Arab Journal of Urology (2017) 15, 7–16

Arab Journal of Urology
(Official Journal of the Arab Association of Urology)

www.sciencedirect.com

STONES/ENDOUROLOGY
ORIGINAL ARTICLE

How practical is the application of percutaneous nephrolithotomy scoring systems? Prospective study comparing Guy's Stone Score, S.T.O.N.E. score and the Clinical Research Office of the Endourological Society (CROES) nomogram

Anurag Singla [a], Nikhil Khattar [a,*], Rishi Nayyar [a], Shibani Mehra [b], Hemant Goel [a], Rajeev Sood [a]

[a] *Department of Urology, Postgraduate Institute of Medical Education and Research (PGIMER) and Dr. Ram Manohar Lohia Hospital, Baba Kharak Singh Marg, New Delhi, Delhi, India*
[b] *Department of Radio-diagnosis, Postgraduate Institute of Medical Education and Research (PGIMER) and Dr. Ram Manohar Lohia Hospital, Baba Kharak Singh Marg, New Delhi, Delhi, India*

Received 5 September 2016, Received in revised form 13 November 2016, Accepted 27 November 2016
Available online 12 January 2017

KEYWORDS
Clinical Research Office of the Endourological Society (CROES);
Percutaneous nephrolithotomy (PCNL);
Renal stone;
Guy's Stone Score;
S.T.O.N.E. score

Abstract ***Objective:*** To prospectively compare the Guy's Stone Score (GSS), S.T. O.N.E. [stone size (S), tract length (T), obstruction (O), number of involved calices (N), and essence or stone density (E)] score and the Clinical Research Office of the Endourological Society (CROES) nephrolithometric nomogram to predict percutaneous nephrolithotomy (PCNL) success rate and assess the correlation with perioperative complications.

Patients and methods: We prospectively evaluated all consecutive PCNL patients at our institute between 1 November 2013 and 31 May 2015. The above scoring systems were applied to preoperative non-contrast computed tomography and the practical difficulties in such applications were noted. Perioperative complications and the stone-free rate (SFR) were also recorded. Receiver operating characteristic curves were drawn and the areas under curves were compared and appropriate statistical analysis done.

* Corresponding author at: Department of Urology, PGIMER and Dr. Ram Manohar Lohia Hospital, Baba Kharak Singh Marg, New Delhi, Delhi 110001, India. Fax: +91 1123360067.
E-mail address: drkhattar@gmail.com (N. Khattar).

Peer review under responsibility of Arab Association of Urology.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.aju.2016.11.005

Results: In all, 48 renal units were included in the study. The overall SFR was 62.2%. The presence of staghorn stones (β = 27.285, 95% confidence interval 1.19–625.35; P = 0.039) was the only significant variable associated with the residual stones on multivariate analysis. Stone-free patients had significantly lower median GSS (2 vs 4) and S.T.O.N.E. scores (6 vs 10) and higher median CROES scores (83% vs 63%) (all $P < 0.001$) compared to residual-stone patients. All scoring systems were significantly associated with SFR (all $P < 0.001$). There was no significant difference in the areas under curves of the scoring systems (0.858, 0.923, and 0.931, respectively). Furthermore, all scoring systems had weak correlations with Clavien–Dindo classified complications (r = 0.29, P = 0.045; r = 0.40, P = 0.005 and r = –0.295, P = 0.04, respectively). We found no standardisation for the measurement of stone dimensions, tract length, Hounsfield units, and staghorn definition.

Conclusions: All scoring systems equally predicted SFR and had a weak correlation with Clavien–Dindo complications. Standardisation is needed for the variables in which they have been found deficient.

ABBREVIATIONS

ACS, acute angle, complicated calyx and stone size;
AUC, area under curve;
BMI, body mass index;
CCI, Charlson Comorbidity Index;
CROES, Clinical Research Office of the Endourological Society;
3D, three-dimensional;
GSS, Guy's Stone Score;
HU, Hounsfield unit;
IQR, interquartile range;
KUB, plain abdominal radiograph of the kidneys, ureters and bladder;
NCCT, non-contrast CT;
PCNL, percutaneous nephrolithotomy;
ROC, receiver operating characteristic;
SFR, stone-free rate;
SFS, stone-free status;
S.O.N., stone size, obstruction and number of involved calyces;
SPSS, Statistical Package for the Social Sciences;
SSD, skin-to-stone distance;
S.T.O.N.E., stone size (S), tract length (T), obstruction (O), number of involved calices (N), and essence or stone density (E);
SWL, shockwave lithotripsy;
US, ultrasonography

Introduction

Urinary stone disease is a prevalent problem throughout the world, with an incidence of 5–10% in the general population [1] and of which 15–20% of patients with renal stones require invasive intervention [2]. The goal of any such intervention is to achieve maximum stone clearance with minimum morbidity. Among the several treatment options, percutaneous nephrolithotomy (PCNL) has the highest stone clearance rates [3]. It is now the treatment of choice for large and complex renal stones, including staghorn stones [3]. But as with any surgery, PCNL too is not 100% successful and complication free despite its minimally invasive nature. The success of PCNL is defined by a stone-free rate (SFR), which has been reported to vary from 56% to 98% [4–8].

Attempts have been made by various authors to predict SFRs preoperatively by using various clinico-radiological parameters and designed various scores and tools but as such, currently, no single model has been standardised as the 'gold standard'. Further, to date, there has been no prospective study to directly compare these tools in the same patient cohort and to assess practical application of these scoring systems in various clinical scenarios. Thus, we prospectively applied three of such models: Guy's Stone Score (GSS) [9], S.T.O.N.E. [stone size (S), tract length (T), obstruction (O), number of involved calices (N), and essence or stone density (E)] nephrolithometry score [10], and the Clinical Research Office of the Endourological Society (CROES) nephrolithometric nomogram [11], on preoperative non-contrast CT (NCCT) and compared them in the same cohort to determine which was most predictive of the SFR and was easiest to use clinically. We also assessed if they could predict perioperative complications too during PCNL. We present the various difficulties and discrepancies encountered in the calculation and the application of these scoring systems.

Patients and methods

After obtaining Institutional Review Board approval and patient's informed consent, we prospectively included all consecutive patients with renal stones in our study who underwent PCNL between 1 November 2013 and 31 May 2015. Only the patients who underwent bilateral concomitant PCNL and patients with bleeding diathesis/uncorrected coagulopathy were excluded from the study.

Preoperative NCCT was performed and the GSS, S.T.O.N.E. score and CROES scores were calculated for each patient as described by Thomas et al. [9], Okhunov et al. [10], and Smith et al. [11], respectively. All difficulties in reaching a conclusive value for a scoring system were noted together with the ways to circumvent such difficulties.

Demographic, clinical, perioperative, and follow-up data were collected in a prospective fashion for all patients and included: age, sex, body mass index (BMI), surgical and medical history [Charlson Comorbidity Index (CCI) [12]], renal anomalies, operative time, number and location of punctures and dilated tracts, complications (modified Clavien-Dindo [13,14]), and other necessary details that were required for the calculation of the scores.

All our patients underwent PCNL in a prone position using a rigid 26-F nephroscope (Karl-Storz) and combined pneumatic-ultrasonic Lithoclast (Swiss-Lithoclast-Master) as per the standard protocol. A 24-F nephrostomy tube was placed at the end of the procedure in most of the cases. A plain abdominal radiograph of the kidneys, ureters and bladder (KUB) was taken for all patients on postoperative day 1 and if significant residual fragments were seen, a second-look PCNL through the same or a new tract was performed after 72 h.

The SFR (success rate) was defined as the absence of residual stone or the presence of asymptomatic clinically insignificant residual fragment of $\leqslant$4 mm on NCCT at 3 months postoperatively. Wherever the application of any variable of a scoring system was not straightforward, the solution was discussed among the authors and applied to all the cases.

The Statistical Package for the Social Sciences (SPSS®) 20 was used for statistical analysis. The test of normality was done using the Kolmogorov–Smirnov and Shapiro–Wilk tests. Continuous variables were presented as means (SDs) in cases of normal distribution and compared using the independent Student's *t*-test, while as median and interquartile range (IQR) in cases of skewed distributions and compared using the Mann–Whitney *U*-test. Categorical variables were presented as numbers with percentages and were compared using chi-square or Fisher's exact tests. Correlations between continuous variables were assessed using Pearson's/Spearman rank correlation coefficients. Logistic regression analysis was performed to correct for any possible confounders and assess the size effect of different variables on stone-free status (SFS). Statistical significance was considered for two-tailed *P* values of < 0.05. Receiver operating characteristic (ROC) curves were drawn for each scoring system and the area under curve (AUC) was calculated for each. The AUCs were compared using the online calculator of significance of difference between the area under two independent ROC curves from the following website: http://vassarstats.net/roc_comp.html, accessed on 22 August 2015. Assuming the incidence of SFR as 77% (based on literature review), and probability of type 1 error as 0.05, the power of this study with 28 patients in the stone-free group and 17 patients in the residual-stone group was calculated as 96.9%.

Results

In all, 48 patients underwent 50 PCNLs during our study period. One patient who underwent bilateral concomitant PCNL was excluded but the other patient who underwent PCNL on bilateral sides sequentially was included in the study. Thus, 48 renal units in 47 patients were treated with PCNL and were included in the present study.

Table 1 lists the descriptive details of all patients with stone characteristics and perioperative data. The mean patient age was 41.9 years and the mean BMI was 24 kg/m^2. Based on the CCI score distribution, nine patients (18.75%) had a CCI score of 1 and six (12.5%) had a CCI score of ⩾2. Treatment of a prior ipsilateral renal stone was present in 18.8% patients. Renal anomalies were present in 11.4% patients (horseshoe kidney was present in one and PUJ obstruction and bifid moieties were present in two patients each). There were staghorn stones in 29.2% of the patients. The median (IQR) GSS, S.T.O.N.E. score and predicted success rate using the CROES nomogram were 2 (1–4), 7 (6–9), and 75 (65.5–89)%, respectively.

SFS

Results were evaluable in 45 renal units as postoperative CT was not available for the rest (Table 1). The stone-free and residual-stone patients had a similar distribution for age, sex, BMI, and stone laterality. Overall, the SFR was 62.2%. The median stone burden in the stone-free and residual-stone patients was 283 and 730.5 mm^2, respectively (P < 0.001). Hydronephrosis was seen in 25% of the stone-free and 70% of the residual-stone patients (P = 0.003). Among the potential variables, tract length, stone density, and stone number were not associated with the presence of residual stones. However, increasing stone burden (P < 0.001), hydronephrosis (P = 0.003), staghorn stones (P < 0.001), increasing number of calyces involved (P < 0.001), and stones at multiple sites (P = 0.013) were significantly associated with the presence of residual stones. In the multivariate logistic model, only the presence of staghorn stones remained significantly associated with the likelihood of residual stones. The number of renal punctures and dilated tracts, and the operative time were also significantly different between the two groups.

When compared with residual fragments, stone-free patients had significantly lower median GSS (2 vs 4) and S.T.O.N.E. scores (6 vs 10) and higher median predicted success rates on the CROES nomogram (83% vs 63%) (all P < 0.001). All the scoring systems were significantly associated with the SFS (all P < 0.001). On comparing the AUCs, all were found to have similar accuracy in predicting SFS (Fig. 1).

Perioperative complications were noted in all 48 renal units and classified as per the Clavien–Dindo system

Table 1 The patients' characteristics, stone characteristics and perioperative data.

Variable	Total (n = 48)	Stone free (n = 28)	Residual stone (n = 17)	P
Age, years, mean (SD; range)	41.9 (14.2; 15–73)	39.3 (15.6)	44.4 (11.3)	0.246
Sex, n (%)				0.789
Male	30 (62.5)	17 (60.7)	11 (39.3)	
Female	18 (37.5)	11 (64.7)	6 (35.3)	
BMI, kg/m^2, mean (SD; range)	24 (5.4; 16–44)	24.2 (6.0)	23.6 (4.6)	0.73
Stone laterality, n (%)				0.299
Right	25 (52.1)	16 (69.6)	7 (30.4)	
Left	23 (47.9)	12 (54.5)	10 (45.5)	
Stone number, n (%)				0.13
Single	27 (56.3)	18 (72)	7 (28)	
Multiple	21 (43.8)	10 (50)	10 (50)	
Stone burden, mm^2, median (IQR; range)	319.1 (249.7–598.9; 78–1633)	283 (181–348.5)	730.5 (365–1010.7)	< 0.001
Tract length, mm, mean (SD; range)	90 (15.8; 61–140)	90 (17.9)	90.9 (13.4)	0.858
Obstruction, n (%)				0.003
None/mild	28 (58.3)	21 (80.8)	5 (19.2)	
Moderate/severe	20 (41.7)	7 (36.8)	12 (63.2)	
Number of calyces involved, n (%)				< 0.001
1–2	33 (68.8)	27 (84.4)	5 (15.6)	
⩾3	15 (31.2)	1 (7.7)	12 (92.3)	
Essence, HU, mean (SD; range)	1023.9 (245; 420–1750)	1023.9 (271.2)	1015.7 (211.9)	0.916
Stone location, n (%)				0.005
Upper/middle/lower pole	9 (18.7)	7 (87.5)	1 (12.5)	
Pelvis	8 (16.7)	8 (100)	0	
Multiple	31 (64.6)	13 (44.8)	16 (55.2)	
Presence of staghorn stone, n (%)				< 0.001
Yes	14 (29.2)	1 (8.3)	11 (91.7)	
No	34 (70.8)	27 (81.8)	6 (18.2)	
GSS, median (IQR; range)	2 (1–4; 1–4)	2 (1–2)	4 (2.5–4)	< 0.001
S.T.O.N.E. score, median (IQR; range)	7 (6–9; 5–11)	6 (6–7)	10 (8–11)	< 0.001
Predicted success rate on CROES nomogram,%, median (IQR; range)	75 (65.5–89; 33–90)	83 (75–90)	63 (53.5–69.5)	< 0.001
Re-look PCNL, n (%)	7 (14.6)	2 (28.6)	5 (71.4)	N.A.
Operative time, min, median (IQR; range)	90 (60–120; 35–270)	72.5 (60–90)	120 (90–142.5)	< 0.001

[13,14]. There were Grade I complications in six patients, Grade II complications in eight patients and Grade IIIa complications in five patients. All the models were found to have a weak correlation with the Clavien-Dindo grade of complications (GSS $r = 0.29$, $P = 0.045$; S.T.O.N.E score, $r = 0.40$, $P = 0.005$; CROES nomogram $r = -0.295$, $P = 0.04$).

Discussion

Many factors have been identified to predict clearance rates after PCNL such as stone burden, composition, number and location, HU, BMI, skin-to-stone distance (SSD), abnormal renal anatomy etc. However, a significant predictor alone is not a predictive tool in itself and thus various models have been developed by various authors combining various such parameters [9–11,15–18]. A few of such models include the GSS [9], Staghorn Morphometry [16], Seoul National University Renal Stone Complexity Scoring system [15], S.T.O.N.E. Nephrolithometry score [10], CROES Nephrolithometeric Nomogram [11], and recently developed acute angle, complicated calyx and stone size (ACS) score [18]. A logistic regression model was also developed by Zhu et al. [19] for minimally invasive PCNL but it might not apply to standard PCNL. Many studies have tried to validate some of these scoring systems [5,20–26] but most have been retrospective studies and the use of different sizes, timing and variable imaging methods for the evaluation of residual fragments after PCNL precludes the standardisation of such scoring systems. There have been studies comparing these tools as well [5,26], but

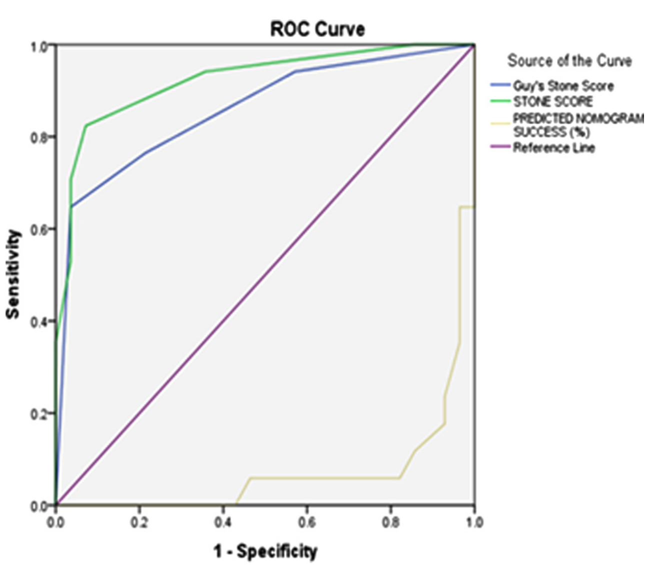

Fig. 1 ROC curves for the scoring systems: the AUCs for GSS, S.T.O.N.E. score and CROES nomogram were 0.858, 0.923 and 0.931 (1.00–0.69), respectively. Increasing values of GSS and S.T.O.N.E. score predict poorer chances whereas increasing scores on CROES nomogram predict better chances of being stone free.

these studies were limited by their retrospective design (Table 2). To date, there has been no prospective study directly comparing these scoring systems. The present study is the first to directly compare three of such tools in a prospective fashion in the same patient cohort.

Age, sex, BMI and stone laterality were comparable in both the stone-free and residual-stone groups in our present study and were comparable with the findings of other authors [5,10]. We found all three scoring systems to be significantly associated with SFS (all $P < 0.001$). These results are consistent with other studies as well [5,9–11,20–22,25,26].

Thomas et al. [9] originally described the GSS on variable imaging modalities and subsequently it has been evaluated by Ingimarsson et al. [22] and Vicentini et al. [21] on preoperative CT showing that the GSS is significantly associated with SFS ($P = 0.01$, $P = 0.03$, and $P < 0.001$, respectively). For the S.T.O.N.E. scoring system, it was originally proposed by Okhunov et al. [10], where they found stone-free patients to have significantly lower scores (6.8 vs 9.7, $P = 0.002$) and also observed a significant correlation of the S.T.O.N.E. score with the postoperative SFS. Both the above findings were consistent with our present study (median score 6 vs 10, $P < 0.001$). The CROES nomogram was developed by Smith et al. [11], which consisted of six variables of which stone burden was the best predictor of the SFR. In our present study, presence of staghorn stones remained the only significant factor for residual stones in the multivariate model.

With respect to comparative studies, Noureldin et al. [26] in their retrospective study compared the GSS and S.T.O.N.E. score and found significantly lower scores (2 vs 2.7, $P < 0.001$; and 7.4 vs 8.3, $P = 0.004$; respectively) in stone-free patients. Both these scoring systems had comparable accuracies in predicting SFS similar to our present study. In another retrospective study by Labadie et al. [5], in which they calculated the GSS, S.T.O.N.E. and CROES nomogram scores based on preoperative CT images, the mean scores were 2.2 vs 2.7, 8.3 vs 9.5, and 222 vs 187, respectively (all $P < 0.001$), consistent with the findings of the present study (2 vs 4, 6 vs 10, and 83% vs 63%, respectively, all $P < 0.001$). All the three scoring models were significantly associated with the SFS ($P = 0.02$, $P = 0.004$, $P < 0.001$, respectively) and had equal accuracies in determining the SFS after PCNL similar to the present study (all $P < 0.001$).

We also found all three scoring systems were correlated with Clavien–Dindo complications grade. This is consistent with Mandal et al. [20] and Vicentini et al. [21] who showed significant associations between the GSS and post-PCNL complications; however, the correlation in our present study was weak and might not be of clinical importance. This might be the effect of the small sample size of our study population. Also, Tho-

Table 2 PCNL scoring systems and their validation and comparative studies of the scoring systems.

PCNL scoring system and reference	No. of patients	Study design	Preoperative imaging method	Postoperative imaging used	Timing of postoperative imaging, weeks	Threshold size of residual fragments, mm	Comments
GSS							
Thomas et al. (2011) [9]	Validated in 100	Prospective	Variable (X-ray, CT, IVU)	X-ray (NCCT/US for radiolucent)	6	<4	Variable imaging
Validation studies of GSS							
Mandal et al. (2012) [20]	200	Prospective	Variable (X-ray, IVU, US, NCCT)	Variable	4	<4	Variable imaging
Ingimarsson et al. (2014) [22]	166	Retrospective	NCCT	X-ray/NCCT	Day 1	<2 on CT and 0 on CT/X-ray	Variable definition and SFR imaging
Vicentini et al. (2014) [21]	155	Partial retrospective	NCCT	NCCT	Day 1 (if ancillary procedure, then on final CT)	<4	All supine PCNL
S.T.O.N.E. score							
Okhunov et al. (2013) [10]	117	Prospective validation	NCCT	Intraoperative nephroscopy or US at 12 weeks	Intraoperative nephroscopy or US at 12 weeks		Vague definition and variable timing of SFR
Validation studies of S.T.O.N.E. score							
Okhunov et al. (2013) [24]	58	Retrospective	NCCT				Checked inter-observer reliability in application of score
Akhavein et al. (2015) [25]	122	Retrospective	NCCT	NCCT	$\leqslant$2 (mostly day 1)	<2	
CROES nomogram							
Smith et al. (2013) [11]	2806 Multicentric (96)	Prospective		X-ray (US/CT patients excluded from the study)		<4	Low sensitivity of X-ray
Comparative studies							
Labadie et al. (2015) [5] (GSS vs S.T.O.N.E. vs CROES)	246 (3 institutes)	Retrospective	NCCT	NCCT	At discharge or at $\leqslant$12	<2	Variable timing
Noureldin et al. (2015) [26] (GSS vs S.T.O.N.E.)	185	Retrospective		X-ray (NCCT for radiolucent)	12	<4	Variable imaging
Present study (GSS vs S.T.O.N.E. vs CROES)	50	Prospective	NCCT	NCCT	12	<4	

US, ultrasonography.

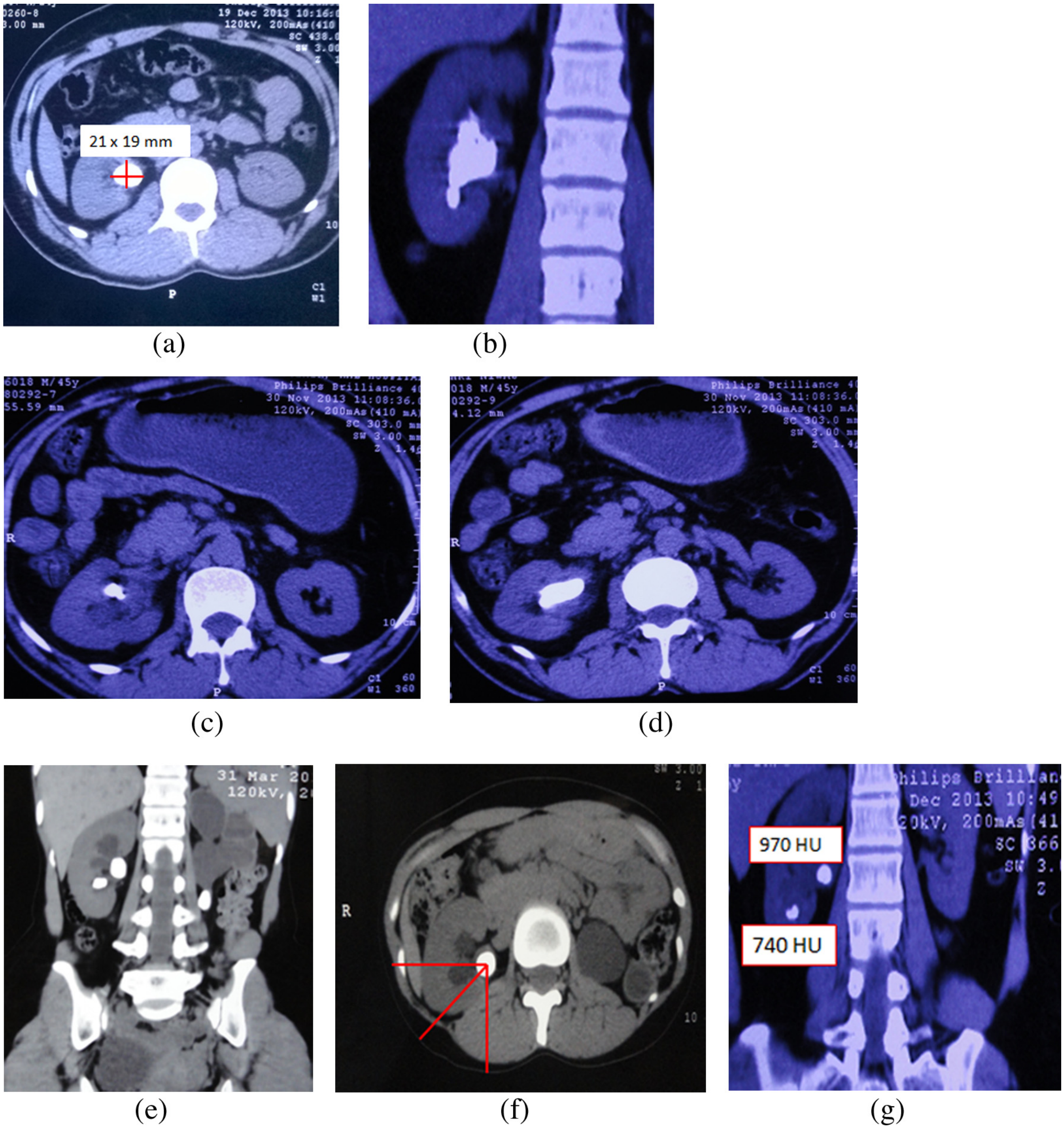

Fig. 2 (a) Dimensions of solitary oval stone reliably measured. (b) Irregular shape staghorn calculus on a coronal section of CT scan. (c) and (d) Same stone in two different axial cuts. Measurement of dimensions of this stone on axial images will produce incorrect stone size. (e) Coronal section of CT scan of another patient depicting two stones: one in pelvis and the other in the inferior calyx. (f) Calculation of tract length/SSD: the stone farthest from the skin (pelvic stone in this case) chosen from the above coronal section. On axial section, centre of the stone marked and three lines drawn at 0, 45 and 90° posteriorly intersecting the skin. These three distances measured and mean value taken. (g) Coronal section of CT scan showing two stones with different HU. Higher HU value was taken for the calculation in view of higher score (more difficult) given for >950 HU in the S.T.O.N.E. scoring system.

mas et al. [9], Okhunov et al. [10] and Noureldin et al. [26] did not find any correlation of the GSS and S.T. O.N.E. score with complications.

Problems faced in the practical application of the scoring systems and how we found our way out!

Calculating stone burden

Nowhere in either of the Okhunov et al. [10,24] studies or in the studies of the authors who validated the S.T. O.N.E. score [5,25,26] is there a description of the formula for stone burden calculation, so they might or might not have used the same equation as Okhunov et al., thus we do not know for sure if the model was actually reproduced in the same way as Okhunov et al. built it. We measured stone burden using the formula: maximum length × maximum width × π × 0.25, on axial cuts, as the same equation has been used by Smith et al. [11] in the CROES nomogram. Both these dimensions could be reliably determined for a smooth ovoid/rounded stone but not large/complex staghorn stones. For example in Fig. 2a the dimensions of a solitary oval stone can be reliably measured (Burden = $\pi \times 0.25 \times 21\ mm \times 19\ mm = 313.6\ mm^2$). Fig. 2b shows an irregular shaped staghorn stone on a coronal CT section. Fig. 2c and d shows the same stone in two different axial cuts calculating dimensions on which will produce an incorrect stone size. Further, the stone burden is best measured as volume on three-dimensional (3D)-reconstructed CT rather surface area especially in large/asymmetric/staghorn stones [27,28]. This is because larger/staghorn stones tend to conform to the pelvicalyceal system and develop protrusions/excrescences that result in irregular, asymmetric, and less geometric configurations. But 3D-CT is costly, complicated and not available in all urological clinics and probably is the reason most models have used surface area and not volume. Another inherent disadvantage is the measurement of exact dimensions of each and every stone especially in multiple secondary stones, which are so closely stacked to each other that making an exact count and dimension measurement is unfeasible.

Tract length and grade of hydronephrosis

Although the term 'tract length' has been used, it cannot be known preoperatively; as it will depend on the calyx punctured and thus can only be ascertained intraoperatively. Measurement is not difficult for a solitary stone but for multiple stones, which stone is to be chosen to measure the SSD presented as a dilemma in few cases. Thus, in cases of multiple stones, we chose the farthest stone from the skin in coronal sections and measured its distance on axial cuts, assuming it to be the most difficult target. For example, Fig. 2e shows a coronal section of a NCCT depicting two stones: one in the pelvis and the second in inferior calyx. The stone farthest from the skin (pelvic stone in this case) chosen from the above coronal section and centre of the stone marked on the axial section and three lines drawn at 0, 45 and 90° posteriorly intersecting the skin. These three distances were measured in millimetres and the mean value taken as the SSD (Fig. 2f).

The grades of hydronephrosis are subjective, non-quantitative and not clearly demarcated especially in mild vs moderate. For example, in Fig. 2f, one may grade it 'mild', while another may grade it 'moderate' hydronephrosis thus adding inaccuracy to the entire predictive model.

Dilemma of pelvic stones

No score has been assigned to a pelvic location in the S.T.O.N.E. scoring model [10]. Thus, presuming the pelvis as one calyx and approaching it with the same difficulty level, we assigned the score '1'. But this is not in concordance with the literature, where pelvic stones have better clearance rates than calyceal stones [29]. Strangely, authors who have used the S.T.O.N.E. score have not commented on this issue to date [10,24–26].

Essence: is it essential?

Regarding stone density/essence, Okhunov et al. [10] have assigned a higher score to >950 HU stones suggesting poorer clearance, which is not in concordance with published data where stones with higher HU values had either equal or greater clearance rates compared to lower HU stones [29,30]. As with stone burden, HU of large/staghorn calculi is also not very accurately calculated owing to their lamellated structures, which leads to varying densities from centre to periphery. Also, nothing has been suggested about which HU to be taken where there are different HU of multiple stones. Fig. 2g depicts a coronal section of NCCT showing two stones with different HU. We took the higher HU value for the calculation in view of the higher score (more difficult) given for >950 HU in the S.T.O.N.E. scoring system. We found 'essence' to be insignificant in predicting SFR.

Staghorn: poorly defined

Even the term 'staghorn' is plagued by an unclear morphology [3]. Rassweiler et al. [31] and Di Silverio et al. [32] classified staghorn stones into: 'borderline', stones occupying pelvis and one calyx; 'partial', pelvis extending into two calyces; 'complete', pelvis extending into all major calyces, filling $\geqslant 80\%$ of the pelvicalyceal system; and 'gigantic', whole pelvicalyceal system have stones and dilation occurs. But, this classification is not based on any specific volume criteria [3], as there is considerable overlap of stone burden between partial and complete staghorn stones. Other definitions of staghorn stone commonly used in literature are: stone located in the pelvis and in at least two of the calyces [33]; partial staghorn, pelvis and at least one infundibulum and calyx; and complete staghorn, occupying more than one portion of the collecting system [22]. We used the latter definition in our study.

Variable imaging modality and timing

Lastly, the imaging methods used (X-ray/ultrasonography/CT), sizes of residual fragments (0, 2, and 4 mm), and timing of imaging (day 1, and 4, 6, and 12 weeks) to document SFR all varied in different studies (Table 2).

Stone size, obstruction and number of involved calyces (S.O.N.) score

During our present analysis, we found that among the five variables of the S.T.O.N.E. score, tract length and essence were not significant predictors of stone clearance. Tract length/SSD may be important only in morbidly obese patients where longer instruments are

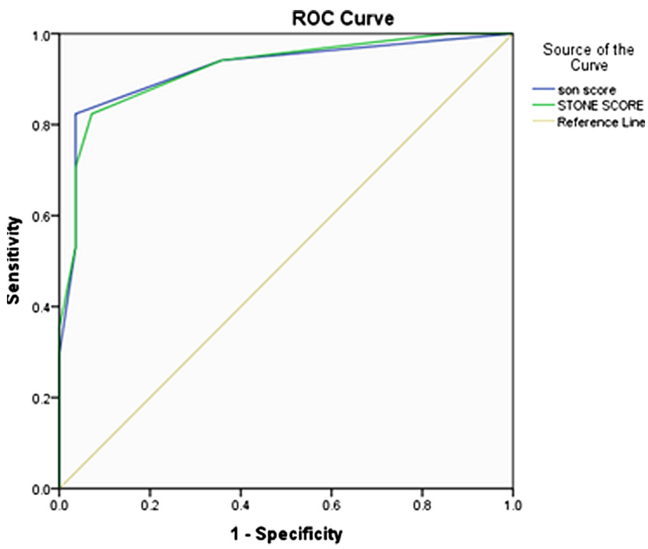

Fig. 3 ROC curves for S.T.O.N.E. and S.O.N. scores. The AUCs for the S.T.O.N.E. and S.O.N. scores were 0.923 and 0.922, respectively.

needed to access the pelvicalyceal system and clear the fragments. Too few obese patients in our present study might also be the reason for inappropriate study of this variable. In the current era, almost all types of stones except matrix are breakable with the current energy sources and might not affect the SFR. Okhunov et al. [10] believe that 'in a tertiary referral centre with substantial experience in PCNL, the tract length and stone density might have a lesser effect on perioperative outcomes when performed by an expert endourologist capable of compensating for factors that can increase the complexity of a case'. Thus, we calculated the S.O.N. score after omitting 'T' and 'E' from S.T.O.N.E. With a $P < 0.001$, it significantly affected the SFS. The AUC was almost consistent with the S.T.O.N.E. score (0.922 and 0.923, respectively; Fig. 3). However, this was not a significant increase ($P = 0.99$). Thus, both models were found to be equally effective in predicting the SFS. But this needs further validation in larger studies. And as with any study, the present study also had some limitations such as small sample size, absence of intraoperative flexible nephroscopy, and a lack of multiple observers in calculating the scores.

Conclusions

All three scoring tools were equally predictive of SFS and correlated weakly with complications after PCNL. A simplification of the S.T.O.N.E. score, i.e. S.O.N. score, is equally predictive and has the potential to bring the scoring system to preoperative counselling room in limited settings/non-obese patients. Further refinement in the definition, calculation and application of individual variables, clear cut definitions for staghorn stones, standardisation of the size of residual fragments, timing and mode of imaging to document SFS, and larger prospective studies are required to determine a single scoring system to be adopted for standardised academic reporting and preoperative prediction for PCNL outcomes.

Source of funding

None.

Conflicts of interest

None.

References

[1] Ramello A, Vitale C, Marangella M. Epidemiology of nephrolithiasis. *J Nephrol* 2000;**13**(Suppl. 3):S45–50.

[2] Dave C, Shetty S, Faraj K, Howes DS, Sinert RH, Craig S, et al. *Nephrolithiasis Treatment & Management*. Available at: http://emedicine.medscape.com/article/437096-treatment. Accessed 27 June 2015.

[3] Preminger GM, Assimos DG, Lingeman JE, Nakada SY, Pearle JS, Wolf Jr JS, et al. Chapter 1: AUA guideline on management of staghorn calculi: diagnosis and treatment recommendations. *J Urol* 2005;**173**:1991–2000.

[4] de la Rosette J, Assimos D, Desai M, Gutierrez J, Lingeman J, Scarpa R, et al. The clinical research office of the endourological society percutaneous nephrolithotomy global study: indications, complications, and outcomes in 5803 patients. *J Endourol* 2011;**25**:11–7.

[5] Labadie K, Okhunov Z, Akhavein A, Moreira DM, Moreno-Palacios J, Del Junco M, et al. Evaluation and comparison of urolithiasis scoring systems in percutaneous kidney stone surgery. *J Urol* 2015;**193**:154–9.

[6] Segura JW, Patterson DE, LeRoy AJ, Williams Jr HJ, Barrett RC, Benson Jr RC, et al. Percutaneous removal of kidney stones: review of 1000 cases. *J Urol* 1985;**134**:1077–81.

[7] Matlaga BR, Kim SC, Lingeman JE. Improving outcomes of percutaneous nephrolithotomy access. *EAU Update Ser* 2005;**3**:37–43.

[8] Akman T, Binbay M, Yuruk E, Sari E, Seyrek M, Kaba M, et al. Tubeless procedure is most important factor in reducing length of hospitalization after percutaneous nephrolithotomy: results of univariable and multivariable models. *Urology* 2011;**77**:299–304.

[9] Thomas K, Smith NC, Hegarty N, Glass JM. The Guy's stone score – grading the complexity of percutaneous nephrolithotomy procedures. *Urology* 2011;**78**:277–81.

[10] Okhunov Z, Friedlander JI, George AK, Duty BD, Moreira DM, Srinivasan AK, et al. S.T.O.N.E. nephrolithometry: novel surgical classification system for kidney calculi. *Urology* 2013;**81**:1154–9.

[11] Smith A, Averch TD, Shahrour K, Opondo D, Daels FP, Labate G, et al. A nephrolithometric nomogram to predict treatment success of percutaneous nephrolithotomy. *J Urol* 2013;**190**:149–56.

[12] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83.

[13] Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;**240**:205–13.

[14] de la Rosette JJ, Opondo D, Daels FP, Giusti G, Serrano A, Kandasami SV, et al. Categorisation of complications and validation of the Clavien score for percutaneous nephrolithotomy. *Eur Urol* 2012;**62**:246–55.

[15] Jeong CW, Jung JW, Cha WH, Lee BK, Lee S, Jeong SJ, et al. Seoul national university renal stone complexity score for predicting stone-free rate after percutaneous nephrolithotomy. *PLoS One* 2013;**8**:e65888.
[16] Mishra S, Sabnis RB, Desai M. Staghorn morphometry: a new tool for clinical classification and prediction model for percutaneous nephrolithotomy monotherapy. *J Endourol* 2012;**26**:6–14.
[17] Mishra S, Sabnis RB, Desai M. Percutaneous nephrolithotomy monotherapy for staghorn: paradigm shift for 'staghorn morphometry' based clinical classification. *Curr Opin Urol* 2012;**22**:148–53.
[18] Gadzhiev N, Brovkin S, Grigoryev V, Dmitriev V, Baketin P, Obidnyak V, et al. Are we ready to predict percutaneous nephrolithotomy (PCNL) stone-free failure? *J Clin Urol* 2016;**9**:11–8.
[19] Zhu Z, Wang S, Xi Q, Bai J, Yu X, Liu J. Logistic regression model for predicting stone-free rate after minimally invasive percutaneous nephrolithotomy. *Urology* 2011;**78**:32–6.
[20] Mandal S, Goel A, Kathpalia R, Sankhwar S, Singh V, Sinha RJ, et al. Prospective evaluation of complications using the modified Clavien grading system, and of success rates of percutaneous nephrolithotomy using Guy's Stone Score: a single-center experience. *Indian J Urol* 2012;**28**:392–8.
[21] Vicentini FC, Marchini GS, Mazzucchi E, Claro JF, Srougi M. Utility of the Guy's stone score based on computed tomographic scan findings for predicting percutaneous nephrolithotomy outcomes. *Urology* 2014;**83**:1248–53.
[22] Ingimarsson JP, Dagrosa LM, Hyams ES, Pais Jr VM. 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.
[23] Choo MS, Jeong CW, Jung JH, Lee SB, Jeong H, Son H, et al. External validation and evaluation of reliability and validity of the S-ReSC scoring system to predict stone-free status after percutaneous nephrolithotomy. *PLoS One* 2014;**9**:e83628.
[24] Okhunov Z, Helmy M, Perez-Lansac A, Menhadji A, Bucur P, Kolla SB, et al. Interobserver reliability and reproducibility of S.T.O.N.E. nephrolithometry for renal calculi. *J Endourol* 2013;**27**:1303–6.
[25] Akhavein A, Henriksen C, Syed J, Bird VG. 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.
[26] 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;**33**:1821–5.
[27] Finch W, Johnston R, Shaida N, Winterbottom A, Wiseman O. Measuring stone volume 3D Software reconstruction or an ellipsoid algebra formula? *BJU Int* 2014;**113**:610–4.
[28] Li H, Chen Y, Liu C, Li B, Xu K, Bao S. Construction of a three dimensional model of renal stones: comprehensive planning for percutaneous nephrolithotomy and assistance in surgery. *World J Urol* 2013;**31**:1587–92.
[29] Anastasiadis A, Onal B, Modi P, Turna B, Duvdevani M, Timoney A, et al. Impact of stone density on outcomes in Percutaneous Nephrolithotomy (PCNL): an analysis of the Clinical Research Office of the Endourological Society (CROES) PCNL global study database. *Scand J Urol* 2013;**47**:509–14.
[30] Gok A, Polat H, Cift A, Yucel MO, Gok B, Sirik M, et al. The hounsfield unit value calculated with the aid of non-contrast computed tomography and its effect on the outcome of percutaneous nephrolithotomy. *Urolithiasis* 2015;**43**:277–81.
[31] Rassweiler JJ, Renner C, Eisenberger F. The management of complex renal stones. *BJU Int* 2000;**86**:919–28.
[32] Di Silverio F, Gallucci M, Alpi G. Staghorn calculi of the kidney: classification and therapy. *Br J Urol* 1990;**65**:449–52.
[33] Desai M, De Lisa A, Turna B, Rioja J, Walfridsson H, D'Addessi A, et al. The clinical research office of the endourological society percutaneous nephrolithotomy global study: staghorn versus non staghorn stones. *J Endourol* 2011;**25**:1263–8.