External validation and comparison of current scoring systems in retrograde intrarenal surgery: multi-institutional study with 949 patients

Ibrahim Halil Bozkurt[1], Nihat Karakoyunlu[2], Omer Koras[3], Serdar Çelik[4], Ertugrul Sefik[1], Mehmet Caglar Cakici[5], Tansu Degirmenci[1], and Muhammed Abdurrahim Imamoglu[6]

[1]University of Health Sciences Izmir Bozyaka Education and Research Hospital
[2]Ankara Diskapi Yildirim Beyzat Training and Research Hospital
[3]Mustafa Kemal University Faculty of Medicine
[4]Izmir Bozyaka Training and Research Hospital
[5]Istanbul Medeniyet Universitesi
[6]HSU, Ankara Dışkapı Training and Research Hospital

December 3, 2020

Abstract

Objectives: To externally validate and compare Resorlu-Unsal stone score(RUSS), modified Seoul National University Renal Stone Complexity Score(S-ReSC), Ito's nomogram and R.I.R.S. scoring systems for predicting capabilities of both the stone-free status and complications in a multi-institutional study. Materials and Methods: We performed a retrospective analysis of 949 patients who were underwent flexible ureterorenoscopy (f-URS) and laser lithotripsy for renal stones in two institutions between March-2015 and June-2020. The RUSS, modified S-ReSC, Ito's nomogram and R.I.R.S. scores were calculated for each patient by same surgeon on imaging methods. Results were compared for their predictive capability of stone-free status and complications. Results: Of 949 patients 603 were male and 346 were female with a mean age of 47.2±14.3 (range 2-84years). Mean stone burden was 102.6±42.2 (48-270mm2). All nomograms predicted stone-free status (AUC were 0.689, 0.657, 0.303 and 0.690, respectively). All four scoring systems predicted complications with AUC values of 0.689, 0.646, 0.286 and 0.664 for RUSS, modified S-ReSC, Ito's nomogram, R.I.R.S., respectively. Although all scoring systems were able to predict complications only Ito's nomogram was able to predict Clavien [?]2 complications. Conclusion: All four scoring systems (RUSS, modified S-ReSC, Ito's nomogram and R.I.R.S.) could predict stone-free status after f-URS, however the AUC values are not satisfactory in our large patient cohort. Although these scoring systems were not developed for predicting post-operative complications, they were associated with complications in our study. However, these four scoring systems have some significant limitations. The ideal scoring system is yet to be developed.

External validation and comparison of current scoring systems in retrograde intrarenal surgery: multi-institutional study with 949 patients

Abstract

Objectives: To externally validate and compare Resorlu-Unsal stone score(RUSS), modified Seoul National University Renal Stone Complexity Score(S-ReSC), Ito's nomogram and R.I.R.S. scoring systems for predicting capabilities of both the stone-free status and complications in a multi-institutional study.

Materials and Methods: We performed a retrospective analysis of 949 patients who were underwent flexible ureterorenoscopy (f-URS) and laser lithotripsy for renal stones in two institutions between March-2015 and June-2020. The RUSS, modified S-ReSC, Ito's nomogram and R.I.R.S. scores were calculated for

each patient by same surgeon on imaging methods. Results were compared for their predictive capability of stone-free status and complications.

Results: Of 949 patients 603 were male and 346 were female with a mean age of 47.2±14.3 (range 2-84years). Mean stone burden was 102.6±42.2 (48-270mm^2). All nomograms predicted stone-free status (AUC were 0.689, 0.657, 0.303 and 0.690, respectively). All four scoring systems predicted complications with AUC values of 0.689, 0.646, 0.286 and 0.664 for RUSS, modified S-ReSC, Ito's nomogram, R.I.R.S., respectively. Although all scoring systems were able to predict complications only Ito's nomogram was able to predict Clavien [?]2 complications.

Conclusion: All four scoring systems (RUSS, modified S-ReSC, Ito's nomogram and R.I.R.S.) could predict stone-free status after f-URS, however the AUC values are not satisfactory in our large patient cohort. Although these scoring systems were not developed for predicting post-operative complications, they were associated with complications in our study. However, these four scoring systems have some significant limitations. The ideal scoring system is yet to be developed.

Key words: Retrograde intrarenal surgery; nomogram; validation; renal stone

What's already known about this topic?

The search for a widely available and generalizable scoring system in retrograde intrarenal surgery continues.

What does this article add?

In this study, we aimed to shed light on the development of the ideal scoring system by emphasizing the lacks of the current scoring systems. We have also showed that these scoring systems can also predict the risk of complications.

Acknowledgements: None

Funding : None

Conflict of interest: All authors declare no conflict of interest

Word count of text: 1951

Word count of Abstract: 250

Introduction

In parallel to minimal invasive surgery trend in urological surgery, flexible ureterorenoscopy (f-URS) and laser lithotripsy gained a significant popularity in the treatment of renal stones. Despite the technological advancements in optical and lithotripsy systems there are still some limiting factors including stone burden and abnormal renal anatomy [1]. To standardize reporting the outcomes in scientific studies and pre-operative counseling of patients some scoring systems were developed but none of them gained popularity and were not widely used. The ideal nomogram should include all the variables that might affect the outcome, besides should be simple, accurate and easy to apply in office conditions. The success of a f-URS can not only be evaluated with stone-free rates, but the ideal success should be stone-free rates without any complication or with acceptable complication rates. So, the ideal scoring system should predict the complications besides the stone free-rates. The first scoring system for predicting success of f-URS was developed by Resorlu et al. in 2012 [2].After than modified Seoul National University Stone Complexity score (S-ReSC) [3], a nomogram by Ito et al. [4] and R.I.R.S. scoring system [5] were developed. We aimed to externally validate and compare these four scoring systems for predicting capabilities of both the stone-free status and complications in a multi-institutional study.

Materials and Methods

We performed a retrospective analysis of 949 patients who were underwent f-URS and laser lithotripsy for renal stones in two institutions (HSU Izmir Bozyaka Training and Research Hospital, HSU Ankara Diskapi Training and Research Hospital) between March 2015 and June 2020. Preoperative evaluation of the patients included medical history, physical examination, complete blood count, urine analysis and culture, serum biochemistry and non-contrast computed tomography (NCCT). All patients were given antibiotic prophylaxis according to current guidelines. Stone characteristics and renal anatomical characteristics were evaluated on NCCT. Stone size was determined by measuring the longest axis at NCCT; in case of multiple stones the sum of each stone size was calculated. Stone surface area and lower pole infindibulo-pelvic angle (IPA) was measured as defined before [6,7]. All procedures were performed by experienced surgeons by using different size and brand ureterorenoscopes (7.5F Flex –X2 (Karl Storz,Tuttlingen Germany), Olympus URF P-5 (Olympus, Tokyo, Japan), 9.5F Pusen (Zhuhai Pusen Medical Technology Co, Ltd., Zhuhai, China). Patients with missing data and/or lacking of 1st month post-operative NCCT were excluded from the study. All procedures were performed under general anesthesia in lithotomy position. A diagnostic ureteroscopy with a semi-rigid ureterorenoscope was always performed at the beginning of surgery both to dilate the ureter and pass a safety guidewire. A ureteral access sheath (9,5-11,5; 12-14 Fr) was placed according to surgeon's preferences. The stones were fragmented by using a holmium: YAG laser. If needed some fragments were removed with tipless nitinol stone baskets. At the end of procedure, a D-J stent was placed according to surgeon's decision. All patients were evaluated with NCCT at 1st month for stone-free status. Residual fragments <2mm were accepted as insignificant.

The RUSS, modified S-ReSC, Ito's nomogram and R.I.R.S. scores were calculated for each patient by same surgeon on imaging methods. Results were compared for their predictive capability of stone-free status and complications.

Summary of scoring systems

Resorlu-Unsal stone score (RUSS): Stone size >20mm, lower pole stone and IPA<45°, stone number in different calyces >1, abnormal renal anatomy (horseshoe kidney or pelvic kidney).Each condition was given 1 point. Total score of 0-4.

Modified S-ReSC: The modified S-ReSC score was assigned according to the number of sites involved in the renal pelvis (#1), superior and inferior major calyceal groups (#2-3), and anterior and posterior minor calyceal groups of the superior (#4-5), middle (#6-7), and inferior calyx (#8-9). If the stone was in the inferior sites (#3, #8-9), one additional point per site was added.

Ito's nomogram: Stone volume ([?]500, 500<x[?]1000, 1000<x[?]2000,>2000 mm^3), lower pole calculi, operator experience (<50, [?]50), hydronephrosis and number of stones. A nomogram score was calculated according to these parameters (total score 0-25). High total score was predictive of successful outcome.

R.I.R.S. score: Renal stone density ([?]1000, >1000), inferior pole stone (non-inferior, inferior with RIPA>30°, inferior with RIPA[?]30°), renal infindibular length ([?]25, >25 mm), stone burden ([?]10, >10 and [?]20, >20 mm). Total score range 4-10 points. Ten points indicates the most complex situation.

Statistical analysis

Allanalysis were performed by using Statistical Package for the Social Sciences (SPSS, Version 22.0; SPSS, Chicago, III). Patients were divided into two groups according to SF status. T-test and Pearson $\chi 2$ test were used to compare variables between SF and non-SF groups. ROC Curve analysis was used for prediction levels of the nomograms. Statistical significance was accepted as $p < 0.05$ for all analyses.

Results

A total of 949 who underwent f-URS and laser lithotripsy between March 2015 and June 2020 were included in the study. Of 949 patients 603 were male and 346 were female with a mean age of 47.2±14.3 (range 2-84 years). Mean stone burden was 102.6±42.2 (48-270 mm^2). In Table 1, the comparison of characteristics of stone-free and non stone-free patients was shown. There were significant differences between two

groups according to age, previous URS or precutaneous nephrolithotomy (PCNL) history, number of stone, stone density, maximum stone diameter, stone burden, lower calyx localization, preoperative hydronephrosis, preoperative J stent placement, operation time, hospitalization time and complications. All nomograms (RUSS, modified S-ReSC, Ito's nomogram, R.I.R.S.) were predicted stone-free status (AUCs were 0.689, 0.657, 0.303 and 0.690, respectively) (Table 2). Only the Ito's nomogram's ROC curve was below reference line, because the higher score indicates residual stone presence in this nomogram however higher score was predictive of stone-free in others (Fig 1). All scoring systems were predicted complications with AUC values of 0.689, 0.646, 0.286 and 0.664 for RUSS, modified S-ReSC, Ito's nomogram, R.I.R.S., respectively (Table 3).Although all scoring systems were able to predict complications only Ito's nomogram was able to predict Clavien [?]2 complications (p=0.016).

Comment

Recent technical improvement including miniaturization, improved deflection and enhanced optical quality have increased the use of flexible ureterorenoscopes in renal stones. Although f-URS is not recommended as a first line treatment modality for renal stones <2cm at current guidelines, a recent meta-analysis resulted an average of 91% stone-free rate with acceptable complication rates and with an average of 1.45 procedures per patient (range 1.11-1.82) [6,8]. As being more widely used, we need some scoring systems that can be used to predict the success and potential complications after f-URS for counseling of the patients preoperatively.

The variables that were used in these four studies were not standard. Although abnormal renal anatomy is one of the parameters of RUSS, patients with musculoskeletal and renal abnormality directly excluded in R.I.R.S. Also, renal anomaly was not studied in modified S-ReSC. IPA was an independent predictive factor in RUSS and R.I.R.S., however it was not studied in modified S-ReSC and nomogram by Ito et al. So, if a variable was not evaluated in a study we cannot know if it is an independent predictive factor or not.

The success was defined as stone-free status in all scoring systems. However, no standard imaging modality was used and also the timing of imaging was different between studies. And thus, there can be a significant bias between studies. In R.I.R.S stone-free status was evaluated with kidney-ureter-bladder (KUB) plane graphy at 1st month and >2mm fragments were accepted as not-stone free, computed tomography (CT) was used whenever necessary. In the nomogram of Ito et al. stone-free status was evaluated with non-contrast CT at 3rd month. In RUSS, treatment success was defined as stone-free or residual fragment [?]1mm at 1st month non-contrast CT imaging. In modified S-ReSC stone free was defined as no evidence of residual stone at 1st month NCCT.

The success of a procedure can be defined as to reach a targeted end-point without any complication or with acceptable complication rates. Thus, an ideal f-URS scoring system scoring system should foresee complications besides stone-free. None of the scoring systems examined the relation of complexity scores and complications. Only in R.I.R.S. authors indicated that the score is correlated with operation time and thus it can be used to predict to leave the operation for a second session and complications can be avoided. All four scoring systems were able to predict complications in the present study, but only nomogram by Ito et al. could predict Clavien [?]2 complications.

Stone burden is one of the most important predictors of stone-free after f-URS as we know from previous studies [8,9]. As expected, stone burden was an independent predictor of success in all scoring systems except modified S-ReSC. This is unsurprising for modified S-ReSC, because the mean stone burden was significantly lower compared to other studies. The authors indicated that patients with high stone burden were underwent percutaneous nephrolithotomy and excluded from the study. Although modified S-ReSC can be predicted stone free status even in our study, it can be no doubtfully said that success of f-URS for 1cm stone cannot be equivalent to 3 cm stone in the same location that both was given same score in S-ReSC.

There are two studies comparing the current nomograms. Erbin et al. compared the RUSS and modified S-ReSC and found out that RUSS, musculoskeletal deformity and stone size were independent factors in logistic regression analysis [10]. And the AUC values were not satisfactory in terms of nomogram's predictive accuracy (0.655 and 0.596 for RUSS and modified S-ReSC, respectively). In another recent study by

Richard et al. RUSS, modified S-ReSC, S.T.O.N.E and Ito's nomogram were compared [11]. The major limitation of that study was evaluation of the stone-free status that they made it via visual inspection using ureterorenoscope and fluoroscopy at the end of the procedure in the absence of clinical complications. They concluded that all scoring systems were predicted the stone-free status but none of them were predictive of complications after RIRS. Different from the former study all nomograms could predict the stone-free status and complications in our study as being the Ito's nomogram most sensitive.

In the present study all four nomograms were underwent external validation in a significantly high number of patient population. Although all nomograms were predictive of stone-free status, the AUC values were lower than that in original study and were not satisfactory. The AUC values in original study and in our study were 0.806 vs 0.657, 0.87 vs. 0,697 and 0.904 vs.0.690 for modified S-ReSC, Ito's nomogram and R.I.R.S., respectively. The AUC value was not calculated in RUSS.

Strengths of our study include high number of study group, standardized post-operative imaging at a determined time period and various cases including different anatomical abnormalities. Limitation of our study is its retrospective design.

We think that none of these four scoring systems is ideal. Because the success definition is not standard and studied variables are limited. To constitute an ideal nomogram a multi-center study examining the all known potential variables with a strict success definition in a defined time period can be planned with high number of patients. The ideal scoring system could also be used in daily clinical practice. In clinical practice it is hard to calculate stone burden by measuring the diameter of all calculi (especially in multiple stones) or to measure the IPA. In our opinion a stone score should be given by radiologists in imaging reports whenever a widely used ideal nomogram can be developed.

Conclusion

All of four scoring systems (RUSS, modified S-ReSC, Ito's nomogram and R.I.R.S.) could predict stone-free status after f-URS, however the AUC values are not satisfactory in our large patient cohort. Although these scoring systems were not developed for predicting post-operative complications, they were associated with post-operative complications in our study. All these four scoring systems have some significant limitations. The ideal scoring system is yet to be developed.

References

1. Rodriguez-Monsalve Herrero M, Doizi S, Keller EX, De Coninck V, Traxer O (2018) Retrograde intrarenal surgery: An expanding role in treatment of urolithiasis. Asian J Urol 5:264-273. doi: 10.1016/j.ajur.2018.06.005.
2. Resorlu B, Unsal A, Gulec H, Oztuna D (2012) A new scoring system for predicting stone-free rate after retrograde intrarenal surgery: The "Resorlu-Unsal Stone Score". Urology 80:512–518.
3. Jung JW, Lee BK, Park YH et al. (2014) Modified Seoul National University renal stone complexity score for retrograde intrarenal surgery. Urolithiasis 42:335-340.
4. Ito H, Sakamaki K, Kawahara T, Terao H, Yasuda K et al (2015) Development and internal validation of a nomogram for predicting stone-free status after flexible ureteroscopy for renal stones. BJU Int 115:446-451. doi: 10.1111/bju.12775.
5. Xiao Y, Li D, Chen L, Xu Y, Zhang D, Shao Y et al. (2017) The R.I.R.S. scoring system: An innovative scoring system for predicting stone-free rate following retrograde intrarenal surgery. BMC Urol 17:105. doi:10.1186/s12894-017-0297-0.
6. Turk C, Neisius A, Petrik A, Seitz C, Skolarikos A, Thomas K (2020) Guidelines on urolithiasis. European Association of Urology. https://uroweb.org/guideline/urolithiasis/2020.
7. Elbahnasy AM, Shalhav AL, Hoenig DM, et al 1998) Lower caliceal stone clearance after shock wave lithotripsy or ureteroscopy: the impact of lower pole radiographic anatomy. J Urol 159:676-682.
8. Geraghty R, Abourmarzouk O, Rai B et al. (2015) Evidence for Ureterorenoscopy and Laser Fragmentation (URSL) for Large Renal Stones in the Modern Era. CurrUrol Rep 16:54.

9. Kılıc O, Akand M, Van Cleynenbreugel B (2017) Retrograde intrarenal surgery for renal stones –Part 2. Turk J Urol 43:252-260. doi: 10.5152/tud.2017.22697.
10. Erbin A, Tepeler A, Buldu I, Ozdemir H et al. (2016) External Comparison of Recent Predictive Nomograms for Stone-Free Rate Using Retrograde Flexible Ureteroscopy with Laser Lithotripsy. J Endourol 30:1180-1184. doi: 10.1089/end.2016.0473
11. Richard F, Marguin J, Frontczak A et al. (2020) Evaluation and comparison of scoring systems for predicting stone-free status after flexible ureteroscopy for renal and ureteral stones. PLoS One 15:e0237068. doi:10.1371/journal.pone.0237068.

Figure legends:

Fig 1. ROC Curves of thenomograms for prediction of SF status.

Hosted file

`Table 1.pdf` available at https://authorea.com/users/381091/articles/496889-external-validation-and-comparison-of-current-scoring-systems-in-retrograde-intrarenal-surgery-multi-institutional-study-with-949-patients

Hosted file

`Table 2.pdf` available at https://authorea.com/users/381091/articles/496889-external-validation-and-comparison-of-current-scoring-systems-in-retrograde-intrarenal-surgery-multi-institutional-study-with-949-patients

Hosted file

`Table 3.pdf` available at https://authorea.com/users/381091/articles/496889-external-validation-and-comparison-of-current-scoring-systems-in-retrograde-intrarenal-surgery-multi-institutional-study-with-949-patients