A Dynamic Model for Predicting Prostate Cancer in Iranian Men Based on a Perceptron Neural Network

Farzad Allameh,[1,*] Hamidreza Qashqai,[2] and Alborz Salavati[3]

[1]Laser Application in Medical Sciences Research Center, Urology Department, Shohada-e-Tajrish Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2]Urology Department, Shohada-e-Tajrish Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3]Urology Research Center, Tehran University of Medical Sciences, Tehran, Iran

[*]*Corresponding author*: Farzad Allameh, Urology Branch, Laser Application in Medical Sciences Research Center, Shohada-e-Tajrish Hospital, Shahid Beheshti University of Medical Sciences, Tajrish Sq, Tehran, Iran. Tel: +98-9123885545, Fax: +98-2122749221, E-mail: farzadallame@gmail.com

Abstract

Objectives: To test a novel neural network prediction model for prostate cancer based on age, rectal examination, prostate specific antigen (PSA) and prostate volume.

Methods: 572 men who underwent trans-rectal ultrasound guided prostate biopsy from February, 2013 to September, 2014 participated in the study. Prostate configuration based on digital rectal examination, serum PSA level, and prostate volume were recorded. Pathologic outcomes were categorized in two groups: adenocarcinoma vs. noncancerous reports. A multi-layer perceptron (MLP) neural network was designed in which total PSA, free PSA, age, rectal examination results and prostate volume were vectors.

Results: 566 men with the average age of 65.9 ± 8.6 years. Average total and free PSA levels were 19.77 ± 50.03 ng/mL and 2.46 ± 8.36 ng/mL respectively. Average free to total PSA ratio was $14.68 \pm 11.24\%$. Prostate size was 58.58 ± 31.64^{CC} on average. Age, total PSA, prostate volume and abnormal DRE were correlated with prostate cancer at biopsy, and the most powerful of all was abnormal DRE with odds ratio of 0.12. Neural networks were formed on a 3-layer perceptron and finally a network of 6 entry, 9 middle, and 2 output nodes was selected with the learning rate of 0.05. The Correct prediction rate for the model was 85.3%.

Conclusions: It seems that our three-layer perceptron neural network model proves better results than the logistic regression model in predicting the presence of prostate cancer based on total and free PSA, DRE result, prostate volume and age.

Keywords: Prostate Cancer, Prediction Model, Neural Network, Iranian Men

1. Background

As in other parts of the world, prostate cancer is among the top three visceral cancers in Iran and accountable for cancer specific deaths as well (1). Prostate cancer is believed to be a slow evolving disease with few markers that make early diagnosis possible. Over the past decade prostate cancer has led to many unnecessary invasive diagnostic and therapeutic procedures, imposing both financial and health wise burdens on society (1, 2). This situation may in part be due to uncertain cut off points and unavailability of complementary tests such as prostate cancer antigen 3 (PCA3). Both PCA3 and, to a lesser extent, a 4k-panel have added value to the digital rectal examination (DRE)- based on European Randomized Study of Screening for Prostate Cancer (ERSPC)risk calculator in detecting prostate cancer in prescreened men (3).

Recently, it has been indicated that the stepwise increase in prostate specific antigen (PSA) from the first day of testing predicted a 3.57 fold increased risk of prostate cancer, and 2.16 fold increased risk of prostate cancer mortality (4). PSA isoforms and PSA kinetics have also been associated with more aggressive phenotypes, but are not routinely employed as part of prediction tools prior to treatment (5).

On the other hand, PSA density has discriminative predictive power for prostate cancer. It has similar sensitivity but greater specificity compared to using total PSA, DRE and the ratio of free to total PSA (PSA f/t) (6). In patients with clinically localized prostate cancer and PSA amount from 4 to10 ng/mL, lower PSA f/t was significantly associated with tumor extracapsular extension, other adverse pathologic features and with biochemical recurrence at the long-term follow-up (7). Another trial showed that both PSA density and PSA transition zone density had significant predictive values in discriminating prostate cancer. Furthermore, it was shown that PSA transition zone density had the strongest predictive value (8).

These findings clearly show the need to an integrated model for the prediction of prostate cancer risk. Therefore, along other applications of neural networks in analyzing data pertaining to prostate cancer, we have tried to test a novel neural network prediction model based on age, rectal examination, PSA level and prostate volume specifically in Iranian men.

Prior studies have investigated the probability and established the efficacy of artificial neural networks in cancer detection in many diseases including breast cancer and prostate cancer (9) with the hope of forming a multivariate model which will be able to answer the critical question: whether or not to perform the prostate biopsy (10). Unnecessary biopsies could yield high complication rates (11) including infections, sepsis, consecutive hospitalizations, and complications. Such consequences are major drawbacks for the urologist (2).

Historically, many indicators such as density, zone specific density, velocity, even ultrasound features such as elasticity and echo texture have been used to assist modeling a disease prediction neural network for prostate cancer (12), but none have been yet recognized as the reliable source for replacing the simple clinical decision making (13).

2. Methods

Patients: 572 consecutive men who were referred for trans-rectal ultrasound (TRUS) guided prostate biopsy from February, 2013 to September, 2014 were enrolled. Subjects were referred to as a result of abnormal prostatic examination (DRE) or raised serum PSA with alarm signs considering international guidelines.

Ethical consideration: Informed consent was obtained from each patient and all participants were assured that their information and corresponding pathology results will be used only in clinical research.

Patient age, total and free PSA, digital rectal examination result which was either recorded as normal or abnormal, and prostate volume (according to TRUS) were recorded anonymously.

All patients were examined by one urologist and when the shape of the prostate was asymmetrical, a prominent nodule was detected or when prostate became hard or stony, DRE was recorded as abnormal.

TRUS guided biopsies were performed in two wards of urology and radiology in Shohada-e-Tajrish hospital and all individuals underwent 12-coreneedle biopsy. One histopathology center was responsible for reporting all the specimens. Unfortunately, 6 patients were excluded from the study due to the failure to follow up. Using SAS and MATLAB a logistic regression was performed to validate the correlation between the above mentioned factors and biopsy results. Pathologic outcomes were categorized in two groups: adenocarcinoma labeled as cancerous versus all other reports which were labeled as non-cancerous. The cut-off point of 0.05 was considered for statistical significance.

2.1. Artificial Neural Network

It is in fact a simulation of the human brain via modeling the neurons in which each neuron works as a processing unit. Multi-layer perceptron (MLP) neural network is one of the most widely used types of networks and its structure includes several layers (input, hidden and output layers) and in each layer a number of activity nodes and functions are defined. The output of each layer is calculated using the sum of the weighed coefficients in that layer and is sent to the next layer via an activity function. There are various methods and algorithms for finding the weights, but in the MLP network, the back propagation (BP) algorithm is used. Additionally, the activity function in the neural network models is similar to link functions in the generalized linear models. As examples of the activity functions the sigmoid and the hyperbolic tangent functions can be mentioned. In a general and simple case in an MLP with one hidden layer, the output value of the ith unit can be calculated as follows:

Equation 1.

$$y_i = f_2 \left(\beta + \sum_{k=1}^{p} W_k f_1 \left(\beta_k + \sum_{j=1}^{m} X_{ij} W_{kj} \right) \right) \qquad (1)$$

where n is the number of observations, p is the number of nodes in the hidden layer, m is the number of nodes in the input layer (the number of independent variables), w_{kj} is the weight of the input x_{ij} in the kth node, w_k is the weight of the kth node, β_0 and β_k^0 are the bias values of the output and hidden layers respectively. Also, f_1 and f_2 are the activity functions of the hidden and output layers of the network, respectively.

Total PSA, free PSA, age, rectal exam result and prostate volume were vectors, and different weight allocations resulted in different networks and prostate pathology results (cancerous vs. non-cancerous) were used as the output (14).

3. Results

Overall, 566 men were enrolled with the average age of 65.9 ± 8.6 years, the youngest being 39 and the oldest being 88 years old. The average total and free PSA levels were 19.77 ± 50.03 ng/mL and 2.46 ± 8.36 ng/mL, respectively with average free to total PSA ratio of 14.68 ± 11.24 %. Prostate size was 58.58 ± 31.64^{CC} in average. Reports of pathology results are summarized in Table 1. The variables are demonstrated in Table 2 with P value calculated for each one separately.

From 276 patients with normal DRE, 195 (70.7%) were in non-cancerous group and from 289 patients with abnormal DRE, 242 (83.7%) patients who had prostate cancer; the p value of Chi-square test was < 0.001.

Table 1. Distribution of Pathology Results in Patients Who Underwent TRUS Guided Biopsy

Pathology Result	No	%
Adenocarcinoma	324	57.2
ASAP	4	0.7
BPH	141	25.0
Chronic inflammation	64	11.3
HG-PIN	2	0.4
Negative for malignancy	31	5.4
Total	566	100.0

Abbreviations: ASAP, Atypical Small Acinar Proliferation, BPH, Benign Prostate Hyperplasia; HG-PIN, High Grade Prostatic Intraepithelial Neoplasm.

Logistic regression confirmed that age, total PSA, prostate volume and abnormal DRE were correlated with prostate cancer in biopsy; the most powerful of all was abnormal DRE with odds ratio of 0.12 that means the probability of detecting adenocarcinoma in patients with abnormal DRE is 88%. However, other odds are calculated per one unit increasing in age, PSA and prostate volume (Table 3).

Neural networks were formed on a 3-layer perceptron comprising 1 to 12 middle layer nodes with threshold of 0.80-0.95, learning rate of 0.01 - 0.40, and sigmoid function. Regarding 12 probable structures, 107 models were processed and finally a network of 6 entry, 9 middle and 2 output nodes was selected with the learning rate of 0.05, and using back propagation it achieved prediction rate of 85.3 %, under receiver operating characteristic (ROC) curve area of 0.9 - 1, wrong prediction rate of 14.7% and the network information criterion of 684.29.

When the neural network model was compared to the logistic regression model, areas under the ROC curve were 0.885 and 0.901 respectively. Correct prediction rates were 85.3% for network and 81.2% for regression model (Table 4).

4. Discussion

After testing the relationship between the input parameters and prostate cancer, these variables were used for creating a logistic regression and a neural network model. Correct prediction rates were greater in neural network guesses compared to the logistic regression model.

The collection of samples from the population of men who referred for prostate biopsy might potentially lower the reliability of our study because the probability of non-cancerous biopsy was lower in our participants compared to the general population. Nevertheless, all results mentioned above are statistically significant; therefore, they could be applicable in general populations. One should have in mind that biopsy is too invasive to be performed on volunteers without indication so we had to gather samples from men who were referred for biopsy.

Cancer detection rate for PSA and DRE screening seems to be different among various ethnicities, for instance it has been reported between 4 to 40% for the same cut-off point of PSA concentration (15, 16). As stated by Pourmand et al. it seems to be about 3.8 % at 2 ng/mL for Iranians. Such studies emphasize that positive predictive value for such tests increases with age thus rendering the decision whether or not performing a biopsy is a complex question.

Historically, prostate biopsy would be performed in case of PSA rise but biopsy complications, regardless of morbidity and costs, made urologists reconsider the flow chart of prostate cancer screening and evaluation and PSA velocity, peripheral or central zone specific PSA density, free to total PSA ratio, etc. have been developed, but regarding biopsy complications, a definitive tool for patient selection is lacking.

In our study we used age, total and free PSA and DRE results beside prostate volume as main measures on which the network would be trained.

Like many other multi-factorial disease, there has been significant use of neural networks with the aim of ameliorating positive predictive value of prostate cancer screening. There has been an effort in Iran by Ghaderzadeh et al. few years ago, albeit there is not any reliable version of clinical decision support system available for Iranian physicians.

According to the results of correct prediction rate by two models and the area under ROC curve, it seems that our 6/9/2 nodes three-layer perceptron neural network proves better results in comparison with the logistic regression model in predicting the presence of prostate cancer based on total and free PSA, DRE result, prostate volume and age, which are factors accepted worldwide in the assessment and selection of patients for prostate biopsy.

Acknowledgments

Table 2. Descriptive Statistics of the Variables

Variables	Range	Mean	SD	P Value
Age, y	39 - 88	65.99	8.61	< 0.001
Total PSA, ng/mL	0.2 - 620	19.77	50.03	< 0.001
Free PSA, ng/mL	0.05 - 107	2.46	8.36	0.090
free/total PSA, %	0.9 - 98	14.68	11.24	< 0.001
Prostate volume, cc	5.1 - 272	58.58	31.64	< 0.001

Abbreviations: PSA, Prostate Specific Antigen; SD, Standard Deviation.

Table 3. Logistic Regression Results with Odds Ratio for One Unit Increasing of Variables

Variables	OR	SD	95% CI	P Value
Age	1.08	0.016	1.05 - 1.12	< 0.001
Total PSAd	1.06	0.017	1.02 - 1.09	< 0.001
Free PSA	0.98	0.029	0.93 - 1.04	0.694
Prostate volume	0.97	0.005	0.96 - 0.98	< 0.001
DRE	0.12	0.255	0.07 - 0.19	< 0.001

Abbreviations: CI, Confidence Interval; DRE, Digital Rectal Examination; OR, Odds Ratio; PSA, Prostate Specific Antigen; SD, Standard Deviation.

Table 4. Values of Sensitivity, and Specificity Areas Under the ROC Curve and Correct Percentage of the Predictions for Comparing Logistic Regression and Neural Networks Models

Model	Sensitivity, %	Specificity, %	Area Under Curve	Correct Percentage, %
Logistic regression	81.2	69.1	0.885	81.2
Neural network	83.9	71.3	0.901	85.3

Abbreviations: ROC, Receiver Operating Characteristic.

Footnotes

References

1. Pourmand G, Ramezani R, Sabahgoulian B, Nadali F, Mehrsai A, Nikoobakht M, et al. Preventing Unnecessary Invasive Cancer-Diagnostic Tests: Changing the Cut-off Points. *Iran J Public Health.* 2012;**41**(2):47–52. [PubMed: 23113134].
2. Pourmand G, Salem S, Mehrsai A, Lotfi M, Amirzargar MA, Mazdak H, et al. The risk factors of prostate cancer: a multicentric case-control study in Iran. *Asian Pac J Cancer Prev.* 2007;**8**(3):422–8. [PubMed: 18159981].
3. Vedder MM, de Bekker-Grob EW, Lilja HG, Vickers AJ, van Leenders GJ, Steyerberg EW, et al. The added value of percentage of free to total prostate-specific antigen, PCA3, and a kallikrein panel to the ERSPC risk calculator for prostate cancer in prescreened men. *Eur Urol.* 2014;**66**(6):1109–15. doi: 10.1016/j.eururo.2014.08.011. [PubMed: 25168616].
4. Orsted DD, Nordestgaard BG, Jensen GB, Schnohr P, Bojesen SE. Prostate-specific antigen and long-term prediction of prostate cancer incidence and mortality in the general population. *Eur Urol.* 2012;**61**(5):865–74. doi: 10.1016/j.eururo.2011.11.007. [PubMed: 22104593].
5. Carter HB. Differentiation of lethal and non lethal prostate cancer: PSA and PSA isoforms and kinetics. *Asian J Androl.* 2012;**14**(3):355–60. doi: 10.1038/aja.2011.141. [PubMed: 22343493].
6. Verma A, St Onge J, Dhillon K, Chorneyko A. PSA density improves prediction of prostate cancer. *Can J Urol.* 2014;**21**(3):7312–21. [PubMed: 24978363].
7. Masieri L, Minervini A, Vittori G, Lanciotti M, Lanzi F, Lapini A, et al. The role of free to total PSA ratio in prediction of extracapsular tumor extension and biochemical recurrence after radical prostatectomy in patients with PSA between 4 and 10 ng/ml. *Int Urol Nephrol.* 2012;**44**(4):1031–8. doi: 10.1007/s11255-012-0135-y. [PubMed: 22315156].
8. Tanaka N, Fujimoto K, Chihara Y, Torimoto M, Hirao Y, Konishi N, et al. Prostatic volume and volume-adjusted prostate-specific antigen as predictive parameters for prostate cancer patients with intermediate PSA levels. *Prostate Cancer Prostatic Dis.* 2007;**10**(3):274–8. doi: 10.1038/sj.pcan.4500957. [PubMed: 17339878].
9. Saritas I, Ozkan IA, Sert IU. Prognosis of prostate cancer by artificial neural networks. *Expert Syst Appl.* 2010;**37**(9):6646–50. doi: 10.1016/j.eswa.2010.03.056.

10. Djavan B, Remzi M, Zlotta A, Seitz C, Snow P, Marberger M. Novel artificial neural network for early detection of prostate cancer. *J Clin Oncol.* 2002;**20**(4):921–9. doi: 10.1200/JCO.2002.20.4.921. [PubMed: 11844812].

11. Glotsos D, Kalatzis I, Theocharakis P, Georgiadis P, Daskalakis A, Ninos K, et al. A multi-classifier system for the characterization of normal, infectious, and cancerous prostate tissues employing transrectal ultrasound images. *Comput Methods Programs Biomed.* 2010;**97**(1):53–61. doi: 10.1016/j.cmpb.2009.07.003. [PubMed: 19647888].

12. Ghaderzadeh M, Sadoughi F, Ketabat A. A computer-aided detection system for automatic classification of prostate cancer from benign hyperplasia of prostate. *Iranian J Med Inform.* 2013;**2**(2).

13. Çınar M, Engin M, Engin EZ, Ziya Ateşçi Y. Early prostate cancer diagnosis by using artificial neural networks and support vector machines. *Expert Syst Appl.* 2009;**36**(3):6357–61. doi: 10.1016/j.eswa.2008.08.010.

14. Goss EP, Vozikis GS. Improving health care organizational management through neural network learning. *Health Care Manag Sci.* 2002;**5**(3):221–7. doi: 10.1023/A:1019760901191. [PubMed: 12363049].

15. Hernandez DJ, Han M, Humphreys EB, Mangold LA, Taneja SS, Childs SJ, et al. Predicting the outcome of prostate biopsy: comparison of a novel logistic regression-based model, the prostate cancer risk calculator, and prostate-specific antigen level alone. *BJU Int.* 2009;**103**(5):609–14. doi: 10.1111/j.1464-410X.2008.08127.x. [PubMed: 19007374].

16. Yousefi Kashi AS, Khaledi S, Houshyari M. CT Simulation to Evaluate of Pelvic Lymph Node Coverage in Conventional Radiotherapy Fields Based on Bone and Vessels Landmarks in Prostate Cancer Patients. *Iran J Cancer Prev.* 2016 June 14;**9**(3):e6233. eCollection 2016.