

Nephrolithometric Scoring Systems to Predict Outcomes of Percutaneous Nephrolithotomy

Simone L. Vernez, BA,¹ Zhamshid Okhunov, MD,¹ Piruz Motamedinia, MD,² Vincent Bird, MD,³ Zeph Okeke, MD,⁴ Arthur Smith, MD⁴

¹Department of Urology, University of California, Irvine, Irvine, CA; ²Department of Urology, Yale School of Medicine, New Haven, CT; ³Department of Urology, University of Florida, Gainesville, FL; ⁴The Arthur DM Smith Institute for Urology, North Shore-LIJ Health System, New Hyde Park, NY

Currently, there is no single agreement upon an ideal predictive model that characterizes the complexity of renal stones and predicts surgical outcomes following percutaneous nephrolithotomy (PCNL). New predictive tools have recently emerged to systematically and quantitatively assess kidney stone complexity to predict outcomes following PCNL: the Guy's Stone Score, the CROES nomogram, S.T.O.N.E. nephrolithometry, and S-ReSC score. An ideal scoring system should include variables that both influence surgical planning and are predictive of postoperative outcomes. This review discusses the strengths, weaknesses, and commonalities of each of the above scoring systems. Additionally, we propose future directions for the development and analysis of surgical treatment for stone disease, namely, the importance of assessing radiation exposure and patient quality of life when counseling patients on treatment options.

[Rev Urol. 2016;18(1):15-27 doi: 10.3909/riu0693]

© 2016 MedReviews®, LLC

KEY WORDS

Percutaneous nephrolithotomy • Urolithiasis • Scoring system • Nomograms

The prevalence of urolithiasis in the United States has increased over the past two decades, reaching nearly 7% in women and 10.3% in men.¹ Though most small stones pass spontaneously or are amenable to shock wave lithotripsy, approximately 10% to 20% require surgical removal. Ureteroscopy and percutaneous nephrolithotomy

(PCNL) are the most commonly utilized modalities. Stone features such as size, extent of calyceal involvement, pelvicaliceal anatomy, and anatomic malformations dictate the feasibility of different treatment modalities and have significant impact on surgical outcomes.²⁻⁶ There is immense heterogeneity in methods for clinical and academic characterization

of nephrolithiasis and for the evaluation of surgical outcomes.

The Guy's Stone Score (GSS), the Clinical Research Office of the Endourological Society (CROES) nomogram, S.T.O.N.E. (stone size, tract length, obstruction, number of involved calices, and essence/stone density) nephrolithometry, and Seoul National University Renal Stone Complexity (S-ReSC) score allow for objective assessment of kidney stones and predict outcomes of PCNL.⁷⁻¹⁰ The widespread use of a standardized nephrolithometric scoring system can improve patient care by informing clinical decision making and patient counseling, in addition to improving academic reporting.

This review discusses the development of the aforementioned nephrolithometric scoring tools, and assesses their individual advantages and disadvantages to seek ways to improve them. The commonalities are discussed and additional factors that should be evaluated for incorporation into a comprehensive scoring system are suggested.

Stone Scoring Tools: Descriptions and Assessments

GSS: Development, Reproducibility, and Validation

The GSS was developed using evidence from published data combined with knowledge and experience of senior endourologists at a single institution.⁷ It consists of four grades based on stone burden and patient anatomy (Table 1). The score was refined using an iterative process after evaluating 10 consecutive cases. The score was prospectively validated in 100 patients who underwent PCNL procedures in a tertiary stone center. The authors used abdominal radiography

preoperatively, and computed tomography (CT) and abdominal radiography to determine stone-free status (SFS), as defined as no stones visible or presence of clinically insignificant residual fragments < 4 mm 6 weeks after PCNL. A multivariate analysis showed that the GSS was the only predictive factor of SFS as compared with stone burden, operating surgeon, weight, age, comorbidity, and urine culture. Moreover, SFS correlated with increasing grade of GSS: I = 81%, II = 74.2%, III = 35%, and IV = 29%. A higher grade correlated with the need for ancillary procedures. The authors did not find any correlation between the stone score and postoperative complications.⁷

External validation in several series demonstrated that GSS effectively predicted SFS.¹¹⁻¹⁴ Importantly, Ingimarsson and colleagues¹⁴ and Vicentini and associates¹³ used preoperative CT to validate the GSS. Additionally, Ingimarsson and colleagues¹⁴ used a more stringent stone-free measurement, < 2 mm fragment, to determine SFS and confirmed a good interrater reliability. Vicentini and associates¹³ also found a positive correlation between GSS and other operative outcomes, such as operative time

across varying guidelines for postoperative imaging.

There are a number of key limitations to this system, however. First, it fails to account for important variables such as calyceal involvement, stone size, density, and composition. These variables determine technical difficulty of PCNL and thus have important implications for procedural success.¹⁵⁻¹⁷ Second, the GSS was developed using variables the expert authors felt were significant rather than by data-driven selection. Data-driven models such as the CROES nomogram have been found to have superior predictability for SFS, and also predict certain postoperative complications, whereas the GSS does not.⁸ Third, the appeal of the GSS is that it is a simple four-grade compartmentalization. However, this is also its weakness, as this simplicity does not account for clinical variability. A patient with spina bifida and full staghorn would receive the same grade IV rating even if he or she only had a 2-cm pelvic stone. The use of just four categories to describe complex stone disease misses important differences in outcomes and complications among different stones and patients with stones. Finally, the overall SFS of 62% was relatively low in compari-

The GSS provides a simple, intuitive, and reproducible tool for predicting SFS following PCNL.

and blood transfusion rates, as well as postoperative complications.

Strengths and Weaknesses.

The GSS provides a simple, intuitive, and reproducible tool for predicting SFS following PCNL. Arguably, its external validation and use of multiple imaging modalities, including abdominal radiograph and CT, broadens its universal appeal and integration

son with global overall SFS,⁸ raising the question of generalizability.¹⁸

CROES Nomogram: Development, Reproducibility, and Validation

The CROES nomogram is a data-driven predictive model created with the primary objective of assessing current indications and treatment outcomes in PCNL.¹⁹ A

TABLE 1

Current Nephrolithometric Scoring Systems

Scoring System	Categorization	Method of Derivation
Guy's Stone Score ^a	<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>	Literature review Expert opinion Iterative process
Nephrolithometric Nomogram ^b	<p>A: Stone Burden – calculated as follows:</p> <ol style="list-style-type: none"> 1. Measure the maximum length of each stone in millimeters 2. Measure the maximum width of each stone in millimeters 3. Calculate the stone burden for each stone = $0.785 \times \text{length} \times \text{width}$ 4. Add individual stone burdens if multiple stones <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</p>	Multiple logistic regression analysis
S.T.O.N.E. Nephrolithometry ^c	<p>Scoring based on 5 variables from preoperative noncontrast computed tomography stone size:</p> <p>S = stone size</p> <ol style="list-style-type: none"> 1: 0-399 mm² 2: 400-799 mm² 3: 800-1599 mm² 4: > 1600 mm² <p>T = tract length</p> <ol style="list-style-type: none"> 1: < 100 mm 2: > 100 mm <p>O = Obstruction</p> <ol style="list-style-type: none"> 1: no or mild dilatation 2: moderate to severe dilatation <p>N = Number of involved calyces</p> <ol style="list-style-type: none"> 1: 1 calyx involved 2: 2-3 calyces involved 3: full staghorn calculus <p>E = Essence (stone density)</p> <ol style="list-style-type: none"> 1: < 950 HU 2: > 950 HU 	Systematic review
Seoul Renal Stone Complexity Score ^d	Identify number of preselected pelvicalyceal locations involved, each of the 9 locations is worth 1 point; total score = cumulative number of locations involved	Not reported

^aFrom Thomas K et al.⁷
^bFrom Smith A et al.⁸
^cFrom Okhunov Z et al.⁹
^dFrom Jeong CW et al.¹⁰

total of 2806 patients from 96 centers worldwide were included in the development of the nomogram. Prior treatments, body mass index (BMI), staghorn stones, renal anomalies, and stone burden, location, and count were all recorded. The authors report that stone burden, calyceal location, stone count, and presence of staghorn stones were the most influential variables predicting SFS. Each variable was associated with a score from 0 to 100 as outlined in the nomogram. The sum of the individual scores was used to predict the SFS, defined as no visible stones or residual fragments < 4 mm on abdominal radiograph 30 days following PCNL. An interpretation scale was created to associate the total score with the percent chance of achieving SFS and bootstrapping techniques were used with 1000 cases to validate the model. Additionally, the authors determined that the nomogram proved clinically useful for directing treatment strategy at a threshold of 60% or more estimated SFS using a decision analysis curve.⁸

The CROES nomogram was externally validated by Labadie and colleagues¹² and Bozkurt and colleagues.²⁰ Significantly, Bozkurt and colleagues found that the CROES nomogram correlated with SFS using CT as well as operative time, estimated blood loss (EBL), length of hospital stay, and complications.

Strengths and Weaknesses.

The CROES nomogram is widely generalizable because it is based on global data generated from both high- and low-volume centers. Moreover, the nomogram uniquely grades risk across a continuous scale rather than dividing stones of varying complexity into discrete groups.

There are, however, two important limitations. First, the CROES

database was not created specifically for the development of a predictive model for classification of stone disease prior to PCNL and, therefore, lacks important variables affecting the outcomes. The authors acknowledge that the study lacked radiologic data on hydronephrosis and other pelvicalyceal abnormalities, each of which may influence outcomes. It also does not take into account the stone density or composition, which have been shown to influence SFS.^{21,22} Second, nomograms are complex, requiring a more time-consuming evaluation of preoperative variables. Many physicians feel that nomograms are impractical in the clinical setting.^{7,23}

S.T.O.N.E. Nephrolithometry: Development, Reproducibility, and Validation

S.T.O.N.E. nephrolithometry was developed via a literature review of English language studies from 1976

the degree of training and levels of expertise with CT imaging clearly impacted the accuracy of grading and assessment of stone complexity. The authors found that quantifying the stone size and number of involved calyces were least reproducible overall.²⁴

S.T.O.N.E. nephrolithometry has been externally validated in a number of studies. Akhavein and colleagues²⁵ validated the system as a reproducible and predictive model for treatment success in 117 patients. A multi-institutional study of 850 patients confirmed that the model was significantly associated with SFS, overall complications, EBL, operation time, and length of stay.²⁶ Studies by Labadie and associates¹² and Noureldin and associates^{27,28} showed that the system was correlated with EBL, operative time, and length of hospital stay, but not complications. S.T.O.N.E. has also been validated among a pediatric population, cor-

S.T.O.N.E. nephrolithometry uniquely includes variables that have been shown to have significant impact on postoperative outcomes.

to 2012 on Medline to identify the most clinically relevant and reproducible variables that had been shown to impact outcomes following PCNL. Its components are stone size, PCNL tract length, presence of obstruction, number of involved calyces, and stone density (Table 1), measured from preoperative CT. A cohort of 117 PCNL patients was used to evaluate the predictive value of the S.T.O.N.E. nephrolithometry. A “low” score of 3 to 5 demonstrated a correlation with SF rates of 94% to 100%, a “moderate” score of 6 to 8 correlated with a SF rate of 83% to 92%, and “high” scores of 9 to 13 correlated with SF rates ranging from 27% to 64%.⁹ A subsequent study demonstrated interobserver reliability, though

relating with SFS, length of hospitalization, and complications.²⁹

Strengths and Weaknesses.

S.T.O.N.E. nephrolithometry uniquely includes variables that have been shown to have significant impact on postoperative outcomes. Furthermore, because it stratifies patients into low-, moderate-, and high-risk groups, it is clinically practical for decision making and surgical planning. Notably, S.T.O.N.E. can be calculated using only a preoperative CT scan, making it ideal for building a retrospective database with limited clinical information.

One limitation is that S.T.O.N.E. was validated with a small cohort. This may limit its applicability

TABLE 2**Validation Studies for Current Nephrolithometric Scoring Systems**

Study	Country	N	Preoperative Imaging	Postoperative Imaging	Definition of Stone Free	Clearance by Stone Complexity	Complications by Stone Complexity	Statistical Method
Guy's Stone Score								
Thomas K et al ⁷	United Kingdom	100	CT, radiograph, IVU	KUB radiograph	< 4-mm fragments	Overall, 62% Grade 1, 81% Grade 2, 72.4% Grade 3, 35% Grade 4, 29%	Not reported; no significant correlation only	Multivariate linear regression analysis
Mandal S et al ¹¹	India	278	Radiograph, IVU, USS, noncontrast CT	KUB radiograph	Complete absence of stones	Overall, 76.1% Grade 1, 100% Grade 2, 74% Grade 3, 56% Grade 4, 0%	Overall, 41.7% Grade 1, 4.8% Grade 2, 9.1% Grade 3, 26.2% Grade 4, 44.4%	Univariate analysis
Vicentini FC et al ¹³	Brazil	147	Noncontrast CT		Asymptomatic fragments < 4 mm	Overall, 71.6% Grade 1, 95.2% Grade 2, 79.5% Grade 3, 59.5% Grade 4, 40.7%	Overall, 18.7% Grade 1, 4.8% Grade 2, 9.1% Grade 3, 26.2% Grade 4, 44.4%	Univariate analysis
Ingimarsson JP et al ¹⁴	Lebanon	66	Noncontrast CT		No fragments/no fragments > 2 mm/ no fragments > 4 mm	Overall, 90% Grade 1, 95% Grade 2, 97% Grade 3, 95% Grade 4, 75%	Not reported; no significant correlation only	Univariate analysis
Labadie K et al ¹²	United States	246	CT scan	Unclear	Not defined	Overall, 56% Grade 1, 70.2% Grade 2, 65.4% Grade 3, 48.1% Grade 4, 35.9%	Overall, 17%	Logistic regression
Noureldin YA et al ²⁷	Canada	185	Unclear	Plain radiographs or non-contrast CT	Asymptomatic residual fragments < 4 mm	Overall, 71.9% Grade 1, 91.5% Grade 2, 80% Grade 3, 47.6% Grade 4, 43.2%	Overall, 16.2% Grade 1, 17% Grade 2, 15% Grade 3, 9.5% Grade 4, 21.6%	Logistic regression analysis

(Continued)

Validation Studies for Current Nephrolithometric Scoring Systems (*continued*)

Study	Country	N	Preoperative Imaging	Postoperative Imaging	Definition of Stone Free	Clearance by Stone Complexity	Complications by Stone Complexity	Statistical Method
Bozkurt IH et al ²⁰	Turkey	437	CT scan	KUB radiograph (unless CT; which is reserved for only symptomatic patients)	Asymptomatic residual fragments < 4 mm	Overall, 75% Grade 1, 92.8% Grade 2, 72.4% Grade 3, 68.5% Grade 4, 47.4%	Overall, 35% Grade 1, 13.0% Grade 2, 2.5% Grade 3, 8.7% Grade 4, 10.3%	Linear and logistic regression analyses
Nephrolithometric Nomogram (CROES)								
Smith A et al ⁸	International (CROES PCNL Group)	2806	Unclear	KUB radiograph	No fragments > 4 mm	Overall, 82% Odds ratios and nomogram reported ROC AUC = 0.76	Not reported	Multivariate logistic regression
Labadie K et al ¹²	United States	246	CT scan	Unclear	Not defined	Overall, 56% 80-129 = 22.7% 130-169 = 46.4% 170-219 = 45.5% ≥ 220 = 72.7%	Overall, 17%	Logistic regression
Bozkurt IH et al ²⁰	Turkey	437	CT scan	KUB radiograph (unless CT; which is reserved for only symptomatic patients)	Asymptomatic residual fragments < 4 mm	Overall, 75% Grade 1, 48.5% Grade 2, 51.4% Grade 3, 69.1% Grade 4, 92.8%	Overall, 35% Grade 1, 4.8% Grade 2, 8.0% Grade 3, 7.5% Grade 4, 14.2%	Linear and logistic regression analyses
S.T.O.N.E. Nephrolithometry								
Okhunov Z et al ⁹	United States	117	Noncontrast CT	Fluoroscopy ± CT	Complete absence of stones on CT	Overall, 80% Low (4-5), 94%-100%; medium (6-8), 83%-92%; high (9-11), 27%-64%; ROC AUC = 0.83	Not reported	Multivariate logistic regression

Okhunov Z et al ²⁴	United States	58	Noncontrast CT	Fluoroscopy ± CT	N/A ^a	N/A ^a	N/A ^a	Cohen's kappa coefficient ^a
Okhunov Z et al ²⁶	United States	706	CT scan	Unclear	Not defined	Overall 67% S.T.O.N.E. score significantly associated with postoperative SFS	S.T.O.N.E. score significantly associated with overall complication rate	Univariate analysis
Labadie K et al ¹²	United States	246	CT scan	Unclear	Not defined	Overall, 56% 5-6 = 70.6% 7-8 = 66.3% 9-13 = 46.2% ROC AUC = 0.67	Overall, 17%	Logistic regression
Akhavain A et al ²⁵	United States	122	CT scan		Residual fragments < 4 mm	Score vs clearance %: 5-7 = 100% 8 = 80% 9 = 94.1% 10 = 79.2% 11 = 66.7% 12 = 45%	Score vs complication %: 5 = 0% 6 = 1.6% 7-8 = 0% 9 = 0.8% 10 = 6.6% 11-12 = 0%	Univariate analysis and logistic regression
Noureldin YA et al ²⁷	Canada	185	Unclear	Plain radiographs or noncontrast CT	Asymptomatic residual fragments < 4 mm	Overall, 71.9% 5 = 100% 6 = 75% 7 = 77.3% 8 = 67.6% 9 = 79.3% 10 = 58.3% 11 = 42.9% 12-13 = 0%	Overall, 16.2% 5 = 20% 6 = 18.2% 7 = 11.4% 8 = 20.6% 9 = 20.7% 10 = 8.3% 11 = 0% 12 = 25% 13 = 0%	Logistic regression analysis

(Continued)

Validation Studies for Current Nephrolithometric Scoring Systems (*continued*)

Study	Country	N	Preoperative Imaging	Postoperative Imaging	Definition of Stone Free	Clearance by Stone Complexity	Complications by Stone Complexity	Statistical Method
Noureldin YA et al ²⁸	Canada	155	Noncontrast CT	KUB and/or non-contrast CT	Residual stones \geq 4 mm considered as non-stone-free	Overall, 71.6% Significant inverse association between S.T.O.N.E. score and clearance AUC = 0.63	Overall, 15.5% No correlation between S.T.O.N.E. score and complications	Multivariate logistic regression
Seoul Renal Stone Complexity Score								
Jeong CW et al ¹⁰	South Korea	155	Noncontrast CT		No evidence of residual stones on postop images after 1 mo	Overall, 72.3% Low (1-2), 96%; medium (3-4), 69%; high (5-9), 28%	Overall, 25.2% Low (1-2), 1.9%; medium (3-4), 28.6%; high (5-9), 34.2%	Univariate analysis; ROCs
Choo MS et al ³¹	South Korea	327	Noncontrast CT		Complete clearance/clearance with clinically insignificant residual fragments (4 mm) at 1 mo follow-up imaging	Overall, 65.4% Low (1-2), 83.9%; medium (3-4), 47.6%; high (5-9), 21.4%	Not investigated	Univariate and multivariate regression analyses; ROCs
Park J et al ³²	South Korea	159	Unclear		No evidence of a stone or with clinically insignificant residual fragment stones < 2 mm	Overall, 73.0% Low (1-2), 86.7%; medium (3-4), 70.2%; high (5-12), 48.6%	Not investigated	Uni- and multivariate linear regression

^aStudy focuses on interobserver reliability.

AUC, area under the curve; CT, computed tomography; IVU, intravenous urogram; KUB, kidneys, ureters, and bladder; ROC, receiver operating characteristic; SFS, stone-free status; S.T.O.N.E., stone size, tract length, obstruction, number of involved calices, and essence/stone density; USS, abdominal and pelvic ultrasound.

to a wider patient population. Moreover, Taily and Razvi³⁰ point out that the low incidence of treatment failure in the initial cohort may have introduced an element of bias into the development and validation of the tool. A larger study is needed for validation.

S.T.O.N.E. would benefit from greater refinement of the methods used to score each of the factors to take into account their relative predictive power. It is not clear that the cutoffs used for tract length and stone density are optimal. Furthermore, the authors acknowledge the need for a standardized method to measure stone size and number of calices involved in order to improve predictive value. A standardized definition of calyx, as well as the imaging plane used to enumerate calyceal involvement, would improve reproducibility. Additionally, the score assigned by degree of hydronephrosis is subjective and would benefit from stricter criteria.

S-ReSC Score: Development, Reproducibility, and Validation

The S-ReSC score was developed based on expert hypotheses about what influences surgical complexity and SFS. The authors assert that the distributional complexity of stones is the most important factor that determines treatment success. They devised a 9-point system, with 1 point assigned to 9 specific pelvic and calyceal locations. A score of 1 to 2 is considered low, 3 to 4 is medium, and ≥ 5 is high. The score is assigned by adding the cumulative sites involved; 155 consecutive cases at one center were used to retrospectively validate S-ReSC. The SFS for low (1 and 2) scores was 96.0%, for medium (3 and 4) was 69.0%, and for high (≥ 5) were 28.9%. Logistic regression showed that S-ReSC significantly predicted SFS. Additionally, multiple

regression analyses showed that the S-ReSC score was associated with stone number, burden, and renometry, variables that are commonly accepted as being predictive of postoperative success. It also showed that the S-ReSC score had comparatively higher predictive value than any variable alone.¹⁰

Three studies each externally validated the S-ReSC score. In a 327-patient cohort at two tertiary centers, Choo and colleagues³¹ demonstrated that it predicted SFS as defined by < 4 mm. Park and colleagues³² validated the S-ReSC score for retrograde intrarenal surgery (RIRS), rather than PCNL. The S-ReSC score was found to be predictive of SFS with low (1-2), intermediate (3-4), and high (5-12) scores correlated with 86.7%, 70.2%, and 48.6% SFS, respectively. Similarly, Jung and associates³³ validated a modified S-ReSC score for predic-

influence success rates in the larger Western patient population (skin-to-stone distance, a surrogate for BMI) may prove useful in broadening its application.

Discussion

A widely applicable, simple disease stratification tool will greatly improve patient counseling, surgical planning, evaluation of outcomes, and academic reporting.^{34,35} It will allow reliable and accurate comparisons of treatment efficacy and quality of surgical care, and facilitate the meaningful comparison of clinical studies.³⁶ Furthermore, it will facilitate risk adjustment,³⁷ enabling physicians to better define the nephrolithiasis disease continuum and identify patients who should be referred to tertiary care centers.³⁸ In order to achieve this, a single, standardized,

The S-ReSC score is a simple, easily applicable scoring tool that predicts SFS. It is based on a single variable—stone location—which is determined with preoperative CT and demonstrates high reproducibility as a result of this simplicity.

tion of SFS after RIRS by adding additional points for hard-to-reach stones in the lower poles.

Strengths and Weaknesses.

The S-ReSC score is a simple, easily applicable scoring tool that predicts SFS. It is based on a single variable—stone location—which is determined with preoperative CT and demonstrates high reproducibility as a result of this simplicity.

Significantly, the S-ReSC score relies on the expert hypothesis that a single variable—stone location—is the ultimate predictive indicator of SFS. However, it is clear from the literature that there are many variables that influence outcomes. Additionally, the score was created and validated initially with a cohort at a single center. Factors that may

widely used system that is easily determined, applicable to individual cases, and indicative of postoperative outcomes is necessary.

We have reviewed several available tools that inform surgical planning and predict PCNL outcomes.⁷⁻¹⁰ No study comparing systems has found a clinically significant difference in their predictive power. Next, we discuss the variables that are common to all of the scoring systems in one form or another, namely, stone burden and location. Additionally, we discuss individual variables not common to all that influence surgical complexity and outcomes. Finally, we discuss future directions for the development and analysis of surgical treatment for stone disease.

Common Variables: Stone Burden

It is well documented that the most important variable to predict outcomes of PCNL is stone burden.^{12,15-17,39,40} In the first study providing a head-to-head comparison of three nephrolithometry systems, GSS, CROES nomogram, and S.T.O.N.E., Labadie and colleagues¹² found that stone burden alone had the same predictive value for stone-free rates as did each of the scoring systems. Most urologists currently utilize stone size to estimate complexity of the procedure and surgical outcomes; however, there is a lack of uniformity in the literature with regard to calculating stone burden.^{41,42} Each of the described models uses a different algorithm for calculating the stone burden. None of these methods has been shown to be superior, so the most reproducible method that accurately reflects clinically significant stone burden should be utilized.

Common Variables: Location

Stone location (or distribution) has an important impact on the achievement of SFS.⁴³ The GSS grade relies

Stone location (or distribution) has an important impact on the achievement of SFS.

heavily on location, as single stones in the lower or middle poles or renal pelvis are assigned the lowest grade, whereas single stones in the upper pole are assigned one grade higher. The CROES nomogram goes further, assigning a different score for the renal pelvis, and lower, middle, and upper poles. S.T.O.N.E. accounts for the number of calyces infiltrated by the stone, but does not specifically include stone location in determining overall score. The S-ReSC score is solely based on stone distribution as determined by the cumulative number of calyces

involved but does not consider their location. The relative weight that each scoring system places on the stone location is a key determinant of how it influences the score.

Additional Variables: Staghorn Stones

Staghorn stones are particularly challenging, requiring more complex procedures, multiple tracts, and, sometimes, multistaged or multimodal treatments.⁴⁴ There is no precise definition for staghorn calculi,^{41,44} leading to interobserver inconsistencies. Mishra and colleagues⁴² created Staghorn Morphometry, a system by which the volume and distribution of staghorn stones can be evaluated to predict the requirement for number of tracts and stages for PCNL procedures. They proposed a scoring system that classifies each stone into one of three types. The authors provide a sophisticated method to

predict the number of tracts and stages required for PCNL for stag-

horn renal calculi; however, it only applies to one type of stone and requires complex imaging software for calculation. Whether specifying staghorn status specifically or accounting for it by proxy via stone burden is required for the best predictive outcome has not been studied.

Additional Variables: Anatomic Malformations

The influence of anatomic abnormalities on PCNL outcomes has been previously evaluated and discussed in several studies.^{2,6}

However, there is no consensus on their relevance, given that a number of studies have demonstrated that abnormal renal anatomy is not associated with inferior surgical outcomes.⁴⁵⁻⁴⁷ Though anatomic variables such as duplicated collecting system, horseshoe kidneys, and calyceal diverticulum were evaluated, during the literature review in the development of S.T.O.N.E., due to the low incidence of these conditions, the study was not powered sufficiently to detect a difference and therefore they were not included in the resulting scoring system. These parameters should be evaluated using a larger cohort to elucidate their impact role in the scoring system.

Additional Variables: Previous Surgery and Experience of Surgeon

Only the CROES nomogram takes previous surgical intervention and

Only the CROES nomogram takes previous surgical intervention and caseload or experience of the surgeon into consideration, having found that both significantly impact treatment success.

caseload or experience of the surgeon into consideration, having found that both significantly impact treatment success. Previous stone treatment has a twofold informational yield. First, it indicates prior therapy failure, demonstrating the complexity of the stone. Second, case volume and experience clearly impact surgical outcomes.

Additional Variables: Tract Length

Skin-to-stone distance is an important parameter to predict success in PCNL. In obese patients in particular, tract length dictates the difficulty of obtaining surgical access. Additionally, skin-to-stone distance alters the effectiveness of real-time fluoroscopy necessary for proper access. The increased skin-to-stone

distance makes puncture, dilation, and securing a tract more difficult. Finally, challenges related to initial access add difficulty to the limitations on length of working sheaths, nephroscopes, and working instruments.⁴⁸ A longer nephrostomy tract decreases maneuverability of

the stone may be more indicative of the postoperative complication risk. Borrowing from the surgical oncology literature and the concept of patient frailty, Leavitt and associates⁵⁰ have illustrated that a patient's inability to independently perform activities of daily living is

should play an important role in clinical decision making. Patients with recurrent urinary tract stones report more bodily pain, depression, and lower general health scores than the general American public.^{51,53-57} Simply measuring physiologic and biologic outcomes is not ideal. A comprehensive universal tool should not only consider procedural complexity and outcomes, but also meaningful measures for patients such as quality of life. A patient may prefer a staged ureteroscopic approach over a single-setting PCNL; these preferences should be considered.

A longer nephrostomy tract decreases maneuverability of the nephroscope and may result in longer procedures with a lower SFS and increased complication rates. S.T.O.N.E. is the only system that includes tract length.

the nephroscope and may result in longer procedures with a lower SFS and increased complication rates.^{5,49} S.T.O.N.E. is the only system that includes tract length. CROES included BMI, a possible correlate, in the initial analysis during development of the system, but ultimately did not find it to be predictive and therefore excluded it from the nomogram.

Future Directions

An all-inclusive system for classifying nephrolithiasis will inform clinical decision making by impacting the risk-benefit analysis of treatment options through objective prediction of outcomes. Additionally, none of the scoring systems were evaluated for their ability to correlate with other guiding factors that inform clinical decision making, such as radiation exposure and quality-of-life outcomes.

Surgical Complications

Risk of surgical complications is an important variable that should guide clinical decision making. In their comparative study, Labadie and colleagues¹² found that, although the GSS and S.T.O.N.E. successfully predicted length of stay and EBL, none of the three scoring systems predicted surgical complications. Patient factors unrelated to

a strong predictor of complications and length of stay following PCNL.

Predicting Radiation Time

A recent multicenter study found that patients with greater stone size, BMI, number of involved calyces, skin-to-stone distance, and S.T.O.N.E. nephrolithometry score have increased radiation exposure during PCNL.⁵¹ In their multivariate regression, Thomas and associates⁷ showed that patient weight, not individual GSS, was correlated with radiation time and exposure. Sfoungaristos and colleagues⁵² found that GSS, CROES score, and S.T.O.N.E. are significantly associ-

Defining the Disease Continuum: Applicability to Multiple Procedures

Each system described above was developed as a way of quantifying stone complexity and predicting surgical outcomes. A system that can be adapted to and therefore be descriptive of multiple surgical procedures is ideal, as it will allow for outcome comparisons and individualized surgical planning encompassing all available modalities. There are a

Understanding the relative predictive value of a universal tool for multiple treatment modalities will broaden its applicability to stone disease and allow for effective comparison between treatment options, informing clinical decision making.

ated with fluoroscopy time. With a predictive classification system, benefits and risks of using alternative imaging modalities during evaluation of disease, as well as during stone removal, may be weighed against increased radiation exposure.

Quality of Life

A recent survey of the literature found health-related outcomes for stone management have revolved around quality-of-life outcomes, suggesting that quality of life

number of additional systems that predict outcomes following other procedures, such as RIRS and extracorporeal shock wave lithotripsy^{58,59}; thus far, however, only the S.T.O.N.E. and the S-ReSC scores have been studied for its applicability to alternate procedures.^{32,60} Understanding the relative predictive value of a universal tool for multiple treatment modalities will broaden its applicability to stone disease and allow for effective comparison between treatment options, informing clinical decision making.

Conclusions

Prior to surgery, patients should be provided with the most accurate information about their estimated SFS and risk of complications. Categorizing of patients into low-, intermediate-, and high-risk surgical groups helps surgeons appropriately counsel patients regarding their likely postoperative clinical course and, in complex scenarios, refer patients to tertiary centers. Current scoring systems have not been prospectively validated, and existing validation papers are limited in utility due to the small, homogenous number of patients studied. Therefore, questions about reproducibility and validity remain unanswered. In addition, it is unclear whether variables such as anatomic malformations and stone composition should be included in a universal scoring system. Existing data are insufficient to state which scoring system is best. The most reproducible and clinically relevant

predictive tool for wide adoption has yet to be identified. ■

References

1. Scales CD Jr, Smith AC, Hanley J, Saigal CS; Urologic Diseases in America Project. Prevalence of kidney stones in the United States. *Eur Urol*. 2012;62:160-165.
2. Binbay M, Akman T, Ozgor F, et al. Does pelvicaliceal system anatomy affect success of percutaneous nephrolithotomy? *Urology*. 2011;78:733-737.
3. Resorlu B, Oguz U, Resorlu EB, et al. The impact of pelvicaliceal anatomy on the success of retrograde intrarenal surgery in patients with lower pole renal stones. *Urology*. 2012;79:61-66.
4. Bagrodia A, Gupta A, Raman JD, et al. Predictors of cost and clinical outcomes of percutaneous nephrostolithotomy. *J Urol*. 2009;182:586-590.
5. Fuller A, Razvi H, Denstedt JD, et al; CROES PCNL Study Group. The CROES percutaneous nephrolithotomy global study: the influence of body mass index on outcome. *J Urol*. 2012;188:138-144.
6. Ooster P, Razvi H, Liatsikos E, et al; CROES PCNL Study Group. Percutaneous nephrolithotomy among patients with renal anomalies: patient characteristics and outcomes; a subgroup analysis of the clinical research office of the endourological society global percutaneous nephrolithotomy study. *J Endourol*. 2011;25:1627-1632.
7. Thomas K, Smith NC, Hegarty N, Glass JM. The Guy's stone score--grading the complexity of percutaneous nephrolithotomy procedures. *Urology*. 2011;78:277-281.
8. Smith A, Averch TD, Shahrour K, et al; CROES PCNL Study Group. A nephrolithometric nomogram to predict treatment success of percutaneous nephrolithotomy. *J Urol*. 2013;190:149-156.
9. Okhunov Z, Friedlander JJ, George AK, et al. S.T.O.N.E. nephrolithometry: a novel surgical classification system for kidney calculi. *Urology*. 2013;81:1154-1159.
10. Jeong CW, Jung JW, Cha WH, et al. Seoul National University renal stone complexity score for predicting stone free rate after percutaneous nephrolithotomy. *PLOS One*. 2013;8:e65888.
11. Mandal S, Goel A, Kathalia R, 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;28:392-398.
12. Labadie K, Okhunov Z, Akhanavein A, et al. Evaluation and comparison of urolithiasis scoring systems used in percutaneous kidney stone surgery. *J Urol*. 2015;193:154-159.
13. Vicentini FC, Giovanni SM, Mazzucchi E, et al. Utility of the Guy's stone score based on computed tomographic scan findings for predicting percutaneous nephrolithotomy outcomes. *Urology*. 2014;83:1248-1253.
14. Ingimarsson JP, Dagrosa LM, Hyams ES, Pais VM Jr. 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;83:45-49.
15. Zhu Z, Wang S, Xi Q, et al. Logistic regression model for predicting stone-free rate after minimally invasive percutaneous nephrolithotomy. *Urology*. 2011;78:32-36.
16. Bagrodia A, Gupta A, Raman JD, et al. Predictors of cost and clinical outcomes of percutaneous nephrostolithotomy. *J Urol*. 2009;182:586-590.
17. Knox ML, Cantor AM, Bryant JE, Burns JR. Predictive factors for percutaneous nephrolithotomy outcomes in neurogenic bladder population. *J Endourol*. 2012;26:823-827.
18. Matlaga BR, Hyams ES. Can the Guy's stone score predict PNL outcomes? *Nature*. 2011;8:363-364.
19. Global PCNL observational study. Clinical Research Office of the Endourological Society website. <http://www.croesoffice.org/OngoingProjects/PCNLStudy.aspx>. Accessed January 29, 2016.

MAIN POINTS

- Although most small kidney stones pass spontaneously or are amenable to shock wave lithotripsy, approximately 10% to 20% require surgical removal. Features such as stone size, composition, extent of calyceal involvement, and anatomy dictate the feasibility of different treatment modalities and have significant impact on surgical outcomes; there is immense heterogeneity in methods for clinical and academic characterization of nephrolithiasis and for the evaluation of surgical outcomes.
- The Guy's Stone Score (GSS), the Clinical Research Office of the Endourological Society (CROES) nomogram, S.T.O.N.E. (stone size, tract length, obstruction, number of involved calices, and essence/stone density) nephrolithometry, and Seoul National University Renal Stone Complexity (S-ReSC) score allow for objective assessment of kidney stones and predict outcomes of percutaneous nephrolithotomy.
- A widely applicable, simple disease stratification tool will greatly improve patient counseling, surgical planning, evaluation of outcomes, and academic reporting by allowing for reliable and accurate comparisons of treatment efficacy and quality of surgical care.
- Categorizing of patients into low-, intermediate-, and high-risk surgical groups may help surgeons appropriately counsel patients regarding their likely postoperative clinical course and, in complex scenarios, refer patients to tertiary centers. Existing data are insufficient to support the adoption of a universal scoring system. The most reproducible and clinically relevant predictive tool for a wide adoption has yet to be identified.

20. 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;29:1006-1010.
21. Gücük A, Uyetürk U, Öztürk U, et al. Does the Hounsfield unit value determined by computed tomography predict the outcome of percutaneous nephrolithotomy? *J Endourol.* 2012;25:792-796.
22. Lehman DS, Hrubby GW, Phillips C, et al. Prospective randomized comparison of a combined ultrasonic and pneumatic lithotrite with a standard ultrasonic lithotrite for percutaneous nephrolithotomy. *J Endourol.* 2008;22:285-289.
23. Ferguson MK, Durkin AE. A comparison of three scoring systems for predicting complications after major lung resection. *Eur J Cardiothorac Surg.* 2003;23:35-42.
24. Okhunov Z, Helmy M, Perez-Lansac A, et al. Interobserver reliability and reproducibility of S.T.O.N.E. nephrolithometry for renal calculi. *J Endourol.* 2013;27:1303-1306.
25. Akhavan A, Henriksen C, Bird VG. Prediction of single procedure success rate using S.T.O.N.E. nephrolithometry surgical classification system with strict criteria for surgical outcomes. *J Urol.* 2013;189(suppl):e627-e628.
26. Okhunov Z, Moreira D, George A, et al. Multicenter validation of S.T.O.N.E. nephrolithometry. *J Urol.* 2014;191(suppl):e839.
27. 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;33:1821-1825.
28. Noureldin YA, Elkoushy MA, Andonian S. External validation of the S.T.O.N.E. nephrolithometry scoring system. *Can Urol Assoc J.* 2015;9:190-195.
29. 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;43:363-367.
30. Taïly T, Razvi H. The S.T.O.N.E. nephrolithometry scoring system: how valid is it? *Can Urol Assoc J.* 2015;9:196.
31. Choo MS, Jeong CW, Jung JH, 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.
32. Park J, Kang M, Jeong CW, et al. External validation and evaluation of reliability and validity of the Modified Seoul National University Renal Stone Complexity Scoring System to predict stone-free status after retrograde intrarenal surgery. *J Endourol.* 2015;29:888-893.
33. Jung JW, Lee BK, Park YH, et al. Modified Seoul National University Renal Stone Complexity score for retrograde intrarenal surgery. *Urolithiasis.* 2014;42:335-340.
34. de la Rosette JJ, Laguna MP, Rassweiler JJ, Conort P. Training in percutaneous nephrolithometry—a critical review. *Eur Urol.* 2008;54:994-1001.
35. Allen D, O'Brien T, Tiptaft R, Glass J. Defining the learning curve for percutaneous nephrolithotomy. *J Endourol.* 2005;19:279-282.
36. De Sio M, Autorino G, Quarto R, et al. Modified supine versus prone position in percutaneous nephrolithotomy for renal stones treatable with a single percutaneous access: a prospective randomized trial. *Euro Urol.* 2008;54:196-202.
37. Opondo D, Gravas S, Joyce A, et al. Standardization of patient outcomes reporting in percutaneous nephrolithotomy. *J Endourol.* 2014;28:767-774.
38. Rassweiler JJ, Renner C, Eisenberger R. The management of complex renal stones. *BJU Int.* 2000;86:919-928.
39. Hyams E, Bruhn A, Lipkin M, et al. Heterogeneity in the reporting of disease characteristics and treatment outcomes in studies evaluating treatments for nephrolithiasis. *J Endourol.* 2010;24:1411-1414.
40. Patel SR, Nakada SY. Quantification of preoperative stone burden for ureteroscopy and shock wave lithotripsy: current state and future recommendations. *Urology.* 2011;78:282-285.
41. Preminger GM, Assimos DG, Lingeman JE, et al; AUA Nephrolithiasis Guideline Panel. Chapter 1: AUA guideline on management of staghorn calculi: diagnosis and treatment recommendations. *J Urol.* 2005;173:1991-2000.
42. 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.
43. Shahrour K, Tomaszewski J, Ortiz T, et al. Predictors of immediate postoperative outcome of single-tract percutaneous nephrolithotomy. *Urology.* 2012;80:19-25.
44. Mishra S, Bhattu AS, Sabnis RB, Desai MR. Staghorn classification: platform for morphometry assessment. *Indian J Urol.* 2014;30:80-83.
45. Penbegul N, Hatipoglu NK, Bodakci MN, et al. Role of ultrasonography in percutaneous renal access in patients with renal anatomic abnormalities. *Urology.* 2013;81:938-948.
46. Ooster PJ, Razvi H, Liatsikos E, et al; Croes PCNL Study Group. Percutaneous nephrolithotomy among patients with renal anomalies: patient characteristics and outcomes; a subgroup analysis of the clinical research office of the endourological society global percutaneous nephrolithotomy study. *J Endourol.* 2011;25:1627-1632.
47. Miller NL, Matlaga BR, Handa SE, et al. The presence of horseshoe kidney does not affect the outcome of percutaneous nephrolithotomy. *J Endourol.* 2008;22:1219-1225.
48. Rais-Bahrami S, Friedlander JJ, Duty BD, et al. Difficulties with access in percutaneous renal surgery. *Ther Adv Urol.* 2011;3:59-68.
49. de la Rosette JJ, Tsakiris P, Ferrandino MN, et al. Beyond prone position in percutaneous nephrolithotomy: a comprehensive review. *Eur Urol.* 2008;54:1262-1269.
50. Leavitt DA, Motamedinia P, Moran S, et al. Can activities of daily living predict for complications following percutaneous nephrolithotomy [published online December 22, 2015]? *J Urol.* doi: 10.1016/j.juro.2015.12.076.
51. Donally CJ 3rd, Gupta A, Bensalah K, et al. Longitudinal evaluation of the SF-36 quality of life questionnaire in patients with kidney stones. *Urol Res.* 2011;39:141-146.
52. Sfoungaristos S, Gofrit ON, Yutkin V, et al. Evaluating parameters affecting fluoroscopy time during percutaneous nephrolithotomy: focus on the predictive role of Guy's, S.T.O.N.E., and CROES scoring systems. *J Endourol.* 2015;29:1366-1370.
53. Rabah DM, Alomar M, Binsaleh S, Arafa MA. Health related quality of life in ureteral stone patients: post-ureterolithiasis. *Urol Res.* 2011;39:385-388.
54. Bryant M, Angell J, Tu K, et al. Health related quality of life for stone formers. *J Urol.* 2012;188:436-440.
55. Angell J, Bryant M, Tu H, et al. Association of depression and urolithiasis. *Urology.* 2012;79:518-525.
56. Bensalah K, Tuncel A, Gupta A, et al. Determinants of quality of life for patients with kidney stones. *J Urol.* 2008;179:2238-2243.
57. Penniston KL, Nakada SY. Development of an instrument to assess the health related quality of life of kidney stone formers. *J Urol.* 2013;189:921-930.
58. Resorlu B, Unsal A, Gulec H, Oztuna D. A new scoring system for predicting stone-free rate after retrograde intrarenal surgery: the "resorlu-unsal stone score". *Urology.* 2012;80:512-518.
59. Kanao K, Nakashima J, Nakagawa K, et al. Preoperative nomograms for predicting stone-free rate after extracorporeal shock wave lithotripsy. *J Urol.* 2006;176(4 Pt 1):1453-1456.
60. Molina WR, Kim FJ, Spendlove J, et al. 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;40:23-29.