International Surgery Journal
Natesan SR et al. Int Surg J. 2022 Jan;9(1):23-26
http://www.ijsurgery.com
pISSN 2349-3305 | eISSN 2349-2902

DOI: https://dx.doi.org/10.18203/2349-2902.isj20214807

Original Research Article

Evaluation of validity of scoring systems in diagnosing acute appendicitis: a single centre prospective observational study

Rakesh Natesan S., Naren Kumar A.*

Department of General Surgery, Shri Satya Sai Medical College and Research Institute, Sri Balaji Vidyapeeth (deemed to be university), Nellikuppam, Tamil Nadu, India

Received: 16 November 2021
Accepted: 01 December 2021

***Correspondence:**
Dr. Naren Kumar A.,
E-mail: surgicalcare.naren@gmail.com

ABSTRACT

Background: Acute appendicitis is a very common cause of acute abdominal pain, requiring surgical intervention with a 7% life time risk. Various clinical scoring systems like Alvarado, appendicitis inflammatory response (AIR), Tzanaki scores enables risk stratification. In this study, we have validated the diagnostic accuracy of various scoring systems like Alvarado, Tzanaki and AIR scores.
Methods: We conducted a prospective observational study for patients admitted in a tertiary care hospital. A total of 81 patients were selected based on the inclusion and exclusion criteria. A detailed clinical history, physical examination, relevant blood investigations and necessary imaging were done for all the patients. Using the above data, the probability of acute appendicitis is calculated using the Alvarado, Tzanaki and AIR scores. The various scores obtained were compared with the histopathological examination (HPE) reports (reference standard) and values like sensitivity, specificity, positive and negative predictive values (PPV and NPV) were calculated.
Results: With regard to scores, Alvarado scores has a sensitivity 95.24%, specificity 77.78%, PPV of 93.75%, NPV of 82.35%. Tzanaki score has a sensitivity of 100%, specificity 23.5%, PPV of 83.12% and NPV of 100%. AIR score has a sensitivity of 95%, specificity of 66.67%, PPV of 89.06% and NPV of 82.35%.
Conclusions: Alvarado score (cut off 7) has the best PPV (93.75%) and hence is the best scoring system to "rule in" a positive diagnosis. However, Tzanaki score (cut off 8) has the best NPV (100%) and hence is the best score to "rule out" a diagnosis of acute appendicitis.

Keywords: Acute appendicitis, Alvarado, AIR, Tzanaki, Diagnostic accuracy, Scoring systems

INTRODUCTION

Acute appendicitis is a very common cause of acute abdominal pain, requiring surgical intervention with a 7% life time risk.[1] Many times, the symptoms of appendicitis overlaps with other causes of acute abdomen and hence makes a definitive diagnosis difficult, especially in the earlier stages of the disease.[2] Patients may be suitably triaged into different managing strategies which include reassurance, conservative in hospital management and operative or radiological intervention. If admitted to hospital, appropriate imaging may be required in some cases prior to proceeding to surgery.[3]

Various clinical scoring systems like Alvarado, AIR, Tzanaki scores enables risk stratification in patients, regardless of age, gender and ethnic differences, presenting with abdominal pain, linking the probability of appendicitis to various treatment strategies as outlined above.[4,5] Further investigations, such as ultrasound and computed tomography (CT) scanning, are recommended when probability of appendicitis is in the intermediate range.[6] However, the time loss, high cost burden and non-availability or scarcity of such resources especially in low resource setup mean that these scoring systems may be a valuable diagnostic aid when appendicitis is suspected to

be the underlying cause of an acute abdomen, where imaging is not feasible or not available.[7]

In recent years, the rate of negative laparotomy and negative appendectomy has increased a lot. From several studies, the rate of negative appendectomy ranges from 15% to 30%.[8] In this study, we are trying to validate the diagnostic accuracy of various scoring systems of proven value and comparing the three most important and useful, easy to measure, scoring systems, i.e., Alvarado, Tzanaki score and AIR scores.

METHODS

This was a prospective observational study to evaluate the validity of different scoring systems in diagnosing acute appendicitis at a tertiary care hospital over a period of one year from June 2017 to June 2018. All patients admitted with right iliac fossa pain with suspicion of acute appendicitis, after informed consent, were enrolled into the study. Ethical approval to conduct the study was obtained from the institutional ethic committee before the commencement of the study Patients with appendicular lump/generalized peritonitis, patients with normal appearing appendix and an alternative diagnosis during operation, patients with previous history of urolithiasis, Pelvic inflammatory disease which can mimic symptoms of appendicitis were excluded. A target sample of 80 patients was required to achieve statistical significance based on the previous census in our study centre.

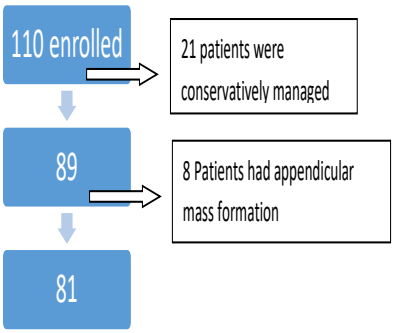

Figure 1: Target Sample size achieved after exclusion.

The intra operative findings were noted. Following surgery, the histopathological reports of the specimens were collected. The various preoperative scores and the HPE reports were correlated to calculate the Sensitivity, Specificity, PPV and NPV. This will prove the validity of scoring systems in the diagnosis of acute appendicitis and provides the most reliable system applicable in south Indian population. Data was analysed using the statistical package for social sciences (SPSS) version 23.0. Descriptive statistics such as frequency and percentage were calculated. Continuous variables are expressed in mean and standard deviation. Association between the groups various study variables predominantly continuous variables was done by repeated measures -analysis of variance test (r-ANOVA). A p=0.05 or less was taken to indicate a significant difference.

RESULTS

In our study, a total of 81 patients were analysed. Our patient group contained males and females in equal predominance with a ratio of 1:1.1. Age wise descriptive data, revealed that majority of patients belonged to the 2nd to 4th decade contributing about 70% (57 patients) (Figure 2). We also compiled the data of incidence of acute and chronic appendicitis based on histopathological report. The data is shown in Table 1. In our analysis of accuracy of Alvarado score predicting appendicitis, taking the HPE as reference standard, we found that 60 patients with acute appendicitis fell in the score >7. However, 3 patients who fell in >7 score had chronic appendicitis as shown in Table 1. In contrast to Alvarado scoring, AIR showed poor specificity as more than 15 patients with >5 score had chronic appendicitis (Table 2). Tzanaki score performed equal to Alvarado score in diagnosing acute appendicitis (Table 3).

Table 4 details comparison of various parameters between the three scoring systems. We see that Alvarado score has very good sensitivity (95.24%) and the best diagnostic accuracy (91.36%) among the three (cutoff value>7). Tzanaki has 100% sensitivity and Negative predictive value making it the best parameter to "rule in" a diagnosis. Appendix Inflammatory response score underperforms in all parameters as compared to other 2.

Table 1: Incidence of appendicitis in our study group.

HPE diagnosis	Frequency	Percent (%)
Acute appendicitis	64	79.0
Chronic appendicitis	17	21.0
Total	81	100.0

Table 2: Alvarado score vs HPE.

Alvarado score		1-4	5- 6	7-8	9-10	Total
HPE	Acute	0	4	24	36	64
	Chronic	8	6	3	0	17
Total		8	10	27	36	81

Table 3: AIR vs HPE.

AIR score		0-4	5-8	9-12	Total
HPE diagnosis	Acute	0	7	57	64
	Chronic	2	12	3	17
Total		2	19	60	81

Table 3: Tzanaki vs HPE.

Tzanaki score		<8	>8	Total
HPE diagnosis	Acute	0	64	64
	Chronic	4	13	17
Total		4	77	81

Table 4: Alvarado vs AIR vs Tzanaki.

Parameters	Alvarado, (%)	AIR, (%)	Tzanaki, (%)
Sensitivity	95.24	95.00	100.
Specificity	77.78	66.67	23.53
Positive predictive value	93.75	89.06	83.12
Negative predictive value	82.35	82.35	100
Diagnostic accuracy	91.36	87.65	83.95

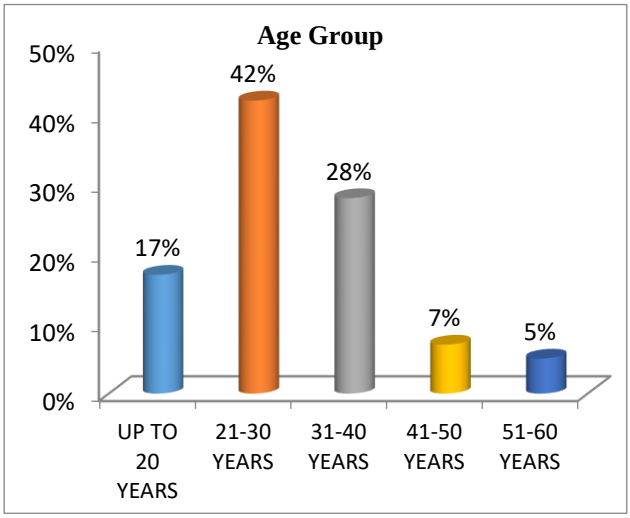

Figure 2: Age-wise distribution of patient cohort.

DISCUSSION

Acute appendicitis is one of the most common surgical emergencies worldwide, the diagnosis of which is still a challenging job for the surgeon. Both delayed diagnoses resulting in perforation (20%) and negative appendectomy (2-30%) have their own complications. Hence there is need for an accurate preoperative diagnosis. Even though CT has a high sensitivity and specificity in diagnosing appendicitis, the cost, availability and time delay become its limitations. Several scoring systems using clinical, laboratory data have been designed to diagnose acute appendicitis. In our study we considered 3 such scoring systems. The aim of our study was to assess and compare the diagnostic accuracy of these scoring systems and to study their role in diagnosing a case of acute appendicitis. In this study 81 patients were subjected to further statistical analysis.

Alvarado score

Alvarado score is found to be more than 7 in majority of the patients and has a sensitivity 95.24%, specificity 77.78%, positive predictive value 93.75%, negative predictive value 82.35%. Since it had high PPV, scores 7 and above is considered high likely to have acute appendicitis as it had low false positivity rate. Likewise, because of its low NPV, scores lower than 7 cannot exclude the diagnosis of acute appendicitis.

A similar study was conducted by Kim et al showed that the diagnosis of acute appendicitis was highly accurate for an Alvarado score above 7 (82.9%). The difference in the sensitivity values between two studies are probably due to the difference in the sample size.[9]

Another study conducted by Schneider et al concludes that lack of sufficient PPV by the Alvarado scoring system makes it not a definitive tool among the paediatric age group. But the scoring system proves to be a good tool among the adult patient group according to our study.[10]

Tzanaki score

Tzanaki score's sensitivity of detecting acute appendicitis with scores 8 or above is 100%, specificity 23.5%, Positive predictive value 83.12% and negative predictive value of about 100%. Since this scoring system has high NPV rate, scores less than 8 ultimately excludes the diagnosis of acute appendicitis and because of its 100% sensitivity, it cannot miss the diagnosis of acute appendicitis when scores are 8 and above.

As compared to our study, Sigdel et al reported, for a sample size of 100 patients with sensitivity, specificity and overall diagnostic accuracy were 91.48%, 66.66% and 90% respectively.[11] In the original study conducted by Tzanakis et al had a sensitivity, specificity, accuracy of 95.4%, 97.4%, 96.5%, and 93%, respectively.[12] Similarly, our study has 100% sensitivity and negative predictive value but contrastingly low specificity. Since this score uses USG findings as one of its criteria there can exist an inter observer variation which in turn can influence the specificity values.

AIR score

AIR score uses C-reactive protein as a parameter and it is a new scoring system found to be better than Alvarado score. But in our study, the sensitivity is 95%, specificity 66.67%, positive predictive value 89.06% and negative predictive value of only 82.35%. Because of high PPV, the AIR score will have low false positivity rate and it has significantly low NPV than Alvarado, which defines it is not beneficial. A study conducted by de Castro et al revealed that AIR score out performs Alvarado score especially in difficult to diagnose patients like women, children and elderly patients.[13] This is in contrast to our study. This could be because of the differences in patient cohort in our study, as it does not include patients in extremes of age and pregnant women.

Over diagnosis

Out of 81 patients underwent appendectomies, only in 17 patients the appendix is not acutely inflamed and are found to be without any evidence of chronic

inflammation and the rest 13 specimens are found to have features suggestive of chronic appendicitis. Hence, the negative appendectomy rate with the use of various clinical scoring systems in our study is zero. This is very less as compared to other studies by Aravindan Narayanan et al and Joshi et al.[14,15]

Best score

In our study, the diagnostic accuracy of Alvarado score was found to be 91.36% in comparison to AIR (87.65%) and Tzanaki (83.95%) making it the most reliable diagnostic parameter. Another observation made was Tzanaki score has a NPV of 100% making it very effective in "ruling out" a diagnosis of the appendicitis (<8).

Limitations

The limitations of our study can be categorized into threefold. Firstly, the small sample size may not be representative of the population and we plan to extend our studies to include more patients to reach a statistical significance in the future.

Secondly, our study did not study the influence of antibiotics on the scoring systems. Since majority of the patients in urban or semi urban population tend to take over the counter painkillers and sometimes even antibiotics before coming to the emergency. This in turn can influence the various clinical signs and laboratory parameters which can affect these scoring systems.

Thirdly, since most of the parameters in these scoring systems involve clinical signs, there can be subjective variability, which in turn can affect the reproducibility of similar results.

CONCLUSION

We conclude from our study that the Alvarado score is the best score in terms of diagnostic accuracy with the cut off value of 7. We also found that Tzanaki score has very high sensitivity and negative predictive value. Hence, we feel that of all the scores, Tzanaki scoring system is best used in emergency situations and also in doubtful cases. We hope that our study has proved the relevance of these age-old systems in the current diagnostic protocol and these scoring systems can be a good guide even in the times to come. We hope that our study can inspire future systematic reviews or meta-analysis of various data available in low resource settings, proving that these scoring systems still can hold their own in this era of modern imaging and diagnostic methods.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. Ergul E. Importance of family history and genetics for the prediction of acute appendicitis. Internet J Surg. 2007;10:2.
2. Bundy DG, Byerley JS, Liles EA, Perrin EM, Katznelson J, Rice HE. Does this child have appendicitis?. JAMA. 2007;298:438-51.
3. Paulson EK, Kalady MF, Pappas TN. Clinical practice. Suspected appendicitis. N Engl J Med. 2003;348:236-42.
4. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med. 2006;144:201-9.
5. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986;15:557-64.
6. Terasawa T, Blackmore CC, Bent S, Kohlwes RJ. Narrative review: computed tomography and ultrasonography to detect acute appendicitis in adults and adolescents. Ann Intern Med. 2004;141:537-46.
7. Howell JM, Eddy OL, Lukens TW, Thiessen ME, Weingart SD, Decker WW. Clinical policy: critical issues in the evaluation and management of emergency department patients with suspected appendicitis. Ann Emerg Med.2010;55:71-116.
8. Repplinger MD, Golden SK, Ebinger A, Svenson JE. Comparison of Appendicitis Clinical Scoring Systems With Physician-Determined Likelihood of Appendicitis. J annemergmed. 2014;7:90.
9. Kim HE, Park SB, Woo SU, Rho HR, Chae GB, Choi WJ. Application of the Alvarado score to the diagnosis of acute appendicitis. J Korean Soc Coloproctol. 2006;22:229-34.
10. Schneider C, Kharbanda A, Bachur R. Evaluating appendicitis scoring systems using a prospective pediatric cohort. Epub. 2007;49(6):778-84.
11. Sigdel G, Lakhey P, Misra P. Tzanakis Score vs Alvarado in Acute Appendicitis. J Nepal Med Assoc. 2010;49(178):96-9.
12. Tzanakis NE, Efstathiou SP, Danulidis K, Rallis GE, Tsioulos DI, Chatzivasiliou A et al. A new approach to accurate diagnosis of acute appendicitis. World J Surg. 2005;29(9):1151-6.
13. De Castro, Steller E, Wagensveld BC, Vrouenraets. Evaluation of the Appendicitis Inflammatory Response Score for Patients with Acute Appendicitis. World J Surg. 2012;36(7):1540-5.
14. Narayanan A, Sundararaman S, Varadhan L, Rajput R, Gupta V, Jones N. What is negative about negative appendicectomy rates? An experience from a district general hospital Int Surg J. 2015;2(2):161-4.
15. Joshi MR, Shaan E, Agarwal S, Kumar S. Negative Appendectomy: An Audit of Resident-Performed Surgery. How Can Its Incidence Be Minimized? Indian J Surg. 2010;10:1094.

Cite this article as: Natesan SR, Kumar AN. Evaluation of validity of scoring systems in diagnosing acute appendicitis: a single centre prospective observational study. Int Surg J 2022;9:23-6.