Check for updates

Predictive value of the stone-free rate after percutaneous nephrolithotomy based on multiple machine learning models

Zhao Rong Liu[1,2], Zhan Jiang Yu[3], Jie Zhou[3] and
Jian Biao Huang[1]*

[1]Jiangxi Cancer Hospital & Institute, Jiangxi Clinical Research Center for Cancer, The Second
Affiliated Hospital of Nanchang Medical College, Nanchang, China, [2]Department of Urology, Yudu
County People's Hospital, Yudu, China, [3]The Second Affiliated Hospital, Jiangxi Medical College,
Nanchang University, Nanchang, China

Purpose: This study aimed to develop three types of machine learning (ML) models based on gradient boosting decision tree (GBDT), random forest (RF), and extreme gradient boosting (XGBoost) to explore their predictive value for the stone-free rate after percutaneous nephrolithotomy (PCNL).

Patients and methods: A retrospective analysis was conducted on 160 patients who underwent PCNL. The patients were randomly divided into a training set and a test set in a 7:3 ratio. Clinical data were collected, and univariate analysis was performed to identify important data significantly associated with the stone-free rate after PCNL. Three ML models (GBDT, RF, and XGBoost) were developed using the training set. The predictive performance of these models was evaluated using the area under the curve (AUC) of the receiver operating characteristic (ROC) on the test set, confusion matrix, specificity, sensitivity, accuracy, and F1 score. For the top-performing prediction model, the study further employed the SHapley Additive exPlanations (SHAP) method to enhance model interpretability by elucidating the contribution of individual features to the prediction outcomes and ranking the relative importance of the predictive data. Finally, the clinical utility of the model was assessed through decision curve analysis (DCA), which quantified the net clinical benefit of applying the model across various risk thresholds.

Results: Postoperative statistics indicated a stone-free rate of 70.6% ($n = 113$) among the patients. The data significantly associated with the absence of residual stones included the number of stones, stone diameter, stone CT value, history of previous stone surgery, stone location, and stone shape ($p < 0.05$). All three models demonstrated strong predictive effects in the validation set, with the GBDT model showing superior performance [AUC: 0.836 (95% CI: 0.785–0.873); accuracy: 0.854; sensitivity: 0.853; specificity: 0.857] compared to the XGBoost [AUC: 0.830 (95% CI: 0.792–0.868); accuracy: 0.771; sensitivity: 0.824; specificity: 0.643] and RF models [AUC: 0.803 (95% CI: 0.763–0.837); accuracy: 0.792; sensitivity: 0.824; specificity: 0.714]. The F1 scores for the GBDT, RF, and XGBoost models were 0.892, 0.836, and 0.849, respectively. The DCA decision curve analysis confirmed that the GBDT model offers a favorable net clinical benefit. In addition, the SHAP analysis identified the number of stones and the stone CT value as the most critical features influencing the model's predictions, contributing significantly to its overall predictive performance.

Conclusion: The prediction models developed based on three machine learning algorithms can accurately predict the stone-free rate after PCNL in patients

with urinary tract stones. Among these, the GBDT model can effectively identify patients who are most likely to achieve successful outcomes from PCNL based on demographic and stone characteristics, thereby assisting in clinical treatment decision-making.

Introduction

As one of the most common conditions in urology, the global incidence of urinary tract stones ranges from 5 to 10%, with a higher proportion observed in North America and Europe (1). The medical costs associated with urinary tract stones are substantial due to emergency medical treatment, surgical interventions, and other related healthcare expenses. Some studies estimate that the annual direct medical costs of urinary tract stones in China may exceed 10 billion yuan (approximately US$1.5 billion) (2, 3). Percutaneous nephrolithotomy (PCNL) is characterized by minimal trauma, rapid recovery, a high stone clearance rate, and fewer complications compared to open surgery, making it widely utilized in clinical practice. It has gradually become the first-line surgical method for complex stones and those with a diameter greater than 2 cm (4, 5). However, factors such as the location and size of the stones may impact the surgical outcome, potentially leading to PCNL failure and resulting in incomplete stone removal, bleeding, or infection. Even with optimal surgical techniques, these situations can occasionally occur (6). Residual stones not only adversely affect patients' quality of life but also pose a risk of requiring secondary surgery, significantly impacting the overall recovery following urinary stone surgery (7). Consequently, accurately identifying the potential for residual stones after PCNL and understanding the postoperative stone-free rate are of utmost importance.

Machine learning, a branch of artificial intelligence, employs algorithms and statistical models to enable computer systems to automatically analyze training data, recognize patterns, and make new decisions without explicit programming. In machine learning, a model is developed by learning from training data and can predict or classify new data (8, 9). Machine learning (ML) models can assist physicians in predicting diagnoses, grading, and prognoses of urological diseases by analyzing patients' clinical data, imaging results, and laboratory indicators (10, 11). A machine learning model that integrates radiomics and clinical data can effectively identify urinary tract infection stones *in vivo*, potentially optimizing the management of urolithiasis and improving patient outcomes (12). Moreover, a model based on machine learning algorithms has proven to be quite accurate in predicting postoperative bleeding after PCNL in the contralateral decubitus position, aiding urologists in making informed treatment decisions (13).

This study developed three new machine learning prediction models for stone-free rates after PCNL, aiming to identify the most suitable patients for the procedure preoperatively, thereby enhancing perioperative management and preventing the occurrence of postoperative stone residues.

Patients and methods

Patients

A total of 160 eligible patients who underwent PCNL at the Second Affiliated Hospital of Nanchang University from April 2021 to November 2022 were included in this retrospective analysis. The inclusion criteria were as follows: (1) patients diagnosed with kidney stones and/or upper ureteral stones by a CT scan within 1 month prior to the procedure; (2) patients who received complete PCNL surgical treatment; and (3) availability of complete clinical data. The exclusion criteria included the following: (1) presence of malignant tumors in the urinary system on the ipsilateral side; (2) severe liver and renal dysfunction or systemic coagulation disorders; and (3) congenital malformations or severe atrophy of the ipsilateral kidney. All surgeries were performed using a standard single-channel technique. The stone-free rate was defined as the absence of residual stones or the presence of remaining stones measuring less than 4 mm, as assessed by a plain X-ray examination of the urinary tract approximately 1 week after surgery. All patients were randomly divided into a training set ($n = 112$) and a test set ($n = 48$) in a ratio of 7:3, using the random seed method in the R language.

Data collection and processing

By querying the hospital's medical record system, we obtained the clinical information of all patients, which included the following: (1) general conditions: age, sex, BMI, hypertension, and diabetes; (2) preoperative conditions: preoperative stent placement or nephrostomy, the extent of preoperative infection and hydronephrosis (defined by preoperative B-ultrasound as mild if the separation of the renal collecting system was 2–3 cm, moderate if 3–4 cm, and severe if greater than 4 cm), surgeon experience (surgeons who have performed more than 100 PCNL procedures were considered experienced, whereas those with fewer than 100 procedures were regarded as less experienced), and history of stone surgery (including PCNL and non-PCNL procedures such as ureteroscopic lithotripsy and open stone surgery); and (3) stone characteristics: affected sides, number of stones, stone location, maximum stone diameter (the diameter of the largest single stone), stone CT value [average CT value representing stone density, measured in Hounsfield units (HU)], and stone shape. Important data were identified through univariate analysis.

Model building and evaluation

Gradient boosting decision tree (GBDT) is a process that combines multiple weak classifiers offline to create a strong classifier. This

algorithm significantly enhances the flexibility and convenience of making predictions based on decision trees. Random forest (RF) is an ensemble learning method that relies on numerous decision trees, with the output category determined by the majority vote from the individual trees. Extreme gradient boosting (XGBoost) efficiently implements GBDT while incorporating numerous improvements in both algorithms and engineering practices. The screened clinical data mentioned above were used to construct the machine learning models. Three types of machine learning algorithms were trained using a binary classification approach. The training set (70%) was utilized to train the models, while the validation set (30%) was employed to evaluate the models' predictive performance. By calculating the prediction accuracy, specificity, sensitivity, F1 score, confusion matrix, and area under the curve (AUC), we compared the predictive values of the GBDT, RF, and XGBoost models. The SHapley Additive exPlanations (SHAP) method was used to evaluate the contribution of individual features within the optimal predictive model, while the clinical net benefit was quantified through decision curve analysis (DCA).

Surgical methods

All patients were placed in the standard lithotomy position before surgery to examine the bladder and ureter on the surgical side. Subsequently, a 5F ureteral catheter was inserted into the ureter on the surgical side, advanced 25–28 cm, secured in place, and connected to an infusion device to induce artificial hydronephrosis; the duration of artificial hydronephrosis induction varied depending on the severity of the patient's hydronephrosis. Following repositioning to the prone position and under ultrasound guidance, the renal calyx was punctured, with the approach determined by preoperative imaging. The supracostal upper pole approach was preferred for pelvic or large stones, while subcostal or lower pole punctures were employed for isolated lower or mid-pole calyx stones. Once the needle core was withdrawn and urine was observed flowing out, a renal puncture guidewire was inserted through the puncture needle sheath. Guided by the wire, an F10 fascia dilator was used to gradually expand the tract to F24 in 2F increments, followed by the insertion of an F24 tearable sheath, which possesses a tearable feature, allowing it to maintain the sinus tract while being flexible enough for appropriate deformation. Subsequently, an L-type nephroscope (24Fr) was introduced to confirm the location of the stones. The stones were then fragmented using a holmium laser, with smaller fragments irrigated out and larger fragments removed using grasping forceps. All procedures adhered to the standard PCNL protocol and were successfully completed through a single puncture channel. An F6 double-J stent was placed postoperatively and maintained for approximately 4 weeks.

Statistical analysis

Data analysis and processing were performed using SPSS version 27.0, R language software (version 4.3.1),[1] and Python language software.[2] All measurement data are expressed as mean ± standard deviation, while categorical data are presented as numbers (%). A single-factor analysis was conducted to compare differences in clinical factors. The *t*-test or Mann–Whitney *U*-test was employed for measurement data, and the chi-squared test or Fisher's exact test was used for categorical data. A *p*-value of <0.05 was considered statistically significant. Lasso regression was performed using the glmnet package in R or the scikit-learn package in Python. The matplotlib package in Python was utilized to generate the receiver operating characteristic (ROC) curve.

Results

Patients

According to postoperative review statistics, among the 160 patients, 47 had residual stones, while 113 had completely cleared stones, resulting in a stone-free rate of 70.6%. The number of stones, stone diameter, stone CT value, history of previous stone surgery, stone location, and stone shape were found to be significantly associated with postoperative stone clearance ($p < 0.05$) (see Table 1).

Machine learning model building

We randomly assigned 70% of the patients to the training set and the remaining 30% to the test set. The stone-free rate was approximately the same in both the training set (70.5%) and the test set (70.8%). The machine learning algorithms were employed to construct predictive models based on the clinical data described above. Utilizing the Python scikit-learn framework, multiple decision trees were constructed, and their prediction results were combined. The model parameters were updated, and the objective function was optimized to enhance the accuracy of the model. This included utilizing second-order derivative information to optimize the loss function, ultimately achieving binary classification predictions of the stone-free rate post-PCNL.

Model evaluation and comparison

The AUCs for the test sets of the GBDT, RF, and XGBoost models were 0.836 (95% CI: 0.785–0.873), 0.803 (95% CI: 0.763–0.837), and 0.830 (95% CI: 0.792–0.868), respectively. The accuracy, specificity, sensitivity, and F1 score of the three models are presented in Table 2. The confusion matrix and ROC curve are depicted in Figure 1. The sensitivity values for the three machine learning models ranged from 0.824 to 0.853, specificity values ranged from 0.643 to 0.857, accuracy values ranged from 0.771 to 0.854, and F1 scores ranged from 0.836 to 0.892. This indicates that all three models demonstrate strong predictive capabilities. After a comprehensive comparison of these metrics, particularly the prediction accuracy, the GBDT model was selected as the final prediction model. To further elucidate the
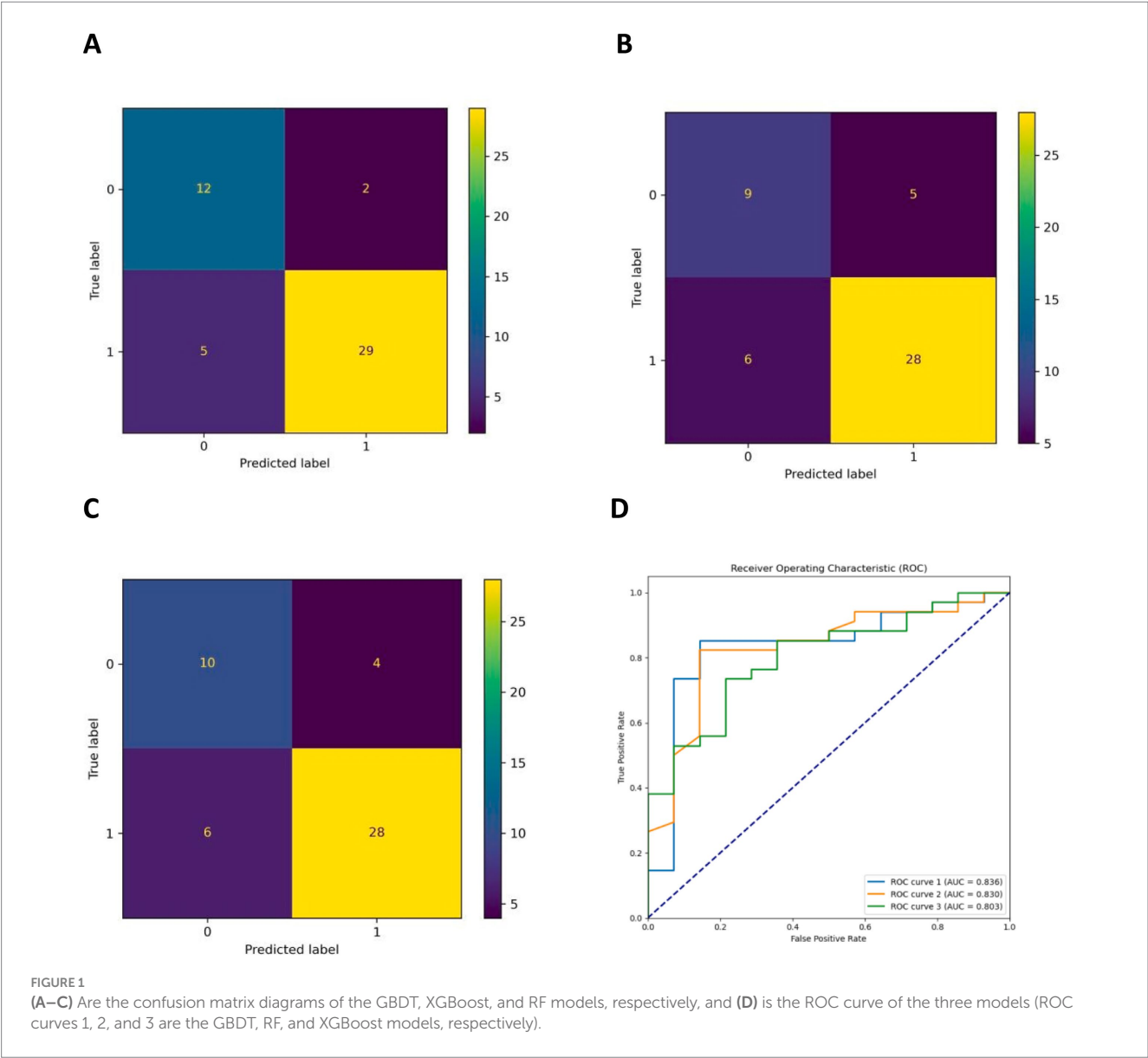
TABLE 1 Basic characteristics and difference analysis.

Data	Total (n = 160)	Non-SF (n = 47)	SF (n = 113)	Statistic	p
Age, year, mean ± SD	54.35 ± 12.55	54.94 ± 11.31	54.11 ± 13.07	$t = 0.38$	0.704
BMI, mean ± SD	23.49 ± 3.47	23.52 ± 3.84	23.48 ± 3.33	$t = 0.06$	0.949
Stones number, M (Q_1, Q_3)	3 (2, 4)	5 (4, 6)	3 (1, 4)	$Z = -6.77$	<0.001
Stone diameter, cm, M (Q_1, Q_3)	2.0 (1.4, 2.6)	2.6 (2.0, 3.5)	1.7 (1.3, 2.3)	$Z = -5.16$	<0.001
Stone CT value, Hu, M (Q_1, Q_3)	1,140 (983, 1,280)	1,240 (1,070, 1,383)	1,080 (950, 1,210)	$Z = -3.78$	<0.001
Sex, n (%)				$\chi^2 = 0.02$	0.893
Male	100 (62.50)	29 (61.70)	71 (62.83)		
Female	60 (37.50)	18 (38.30)	42 (37.17)		
Pre-stenting/nephrostomy, n (%)				$\chi^2 = 1.06$	0.303
No	150 (93.75)	46 (97.87)	104 (92.04)		
Yes	10 (6.25)	1 (2.13)	9 (7.96)		
History of stone surgery, n (%)				$\chi^2 = 6.04$	0.049
No	122 (76.25)	30 (63.83)	92 (81.42)		
PCNL	16 (10.00)	8 (17.02)	8 (7.08)		
Non-PCNL surgery	22 (13.75)	9 (19.15)	13 (11.50)		
Pre-infection, n (%)				$\chi^2 = 2.83$	0.419
−	17 (10.62)	3 (6.38)	14 (12.39)		
+	84 (52.50)	24 (51.06)	60 (53.10)		
++	29 (18.12)	8 (17.02)	21 (18.58)		
+++	30 (18.75)	12 (25.53)	18 (15.93)		
Hypertension, n (%)				$\chi^2 = 1.43$	0.232
No	116 (72.50)	31 (65.96)	85 (75.22)		
Yes	44 (27.50)	16 (34.04)	28 (24.78)		
Diabetes, n (%)				$\chi^2 = 0.00$	1.000
No	148 (92.50)	43 (91.49)	105 (92.92)		
Yes	12 (7.50)	4 (8.51)	8 (7.08)		
Stone laterality, n (%)				$\chi^2 = 0.02$	0.877
Left	73 (45.62)	21 (44.68)	52 (46.02)		
Right	87 (54.38)	26 (55.32)	61 (53.98)		
Stone location, n (%)				$\chi^2 = 7.48$	0.024
Kidney	26 (16.25)	2 (4.26)	24 (21.24)		
Ureter	97 (60.62)	31 (65.96)	66 (58.41)		
Kidneys and ureters	37 (23.12)	14 (29.79)	23 (20.35)		
Hydronephrosis, n (%)				$\chi^2 = 3.31$	0.346
No	13 (8.12)	1 (2.13)	12 (10.62)		
Mild	81 (50.62)	26 (55.32)	55 (48.67)		
Moderate	35 (21.88)	11 (23.40)	24 (21.24)		
Severe	31 (19.38)	9 (19.15)	22 (19.47)		
Experience of operator, n (%)				$\chi^2 = 2.44$	0.118
PCNL <100	32 (20.00)	13 (27.66)	19 (16.81)		
PCNL >100	128 (80.00)	34 (72.34)	94 (83.19)		
Stone shape, n (%)				$\chi^2 = 21.87$	<0.001
Non-staghorn stones	96 (60.00)	15 (31.91)	81 (71.68)		
Staghorn stones	64 (40.00)	32 (68.09)	32 (28.32)		

BMI, body mass index; non-PCNL surgery: ureteroscopic lithotripsy, open stone surgery, and extracorporeal shock wave lithotripsy; pre-infection: + mild infection, ++ moderate infection, and +++ severe infection.

TABLE 2 AUC, accuracy, sensitivity, specificity, and F1 values of the different models in the test set.

Models	Accuracy	Sensitivity	Specificity	F1 score	AUC	95% CI
GBDT	0.854	0.853	0.857	0.892	0.836	0.785–0.873
XGBoost	0.771	0.824	0.643	0.836	0.803	0.736–0.837
RF	0.792	0.824	0.714	0.849	0.830	0.792–0.868

FIGURE 1
(A–C) Are the confusion matrix diagrams of the GBDT, XGBoost, and RF models, respectively, and (D) is the ROC curve of the three models (ROC curves 1, 2, and 3 are the GBDT, RF, and XGBoost models, respectively).

decision-making process of the optimal model, we performed DCA and employed the SHAP method to analyze the top-performing GBDT model, ranking the importance of the predictive data. The SHAP analysis indicated that the number of stones and the stone CT value were the two most influential features affecting the model's predictions (Figure 2), with their significant contributions thoroughly validated by the SHAP values (Figure 2). The DCA curve demonstrates that the GBDT model offers substantial clinical benefits to patients (Figure 3).

Discussion

With the continuous development and advancement of minimally invasive surgical techniques and endoscopic instruments, the surgical treatment of urinary tract stones has gradually evolved into a system primarily based on transurethral ureteroscopy and percutaneous nephrolithotomy (PCNL) (14, 15). PCNL, in particular, is widely favored due to its minimal trauma and rapid recovery. Compared to traditional open surgery, PCNL has significantly improved stone
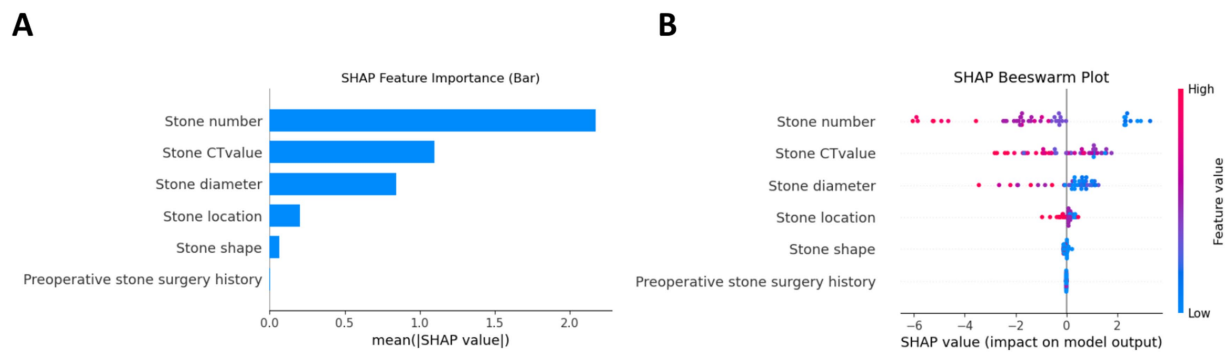
FIGURE 2
Interpretability analysis of the GBDT model using the SHAP method. **(A)** Feature importance ranking by mean absolute SHAP values. **(B)** SHAP summary plot demonstrating individual feature contributions, where each point represents a patient and color intensity corresponds to feature value magnitude and direction of prediction impact.

clearance rates, reduced postoperative complications, and shortened hospital stays. However, despite the maturation of PCNL technology, its effectiveness and safety across different patient groups still require further investigation (16, 17). Factors such as a patient's anatomical characteristics, the size and nature of the stone, and the presence of comorbidities can all impact surgical outcomes. The removal of larger or more complex stones inevitably increases the risk of residual stones (18, 19). Therefore, when selecting treatment options, it is essential to thoroughly evaluate the individual characteristics of each patient and to develop a personalized treatment strategy.

Extensive previous research has successfully developed nomogram-based prediction models to accurately evaluate treatment outcomes following PCNL (20). With the increasing application of artificial intelligence approaches, particularly machine learning methods, in medical research, these advanced techniques have demonstrated remarkable advantages in predicting surgical outcomes. This advancement highlights the significant potential of machine learning-based models in predicting postoperative outcomes of PCNL, with particular focus on residual stones—a crucial clinical endpoint that greatly influences patient management (21–23). The application of machine learning in the diagnosis and treatment of urinary tract stones is emerging as a significant research focus in urology (24, 25). By analyzing extensive clinical data, machine learning can identify risk factors unique to different patients, predict the occurrence, progression, and treatment outcomes of stones, and assist in optimizing treatment plans (26–28). A study by Shee et al. (29) found that a machine learning model based on 24-h urine data could predict stone recurrence with moderate accuracy. In addition, the machine learning model developed by Zhang et al. (30), which utilized various laboratory test indicators and stone characteristics, was able to accurately predict the likelihood of systemic inflammatory response syndrome following PCNL, thereby aiding surgeons in their clinical decision-making. These automated data analysis models founded on machine learning not only enhance work efficiency but also reduce human error, providing clinicians with robust decision-making support (31, 32).

Machine learning models demonstrate theoretical potential for predicting stone-free rates in PCNL procedures, as they primarily analyze stone characteristics and clinical data (13, 33). While core predictive features (stone morphology, density, and anatomical
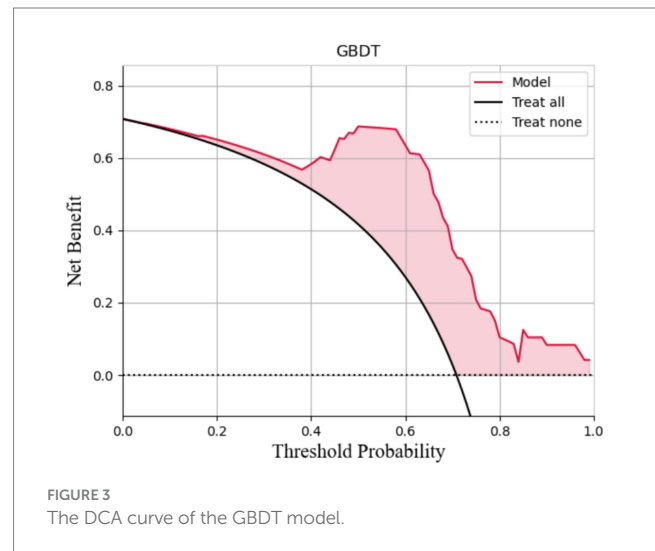

FIGURE 3
The DCA curve of the GBDT model.

parameters) are fundamentally important across different surgical approaches, it should be emphasized that our current model was exclusively developed and validated using prone-position PCNL data. Therefore, while these algorithms may conceptually maintain relevance for supine PCNL and miniaturized techniques, their actual clinical utility for these specific variants remains to be prospectively validated. This important limitation must be considered when interpreting the model's generalizability across different PCNL approaches. Future studies specifically evaluating the algorithm's performance in supine and mini-PCNL settings are needed to confirm its broader applicability.

This study utilized meaningful clinical data to develop prediction models based on GBDT, XGBoost, and RF algorithms. By calculating the accuracy, specificity, sensitivity, F1 score, ROC curve, and confusion matrix of the three models, we investigated the applicability of machine learning methods in preoperative predictions of stone-free rates after PCNL. Among these three models, the GBDT model demonstrated the highest AUC value and accuracy. The SHAP method was utilized to evaluate the contribution of features within the optimal predictive model, and the favorable clinical net benefit of the GBDT model was established through DCA. Given the uncertainties

associated with PCNL, a clinical model that integrates multiple conventional parameters may prove to be more effective than relying on any single parameter (34, 35). To achieve this goal, we employed advanced machine learning techniques, which have been applied in preventing and managing stone treatment outcomes, to build predictive models based on common clinical parameters that are both simple and technically undemanding, making them suitable for use at the grassroots level. This approach enhances the potential for hospital implementation, significantly broadening the applicability of this study.

However, this study has several limitations. As a retrospective analysis, despite the implementation of stringent inclusion and exclusion criteria, there remains the potential for selection bias. The sample size was relatively small, which may limit the generalizability of the results, and the model was developed using data from a single center, potentially restricting its broader applicability. Importantly, the reliance on kidney-ureter-bladder (KUB) imaging for postoperative assessment introduces a significant limitation, as radiolucent stones (e.g., uric acid calculi) cannot be reliably detected. This not only affects the accuracy of residual stone evaluation but also confines the machine learning model's training and validation to radiopaque stones, thereby limiting its generalizability across all stone types. Furthermore, the model was trained and tested exclusively on data from patients undergoing prone-position PCNL, and its performance may not extend to other surgical positions (e.g., supine or modified positions), which could influence stone localization, renal anatomy, and surgical outcomes. Future research should involve multi-center prospective studies incorporating a broader range of data, including diverse stone compositions, calyceal locations, and patient positioning, as well as more rigorous outcome assessments (e.g., advanced imaging modalities or standardized protocols). In addition, exploring advanced computational techniques, such as deep learning and artificial neural networks, may further enhance predictive accuracy and clinical utility.

Conclusion

The prediction models based on GBDT, XGBoost, and RF showed strong effectiveness in predicting the stone-free rate after PCNL, with the GBDT model performing the best. It holds promise for aiding clinical decision-making and enabling personalized prevention and treatment strategies for patients with different residual stone risks.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the Second Affiliated Hospital of Nanchang University. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/

next of kin because the use of patient case data will not have any impact on patient treatment and physical health.

Author contributions

ZL: Conceptualization, Data curation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing. ZY: Data curation, Writing – original draft, Investigation, Validation, Writing – review & editing. JZ: Data curation, Formal analysis, Writing – review & editing, Conceptualization, Writing – original draft. JH: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

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.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2025.1559613/full#supplementary-material

References

1. Thongprayoon C, Krambeck AE, Rule AD. Determining the true burden of kidney stone disease. *Nat Rev Nephrol*. (2020) 16:736–46. doi: 10.1038/s41581-020-0320-7

2. Peerapen P, Thongboonkerd V. Kidney stone prevention. *Adv Nutr*. (2023) 14:555–69. doi: 10.1016/j.advnut.2023.03.002

3. Tan S, Yuan D, Su H, Chen W, Zhu S, Yan B, et al. Prevalence of urolithiasis in China: a systematic review and meta-analysis. *BJU Int*. (2024) 133:34–43. doi: 10.1111/bju.16179

4. Akram M, Jahrreiss V, Skolarikos A, Geraghty R, Tzelves L, Emilliani E, et al. Urological guidelines for kidney stones: overview and comprehensive update. *J Clin Med*. (2024) 13:1114. doi: 10.3390/jcm13041114

5. Desai J, Shah HN. Mini percutaneous kidney stone removal: applicable technologies. *Urol Clin North Am*. (2022) 49:161–73. doi: 10.1016/j.ucl.2021.08.003

6. Gadzhiev N, Malkhasyan V, Akopyan G, Petrov S, Jefferson F, Okhunov Z. Percutaneous nephrolithotomy for staghorn calculi: troubleshooting and managing complications. *Asian J Urol*. (2020) 7:139–48. doi: 10.1016/j.ajur.2019.10.004

7. Doykov M, Kostov G, Doykova K. Factors affecting residual stone rate, operative duration, and complications in patients undergoing minimally invasive percutaneous nephrolithotomy. *Medicina*. (2022) 58:422. doi: 10.3390/medicina58030422

8. Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: potential resources for the infection clinician. *J Infect*. (2023) 87:287–94. doi: 10.1016/j.jinf.2023.07.006

9. Erickson BJ. Basic artificial intelligence techniques: machine learning and deep learning. *Radiol Clin North Am*. (2021) 59:933–40. doi: 10.1016/j.rcl.2021.06.004

10. Yang B, Veneziano D, Somani BK. Artificial intelligence in the diagnosis, treatment and prevention of urinary stones. *Curr Opin Urol*. (2020) 30:782–7. doi: 10.1097/MOU.0000000000000820

11. Hameed BMZ, Dhavileswarapu AVLS, Raza SZ, Karimi H, Khanuja HS, Shetty DK, et al. Artificial intelligence and its impact on urological diseases and management: a comprehensive review of the literature. *J Clin Med*. (2021) 10. doi: 10.3390/jcm10091864

12. Zheng J, Yu H, Batur J, Shi Z, Tuerxun A, Abulajiang A, et al. A multicenter study to develop a non-invasive radiomic model to identify urinary infection stone *in vivo* using machine-learning. *Kidney Int*. (2021) 100:870–80. doi: 10.1016/j.kint.2021.05.031

13. Meng R, Wang W, Zhai Z, Zuo C. Machine learning algorithm to predict postoperative bleeding complications after lateral decubitus percutaneous nephrolithotomy. *Medicine*. (2024) 103:e37050. doi: 10.1097/MD.0000000000037050

14. Zeng G, Traxer O, Zhong W, Osther P, Pearle MS, Preminger GM, et al. International alliance of urolithiasis guideline on retrograde intrarenal surgery. *BJU Int*. (2023) 131:153–64. doi: 10.1111/bju.15836

15. Kim CH, Chung DY, Rha KH, Lee JY, Lee SH. Effectiveness of percutaneous nephrolithotomy, retrograde intrarenal surgery, and extracorporeal shock wave lithotripsy for treatment of renal stones: a systematic review and meta-analysis. *Medicina*. (2020) 57:26. doi: 10.3390/medicina57010026

16. Grosso AA, Sessa F, Campi R, Viola L, Polverino P, Crisci A, et al. Intraoperative and postoperative surgical complications after ureteroscopy, retrograde intrarenal surgery, and percutaneous nephrolithotomy: a systematic review. *Minerva Urol Nephrol*. (2021) 73:309–32. doi: 10.23736/S2724-6051.21.04294-4

17. Setthawong V, Srisubat A, Potisat S, Lojanapiwat B, Pattanittum P. Extracorporeal shock wave lithotripsy (ESWL) versus percutaneous nephrolithotomy (PCNL) or retrograde intrarenal surgery (RIRS) for kidney stones. *Cochrane Database Syst Rev*. (2023) 8:CD007044. doi: 10.1002/14651858.CD007044.pub4

18. Soderberg L, Ergun O, Ding M, Parker R, Borofsky MS, Pais V, et al. Percutaneous nephrolithotomy versus retrograde intrarenal surgery for treatment of renal stones in adults. *Cochrane Database Syst Rev*. (2023) 11:CD013445. doi: 10.1002/14651858.CD013445.pub2

19. Danilovic A, Torricelli FCM, Marchini GS, Batagello C, Vicentini FC, Traxer O, et al. Residual stone fragments after percutaneous nephrolithotomy: shockwave lithotripsy vs retrograde intrarenal surgery. *J Endourol*. (2021) 35:609–14. doi: 10.1089/end.2020.0868

20. Mazzon G, Claps F, Pavan N, Choong S, Zeng G, Wu W, et al. Reliability of nephrolithometric nomograms in patients treated with minimally invasive percutaneous nephrolithotomy: a precision study. *Asian J Urol*. (2023) 10:70–80. doi: 10.1016/j.ajur.2022.02.010

21. Li J, Du Y, Huang G, Huang Y, Xi X, Ye Z. Constructing a machine learning model for systemic infection after kidney stone surgery based on CT values. *Sci Rep*. (2025) 15:4327. doi: 10.1038/s41598-025-88704-y

22. Ma L, Qiao Y, Wang R, Chen H, Liu G, Xiao H, et al. Machine learning models decoding the association between urinary stone diseases and metabolic urinary profiles. *Metabolites*. (2024) 14:674. doi: 10.3390/metabo14120674

23. Altunhan A, Soyturk S, Guldibi F, Tozsin A, Aydın A, Aydın A, et al. Artificial intelligence in urolithiasis: a systematic review of utilization and effectiveness. *World J Urol*. (2024) 42:579. doi: 10.1007/s00345-024-05268-8

24. Abraham A, Kavoussi NL, Sui W, Bejan C, Capra JA, Hsi R. Machine learning prediction of kidney stone composition using electronic health record-derived features. *J Endourol*. (2022) 36:243–50. doi: 10.1089/end.2021.0211

25. Huettenbrink C, Hitzl W, Distler F, Ell J, Ammon J, Pahernik S. Personalized prediction of patient radiation exposure for therapy of urolithiasis: an application and comparison of six machine learning algorithms. *J Pers Med*. (2023) 13:643. doi: 10.3390/jpm13040643

26. Yildirim K, Bozdag PG, Talo M, Yildirim O, Karabatak M, Acharya UR. Deep learning model for automated kidney stone detection using coronal CT images. *Comput Biol Med*. (2021) 135:104569. doi: 10.1016/j.compbiomed.2021.104569

27. Zhu B, Nie Y, Zheng S, Lin S, Li Z, Wu W. CT-based radiomics of machine-learning to screen high-risk individuals with kidney stones. *Urolithiasis*. (2024) 52:91. doi: 10.1007/s00240-024-01593-0

28. Nakamae Y, Deguchi R, Nemoto M, Kimura Y, Yamashita S, Kohjimoto Y, et al. AI prediction of extracorporeal shock wave lithotripsy outcomes for ureteral stones by machine learning-based analysis with a variety of stone and patient characteristics. *Urolithiasis*. (2023) 52:9. doi: 10.1007/s00240-023-01506-7

29. Shee K, Liu AW, Chan C, Yang H, Sui W, Desai M, et al. A novel machine-learning algorithm to predict stone recurrence with 24-hour urine data. *J Endourol*. (2024) 38:809–16. doi: 10.1089/end.2023.0457

30. Zhang T, Zhu L, Wang X, Zhang X, Wang Z, Xu S, et al. Machine learning models to predict systemic inflammatory response syndrome after percutaneous nephrolithotomy. *BMC Urol*. (2024) 24:140. doi: 10.1186/s12894-024-01529-1

31. Sánchez C, Larenas F, Arroyave JS, Connors C, Giménez B, Palese MA, et al. Artificial intelligence in urology: application of a machine learning model to predict the risk of urolithiasis in a general population. *J Endourol*. (2024) 38:712–8. doi: 10.1089/end.2023.0702

32. Kavoussi NL, Floyd C, Abraham A, Sui W, Bejan C, Capra JA, et al. Machine learning models to predict 24 hour urinary abnormalities for kidney stone disease. *Urology*. (2022) 169:52–7. doi: 10.1016/j.urology.2022.07.008

33. Nedbal C, Cerrato C, Jahrreiss V, Castellani D, Pietropaolo A, Galosi AB, et al. The role of 'artificial intelligence, machine learning, virtual reality, and radiomics' in PCNL: a review of publication trends over the last 30 years. *Ther Adv Urol*. (2023) 15:17562872231196676. doi: 10.1177/17562872231196676

34. Jannello LMI, Turetti M, Silvani C, Galbiati G, Garbagnati S, Pozzi E, et al. Urologists are optimistic surgeons: prevalence and predictors of discordance between intraoperative stone-free rate and cross-sectional imaging evaluation after vacuum-assisted mini-percutaneous nephrolithotomy. *World J Urol*. (2022) 40:2331–8. doi: 10.1007/s00345-022-04091-3

35. Hartung FO, Müller KJ, Herrmann J, Grüne B, Michel MS, Rassweiler-Seyfried MC. Comparison of endoscopic versus CT assessment of stone-free status after percutaneous nephrolithotomy (PCNL). *Urolithiasis*. (2023) 51:120. doi: 10.1007/s00240-023-01495-7