Prenatal Diagnosis of Aneuploidy Using Artificial Neural Networks in Relation to Health Economics

A.C. Neocleous[1,2], C.K. Neocleous[3], N. Petkov[2], K.H. Nicolaides[4], and C.N. Schizas[1]

[1] University of Cyprus, Computer science, Nicosia, Cyprus
[2] University of Groningen, Johann Bernoulli Institute for Mathematics and Computer Science, Groningen, The Netherlands
[3] Cyprus University of Technology, Department of Mechanical Engineering and Materials Science and Engineering
[4] Harris Birthright Research Center for Fetal Medicine, King's College Hospital, London, England, UK

Abstract— **The early detection of fetal chromosomal abnormalities such as aneuploidies, has been an important subject in medicine over the last thirty years. A pregnant woman is advised by the doctor to perform an amniocentesis test, after the identification of increased risk for fetal aneuploidy. Even though the amniocentesis test is almost perfectly accurate, it has several drawbacks. It is an invasive test with around 1% risk for miscarriage; it is financially expensive and requires laboratories and special equipment. In this work we propose a non-invasive method for aneuploidy detection using a dataset with pre-natal examinations of pregnant women and artificial neural networks. We have used a dataset with 50,517 euploid and 691 aneuploid cases. Biological markers of the mother such as the age, blood proteins and ultrasonographic information from the fetus are used as input to the networks. A training set is used to construct neural networks and a test set is used for validation. Each unknown case is assigned into a class between "euploid" and "aneuploid" using a cut-off value on the network output. We create a ROC curve by computing the sensitivity and the specificity for a set of different cut-off values. From the ROC curve, we indicate the importance of the cut-off values in terms of health economics and social affection. It is shown that by increasing the cut-off value, the false positive rate reduces with the cost of an increased false negative rate.**

Keywords— **Non-invasive diagnosis, computational intelligence, chromosomal abnormalities, artificial neural networks, ROC curve.**

I. Introduction

In the hospitals of the Fetal Medicine Foundation[1] (FMF), all pregnant women follow a prenatal examination. This is done for estimating the risk of several diseases, pregnancy complications, fetal chromosomal abnormalities, fetal cardiac diseases and others. The standard examination includes a) blood analysis, b) ultrasonographic screening of the fetus and c) other biomarkers such as maternal age, weight, heart rate and history of previous pregnancies.

The doctors use three diagnosis tests for the detection of chromosomal abnormalities and other fetal dysfunctions. The first test that was suggested by Prochownick and his group over than hundred years ago is the amniocentesis test. A sample of about 20 ml of the amniotic fluid is extracted with a needle and the fetal DNA is examined for chromosomal abnormalities and other fetal infections. The amniocentesis is an invasive test with induced risk for miscarriage around 1.0%.

In the last decade, there was an increased interest to explore non-invasive methods for the detection of chromosomal abnormalities. Statistical methods were initially applied to datasets mainly by the group of Nicolaides et.al. [1-2], and Dugoff [3].

In [4], we proposed a non-invasive method for the early detection of chromosomal abnormalities using artificial neural networks. We have trained feed-forward neural networks with nine units in the input layer, one and two hidden layers and one unit in the output layer.

In this paper we use the results of our experiments published in [4], to present the ROC curves for the two classes "euploid" and "aneuploid". These curves are used to interpret the significance of the cut-off value that is used to classify an unknown case into one of the two classes mentioned above. We discuss how different cut-off values

The original version of this chapter was inadvertently published with an incorrect chapter pagination 930–934 and DOI 10.1007/978-3-319-32703-7_181. The page range and the DOI has been re-assigned. The correct page range is 936–940 and the DOI is 10.1007/978-3-319-32703-7_182. The erratum to this chapter is available at DOI: 10.1007/978-3-319-32703-7_260

E. Kyriacou et al. (eds.), *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016*,
IFMBE Proceedings 57,
DOI: 10.1007/978-3-319-32703-7_182

can affect social and economic aspects. The decisions that are made based on the output value of such methods may require additional examinations such as amniocentesis, or abortion. The optimal cut-off value minimizes the number of invasive tests and at the same time, it allows for minimum births of aneuploidy.

The rest of the paper is organized as follow. In Section II we present our methodology including the architecture of the neural network schemes and the ROC curves. In Section III we present our results and we discuss the outcome of this work in Section IV. Finally, we conclude in Section V.

II. Methods

A. Data

Our dataset contains with 51,208 medical records derived from pre-natal examinations, done in the first trimester of pregnancy. The majority of the cases are euploid confirmed (99% of the total dataset). From the total 51,208 cases, 50,517 are euploid, 408 are T21 and 283 are other chromosomal abnormalities.

B. Features used

The following nine markers were used as input to the first neural network layer: CRL, NT, PAPP-A, b-hCG, nasal bone, tricuspid flow, ductus venosus flow, maternal age, history of pregnancies with previous T21.

From the maternal blood, the most informative markers are the b-hCG and PAPP-A. In pregnancies with fetal chromosomal abnormalities, it was shown from previous work [4] that there is an increase of b-hCG and decrease of PAPP-A levels.

From the ultrasonographic screening, the fetal crown rump length (CRL), the nasal bone, the nuchal tranclucency, the ductus venosus and the tricuspit flows are measured. The CRL is a physical distance between the crown and the rump of the fetal. The nasal bone is a binary marker that is set by the doctor as normal if it is present or abnormal if it is absent. The volume of the fluid under the skin at the back of the fetus is called nuchal tranclucency (NT) and it increases with fetal chromosomal abnormalities. The Tricuspid flow is a binary measurement of its normality that is set by the doctor. The ductus venosus flow is a non-binary marker derived directly from the ultrasonographic screening.

C. Neural network schemes

We used the training set to apply a grid search with 36 neural networks. We used 12 different architectures for a three-fold cross validation (Table 1). In this paper we present the ROC curves of the first training set with 33,619 euploid, 279 T21, 106 T18, 38 T13, 22 triploid and 41 Turner. The test set is contains 16,898 euploid, 129 T21, 39 T18, 14 T13, 10 triploid and 13 Turner.

The selected architectures are feed-forward networks with nine units in the input layer, one and two hidden layers and one unit in the output layer. The parameters of the networks are summarized in Table 1. All networks were trained for 800 epochs, a momentum of 0.3 and learning rate of 0.1. The output layer is normalized with a sigmoid function in the range 0 and 1.

A. Binary classification

In a medical diagnostic system where the task is the identification of the presence or absence of a disease, we denote with X and Y the diagnostic test results of the normal and the abnormal populations respectively. Examples of the distributions of X and Y are shown with solid and dashed lines in Fig. 1. A positive classification of an unknown case is considered if the diagnostic test result of the system exceeds the cut-off point c that is shown with vertical line in Fig. 1.

A true positive (TP) classification is considered when the test result of an abnormal case exceeds the value of cut-off point c and classified as positive. Similarly, a true negative (TN) classification is defined when a normal case is classified as

Table 1 Architecture structures of the neural networks built. In the first column we present the network ID and in the following four columns show the number of neurons and the activation functions of the hidden layers used.

Network	Hiden layer 1		Hiden layer 2	
	Neurons #	Activation	Neurons #	Activation
FF 1	30	Logistic	-	-
FF 2	40	Logistic	-	-
FF 3	50	Logistic	-	-
FF 4	30	Tanh	-	-
FF 5	40	Tanh	-	-
FF 6	50	Tanh	-	-
FF 7	10	Logistic	10	Logistic
FF 8	20	Logistic	20	Logistic
FF 9	30	Logistic	30	Logistic
FF 10	10	Tanh	10	Tanh
FF 11	20	Tanh	20	Tanh
FF 12	30	Tanh	30	Tanh

negative. A negative classification of an abnormal case is considered false negative (FN) and it is shadowed with vertical lines in Fig.1. The false positive (FP) classification is when a normal case is classified as positive (shown in Fig. 1 with horizontal lines).

A. Receiver operating characteristic curves

a) Overview:

ROC curves (Fig. 2) give information regarding the quality of the performance of medical diagnostic systems. Such information includes the measurement of the area under the ROC curve and the ability to select for optimal cut-off point. The area under the ROC curve is related to the accuracy of the system and it is used for comparison measurements between several classifiers or diagnostic systems.

ROC curves are usually visualized as a 2D plot of the true positive rate (sensitivity) in y axis and the false positive rate (1 – specificity) in x axis, for a set of different cut-off points c. The sensitivity and the specificity of a medical diagnostic system use the values of the TP, TN, FP and FN and are computed as shown in Equations 1 and 2.

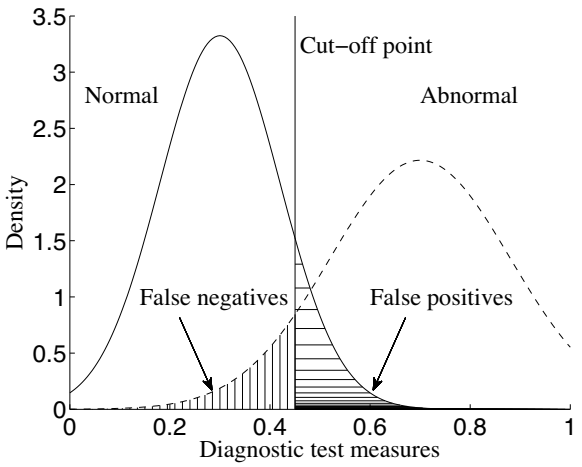

Fig. 1 Distributions of the diagnostic test measures for the normal and the abnormal populations. A cut-off point is used to classify an unknown case into normal or abnormal.

$$Sensitivity = \frac{TN}{TN + FP} \tag{1}$$

$$Specificity = \frac{TN}{TN + FP} \tag{2}$$

Additional important information that yields from the ROC curves is the automatic determination of an optimal *c* point. While increasing the value of *c*, the false positive rate decreases, and the false negative rate increases. In many medical applications, the value of *c* may have a critical importance regarding the cost of the false negative and false positive decisions.

b) Area under ROC curve:

One way to measure the accuracy of a diagnostic system is to calculate the area under the receiver operating characteristic curve (AUC). The AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance.

The methods for estimating the AUC are classified as parametric or non-parametric. A non-parametric method for estimating the AUC is pro-

posed by Bamber [5] and it is equal to the Mann–Whitney U statistic. Another non-parametric method is to fit a smooth ROC curve using kernel smoothing and then estimate the AUC by integration. Parametric methods assume that the output values of both *X* and *Y* follow a normal distribution. Then the AUC is estimated using standard parametric methods with the mean and the standard deviation of the diagnostic test results. The value of the AUC is equal to 1 when perfect classification is achieved.

c) Optimal cut-off point:

A common approach to choose for an optimal cut-off point is to identify the point that reaches the highest sensitivity with the lowest false positive rate (1 – specificity). Another similar method is to minimize the distance *d* between the ROC curve and the point [0, 1], as shown in Fig. 2. In this example, the point [x, y] is the closest to the point [0, 1]. The Youden index maximizes the vertical distance from line of equality (diagonal line) to the ROC curve, as indicated in Fig. 2.

Other approaches proposed in [6], take into account parameters such as cost of false and true positive and negative decisions. A parameter S is calculated as shown in Eq. 3. The point where the line with slope S touches the curve is the optimal cut-off point. The FPc and FNc in Eq. e are the costs of a false positive and a false negative decision. Respectively, the TPc and TNc are the costs of true positive and true negative decisions, if any.

$$S = \left(\frac{FP_c - TN_c}{FN_c - TP_c}\right)\left(\frac{1-P}{P}\right) \qquad (3)$$

Where P denotes the prevalence in the target population [7].

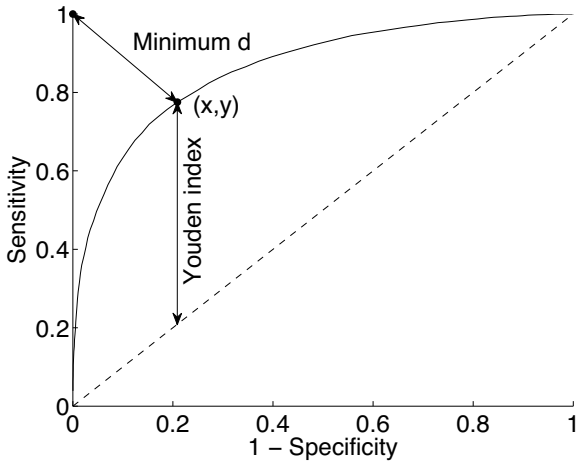

Fig. 2 Example of a ROC curve. Two methods for optimal cut-off value are illustrated. The first method uses the minimum distance between the curve and the point [0, 1]. The Youden index maximizes the vertical distance between the curve and the line of equality (dashed line).

III. Results

In this Section, we present the results of the network FF6 as shown in Table 1, trained with 9 units in the input layer, one hidden layer with 50 units and one unit in the output layer.

In Table 2, we present the distribution of the output of the nine-input network in the range 0 and 1 with step of 0.1, as shown in the first column of Table 2. The number of euploidies and aneuploidies for each step are shown in the second and third columns respectively. The sensitivity and the specificity are shown in the last two columns. In Table 3, we present the results of the same network with an arbitrary cut-off point of 0.5. It is shown that in the range 0–0.5, the system correctly identified as abnormal all 129 cases of T21 with an FPR of 3.9%.

The ROC curve for the two classes "euploid" and "aneuploid" is shown in Fig. 3. The optimal cut-off point using the minimum distance from the point of [0, 1] and the Youden index, is shown with a dot in Fig.3 and has a value of 0.9. The sensitivity at *c*=0.9 is 0.81 and the specificity is 0.92.

IV. Conclusions

In the problem of the non-invasive methods for aneuploidy detections, the cost of a false positive prediction includes several risks for miscarriage and psychological disturbances. If the fetus has trisomy 21, it will be born having all the social and health consequences. If it has one of the other chromosomal abnormalities, the mother will most likely deliver early with the risk of developing other pregnancy complications such as pre-eclampsia or infections.

The cut-off point for classifying a case as normal or abnormal becomes a difficult task to be decided. For medical diagnosis problems, the decisions that are taken after the classification may follow health, social and financial costs.

References

1. Nicolaides H. K, (2011) Screening for Fetal Aneuploidies at 11 to 13 Weeks. Prenat. Diag., vol. 31, No.2, pp. 7-15
2. Nicolaides H. K, Spencer K, Avgidou K, et al (2005) Multicenter Study of First-Trimester Screening for Trisomy 21 in 75821 Pregnancies: Results and Estimation of the Potential Impact of Individual Risk-Orientated Two-Stage First-Trimester Screening. Ultrasound Obstet. Gynecol. vol. 25, No. 3, pp. 221-226
3. Dugoff L, (2010) First-and Second-Trimester Maternal Serum Markers for Aneuploidy and Adverse Obstetric Outcomes. Amer. Coll. of Obst. And Gynecol. vol. 115, No. 5, pp. 1052-61
4. Neocleous C. A, Nicolaides H. K, Schizas N. C, (2015) First Trimester Non-invasive Prenatal Diagnosis: A Computational Intelligence Approach. IEEE Journal of Biomedical and Health Informatics
5. Bamber D, (1975) The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. Journal of Mathematical Psychology vol. 12, pp. 387-415
6. Zweig H. M, Campbell G, (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry vol. 39, pp. 561-577
7. Greiner M, Pfeiffer D, Smith D, (2000) Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Preventive Veterinary Medicine vol. 45 pp. 23-41

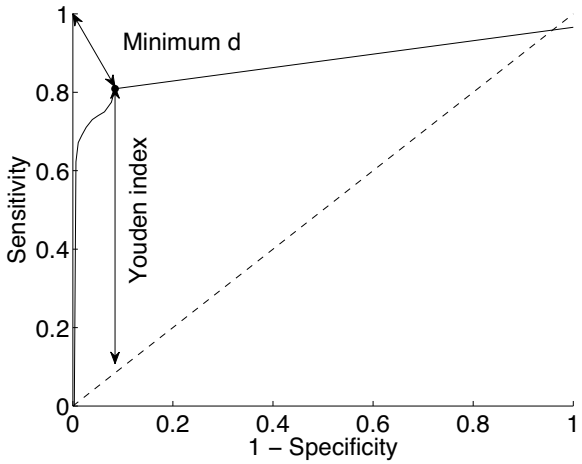

Fig. 3 The ROC curve from the output of the nine-units input network. The minimum distance between the curve and the point [0, 1] is shown with a dot. The same position is found from the Youden index.

Table 2 Distribution of the output of the nine-neuron network.

Output	Euploidy	Aneuloidy	Sensitivity	Specificity
0 to 1	104	127	0	1
> 0.1 to 0.2	80	11	0.62	0.99
> 0.2 to 0.3	129	4	0.67	0.99
> 0.3 to 0.4	138	4	0.69	0.98
> 0.4 to 0.5	203	4	0.71	0.97
> 0.5 to 0.6	192	2	0.73	0.96
> 0.6 to 0.7	217	2	0.74	0.95
> 0.7 to 0.8	231	5	0.75	0.94
> 0.8 to 0.9	123	7	0.77	0.92
> 0.9 to 1	15481	39	0.81	0.92

Table 3 Results of the nine-units input network. The AUC is calculated for euploid and every abnormality separately.

	Predicted	Population	Detection rates	AUC
Euploid	16244	16898	96.13%	
T21 DR	129	129	100.00%	0.99
T18	12	39	30.77%	0.86
T13	3	14	21.43%	0.85
Triploidy	0	10	0.00%	0.47
Turner	6	13	46.15%	0.94