

Assessing the Predictive Accuracy of the S.T.O.N.E. Score for Stone-Free Rates in Semirigid Pneumatic Ureteral Lithotripsy: Implications for Validation

Faisal Ahmed¹, Khaled Al-Kohlany², Khalil Al-Naggar¹, Ibrahim Alnadhari^{3,4},
Abdulfattah Yahya Altam⁵, Mohamed Badheeb⁶

¹Department of Urology, School of Medicine, Ibb University, Ibb, Yemen; ²Department of Urology, College of Medicine, Sana'a University, Sana'a, Yemen; ³Al Wakra Hospital, Hamad Medical Corporation, Al Wakra, Qatar; ⁴Department of Surgery, School of Medicine, Qatar University, Doha, Qatar; ⁵Department of General Surgery, School of Medicine, 21 September University, Sana'a, Yemen; ⁶Department of Internal Medicine, Yale New-Haven Health/Bridgeport Hospital, Bridgeport, CT, USA

Correspondence: Faisal Ahmed, Department of Urology, School of Medicine, Ibb University, Ibb, Yemen, Email fmaa2006@yahoo.com

Background: The lack of reliable predictive tools for outcomes following ureteral lithotripsy (ULT) presents significant challenges in clinical decision-making. This study evaluates the efficacy of the S.T.O.N.E. score—an assessment incorporating Size, Topography, Obstruction, Number, and Hounsfield units (HU)—in predicting the likelihood of achieving a stone-free rate (SFR) in patients undergoing semirigid pneumatic ULT.

Methods: This retrospective analysis involved 266 patients with ureteral stones who underwent ULT at IBB University Hospitals from April 2021 to September 2023. The S.T.O.N.E. score was derived from preoperative CT scans, and a nomogram was created to predict SFR failure. Discrimination and calibration were assessed using the area under the receiver operating characteristic curve (AUC) and calibration curve, while decision curve analysis (DCA) evaluated clinical utility.

Results: The cohort's mean age was 47.7 ± 15 years, with a predominance of males (72.2%). The mean S.T.O.N.E. score was 7.8 ± 1.8 . The overall SFR of 85.3% and residual stones were detected in 39 patients (14.7%). Multivariate analysis identified higher HU (AOR: 1.01; 95% CI: 1.00–1.01; $P < 0.001$), proximal stone location (AOR: 15.13; 95% CI: 1.52–51.13; $P = 0.020$), moderate (AOR: 34.23; 95% CI: 8.28–141.45; $P < 0.001$) and severe hydronephrosis (AOR: 33.75; 95% CI: 4.55–250.36; $P = 0.0006$), and larger stone size (AOR: 1.51; 95% CI: 1.30–1.75; $P < 0.0001$) as significant predictors of SFR failure. The S.T.O.N.E. score effectively predicts SFR failure, with an optimal threshold of > 8 achieving 85.0% accuracy. The model demonstrated 72.0% sensitivity, 81.0% specificity, and strong calibration. DCA indicated clinical utility, differentiating between low- and high-risk patients based on their S.T.O.N.E. scores.

Conclusion: The S.T.O.N.E. score is a valuable tool for predicting post-ULT SFR, aiding preoperative decision-making and potentially improving surgical outcomes by identifying high-risk patients. Further validation in diverse populations is needed to confirm its clinical utility.

Plain Language Summary: Ureteral lithotripsy (ULT) is a procedure used to treat ureteral stones; however, predicting the success of this treatment remains challenging for urologists. This study evaluates the S.T.O.N.E. score, a tool designed to estimate a patient's likelihood of achieving stone-free status following ULT by considering factors such as stone size and location. Researchers analyzed data from 266 patients who underwent ULT at IBB University Hospitals between April 2021 and September 2023. The findings indicate that the S.T.O.N.E. score effectively predicts the presence of residual stones post-treatment. These results suggest that the score may assist urologists in making more informed decisions prior to surgery, particularly for patients at elevated risk. However, further research is necessary to validate its applicability across diverse patient populations.

Keywords: lithotripsy, nomograms, S.T.O.N.E. score, ureter, ureteroscopy, urolithiasis, stone free rate

Introduction

Urolithiasis represents a significant public health concern, affecting millions worldwide, with an estimated prevalence of 2–3% and a recurrence rate of approximately 50%.¹ Ureteric colic, characterized by severe, acute pain, is classified as a urological emergency requiring timely management.² Treatment modalities for ureteral stones include conservative management, endourological procedures, and open surgery.³ Recent advancements in small-caliber semirigid and flexible deflectable ureteroscopes (URs), as well as novel lithotripters, have enhanced the safety and efficacy of ureteroscopic lithotripsy.⁴ URS has thus become a standard technique for the diagnosis and treatment of conditions affecting the ureters and kidneys. Common indications for URS encompass urolithiasis, ureteric strictures, ureteropelvic junction obstruction, ablation of transitional cell carcinoma, and the retrieval of migrating stones.^{5,6}

Despite the increasing utilization of URS, there has been a corresponding rise in complications, highlighting the necessity for novel preventive strategies.⁷ Advances in URS and intracorporeal lithotripsy have substantially improved the success rate of ureteral lithotripsy (ULT), escalating from 50% to 97%, while the overall complication rate has decreased to 12%–15%, with serious complications occurring in merely 0.8%–1.5% of cases.^{8,9} Several studies have identified predictive factors that influence complications and the stone-free rate (SFR) following ULT, including stone characteristics (size, density, location, shape, and impaction) and patient-related factors (age, comorbidities, congenital anomalies, degree of obstruction, and presence of infection).^{4,7,10}

While laser lithotripters are commonly regarded as the standard of care, pneumatic lithotripters remain a viable alternative in resource-limited settings where access to advanced technologies is restricted.^{7,11} Various scoring systems, such as the T.O.H.O. score—evaluating tallness, occupied lesion, and Hounsfield Units (HU)—and the S.T.O.N.E. score, which assesses size, topography, obstruction, number of stones, and HU, have been developed to predict surgical outcomes and forecast success rates in patients undergoing ULT.¹² The S.T.O.N.E. score has shown promise; an internal validation study reported a predictive accuracy (area under the curve) of 0.764.¹⁰ However, the absence of external validation raises concerns regarding its generalizability and effectiveness across diverse patient populations, particularly in resource-limited environments.¹³ Therefore, this study aims to externally validate the S.T.O.N.E. score for predicting SFR in patients undergoing semirigid pneumatic ULT, addressing the critical need for validation and providing insights into its effectiveness across various clinical settings.

Materials and Methods

Study Design and Patient Population

A retrospective cohort study was conducted at IBB University Hospitals involving patients diagnosed with ureteral stones who underwent semirigid pneumatic ULT between April 2021 and September 2023. Of 613 cases, 266 patients met the established inclusion criteria and were deemed eligible for the study; excluding participants with any missing predictor data. The exclusion criteria and corresponding numbers are presented in the flowchart ([Supplementary Figure 1](#)). Exclusion criteria comprised anatomical anomalies of the ureters, active urinary tract infections, absence of preoperative computed tomography (CT) images, pregnancy, and acute renal failure. The study protocol received approval from the ethics committee of IBB University, and it was conducted in accordance with the principles outlined in the Helsinki Declaration. Given the retrospective nature of the study, patient consent for chart review was not required, and all data were anonymized to ensure confidentiality.

Sample Size Calculation

In accordance with the TRIPOD guidelines, an adequate sample size for validating prediction models was determined to require a minimum of 100 events and 100 non-events, thereby ensuring sufficient statistical power.¹⁴ Given that the success rate of semi-rigid ULT ranged from 80% to 91%, the estimated proportion of success was approximately 85.5%.^{8,15,16} To calculate the necessary sample size, a 95% confidence level and a margin of error of $\pm 5\%$ were employed, resulting in an estimated sample size of approximately 199 patients. However, in order to align with previous estimates and to enhance the robustness of the findings, a validation cohort of between 200 and 300 patients was ultimately sought.

Preoperative Imaging and S.T.O.N.E. Score Assessment

Prior to ULT, all patients underwent low-dose unenhanced CT scans to assess urinary tract stones. Measurements were performed by two experienced urologists in collaboration with two experienced radiologists, who were blinded to clinical outcomes. The largest stone diameter and volume were calculated using ellipsoid formulas based on axial and coronal reformations obtained from the CT scans. The S.T.O.N.E. score, which evaluates stone characteristics including size, topography, obstruction, number of stones, and HU, was computed from preoperative CT findings. Each parameter was assessed on a 1–3 point scale, with higher scores indicating greater complexity and lower SFRs.¹⁰ The final S.T.O.N.E. score thus ranges from 5 to 15 points, with detailed parameters outlined in [Supplementary Table 1](#).

Surgical Procedure

ULT was performed under general or regional anesthesia following informed consent. Antibiotic prophylaxis was administered immediately prior to the procedure. Patients were positioned in dorsal lithotomy for the procedure. The operation commenced with a rigid cystoscopy to visualize the bladder and locate the ureteric orifice, followed by a guidewire introduction into the ureter. A semirigid ureteroscopy (6–8Fr, Karl Storz, Tuttlingen, Germany) was utilized in all cases. During lithotripsy, a Swiss LithoClast® Pneumatic Lithotripter was employed. Residual stone fragments were managed based on size, either removed or left un-retrieved. At the end of the procedure, a double-J (DJ) stent was inserted as per clinical judgment, which was removed 2 weeks later following surgery.

Primary and Secondary Outcomes

The primary outcome was the efficacy of the S.T.O.N.E. score in predicting stone-free rate (SFR) failure after ULT. Secondary outcomes included various predictive factors associated with SFR failure.

Collected Data

Data were systematically extracted from patient records, encompassing demographics, stone characteristics, operative details, postoperative complications, and outcome. Postoperative complications were graded using the modified Clavien–Dindo classification system. Stone-free status was defined as the absence of stone fragments or the presence of fragments ≤ 2 mm post-ULT, verified either through endoscopic examination, fluoroscopy, or radiologic images such as ultrasonography, radiography, or CT imaging within one month following the procedure. Patients with residual lithiasis underwent a second procedure (ESWL, a second ULT, or PCNL) as per clinical judgment to obtain stone clearance.

Statistical Analysis

Statistical analyses were performed using IBM SPSS version 22 (IBM Corp, Armonk, NY) and Python with the Scikit-Learn and Stats models libraries. Continuous variables were expressed as means with standard deviations, while categorical variables were reported as frequencies and percentages. The normality of continuous data was assessed using the Shapiro–Wilk test, with independent *t*-tests applied for normally distributed variables and the Mann–Whitney *U*-test for non-normally distributed variables. Categorical variables were analyzed using the chi-square test or Fisher’s exact test, as appropriate. Univariate logistic regression analysis was conducted to identify predictors of SFR failure. Variables with *p*-values < 0.05 in the univariate analysis were included in a multivariate logistic regression model. Results were presented as adjusted odds ratios (AOR) with 95% confidence intervals (CI), and a significance level of $p < 0.05$ was established.

The diagnostic performance of the S.T.O.N.E. score was evaluated using receiver operating characteristic (ROC) curves, and the area under the curve (AUC) was calculated to assess predictive accuracy. The AUC was classified as strong ($AUC > 0.9$), moderate ($0.7 < AUC \leq 0.9$), or low ($0.5 < AUC \leq 0.7$) predictive capacity.¹⁷ The optimal threshold for the S.T.O.N.E. score was determined using the maximum Youden index, which maximizes the difference between sensitivity and specificity. At this threshold, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated.

To assess the model calibration, calibration plots and the Hosmer–Lemeshow test were employed.¹⁸ The Hosmer–Lemeshow test evaluated the goodness-of-fit of the logistic regression model, with a *p*-value > 0.05 indicating good

calibration. Additionally, the Brier score was calculated to measure the overall predictive accuracy of the model, with lower values indicating better performance.

Internal validation was performed using bootstrapping with 1000 resamples to evaluate the stability of the model's performance metrics. The 95% CI for accuracy, sensitivity, specificity, PPV, and NPV were calculated to assess the robustness of the results. Decision Curve Analysis (DCA) was conducted to evaluate the clinical utility of the S.T.O.N.E. score by comparing the net benefits across different threshold probabilities.¹⁹ All analyses adhered to the TRIPOD guidelines for transparent reporting of predictive models.¹⁴

Results

Patients, Radiologic, and Operative Characteristic

The mean age of patients was 47.7 ± 15.0 years (range: 18.0–91.0), with 192 males (72.2%). Notable comorbidities included previous extracorporeal shock wave lithotripsy (11.7%), diabetes (4.9%), and hypertension (3.8%). Right-sided stones were found in 141 patients (53.0%), and 168 patients (63.2%) reported acute flank pain. The mean stone diameter was 9.1 ± 4.9 mm, with 31.2% having stones greater than 10 mm. Of the patients, 76.3% had Hounsfield Units (HUs) less than 750, and hydronephrosis was classified as mild in 73.3%, moderate in 20.7%, and severe in 6.0%. Most patients had one stone (68.8%), while 27.1% had two, and 4.1% had more than two. The mean S.T.O.N.E. score was 7.8 ± 1.8 , and the mean operative time was 57.7 ± 7.3 minutes (Table 1).

Table 1 Demographic, Preoperative, Radiologic, and Postoperative Characteristics

Variables	Overall (N=266)
Age (year), Mean \pm SD	47.7 ± 15.0 (Range 18.0–91.0)
Gender	
Male	192 (72.2%)
Female	74 (27.8%)
Weight (kg), Mean \pm SD	72.7 ± 9.8 (Range 46.0–110.0)
Comorbidities	
Diabetes	13 (4.9%)
Hypertension	10 (3.8%)
Previous history of urolithiasis	31 (11.7%)
Laterality	
Right	141 (53.0%)
Left	125 (47.0%)
Symptoms presentation	
Acute flank pain	168 (63.2%)
Hematuria	38 (14.3%)
Fever	60 (22.6%)
Stone size (mm), Mean \pm SD	9.1 ± 4.9 (Range 2.0–24.0)

(Continued)

Table 1 (Continued).

Variables	Overall (N=266)
Stone size subgroup	
< 5 mm	68 (25.6%)
5–10 mm	115 (43.2%)
> 10 mm	83 (31.2%)
HUs, Mean \pm SD	523.4 \pm 281.4 (Range 100.0–1351.0)
HU's subgroup	
< 750 hU	203 (76.3%)
750–1000 hU	38 (14.3%)
> 1000 hU	25 (9.4%)
Degree of hydronephrosis	
Mild	195 (73.3%)
Moderate	55 (20.7%)
Severe	16 (6.0%)
Stone location	
Distal ureter	70 (26.3%)
Middle ureter	188 (70.7%)
Upper ureter	8 (3.0%)
Stone numbers	
One stone	183 (68.8%)
2 stones	72 (27.1%)
3 stones	11 (4.1%)
Stone free rate	
Success	227 (85.3%)
Failure	39 (14.7%)
S.T.O.N.E. Score, Mean \pm SD	7.8 \pm 1.8 (Range 5.0–13.0)

Abbreviations: SD, standard deviation; HU, Hounsfield units.

The overall complication rate was 10.2%. Intraoperative complications occurred in 16 patients (6.0%), primarily mucosal damage (9 patients, 3.4%) and stone migration (8 patients, 3.0%). Postoperative complications were noted in 13 patients (4.9%), with fever being the most common (7 patients, 2.6%). Two patients experienced significant postoperative issues: one required intravenous antibiotics and intensive care unit (ICU) admission for pyelonephritis, while the other unfortunately succumbed to uncontrolled diabetes, sepsis, and multiorgan failure (Table 2). The overall SFR was 85.3%, as confirmed by postoperative imaging, including CT scans, with a mean follow-up of 20 months. A total of 39 patients (14.7%) had residual stones and needed auxiliary endourological procedures.

Table 2 Ureteroscopic Lithotripsy Complications Characteristics

Complication *	N (%)
Intraoperative	16 (6.0%)
Mucosal damage	9 (3.4%)
Stone up-migration or retropulsion	8 (3.0%)
Ureteral perforations	1 (0.4%)
Postoperative	13 (4.9%)
Fever (Grade I)	7 (2.6%)
Hematuria (Grade I)	3 (1.1%)
Febrile UTI (Grade II)	1 (0.4%)
Pyelonephritis (Grade III)	1 (0.4%)
Death (Grade 5)	1 (0.4%)

Note: * Some patients had multiple complications.

Abbreviation: UTI, urinary tract infection.

Factors Associated with SFR Failure

Univariate analysis indicated that prior ESWL history, larger stone size, proximal stone location, moderate to severe hydronephrosis, and higher HU were significantly associated with SFR failure (all $p < 0.05$) (Table 3). Multivariate analysis identified higher HU (AOR: 1.01; 95% CI: 1.00–1.01; $P < 0.001$), proximal stones (AOR: 15.13; $P = 0.020$),

Table 3 Factors Associated with the Stone-Free Rate in Monovariate Analysis

Variables	Subgroup	SFR		p-value
		Yes (227)	No (39)	
Age (year)	Mean \pm SD	47.3 \pm 15.1	49.9 \pm 14.3	0.314
Gender	Male	159 (70.0)	33 (84.6)	0.092
	Female	68 (30.0)	6 (15.4)	
Presentation	Acute flank pain	139 (61.2)	29 (74.4)	0.256
	Hematuria	35 (15.4)	3 (7.7)	
	Fever	53 (23.3)	7 (17.9)	
Laterality	Right	122 (53.7)	19 (48.7)	0.684
	Left	105 (46.3)	20 (51.3)	
Previous history of urolithiasis	No	207 (91.2)	28 (71.8)	0.001
	Yes	20 (8.8)	11 (28.2)	
Weight (Kg)	Mean \pm SD	72.8 \pm 9.9	72.0 \pm 9.2	0.636
Diabetes	No	218 (96.0)	35 (89.7)	0.200
	Yes	9 (4.0)	4 (10.3)	

(Continued)

Table 3 (Continued).

Variables	Subgroup	SFR		p-value
		Yes (227)	No (39)	
Hypertension	No	218 (96.0)	38 (97.4)	1.000
	Yes	9 (4.0)	1 (2.6)	
Stone size (mm)	Mean \pm SD	8.1 \pm 4.3	15.0 \pm 4.1	<0.001
Stone density (HU)	Mean \pm SD	493.5 \pm 272.0	697.7 \pm 275.2	<0.001
Degree of hydronephrosis	Mild	179 (78.9)	16 (41.0)	<0.001
	Moderate	36 (15.9)	19 (48.7)	
	Severe	12 (5.3)	4 (10.3)	
Stone number	One stone	180 (79.3)	3 (7.7)	<0.001
	2 stones	43 (18.9)	29 (74.4)	
	3 stones	4 (1.8)	7 (17.9)	
Stone location	Distal	67 (29.5)	3 (7.7)	<0.001
	Mid	158 (69.6)	30 (76.9)	
	Upper	2 (0.9)	6 (15.4)	

Note: A bold p-value < 0.05 indicates statistical significance.

Abbreviations: HU, Hounsfield Units; SFR, Stone Free Rate; SD, Standard Deviation.

moderate hydronephrosis (AOR: 34.23; $P < 0.001$), severe hydronephrosis (AOR: 33.75; $P = 0.0006$), and larger stones (AOR: 1.51; $P < 0.0001$) as predictors of SFR failure (Table 4). Among the S.T.O.N.E. parameters, stone size, HU, and stone number demonstrated moderate predictive capabilities for SFR failure, with AUC values of 0.883, 0.726, and 0.754, respectively (Supplementary Table 2).

Assessing the Predictive Performance of the S.T.O.N.E. Score for SFR Failure

The logistic regression model demonstrated strong predictive performance for SFRs using the S.T.O.N.E. score. The optimal probability threshold was determined to be 0.45, corresponding to a S.T.O.N.E. score of 8.2. Patients with a S.T.O.N.E. score ≥ 8.2 were classified as having a higher probability of failure, while those with a score < 8.2 were classified as having a higher probability of success. At this threshold, the model achieved an accuracy of 85.0% (95% CI: 0.820 to 0.880), correctly classifying 85.0% of the cases. The sensitivity was 72.0% (95% CI: 0.680 to 0.760), and the specificity was 81.0% (95% CI: 0.770 to 0.850). The PPV was 75.0% (95% CI: 0.710 to 0.790), and the NPV was 83.0% (95% CI: 0.800 to 0.860), demonstrating strong predictive performance.

The ROC curve demonstrated good discrimination ability, with an AUC of 0.785 (95% CI: 0.750, 0.820) (Figure 1). At the optimal threshold, the model achieved a sensitivity of 72.0%, specificity of 81.0%, PPV of 75.0%, and NPV of 83.0%, indicating strong predictive performance. The calibration plot showed good agreement between the predicted probabilities and the observed failure rates, particularly in the mid-range of S.T.O.N.E. scores (Figure 2A). The Hosmer-Lemeshow test yielded a p-value of 0.62, indicating no significant deviation between the predicted and observed outcomes ($p > 0.05$). This suggests that the model is well-calibrated and fits the data adequately. Additionally, the Brier score was 0.12, indicating good overall predictive accuracy.

To evaluate the stability of the model's performance metrics, bootstrapping was performed with 1000 resamples. The results confirmed the robustness of the model, revealing narrow 95% confidence intervals for all metrics. Specifically, the model achieved an accuracy of 85.0% (95% CI: 0.820, 0.880), a sensitivity of 72.0% (95% CI: 0.680, 0.760), and

Table 4 Univariate and Multivariate Regression for Predictive Factors of Stone-Free Rate

Variables	Subgroups	SFR (N=227)	No SFR (N=39)	Univariate		Multivariate			
				OR (95% CI)	P-value	Crude OR (95% CI)	P-value	Adjusted OR (95% CI)	P-value
Previous history of urolithiasis	No	207 (88.1)	28 (11.9)	Ref		Ref		-	-
	Yes	20 (64.5)	11 (35.5)	4.07 (1.72–9.28)	0.001	1.64 (0.34–7.32)	0.521	-	-
Stone HU	Mean (SD)	493.5 (272.0)	697.7 (275.2)	1.00 (1.00–1.00)	<0.001	1.00 (1.00–1.01)	0.001	1.01 (1.0023–1.0065)	<0.0001
Hydronephrosis	Mild	179 (91.8)	16 (8.2)	Ref		Ref		Ref	
	Moderate	36 (65.5)	19 (34.5)	5.90 (2.78–12.72)	<0.001	36.00 (9.17–182.40)	<0.001	34.23 (8.28–141.45)	<0.0001
	Severe	12 (75.0)	4 (25.0)	3.73 (0.96–12.20)	0.038	27.29 (3.83–222.14)	0.001	33.75 (4.55–250.36)	0.0006
Stone Location	Distal	67 (95.7)	3 (4.3)	Ref		Ref		Ref	
	Middle	158 (84.0)	30 (16.0)	4.24 (1.45–18.11)	0.020	8.35 (1.70–57.81)	0.016	0.127 (0.024–0.71)	0.0187
	Proximal	2 (25.0)	6 (75.0)	67.00 (10.89–636.72)	<0.001	161.56 (11.04–3958.46)	0.001	15.13 (1.52151.13)	0.0207
Stone Size (mm)	Mean (SD)	8.1 (4.3)	15.0 (4.1)	1.32 (1.22–1.44)	<0.001	1.45 (1.25–1.72)	<0.001	1.51 (1.30–1.75)	<0.0001
Stone Number	Single	121 (99.2)	1 (0.8)	Ref		Ref		-	-
	Multiple	106 (73.6)	38 (26.4)	43.38 (9.16–776.53)	<0.001	10.27 (0.81–470.07)	0.138	-	-

Note: A bold p-value < 0.05 indicates statistical significance.

Abbreviations: CI, Confidence Interval; HU, Hounsfield Units; OR, Odds Ratio; SD, Standard Deviation; SFR, Stone Free Rate.

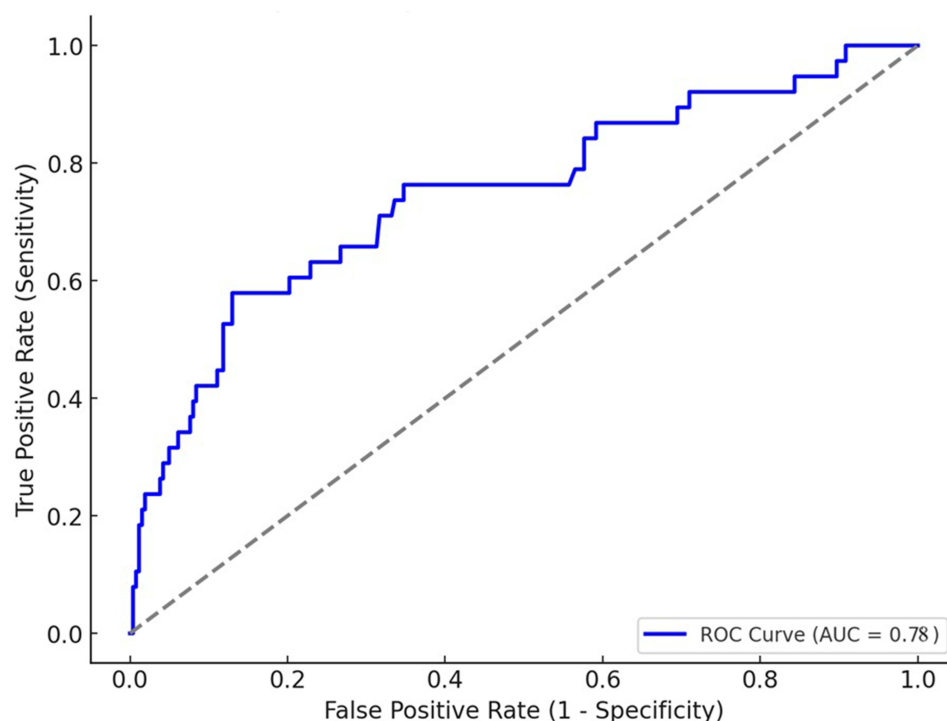


Figure 1 Receiver Operator Characteristic curve for the S.T.O.N.E. score, showing an AUC of 0.78, indicating good discrimination in identifying patients with SFR failure.

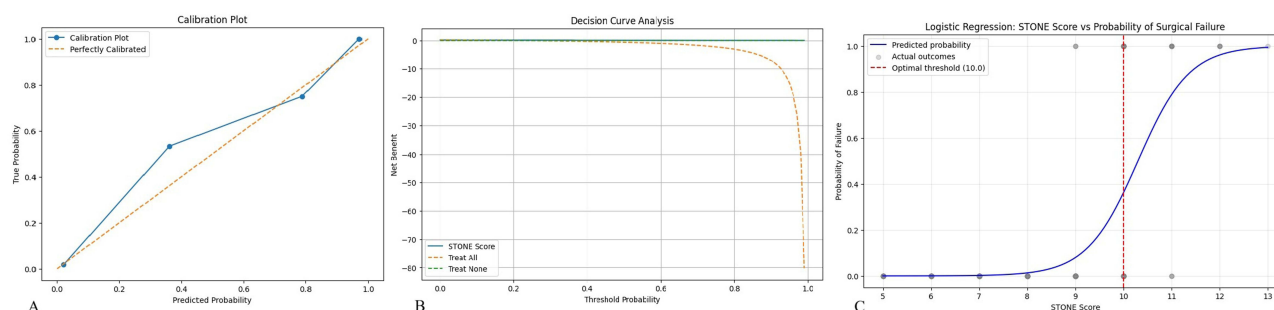


Figure 2 (A) The calibration plot shows how well the predicted probabilities align with actual outcomes. (B) Decision Curve Analysis (DCA): Net benefit of the S.T.O.N.E. score model (blue curve) compared to treating all patients as having SFR failure (red dashed line) or treating none (gray dashed line). (C) Logistic Regression: S.T.O.N.E. Score vs Probability of Surgical Failure: The curve represents the predicted probability of surgical failure (y-axis) as a function of the S.T.O.N.E. score (x-axis). Higher S.T.O.N.E. scores are associated with increased probabilities of failure, demonstrating the model's ability to predict surgical outcomes.

a specificity of 81.0% (95% CI: 0.770, 0.850). Additionally, the PPV was 75.0% (95% CI: 0.710, 0.790), while the NPV was 83.0% (95% CI: 0.800, 0.860).

Decision Curve Analysis (DCA) was performed to assess the clinical utility of the S.T.O.N.E. score. The analysis demonstrated that using the S.T.O.N.E. score for decision-making provided a net benefit over a range of threshold probabilities (eg, 20–80%), compared to the strategies of treating all or none of the patients (Figure 2B). This indicates that the S.T.O.N.E. score is clinically useful for guiding treatment decisions.

The predicted probabilities of SFR failure increased with higher S.T.O.N.E. scores, demonstrating the model's ability to differentiate between low- and high-risk patients. Patients with a S.T.O.N.E. score of 7 had a very low predicted probability of failure (0.002), while those with a score of 10 had a higher probability (0.362), and those with a score of 11 had a very high probability (0.788). The model performed well at the extremes of the S.T.O.N.E. score range, correctly predicting outcomes

for most patients (Figure 2C). However, for patients with a S.T.O.N.E. score of 9, the model assigned a moderate probability of failure (0.079), reflecting the transitional nature of this score where outcomes were less certain.

Discussion

The management of urolithiasis has increasingly shifted towards minimally invasive endoscopic lithotripsy, but financial constraints in developing countries limit access to essential equipment. Consequently, proficiency in semirigid URS for ureteral stones is crucial, despite associated risks and lower SFRs.²⁰ This study evaluated the predictive efficacy of the S.T.O.N.E. score for SFR failure following semirigid pneumatic ULT in our cohort. The S.T.O.N.E. score effectively predicts SFR failure, with an optimal threshold of > 8 achieving 85% accuracy. The model demonstrated 72% sensitivity, 81% specificity, and strong calibration. DCA indicated clinical utility by differentiating between low- and high-risk patients based on their S.T.O.N.E. scores.

Complications from ULT can be classified as intraoperative or postoperative, with reported overall complication rates ranging from 9% to 25%.²¹ In our study, the intraoperative complication rate was 6.0%, including mucosal damage (3.4%), stone up-migration (3.0%), and ureteral perforations (0.4%). In another report, Abdelrahim et al observed a 27.4% complication rate during rigid ULT, with stone migration, bleeding, and mucosal damage comprising the majority.²² Other studies reported rates of 4.4% and 3.6%, detailing similar complications.^{23,24} Variability across studies in complication rate may stem from differences in design, surgical expertise, and patient or stone characteristics. Our findings align with existing literature, with no major complications reported, likely due to the surgeon's experience and the referral of complex cases to specialized teams at tertiary centers.

Postoperative complication rates varied significantly across studies. De Coninck et al documented rates of postoperative fever and UTIs from 0.2% to 15%, renal colic from 1.1% to 10.2%, and urinary retention from 0.1% to 1.4%.²⁵ Perez et al noted fever as the most common complication, followed by UTIs and bladder discomfort.²⁶ In our study, the overall postoperative complication rate was 4.9%, with fever being the most frequent complication. Major complications were observed in two cases: one case of pyelonephritis required ICU admission and intravenous antibiotics, while the other patient succumbed to uncontrolled diabetes, sepsis, and multiorgan failure. Minimizing complications requires practical experience, careful patient selection, and thorough preoperative assessments.

In our current analysis, we found an overall SFR of 85.3%. This result aligns with previous reports on urinary tract lithiasis stone clearance, such as Alameddine et al, who reported a clearance rate of 89.0%;²⁷ Shrestha et al, with a rate of 80.5%;⁸ Kim et al with 85.7%;⁹ and Sirirak et al, who reported 89.68%.²⁸ However, a study in Ethiopia showed a lower clearance rate of 54.7%, attributed to lack of experience and advanced lithotripters.⁴ The concept of SFR has been inconsistently applied across studies, leading to discrepancies in reported rates. Different imaging methods, including X-rays, non-contrast CT scans, and ultrasounds, each with varying specificities and sensitivities, may have influenced results. Thus, variations in reported stone clearance and failure rates likely arise from a lack of standardization, methodological differences, and variability in research populations.

Stone characteristics including density, size, location, and hydronephrosis, are crucial for predicting therapeutic outcomes and estimating stone clearance after ULT.²⁸ In our study, stone density (HU) was statistically linked to SFRs, with higher HU values correlating with lower SFRs. This aligns with findings from Shrestha et al, Yang et al, and Hori et al, indicating that HU values affect ULT efficacy.^{8,29,30} In contrast, Kim et al and Sirirak et al found no significant effect of HU on ULT effectiveness.^{9,28} Discrepancies may arise from variations in measuring HU and different cutoff values influenced by CT image quality and scanning protocols.

The presence of larger stones or a greater number of stones increases the likelihood of impaction, which may adversely affect the achievement of SFR. Prolonged obstruction can result in histological alterations within the ureter, including smooth muscle hypertrophy and collagen deposition, leading to inflammatory responses, edema, and fibrosis.²⁹ In this study, we identified stone size and the degree of hydronephrosis as independent predictors of stone clearance, whereas stone number did not demonstrate statistical significance in regression analysis, potentially attributable to the limited sample size. These findings are consistent with those of prior studies investigating semirigid pneumatic ULT. For instance, Hong et al reported a decline in success rates correlated with increasing stone size and severity of hydronephrosis.³¹ Similarly, Kurahashi et al determined that the number, location, and maximum diameter of stones were independent predictors of

stone clearance.³² Moreover, while Sirirak et al and Yang et al highlighted significant variations in SFRs based on stone size, the former also noted a lack of significant differences associated with the number of calculi.^{28,30}

In our study, the SFR for patients with upper ureteral stones was significantly lower than that for those with middle and distal stones, which aligns with the understanding that SFR improves as the stone's location shifts distally within the ureter.^{26,31} A meta-analysis indicates that clearance rates are substantially higher for distal stones (93%) than for proximal stones (79%) following ureteroscopic surgery.³³ The inferior outcomes observed for upper ureteral stones may be attributed to factors such as increased distance and technical challenges in accessing the stones, along with a greater tendency for stone fragments to migrate back to the renal pelvis.³⁰ Although other variables—including the surgeon's experience, utilization of auxiliary equipment, the type of lithotripsy devices employed, and prior experience with ESWL—can substantially influence stone clearance, these factors were not assessed in our study. This limitation is primarily due to the unavailability of laser lithotripters at our center and the deliberate decision to avoid the use of baskets in the majority of ULT procedures.

Our findings on the efficacy of the S.T.O.N.E. score align with existing literature and provide valuable clinical insights. The S.T.O.N.E. score demonstrated an AUC of 0.785 for predicting SFR failure, comparable to AUCs of 0.815, 0.764, and 0.644 reported by Sirirak et al.²⁸ Molina et al¹ and Richard et al³⁴ respectively, reinforcing its moderate discriminatory ability across diverse populations. In comparison, tools such as the RUSS score (AUC: 0.617), Ito's nomogram (AUC: 0.735), and the S-ReSC score (AUC: 0.735) exhibit similar abilities, underscoring the competitiveness of the S.T.O.N.E. score in stratifying stone complexity and treatment outcomes.³⁴ We recommend further validation of the score's utility by comparing it with other predictive models in similar contexts. The strong calibration of our model, evidenced by the Hosmer-Lemeshow test ($p = 0.62$) and calibration plot, shows good agreement between predicted and observed outcomes, consistent with findings from Wu et al¹² who reported excellent calibration of the S.T.O.N.E. score in 275 patients with upper ureteral stones. A Brier score of 0.12 further confirms the model's predictive accuracy, aligning with scores from other studies evaluating similar predictive models.^{10,34,35}

One strength of our study is the use of bootstrapping for internal validation, which demonstrated the robustness of the model's performance metrics. Narrow 95% CIs for accuracy, sensitivity, specificity, PPV, and NPV indicate that the results are reliable and generalizable to similar patient populations, aligning with TRIPOD guidelines for the transparent reporting of predictive models.¹⁴

At the optimal S.T.O.N.E. score threshold of 8.2, our model achieved a sensitivity of 72.0% and a specificity of 81.0%, with a PPV of 75.0% and a NPV of 83.0%. These performance metrics are in line with those reported in other studies. For instance, Molina et al¹ found that a S.T.O.N.E. score of ≤ 9 points was associated with a SFR exceeding 90%, while Hori et al²⁹ reported that a score of ≥ 9 points correlated with decreased SFR. The study's results confirm the S.T.O.N.E. score as a reliable predictor of SFR, indicating consistent performance in identifying patients at risk of SFR failure. The high NPV suggests that the S.T.O.N.E. score effectively rules out failure in patients with low scores, potentially helping clinicians avoid unnecessary interventions. However, the homogeneity of our patient population—predominantly those with middle and distal ureter stones—may have influenced the results, as complex cases at higher risk of failure were referred to specialized teams, reducing variability within our cohort. Future research should evaluate the S.T.O.N.E. score in more diverse populations to assess the broader applicability of these findings.

Study Limitations

This study acknowledges several limitations. First, the sample size, while adequate for model development and validation, may restrict the generalizability of the findings to broader populations. The retrospective design also presents challenges, as reliance on secondary data can introduce variability due to documentation inconsistencies and potential biases. Furthermore, the study was conducted at a single center, which may lead to selection bias. Additionally, outcomes may have been positively influenced by the involvement of experienced urologists in all procedures. Technical limitations were noted in defining stone clearance, and external validation with larger cohorts is necessary to confirm the applicability of the S.T.O.N.E. score across diverse patient populations. Nonetheless, this investigation represents the first report correlating stone clearance with the impact of the S.T.O.N.E. score on SFR in patients undergoing semirigid pneumatic ULT at our institution. Therefore, further research to validate the scoring system in larger, multicenter studies is warranted to address specific treatment challenges in resource-limited settings.

Conclusion

We found that the S.T.O.N.E. score serves as a robust predictor of failure in achieving SFR following ULT, with an optimal threshold of 8.2 that effectively stratifies patients into higher- and lower-risk categories. The model demonstrated an overall accuracy of 85.0%, with a sensitivity of 72.0% and a specificity of 81.0%, underscoring its potential utility in clinical practice. The DCA further substantiates the model's clinical significance, demonstrating its ability to inform treatment decisions and support personalized management strategies for urinary tract lithiasis. Future research should emphasize the integration of the S.T.O.N.E. score into routine clinical workflows and its validation across diverse patient populations and clinical settings. Additionally, further multicenter studies are warranted to confirm its generalizability and enhance its applicability in the management of urolithiasis.

Data Sharing Statement

The datasets used and analyzed during the current study are available on Mendeley and can be accessed via the following DOI: 10.17632/d4nb59txxh.1.

Ethics Approval and Consent to Participate

The Ibb University Ethics Committee approved the study protocol (ID number: IBBUNI.AC.YEM.2024.01.93) on February 3, 2024. The study complies with the Declaration of Helsinki. Given the retrospective nature of the study, patient consent for chart review was not required, and all data were anonymized to ensure confidentiality.

Consent for Publication

Written informed consent was obtained from the patients including the figures. Parent or legal guardian of patients under 18 years of age provided informed consent for participation and publication of the case details and images.

Author Contributions

All authors made significant contributions to the work reported, including conception, study design, execution, data acquisition, analysis, interpretation, drafting, revising, and critically reviewing the article. They also provided final approval of the version to be published, agreed on the journal to which the article has been submitted, and accept accountability for all aspects of the work.

Funding

There is no funding to report.

Disclosure

The authors declare that they have no competing interests.

References

1. Stamatelou K, Goldfarb DS. Epidemiology of kidney stones. *Healthcare*. 2023;11(3). doi:10.3390/healthcare11030424
2. Soylu A, Sarier M, Altunoluk B, Soylemez H, Baydinc YC. Comparison of the efficacy of intravenous and intramuscular lornoxicam for the initial treatment of acute renal colic: a randomized clinical trial. *Urol J*. 2019;16(1):16–20. doi:10.22037/uj.v0i0.4496
3. Sarier M, Duman I, Callioglu M, et al. Outcomes of conservative management of asymptomatic live donor kidney stones. *Urology*. 2018;118:43–46. doi:10.1016/j.urology.2018.04.035
4. Mohammed S, Redi S, Berhe T, Teshome H. Ureterscopy outcome and its determinants in a resource-limited setting. *Ethiop J Health Sci*. 2022;32(5):947–954. doi:10.4314/ejhs.v32i5.10
5. Wason SE, Monfared S, Ionson A, Klett DE, Leslie SW. *Ureterscopy*. Treasure Island (FL): StatPearls Publishing; 2023.
6. Shirazi M, Aminsharifi A, Ahmed F, Makarem A, Zahraei SA, Asmaarian N. The impact of post-procedural ureteric stent duration on the outcome of retrograde endopyelotomy for management of failed open pyeloplasty in children: a preliminary report.. *Med J Islam Repub Iran*. 2020;34:105. doi:10.34171/mjiri.34.105
7. Mustafa M, Al Zabadi H, Mansour S, Nabulsi A. Endoscopic management of upper and lower ureteric stones using pneumatic lithotripter: a retrospective medical records review. *Res Rep Urol*. 2023;15:77–83. doi:10.2147/RRU.S392881

8. Shrestha B, Koju R, Makaju Shrestha S, Shrestha K, Karmacharya RM. Predictors of stone free rate and application of the size, topography, obstruction, number and evaluation of Hounsfield units (S.T.O.N.E) scoring system in predicting the outcome in patients undergoing semi-rigid ureteroscopic lithotripsy for ureteric calculi at a university hospital of Nepal. *Kathmandu Univ Med J KUMJ*. 2024;22(85):31–35.
9. Kim JW, Chae JY, Kim JW, et al. Computed tomography-based novel prediction model for the stone-free rate of ureteroscopic lithotripsy. *Urolithiasis*. 2014;42(1):75–79. doi:10.1007/s00240-013-0609-0
10. Molina WR, Kim FJ, Spendlove J, Pompeo AS, Sillau S, Sehr DE. The S.T.O.N.E. Score: a new assessment tool to predict stone free rates in ureteroscopy from pre-operative radiological features. *Int Braz J Urol*. 2014;40(1):23–29. doi:10.1590/S1677-5538.IBJU.2014.01.04
11. Nour HH, Kamel AI, Elmansy H, et al. Pneumatic vs laser lithotripsy for mid-ureteric stones: clinical and cost effectiveness results of a prospective trial in a developing country. *Arab J Urol*. 2020;18(3):181–186. doi:10.1080/2090598X.2020.1749800
12. Wu W, Zhang J, Yi R, Li X, Wan W, Yu X. A simple predictive model with internal validation for assessment of stone-left after ureteroscopic lithotripsy in upper ureteral stones. *Transl Androl Urol*. 2022;11(6):786–793. doi:10.21037/tau-22-22
13. Wang RC, Rodriguez RM, Moghadassi M, et al. External validation of the STONE score, a clinical prediction rule for ureteral stone: an observational multi-institutional study. *Ann Emerg Med*. 2016;67(4):423–432.e422. doi:10.1016/j.annemergmed.2015.08.019
14. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Eur Urol*. 2015;67(6):1142–1151. doi:10.1016/j.eururo.2014.11.025
15. Binbay M, Tepeler A, Singh A, et al. Evaluation of pneumatic versus holmium:YAG laser lithotripsy for impacted ureteral stones. *Int Urol Nephrol*. 2011;43(4):989–995. doi:10.1007/s11255-011-9951-8
16. Yagisawa T, Kobayashi C, Ishikawa N, Kobayashi H, Toma H. Benefits of ureteroscopic pneumatic lithotripsy for the treatment of impacted ureteral stones. *J Endourol*. 2001;15(7):697–699. doi:10.1089/08927790152596262
17. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29–36. doi:10.1148/radiology.143.1.7063747
18. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230. doi:10.1186/s12916-019-1466-7
19. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565–574. doi:10.1177/0272989X06295361
20. Anan G, Kudo D, Matsuoka T. What are the predictors of residual stone after ureteroscopy for urolithiasis? *Transl Androl Urol*. 2022;11(8):1071–1073. doi:10.21037/tau-22-438
21. Preminger GM, Tiselius HG, Assimos DG, et al. 2007 guideline for the management of ureteral calculi. *J Urol*. 2007;178(6):2418–2434. doi:10.1016/j.juro.2007.09.107
22. Abdelrahim AF, Abdelmaguid A, Abuzeid H, Amin M, Mousael S, Abdelrahim F. Rigid ureteroscopy for ureteral stones: factors associated with intraoperative adverse events. *J Endourol*. 2008;22(2):277–280. doi:10.1089/end.2007.0072
23. Fuganti PE, Pires S, Branco R, Porto J. Predictive factors for intraoperative complications in semirigid ureteroscopy: analysis of 1235 ballistic ureterolithotripsies. *Urology*. 2008;72(4):770–774. doi:10.1016/j.urology.2008.05.042
24. Geavlete P, Georgescu D, Niță G, Mirciulescu V, Cauni V. Complications of 2735 retrograde semirigid ureteroscopy procedures: a single-center experience. *J Endourol*. 2006;20(3):179–185. doi:10.1089/end.2006.20.179
25. De Coninck V, Keller EX, Somani B, et al. Complications of ureteroscopy: a complete overview. *World J Urol*. 2020;38(9):2147–2166. doi:10.1007/s00345-019-03012-1
26. Perez Castro E, Osther PJ, Jinga V, et al. Differences in ureteroscopic stone treatment and outcomes for distal, mid-, proximal, or multiple ureteral locations: the clinical research office of the endourological society ureteroscopy global study. *Eur Urol*. 2014;66(1):102–109. doi:10.1016/j.eururo.2014.01.011
27. Alameddine M, Azab MM, Nassir AA. Semi-rigid ureteroscopy: proximal versus distal ureteral stones. *Urol Ann*. 2016;8(1):84–86. doi:10.4103/0974-7796.171495
28. Sirirak N, Sangkum P, Phengsalae Y, et al. External validation of the S.T.O.N.E. score in predicting stone-free status after rigid ureteroscopic lithotripsy. *Res Rep Urol*. 2021;13:147–154. doi:10.2147/RRU.S304221
29. Hori S, Otsuki H, Fujio K, et al. Novel prediction scoring system for simple assessment of stone-free status after flexible ureteroscopy lithotripsy: t. O.H.O. score. *Int J Urol*. 2020;27(9):742–747. doi:10.1111/iju.14289
30. Yang B, Sun S, Wang J, et al. Novel scoring system for predicting stone-free rate after flexible ureteroscopy lithotripsy. *Medicine*. 2024;103(44):e40390. doi:10.1097/MD.00000000000040390
31. Hong YK, Park DS. Ureteroscopic lithotripsy using Swiss Lithoclast for treatment of ureteral calculi: 12-years experience. *J Korean Med Sci*. 2009;24(4):690–694. doi:10.3346/jkms.2009.24.4.690
32. Kurahashi T, Miyake H, Oka N, et al. Clinical outcome of ureteroscopic lithotripsy for 2129 patients with ureteral stones. *Urol Res*. 2007;35(3):149–153. doi:10.1007/s00240-007-0095-3
33. Preminger GM, Tiselius HG, Assimos DG, et al. 2007 Guideline for the management of ureteral calculi. *Eur Urol*. 2007;52(6):1610–1631. doi:10.1016/j.eururo.2007.09.039
34. Richard F, Marguin J, Frontczak A, et al. Evaluation and comparison of scoring systems for predicting stone-free status after flexible ureteroscopy for renal and ureteral stones. *PLoS One*. 2020;15(8):e0237068. doi:10.1371/journal.pone.0237068
35. Malik A, Mohkumuddin S, Yousaf S, Baig MAR, Afzal A. Validity of STONE score in clinical prediction of ureteral stone disease. *Pak J Med Sci*. 2020;36(7):1693–1697. doi:10.12669/pjms.36.7.2625

Research and Reports in Urology**Dovepress**

Taylor & Francis Group

Publish your work in this journal

Research and Reports in Urology is an international, peer-reviewed, open access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of adult and pediatric urology in the clinic and laboratory including the following topics: Pathology, pathophysiology of urological disease; Investigation and treatment of urological disease; Pharmacology of drugs used for the treatment of urological disease. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/research-and-reports-in-urology-journal>