International Surgery Journal
Hassan M et al. Int Surg J. 2022 Dec;9(12):1937-1942
http://www.ijsurgery.com
pISSN 2349-3305 | eISSN 2349-2902

DOI: https://dx.doi.org/10.18203/2349-2902.isj20223152

Original Research Article

Evaluation of Alvarado score and appendicitis inflammatory response score as diagnostic tools for acute appendicitis

Mohamed Hassan[1]*, Milad Jeilani[1], Ahmed A. Saad[1], Sheeraz Iqbal[1], Mohamed Boshnaq[2,3]

[1]Department of General Surgery, Tunbridge Wells Hospital, Tunbridge Wells, Kent, United Kingdom
[2]Department of General Surgery, Queen Elizabeth The Queen Mother Hospital Margate, Kent, United Kingdom
[3]Department of Colorectal Surgery, Ain Shams University, Cairo, Egypt

Received: 03 October 2022
Revised: 04 November 2022
Accepted: 09 November 2022

***Correspondence:**
Dr. Mohamed Hassan,
E-mail: drmohamedgad2009@gmail.com

ABSTRACT

Background: The study aims to evaluate the diagnostic accuracy of AIR and Alvarado scores for acute appendicitis cases.
Methods: We conducted a retrospective study with 73 patients to compare the efficacy of the appendicitis inflammatory response (AIR) score with the Alvarado score in the diagnosis of acute appendicitis from January 2021 to March 2021. The study included all consecutive patients referred with suspected appendicitis or acute RIF pain, to the surgical unit at Tunbridge Wells Hospital.
Results: Based on histopathology findings of 73 patients, 59 (80.8%) patients had acute appendicitis. Using the AIR score, 5 (8.5%) patients scored high, 41 (69.5%) scored moderate, and 13 (22.0%) scored low in the histologically-positive group. In the histology-negative group, none (0.0%) scored high, 2 (14.3%) scored moderate, and 12 (85.7%) scored low, with only 2 patients being false positives. The correlation between the AIR score and histopathology results was thus highly significant (p value=0.000). Using the Alvarado score, in the histology-positive group, 33 (55.9%) scored high, 17 (28.8%) moderate and 9 (15.3%) low. In the histology-negative group, 2 (14.3%) patients scored high, 6 (42.9%) scored moderate, and 6 (42.9%) patients scored low. 8 patients were false positives. We found a significant correlation between the Alvarado score and histopathology results (p value=0.011). The sensitivity of the AIR and Alvarado scores were 77.97% and 67.80% respectively and the specificity was 85.71% and 78.57% respectively.
Conclusions: The AIR score had higher specificity, sensitivity, positive predictive value, and a lower rate of false positives.

Keywords: Appendicitis, Alvarado, AIR

INTRODUCTION

Acute appendicitis is one of the most common abdominal emergencies worldwide, with a lifetime incidence of 7-9%.[1-3] Upon diagnosis, urgent surgical intervention is required to avoid the risk of severe and progressive inflammation that can cause perforation, abscess formation, peritonitis and intra-abdominal adhesions.[4] Acute appendicitis is suspected in patients who present acutely with right lower quadrant pain or tenderness and leukocytosis. However, it is only confirmed on histological examination of the appendix specimen. Therefore, acute appendicitis is still clinically diagnosed on the basis of history, physical examination, laboratory tests and imaging. However, the diagnosis can be challenging because of vague or non-specific symptoms and atypical presentations, particularly in younger patients.[5-8] Given the difficulties of clinical diagnosis and

the potentially life-threatening complications of untreated appendicitis, there is a historically high prevalence of negative appendectomies, which is associated with a higher financial burden, hospitalisation rates, complication rates and mortality.[9-11] To avoid both unnecessary negative appendectomies and the complications of untreated appendicitis, surgeons have sought to increase their diagnostic accuracy to better differentiate between patients requiring surgical management and those requiring conservative management. Imaging modalities such as computed tomography (CT) and ultrasound have been shown to improve diagnostic accuracy and reduce negative appendectomy rates.[11-12] However, they are limited by inaccuracies, inconclusive results, inaccessibility outside of business hours, cost, operator skill with ultrasound and radiation risk with CT. Thus, various clinical scoring systems have been created as diagnostic aids and to reduce imaging use. These include the appendicitis inflammatory response score (AIR), the Alvarado score, the pediatric appendicitis score (PAS), the adult appendicitis score (AAS), the Lintula score, the RIPASA score and the pediatric appendicitis risk calculator (pARC).[13-19] In this study, we utilise the widely used Alvarado score (Table 1) and the newer AIR score (Table 2) to explore the demographic and clinical characteristics of suspected appendicitis cases over a three-month period at Tunbridge Wells Hospital, Tunbridge Wells, England, UK. We evaluate the diagnostic accuracy of the AIR and Alvarado scores for acute appendicitis and comment on their clinical usefulness.

METHODS

We conducted a retrospective observational study in general surgery department at Tunbridge Wells Hospital, England, UK. from January 2021 to March 2021 to compare the efficacy of the appendicitis inflammatory response (AIR) score with the Alvarado score in the diagnosis of acute appendicitis. Both scores allow for risk stratification (Table 3).

Inclusion and exclusion criteria

The inclusion criteria were all consecutive patients referred by a general practitioner or emergency physician to the on-call surgeon's team with suspected appendicitis or acute RIF pain, identified at the point of admission to the surgical unit. The exclusion criteria were patients presenting with any form of non-right iliac fossa pain, emergency laparotomy which included appendectomy, elective appendectomy, lump in the right iliac fossa, immunocompromised patients and pregnant patients. Ethical committee approval was not required as the data was retrospective. Data were analyzed using SPSS software and were expressed as numbers and percentages. Chi-square and Mann-Whitney tests were applied for comparing the data of both the groups and P value less than 0.05 was considered statistically significant.

The recording of the patient's history and the physical examination were completed per the designated forms. A full range of routine haematological investigations was performed, and the scores were calculated. All cases were thoroughly examined by a senior faculty member and referred for surgical intervention. Lab tests were performed, and imaging studies (CT or ultrasound) were conducted at the surgeon's discretion in some cases. The demographic data and clinical examination findings (signs and symptoms) were noted in a separate case record form. Histopathology was evaluated on the excised appendix for confirmation of diagnosis. Upon receiving the results, the scores were correlated with the histopathological reports. The optimal cut-off threshold was four for AIR and seven for Alvarado. Our analyses included the determination of sensitivity, specificity and positive predictive value (PPV), as well as a negative predictive value (NPV) for the AIR and Alvarado scores.

Table 1: Scoring scheme for the Alvarado score.

The Alvarado score	Score
Signs	
Right lower quadrant tenderness	2
Elevated temperature (>37.3)	1
Rebound tenderness	1
Symptoms	
Anorexia	1
Nausea	1
Migration of pain to right lower quadrant	1
Laboratory values	
Leucocytosis (>10.000)	2
Left Shift (>75% neutrophils)	1
Total	10

Table 2: Scoring scheme for the AIR score.

Parameters	Score
Appendicitis inflammatory score	
Vomiting	1
Pain in the right lower quadrant	1
Rebound tenderness	
Light	1
Medium	2
Strong	3
Polymorphonuclear leukocytes	
70%-84%	1
≥85%	2
White blood cell count	
10.000-14.999 cells/l	1
≥ 15.000 cells/l	2
C-reactive protein	
10-49 mg/l	1
≥50 mg/l	2
Total	12

RESULTS

There were 73 patients eligible for inclusion during the study period. Patient demographics included 33 (46.4%) males and 40 (54.8%) females, with a mean age of 31±17 (Table 4). There was no association of the AIR and Alvarado scores with age, sex and co-morbidities.

Table 3: Risk stratification based on AIR and Alvarado scores.

Parameters	
Interpretation of the cumulative Alvarado score	
0-4	Not likely appendicitis
5-6	Equivocal
7-8	Probable appendicitis
9-10	Highly likely appendicitis
0-4	Not likely appendicitis
5-6	Equivocal
Interpretation of the cumulative AIR score	
0-4	Low probability of appendicitis
5-8	Indeterminate probability of appendicitis
9-12	High probability of appendicitis
0-4	Low probability of appendicitis

Table 4: Relation of histopathology results with demographic data and characteristics of the studied patients.

Variables		Histopathology	
		Negative (N=14)	**Positive (N=59)**
Age	Median (IQR)	24.21 (15-29.39)	29.54 (18.77-46.21)
Gender N (%)	Females	10 (71.4)	29 (50.8)
	Males	4 (28.6)	30 (50.8)

Based on the operative findings, only 6 (8.2%) out of 73 patients had a non-inflamed appendix. Moreover, based on the histopathology findings, 59 (80.0%) out of 73 patients had acute appendicitis (Table 5). The majority of patients scored in the range of 5-8. Patients were divided into two groups based on histopathological reports, positive and negative. In the negative group, the median AIR score was three, and in the positive group, it was six.

Table 5: Histopathology and operative findings among the studied patients (n=73).

Variables		N (%)
Histopathology	Negative	14 (19.2)
	Positive	59 (80.8)
Operative findings	Inflamed appendix	59 (80.9)
	Non-inflamed appendix	6 (8.2)
	Perforated appendix	3 (4.1)
	Gangrenous appendix	5 (6.8)

Correlation between AIR score and histopathological findings

In the positive group, five (8.5%) patients had high AIR scores, 41 (69.5%) had moderate AIR scores, and 13 (22.0%) had low AIR scores. In the negative group, AIR scores were low for 12 (85.7%) patients, moderate for two (14.3%) patients, and high for none (0.0%) of the patients, with only two patients being false positives. The correlation between the AIR score and histopathology results was highly significant (p=0.001) (Table 6).

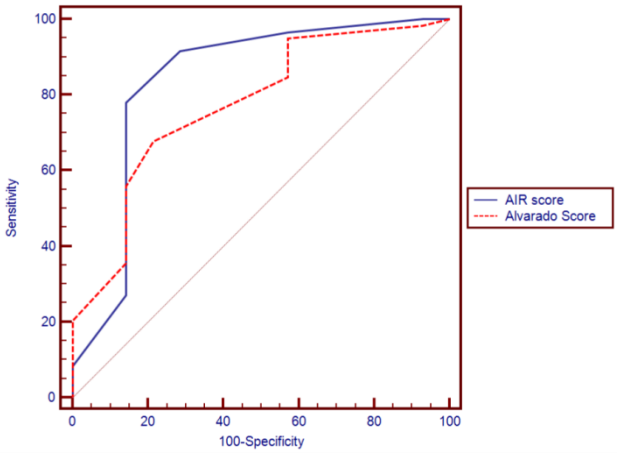

Figure 1: Receiver operating characteristic curve (ROC) for AIR score and Alvarado score in differentiation between negative and positive histopathology groups.

Correlation between Alvarado score and histopathological findings

In the positive group, the Alvarado scores were high for 33 (55.9%) patients, moderate for 17 (28.8%) patients, and low for 9 (15.3%) patients. In the negative group, six (42.9%) patients had low Alvarado scores, six (42.9%) patients had moderate Alvarado scores, and two (14.3%) patients had high Alvarado scores, with eight patients being false positives (Table 6). There was a significant correlation between the Alvarado scores and the histopathology results (p=0.011). The sensitivity of the AIR and Alvarado scores was 77.97% and 67.80%, respectively, and the specificity was 85.71% and 78.57%, respectively (Figure 1). The positive predictive value (PPV) for the AIR and Alvarado scores was 95.8% and 93.0%, respectively, and the negative predictive value (NPV) was 48.0% and 36.7%, respectively. The AIR score had higher specificity, sensitivity and positive predictive value, and a lower rate of false positives (Table 7).

DISCUSSION

A clinical score should be easy to use, predict clinical outcomes and avoid unnecessary investigations and treatments. Keeping this in mind, several scores that

predict acute appendicitis have been designed so far. These scores also aim to reduce the reliance on imaging and negative appendicectomy rates (NAR) while keeping perforation rates at an acceptable level. It has traditionally been thought that a lower NAR comes at the cost of higher perforation rates and, therefore, higher morbidity due to a delay in diagnosis or treatment.

Table 6: Correlation of histopathology results with AIR and Alvarado scores of the studied patients.

Parameters		Histopathology		Test value	P value	Significance
		Negative (N=14)	Positive (N=59)			
AIR score	Median (IQR)	3 (2-4)	6 (5-8)	-3.947**	0.000	HS
	Range	1–8	2–9			
Category N (%)	Low	12 (85.7)	13 (22.0)	20.439*	0.000	HS
	Moderate	2 (14.3)	41 (69.5)			
	High	0 (0.0)	5 (8.5)			
Alvarado Score	Median (IQR)	5 (3-5)	7 (5-8)	-3.185**	0.001	HS
	Range	2–8	2–10			
Category	Low	6 (42.9)	9 (15.3)	8.997*	0.011	S
	Moderate	6 (42.9)	17 (28.8)			
	High	2 (14.3)	33 (55.9)			

*Chi-square test; **Mann-Whitney test, HS-highly significant, S-significant

Table 7: Sensitivity and Specificity of AIR and Alvarado scores.

Scores	Cut off point	AUC	Sensitivity	Specificity	PPV	NPV
AIR score	>4	0.838	77.97	85.71	95.8	48.0
Alvarado Score	> 5	0.772	67.80	78.57	93.0	36.7

However, a sizeable US population-based study contests this association, and one American centre reports an NAR of 1.7% with the routine use of pre-operative CT.[20,21] Our study gave an NAR of 19.2%. Although there is no official 'acceptable' NAR in the UK, ours appears consistent with and is probably below the average for the country.[22] Our perforation rate was 4.1%. The older and more well-established Alvarado score was developed from a retrospective analysis of patients that underwent appendicectomy for suspected appendicitis and has undergone a number of validation studies.[14,23,24] However, its reliance on dichotomous and subjective clinical signs and symptoms has seemingly weakened its discriminative ability. AIR, in comparison, was more recently developed by Andersson and Andersson from a prospective cohort, with the grading of variables and more reliance upon objective inflammatory results, particularly CRP, which has been shown to be especially correlative for acute appendicitis.[13,25]

Many studies that have compared the two scores show AIR to be more accurate than Alvarado over a range of patient demographics.[26-29] In Andersson and Andersson's original prospective study of 545 patients, the area under the curve (AUC) was reported to be 0.93 vs 0.88 for AIR and Alvarado, respectively.[13] Jose et al cross-sectional observational study of 130 patients elicited an AUC of 0.90 vs. 0.82 for AIR and Alvarado, respectively.[26] With an AIR cut-off of five, they found a sensitivity of 98% and a specificity of 36%. Alvarado's cut-off of six gave a sensitivity of 72% and a specificity of 79%. De Castro et al noted an AUC of 0.96 vs 0.82 for AIR and Alvarado, respectively, in 941 patients, with the AIR score outperforming the Alvarado score in patients for whom the diagnosis is usually more difficult, namely children, women and the elderly.[27] A cut-off of four points gave similar sensitivity for AIR and Alvarado (0.93 vs 0.90, respectively) but much higher specificity (0.85 vs. 0.55, respectively), with corresponding NPV of 0.95 vs. 0.90 for AIR and Alvarado, respectively.

Kollár et al evaluated the AIR score and compared its performance in predicting the risk of appendicitis to both the Alvarado score and the clinical impression of a senior surgeon.[28] In their low-risk group, they found AIR to be accurate at excluding appendicitis, and in their high-risk group, they discovered that AIR was more accurate at predicting appendicitis than Alvarado. They proposed that AIR could be used as the basis for selective CT imaging for intermediate-risk patients. Our results are consistent with these studies, with an AUC of 0.84 vs 0.77 for AIR and Alvarado, respectively. Moreover, the NPV in our study was 48.0% and 36.7% for AIR and Alvarado, respectively, and the PPV was 95.8% and 93.0%, respectively. However, other studies question the reliability of these scores, especially in paediatric populations. Vaziri's multicentric prospective study of 661 children showed that AIR was more sensitive and specific than Alvarado only among their low-risk group (95% vs. 90% sensitivity, and 74% vs. 70% specificity, respectively).[29,30] High-risk patients in both scores had a sensitivity and specificity of less than 50%, which was thought to be insufficient to aid diagnosis or decide to proceed with surgery.

Limitations

The study's limitations include the relatively small number of patients, its cross-sectional and observational nature, and the inclusion of only those patients that underwent appendicectomy. Further studies that are more extensive, prospective and look at specific populations in more depth are required. This would allow for the establishment of optimal cut-off points for particular scores and populations, thus enabling the use of selective imaging based on risk stratification.

CONCLUSION

The diagnosis for patients with suspected acute appendicitis is still challenging. The results of our study demonstrate that the AIR score is a superior diagnostic tool to the Alvarado score, outperforming it with higher sensitivity and specificity. Thus, AIR could help identify patients who require immediate surgery or further evaluation. In these cases, a follow-up will help determine whether surgical intervention is necessary.

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

REFERENCES

1. Henderson J, Goldacre MJ, Fairweather JM, Marcovitch H. Conditions accounting for substantial time spent in hospital in children aged 1-14 years. Arch Dis Child. 1992;67:83-86.
2. Addiss DG, Shaffer N, Fowler BS, Tauxe RV. The epidemiology of appendicitis and appendectomy in the United States. Am J Epidemiol. 1990;132:910-25.
3. Lee JH, Park YS, Choi JS. The epidemiology of appendicitis and appendectomy in South Korea: national registry data. J Epidemiol. 2010;20:97-105.
4. Kabir SA, Kabir SI, Sun R, Jafferbhoy S, Karim A. How to diagnose an acutely inflamed appendix; a systematic review of the latest evidence. Int J Surg Lond Engl. 2017;40:155-62.
5. Bundy DG, Byerley JS, Liles EA, Perrin EM, Katznelson J, Rice HE. Does this child have appendicitis? JAMA. 2007;298:438-51.
6. Irish MS, Pearl RH, Caty MG, Glick PL. The approach to common abdominal diagnosis in infants and children. Pediatr Clin North Am. 1998;45:729-72.
7. Marzuillo P, Germani C, Krauss BS, Barbi E. Appendicitis in children less than five years old: a challenge for the general practitioner. World J Clin Pediatr. 2015;4:19-24.
8. Paulson EK, Kalady MF, Pappas TN.: Clinical practice. Suspected appendicitis. N Engl J Med. 2003; 348:236-42.
9. Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: nationwide analysis. Arch Surg Chic Ill. 2002;137:799-804.
10. Flum DR, Morris A, Koepsell T, Dellinger EP. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. JAMA. 2001;286:1748-53.
11. Sharma R, Kasliwal DK, Sharma RG.: Evaluation of negative appendicectomy rate in cases of suspected acute appendicitis and to study the usefulness of ultrasonography in improving the diagnostic accuracy. Indian J Surg. 2007;69:194-7.
12. Chan J, Fan KS, Mak TLA, Loh SY, Ng SWY, Adapala R.: Pre-operative imaging can reduce negative appendectomy rate in acute appendicitis. Ulster Med J. 2020;89:25-8.
13. Andersson M, Andersson RE. The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. World J Surg. 2008;32:1843-9.
14. Alvarado A.: A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986;15:557-64.
15. Samuel M.: Pediatric appendicitis score. J Pediatr Surg. 2002;37:877-81.
16. Sammalkorpi HE, Mentula P, Leppäniemi A. A new adult appendicitis score improves diagnostic accuracy of acute appendicitis-a prospective study. BMC Gastroenterol. 2014;14:114.
17. Lintula H, Kokki H, Pulkkinen J, Kettunen R, Gröhn O, Eskelinen M. Diagnostic score in acute appendicitis. Validation of a diagnostic score (Lintula score) for adults with suspected appendicitis. Langenbecks Arch Surg. 2010;395:495-500.
18. Chong CF, Adi MIW, Thien A. Development of the RIPASA score: a new appendicitis scoring system for the diagnosis of acute appendicitis. Singapore Med J. 2010;51:220-5.
19. Kharbanda AB, Vazquez-Benitez G, Ballard DW, et al.: Development and validation of a novel pediatric appendicitis risk calculator (pARC). Pediatrics. 2018; 141:e20172699.
20. Livingston EH, Woodward WA, Sarosi GA, Haley RW. Disconnect between incidence of nonperforated and perforated appendicitis: implications for pathophysiology and management. Ann Surg. 2007; 245:886-92.
21. Raja AS, Wright C, Sodickson AD. Negative appendectomy rate in the era of CT: an 18-year perspective. Radiology. 2010, 256:460-5.
22. National Surgical Research Collaborative. Multicentre observational study of performance variation in provision and outcome of emergency appendicectomy. Br J Surg. 2013;100:1240-52.
23. McKay R, Shepherd J. The use of the clinical scoring system by Alvarado in the decision to perform computed tomography for acute appendicitis in the ED. Am J Emerg Med. 2007;25:489-93.

24. Coleman JJ, Carr BW, Rogers T. The Alvarado score should be used to reduce emergency department length of stay and radiation exposure in select patients with abdominal pain. J Trauma Acute Care Surg. 2018;84:946-50.
25. Andersson REB.: Meta-analysis of the clinical and laboratory diagnosis of appendicitis. Br J Surg. 2004, 91:28-37.
26. Jose T, Rajesh PS. Appendicitis inflammatory response score in comparison to Alvarado score in acute appendicitis. Surg JNYN. 2021;7:e127-31.
27. de Castro SMM, Ünlü C, Steller EP, van Wagensveld BA, Vrouenraets BC.Evaluation of the appendicitis inflammatory response score for patients with acute appendicitis. World J Surg. 2012;36:1540-45.
28. Kollár D, McCartan DP, Bourke M, Cross KS, Dowdall J.: Predicting acute appendicitis? A comparison of the Alvarado score, the Appendicitis Inflammatory Response Score and clinical assessment. World J Surg. 2015;39:104-9.
29. Macco S, Vrouenraets BC, de Castro SMM. Evaluation of scoring systems in predicting acute appendicitis in children. Surgery. 2016;160:1599-604.
30. Vaziri M, Nafissi N, Jahangiri F, Nasiri M. Comparison of the appendicitis inflammatory response and Alvarado scoring systems in the diagnosis of acute appendicitis in children. J Med Life. 2021;14:75-80.

Cite this article as: Hassan M, Jeilani M, Saad AA, Iqbal S, Boshnaq M. Evaluation of Alvarado score and appendicitis inflammatory response score as diagnostic tools for acute appendicitis. Int Surg J 2022;9:1937-42.