

## Prognostic Value of Artificial Neural Network in Predicting Bladder Cancer Recurrence After BCG Immunotherapy

### Research Article

Ajili F<sup>1</sup>, Issam B M<sup>2</sup>, Kourda N<sup>3</sup>, Darouiche A<sup>4</sup>, Chebil M<sup>4</sup>, Boubaker S<sup>1</sup>

<sup>1</sup>Laboratory of Human and Experimental Pathology, Pasteur Institute of Tunis, Tunisia. 13, Place Pasteur, 1002 Tunis, Tunisia.

<sup>2</sup>Higher School of Sciences and Techniques of Tunis, University of Tunis, Tunisia, Signal, Image and Intelligent Control of Industrial Systems.

<sup>3</sup>Department of Pathology, Charles Nicole Hospital, Tunis, Tunisia.

<sup>4</sup>Department of Urology, Charles Nicole Hospital, Tunis, Tunisia. Bur. Entrées: Bd du 9 Avril 1006, Tunis Bab Souika, TUNIS, 1006, Tunisia

### Abstract

**Background:** Artificial neural network (ANN) has been used in medicine to predict either the treatment or the investigative outcomes. The aim of this study was to validate the use of ANN models for predicting recurrence in non muscle invasive bladder cancer (NMIBC) treated by Bacillus Calmette Guerin (BCG) immunotherapy.

**Materials and methods:** In this study, we developed a Multilayer Perceptron (MLP) based ANN to detect recurrence in NMIBC through the analysis of histopathologic data. The study includes 308 patients (mean age, 63.92 years; range, 31–92 years) who were treated with transurethral resection followed by BCG-immunotherapy. Time follow-up was 30 months.

**Results:** In the test group, 39 out of 40 cases were correctly classified by the MLP base neural network with an optimum Mse error (0.02634). Only one case was classified as false positive, with no false negative results. The correlation between the ANN output and histopathologic results are summarized in Table 2. The sensitivity, specificity, positive predictive and negative predictive values calculated from the output data were 96.66%, 100%, 100%, and 90.9%, respectively (Table 2). Network can predict the outcome of 79% (34\*100/35) of patients in the testing data set correctly.

**Conclusion:** The proposed algorithm produced high sensitivity and specificity in predicting the recurrence in NMIBC after BCG immunotherapy compared to conventional statistical analysis. Therefore the use of ANNs will increasingly become the method of choice to calibrate complex medical models.

**Keywords:** Bladder Cancer; BCG-Immunotherapy; Prognostic; Artificial Neural Network; Recurrence.

### \*Corresponding Author:

Dr. Ajili Faouzia, PhD

Laboratory of Human and Experimental Pathology, Pasteur Institute of Tunis, Institut Pasteur de Tunis 13, Place Pasteur, 1002 Tunis, Tunisia.

Tel/Fax: 00 216.71.789 608/ 00 216.71.791 833

E-mail: Faouziaagili@yahoo.fr

**Received:** April 05, 2013

**Accepted:** April 26, 2013

**Published:** April 30, 2013

**Citation:** Ajili F et al (2013) Prognostic value of artificial neural network in predicting bladder cancer recurrence after BCG immunotherapy. Int J Cancer Stud Res. 2(2), 26-30. doi: <http://dx.doi.org/10.19070/2167-9118-130004>

**Copyright:** Ajili Faouzia © 2013. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

### Introduction

Non-muscle-invasive bladder cancer (NMIBC) is a slow-growing tumor. These tumors recur in 48% to 70% of patients after tran-

surethral bladder tumor resections, and 10% to 48% of these recurrent tumors carry the potential to become muscle invasive and metastatic [1-3]. Until now, intravesical bacillus Calmette-Guerin (BCG) instillations have proven to be the most successful adjuvant treatment for patients with intermediate- and high-risk non-muscle-invasive bladder cancer; however, no markers are available to predict BCG immunotherapy response. For these reasons, many studies have been conducted to develop scales to predict the recurrence and progression of these tumors [4-9].

Predictive models are used in a variety of medical domains for diagnostic and prognostic tasks. These models are built from experience, which constitutes data acquired from actual cases. The data can be preprocessed and expressed in a set of rules, such as training data for statistical and medical learning models. Among the options in the latter category, the most popular models in medicine are logistic regression (LR) and artificial neural networks (ANN). This latter model is an artificial intelligence tool that identifies arbitrary nonlinear multiparametric discriminant functions directly from clinical data. The use of ANNs has gained increasing popularity for applications where description of the dependency between dependent and independent variables is either unknown or very complex. This learning technique can be roughly

described as a universal algebraic function that will distinguish signal from noise directly from clinical data. The application of ANNs to complex relationships makes them highly attractive for the study of complexed medical decision making. In medicine, ANN has been used to predict either the treatment or the investigative outcomes. In this study, we looked at the use of artificial neural network to predict the early recurrence of non muscle invasive bladder cancer. This will then serve as an aid in determining the most appropriate adjuvant treatment after transurethral resection (TUR) and to prevent recurrence and progression after surgical resection.

**Materials**

During an 8-year period from 2002 to 2010, a total of 308 consecutive patients with histopathologically confirmed NMIBC including pTa and pT1 stages with or without concomitant CIS, were recruited from the Department of urology of Charles Nicole Hospital, Tunisia. All the tumours were totally endoscopically removed. All the patients had received six weekly instillations of BCG (BCG Pasteur strain, 75 mg in 50 ml saline), three to six weeks after the last transurethral resection. After the last instillation, urinary cytology and cystoscopic examination were performed. If these examinations were negative, patients continue the treatment of maintenance which consists in 24 additional monthly instillations. The follow up was performed for 30 months. The resulting grade was evaluated according to the 2004 WHO grading system [10]. The most widely used and universally accepted staging system is the tumour-node-metastases (TNM) system [11]. All patients were informed about the aim of the study and signed a written consent stating their agreement to participate in the trial. Information was obtained on professional data. For each patient, data were collected on multiplicity (single or multiple), histological grade (high grade or low grade) and stage (pTa or pT1). Patients underwent urine cytology and cystoscopy every 3 months for 2 years during treatment and every 6 months thereafter. Responders to BCG-immunotherapy were defined as patients who did not show cystoscopic or cytological evidence for tumor recurrence during the follow-up. Recurrence was defined as reappearance of tumor after the initial treatment with at least one tumor-free cystoscopy interval. The end-point for follow-up was either development of recurrence or the termination date of the study.

**Methods**

**Artificial Neural Networks (ANNs)**

Artificial Neural Networks (ANNs) are non-linear mapping structures based on the function of the human brain. They are powerful tools for modeling, especially when the underlying data relationship is unknown. ANNs can identify and learn correlated patterns between input data sets and corresponding target values. After training, ANNs can be used to predict the outcome of new independent input data. ANNs imitate the learning process of the human brain and can process problems involving non-linear and complex data even if the data are imprecise and noisy.

Neural networks have emerged as an important tool for classification. The recent vast research activities in neural classification have established that neural networks are a promising alternative to various conventional classification methods. The advantage of neural networks lies in the following theoretical aspects. First, neural networks are data driven self-adaptive methods in that they can adjust themselves to the data without any explicit specification of functional or distributional form for the underlying model. Second, they are universal functional approximators in that neural networks can approximate any function with arbitrary accuracy [12,13,14].

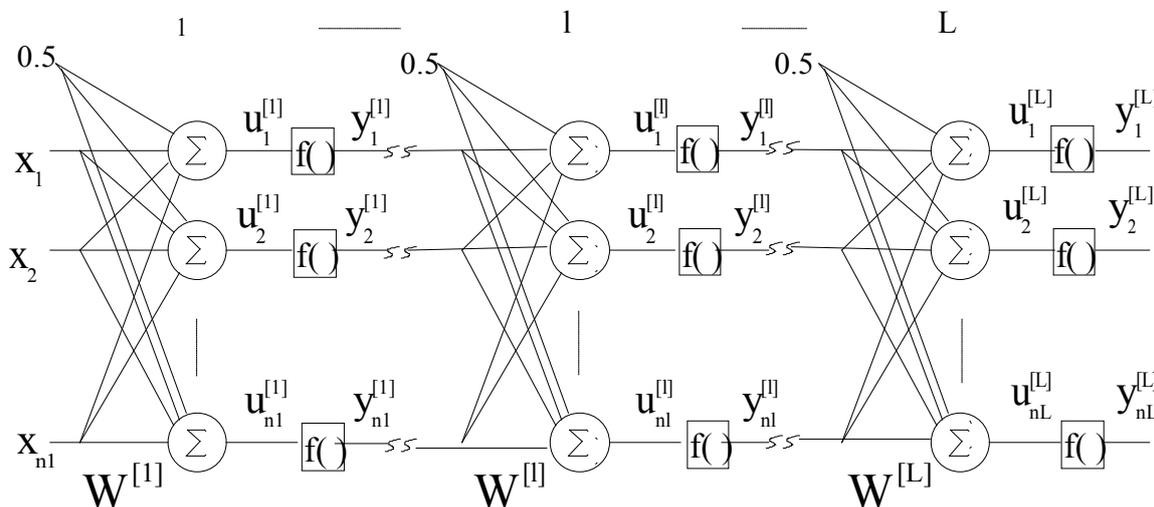
In this study, the ANN is implemented in the MATLAB software package (MATLAB version 6.5 with neural network toolbox), and is composed of a multilayer perceptron (MLP) was employed and trained with Back-Propagation (BP) algorithm.

**Multilayer Perceptron (MLP)**

An MLP consists of three layers: an input layer, an output layer, and one or more hidden layers (HL). Each layer is composed of a predefined number of neurons (Fig 1).

MLP has one or more HLs with a linear combination function is the inner product of inputs, weights and a bias. Activation function is logistic or tanh function. Inputs are fully connected to the first HL, each HL is fully connected to the next, and the last HL is fully connected to the outputs. MLPs use supervised learning or

**Figure1: Fully connected feed forward multilayer perceptron.**



backpropagation. Designing and training an MLP requires:

- i) Selecting number of HLs.
- ii) Number of neurons to be used in each HL.
- iii) Avoiding local minima.
- iv) Converging to an optimal solution in a reasonable period of time.
- v) Validating NN to test for overfitting.

MLPs are used in classification problems, in fitness approximations [15,16].

**Review of the Back propagation algorithm**

Back propagation training algorithm (BP) which is an iterative gradient descent algorithm is a simple way to train multilayer feed forward neural networks. The BP algorithm has become the standard algorithm used for training multilayer perceptron as shown in Fig.1. It is a generalized Least Mean Squares (LMS) algorithm that minimizes the sum of the squares of the errors between the actual and the desired outputs, Ep [17].

The linear and nonlinear neuron outputs for the current pattern are given respectively by:

$$u_j^{[s]} = \sum_{i=0}^{n_{s-1}} w_{ji}^{[s]} y_i^{[s-1]}$$

In the BP algorithm framework, each computational unit computes the same activation function. The computation of the sensitivity for each neuron requires the derivative of activation function, therefore this function must be continuous. The activation function is normally a sigmoid function chosen between the two

functions  $f(x) = \frac{1}{1+e^{-\lambda x}}$  and  $(x) = \frac{1-e^{-\lambda x}}{1+e^{-\lambda x}}$ . The steepness parameter  $\lambda$  determines the active region (region in which the derivative of sigmoid function is not very small) of activation function. As the steepness parameter decreases from positive infinity to zero, the sigmoid function changes from a unit step function to constant value 0.5.

In our application we used:

$$f(u_j^{[s]}) = \frac{1-e^{-\lambda u_j^{[s]}}}{1+e^{-\lambda u_j^{[s]}}} = y_j^{[s]}$$

Now:

$$E_p = \sum_{j=1}^{n_L} \frac{1}{2} (e_j^{[L]})^2$$

where the nonlinear error signal is:

$$e_j^{[L]} = d_j^L - y_j^L$$

$d_j^L$  and  $y_j^L$  are respectively the desired and the current outputs, for the jth unit. In (3), P denotes the pth pattern, nL is the number of the output units. The gradient descent method is given by:

$$\Delta w_{ji}^{[s]} = \mu \left( - \frac{\partial E_p}{\partial w_{ji}^{[s]}} \right)$$

Where  $w_{ji}^{[s]}$  is the weight of the i<sup>th</sup> unit in the (S-1)<sup>th</sup> layer to the j<sup>th</sup> unit in the s<sup>th</sup> layer. Since the BP algorithm is treated in the literature [18,19], only a summary of the important steps is given here:

- Compute the error signals for the output layer using:

$$\delta_j^{[L]} = e_j^{[L]} \cdot f'(u_j^{[L]})$$

- Compute the error signals for the hidden layers, i.e. for s=L-1 to 1, by:

$$\delta_j^{[L]} = f'(u_j^{[L]}) \sum_{r=1}^{n_{L+1}} (\delta_j^{[L+1]} w_{jr}^{[L+1]})$$

- Update the weights according to the following equation:

$$w_{ji}^{[s]}(k+1) = w_{ji}^{[s]}(k) + \mu \cdot \delta_j^{[s]} y_i^{[s-1]}$$

where  $\mu$  is the learning coefficient and  $f'$  is the first derivative of  $f$ .

As we note from (3) and (4), the BP algorithm minimizes the squared error at the output of the activation function with respect to the weights.

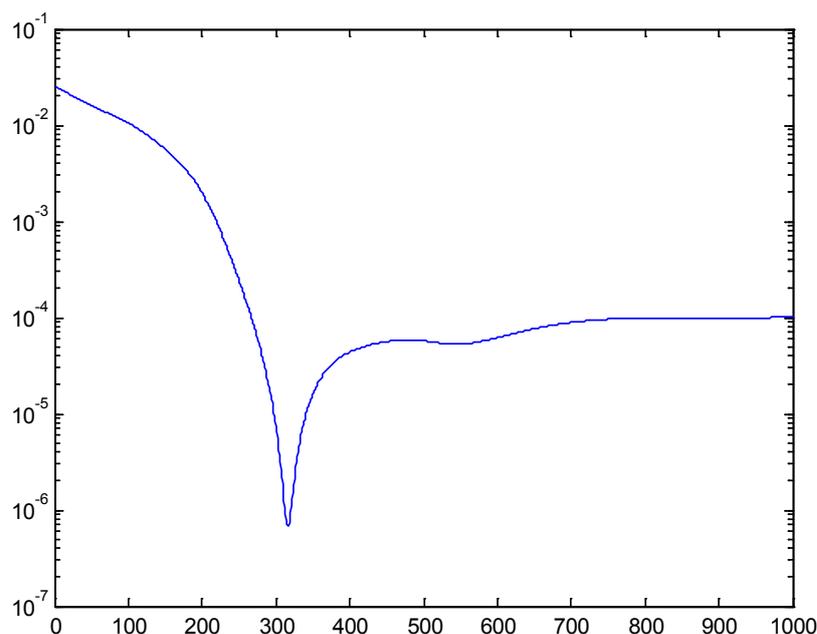
**ANN analysis**

Age, gender, tumour stage and grade, carcinoma in situ, tumor size, multiplicity and CD34 expression were used as input data for the ANN. Time to event and censoring information were not used by the neural network. Target variables were tumour recurrence. For each endpoint, 500 different ANN were constructed using the three-layer feed-forward multilayer perceptron architecture. This is the most common ANN type, which consists of an input layer, a hidden layer, and the output neuron. Each input neuron (perceptron) receives data from one input variable. 70% of the cases (patients) were selected by random and used for the training process, and the remaining 30% served as independent validation data set. This random split was repeated for every new ANN model. During training, the network calculates a decision (classification) for each case of the training data set. Then, the prediction is compared with the true category of each case, and the classification error is calculated. In order to minimize the error, the neurons and interconnections (“weights”) of the ANN are adjusted, and the next training cycle begins. This iterative training process is repeated until the classification error reaches a minimum. After training is finished, the network is tested by using the validation data set.

**Results**

In this study, data from 268 to 308 subjects were used for training, and the rest of them were used for control and testing procedures. To reduce network training epoch number and to obtain a better network result or generalization data, all of data were simply pre-classified with respect to patient age, gender, smoking, tumor

**Figure 2: The error variations with respect to training epoch numbers (sum squared error, Mse, was selected 0.0001. If Mse value is equal to 0.0001, training iterations is stopped).**



stage, tumor grade, tumor size, and multiplicity. The network outputs represented by unit basis are as follows:

- 1: Response to BCG immunotherapy
- 0: No response to BCG immunotherapy

The trained networks have 7 input units, one hidden layer with 20 units and 1 output neuron as shown in Table 1. In order to make this classification, four-layered perceptron was employed and trained by the back propagation learning algorithm. The sigmoid

function  $f(x) = \frac{1 - e^{-\lambda x}}{1 + e^{-\lambda x}}$  was selected as the transfer function that is used to determine the outputs according to the neuron inputs.

The testing mean square error (mse) was obtained 0.02634 from optimized multilayer perceptron feed forward network with training mean square error about 0.0001 at 1000 epochs.

After the training and testing procedures are completed, the results from the test data were analyzed in correlation with the histopathologic results, which are considered as the golden standard in clinical decision making. For a practical applicability in the clinical decisions, the following statistical parameters were calculated:

- Sensitivity = True positive / (True positive + False negative)
- Specificity = True negative / (True negative + False positive)
- Positive predictive value = True positive / (True positive + False positive)
- Negative predictive value = True negative / (True negative + False negative)

### Training Neural Networks

In the test group, 39 out of 40 cases were correctly classified by the MLP base neural network with an optimum Mse error (0.02634) (Fig.2). Only one case was classified as false positive, with no false negative results. The correlation between the ANN output and

histopathologic results are summarized in Table 2. The sensitivity, specificity, positive predictive and negative predictive values calculated from the output data were 96.66%, 100%, 100%, and 90.9%, respectively (Table 2). Network can predict the outcome of 79% (34\*100/35) of patients in the testing data set correctly.

### Discussion

Tumor recurrence following transurethral resection (TUR) is a significant problem for both the patient and the clinician. Intravesical bacillus Calmette-Guerin (BCG) immunotherapy represents a highly successful therapy for patients with non-muscle-invasive bladder cancer (NMIBC) tumors [20-23]. These tumors shows sometimes repeated superficial recurrences, but in some cases, they progresses into infiltrative or metastatic cancer despite diverse treatments. A significant portion of these patients fail to respond to BCG therapy; their tumors not only persist or recur, but also they may become invasive or metastatic [22-25].

Artificial neural networks (ANNs) are now the most popular artificial learning tool in biotechnology. Recently, it have been used by various researchers to predict bladder cancer progression [26,27,28]. However, no study has been interested to predict bladder cancer response to BCG immunotherapy. ANNs have an increasing popularity as a computational tool which is being evaluated in various medical areas including early detection, risk assessment, classification, and prognosis of disease states [29]. A significant advantage of this method is its ability to reveal complex relationships between available clinical data and the final outcome.

In this study, we developed a MLP based ANN to detect recurrence after BCG immunotherapy in NMIBC through the analysis of histopathologic data. To the best of our knowledge, our study is the first attempt to evaluate the role of ANN in predicting recurrence in NMIBC after BCG immunotherapy. Hereby, we found that the sensitivity, specificity, positive predictive and negative predictive values calculated from the output data were 96.66%, 100%,

100%, and 90.9%, respectively. This shows that the proposed algorithm produced high sensitivity and specificity in predicting the histopathologic results, which shows that this method has a promising value in estimation of NMIBC recurrence.

On line with our work, many studies have shown a significant improved accuracy in diagnosing, staging and predicting post-treatment results for ANNs compared to conventional statistical analysis. Indeed, Qureshi et al. [30] evaluated the ability of an ANN to predict bladder cancer recurrence within 6 months of diagnosis and stage progression in patients with Ta/T1 bladder cancer. ANN predictions were compared with those of four consultant urologists. The accuracy of the neural network in predicting stageprogression and recurrence within 6 months for Ta/T1 tumors and 32-month cancerspecific survival for T2–T4 cancers was 80%, 75%, and 82%; respectively; with corresponding figures for clinicians being 74%, 79% and 65%. In the other hand Parekattil and al [31] showed in a clinical trial on bladder cancer that their ANNs model was more accurate in identifying patients who required cystoscopy thereby providing possible savings. To summarize we can suggest that ANN models can be valuable tools in reducing the workload on the clinicians by detecting artefact and providing decision support.

## Conclusion

The proposed algorithm produced high sensitivity and specificity in predicting the recurrence in NMIBC after BCG immunotherapy compared to conventional statistical analysis. Therefore the use of ANNs will increasingly become the method of choice to calibrate complex medical models. Indeed, ANNs will allow a patient-oriented individual decision tool, relying on a waste base of information and decisions. They are able to combine biochemical markers, imaging studies and other tools with personal experience and act as a thinking module. However, prospective validation using a standardized multi-institutional database to overcome the pitfalls of ANNs are required. Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## References

- Sylvester RJ, van der Meijden AP, Oosterlinck W, Witjes JA, Bouffoux C, Denis L, et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. *Eur Urol* 2006;49:466-77.
- Oosterlinck W, Lobel B, Jakse G, Malmström PU, Stöckle M, Sternberg C. Guidelines on bladder cancer. *Eur Urol* 2002;41: 105-12.
- Millan-Rodriguez F, Chechile-Toniolo G, Salvador-Bayarri J, Palou J, Algaba F, Vicente-Rodriguez J. Primary superficial bladder cancer risk groups according to progression, mortality and recurrence. *J Urol* 2000;164:680-4.
- Kurth KH, Denis L, Bouffoux C, Sylvester R, Debruyne FM, Pavone-Macaluso M, et al. Factors affecting recurrence and progression in superficial bladder tumours. *Eur J Cancer* 1995; 31A: 1840-6.
- 5.Millan-Rodriguez F, Chechile-Toniolo G, Salvador-Bayarri J, Palou J, Vicente-Rodriguez J. Multivariate analysis of the prognostic factors of primary superficial bladder cancer. *J Urol* 2000; 163:73-8.
- Solsona E, Iborra I, Dumont R, Rubio-Briones J, Casanova J, Almenar S. The 3-month clinical response to intravesical therapy as a predictive factor for progression in patients with high risk superficial bladder cancer. *J Urol* 2000;164:685-9.
- Kaasinen E, Rintala E, Hellström P, Viitanen J, Juusela H, Rajala P, et al. Factors explaining recurrence in patients undergoing chemoimmunotherapy regimens for frequently recurring superficial bladder carcinoma. *Eur Urol* 2002;42:167-74.
- Kwak C, Ku JH, Park JY, Lee E, Lee SE, Lee C. Initial tumor stage and grade as a predictive factor for recurrence in patients with stage T1 grade 3 bladder cancer. *J Urol* 2004;171:149-52.
- Choi CK, Kim JS, Rim JS. Application of scoring system reflecting various prognostic factors to the prediction of recurrence in superficial bladder carcinoma. *Korean J Urol* 1999;40:878-85.
- G. Sauter, F. Algaba, M.B. Amin, et al. Non invasive urothelial tumors, in: J.N. Elbe, G. Sauter, J.I. Epstein, I.A. Sesterhenn (Eds.), *Pathology and Genetics. Tumors of the Urinary System and Male Genita Organs*, IARC Press, Lyon, 2004, pp. 110–111.
- Sobin DH, Wittekind CH, editors. *TNM classification of malignant tumours*, 6th ed. New York, New York: Wiley-Liss; 2002. p. 199–202.
- G. Cybenko. Approximation by superpositions of a sigmoidal function. *Math. Contr. Signals Syst.* 1989,(2):303–314.
- P. A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1982.
- K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, 1989 (2):359–366.
- Judd, J.S. *Neural Network Design and the Complexity of Learning*, Cambridge, MA: The MIT Press. 1990.
- Hagan, T.M., Demuth,H.B., Beale, M. *Neural Network Design*. Cengage Learning India, 1996.
- S. Abidi, F. Fnaiech and M. Najim: A Fast Feed-Forward Training algorithm using A Modified Form of the standard Back Propagation Algorithm. *IEEE Trans. Neural Network* 2001 (12); No 2.
- A. Cichocki. R. Unbehauen, *Neural network for optimization and signal processing*. John Wiley & Sons Ltd. Baffins Lane, Chichester. West Sussex PO19 1 UD, England 1993.
- R. P. Lippman. An introduction to computing with neural networks. *IEEE ASSP Mag.* 1987 (4). No, 2 April.
- Morales A, Eideger D, Bruce AW. Intracavitary Bacillus CalmetteGuerin in the superficial bladder tumor. *J Urol* 1976;116:180–3.
- Lamm DL. Preventing progression and improving survival with BCG maintenance. *Eur Urol* 2000;37(Suppl 1):9–15.
- Patard JJ, Rodriguez A, Leary E, et al. Intravesical bacillus Calmette Guerin treatment improves patient survival in T1G3 bladder tumors. *Eur Urol* 2002;41:635–42.
- Saint F, Salomon L, Quintela R, et al. Do prognostic parameters of remission versus relapse after bacillus Calmette–Guerin (BCG) immunotherapy exist? Analysis of a quarter century of literature. *Eur Urol* 2003;43:351–61.
- Sylvester RJ, van der Meijden APM, Oosterlinck W, et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. *Eur Urol* 2006;49:466–77.
- Fernandez-Gomez J, Solsona E, Unda M, et al. Prognostic factors in patients with non-muscle-invasive bladder cancer treated with bacillus Calmette-Guerin: multivariate analysis of data from four randomized CUETO trials. *Eur Urol* 2008;53:992–1002.
- Abbod, M. F., Linkens, D. A., Catto, J. W., & Hamdy, F. C. Comparative study of intelligent models for the prediction of bladder cancer progression. *Oncology Reports*; 2006 (15), 1019–1022.
- Catto, J. W., Linkens, D. A., Abbod, M. F., Chen, M., Burton, J. L., Feeley, K. M., et al. (2003). Artificial intelligence in predicting bladder cancer outcome: a comparison of neuro-fuzzy modeling and artificial neural networks. *Clinical Cancer Research*, 9(11), 4172–4177.
- Fujikawa, K., Matsui, Y., Kobayashi, T., Miura, K., Oka, H., Fukuzawa, S., et al. Predicting disease outcome of non-invasive transitional cell carcinoma of the urinary bladder using an artificial neural network model: results of patient follow-up for 15 years or longer. *International Journal of Urology*. 2003, 10(3), 149–152.
- Kapetanovic, I. M., Rosenfeld, S., & Izmirlan, G. Overview of commonly used bioinformatics methods and their applications. *Annals of the New York Academy of Science*. 2004, 1020, 10–21.
- Qureshi KN, Naguib RNG, Hamdy FC, Neal DE, Mellon JK. Neural network analysis of clinicopathological and molecular markers in bladder cancer. *J Urol* 2000;163:630–3.
- Parekattil, S. J., Fisher, H. A., & Kogan, B. A. (2003). Neural network usingcombined urine nuclear matrix protein-22, monocyte chemoattractant protein-1 and urinary intercellular adhesion molecule-1 to detect bladder cancer. *The Journal of Urology*, 169(3), 917–920.