



Artificial Intelligence Combined With Big Data to Predict Lymph Node Involvement in Prostate Cancer: A Population-Based Study

Liwei Wei^{1†}, Yongdi Huang^{2†}, Zheng Chen^{1†}, Hongyu Lei², Xiaoping Qin¹, Lihong Cui^{2*} and Yumin Zhuo^{1*}

¹ Department of Urology, the First Affiliated Hospital of Jinan University, Guangzhou, China, ² College of Mathematics and Physics, Beijing University of Chemical Technology, Beijing, China

OPEN ACCESS

Edited by:

Benyi Li,
University of Kansas Medical Center,
United States

Reviewed by:

Junru Chen,
Sichuan University, China
Jun Hao,
University Health Network, Canada

*Correspondence:

Lihong Cui
mathcui@163.com
Yumin Zhuo
tzuoyumin@126.com

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Genitourinary Oncology,
a section of the journal
Frontiers in Oncology

Received: 23 August 2021

Accepted: 22 September 2021

Published: 14 October 2021

Citation:

Wei L, Huang Y, Chen Z, Lei H,
Qin X, Cui L and Zhuo Y (2021)
Artificial Intelligence Combined With
Big Data to Predict Lymph Node
Involvement in Prostate Cancer: A
Population-Based Study.
Front. Oncol. 11:763381.
doi: 10.3389/fonc.2021.763381

Background: A more accurate preoperative prediction of lymph node involvement (LNI) in prostate cancer (PCa) would improve clinical treatment and follow-up strategies of this disease. We developed a predictive model based on machine learning (ML) combined with big data to achieve this.

Methods: Clinicopathological characteristics of 2,884 PCa patients who underwent extended pelvic lymph node dissection (ePLND) were collected from the U.S. National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) database from 2010 to 2015. Eight variables were included to establish an ML model. Model performance was evaluated by the receiver operating characteristic (ROC) curves and calibration plots for predictive accuracy. Decision curve analysis (DCA) and cutoff values were obtained to estimate its clinical utility.

Results: Three hundred and forty-four (11.9%) patients were identified with LNI. The five most important factors were the Gleason score, T stage of disease, percentage of positive cores, tumor size, and prostate-specific antigen levels with 158, 137, 128, 113, and 88 points, respectively. The XGBoost (XGB) model showed the best predictive performance and had the highest net benefit when compared with the other algorithms, achieving an area under the curve of 0.883. With a 5%–20% cutoff value, the XGB model performed best in reducing omissions and avoiding overtreatment of patients when dealing with LNI. This model also had a lower false-negative rate and a higher percentage of ePLND was avoided. In addition, DCA showed it has the highest net benefit across the whole range of threshold probabilities.

Conclusions: We established an ML model based on big data for predicting LNI in PCa, and it could lead to a reduction of approximately 50% of ePLND cases. In addition, only ≤3% of patients were misdiagnosed with a cutoff value ranging from 5% to 20%. This promising study warrants further validation by using a larger prospective dataset.

Keywords: prostate cancer, machine learning, lymph node involvement, predictive model, SEER database

INTRODUCTION

Prostate cancer (PCa) is the most common type of malignant tumor in American men and accounts for nearly 15% of all cancer cases. Recurrence and metastasis are the most common causes of death in PCa patients (1). Radical prostatectomy (RP) is the gold-standard treatment for patients with PCa and those with either organ-confined or locally advanced PCa can benefit from it (2–4). Because of the biological characteristics of this disease and its response to effective treatment, patients with PCa generally have an excellent long-term prognosis with the vast majority of patients surviving over 5 years (5). However, if a PCa patient is diagnosed with lymph node involvement (LNI), the probability of tumor recurrence will increase, and the prognosis will deteriorate significantly (6, 7).

For a lymph node (LN)-positive PCa patient, extended pelvic lymph node dissection (ePLND) can often offer a better cancer-specific outcome either with or without adjuvant androgen deprivation therapy (ADT) (8, 9). As LNI is a significant component in the prognosis of a patient, it can influence any clinical decisions made by the surgeon. Before surgery, it is essential to know the precise clinical status of LNI in patients with PCa. To this end, researchers have reported several approaches and tools that can help the clinician to estimate the occurrence of LNI, including prostate-specific membrane antigen-positron emission tomography (PSMA-PET) scans, multiparametric magnetic resonance imaging (mpMRI), and the use of some emerging molecular biomarkers (10–13). However, these imaging techniques are not accurate and most of the molecular biomarkers are unproven. Therefore, the most commonly used tools available are Briganti, Partin, and Memorial Sloan Kettering Cancer Center (MSKCC) nomograms which have an accuracy of less than 80% (14–17). Machine learning (ML) is an emerging intersection approach involving many fields that allows for accurate prediction of outcomes from multiple unrelated datasets, which would otherwise be discrete and difficult to associate (18).

With the rapid development of evidence-based medicine, vast and complex medical datasets need more advanced techniques for their interpretation, and ML is becoming a promising option for the diagnosis and prognosis/prediction of many diseases (19, 20). In addition, ML has demonstrated excellent performance of predictive abilities and a good potential application in several areas of medicine (21). Our goal was to develop a new decision-support ML model based on big data for predicting the risk of LNI in PCa patients. We used area under the curves (AUCs), calibration plots, and decision curve analysis (DCA) to evaluate the performance of the model. We further validated the accuracy of our ML model by using a validation set.

MATERIAL AND METHODS

Data Source and Study Population

In this study, we used the Surveillance, Epidemiology, and End Results (SEER) database (<https://seer.cancer.gov/>) from the National Cancer Institute, a freely available cancer registry in

the United States. We obtained permission to access the files of the SEER database and all authors followed the SEER database regulations throughout the study. No personally identifiable information was used in this study and informed consent was not necessary from individual participants. The Medical Ethics Committee at Jinan University's First Affiliated Hospital examined and approved this work.

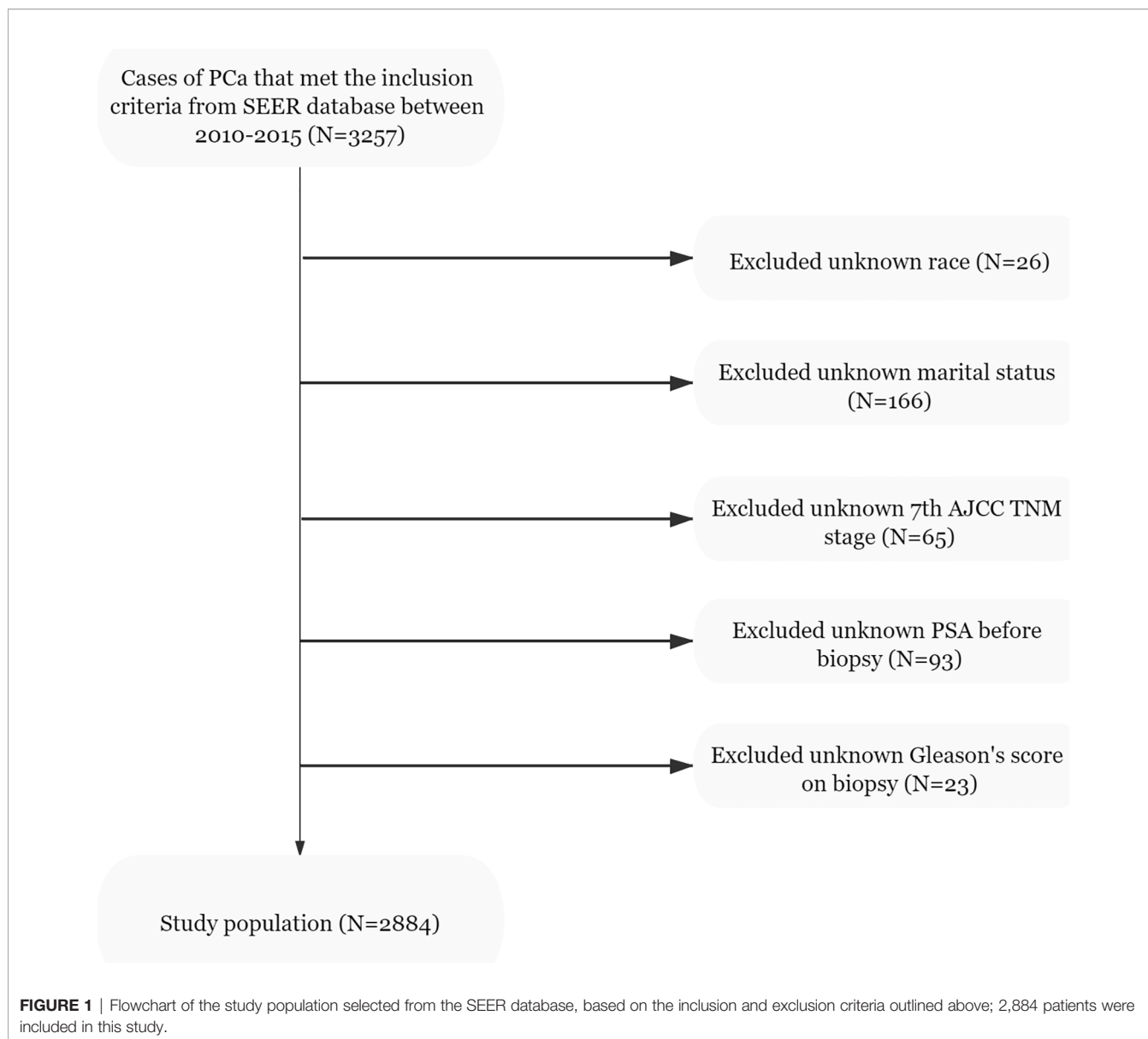
Data of the patients were downloaded from the SEER 18 Regs Research Data Nov 2018 Sub (1975–2016) by using the SEER*Stat 8.3.9.1 software. The selection criteria included men aged 35–90 years who were diagnosed with histologically confirmed prostatic adenocarcinoma (site code C61.9, morphology code 8140/3). The patients were diagnosed between 2010 and 2015 and PCa was the first malignant tumor found. All cases were treated with radical prostatectomy (RP) and ePLND without neoadjuvant systematic therapy. However, as the data obtained did not include ePLND as defined by anatomical location, we referred to the literature and defined this as more than or equal to 10 LNs being removed from the patients (22, 23). The T stage of the tumors was derived from preoperative examination and postoperative specimens. The number of needle cores examined was between 4 and 24, and the autopsy cases were excluded. The downloaded data of the patients were obtained from 3,257 cases. The exclusion criteria were as follows: unknown information from the American Joint Committee on Cancer seventh TNM stage, race, marital status, prostate-specific antigen (PSA) levels before biopsy, and the absence of Gleason scores (GS). From the 3,257 initial cases considered, 2,884 patients remained in the final cohort for use in model development. The population data selection procedure used is shown in **Figure 1**. The validation set was derived from the same SEER database and cases were diagnosed in 2016. Besides the above criteria, an extra one from the Briganti nomogram was chosen: patients with PSA >50 ng/ml were excluded from the validation set. A total of 535 patients were included in the analysis as the validation set.

Variable Selection

Several readily available clinical and demographic characteristics were chosen as independent variables for analysis. Since the SEER database does not include tumor size based on preoperative imaging, this parameter is prone to some errors. However, the tumor size of postoperative specimens in most patients did not differ significantly from the tumor size evaluated by imaging, and therefore, we considered this error to be within an acceptable range (24–27). At last, eight demographic and clinicopathological variables, namely, age at diagnosis, race, marital status, T stage, tumor size, PSA levels before biopsy, GS on biopsy, and percentage of positive cores (PPC), were selected as independent variables for analysis.

Model Establishment and Development

All statistical analyses in the study were performed using SPSS (version 22, IBM SPSS Software Foundation) and Python (version 3.8.1, Python Software Foundation). All variables were tested for Pearson correlations and the results are presented as a heat map (**Figure 2**). All patients studied were randomly divided into a training set and a test set at a ratio of 7:3 (**Table 1**). The

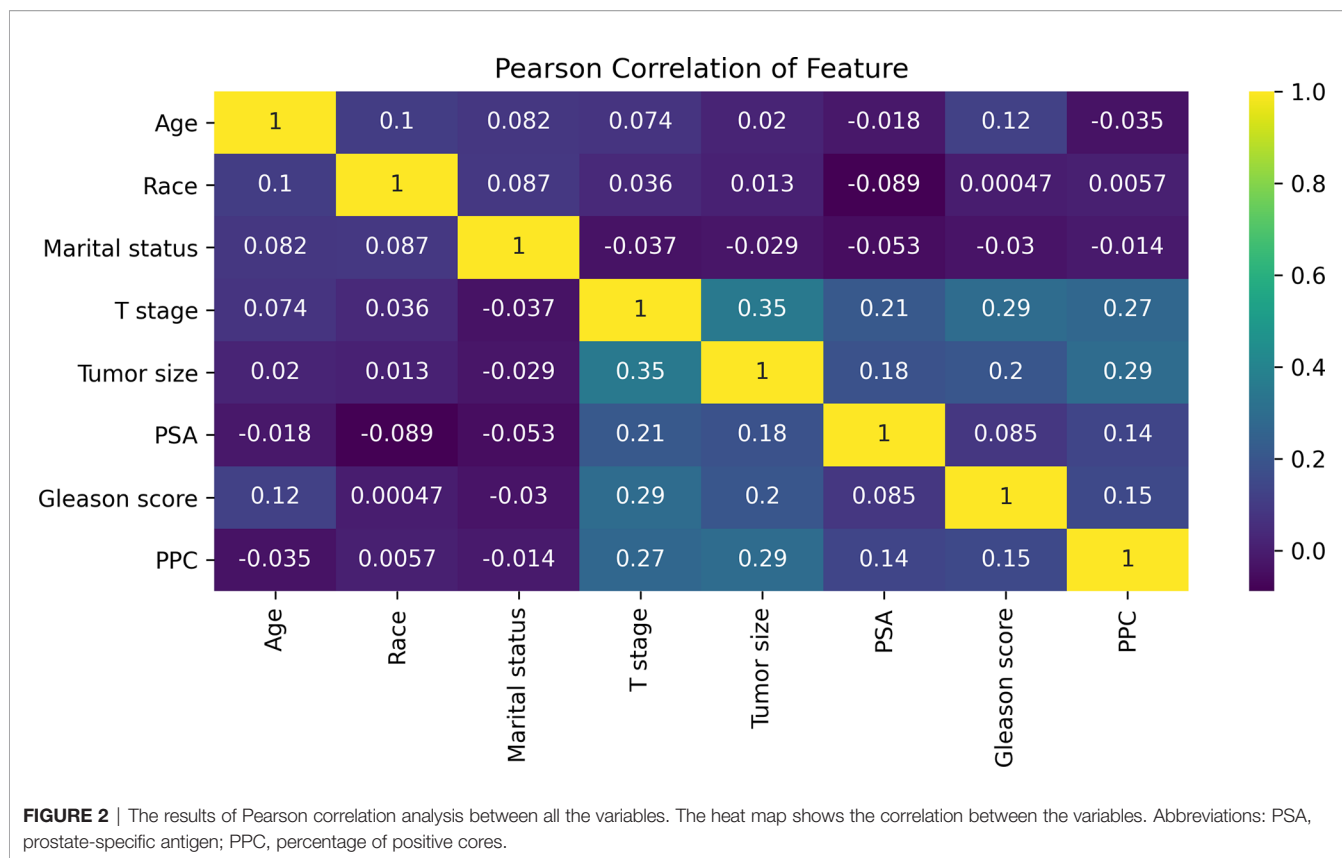


chi-square test was used to analyze the differences between the training and test sets. The training set was used to establish the extreme gradient boosting (XGB) and multivariate logistic regression (MLR) models, and the test set was then applied to evaluate both of them. We used 600 trees in XGB to build the ML model. For MLR, we used an entry variable selection method to establish the model. Then, the 5%, 10%, 15%, and 20% cutoff values of the XGB and MLR models were calculated, and these were used respectively to test their clinical value for directing the possible treatment options.

Model Improvement

To ensure that the model was stable, a 10-fold cross-validation was adopted to evaluate the predictive capability of the model. The training set was randomly divided into 10 groups. In each iteration of 10-fold cross-validation, nine groups were randomly

selected for training, and the remaining group was used as the test set. This means that each group was chosen as the test set in turn, which ensured that the evaluation results were not accidental. Then we averaged the results of the 10 evaluations in order to reduce the errors caused by any unreasonable selections made in the test set. To achieve the overall optimum value in the XGB model, we used the learning curve method to find the optimal parameters. The learning curve is shown in **Figure 3** where the abscissa axis represents the number of trees and different learning rates, and the ordinate axis represents the average AUC of the 10-fold cross-validation. The final optimal parameter combination was as follows: the number of trees (“*n* tree”) = 851 (**Figure 3A**), the learning rates (“eta”) = 0.16 (**Figure 3B**), the maximum length from the root node to leaf node (“max depth”) = 6, the sum of weights of the minimum leaf node samples (“min child weight”) = 1, and the L2 regularization



parameters (“reg lambda”) = 120. All other parameters were selected as default values for the calculations.

Evaluation of Model Performance

The performance analysis used comprised of three components. Firstly, model discrimination was quantified with receiver operating characteristic (ROC) curve analysis, and its predictive accuracy was assessed with the AUCs obtained. Secondly, we used calibration plots to evaluate the performance of the model, which indicated the calibration and how far the predictions of the model deviated from the actual event. Thirdly, clinical usefulness and net benefits were assessed with DCA which could estimate the net benefit by calculating the difference between the true- and false-positive rates and weighted these by the odds of the selected threshold probability of risks involved. Also, additional ML algorithms such as decision tree (DT) and support vector machine (SVM) were introduced for comparison. ROC curves and calibration plots were used to further evaluate the appropriateness and generalizability of our model.

RESULTS

Demographic and Clinicopathological Characteristics

A total of 2,884 PCa patients were analyzed in this study. Three hundred forty-four patients had LNI (11.9%) and 2,540 (88.1%)

did not have and these were classified as none LNI (nLNI). All patients were completely randomized with a ratio of 7:3 into a training set ($n = 2,018$) and a test set ($n = 866$). The demographic and clinicopathological variables of the patients are detailed in **Table 1**.

Model Analysis and Variable Feature Importance of the Prediction

The Pearson correlation analysis was performed for all the factors. A correlation heat map showed weak correlations between several clinicopathological variables (T stage, tumor size, PSA, GS, and PPC), and there was a moderate correlation between T stage and tumor size (**Figure 2**). For the MLR model, five of the eight variables were identified as independent risk factors, namely, T stage ($p < 0.001$), tumor size ($p < 0.001$), PSA before biopsy ($p < 0.001$), GS ($p < 0.001$), and PPC ($p = 0.006$) (**Table 2**). For the XGB model, we identified the feature of importance by the size of the gain value for each variable, with the higher values indicating more importance for the prediction target: GS (158 points), T stage (137 points), PPC (128 points), tumor size (113 points), PSA (88 points), race (64 points), age at diagnosis (51 points), and marital status (36 points) (**Figure 4**).

Model Performance

ROC curves, calibration plots, and DCA for the training set ($n = 2,018$) and the test set ($n = 866$) were constructed in order to determine the accuracy of our models. The XGB model had the

TABLE 1 | Clinical and pathological characteristics of the training and test sets.

Variables	Training set		Test set		p-value
	NLNI (n = 1,787) (%)	LNI (n = 231) (%)	NLNI (n = 753) (%)	LNI (n = 113) (%)	
Age at diagnosis					0.072
<62	801 (44.8)	111 (48.1)	362 (48.1)	61 (46.0)	
≥62	986 (55.2)	120 (51.9)	391 (51.9)	52 (54.0)	
Race					0.578
Black	203 (11.3)	24 (10.4)	99 (13.2)	15 (13.4)	
Other	138 (7.7)	13 (5.6)	37 (4.9)	8 (7.0)	
White	1,446 (81.0)	194 (84.0)	617 (81.9)	90 (79.6)	
Marital status					0.683
Unmarried	178 (9.9)	33 (14.3)	79 (10.5)	16 (14.2)	
Married	1,609 (90.1)	198 (85.7)	674 (89.5)	97 (85.8)	
T stage					0.371
T1	1 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
T2	1,092 (61.2)	26 (11.2)	458 (60.8)	12 (10.6)	
T3	693 (38.8)	199 (86.1)	291 (38.7)	96 (84.9)	
T4	1 (0.0)	6 (2.7)	4 (0.5)	5 (4.5)	
Tumor size					0.415
0–10 mm	260 (14.5)	8 (3.5)	104 (13.8)	3 (2.7)	
10–20 mm	812 (45.4)	51 (22.1)	344 (45.7)	22 (19.5)	
20–30 mm	442 (24.7)	69 (29.9)	186 (24.7)	39 (34.5)	
>30 mm	273 (15.4)	103 (44.5)	119 (15.8)	49 (43.3)	
PSA before biopsy					0.859
0–4 ng/ml	205 (11.5)	15 (6.5)	82 (10.9)	6 (5.3)	
4–10 ng/ml	1,144 (64.0)	102 (44.1)	495 (65.7)	43 (38.1)	
10–20 ng/ml	316 (17.7)	57 (24.7)	139 (18.5)	29 (25.7)	
>20 ng/ml	122 (6.8)	57 (24.7)	37 (4.9)	35 (30.9)	
Primary biopsy Gleason score					0.223
≤6	323 (18.1)	8 (3.5)	112 (14.9)	5 (4.4)	
7	1073 (60.0)	102 (44.2)	446 (59.2)	50 (44.2)	
8	249 (13.9)	57 (24.7)	127 (16.9)	29 (25.7)	
≥9	142 (8.0)	64 (27.6)	68 (9.0)	29 (25.7)	
Positive biopsy percentage					0.650
0–25%	544 (30.4)	27 (11.7)	233 (30.1)	9 (8.0)	
25–50%	668 (37.4)	69 (29.9)	279 (37.1)	37 (32.7)	
50–75%	359 (20.1)	55 (23.8)	145 (19.3)	26 (23.0)	
75–100%	216 (12.1)	80 (34.6)	96 (13.5)	41 (36.3)	

best performance in both training and test sets (AUC = 0.907 and 0.883, respectively), compared with SVM (AUC = 0.837 and 0.831, respectively), DT (AUC = 0.873 and 0.853, respectively), and MLR (AUC = 0.769 and 0.763, respectively) (**Figures 5A, B**).

The sensitivity, specificity, and cutoff values of the predictions of the two models were also calculated for the patients having LNI or not. In addition, the predictive accuracy and error when the patients were predicted to be at risk of LNI are given in **Table 3**.

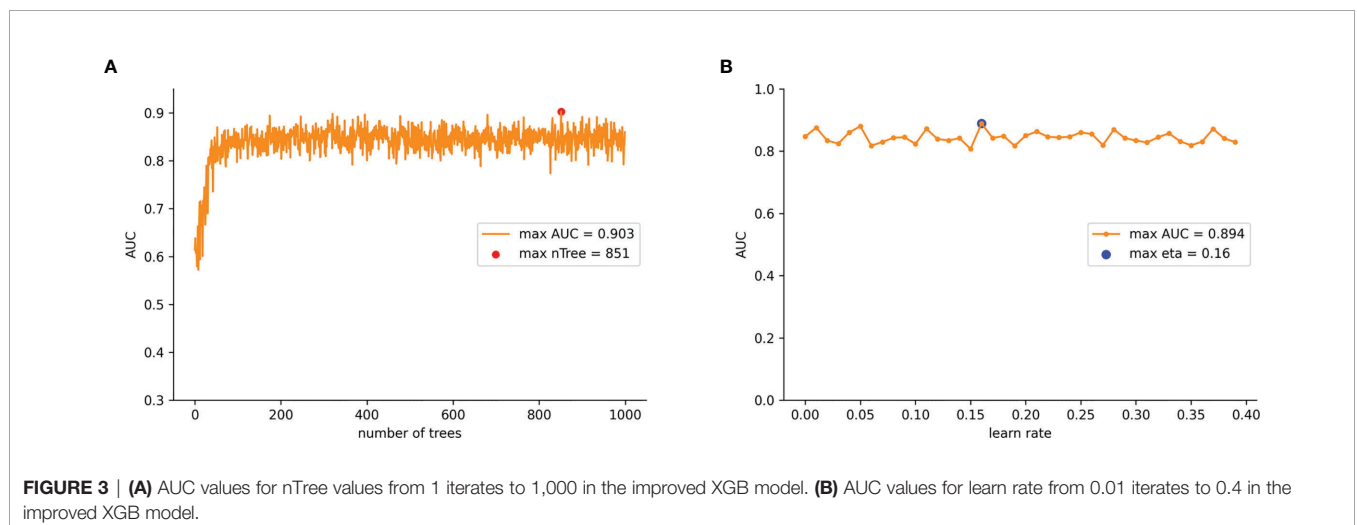


TABLE 2 | Multivariable logistic regression model with the entered variable selection.

Variables	OR (95% CI)	p-value
Age at diagnosis		
<62	Reference	
≥62	0.771 (0.592–1.003)	0.053
Race		
Black	Reference	
Other	1.047 (0.551–1.990)	0.889
White	1.251 (0.826–1.893)	0.290
Marital status		
Unmarried	Reference	
Married	0.883 (0.598–1.305)	0.533
T stage		
T1	–	
T2	Reference	
T3	6.363 (4.409–9.181)	<0.001
T4	32.343 (9.977–104.839)	<0.001
Tumor size		
0–10 mm	Reference	
10–20 mm	1.376 (0.691–2.737)	0.363
20–30 mm	2.436 (1.233–4.816)	0.010
>30 mm	4.033 (2.044–7.957)	<0.001
PSA before biopsy		
0–4 ng/ml	Reference	
4–10 ng/ml	1.123 (0.666–1.895)	0.663
10–20 ng/ml	1.785 (1.026–3.105)	0.040
>20 ng/ml	4.118 (2.310–7.338)	<0.001
Gleason's score		
≤6	Reference	
7	2.414 (1.315–4.431)	0.004
8	4.393 (2.321–8.315)	<0.001
≥9	5.696 (2.986–10.867)	<0.001
Positive biopsy percentage		
0–25%	Reference	
25–50%	1.679 (0.929–2.294)	0.662
50–75%	1.961 (0.961–2.379)	0.083
75–100%	4.368 (2.413–7.533)	0.006

The results showed that whether 5%, 10%, 15%, or 20% was chosen as the cutoff value, the XGB model was better than the MLR model in reducing omissions and avoiding overtreatment of patients, with a lower false-negative rate and a higher percentage of ePLND avoided. With a 5%–20% cutoff value, the XGB model could keep the risk of missing patients below 3% (1.2%–2.9%).

The calibration plots of the two case sets indicated that the predictive probabilities against observed LNI rates showed excellent concordance in the XGB model, followed by the SVM and DT models, respectively. The calibration of the MLR model tended to underestimate the LNI risk across the entire range of predicted probabilities compared with the other two models (Figure 6).

DCA of the four models was subsequently constructed in our study (Figure 7). The y-axis of the decision curve represents the net benefit which is a decision-analytic measure to judge whether any particular clinical decision results in more benefit than harm. Each point on the x-axis represents a threshold probability that differentiates between those patients with and without LNI (LNI vs. nLNI). This shows that all the models achieved net clinical benefit against a treat-all-or-none plan. With a risk threshold of less than 80%, the ML models showed a greater net benefit for patient interventions in the test set than the MLR model, and the XGB model had the highest net benefit across the whole range of threshold probabilities.

Model Validation

The validation set was used for model validation and the clinical and pathological characteristics of the validation sets are detailed in Table 4. In addition, ROC curves and calibration plots were constructed to compare the accuracy

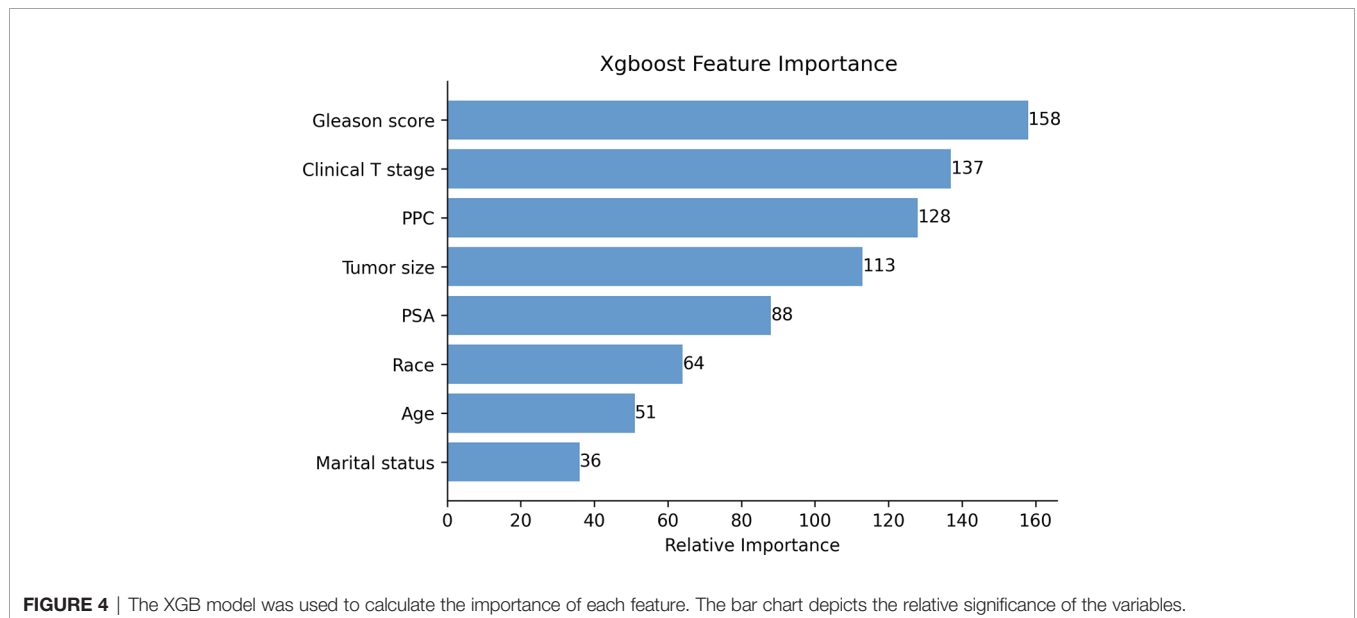
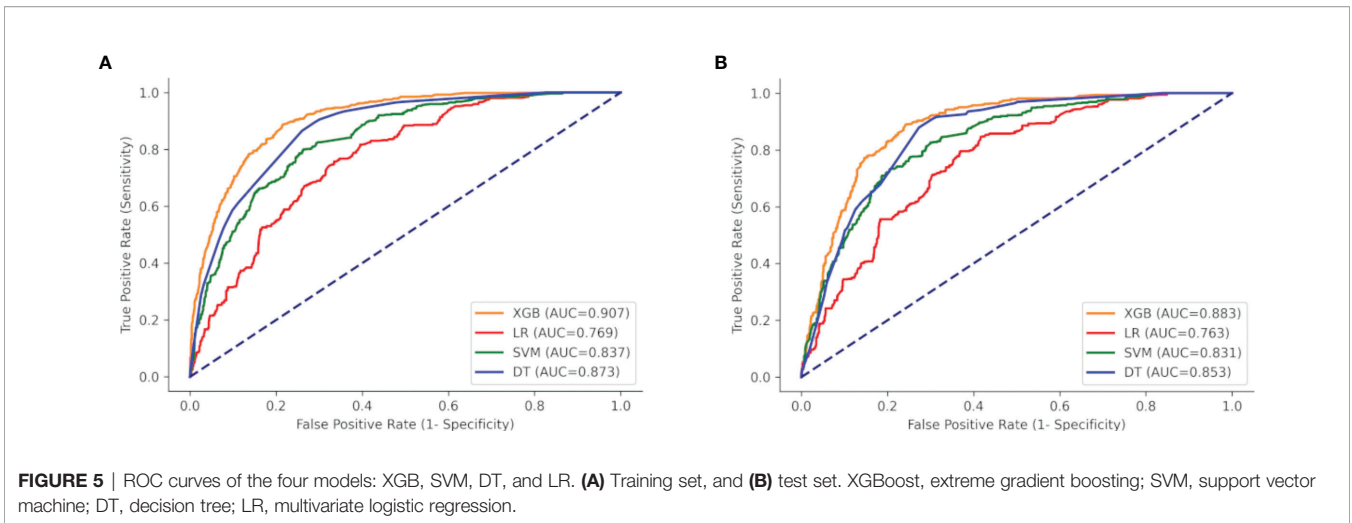


FIGURE 4 | The XGB model was used to calculate the importance of each feature. The bar chart depicts the relative significance of the variables.

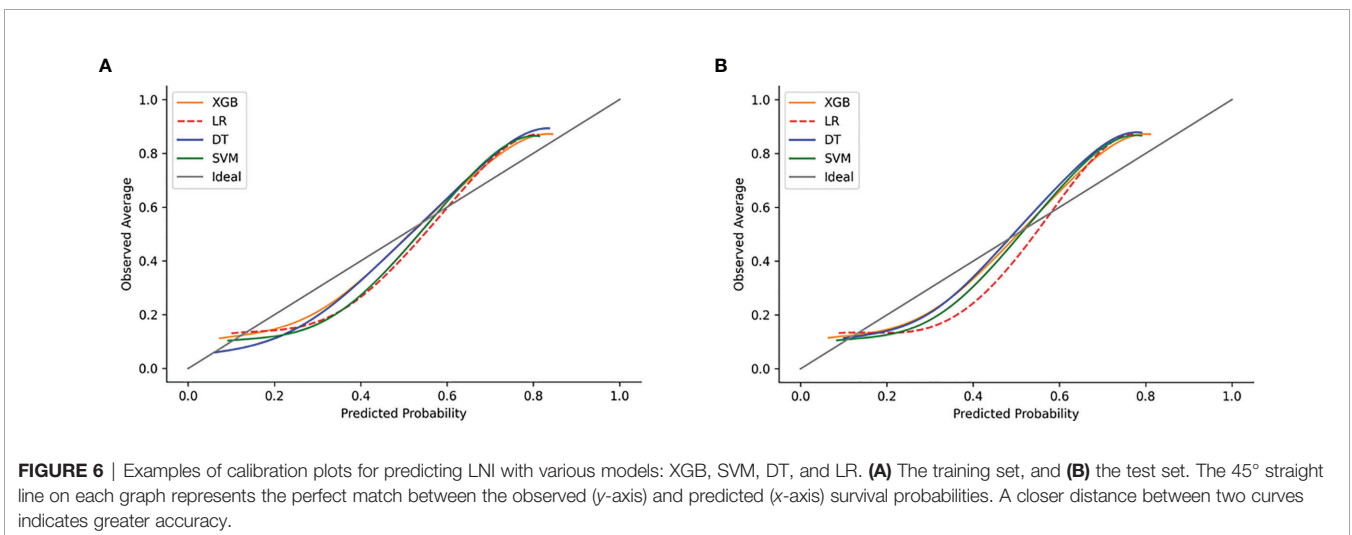


of our XGB model and the Briganti nomogram. According to ROC curves, the accuracy of the XGB model is higher than that of the nomogram (AUC: 0.850 vs. 0.816) (**Figure 8A**). Moreover, the calibration plots indicated that the XGB model

has better consistency across the 0% to 60% range of prediction probability. Instead, the calibration of the nomogram tended to underestimate the LNI risk in the same range (**Figure 8B**).

TABLE 3 | Analyses of the cutoff values of the XGB model based on the test set.

Treatment option	False positive, n/N (%)	False negative, n/N (%)	True positive, n/N (%)	True negative, n/N (%)
5% cutoff				
XGB	287/438 (65.5)	5/428 (1.2)	151/438 (34.5)	423/428 (98.9)
LR	459/630 (72.9)	6/236 (2.5)	171/630 (27.1)	230/236 (97.5)
10% cutoff				
XGB	186/322 (57.8)	9/544 (1.7)	136/322 (42.2)	535/544 (98.3)
LR	244/361 (67.6)	20/505 (4.0)	117/361 (32.4)	485/505 (96.0)
15% cutoff				
XGB	133/262 (50.8)	15/604 (2.5)	129/262 (49.2)	589/604 (97.5)
LR	139/222 (62.6)	34/644 (5.3)	83/222 (37.4)	610/644 (94.7)
20% cutoff				
XGB	81/190 (42.6)	20/676 (2.9)	109/190 (57.4)	656/676 (97.1)
LR	88/151 (58.3)	47/715 (6.6)	63/151 (41.7)	668/715 (93.4)



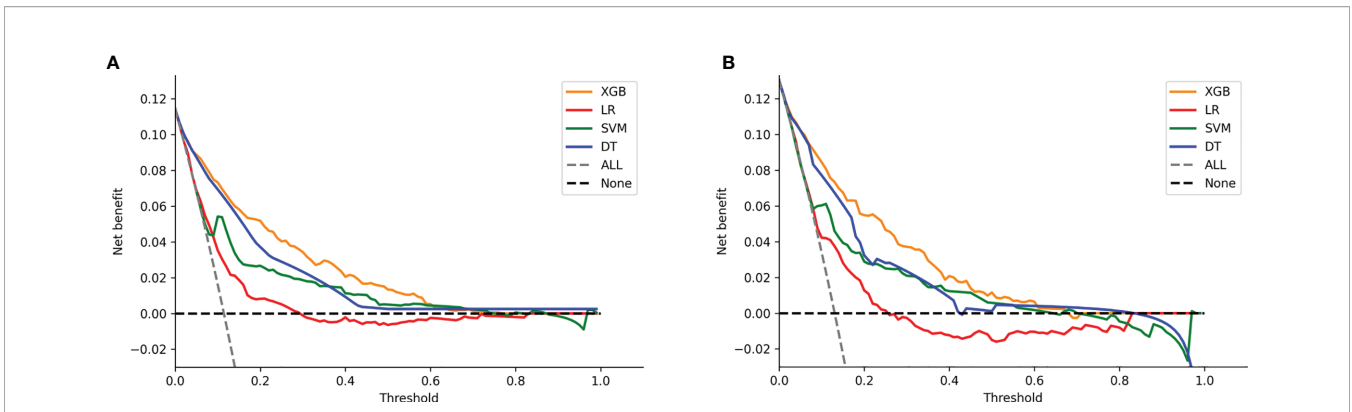


FIGURE 7 | Decision curve analysis graph showing the net benefit against threshold probabilities based on decisions from model outputs. Three curves were obtained based on predictions of the four different models, and the two curves obtained were based on two kinds of extreme decisions. The curves referred to as “All” represent the prediction that all the patients would progress to LNI, and the curves referred to as “None” represent the prediction that all the patients were nLNI. **(A)** The training set, and **(B)** the test set.

DISCUSSION

In this study, we developed a more accurate model to predict the risk of LNI in patients with PCa by combining eight clinicopathologic parameters. To our knowledge, this is the first ML model for predicting LNI established by using big data and readily available clinicopathological parameters.

LNI is found in up to 15% of PCa patients upon postoperative pathological examination and is associated with the recurrence and prognosis of PCa (28). As a standard treatment for PCa patients with LNI, ePLND can accurately help diagnose occult micrometastases, allowing PCa patients to get effective treatment and also identifying a more accurate stage of the disease. This is important for postsurgical follow-up and the subsequent selection of adjuvant and salvage therapies (29, 30). However, the rate of detection of an earlier stage or the presence of low-risk tumors at the time of diagnosis rises with PSA screening which leads to a higher rate of surgeries in patients with low- and intermediate-risk tumors. In addition, the likelihood of finding postsurgery-positive LNs decreases.

Our study found that the detection rate of postsurgical positive LN was lower than that reported in the literature, at approximately 12% (28). This indicated that some patients were overtreated, resulting in complications such as pelvic lymphocele, ileus, thrombosis, scrotal swelling, nerve injury, and so on (31, 32).

Therefore, determining the LN stage of PCa in patients before surgery is critical in determining whether they should receive ePLND. As routine imaging examinations such as CT scans and MRIs are currently ineffective at detecting nodal metastases (33), and there are only a few promising biomarkers in the preclinical stages of PCa (12, 13), ePLND is the only accurate way of detecting nodal metastases. In order to weigh the benefits and drawbacks of ePLND, the National Comprehensive Cancer Network (NCCN) and the European Association of Urology (EAU) guidelines both recommend the use of nomograms, such as Briganti, MSKCC, and Partin nomograms. These nomograms are largely based on clinicopathologic characteristics that can be easily acquired in the actual clinical procedure (34, 35). These together with

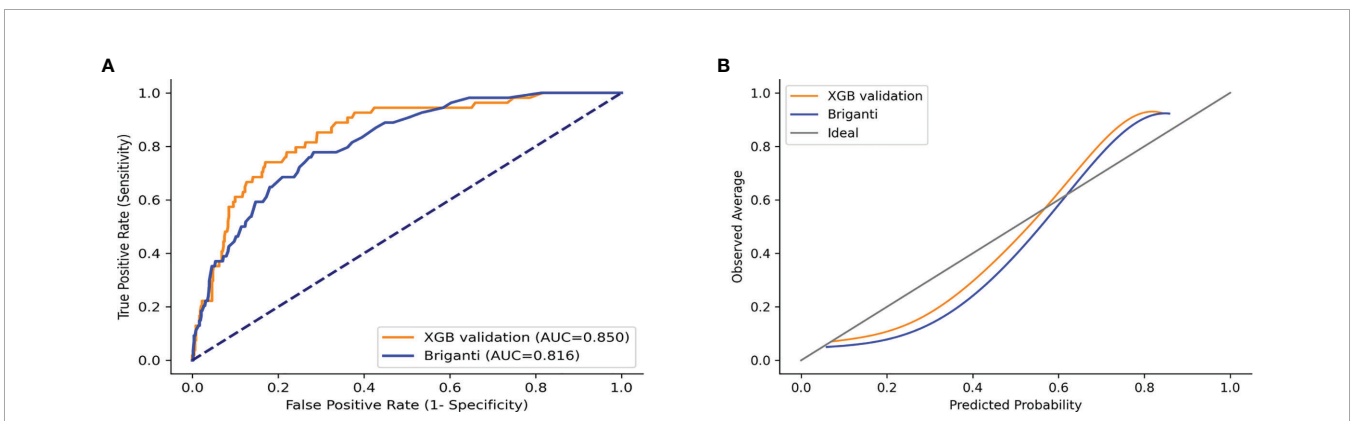


FIGURE 8 | ROC curves and calibration plots of XGB and Briganti nomogram for the validation set. **(A)** ROC curves. **(B)** Calibration plots.

TABLE 4 | Clinical and pathological characteristics of the test and validation sets.

Variables	Validation set		Test set		p-value
	NLNI (n = 481) (%)	LNI (n = 54) (%)	NLNI (n = 753) (%)	LNI (n = 113) (%)	
Age at diagnosis					0.006
<62	198 (41.2)	23 (42.6)	362 (48.1)	61 (46.0)	
≥62	283 (58.8)	31 (57.4)	391 (51.9)	52 (54.0)	
Race					0.316
Black	52 (10.8)	7 (13.0)	99 (13.2)	15 (13.4)	
Other	27 (5.6)	4 (7.4)	37 (4.9)	8 (7.0)	
White	402 (83.6)	43 (79.6)	617 (81.9)	90 (79.6)	
Marital status					0.683
Unmarried	53 (11.1)	2 (3.7)	79 (10.5)	16 (14.2)	
Married	428 (88.9)	52 (96.3)	674 (89.5)	97 (85.8)	
T stage					0.154
T1	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
T2	299 (62.2)	7 (13.0)	458 (60.8)	12 (10.6)	
T3	182 (37.8)	47 (87.0)	291 (38.7)	96 (84.9)	
T4	0 (0.0)	0 (0.0)	4 (0.5)	5 (4.5)	
Tumor size					<0.001
0–10 mm	75 (15.6)	0 (0.0)	104 (13.8)	3 (2.7)	
10–20 mm	249 (51.8)	16 (29.6)	344 (45.7)	22 (19.5)	
20–30 mm	101 (21.0)	22 (40.8)	186 (24.7)	39 (34.5)	
>30 mm	56 (11.6)	16 (29.6)	119 (15.8)	49 (43.3)	
PSA before biopsy					0.063
0–4 ng/ml	69 (14.3)	3 (5.5)	82 (10.9)	6 (5.3)	
4–10 ng/ml	312 (64.9)	25 (46.3)	495 (65.7)	43 (38.1)	
10–20 ng/ml	70 (14.6)	13 (24.1)	139 (18.5)	29 (25.7)	
>20 ng/ml	30 (6.2)	13 (24.1)	37 (4.9)	35 (30.9)	
Primary biopsy Gleason score					0.402
≤6	59 (12.3)	1 (1.9)	112 (14.9)	5 (4.4)	
7	328 (68.2)	24 (44.4)	446 (59.2)	50 (44.2)	
8	50 (10.4)	12 (22.2)	127 (16.9)	29 (25.7)	
≥9	44 (9.1)	17 (31.5)	68 (9.0)	29 (25.7)	
Positive biopsy percentage					<0.001
0–25%	151 (31.4)	12 (22.2)	233 (30.1)	9 (8.0)	
25–50%	218 (45.3)	22 (40.8)	279 (37.1)	37 (32.7)	
50–75%	73 (15.2)	4 (7.4)	145 (19.3)	26 (23.0)	
75–100%	39 (8.1)	16 (29.6)	96 (13.5)	41 (36.3)	

several clinicopathologic characteristics which can be readily acquired during the clinical procedure may predict the risk of LNI before surgery. Because of the convenience and practicality of nomograms, they are the most commonly used tools for LNI prediction. However, the performance of these old version nomograms is not always reliable, with a prediction accuracy below 80% (36), and the new version nomograms are less used due to the acquisition of some non-conventional variables (37). Hence, a more advanced prediction model based on the basic variables is needed. XGB is a gradient boosting algorithm that is one of the most powerful techniques for constructing prediction models, and it has been widely used in various medical studies (38, 39).

We found that the five independent risk factors identified by the MLR model were identical to the top 5 most important factors calculated by the XGB model, including T stage, tumor size, PSA before biopsy, GS, and PPC. GS is an evaluation method for determining the state of differentiated PCa tumor cells, and a higher score represents a less differentiated tumor. This parameter is highly correlated with the aggressiveness of the malignancy, and highly aggressive tumors progress more rapidly and are usually associated with LN metastasis. In our

study, the XGB model showed that the weight value of GS was the highest, showing the importance of this parameter and indicated that it contributed most to the results obtained (40, 41). As our results demonstrated, the XGB model assigned weighted values to all variables and arranged them by order of importance, thus allowing for more variables to be involved in the analysis and helping physicians to better understand the risk factors.

In this study, the predictive accuracy of our XGB model was the highest in both training and test sets (AUC = 0.908 and 0.883, respectively). Compared to this, the MLR model was less accurate in both training and test sets (AUC = 0.769 and 0.763, respectively). The XGB model was also more accurate than the Briganti nomogram in the validation set (AUC: 0.850 vs. 0.816). Previous studies have shown that the predictive accuracy of the Briganti nomogram was 0.798, which was close to our result (0.798 and 0.816) (42). Considering that the nomogram is a visualization tool used in the MLR model, it is established from the MLR algorithm, and this result was not surprising. In addition, the calibration plots for the validation set indicated that the XGB model has a better consistency across the 0% to 60% range of prediction probability. In actual clinical practice,

physicians are more concerned with accurately identifying LNI in low-risk PCa patients. It is exciting that our model has better consistency within the low-risk range, and conversely, the nomogram tends to underestimate the risk of LNI in this range. This indicates that our model is better to avoid omissions within the low-risk range. Our results indicate that the MLR model has a weakness in its accuracy when analyzing the linkages seen in multiple data, whereas the XGB model excels at accurately predicting outcomes from multiple unrelated datasets.

Other researchers have tried to use ML to predict LNI. Hou et al. used an ML algorithm combined with mpMRI to build a model for predicting lymph node metastasis in PCa patients. It had a very high prediction accuracy (AUC = 0.906), but the sample size was small and its practical use might be limited as mpMRI was not an easily available parameter (43). Instead, our model was established using a more advanced algorithm based on a big sample and using basic parameters. Similar results were obtained from the calibration plots in both the test and validating sets, which predicted probabilities against observed average LNI rates, indicating that the XGB model had excellent consistency with the MLR model. In addition, the XGB AUC curve obtained here was closer to the ideal line. The MLR model tended to underestimate the LNI risk across the entire range of predicted probabilities, indicating that the use of nomograms based on the MLR model might result in a higher false-negative rate and lead to some LNI patients being omitted.

The significance of determining a cutoff value for a predictive model is to guide the physician during clinical decisions, but it also implies that a certain number of patients below that cutoff value may be missed. The current EAU guidelines suggest that the indication for ePLND is based on the risk of LN metastasis >5% using the Briganti nomogram (44). Using this cutoff, our model could spare about 50% (428/866) of ePLND and only 1.2% (5/428) of patients would be missed. Compared with the MLR model, XGB had a higher positive predictive value (34.5% vs. 27.1%) and a negative predictive value (98.9% vs. 97.5%), but the percentage of overtreated patients was lower (65.5% vs. 72.9%). In addition, our model had low false-negative rates in all the 5%–20% cutoff values. Choosing 20% as the cutoff value can largely reduce the number of ePLND and the possibility of missing patients would be as low as 2.9%. These results imply that our model has a low missing rate, and more LNI patients would be identified. This is consistent with the results of previous studies (44).

Considering that physicians focus on different goals in different situations during clinical practice, DCA was developed by Vickers and Elkin to evaluate clinical effectiveness. The method uses the net benefits at different thresholds to evaluate the clinical utility (45). In our study, within a threshold of 5% to 20%, the net benefit of our XGB model was higher than the other models, which indicates that the value of benefits (i.e., the correct identification of LNI) minus the drawbacks (i.e., overtreatment and omissions) is the biggest.

However, this study still has several limitations. Firstly, the model is based on the SEER database which collects data from

the North American population, so there may be gaps in population applicability, necessitating the use of broader populations in future studies. Secondly, tumor size as determined by preoperative imaging was not available for our data and this could cause some errors. We will further improve our model with more complete external validation data in future studies. Thirdly, we excluded patients with <10 LNs examined so as to avoid patients treated with PLND. This is not the ideal definition of ePLND.

This is the first model to predict LNI in PCa patients based on standard clinicopathological features and a novel AI algorithm. Our model performed exceptionally well in predicting LNI in presurgical PCa patients and could potentially assist clinicians to make more accurate and personalized medical decisions. The practical use of this model would be to help surgeons predict the probability of LNI in PCa patients and so help guide surgical alternatives. For patients, it could make some who are more at high risk to be more vigilant, thereby improving their prognosis of PCa.

CONCLUSIONS

We established an ML model based on big data for predicting LNI in PCa patients. Our model had excellent predictive accuracy and practical clinical utility, which might help guide the decision of the urologist and help patients to improve their long-term prognosis. This study follows the trend toward precision medicine for all future patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The software packages of AI algorithms are available at Scikit-learn (<https://scikit-learn.org/stable/>) and matplotlib (<https://matplotlib.org/>). The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Medical Ethics Committee at Jinan University's First Affiliated Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

YZ, LW, and YH conceived and designed the study. YZ and LC provided administrative support. LW, ZC, and XQ collected and

assembled the data. YH, LW, HL, and LC analyzed and interpreted the data. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Natural Science Foundation of China (81902615) and the Leading Specialist Construction Project-Department of Urology, the First Affiliated Hospital, Jinan University (711006).

REFERENCES

- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer Clin* (2021) 71(1):7–33. doi: 10.3322/caac.21654
- Bill-Axelsson A, Holmberg L, Garmo H, Taari K, Busch C, Nordling S, et al. Radical Prostatectomy or Watchful Waiting in Prostate Cancer - 29-Year Follow-Up. *N Engl J Med* (2018) 379(24):2319–29. doi: 10.1056/NEJMoa1807801
- Gandaglia G, De Lorenzis E, Novara G, Fossati N, De Groot R, Dovey Z, et al. Robot-Assisted Radical Prostatectomy and Extended Pelvic Lymph Node Dissection in Patients With Locally-Advanced Prostate Cancer. *Eur Urol* (2017) 71(2):249–56. doi: 10.1016/j.eururo.2016.05.008
- Moltzahn F, Karnes J, Gontero P, Kneitz B, Tombal B, Bader P, et al. Predicting Prostate Cancer-Specific Outcome After Radical Prostatectomy Among Men With Very High-Risk Ct3b/4 PCa: A Multi-Institutional Outcome Study of 266 Patients. *Prostate Cancer Prostatic Dis* (2015) 18(1):31–7. doi: 10.1038/pcan.2014.41
- Eloranta S, Smedby KE, Dickman PW, Andersson TM. Cancer Survival Statistics for Patients and Healthcare Professionals - a Tutorial of Real-World Data Analysis. *J Intern Med* (2021) 289(1):12–28. doi: 10.1111/joim.13139
- Cheng L, Zincke H, Blute ML, Bergstralh EJ, Scherer B, Bostwick DG. Risk of Prostate Carcinoma Death in Patients With Lymph Node Metastasis. *Cancer* (2001) 91(1):66–73. doi: 10.1002/1097-0142(20010101)91:1<66::aid-cncr9>3.0.co;2-p
- Wilczak W, Wittmer C, Clauditz T, Minner S, Steurer S, Buscheck F, et al. Marked Prognostic Impact of Minimal Lymphatic Tumor Spread in Prostate Cancer. *Eur Urol* (2018) 74(3):376–86. doi: 10.1016/j.eururo.2018.05.034
- Boorjian SA, Thompson RH, Siddiqui S, Bagniewski S, Bergstralh EJ, Karnes RJ, et al. Long-Term Outcome After Radical Prostatectomy for Patients With Lymph Node Positive Prostate Cancer in the Prostate Specific Antigen Era. *J Urol* (2007) 178(3):864–71. doi: 10.1016/j.juro.2007.05.048
- Mandel P, Rosenbaum C, Pompe RS, Steuber T, Salomon G, Chun FK, et al. Long-Term Oncological Outcomes in Patients With Limited Nodal Disease Undergoing Radical Prostatectomy and Pelvic Lymph Node Dissection Without Adjuvant Treatment. *World J Urol* (2017) 35(12):1833–9. doi: 10.1007/s00345-017-2079-4
- Koerber SA, Stach G, Kratochwil C, Haefner MF, Rathke H, Herfarth K, et al. Lymph Node Involvement in Treatment-Naive Prostate Cancer Patients: Correlation of PSMA PET/CT Imaging and Roach Formula in 280 Men in Radiotherapeutic Management. *J Nucl Med* (2020) 61(1):46–50. doi: 10.2967/jnumed.119.227637
- Hatano K, Tanaka J, Nakai Y, Nakayama M, Kakimoto KI, Nakanishi K, et al. Utility of Index Lesion Volume Assessed by Multiparametric MRI Combined With Gleason Grade for Assessment of Lymph Node Involvement in Patients With High-Risk Prostate Cancer. *Jpn J Clin Oncol* (2020) 50(3):333–7. doi: 10.1093/jco/hyz170
- Gao X, Li LY, Rassler J, Pang J, Chen MK, Liu WP, et al. Prospective Study of CRMP4 Promoter Methylation in Prostate Biopsies as a Predictor For Lymph Node Metastases. *J Natl Cancer Inst* (2017) 109(6):djw282. doi: 10.1093/jnci/djw282
- Lunger L, Retz M, Bandur M, Souchay M, Vitzthum E, Jager M, et al. KLK3 and TMPRSS2 for Molecular Lymph-Node Staging in Prostate Cancer Patients Undergoing Radical Prostatectomy. *Prostate Cancer Prostatic Dis* (2021) 24(2):362–9. doi: 10.1038/s41391-020-00283-3
- Briganti A, Chun FK, Salonia A, Zanni G, Scattoni V, Valiquette L, et al. Validation of a Nomogram Predicting the Probability of Lymph Node Invasion Among Patients Undergoing Radical Prostatectomy and an Extended Pelvic Lymphadenectomy. *Eur Urol* (2006) 49(6):1019–26; discussion 1026–7. doi: 10.1016/j.eururo.2006.01.043
- Cagiannos I, Karakiewicz P, Eastham JA, Ohori M, Rabbani F, Gerigk C, et al. A Preoperative Nomogram Identifying Decreased Risk of Positive Pelvic Lymph Nodes in Patients With Prostate Cancer. *J Urol* (2003) 170(5):1798–803. doi: 10.1097/01.ju.0000091805.98960.13
- Eifler JB, Feng Z, Lin BM, Partin MT, Humphreys EB, Han M, et al. An Updated Prostate Cancer Staging Nomogram (Partin Tables) Based on Cases From 2006 to 2011. *BJU Int* (2013) 111(1):22–9. doi: 10.1111/j.1464-410X.2012.11324.x
- Oderda M, Diamand R, Albisinni S, Callaris G, Carbone A, Falcone M, et al. Indications for and Complications of Pelvic Lymph Node Dissection in Prostate Cancer: Accuracy of Available Nomograms for the Prediction of Lymph Node Invasion. *BJU Int* (2021) 127(3):318–25. doi: 10.1111/bju.15220
- May M. Eight Ways Machine Learning is Assisting Medicine. *Nat Med* (2021) 27(1):2–3. doi: 10.1038/s41591-020-01197-2
- Kather JN, Pearson AT, Halama N, Jager D, Krause J, Loosen SH, et al. Deep Learning can Predict Microsatellite Instability Directly From Histology in Gastrointestinal Cancer. *Nat Med* (2019) 25(7):1054–6. doi: 10.1038/s41591-019-0462-y
- Diller GP, Kempny A, Babu-Narayan SV, Henrichs M, Brida M, Uebing A, et al. Machine Learning Algorithms Estimating Prognosis and Guiding Therapy in Adult Congenital Heart Disease: Data From a Single Tertiary Centre Including 10 019 Patients. *Eur Heart J* (2019) 40(13):1069–77. doi: 10.1093/eurheartj/ehy915
- Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med* (2019) 380(14):1347–58. doi: 10.1056/NEJMra1814259
- Fergany A, Kupelian PA, Levin HS, Zippe CD, Reddy C, Klein EA. No Difference in Biochemical Failure Rates With or Without Pelvic Lymph Node Dissection During Radical Prostatectomy in Low-Risk Patients. *Urology* (2000) 56(1):92–5. doi: 10.1016/s0090-4295(00)00550-1
- Joslyn SA, Konety BR. Impact of Extent of Lymphadenectomy on Survival After Radical Prostatectomy for Prostate Cancer. *Urology* (2006) 68(1):121–5. doi: 10.1016/j.urology.2006.01.055
- Qi TY, Chen YQ, Jiang J, Zhu YK, Yao XH, Qi J. Contrast-Enhanced Transrectal Ultrasonography: Measurement of Prostate Cancer Tumor Size and Correlation With Radical Prostatectomy Specimens. *Int J Urol* (2013) 20(11):1085–91. doi: 10.1111/iju.12125
- Wang M, Janaki N, Buzzy C, Bukavina L, Mahran A, Mishra K, et al. Whole Mount Histopathological Correlation With Prostate MRI in Grade I and II Prostatectomy Patients. *Int Urol Nephrol* (2019) 51(3):425–34. doi: 10.1007/s11255-019-02083-8
- Delongchamps NB, Beuvon F, Eiss D, Flam T, Muradyan N, Zerbib M, et al. Multiparametric MRI is Helpful to Predict Tumor Focality, Stage, and Size in Patients Diagnosed With Unilateral Low-Risk Prostate Cancer. *Prostate Cancer Prostatic Dis* (2011) 14(3):232–7. doi: 10.1038/pcan.2011.9
- Nakashima J, Tanimoto A, Imai Y, Mukai M, Horiguchi Y, Nakagawa K, et al. Endorectal MRI for Prediction of Tumor Site, Tumor Size, and Local

ACKNOWLEDGMENTS

The authors are grateful to Dr. Dev Sooranna of Imperial College London for English language edits of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2021.763381/full#supplementary-material>

- Extension of Prostate Cancer. *Urology* (2004) 64(1):101–5. doi: 10.1016/j.urol.2004.02.036
28. von Bodman C, Godoy G, Chade DC, Cronin A, Tafe LJ, Fine SW, et al. Predicting Biochemical Recurrence-Free Survival for Patients With Positive Pelvic Lymph Nodes at Radical Prostatectomy. *J Urol* (2010) 184(1):143–8. doi: 10.1016/j.juro.2010.03.039
 29. Abdollah F, Sun M, Thuret R, Budaus L, Jeldres C, Graefen M, et al. Decreasing Rate and Extent of Lymph Node Staging in Patients Undergoing Radical Prostatectomy may Undermine the Rate of Diagnosis of Lymph Node Metastases in Prostate Cancer. *Eur Urol* (2010) 58(6):882–92. doi: 10.1016/j.eururo.2010.09.029
 30. Withrow DR, DeGroot JM, Siemens DR, Groome PA. Therapeutic Value of Lymph Node Dissection at Radical Prostatectomy: A Population-Based Case-Cohort Study. *BJU Int* (2011) 108(2):209–16. doi: 10.1111/j.1464-410X.2010.09805.x
 31. Fossati N, Passoni NM, Moschini M, Gandaglia G, Larcher A, Freschi M, et al. Impact of Stage Migration and Practice Changes on High-Risk Prostate Cancer: Results From Patients Treated With Radical Prostatectomy Over the Last Two Decades. *BJU Int* (2016) 117(5):740–7. doi: 10.1111/bju.13125
 32. Ploussard G, Briganti A, de la Taille A, Haese A, Heidenreich A, Menon M, et al. Pelvic Lymph Node Dissection During Robot-Assisted Radical Prostatectomy: Efficacy, Limitations, and Complications—a Systematic Review of the Literature. *Eur Urol* (2014) 65(1):7–16. doi: 10.1016/j.eururo.2013.03.057
 33. Budiharto T, Joniau S, Lerut E, Van den Bergh L, Mottaghly F, Deroose CM, et al. Prospective Evaluation of 11C-Choline Positron Emission Tomography/Computed Tomography and Diffusion-Weighted Magnetic Resonance Imaging for the Nodal Staging of Prostate Cancer With a High Risk of Lymph Node Metastases. *Eur Urol* (2011) 60(1):125–30. doi: 10.1016/j.eururo.2011.01.015
 34. Mohler JL, Antonarakis ES, Armstrong AJ, D'Amico AV, Davis BJ, Dorff T, et al. Prostate Cancer, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* (2019) 17(5):479–505. doi: 10.6004/jnccn.2019.0023
 35. Mottet N, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, 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):243–62. doi: 10.1016/j.eururo.2020.09.042
 36. Cimino S, Reale G, Castelli T, Favilla V, Giardina R, Russo GI, et al. Comparison Between Briganti, Partin and MSKCC Tools in Predicting Positive Lymph Nodes in Prostate Cancer: A Systematic Review and Meta-Analysis. *Scand J Urol* (2017) 51(5):345–50. doi: 10.1080/21681805.2017.1332680
 37. Gandaglia G, Ploussard G, Valerio M, Mattei A, Fiori C, Fossati N, et al. A Novel Nomogram to Identify Candidates for Extended Pelvic Lymph Node Dissection Among Patients With Clinically Localized Prostate Cancer Diagnosed With Magnetic Resonance Imaging-Targeted and Systematic Biopsies. *Eur Urol* (2019) 75(3):506–14. doi: 10.1016/j.eururo.2018.10.012
 38. Huang Y, Chen H, Zeng Y, Liu Z, Ma H, Liu J. Development and Validation of a Machine Learning Prognostic Model for Hepatocellular Carcinoma Recurrence After Surgical Resection. *Front Oncol* (2020) 10:593741. doi: 10.3389/fonc.2020.593741
 39. Xiang L, Wang H, Fan S, Zhang W, Lu H, Dong B, et al. Machine Learning for Early Warning of Septic Shock in Children With Hematological Malignancies Accompanied by Fever or Neutropenia: A Single Center Retrospective Study. *Front Oncol* (2021) 11:678743. doi: 10.3389/fonc.2021.678743
 40. Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part B: Prostate and Bladder Tumours. *Eur Urol* (2016) 70(1):106–19. doi: 10.1016/j.eururo.2016.02.028
 41. Egevad L, Delahunt B, Yaxley J, Samarasinghe H. Evolution, Controversies and the Future of Prostate Cancer Grading. *Pathol Int* (2019) 69(2):55–66. doi: 10.1111/pin.12761
 42. Bandini M, Marchioni M, Pompe RS, Tian Z, Gandaglia G, Fossati N, et al. First North American Validation and Head-to-Head Comparison of Four Preoperative Nomograms for Prediction of Lymph Node Invasion Before Radical Prostatectomy. *BJU Int* (2018) 121(4):592–9. doi: 10.1111/bju.14074
 43. Hou Y, Bao ML, Wu CJ, Zhang J, Zhang YD, Shi HB. A Machine Learning-Assisted Decision-Support Model to Better Identify Patients With Prostate Cancer Requiring an Extended Pelvic Lymph Node Dissection. *BJU Int* (2019) 124(6):972–83. doi: 10.1111/bju.14892
 44. Malkiewicz B, Ptaszkowski K, Knecht K, Gurwin A, Wilk K, Kielb P, et al. External Validation of the Briganti Nomogram to Predict Lymph Node Invasion in Prostate Cancer-Setting a New Threshold Value. *Life (Basel)* (2021) 11(6):479. doi: 10.3390/life11060479
 45. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wei, Huang, Chen, Lei, Qin, Cui and Zhuo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.