e-CliniC 2025; Vol. 13, No. 3: 433-439
DOI: https://doi.org/10.35790/ecl.v13i3.62803
URL Homepage: https://ejournal.unsrat.ac.id/index.php/eclinic

Comparison of Scoring Systems in Predicting Stone Free Rate (SFR) in Patients with Fluoroless Retrograde Intrarenal Surgery (fRIRS)

Eko Arianto,[1] Christof Toreh,[1] Ari Astram,[1] Frendy Wihono,[1] Bryan P. Panelewen[2]

[1]Division of Urology, Department of Surgery, Faculty of Medicine, Universitas Sam Ratulangi, Manado, Indonesia
[2]Department of Surgery, Faculty of Medicine, Universitas Sam Ratulangi, Manado, Indonesia
Received: June 29, 2025; Accepted: July 22, 2025; Published online: July 25, 2025

Abstract: Between March 2022 and October 2023, a prospective study evaluated 216 patients who underwent Fluoroless Retrograde Intrarenal Surgery (fRIRS), including right, left, and bilateral procedures. The study compared the predictive abilities of Guy's Stone Score (GSS), Seoul National University Renal Stone Complexity Score (S-ReSC), and Resorlu-Unsal Stone Score (RUSS) for stone-free rate (SFR). The patient cohort consisted of 164 men (76.3%) and 52 women (23.7%), with an average age of 57.3 years. The distribution of procedures was 49.07% right fRIRS, 49.5% left fRIRS, and 1.43% bilateral fRIRS. GSS scores were evenly split, with 50% of patients in Grade I and 50% in Grade II. S-ReSC scores were predominantly high (65.3%), with 20.4% moderate and 14.4% low. RUSS scores were distributed as 1 (43 patients), 2 (112 patients), and 3 (61 patients). The overall SFR was 82.1%, indicating a high success rate in eliminating residual stones post-procedure. In conclusion, GSS was the most accurate scoring system for predicting SFR, with an AUC of 75%, compared to S-ReSC (72%) and RUSS (70%). All three scoring systems demonstrated good predictive ability, with AUC values ≥70%. Fluoroless RIRS, a minimally invasive surgical option for kidney stones, has gained popularity due to technological advancements. While GSS showed the highest accuracy, the study highlighted the utility of all three scoring systems in preoperative planning and outcome prediction. These findings underscore the importance of scoring systems in optimizing surgical management and improving patient outcomes in kidney stone treatment.
Keywords*:* stone free rate; retrograde intrarenal surgery; Resorlu-Unsal score; Guy'stone score; Seoul nation university renal stone complexity score

INTRODUCTION

Kidney stones (nephrolithiasis) are the cause of most urinary tract disorders. Typical kidney stones are found in the calyces or pelvis, and when out, they stop and block the ureter (ureter stones) and bladder (bladder stones).[1] In the last three decades, the prevalence of kidney stones has increased, and the incidence of nephrolithiasis in developed and developing countries is different, because the incidence of nephrolithiasis in developed countries is generally asymptomatic, where salt and protein consumption increases, while malnutrition and limited water sources contribute to the increase in nephrolithiasis in developing countries.[2,3]

Operative management of kidney stones is usually performed on stones measuring more than 0.5 cm.[4] As an operative management option for kidney stone surgery fluoroless retrograde intrarenal surgery (fRIRS) has been widely performed in recent decades, along with developments in technology and science.[5] Several research groups have proven the success of fRIRS surgery in treating large kidney stones.[6] In addition to fRIRS, minimally invasive surgery methods have also been developed, such as percutaneous nephrolithotomy (PNCL), which is the gold standard and first-line treatment for large and complex kidney stones.[7]

There are some available scoring systems namely Guy's Stone Score (GSS), Resorlu-Unsal Stone Score (RUSS), Seoul National University Renal Stone Complexity Score (S-ReSC), fRIRS, and S.T.O.N.E.[8-16] Guy's Stone Score is an easy-to-understand, valid and reliable method to describe the stone condition and predict the stone-free rate (SFR).[6] Resorlu-Unsal Stone Score is a simple scoring system that can predict the postoperative SFR after fRIRS. All components of fRIRS are obtained through CT urography to predict the characteristics of stones depicted in the reproductive system. Moreover, the S.T.O.N.E score is a scoring system proposed to predict the patient's SFR based on the preoperative characteristics available on CT-KUB, meanwhile S-ReSC system assessment can predict SFR after PCNL.[7]

Based on the background mentioned, the purpose of this study was to compare the scoring systems of GSS, RUSS, S-ReSC, fRIRS, and S.T.O.N.E as the gold standard for predicting the SFR in patients having kidney stones with RIRS.

METHODS

Between March 2022 and October 2023, 216 patients who underwent fRIRS surgery in the form of right fRIRS, left fRIRS, and bilateral fRIRS, performed by experienced urologists, were prospectively evaluated. This study was approved by the Ethics Committee of Dr. J. H. Awaloei Hospital. Data were collected from all patients with kidney stones undergoing Fluoroless RIRS who met the inclusion and exclusion criteria and were listed in the medical records. The inclusion criteria were patients aged over 18 years, who had complete examination data in the medical records, patients diagnosed with kidney stones with non-contrast urology CT scans who underwent fRIRS. The exclusion criteria were the size of the patient's kidney stone measuring >2 cm, and patients with congenital abnormalities and anatomical abnormalities of the urinary tract. This study used retrospective data collected through medical records and did not require additional patient intervention; therefore, this group of subjects was no longer exposed to risks other than those experienced during the course of treatment, and informed consent was clearly not required.

The ability of GSS, RUSS, S-ReSC, fRIRS, and S.T.O.N.E to predict the SFR rate was determined through receiver operating characteristic (ROC) curve analysis, and the results were expressed as sensitivity, specificity, and accuracy figures based on the optimal cutoff point according to the J Youden index.

RESULTS

Table 1 shows that the current study population consisted of 216 patients with a mean age of 55 years, with a variation of approximately 13 years, including 164 (76.3%) men and 52 (23.7%) women, who underwent fRIRS surgery at Dr. J. H. Awaloei Hospital between March 2022 and October 2023. The procedures used in these samples were the right fRIRS, left fRIRS, and

bilateral fRIRS. Regarding the results of the GSS score assessment, 108 patients (50%) had grade I disease and 108 (50%) had grade II disease. The distribution of the S-ReSC score tended to be centered at a high level with percentages of 65.3% and 20.4% at the medium level, while 14.4% were at a low level. Regarding the RUSS score, 43 patients had a score of 1, 112 patients had a score of 2, and 61 patients had a score of 3. The SFR was defined as the percentage of patients without residual stones after the procedure. In this study, the SFR was 82.1%. Further analysis was conducted on three scoring systems: GSS, RUSS, and S-ReSC. However, the fRIRS and S.T.O.N.E scoring systems could not be used in this study.

Table 1. Research characteristics

Characteristics	n	%	Mean ± SD
Gender			
Men	164	76.3	
Women	52	23.7	
Age (years)			57.3 ± 10.9
Action			
Right fRIRS	106	49,07	
Left fRIRS	107	49.5	
Bilateral fRIRS	3	1.43	
Scoring			
GSS			
I	108	50	
II	108	50	
SReSC			
Low	30	14.3	
Medium	45	20.4	
High	141	65.3	
RUSS			
1	43	19.6	
2	112	51.5	
3	61	28.9	
Stone free rate			
Yes	178	82.1	
No	38	17.9	

Overall, the GSS demonstrated a higher level of accuracy in predicting SFRs compared to the S-ReSC and the RUSS. The GSS prediction results revealed a sensitivity of 76% and a specificity of 75%, with a positive predictive value (PPV) of 93.2% and a negative predictive value (NPV) of 41.1%. These findings suggest that GSS is a reliable tool for identifying patients likely to achieve a stone-free status after surgery. However, it is important to interpret these results with caution, as the analysis was conducted on a relatively limited patient population. The number of patients who achieved a SFR in this study may not provide statistically stable or generalizable results, and further research with larger sample sizes is needed to confirm these findings.

In contrast, the RUSS, which ranges from 1 to 3, showed a less even distribution across the patient population. This uneven distribution may have impacted its predictive accuracy. The sensitivity of RUSS in predicting SFR was 75.4%, while its specificity was 62.2%. The PPV for RUSS was 90%, and the NPV was 36.2%. While these values indicate that RUSS has some utility in predicting surgical outcomes, its uneven score distribution and lower specificity compared to GSS suggest that it may be less reliable in certain clinical scenarios. It is worth noting that these predictive values could vary under different sample conditions or in studies with different patient demographics, highlighting the need for further validation in diverse populations.

Similarly, the S-ReSC score demonstrated a sensitivity of 71% and a specificity of 77.5%,

with a PPV of 93.4% and an NPV of 37.3%. While the specificity of S-ReSC was slightly higher than that of GSS, its sensitivity was lower, indicating that it may be less effective at correctly identifying patients who will achieve a stone-free status. The high PPV of S-ReSC suggests that it is useful for confirming positive outcomes, but its lower NPV indicates limitations in ruling out negative outcomes. Like RUSS, the performance of S-ReSC may vary depending on the study population and sample conditions, and additional research is needed to fully assess its reliability and applicability in clinical practice.

Table 2 provides a detailed analysis of the sensitivity and specificity of the three scoring systems—GSS, S-ReSC, and RUSS—in predicting SFRs among patients who underwent fRIRS. The sensitivity of GSS was found to be 76%, which means that out of the 216 patients who underwent fRIRS, the GSS scoring system correctly identified 167 patients as likely to achieve a stone-free status. However, in 52 cases, the GSS scoring system failed to identify patients who were actually stone-free after the procedure, resulting in false-negative outcomes. The specificity of GSS was 75%, indicating that the scoring system correctly identified 164 patients as not achieving a stone-free status. However, in 55 cases, the GSS scoring system incorrectly classified patients as stone-free when they were not, leading to false-positive results. These findings suggest that while GSS is a relatively accurate tool for predicting SFR, there is still room for improvement in reducing both false negatives and false positives.

In comparison, the S-ReSC scoring system demonstrated a sensitivity of 71%. This means that out of the 216 patients, S-ReSC correctly identified 155 patients as likely to achieve a stone-free status. However, in 64 cases, the S-ReSC scoring system failed to identify patients who were actually stone-free, resulting in false-negative outcomes. The specificity of S-ReSC was 77.5%, indicating that it correctly identified 169 patients as not achieving a stone-free status. However, in 50 cases, the S-ReSC scoring system incorrectly classified patients as stone-free when they were not, leading to false-positive results. While the specificity of S-ReSC was slightly higher than that of GSS, its lower sensitivity suggests that it may be less effective at correctly identifying patients who will achieve a stone-free status.

The RUSS scoring system, on the other hand, demonstrated a sensitivity of 75.4%. This means that out of the 216 patients, RUSS correctly identified 165 patients as likely to achieve a stone-free status. However, in 53 cases, the RUSS scoring system failed to identify patients who were actually stone-free, resulting in false-negative outcomes. The specificity of RUSS was 62.5%, indicating that it correctly identified 136 patients as not achieving a stone-free status. However, in 83 cases, the RUSS scoring system incorrectly classified patients as stone-free when they were not, leading to false-positive results. The lower specificity of RUSS compared to GSS and S-ReSC suggests that it may be less reliable in ruling out negative outcomes, particularly in cases where patients do not achieve a stone-free status.

Table 2. Sensitivity and specificity analysis

Scoring	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
GSS	76	75	93.2	41.1
S-ReSC	71	77.5	93.4	37.3
RUSS	75.4	62.5	90	36.2

Figures 1 illustrates the comparison of the three scoring systems based on the ROC curve. The ROC curve analysis provides a visual representation of the trade-off between sensitivity and specificity for each scoring system. The area under the curve (AUC) for GSS, S-ReSC, and RUSS further highlights their predictive accuracy. GSS, with an AUC of 75%, demonstrated the highest overall accuracy among the three scoring systems, followed by S-ReSC (72%) and RUSS (70%). These results suggest that while all three scoring systems have utility in predicting SFR, GSS may be the most reliable tool for preoperative assessment and outcome prediction in patients undergoing fRIRS.

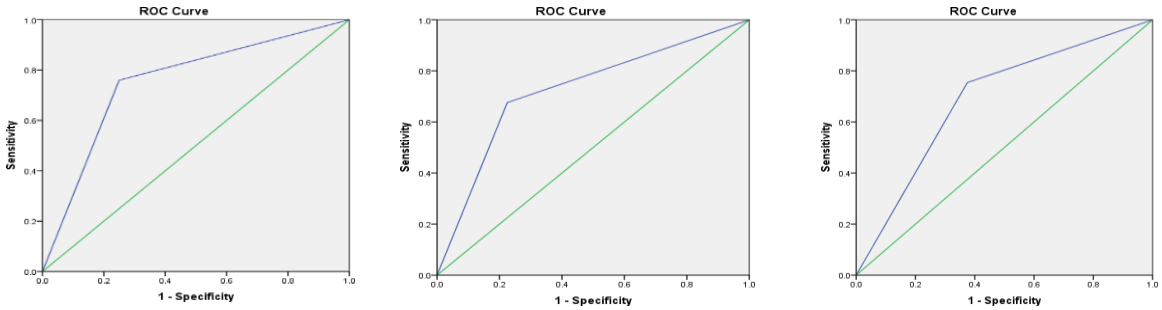

Figure 1. Comparison of the three scoring systems based on ROC, caption: A, ROC curve GSS; B, ROC Curve S-ReSC; C, ROC Curve RUSS

The ability of GSS to predict the SFR was found to be good, with an AUC of 0.75 (75%). An AUC value of 70% or higher is generally considered indicative of a good predictive ability, and GSS comfortably meets this threshold. This suggests that GSS is a reliable tool for estimating the likelihood of achieving a stone-free status after fRIRS. The AUC value reflects the overall accuracy of the scoring system, taking into account both sensitivity and specificity, and the result for GSS underscores its utility in clinical practice for preoperative planning and outcome prediction.

Similarly, the RUSS also demonstrated good predictive ability, with an AUC value of 0.70 (70%). While this is slightly lower than the AUC for GSS, it still meets the ≥70% threshold, indicating that RUSS is a useful scoring system for predicting stone-free rates. The AUC value for RUSS suggests that it has a reasonable balance between sensitivity and specificity, making it a viable option for assessing the complexity of kidney stones and estimating surgical outcomes. However, the slightly lower AUC compared to GSS may indicate that RUSS is less robust in certain clinical scenarios or patient populations, and further research may be needed to optimize its predictive accuracy.

The S-ReSC also performed well, with an AUC of 0.72 (72%). Like GSS and RUSS, S-ReSC meets the ≥70% threshold, confirming its good predictive ability for stone-free rates. The AUC value for S-ReSC indicates that it is a reliable tool for evaluating stone complexity and predicting surgical outcomes, particularly in the context of percutaneous nephrolithotomy (PCNL). However, its slightly lower AUC compared to GSS suggests that it may be less accurate in certain cases, particularly when applied to fRIRS procedures. Despite this, S-ReSC remains a valuable scoring system, especially in cases where stone distribution and anatomical complexity are key factors in determining surgical success.

Table 3 provides a summary of the AUC values and 95% confidence intervals for the three scoring systems, offering a comparative overview of their predictive abilities. The AUC values for GSS (0.75), RUSS (0.70), and S-ReSC (0.72) all fall within the range considered to indicate good predictive performance. However, the differences in AUC values highlight the relative strengths and weaknesses of each scoring system. The GSS, with the highest AUC, appears to be the most accurate and reliable tool for predicting stone-free rates in patients undergoing fRIRS. RUSS and S-ReSC, while still effective, may require further refinement or contextual adjustments to improve their predictive accuracy.

Table 3. Comparison of AUC and 95% confidence intervals of various scoring system

Scoring	AUC	95% confident interval
GSS	0.75	0.66 – 0.84
S-ReSC	0.72	0.64 – 0.81
RUSS	0.70	0.59 – 0.78

DISCUSSION

Several scoring systems and nomograms have been developed using these factors to predict the

success of interventions in kidney stone management, such as GSS, RUSS, S-ReSC, fRIRS, and S.T.O.N.E. In this study, we compared the GSS, RUSS, S-ReSC, fRIRS, and S.T.O.N.E to predict the SFR in post-RIRS patients.[5] We involved 216 patients who underwent RIRS procedures in the form of right fRIRS, left fRIRS, and bilateral fRIRS. The SFR in this study reached 82.1%, compared to the incidence of no SFR, which reached 17.9%. This study found that the GSS score showed a higher level of accuracy (76%) than the S-ReSC and RUSS. These results need to be carefully considered considering that the analysis was conducted on a limited population, and the number of patients who then experienced a SFR has not been able to provide stable statistics. The S-ReSC score was the lowest in this study (71%).

Sensitivity indicates the ability of a test to be positive for a sick person. The higher the sensitivity of a test, the more positive the test results in sick individuals or fewer false negatives. Specificity indicates the ability of a test to state that a non-sick person is negative. The higher the specificity of a test, the more negative the test results in non-sick individuals or fewer false positives. In this study, it was found that GSS, S-ReSC and RUSS scoring had similar abilities in predicting SFR. The GSS prediction results showed sensitivity and specificity of 76% and 75% respectively with PPV 93.2%, NPV 41.1%. The RUSS score was only limited between 1-3 with its distribution not too even so that the sensitivity of the SFR prediction recorded was 75.4%, specificity 62.2% with PPV 90%, NPV 36.2%. While the S-ReSC score appeared to have a sensitivity of 71%, specificity reaching 77.5% with PPV value 93.4%, NPV 37.3%. In this study, it was also found that the ability of GSS in predicting SFR was good with AUC of 0.75 (75%) or ≥70%. Similar to GSS scoring, the RUSS obtained an AUC value of 0.70 (70%) or ≥70%, which indicates that this scoring has good ability in predicting stone free rate. Similar to other scoring systems, the S-ReSC scoring system obtained an AUC of 0.72 (72%) or ≥70%, which indicates that this scoring is good in predicting SFR.

In recent years, predicting SFR and risk of complications before surgery has become quite interesting in the field of endourology. Several nomograms have been developed to predict postoperative success rates, namely SWL, URS, PCNL, and fRIRS. The GSS is a simple and reliable scoring system to predict success rates considering stone location and renal anatomy. Higher scores indicate a low SFR. The SFR is also independent of stone burden, surgeon experience, age, weight, and patient comorbidities. In a study using PCNL, the SFR reached between 0 and 100% for GSS. In the study, the incidence of residual stones was higher in patients with GSS Grade 2.[16] One study revealed that the ROC curve analysis revealed that S-ReSC was the least sensitive scoring system in predicting stone free status, compared to other nomograms.

Another study with fRIRS reported an overall SFR of 59% (92/157) with a complication rate of 22% (35/157). Stone burden <542mm3 was significantly associated with SFR (p=0.001). In univariate analysis, all scoring systems were identified as significant factors in relation to SFR. The GSS, CROES score, and S-ReSC score were associated with complications (p<0.02). Multivariate logistic regression analysis of the study showed that CROES score was identified as a significant factor in terms of SFR and complications (p<0.01). The area under the ROC curve for stone burden, Guy's score, STONE score, CROES score, and S-ReSC score showed good results (0.737, 0.674, 0.762, 0.746, and 0.710, respectively). In addition, another study also revealed that the initial stone-free and complication rates after PCNL were 78.7 (111/141) and 17.0% (24/141). Univariate analysis in a previous study reported that all three scoring systems were identified as significant factors in relation to SFR. Multivariate logistic regression analysis of the study showed that GSS and stone burden ≥385 mm2, which had a significant correlation with stone-free status [odds ratio (OR)=3.220, p=0.001 and OR=6.451, p=0.002, respectively]. Guy's stone score (OR=1.879, p=0.013) was an independent risk factor for the development of complications. The AUCs for GSS, S.T.O.N.E., and CROES and stone burden showed good results (0.821, 0.816, 0.820, and 0.800, respectively). Pairwise comparison of ROC curves in the study showed that there was no significant difference between each final score and stone burden. These results support our findings that GSS has good ability in predicting SFR.

CONCLUSION

Fluoroless Retrograde Intrarenal Surgery (fRIRS) has emerged as a highly effective, minimally invasive option for kidney stone management, leveraging advancements in medical technology to reduce radiation exposure and improve patient outcomes. This study evaluated the predictive accuracy of three scoring systems—Guy's Stone Score (GSS), Seoul National University Renal Stone Complexity Score (S-ReSC), and Resorlu-Unsal Stone Score (RUSS)—for stone-free rates (SFR) post-fRIRS. GSS demonstrated the highest accuracy with an AUC of 75%, outperforming S-ReSC (72%) and RUSS (70%), both of which also showed good predictive performance. While GSS is the most reliable tool for preoperative planning, S-ReSC and RUSS remain valuable alternatives in specific clinical contexts. Future research should focus on validating these systems in diverse populations and enhancing their predictive capabilities to further optimize kidney stone management.

Conflicts of Interest

The authors declare no conflict of interest.

REFERENCES

1. Mayans L. Nephrolithiasis. Prim Care. 2019;46(2):203-12. Doi:10.1016/j.pop.2019.02.001.
2. Ziemba JB, Matlaga BR. Epidemiology and economics of nephrolithiasis. Investig Clin Urol. 2017;58(5):299-306. Doi:10.4111/icu.2017.58.5.299
3. Atmoko W, Savira M, Fajriani R, Wistara SS, Asmarinah, Harahap AR, et al. Stone recurrence among Indonesian kidney stone formers: a comprehensive analysis of genetic polymorphism, demographic, and clinical factors. Med J Indones. 2024;33(3):190-7. Doi:10.13181/mjloa.247596.
4. Schlomer BJ. Urologic treatment of nephrolithiasis. Curr Opin Pediatr. 2020;32(2):288-94. Doi:10.1097/MOP.0000000000000849.
5. Karsiyakali N, Karabay E, Erkan E, Kadihasanoglu M. Evaluation of nephrolithometric scoring systems to predict outcomes of retrograde intrarenal surgery. Urol J. 2020;17(4):352-7. Doi:10.22037/uj.v0i0.5256
6. Chatterjee S, Abhishek A, Samanta K, Ghosh A, Mandal SN, Karmakar D. Evaluation of various nephrometric scoring systems for nephrolithiasis to predict stone-free status after percutaneous nephrolithotomy: a comparative study. Urologia. 2022;89(3):418-23. Doi: 10.1177/03915603211030162
7. Jiang K, Sun F, Zhu J, Luo G, Zhang P, Ban Y, Shan G, Liu C. Evaluation of three stone-scoring systems for predicting SFR and complications after percutaneous nephrolithotomy: a systematic review and meta-analysis. BMC Urol. 2019;19(1):57. Doi: 10.1186/s12894-019-0488-y
8. Weiss B, Shah O. Evaluation of dusting versus basketing – can new technologies improve stone-free rates? Nat Rev Urol. 2016;13(12):726-33. doi: 10.1038/nrurol.2016.172
9. Cheungpasitporn W, Rossetti S, Friend K, Erickson SB, Lieske JC. Treatment effect, adherence, and safety of high fluid intake for the prevention of incident and recurrent kidney stones: a systematic review and meta-analysis. J Nephrol. 2016;29(2):211-19. Doi: 10.1007/s40620-015-0210-4
10. Estrade V, Meria P, Almeras C. 2022 Recommendations of the AFU Lithiasis Committee: Combined approach for the management of kidney and ureteral stones (Endoscopic Combined IntraRenal Surgery, ECIRS). Prog Urol. 2023;33(14):871-4. Doi: 10.1016/j.purol.2023.08.009
11. Koo K, Matlaga BR. New imaging techniques in the management of stone disease. Urol Clin North Am. 2019 ;46(2):257-63. Doi:10.1016/j.ucl.2018.12.007
12. Brisbane W, Bailey MR, Sorensen MD. An overview of kidney stone imaging techniques. Nat Rev Urol. 2016;13(11):654-662. Doi:10.1038/nrurol.2016.154
13. Sfoungaristos S, Gofrit ON, Mykoniatis I, Landau EH, Katafigiotis I, Pode D, et al. External validation of Resorlu-Unsal stone score as predictor of outcomes after retrograde intrarenal surgery. Int Urol Nephrol. 2016;48(8):1247-1252. Doi:10.1007/s11255-016-1311-2
14. Singla A, Khattar N, Nayyar R, Mehra S, Goel H, Sood R. How practical is the application of percutaneous nephrolithotomy scoring systems? Prospective study comparing Guy's Stone Score, S.T.O.N.E. score and the Clinical Research Office of the Endourological Society (CROES) nomogram. Arab J Urol. 2017 12;15(1):7-16. Doi:10.1016/j.aju.2016.11.005
15. Mazzon G, Choong S, Celia A. Stone-scoring systems for predicting complications in percutaneous nephrolithotomy: A systematic review of the literature. Asian J Urol. 2023;10(3):226-38. Doi: 10.1016/j.ajur.2023.01.005
16. Özman O, Akgül HM, Başataç C, Sancak EB, Çınar Ö, Çakır H, Yazıcı CM, Akpınar H, Önal B; RIRSearch Study Group. Recent scoring systems predicting stone-free status after retrograde intrarenal surgery; a systematic review and meta-analysis. Cent European J Urol. 2022;75(1):72-80. Doi:10.5173/ceju.2022.0277