# Artificial neural networks for non-invasive chromosomal abnormality screening of fetuses

C.K. Neocleous, K.H. Nicolaides, K.C. Neokleous and C.N. Schizas

***Abstract*—A large number of different neural network structures have been constructed, trained and tested to a large data base of pregnant women characteristics, aiming at generating a classifier-predictor for the presence of chromosomal abnormalities in fetuses, namely the Trisomy 21 (Down syndrome), Trisomy 18 (Edwards syndrome), Trisomy 13 (Patau syndrome) and the Turner syndrome.**

**The database was composed of 31611 cases of pregnant women. 31135 women did not show any chromosomal abnormalities, while the remaining 476 were confirmed as having a chromosomal anomaly of T21, T18, T13, or Turner Syndrome.**

**From the total of 31611 cases, 8191 were kept as a totally unknown database that was only used for the verification of the predictability of the network. In this set, 7 were of the Turner syndrome, 14 of the Patau syndrome, 42 of the Edwards syndrome and 71 of the Down syndrome.**

**For each subject, 10 parameters were considered to be the most influential at characterizing the risk of occurrence of these types of chromosomal anomalies.**

**The best results were obtained when using a multi-layer neural structure having an input, an output and three hidden layers. For the case of the totally unknown verification set of the 8191 cases, 98.1% were correctly identified. The percentage of abnormal cases correctly predicted was 85.1%. The unknown T21 cases were predicted by 78.9%, the T18 by 76.2%, the T13 by 0.0% and the Turner syndrome by 42.9%.**

## I. Introduction

Effective first-trimester screening for fetal trisomy 21 and most other major chromosomal abnormalities is achieved by a combination of certain maternal and feto-placental characteristics. The risk for aneuploidies increases with maternal age, is higher in women with previous affected pregnancies, increases with fetal nuchal translucency (NT) thickness. and is higher in those with absence of the fetal nasal bone and abnormal flow through the ductus venosus and across the tricuspid valve. It is also related to the maternal serum concentration of the placental products free ß-human chorionic gonadotrophin (ß-hCG) and pregnancy associated plasma protein-A (PAPP-A). Serum PAPP-A is decreased in trisomies 21, 18 and 13 and Turner syndrome but serum free b-hCG is increased in trisomy 21, decreased in trisomies 18 and 13 and not altered in Turner's syndrome. In addition, trisomy 13 and Turner's syndrome, unlike trisomies 21 and 18, are associated with fetal tachycardia. There is also an association between aneuploidies and certain fetal anatomical abnormalities detectable at the first-trimester scan, including holoprosencephaly, cardiac defects, diaphragmatic hernia and exomphalos.

The traditional approach to screening for each aneuploidy is firstly, to estimate the patient-specific risk for each aneuploidy by multiplying the a-priori risk with the likelihood ratio for each sonographic and biochemical marker and secondly, to calculate the detection and false positive rates by taking the proportions with risks above a given risk threshold. The a-priori risk is the maternal age-related risk for each aneuploidy and this is increased by about 75% in those cases with a previous history of aneuploidy. The likelihood ratios for categorical variables, such as fetal abnormalities, absent nasal bone and abnormal tricuspid or ductus venosus flow, are the ratios of the prevalence of each marker in each type of fetal aneuploidy to the prevalence in euploid fetuses. In the case of continuous variables, such as fetal NT thickness, fetal heart rate, maternal serum free ß-hCG and PAPP-A, the likelihood ratios are derived from the overlap of the Gaussian distribution of each marker in each type of fetal aneuploidy with the respective Gaussian distribution in euploid fetuses.

The Fetal Medicine Foundation (FMF), which is a UK registered charity, has established a process of training and quality assurance for the appropriate introduction of NT screening into clinical practice.

In the present study, an alternative to the traditional statistical approaches for chromosomal abnormality risk estimation is used. For such estimation, the paradigm of artificial neural networks (ANN) has been used.

The aim was to develop and apply appropriate ANN structures for the early identification of chromosomally abnormal fetuses. The data were on maternal and feto-placental characteristics, obtained at the time of traditional screening at 11-13 weeks.

Indeed, a large number of different neural network structures have been constructed and trained to a large data base of pregnant women, aiming at generating a neural classifier-predictor for the presence of chromosomal abnormalities in fetuses. The abnormalities of interest were the Trisomy 21 (Down syndrome), Trisomy 18 (Edwards syndrome), Trisomy 13 (Patau syndrome) and the Turner syndrome.

C.K. Neocleous (corresponding author) is with the Department of Mechanical Engineering, Cyprus University of Technology, Lemesos, CYPRUS. Tel: +357 25002580; Email: costas.neocleous@cut.ac.cy

K.C. Neokleous and C.N. Schizas are with the Department of Computer Science, University of Cyprus, 75 Kallipoleos, 1678, POBox 20537, Nicosia, CYPRUS.

K.H. Nicolaides is with the Harris Birthright Research Centre for Fetal Medicine, King's College Hospital Medical School, Denmark Hill, SE5 8RX, London, United Kingdom

The neural structures that were attempted were mainly of the feedforward class, both of standard multi-layer type, as well as multi-slab topologies.

Once acceptably reliable results are obtained through the examination of a large number of ANN structures and neuron parameters, an automatic and user friendly tool can be built. Such a tool may improve the detection of chromosomal defects, and be a user-friendly and reliable predictor or method for the effective and early identification of an abnormality. The tool would be of great help to obstetricians and of course to pregnant women and the unborn children.

In recent years, ANNs and other computationally intelligent techniques have been used as medical diagnosis tools, aiming at achieving effective medical decisions that could be incorporated in appropriate medical support systems, for diagnosis, in signal processing/analysis, and in medical image analysis and radiology [1]-[3]. Neural networks in particular have proved to be quite effective and also have resulted in some relevant patents [4]-[8].

## II. Data and Methods

The database that was used for the training and evaluation of the neural network was made of 31611 cases of pregnant women. These were provided by the Fetal Medicine Foundation (FMF) of London, and were obtained from the greater London area and South-East England from pregnant women attending routine clinical and ultrasound assessment for the risk of chromosomal abnormalities.

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

The transabdominal ultrasound examination was performed in order to diagnose any major fetal defects and also for the measurement of the fetal crown–rump length (CRL), the NT thickness, and the fetus heart rate (FHR). Examination of the nasal bone and Doppler assessment of flow in the ductus venosus and across the tricuspid valve were routinely performed by sonographers who had received the appropriate FMF Certificates of Competence. The fetal nasal bone was classified as present or absent, the a-wave of the flow velocity waveform in the ductus venosus was classified as reversed or normal, and the flow velocity waveform across the tricuspid valve was classified as regurgitant or normal. The pregnancy was dated according to the last menstrual period, but if the dates were uncertain or the estimated gestation by CRL was discordant by more than 7 days from the estimated gestation dates, the CRL was used to date the pregnancy. During the examination, the pulsed-wave Doppler was routinely used to obtain 6–10 cardiac cycles during fetal quiescence and the FHR was calculated by the ultrasound machine software. Automated machines that provide reproducible results within 30 min were used to measure PAPP-A and free β-hCG (Delfia express system, Perkin Elmer, Waltham, MA, USA). The measured PAPP-A and free b-hCG were converted into multiples of the median (MoM) for gestational age adjusted for maternal weight, ethnicity, smoking status, method of conception, parity and machine for the assays.

Also, maternal demographic characteristics, ultrasonographic measurements and biochemical results were appropriately recorded. Karyotype results and details on pregnancy outcomes were added into the database as soon as they were available.

For each case, a total of 16 parameters were identified and measured or recorded. However, because some data were redundant, and also because there were very few cases of pregnant women exhibiting some of the parameters (mainly on whether there were previous occurrences of the chromosomal anomalies), these were excluded in the study, and ultimately only 10 parameters were considered to be the most influential at characterizing the risk of occurrence of chromosomal anomalies.

The 10 parameters used for the network training are shown in Table 1, where the explanation for each of the characteristics has been given in previous paragraphs.

TABLE I
INPUT PARAMETERS USED IN THE NEURAL NETWORK STRUCTURES

| | |
|---|---|
| 1 | **MA**, Maternal age |
| 2 | **CRL**, **C**rown **R**ump **L**ength (mm) |
| 3 | **GA**, **G**estation **A**ge when the CRL was measured (days) |
| 4 | **NT**, **N**uchal **T**ranslucency (mm) |
| 5 | **FHR**, **F**etus **H**eart **R**ate |
| 6 | **NB**, **N**asal **B**one (normal, abnormal) |
| 7 | **TF**, **T**ricuspid **F**low from RA to RV (normal, abnormal) |
| 8 | **DV**, **D**uctus **V**enosus flow (normal, abnormal) |
| 9 | Serum marker **PAPP-A** |
| 10 | **β-hCG**, Human chorionic gonadotropin |

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

From the total of 31611 records, 31135 (98.5%) women did not show any chromosal abnormalities - something which is fortunate, but as far as the potential for building an effective neural network predictor rather undesirable. The remaining 476 cases (1.5%) were confirmed as having a chromosomal anomaly of T21, T18, T13 or Turner Syndrome. This is a highly unbalance data set, that creates difficulties when attempting to develop an effective predictor.

A guidance set of 109 cases was extracted and used to test the progress of training at the end of each epoch. Thus it was not directly being used for the training of the network. It was rather a testing paradigm that was mainly used for helping the system to always keep (save) the best performing synaptic weight distribution. This guidance data set included 40 cases (45%) of fetuses that were chromosomally abnor-

mal. Also, out of the total of 31611 cases, a data set having 8191 (25.9%) cases of pregnant women was kept as a totally unknown database that was used for the verification of the predictability of each of the attempted neural network. Thus these cases were never given to the network during the training period. They were totally unknown to the best selected neural structure, and thus they were a reliable way for ascertaining the predictability of the network. In this set, 7 (0.08%) cases were of the Turner syndrome, 14 (0.17%) of the Patau syndrome, 42 (0.51%) of the Edwards syndrome and 71 (0.85%) of the Down syndrome.

Thus, the data set is highly unbalanced, and consequently very difficult to study. The nature of this problem makes the use of artificial neural networks more interesting since their ability to handle such complex problems can be demonstrated.

## III. The Neural Classifier/Predictor

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

After numerous trials, it was ultimately possible to select and use a feedforward multilayer neural structure having five layers. The particulars of the neural structure are depicted in Figure 1.

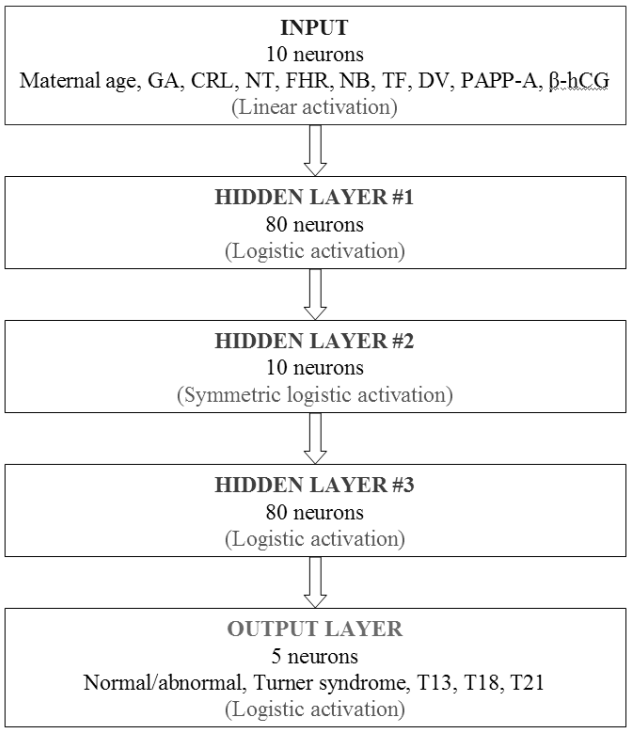


Fig. 1. The neural structure that was ultimately selected and used for the first trimester diagnosis of chromosomal defects

Based on extensive previous experience, all the weights were initialized to 0.3.

The learning scheme used was the standard backpropagation with momentum. The learning rate was the same for all connections, having a value of 0.1. Similarly, the momentum rate was 0.2, applied to all links.

The guidance test set 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 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 and kept as a better performing set. Thus, at the end, the weight distribution found was the one that resulted in the best performance in this guidance test set.

TABLE 2
SUMMARY OF RESULTS FOR ALL THE CASES IN THE UNKNOWN VERIFICATION DATA SET

| **Total number of cases** | **8191** |
|---|---|
| Number of cases that DID NOT HAVE chromosomal anomalies | 8057 |
| Number of cases that HAD chromosomal anomalies | 134 |
| Number of cases that DID NOT HAVE chromosomal anomalies and were correctly predicted (True Negative) | 7966 |
| Number of cases that DID NOT HAVE chromosomal anomalies and were NOT correctly predicted (False Positive) | 91 |
| Percentage of cases that DID NOT HAVE chromosomal anomalies and were correctly predicted | 98.9 |
| Sensitivity (%) | 88.8 |
| Specificity (%) | 97.9 |

The training progress was monitored in order to observe whether there was improvement during the application of the training and test set data. For most of the network structures attempted, there was little generalization improvement after about 1500 epochs.

TABLE 3
SUMMARY OF RESULTS FOR EACH CASE OF CHROMOSOMAL ANOMALY IN THE UNKNOWN VERIFICATION DATA SET

| | **TURNER SYNDROME** | **T13** | **T18** | **T21** |
|---|---|---|---|---|
| **Total number of cases** | **7** | **14** | **42** | **71** |
| Percentage of cases that were correctly predicted | 42.9 | 0.0 | 76.2 | 78.9 |
| Sensitivity (%) | 14.3 | 0.0 | 73.8 | 80.3 |
| Specificity (%) | 100.0 | 100.0 | 99.9 | 98.6 |

Different sets of inputs – among the 10 available parameters – were used to find an effective neural structure that would screen for chromosomal defects to an acceptable level. The inputs that were ultimately selected are those shown in Table 1.

## IV. Results and conclusions

The results on the performance of the neural network for the totally unknown verification set of 8191 cases are summarized in Table 2.

In Tables 2 and 3, the sensitivity and specificity quoted are those that were obtained by the following expressions:

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

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

where,

| **Presence of chromosomal anomaly in patient** | **Network prediction of chromosomal anomaly in patient** | **Classification** |
|---|---|---|
| YES | YES | True Positive |
| NO | YES | False Positive |
| YES | NO | False Negative |
| NO | NO | True Negative |

The results on the performance of the neural network for the totally unknown verification set for each separate class of chromosomal abnormality are summarized in Table 3.

The results shown in Tables 2 and 3 are very encouraging, because they give a high screening/diagnostic yield for the totally unknown data set. This is particularly true for the Down syndrome.

For the other syndromes the diagnostic yield is low, but this is mainly due to the fact that the number of abnormal chromosomal cases of pregnant women in the data sets is small. This of course is fortunate, and there is no way to generate artificial data for such serious genetic anomalies.

Experiments are currently being done to screen separately for each anomaly through dedicated networks. The results so far are encouraging. These will be reported in future work.

## Acknowledgment

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

This research is partly supported by the University of Cyprus International Affairs Committee.

## References

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

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

[3] 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.

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

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

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

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

[8] 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.