Author's Accepted Manuscript

Prediction of pediatric PCNL outcomes using contemporary scoring systems

Hussein abdelhameed Aldaqadosi , Hosni Khairy , Yousef Kotb , Hussein Aly Hussein , Hossam Shaker , Nikolaos Dikaos

PII: S0022-5347(17)57833-7
DOI: [10.1016/j.juro.2017.04.084](https://doi.org/10.1016/j.juro.2017.04.084)
Reference: JURO 14717

To appear in: *The Journal of Urology*
Accepted Date: 16 April 2017

Please cite this article as: Aldaqadosi Ha, Khairy H, Kotb Y, Hussein HA, Shaker H, Dikaos N, Prediction of pediatric PCNL outcomes using contemporary scoring systems, *The Journal of Urology*® (2017), doi: 10.1016/j.juro.2017.04.084.

DISCLAIMER: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our subscribers we are providing this early version of the article. The paper will be copy edited and typeset, and proof will be reviewed before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to The Journal pertain.

Embargo Policy

All article content is under embargo until uncorrected proof of the article becomes available online.

We will provide journalists and editors with full-text copies of the articles in question prior to the embargo date so that stories can be adequately researched and written. The standard embargo time is 12:01 AM ET on that date. Questions regarding embargo should be directed to jumedia@elsevier.com.

Prediction of pediatric PCNL outcomes using contemporary scoring systems

Hussein abdelhameed Aldaqadosi ¹, Hosni Khairy ², Yousef Kotb ³, Hussein Aly Hussein ², Hossam Shaker ¹, Nikolaos Dikaïos ⁴

1 Fayoum Faculty of Medicine, Egypt

2 Cairo Faculty of Medicine, Egypt

3 Ain Shams Faculty of Medicine, Egypt

4 Centre for Vision, Speech and Signal Processing, University of Surrey, UK

Corresponding Author:

Hussein Abdelhameed Aldaqadosi

Fayoum Faculty of Medicine,

Kimn Fares, 63514

Fayoum, Egypt

Telephone: 00201005168856

Fax number: 084 6302350

Email: aldaqados@yahoo.com

Hosni Khairy: dr_hsoni@yahoo.com

Yousef Kotb: yusufkotb@yahoo.com

Hussein Aly Hussein: drhusseinali2008@yahoo.com

Hossam Shaker: hossamshaker2008@yahoo.com

Nikolaos Dikaïos: n.dikaïos@surrey.ac.uk

Running head: pediatric PCNL

key words: Pediatric, PCNL, outcomes, complications, scoring systems

Purpose

To evaluate the applicability of contemporary percutaneous nephrolithotomy (PCNL) scoring systems in pediatric patients and to compare their predictive power for postoperative outcomes.

Materials and methods

The records of 125 pediatric patients who were diagnosed with renal calculi and managed with PCNL between March 2011 and April 2016 were retrospectively analyzed. The predictive scoring systems; The Guy's Stone Score (GSS), S.T.O.N.E. nephrolithometry and, Clinical Research Office of the Endourological Society (CROES) were calculated for all patients included in the study. Patient demographics, stone free rate (SFR), and complications were all reported and analyzed.

Results

In patients with residual stones (group I) vs those who were (group II) stone free the median (IQR) of GSS was 2 (2-3) and 2 (1-2), CROES nomogram score was 215 (210-235) and 257 (240-264), and S.T.O.N.E. nephrolithometry score was 8 (7-9) and 5 (5-6), respectively (each $p < 0.0001$). S.T.O.N.E. nephrolithometry score revealed the highest accuracy in predicting SFR. GSS was significantly correlated with complications but the CROES nomogram and S.T.O.N.E nephrolithometry were not significantly correlated with complications.

Conclusion

The scoring systems could be used in predicting PCNL success in pediatric setting. However, further studies are required to make modifications in the scoring systems in pediatrics. The main variables in the scoring systems as stone burden, tract length and case volume were measured using records from adult patients. Besides these variables, the relatively small pelvicalyceal system and the higher incidence of anatomic malformations in pediatrics could potentially affect PCNL outcomes.

Introduction

Pediatric urolithiasis is often complex and mostly attributed to underlying metabolic or urinary tract anatomical abnormalities. There is a considerable regional variation in the incidence of pediatric urolithiasis. It is relatively higher in developing countries than in the developed countries, with an incidence of 5–15% and 1–5%, respectively¹⁻³. Surgical management of renal calculi has evolved over the past few decades with the widespread use of extracorporeal shockwave lithotripsy (ESWL) and percutaneous nephrolithotomy (PCNL). Nowadays, these minimally invasive techniques are well established treatment modalities for pediatric nephrolithiasis with confirmed safety and efficacy^{4,5}. PCNL is commonly selected as a first-line treatment for pediatric patients with renal calculi larger than 2 cm⁶. In 1985 Woodside and associates⁷ described the first pediatric PCNL, and they used adult instrument. Since then instruments miniaturization and many technologic modifications have been done on pediatric PCNL. The Mini-PCNL technique has been developed to reduce complications while not affecting stone clearance rate⁸.

Several grading systems have been evolved to anticipate the procedure outcomes and potential risk preoperatively. Validated and reliable predictive scoring systems can powerfully assist in clinical decision taking and patient counseling. The most popular predictive scoring systems for PCNL in literature are the Guy's Stone Score (GSS), the Clinical Research Office of the Endourological Society (CROES) nomogram, and S.T.O.N.E. (stone size, tract length, obstruction, number of involved calices, and essence/ stone density) nephrolithometry (supplementary table). The GSS consists of four grades according to stone burden, urinary tract anatomy, and staghorn status. The CROES nomogram is based on influential variables predicting SFR (stone burden, stone count, location, and presence of staghorn stones) in addition to prior treatment and case volume/year. Each factor had a score from 0 to 100 as delineated in the nomogram. The sum of the different scores was used to predict the SFR. S.T.O.N.E. nephrolithometry components are stone size, PCNL tract length, presence of obstruction, number of involved calyces, and stone essence (density) measured from preoperative CT⁹⁻¹¹. All these scoring systems are highly accurate and reliable in predicting PCNL outcomes

¹²⁻¹⁴. They have been developed for adult patients. To our knowledge, there is a relative lack of studies evaluating these scoring systems in pediatrics. In a recent study Utangac et al ¹⁵ compared GSS and CROES scoring system in pediatric Mini-PCNL. Doulian et al ¹⁶ evaluated the S.T.O.N.E. nephrolithometry in pediatrics.

No previous study has compared the three scoring systems to predict the SFR in the same cohort of pediatric patients. The aim of our study was to compare the GSS, CROES and S.T.ON.E. nephrolithometry to predict pediatric PCNL outcomes.

Materials and methods

The records of 125 pediatric patients who were diagnosed with renal calculi and managed with PCNL between March 2011 and April 2016 were retrospectively analyzed. The study was conducted in three pediatric/adult; tertiary care university hospitals. We excluded patients with a history of open stone surgery, DJ stent, or nephrostomy tube in the ipsilateral kidney and those with no available preoperative computerized tomography (CT) scans. The patients were evaluated with laboratory tests (urinalysis, urine culture and sensitivity test, serum creatinine, blood urea, complete blood count, and coagulation profile) and imaging studies (urinary tract ultrasonography, plain radiography of kidneys, ureters, and bladder (KUB) and CT). Stone surface area was calculated according to the formula: $\text{length} \times \text{width} \times 3.14 \times 0.25$ ¹⁷. The demographic characteristics, age, gender, stone location and burden, operative, fluoroscopy, and hospitalization time, hemoglobin changes, number of access, stone-free rate, and complications were all reported and analyzed. We used the modified Clavien system to grade perioperative complications. Our study was approved by the local ethics committee.

Surgical Technique

All PCNL procedures were performed under general anesthesia. An ureteral catheter (5 F) was inserted under fluoroscopy in lithotomy position. PCNL access was accomplished in prone position after retrograde pyelography. The pelvicyceal system was punctured using 18-gauge needle under fluoroscopic control. The working tract was established using Amplatz dilators up to 20 F or 24 F according to patient age and degree of hydronephrosis. We used a 9.5 F Storz semi rigid ureteroscope or a 24 F Storz nephroscope without its own outer according to tract size. Stone disintegration

was carried out with a pneumatic lithotripter. Stone-free status was evaluated by the nephroscope and fluoroscopy. A nephrostomy tube was introduced in the tract according to operator decision based on degree of bleeding and residual stones. On the first postoperative day KUB was performed. The nephrostomy tube if inserted was removed on the second day. The patients were examined every week during the first month then every three months during the first year.

Measurements

The scoring systems; GSS, the CROES, and S.T.O.N.E. nephrolithometry were calculated for all patients included in the study according to preoperative CT images. The scores were calculated by two experienced assistant professors of urology as outlined in Thomas et al, Smith et al, and Okhunov et al studies⁹⁻¹¹.

Outcomes

The primary outcome of our study was SFR after PCNL. SFR was determined by CT scan 4 weeks after PCNL, and it was defined as the absence of any significant stone fragment >4 mm. The secondary outcome was assessment of PCNL related complications.

Statistical analysis of data:

The collected data was organized, tabulated and statistically analyzed using SPSS version 18 (SPSS Inc, USA). For quantitative data Mann Whitney U was used to analyze categorical variables. For qualitative data, the number and percent distribution was calculated, and chi square was used as a test of significance. For interpretation of results of tests of significance, significance was adopted at P value ≤ 0.05 . Individual linear discriminant analysis models were derived for each contemporary PCNL scoring system. The receiver operator characteristic area under the curve (ROC AUC) was calculated for each model. Model threshold probabilities for prediction of postoperative outcomes were based on the Youden index (J); where J is a function of sensitivity and specificity, and is defined as the maximum vertical distance between the ROC curve and the diagonal line¹⁸⁻²⁰. J occurs at the threshold probability for optimal diagnostic models classification ability. A second threshold was also evaluated, set at the specificity closer to 50% from the ROC analysis. The accuracy of each contemporary PCNL scoring system was determined by calibration plot, which

compares predicted probabilities to the observed fraction of positives. Decision curve analysis was used to illustrate the net benefits of the univariate prediction models when different threshold probabilities were considered.

Results

The study included 125 pediatric patients who underwent PCNL. Study cohort composed of 77 (61.6 %) males and 48 (38.4 %) females. Based on postoperative SFR we divided the patients into two groups; group I comprised 33 (26.4%) patients who were not stone free and group II consisted of 92 (73.6%) stone free patients. In our study, the stone free rate was (73.6%). This percentage increased to 96.8% with adjunctive SWL (22 patients) and ureteroscopy (7 patients). No significant differences were detected between the two groups in terms of age, sex, stone laterality, mean hemoglobin change, and mean hospital stay. Patient demographics and stone characteristics are listed in table 1 and figure 1. The Median (IQR) of stone burden was significantly higher in group I than in group II 400 (270-490) vs. 320 (292-380) respectively (p value < 0.002).

In group I and group II the median (IQR) of GSS was 2 (2-3) and 2 (1-2), CROES nomogram score was 215 (210-235) and 257 (240-264), and S.T.O.N.E. nephrolithometry score was 8 (7-9) and 5 (5-6), respectively (each $p < 0.0001$). (table1). Table 2 and the figure 2 show AUC for the three scoring system and the stone burden. S.T.O.N.E. nephrolithometry score revealed the highest accuracy in predicting SFR. The estimated AUC for S.T.O.N.E. nephrolithometry was 0.92 compared to 0.70, 0.78 and 0.68 for GSS, CROES score and stone burden. The derived sensitivity and specificity for each model at the Youden index and the set 50% specificity probability threshold is given in tables 3 and 4 respectively. On calibration plot (figure 3), the S.T.O.N.E. score was closer to the ideal predictor, indicating better accuracy. Decision-curve analysis (figure 4) demonstrated better prediction of postoperative outcomes based on the STONE score in comparison with other scores.

In our cohort, a total of 24 (19.2%) complications were reported postoperatively. According to the modified Clavien system the complications (table 5) were classified as following: grade I reported in 13 (10.4%) patients, 8 (6.4%) were

grade II, grade III developed in 2 (1.6%), and one (0.8%) was grade IV. According to table 6, GSS was significantly correlated with complications ($P= 0.017$) but the CROES nomogram S.T.O.N.E nephrolithometry were not significantly correlated with complications.

Discussion

The first PCNL in adults was described by Fernstrom and Johansson ²¹ in 1976, and the first PCNL in pediatrics was described by Woodside and colleagues ⁷ in 1985. Nowadays, there is an increasing experience in PCNL, miniaturization of instruments, and technology advancement. These factors led to the widespread of PCNL as an effective and safe option for the management of pediatric urolithiasis. Pediatric patients have a small renal reserve and quietly feeble renal parenchyma, therefore pediatric PCNL can be a life threatening procedure. Mini-PCNL using 14 French to 16 French working tract, is effective and safe for pediatric renal stones management ²².

The use of statistics and mathematics in modern medicine permits transformation of medical findings into a prediction score. With the prevalent application of PCNL several groups have attempted to develop scoring systems to anticipate SFR after PCNL. Predictive tools are needed for objective assessment of renal stones, optimize medical decision and assist patient counseling. Furthermore, they allow the standardization of the PCNL outcomes to compare different studies ²³. Nowadays, authors advise the use of prognostic tools that can predict the postoperative PCNL outcomes based on preoperative imaging modalities ^{9-12, 24-26}. Some of the most common prognostic tools are the GSS, CROES nomogram, and S.T.O.N.E. nephrolithometry scoring systems ⁹⁻¹¹ (supplementary table).

The predictive accuracy of GSS, CROES nomogram and S.T.O.N.E. nephrolithometry scoring systems has been compared by numerous published studies. These systems are statistically significant tools for predicting SFR following PCNL ^{12, 13, 24-27}. The three scoring systems have been externally validated in a number published of studies. While The GSS is simple, reproducible, and easy to implement in clinical practice, it does not incorporate important variables as stone density and calyceal involvement. These variables have important influence on SFR. Also, GSS have four grades limiting its ability to describe complex stone disease. The CROES

nomogram is complex and incorporates comprehensive data so that it is impractical in clinical setting. In addition, it lacks critical variables as hydronephrosis and stone density. These factors have been shown to affect SFR ²⁸.

Labadie et al ²⁴ studied the GSS, S.T.O.N.E., and CROES scoring systems and reported that they had similar efficacy in predicting PCNL outcomes. In 2015, Bozkurt and colleagues ²⁵ compared GSS and CROES scoring systems and concluded that both scores were comparable in predicting SFR after PCNL. In a cohort of 185 patients Noureldin and associates ²⁹ compared GSS and STONE scores. They concluded that both systems were significantly predictive of postoperative SFR.

When favoring the ideal scoring system, it should be reproducible, easily applicable and suitable for recording and comparison. S.T.O.N.E. nephrolithometry is the most comprising scoring system and practical in the clinical setting. Furthermore, it is the only scoring system developed based on CT images. The contemporary imaging modality routinely used in almost all patients with urolithiasis is CT, so that S.T.O.N.E. nephrolithometry is best suited for assessment of PCNL outcomes. S.T.O.N.E. nephrolithometry is easier than the CROES nomogram to implement in clinical practice and comprises more accurate risk stratification variables than GSS ²⁴.

The three scoring systems were developed and used in adults. To our knowledge, there is a relative lack of studies evaluating these scoring systems in pediatrics. S.T.O.N.E. nephrolithometry was first applied by Doulian et al¹⁶ in children with upper urinary tract stones managed by mini-PCNL. The authors concluded that (S.T.O.N.E. nephrolithometry is applicable in children with renal calculi and predict the procedure success rate, hospitalization, and even postoperative complications). They suggested modification of this scoring system to be more suitable for the pediatric renal anatomy. Goyal and colleagues ³⁰ studied factors associated with pediatric PCNL complications, among these factors was GSS. In a recent study, Utangac et al ¹⁵ reported that both GSS and CROES scoring system can predict SFR after pediatric mini-PCNL.

In the present work we compared GSS, CROES and S.T.O.N.E. scoring systems in a single cohort of pediatric patients who underwent PCNL. To our knowledge this study presents the first comparison of the three scoring systems in the same cohort of pediatric patients. We analyzed and compared the three scoring systems and in

addition to stone burden to determine the predictor system for pediatric PCNL success. S.T.O.N.E. nephrolithometry score revealed the highest accuracy in predicting SFR after pediatric PCNL. It had the highest AUC compared to GSS, CROES and stone burden (0.92 vs. 0.70, 0.78 and 0.68 respectively). GSS was significantly correlated with complications but the S.T.O.N.E. nephrolithometry and CROES nomogram were not significantly correlated with complications.

S.T.O.N.E. nephrolithotomy was developed based on data of adult PCNL. It has been validated among pediatric patients¹⁶. In our study we noticed that the tract length parameter in S.T.O.N.E. nephrolithotomy requires modification in pediatric patients, as all cases scored one point. In addition, stone burden and case volume are essential variables in the scoring systems. These factors were estimated using records from adult patients. Along with these factors, the age of the pediatric patients is a great challenge increasing the complexity of the PCNL in pediatrics. Although anatomic malformation is not a variable used by the scoring systems, pediatric patients have a higher incidence of anatomic malformations which could potentially affect PCNL outcomes. All of these parameters declare that the scoring systems require modification in pediatrics. More studies are necessary to support our conclusion and make modification.

The main limitations of this study are the retrospective form and multi-centered type of the study. In order to overcome these limitations we standardized data collection process and rigid outcomes definitions were used. The strengths to our study included; being the first comparison of the three scoring systems in the same cohort of pediatric patients and rigid outcomes definitions were used.

Conclusion

The scoring systems could be used in predicting PCNL success in pediatric setting. However, further studies are required to make modifications in the scoring systems in pediatrics. The main variables in the scoring systems as stone burden, tract length and case volume were measured using records from adult patients. Besides these variables, the relatively small pelvicalyceal system and the higher incidence of anatomic malformations in pediatrics could potentially affect PCNL outcomes.

References

1. Mahmud M and Zaidi Z. Percutaneous nephrolithotomy in children before school age: Experience of a Pakistani centre. *BJU Int* 2004; 94:1352–4
2. Shah AM, Kalmunkar S, Puneekar SV et al. Spectrum of paediatric urolithiasis in Western India. *Indian J Ped* 1991; 58: 543–9
3. Yoshida O and Okada Y. Epidemiology of urolithiasis in Japan: a chronological and geographical study. *Urol Int* 1990; 45: 104–11
4. Rizvi SAH, Naqvi SAA, Hussain Z et al. The Management of stone disease. *BJU Int* 2002; 89 (Suppl. 1): 62–8
5. Aldaqadossi HA, Kotb Y, Mohi K. Efficacy and Safety of Percutaneous Nephrolithotomy in Children with Previous Renal Stone Operations. *J Endourol.* 2015 Aug;29(8):878-82
6. Tekgül S, Dogan HS, Hoebeke P et al. Guidelines on Paediatric Urology. European Society for Paediatric Urology. European Association of Urology. 2016
7. Woodside JR, Stevens GF, Stark GL et al. Percutaneous stone removal in children. *J Urol* 1985; 134:1166–7.
8. Helal M, Black T, Lockhart J et al. The Hickman peelaway sheath: alternative for pediatric percutaneous nephrolithotomy. *J Endourol.* 1997; 11:171-2.
9. Thomas K, Smith NC, Hegarty N et al. The Guy's stone score—grading the complexity of percutaneous nephrolithotomy procedures. *Urology.* 2011; 78:277–81.
10. Smith A, Averch TD, Shahrour K et al. A nephrolithometric nomogram to predict treatment success of percutaneous nephrolithotomy. *J Urol.* 2013; 190:149–56
11. Okhunov Z, Friedlander JI, George AK et al. S.T.O.N.E. nephrolithometry: novel surgical classification system for kidney calculi. *Urology.* 2013; 81:1154–9
12. Vicentini FC, Marchini GS, Mazzucchi E et al. Utility of the Guy's stone score based on computed tomographic scan findings for predicting percutaneous nephrolithotomy outcomes. *Urology.* 2014; 83:1248–53
13. Ingimarsson JP, Dagrosa LM, Hyams ES et al. External validation of a preoperative renal stone grading system: reproducibility and inter-rater concordance of the Guy's stone score using preoperative computed tomography and rigorous postoperative stone-free criteria. *Urology.* 2014; 83:45–9
14. Akhavein A, Henriksen C, Syed J et al. Prediction of single procedure success rate using S.T.O.N.E. nephrolithometry surgical classification system with strict criteria for surgical outcome. *Urology.* 2015; 85:69–73
15. Utangac MM, Tepeler A, Daggulli M et al. Comparison of Scoring Systems in Pediatric Mini-Percutaneous Nephrolithotomy. [dx.doi.org/10.1016/j.urol.2016.03.038](https://doi.org/10.1016/j.urol.2016.03.038)

16. Doulian S, Hasimu S, Jun D et al. The application of S.T.O.N.E. nephrolithometry in pediatric patients with upper urinary tract calculi treated with mini-percutaneous nephrolithotomy. *Urolithiasis*. 2015; 43:363-7.
17. Tiselius HG and Andersson A. Stone burden in an average Swedish population of stone formers requiring active stone removal: how can the stone size be estimated in the clinical routine? *Eur Urol*. 2003;43:275-81
18. Youden WJ. Index for rating diagnostic tests; *Cancer*. 1950. Vol. 3. pp. 32–35.
19. Reiser B. Measuring the effectiveness of diagnostic markers in the presence of measurement error through the use of ROC curves. *Statistics in Medicine*. 2000;19:2115–2129.
20. Schisterman EF and Perkins NJ. Confidence Intervals for the Youden Index and Corresponding Optimal Cut-point. *Communications in Statistics: Simulations and Computations*. 2007;36:549–563.
21. Fernstrom I and Johansson B. Percutaneous pyelolithotomy. A new extraction technique. *Scand J Urol Nephrol* 1976; 10: 257–9.
22. Wah TM, Kidger L, Kennish S et al. MINI PCNL in a pediatric population. *Cardiovasc Intervent Radiol*. 2013; 36(1):249–54
23. Sfoungaristos S, Gofrit ON, Pode D et al. Percutaneous nephrolithotomy for staghorn stones: Which nomogram can better predict postoperative outcomes? *World J Urol*. 2016 Aug;34(8):1163-8
24. Labadie K, Okhunov Z, Akhavein A et al. Evaluation and comparison of urolithiasis scoring systems used in percutaneous kidney stone surgery. *J Urol*. 2015;193:154-9.
25. Bozkurt IH, Aydogdu O, Yonguc T et al. Comparison of guy and clinical research office of the endourological society nephrolithometry scoring systems for predicting stone-free status and complication rates after percutaneous nephrolithotomy: a single center study with 437 cases. *J Endourol*. 2015; 29:1006-10.
26. Sinha RK, Mukherjee S, Jindal T et al. Evaluation of stone-free rate using Guy's stone score and assessment of complications using modified Clavien grading system for percutaneous nephrolithotomy. *Urolithiasis*. 2015; 43:349-53.
27. Akhavein A, Henriksen C, Syed J et al. Prediction of single procedure success rate using S.T.O.N.E. nephrolithometry surgical classification system with strict criteria for surgical outcome. *Urology*. 2015; 85:69–73

28. Vernez SL, Okhunov Z, Motamedinia P et al. Nephrolithometric Scoring Systems to Predict Outcomes of Percutaneous Nephrolithotomy. *Rev Urol.* 2016; 18(1): 15–27.
29. Noureldin YA, Elkoushy MA, Andonian S. Which is better? Guy's versus S.T.O.N.E. nephrolithometry scoring systems in predicting stone-free status post-percutaneous nephrolithotomy. *World J Urol.* 2015 Nov; 33(11):1821-5
30. Goyal NK, Goel A, Sankhwar SN et al. A critical appraisal of complications of percutaneous nephrolithotomy in paediatric patients using adult instruments. *BJU Int.* 2014; 113:801-10.

	Group 1(Non stone-free)	Group 2 (Stone free)	P value
No. (%)	33	92	
Gender			
Male	19	58	0.579##
Age, median (IQR) (SE)	7 (7-7) 0.91	6 (6-8) 0.55	0.054#
BMI, median (IQR) (SE)	18 (17-18) 2.36	17 (17-18) 1.04	0.059#
Side			
Right	16	51	0.492##
Stone number (n), median (IQR) (SE)	2 (1.5-2) 0.21	1 (1-2) 0.11	<0.0001#*
Stone burden, median (IQR) (SE)	400 (270-490) 18.63	320 (292-380) 6.11	0.002#*
Stone location			0.024###*
Renal pelvis	9	53	
Lower pole	7	14	
Middle pole	4	3	
Upper pole	0	1	
Multiple locations	13	21	
Hydronephrosis	21	65	0.456##
Mean operative time (min), median (IQR) (SE)	95 (89-97) 6.01	87 (83-100) 3.29	0.118#
Mean fluoroscopy time (min), median (IQR) (SE)	4 (3.5-4.5) 0.42	3 (3-3) 0.22	<0.0001#*
Access number (n), median (IQR) (SE)	1 (1-1) 0.07	1 (1-1) 0.02	0.589#
Mean hemoglobin change (mg/dL), median (IQR) (SE)	1 (1-2) 0.09	1 (1-1) 0.03	0.218#
Mean hospital stay (days), median (IQR) (SE)	3 (3-4) 0.24	3 (3-3) 0.16	0.696#
Guy's score, median (IQR) (SE)	2 (2-3) 0.13	2 (1-2) 0.07	<0.0001#*
CROES score, median (IQR) (SE)	215 (210-235) 4.37	257 (240-264) 2.14	<0.0001#*
STONE score, median (IQR) (SE)	8 (7-9) 0.18	5 (5-6) 0.09	<0.0001#*

Mann-Whitney-U test ## Chi squared test

*statistically significant

Table 1: Patient demographic and clinical data

	AUC	Std. Error	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
Guy's score	0.70	0.06	0.60	0.81
STONE score	0.92	0.03	0.87	0.97
Stone burden	0.68	0.07	0.55	0.82
CROES score	0.78	0.05	0.68	0.88

Table 2: Univariate ROC analysis of contemporary PCNL scoring systems for prediction of postoperative outcomes

	Cut-off point	Sensitivity	Specificity	PPV	NPV	Total accuracy
Guy's score	1.5	81.8	43.5	34.2	87.0	53.6
STONE score	5.5	100.0	51.1	42.3	100.0	64.0
Stone burden	310	72.7	44.6	32.0	82.0	52.0
CROES score	257	81.8	51.1	37.5	88.7	59.2

Table 3: Performance of contemporary PCNL scoring systems at set 50% specificity derived probability threshold.

	Cut-off point	Sensitivity	Specificity	PPV	NPV	Total accuracy
Guy's score	2.5	93.5	39.4	81.1	68.5	79.2
STONE score	6.5	87.9	76.1	56.9	94.6	79.2
Stone burden	390	60.6	79.4	51.3	84.9	74.4
CROES score	225	75.8	80.4	58.1	90.2	79.2

Table 4: Performance of contemporary PCNL scoring systems at Youdens index derived probability threshold.

	Patients no.	Grade
Urinary leakage managed managed by watchful waiting	7	I
Bleeding managed by nephrostomy tube clamping	6	I
Infections managed by antibiotics	5	II
Bleeding managed by blood transfusion	3	II
Urine leakage managed by DJ stent	2	III
Intensive care unit admission due to septicemia	1	IV

Table 5: Complications of Percutaneous Nephrolithotomy according to Clavien Classification

	Patients without postop. complications (NO. =101)	Patients with postop. complications (NO. =24)	P value
Guy's score	2 (1-2)	2 (1-3)	0.017*
CROES score	254 (215-264)	256 (214.25-263)	0.890
STONE score	6 (5-7)	6 (5-7.75)	0.437
Stone burden	320 (280-395)	340 (310-407.5)	0.182

Mann-Whitney –U test

*statistically significant

Table 6: Complications rate according to scoring systems and stone burden

Figure 1: A graphical presentation of age distribution

Figure 2: ROC curves of contemporary PCNL scoring systems

Figure 3: Calibration plots of the average predicted value vs observed positive fractions rate for each contemporary PCNL scoring system. Diagonal line represents the ideal predictor.

Figure 4: Decision curve analysis for each contemporary PCNL scoring system. The y axis measures net benefit, calculated by summing the true positive findings and subtracting the false positive findings.

1. PCNL: Percutaneous nephrolithotomy
2. GSS: The Guy's Stone Score
3. S.T.O.N.E. nephrolithometry: Stone size, tract length, obstruction, number of involved calices, and essence/stone density (nephrolithometry)
4. CROES: Clinical Research Office of the Endourological Society
5. SFR: Stone free rate
6. ROC curve: Receiver Operating Characteristic curve
7. CT scans: Computerized tomography scans.

Scoring system	Categorization
Guy's Stone Score	<p>Grade I: A solitary stone in the mid/lower pole, or renal pelvis with simple anatomy</p> <p>Grade II: A solitary stone in the upper pole with simple anatomy, multiple stones in a patient with simple anatomy, or any solitary stone in a patient with abnormal anatomy</p> <p>Grade III: Multiple stones in a patient with abnormal anatomy, stones in a calyceal diverticulum, or partial staghorn calculus</p> <p>Grade IV: Staghorn calculus or any stone in a patient with spina bifida or spinal injury</p>
Nephrolithometric Nomogram	<p>A: Stone Burden (mm^2)</p> <p>B: Calyceal location– position in renal pelvis or multiple calyces involved, including staghorn calculi</p> <p>C: Stone count– single or multiple</p> <p>D: Case volume / year</p> <p>E: Prior treatment</p>
S.T.O.N.E. Nephrolithometry	<p>S = stone size (mm^2) 1: 0-399 2: 400-799 3: 800-1599 4: >1600</p> <p>T = tract length 1: ≤ 100 mm 2: > 100 mm</p> <p>O = Obstruction 1: no or mild dilatation 2: moderate to severe dilatation</p> <p>N = Number of involved calyces 1: 1 calyx involved 2: 2-3 calyces involved</p> <p>3: full staghorn calculus</p> <p>E = Essence (stone density) 1: < 950 HU 2: > 950 HU</p>

Summary of scoring systems analyzed in our study