Ann Med Res 2020;27(12):3166-70

Annals of Medical Research

DOI: 10.5455/annalsmedres.2020.06.579

Original Article

Comparison of diagnostic scoring systems with imaging methods for the diagnosis of acute appendicitis

Birkan Birben[1], Bedriye Muge Sonmez[2], Sadettin Er[1], Sabri Ozden[3], Murat Tugra Kosa[2], Mesut Tez[1]

[1]Department of General Surgery, Ankara City Hospital, Ankara, Turkey
[2]Department of Emergency Medicine, Ankara City Hospital, Ankara, Turkey
[3]Department of General Surgery, Konya Training And Research Hospital, Konya, Turkey

Abstract

Aim: Acute appendicitis scores have been developed to better analyze the symptoms and signs of acute appendicitis. In this study, we compared the success rates of different scoring systems with imaging methods in the diagnostic confirmation of acute appendicitis.
Materials and Methods: Patients aged above 18 years, who presented to the emergency department with right lower quadrant pain and were suspected to have acute appendicitis, were prospectively and observationally evaluated. The demographic characteristics, imaging modalities, Alvarado score, acute inflammatory response score, and adult appendicitis score were assessed.
Results: 237 patients, 46.8% female and 53.2% male, mean age of 34±13 (18-95) years. Appendectomy was performed in 144 (61%) patients with a prediagnosis of acute appendicitis. The pathological results were appendiceal cancer in two patients, lymphoid hyperplasia in 12, and acute appendicitis in 130. Imaging methods were found to be more specific than scores, and positive predictive values in scoring systems were more sensitive than imaging methods in the diagnosis of acute appendicitis.
Conclusion: There is still no effective and guiding scoring system for the diagnosis of acute appendicitis. Due to the low negative predictive values of the available scoring systems, patients should be evaluated with detailed anamnesis, examination and laboratory findings, and computed tomography should be performed if there is clinical suspicion of acute appendicitis.

Keywords: Acute inflammatory response score; adultappendicitis score; alvarado score; appendicitis; computed tomography

INTRODUCTION

Acute appendicitis (AA) is an inflammatory condition caused by an increase in intra-lumen pressure and bacterial translocation secondary to the obstruction of the appendix lumen. Almost 2% of patients who refer to emergency service (ER) with abdominal pain, have AA (1). Diagnosis of AA based upon the patients' symptoms, history, and physical examination findings. The similarity of the symptoms to other diseases often complicates and delays diagnosis, which may lead to the development of appendectomy perforation (2). In operations performed with the suspicion of AA, the negative appendectomy rate is 10-15%. With the advances in imaging techniques, this ratio has decreased. AA scoring systems are designed to assist in the follow-up, advanced examination and direct the surgical planning of patients with suspected appendicitis (3,4). Many scoring systems are used to better evaluate the diagnosis. The main purpose of this study was to compare Alvarado score, acute inflammatory response score (AIR), and adult appendicitis score (AAS) with imaging methods in our cohort for determining which scoring system is the most diagnostically accurate method.

MATERIALS and METHODS

Patient data

The local ethics committee has been approved for this study (approval number: E-18-1994). All patients that presented to the ER of a tertiary reference hospital with the complaint of right lower abdominal pain and suspected of having AA between February 2018 and December 2018 were prospectively examined and included to the study. A form was created which including parameters of three scoring systems; Alvarado scoring system, AIR and AAS and filled out at patients' first presentation to ER by staff physician. Patients who were under age 18 have excluded from study. Experienced surgeons evaluated the patients with their clinical laboratory and imaging findings. After getting the diagnosis of AA, the patients were hospitalized and undergone surgery. All patients have at least one imaging study ultrasonography (USG) or computed tomography (CT). Ultrasonography (USG) or computed tomography (CT) or both have been applied to all patients.

Received: 08.06.2020 **Accepted:** 23.11.2020 **Available online:** 24.12.2020
Corresponding Author: Birkan Birben, Department of General Surgery, Ankara City Hospital, Ankara, Turkey
E-mail: birkanbirben53@gmail.com

The demographic characteristics and USG, CT and pathology results of the patients were evaluated together with their Alvarado, AIR and AAS appendicitis scores. Data were collected from the forms that were fulfilled at preseantaton of patients to ER. Scoring systems was calculated after study period by authors retrospectively. Table 1 summarizes the scoring details of each system used in this study. The gold standard diagnosis method for AA has been determined as the final pathology.

Statistical analysis

All statistical analyses were performed using SPSS 16 software. Normality has been defined with the Shapiro-Wilk test. All values except specific ones are stated with mean ± standard deviation or counts (percentage). Model discrimination was measured by the area under the receiver–operator characteristic curve (AUC) to evaluate how well the model distinguished patients experienced the events from those who did not. The best model for discrimination of patients who have experienced the events or not was determined by the area under the receiver–operator characteristic curve (AUC). The discrimination of diagnostic models has defined via AUC values; 1 as perfect, >0.8 as good, 0.6-0.8 as moderate, and <0.6 as poor.

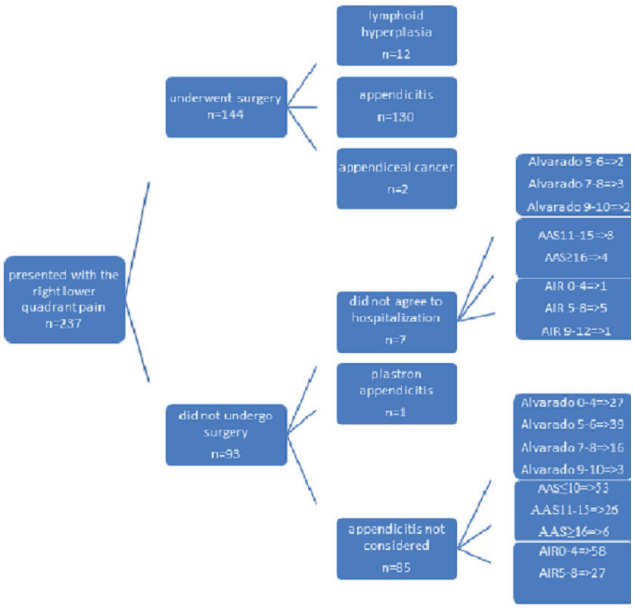
Table 1. Appendicitis scoring

Diagnosis	Alvarado Score	AIR	AAS
Migration of pain to the right lower quadrant	1		2
Right lower quadrant pain		1	2
Lack of appetite	1		
Nausea or vomiting	1		
Vomiting		1	
Right lower quadrant tenderness	2		Females aged 16-49 years 1 Others 3
Rebound tenderness	1		
Defense/rebound tenderness			
mild		1	2
moderate		2	4
severe		3	
Temperature ^{0}C			
>37.5	1		
>38.5		1	
Neutrophil (%)			
70-84		1	
≥85		2	
>62<75			2
≥75 <83			3
≥83			4
CRPg/l			
10-49		1	
≥50		2	
CRPg/l symptom<24h			
≥4 <11			2
≥11 <25			3
≥25 <83			5
≥83			1
CRPg/l symptom>24h			2
≥12 <53			2
≥53 <152			1
≥152			
Leukocytosis x10^9			
>10	2		
≥7.2 <10.9			1
≥10.9 <14			2
≥14			3
>10 <14.9		1	
≥15		2	
Leukocytosis with left shift	1		
Total score	10	12	23

Alvarado scoring: 0-4 not likely appendicitis, 5-6 equivocal, 7-8 probably appendicitis, 9-10 highly likely appendicitis; AIR (Acute Inflammatory Response score scoring): 0-4 low probability, 5-8 indeterminate group, 9-12 high probability of appendicitis; AAS (Adult Appendicitis Score) scoring: ≤10 low, 11-15 intermediate, ≥16 high probability of appendicitis

RESULTS

There were 237 patients in total. 46.8% of patients were female and 52.3% of patients were male. The patients mean age was 34 ± 13 (18-95) years. Figure 1 presents the detailed data of the patients that presented to the ER with right lower quadrant pain. Appendectomy was performed in 144 patients (61%) with the prediagnosis of AA. The pathological results were appendiceal cancer in two patients, lymphoid hyperplasia in 12, and AA in 130. Seven patients who were clinically and radiologically considered to have AA did not agree to surgery. One patient received medical treatment due to plastron appendicitis. For eighty-five patients (36%), the symptoms were not found to be related to AA following sequential physical examination and necessary imaging methods, with some having non-specific abdominal pain and others having alternative diagnoses. Table 2 shows the comparison of Alvarado, AIR and AAS appendicitis scores. At least one of USG and CT imaging methods was performed in all patients. In 199 patients, the first imaging modality applied was USG, and 137 patients underwent CT.



Alvarado: Alvarado Score, AAS: Adult Appendicitis Score, AIR: Acute Inflammatory Response score

Figure 1. Detailed data of the patients presenting with the right lower quadrant pain

Table 2. Comparison of appendicitis scoring systems

	Alvarado Score				AAS			AIR		
	Normal 41 (17.3%)	Compatible 87 (36.7%)	Probable 76 (32%)	Very probable 33 (13.9%)	Low 81 (34.3%)	Intermediate 91 (38.6%)	High 64 (27.1%)	Low 92 (39%)	Intermediate 119 (50.4%)	High 25 (10.6%)
Pathology										
Lymphoid hyperplasia	3 (7.3%)	6 (6.8%)	3 (3.9%)		5 (6.2%)	5 (5.5%)	2 (3.1%)	7 (7.6%)	4 (3.4%)	1 (4%)
Appendicitis	11 (26.8%)	38 (43.6%)	53 (69.7%)	28 (84.8%)	22 (27.2%)	55 (60.4%)	52 (81.3%)	25 (27.2%)	81 (68.1%)	23 (92%)
Cancer		1 (1.1%)	1 (1.3%)		1 (1.2%)	1 (1.1%)		1 (1.1%)	1 (0.8%)	
No surgery	27 (65.8%)	42 (48.5%)	19 (25.1%)	5 (15.2%)	53 (65.4%)	30 (33%)	10 (15.6%)	59 (64.1%)	33 (27.7%)	1 (4%)

AAS: Adult Appendicitis Score, AIR: Acute Inflammatory Response score

Of the 92 (46.2%) patients with a normal appendix on USG, 28 (30.4%) had AA, two (2.2%) had lymphoid hyperplasia, and one (1.1%) had appendiceal cancer. One patient with suspicion of AA did not agree to hospitalization. AA was not considered in 60 patients. Of the 27 patients (13.6%) suspected to have AA based on the USG results, 11 were found to have AA (40.7%), three (11.1%) lymphoid hyperplasia, and one (3.7%) appendiceal cancer. Plastron appendicitis was considered in one patient, who subsequently received medical therapy. Three patients suspected to have AA on USG did not give consent for hospitalization. AA was not considered in eight patients (29.6%). Of the 80 patients (40.2%) evaluated to have AA according to the USG results, 73 (6.3%) were confirmed to have AA and five (6.3%) were determined to have lymphoid hyperplasia. One patient considered to have AA was not hospitalized. An AA diagnosis was not considered in one patient (1.3%).

Of the 51 patients (37.2%) reported to have a normal appendix according to the CT results, one was found to have AA (2%). This imaging method revealed suspicion of AA in 23 patients (16.8%), of whom seven (30.5%) had lymphoid hyperplasia and one (4.3%) had appendiceal cancer.

Table 3. Sensitivity, specificity, positive predictive value and negative predictive value of the scoring systems and imaging methods

	Sensitivity	CI 95%	Specificity	CI 95%	PPV	CI 95%	NPV	CI 95%
Alvarado Score	61.3%	(52.94-70.18%)	69.23%	(38.57-90.91%)	95.29%	(89.86-97.88%)	15.25%	(10.55-21.55%)
AAS	82.95%	(72.33-88.99%)	42.86%	(17.66-71.14%)	93.04%	(89.41-95.49%)	%21.43%	(11.77-35.79%)
AIR	80.62%	(72.74-87.05%)	57.14%	(28.86-82.34%)	94.55%	(90.39-96.96%)	%24.24%	(15.27-36.23%)
CT	98.51%	(91.96-99.96%)	83.33%	(71.48-91.71%)	86.84%	(78.93-92.08%)	%98.04%	(87.69-99.72%)
USG	75.21%	(66.38-82.73%)	77.78%	(67.17-86.27%)	83.02%	(76.25-88.16%)	%68.48 %	(60.81-75.26%)

AAS: Adult Appendicitis Score, AIR: Acute Inflammatory Response score, CT: Computed Tomography, USG: Ultrasonography, CI: Confidence Interval, PPV: Positive Predictive Value, NPV: Negative Predictive Value

AA diagnosis was not considered in eight patients (34.7%). Among 63 patients (46%) that were suspected to have AA on CT, AA was confirmed in 55 (87.3%), lymphoid hyperplasia was present in one case (1.6%), and appendiceal cancer in another case (1.6%). Four patients considered to have AA did not accept undergoing surgery. One patient was given medical treatment with the diagnosis of plastron appendicitis. AA diagnosis was not considered in one patient (1.6%).

Table 3 shows the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the scoring systems and imaging methods. When the data were evaluated, it was observed that CT was more sensitive and specific in the diagnosis of AA. Imaging methods were more specific than scores, and the PPV of scoring systems were more sensitive than that of imaging methods. In addition, the NPV of imaging method was more sensitive compared to the scoring systems. All data were evaluated using a receiver operating characteristic (ROC) curve analysis, and the results are given in Table 4. In the clinical diagnosis of AA, the AIR score was found to be more significant than the remaining scores.

Table 4. The results of ROC analysis

	AUC	CI 95%
CT	0.736	(0.628-0.844)
USG	0.540	(0.426-0.655)
Alvarado Score	0.719	(0.617-0.821)
AAS	0.738	(0.639-0.836)
AIR	0.781	(0.690-0.873)

AUC: Area Under Curve, CI: Confidence Interval, AAS: Adult Appendicitis Score, AIR: Acute Inflammatory Response score

DISCUSSION

The lifetime prevalence of AA is 7%. Despite this considerable frequency, the similarity of the symptoms with other diseases complicates the diagnostic process (2). Castro et al. (3) emphasized that 46% of the patients presenting to the ER with abdominal pain and evaluated with AA suspicion were treated surgically. However, in our study, 61% of the patients underwent surgery. The higher rate of patients who underwent surgery in our study was attributed to the routine application of imaging in all patients.

Scoring systems have been developed to reduce the rates of negative appendectomy, as well as appendix perforation. Negative appendectomy rates are reported as 15% in the literature. This rate is further decreased to 10% with the addition of USG and 5-10% with CT (3). In the current study, at least one of these two imaging modalities was used, and the negative appendectomy rate was 9.7%.

The Alvarado system is one of the most used scoring systems in the diagnosis of AA. In their analysis, Ohle et al. (5) stated that at the cut-off value of 7, this scoring system had 82% sensitivity and 81% specificity. In another study, Frountzas et al. (6) reported the sensitivity, specificity and area under the curve (AUC) of the Alvarado score as 69%, 77% and 0.7944, respectively. In a similar study, Maghrebi et al. (7) determined that at a cut-off value 8, the sensitivity of the Alvarado score was 81.25%. In the current study, the results showed that at the cut-off value of 7, the Alvarado score had lower sensitivity, specificity and AUC values than reported in the literature.

Considering that the Alvarado score is not sufficient for the diagnosis of AA, researchers have evaluated new scoring systems with different parameters, one of which is AAS. Sammalkorpi et al. (8) found that at a cut-off value of 11, AAS had a sensitivity of 94.7% and specificity of 60.2%. In our study, when the cut off value was accepted as 11, the results were lower compared to the literature. Another scoring system is AIR. Karami et al. (9), comparing different scoring systems, found that the sensitivity and specificity were 78.41% and 91.67%, respectively for the AIR score at a cut-off value of 4, and 78.41% and 100%, respectively for the Alvarado score at a cut-off value of 7. The authors noted that according to the results of ROC curve analysis, the AUC of the Alvarado score was 0.906 and the AUC of

the AIR score was 0.867. In our study, when the cut-off value was accepted as 4, the specificity and AUC values were lower than reported in the literature.

Clinicians frequently use CT and USG as an imaging method in the diagnosis of AA. Al-khayal et al. (10) indicated that USG had a specificity of 95.9% and a sensitivity of 83.7% in the diagnosis of AA. Eng et al (11) showed that USG had 90.9% specificity and 83.1% sensitivity. In the same study CT had 93.6% specificity and 89.9% sensitivity. In the current study, USG had a sensitivity of 75.21%, a specificity of 77.78%, and AUC of 0.540, whereas these values were 98.51%, 83.33%, and 0.736, respectively for CT. Compared to CT, the non-invasive nature and ease of use were considered to be the advantages of USG while it had the disadvantage of lower sensitivity and specificity. Otherwise, although CT had the advantage of high specificity and sensitivity, ionizing radiation is a disadvantage. Smith-Bindman et al. (12) indicated the increase in the rate of cancer especially in young people who have exposed to CT-related radiation.

In the literature, at the cut-off value of 7 and above, the Alvarado score was reported have a PPV of 97.3% and NPV of 10.94% (13). In a study on AIR, at the cut-off value of 5 and above, PPV and NPV were calculated as 49% and 94%, respectively (14). In another study evaluating AIR, it was noted that PPV was 98.57% and NPV was 36.67% (9). In the current study, it was seen that the AIR score was better in the diagnosis of AA compared to other scoring and imaging methods. Although the Alvarado score is used frequently, its success rate was found to be lower in the diagnosis of AA. Although the success rate of the AAS score was better than that of Alvarado, the former was more difficult to use with a more complicated calculation system. The scoring systems evaluated in this study were determine to have higher PPVs but much lower NPVs compared to the imaging methods.

CONCLUSION

Despite the better values obtained from the AIR score compared to other systems, we consider that there is still a lack of an effective scoring system that can serve as a guide in the diagnosis of AA. Due to the low negative predictive values of the available scoring systems, patients should be evaluated with detailed anamnesis, examination and laboratory findings, and CT may be the preferred imaging option if there is clinical suspicion of acute appendicitis.

Conflict of interest : The authors declare that they have no competing interest.

Financial Disclosure: There are no financial supports.

Ethical approval: Ethical approval was obtained from Ankara Numune Training and Research Hospital. Approval number: E-18-1994

REFERENCES

1. Yesiltas M, Karakas DO, Gokcek B, et al. Can Alvarado and Appendicitis Inflammatory Response scores evaluate the severity of acute appendicitis? Ulusal travma ve acil cerrahi dergisi. TJTES 2018;24:557-62.
2. Shogilev DJ, Duus N, Odom SR, et al. Diagnosing appendicitis: evidence-based review of the diagnostic approach in 2014. West J Emerg Med 2014;15:859.
3. De Castro S, Unlu C, Steller EP, et al. Evaluation of the appendicitis inflammatory response score for patients with acute appendicitis. World J Surg 2012;36:1540-5.
4. Altieri MS, Yang J, Nie L, et al. Incidence of cholecystectomy after bariatric surgery. Surg Obes Relat Dis 2018;14:992-6.
5. Ohle R, O'Reilly F, O'Brien KK, et al.The Alvarado score for predicting acute appendicitis: a systematic review. BMC Med 2011;9:139.
6. Frountzas M, Stergios K, Kopsini D, et al. Alvarado or RIPASA score for diagnosis of acute appendicitis? A meta-analysis of randomized trials. Int J Surg 2018;56:307-14.
7. Maghrebi H, Maghraoui H, Makni A, et al. Role of the Alvarado score in the diagnosis of acute appendicitis. Pan Afr Med J 2018;29:56.
8. Sammalkorpi H, Mentula P, Savolainen H, et al. The introduction of Adult Appendicitis Score reduced negative appendectomy rate. Scand J Surg 2017;106:196-201.
9. Karami MY, Niakan H, Zadebagheri N, et al. Which One is Better? Comparison of the Acute Inflammatory Response, Raja Isteri Pengiran Anak Saleha Appendicitis and Alvarado Scoring Systems. Ann Coloproctol 2017;33:227.
10. Al-Khayal KA, Al-Omran MA. Computed tomography and ultrasonography in the diagnosis of equivocal acute appendicitis. A meta-analysis. Saudi Med J 2007;28:173-80.
11. Eng KA, Abadeh A, Ligocki C, et al. Acute appendicitis: a meta-analysis of the diagnostic accuracy of US, CT, and MRI as second-line imaging tests after an initial US. Radiology 2018;288:717-27.
12. Smith-Bindman R, Lipson J, Marcus R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Arch Intern Med 2009;169:2078-86.
13. do Nascimento RR, Souza JCG, Alexandre VB, et al. Association between the Alvarado score and surgical and histopathological findings in acute appendicitis. Rev Col Bras Cir 2018;45.
14. Scott A, Mason S, Arunakirinathan M, et al. Risk stratification by the Appendicitis Inflammatory Response score to guide decision-making in patients with suspected appendicitis. Br J Surg 2015;102:563-72.