	FF: 7lin-30log-1log	**98.5**	**98.7**	**84.2**		
6:MA,Prev,CRL,NT	FF: 6lin-15log--15log-1log	**98.7**	**99.0**	**63.2**		
6:MA,Prev,CRL,NT	FF: 6lin-50tanh-1log	**98.5**	**98.8**	**67.1**		
6:MA,Prev,CRL,NT	FF: 6lin-30log-1log	**98.5**	**98.9**	**63.2**		

The various abbreviations are explained by example cases as follows:

The input parameter notation **8:MA,Prev,CRL,NT,PAPP-A,b_hCG** in the network input means that there are 8 input parameters. These are the mother's age, the existence of previous trisomies (T13, T18, T21), the crown rump length, the nuchal translucency, and the biomarkers PAPP-A and b_hCG.

The topology notation **FF:8lin-15log-20tanh-1log** means that the network is feedforward (FF) of four layers. The first layer has 8 neuronal units of linear activation, the second layer has 15 units of logistic sigmoid activation, the third layer has 20 units of hyperbolic tangent activation, and the output layer is of one unit having logistic sigmoid activation.

Ultimately, the best performing feedforward multilayer neural structure had four layers. The particulars of this neural structure are depicted in Figure 1.

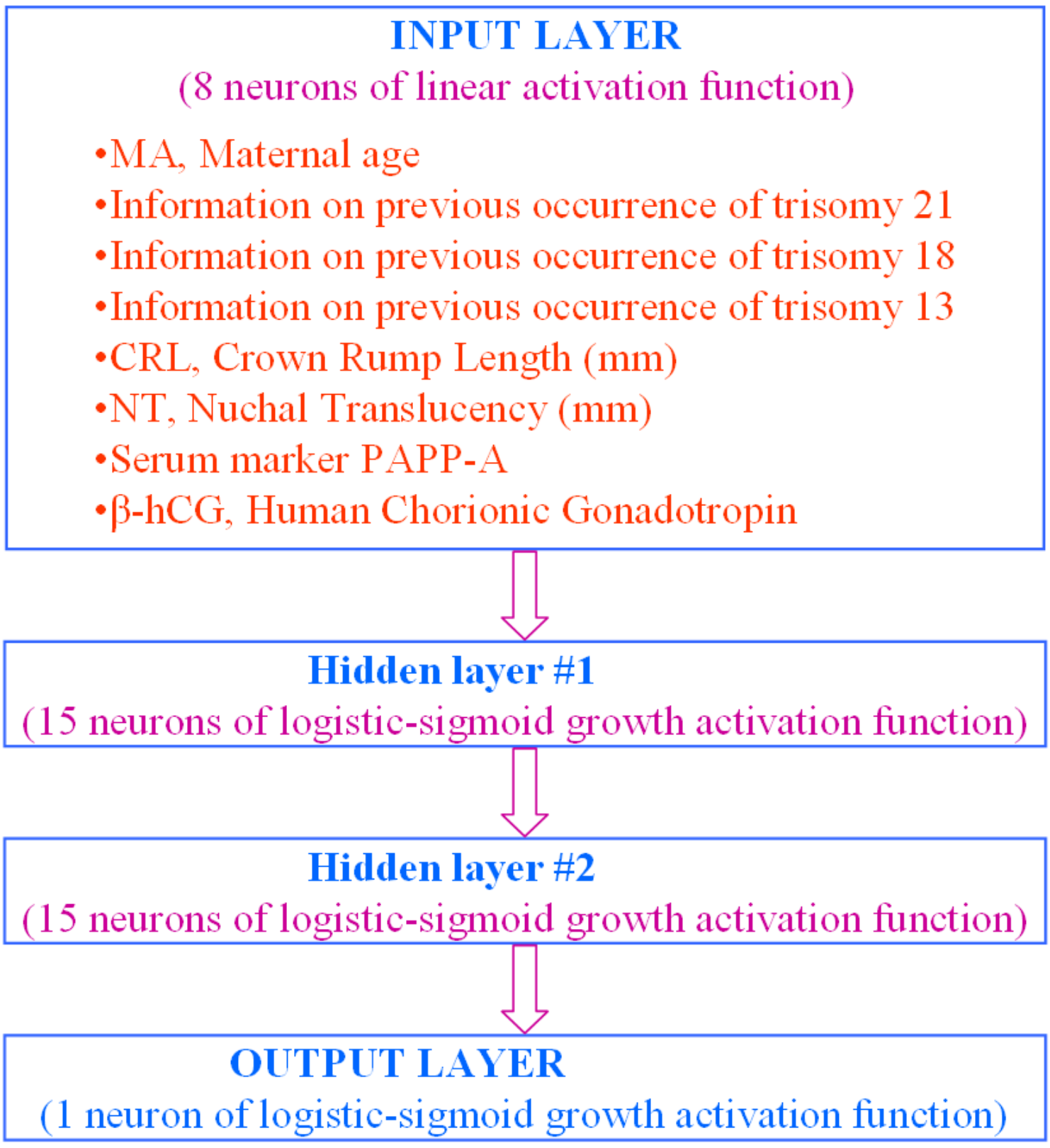


Figure 1. The network structure

Based on extensive previous experience of the authors, all the weights were initialized to 0.1. The learning rate was the same for all connections, having a value of 0.1. Similarly, the momentum rate was 0.4, and all these settings were applied to all the links. The learning scheme that was used was the standard backpropagation with momentum.

A guidance test set composed of very few representative test cases (31) was applied to each attempted network at the end of each epoch. This was done in order to test the progress of training, and thus to keep the best performing weight distribution. If the results of the testing at the end of a particular epoch were better than those at the previous epoch,

the weights were saved as a better performing network set. Thus, at the end of each training procedure, for each network topology, the weight distribution found was the one that resulted in the best performance in this guidance test set.

RESULTS AND CONCLUSIONS

A summary of the results of the performance of each attempted neural network, for the totally unknown verification set, is depicted in Table 2, where the sensitivity and specificity are also presented. These are used in their usual definitions shown below:

$$\text{SENSITIVITY} \equiv \frac{\text{TRUE POSITIVE}}{\text{TRUE POSITIVE + FALSE NEGATIVE}}$$

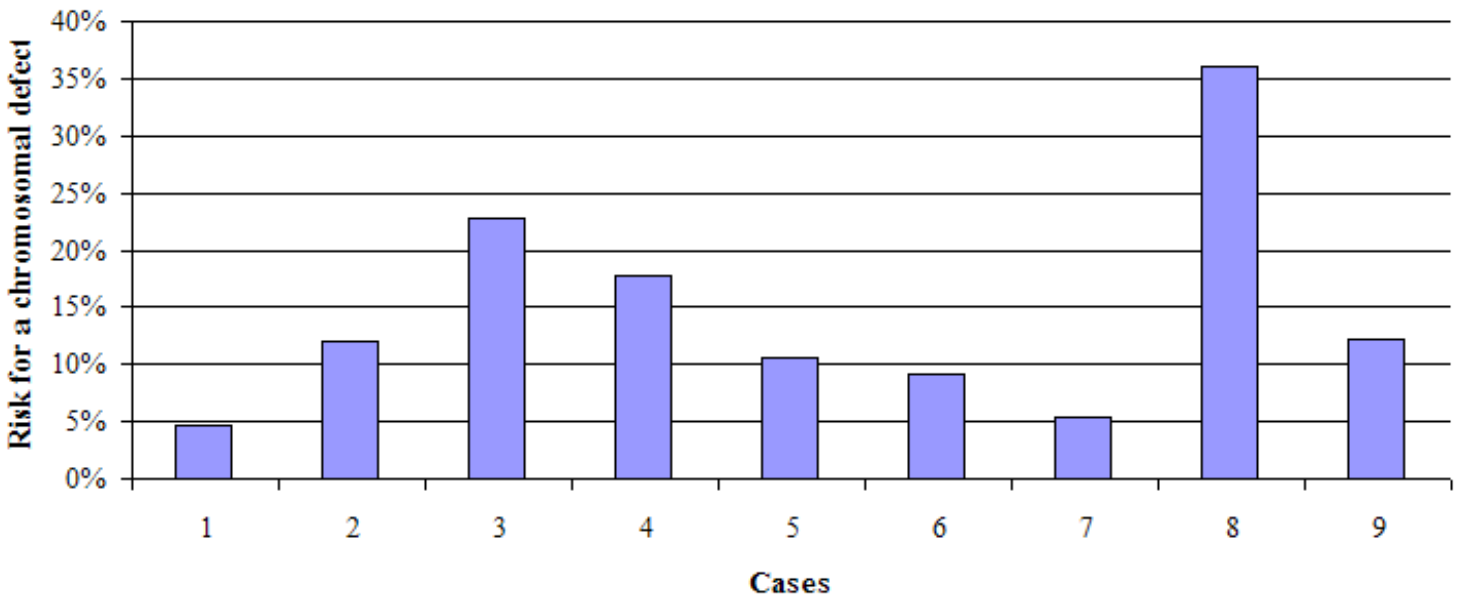
$$\text{SPECIFICITY} \equiv \frac{\text{TRUE NEGATIVE}}{\text{FALSE POSITIVE + TRUE NEGATIVE}}$$

where,

Classification	Known presence of a chromosomal anomaly in the patient	Network output prediction on the risk of chromosomal anomaly in the patient
True Positive	YES	YES
False Positive	NO	YES
False Negative	YES	NO
True Negative	NO	NO

The threshold that was used in order to consider the network output as true was 50%.

TABLE 3
FALSE NEGATIVE RESULTS.
NEURAL ESTIMATIONS FOR A CHROMOSOMAL DEFECT



The results of the best neural network, for the 9 cases (out of 76 abnormal cases in the verification set) that gave FALSE NEGATIVE are shown in Table 3.

These are considered FALSE if the output neuron activation, which is an indication of the risk of a chromosomal defect, was less than 50%. The results are for a standard feedforward neural network of 8 inputs, 15 neurons in the first hidden layer and 15 neurons in a second hidden layer. In all the layers but the input one, the activations were of the logistic-sigmoid growth function.

Similarly, in Figure 2 the distribution of the risk predictions for the 168 cases of FALSE POSITIVE are shown.

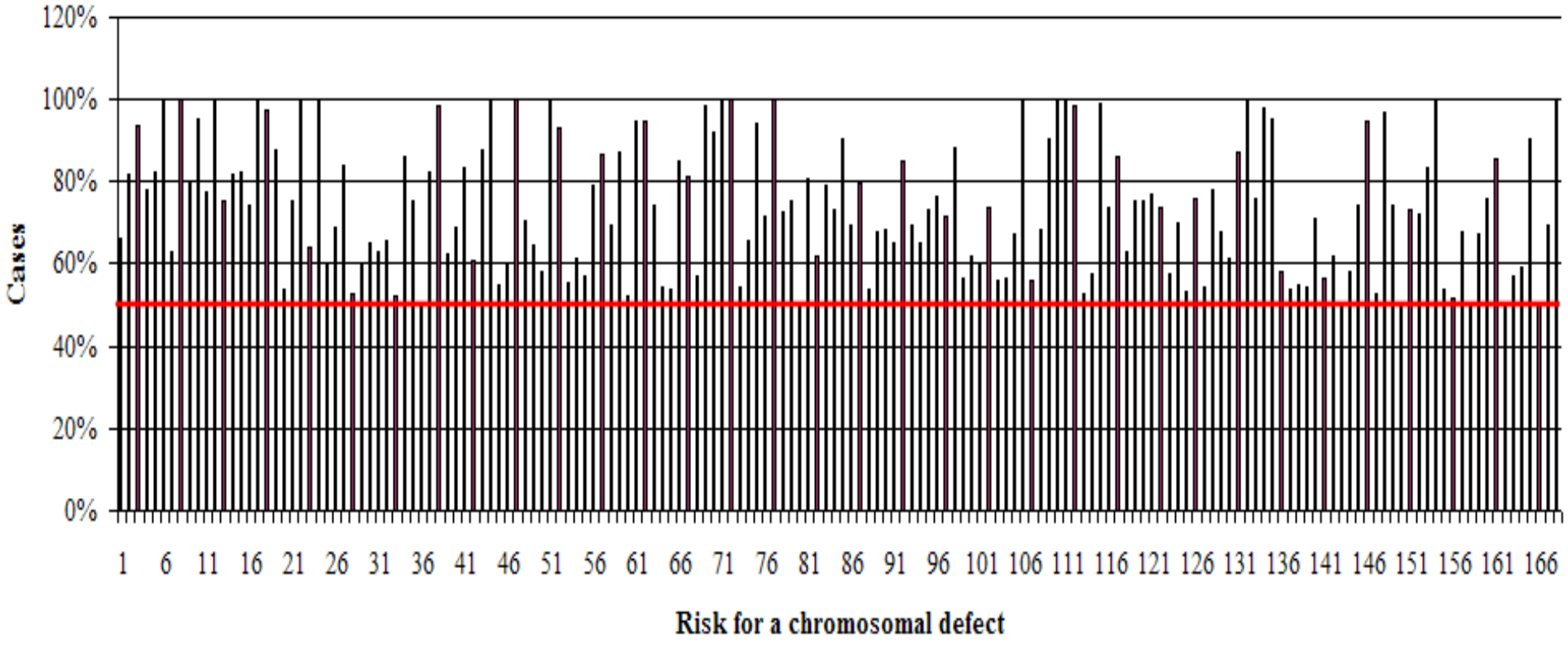


Figure 2. Distribution of the network outputs for the 168 cases of FALSE POSITIVE.

When the PAPP-A was excluded from the input layer, and thus not used for training, the best diagnostic yield for the abnormal cases dropped down to 65.8% for all the neural topologies that were tried, as is shown in Table 2.

Similarly, when the b-hCG was excluded from the input layer, the best diagnostic yield for the abnormal cases dropped down to 84.2% for a three-layer neural topology of 7 inputs and 30 neurons in the hidden layer (Table 2). All but the input layer had logistic-sigmoid growth activation functions.

When both PAPP-A and the b-hCG were excluded, the best diagnostic yield for the abnormal cases was 67.1% (Table 2).

From Figure 2, it is noted that in 43 cases out of the 168 that have been classified as False Positive, the risk for chromosomal anomalies is in the range from 50% to 60%. That is they are quite close to predicting them well if the 50% threshold is used.

From the previous comments it is noted that the pregnancy associated plasma protein-A is a highly important diagnostic factor, which is necessary for proper chromosomal anomaly diagnosis through the use of artificial neural networks.

The results shown in Table 3 and Figure 2 are very encouraging, because they give a high screening/diagnostic yield for the totally unknown data set.

Currently, experiments are being done in order to improve the diagnostic yield, by using smaller number of input parameters. Also, to identify the problematic cases, and to re-examine them together with the specialist medical doctors, aiming at making better training and verification data samples.

ACKNOWLEDGMENTS

The Cyprus Research Promotion Foundation.
The FMF foundation.
The Cyprus University of Technology.
The University of Cyprus.

REFERENCES

[1] Nicolaides K. (2004a). Nuchal translucency and other first-trimester sonographic markers of chromosomal abnormalities. *Am. J. Obstet. Gynecol*. 1191:45-67.

[2] Nicolaides K. (2004b). The 11-13+6 weeks scan. *Fetal Medicine Foundation*, London.

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

[4] Brause R. (2001). *Medical Analysis and Diagnosis by Neural networks.* Computer Science Department, Frankfurt am Main, Germany.

[5] Patel J., Goyal R. (2007). Applications of artificial neural networks in medical science. *Curr. Clin. Pharmacol*. 2(3):217-26.

[6] Neocleous C., Anastasopoulos P., Nicolaides K., Schizas C., Neokleous K, (2009). Neural networks to estimate the risk for preeclampsia occurrence. *Proceedings of the International Joint Conference on Neural Networks*, Atlanta, USA.

[7] Neocleous C., Nikolaides K., Neokleous K., Schizas C. (2010). Artificial neural networks for non-invasive chromosomal abnormality screening of fetuses. *Proceedings of the WCCI 2010*. Barcelona, Spain.