Original Article

External validation of Stone Probability Score for screening Loei urolithiasis patients

Krit Santanapipatkul[1], Abhisit Sukjaem[2], Kunyabutt Sutichavengkul[3], Thanaporn Keawruang[3], Pongsathon Wongondee[3]

[1]Division of Urology, Department of Surgery, [2]Department of Anesthesiology, [3]Loei Hospital, Loei, Thailand

Keywords:
External validation, predictive model, probability stone formation, stone probability score, screening urolithiasis patients

Abstract

Objective: Urolithiasis is prevalent in Thailand, particularly in Loei region. Early detection is vital for effective management. In 2022, Santanapipatkul et al. developed a Stone Probability Score for screening urolithiasis patients at Loei, however, external validation of this predictive model is necessary to ensure its reliability and applicability. The objective of this study is to externally validate the Stone Probability Score for Screening Urolithiasis Patients at Loei developed in 2022.

Materials and Methods: The external validation was conducted using cross-sectional data from urolithiasis patients at Loei Hospital between September 28, 2022, and December 31, 2023. Logistic regression analysis was employed to evaluate the performance of the predictive model with regard to discrimination, calibration, and multicollinearity.

Results: This validation study included 347 patients and an accuracy of 92.6% (95%CI 88.9-96.3) was achieved with an AUC pertinent to discrimination measurement resulting in a sensitivity of 96.9% (95%CI 94.2-98.6), a specificity of approximately 53.4% (95%CI 39.9-66.7), positive predictive value of 91.2% (95%CI 87.5-94.1) and negative predictive value of 77.5% (95%CI 61.5-89.2). The performance of the model was found to be consistent after external validation in three models in comparison with the previous study.

Conclusion: The external validation of SPS for Screening Loei Urolithiasis Patients exhibited excellent discrimination and calibration. The overall performance of the models was validated with high accuracy. This model can be used as a screening tool to identify individuals at risk of developing urolithiasis in the Loei region.

Insight Urol 2024;45(2):96-102. doi: 10.52786/isu.a.91

Introduction

Urolithiasis is a common urological problem in Thailand, being particularly prevalent in the northeast region.[1] The major problems associated with urolithiasis are seen in patients with complications as a result of urolithiasis. Early detection and intervention are crucial in the management of urolithiasis, as timely treatment can prevent complications.

Current practice is to separate symptoms into 2 stages: evaluation of metabolism in the 1st

Corresponding author: Krit Santanapipatkul
Address: Division of Urology, Department of Surgery, Loei Hospital, Mueang Loei, Loei 42000, Thailand
E-mail: jozaeuro@yahoo.com
Manuscript received: May 18, 2024
Revision received: November 24, 2024
Accepted after revision: December 9, 2024

episode with recurrent stone formation as secondary a secondary stage. However, it is evident that current standard protocols for screening methods for stones in the urinary tract have yet to be established to facilitate primary prevention. Additionally, metabolic evaluation remains costly and cumbersome and is often only feasible in select hospital settings in Thailand, primarily in medical school-affiliated institutions.[2]

In 2022, Santanapipatkul et al. developed a predictive model for stone formation known as the Stone Probability Score (SPS) for screening urolithiasis patients in Loei. The score included eight parameters: BMI, being of male gender, a family history of urolithiasis, working in farming, working time > 8 hours/day, regular consumption of water spinach, regular consumption of bamboo shoots, eating Laab/Koi more than or equal to 3 times per week. The cut-off of probability for stone formation was 0.46 with a sensitivity of 72.41%, a specificity of 79.52%, and an area under the ROC curve of 0.83.[3] This study aimed to externally validate SPS for screening Loei urolithiasis patients by Santanapipatkul et al. using external patient data.

Materials and Methods

Study design

This study is an external validation of the SPS designed in 2022 by Santanapipatkul et al. to enable the identification of urolithiasis patients in Loei. Cross-sectional data were collected from patients at Loei Hospital between September 28, 2022, and December 31, 2023. Approval for the study protocol was obtained from the Institutional Review Board of Loei Hospital (EC011/2565).

Study patients

Patients diagnosed with urolithiasis at a hospital in the Loei region between September 28, 2022, and December 31, 2023 were enrolled onto the study. A total of 347 patients were included in the study, categorized into two groups. Group 1 consisted of 300 non-migrant urolithiasis patients, who permanently reside in Loei province, while Group 2 comprised 47 immigrant urolithiasis patients, who had migrated into Loei province. Urolithiasis patients were identified based on imaging results, including ultrasound, plain KUB, IVP, or CT scan.

Data collection

The data were collected by face-to-face interview and included the following information: [1] Baseline characteristics and demographic data including sex, body mass index, comorbidities, family history of urolithiasis, past stone treatment, occupation, education, monthly income, working hours, and workplace; [2] Dietary habits including consumption of different types of food and water associated with urolithiasis.

Evaluation variables

The workplace was categorized into two groups: outdoor occupations, involving exposure to sunlight, and indoor occupations, such as office or hospital work. The amount of daily water intake was estimated by asking participants, "How many glasses of water do you drink per day?" (1 glass = 250 ml). Dietary risk factors for urolithiasis were assessed through questions about the frequency of consuming specific foods, categorized into four groups: [1] never (never consumed), [2] seldom (consumed monthly), [3] sometimes (consumed weekly), and [4] always (consumed daily). Patients were asked to recall their consumption of these specific foods over the past year.

Study determinant

The mainly focus of our validation study is to assess the SPS developed by Santanapipatkul et al. for the identification of urolithiasis patients in Loei. This SPS comprises 8 predictors, including BMI, male gender, family history of urolithiasis, occupation as a farmer, working more than 8 hours per day, regular consumption of water spinach and bamboo shoots, and eating Laab/Koi 3 or more times per week. If the probability of stone formation calculated by the SPS is greater than or equal to 0.46, we recommend further investigation for urolithiasis.

Study outcome

The primary outcome of the study was to assess the performance of the predictive model in identifying patients at risk of developing urolithiasis in the Loei region.

Statistical analysis

Statistical analyses conducted using STATA carried out in 3 models. The first model included all 347 patients enrolled onto the study model. These were then categorized into a secondary model. The secondary model consisted of 300

non-migrant urolithiasis patients who permanently reside in Loei province while the last model comprised 47 immigrant urolithiasis patients.

Logistic regression was used to evaluate the performance of all predictive models. The goodness of fit assessed by using Hosmer-Lemeshow test if $p > 0.05$ indicated that the model has high reliability perfect. The multicollinearity of the model was considered using the variance inflation factor (VIF) together with the conditional index. mean VIF of more than 1 is indicative of moderate multicollinearity. A conditional index of >30 the multi-collinearity is strong.

The external validation was performed by considering overall performance including discrimination measures then accuracy assessments, specifically sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), and area under the curve (AUC). If $0.7 \leq AUC < 0.8$ the level of discrimination was acceptable, $0.8 \leq AUC < 0.9$ indicated excellent discrimination and $AUC \geq 0.9$ outstanding discrimination. We also used the Brier score, a quadratic score rule, where the squared differences between actual binary outcome range from 0 indicates a perfect model to 0.25 for a noninformative model. Calibrations refers the agreement between observed outcomes and prediction that consider calibration in the large (CITL), Calibration slope with graphical assessment of calibration is calibration plot.

However, the provided information helps to clearly understand how the model's accuracy was evaluated and referenced to assess its predictive performance in the study.

Results

Three hundred and forty-seven participants were enrolled onto the study, comprising 223 (64%) males and 124 (36%) females. The average BMI was 23.7 kg/m^2 (±3.61) for males, 24.4 kg/m^2 (±4.3) for females, and 23.9 kg/m^2 (±3.9) overall. A notable portion of participants had a family history of urolithiasis (62% males, 60% females). The majority of both males (83%) and females (89%) worked as farmers. Around 60% of males and 56% of females reported working more than 8 hours per day. Consumption of water spinach varied among participants, with "sometimes" being the most common frequency reported. Similarly, bamboo shoot consumption also varied, with "sometimes" being the most frequently reported frequency. About 26% of males and 8% of females reported consuming Laab/Koi 3 or more times per week (Table 1).

Table 1. Demographic data and type of food-related to urolithiasis

Valiables	Group (n=347)	
	Male	Female
Sex n (%)	223 (64)	124 (36)
Age; year (mean±SD)	57.9±11.3	55.5±12.9
	57.0±11.9	
BMI; kg/m^2 (mean±SD)	23.7±3.61	24.4±4.3
	23.9±3.9	
Underlying disease n (%)	134 (60)	80 (65)
Diabetes mellitus	30 (13)	32 (26)
Hypertension	74 (33)	53 (43)
Dyslipidemia	27(12)	22 (18)
Gout	24 (11)	4 (3)
Chronic kidney disease*	69 (31)	31 (25)
Education n (%)		
Equal to or lower than Bachelor's degree	220 (99)	123 (99)
Higher than Bachelor's degree	2 (1)	1 (0.8)
Occupations n (%)		
Farmer	187 (83)	111 (89)
Self-employed	16 (7)	4 (3.2)
Others	3 (1)	1 (0.8)
History of family urolithiasis n (%)		
Yes	139 (62)	75 (60)
No	84 (38)	49 (40)
Working times n (%)		
> 8 hours/day	133 (60)	70 (56)
≤ 8 hours/day	90 (40)	54 (44)
Eate Laab/Koi 3 or more times/ week n (%)		
Yes	58 (26)	0 (8)
No	165 (74)	114 (92)
Water spinach n (%)		
Never	39 (17))	25 (20)
Seldom	84 (38)	47 (38)
Sometimes	81 (36)	43 (35)
Always	19 (9)	9 (7)
Bamboo shoot		
Never	10 (5)	9 (7)
Seldom	25 (11)	18 (15)
Sometimes	119 (53)	70 (56)
Always	69 (31)	27 (22)

SD = standard deviation, BMI = body mass index
*Chronic kidney disease is defined by a glomerular filtration rate less than 60 ml/min/1.73 m^2 for at least 3 months.[4]

The following predictive model that separated 3 model external validation were

Model 1: Overall total of 347 participated in the study:

SPS = 1/ {1+ EXP (-[-1.002*Male) + (0.072*BMI) + (1.172*Family history of urolithiasis) + (1.893*Working time > 8 hours/day) + (1.279*Eats Laab/Koi 3 or more times/week) + (0.170*always eat water spinach) + (2.278*sometimes eats Bamboo shoots)])}.

Model 2: 300 non-migrant urolithiasis patients:

SPS = 1/ {1+ EXP (-[-1.294*Male) + (0.083*BMI) + (0.897*Family history of urolithiasis) + (1.928*Working time > 8 hours/day) + (1.343*Eats Laab/Koi 3 or more times/week) + (0.234*always eat water spinach) + (2.211*sometimes eats Bamboo shoots)])}.

Model 3: 47 immigrant urolithiasis patients, who migrated to Loei province:

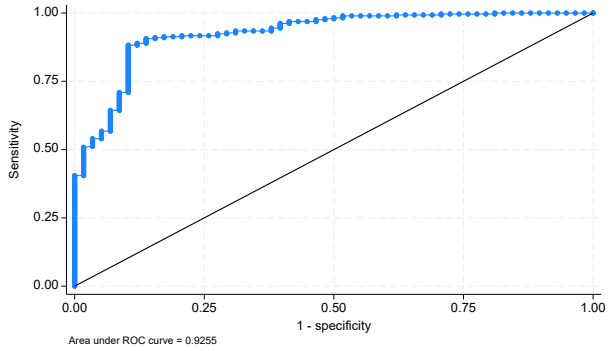
SPS = 1/ {1+ EXP (-[0.529*Male) + (-0.082*BMI) + (71.289*Family history of urolithiasis) + (1.162*Working time > 8 hours/day) + (17.200*Eats Laab/Koi 3 or more times/week) + (0.301*always eat water spinach) + (36.621*sometimes eats Bamboo shoots)])}.

For each individual model in our study, the predicted probability of SPS was calculated using these formulae, leading to the AUC, goodness of

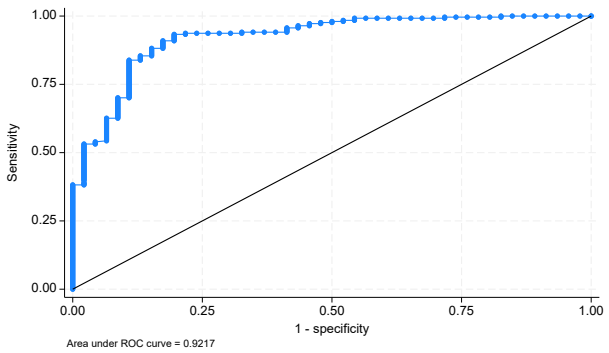
Table 2. The accuracy assessment and goodness of fit of the predictive models in the validation of SPS.

Number in cohort	Model	AUC (%)	Standard error	Asymptotic normal (95% confidence interval)	P-value	VIF	CI
347	1	92.6	0.019	88.9-96.3	0.336[a]	1.05[b]	20.67[c]
300	2	92.3	0.022	87.9-96.5	0.648[a]	1.05[b]	20.70[c]
47	3	95.5	0.029	89.8-1.00	0.941[a]	1.09[b]	20.62[c]

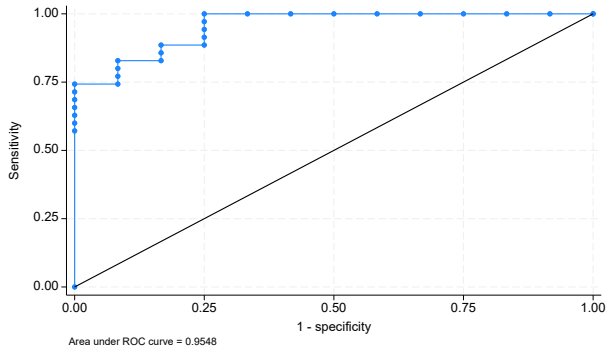
[a] = Hosmer-Lemeshow test, [b] = Variance inflation factor, [c] = Conditional index
SPS = Stone Probability Score, AUC = area under the curve, VIF = variance inflation factor, CI = conditional index



(A)



(B)



(C)

Figure 1. Area under Receiver operating characteristic curve for Model 1-3, A. Area under Receiver operating characteristic curve for Model 1 (347 Loei urolithiasis showing 92.6 % (95% CI, 89.9-96.3)), B. Area under Receiver operating characteristic curve for Model 2 (300 non-migrant urolithiasis patients who permanently reside in Loei province showing 92.3(95% CI 87.9-96.5)), C. Area under Receiver operating characteristic curve for Model 3 (47 immigrant urolithiasis patients showing 95.5(95% CI 89.8-100))

Table 3. Performance measure of external validation all 3 models.

Performance measure	External validation		
	Model 1 (n=347)	Model 2 (n=300)	Model 3 (n=47)
Overall			
Brier score	0.072	0.068	0.058
Cragg & Uhler's R^2(%)	79.6	56.7	79.6
Discrimination			
C stat	92.6	92.3	95.5
Discrimination slope	0.486	0.475	0.695
Sensitivity (%)	96.9 (94.2-98.6)	97.6 (94.9-99.1)	100 (90-100)
Specificity (%)	53.4 (39.9-66.7)	50 (34.9-65.1)	75 (42.8-94.5)
PPV (%)	91.2 (87.5-94.1)	91.5 (87.5-94.5)	92.1 (78.6-98.3)
NPV (%)	77.5 (61.5-89.2)	79.3 (60.3-92)	100 (66.4-100)
Calibration			
Calibration in the large	0	-0	0
Calibration slope	1	1	1
H-L test	0.336	0.648	0.941

PPV = positive predictive value, NPV = negative predictive value

fit by Hosmer-Lemeshow statistical and multicollinearity issue about VIF and CI as shown in table 2 and fig. 1. The AUCs in all 3 models are much higher in our validation.

External validation

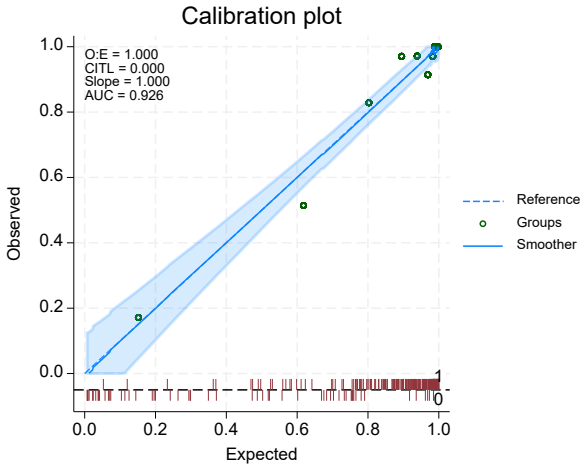
From the table 2, predictive accuracy of the three models as described by the system proposed by Santanapipatkul et al was significantly higher in our study. Firstly, a total of 347 patients in the original data set was used in the first model in table 1 showed an AUC of approximately 92.6% (95% CI 88.9-96.3). Secondly, the first model was split into 2 groups of urolithiasis patients, 300 non-migrant who permanently reside in Loei province showed an AUC of 92.3 (95% CI 87.9-96.5) and the final model had 47 immigrant urolithiasis patients and showed 95.5 (95% CI 89.8-100) if an AUC greater than 0.9 suggests outstanding discrimination between the presence and absence of SPS.

From the table 3, overall model performance in the study of 347 patients was greater than in the other two models. We evaluated the overall performance using the Brier score, which was found to be 0.072, according to Cragg & Uhler's R^2 about 79.6%. For the discrimination measure have shown sensitivity approximately 96.9% (95%CI 94.2-98.6) while specificity may be 53.4% (39.9-66.7) when we consider about PPV is 91.2% (95%CI 87.5-94.1) and NPV probably 77.5% (95%CI 61.5-89.2).

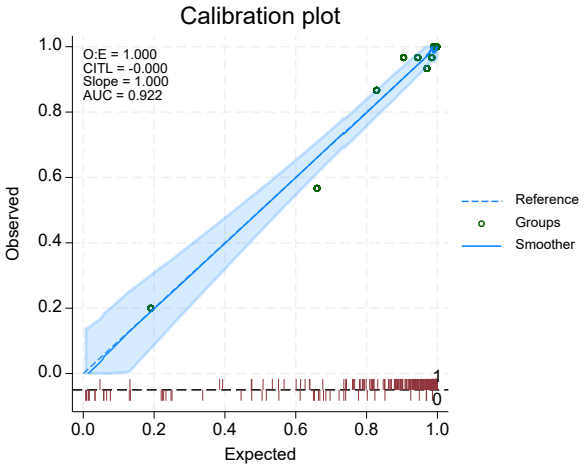
The secondary model was split and from the group of 300 patients the Brier score was 0.068 whereas Cragg & Uhler's R^2 less than the first model about 56.7%. This model has highly sensitivity 97.6% (95%CI 94.9-99.1) but the specificity was poorer at about 50% (95%CI 34.9-65.1) while the PPV and NPV were 91.5% (95%CI 87.5-94.5), 79.3% (95%CI 60.3-92) respectively.

Finally, the model from the 47 immigrants with urolithiasis had a Brier score of 0.058 but coincidentally the Cragg & Uhler's R^2 is the same as in the first model at about 79.6%. The discrimination measure showed that the 100% of sensitivity (95%CI 90-100) was not consistent with the specificity at only 75% (95%CI 42.8-94.5) with a PPV of 92.1 (95%CI 78.6-98.3) and NPV of 100% with a 95%CI 66.4-100.

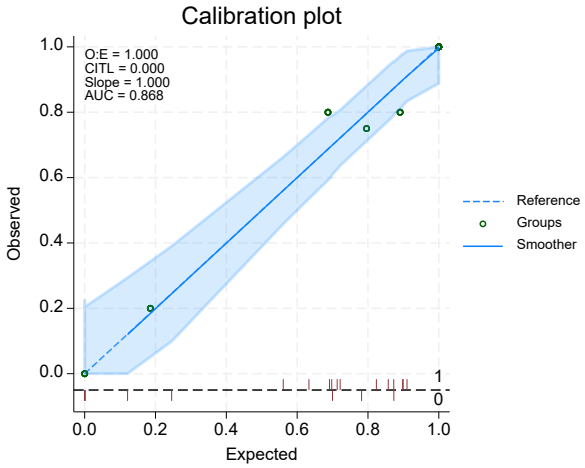
The calibration of all three models by the E:O ratio is 1.000 meaning the ratio of predicted accuracy (E) to actual accuracy (O) is 1.000 indicating that the model predicts accurately or predictions closely match the actual outcomes. CITL is 0.000 meaning a phenomenon that is fully independent of discrimination value is 0.000. This could suggest that the first model and last model are well-calibrated including being balanced and unbiased on a large scale. However, the second model was -0.000 but the discrimination will be unaffected and calibration slope is 1.000 indicating the model calibration is perfect without any bias or adjustment needed as shown in Fig. 2.



(D)



(E)



(F)

Figure 2. Calibration plot, D. Calibration plot in 347 Loei residents with urolithiasis, E. Calibration plot in 300 Non-migrant Loei residents with urolithiasis, F. Calibration plot in 47 Immigrant Loei residents with urolithiasis

Discussion

Loei has a very high prevalence of urolithiasis patients[5] which is negatively associated with health and expenditures. To decrease the cost and prevalence of urolithiasis, prevention and control of stone formation are important factors.

Todate many studies have identified predictors for the recurrence of stone after the first episode[6,7] However, there is limited research into screening for primary prevention of stone formation. One notable study in this regard is by Okita et al., which introduced a nomogram for the prediction of potential candidates for stone episodes.[8] This nomogram suggests an optimal cut-off value for the probability of a stone episode (>28%), demonstrating a sensitivity of 79%, a specificity of 76%, and an area under the curve of 0.860, indicating satisfactory discrimination for screening of a population. Based on the current data, it is apparent that multiple risk factors are linked to stone formation. This observation indicates that each geographical area may have distinct risk factors contributing to the occurrence of stones.[9] Hence, we conducted a study utilizing the SPS for screening Loei urolithiasis patients.

This study found that among the 347 Loei patients with urolithiasis, the predictive model demonstrated good discrimination ability and calibration in an external validation setting. The accuracy of the model showed an AUC of 92.6% (95%CI 88.9-96.3). Overall performance measured by Brier score was 0.0072 while Cragg & Uhler R^2 about 79.6% with discrimination measurement, a sensitivity of 96.9% (95%CI 94.2-98.6), a specificity of approximately 53.4% (95%CI 39.9-66.7), a positive predictive value of 91.2% (95%CI 87.5-94.1) and a negative predictive value of 77.5% (95%CI 61.5-89.2).

Additionally, when analyzing subgroups within the Loei urolithiasis population, including non-migrants we found the Brier score to be 0.068 whereas Cragg & Uhler's R^2 was less than

the first model at 56.7%. This model had a high sensitivity of 97.6% (95%CI 94.9-99.1) but poor specificity of about 50% (95%CI 34.9-65.1) while the PPV and NPV were 91.5% (95%CI 87.5-94.5), and 79.3% (95%CI 60.3-92), respectively. Analysis of the immigrant Loei patients showed a Brier score of 0.058 but coincidentally the Cragg & Uhler's R^2 is the same as in the first model at about 79.6%. The discrimination measures were 100% sensitivity (95%CI 90-100) and a lower specificity of 75% (95%CI 42.8-94.5), a PPV of 92.1 (95%CI 78.6-98.3) but a NPV of 100% but wide of 95%CI 66.4-100. In summary, we suggest that these models still exhibited good discrimination ability and calibration.

These results suggest that the SPS may be applied effectively in the northeastern region of Thailand, especially in the Loei province. The extended area outside Loei is due to similar dining habits and traditions in other provinces in the northeast.

One limitation of our study is that its cross-sectional nature only provides a snapshot of a population at a specific point in time, without a longitudinal follow-up., however, the aim of the study was to validate a scoring system which had been developed in previous research so that in this case the cross-sectional approach was valid.

Conclusions

The external validation of SPS for the Screening of Urolithiasis Patients in Loei showed that the screening process exhibits excellent discrimination and calibration. The overall performance of the models was validated with a high degree of accuracy. This model can be used as a screening tool to identify individuals at risk of developing urolithiasis in the Loei region.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Yanagawa M, Kawamura J, Onishi T, Soga N, Kameda K, Sriboonlue P, et al. Incidence of urolithiasis in northeast Thailand. Int J Urol 1997;4:537-40.
2. Goldfarb DS, Arowojolu O. Metabolic evaluation of first-time and recurrent stone formers. Urol Clin North Am 2013;40:13-20.
3. Santanapipatkul K, Jantakun W. Development and validation of a score for the screening of Loei urolithiasis patients. Insight Urol 2022;43:13-24.
4. Levey AS, Eckardt K, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. Definition and classification of chronic kidney disease: A position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int 2005;67:2089-100.
5. Santanapipatkul K, Jantakun W, Tanthanuch M. Urinary tract stone analysis in Loei province. Thai J Urol 2019;40:9-18.
6. Eisner BH, Goldfarb DS. A nomogram for the prediction of kidney stone recurrence. J Am Soc Nephrol 2014;25:2685-7.
7. Rule AD, Lieske JC, Li X, Melton LJ 3rd, Krambeck AE, Bergstralh EJ. The ROKS nomogram for predicting a second symptomatic stone episode. J Am Soc Nephrol. 2014;25:2878-86.
8. Okita K, Hatakeyama S, Imai A, Tanaka T, Hamano I, Okamoto T, et al. Stone episode prediction: development and validation of the prediction nomogram for urolithiasis. Int J Urol 2020;27:344-9.
9. Liu Y, Chen Y, Liao B, Luo D, Wang K, Li H, et al. Epidemiology of urolithiasis in Asia. Asian J Urol 2018;5:205-14.