

# Modeling the Factors Affecting Gleason Score with Artificial Neural Networks and Indirectly Determining Prostate Cancer Risk Factors

**Type:** Research Article

**Received:** May 30, 2024

**Published:** June 27, 2024

**Citation:**

Zeynep Kucukakcali., et al.  
"Modeling the Factors Affecting  
Gleason Score with Artificial  
Neural Networks and Indirectly  
Determining Prostate Cancer  
Risk Factors". PriMera Scientific  
Medicine and Public Health 5.1  
(2024): 03-11.

**Copyright:**

© 2024 Zeynep Kucukakcali., et al. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Zeynep Kucukakcali\* and Ipek Balikci Cicek**

*Department of Biostatistics and Medical Informatics, Inonu University Faculty of Medicine, 44280, Malatya, Turkey*

**\*Corresponding Author:** Zeynep Kucukakcali, Department of Biostatistics and Medical Informatics, Inonu University Faculty of Medicine, 44280, Malatya, Turkey.

## Abstract

**Aim:** Prostate cancer, one of the most common cancer among men and a cancer that can vary significantly in its aggressiveness, will cause more deaths in the future with its increasing incidence. Gleason score has been defined as the most reliable and autonomous predictor of prostate cancer outcomes. The study aim to determine the variables affecting Gleason score and indirectly to establish prognostic indicators for prostate cancer.

**Material and methods:** The variables in the data set were analyzed according to the dependent variable categories Independent sample test and Mann Whitney U test were used in statistical analyses and  $p < 0.05$  was considered significant. Analyses were performed with IBM SPSS 26.0. In the modeling phase, the relationship between the grouped form of Gleason score and other variables was examined with Multilayer Perceptron and Radial Basis Function Neural Network methods. The dataset was divided into training and test datasets in a 70:30 ratio. The results are reported using accuracy, balanced accuracy, sensitivity, specificity, PPV, NPV and F1 score as performance metrics.

**Results:** The data set used in the study consists of variables belonging to 97 patients. The mean age of the patients was 63.87 years. Patients were divided into two groups: those with a Gleason value of 7 and above and those with a Gleason value below 7. There were 35 patients with a Gleason value below 7 and 62 patients with a Gleason value above 7. According to the results obtained from the modeling, the best result was obtained from the Multilayer Perceptron model. accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value and F1 score were 96%, 100%, 93.3%, 90.9%, 100%, 95.2%, respectively.

**Conclusion:** The study obtained highly accurate classification results when modeling with Gleason score categories and other independent variables. This shows that machine learning models can be used effectively and successfully in medical data. Furthermore, important variables were identified and their indirect associations with prostate cancer were revealed. In the future, more

detailed research on prostate cancer can be conducted by focusing on these variables.

**Keywords:** Prostate cancer; Gleason score; modelling; classification

## Introduction

A malignancy that develops in the prostate gland, an organ of the male reproductive system, is known as prostate cancer. It is one of the most prevalent cancers among males, and its degree of aggressiveness can differ significantly. The prostate gland contributes to the production of a portion of the fluid comprising sperm. Prostate cancer generally arises from the uncontrolled proliferation of cells within the prostate organ. These abnormal cells may eventually aggregate into a tumor and, in certain instances, metastasize to other anatomical sites [1].

Prostate cancer, with 1,276,000 new cases and 359,000 deaths in 2018, ranks as the second most often diagnosed disease and the sixth highest cause of death from cancer among men worldwide. By 2040, the incidence of prostate cancer is projected to increase to 2.3 million new cases, and the number of deaths caused by the disease is estimated to reach 740,000. This rise in prostate cancer cases and deaths can be attributed to both population expansion and the aging population [2]. Prostate cancer may not show any symptoms in its initial stages and tends to progress slowly, often not requiring significant therapy. Nevertheless, the most prevalent grievance is urinary difficulties, heightened frequency, and nocturia, all of which can be attributed to prostate enlargement. Urinary incontinence and back stiffness may occur in more advanced stages of bone metastatic disease, as the crucial skeleton is the most frequent site of the disease. Elevated plasmatic levels of prostate-specific antigen (PSA > 4 ng/mL), a glycoprotein regularly generated by prostate tissue, are often used to identify many prostate cancers. Nevertheless, tissue biopsy is considered the most reliable method to definitively confirm the existence of malignancy, as there have been instances where men without cancer have shown elevated levels of PSA [3].

Multiple factors, including genetic susceptibility, age, and hormone imbalances impact prostate cancer growth. The condition may manifest with symptoms such as urinary difficulties, hematuria, impotence, and discomfort in the buttocks, back, or chest. Nevertheless, during its initial phases, prostate cancer may not exhibit any symptoms, underscoring the significance of regular screenings and check-ups for timely identification [4].

The diagnosis of prostate cancer typically entails a comprehensive approach that includes physical examinations, blood tests to assess prostate-specific antigen (PSA) levels, and imaging techniques such as ultrasound, MRI, or biopsy. Possible treatment modalities for prostate cancer encompass active surveillance, surgical intervention, radiation therapy, hormone manipulation, chemotherapy, immunotherapy, or targeted therapy, contingent upon the cancer's stage and level of aggressiveness [5].

The intricate and diverse nature of prostate cancer presents difficulties in terms of categorization and the implementation of therapy beginning techniques. In the case of a very intricate malignancy like prostate carcinoma, doing a pathologic examination, which involves assessing biopsy cores, has been proven to enhance the chances of finding the disease and improve the precision of pathologic grading [6]. The Gleason score was identified as the most reliable and autonomous predictor of prostate cancer outcomes [7].

Hence, this investigation aims to identify the variables that influence the Gleason score and establish prognostic indicators for prostate cancer by indirectly linking them to the Gleason score. In order to achieve this objective, the data set will be modeled using neural network techniques to identify the components.

### ***The Data Set Included in The Study And Its Characteristics***

The data utilized in this investigation were obtained from a study by Stamey et al [8]. The dataset comprises information obtained from a cohort of 97 males who were scheduled to undergo radical prostatectomy. The variables examined in the study are as follows:

*lcavol*: The logarithm of the cancer volume, measured in milliliters (cc). The cancer region was ascertained from digital pictures and then multiplied by a given thickness to compute its volume.

*lweight*: Prostate weight is quantified in grams and represented by the logarithm.

*Age*: The patient's age, measured in years.

*lbph*: The logarithm of the size of benign prostatic hyperplasia (BPH), a noncancerous enlargement of the prostate gland, measured as an area in a digital picture and expressed in square centimeters.

*svi*: Seminal vesicle invasion is a binary indicator (0/1) that shows whether prostate cancer cells have invaded the seminal vesicle.

LCP refers to the logarithm of capsular penetration, which quantifies the degree of cancer infiltration into the fibrous tissue surrounding the prostate gland. It is calculated as the linear distance of penetration in centimeters.

*Gleason*: The Gleason score is a metric used to assess a tumor's aggressiveness level. The Gleason grading method categorizes the two greatest cancerous regions in tissue samples by assigning a score ranging from 1 to 5. The grades are added together to calculate the Gleason score.

*pgg45*: The proportion of Gleason scores that are classified as 4 or 5.

LPSA refers to the logarithm of the concentration of prostate-specific antigen (PSA) measured in nanograms per milliliter.

### **Neural networks, Multilayer Perceptron and Radial Basis Function Neural Network**

Neural networks are a significant subject in the realm of artificial intelligence, encompassing various classifications. The Multilayer Perceptron (MLP) and Radial Basis Function Neural Network (RBFNN) are the predominant neural network models utilized in this domain. MLP, short for Multilayer Perceptron, is a neural network architecture commonly employed for solving classification and regression tasks [9]. Conversely, the Radial Basis Function Neural Network (RBFNN) is a specific sort of neural network that has an input layer, a hidden layer, and an output layer. It excels at solving pattern recognition and prediction issues [10]. MLP models are frequently employed for acquiring intricate connections and demonstrate notable efficacy in solving regression and classification tasks [11]. Furthermore, artificial neural networks are often used in time series analysis and financial forecasting [12]. Deep learning techniques, namely multilayer neural networks, are highly successful for tackling intricate issues [13]. MLPNN is a model of an artificial neural network that has a feed forward structure. It is composed of the input layer, the output layer, and the hidden layer(s) that are located between these two layers. Multiplying the connection weights between the hidden layer and the output layer in the same manner allows for the collection of the inputs that are received by the neurons in the hidden layer and their subsequent transmission to the output layer. A neuron in the output layer is responsible for collecting these inputs and producing an output in accordance with them [14].

RBFNN operates by calculating the distances between neurons in the input layer and is commonly employed for the identification of intricate patterns within a dataset. Neural networks have the potential to achieve success in tasks involving pattern recognition and prediction [10]. The Radial Basis Function neural network (RBFNN) is composed of three layers: an input layer, a single hidden layer that utilizes radial functions, which are responsible for the network's name, and an output layer. The working principle of RBFNN involves determining the appropriate width and center values in the hidden layer based on the input data. This is done by creating linear combinations of the outputs produced by these functions in the output layer and establishing the relationship between input and output [14].

As a conclusion, neural networks are an effective instrument that may be utilized to solve a variety of issues that arise within the realm of artificial intelligence. In the event, that they are trained appropriately, neural network models such as MLP and RBFNN can be applied to a variety of applications and can produce beneficial outcomes.

### Biostatistical Analyses and Modelling Phase

The variables in the data set used in the study were subjected to analysis with the categories of the dependent variable and the data were summarized under the assumption of normal distribution with mean  $\pm$  standard deviation for normally distributed variables and median (minimum-maximum) for non-normally distributed variables. Normal distribution was checked by Shapiro Wilk test. In statistical analyses, independent samples t test and Mann Whitney U test were used where appropriate.  $p < 0.05$  was considered statistically significant. Analyses were performed using IBM SPSS 26.0.

In the modeling phase, the relationship between the grouped form of the dependent variable Gleason score and other variables was examined using Multilayer Perceptron and Radial Basis Function Neural Network methods. First, the data set was divided into training and test data sets in a ratio of 70 to 30 to ensure internal validity using the retention method. The results obtained as a result of modeling are given by performance metrics. At this stage, the results are given separately for the training and test dataset and the results are given with the performance metrics of accuracy, balanced accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), F1 score.

### Results

The data set used in the study consists of variables belonging to 97 patients. The mean age of the patients was 63.87 years and the patients were divided into 2 groups as those with Gleason values of 7 and above and those with Gleason values below 7. There are 35 patients with a Gleason value below 7 and 62 patients with a Gleason value above 7. The results obtained from the statistical analysis of the independent variables, with the grouped Gleason value being the dependent variable, are given in Table 1.

Variable	Gleason score		p
	Less than 7	7 and greater than 7	
	Mediaan (Minimum-maximum)		
lbph	-1,386(-1,386-2,326)	0,438(-1,386-2,326)	0.227*
lcp	-1,386(-1,386-1,619)	0,336(-1,386-2,904)	<0.001*
pgg45	0(0-0)	30(4-100)	<0.001*
	Mean $\pm$ Standart Deviation		
lcavol	0,565 $\pm$ 1,13	1,793 $\pm$ 0,96	<0.001**
lweight	3,558 $\pm$ 0,387	3,669 $\pm$ 0,448	0.224**
age	61,114 $\pm$ 8,018	65,419 $\pm$ 6,679	0.006**
lpsa	1,738 $\pm$ 1,019	2,896 $\pm$ 1,013	<0.001**

\*:Mann Whitney U test; \*\*:Independent Sample t test.

**Table 1:** Statistical Results of The Dependent Variable and Other Variables.

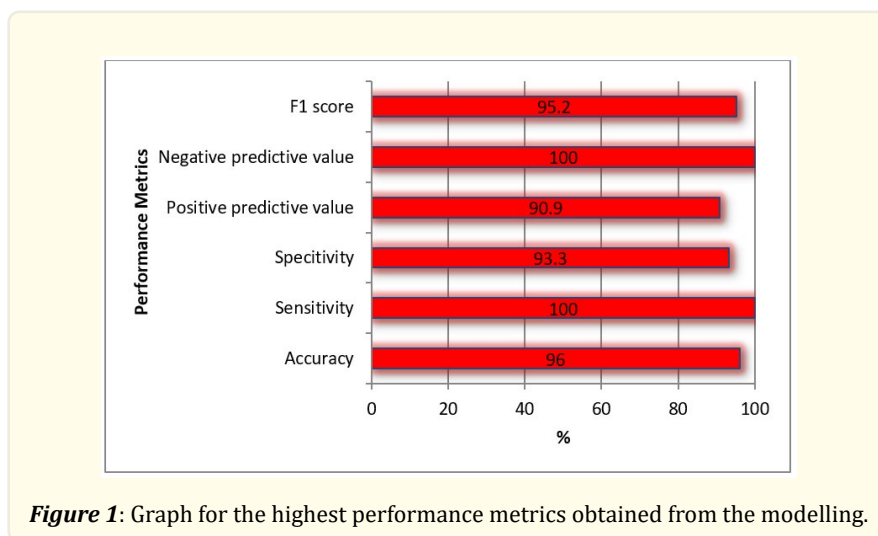
The values of the performance metrics obtained in the modellings between the dependent variable and other variables using the Multilayer Perceptron and Radial Basis Function Neural Network method are given for test data set and the results of these values are given in Table 2.

<b>Modeller</b>	<b>Metric</b>	<b>Value(%)</b>
Multilayer Perceptron	Accuracy	96.0 (88.3-100)
	Sensitivity	100 (69.2-100)
	Specitivity	93.3 (68.1-99.8)
	Positive predictive value	90.9 (58.7-99.8)
	Negative predictive value	100 (76.8-100)
	F1 score	95.2 (86.9-100)
Radial Basis Function Neural Network	Accuracy	88.6 (78.0-99.1)
	Sensitivity	88.3 (51.6-97.9)
	Specitivity	91.3 (72.0-98.9)
	Positive predictive value	88.3 (51.6-97.9)
	Negative predictive value	91.3 (72.0-98.9)
	F1 score	83.3 (71.0-95.7)

**Table 2:** Performance metrics for utilized modellings.

The accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value and F1 score obtained from the Multilayer Perceptron model with the highest result are 96%, 100%, 93.3%, 90.9%, 100%, 95.2% respectively.

Figure 1 displays the graphical depiction of the performance metrics acquired from the Multilayer Perceptron model, which has the highest performance metrics.



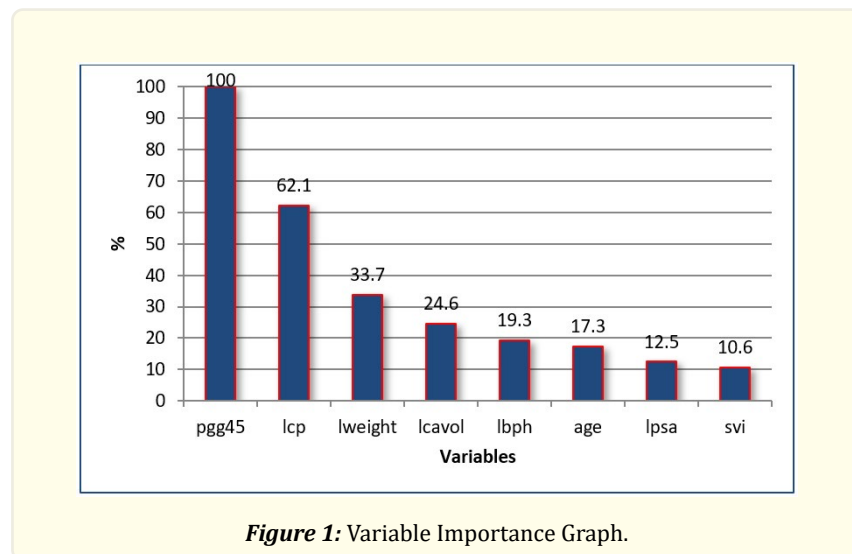
**Figure 1:** Graph for the highest performance metrics obtained from the modelling.

Variable importance values are a metric that measures how effective certain input variables are in the predictions of a model. The significance values calculated for each attribute are used to understand the complex internal structure of the model and to explain the predictions of the model. The variable importance values obtained Multilayer Perceptron model as a result of the modelling is given in Table 3.

<i>Variables</i>	<i>Normalized Importance</i>
pgg45	100.0%
lcp	62.1%
lweight	33.7%
lcavol	24.6%
lbph	19.3%
age	17.3%
lpsa	12.5%
svi	10.6%

The variable importance graph of the modelling is given in figure 1.

**Table 3:** Variable Importance Values Obtained as A Result of Multilayer Perceptron Model.



**Figure 1:** Variable Importance Graph.

## Discussion

Prostate cancer is a major global public health concern that affects men worldwide and contributes to illness and death rates. The incidence of prostate cancer exhibits regional disparities, underscoring the significance of comprehensive epidemiological investigations to comprehend its effects on diverse populations [15-17]. Obesity, dietary habits, and genetic predispositions are recognized as risk factors that affect the development and progression of prostate cancer, highlighting the intricate nature of addressing this public health issue [18, 19]. Lifestyle factors, including diet and comorbidities like diabetes, have been associated with prostate cancer incidence and outcomes, emphasizing the need for preventive strategies and early detection initiatives [20]. Efforts to address prostate cancer as a public health issue involve conducting research on chemoprevention, genetic markers, and tailored screening methods to improve patient outcomes and decrease the impact of the disease [21-23]. Therefore, studies on specific biomarkers that can reveal the causes of prostate cancer and the development of personalized medicine steps for the disease by elucidating the genomic infrastructure are important. Prostate cancer is a significant worldwide public health issue that necessitates the identification of early detection markers, comprehensive research involving multiple disciplines, strong preventative efforts, and customized interventions to minimize its effects on individuals and healthcare systems.

The Gleason score plays a crucial role in evaluating and treating prostate cancer. The Gleason score system, an established prognostic indicator, assesses the architectural patterns of prostate cancer cells. It is used to guide treatment decisions and predict clinical outcomes [24]. Research has demonstrated that the Gleason score is a robust indicator of the advancement of the disease, the likelihood of death, and the possibility of the cancer returning in individuals with prostate cancer [24, 25]. The Gleason score categorizes prostate cancer into distinct risk categories, facilitating the classification of tumors according to their level of aggressiveness and likelihood of spreading to other parts of the body [26]. Distinctive patterns within the Gleason score, such as cribriform growth, have been recognized as strong indicators for the likelihood of distant metastases and mortality specifically caused by the disease in individuals with Gleason score 7 prostate cancer [27]. Furthermore, attempts have been undertaken to distinguish between prostate cancer patients with a Gleason score of 7 by analyzing histopathological images and genomic data to acquire a deeper understanding of the variation in the illness and its prognosis [28].

To summarize the Gleason score is an essential tool used to categorize the risk, evaluate the prognosis, and develop treatment strategies for individuals diagnosed with prostate cancer. The significance of its application in clinical practice is shown by its capacity to accurately capture the histological characteristics of prostate cancers and forecast disease outcomes.

Based on these results, this study aims to identify risk factors for prostate cancer by determining the factors affecting Gleason score with machine learning models. For this purpose, patients were divided into two groups those with a Gleason value of 7 and above and those with a Gleason value below 7. Modeling was done with the data set obtained. Multilayer Perceptron model and Radial Basis Function Neural Network model were used in the modeling phase. According to the performance metrics obtained from the modeling, the performance metrics obtained from the MLP model were higher. The performance metrics obtained as accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value are 96%, 100%, 93.3%, 90.9%, 100%, 95.2% respectively.

Variable importance values, which measure how effective certain input variables are in the predictions of a model, are given based on the MLP model where high performance metrics are obtained. In line with the results obtained from the MLP model, it is seen that the variables that explain the dependent variable most in the modeling established with the categories of Gleason score, which is the dependent variable, are pgg45, lcp, lweight, lcaivol, lbph, age, lpsa, svi, respectively. By evaluating the importance of the variables with the given variable importance values, the variables that have the most impact on the performance of the model can be prioritized, which allows for more accurate predictions for the disease and a better understanding of the underlying relationships. It is widely recognized in the literature that pgg45 is an important marker for predicting prostate cancer progression and aggressiveness. High pgg45 rates are generally associated with higher Gleason scores and poor prognostic outcomes. Therefore, identifying pgg45 as the most important variable in the model is consistent with the findings in the literature and makes an important contribution in improving the accuracy of the model [29]. Clinical stage (lcp) plays a critical role in the diagnosis and treatment of prostate cancer. Literature shows that clinical stage has a strong association with cancer spread and patient prognosis. Therefore, specifying lcp as the second most important variable in the model is consistent with clinical observations and positively affects the prediction performance of the model [29]. Prostate weight has been included in the literature as an auxiliary factor in the assessment of prostate cancer risk. Larger prostates may have an impact on cancer detection and Gleason score, especially due to the possibility of confusion with benign prostatic hyperplasia (BPH). In this context, it makes sense to include it as an important variable in the model. Tumor volume is an important indicator in assessing the aggressiveness of cancer and its potential to spread. In the literature, larger tumor volumes have been reported to be associated with higher Gleason scores and worse clinical outcomes. Therefore, the high significance value of lcaivol in the model is consistent with the findings in the literature [30]. lbph may play a role in differentiating BPH from prostate cancer and in the assessment of cancer progression. Including lbph as an important variable in the model is clinically relevant and may improve the performance of the model. Age is an important determinant of prostate cancer risk. As age increases, the incidence of prostate cancer also increases; therefore, age is expected to be an important variable in the model and is consistent with the findings in the literature [31]. PSA level is a widely used biomarker in prostate cancer screening and is correlated with Gleason score. The inclusion of lpsa as an important variable in the model reflects the prognostic value of PSA and is consistent with the findings in the literature [32]. Seminal vesicle invasion is an advanced sign of prostate cancer invasion and is associated with high Gleason scores. The identification

of svi as an important variable in the model is meaningful in terms of reflecting the severity of the disease. In conclusion, the variable importance values obtained from the MLP model are consistent with the literature and these variables have a significant impact on model performance. This increases the predictive power of the model, allowing more accurate and reliable predictions to be made about the disease.

In conclusion, in light of the findings of the study, very successful classification results were obtained from the modeling using Gleason score categories and other independent variables. These results prove that machine learning models can be used very successfully in medical data and provide high accuracy. In addition, with the variable importance values obtained, the variables that affect the Gleason score categories the most were determined and their indirect relationship with prostate cancer was revealed. Different studies can be conducted with these variables and evaluations can be made for prostate cancer.

## References

1. Habib A., et al. "Risk factors associated with prostate cancer". *Journal of Drug Delivery and Therapeutics* 11.2 (2021): 188-93.
2. Culp MB., et al. "Recent Global Patterns in Prostate Cancer Incidence and Mortality Rates". *European urology* 77.1 (2020): 38-52.
3. Rawla P. "Epidemiology of prostate cancer". *World journal of oncology* 10.2 (2019): 63-89.
4. Archer M, Dogra N and Kyrianiou N. "Inflammation as a driver of prostate cancer metastasis and therapeutic resistance". *Cancers* 12.10 (2020): 2984.
5. Tinay İ and Türkeri L. "A New Era in Metastatic Prostate Cancer: "The Combination of Chemotherapy and Hormonal Treatment as Initial Treatment". *Bulletin of Urooncology* (2016).
6. Bastian PJ., et al. "Characteristics of Insignificant Clinical T1c Prostate Tumors". *Cancer* (2004): 2001-5.
7. Mahal BA., et al. "Prostate Cancer-Specific Mortality Across Gleason Scores in Black vs Nonblack Men". *Jama* (2018): 2479-2481.
8. Stamey TA., et al. "Prostate Specific Antigen in the Diagnosis and Treatment of Adenocarcinoma of the Prostate. II. Radical Prostatectomy Treated Patients". *The Journal of Urology* 141.5 (1989): 1076-83.
9. Ramchoun H., et al. "Multilayer Perceptron: Architecture Optimization and Training". *International Journal of Interactive Multimedia and Artificial Intelligence* (2016).
10. Wang Z, Liu Z and Zheng CD. *Introduction to Neural Networks* (2015).
11. Yi W, Zhao Y and Chan APC. "A Tailored Artificial Intelligence Model for Predicting Heat Strain of Construction Workers". *Iop Conference Series Earth and Environmental Science* (2022).
12. Tekin TG and Patir S. "Estimating the American Dollar Exchange Rate by Artificial Neural Networks: 2009 - 2021 Period". *R&s - Research Studies Anatolia Journal* (2023).
13. Yıldırım DK, Cemek B and Kucuktopcu E. "Daily Evaporation Estimation with Fuzzy Artificial Neural Networks and Multilayer Artificial Neural Networks". *Earth Water Magazine* (2019).
14. Çiçek İB and Küçükakçali Z. "Classification of prostate cancer and determination of related factors with different artificial neural network". *Middle Black Sea Journal of Health Science* 6.3 (2020): 325-32.
15. Zhou CK., et al. "Prostate Cancer Incidence in 43 Populations Worldwide: An Analysis of Time Trends Overall and by Age Group". *International Journal of Cancer* 138.6 (2015): 1388-400.
16. Djiwa T., et al. "Prostate Cancers in Men Under the Age of 50: About a Series in Togo, Sub-Saharan Africa". *BMC Cancer* 22.1 (2022): 1341.
17. Allott EH, Masko EM and Freedland SJ. "Obesity and Prostate Cancer: Weighing the Evidence". *European Urology* 63.5 (2013): 800-9.
18. Sheng T., et al. "No Association Between Fiber Intake and Prostate Cancer Risk: A Meta-Analysis of Epidemiological Studies". *World Journal of Surgical Oncology* 13.1 (2015): 264.
19. Allott EH and Hursting SD. "Obesity and Cancer: Mechanistic Insights from Transdisciplinary Studies". *Endocrine Related Cancer* 22.6 (2015): R365-R86.
20. Moradi A, Zamani M and Moudi E. "A Systematic Review and Meta-Analysis on Incidence of Prostate Cancer in Iran". *Health Pro-*



- motion Perspectives 9.2 (2019): 92-8.
21. McDonald MJ and Parsons JK. "The Case for Tailored Prostate Cancer Screening: An NCCN Perspective". *Journal of the National Comprehensive Cancer Network* 13.12 (2015): 1576-83.
  22. Butler EN, et al. "Fatal Prostate Cancer Incidence Trends in the United States and England by Race, Stage, and Treatment". *British Journal of Cancer* 123.3 (2020): 487-94.
  23. Porcacchia AS, et al. "Sleep Disorders and Prostate Cancer Prognosis: Biology, Epidemiology, and Association with Cancer Development Risk". *European Journal of Cancer Prevention* 31.2 (2021): 178-89.
  24. Kikugawa T. "Primary Gleason Pattern Does Not Affect Recurrence-Free Survival in Patients Receiving Radiotherapy for Prostate Cancer". *Journal of Nuclear Medicine & Radiation Therapy* 05.03 (2014).
  25. Sinha AA, et al. "Ratio of Cathepsin B to Stefin a Identifies Heterogeneity Within Gleason Histologic Scores for Human Prostate Cancer". *The Prostate* 48.4 (2001): 274-84.
  26. Epstein JI, et al. "A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score". *European Urology* 69.3 (2016): 428-35.
  27. Kweldam CF, et al. "Cribriform Growth Is Highly Predictive for Postoperative Metastasis and Disease-Specific Death in Gleason Score 7 Prostate Cancer". *Modern Pathology* 28.3 (2015): 457-64.
  28. Ren J, et al. "Differentiation Among Prostate Cancer Patients with Gleason Score of 7 Using Histopathology Image and Genomic Data". *Proc SPIE Int Soc Opt Eng* (2018).
  29. Epstein JI, et al. "The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma". *The American journal of surgical pathology* 29.9 (2005): 1228-1242.
  30. Humphrey PA. "Gleason grading and prognostic factors in carcinoma of the prostate". *Modern pathology* 17.3 (2004): 292-306.
  31. Humphrey PA. "Prostate Pathology". Chicago: ASCP Press (2003).
  32. Thompson Ian M., et al. "Prevalence of prostate cancer among men with a prostate-specific antigen level  $\leq$  4.0 ng per milliliter". *New England Journal of Medicine* 350.22 (2004): 2239-2246.