TURKISH JOURNAL OF EMERGENCY MEDICINE

Original Article

Access this article online

Quick Response Code:

Website: https://turkjemergmed.com/

DOI: 10.4103/tjem.tjem_277_25

Evaluation of STONE, modified STONE, and CHOKAI scores for predicting ureteric stone disease in Indian emergency patients: A prospective observational study

Milan Manu[1], Naman Agrawal[1]*, Roshan Mathew[1], Saroj Kumar Pati[2], Amit Kumar Mishra[3], Debendra Kumar Tripathy[1]

Departments of [1]Trauma and Emergency, [2]Radiology and [3]Community Medicine, All India Institute of Medical Science, Raipur, Chhattisgarh, India
**Corresponding author*

Abstract:

OBJECTIVE: Acute flank pain is a common complaint in the emergency department (ED), with urolithiasis being a major cause. This prospective observational study evaluated the performance of the STONE, modified STONE, and CHOKAI scores in predicting ureteric stones in Indian ED patients with acute flank pain.

METHODS: The study included adult patients who underwent noncontrast computed tomography for suspected urolithiasis. Clinical scores were calculated independently, and their diagnostic accuracy was assessed using receiver operating characteristic (ROC) curve analysis.

RESULTS: Ureteric stones were diagnosed in 58.3% of the patients. The CHOKAI score demonstrated the highest accuracy (area under the curve [AUC] 0.89), followed by the modified STONE (AUC 0.84) and STONE (AUC 0.65) scores. Optimal cutoffs were identified using Youden's index. DeLong's test revealed that CHOKAI and modified STONE scores outperformed the STONE score, while the difference between CHOKAI and modified STONE was not significant. The Hosmer–Lemeshow test showed good calibration for the CHOKAI and modified STONE scores.

CONCLUSIONS: The CHOKAI score demonstrated the highest diagnostic accuracy in our cohort and may be the preferred tool for predicting ureteric stones in this clinical setting. Further validation in larger multicenter studies is warranted.

Keywords:

Abdominal pain, scores, urolithiasis

Submitted: 26-07-2025
Revised: 04-09-2025
Accepted: 03-10-2025
Published: 01-01-2026

ORCID:
MM: 0009-0007-7117-1500
NA: 0000-0001-8239-3351
RM: 0000-0001-7047-0399
SKP: 0000-0002-9120-5452
AKM: 0000-0003-3868-3452
DKT: 0000-0003-0346-1374

Address for correspondence:
Dr. Naman Agrawal, Department of Trauma and Emergency, All India Institute of Medical Science, Raipur, Chhattisgarh, India.
E-mail: namanemaiims@gmail.com

Introduction

Acute flank pain is one of the leading emergency department (ED) complaints that contributes to a sizeable health burden at the national and global levels.[1] Urolithiasis is a common cause of acute flank pain. Studies have shown that the prevalence of urolithiasis in Indian patients is approximately 12%, with a significant proportion of patients at risk for renal damage or functional loss if left untreated.[2] The high sensitivity and specificity of noncontrast computed tomography (NCCT) make it the gold standard for diagnosing ureteric calculi in the ED, although radiation exposure and

How to cite this article: Manu M, Agrawal N, Mathew R, Pati SK, Mishra AK, Tripathy DK. Evaluation of STONE, modified STONE, and CHOKAI scores for predicting ureteric stone disease in Indian emergency patients: A prospective observational study. Turk J Emerg Med 2026;26:55-61.

Box-ED Section

What is already known on the study topic?

- STONE, modified STONE, and CHOKAI scores in detecting ureteric stones have been studied in Western populations

What is the conflict on the issue? Is this important for readers?

- The STONE score uses race (Black/non-Black), which loses its applicability in Indian patients. The modified STONE and CHOKAI scores address these gaps using sonographic pelvicalyceal obstruction instead of race and history of prior renal stones for their assessment. However, these have not been validated in Asian populations.

How is this study structured?

- This was a single-center, prospective cohort study that included data from 72 patients in India.

What does this study tell us?

- The CHOKAI and modified STONE scores effectively predicted ureteric stones in Indian patients presenting with acute flank pain
- The CHOKAI score demonstrated the highest diagnostic accuracy, making it the preferred tool for Indian clinical settings.

costs require doctors to use NCCT more selectively, with the risk of missing more serious causes of flank pain.[3] This clinical challenge is amplified in the Indian context, given the high burden of urolithiasis, out-of-pocket payments for health care, and resource-limited settings. Various bedside clinical scores have been developed to predict urolithiasis, thus enabling healthcare providers to triage imaging needs and decrease the number of unnecessary scans.[4,5]

Despite the growing interest in this field, existing prediction scores have limitations. The STONE score uses race (Black/non-Black), which loses its applicability in Indian patients.[4] The Modified STONE and CHOKAI scores address these gaps using sonographic pelvicalyceal obstruction instead of race and history of prior renal stones for their assessment.[6,7] Most studies validating these scores have come from Western and East Asian populations, but South Asian data remain scarce. Indian populations exhibit distinct epidemiological patterns, variations in access to health care, and patient symptoms. Therefore, there is a need to evaluate various scoring systems to predict urolithiasis in Indian patients presenting with acute flank pain.

This study investigated the performance of the STONE, modified STONE, and CHOKAI scores in detecting ureteric stones in patients who visited the ED with acute flank pain at a tertiary care hospital in India. This study aimed to guide evidence-based imaging decisions, while ensuring patient safety.

Methods

This study was performed in the ED of a tertiary care academic hospital. The ED treats approximately 50,000 patients annually, including a mix of medical, surgical, and trauma cases. Urolithiasis represents a major proportion of cases presenting with acute flank pain. The current standard for flank pain includes imaging with NCCT after the initial clinical evaluation of the patient.

This single-center, prospective, observational study was conducted over 18 months at our institution. Ethical clearance was obtained from the AIIMS Raipur Institutional Ethics Committee (approval no. AIIMS/RPR/IEC/2023/473 dated 14.12.2023) before study initiation. Written informed consent was obtained from all participants before their inclusion in the study. According to Song *et al*., the sensitivity of the modified STONE score was 85.9%. Considering a relative precision of 10% and a confidence level of 95%, the minimum number of participants required for the study was 64. Considering a nonresponse rate of 10%, the final sample size was 72.[8] This study was designed and reported in accordance with the Standards for Reporting of Diagnostic Accuracy Studies 2015 guidelines.

The study included a convenience sample of adult patients (aged ≥18 years) who visited the ED with acute flank pain as a chief complaint and underwent NCCT kidney, ureter, and bladder (KUB) as part of their standard diagnostic evaluation. Pregnant women, patients with morbid obesity (BMI ≥40), patients with a history of urological surgery or interventions, patients with known chronic kidney disease, postrenal transplant recipients, and critically ill patients requiring immediate resuscitation were excluded from the study.

The emergency team performed a detailed clinical evaluation, including a history, physical examination, and point-of-care ultrasound (POCUS). Demographic information, symptom characteristics, and laboratory results were collected using a structured data collection sheet. STONE, modified STONE, and CHOKAI scores were calculated for each patient [Table 1] independently by the investigators using prospectively collected data without knowledge of NCCT findings. Emergency physicians performed bedside ultrasonography using a portable device (SONOSITE M-Turbo) to detect hydronephrosis. Bedside ultrasound examinations were performed by one of five emergency physicians who had completed standardized training in POCUS for renal applications. All patients underwent NCCT KUB

Table 1: Description of scores

Characteristics	Score
STONE score	
Sex	
Female	0
Male	2
Duration of pain presentation (h)	
>24	0
6–24	1
<6	3
Race	
Black	0
Nonblack	3
Nausea and vomiting	
None	0
Nausea alone	1
Vomiting	2
Hematuria by dipstick	
Absent	0
Present	3
Total score	0–13
Modified STONE score	
Sex	
Female	0
Male	2
Duration of pain presentation (h)	
>24	0
6–24	1
<6	3
Obstruction (pelvicalyceal dilatation)	
No	0
Yes	3
Nausea and vomiting	
None	0
Nausea only	1
Vomiting	2
Hematuria on urine dipstick	
Absent	0
Present	3
Total score	0–13
CHOKAI score	
Distension of the kidney capsule (nausea and vomiting)	
Yes	1
No	0
Hydronephrosis	
Yes	4
No	0
Occult blood in urine	
Yes	3
No	0
Kidney stone history	
Yes	1
No	0
Sex	
Female	0
Male	1
Age (years)	
<60	1
>60	0
Diminution of pain within 6 h	
Yes	2
No	0
Total score	0–13

during the same ED visit. Upon undergoing NCCT KUB, the radiologist evaluated the presence of ureteric stones, their dimensions and exact positions, and alternate diagnoses, if any. The radiologists were blinded to the clinical data. All data points required for calculating the scores and assessing the outcomes were completed for the patients included. No missing data were obtained from the analysis.

Data were analyzed using the IBM SPSS Statistics version 26 (IBM Corp., Armonk, NY, USA). Continuous variables were summarized using mean ± standard deviation if normally distributed, or median with interquartile range if not normally distributed. Categorical variables were expressed as frequencies and percentages. Differences between patients with and without confirmed ureteric stones were assessed using the Chi-square test or Fisher's exact test for categorical variables, and the independent *t*-test or Mann–Whitney *U*-test for continuous variables. The diagnostic performances of the STONE, Modified STONE, and CHOKAI scores were evaluated using the area under the receiver operating characteristic (ROC) curve. Youden's index was used to determine the optimal cutoff points. The area under the curve (AUC) between the scores was compared using DeLong's test to evaluate the differences in diagnostic accuracy. The calibration of the predicted probabilities was assessed using the Hosmer–Lemeshow goodness-of-fit test. Statistical significance was set at $P < 0.05$.

Results

The study was conducted over an 18-month period, from January 2023 to June 2024, in the ED. A total of 72 patients with acute flank pain were enrolled in this study [Figure 1]. Of these, 42 patients (58.3%) were diagnosed with ureteric stones, while 30 (41.7%) were diagnosed with other conditions. The median age of the participants was 30 (26–37) years. Of all patients, 55.6% were men, with a higher proportion of men in the ureteric stone group. Approximately one-third of the patients had a history of renal stone disease. The average duration of flank pain in the cohort was 20 h, with most patients (61.1%) presenting with 6–24 h of onset. Very few patients presented with pain within 6 h

of pain onset. Nausea and vomiting were more common in patients with ureteric stones than in those without stones (78.6% vs. 53.3%, respectively). Hematuria on urine dipstick examination was more common in patients with stones (47.6%). Hydroureteronephrosis (HDUN) on POCUS was noted in a significantly higher proportion of patients in the stone-positive group than in those without ureteric stones (78.6% vs. 16.7%, respectively). Most patients in this study were discharged from the ED after symptom relief. The baseline characteristics of the study population are shown in Table 2.

The median STONE, modified STONE, and CHOKAI scores were significantly higher in patients diagnosed with ureteral stones than in those without ureteral stones. ROC curve analysis revealed that the CHOKAI score had the highest diagnostic accuracy, with an AUC of 0.89 (95% CI, 0.81–0.96), followed by the modified STONE score (AUC: 0.84 (95% CI, 0.74–0.93) and the STONE score (AUC: 0.65 (95% CI, 0.5–0.8)). The optimal thresholds of the three scores were identified using Youden's index to optimize accuracy. Comparing the performance of the scores using DeLong's test revealed that the CHOKAI and modified STONE scores were superior to the STONE score. However, the difference between the CHOKAI and modified STONE scores was not statistically significant. Calibration analysis using the Hosmer–Lemeshow test revealed a good model fit for the CHOKAI and modified STONE scores. The findings are presented in Table 3 and illustrated in Figure 2.

Table 2: Baseline characteristics

Variables	Diagnosis of ureteric stones		Total (*n*=72)
	Yes (*n*=42)	No (*n*=30)	
Age (years)	31.5 (27.2–40.0)	29.5 (26.0–32.7)	30 (26.0–37.0)
Gender (%)			
Males	61.90	53.30	55.6
Females	38.10	46.70	44.4
Duration of flank pain (h)	17 (10–48)	21 (8–48)	20 (10–48)
<6	2.4	3.3	2.8
6–24	66.7	53.3	61.1
>24	31	43.3	36.1
Hematuria present (%)	47.6	33.3	41.7
Nausea and vomiting (%)			
Nil	21.4	46.7	31.9
Only nausea	35.7	10	25
Vomiting	42.9	43.3	43.1
History of kidney stone (%)	31	36.7	33.3
Hydronephrosis on POCUS (%)	78.6	16.7	52.8
STONE score	7 (2.0)	6 (2.0)	7 (3.0)
Modified STONE score	7 (2.0)	3 (3.0)	5 (4.0)
CHOKAI score	9 (2.0)	5 (2.0)	8 (3.5)
Alternate diagnoses (%)			
Pyelonephritis	-	9 (30)	
Renal calculi	-	9 (30)	
Ovarian cyst	-	1 (3.3)	
No diagnoses reached	-	11 (36.7)	
ED disposition (%)			
Admission	0	26.7	11.1
Discharge	100	73.3	88.9

ED: Emergency department, POCUS: Point-of-care ultrasound

Discussion

In our study, the CHOKAI score demonstrated the highest diagnostic accuracy for predicting ureteric calculi in patients presenting with acute flank pain, followed by the modified STONE and STONE scores. All three scores may prove helpful in patients with flank pain, but the CHOKAI score appears to be the most effective in the Indian population.

The findings of this study suggest that clinical scores incorporating a broader range of clinical and imaging findings may be more useful in the evaluation of suspected ureteric colic. The superior performance of the CHOKAI and modified STONE scores highlights the value of integrating POCUS and patient history into clinical prediction tools. The STONE score, which includes race as a variable instead of POCUS findings of HDUN, had a poor performance, confirming the futility of race as a discriminatory variable.

Our results are consistent with those of previous studies by Rehman *et al*.[9] and Acar *et al*.,[10] who concluded that all three clinical scoring systems – STONE, modified STONE, and CHOKAI – demonstrate fair to good predictive value in identifying ureteric stones in patients presenting with acute flank pain. The high performance of the CHOKAI score in our study is in alignment with the findings of Fukuhara *et al*.[11] who first developed the CHOKAI score in Japanese patients. Rohat *et al*. also found that CHOKAI outperformed STONE in diagnosing ureteric stones, with significantly higher scores in stone-positive patients.[12] In both our study

Table 3: Diagnostic performance of clinical scores

Score	AUC	95% CI	Optimal threshold	Sensitivity (95% CI)	Specificity (95% CI)	Youden's index	Accuracy (%)
STONE score	0.65	0.5–0.8	6	76 (63–89)	47 (29–64)	0.23 (0.008–0.49)	64
Modified STONE score	0.84	0.74–0.93	5	86 (75–96)	70 (54–86)	0.56 (0.36–0.75)	79
CHOKAI score	0.89	0.81–0.96	8	88 (78–98)	80 (66–94)	0.68 (0.50–0.85)	85

Pairwise comparison of AUC (Delong's test). CHOKAI versus. STONE *P*=0.002, Modified STONE versus stone *P*=0.011, CHOKAI versus modified Stone *P*=0.245. CI: Confidence interval. AUC: Area under the curve

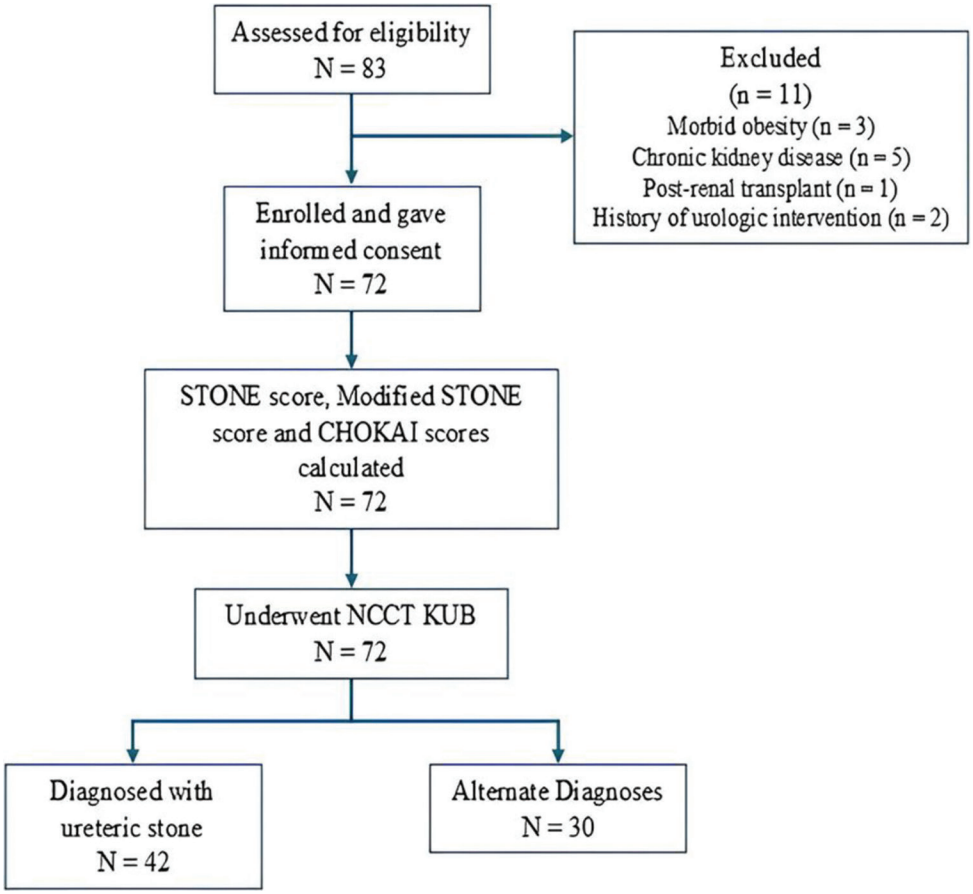

Figure 1: The Standards of Reporting of Diagnostic Accuracy Studies diagram (NCCT KUB: Noncontrast computer tomography of kidney, ureter, and bladder)

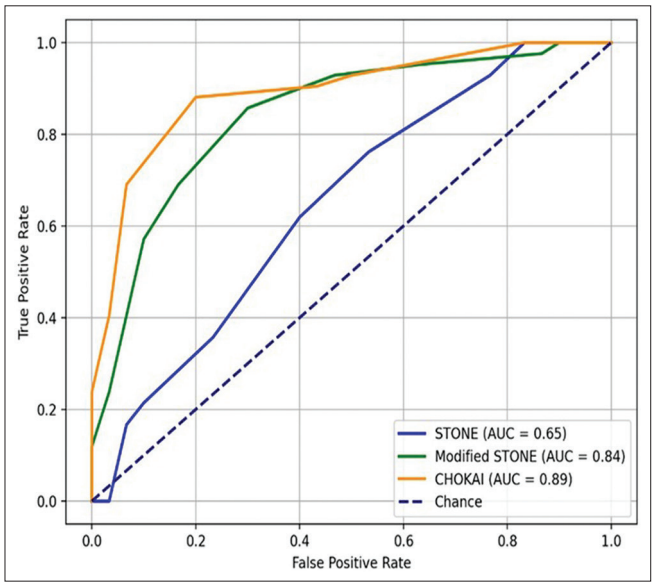

Figure 2: Receiver operator curves for STONE, Modified STONE, and CHOKAI scores AUC: Area under the curve)

and the study by Eraybar *et al*., the CHOKAI score demonstrated superior diagnostic performance over the STONE score in predicting ureteric stones, with AUCs of 0.89 and 0.788, respectively. While Eraybar *et al*. found the optimal CHOKAI cutoff to be > 7, our study identified 8 as the optimal threshold, reflecting strong but slightly differing predictive thresholds in distinct populations.[13] Uzun *et al*.[7] demonstrated that the modified STONE score performed better than the original STONE score when sonographic findings were incorporated instead of race. Moreover, the AUC values of the scores in our study were comparable to those in the Turkish and Japanese cohorts.[10,11] The role of emergency physician-performed POCUS in detecting HDUN also parallels the results of Pathan *et al*., reinforcing the value of ultrasound as an adjunct in the diagnostic process.[14]

Some previous studies reported a higher discriminatory power for the original STONE score than that in our study. Wang *et al*. observed a higher AUC for the STONE score (0.78) in a large US cohort.[5] Kim *et al*. found an AUC of 0.92 for the STONE score in the Korean population.[15] In addition, an Asian ED cohort highlighted only moderate STONE accuracy (AUC 0.78), but it was still above the 0.65 noted in our population.[16] These differences may be attributed to several methodological and demographic factors. Race is a prominent component of the original STONE score. As the Indian sample lacked black patients, that variable offered no discriminative value and likely drove the performance of the STONE

score downward, as against the US and Korean cohorts.[5,15] Notably, cohorts with a higher prevalence of urolithiasis (e.g., 79% in Korea) inherently boost positive predictive values.[15]

Our study stands apart because of its focus on the Indian ED population, a demographic largely underrepresented in the prior literature. Unlike previous studies, our prospective, single-center methodology allowed for controlled data collection and minimized documentation bias. Furthermore, we incorporated a direct comparison of the three scoring systems in the same cohort, which has not been widely performed.

Limitation

The primary limitation of our study is its relatively small sample size ($n = 72$), which was recruited from a single center. This may limit the precision of our diagnostic accuracy estimates and constrain the generalizability of our findings to the broader Indian population. The recruitment rate was influenced by stringent exclusion criteria designed to ensure a homogenous cohort and the logistical challenges (lack of availability of radiologist) inherent in prospective research in a busy ED. However, as the first prospective validation of these scores in this demographic, this study provides crucial preliminary evidence and a strong rationale for future adequately powered multicenter studies. Our study was designed as an initial, exploratory investigation into the performance of these scores in a South Asian population, where such data is critically lacking. While the study may be underpowered for highly precise estimation of individual sensitivity and specificity values, it was sufficiently powered to demonstrate the primary outcome. Interobserver variability in POCUS interpretation was not formally assessed, although all participating physicians underwent standardized training to mitigate this variability. Sequira *et al*. highlighted the significant regional variation in urolithiasis prevalence across India, which is influenced by factors such as climate, diet, and water composition. These differences may impact the performance and optimal cutoffs of diagnostic scores such as STONE and CHOKAI, underscoring the need for population-specific validation, as demonstrated in our study.[17]

It is necessary to evaluate whether the scores are valid across diverse Indian EDs with varying patient demographics and ultrasound expertise among doctors. The implementation of these scores in the management of acute flank pain further requires cost-effectiveness analyses and clinical outcome studies to establish their ability to reduce unnecessary imaging while avoiding critical diagnosis misses.

Conclusions

In our study cohort, the CHOKAI score demonstrated the most favorable balance of sensitivity and specificity, suggesting that it may be the most effective tool among the three for this specific patient population. Additional multicenter research with larger participant groups should be conducted to confirm these results and evaluate their potential use in clinical pathways throughout India.

Author contributions statement

- MM, NA, RM, DKT: Writing – review and editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization
- MM, NA, AKM, SKP: Investigation, Formal analysis, Data curation.

Conflicts of interest

None Declared.

Ethical approval

Institutional Ethics Committee – All India Institute of Medical Sciences, Raipur, CG, India; Approval number AIIMS/RPR/IEC/2023/473 (date 14/12/2023).

Funding

None.

References

1. Brewer BJ, Golden GT, Hitch DC, Rudolf LE, Wangensteen SL. Abdominal pain. An analysis of 1,000 consecutive cases in a University Hospital emergency room. Am J Surg 1976;131:219-23.
2. Guha M, Banerjee H, Mitra P, Das M. The demographic diversity of food intake and prevalence of kidney stone diseases in the Indian continent. Foods 2019;8:37.
3. Imaging of Flank Pain: Readdressing State-of-the-art | Emergency Radiology. Available from: https://link.springer.com/article/10.1007/s10140-016-1443-9. [Last accessed on 2025 Jun 27].
4. Moore CL, Bomann S, Daniels B, Luty S, Molinaro A, Singh D, *et al*. Derivation and validation of a clinical prediction rule for uncomplicated ureteral stone – the STONE score: Retrospective and prospective observational cohort studies. BMJ 2014;348:g2191.
5. Wang RC, Rodriguez RM, Moghadassi M, Noble V, Bailitz J, Mallin M, *et al*. External validation of the STONE score, a clinical prediction rule for ureteral stone: An observational multi-institutional study. Ann Emerg Med 2016;67:423-32.e2.
6. Fukuhara H, Kobayashi T, Takai S, Tawara T, Kikuta M, Sugiura A, *et al*. External validation of the CHOKAI score for the prediction of ureteral stones: A multicenter prospective observational study. Am J Emerg Med 2020;38:920-4.
7. Uzun A, Korkut M, Kartal M, Bedel C. Evaluation of modified STONE score in patients presenting to the emergency department with flank pain. Urol Sci 2020;31:221.
8. Song C, Lim J, Oh SH, Kim HJ. External validation of the STONE score and the modified STONE score for the patients with suspicious urinary stone in the emergency department. J Korean Soc Emerg Med 2022;33:573-80.
9. Rehman MA, Ahmed T, AlZaabi A, Almarzooqi MA, Husain MM, AlNuaimi R, *et al*. Comparing CHOKAI and STONE scores in predicting ureteral stones in the emergency department: a prospective multicenter diagnostic study. J Emerg Med 2025;78:302-12.
10. Acar YA, Uysal E. External validation of STONE, modified STONE, and CHOKAI scores for the diagnosis of ureteral stones in the Turkish population. Hong Kong J Emerg Med 2023;30:23-9.

11. Fukuhara H, Ichiyanagi O, Midorikawa S, Kakizaki H, Kaneko H, Tsuchiya N. Internal validation of a scoring system to evaluate the probability of ureteral stones: The CHOKAI score. Am J Emerg Med 2017;35:1859-66.
12. Rohat AK, Kurt E, Şenel Ç. The comparison of two prediction models for ureteral stones: CHOKAI and STONE scores. Am J Emerg Med 2021;44:187-91.
13. Eraybar S, Yuksel M. The prospective evaluation of the effectiveness of scoring systems in the emergency department in cases with suspected ureteral stones: STONE? CHOKAI? Am J Emerg Med 2021;49:94-9.
14. Pathan SA, Mitra B, Mirza S, Momin U, Ahmed Z, Andraous LG, *et al*. Emergency physician interpretation of point-of-care ultrasound for identifying and grading of hydronephrosis in renal colic compared with consensus interpretation by emergency radiologists. Acad Emerg Med 2018;25:1129-37.
15. Kim B, Kim K, Kim J, Jo YH, Lee JH, Hwang JE, *et al*. External validation of the STONE score and derivation of the modified STONE score. Am J Emerg Med 2016;34:1567-72.
16. Ang JS, Wong SY, Ooi CK. Use of STONE score to predict urolithiasis in an Asian emergency department. J Acute Med 2022;12:53-9.
17. Sequira L, Hebbar S, Ravishanker N. Prevalence and associated risk factors of urolithiasis in India, a systematic review. Biomed India 2023;43:572-6.