Original Article
Special Topic: Advances in BPH Research
Current Urology
OPEN

Clinical-radiomics combination model for predicting the short-term efficacy of bipolar transurethral enucleation of the prostate in patients with benign prostatic hyperplasia

Tianyou Zhang[a,b,#], Zijun Mo[a,b,#], Jiayu Huang[a,b,#], Jun Wang[c,#], Yiran Tao[a,b], Lei Ye[a,b], Wenwen Zhong[a,b], Bing Yao[a,b], Hu Qu[a,b], Bo Ma[a,b], Dejuan Wang[a,b], Jiahui Mo[d], Chunwei Ye[e,*], Junying Zhu[b,f,*], Jianguang Qiu[a,b,*]

[a]Department of Urology, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China; [b]Department of Urology, Biomedical Innovation Center, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China; [c]State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, China; [d]Department of Urology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China; [e]Department of Urology, The Second Affiliated Hospital of Kunming Medical University, Kunming, China; [f]Department of Radiology, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

Abstract

Background: Bipolar transurethral enucleation of the prostate (B-TUEP) is a well-established surgical treatment for benign prostatic hyperplasia (BPH); however, its efficacy may vary depending on patient characteristics. Magnetic resonance imaging (MRI) with radiomics analysis can offer comprehensive and quantitative information about prostate characteristics that may relate to surgical outcomes. This study aimed to explore the value of MRI and radiomics analysis in predicting the short-term efficacy of B-TUEP for BPH.

Materials and methods: A total of 137 patients with BPH who underwent B-TUEP at 2 institutions were included. Radiological features were measured in the MRIs, and the radiomics score was developed from 1702 radiomics features extracted from the prostate and transitional zone regions of interest. Three prediction models were developed and validated based on clinical-radiological features, radiomic features, and their combinations. The models were evaluated using the area under the receiver operating characteristic curve, calibration curve, and decision curve analysis.

Results: The combination model exhibited the highest area under curve in both the training set (0.838) and the external validation set (0.802), indicating superior predictive performance and robustness. Furthermore, the combination model demonstrated good calibration ($p > 0.05$) and optimal clinical utility. The combination model indicated that a higher maximum urine flow rate, lower transitional zone index, and higher radiomics score were associated with an increased risk of poor efficacy.

Conclusions: Magnetic resonance imaging with radiomic analysis can offer valuable insights for predicting the short-term efficacy of B-TUEP in patients with BPH. A combination model based on clinical and radiomics features can assist urologists in making more precise clinical decisions.

Keywords: Benign prostatic hyperplasia; Radiomics; Prediction model; Surgical efficacy; Transurethral enucleation of the prostate

1. Introduction

Benign prostatic hyperplasia (BPH) is prevalent among middle-aged and elderly men in urological clinical practice. For patients with BPH unresponsive to lifestyle changes and medication, alongside severe urinary obstruction, surgical intervention is typically warranted.[1,2] As medical technology advances and the principles of "minimally invasive" and "enhanced recovery after surgery" gain broader acceptance, endoscopic procedures have emerged as the primary approach for managing benign prostatic obstruction (BPO), notably transurethral resection of the prostate (TURP) and transurethral enucleation of the prostate (TUEP).

In clinical practice, treatment strategies primarily hinge on the severity of lower urinary tract symptoms (LUTSs), rather than solely on the etiology of BPH, characterized by prostate enlargement.[3] Lower urinary tract symptoms have a multifaceted etiology. Despite surgical interventions effectively addressing BPH as the underlying cause, some patients may not experience optimal postoperative improvement in LUTSs or sustained therapeutic effects, even after successful removal of the enlarged prostate via surgery.

*Corresponding Author: Jianguang Qiu, Department of Urology, The Sixth Affiliated Hospital, Sun Yat-sen University, No. 26 Yuancun Erheng Rd, Guangzhou, 510655, Guangdong, China. E-mail address: qiujg@mail.sysu.edu.cn (J. Qiu); Junying Zhu, Department of Radiology, The Sixth Affiliated Hospital, Sun Yat-sen University, No. 26 Yuancun Erheng Road, Guangzhou, 510655, Guangdong, China. E-mail address: zhujy65@mail.sysu.edu.cn (J. Zhu); Chunwei Ye, Department of Urology, The Second Affiliated Hospital of Kunming Medical University, No. 374 Dianmian Road, Kunming, 653100, Yunnan, China. E-mail address: yechunwei@kumu.edu.cn (C. Ye).

#Tianyou Zhang, Zijun Mo, Jiayu Huang, and Jun Wang have contributed equally to this work and share first authorship.

Supplemental Digital Content is available for this article.

Current Urology, (2025) 19, 1, 30–38

Received August 19, 2023; Accepted February 29, 2024.

http://dx.doi.org/10.1097/CU9.0000000000000256

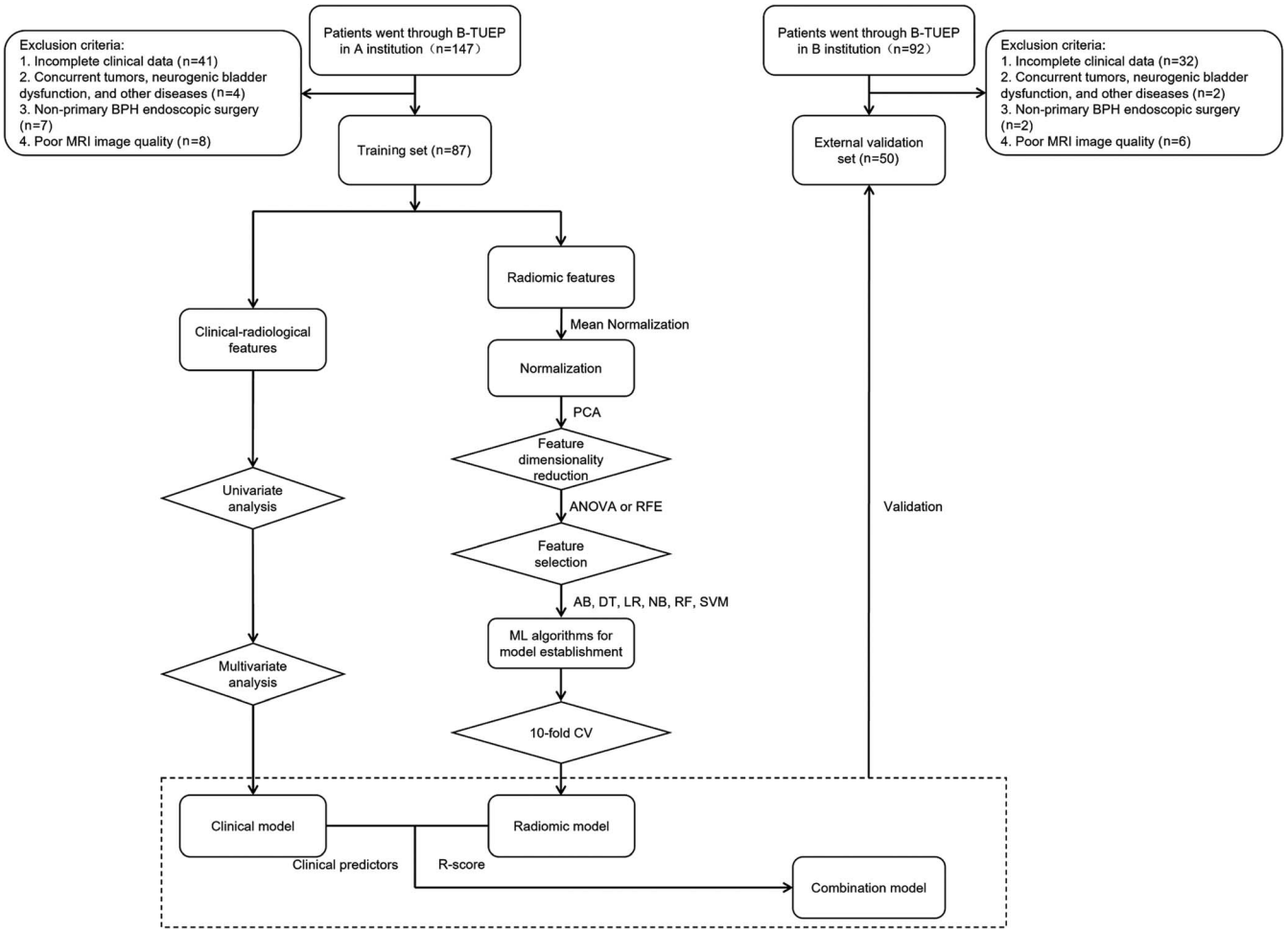

Figure 1. Workflow of patients' selection, and models development and validation. AB = Adaptive Boosting; ANOVA = analysis of variance; BPH = benign prostatic hyperplasia; B-TUEP = bipolar transurethral enucleation of the prostate; CV = cross-validation; DT = Decision Tree; LR = Logistic Regression; ML = machine learning; MRI = magnetic resonance imaging; NB = Naive Bayes; PCA = principle component analysis; RF = Random Forest; RFE = recursive feature elimination; SVM = Support Vector Machines.

Transrectal ultrasound (TRUS) serves as a common and effective diagnostic tool for BPH; however, its utility is hindered by poor interreader reliability and limited diagnostic information.[4,5] Although magnetic resonance imaging (MRI) is not obligatory for BPH assessment, it offers advantages such as superior soft tissue contrast and high-resolution imaging, thereby enhancing measurement precision.[6] Moreover, magnetic resonance (MR) radiomic analysis, leveraging machine learning, can comprehensively assess the anatomical and tissue characteristics of the prostate.

Radiomics involves the transformation of medical images (eg, computed tomography and MR) into high-throughput quantitative data, amenable to in-depth analysis.[7,8] By correlating clinical images with clinical, histological, and biological features, it becomes feasible to develop or refine tools supporting clinical decision making.[9] Although extensive research demonstrates the value of multimodal analysis integrating radiomics in clinical decision-making, its application in surgical decision-making for BPH remains unexplored.

This study investigates the predictive value of MRI in assessing the therapeutic efficacy of bipolar transurethral enucleation of the prostate (B-TUEP). In addition, we aim to establish and assess a predictive model for the short-term therapeutic efficacy of B-TUEP preoperatively, thereby facilitating enhanced surgical decision-making.

2. Material and methods

2.1. Patients

This study was conducted following approval from the ethics review committees of 2 medical institutions. Given its retrospective nature, the necessity for obtaining informed consent from patients was waived.

The study recruited patients diagnosed with BPH who underwent B-TUEP performed by experienced surgical teams at the Sixth Affiliated Hospital of Sun Yat-sen University (referred to as institution A, the training set) and the Second Affiliated Hospital of Kunming Medical University (referred to as institution B, the external validation set) between January 2019 and July 2022. Exclusion criteria comprised the following: (1) incomplete clinical data, including missing preoperative tests, urodynamic examination, MRI examination within 14 days before surgery, and postoperative follow-up data; (2) concurrent bladder cancer, prostate cancer, neurogenic bladder dysfunction, active bleeding, uncontrolled hypertension, or diabetes; (3) nonprimary BPH endoscopic surgery; and (4) poor MRI image quality. The workflow diagram is depicted in Figure 1.

For this investigation, we retrospectively gathered preoperative clinical data and T2-weighted MRI image data of patients, which included age; body mass index; preoperative International Prostate

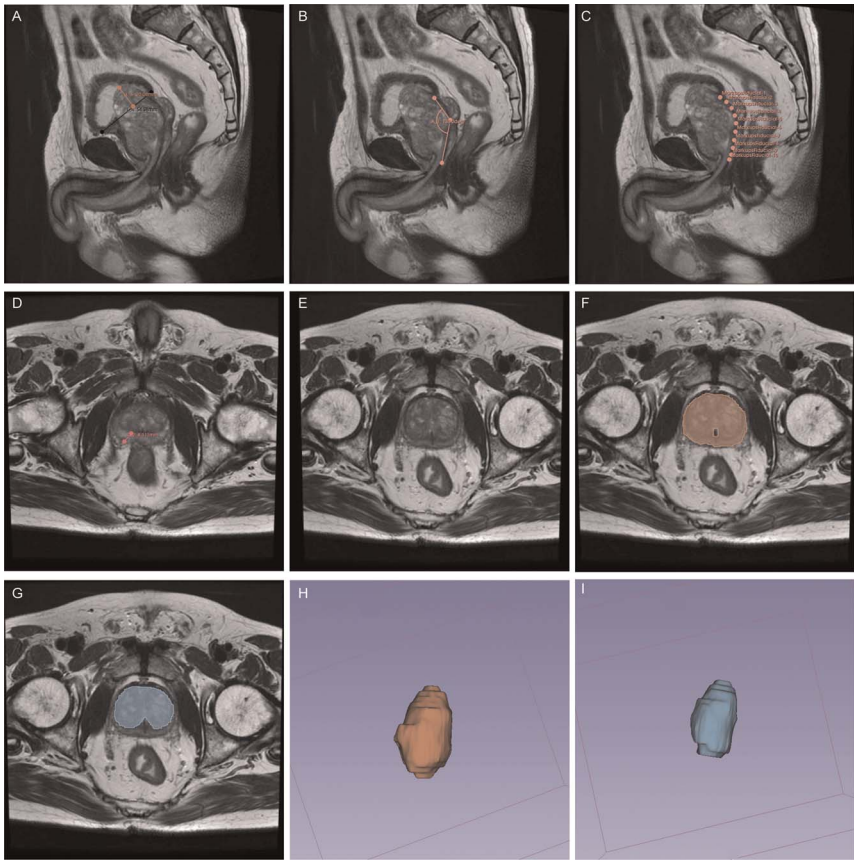

Figure 2. Schematic diagram of radiographic measurement and image segmentation. (A) Intravesical prostatic protrusion. (B) Prostatic urethral angulation. (C) Urethral length. (D) Peripheral zone thickness. (E) Original image of prostate MRI. (F) Region of interest delimitation of prostate. (G) Region of interest delimitation of TZ. (H) Three-dimensional model of prostate. (I) Three-dimensional model of transitional zone. MRI = magnetic resonance imaging; TZ = transition zone.

Symptom Score (IPSS) and quality of life (QoL); history of medication treatment for BPH, hypertension, and diabetes; presence of bladder stones and urinary system infection; preoperative total prostate-specific antigen, free prostate-specific antigen levels, maximum flow rate (Q_{max}), and postvoid residual volume (PVR); clinical-radiological features; and symmetric enlargement of bilateral lobes.

2.2. Radiographic measurement

The features extracted from the original T2 weight image (T2WI) MRI files imported into 3D Slicer (Brigham and Women's Hospital, USA, version 5.2.0, https://www.slicer.org/) encompassed intravesical prostatic protrusion (IPP), peripheral zone thickness (PT), prostatic urethral angulation, and urethral length (UL) (Fig. 2). Prostate volume (PV) and transitional zone volume (TZV) were derived from the radiomics feature "Mesh Volume" of the regions of interest (ROIs) within the prostate and transitional zone (TZ), respectively. The transitional zone index (TZI) was computed as the ratio of TZV to PV. Intravesical prostatic protrusion was defined as the vertical distance from the highest point of the prostate protrusion into the bladder in the sagittal plane to the level of the bladder neck. Peripheral zone thickness was defined as the greatest distance between the inner and outer edges of the peripheral zone when drawing a straight line from the center of the TZ to the outer edge of the peripheral zone in a cross-sectional view, illustrating the maximum TZ area.[10] Prostatic urethral angulation was characterized as the angle formed between the proximal and distal urethra, as indicated by the verumontanum. Urethral length represented the length of the urethra from the bladder neck to the tip of the prostate, namely, the prostatic urethra. Using the "Curve Maker" extension plugin in the software, a minimum of 10 anchor points were delineated along the prostatic urethra in the sagittal image, and the length of the curve connecting these points was measured as UL.

2.3. Follow-up

All patients underwent follow-up 4 to 6 weeks after surgery, with assessments including IPSS, QoL scores, Q_{max}, and PVR.

The evaluation criteria for efficacy were based on the standards developed by Homma et al.[11] (Supplementary Table 1, http://links.lww.com/CURRUROL/A60), encompassing symptoms (IPSS), function (Q_{max}), and QoL (QoL score). The outcome of this study was a binary variable (0, good; 1, poor). Good efficacy was defined as 2 or more indicators categorized as excellent or good; otherwise, efficacy was considered poor.

2.4. Clinical analysis

To compare the characteristics and differences between the 2 groups, Student *t* test or Mann-Whitney rank-sum test was used for continuous variables, whereas Pearson χ^2 test or Fisher exact test was used for categorical variables. Univariate logistic regression analysis was conducted to ascertain the significance of each feature in the training set to identify potential risk factors associated with the efficacy of B-TUEP. Features showing significant associations with the efficacy of B-TUEP were included in a stepwise multivariate logistic analysis to explore their predictive potential. Statistical analyses were performed using SPSS software (version 25.0; IBM, Chicago, IL), with significance defined as a p value of <0.05.

2.5. Radiomic analysis

Retrospectively collected T2WI sequence images, including transverse, coronal, and sagittal planes, were exported in the Digital Imaging and Communications in Medicine format.

To investigate the influence of both the entire prostate and TZ on surgical efficacy simultaneously, separate ROIs were delineated for the prostate and TZ for image segmentation. The original MRI images were imported into 3D Slicer software, and the prostate and TZ ROIs were manually delineated layer by layer in the transverse plane, with verification and adjustments made in the coronal and sagittal planes. Region of interest segmentation was performed by an experienced urologist with training in prostate MR and corrected by a senior radiologist. Radiomic features were extracted from each ROI using the Pyradiomics package in Python (Python Software Foundation, Netherlands, version 3.7.6, https://www.python.org), which included 107 original and 744 wavelet features (https://pyradiomics.readthedocs.io/en/latest/features.html). The scans were resampled to a voxel size of 1 mm × 1 mm × 1 mm to minimize variability among different images. The entire radiomic analysis process was integrated into an open-source software package for the FAE software deposited on the Github Web site (Shanghai Key Laboratory of Magnetic Resonance Imaging, East China Normal University, Version 0.5.4, https://github.com/salan668/FAE).

Mean normalization was used to eliminate the adverse effects caused by singular sample data and enhance comparability between feature vectors. Principal component analysis (PCA) was used for feature dimensionality reduction, transforming potentially correlated original feature vectors into a set of independent feature vectors through orthogonal transformation. Feature selection was performed using analysis of variance (ANOVA) and recursive feature elimination methods.

2.6. Development, evaluation, and validation of prediction models

In the clinical prediction model, variables with a p value of <0.1 were selected as predictors based on the results of multivariate analysis. Binary logistic regression was performed using these predictors to obtain a logistic regression probability formula, and a nomogram was plotted accordingly.

For the radiomic prediction model, models were developed for the radiomic features of the prostate, TZ, and their combination. Six commonly used machine learning algorithms—Adaptive Boosting, Decision Tree, Logistic Regression (LR), Naive Bayes, Random Forest, and Support Vector Machines—were used to develop radiomic models using features selected through ANOVA or recursive feature elimination. In total, there were 36 modeling method combinations (Supplementary Table 2, http://links.lww.com/CURRUROL/A60). A 10-fold cross-validation (CV) was applied during the training process to mitigate overfitting and selection bias. The methodology with the best CV performance (with the maximum area under curve [AUC]) was selected as the most robust model. The training set was used to train the final model using this combination of methodologies, resulting in a radiomic score (R score).

For the clinical radiomics combination model, we combined the R score as an independent variable with the clinical model variables, used logistic regression to develop the model, and plotted a prediction nomogram.

The performance of the 3 models was evaluated using the area under the receiver operating characteristic (ROC) curve, which can account for variations in the true-positive rate and false-positive rate, providing a comprehensive assessment of the model. Calibration curves were plotted using the Hosmer-Lemeshow test, and the predictive ability of the model was evaluated using the

Table 1

Univariate and multivariate logistic regression analysis for B-TUEP efficacy about clinical-radiological features in training set.

	Univariate		Multivariate	
Variable	**OR (95% CI)**	***p***	**OR (95% CI)**	***p***
Age, yr	0.951 (0.893–1.012)	0.115		
BMI, kg/m^2	1.029 (0.878–1.206)	0.726		
Preoperative IPSS	0.887 (0.806–0.975)	**0.013***	0.927 (0.828–1.038)	0.191
Preoperative QoL	0.867 (0.538–1.398)	0.558		
Medication treatment (yes)	1.503 (0.559–2.447)	0.397		
Hypertension (yes)	1.108 (0.430–2.857)	0.832		
Diabetes (yes)	0.665 (–0.716 to 2.046)	0.563		
Bladder calculi (yes)	1.007 (0.339–2.996)	0.990		
Urinary infection (yes)	0.690 (–0.247 to 1.626)	0.437		
Preoperative t-PSA, ng/mL	0.937 (0.847–1.028)	0.160		
Preoperative f-PSA, ng/mL	0.725 (0.264–1.186)	0.172		
Q_{max}, mL/s	1.152 (1.053–1.250)	**0.005***	1.125 (1.002–1.263)	**0.046***
PVR, mL	0.997 (0.992–1.001)	0.152		
PV, mL	0.975 (0.956–0.994)	**0.010***	0.951 (0.845–1.070)	0.406
TZV, mL	0.967 (0.944–0.990)	**0.004***	1.074 (0.927–1.244)	0.342
TZI, %	0.928 (0.891–0.965)	**<0.001***	0.910 (0.826–1.001)	**0.054**†
IPP, mm	0.870 (0.779–0.961)	**0.003***	0.948 (0.853–1.053)	0.321
PUA, degree	1.016 (0.978–1.053)	0.419		
UL, mm	0.950 (0.909–0.991)	**0.014***	0.992 (0.934–1.054)	0.797
PT, mm	1.225 (1.071–1.378)	**0.010***	1.024 (0.806–1.301)	0.847
Symmetric enlargement (no)	2.218 (0.856–5.752)	0.101		

B-TUEP = bipolar transurethral enucleation of the prostate; BMI = body mass index; CI = confidence interval; f-PSA = free prostate-specific antigen; IPP = intravesical prostatic protrusion; OR = odds ratio; PT = peripheral zone thickness; PUA = prostatic urethral angulation; PV = prostate volume; PVR = postvoid residual; Q_{max} = maximum flow-rate; QoL = quality of life; t-PSA = total prostate-specific antigen; TZI = transitional zone index; TZV = transitional zone volume; UL = urethral length.

Boldface format is the p values with significant difference.

*$p < 0.05$.

†$p < 0.1$.

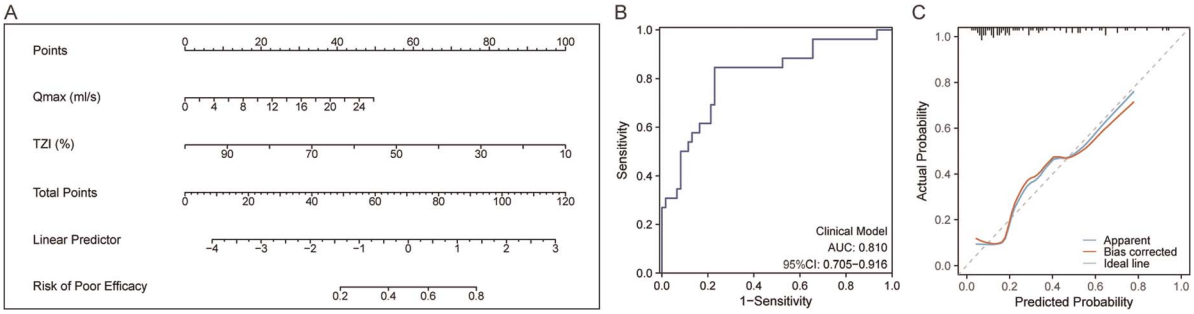

Figure 3. Nomogram and predictive performance of clinical prediction model. (A) Nomogram of clinical prediction model. (B) Receiver operating characteristic curve for clinical prediction model. (C) Calibration curve for clinical prediction model. The AUC of 0.810 indicated good predictive performance, and the model fitted well as the *p* value >0.05 of goodness-of-fit test. AUC = area under the curve; CI = confidence interval; TZI = transitional zone index; Qmax = maximum flow-rate.

concordance index (C-index). A decision curve was constructed to evaluate the clinical usefulness and net benefits of the model.

To validate and compare the robustness of the prediction models, the AUCs were calculated to evaluate the performances of the 3 models in the external validation set.

3. Results

3.1. Baseline clinical-radiological characteristics

A total of 137 patients with BPH who underwent B-TUEP were included in this study. The training set comprised 87 patients from institution A, whereas the external validation set comprised 50 patients from institution B. Within the training set, 26 patients exhibited poor efficacy (29.9%), whereas, in the validation set, 15 patients showed poor efficacy (30.0%). Patients with poor efficacy demonstrated lower preoperative IPSS, PV, TZV, TZI, IPP, and UL and higher Q_{max} and PVR. The risk coefficients estimated through univariate analysis are summarized in Table 1.

3.2. Development of clinical prediction model

Factors potentially influencing the short-term efficacy of B-TUEP (preoperative IPSS, Q_{max}, PV, TZV, TZI, IPP, UL, and PT) were subjected to multivariate analysis (Table 1), which identified Q_{max} and TZI as independent predictors of short-term efficacy. An increased risk of poor efficacy was associated with higher Q_{max} (odds ratio, 1.125; $p = 0.046$) and lower TZI (odds ratio, 0.910; $p = 0.054$). Using short-term efficacy of B-TUEP (0, good; 1, poor)

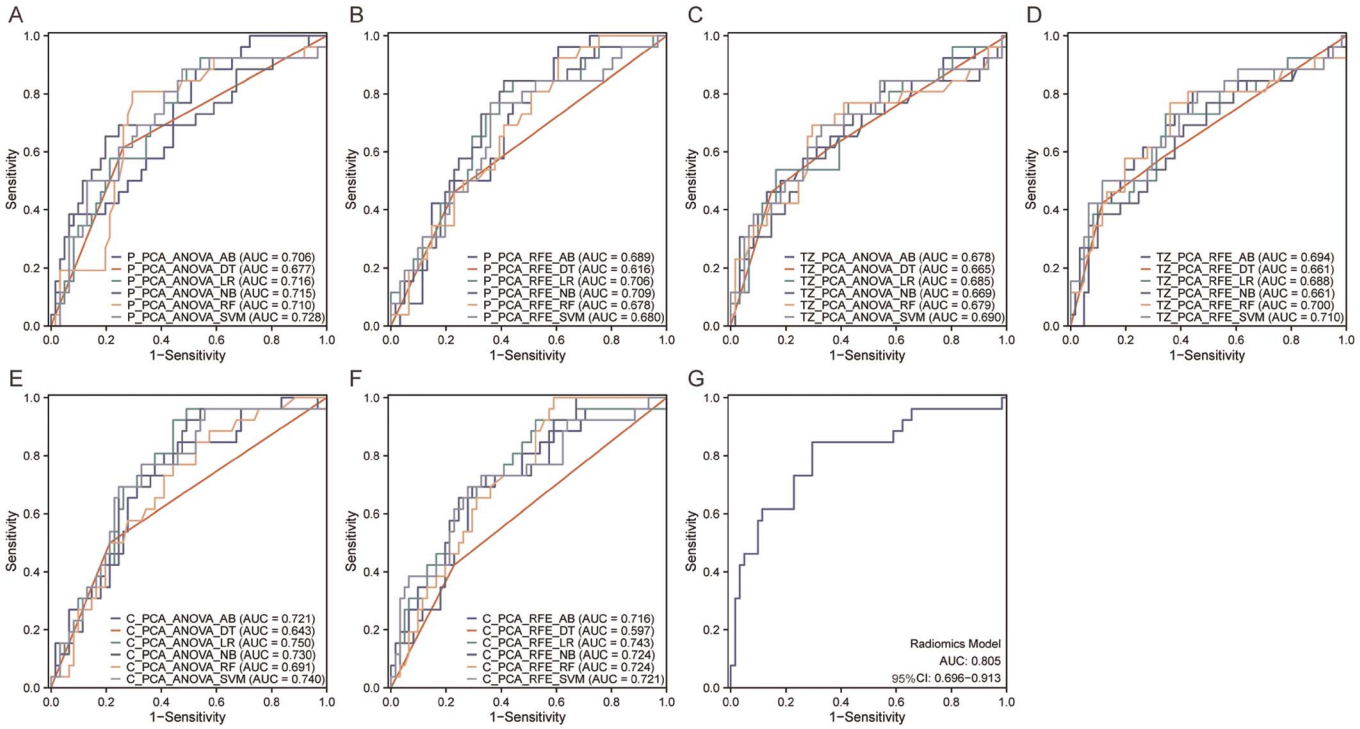

Figure 4. Ten-fold CV and predictive performance of radiomics model. (A–F) Receiver operating characteristic curves for 36 modeling method combinations in 10-fold CV. (G) Receiver operating characteristic curve for radiomics model. Among the 36 combinations, the C_PCA_ANOVA_LR model had the highest AUC of 0.749, so that, finally, the radiomics model was developed with features extracted from combination of ROIs of prostate and TZ, PCA for dimensionality reduction, ANOVA for features selection, and logistic regression algorithms for modeling. ANOVA = analysis of variance; AUC = area under the curve; CV = cross-validation; CI = confidence interval; PCA = principal component analysis; ROI = region of interest; TZ = transitional zone.

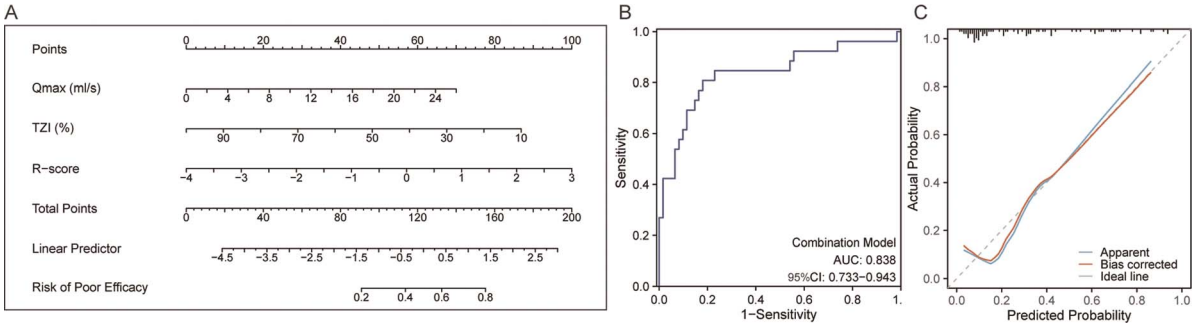

Figure 5. Nomogram and predictive performance of combination prediction model. (A) Nomogram of combination prediction model. (B) Receiver operating characteristic curve for combination prediction model. (C) Calibration curve for combination prediction model. The AUC of 0.838 indicated good predictive performance, and the model fitted well as the p value <0.05 of goodness-of-fit test. AUC = area under the curve; CI = confidence interval; Qmax = maximum flow-rate; TZI = transitional zone index.

as the outcome and Q_{max} and TZI as predictors, a predictive probability formula was established as follows:

$$\text{Logit}(p) = 2.629 + 0.123 \times Q_{\max} - 0.072 \times \text{TZI}$$

The nomogram of the clinical prediction model is presented in Figure 3A. To use it, identify the points on the scale corresponding to Q_{max} and TZI values of the patient, draw a straight line to obtain the scores, sum the scores of each variable to obtain the total score, draw a straight line downward from the point on the scale corresponding to the total score, and determine the predictive probability of poor efficacy of B-TUEP where it intersects the scale for risk probability.

The ROC curve of the clinical model in the training set is depicted in Figure 3B, with an AUC of 0.810 indicating good predictive performance. The calibration curve is illustrated in Figure 3C, demonstrating that the model exhibited good predictive ability, with a C-index of 0.810 (95% confidence interval, 0.704–0.916) compared with the ideal model. The model fit well, as evidenced by the goodness-of-fit test (p = 0.393, <0.05).

3.3. Development of radiomic prediction model

A summary of the ROC curves of the 10-fold CV for the 36 modeling method combinations is shown in Figures 4A to F. Among the 36 combinations, the C_PCA_ANOVA_LR model had the highest AUC of 0.750, indicating that the model developed by combining the ROI, PCA for dimensionality reduction, ANOVA for feature selection, and the LR algorithm for modeling had the most robust predictive performance in the training set. All training set samples were used to develop C_PCA_ANOVA_LR as the final radiomic model, which included PCA_feature_19, PCA_feature_2, and PCA_feature_87 as predictors and a radiomic signature was obtained. The R score formula is as follows:

$$\text{R\_score} = -1.071 + 0.1303 \times \text{PCA}_{\text{feature}_{19}} - 0.313 \times \text{PCA}_{\text{feature}_2} + 1.204 \times 10^{-14} \times \text{PCA}_{\text{feature}_{87}}$$

The ROC curve of the radiomic model in the training set is shown in Figure 4G, with an AUC of 0.805, which indicated good predictive performance.

Table 2

Baseline characteristics of the patients in training set and validation set.

Variable	Training set	Validation set	$t/\chi^2/Z$	p
No. cases	87	50		
Efficacy, n (%)			χ^2, 0.00	0.99
Good	61 (70.1)	35 (70.0)		
Poor	26 (29.9)	15 (30.0)		
Age, yr	69.08 ± 7.69	69.96 ± 7.79	t, −0.64	0.52
BMI, kg/m^2	22.59 ± 2.92	22.64 ± 2.44	t, −0.10	0.92
Preoperative IPSS	20.17 ± 5.36	20.78 ± 4.3	t, −0.68	0.50
Preoperative QoL	4.09 ± 0.96	4.02 ± 1.17	t, 0.37	0.71
Medication treatment, n (%)			χ^2, 0.26	0.61
Yes	31 (35.6)	20 (40.0)		
None	56 (64.4)	30 (60.0)		
Hypertension, n (%)			χ^2, 0.69	0.41
Yes	32 (36.8)	22 (44.0)		
No	55 (63.2)	28 (56.0)		
Diabetes, n (%)			χ^2, 0.58	0.45
Yes	13 (14.9)	10 (20.0)		
No	74 (85.1)	40 (80.0)		
Bladder calculi, n (%)			χ^2, 0.82	0.37
Yes	13 (14.9)	10 (20.0)		
No	67 (77.0)	35 (70.0)		
Urinary infection, n (%)			χ^2, 2.92	0.09
Yes	39 (44.8)	30 (60.0)		
No	48 (55.2)	20 (40.0)		
Preoperative t-PSA, ng/mL	7.15 ± 6.13	9.99 ± 7.35	t, −2.42	**0.02***
Preoperative f-PSA, ng/mL	1.64 ± 1.89	1.58 ± 1.07	t, 0.23	0.82
Q_{max}, mL/s	7.24 ± 5.01	7.1 ± 3.63	t, 0.17	0.87
PVR, mL	129.14 ± 115.81	146.3 ± 125.09	t, −0.81	0.42
PV, mL	73.46 ± 37	78.03 ± 41.97	t, −0.66	0.51
TZV, mL	51.24 ± 34.97	56.72 ± 38.73	t, −0.85	0.40
TZI, mL	64.60 ± 15.70	67.63 ± 13.57	t, −1.14	0.26
IPP, mm	12.85 ± 7.48	12.23 ± 7.17	t, 0.48	0.63
PUA, degree	134.33 ± 12.47	134.75 ± 13.48	t, −0.18	0.85
UL, mm	54.18 ± 13.40	53.38 ± 11.26	t, 0.35	0.72
PT, mm	9.39 ± 3.21	10.08 ± 2.36	t, −1.32	0.19
Symmetric enlargement, n (%)			χ^2, 1.97	0.16
Yes	58 (66.7)	39 (78)		
No	29 (33.3)	11 (22.0)		

BMI = body mass index; f-PSA = free prostate-specific antigen; IPSS = International Prostate Symptom Score; IPP = intravesical prostatic protrusion; PSA = prostate-specific antigen; PT = peripheral zone thickness; PUA = prostatic urethral angulation; PV = prostate volume; PVR = postvoid residual; Q_{max} = maximum flow-rate; QoL = quality of life; t-PSA = total prostate-specific antigen; TZI = transitional zone index; TZV = transitional zone volume; UL = urethral length; PT = peripheral zone thickness; PVR = postvoid residual volume.
Boldface format is the p values with significant difference.
Fisher exact test was used when the number of events was <5.
*$p < 0.05$.

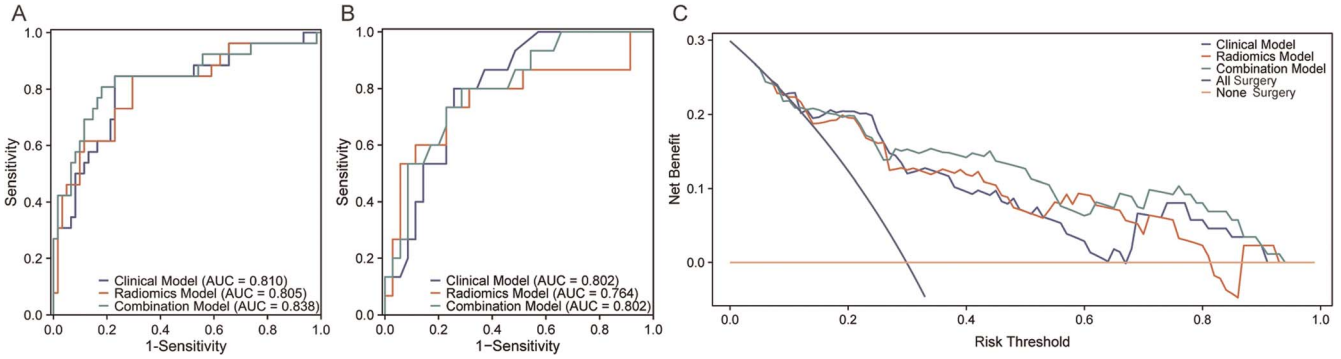

Figure 6. Comparison of the 3 prediction models. (A) Receiver operating characteristic curves for the 3 prediction models in training set. (B) Receiver operating characteristic curves for the 3 prediction models in validation set. (C) Decision curve analysis for the 3 prediction models. The combined model had highest AUC both in the training set and the validation set, which indicated that the model combined clinical and radiomic features had the best prediction ability and robust. AUC = area under the curve.

3.4. Development of combination prediction model

Using the short-term efficacy of B-TUEP (0, good; 1, poor) as the outcome and the Q_{max}, TZI, and R score as predictors, a predictive probability formula was established as follows:

$$\text{Logit}(p) = 1.319 + 0.116 \times Q_{\text{max}} - 0.042 \times \text{TZI} + 0.616 \times \text{R\_score}$$

The nomogram of the clinical prediction model is shown in Figure 5A. The combination model indicated that an increased risk of poor efficacy was associated with a higher Q_{max}, lower TZI, and higher R scores.

The ROC curve of the combination model in the training set is shown in Figure 5B, with an AUC of 0.838, which indicated good predictive performance. The calibration curve is shown in Figure 5C, which indicates that the model has good predictive ability, with a C-index of 0.838 (95% confidence interval, 0.733–0.943) compared with the ideal model, and the model fitted well, as indicated by the goodness-of-fit test (p = 0.796, >0.05).

3.5. Validation and performance comparison of models

The baseline characteristics of patients in both the training and validation sets exhibited no significant differences, except for preoperative total prostate-specific antigen levels, as outlined in Table 2. The ROC curves summarizing the performance of the 3 prediction models in the training set are depicted in Figure 6A, whereas those for the external validation set are presented in Figure 6B. Notably, the combined model demonstrated the highest AUC in both the training set (0.838 > 0.810 > 0.805) and the external validation set (0.802 = 0.802 > 0.764), underscoring the superior predictive capability and robustness of the model integrating clinical and radiomic features.

3.6. Clinical usefulness of the prediction models

The clinical utility of prediction models was evaluated using decision curves, as depicted in Figure 6C. The analysis revealed that, when the risk threshold exceeded approximately 15%, all 3 models outperformed the binary decisions of "all surgery" or "none surgery." Furthermore, when the threshold surpassed about 28%, the combined model demonstrated superior net benefit compared with the individual models.

4. Discussion

This study investigated the correlation between clinical, radiological, and radiomic characteristics and surgical outcomes in BPH patients undergoing B-TUEP. We devised and validated a predictive model merging clinical and radiomic features to aid in clinical decision making regarding BPH surgical interventions.

In patients with LUTSs stemming from BPH, prostate enlargement induces anatomical obstruction of the bladder outlet, instigating pathophysiological alterations. Conversely, voiding and storage functions are regulated by a constellation of structures including the prostate, bladder, urethra, pelvic floor, and nerves, eliciting a spectrum of clinical manifestations.

Several predictive models combining clinical and radiological parameters have been proposed for assessing surgical outcomes in BPH patients. Nevertheless, these findings exhibit discrepancies. Tian et al.[10] identified age, IPSS, IPP, transition zone thickness, PT, and bladder wall thickness as predictors for TURP efficacy, yielding AUCs of 0.860 and 0.806 in the primary and validation cohorts, respectively. Chen et al.[12] considered age, IPSS-voiding/storage ratio, PVR, and Q_{max} as predictors for holmium laser enucleation of the prostate efficacy, but model evaluation and validation were lacking. Other studies focused on individual predictors such as TZV and TZI.[13,14] Divergent findings can be attributed to disparate efficacy evaluation criteria. This study adopted stringent criteria supported by empirical data for outcome assessment.[10] Furthermore, the variability and limited information content of TRUS measurements contribute to inconsistent results. Transrectal ultrasound relies on 2-dimensional imaging, with measurement accuracy contingent upon the operator's expertise. Prostatic compression due to the probe and variations in prostate morphology can further impede accuracy. In addition, TRUS inadequately evaluates overall prostate enlargement, thereby constraining information availability during model development.

Radiomic analysis has been widely used in tumor diagnosis, efficacy prediction, and prognosis assessment.[15–18] However, to date, no studies have applied this methodology to predict surgical efficacy in BPH. Although MRI is not imperative for BPH evaluation, it offers multidimensional, high-resolution images with high soft tissue contrast (T2WI).[6] Moreover, it aids in reducing measurement errors across various directions and planes, thereby enhancing measurement repeatability. In addition, through radiomic analysis, MR images can be transformed into high-throughput radiomic feature data, enabling a comprehensive quantitative characterization

of morphological, grayscale, and textural features of the study subjects. This facilitates a thorough quantitative description.

Maximum flow rate and TZI were identified as independent predictive factors of the short-term efficacy of B-TUEP in this study. Urine flow rate assessment is a commonly used noninvasive urodynamic test for evaluating urination function.[1] It is generally accepted that Q_{max} can distinguish patients with urinary outflow obstruction from the normal population and can be used to monitor the efficacy of medication or surgery.[1,19] However, the accuracy of urine flow rate assessment as a diagnostic test is greatly affected by the threshold and cannot differentiate the underlying mechanisms.[20,21] Maximum flow rate was also considered a significant predictor for efficacy prediction in other studies.[10,12] McNeal et al.[22,23] conducted thorough research on the anatomy and histology of the prostate, proposing that most BPH lesions originated from the TZ and a few from the periurethral glands. Several studies have found that the TZV in BPH patients is much larger than in normal individuals.[24] Kaplan et al.[13] introduced the concept of the "transitional zone index," discovering that TZI had a higher association with symptoms and severity compared with PV and TZV, and the higher the TZI, the more severe the LUTS/BPH. Ohtani et al.[14] proposed that TZV and TZI could be new and effective criteria for making surgical treatment choices, consistent with the findings of this study. Thus, the results of the clinical prediction model indicate that, for patients with higher Q_{max} and lower TZI, caution should be exercised regarding other potential causes of LUTSs, such as abnormalities in pelvic floor structure and bladder overactivity. As B-TUEP can only address BPO and cannot resolve nonobstructive causes, the efficacy of surgical treatment may be poorer in this patient population. Therefore, further investigations, comprehensive evaluations of the benefits of B-TUEP surgery, and the development of more comprehensive treatment plans are warranted.

According to the results of the 10-fold CV analysis (Fig. 5), the AUC of CV modeling using radiomics features from both the prostate and TZ regions generally surpassed that of using radiomics features solely from the prostate or TZ. This observation suggests that the effectiveness of B-TUEP may be associated not only with the anatomical and histological characteristics of the TZ surrounding the urethra but also with those of the entire prostate. Histological types of BPH can be categorized into 5 types based on varying ratios of stromal and epithelial cells.[25] Some studies have indicated that these histological differences stem from variations in pathogenesis, potentially leading to delays or failures in medication efficacy.[26] Therefore, we hypothesize that the histological variances in BPH may also correlate with the efficacy of B-TUEP.

A comparison of the performances of the three predictive models in the training set (Fig. 6) revealed that the combination model outperformed the clinical model, with the radiomics prediction model exhibiting comparatively lower performance. It is important to note that LUTSs are rooted in a complex etiology. In patients with BPH, long-term BPO induces adaptive changes in bladder function alongside increased urethral resistance, leading to voiding dysfunction. The radiomics analysis solely investigated the relationship between prostate anatomical and histological characteristics and B-TUEP efficacy, without considering clinical symptoms and function. In contrast, the clinical model incorporates parameters such as Q_{max}, which comprehensively reflects the combined influences of bladder detrusor function and urethral resistance. Consequently, the clinical model encompasses an additional dimension of clinical information compared with the radiomics model, resulting in superior overall performance. However, the high-throughput information regarding the prostate encapsulated in the radiomics model is challenging to replicate using the clinical model alone. Thus, integrating the R score, which encapsulates prostate anatomical and histological information, with the predictive factors of the clinical model to formulate a combined predictive model can enhance predictive performance and facilitate urologists in making more precise clinical decisions.

The challenges and prospects warrant discussion. In clinical practice, it would be more convenient to use a clinical model that performs similarly well in predicting surgical efficacy. Moreover, we anticipate the development of an automated tool capable of reproducing the radiomic analysis process outlined in this study, thereby yielding predicted results directly upon input of clinical and MRI data.[27] In addition, although MRI is not essential for BPH assessment, we have investigated its potential application value in predicting B-TUEP efficacy and have obtained promising results. We aim to emphasize the relevance of MRI applications to urologists and anticipate further exploration by the urological community to comprehensively and objectively evaluate MRI's application value.[28,29] Given that only patients who underwent B-TUEP were included in this study, it would be necessary and valuable to gather more data from other surgical treatments for BPH (eg, TURP, holmium laser enucleation of the prostate, simple prostatectomy) to further validate and extend the prediction model.[28,29]

This study had several advantages. We integrated clinical factors and quantitative MRI radiomic features of the prostate to comprehensively explore the factors predicting B-TUEP efficacy and establish prediction models. Moreover, unlike most current radiomics studies that are single center, this study incorporated imaging data from 2 institutions, enhancing the robustness of the prediction model. In addition, to mitigate overfitting and selection bias, various combinations of methods for radiomics modeling were attempted, evaluated through a 10-fold CV, and the most robust combination was ultimately selected.

Limitations

This study also encountered limitations. Apart from the selection bias inherent in retrospective studies, the small size of the training set and the reliance on a single validation set may introduce statistical bias. Therefore, future studies should encompass larger training sets, multiple validation sets, and prospective validation sets. Furthermore, manual image segmentation was used in this study, potentially introducing subjective selection bias and inconsistent findings. Nevertheless, manual segmentation remains the standard method because of the lack of reliable fully or semiautomatic methods.[7] Moreover, the abstract nature of radiomics features extracted via computer algorithms poses challenges in elucidating their clinical significance, thereby limiting the generalizability of radiomics research findings.[30,31]

5. Conclusions

In summary, MRI and radiomic analyses offer valuable insights for predicting the short-term efficacy of B-TUEP in patients with BPH. A combined model integrating clinical and radiomics features can assist urologists in making more precise clinical decisions.

Acknowledgments

None.

Statement of ethics

This study was conducted following approval from the Institutional Committee of the Sixth Affiliation Hospital of Sun Yat-sen

University and the Institutional Committee of the Second Affiliated Hospital of Kunming Medical University. Given its retrospective nature, the necessity for obtaining informed consent from patients was waived. All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Conflict of interest statement

The authors declare that they have no conflicts of interest.

Funding source

This study was supported by a grant from the Basic and Applied Basic Research Foundation of Guangdong Province, China (grant no. 2019A1515010386).

Author contributions

TZ: Research design, data analysis and original draft writing;
ZM, JH: Data collection and analysis;
JW: performance of research;
YT, LY, WZ, BY, BM, JM: Providing resource and data collection;
HQ, DW: Project administration;
CY, JZ: Supervision and review;
JQ: Research design and editing.

Data availability

We would like to share the data collected for this study to investigators for individual participant data meta-analysis, including individual participant data that underlie the results reported in this article, after deidentification (text, tables, figures, and appendices). Data availability begins 9 months following article publication and ends 36 months following article publication. Proposals should be directed to qiujg@mail.sysu.edu.cn. To gain access, data requestors will need to sign a data access agreement.

References

[1] Gravas S, Bach T, Drake M, et al. Management of non-neurogenic male lower urinary tract symptoms lastname (LUTS), incl. benign prostatic obstruction (BPO): EAU Guidelines. Proceedings of the EAU Annual Congress. Amsterdam, The Netherlands, 2022.
[2] Lerner LB, McVary KT, Barry MJ, et al. Management of lower urinary tract symptoms attributed to benign prostatic hyperplasia: AUA GUIDELINE PART I—Initial work-up and medical management. *J Urol* 2021;206(4):806–817.
[3] Foo KT. What is a disease? What is the disease clinical benign prostatic hyperplasia (BPH)? *World J Urol* 2019;37(7):1293–1296.
[4] Shim M, Bang WJ, Oh CY, Lee YS, Cho JS. Correlation between prostatic urethral angulation and symptomatic improvement after surgery in patients with lower urinary tract symptoms according to prostate size. *World J Urol* 2020;38(8):1997–2003.
[5] Lu W-W, Zhang D, Ni X-J. A review of the role of ultrasound radiomics and its application and limitations in the investigation of thyroid disease. *Med Sci Monit* 2022;28:e937738.
[6] Turkbey B, Fotin SV, Huang RJ, et al. Fully automated prostate segmentation on MRI: Comparison with manual segmentation methods and specimen volumes. *AJR Am J Roentgenol* 2013;201(5):W720–W729.
[7] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* 2016;278(2):563–577.
[8] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14(12):749–762.
[9] Suarez-Ibarrola R, Hein S, Reis G, Gratzke C, Miernik A. Current and future applications of machine and deep learning in urology: A review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. *World J Urol* 2020;38(10):2329–2347.
[10] Tian Y, Zhang H, Cao Y, Yang L, Luo G. The P.R.OS.T.A.T.E nomogram for the preoperative prediction of clinical efficacy of transurethral resection of the prostate in benign prostatic hyperplasia patients. *Clin Interv Aging* 2022;17:845–855.
[11] Homma Y, Kawabe K, Tsukamoto T, et al. Estimate criteria for efficacy of treatment in benign prostatic hyperplasia. *Int J Urol* 1996;3(4):267–273.
[12] Chen X, Man Q, Wei X, et al. Predictive value of preoperative comprehensive evaluation on the efficacy of HoLEP. *Transl Androl Urol* 2020;9(4):1603–1610.
[13] Kaplan SA, Te AE, Pressler LB, Olsson CA. Transition zone index as a method of assessing benign prostatic hyperplasia: Correlation with symptoms, urine flow and detrusor pressure. *J Urol* 1995;154(5):1764–1769.
[14] Ohtani T, Hayashi Y, Kishino TE, et al. A new parameter in decision making for transurethral electroresection of benign prostate hyperplasia. *Eur Urol* 1999;35(3):185–191.
[15] Liu YF, Shu X, Qiao XF, et al. Radiomics-based machine learning models for predicting P504s/P63 immunohistochemical expression: A noninvasive diagnostic tool for prostate cancer. *Front Oncol* 2022;12:911426.
[16] Mühlbauer J, Egen L, Kowalewski KF, et al. Radiomics in renal cell carcinoma—A systematic review and meta-analysis. *Cancers (Basel)* 2021;13(6):1348.
[17] Kozikowski M, Suarez-Ibarrola R, Osiecki R, et al. Role of radiomics in the prediction of muscle-invasive bladder cancer: A systematic review and meta-analysis. *Eur Urol Focus* 2022;8(3):728–738.
[18] Lim EJ, Castellani D, So WZ, et al. Radiomics in urolithiasis: Systematic review of current applications, limitations, and future directions. *J Clin Med* 2022;11(17):5151.
[19] Siroky MB, Olsson CA, Krane RJ. The flow rate nomogram: II. Clinical correlation. *J Urol* 1980;123(2):208–210.
[20] Reynard JM, Yang Q, Donovan JL, et al. The ICS-'BPH' study: Uroflowmetry, lower urinary tract symptoms and bladder outlet obstruction. *Br J Urol* 1998;82(5):619–623.
[21] Siroky MB, Olsson CA, Krane RJ. The flow rate nomogram: I. Development. *J Urol* 1979;122(5):665–668.
[22] McNeal JE. Normal histology of the prostate. *Am J Surg Pathol* 1988;12(8):619–633.
[23] McNeal JE, Redwine EA, Freiha FS, Stamey TA. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *Am J Surg Pathol* 1988;12(12):897–906.
[24] Greene DR, Egawa S, Hellerstein DK, Scardino PT. Sonographic measurements of transition zone of prostate in men with and without benign prostatic hyperplasia. *Urology* 1990;36(4):293–299.
[25] Roehrborn CG. Pathology of benign prostatic hyperplasia. *Int J Impot Res* 2008;20(Suppl 3):S11–S18.
[26] Devlin CM, Simms MS, Maitland NJ. Benign prostatic hyperplasia — What do we know? *BJU Int* 2021;127(4):389–399.
[27] Hamm CA, Baumgärtner GL, Biessmann F, et al. Interactive explainable deep learning model informs prostate cancer diagnosis at MRI. *Radiology* 2023;307(4):e222276.
[28] Pandolfo SD, Del Giudice F, Chung BI, et al. Robotic assisted simple prostatectomy versus other treatment modalities for large benign prostatic hyperplasia: A systematic review and meta-analysis of over 6500 cases. *Prostate Cancer Prostatic Dis* 2023;26(3):495–510.
[29] Autorino R, Amparore D, Loizzo D, Pandolfo SD, Checcucci E, Porpiglia F. Robot-assisted simple prostatectomy is better than endoscopic enucleation of the prostate. *Eur Urol Focus* 2022;8(2):368–370.
[30] Ming Y, Qu H, Bertini E. RuleMatrix: Visualizing and understanding classifiers with rules. *IEEE Trans Vis Comput Graph* 2019;25(1):342–352.
[31] Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy, October 22–29, 2017. pp. 618–626.

How to cite this article: Zhang T, Mo Z, Huang J, Wang J, Tao Y, Ye L, Zhong W, Yao B, Qu H, Ma B, Wang D, Mo J, Ye C, Zhu J, Qiu J. Clinical-radiomics combination model for predicting the short-term efficacy of bipolar transurethral enucleation of the prostate in patients with benign prostatic hyperplasia. *Curr Urol* 2025;19(1):30–38. doi: 10.1097/CU9.0000000000000256