International Surgery Journal
Kumar A et al. Int Surg J. 2019 Aug;6(8):2843-2848
http://www.ijsurgery.com
pISSN 2349-3305 | eISSN 2349-2902

DOI: http://dx.doi.org/10.18203/2349-2902.isj20193328

Original Research Article

Efficacy and comparison of appendicitis inflammatory response score with Alvarado score in predicting the diagnosis of acute appendicitis

Abhishek Kumar*, K. G. Byakodi, Prathamesh P. Chandrapattan, Vasant Teggimani

Department of General Surgery, KIMS, Hubli, Karnataka, India

Received: 06 December 2018
Revised: 04 July 2019
Accepted: 05 July 2019

***Correspondence:**
Dr. Abhishek Kumar,
E-mail: kumar.abhishek595@gmail.com

ABSTRACT

Background: Aute appendicitis being the most common surgical emergency poses a significant diagnosing dilemma. Early diagnosis has very favourable outcomes and if diagnosis is delayed leads to significant morbidity and mortality. USG has some limitations and best investigation computed tomography (CT) is being overused which leads to unnecessary exposure to radiation especially in children and young adult. Scoring systems are a valuable aid when it comes to diagnosing appendicitis.
Methods: Study done in department of surgery in KIMS, Hubli. The study period was December 2015– September 2016, and 107 patients with right lower abdomen pain were studied. History, clinical examination, biochemical, haematological, radiological investigations were done. Alvarado and AIR scores were calculated. Considering HPR report as gold standard the scores were compared.
Results: There were 70 male patients (66%) and 37 female patients (34%). The mean age in our study population was 28.1 years ±13.57 years. The overall area under the receiver operating characteristic (ROC) curve of the AIR score was 0.967 and significantly better than the area under the curve of 0.825 of the Alvarado score (p=0.05).
Conclusions: This study externally validates the AIR score for patients with acute appendicitis. The scoring system has a high discriminating power compared to Alvarado score especially in cases in which appendicitis is difficult to diagnose such as females, children, elderly, in advanced cases and very effective supplement to available radiological investigations.

Keywords: Acute appendicitis, Alvarado score, AIR score, Histopathology, Receiver operating characteristic

INTRODUCTION

The first question that strikes our mind is in this modern era of medicine, when we are supplemented by vast number of radiological investigations what the need of a clinical is scoring system to predict a frequent surgical problem like acute appendicitis. The answer that is most relevant is in a country like India where approximately 75% population still lives in villages, and remote areas who have no or little access to basic medical and surgical facilities. Hence in such a scenario diagnosing appendicitis at an early stage become a challenge and these scoring system play a vital role not only in decreasing the morbidity of the disease by reducing the incidence of complications but also overall diminution of the burden of the disease.

Tait performed the first appendectomy for appendicitis in England in 1880.[1] But even after 130 years this most common surgical emergency is a diagnostic dilemma. Clinical diagnosis alone leads to a negative appendectomy rate of 15 to 30%.[1,2] Negative laprotomy rate declined to approximately 10% with the routine use of ultrasonography (US).[2]

The higher sensitivity of computed tomography (CT) seems to have had an even greater effect on the negative laparotomy rate, which has decreased even further to 5–10%.[2]

Abdominal organs are sensitive to ionizing radiation, and suspected appendicitis is most frequent in young patients for whom the considerations of radiation-induced risks are most important.

Diagnostic scoring was originally invented before the era of modern imaging technologies as an independent diagnostic tool. However, scoring and imaging should optimally be used as complementary methods in a diagnostic algorithm. An ideal scoring system would work as a tool that speeds up and increases the accuracy of decision making, and at the same time reduces the need of potentially harmful and expensive imaging.

Table 1: Alvarado and AIR score.

	Alvarado	AIR
Symptoms		
Abdominal pain that migrates to the right iliac fossa	1	
Anorexia (loss of appetite)	1	
Nausea or vomiting	1	
Vomiting		1
Pain in RIF		1
Signs		
Tenderness in the right iliac fossa	2	
Rebound tenderness	1	
Light		1
Medium		2
Strong		3
Fever of 37.3 °C or more	1	
Temp >38.5 °C		1
Investigations		
Leukocytosis >10,000	2	
10.0-14.9X10^9		1
>15.0 X 10^9		2
Neutrophilia > 70%	1	
Polymorphonuclear leukocytes		
70-84%		1
≥ 85%		2
C-reactive protein concentration		
10-49 g/l		1
>50 g/l		2
Total	10	12

Alvarado (score ≤4 - low likelihood of appendicitis; 5 - 7 - consider further imaging; >7 and high - likelihood of appendicitis). AIR score (score ≤4 - low likelihood of appendicitis; 5-8 - intermediate, consider further imaging, observation; >8 and high - likelihood of appendicitis).

METHODS

It is a prospective hospital based observational study. All the patients coming to KIMS, Hubli, hospital with non traumatic right lower quadrant abdominal pain and suspected appendicitis during December 2015–September 2017 were included.

Selection criteria

Inclusion criteria were patients of age group of 2-70 years. Exclusion criteria were extremes of age of children below 2 years and adults >70 years.

History and physical examination abdominal and relevant investigations for evaluation were done. Variables necessary to evaluate the scoring system was registered and included. In the paediatric population, the child's history is obtained from parents since the patient are too young to give complete history.

Using these variables the Alvarado and AIR scores were calculated and the diagnosis was predicted or the diagnosis was ruled out. Further confirmation of diagnosis was done using the help of radiological investigations like USG, intra operative findings and the confirmatory gold standard for this study was histopathological report. Appendicitis was pathologically diagnosed when infiltration of the muscularis propria by neutrophil granulocytes was seen.[6,7]

Statistical analysis was performed with SPSS statistical software (SPSS Inc, Chicago, IL). A $p<0.05$ was considered statistically significant.

Pearson's chi-square test was used to test if differences between dichotomous groups were significant. Fisher's exact test was used when a table had a cell with an expected frequency of less than 5. The area under the receiver operating characteristic (ROC) curves was used to examine the performance characteristics of the two scoring systems.

RESULTS

The present study includes 107 cases, with suspicion of appendicitis clinically. The incidences in different group are tabulated as follow.

Table 2: Age distribution of the patients.

Age group (in years)	Number	Percentage (%)
11-20	38	35
21-30	33	31
31-40	23	21
41-50	5	5
>50	8	7
Total	107	100

The mean patient age was 28.1 years, with a range of 13–70 years (SD±13.57). Majority of patients were in 11-25 age group (56%). Female to male ratio 1:1.9, majority of study participants were males (66%).

Pain in right lower quadrant of abdomen was uniformly present in all subjects (100%), nausea and vomiting were the second and third most common symptoms (70 and 54%) respectively. A majority of patients had complaints of loss of appetite (50%). Migration of pain from umbilicus to right iliac fossa which is a variable in traditional Alvarado's score was seen in few patients (14%).

In 7 cases ultrasonography was not done (6%) and in 4 cases there was no evidence of appendicitis (4%), in 69 cases ultrasonography gave features suggestive of inflammation of appendicitis (65%) and in 27 cases there was evidence of appendicitis with features of pus collection, gangrenous changes, perforation, formation, (features suggestive of advance appendicitis (25%).

Table 3: Histopathological report.

	Frequency	Percentage (%)
Normal appendix	7	7
Acute appendicitis	69	64
Acute suppurative appendicitis	13	12
Acute perforative appendicitis	9	8.5
Chronic appendicitis	9	8.5
Total	107	100

According to histopathological reports there was no evidence of appendicitis in 7 cases (7%). In 69 cases there were features suggestive of acute appendicitis (64%). In 13 cases there were features of acute suppurative appendicitis (12%). In 9 cases there were acute features of acute gangrenous appendicitis (8.5%). Acute suppurative and acute gangrenous appendicitis were considered as advanced appendicitis (24 patients). In 9 cases there were features of chronic appendicitis (8.5%).

Table 4: Distribution of number of patients according to Alvarado scoring groups.

Alvarado scoring groups	Frequency	Percentage (%)
Normal (0-4)	15	14
Intermediate (5-7)	71	66
Suggestive of appendicitis (8-10)	21	20
Total	107	100

According to Alvarado score 15 cases (14%) were ruled out of the diagnosis of acute appendicitis (scores 0-4), 71 cases (66%) had an intermediate scores (5-7) and a moderate risk of suffering from acute appendicitis and 21 cases (20%) had strong suspicion of acute appendicitis with (scores 8-10).

Table 5: Distribution of cases according to AIR score.

	Frequency	Percentage (%
Normal (0-4)	18	17
Intermediate (5-8)	45	42
Appedicitis (9-12)	44	41
Total	107	100

According to AIR score 18 cases (17%) were ruled out of the diagnosis of appendicitis scores being (0-4), 45 cases had a moderate suspicion of acute appendicitis, which required admission and further observation which had scores (5-8) and 44 cases had a strong suspicion of appendicitis and required surgical intervention having scores (>8).

Following the statistical analysis done in other similar studies done worldwide for our statistical analysis of both the scoring groups we have divided the Alvarado group as score more than 4 Alvarado 2 (score >4) and score more than 7 Alvarado 1 (score >7) group, for AIR group we have divided the groups into score more than 4, AIR 2 (score >4) and more than 8 AIR 1 (score>8) groups. Comparison of respective groups are done as per done in other worldwide studies and the results are shown below for our studies.

Table 6: Comparative data for Alvarado1 and AIR1 groups.

	Alvarado 1 (95% CI) (>7)	Air 1 (95% CI) (>8)
Sensitivity (%)	21% (13.49% to 30.29%)	44% (38.04 to 54.90)
Specificity (%)	100 (59.04% to 100.00%)	100
AUC	0.605 (0.506 to 0.698)	0.720 (0.625 to 0.825)
PPV (%)	100 (71.51-100.00)	100
NPV (%)	8.14 (7.42% to 8.93%)	11.11% (9.508 to 12.946)
Accuracy	26%	43%
LK+	VH	VH
LK-	0.79 (0.714 to 0.874)	0.56 (0.47 to 0.677)

A comparison between the stongly suspected groups of Alvarado and AIR had a lower sensitivity for appendicitis for the Alvarado score compared with the AIR score (21%vs. 44%). However, this was associated with specificity (1.00 vs. 1.00, respectively). These scores translate to a positive predictive value of 1.00 and 1.00 for the AIR and the Alvarado scores, respectively. The AIR classified 44 patients to the high-risk group. All of them had appendicitis. The corresponding figure for the Alvarado score was 21 patients, all whom had appendicitis.

Table 7: Comparative data for Alvarado2 and AIR2 groups.

	Alvarado 2 (95%CI) (>4)	AIR 2 (95%CI) (>4)
Sensitivity (%)	84 (75.31 to 90.56)	88% (79.02 – 93.643)
Specificity (%)	71.43 (20.092 to 96.31)	85.714 (42.128 – 99.638)
AUC	0.777 (0.686 to 0.852)	0.869 (0.79 - 0.926)
PPV (%)	97.67 (92.68 to 99.02)	98.88 (93.790 - 98.815)
NPV (%)	23.81 (14.03 to 34.45)	33.33 (21.39 to 47.125)
Accuracy	83.18%	91%
LK+	2.940 (0.908 to 9.05)	6.160 (1.002 to 37.006)
LK-	0.224 (0.117 to 0.443)	0.140 (0.076 to 0.254)

Table 8: Results of comparison of area under the curves (ROC) of Alvarado and AIR scores in gender groups and age cohorts.

Variable		Number	Alvarado		AIR	
			AUC	P value	AUC	P value
Gender	Male	70	0.858	0.002	0.955	<0.005
	Female	37	0.813	0.007	0.978	<0.005
Age groups (in years)	11-25 yrs	60	0.938	<0.005	0.958	<0.005
	26-40 yrs	34	0.755	0.072	0.969	0.001
	>40 yrs	13	0.455	0.844	0.999	0.03
Total		107	0.825	<0.005	0.967	<0.005

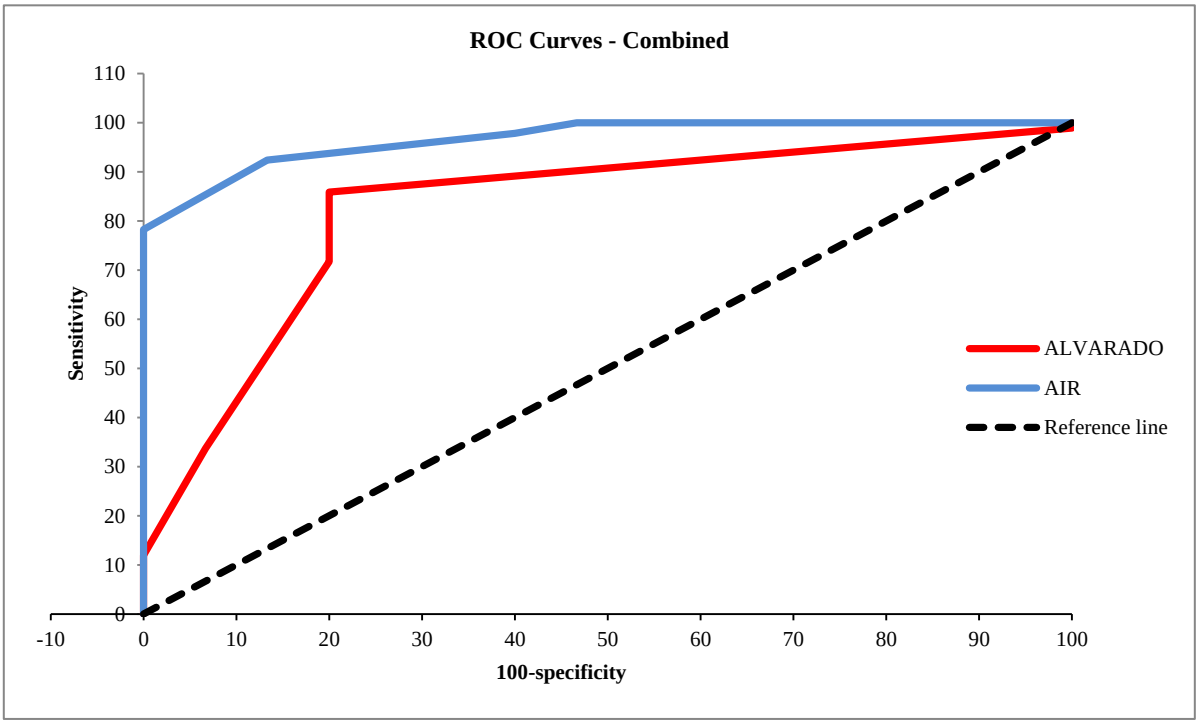

Figure 1: ROC curves combined.

Discriminatory capacity of AIR score better than ALVARADO score overall and especially in difficult cases such as women,children, old age and advanced cases such as peroration and abscess.

A score of greater than 4 points gave a similar sensitivity for the AIR score and the Alvarado score (0.88 vs. 0.84, respectively) but gave a much higher specificity (0.857 vs 0.714, respectively) (Table 7). This corresponds to a negative predictive value of 0.33 for the AIR score compared to 0.23 for the Alvarado score. The area under the ROC curve of the AIR score was 0.967 and significantly better than the area under the curve of 0.825 of the Alvarado score (p=0.05). The AIR score also outperformed the Alvarado score in the analysis of the more difficult to diagnose patients, including women, children, and the elderly.

DISCUSSION

In this prospective study, an attempt was made to evaluate the efficiency of appendicitis inflammatory response score and compare it with Alvarado score. The present study shows that the AIR score has a good statistical discrimination for patients with acute appendicitis and outperforms the Alvarado score. The discriminatory property of the AIR score remains high in the more difficult to diagnose patients (e.g., women, children, and the elderly) (Figure 1).

Appendicitis inflammatory response score outperformed Alvarado score displaying higher sensitivity and specificity. This scoring system has very high sensitivity (88%) and specificity (88%) when the cut off is kept at 4 (Table 7). This scoring system predicts the positive cases and rules out cases which do not have appendicitis equally well when the cut off is kept at 4.

Whereas if the cut off is kept at 8 AIR Score has low sensitivity (44%) but very high specificity (100%) (Table 6). At this cut off this score cannot be used as a modality to screen the disease among population but its reliability is very high in which ever cases it rules out from the diagnosis of appendicitis. The discriminatory capacity of AIR Score is also very high. It has a good discriminatory capacity in severe form of appendicitis such as gangrenous, perforative appendicitis etc. It also has a very high discriminatory capacity in extreme of ages such as children, and elderly and in which diagnosis becomes a challenge.

This scoring system can be used to stratify the patients into three groups on the probability of having appendicitis. High (score >8), intermediate (5-8), and low risk (<-4) for appendicitis. Ideally, the patients in the low-risk group can be discharged, and patients in the high-risk group can be directly scheduled for surgery. The patients in the intermediate risk group benefit most from further investigations such as imaging.

Table 9: Comparison of present study with other similar studies in diagnostic characteristics of Alvarado score and AIR score according to cut off points.[4,5,8,9]

Diagnostic characteristics	Present study		De Castro et al[4]	
	Alvarado	AIR	Alvarado	AIR
	>4 (ALV 2)	>4 (2)	>4 (ALV 2)	>4 (2)
	>8 (ALV 1)	>8 (1)	>8 (ALV 1)	>8 (1)
Sensitivity	84	88	90	93
	21	44	29	10
Specificity	71	86	55	85
	100	100	95	100
PPV	97	99	53	79
	100	100	77	100
NPV	24	34	98	99
	8	11	90	88
P value	<0.05	<0.05	<0.01	<0.01

Table 10: Comparison of discriminatory capacity of Alvarado, AIR scores according to gender and various age groups using receptop operative curves (ROC) of present study with other studies.[4,5,8-11]

Variable		Present study		De Castro et al[4]		Kollar et al[10]	
		Alvarado	AIR	Alvarado	AIR	Alvarado	AIR
Gender	Male	0.858	0.955	0.79	0.95	0.804	0.85
	Female	0.813	0.978	0.82	0.96	0.883	0.795
Age (in years)	11-25	0.938	0.958	0.80	0.96	0.883	0.873
	26-40	0.755	0.969	0.88	0.97	0.805	0.853
	>40	0.455	0.999	0.75	0.92	0.65	0.880
Total		0.825	0.967	0.82	0.96	0.863	0.825

Sensitivity of Alvarado and AIR at (cut off > 4) is 84 & 88 respectively which is comparable to similar study done by De Castro et al respectively. Sensitivity of Alvarado and AIR at cut off (>8) is 21% & 44% respectively which is comparable to similar study done by De Castro et al 29% & 10%.[4] Specificity of Alvarado and AIR at (cut off >4) is 86 & 71 respectively which is comparable to similar study done by De Castro et al 85 & 55.[4] Specificity of Alvarado and AIR at (cut off > 8) is 100 & 100 respectively which is comparable to similar study done by De Castro et al 100 & 95 respectively.[4]

Comparison of area under the curve (AUC) values from analysis of receiver operator characteristic (ROC) curves shows that the discriminatory capacities of all the three

methods of assessment are similar in present study when in comparison with studies done by De Castro et al and Kollar et al.[4,10] The AUC for male and female gender for present study was 0.955 and 0.978 which was comparable with De Castro et al study 0.95 & 0.96.[4] Thus the AIR score has better discriminatory capacity in females as compared to Alvarado score. Similarly the AIR score outperforms the Alvarado score in discriminating appendicitis in younger and elderly age groups. AUC for various age groups (11-25), (25-40), >40 was 0.958, 0.969, 0.999 which was similar to De Castro et al 0.96, 0.97, 0.92.[4] The present study with almost identical ROC curves for the AIR and Alvarado scores when compared to other studies.

This objective validated scoring system can be a good adjunct to the current modalities of diagnosing appendicitis such as ultrasonography and CT scan. It can also legally strengthen the decision making in emergency room and could avoid malpractice liability as most diagnosis involves misdiagnosis or delayed diagnosis. Such a scoring system is important for better outcome.

This score could aid in selecting patients who require timely surgery or those who require further evaluation. Finally, the score could safely avoid hospitalization and unneeded investigations in patients in whom the diagnosis is unlikely. Such a scoring system is important for future research to better compare results.

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

REFERENCES

1. Zinner MJ, Ashley SW. Maingot's abdominal operations. 12th Edition. İstanbul: Nobel Tıp Kitabevleri Ltd Şti; 2008: 939–968.
2. Cuschieri J, Florence M, Flum DR, Jurkovich GJ, Lin P, Steele SR, et al. Negative appendectomy and imaging accuracy in the Washington state surgical care and outcomes assessment program. Ann Surg 2008;248:557–63.
3. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986;15:557-64.
4. De Castro SMM, Ünlü C, Steller EPH, Wagensveld BA, Vrouenraets BC. Evaluation of the appendicitis inflammatory response score for patients with acute appendicitis. World J Surg. 2012;36(7):1540-5.
5. 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.
6. McBurney C. Experience with early operative interference in cases of disease of the vermiform appendix. New York Medical J. 1889;50:676-84.
7. Marudanayagam R, Williams GT, Rees BI. Review of the pathological results of 2660 appendicectomy specimens. J Gastroenterol. 2006;41:745–9.
8. Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. Br J Surg. 2004;91:28-37.
9. 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;32:1843-49.
10. 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(1):104-9.
11. Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. Br J Surg. 2004;91:28-37.

Cite this article as: Kumar A, Byakodi KG, Chandrapattan PP, Teggimani V. Efficacy and comparison of appendicitis inflammatory response score with Alvarado score in predicting the diagnosis of acute appendicitis. Int Surg J 2019;6:2843-8.