From Single Biomarkers to Combinations: Identifying Optimal Cutoffs for Prostate Cancer Detection

Xue Qin, Qiwei He (Supervisor), Gina D'Angelo (Mentor)

Georgetown University, Astra Zeneca

Abstract

Prostate cancer is one of the most common cancer types, making early detection methods crucial for improving patient outcomes. Biomarkers play a vital role in identifying the presence of disease at an early stage and assessing an individual's risk of developing cancer. Compared to individual biomarker cutoffs, finding biomarker combination cutoffs are essential to increase the accuracy of the diagnosis outcomes. This study will focus on comparing various methods for determining biomarker cutoffs, evaluating their predictive performance, and identifying the method that shows the most potential for accurately estimating cutoffs to aid in diagnostic detection and outcomes.

Key Words: Biostatistics, Biomarker, Cutoff, Prostate Cancer

From Single Biomarkers to Combinations: Identifying Optimal Cutoffs for Prostate Cancer Detection

1. Introduction

According to clinical statistics from 2016 in the USA, more than 1,685,210 new cancer cases were projected to occur. The most common cancer types included prostate, lung and bronchus, and colorectal cancers, which accounted for about 44% of cases in men (Siegel, 2016). This highlights the importance of developing methods for early disease detection to improve survival rates. Biomarkers are essential in helping to detect the presence of a disease at an early stage and certain biomarkers can indicate Individual's risk of developing a disease. As defined in the book, Biomarkers in Cancer Screening and Early Detection, "a biomarker is a particular characteristic, or a molecular fingerprint, which indicates manifestation of a physiological state, and which can be objectively quantified to distinguish a normal state from a pathological condition (e.g., cancer) or a response to a therapeutic intervention" (Srivastava, 2017). Clinical trials often rely on biomarkers for stratification and enrichment, particularly to identify subgroups of patients who are at higher risk of disease or more likely to respond to treatment. Since biomarkers are often continuous variables, it is important to establish cut-off points for effective stratification. Cancer is a complex, heterogeneous disease with numerous genetic and epigenetic changes. This complexity makes detection, diagnosis, and treatment challenging, thus using tools like biomarkers can facilitate early detection, help avoid overdiagnosis, and test drug sensitivity and monitoring treatment (Srivastava, 2017), which help diagnose the patient who will be more likely to benefit from the drug.

1.1 Traditional Method Finding Cutoffs for Individual Biomarker

While many current approaches focus on finding a single biomarker cut-off, there is a need for methods that evaluate multiple biomarkers simultaneously to improve predictive accuracy and clinical outcomes. The traditional method for identifying the individual biomarker cutoff involved applying the ROC-based methods, like Youden Index and Point Closest to (0,1). To explore the optimal cutoff point, based on the sensitivity and specificity of different biomarker levels, calculate the Youden Index for each cutoff value, and maximize the Youden Index would be the optimal cutoffs (Chen, 2015). Total prostate-specific antigen (TPSA) and ratio of free-to-total PSA (RPSA) are the essential biomarkers that can be used for diagnosing

Prostate cancer (Etzioni, 1999). While ROC-based methods can help determine individual biomarker cutoffs, relying on them alone may not provide an accurate assessment of cancer stage or severity. Based on the past research, total-PSA performs better as the time to diagnose diseases, while its diagnostic performance declines over time. Conversely, the free-to-total PSA ratio performs relatively better at earlier time points, but total PSA becomes superior closer to the time of diagnosis (Etzioni, 1999). Therefore, the individual biomarker has deficiencies in identifying the outcome at different times. Relying on individual PSA cutoff points has limitations at predicting the diagnosis.

1.2 Machine Learning Method Finding Cutoffs for Biomarker Combinations

Finding Biomarker combination cutoffs are essential to increase the accuracy of the diagnosis outcomes. Logistic Regression is another solution that is useful and can be used to directly determine the cut-off values for biomarker combinations and increase the accuracy. Based on Z.zhang's research, the application of the logistic regression model compared to the traditional approach, would reduce the bias and variance (Baker, 2011). They used the Molecular Epidemiology of Colorectal Cancer(MECC) study to do simulation experiments to compare the method result and showed improved estimation accuracy particularly in scenarios involving complex interactions between genetic markers and environmental factors (Baker, 2011). The use of logistic regression also facilitates the exploration of interaction effects, which are also crucial in understanding complex diseases like cancer. However, Logistic Regression is not the only method for estimating optimal cutoffs. In Balkishan Sharma's research, several statistical methods were compared to determine the best cutoffs. One example used Logistic Regression to determine cut-off scores for variables like the Berg Scale, ABC Scale, and simple reaction time to identify individuals at risk of falling (Sharma, 2014). The sensitivity is 91% and specificity is 97%, which showed the effectiveness of accurately predicting the cutoffs (Sharma, 2014). While Logistic Regression performs well when the relationship between variables is linear, it may struggle with non-linear relationships. When the relationship between biomarkers and disease is non-linear, alternative methods may need to be explored to improve cutoff estimation.

Methods like Logic Regression can also be used to identify optimal combinations of biomarkers. Logic Regression is a statistical technique that searches for the best "and-or" combinations of biomarker thresholds to maximize diagnostic accuracy. In Ingo Ruczinski's study, he stated the outperformed results with Logic Regression than other techniques like CART

and MARS with binary data modeling (Ruczinski, 2003). The study also used Logic Regression to explore associations between stroke locations (infarcts) across different brain regions and cognitive function, specifically measured by the Mini-Mental State Examination (MMSE) scores. The authors demonstrated the method's advantage in handling combinations of brain regions that might impact cognitive scores rather than looking at individual factors alone (Ruczinski, 2003). The Logic Regression had also applied to diagnosing Prostate cancer. In previous research, using a combination of TPSA and RPSA cut-offs, the diagnostic result showed a slight improvement over using TPSA alone, reducing the false-positive rate and increasing sensitivity (Etzioni, 2003) . This research indeed demonstrates the potential of Logic Regression to identify effective cut-off points for multiple biomarkers. However, the diagnosis result accuracy improvement is modest, which suggests there is room for further exploration and research into improving this method to achieve greater accuracy results.

Some machine learning approaches like Decision Tree also have been applied to disease diagnosis that can be used in this research to potentially increase the prediction accuracy. The Decision tree is more flexible and interpretable than linear combinations of biomarkers. According to Yuxin Zhu and Mei-Cheng Wang' research, they used a rank-based estimation method to search for optimal cutoff points, and Decision Tree maximizes the rank correlation between the biomarkers and the outcome. This method showed the ability to improve prediction accuracy by optimizing the cutoff points (Zhu, 2020). Another research also showed a high accurate outcome identifying the biomarker combinations. Daid Smadjah holds research in using decision tree classification to detect the subclinical keratoconus (Zhu, 2020). By applying decision tree classification, it achieved 100% sensitivity and 99.5% specificity in distinguishing between normal and KC eyes and It also showed 93.6% sensitivity and 97.2% specificity in differentiating normal eyes from those with subclinical KC (Smadja, 2013). Although the results show high scores in both sensitivity and specificity, this method may lead to a potential overfit in prediction, so the doubt of application of different dataset is questionable.

1.3 Objectives

The research mentioned above has demonstrated effectiveness in identifying both single biomarker cutoffs and combinations of biomarker cutoffs, ultimately providing more precise predictive outcomes for prostate cancer. However, determining which method or combination of

methods is optimal for identifying biomarker cutoffs with the highest accuracy remains an area for further exploration. This study will focus on the following objectives:

1. We will compare various methods for determining cutoffs, evaluating the predictive results from each, and identifying the method that shows the most potential for accurately estimating cutoffs to aid in diagnostic detection.
2. We will find the cutoff point for each biomarker and evaluate biomarker performance of prostate cancer prediction.

2. Methodology

2.2 Participants and Material

For this study, we used the dataset from the Diagnostic and Biomarkers Statistical (DABS) Center Caret PSA Dataset, which includes patient records related to the diagnosis of prostate cancer. The dataset also includes key indicators used in diagnosing the disease. Prostate-Specific Antigen (PSA), a protein produced by the prostate, is one of the main indicators tracked, as it helps monitor the progression of known prostate cancer.

The dataset consists of 683 records with 6 variables as shown in the summary table1 below: 'id,' 'd,' 't,' 'fpsa,' 'tpsa,' and 'age.' The 'id' represents the ID of each patient, and one ID may have multiple records. The variable 'd' represents the prostate cancer diagnosis, and there are 141 unique IDs, indicating that 141 patients were recorded. The variable 't' represents the time (in years) relative to the prostate diagnosis, which can have both negative and positive values. 'FPSA' is the free prostate-specific antigen level in the blood, and 'TPSA' is the total prostate-specific antigen level in the blood. 'Age' refers to the patient's age at the time of the blood draw, with recorded ages ranging from 46 to 80 years old, and an average age of 64.86.

The outcome variable is 'd,' representing the presence of prostate cancer, where 1 likely indicates the presence of prostate cancer, and 0 indicates its absence. Among the variables, two biomarkers for prostate cancer—total serum PSA and the ratio of free to total PSA—will be evaluated and used in the model training.

[----- Insert Table 1 here -----]

2.3 Design and Procedure

2.3.1 Data Preprocessing

Based on the initial statistical summary, the entire dataset would be retained for model fitting, and no additional data cleaning process was required. The dataset will be filtered with only the last record for each patient to eliminate correlations between multiple records for the same patient, which could potentially affect the model results. After filtering the dataset, there are 71 cases of prostate cancer and 70 cases without prostate cancer, resulting in a balanced outcome variable.

2.3.2 Research Design

The research aims to identify the optimal method for predicting diagnostic outcomes using both individual biomarkers and combinations of biomarkers. The study is divided into two parts based on this goal:

1. Individual Biomarkers: We will evaluate the predictive power of each biomarker separately using ROC-based methods, including the Youden Index and the point closest to (0,1) on the ROC curve.
2. Biomarker Combinations: We will assess the predictive accuracy of biomarker combinations using machine learning methods such as logic regression, logistic regression, and decision trees.

Finally, the results from all models will be compared to determine the most effective approach for predicting diagnostic outcomes.

2.2.3 Methods For Finding the Optimal Cutoff of Individual Biomarker

All the data will first be fitted using the traditional Receiver Operating Characteristic (ROC)-based method by calculating the Youden Index or the point closest to (0,1) to determine the optimal cutoff for each individual biomarker. The R package cutpointr() will be used to find individual cutoffs based on different metrics.

The Youden Index is one of the metrics used to determine cutoffs. By creating an ROC curve for each potential cutoff, the specificity and sensitivity values are obtained, and the Youden Index is calculated using the following formula:

$$J = Sensitivity + Specificity - 1 = Recall_1 + Recall_0 - 1$$

The optimal cutoff is the point where the Youden Index is maximized. This maximization indicates the point at which the combined performance of sensitivity and specificity is highest,

providing the best balance between true positive and true negative rates (American Cancer Society Journals, 2006).

Similar to the ROC-based Youden Index method, the ROC-based Point Closest to (0,1) is another method used to identify the optimal cutoff for individual biomarkers. This method also requires creating an ROC curve for each potential cutoff to determine the true positive rate and false positive rate. The method identifies the point (0,1) on the ROC curve, representing 100% sensitivity (true positive rate = 1) and 100% specificity (false positive rate = 0). For each cutoff value, the model calculates the Euclidean distance from the point (Sensitivity, 1 - Specificity) on the ROC curve to the ideal point (0,1) using the following formula:

$$Distance = \sqrt{(1 - Sensitivity)^2 + (0 - (1 - Specificity)^2)}$$

The optimal cutoff is the one that minimizes the distance to (0,1) (Hajian-Tilaki, 2018). This cutoff represents the best balance between sensitivity and specificity, offering the optimal trade-off for the diagnostic test. The results from identifying the optimal cutoffs for individual biomarkers will be compared with future model results for cutoff values derived from biomarker combinations.

2.3.4 Methods For Finding the Optimal Cutoffs of Multiple Biomarkers

The research employs three key methods to identify optimal cutoffs for multiple biomarkers: Logic Regression, Logistic Regression, and Decision Trees. Each of these techniques offers unique strengths in handling the biomarker combination, enabling a robust analysis that enhances the accurate predictions.

Logic regression is particularly well-suited for situations where combining multiple biomarkers in predicting outcomes. Unlike traditional regression methods, Logic Regression searches for combinations of conditions using logical operators (AND, OR, NOT) to predict the outcome (Ruczinski, 2003). For example, a condition might be: FPSA > 1 AND TPSA > 2, meaning under both of the conditions may increase the accuracy of outcome. The R package LogicReg() will be used to find the optimal cutoffs. By applying logic regression to the dataset, the algorithm will search for combinations of biomarkers that provide the best predictive power by combining the biomarkers using logical rules.

Logistic Regression is another useful algorithm, particularly for continuous variables, to predict a binary outcome. In logistic regression, the log-odds of the outcome are modeled as a linear combination of the biomarkers, with the formula calculated as follows:

$$log\ odds = \beta_0 + \beta_1 \times Biomarker_1 + \beta_2 \times Biomarker_2 + \ldots$$

Then the logistic function is applied to convert the log-odds into probabilities between 0 and 1 (Elkahwagy, 2024) as below:

$$Probability\ of\ Disease = \frac{1}{1+e^{-(\beta 0 + \beta 1 \times Biomarker1 + \beta 2 \times Biomarker2 + \ldots)}}$$

After fitting the Logistic Regression model, the coefficients indicate how strongly each biomarker contributes to the outcome (Elkahwagy, 2024).

Decision Tree is the other powerful machine learning method that can classify the outcome based on continuous variables. The decision tree algorithm works by recursively splitting the data into subsets based on the values of the biomarkers, and the splitting process continues until it creates a tree structure where each branch leads to a specific outcome classification (Colledani, 2023). Each split is selected to maximize the difference between the outcome groups. At each node, the tree selects a biomarker and a cutoff value that results in the best split of the data.

2.4 Model Training and Evaluation

The dataset will be split into training and testing sets, with 70% used for training and 30% for testing, and then fitted into models that identify cutoffs for multiple biomarkers. Additionally, cross-validation will be employed to validate the robustness of the cutoffs across different subsets of the data.

After fitting the models, their performance will be evaluated using metrics such as accuracy, sensitivity, specificity, and the area under the ROC curve (AUC). The ROC curve illustrates the tradeoff between sensitivity and specificity, making it an effective tool for evaluating the performance of a diagnostic test that classifies subjects into two categories, such as diseased and non-diseased (Sharma, 2014). Each model's performance will be assessed and compared, with particular attention to the comparison between the cutoffs for individual biomarkers and the cutoffs for biomarker combinations.

3. Results

3.1 Traditional Method for Identify Individual Biomarker Cutoffs

The methods were applied to individual biomarkers and found the individual optimal cutpoints. By maximizing the Youden Index, the optimal cutoff point for FPSA is 0.37 and for TPSA is 0.47, and similarly, minimizing the distance to point (0,1) on the ROC curve, the

optimal cutoff point for FPSA is 0.47 and TPSA is 2.46. For biomarker TPSA, the optimal cutpoint remains the same (2.46) across both methods, but FPSA differs significantly. In terms of accuracy score, for biomarker FPSA, the Point-Closest-to-(0,1) method achieved higher accuracy (71.01%) compared to the Youden Index Method (66.76%). For biomarker TPSA, accuracy is identical (77.01%) under both methods.

The Sensitivity and Specificity score are also crucial to identify the best method. For biomarker FPSA, the Youden Index method achieves higher sensitivity (81.66%) but lower specificity (59.25%), ROC01 method gives a significantly higher sensitivity score. The Sensitivity and Specificity scores are identical for indicator TPSA, which are 75.11% and 77.87%. As for the AUC values, they remain constant for both methods across both predictors and the high AUC values also indicate decent classification performance for both biomarkers.

From the final results, Point-Closest-to-(0,1) appeared to provide higher overall accuracy, while Youden Index maximized the trade-off between sensitivity and specificity. The higher the Youden Index results show the better the prediction result, the TPSA would give the higher Youden index result, 0.53, and the lower the distance to (0,1) on the ROC Curve indicated a better prediction result, and TPSA also gave a better result, 0.47. Those indicated the TPSA would be a better predictor than FPSA while individually predicting the diagnosis outcome.

[----- Insert Table 2 here -----]

3.2 Logic Regression

3.2.1 AND Condition

We first used ‘AND’ logic conditions to combine FPSA and TPSA, and ran loops finding all the possible combinations that can maximize the accuracy score. We use the condition (FPSA > FPSA_thredshold ∧ TPSA > TPSA_thredshold) and loop over FPSA_thredshold range from 0 to 5, by 0.01 as interval and TPSA_thredshold from 0 to 10 by 0.01 as interval. In the heatmap based on accuracy score, the highest accuracy score would be closer to yellow and the lower accuracy score would be closer to blue, which suggested that the optimal accuracy score would be where FPSA is in the (0,1) range and TPSA would be in the (2.5,4) range. After the iteration, the optimal condition would be (FPSA > 0 ∧ TPSA > 3.61), and this would achieve an accuracy score of 82.26%.

[----- Insert Figure 1 here -----]

3.2.2 OR Condition

Similar to the 'AND' Condition, the 'OR' condition would be (FPSA > FPSA_thredshold ∨ TPSA >TPSA_thredshold), and the loop over the same threshold range as 'OR' condition. The heatmap made it harder to identify the optimal range. After all the interactions, the optimal condition would give an 82.97% accuracy score, with the condition (FPSA > 1.10 ∨ TPSA >3.61).

[----- Insert Figure 2 here -----]

3.2.3 AND, OR Combination Condition

Based on the previous research associated with the two biomarkers, we found the range of TPSA would be more essential than the FPSA. After researching several condition combinations, we have the optimal condition combination, which was (TPSA ≥ TPSA_thredshold_lower ∧ TPSA < TPSA_thredshold_upper ∧ FPSA > FPSA_thredshold) ∨ (TPSA > TPSA_thredshold_upper). We iterate the FPSA threshold in range (0, 0.5) by 0.01 interval, and TPSA threshold in range (0, 10) by 0.01 interval. Based on the highest accuracy score 79.94%, the final logic equation would be ((TPSA ≥ 3.21 ∧ TPSA < 6.27) ∧ FPSA > 0) ∨ (TPSA > 6.27).

[----- Insert Figure 3 here -----]

3.3 Logistic Regression

The Logistic Regression was fitted with two continuous variables, TPSA and FPSA, on the train data, and then tested on the test data. The model was shown as following:

$$y = -0.18FPSA + 0.23TPSA - 0.87$$

The model summary showed that the p-value of FPSA is approximately 0.13, which is greater than 0.05, indicating that FPSA was not statistically significant at the 0.05 level of significance. In contrast, the p-value of TPSA was 0.006, making it statistically significant. This aligns with previous findings, demonstrating that TPSA had a stronger impact on the outcome than FPSA.

[----- Insert Table 3 here -----]

The results of the Logistic Regression model are robust, with a high sensitivity of 91.3% and specificity of 89.5%, indicating its ability to effectively identify both positive and negative cases. With an accuracy of 88.09% and a high AUC score from the ROC curve, the model showed excellent ability to distinguish between the positive and negative classes. After fitting the

logistic Regression, we converted the predicted continuous outcome into binary based on the optimal cutoff, 0.44, by maximizing the Youden Index. Comparing the converted binary outcome with the actual labels, the confusion matrix reflected a balanced classification, with only 2 false positives and 3 false negatives. The results indicated the logistic regression performed excellently in predicting the classification.

[----- Insert Figure 4 here -----]

[----- Insert Figure 5 here -----]

3.4 Decision Tree

The decision tree model predicted outcome (d) also based on two variables, TPSA and FPSA. After hyperparameter tuning, we set the minimum number of observations in a node to 10, the complexity parameter to 0.01 and the maximum depth of the tree to 4. The first split is on TPSA < 2.8, which divides the dataset into two branches: one dominated by d = 0 (negative class) and another split into subgroups based on TPSA and FPSA thresholds. Based on the Decision tree results, nodes with TPSA < 7.2 and FPSA >= 0.53 played more significant roles in classifying observations into positive classes.

[----- Insert Figure 6 here -----]

The model achieved an accuracy of 83.3% on the test data, showing reasonable predictive power. The ROC curve demonstrated the model's good discriminatory capability, with a high AUC value of 89.90%, which indicated the high distinguishing ability. The darker blue color in the confusion matrix highlighted 17 true negatives and 18 true positives, and the lighter blue color showed 4 false positives and 3 false negatives, which suggested overall good performance in predicting outcomes.

[----- Insert Figure 7 here -----]

[----- Insert Figure 8 here -----]

4. Discussion and Conclusion

The research investigated the optimal method for identifying cutoff points of biomarker combinations to predict final censoring outcomes. While much past research focuses on single biomarkers, the results demonstrated that using biomarker combinations improved prediction accuracy compared to individual biomarkers. Compared with all the methods, Logistic Regression showcases the strongest ability to distinguish the diseased and non-diseased groups,

with high accuracy, sensitivity, and specificity scores. After researching all the methods, we concluded that biomarker TPSA would be a more significant predictor of the diagnosis outcome than biomarker FPSA.

Despite the interesting findings in the current study, some limitations must be addressed. First, the dataset consists of 683 data points, and after keeping one record for individual patients, 141 records were used. After splitting 70% into the training set and 30% into the test set, the data size became even smaller. Also, The small data size may limit the results with overfitting, resulting in poor performance of reflecting a broader population (Bailly et al., 2022). With more data, the cutoff point would be more precise and accurate without bias. For example, in this Caret PSA dataset, some severe patients with extremely high TPSA or FPSA values would impact the model's performance in finding the precise cutoff points.

Secondly, the models are solely applied to the DABS Center Caret PSA Dataset, so the application results to other datasets or diseases are uncertain. With the small dataset issues mentioned above, the models were likely to overfit the given dataset with poor performance in other new datasets. Moreover, the dataset used was balanced with 71 cases of prostate cancer and 70 controls, which might be different from the actual situation with mostly imbalanced data. Thirdly, the study did not involve the time effect on the diagnosis outcome and only kept the last records of patients. However, the time effect can be also important and impacts the biomarker level change. The study only considered two biomarkers but did not include other biological variables. The dataset only included the feature with patient age, but other biological features that may cause the biomarker level differences should also be considered as factors that impact the diagnosis outcome.

This research indeed highlighted biomarker TPSA's strong diagnostic potential, in the future, more biomarkers could be combined with TPSA beyond FPSA to provide possible more accurate prediction results. Also, some features other than biomarkers can be included in the machine learning model training to predict the censoring outcome, like age, weight, height, family history, geographic information, etc. Furthermore, the results showed less contribution of biomarker FPSA for prostate cancer detection, thus other biomarker levels can be considered in future datasets with prostate cancer diagnosis. Future research could explore the Logistic Regression on different datasets of other diseases to see the model performance and outcomes.

Even the biomarker cutoff shows a strong ability to detect the disease, however, the application of machine learning to determine the biomarker cutoff point that can be used in medical diagnosis is still a long way to go. Finding precise biomarker cutoffs is meaningful and meanwhile challenging for better accurate diagnosis outcomes. The biomarker cutoff can set up an alert for patients early to detect their disease and get treatment accordingly. The machine learning models help find the biomarker combination cutoffs which can enhance diagnostic outcomes and help patients become aware of their condition at an earlier stage.

References

American Cancer Society Journals. (2006). *Index for rating diagnostic tests*. Wiley Online Library.

Baker, S. G. (2011). Logistic regression analysis of biomarker data subject to pooling and dichotomization. *Statistics in Medicine, 30*(25), 2901-2913. https://doi.org/10.1002/sim.4367

Bailly, A., Blanc, C., Francis, É., Guillotin, T., Jamal, F., Wakim, B., & Roy, P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine, 213*, 106504. https://doi.org/10.1016/j.cmpb.2021.106504

Chen, F., Xue, Y., Tan, M. T., & Chen, P. (2015). Efficient statistical tests to compare Youden index: Accounting for contingency correlation. *Statistics in Medicine, 34*(9), 1560–1576. https://doi.org/10.1002/sim.6432

Colledani, D., Anselmi, P., & Robusto, E. (2023). Machine learning-decision tree classifiers in psychiatric assessment: An application to the diagnosis of major depressive disorder. *Psychiatry Research, 322,* 115127. https://doi.org/10.1016/j.psychres.2023.115127

Elkahwagy, D. M. A. S., Kiriacos, C. J., & Mansour, M. (2024). Logistic regression and other statistical tools in diagnostic biomarker studies. *Clinical and Translational Oncology, 26*(10), 2172-2180. https://doi.org/10.1007/s12094-024-03413-8

Etzioni, R., Kooperberg, C., Pepe, M., Smith, R., & Gann, P. H. (2003). Combining biomarkers to detect disease with application to prostate cancer. *Biostatistics, 4*(4), 523–538. https://doi.org/10.1093/biostatistics/kxg010

Etzioni, R., Pepe, M., Longton, G., Hu, C., & Goodman, G. (1999). Incorporating the time dimension in receiver operating characteristic curves: A case study of prostate cancer. *Medical Decision Making, 19*(2), 242-251. https://doi.org/10.1177/0272989X9901900212

Hajian-Tilaki, K. (2018). The choice of methods in determining the optimal cut-off value for quantitative diagnostic test evaluation. Statistical methods in medical research, 27(8), 2374-2383.

Sharma, B., & Jain, R. (2014). Right choice of a method for determination of cut-off values: A statistical tool for a diagnostic test. *Asian Journal of Medical Sciences, 5*(3), 30-34.

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians, 66*(1), 7–30. https://doi.org/10.3322/caac.21332

Smadja, D., Touboul, D., Cohen, A., Doveh, E., Santhiago, M. R., Mello, G. R., Krueger, R. R., & Colin, J. (2013). Detection of subclinical keratoconus using an automated decision tree classification. *American Journal of Ophthalmology, 156*(2), 237-246. https://doi.org/10.1016/j.ajo.2013.03.034

Srivastava, S. (Ed.). (2017). *Biomarkers in cancer screening and early detection*. John Wiley & Sons, Incorporated. ProQuest Ebook Central. http://ebookcentral.proquest.com/lib/georgetown/detail.action?docID=4883061

Ruczinski, I., Kooperberg, C., & LeBlanc, M. (2003). Logic regression. *Journal of Computational and Graphical Statistics, 12*(3), 475-511. https://doi.org/10.1198/1061860032238

Zhu, Y., & Wang, M.-C. (2020). Obtaining optimal cutoff values for tree classifiers using multiple biomarkers. *Biometrics, 78*(1), 128-140. https://doi.org/10.1111/biom.13409

Appendix

Variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
id	1.00	39.00	73.00	72.69	107.00	141.00
d	0.00	0.00	0.00	0.34	1.00	1.00
t	-9.01	-3.73	-1.42	-1.37	0.83	7.42
FPSA	0.00	0.25	0.41	0.83	0.70	100.00
TPSA	0.03	1.05	1.80	4.80	3.96	99.98
age	46.75	61.07	65.16	64.86	69.00	80.83

Table1. Statistical Summary

Biomarker	Index	Method	Optimal Cutpoint	Index Results	Accuracy	Sensitivity	Specificity	AUC
FPSA	Youden Index	maximize_metric	0.37	0.41	66.76%	81.66%	59.25%	77.35%
	Distance to (0,1)	minimize_metric	0.47	0.42	71.01%	68.56%	72.25%	77.35%
TPSA	Youden Index	maximize_metric	2.46	0.53	77.01%	75.11%	77.97%	83.75%
	Distance to (0,1)	minimize_metric	2.46	0.33	77.01%	75.11%	77.97%	83.75%

Table 2: Results of Youden Index and ROC-Point-Closest-to-(0,1) Methods

Term	Estimate	Std. Error	z-value	Pr(>|z|)
(Intercept)	-0.87	0.33	-2.61	0.00904
FPSA	-0.18	0.12	-1.53	0.12933
TPSA	0.23	0.08	2.75	0.00588

Table 3: Summary of Logistic Regression Model

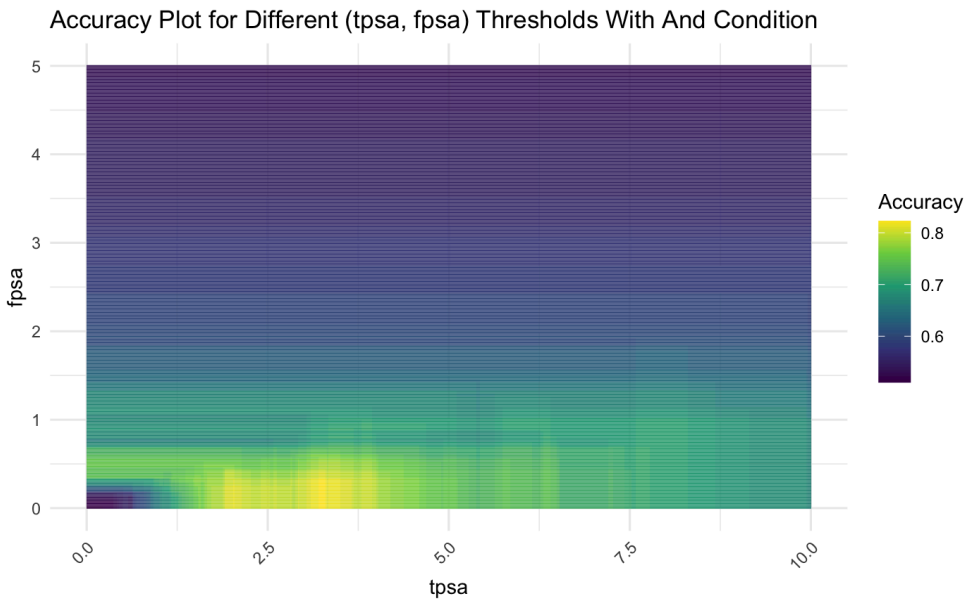

Figure 1: Heatmap with Different TPSA and FPSA Combination for AND condition

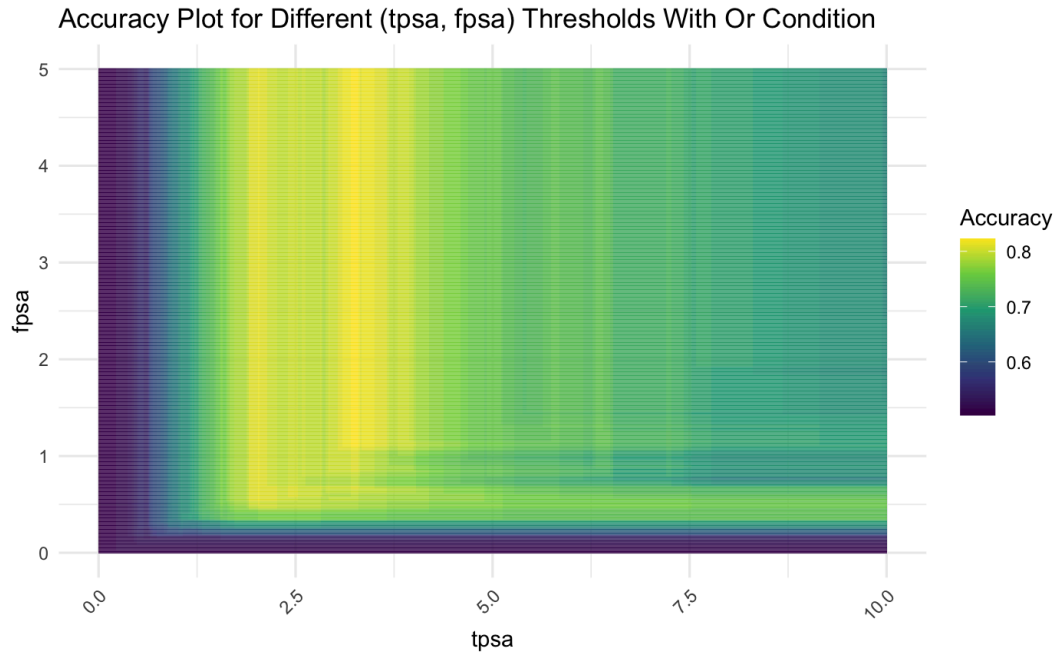

Figure 2: Heatmap with Different TPSA and FPSA Combination for OR Condition

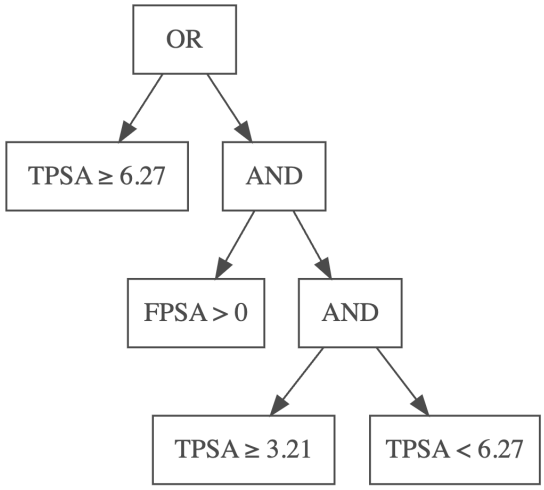

Figure 3: Logic Tree with Different TPSA and FPSA Combination for OR Condition

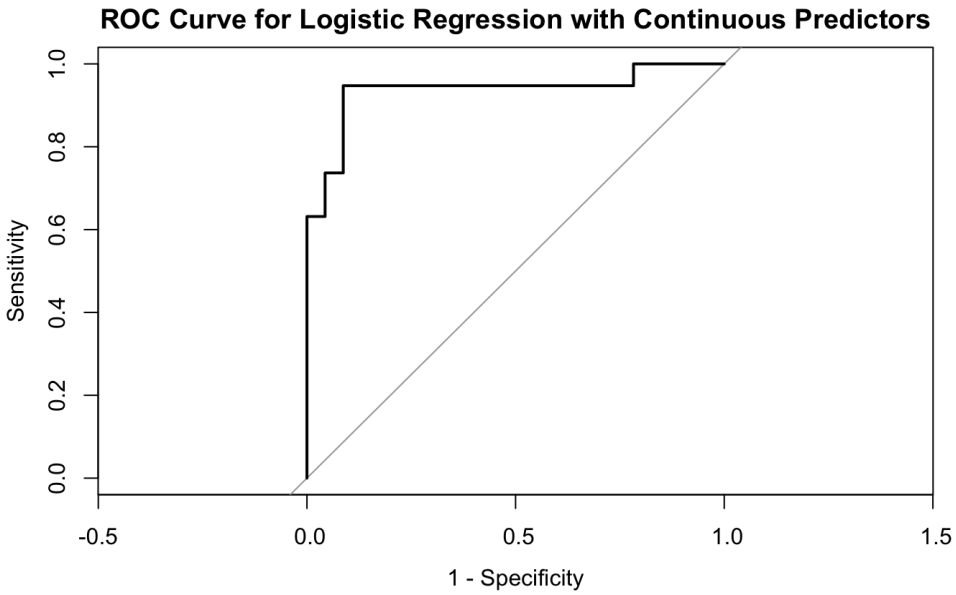

Figure 4: Logistic Regression ROC Curve

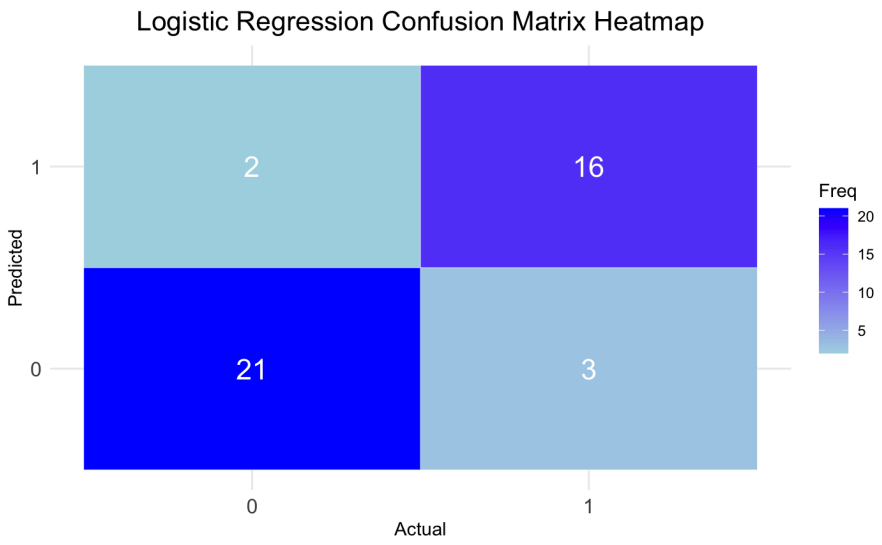

Figure 5: Logistic Regression Confusion Matrix Heatmap

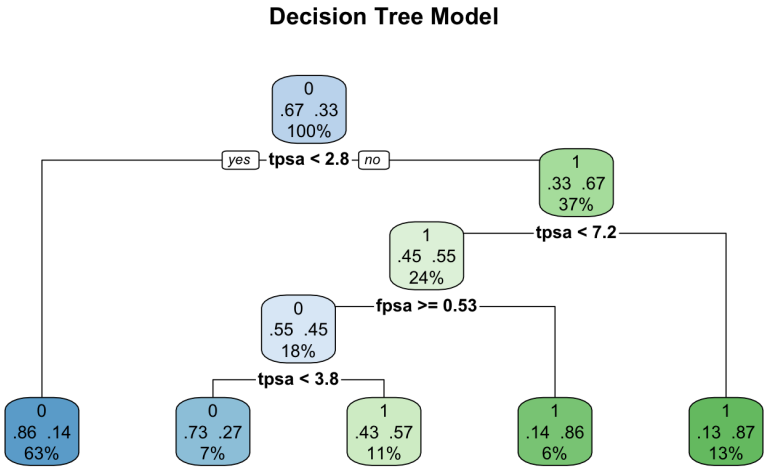

Figure 6: Decision Tree Model

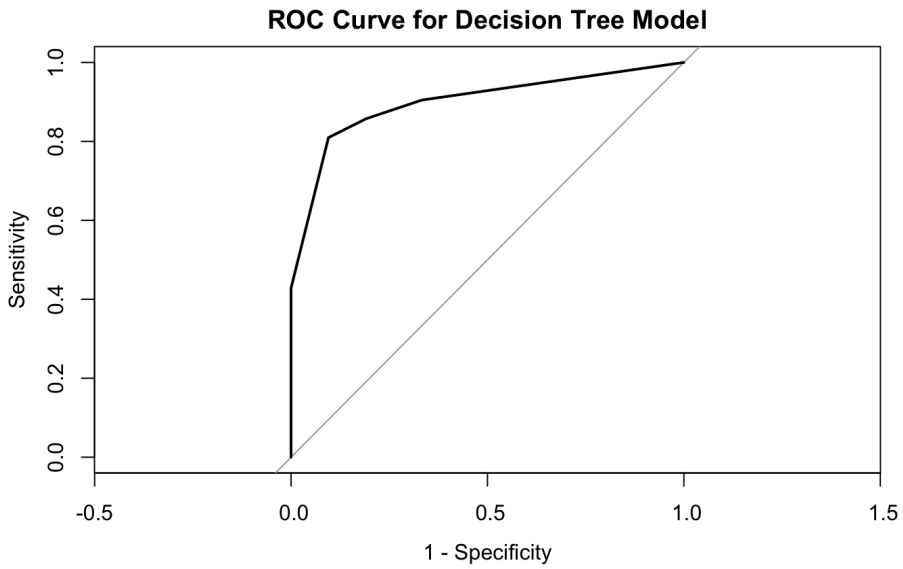

Figure 7: Decision Tree ROC Curve

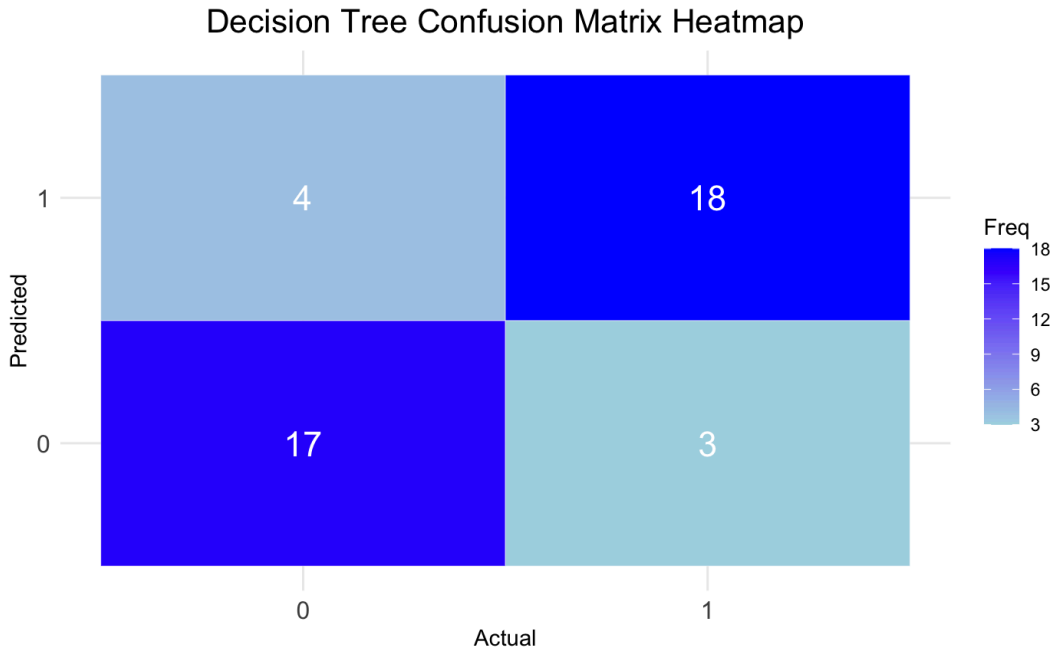

Figure 8: Decision Tree Confusion Matrix Heatmap