Received: 13 June 2024 | Revised: 8 August 2024 | Accepted: 16 August 2024

DOI: 10.1111/cas.16327

ORIGINAL ARTICLE

Cancer Science WILEY

Prostate cancer treatment recommendation study based on machine learning and SHAP interpreter

Shengsheng Tang[1] | Hongzheng Zhang[1] | Junhao Liang[1] | Shishi Tang[1] | Lin Li[1] | Yuxuan Li[1] | Yuan Xu[2] | Daohu Wang[3] | Yi Zhou[1]

[1]Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, Guangdong, China

[2]Medical Big-Data Center, The Second Affiliated Hospital of Nanchang University, Nanchang, Jiangxi, China

[3]Department of Urology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong, China

Correspondence
Yi Zhou, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, Guangdong 510080, China.
Email: zhouyi@mail.sysu.edu.cn

Daohu Wang, Department of Urology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong 510080, China.
Email: wangdaoh@mail.sysu.edu.cn

Funding information
the Key Research and Development Program of China, Grant/Award Number: 2022YFC3601600; the Guangdong Natural Science Foundation, Grant/Award Number: 2024A1515011989; the Guangzhou Science and Technology Plan Project, Grant/Award Number: 202206010028; the Fundamental Research Funds for the Central Universities of Sun Yat-sen University, Grant/Award Number: 24xkjc025; the Key Research and Development Programs of Jiangxi Provincial Department of Science and Technology, Grant/Award Number: 20202BBGL73074

Abstract

This study utilized data from 140,294 prostate cancer cases from the Surveillance, Epidemiology, and End Results (SEER) database. Here, 10 different machine learning algorithms were applied to develop treatment options for predicting patients with prostate cancer, differentiating between surgical and non-surgical treatments. The performances of the algorithms were measured using the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value, negative predictive value. The Shapley Additive Explanations (SHAP) method was employed to investigate the key factors influencing the prediction process. Survival analysis methods were used to compare the survival rates of different treatment options. The CatBoost model yielded the best results (AUC = 0.939, sensitivity = 0.877, accuracy = 0.877). SHAP interpreters revealed that the T stage, cancer stage, age, cores positive percentage, prostate-specific antigen, and Gleason score were the most critical factors in predicting treatment options. The study found that surgery significantly improved survival rates, with patients undergoing surgery experiencing a 20.36% increase in 10-year survival rates compared with those receiving non-surgical treatments. Among surgical options, radical prostatectomy had the highest 10-year survival rate at 89.2%. This study successfully developed a predictive model to guide treatment decisions for prostate cancer. Moreover, the model enhanced the transparency of the decision-making process, providing clinicians with a reference for formulating personalized treatment plans.

KEYWORDS
machine learning, prostate cancer, SEER database, SHAP interpreter, survival analysis

Abbreviations: CPP, cores positive percentage; Grade, cancer grade; GS_sec, clinical secondary Gleason grade; M, metastasis stage; N, node stage; PG_pri, clinical primary Gleason grade; PSA, prostate-specific antigen; Stage, cancer stage; T, tumor stage.

Shengsheng Tang and Hongzheng Zhang contributed equally to this work.

1 | INTRODUCTION

Prostate cancer is a malignant tumor that is prevalent in men worldwide. GLOBOCAN 2022[1–4] reported that it is considered a significant public health issue, with the second-highest incidence and fifth-highest mortality rates among men.

Clinical research and practical experience have demonstrated that surgical treatment can prevent or delay tumor recurrence and progression, leading to a significant improvement in the disease-free and overall survival rates of patients. Besides surgical treatment, there are also some non-surgical options that can help, such as radiation therapy (external radiation radiotherapy and brachytherapy), chemotherapy, hormone therapy, and targeted drug therapy.[5,6] It is imperative to choose the most suitable treatment options for prostate cancer patients because of the wide range of disease progression stages, individual physiological conditions, and quality of life aspirations. The decision has a direct impact on disease control and prognosis, as well as on minimizing healthcare resource consumption and alleviating patient burden in daily life and the economy.

In recent years, there have been significant advances in prostate cancer diagnosis and treatment research. Machine learning technology has demonstrated excellent potential in personalized treatment plan design.[7–12] By systematically considering individual factors such as age, health status, and lifestyle, machine learning models aid researchers in discerning subtle differences in tumor characteristics among prostate cancer patients. This provides strong technical support for creating personalized treatment plans. Despite the use of machine learning techniques in some studies to propose personalized treatment plans for prostate cancer,[13–18] this field still struggles with several challenges. First, small sample sizes and data sources being concentrated in individual research centers have generally resulted in the weakening of the representativeness and generalizability of past studies. Second, the variables used in past studies are not comprehensive in terms of dimension and breadth.[14,15] This limitation affects the predictive performance of the model to a certain extent, which in turn affects the reliability of the experimental results and the understanding of the model of complex treatment decisions.[17] Finally, machine learning models currently in use tend to rely on a single or limited combination of types. Comprehensive comparisons of multivariate models are lacking, which may restrict the generalizability and adaptability of the models necessary in real and complex clinical settings.[14,16]

This study aims to construct a more accurate and reliable prostate cancer treatment recommendation model by using a large-scale, multicenter dataset of prostate cancer patients, combined with machine learning algorithms and Shapley Additive Explanations (SHAP) interpreter technology.[19] Additionally, we employ the SHAP algorithm to investigate the impact of various variables on the choice of treatment methods and assess the prognostic survival rates of different therapies. The research provides important references for clinical decision-making, promoting the optimization of treatment effects and improving the quality of life for prostate cancer patients.

2 | MATERIALS AND METHODS

2.1 | Data source and study population

This study was based on the Surveillance, Epidemiology, and End Results (SEER) database, which is maintained by the National Cancer Institute in the United States. Previous studies and clinical treatment guidelines[20–23] have pointed out that the patient's age, race, cancer stage, Gleason score, TNM stage, PSA level, etc. are all important indicators that affect the choice of treatment methods. Therefore, this study used SEER*Stat software to extract relevant demographic information and clinical characteristics records of patients initially diagnosed with primary malignant prostate cancer from the SEER database between 2010 and 2015.

We developed a series of inclusion criteria to ensure the quality and applicability of the included data: (1) All included cases were confirmed as prostate cancer based on the pathological diagnosis. (2) It contains complete basic patient information, covering a range of core clinical variables, such as age, gender, and detailed grading information about cancer. (3) The study was restricted to males aged 40 years and above.

After filtering the data and excluding missing values, the characteristic variables are obtained:

(1) Demographic information, including age at first diagnosis, total annual household income, total population at home and ethnicity.

(2) Clinical characteristics, including tumor histological type, overall cancer stage, cancer grading, Gleason score (major and minor scores), TNM stage (T stage indicating primary tumor status, N-stage indicating lymph node metastasis, and M stage indicating distant metastasis), PSA level, and the proportion of prostate cancer cells in the biopsy tissue (core positive proportion).

(3) Outcome characteristics, including the specific treatment modality received by the patient, the actual survival time, and the survival status. Figure 1 shows the process of establishing a prostate cancer treatment decision model and the principles of survival analysis.

2.2 | Data preprocessing and feature selection

For more detailed and targeted data analysis, the categorical variables were analyzed using code conversion. Adult male subjects were classified into 10 continuous age groups based on their actual age at diagnosis, which ranged from 40 to 44 years to 85 years and older. We divide annual household income into 10 levels, with the lowest level below $35,000 and the highest level being $75,000 and above. The study subjects' regions were classified into five major categories based on different population sizes: mega-urban areas (population over 1 million), large urban areas (population between 250,000 and 1 million), medium urban areas (population less than 250,000), near-urban non-urban areas, and distant non-urban areas. In addition, the subjects were divided into three main racial categories: White, Black, and Other.

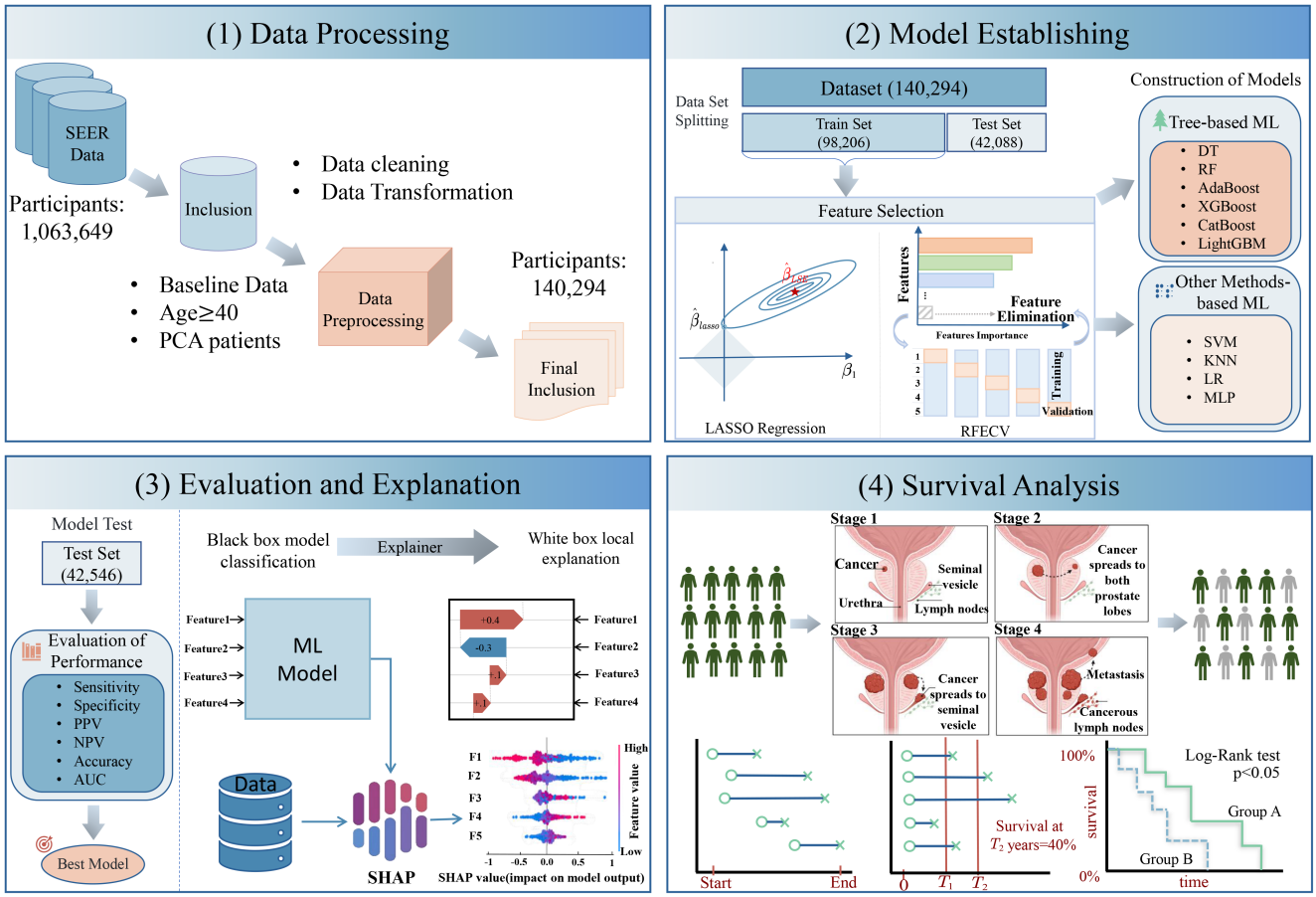

FIGURE 1 Framework diagram. AUC, area under the receiver operating characteristic curve; F, feature; LASSO, least absolute shrinkage and selection operator; NPV, negative predictive value; PPV, positive predictive value.

In defining the clinical characteristics of prostate cancer patients, we adopted the criteria published by the American Joint Committee on Cancer (AJCC 7th edition, 2009) for unified classification and description. This study aimed to predict treatment options for prostate patients based on their specific characteristics, differentiating between surgical and non-surgical treatments. Surgical treatment included radical prostatectomy and cryotherapy.

We employed two methods, least absolute shrinkage and selection operator (LASSO) regression and recursive feature elimination with cross-validation (RFECV), for feature selection to ensure the rationality of the selected features. LASSO regression introduces an L1 regularization penalty term on top of the classical linear regression model, which can drive certain less important feature coefficients close to zero, thereby achieving feature selection and dimensionality reduction. The advantage of this method is its ability to identify and eliminate redundant features with low contribution to prediction performance.

RFECV iteratively removes features deemed least important and utilizes cross-validation to evaluate the performance of the selected feature subsets at each step, determining the optimal number of features. The primary advantage of RFECV is its automated feature selection process, which reduces the subjectivity of feature selection and enhances the accuracy and generalization capability of the model.

2.3 | Model establishment and performance

In this study, we selected 10 well validated supervised machine learning algorithms that have proven utility in medical research.[10,24–29] These algorithms encompass linear, tree-based, ensemble, and neural network models, including support vector machine (SVM), k-nearest neighbors (KNN), decision tree (DT), logistic regression (LR), random forest (RF), adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), categorical boosting (CatBoost), and multi-layer perceptron (MLP). Our goal was to address the limitations of model simplicity in current research by comprehensively exploring data characteristics and capturing complex relationships.

The KNN algorithm provides an easily understandable and implementable solution for classification tasks, owing to its simple and intuitive principles and robustness to outliers. SVM excel in handling small sample sizes and high-dimensional data, particularly in nonlinear problems, by utilizing kernel techniques to efficiently

delineate decision boundaries. LR is renowned for its practical interpretability and model transparency, often serving as a baseline model. Ensemble learning strategies, such as RF and Gradient Boosting Machine, are highly regarded for their excellent performance and robustness on complex datasets. By combining the predictions of multiple weak learners, these methods significantly enhance model accuracy and generalization capability. Additionally, MLP leverages deep neural network architectures to automatically learn complex patterns within data, thereby improving classification accuracy.

These algorithms were programmed using Python (version 3.10.9) and implemented with open-source machine learning libraries (scikit-learn). The dataset was randomly shuffled and split into training and test sets in a 7:3 ratio. The parameters of the ML methods were determined through grid search, where each algorithm's performance was repeatedly evaluated by setting parameters to different values within a feasible range to identify those that yielded the best results (Table S1). The evaluation metrics included: Area Under the Curve (AUC), sensitivity, specificity, accuracy, negative predictive value (NPV), and positive predictive value (PPV).

2.4 | Model explanation and survival analysis

We analyzed and discussed the model with the highest overall performance using an interpretive approach based on SHAP to uncover its predictive mechanisms and feature contributions. SHAP theory, derived from cooperative game theory, provides a rigorous and highly interpretable tool capable of quantifying the specific influence and relative importance of each feature on the model's prediction outcomes. Subsequently, Kaplan–Meier survival analysis was used to evaluate the effects of different treatment methods on the survival time of prostate cancer patients.

The study presents continuous variables as means and standard deviations and uses the *t*-test or Mann–Whitney test to determine differences between groups. Categorical variables are expressed as numbers and percentages, and the chi-squared test or Fisher's exact test is used for group comparisons. The statistical significance was determined by $p < 0.05$.

3 | RESULTS

3.1 | Baseline characteristics

In this study, we include a total of 140,294 patients with prostate cancer for detailed retrospective analysis. Based on the treatment chosen, the study population was divided into two groups. The surgical treatment group accounted for 35% of the total sample volume and included patients who had undergone radical prostatectomy, cryotherapy of the prostate. The non-surgical treatment group comprises the remaining 65% and includes individuals managed with endocrine therapy, radiation therapy, and active surveillance.

Table 1 shows the characterization of patients with prostate cancer treated surgically versus non-surgically. Compared with the patients treated without surgery, the patients in the surgery group are concentrated in their age of about 60 years old, with a lower PSA level, and most of them show low-to-medium grade characteristics in terms of tumor grade, stage, tumor size, and lymph node status.

3.2 | Feature extraction

Figure 2A,B illustrates the regularization path diagram and the mean squared error curve of the LASSO regression model. It was observed that, when the logarithm of λ equaled −4.0, the model's mean squared error reached its minimum level, indicating the effectiveness of including all 14 features in model construction.

Furthermore, Figure 2C presents the results of the RFECV method for feature selection. This method utilized five-fold cross-validation based on an RF classifier, using the average AUC as the evaluation criterion to automatically select the optimal number of features. As the number of features increases, Figure 2C demonstrates an upward trend in the model's average AUC, thus reaffirming the importance of these 14 features in constructing a high-performance predictive model.

3.3 | Performance analysis

Our results showed that XGBoost, LightGBM and CatBoost have the highest AUC values, achieving a combined 0.939 as an optimal performance metric. In terms of sensitivity as an evaluation criterion, SVM, AdaBoost and LR demonstrated the ability to correctly identify true positives, achieving rates of 0.931, 0.925, and 0.902, respectively. Conversely, in terms of specificity, CatBoost and RF emerged as superior performers (Table 2).

To further compare the performance differences between the XGBoost, LightGBM, and CatBoost models, we adjusted the number of weak classifiers and analyzed their AUC results on the test set and their scores during the 10-fold cross-validation phase (Figure 3). Our study found that, as the number of weak classifiers was increased, both LightGBM and XGBoost experienced a significant decline in AUC values on the validation set, indicating a potential overfitting issue in these models. In contrast, CatBoost's scores remained stable, suggesting its strong generalization ability to better handle unknown data. Furthermore, in the 10-fold cross-validation score graph, XGBoost and CatBoost performed similarly when the number of weak classifiers was less than 100. However, when the number of weak classifiers exceeded 100, CatBoost's scores were significantly higher than the other two methods. Based on both generalization capability and model accuracy, we concluded that CatBoost was the most effective option.

TABLE 1 Characteristics of prostate cancer patients undergoing surgical and non-surgical treatments.

Characteristic	Surgery[a], N = 49,394	No surgery[a], N = 90,900	p-value[b]
Age group, years			
40–44	458 (0.77%)	235 (0.26%)	<0.001
45–49	1989 (3.34%)	1279 (1.41%)	
50–54	6136 (10.32%)	4676 (5.14%)	
55–59	10,258 (17.25%)	10,522 (11.58%)	
60–64	12,451 (20.92%)	16,781 (18.46%)	
65–69	11,963 (20.11%)	21,792 (23.97%)	
70–74	4911 (8.26%)	17,419 (19.16%)	
75–79	1058 (1.78%)	11,401 (12.54%)	
80–84	148 (0.25%)	4807 (5.29%)	
85+	22 (0.04%)	1988 (2.19%)	
Median annual household income, $			
<35,000	670 (1.13%)	1709 (1.88%)	<0.001
35,000–39,999	1207 (2.03%)	2834 (3.12%)	
40,000–44,999	2012 (3.38%)	4232 (4.66%)	
45,000–49,999	3069 (5.16%)	5788 (6.37%)	
50,000–54,999	4713 (7.93%)	7611 (8.37%)	
55,000–59,999	4037 (6.78%)	7617 (8.38%)	
60,000–64,999	9680 (16.28%)	16,171 (17.79%)	
65,000–69,999	5769 (9.70%)	10,862 (11.95%)	
70,000–74,999	3614 (6.08%)	6198 (6.82%)	
70,000+	14,623 (24.57%)	27,878 (30.67%)	
Residential population (million)			
Metro (pop > 1)	28,392 (57%)	54,206 (60%)	<0.001
Metro (0.25 < pop < 1)	11,527 (23%)	19,629 (22%)	
Metro (pop < 0.25)	3914 (7.9%)	6931 (7.6%)	
Nonmetro (adjacent)	3171 (6.4%)	6071 (6.7%)	
Nonmetro (not)	2390 (4.8%)	4063 (4.5%)	
Race/ethnicity			
White	40,149 (81%)	68,175 (75%)	<0.001
Black	6158 (12%)	15,550 (17%)	
Other races	3087 (6.3%)	7175 (7.9%)	
Histological type			
Adenoma	49,373 (100%)	90,745 (100%)	<0.001
Neoplasm	13 (<0.1%)	58 (<0.1%)	
Carcinoma	8 (<0.1%)	97 (0.1%)	
Clinical grade			
I	2441 (4.9%)	13,007 (14%)	<0.001
II	19,495 (39%)	38,606 (42%)	
III	27,406 (55%)	39,175 (43%)	
IV	52 (0.1%)	112 (0.1%)	
Clinical stage			
In situ	0 (0%)	2 (<0.1%)	<0.001
Localized	32,018 (65%)	82,295 (91%)	
Regional	17,229 (35%)	3160 (3.5%)	
Distant	147 (0.3%)	5443 (6.0%)	
Tumor stage			
T0	0 (0.0%)	1 (<0.1%)	<0.001
T1	748 (1.52%)	62,243 (68%)	
T2	34,529 (70.23%)	24,491 (27%)	
T3	13,932 (28.32%)	2931 (3.2%)	
T4	185 (0.38%)	868 (1.0%)	
TX	0 (0.0%)	366 (0.4%)	
Node stage			
N0	47,052 (95%)	84,980 (93%)	<0.001
N1	2078 (4.2%)	2653 (2.9%)	
NX	264 (0.5%)	3267 (3.6%)	
Metastasis stage			
M0	49,257 (99.72%)	85,496 (94%)	<0.001
M1	137 (0.28%)	5404 (5.9%)	
Clinical primary Gleason grade			
1	1 (<0.1%)	5 (<0.1%)	<0.001
2	27 (<0.1%)	131 (0.1%)	
3	34,579 (70%)	60,317 (66%)	
4	14,016 (28%)	26,848 (30%)	
5	771 (1.6%)	3599 (4.0%)	
Clinical secondary Gleason grade			
1	1 (<0.1%)	0 (0%)	<0.001
2	23 (<0.1%)	113 (0.1%)	
3	25,307 (51%)	49,965 (55%)	
4	21,149 (43%)	32,203 (35%)	
5	2914 (5.9%)	8619 (9.5%)	
PSA level			
Mean (SD)	8.586 (9.505)	14.819 (21.530)	<0.001
Core positive proportion			
Mean (SD)	0.404 (0.262)	0.404 (0.299)	<0.001

Abbreviations: Mean, mean value; pop, population; PSA, prostate-specific antigen; SD, standard deviation.

[a]*n* (%).

[b]Wilcoxon rank sum test; Pearson's chi-squared test.

3.4 | Explanatory analysis

In this study, we utilized SHAP tools to generate bee swarm plots and bar charts for three models with identical AUC values (XGBoost, LightGBM, and CatBoost) (Figure S1), providing a clear and detailed explanation of the models' prediction processes. Radar chart (Figure 4, left) illustrates the relative impact of different features on

the prediction outcomes of the three models. The bee swarm plots on the right side of Figure 4 illustrate the impact of different feature values on model prediction outputs.

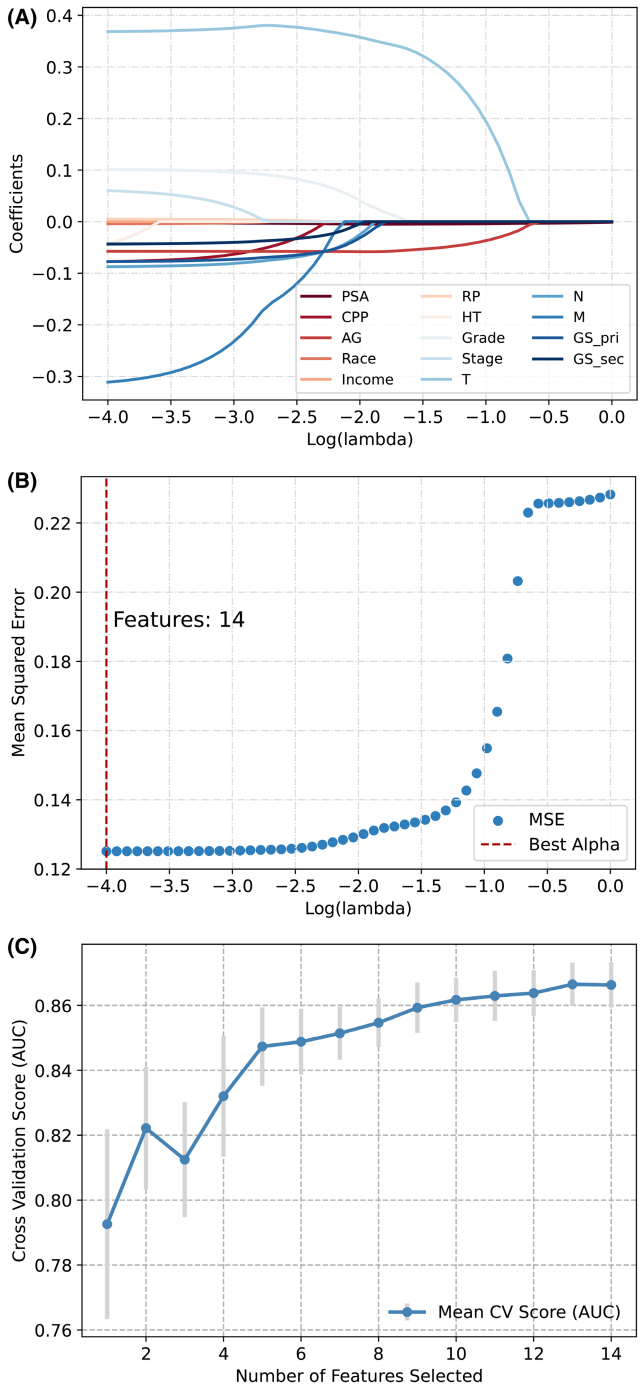

FIGURE 2 Feature selection using least absolute shrinkage and selection operator (LASSO) regression in the training set. (A) Distribution of LASSO coefficients for 14 features. In the LASSO algorithm, the trajectory of each relevant feature coefficient is observed in the LASSO coefficient plot as lambda varies. (B) Conducting 10-fold cross-validation to select the optimal parameter (lambda) for the LASSO model. (C) Five-fold cross-validation, the number of optimal features of the recursive feature selection cross-validation (RFECV) method.

The study found that T stage (mean SHAP value of 1.72), cancer stage (mean SHAP value of 0.62), age (mean SHAP value of 0.47), and CPP (mean SHAP value of 0.46) had a dominant influence on the models' predictive performance. Further analysis revealed that lower levels of T stage, cancer stage, and cancer grade led to predictions favoring non-surgical treatments. Conversely, higher values of CPP, age, PSA, and Gleason score led to predictions favoring non-surgical treatments. This aligns with the standard recommendations of major guidelines that use conventional clinicopathological characteristics (clinical stage, cancer grade, Gleason score, etc.) to guide treatment,[30,31] validating the feasibility of our models.

Notably, in our study, the impact of PSA and Gleason score on treatment prediction was significantly lower than that of the T stage and cancer stage. This contrasts with existing research, which identifies PSA and Gleason score as independent predictors of metastatic disease progression in patients with biochemical recurrence,[22,32,33] especially those with higher PSA levels and Gleason scores, who may have a biochemical recurrence rate exceeding 50% within 5 years.[34] This discrepancy may be due to differences in our dataset, experimental design, and prediction objectives.

3.5 | Survival analysis

The survival curves established for each age group (Figure 5) reveal the impact of different treatment methods on the survival prognosis of patients at different ages. Specifically, the surgical treatment group performed well in the 5-year and 10-year survival rates, which are detailed in Table 3. Through a detailed comparative analysis of these survival curves, we revealed the following key conclusions:

The survival rate of the non-surgical treatment group was significantly lower than those who had undergone surgical treatments in the comparative study of survival rates for prostate cancer treatment plans. Across all age groups, patients who had undergone radical prostatectomy demonstrated the highest survival rates in both 5 and 10 years. The survival rates of patients who had undergone cryotherapy were similar to those who had not undergone surgery.

Cryotherapy was found to have a greater survival benefit for elderly prostate cancer patients, particularly those aged 75 and above, compared with other surgical methods. This suggested that this therapy might have a better survival-improving effect in older patients.

With age, the survival curve for all treatments decreased. However, to fully interpret these survival curve data, we need to fully consider many potential confounding variables including, but not limited to, the patient's general health status, the stage and grade of the tumor, and other factors.

4 | DISCUSSION

This study used the data for 140,294 prostate cancer patients in the SEER database to construct an accurate, reliable, and personalized

TABLE 2 Evaluation metrics for 10 machine learning models.

Models	AUC_ train	AUC_ test	Sensitivity	Specificity	Accuracy	NPV	PPV
SVM	0.926	0.922	0.931	0.805	0.850	0.954	0.728
KNN	0.976	0.899	0.856	0.839	0.845	0.912	0.749
DT	0.938	0.809	0.767	0.849	0.820	0.867	0.740
LR	0.921	0.917	0.902	0.834	0.859	0.938	0.753
RF	0.985	0.929	0.852	0.876	0.868	0.913	0.795
AdaBoost	0.935	0.928	0.925	0.815	0.855	0.951	0.738
XGBoost	0.969	0.939	0.880	0.875	0.877	0.928	0.798
LightGBM	0.958	0.939	0.890	0.867	0.876	0.934	0.790
CatBoost	0.968	0.939	0.877	0.878	0.877	0.927	0.801
MLP	0.944	0.937	0.895	0.862	0.874	0.936	0.784

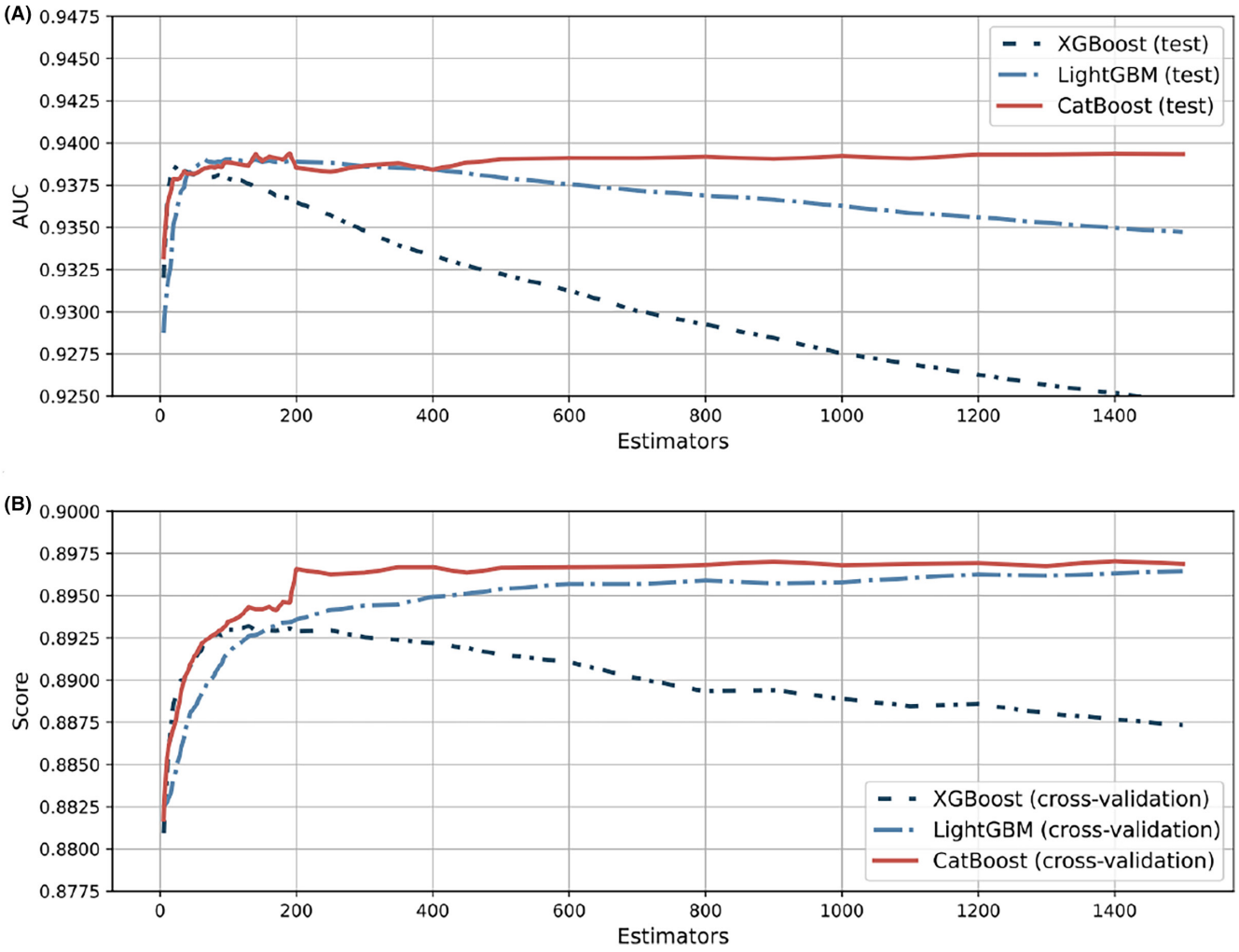

FIGURE 3 Area under the receiver operating characteristic (ROC) curve and scores of the 10-fold cross-validation test set of XGBoost, LightGBM and CatBoost with different numbers of weak classifiers. The performance of three models (XGBoost, LightGBM, and CatBoost) at different numbers of estimators. (A) Displays the AUC values on the test set for these models, while (B) shows their scores on the 10-fold cross-validation set.

prostate cancer treatment recommendation model based on a variety of machine learning algorithms. At the same time, the SHAP algorithm was used to analyze in depth the importance and influence mechanism of variables. After rigorous and meticulous performance evaluation, the CatBoost model performed well in predicting the most suitable treatment regimen for prostate cancer patients, with an AUC value of 0.939, sensitivity, and specificity of 0.877 and 0.878, respectively. The analysis method constructed by the

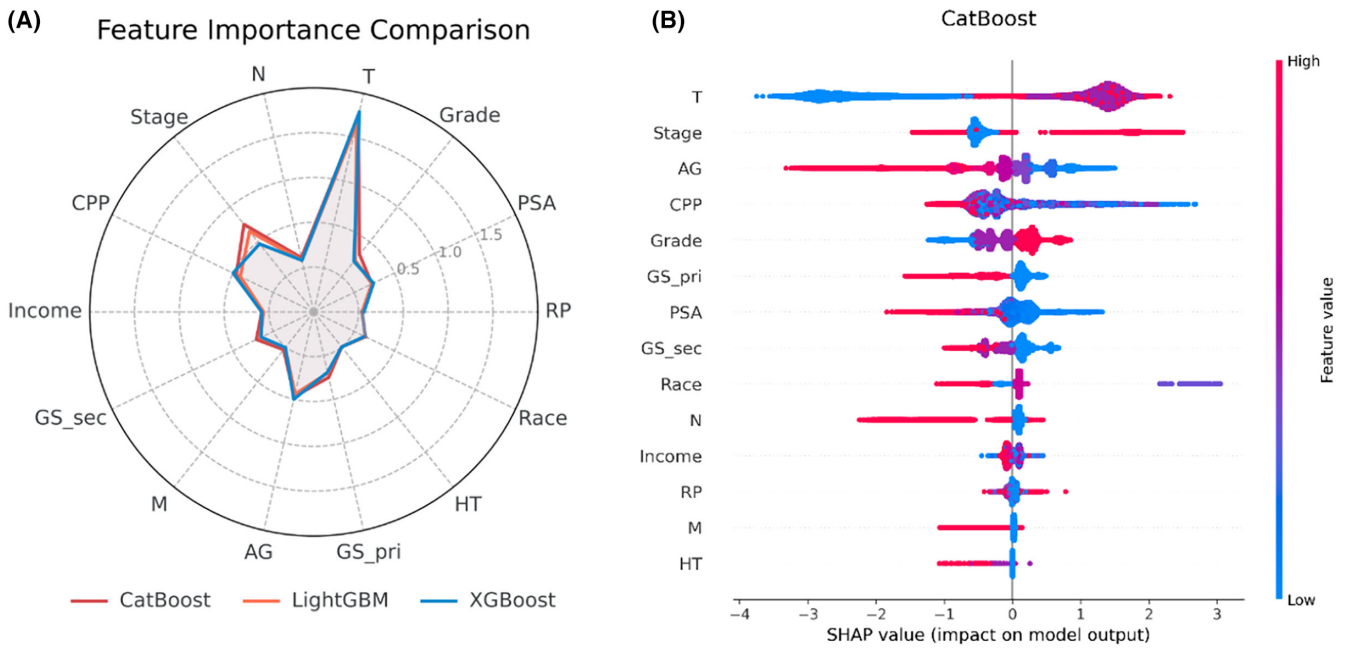

FIGURE 4 SHAP value model feature importance analysis. (A) Illustrates the SHAP values of CatBoost, LightGBM, and XGBoost on different features, representing the relative importance of features under different models. Each feature is assigned to a quadrant, with colors representing different models. (B) Demonstrates the impact of each feature on prediction results in the CatBoost model. The horizontal axis represents the SHAP values, where the SHAP value of each feature indicates its impact on the model output. The vertical axis represents feature ranking, with higher ranks indicating greater influence of the feature on the model output.

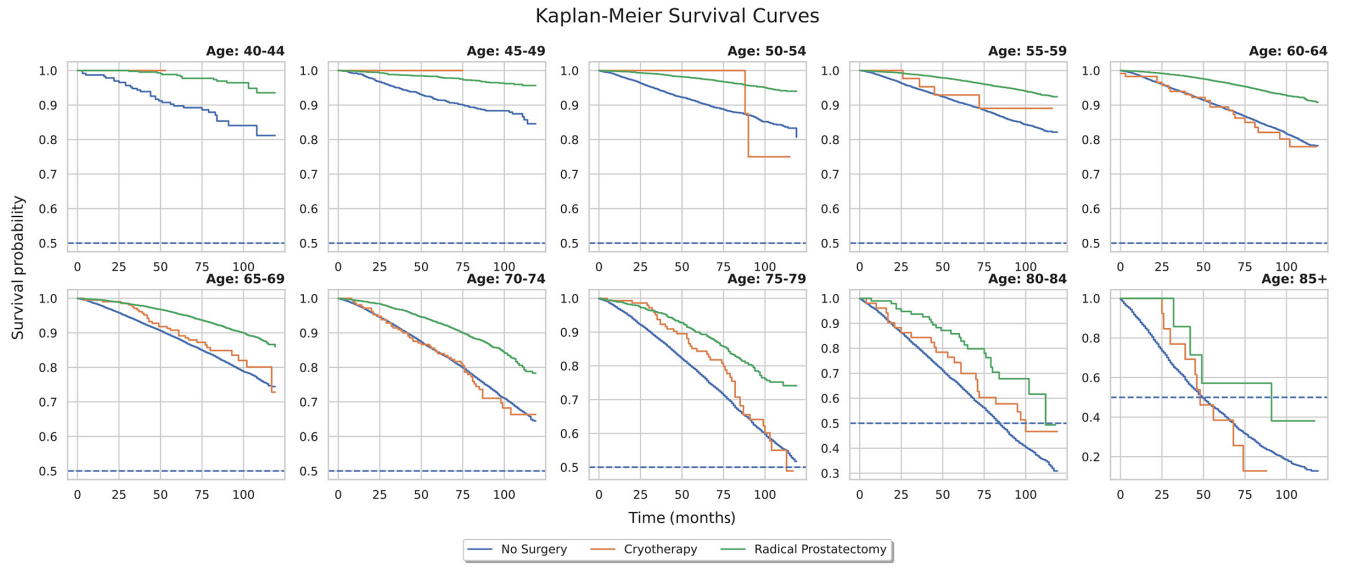

FIGURE 5 Survival curves for different treatment modalities among prostate cancer patients of different age groups. Comprising 10 subplots, each subplot represents the changes in survival probability over time for patients in different age groups. The horizontal axis of each subplot represents time (in months), while the vertical axis represents survival probability. Survival probability refers to the likelihood of patients surviving during a given time period. The curves in each subplot represent the changes in survival probability over time for patients undergoing different surgical methods. The colors of the curves denote different surgical methods, including no surgery, cryotherapy, radical prostatectomy.

CatBoost model holds substantial guiding value for clinical practice, providing strong support for doctors to make more accurate and reasonable decisions when formulating individualized treatment plans for prostate cancer.

In recent years, some researchers have carried out machine learning-based research on prostate cancer treatment recommendation models, but have mainly focused on the prediction of non-surgical strategies. Hu et al. used an LR model to establish a model to

TABLE 3 Five-year and 10-year survival rates for different treatments.

Treatment	5 years[a]	10 years[a]
No surgery	0.846 (0.844, 0.849)	0.678 (0.672, 0.684)
Cryotherapy	0.861 (0.837, 0.886)	0.644 (0.578, 0.718)
Radical prostatectomy	0.962 (0.960, 0.964)	0.892 (0.889, 0.899)

[a]The 5-year and 10-year survival rates for various treatment modalities were estimated using the Kaplan–Meier method.

distinguish between active surveillance and watchful waiting treatment strategies in conservative treatment, with an AUC of 0.73, an *F*-value of 0.79, and an accuracy of 0.71.[15] Alatas et al. considered five influencing factors (Gleason score, biopsy results, PSA level, SUVmax and age) on 88 patient data and used the XGBoost model to conduct recommendation predictions for hormone therapy, radiotherapy and combination therapy.[13] Compared with previous studies, this study not only greatly increases the size of the experimental population but also considers more factors influencing treatment decisions, and carries out comprehensive and detailed comparative optimization of a variety of machine learning models.

Through in-depth research, we revealed that T stage, cancer stage, and cancer grade at lower stages have a significant impact on the choice of non-surgical treatment options for prostate cancer patients, which deepens the understanding of early prostate cancer treatment decisions. This view is supported by empirical studies, including a 29-year follow-up study[35] that the presence of extracapsular invasion or a Gleason score of 8 to 9 predicts a higher risk of fatality in patients who have undergone radical prostatectomy. Therefore, when patients are in the early stages of the disease and there is no significant spread of the disease, non-surgical treatments, including active monitoring and observational waiting, may be a more sensible initial treatment strategy.

Age is a key consideration in prostate cancer treatment decisions. In our study, younger patients are more inclined to opt for surgery. Cryotherapy is becoming the preferred alternative treatment options for older patients due to their low risk of invasiveness and complications.[36] Several studies have noted a poor prognosis in older patients who had undergone radical prostatectomy.[35,37–41] As an example, Bill-Axelson et al. found that men aged 65 years and older had a lower response to radical prostatectomy compared with men younger than 65 years of age.[41] In a large-scale analysis of 160,787 men treated with radical prostatectomy in the United States, patients who were older at diagnosis had higher rates of prostate cancer mortality.[42] In patients with prostate cancer aged 75 years and older, cryotherapy has shown a survival advantage over other surgical modalities at a specific age,[43,44] which is consistent with the results in our study. However, there is some controversy in the academic community about the effect of age on the risk of death from prostate cancer. The study by Pettersson A et al. did not find a strong inherent effect of age on the risk of death from prostate cancer, but instead noted that, in current clinical practice, older patients with prostate cancer may not receive adequate diagnostic testing and subsequent curative treatment.[45] A study of 29 years of follow-up further revealed that neither age nor treatment regimen had demonstrated that age plays a decisive role in the outcome of radical prostatectomy.[35]

Our study revealed that radical prostatectomy as a mainstream treatment for prostate cancer is significantly more than 50% in young and middle-aged patients, especially in the age group of 40–54 years, highlighting the centrality of the procedure in the treatment of patients in this specific age group. This not only reflects the affirmation of the therapeutic effect of radical prostatectomy in young and middle-aged patients, but also suggests that clinicians tend to choose radical resection as an effective means to improve survival and quality of life after a comprehensive evaluation of the patient's age, disease stage, and physical condition.[32,46] We also found that patients with prostate cancer treated with radical prostatectomy showed significant advantages in 5-year and 10-year survival, regardless of age stratification. A long-term follow-up also confirmed that when 80% of participants died, the radical prostatectomy group had lower overall mortality, prostate cancer mortality, and metastatic risk, and that such procedures increased the average life expectancy of patients by about 2.9 years.[35] A long-term survival study further confirmed that radical prostatectomy was associated with higher survival rates in patients with early-stage prostate cancer, particularly in cases with a good prognosis, in which the probability of survival was relatively high for 10 years or more after surgery.[47]

In the study of the treatment effect of prostate cancer, we used survival curve analysis to reveal a significant advantage of surgical treatment over non-surgical treatment in a 10-year survival rate, which is consistent with the results of previous studies.[48,49] A 10-year study[47] revealed that surgery and radiotherapy had a more positive performance than active surveillance in inhibiting the risk of disease progression and metastasis. The fact that patients with prostate cancer who had not undergone surgery generally had a shorter survival rate reveals the importance of factors such as late diagnosis, severe comorbidities, and surgical inappropriateness in treatment decisions.

Therefore, in the process of making decisions on the treatment of prostate cancer, it is necessary to carefully evaluate the differences in the effects of various treatment modalities and realize the individualization and precision of the treatment plan based on the specific situation of the patient. Following current internationally accepted clinical practice guidelines,[50,51] for young prostate cancer patients, surgery may be a more appropriate option due to its long-term efficacy and quality of life advantages compared with endocrine therapy alone, although endocrine therapy can inhibit tumor proliferation to some extent by blocking androgen signaling. In the elderly

population of prostate cancer patients, it is particularly important to strictly implement prostate cancer screening and treatment guidelines to determine the most suitable individualized treatment strategy for them, given the increasing trend of life expectancy.

In this study, machine learning methods and SHAP tools are used to construct a predictive model, which effectively realizes the interpretability of the model. However, there are some limitations to the study. On the one hand, in the current data set, all biomedical indicators that may influence treatment decisions, such as the patient's overall health, comorbidities, quality of life, etc., may not be adequately covered. On the other hand, sample size, especially in the patient population of rare or specific types of prostate cancer, may not be sufficient to reveal all potential correlations. Although SHAP analysis helps us understand the importance of feature values for prediction outcomes, this importance should not be interpreted as a causal relationship. Focusing on the future, the research team plans to expand data sources, including but not limited to imaging data, genomic information and other detailed clinical data, to build a more accurate and comprehensive prediction model, hoping to provide a more accurate and comprehensive basis for treatment decisions for prostate cancer, and further improve the scientificity and practicability of individualized treatment and prognosis assessment.[52]

Based on the SEER database, this study applies a variety of machine learning techniques to predict the optimal treatment for prostate cancer patients. Among them, the CatBoost model performs best with an AUC value of 0.939, showing its advantages in prostate cancer treatment prediction. The SHAP tool is used to reveal the internal mechanism of the model, and the key role of clinical indicators such as T stage, cancer stage, core positive ratio, and age in the treatment decision of prostate cancer is clarified. In addition, by plotting and comparing the survival curves under different surgical methods, the difference in the impact of each treatment regimen on the survival prognosis of patients is visually displayed. Combined with machine learning and SHAP explanatory analysis, this study successfully constructs a model that accurately predicts the treatment mode of prostate cancer, which strongly supports clinically accurate and personalized diagnosis and treatment decisions.

AUTHOR CONTRIBUTIONS

Shengsheng Tang: Conceptualization; data curation; investigation; methodology; writing – original draft; writing – review and editing. **Hongzheng Zhang:** Conceptualization; data curation; investigation; methodology; writing – original draft; writing – review and editing. **Junhao Liang:** Investigation; methodology; writing – original draft. **Shishi Tang:** Visualization; writing – review and editing. **Lin Li:** Writing – review and editing. **Yuxuan Li:** Writing – review and editing. **Yuan Xu:** Methodology. **Daohu Wang:** Conceptualization. **Yi Zhou:** Conceptualization; funding acquisition; writing – review and editing.

ACKNOWLEDGMENTS

This research was supported by the Key Research and Development Program of China (2022YFC3601600), the Guangdong Natural Science Foundation (2024A1515011989), the Guangzhou Science and Technology Plan Project (202206010028), the Fundamental Research Funds for the Central Universities of Sun Yat-sen University (24xkjc025), and the Key Research and Development Programs of Jiangxi Provincial Department of Science and Technology (20202BBGL73074). We also extend our sincere gratitude to the SEER program in the United States for providing the data used in this study. Furthermore, we would like to express our heartfelt thanks to the editor and anonymous reviewers, whose valuable feedback significantly improved this paper.

CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

ETHICS STATEMENT

Approval of the research protocol by an Institutional Reviewer Board: N/A.

Informed Consent: N/A.

Registry and the Registration No. of the study/trial: N/A.

Animal Studies: N/A.

ORCID

Yi Zhou https://orcid.org/0000-0002-4774-2234

REFERENCES

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin*. 2023;73(1):17-48.
2. Miller KD, Nogueira L, Devasia T, et al. Cancer treatment and survivorship statistics, 2022. *CA Cancer J Clin*. 2022;72(5):409-436.
3. Valerio M, Ahmed HU, Emberton M, et al. The role of focal therapy in the management of localised prostate cancer: a systematic review. *Eur Urol*. 2014;66(4):732-751.
4. Parker C, Castro E, Fizazi K, et al. Prostate cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2020;31(9):1119-1134.
5. Litwin MS, Tan H. The diagnosis and treatment of prostate cancer. *JAMA*. 2017;317(24):2532.
6. Chen AB, Niu J, Cronin AM, Shih YCT, Giordano S, Schrag D. Variation in use of high-cost technologies for palliative radiation therapy by radiation oncologists. *J Natl Compr Cancer Netw*. 2021;19(4):421-431.
7. Bulten W, Kartasalo K, Chen PHC, et al. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat Med*. 2022;28(1):154-163.
8. McIntosh C, Conroy L, Tjong MC, et al. Clinical integration of machine learning for curative-intent radiation treatment of patients with prostate cancer. *Nat Med*. 2021;27(6):999-1005.
9. Ström P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol*. 2020;21(2):222-232.
10. Goldenberg SL, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. *Nat Rev Urol*. 2019;16(7):391-403.
11. Wang Y, Wang D, Geng N, Wang Y, Yin Y, Jin Y. Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection. *Appl Soft Comput*. 2019;77:188-204.
12. Kench JG, Amin MB, Berney DM, et al. WHO classification of Tumours fifth edition: evolving issues in the classification, diagnosis, and prognostication of prostate cancer. *Histopathology*. 2022;81(4):447-458.

13. Alataş E, Tanyıldızı Kökkülünk H, Tanyıldızı H, Alcın G. Treatment prediction with machine learning in prostate cancer patients. *Comput Methods Biomech Biomed Engin*. 2023;1-9. doi:10.1080/10255842.2023.2298364
14. Conroy L, Khalifa A, Berlin A, McIntosh C, Purdie TG. Performance stability evaluation of atlas-based machine learning radiation therapy treatment planning in prostate cancer. *Phys Med Biol*. 2021;66(13):134001.
15. Hu S, Chang CP, Snyder J, et al. Comparing active surveillance and watchful waiting with radical treatment using machine learning models among patients with prostate cancer. *JCO Clin Cancer Inform*. 2023;7:e2300083.
16. Syed K, Sleeman W, Soni P, et al. Machine-learning models for multicenter prostate cancer treatment plans. *J Comput Biol*. 2021;28(2):166-184.
17. Auffenberg GB, Ghani KR, Ramani S, et al. askMUSIC: leveraging a clinical registry to develop a new machine learning model to inform patients of prostate cancer treatments chosen by similar men. *Eur Urol*. 2019;75(6):901-907.
18. Samarinas M, Mallios N, Papageorgiou E. Comparison of Machine Learning Techniques using the WEKA Environment for Prostate Cancer Therapy Plan. in 2011 IEEE 20th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises. 2011.
19. Rodriguez-Perez R, Bajorath J. Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. *J Comput Aided Mol Des*. 2020;34(10):1013-1026.
20. Cooperberg MR, Vickers AJ, Broering JM, Carroll PR, for the Cancer of the Prostate Strategic Urologic Research Endeavor (CaPSURE) Investigators. Comparative risk-adjusted mortality outcomes after primary surgery, radiotherapy, or androgen-deprivation therapy for localized prostate cancer. *Cancer*. 2010;116(22):5226-5234.
21. Herlemann A, Cowan JE, Washington SL 3rd, et al. Long-term prostate cancer-specific mortality after prostatectomy, brachytherapy, external beam radiation therapy, hormonal therapy, or monitoring for localized prostate cancer. *Eur Urol*. 2024;85(6):565-573.
22. Hoffman RM, Koyama T, Fan KH, et al. Mortality after radical prostatectomy or external beam radiotherapy for localized prostate cancer. *J Natl Cancer Inst*. 2013;105(10):711-718.
23. Sooriakumaran P, Nyberg T, Akre O, et al. Comparative effectiveness of radical prostatectomy and radiotherapy in prostate cancer: observational study of mortality outcomes. *BMJ*. 2014;348:g1502.
24. Shimizu H, Nakayama KI. Artificial intelligence in oncology. *Cancer Sci*. 2020;111(5):1452-1460.
25. Nguyen QTN, Nguyen PA, Wang CJ, et al. Machine learning approaches for predicting 5-year breast cancer survival: a multicenter study. *Cancer Sci*. 2023;114(10):4063-4072.
26. Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med*. 2023;388(13):1201-1208.
27. Wang G, Teoh JYC, Choi K-S. Diagnosis of prostate cancer in a Chinese population by using machine learning methods. in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018.
28. Chiu PK, Shen X, Wang G, et al. Enhancement of prostate cancer diagnosis by machine learning techniques: an algorithm development and validation study. *Prostate Cancer Prostatic Dis*. 2022;25(4):672-676.
29. Li L, Cheng Y, Ji W, et al. Machine learning for predicting diabetes risk in western China adults. *Diabetol Metab Syndr*. 2023;15(1):165.
30. Evans AJ. Treatment effects in prostate cancer. *Mod Pathol*. 2018;31(S1):S110-S121.
31. Cornford P, van den Bergh RCN, Briers E, et al. EAU-EANM-ESTRO-ESUR-ISUP-SIOG guidelines on prostate cancer—2024 update. Part I: screening, diagnosis, and local treatment with curative intent. *Eur Eurol*. 2024;86:148-163.
32. Teo MY, Rathkopf DE, Kantoff P. Treatment of advanced prostate cancer. *Annu Rev Med*. 2019;70:479-499.
33. Antonarakis ES, Feng Z, Trock BJ, et al. The natural history of metastatic progression in men with prostate-specific antigen recurrence after radical prostatectomy: long-term follow-up. *BJU Int*. 2012;109(1):32-39.
34. Sundi D, Wang VM, Pierorazio PM, et al. Very-high-risk localized prostate cancer: definition and outcomes. *Prostate Cancer Prostatic Dis*. 2014;17(1):57-63.
35. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in prostate cancer—29-year follow-up. *N Engl J Med*. 2018;379(24):2319-2329.
36. Abt D, Müllhaupt G, Hechelhammer L, et al. Prostatic artery embolisation versus transurethral resection of the prostate for benign prostatic hyperplasia: 2-yr outcomes of a randomised, open-label. *Single-Centre Trial Eur Urol*. 2021;80(1):34-42.
37. Briganti A, Spahn M, Joniau S, et al. Impact of age and comorbidities on long-term survival of patients with high-risk prostate cancer treated with radical prostatectomy: a multi-institutional competing-risks analysis. *Eur Urol*. 2013;63(4):693-701.
38. Richstone L, Bianco FJ, Shah HH, et al. Radical prostatectomy in men aged >or=70 years: effect of age on upgrading, upstaging, and the accuracy of a preoperative nomogram. *BJU Int*. 2008;101(5):541-546.
39. Sanchez-Salas R, Prapotnich D, Rozet F, et al. Laparoscopic radical prostatectomy is feasible and effective in 'fit' senior men with localized prostate cancer. *BJU Int*. 2010;106(10):1530-1536.
40. Sun L, Caire AA, Robertson CN, et al. Men older than 70 years have higher risk prostate cancer and poorer survival in the early and late prostate specific antigen eras. *J Urol*. 2009;182(5):2242-2248.
41. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in early prostate cancer. *N Engl J Med*. 2014;370(10):932-942.
42. Abdollah F, Sun M, Thuret R, et al. A competing-risks analysis of survival after alternative treatment modalities for prostate cancer patients: 1988-2006. *Eur Urol*. 2011;59(1):88-95.
43. Ramsay CR, Adewuyi TE, Gray J, et al. Ablative therapy for people with localised prostate cancer: a systematic review and economic evaluation. *Health Technol Assess*. 2015;19(49):1-490.
44. Zheng Y, Ye Y, Chen L, et al. Prevalence and outcomes of focal ablation versus prostatectomy for elderly patients with prostate cancer: a population-based study. *J Natl Cancer Cent*. 2022;2(1):25-32.
45. Pettersson A, Robinson D, Garmo H, Holmberg L, Stattin P. Age at diagnosis and prostate cancer treatment and prognosis: a population-based cohort study. *Ann Oncol*. 2018;29(2):377-385.
46. Desai K, McManus JM, Sharifi N. Hormonal therapy for prostate cancer. *Endocr Rev*. 2021;42(3):354-373.
47. Donovan JL, Hamdy FC, Lane JA, et al. Patient-reported outcomes 12 years after localized prostate cancer treatment. *NEJM Evid*. 2023;2(4):EVIDoa2300018.
48. Sooriakumaran P, Nyberg T, Akre O, et al. Survival among men at high risk of disseminated prostate cancer receiving initial locally directed radical treatment or initial androgen deprivation therapy. *Eur Urol*. 2017;72(3):345-351.
49. Hwang WL, Tendulkar RD, Niemierko A, et al. Comparison between adjuvant and early-salvage postprostatectomy radiotherapy for prostate cancer with adverse pathological features. *JAMA Oncol*. 2018;4(5):e175230.
50. Mottet N, van den Bergh RCN, Briers E, et al. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer-2020 update. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol*. 2021;79(2):243-262.
51. Cornford P, van den Bergh RCN, Briers E, et al. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer. Part II-2020 update: treatment of relapsing and metastatic prostate cancer. *Eur Urol*. 2021;79(2):263-282.

52. Liu M, Luo J, Li L, et al. Design and development of a disease-specific clinical database system to increase the availability of hospital data in China. *Health Inf Sci Syst*. 2023;11(1):11.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Tang S, Zhang H, Liang J, et al. Prostate cancer treatment recommendation study based on machine learning and SHAP interpreter. *Cancer Sci*. 2024;115:3755-3766. doi:10.1111/cas.16327