Expert Systems with Applications 37 (2010) 6646–6650

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Prognosis of prostate cancer by artificial neural networks

Ismail Saritas [a,*], Ilker Ali Ozkan [a], Ibrahim Unal Sert [b]

[a] Department of Electronic and Computer Education, Technical Education Faculty, Selcuk University, Konya, Turkey
[b] Department of Urology, Meram Medicine Faculty, Selcuk University, Konya, Turkey

ARTICLE INFO

Keywords:
Artificial neural network
Prostate cancer
Prostate-specific antigen
Prognosis of prostate cancer

ABSTRACT

In this study, an artificial neural network has been devised that yields a prognostic result indicating whether patients have cancer or not using their free prostate-specific antigen, total prostate-specific antigen and age data. Though this system does not diagnose cancer conclusively, it helps the doctor in deciding whether a biopsy is necessary by providing information about whether the patient has prostate cancer or not. Data from 121 patients who were definitively diagnosed with cancer after biopsy were used in devising the system. The results of the definitive diagnoses of the patients and the results of the ANN that was performed were analysed using confusion matrix and ROC analyses. As a result of ANN, which was implemented on the basis of these analyses, success rates of 94.11% and 94.44% were achieved for prognosis of disease and validity, respectively. The ANN, which yielded these high rates of reliability, will help doctors make quick and reliable diagnoses without any risks and make it a better option to monitor patients with low prostate cancer risk on whom biopsies must not be carried out through a policy of wait and see.

1. Introduction

Prostate cancer is one of the most common types of cancer found in men. Risk factors for prostate cancer include age, the family's cancer history and ethnic background. The definitive diagnosis of prostate cancer can be made through a biopsy. An initial diagnosis is made after patients' transrectal ultrasonography, rectal examination results and the amount of prostate-specific antigen (PSA) are assessed by a specialist doctor. The PSA level in blood has become one of the most common methods as a result of the studies conducted in recent years for early diagnosis of prostate cancer (Catalona et al., 1998; Shin Egawa et al., 1997; Van Cangh et al., 1996a).

PSA levels below 4 ng/ml in blood are considered normal, while levels between 4 and 10 ng/ml are considered limit values and levels above 10 ng/ml are high. It is stated that the higher the PSA level is, the higher the prostate cancer risk is (Catalona et al., 1998; Metlin, Lee, & Drago, 1991; Nguyen & Kreinovich, 2001; Seker, Odetayo, Petrovic, & Naguib, 2003; Van Shin Egawa et al., 1997; Van Cangh et al., 1996a,b). However, PSA values may not yield conclusive results about existence of prostate cancer because PSA levels can be increased by inflammation of prostate and benign prostate hyperplasia (BPH). Therefore, patients are also given rectal examination. If anomalies are observed at the end of the rectal examination, even if PSA results may seem normal, it is recommended that a prostate biopsy be performed and definitive diagnosis be made.

In general, prostate cancer is a disease that can be diagnosed with prostate biopsy in accordance with the suspicions that arose as a result of PSA test, rectal examination, and transrectal findings (Metlin et al., 1991; Nguyen & Kreinovich, 2001; Seker et al., 2003).

The studies that were conducted in this regard established that prostate cancer is related to age and the age–PSA relationship obtained in these studies has been given in Table 1 (Brawer & Kirby, 1999; Prostate Cancer Symptoms, 2009).

There is a need for biopsy and definitive diagnosis besides transrectal ultrasonography and rectal examination for diagnosis of prostate cancer. Despite the need for biopsy for conclusive diagnosis, patients with low cancer risk avoid this process due to possible complications that may arise, the risk of rectal mucosa being damaged and its high costs. Therefore, before they agree to biopsy, patients may prefer a different optimum method that may yield a more accurate result. To attain this goal, a prediction method has been developed that determines patients' cancer risk on the basis of their age, ethnic background, history of cancer in the family and PSA levels (Prostate Online Calculator, 2009); besides, various logical models and ANN have been devised since 1998 (Cınar et al., 2009). These models, which use clinical and laboratory data, prevent patients with benign prostate hyperplasia from undergoing unnecessary biopsy operation as well as providing satisfactory information about diagnosis of prostate cancer. For this purpose,

* Corresponding author. Address: Selcuk University, Technical Education Faculty, Campus 42075, Selcuklu-Konya, Turkey. Tel.: +90 3322233354; fax: +90 3322412179.
E-mail addresses: isaritas@selcuk.edu.tr, ismail@selcuk.edu.tr (I. Saritas).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.03.056

Table 1
The normal PSA values that depend on age.

Age (year)	PSA (ng/ml)
40–49	<2.5
50–59	<3.5
60–69	<4.5
>70	<6.5

studies are underway on different fields of artificial intelligence such as fuzzy expert system that can identify cancer risk by using relationships between prostate volume (PV), PSA level and age (Abbod, Keyserlingk, Linkens, & Mahfouf, 2001; Allahverdi, 2002; Allahverdi & Yaldiz, 1998; Boegla, Adlassniga, Hayashic, Rothenfluhd, & Leiticha, 2002; Saritas, 2003; Saritas, Allahverdi, & Sert, 2003) and ANN that can detect prostate cancer risk at an early phase (Cınar et al., 2009; Lorenz, Blum, Ermert, & Senge, 1997; Ronco & Fernandez, 1999).

ANN is better able to predict patients' prostate biopsy results than traditional statistics and assess large numbers of variables ranging from non-linear relationships to logical regression.

The purpose of this study is to develop a model that can determine whether patients have prostate cancer or not on the basis of data about their tPSA, fPSA and age prior to prostate biopsy.

2. Materials and method

The present study used only a small portion of the biopsies performed in the urology clinic at Meram Faculty of Medicine of Selcuk University and the laboratory data belonging to 121 male patients who were diagnosed and treated in the urology clinic of the Faculty of Medicine of Ankara University between the years of 2002 and 2004. The tPSA (0.28–150 ng/ml), fPSA (0.04–25 ng/ml), and age (44–89) data belonging to these patients and the results of the definitive diagnosis after the biopsy indicating whether they had cancer or not were used in the study.

MATLAB Neural Network Toolbox software was used in order to devise an ANN and conduct analyses.

2.1. Artificial neural network

ANN are computer systems that have been devised to automatically perform abilities such as deriving new information through learning, forming new information and discovering, all of which are characteristics of the human brain.

Generally, it consists of three layers, i.e. an input layer, one or more hidden layers and an output layer. Each layer has a certain number of components attached to one another called neurons or nodes. Each of the neurons is connected to the other with weights and accompanying communication networks. Signals move through neurons over weights. Each neuron receives multiple inputs from other neurons depending on their weights and generates an output signal that may also be generated by other neurons (Cınar et al., 2009; Lorenz et al., 1997; Ronco & Fernandez, 1999).

In order to devise an ANN, the network is processed in two levels, i.e. training and testing. In the level of training, the network is trained for an output prediction on the basis of input data. In the testing level, on the other hand, it is tested to stop or save the training and is used to predict an output.

When the tested error reaches the desired tolerance value, the training of the network is stopped (Lorenz et al., 1997; Ronco & Fernandez, 1999).

The back propagation (BP) algorithm is the most popular algorithm that has the widest area of use. BP is composed of two phases, namely procedures of feed forward and back propagation.

During feed forward, information that is processed from the input layer to the output layer is generated. In the case of back propagation, on the other hand, the difference between the network output value obtained from the feed forward procedure and the desired output value is compared to the desired difference tolerance and the error in the output layer is calculated. The error thus obtained is propagated backward in order to update the links in the input layer (Cınar et al., 2009; Ronco & Fernandez, 1999).

The BP training algorithm is a gradient descent. The BP algorithm functions to improve the performance of the network by reducing the total error through changing weights along its gradient. When the tested mean squared errors (MSE) stop decreasing and they begin to increase, which is a sign of over-training, the training is stopped (Lorenz et al., 1997; Ronco & Fernandez, 1999).

$$\mathrm{MSE}\% = \frac{1}{n}\sum_{i=1}^{n}(d_i - O_i)^2 \tag{1}$$

Here d_i is targeted or real value, O_i is network output or predicted value, and n is the output data number.

2.2. Application of artificial neural network to patient data

The purpose of this devised ANN is to determine existence of prostate cancer (PCa) on the basis of the tPSA, fPSA and age data. Data belonging to the 121 patients were divided into two, namely a training set and a test set. Ninety two data (70% of the total data) selected randomly from these data were used for the training data set of the ANN and the remaining 29 data (30% of the total data) were used for the test set.

In this study, a feed forward network structure that contains an input layer, a hidden layer and an output layer (Fig. 1) was used. After the ANN structure was designed, the data obtained in the experimental study were normalized in the 0–1 value set using Eq. (2) in order to improve the characteristics of the training. The back propagation algorithm was used in the training procedure. Different transfer functions (Purelin, Tansig, Logsig, etc.) were used and tried in the neurons in the hidden and output layers and Logarithmic-Sigmoid (Logsig) was selected as the transfer function that yielded the best result.

$$x_{\mathrm{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{2}$$

The training data set was used to determine ANN neuron and bias weight values. Training was repeated to obtain the lowest level of error by changing the number of neurons and the epoch number. Then, the trained algorithm was applied on the test data set.

First, ANN was trained by changing the number of neurons in the hidden layer (2–100) in order to determine the artificial neural

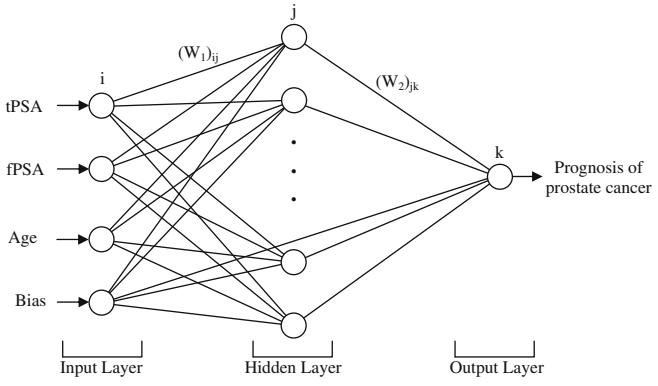

Fig. 1. The structure of ANN.

network algorithm that yielded the best result. As can be seen in Fig. 2, the best classification was yielded by the ANN structure that has eight neurons in its hidden layer.

The network structure with a hidden layer of eight neurons that was thus obtained was tested with different epoch numbers. The best epoch number was determined as 2500. At the end of these procedures, the network structure that yielded the best classification is given in Table 2.

The best approach that had the minimum number of errors was formed using the BP algorithm with eight neurons. A mathematical formula that can yield a prostate cancer diagnosis was obtained by using the weight values obtained in ANN (Eq. (3), PPC is prognosis of prostate cancer).

Table 3
Weight values for Eq. (7).

Number of neurons in the hidden layer	$W_{1(i,j)}$ for tPSA	$W_{1(i,j)}$ for fPSA	$W_{1(i,j)}$ for age
1	−19.2891	4.2242	8.4395
2	8.4335	−16.6372	−5.3583
3	17.9918	9.2384	15.5443
4	6.6621	−11.8568	18.7641
5	−30.6352	1.5713	−16.256
6	8.3159	8.4314	−30.2445
7	−8.7948	−3.3351	−6.1601
8	−2.4227	−11.468	0.2567

$$PPC = \frac{2}{1 + e^{(-2\times(-9.6980\times F_1 - 13.3551\times F_2 + 23.5396\times F_3 - 23.5578\times F_4 - 18.7654\times F_5 - 13.3128\times F_6 + 8.3472\times F_7 - 24.9484\times F_8 + 5.4932\times \text{bias}))}} - 1 \quad (3)$$

The transfer function used for this approach is the Tansig transfer function given in Eq. (4).

$$F_j = \frac{2}{1 + e^{(-2^* NET_j)}} - 1 \quad (4)$$

For Eq. (4), NET_j is calculated by using Eq. (5) depending on tPSA, fPSA and age.

$$NET_j = (W_1)_{i,1} * \text{tPSA} + (W_1)_{i,2} * \text{fPSA} + (W_1)_{i,3} * \text{Age} + (W_b)_{i,4} * \text{bias} \quad (5)$$

The constants whose weight values are $(W_1)_{i,j}$ for the BP algorithm, which has eight neurons in its hidden layer, are given in Table 3. In the equation given above, NET_j is the total of the input parameters and the multiplication of their weights. The subscripts i and j are input and hidden neuron numbers, respectively.

tPSA, fPSA and age were used in the ANN input layer, eight neurons were used in the hidden layer were and the prostate cancer diagnosis was used in the output layer. Eight equations for each of NET_1–NET_8 and F_1–F_8, were used as total and activation functions, respectively.

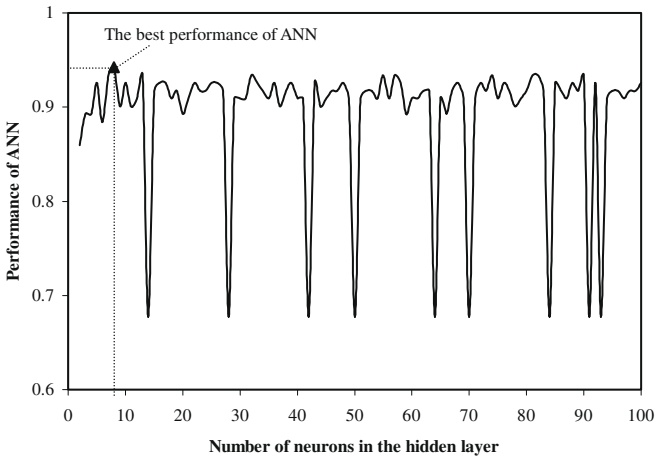

Fig. 2. The relationship between the number of neurons in the hidden layer and the ANN performance.

Table 2
The parameters and properties used in ANN.

Parameters	Properties
Number of neurons in the input layer	3
Number of the hidden layers	1
Number of neurons in the hidden layer	8
Number of neurons in the output layer	1
Learning rate (α)	0.3
Coefficient of momentum (β)	0.3
Learning algorithm	Gradient descent (Traingd)
Transfer function	Logarithmic-Sigmoid (Logsig)

3. Conclusion and suggestions

Confusion (Wichard, Cammann, Stephan, & Tolxdorff, 2008) matrix and ROC (Turker, Tokan, & Yıldırım, 2005) analyses were conducted to see the success rate of the study that was carried out.

3.1. Confusion matrix analysis

The ANN that was devised classified 88 of the 92 training data, 26 of the 29 test data and in total 114 of the 121 data as successful. The confusion matrix obtained through all of the training and test data sets is given in Table 4.

When the diagnostic test conducted on all of the data in Table 4 is examined, the following results are reached:

1. Sensitivity is the ability to distinguish the sick from the true ill.

$$\text{Sensitivity} = \frac{A}{A + C} \quad (6)$$

From Eq. (6), sensitivity was found to be = 0.9756.

2. Specificity is the ability to distinguish the healthy from the true healthy.

$$\text{Specificity} = \frac{D}{D + B} \quad (7)$$

From Eq. (7), specificity was found to be = 0.8717.

Table 4
Confusion matrix.

Classification	Sick	Healthy	Sum of rows
Data of reference			
Sick	80 (A)	5 (B)	85
Healthy	2 (C)	34 (D)	36
Sum of columns	82	39	121

3. Accuracy is the test's total accurate diagnoses of the sick and the healthy

$$\text{Accuracy} = \frac{A+D}{A+B+C+D} \tag{8}$$

From Eq. (8), accuracy was found to be = 0.9421.

4. The positive predictive value of the result (PPV) is the probability of being truly sick when the diagnostic test passes the judgment of sick.

$$\text{PPV} = \frac{A}{A+B} \tag{9}$$

From Eq. (9), PPV was found to be = 0.9411.

5. The negative predictive value of the result (NPV) is the probability of being really sick when the diagnostic test passes the judgment of healthy.

$$\text{NPV} = \frac{D}{D+C} \tag{10}$$

From Eq. (10), NPV was found to be = 0.9444.

3.2. ROC analysis

The basic idea behind medical tests is to calculate the probability of patient's being sick on the basis of the patient's test results. ROC analysis is used to determine the real accuracy of the results of the medical diagnosis. Among the terms used in this analysis, sensitivity is the rate of patients with positive tests (sick) to patients (True Positive), and specificity is the rate of healthy individuals (False Positive) whose tests are negative (healthy) to the number of healthy individuals (Metz, 1978). ROC analysis is a standard approach used to determine the sensitivity and specificity of the diagnosis. For this purpose, ROC curves are used to define the relationship between the sensitivity and specificity of the diagnosis. The axes of the curves are TP (labeling the sick ill) and FP (labeling the healthy sick). The curves are between the limits of 0 and 1, and while proximity to coordinate y and upper limit indicates successful test, curves that have a gradient of 45° point to an unsuccessful test.

Thus, the achievement of the test can be determined by examining the ROC curves. In a successful test, the area that remains under the curves is expected to be large. The ROC curve obtained for the training data set is given in Fig. 3, the ROC curve for the test data set is given in Fig. 4 and the ROC curve for the data set for all patients is given in Fig. 5.

ANN predictive values are obtained in a quick and simple manner with the patients' for tPSA, fPSA and age data by using the equations obtained in the study (Eqs. (3)–(5)).

According to the results of Confusion matrix analysis, it is observed that the rate of distinguishing the sick from the true sick is 97.56%, the ability to distinguish the healthy from the true healthy is 87.17% and the rate of total accurate diagnoses of sick and healthy is 94.21%. Accordingly, when the ANN predictive value passes the judgment of sick, the probability of being truly sick is 94.11%. Likewise, when the ANN predictive value passes the judgment of healthy, the probability of being truly healthy is 94.44%.

According to ROC analysis, the area that remains under the ROC curve is 0.9305 for the ANN training data set (Fig. 3); the area that remains under the ROC curve is 0.8947 for the ANN test data set (Fig. 4), and the area that remains under the ROC curve is 0.9237 for the ANN all data set (Fig. 5). The fact that these areas are large (close to 1) indicates that the study that was conducted was successful in predicting prostate cancer diagnosis.

According to the two analyses, the tPSA, fPSA and age data, which were used as input parameters, can be used in the accurate prediction of prostate cancer.

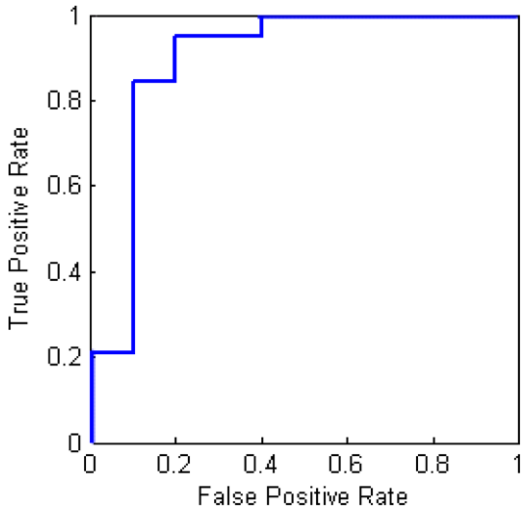

Fig. 4. Graph for the ROC analysis of test data.

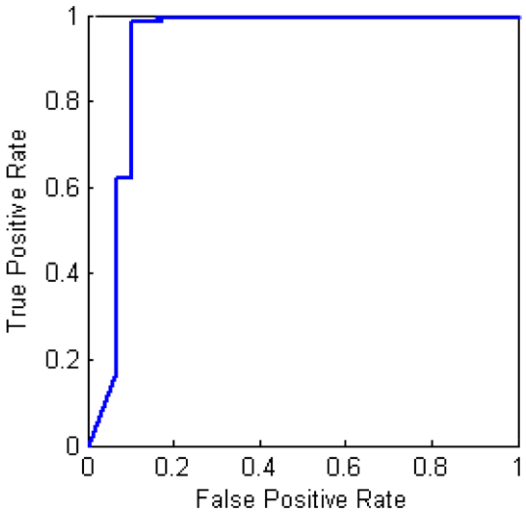

Fig. 3. Graph for the ROC analysis.

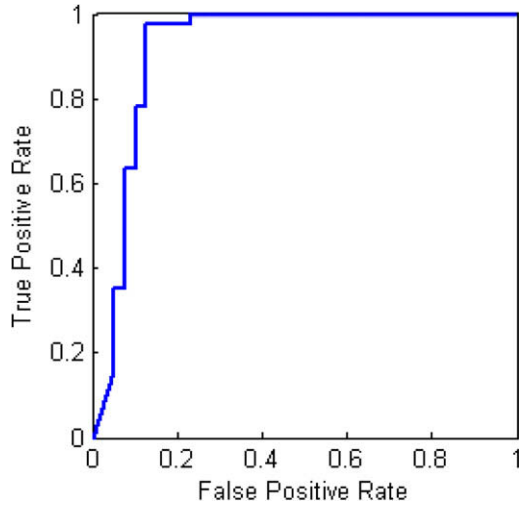

Fig. 5. Graph for the ROC analysis of all patients' data.

Since better results were observed in the present study than the other studies mentioned in the literature that were conducted previously using different methods, it can be used as an important instrument that eliminates unnecessary biopsy.

This study will contribute to science by making a more reliable observation of patients with a low risk of prostate cancer that must not be given a biopsy operation through a policy of wait and see.

Moreover, a better prediction of prostate cancer will be possible by increasing the patient data, adding appropriate input parameters and using artificial intelligence methods that can work together with ANN.

Acknowledgements

This study was supported by Selcuk University Coordination Office of Scientific Research Projects (BAP). Moreover, we would like to extend our heartfelt thanks to Prof. Dr. Sümer Baltacı of Ankara University, Faculty of Medicine, and Department of Urology, who assist in the obtainment of the data.

References

Abbod, M. F., Keyserlingk, D. G., Linkens, D. A., & Mahfouf, M. (2001). Survey of utilization of fuzzy technology in medicine and healthcare. *Fuzzy Sets and Systems, 120*, 331–349.

Allahverdi, N. (2002). *Expert systems. An artificial intelligence application*. Istanbul: Atlas Press. 248p.

Allahverdi, N., & Yaldiz, S. (1998). Expert system applications in medicine and example of design of a pre-diagnosis expert system. In *Proceedings of the second Turkish-German joint computer application days, 15–16 October 1998, Konya* (pp. 175–192).

Boegla, K., Adlassniga, K. P., Hayashic, Y., Rothenfluhd, T. E., & Leiticha, H. (2002). Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system. *Artificial Intelligence in Medicine, 676*, 1–26.

Brawer, M. K., & Kirby, R. (1999). *Prostate specific antigen* (2nd ed.). Health Press. 92p.

Catalona, W. J., Partin, A. W., Slawin, K. M., Brawer, M. K., Flanigan, R. C., Patel, A., et al. (1998). Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: A prospective multicenter clinical trial. *Journal of American Medical Association (JAMA), 279*, 1542–1547.

Cinar, M., Engin, M., Engin, E. Z., & Ates, Y. Z. (2009). Early prostate cancer diagnosis by using artificial neural networks. *Expert Systems with Applications*, 6357–6361.

Lorenz, A., Blum, M., Ermert, H., & Senge, Th. (1997). Comparison of different neuro-fuzzy classification systems for the detection of prostate cancer in ultrasonic images. *Ultrasonics Symposium, 2*, 1201–1204.

Metlin, C., Lee, F., & Drago, J. (1991). The American Cancer Society National prostate cancer detection, project: Findings on the detection of early prostate cancer in 2425 men. *Cancer, 67*, 2949–2958.

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 283–298.

Nguyen, H. P., & Kreinovich, V. (2001). Fuzzy logic and its applications in medicine. *International Journal of Medical Informatics, 62*, 165–173.

Prostate Cancer Symptoms. <http://www.medicinenet.com/Prostate_Cancer/page3.htm>. Arrived March 2009.

Prostate Online Calculator. <http://www.compass.fhcrc.org/edrnnci/bin/calculator/main.asp?t=prostate&sub=s1&m=&v=prostate&x=Prostate%20Cancer>. Arrived February 2009.

Ronco, A. L., & Fernandez, R. (1999). Improving ultrasonographic diagnosis of prostate cancer with neural networks. *Ultrasound in Medicine and Biology, 25*(5), 729–733.

Saritas, I. (2003). *Fuzzy control in medicine*. Master thesis. Selcuk University.

Saritas, I., Allahverdi, N., & Sert, I. U. (2003). A fuzzy expert system design for diagnosis of prostate cancer. In: *Proceedings of the international conference on computer systems and technologies – CompSysTech'2003, Sofia, Bulgaria* (pp. 345–351).

Seker, H., Odetayo, M., Petrovic, D., & Naguib, R. N. G. (2003). A fuzzy logic based method for prognostic decision making in breast and prostate cancers. *IEEE Transactions on Information Technology in Biomedicine, 7*, 114–122.

Shin Egawa, M. D., Shigehiro Soh, M. D., Makoto Ohori, M. D., Toyoaki Uchida, M. D., Kazuo Gohji, M. D., Akio Fujii, M. D., et al. (1997). The ratio of free to total serum prostate specific antigen and its use in differential diagnosis of prostate carcinoma in Japan. *Cancer, 79*, 90–98.

Turker, N., Tokan, F., & Yıldırım, T. (2005). Determination of artificial neural networks performances by ROC analysis on the diagnosis of heart disease. *Biomedical Engineering National Congress*, 206–208.

Van Cangh, P. J., De Nayer, P., De Vischer, L., Sauvage, P., Tombal, B., Lorge, F., et al. (1996a). Free to total prostate-specific antigen (PSA) ratio is superior to total-PSA in differentiating benign prostate hypertrophy from prostate cancer. *The Prostate, 29*, 30–34.

Van Cangh, P. J., De Nayer, P., De Vischer, L., Sauvage, P., Tombal, B., Lorge, F., et al. (1996b). Free to total prostate-specific antigen (PSA) ratio improves the discrimination between prostate cancer and benign prostatic hyperplasia (BPH) in the diagnostic gray zone of 1.8–10 ng/ml total PSA. *Urology, 48*, 67–70.

Wichard, J. D., Cammann, H., Stephan, C., & Tolxdorff, T. (2008). Classification models for early detection of prostate cancer. *Hindawi Publishing Corporation Journal of Biomedicine and Biotechnology*.