

STONES/ENDOUROLOGY

ORIGINAL ARTICLE

Implications of different nephrolithometry scoring systems on clinical practice of endourologists: An international web-based survey



Mohamed A. Elkoushy^{a,*}, Adel H. Metwally^a, Yasser A. Noureldin^b

^a Department of Urology, Suez Canal University, Ismailia, Egypt

^b Department of Urology, Benha University, Benha, Egypt

Received 29 February 2016, Received in revised form 13 April 2016, Accepted 29 April 2016

Available online 10 June 2016

KEYWORDS

Endourology;
Percutaneous
nephrolithotomy;
Outcome;
Nomogram;
Clinical practice pat-
tern

ABBREVIATIONS

CROES, the Clinical
Research Office of the
Endourology Society;
NCCT, non-contrast
CT;
NLSS, nephrolithome-
try scoring systems;
PCNL, percutaneous
nephrolithotomy;
SFS, stone-free status;

Abstract Objectives: To assess the implications of different nephrolithometry scoring systems (NLSS) on clinical practice of endourologists to predict stone-free status (SFS) after percutaneous nephrolithotomy (PCNL).

Methods: A web-based survey was sent to members of the Endourological Society. Demographic and practice pattern data were collected. Multiple-choice and open-ended questions were used to assess awareness about the NLSS and their authentic use in clinical practice. Surgeon preferences and limitations of NLSS and how to overcome them were asked.

Results: In all, there were 162 responses, including 17 (10.5%) respondents who were not aware of NLSS. Most respondents (82.1%) denied the efficacy of NLSS in predicting SFS after PCNL. Of 145 respondents who were aware of NLSS, 85.5% did not use them in clinical practice. Endourologists aged 40–60 years ($P < 0.001$), in practice for 10–20 years ($P = 0.003$), those performing 100–200 PCNLS/year ($P = 0.02$), and those from North America ($P < 0.001$) seemed to use NLSS more frequently. In all, 50% of respondents preferred not to use any NLSS, while 29% chose the S.T.O.N.E followed by the Guy's Stone Score (10.3%) and The Clinical Research Office of the Endourology Society (CROES) nomogram (8.3%). Inconsistency and variability among different NLSS were the main drawbacks reported by 82% of 89 respondents. The need for high-level evidence for NLSS through direct randomised prospective comparison was recommended by 24.8% of respondents who answered that question.

* Corresponding author at: Suez Canal University, Ismailia, Egypt. Tel.: +20 010 001 35102.

E-mail addresses: mohamed_elkoushy@med.suez.edu.eg, melkoushy@yahoo.com (M.A. Elkoushy).

Peer review under responsibility of Arab Association of Urology.



Production and hosting by Elsevier

S-ReSC, Seoul National University Renal Stone Complexity; S.T.O.N.E., Stone Size, Tract length, Obstruction/hydronephrosis, Number of involved calyces, Essence/Hounsfield units

Conclusion: There is a lack of compliance and acceptance of different NLSS in clinical practice among endourologists. Inconsistency and inaccuracy in predicting SFS after PCNL limits their incorporation into clinical practice. However, the results of this study might not be generalisable due to the selection bias resulting from the geographical distribution of the respondents and the heterogeneity in surgical expertise. Therefore, randomised prospective direct comparisons and validation of these systems are recommended.

© 2016 Arab Association of Urology. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Percutaneous nephrolithotomy (PCNL) has become the standard procedure for large renal stones in most recent guidelines [1]. In the past decade, PCNL has been performed increasingly in the USA, a trend which is assumed to continue [2]. However, inconsistent definitions and accepted standardised reporting of PCNL outcomes may impact the quality of urological research and influence any comparative evaluation for the management of urolithiasis between different series [3,4]. Moreover, preoperative patient counselling necessitates the development of an integrated scoring system to assess PCNL complexity for optimal decision-making [3].

Therefore, different scoring systems and nomograms have been introduced to overcome these limitations for systematic and quantitative assessment of the outcomes of PCNL including the Guy's Stone Score [5], S.T.O.N.E. (Stone Size, Tract length, Obstruction/hydronephrosis, Number of involved calyces, Essence/Hounsfield units) nephrolithometry scoring system [6], The Clinical Research Office of the Endourology Society (CROES) nomogram [7], and the Seoul National University Renal Stone Complexity (S-ReSC) score [8,9]. All these nephrolithometry scoring systems (NLSS) aim for preoperative prediction of stone-free status (SFS) and complications through assessment of the complexity of different cases undergoing PCNL. NLSS do not only consider the imaging criteria of stones and renal anatomy, but also relevant patient characteristics such as body mass index, previous renal surgery, and surgeon experience.

However, these NLSS were validated only in retrospective studies [10–15] and they all have their own limitations. In addition, these systems seem to suffer from an intrinsic bias favouring predictive efficacy that is dependent on the evaluated population of patients [11]. Furthermore, SFS may also be affected by surgeon experience, heterogeneity of practice, surgical modifications among different centres, and involvement of post-graduate trainees, which should be considered as well.

Currently, there are no sufficient high-level evidence data regarding the applicability and reliability of these

NLSS in achieving their objectives. The question now is whether the validation of different NLSS have been translated into a change in the practice patterns among practitioners. In other words, are practising endourologists convinced of the utility of these NLSS and their ability to predict post-PCNL outcomes? The present worldwide survey was planned and conducted as a trial to answer this question and to assess the implications of NLSS on clinical practice patterns among endourologists.

Methods

An internet-based survey was sent to the members of the Endourology Society using their e-mail addresses. The website remained open for 2 months (9 September to 9 November 2015) to give the respondents an opportunity to visit and complete the online survey. The survey was sent by e-mail independently and was planned to be concise enough to obtain the answers to the research questions. On 9 October 2015, friendly reminder e-mails were re-sent to encourage those who missed the first invitation to complete the survey. The survey was set to receive only one response per computer to avoid the potential response bias of duplicate responses, especially when sending such reminders.

Demographic and practice pattern data were collected, including age, place and years of practice, and the number of PCNLs performed within the previous 12 months. In addition, multiple-choice and open-ended questions were used to assess awareness of the endourologists about the current NLSS, authentic use in clinical practice, surgeon preferences, and reliability of these systems in accurately predicting SFS after PCNL. Also, the respondents were queried about the limitations of the NLSS, reasons for not using them, and how to overcome such obstacles (Appendix 1). Participants were given the option to check more than one option for question number 7.

NLSS

The Guy's Stone Score, S.T.O.N.E. nephrolithometry score and CROES nomogram consist of 11 variables,

including four shared parameters (stone size, location, number, and staghorn status). Other variables which are considered by these three NLSS are stone density, renal anatomy, tract length, obstruction of renal pelvis, case load/year, prior treatment history, and the presence of spina bifida or spinal injury.

The Guy's Stone score categorises PCNL cases into four grades, increasing in complexity from Grade 1 to Grade 4, depending on patients' past medical history and non-contrast CT (NCCT) [5]. The S.T.O.N.E. nephrolithometry scoring system assesses PCNL complexity by nine different possible scores, ranging from 5 to 13, according to five NCCT-calculated parameters including stone size, tract length, obstruction/hydronephrosis, number of involved calyces, essence/Hounsfield units [6].

However, the CROES nephrolithometric nomogram predicts treatment success using plain X-ray of the kidneys, ureters and bladder considering the stone burden, count, and location, in addition to case volume and prior stone treatment [7]. The total score to predict the chance of treatment success is the sum of individual scores derived from each predicting variable; the higher the score, the higher the chance of treatment success, while a patient with a low score has a low chance of achieving a SFS [7].

On the other hand, the S-ReSC score is calculated by counting the number of sites involved in the renal pelvis and calyces, regardless of stone parameters or renal anatomy. This system assigns a score from 1 to 9 depending mainly on the number of sites involved; the renal pelvis (1), superior and inferior major calyces (2–3), anterior and posterior minor calyces of the superior (4–5), middle (6–7), and inferior calyces (8–9).

Statistical analysis

Data were analysed using the commercially available Statistical Package for Social Science (SPSS) for windows, version 20 (Chicago, IL, USA). Descriptive data were presented in terms of number of responses, percentages, medians and interquartile range (IQR). Fisher's exact test was used for comparing discrete variables with a two-tailed $P < 0.05$ indicating significant differences between groups.

Results

A total of 162 responses were obtained within the 2-month period from September to November 2015. Most of the responses came from North America, Africa, and Europe (31.5%, 19.7%, and 16.7%, respectively). Most endourologists were aged 40–60 years (59.9%) and had practised for 10–20 years (45.1%) (Table 1). Over the previous 12-months, the respondents performed a median (IQR) of 152 (63–182) PCNLs. There was a

wide range of PCNL caseloads with only 12.3% of respondents performing a high volume of cases.

In all, 17 respondents (10.5%) were not aware of NLSS and were excluded from data analysis of questions numbers 6–10, bringing the total number of respondents to these questions to 145 (89.5%) respondents. The 10.5% of respondents who were not aware of NLSS came from Africa and Asia (5.6% and 4.9%, respectively). Most of them (9.2%) were aged < 40 years and had practised for < 10 years (8.6%), while their PCNL case load was zero or < 100 procedures in the previous 12 months (6.2% and 4.3%, respectively).

Surprisingly, most respondents (82.1%) denied the efficacy of NLSS in accurately predicting SFS after PCNL, including 33.8% who were not sure (Fig. 1). This was translated into 124 of 145 respondents (85.5%) who did not use NLSS in their daily clinical practice. Respondents who did not use NLSS are characterised in Fig. 2, including 66% of those aged 40–60 years vs $> 91\%$ in the other two age groups ($P < 0.001$), 64.4% of those in practice for 10–20 years vs 82–90% in the other two groups ($P = 0.003$), 65.5% in those performing 100–200 PCNLs vs 70–100% in the other two groups ($P = 0.02$), and 54.9% in North America vs 74.1–93.8% for the other continents ($P < 0.001$).

When asking about the explanation of the latter response, the responses were obtained from only 72 (58.1%) of those who did not use NLSS. The responses obtained were: 80.5% of respondents reported that some of these NLSS are time-consuming, while others found that NLSS did not consider some factors that could impact postoperative outcomes such as surgeon

Table 1 Baseline and demographic characteristics of the respondents.

Variable	N (%)
Respondents	162 (100)
Age, years	
< 40	42 (25.9)
40–60	97 (59.9)
> 60	23 (14.2)
Place of practice	
Africa	32 (19.7)
Asia	25 (15.4)
Australia	9 (5.6)
Europe	27 (16.7)
North America	51 (31.5)
South America	18 (11.1)
Duration of practice, years	
< 10	50 (30.9)
10–20	73 (45.1)
> 20	39 (24.1)
Number of PCNL performed within the previous 12 months	
None	8 (4.9)
< 100	76 (46.9)
100–200	58 (35.8)
> 200	20 (12.3)

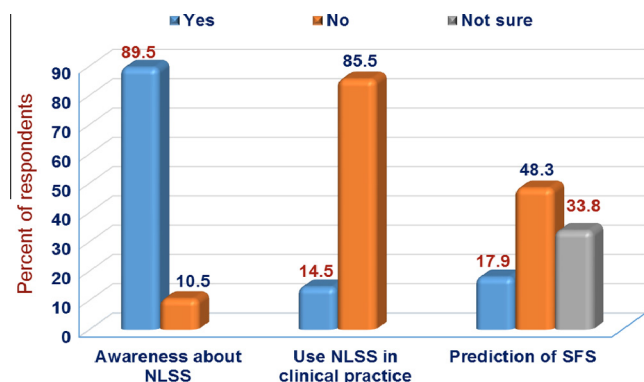


Figure 1 Implications of NLSS on clinical practice of respondents.

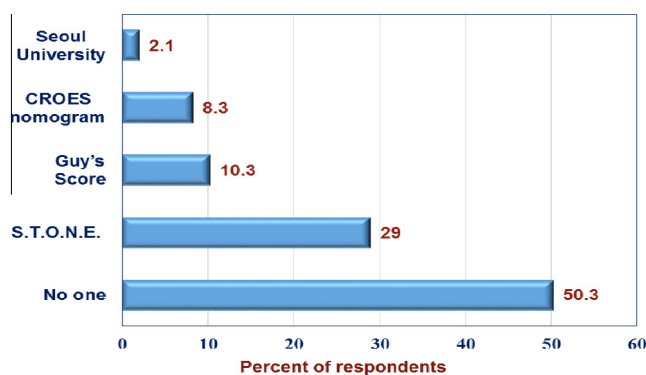


Figure 3 Which NLSS do you use or prefer to use?

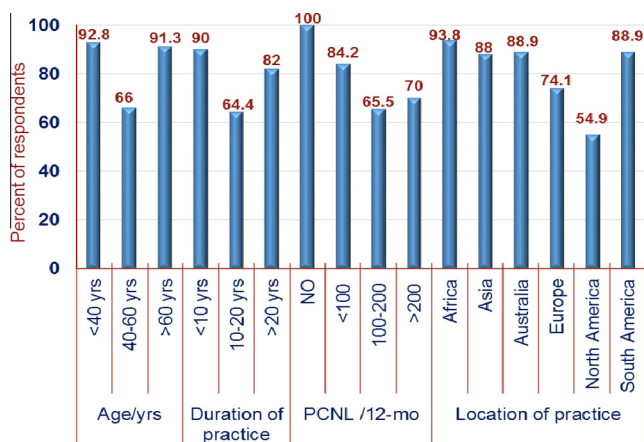


Figure 2 Characterisation of respondents who did not use NLSS in clinical practice.

experience. Increased work load was reported by 18.1%, in addition to 12.5% who cited the complexity of these systems and lack of their clinical utility for not using NLSS.

When questioned about which NLSS is currently or is preferred to be used, 50.3% of respondents chose no one, while 29% choose the S.T.O.N.E nephrolithometry score followed by the Guy's Stone Score and CROES nomogram (10.3% and 8.3%, respectively; Fig. 3). Only three respondents (2.1%) recommended the S-ReSC and all of them came from Asia.

In all, 89 (61.4%) respondents provided several drawbacks or obstacles to using NLSS in clinical practice. Inconsistency and variability among different systems were reported by 82% of the respondents. When questioned whether they had any recommendations to overcome these limitations and/or improve the validity of NLSS, only 36 (24.8%) participants responded to that question and most of them recommended the need for high-level evidence for these NLSS through direct prospective comparison between these systems in a randomised controlled trial. Others mandated a large multicentre series to ensure matched or similar groups of patients.

Several NLSS have been developed in the past few years to predict the outcomes of PCNL. However, these NLSS lack high-level evidence to show their reliability and reproducibility, despite both the Guy's Stone Score and S.T.O.N.E nephrolithometry score having been tested for inter-observer concordance, albeit retrospectively. Therefore, these NLSS are limited by the data captured, as they are unable to take into account factors that may influence the outcome of PCNL such as surgeons' experience and skills, together with the emerging techniques that would improve the SFS such as retrograde intrarenal surgery combined with PCNL.

In the current era of NLSS, there are no data available about their implementation in clinical practice and whether practising endourologists are convinced of their utility and ability to predict the outcomes of PCNL. The present study was conducted to address this issue and to assess the implications of these NLSS on clinical practice patterns among endourologists.

In the present survey, 10% of the respondents were not aware of NLSS, while 85.5% of the remaining respondents do not use them in their daily clinical practice. This could be explained by the finding that most respondents (82.1%) denied the efficacy of NLSS to accurately predict post-PCNL SFS. Moreover, 80% of those who do not use NLSS found them time-consuming, inconsistent and/or inaccurate. A lack of level-1 evidence to support their use in clinical practice, increased workload, and complexity of these systems may also explain the high percentage of those who do not use them.

Characterisation of respondents who use/do not use NLSS in clinical practice showed that middle-aged respondents (aged 40–60 years), those in practice for 10–20 years, and those performing 100–200 PCNLs in the previous 12-months used NLSS significantly more than either extreme. It seems also that respondents from North America use NLSS significantly more frequently than their colleagues from other continents.

While half of the respondents would not like to use NLSS in clinical practice, 29% favoured the S.T.O.N.E, followed by the Guy's Stone Score and CROES

Discussion

nomogram. Those who preferred to use the S.T.O.N.E nephrolithometry score found it easier and more practical than the other scoring systems. Despite the Guy's Stone score, S.T.O.N.E. nephrolithometry score and CROES nomogram being equally predictive of post-PCNL SFS in one study, Labadie et al. [11] considered that the S.T.O.N.E. nephrolithometry scoring system provides more accurate data stratification than the Guy's Stone score and offers an easier application than the CROES nomogram. Despite a scale incorporation into the latter nomogram to help practitioners to give a percentage chance of treatment success, it was found by some respondents to be the most difficult to apply in clinical practice. Furthermore, a multicentre study including 850 patients found that the S.T.O.N.E. nephrolithometry score accurately predicted surgical outcomes after PCNL, including SFS and overall complications [16]. However, these results should be cautiously interpreted as participation of different institutions with different raters in such a retrospective approach would impact its outcomes. Conversely, the S-ReSC was recommended by only three Asian respondents, despite appearing to be much easier. However, it does not account for the stone characteristics specifically the stone size, which remains the leading predictor of perioperative outcome in all the published reports.

In the present study, inconsistency and variability among different NLSS were reported to be the main problems of using them in clinical practice. Post-PCNL complications may be considered as a good example of such variability where some authors found significant positive association between the Guy's Stone scoring system and post-PCNL complications [17,18], while others did not detect a similar correlation [5,14]. However, it is unfortunate that the respondents did not make a distinction between the different NLSS, as it seems obvious that all NLSS were not considered to be equally time consuming, complex, and/or inaccurate, from the respondents' point of view.

When questioned about their advice to overcome these limitations and to improve the validity of NLSS, most respondents recommended direct prospective comparison between these NLSS in randomised controlled trials. A large multicentre series is awaited to ensure matched or similar groups of patients. Others recommended a single, reliable, and reproducible NLSS, which is adequately comprehensive for thorough reporting and comparison. This would help to answer the current question regarding how clinicians incorporate research advances into clinical practice as such trials with a rising tide of **consensus** invariably lead to practice changes to deliver optimum care to the patients. Nevertheless, incorporation of research advances into practice will typically require more time and more proof.

Being a web-based survey, the present study may be limited by the low response rate, despite it being sent

independently to the e-mail addresses of the endourologists, which may be potentially outdated. A response bias may be expected by respondents who may have been more inclined to complete the survey due to their interest in the subject. Moreover, the unequal worldwide distribution of the respondents may represent a dominant practice from a certain continent such as North America. This has been indicated and deserves to be highlighted. Furthermore, lack of data about academic versus community practice of the respondents might limit interpretation of some findings. Unequal distribution of respondents over different continents might also limit the generalisability of the present results. Nevertheless, the present survey appears to reveal a disparity between research and real clinical settings, highlighting the necessity for high-level evidence rather than retrospective trials when trying to validate a new technique or evaluation tools.

Conclusion

There is a lack of compliance and acceptance of different NLSS among endourologists in clinical practice. Endourologists aged 40–60 years, in practice for 10–20 years, those performing 100–200 PCNLs/year, and those from North America seem to use NLSS significantly more frequently than their colleagues. Inconsistency and inaccuracy of some NLSS to predict post-PCNL outcomes limits their incorporation into clinical practice. However, the results of the present study might not be generalisable due to the selection bias resulting from the geographical distribution of the respondents and the heterogeneity in surgical expertise. Therefore, randomised controlled trials are recommended for direct comparison and validation of these NLSS.

Compliance with ethical standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments.

This article is an online survey and does not contain any studies with human participants or animals performed by any of the authors.

Author's contributions

Mohamed A. Elkoushy:

Project development
Data collection and management
Data analysis
Manuscript writing and editing

Adel H. Metwally:

Data collection and management

Manuscript writing and editing
Yasser A. Noureldin:
 Data management
 Manuscript writing and editing

Source of funding

None.

Appendix 1: Online questionnaire

Conflicts of interest

All the authors declare no conflicts of interest.

1- How old are you?

- <40 years
- 40–60 years
- >60 years

2- Where do you practice?

- Africa
- Asia
- Australia
- Europe
- North America
- South America

3- How long have you been practicing?

- <10 years
- 10–20 years
- >20 years

4- Approximately how many PCNL you performed within the last 12 months?

- None
- <100 procedures
- 100–200 procedures
- >200 procedures

5- Did you hear about the nephrolithometry scoring systems to predict post-PCNL stone-free status?

- Yes
- No
- If No, thanks, you do not have to complete the survey!
- If Yes, please specify which scoring systems do you know, then go to the second page

6- Do you really use any of them in your current practice?

- Yes
 - No
- If no, please specify, why?

7- Which scoring system are you using or prefer to use?

- No one
 - Guy's scoring system
 - S.T.O.N.E. nephrolithometry scoring system
 - CROES nephrolithometric nomogram
 - Seoul National University Renal Stone Complexity (S-ReSC)
- Please specify, why?

8- Do you think these scoring systems can accurately predict post-PCNL stone-free status?

- Not sure
- Yes
- No

Please specify, why?

9- In your opinion, what are the drawbacks/ obstacles of using these scoring systems?

10- Do you have any suggestion to improve the validity of these scoring systems?

Thank you

References

- [1] Türk C, Knoll T, Petrik A, Sarica K, Skolarikos A, Straubet M, et al. *Guidelines on Urolithiasis*. European Association of Urology 2014. Available at: http://uroweb.org/wp-content/uploads/22-Urolithiasis_LR.pdf. [Accessed 19 November 2015].
- [2] Mirheydar HS, Palazzi KL, Derweesh IH, Chang DC, Sur RL. Percutaneous nephrolithotomy use is increasing in the United States: an analysis of trends and complications. *J Endourol* 2013;**27**:979–83.
- [3] Opondo D, Gravas S, Joyce A, Pearle M, Matsuda T, Sun YH, et al. Standardization of patient outcomes reporting in percutaneous nephrolithotomy. *J Endourol* 2014;**28**:767–74.
- [4] Hyams ES, Bruhn A, Lipkin M, Shah O. Heterogeneity in the reporting of disease characteristics and treatment outcomes in studies evaluating treatments for nephrolithiasis. *J Endourol* 2010;**24**:1411–4.
- [5] Thomas K, Smith NC, Hegarty N, Glass JM. The Guy's stone score—grading the complexity of percutaneous nephrolithotomy procedures. *Urology* 2011;**78**:277–81.
- [6] Okhunov Z, Friedlander JI, George AK, Duty BD, Moreira DM, Srinivasan AK, et al. S.T.O.N.E. nephrolithometry: novel surgical classification system for kidney calculi. *Urology* 2013;**81**:1154–9.
- [7] Smith A, Averch TD, Shahrour K, Opondo D, Daels FP, Labate G, et al. A nephrolithometric nomogram to predict treatment success of percutaneous nephrolithotomy. *J Urol* 2013;**190**:149–56.
- [8] Jeong CW, Jung JW, Cha WH, Lee BK, Lee S, Jeong SJ, et al. Seoul National University Renal Stone Complexity Score for Predicting Stone-Free Rate after Percutaneous Nephrolithotomy. *PLoS ONE* 2013;**8**:e65888.
- [9] Jung JW, Lee BK, Park YH, Lee S, Jeong SJ, Lee SE, et al. Modified Seoul National University Renal Stone Complexity score for retrograde intrarenal surgery. *Urolithiasis* 2014;**42**:335–40.
- [10] Noureldin YA, Elkoushy MA, Andonian S. External validation of the S.T.O.N.E. nephrolithometry scoring system. *Can Urol Assoc J* 2015;**9**:190–5.
- [11] Labadie K, Okhunov Z, Akhavein A, Moreira DM, Moreno-Palacios J, Del Junco M, et al. Evaluation and comparison of urolithiasis scoring systems in percutaneous kidney stone surgery. *J Urol* 2015;**193**:154–9.
- [12] Okhunov Z, Helmy M, Perez-Lansac A, Menhadji A, Bucur P, Kolla SB, et al. Interobserver reliability and reproducibility of S. T.O.N.E. Nephrolithometry for renal calculi. *J Endourol* 2013;**27**:1303–6.
- [13] Ingimarsson JP, Dagrosa LM, Hyams ES, Pais Jr VM. External validation of a preoperative renal stone grading system: reproducibility and inter-rater concordance of the Guy's stone score using preoperative computed tomography and rigorous postoperative stone-free criteria. *Urology* 2014;**83**:45–9.
- [14] Noureldin YA, Elkoushy MA, Andonian S. Which is better? Guy's versus S.T.O.N.E. nephrolithometry scoring systems in predicting stone-free status post-percutaneous nephrolithotomy. *World J Urol* 2015;**33**:1821–5.
- [15] Park J, Kang M, Jeong CW, Oh S, Lee JW, Lee SB, et al. External validation and evaluation of reliability and validity of the Modified Seoul National University Renal Stone Complexity Scoring System to predict stone-free status after retrograde intrarenal surgery. *J Endourol* 2015;**29**:888–93.
- [16] Okhunov Z, Moreira D, George A, Arvin G, Arash A, Sammy E, et al. Multicenter validation of S.T.O.N.E. nephrolithometry. *J Urol* 2015;**191**:e839, abstract.
- [17] Vicentini FC, Marchini GS, Mazzucchi E, Claro JF, Srougi M. Utility of the Guy's stone score based on computed tomographic scan findings for predicting percutaneous nephrolithotomy outcomes. *Urology* 2014;**83**:1248–53.
- [18] Mandal S, Goel A, Kathpalia R, Sankhwar S, Singh V, Sinha RJ, et al. Prospective evaluation of complications using the modified Clavien grading system, and of success rates of percutaneous nephrolithotomy using Guy's Stone Score: a single-center experience. *Indian J Urol* 2012;**28**:392–8.