Research Article

Comparison Of Guy's Stone Score, Stone Score, And Seoul National University Renal Stone Complexity Score On CT Scan As Predictors Of Stone-Free Rate After Percutaneous Nephrolithotomy: A Retrospective Study

Hammam Muhammad Al Afghani[1*], M. Hidayat Surya Atmaja[1], Widiana Ferriastuti[1]

[1] Department of Radiology, Faculty of Medicine, Universitas Airlangga, RSUD DR. Soetomo Surabaya, Indonesia

Abstract

Background: Percutaneous nephrolithotomy (PCNL) is the primary treatment of choice for large (>2 cm) or complex renal calculi. Procedural success is highly dependent on preoperative factors, which can be evaluated via non-contrast computed tomography (NCCT). Several CT-based predictive scoring systems have been developed, including the Guy's Stone Score (GSS), STONE Score, and Seoul National University Renal Stone Complexity Score (S-ReSC). Direct comparative studies evaluating the performance of all three scoring systems simultaneously in the same population remain limited, particularly in Indonesia. **Objective:** To compare the predictive ability of GSS, STONE Score, and S-ReSC in predicting the stone-free rate (SFR) following PCNL at Dr. Soetomo General Hospital. **Methods:** This was a retrospective observational analytic study involving 44 patients who underwent unilateral single-tract PCNL at Dr. Soetomo General Hospital between January and December 2025, who met the inclusion and exclusion criteria. The independent variables were the GSS, STONE Score, and S-ReSC assessed on preoperative NCCT. The dependent variable was SFR (defined as residual fragments <4 mm) evaluated via ultrasonography and/or plain kidney-ureter-bladder (KUB) radiography postoperatively. Data were analyzed using the Mann-Whitney test, bivariate and multivariate logistic regression, and receiver operating characteristic (ROC) curve analysis to determine the area under the curve (AUC). **Results:** The overall SFR was 50.0% (22/44 patients). Bivariate analysis demonstrated significant associations with SFR for all three scores (GSS p=0.013; STONE p=0.004; S-ReSC p<0.001). In the multivariate model, only S-ReSC remained a significant independent predictor (OR=2.32; p=0.025; 95% CI: 1.11–4.85). The S-ReSC yielded the highest AUC (0.840; 95% CI: 0.709–0.947), compared to the STONE Score (0.777; 95% CI: 0.625–0.902) and GSS (0.748; 95% CI: 0.602–0.874). The optimal cut-off value for S-ReSC was ≥5, providing a sensitivity of 81.8% and specificity of 81.8%. **Conclusion:** The S-ReSC demonstrated superior predictive performance compared to the GSS and STONE Score in predicting SFR after PCNL and was the only independent predictor in the multivariate model. An S-ReSC cut-off of ≥5 can serve as a valuable clinical reference for preoperative risk stratification and procedural planning.

Keywords: Percutaneous Nephrolithotomy, Guy's Stone Score, STONE Score, S-ReSC, Stone-Free Rate.

1 Introduction

Urinary tract stone disease (urolithiasis) is one of the most commonly encountered urological conditions globally [1]. The 2021 Global Burden of Disease study recorded a significant increase of 26.7% to 106 million new cases worldwide since 2000 [2]. In Indonesia, the average prevalence of kidney stones is around 0.6% based on 2013 Basic Health Research (Riskesdas) data. In its management, there has been a paradigm shift from open surgery toward minimally invasive procedures, in which Percutaneous Nephrolithotomy (PCNL) has become the primary treatment of choice for large (>2 cm) or complex kidney stones due to its high effectiveness in achieving a stone-free rate (SFR) [3].

Evaluation of SFR (4 mm) performed four weeks after the PCNL procedure is the best modality for detecting residual fragments [4]. SFR evaluation can be performed via non-contrast CT scan, ultrasonography, and plain kidney ureter bladder (KUB) radiography, with ultrasonography and KUB having sensitivity in detecting stones ≥4 mm [5]. The success of the procedure is highly dependent on preoperative factors, including stone characteristics and kidney anatomy that can be evaluated through non-contrast CT scan as the current gold standard imaging modality [6], [7], [8]. Various CT-based predictive scoring systems have been developed to estimate PCNL outcomes, including the Guy's Stone Score (GSS), STONE Score, and Seoul National University Renal Stone Complexity Score (S-ReSC) [9], [10].

The STONE nephrolithometry score system analyzes five parameters on non-contrast CT scan, namely stone size (S), skin-to-stone distance (T), degree of obstruction (O), number of involved calyces (N), and stone density (E) [11]. Meanwhile, the Guy's Stone Score (GSS) classifies kidney stones into four grades (I to IV) based on characteristics such as number, location within the calyceal system, and the presence or absence of renal anatomical abnormalities [12]. The S-ReSC scoring system, on the other hand, counts the number of anatomical locations in the pelvicalyceal system occupied by stones. The pelvicalyceal system is divided into 9 potential locations consisting of the renal pelvis (#1), the superior (#2) and inferior (#3) major calyx groups, and six minor calyx groups divided into anterior and posterior divisions in each of the superior (#4 and #5), middle (#6 and #7), and inferior (#8 and #9) calyx groups. This scoring system is expected to be able to predict the success rate and technical difficulty of the PCNL procedure [13].

Direct comparative studies examining the performance of the three scoring systems above simultaneously in the same population remain limited, particularly in Indonesia. RSUD Dr. Soetomo, as a tertiary referral center with a high volume of PCNL cases, provides a unique opportunity to conduct such a comparative evaluation among Indonesian patient population [14], [15], [16].

Based on this background, this study aims to examine the comparison of GSS, STONE Score, and S-ReSC as predictors of SFR after PCNL at RSUD Dr. Soetomo. The results of this study are expected to provide theoretical benefits for the development of urological science, as well as practical benefits as an evidence-based guide for clinicians in preoperative planning, patient risk stratification, pre-procedure counseling, and improvement of healthcare service quality at tertiary referral institutions.

2 Methods

2.1 Type of Research

This study is an observational analytic study with a retrospective design. The study was conducted using secondary data derived from patient medical records to analyze the relationships and predictive ability of several kidney stone complexity scoring systems on the stone-free rate (SFR) in patients undergoing Percutaneous Nephrolithotomy (PCNL).

2.2 Research Population and Sample

The study population consisted of all kidney stone patients who underwent PCNL procedures at RSUD Dr. Soetomo. The study sample consisted of patients who underwent PCNL procedures during the period January 1, 2025 to December 31, 2025 and met the established inclusion and exclusion criteria. Inclusion criteria included patients undergoing single-tract and unilateral PCNL procedures, aged over 18 years, having undergone non-contrast abdominal CT scan prior to the surgical procedure, and having urological ultrasonography and/or plain BOF X-ray results as postoperative evaluation. Exclusion criteria involved incomplete medical records, history of nephrocalcinosis, history of surgery on the ipsilateral kidney, active urinary tract infection, and history of polycystic kidney disease.

2.3 Sample Size

The study sample size was calculated using the correlation research formula to determine the minimum sample number. The calculation was based on a significance level (α) of 0.05, study power (β) of 0.10, and expected correlation coefficient (r) of 0.55 based on a previous study by I Dewa Gede Reza Sanjaya et al. [17]. Based on this calculation, a minimum sample requirement of 31 study subjects was obtained.

2.4 Research Variables

The independent variables in this study included the Guy's Stone Score (GSS), Stone Score, and Seoul National University Renal Stone Complexity Score (S-ReSC). Meanwhile, the dependent variable was the stone-free rate (SFR) after PCNL treatment.

2.5 Research Instruments

The instruments used in this study included a 128-slice CT scan Philips Ingenuity MRC 880 series, a 128-slice CT scan Toshiba TSX-101A, a Philips ClearVue 550 ultrasonography unit, and a Philips CombiDiagnost R90 RF fluoroscopy X-ray machine. All of these instruments were used as means to assess kidney stone conditions before the procedure and to evaluate treatment outcomes after the PCNL procedure.

2.6 Data Collection

Data collection was carried out retrospectively using secondary data derived from non-contrast abdominal CT scan results, urological ultrasonography, and patient Buik Overzicht Foto (BOF) X-ray images. The entire measurement and interpretation process was carried out under the supervision of a consultant abdominal radiologist to ensure the validity of the study results.

2.7 Data Analysis

Data analysis was performed using IBM SPSS version 26 software to evaluate the ability of the Guy's Stone Score (GSS), Stone Score, and S-ReSC to predict the stone-free rate (SFR) in patients undergoing PCNL at RSUD Dr. Soetomo. The independent variables, namely the GSS, Stone Score, and S-ReSC scores, had an ordinal scale, while the dependent variable, namely the stone-free rate (SFR), had a nominal scale with two categories.

The analysis began with a univariate analysis to describe the characteristics of the study data. Variables with nominal and ordinal scales were presented as frequency distributions and percentages, while numerical variables were presented as means and standard deviations if normally distributed, or medians with ranges if not normally distributed.

Next, bivariate analysis was conducted using the Mann-Whitney test to compare the GSS, Stone Score, and S-ReSC scores against the stone-free rate (SFR). In addition, logistic regression analysis was performed to identify the relationship between the independent variables and the stone-free rate. Variables with a p value < 0.250 in the initial analysis were included in the logistic regression analysis, while the level of statistical significance for other tests was set at $p < 0.05$.

Multivariate analysis was conducted using logistic regression to model the relationship between the GSS, Stone Score, and S-ReSC scores as the main predictors of the stone-free rate (SFR). Before the analysis was conducted, a multicollinearity assumption check was performed, while model feasibility was evaluated using the Hosmer-Lemeshow test to assess model fit with the data. All statistical tests in the multivariate analysis used a significance level of $p < 0.05$.

2.8 Research Ethics

This study was conducted after obtaining ethical approval from the Ethics Committee of the Faculty of Medicine, RSUD Dr. Soetomo Surabaya. All data used were secondary data from patient medical records and were kept confidential in accordance with research ethics principles and applicable regulations.

3 Results

3.1 Descriptive Demographic Data

This study is a retrospective observational analytic study. The study population consisted of kidney stone patients at RSUD Dr. Soetomo who underwent percutaneous nephrolithotomy procedures. The study sample consisted of patients who underwent percutaneous nephrolithotomy procedures at RSUD Dr. Soetomo during the period January 1, 2025 to December 31, 2025 who met the inclusion and exclusion criteria, resulting in a sample size of 44 samples. From the 44 samples, descriptive data on general characteristics/demographics were examined, namely: Gender, Age, Education, Occupation, and clinical data, namely: Procedure, Evaluation (BOF/USG), and comorbidity. The following is descriptive data on both general and clinical demographic characteristics of the study, presented as frequency and percentage for categorical data and range, median, mean, and standard deviation for numerical data:

Table 1: Descriptive Characteristic Data

Characteristic	Frequency (n=44)	Range (Median)	Mean ± SD
General Characteristics			
Gender			
Male	24 (54.5%)		
Female	20 (45.5%)		
Age	29 – 71 (56.0)	52.39 ± 12.949	
Education			
No formal education	2 (4.5%)		
Elementary school	10 (22.7%)		
Junior high school	7 (15.9%)		
Senior high school	9 (20.5%)		
Diploma	2 (4.5%)		
Bachelor's degree	14 (31.8%)		
Occupation			

Characteristic	Frequency (n=44)	Range (Median)	Mean ± SD
Housewife/Unemployed	13 (29.5%)		
Private employee	11 (25.0%)		
Self-employed	11 (25.0%)		
Civil servant/Military/Police/Health worker	9 (20.5%)		
Clinical Characteristics			
Procedure			
PCNL D	23 (52.3%)		
PCNL S	21 (47.7%)		
PostOps Evaluation			
BOF	36 (81.8%)		
USG	1 (2.3%)		
BOF & USG	7 (15.9%)		
Comorbidity			
Type 2 diabetes mellitus	1 (2.3%)		
HT	5 (11.4%)		
CHD	0 (0%)		
CKD	2 (4.5%)		
HN	8 (18.2%)		

Based on Table 1, the characteristics of the 44 respondents in this study are described below. Based on general characteristics, gender distribution showed a fairly balanced distribution between genders, although dominated by males with 24 people (54.5%), while females numbered 20 people (45.5%). This indicates that kidney stone cases undergoing percutaneous nephrolithotomy procedures in this study occurred slightly more frequently in males compared to females. Based on patient age, the range was 29 to 71 years with a median age of 56.0 years, and a mean age of 52.39 ± 12.949 years. This data shows that the majority of patients were in the adult to elderly age group, which clinically is indeed the age group with higher risk of experiencing kidney stones due to the accumulation of metabolic and lifestyle factors. Based on education level, respondent distribution was quite varied. The group with a bachelor's degree was the most common, with 14 people (31.8%), followed by elementary school graduates with 10 people (22.7%), senior high school with 9 people (20.5%), and junior high school with 7 people (15.9%). Meanwhile, respondents with no schooling and diploma education each numbered 2 people (4.5%). This variation indicates that the incidence of kidney stones is not limited to a particular education level, although the largest proportion was found in the group with higher education. Based on occupation, the housewife or unemployed group was the most dominant with a total of 13 people (29.5%). Next, private employees and self-employed individuals each numbered 11 people (25.0%), and the civil servant/military/police/health worker group numbered 9 people (20.5%). This distribution reflects that kidney stone conditions can occur across various occupational backgrounds, without overly striking dominance in any particular sector.

In terms of clinical characteristics, the type of percutaneous nephrolithotomy procedure was divided into two relatively balanced groups, namely PCNL D with 23 patients (52.3%) and PCNL S with 21 patients (47.7%). In postoperative evaluation, most patients were evaluated using BOF, with 36 people (81.8%), followed by the combination of BOF and USG with 7 people (15.9%), and only 1 patient (2.3%) using USG alone. Based on comorbidities, hydronephrosis was the most frequently found condition, occurring in 8 patients (18.2%), followed by hypertension in 5 patients (11.4%), and chronic kidney disease in 2 patients (4.5%). Type 2 diabetes mellitus was found in 1 patient (2.3%), while no cases of coronary heart disease were found.

3.2 Descriptive Data on Research Variables

In this study, the research variables included the independent variables, namely the Guy's Stone Score, Stone Score, and Seoul National University Renal Stone Complexity Score (S-ReSC), and the dependent variable, namely the stone-free rate (SFR).

3.3 Descriptive Data on Guy's Stone Score

The Guy's Stone Score is a score used to categorize the complexity of kidney stones based on size, location, and number of stones, in which the assessment of stone size, location, and number is performed by a radiologist and divided into 4 grades, namely Grade I, Grade II, Grade III, and Grade IV. The following is descriptive data on the Guy's Stone Score:

Table 2: Descriptive Data on Guy's Stone Score

Guy's Stone Score Grade	Frequency (n=44)
Grade I	5 (11.4%)
Grade II	14 (31.8%)
Grade III	16 (36.4%)
Grade IV	9 (20.5%)

Based on Table 2, the distribution of Guy's Stone Score among 44 patients shows a fairly diverse variation in kidney stone complexity levels. The analysis results show that most patients fall into the moderate to high complexity category. The Grade III group was the most dominant, with 16 patients (36.4%), indicating that the majority of patients had kidney stones with fairly high complexity levels, such as involvement of multiple locations or stone characteristics that are more difficult to manage. Next, Grade II ranked second with 14 patients (31.8%), representing moderate stone complexity. At the highest complexity level, Grade IV, there were 9 patients (20.5%). This indicates that around one-fifth of the total sample had very complex kidney stone conditions, which clinically could pose greater challenges in percutaneous nephrolithotomy procedures and could affect the success of the stone-free rate. Meanwhile, Grade I, which represents low complexity, was found in only 5 patients (11.4%), showing that the proportion of patients with simple stone conditions is relatively small in this study population. Overall, the descriptive picture shows that most patients in this study fall into the Guy's Stone Score Grade II to Grade IV category, indicating a dominance of cases with moderate to high kidney stone complexity.

3.4 Descriptive Data on Stone Score

The Stone score uses five main parameters obtained from non-contrast CT scan examination, namely stone size (S), skin-to-stone distance (T), degree of obstruction (O), number of involved calyces (N), and stone density in Hounsfield Units (E), assessed by a radiologist and divided into 3 risk levels, namely 1 = low risk (score 0-5), 2 = moderate risk (score 6-9), and 3 = high risk (score 10-13). The following is descriptive data on the Stone score:

Table 3: Descriptive Data on Stone Score

Stone Score	Descriptive (n=44)
Range (Median)	6 – 13 (9.00)
Mean ± SD	8.64 ± 1.806
Low Risk	0 (0%)
Moderate Risk	30 (68.2%)
High Risk	14 (31.8%)

Based on Table 3, the distribution of Stone Score among 44 patients shows a picture of kidney stone complexity levels tending to fall into the moderate to high risk category. Descriptively, the Stone Score value in this study ranged from 6 to 13 with a median value of 9.00. The mean score obtained was 8.64 with a standard deviation of 1.806, indicating that patients generally had scores concentrated at moderate values with relatively moderate variation. The median value of 9 further reinforces that most patients were at the upper boundary of the moderate risk category.

Based on risk categorization, no patients with low risk were found (0%), indicating that there were no cases with very simple stone complexity in this study population. Most patients fell into the moderate risk category, with 30 people (68.2%), reflecting that the majority of patients had stone characteristics with moderate complexity based on the combination of parameters assessed. Meanwhile, 14 patients (31.8%) fell into the high risk category, indicating a fairly significant proportion of patients with complex kidney stone conditions that are potentially more difficult to manage.

3.5 Descriptive Data on S-ReSC Score (Seoul National University Renal Stone Complexity Score)

The Seoul National University Renal Stone Complexity Score (S-ReSC) system counts the number of anatomical locations in the pelvicalyceal system occupied by stones, without considering the size or number of stones at each location, in which the score is then categorized into three risk categories: low score group (1-2), moderate score group (3-4), and high score group (5-9). The following is descriptive data on the Seoul National University Renal Stone Complexity Score (S-ReSC):

Table 4: Descriptive Data on S-ReSC

S-ReSC	Descriptive (n=44)
Range (Median)	1 – 9 (4.5)
Mean ± SD	4.70 ± 2.247
Low Risk	9 (20.5%)
Moderate Risk	13 (29.5%)
High Risk	22 (50.0%)

Based on Table 4, the distribution of Seoul National University Renal Stone Complexity Score (S-ReSC) among 44 patients shows a fairly wide variation in kidney stone complexity levels, with a tendency toward dominance in the high risk category. Descriptively, the S-ReSC score in this study ranged from 1 to 9 with a median value of 4.5. The mean score recorded was 4.70 with a standard deviation of 2.247, indicating fairly large data variation among patients. The median value, which falls between the moderate and high categories, indicates that most patients had involvement of more than several pelvicalyceal locations.
Based on risk categorization, the high risk group was the most dominant, with 22 patients (50.0%). This indicates that half of the study population had kidney stones distributed across many anatomical locations, which clinically reflects high complexity and potential difficulty in the process of thorough stone clearance. Next, the moderate risk group was found in 13 patients (29.5%), indicating moderate involvement of location numbers. Meanwhile, the low risk group numbered 9 patients (20.5%), representing limited stone location involvement.

3.6 Descriptive Data on Stone Free Rate (SFR)

Stone Free Rate (SFR) refers to a patient's condition without residual kidney stones with a total size <4 mm after PCNL within a time span of <4 weeks, evaluated using USG and BOF X-ray, with measurement categorized as follows: <4 mm = Yes; >4 mm = No (Residual). The following is descriptive data on the Stone Free Rate (SFR):

Table 5: Descriptive Data on Stone Free Rate (SFR)

Stone Free Rate (SFR)	Frequency (n=44)
Yes (< 4 mm)	22 (50.0%)
No (> 4 mm)	22 (50.0%)

Based on Table 5, the distribution of Stone Free Rate (SFR) among 44 patients shows balanced results between success and failure in achieving stone-free condition after percutaneous nephrolithotomy treatment. Stone Free Rate in this study is defined as a condition in which no residual kidney stones are found, or only residual fragments with a size of less than 4 mm are present within less than 4 weeks after the procedure, evaluated using ultrasonography and BOF X-ray. Descriptive results show that 22 patients (50.0%) successfully achieved stone-free condition, that is, with residual stones less than 4 mm. Meanwhile, the same number, 22 patients (50.0%), did not achieve stone-free condition, meaning residual stones larger than 4 mm remained.

3.7 Comparative Test Analysis Based on Stone Free Rate (SFR)

The comparative bivariate analysis conducted in this study aims to determine whether significant differences exist in the Guy's Stone Score, the Stone Score, and the Seoul National University Renal Stone Complexity Score (S-ReSC) according to the occurrence of a stone-free rate (SFR), categorized as either present or absent.

Table 6: Comparative Test (Bivariate) Based on Stone Free Rate (SFR) Occurrence

Variable	Stone Free Rate (SFR)		p value
	Present (< 4 mm) (n=22)	Absent (> 4 mm) (n=22)	
Derajat Guy's Stone Score			
Grade I	5 (22.7%)	0 (0%)	0.013
Grade II	8 (36.4%)	6 (27.3%)	
Grade III	8 (36.4%)	8 (36.4%)	
Grade IV	1 (4.5%)	8 (36.4%)	

Stone Score			
Low Risk	0 (0%)	0 (0%)	0.004
Moderate Risk	20 (90.9%)	10 (45.5%)	
High Risk	2 (9.1%)	12 (54.5%)	
S-ReSC			
Low Risk	8 (36.4%)	1(4.5%)	0.001
Moderate Risk	10 (45.5%)	3 (13.6%)	
High Risk	4 (18.2%)	18 (81.8%)	

*Comparative test is stated as significantly different at p < 0.05

Based on Table 6, the results of the comparative test analysis using the Mann-Whitney test show significant differences between the Guy's Stone Score, Stone Score, and Seoul National University Renal Stone Complexity Score (S-ReSC) with regard to the occurrence of Stone Free Rate (SFR), with all p values < 0.05.

For the Guy's Stone Score variable, a p value of 0.013 was obtained (p<0.05), indicating a significant difference between the group of patients who achieved stone-free status and those who did not. Data distribution shows that in the stone-free group there were no patients with Grade I, while in the non-stone-free group there were 0 patients (0%). For Grade II, the distribution was relatively similar between the two groups, with 8 patients (36.4%) in the stone-free group and 6 patients (27.3%) in the non-stone-free group. For Grade III, the distribution was the same in both groups, at 8 patients (36.4%) each. However, for Grade IV, a marked difference was observed, where the non-stone-free group had a higher proportion, 8 patients (36.4%), compared to only 1 patient (4.5%) in the stone-free group. This indicates a difference in the distribution of stone complexity levels between the two groups that contributes to SFR outcomes.

For the Stone Score variable, a p value of 0.004 was obtained (p<0.05), indicating a significant difference between the stone-free and non-stone-free groups. No patients with low risk were found in either group. In the moderate risk category, the stone-free group was more dominant with 20 patients (90.9%) compared to 10 patients (45.5%) in the non-stone-free group. Conversely, in the high risk category, the non-stone-free group was larger, with 12 patients (54.5%) compared to only 2 patients (9.1%) in the stone-free group. This distribution shows a statistically significant difference in risk patterns between the two groups.

For the S-ReSC variable, a p value of 0.001 was obtained (p<0.05), indicating a highly significant difference between the stone-free and non-stone-free groups. In the low risk category, the stone-free group was larger, with 8 patients (36.4%) compared to only 1 patient (4.5%) in the non-stone-free group. In the moderate risk category, the stone-free group was also more dominant, with 10 patients (45.5%) compared to 3 patients (13.6%) in the non-stone-free group. Conversely, in the high risk category, the non-stone-free group was very dominant, with 18 patients (81.8%) compared to only 4 patients (18.2%) in the stone-free group. This indicates that the distribution of complexity based on S-ReSC differs significantly between the two groups.

Based on the test results, it is shown that there are statistically significant differences in all three scoring systems with regard to the occurrence of Stone Free Rate. This finding indicates that the Guy's Stone Score, Stone Score, and S-ReSC play an important role as parameters in distinguishing the likelihood of achieving stone-free condition after PCNL treatment.

3.8 Analysis of Effect on Stone Free Rate (SFR)

The effect test analysis in this study examines whether the Guy's Stone Score, Stone Score, and Seoul National University Renal Stone Complexity Score (S-ReSC) have an effect or impact on the occurrence of Stone Free Rate (SFR). Effect analysis was conducted using bivariate and multivariate approaches with logistic regression tests.

Table 7: Effect Test on Stone Free Rate (SFR) Occurrence

Variable	Bivariate			Multivariate		
	β	p value	Exp(B) (CI95%)	β	p value	Exp(B) (CI95%)
Guy's Stone Score	1.21	0.005	3.35 (1.44 – 7.77)	-1.05	0.276	0.35 (0.005 – 2.32)
Stone Score	0.67	0.004	1.96 (1.24 – 3.10)	0.47	0.222	1.60 (0.75 – 3.40)
S-ReSC	0.71	0.002	2.03 (1.31 – 3.14)	0.84	0.025	2.32 (1.11 – 4.85)

*Test is stated as significant at p < 0.05

Based on Table 7, the results of the effect test analysis on the occurrence of Stone Free Rate (SFR) using logistic regression show differences in findings between the bivariate and multivariate analyses for the three variables studied, namely the Guy's Stone Score, Stone Score, and Seoul National University Renal Stone Complexity Score (S-ReSC).

In the bivariate analysis, all three variables showed a significant effect on the occurrence of SFR with p values < 0.05. The Guy's Stone Score had a coefficient β value of 1.21 with a p value of 0.005 and an Exp(B) value of 3.35 (CI95%: 1.44–7.77). This indicates that an increase of one complexity level in the Guy's Stone Score increases the odds of a given SFR outcome by approximately 3.3 times. Next, the Stone Score showed a significant effect with a β value of 0.67, p value of 0.004, and Exp(B) of 1.96 (CI95%: 2.240–64.285), meaning an increase in score is related with an increase in odds by up to 1.9 times. Meanwhile, the S-ReSC also showed a significant effect with a β value of 0.71, p value of 0.002, and Exp(B) of 2.03 (CI95%: 2.486–28.574), indicating an increase in odds of approximately 2 times.

In the multivariate analysis, which considers the effect of all three variables simultaneously, only the S-ReSC remained significantly associated with the occurrence of SFR. The S-ReSC had a β value of 0.84 with a p value of 0.025 and an Exp(B) of 2.32 (CI95%: 1.11–4.85), indicating that patients with higher S-ReSC scores had approximately 2.3 times greater odds of a given SFR outcome compared to the reference group, after controlling for other variables. The Guy's Stone Score no longer showed a significant effect in the multivariate analysis, with a β value of -1.05, p value of 0.276, and Exp(B) of 0.35 (CI95%: 0.005–2.32). This indicates that after controlling for other variables, the effect of the Guy's Stone Score became statistically non-significant. Similarly, the Stone Score showed a fairly strong tendency of effect with a β value of 0.47 and Exp(B) of 1.60 (CI95%: 0.75–3.40), but was not statistically significant, with a p value of 0.222 slightly above the significance threshold.

3.9 Discriminative Ability Analysis with ROC Curve

Table 8: AUC Values and Best Cut-off Points for Each Score

Variable	AUC	95% CI	Best Cut-off	Sensitivity	Specificity
Guy's Stone Score	0.748	0.602–0.874	≥ 4	36.4%	95.5%
STONE Score	0.777	0.625–0.902	≥ 10	54.5%	90.9%
S-ReSC Score	0.840	0.709–0.947	≥ 5	81.8%	81.8%

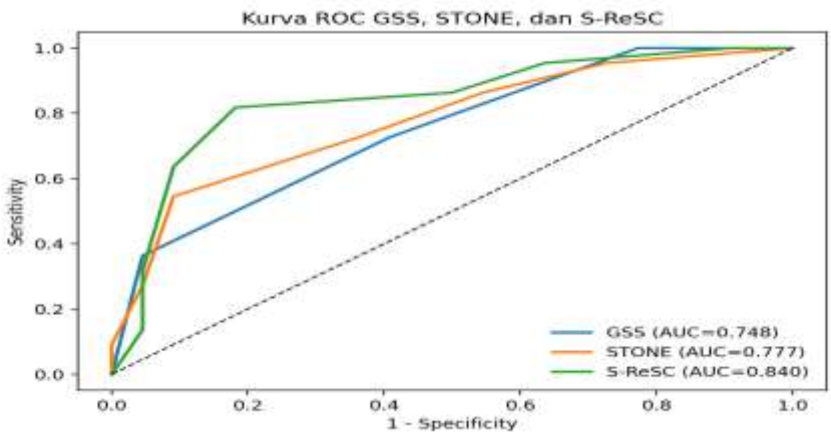

Figure 1: ROC Curve of Guy's Stone Score, STONE Score, and S-ReSC Score

The highest AUC value was obtained by S-ReSC (0.840), followed by STONE (0.777) and GSS (0.748). The best cut-off points based on the Youden index were ≥4 for GSS, ≥10 for STONE, and ≥5 for S-ReSC. The S-ReSC cut-off of ≥5 provided a sensitivity of 81.8% and specificity of 81.8%, indicating the best balance between the ability to detect patients at risk of residual stones and the ability to identify patients with potential for stone-free outcome. Overall, these findings indicate that the S-ReSC has the best discriminative ability in this dataset of 44 patients.

The examination reveals that although all three scores had a significant individual effect on the occurrence of Stone Free Rate, only the S-ReSC proved to be a significant independent predictor in the multivariate model. This suggests that stone location within the pelvicalyceal system, assessed via S-ReSC, is more significant for achieving stone-free outcomes after PCNL than other complexity factors.

4 Discussion

4.1 General Overview of the Study Population

This study involved 44 patients who underwent PCNL at RSUD Dr. Soetomo. The median patient age was 56 years, with a relatively balanced gender distribution and a low proportion of hydronephrosis. This composition indicates that the study subjects represent an adult population with fairly complex kidney stone cases, as is commonly encountered at tertiary referral hospitals. The dominance of GSS III, the absence of the low risk STONE category, and the high proportion of high risk S-ReSC show that patients admitted to the PCNL service in this study generally had a not-simple stone burden.

The distribution of outcomes shows that 50.0% of patients achieved stone-free status and 50.0% did not. This figure needs to be read in the context of case complexity, because the high proportion of multicalyceal stones, partial or complete staghorn stones, and stone distribution across the pelvicalyceal system can reduce the likelihood of complete stone clearance in a single procedure. Thus, this study population is well-suited to evaluate the usefulness of preoperative scoring systems in estimating the likelihood of procedural success before the procedure is performed.

4.2 Guy's Stone Score as a Predictor of SFR

In this study, the Guy's Stone Score showed a clear relationship with SFR outcome. The median GSS in the stone-free group was lower than in the non-stone-free group, and the Mann-Whitney test showed a significant difference. Categorically, all patients with GSS grade I achieved stone-free status, while the large majority of patients with GSS grade IV did not achieve stone-free status. Bivariate logistic regression also showed that each one-grade increase in GSS increased the odds of non-stone-free status by more than three times.

This finding is consistent with the basic concept of GSS developed by Thomas et al., namely that stone anatomical complexity, dominant stone location, and calyceal system involvement are directly related to the technical difficulty of the PCNL procedure and the likelihood of residual stones after the procedure [12]. Compared to the previous 28-patient draft, the 44-patient dataset in this study produced a stronger statistical signal. This is logical because the distribution of GSS grades becomes wider, making differences between outcome groups easier to detect.

However, in the multivariate model, GSS was no longer significant. This result indicates that some of the information captured by GSS may overlap with STONE and especially S-ReSC. Clinically, GSS remains very useful because it is simple, quick to apply, and easy to understand during preoperative counseling. However, when the main goal is independent prediction after accounting for other scoring systems, GSS in this study was not as strong as S-ReSC. This finding is still consistent with the study by Kumar et al., which showed that GSS remains informative, but its performance may vary according to population and distribution of stone complexity [7].

4.3 STONE Score as a Predictor of SFR

The STONE score also showed good performance in this study. The median STONE score increased from 8 in the stone-free group to 10 in the non-stone-free group, and this difference was significant on the Mann-Whitney test. In bivariate logistic regression, each one-point increase in STONE increased the odds of non-stone-free status by nearly two times. In addition, ROC analysis showed an AUC of 0.777 with the best cut-off point at a score of ≥10, which provided high specificity for identifying patients at risk of residual stones.

This result makes sense because the STONE system combines five important dimensions that are technically relevant to PCNL, namely stone size, tract length, obstruction, number of involved calyces, and essence or stone density. Okhunov et al. [11] developed this system precisely to quantify the level of procedural difficulty in more detail than a simple categorical classification. In this study, the ability of STONE to distinguish outcomes was already clearly evident in the bivariate and ROC analyses.

However, like GSS, STONE also lost significance in the multivariate model. One of the strongest explanations is collinearity, because STONE components such as stone surface area and calyceal involvement are also conceptually reflected in GSS and S-ReSC. Nevertheless, STONE remains relevant in practice because it provides a fairly comprehensive quantitative approach. This finding is consistent with the comparative study by Kumar et al. and also the more recent validation by Chen et al., which shows that STONE is a viable tool for predicting PCNL success, although the best score may differ between populations [7], [18].

4.4 S-ReSC Score as a Predictor of SFR

The S-ReSC showed the best and most consistent performance at every stage of analysis. The median S-ReSC increased markedly from 3.5 in the stone-free group to 6.0 in the non-stone-free group. In bivariate logistic regression, each one-point increase in S-ReSC increased the odds of non-stone-free status by approximately two times. After all three scores were entered together into the multivariate model, only S-ReSC remained significant as an independent predictor. ROC analysis also placed S-ReSC as the score with the highest AUC, namely 0.840, along with the most balanced combination of sensitivity and specificity at the cut-off point of ≥5.

Pathophysiologically and technically, the dominance of S-ReSC is very rational. Unlike systems that place more emphasis on size or general classification, S-ReSC focuses on the number of anatomical locations in the pelvicalyceal system occupied by stones. The wider the distribution of stones within the kidney compartments, the more difficult it is for the procedure to achieve complete clearance, especially if percutaneous access is limited or if fragments are located in hard-to-reach calyces. Thus, S-ReSC is likely closer to the actual mechanism that determines residual stones after PCNL.

The results of this study are highly consistent with the developmental study by Jeong et al., which showed that S-ReSC has good predictive ability for stone-free rate after PCNL [13]. In this study population, S-ReSC even became the sole independent predictor after joint adjustment, indicating that stone anatomical distribution is the strongest determinant of procedural success. This finding is also consistent with recent comparative studies that place S-ReSC as one of the high-performing systems in predicting PCNL outcomes [18], [19].

4.5 Comparison of the Three Scoring Systems

When the three systems are compared directly, the order of performance in this study is S-ReSC > STONE > GSS. This order is consistent when viewed from the strength of the bivariate p-values, the stability of the odds ratio values after adjustment, and the discriminative ability based on AUC. GSS remains useful as a simple classification system, STONE provides the advantage of comprehensive multi-parameter quantitative assessment, while S-ReSC appears strongest in capturing the anatomical aspects that truly determine whether a patient will achieve stone-free status.

Clinically, this result is important for preoperative planning. If the main goal is to estimate the likelihood of stone-free status after PCNL, then S-ReSC can be considered the most informative system in the RSUD Dr. Soetomo population. However, the choice of score should not be exclusive. In daily practice, GSS is useful for quick classification, STONE helps with complexity quantification, and S-ReSC adds specific information regarding the anatomical distribution of stones. Integrating all three can actually improve the quality of counseling, access strategy selection, and anticipation of the need for additional procedures.

Differences in performance between scores can also be influenced by population characteristics, PCNL technique, operator experience, and postoperative evaluation modality. Therefore, the results of this study are most appropriately applied to populations with characteristics similar to this study's sample, namely referral hospital patients with complex stone cases and postoperative stone-free evaluation based on BOF or USG.

4.6 Strengths and Limitations of the Study

This study has several strengths. First, this study directly compares the three most commonly used nephrolithometry scoring systems in the international literature for predicting PCNL success in a single cohort. This direct (head-to-head) comparison provides higher internal validity compared to indirect comparisons between studies using different populations. Second, all 44 patients were included without exclusion for the main variables, minimizing selection bias. Third, statistical analysis was conducted in stages starting from descriptive, bivariate, multivariate, to ROC analysis, thereby providing a thorough evaluation of the predictive ability of each score. Fourth, presentation of individual data in anonymous form increases transparency and facilitates internal and external audit of the analysis results.

Several limitations of this study need to be explicitly stated. First, the retrospective design carries the risk of selection bias and limited control over unrecorded confounding variables, such as operator experience, PCNL technique used (standard vs mini-PCNL), number of access tracts, operative duration, and use of flexible nephroscopy.

Second, the sample size of 44 patients is still relatively limited for a multivariate model with three highly correlated predictors. The events per variable (EPV) ratio in this study is approximately 7.3 (22 events for 3 predictors), which is below the ideal threshold of 10–20 EPV for a stable logistic regression. Third, postoperative stone-free evaluation was not uniform and did not entirely use CT scan. Most used BOF, which has lower sensitivity in detecting small residual fragments or radiolucent stones. This may lead to overestimation of the stone-free rate in some patients.

Fourth, this study was conducted at a single institution, so the generalizability of the results to other populations needs to be considered carefully. Fifth, interobserver variability in calculating nephrolithometry scores was not

formally evaluated; all scores were calculated by a single researcher, which minimizes intra-rater variability but does not test inter-reader reproducibility.
Fifth, this study has not conducted a direct statistical comparison between AUCs using the DeLong method, so the conclusion regarding the superiority of S-ReSC remains exploratory.

4.7 Synthesis of the Results

Based on the analysis of 44 patients, all three nephrolithometry scoring systems, GSS, STONE, and S-ReSC, were associated with stone-free rate outcome after PCNL. All three showed significant differences in bivariate analysis and had fairly good discriminative ability. However, after being analyzed simultaneously in multivariate logistic regression, S-ReSC was the sole independent predictor of non-stone-free outcome. Practically, this finding indicates that the higher the stone complexity on preoperative CT scan, the greater the risk that the patient will not achieve stone-free status after PCNL. Among the three systems compared, S-ReSC appears to be the most stable and most informative in this study population, so it can be considered as the main score for preoperative risk stratification in samples with similar characteristics.

5 Conclusion

Based on the analysis of 44 patients who underwent percutaneous nephrolithotomy (PCNL) procedures at RSUD Dr. Soetomo, it can be concluded that the distribution of study subjects was dominated by patients with moderate to high kidney stone complexity, as shown by Guy's Stone Score (GSS) grade III at 36.4%, STONE score moderate risk category at 68.2%, and S-ReSC high risk category at 50.0%. The overall stone-free rate (SFR) in this study reached 50.0% (22 of 44 patients), which is consistent with the characteristics of the complex kidney stone patient population at a tertiary referral hospital.
In bivariate analysis, the Guy's Stone Score showed a significant relationship with the stone-free rate, with a Mann-Whitney test result of $p = 0.003$, bivariate logistic regression odds ratio (OR) of 3.35, and area under the curve (AUC) of 0.748. However, this relationship was no longer significant in the multivariate analysis ($p = 0.276$), which is thought to be due to overlapping predictive information with other scoring systems. The STONE score also showed a significant relationship with the stone-free rate in bivariate analysis (Mann-Whitney $p = 0.002$; OR = 1.96; AUC = 0.777), but lost statistical significance in the multivariate analysis ($p = 0.222$).
In contrast, the S-ReSC consistently showed a significant relationship with the stone-free rate at every stage of analysis. This scoring system had the lowest p value in the bivariate test (Mann-Whitney $p < 0.001$) and became the sole independent predictor in the multivariate logistic regression model with an adjusted OR of 2.32 ($p = 0.025$). In addition, the S-ReSC also showed the best predictive performance compared to the Guy's Stone Score and the STONE score, as shown by the highest AUC value of 0.840 and the best balance of sensitivity and specificity at the cut-off point of ≥5, each at 81.8%. Thus, in this study population, the S-ReSC is the nephrolithometry scoring system with the best ability to predict the success of the stone-free rate after PCNL treatment.
Academically, a prospective study with a larger sample size, at least 80-100 patients, is needed to confirm the superiority of S-ReSC as a predictor of stone-free rate after PCNL, while also allowing comparison of discriminative ability between scoring systems through the DeLong test with adequate statistical power. A multicenter study involving various referral hospitals in Indonesia also needs to be done so that the study results can be generalized to a broader population with variations in kidney stone characteristics and PCNL techniques. Further, future studies should evaluate interobserver variability in calculating the Guy's Stone Score, STONE score, and S-ReSC in order to assess the reproducibility and reliability of each scoring system in routine clinical practice. The development of a predictive model that integrates operative factors, such as number of access tracts, use of flexible nephroscopy, operative duration, and operator experience, also needs to be considered to determine whether S-ReSC remains the best predictor after adjustment for technical procedural factors.
From a clinical perspective, the results of this study indicate that S-ReSC deserves consideration as the main scoring system in preoperative risk stratification for patients who will undergo PCNL, particularly for patients with complex kidney stones at tertiary referral hospitals. An S-ReSC value of ≥5 can be used as a cut-off point to identify patients with a higher risk of not achieving stone-free condition, so that patients can be given counseling regarding the possible need for additional procedures, such as second-look PCNL, adjuvant shock wave lithotripsy (SWL), or flexible ureteroscopy (URS), as well as undergoing more intensive postoperative evaluation. Nevertheless, the three scoring systems should not be used separately, but rather complementarily, with the Guy's Stone Score for quick classification, the STONE score for comprehensive assessment of stone

characteristics, and S-ReSC as the main predictor of clinical outcome, so that preoperative planning can be carried out more optimally.

From a practical and institutional perspective, RSUD Dr. Soetomo is advised to integrate S-ReSC assessment into the preoperative CT scan reporting format for kidney stone patients. Radiologists are expected to routinely calculate and include the S-ReSC score as part of a structured radiology report so that it can be used by urologists in procedure planning. Further, standardization of the postoperative stone-free evaluation protocol is needed. Given the limitations of plain abdominal X-ray (BOF) in detecting residual stone fragments, the use of low-dose non-contrast CT scan (low-dose NCCT) as a postoperative evaluation modality needs to be considered, especially in patients with high S-ReSC scores who have a greater risk of experiencing residual stones. The development of a kidney stone CT scan reporting template that explicitly includes the Guy's Stone Score, S-ReSC, and STONE score components is also expected to improve reporting consistency among radiologists and strengthen interdisciplinary communication within the Department of Radiology and the Department of Urology.

References

[1] C. Türk et al., "EAU Guidelines on Diagnosis and Conservative Management of Urolithiasis," Eur. Urol., vol. 69, no. 3, pp. 468–474, 2016, doi: 10.1016/j.eururo.2015.07.040.

[2] A. F. Awedew et al., "The global, regional, and national burden of urolithiasis in 204 countries and territories, 2000–2021: a systematic analysis for the Global Burden of Disease Study 2021," eClinicalMedicine, vol. 78, p. 102924, Dec. 2024, doi: 10.1016/j.eclinm.2024.102924.

[3] E. Alma, M. H. Özbilen, A. Altunkol, H. Anıl, and H. Ercil, "Comparison of three different scoring systems in predicting success of retrograde intrarenal surgery in kidney stones larger than 20 millimeters," World J. Urol., vol. 43, no. 1, p. 50, 2024, doi: 10.1007/s00345-024-05415-1.

[4] F. Panthier, V. Gauhar, E. Ventimiglia, J.-L. Kwok, E. X. Keller, and O. Traxer, "Rethinking Stone-free Rates and Surgical Outcomes in Endourology: A Point of View from PEARLS Members," Eur. Urol., vol. 86, no. 3, pp. 198–199, 2024, doi: 10.1016/j.eururo.2024.06.001.

[5] M. I. Gokce, E. Ozden, E. Suer, B. Gulpinar, O. Gulpınar, and S. Tangal, "Comparison of imaging modalities for detection of residual fragments and prediction of stone related events following percutaneous nephrolitotomy," Int. braz j urol, vol. 41, no. 1, pp. 86–90, 2015, doi: 10.1590/s1677-5538.ibju.2015.01.12.

[6] H. Zhao, W. li, J. Li, L. Li, H. Wang, and J. Guo, "Predicting Stone-Free Status of Percutaneous Nephrolithtomy With Machine Learning System: Comparative Analysis With Guy's Stone Score and S.T.O.N.E Score System," 2021, Research Square Platform LLC. doi: 10.21203/rs.3.rs-823907/v1.

[7] U. Kumar et al., "STONE score versus Guy's Stone Score - prospective comparative evaluation for success rate and complications in percutaneous nephrolithotomy," Urol. Ann., vol. 10, no. 1, p. 76, 2018, doi: 10.4103/ua.ua_119_17.

[8] T. Christin, Z. Ali, L. Legiran, and F. Ferawaty, "Overview Of Peripheral Neuropathy In Chronic Kidney Disease Patients Undergoing Hemodialysis At Dr. Mohammad Hoesin Hospital Palembang," Pharmacol. Med. Reports, Orthop. Illn. Details, vol. 2, no. 3, pp. 1–18, 2023, doi: 10.55047/comorbid.v2i3.890.

[9] K. Farooq et al., "Comparison of STONE Score, Guy's Stone Score, CROES Nomogram, and Seoul National University Renal Stone Complexity Score in Prognosticating Outcomes of Multiple-Tract Mini-Percutaneous Nephrolithotomy: A Retrospective Study," Cureus, vol. 16, no. 2, p. e54790, 2024, doi: 10.7759/cureus.54790.

[10] T. Çetin et al., "Guy's, S.T.O.N.E., CROES Nomograms in Percutaneous Nephrolithotomy Can Predict the Stone-Free Rate Similarly: A Retrospective Study of Thousand Patients," J. Urol. Surg., vol. 11, no. 3, pp. 179–186, 2024, doi: 10.4274/jus.galenos.2024.2024-3-11.

[11] Z. Okhunov et al., "S.T.O.N.E. Nephrolithometry: Novel Surgical Classification System for Kidney Calculi," Urology, vol. 81, no. 6, pp. 1154–1160, 2013, doi: 10.1016/j.urology.2012.10.083.

[12] K. Thomas, N. C. Smith, N. Hegarty, and J. M. Glass, "The Guy's Stone Score—Grading the Complexity of Percutaneous Nephrolithotomy Procedures," Urology, vol. 78, no. 2, pp. 277–281, Aug. 2011, doi: 10.1016/j.urology.2010.12.026.

[13] C. W. Jeong et al., "Seoul National University Renal Stone Complexity Score for Predicting Stone-Free Rate after Percutaneous Nephrolithotomy," PLoS One, vol. 8, no. 6, p. e65888, 2013, doi: 10.1371/journal.pone.0065888.

[14] R. Ugalde-Resano, C. I. Villeda-Sandoval, E. Kobashi-Sandoval, J. A. Rivera-Ramírez, M. A. Vargas-Robles, and C. E. Méndez-Probst, "Comparison of the most popular methods for predicting stonefree

rate after percutaneous nephrolithotomy," Rev. Mex. Urol., vol. 77, no. 4, pp. 279–288, 2017, doi: 10.48193/rmu.v77i4.40.

[15] A. Choudhary, S. Jayadeva-Reddy, S. Shetty, B. Sourabh-Reddy, A. Singh, and M. Irappa-Wali, "Evaluation of the STONE nephrolithometry score in predicting surgical outcomes of percutaneous nephrolithotomy: results of a prospective study at a university hospital," Rev. Mex. Urol., vol. 83, no. 3, pp. 1–13, 2023, doi: 10.48193/revistamexicanadeurologa.v83i3.922.

[16] M. R. Al-Azab and B. S. Al-Romeedy, "Servant Leadership and Tourism Businesses' Outcomes: A Multiple Mediation Model," Tour. Rev., vol. 79, no. 1, pp. 184–204, 2023, doi: 10.1108/tr-11-2022-0538.

[17] I. D. G. R. Sanjaya, D. M. Soebadi, T. Djojodimedjo, and B. Soeprijanto, "Comparison Of S.T.O.N.E. Score And Guy Score In Predicting Stone Free Rates Of Percutaneous Nephrolithotomy Procedures In The Soetomo General Hospital," Indones. J. Urol., vol. 28, no. 1, pp. 91–97, Jan. 2021, doi: 10.32421/juri.v28i1.703.

[18] S. Chen et al., "Current Status and Prospects of Pediatric Stone Disease: A Bibliometric and Visualization Study," Cureus, vol. 16, no. 3, p. e56549, Mar. 2024, doi: 10.7759/cureus.56549.

[19] A. M. Al Adl, A. Mohey, A. Abdel Aal, H. A. F. Abu-Elnasr, T. El Karamany, and Y. A. Noureldin, "Percutaneous Nephrolithotomy Outcomes Based on S.T.O.N.E., GUY, CROES, and S-ReSC Scoring Systems: The First Prospective Study," J. Endourol., vol. 34, no. 12, pp. 1223–1228, Dec. 2020, doi: 10.1089/end.2019.0856.