



# Predictive value of scoring systems for the diagnosis of acute appendicitis in emergency department patients: Is there an accurate one?

Hong Kong Journal of Emergency Medicine  
2020, Vol. 27(5) 262–269  
© The Author(s) 2019  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1024907919840175  
journals.sagepub.com/home/hkj  


Rohat Ak<sup>1,2</sup>, Fatih Doğanay<sup>2,3</sup>, Ebru Unal Akoğlu<sup>2</sup>, Haldun Akoğlu<sup>4</sup>, Aslı Bahar Uçar<sup>2,5</sup>, Erdem Kurt<sup>2,6</sup>, Cansu Arslan Turan<sup>2</sup> and Ozge Onur<sup>4</sup>

## Abstract

**Background:** Acute appendicitis is one of the challenging surgical conditions presented in the emergency departments. Clinical scoring systems were developed to reduce the negative appendectomy rate and also to avoid unnecessary diagnostic evaluation.

**Objectives:** The primary aim was to compare the clinical adequacy of the Alvarado, Acute Inflammatory Response, and the Raja Isteri Pengiran Anak Saleha Appendicitis scores in patients with right lower quadrant pain for the diagnosis of acute appendicitis.

**Methods:** This was a prospective and observational study. All patients over the age of 18 years who presented with a complaint of right lower quadrant pain were enrolled. The Alvarado, Acute Inflammatory Response, and Raja Isteri Pengiran Anak Saleha Appendicitis scoring systems were compared. The patients were either admitted or followed-up as out-patient. Face-to-face or telephone follow-up visits were arranged for the patients who did not have surgery and who were not admitted.

**Results:** 232 patients were included and 14 patients were excluded from the study. Of the 218 patients, 114 patients underwent surgery. Of the 114 patients, 107 patients were pathologically diagnosed with acute appendicitis. It was determined that Raja Isteri Pengiran Anak Saleha Appendicitis score was the most valuable score with 0.88 accuracy, followed by Acute Inflammatory Response (area under the curve=0.79) and Alvarado (area under the curve=0.71) scores.

<sup>1</sup>Department of Emergency Medicine, Kartal Dr. Lütfi Kırdar Education and Research Hospital, Istanbul, Turkey

<sup>2</sup>Department of Emergency Medicine, Fatih Sultan Mehmet Education and Research Hospital, Istanbul, Turkey

<sup>3</sup>Emergency Room, Erciş Public Hospital, Van, Turkey

<sup>4</sup>Department of Emergency Medicine, Marmara University Pendik Education and Research Hospital, Istanbul, Turkey

<sup>5</sup>Emergency Room, Acibadem Hospital, Istanbul, Turkey

<sup>6</sup>Emergency Room, Kahta Public Hospital, Adiyaman, Turkey

## Corresponding author:

Rohat Ak, Department of Emergency Medicine, Kartal Dr. Lütfi Kırdar Education and Research Hospital, Şemsi Denizler Cad. E-5 Karayolu Cevizli Mevkii, Kartal, 34890 Istanbul, Turkey.

Email: rohatakmd@gmail.com



**Conclusion:** The accuracy of Raja Isteri Pengiran Anak Saleha Appendicitis scoring system was higher for the diagnosis of acute appendicitis than the other scores. The cut-off of the Raja Isteri Pengiran Anak Saleha Appendicitis score from a 7.5-point threshold provides a practical, non-invasive, rapid diagnostic method that increases acute appendicitis discriminative power in patients presenting with right lower quadrant pain.

### Keywords

Appendicitis, Alvarado, Acute Inflammatory Response, Raja Isteri Pengiran Anak Saleha Appendicitis, abdominal pain

## Introduction

Acute appendicitis (AA), the inflammation of appendix tissue, is one of the most presenting causes of abdominal pain in emergency departments (EDs). The lifetime incidence is 7% in the general population.<sup>1</sup> The diagnosis of AA remains a difficulty because of the atypical presentation, especially in young adults, elders, and females. It is easy to misdiagnose gynecological and urogenital conditions as AA because of the similar presentation symptoms and findings.<sup>2</sup>

Scoring systems were developed to choose the appropriate clinical and surgical treatment, and to reduce the negative appendectomy rates.<sup>3-5</sup> The Alvarado scoring system, a composite of symptoms, clinical, and laboratory findings, has been the most popular.<sup>6,7</sup> It has been shown that, especially in Eastern populations, the Alvarado score has a lower diagnostic accuracy.<sup>8,9</sup> The Acute Inflammatory Response (AIR) score is developed from the same clinical criteria of the Alvarado score with the addition of C-reactive protein (CRP) and the complete blood count tests, to improve the discriminating power.<sup>10,11</sup> More recently, a new scoring system, called the Raja Isteri Pengiran Anak Saleha Appendicitis (RIPASA) score was developed. This score includes 14 clinical parameters and has a higher sensitivity, specificity, and diagnostic accuracy than the Alvarado scoring system, especially in Asian populations who have a completely different ethnic origin and diet.<sup>12,13</sup> The objective of this study was to evaluate the Alvarado, the AIR, and the RIPASA scoring systems and to compare their performance to predict the risk for AA.

## Methods

### Study design

This prospective observational diagnostic accuracy study with cohort design was conducted at Fatih Sultan Mehmet Education and Research Hospital, Istanbul, Turkey, with an annual patient load of 220,000. Our prospective cohort consisted of 232 consecutive patients with right lower quadrant pain who presented to the ED between 1 January and 1 July 2016. Ethical permission was obtained prior to the beginning of this study.

### Selection of participants

All consecutive adult patients (aged over 18 years), who were admitted to the ED with right lower quadrant pain, were eligible for the study. Patients who met any of the following criteria were included in the study: (1) admitted during the shifts of the researchers and (2) without the history of appendectomy. Patients (1) who were diagnosed with pregnancy during the ED work-up and (2) who were lost to follow-up were excluded from the final analysis.

### Index tests

The Alvarado score contains 8, AIR score contains 7, and RIPASA score contains 16 variables. The scores for each of the variables are shown in Supplementary Tables E1 to E3. The optimal cut-off thresholds were accepted as  $\geq 5$  for a high probability of AA in the Alvarado scoring system,  $\geq 7.5$  for the RIPASA scoring system, and  $\geq 5$  for the AIR scoring system.<sup>13,14</sup> Scoring charts were filled by the researchers at the time of presentation. All the data necessary for the calculation of scores were collected by the researchers observationally from the patient charts or interviews.

### Reference standard

The reference standard of this study was as a composite outcome and defined as the diagnosis of AA by surgery and histopathological reports, or during the follow-up. The surgeon's clinical judgment, based on all the findings of clinical, laboratory, and radiological investigation, was the essential decision way for appendectomy or conservative treatment and follow-up. RIPASA, Alvarado, and AIR scores were only calculated for the study purpose. At the end of the study, an independent panel of experts blinded to the calculated scores of the patients made definitive AA diagnosis by analyzing the radiological imaging reports, surgery reports, histopathological results, patient charts, outpatient clinic follow-up form, and the results of the follow-up by phone.

### Data collection and study protocol

The baseline characteristics of all patients including age, sex, vital findings, medical history, clinical symptoms, physical examination findings, and radiological imaging findings were recorded for all right lower quadrant pain patients presenting to the ED. The primary survey of all patients was managed by the emergency physicians (EPs). Blood biochemistry, urine analysis, and, if necessary, radiological tests were obtained as a part of the routine workup. When an eligible patient was presented according to the criteria mentioned above, a researcher on duty was alerted for the evaluation of the patient for study inclusion. All patients were managed without any intervention. The researchers never intervened into the decision of primary or consulting physicians.

### Follow-up

After the evaluation process, patients were followed-up in three different ways, solely based on the test results and gestalt of the primary or consulting physician: (1) very low probability for AA (other non-surgical diagnoses were more probable, medical treatment was introduced, patients were scheduled for an outpatient follow-up the day after and in 2 weeks), (2) low probability for AA (the diagnosis of AA could not be excluded, but the risk was low, and the patient was discharged with medical treatment by the consulting surgeon with an outpatient follow-up the day after and in 2 weeks), (3) moderate-to-high probability for AA (those patients were admitted to a hospital ward for further evaluation or surgery). Patients who had definitive diagnosis of AA, but scheduled for conservative treatment, were called for follow-up in 2 weeks, and patients who did not come were reached by telephone. Patients, who were reached by telephone, were asked whether they presented to another health institution with the same complaint in the following 2 weeks or underwent surgery in another hospital. Those who did not come to follow-up and those who cannot be reached by telephone were excluded from the study.

### Outcome

Patients were categorized into two groups: AA group (Group AA) which included the patients who underwent surgery and the diagnosis of AA was confirmed with histopathological reports. Non-appendicitis group (Group N-A) which included the patients to whom surgical intervention was performed and had negative appendectomy, who had pathologies other than AA, who were followed-up conservatively.

### Statistical analysis

The analyses were conducted with MedCalc Statistical Software version 18.6 (MedCalc Software bvba, Ostend,

Belgium; <http://www.medcalc.org>; 2018). Continuous variables were reported with mean values, standard deviations, and 95% confidence intervals (CIs) or medians and interquartile ranges according to the distribution pattern of the variable assessed by the Shapiro–Wilks test. Mean values were compared among independent groups with Student's t-test or analysis of variance (ANOVA) for normal distributions, and medians with Mann–Whitney U test or Kruskal–Wallis test for non-normal distributions. Dependent groups were compared using paired sample t-test, or Wilcoxon sign test for normal and non-normal distributions, respectively. Proportions were compared among discrete groups by chi-square test and Fisher's exact test with Continuity (Yates) Correction. Receiver operating characteristic (ROC) curves were used to identify the optimal cut-off points. Contingency tables were used to calculate sensitivity, specificity, and the diagnostic accuracy values of the scoring systems.

The sample size calculation was based on a significance level of 0.05. A sample of 144 patients to achieve 80% power was calculated. It was hypothesized that 20% were lost to follow-up, and the estimated sample size was calculated as 175. For the patients who were lost to follow-up, a sensitivity analysis was performed. In this study, the accepted type 1 error was 5%.

### Results

A total of 232 patients with right lower quadrant pain were included, 14 patients were excluded; therefore, 218 patients completed the follow-up of the study. Of 218 patients, 114 patients were hospitalized and underwent to surgery since they were presumed to have moderate-to-high pretest probability for AA. Demographics, symptoms, physical examination findings, duration of symptoms, and laboratory findings are summarized in Table 1, and the final diagnosis of the patients in the study is shown in Table 2. A diagnostic radiological test—either computed tomography (CT) or ultrasound—was performed on all of the patients. Histopathologically, 107 of 114 patients were diagnosed with confirmed AA, and the rest were diagnosed as (n=7) negative appendectomies (lymphoid hyperplasia in 4 patients and fibrous obliteration in 3 patients). The negative appendectomy rate was 7%. The flow diagram of the patients is presented in Figure 1.

The areas under the curve (AUCs) of the scores were compared, and it was found that RIPASA score was the most valuable score with an accuracy of 0.88, followed by AIR (AUC=0.79) and Alvarado scores (AUC=0.71) (Table 3). The comparison of the diagnostic accuracies of the RIPASA, AIR, and Alvarado scores was also statistically significant and shown in Figure 2 and Table 4.

Based on the optimal cut-off values with the highest discriminative power, sensitivity and specificity of AIR score for  $\geq 5$  points was 94.39% and 26.13%; the

**Table 1.** The demographics and clinical variables.

	All patients	Group AA (n= 107)	Group N-A (n= 111)	p
Age (year), median (IQR)	33.00 (25.00 to 44.00)	30 (23.00 to 40.75)	34 (26.00 to 46.75)	<b>0.0120<sup>a</sup></b>
Male, n (%)	105 (48.2)	69 (65.7)	36 (34.3)	<b>&lt;0.0001<sup>a</sup></b>
Vital signs				
Fever (C), median (IQR)	36.40 (36.20 to 36.60)	36.40 (36.20 to 36.50)	36.40 (36.20 to 36.60)	0.3928 <sup>a</sup>
Clinical features				
Loss of appetite, n (%)	98 (45.0)	60 (61.2)	38 (38.8)	<b>0.0012<sup>b</sup></b>
N&V, n (%)	125 (57.3)	73 (58.4)	52 (41.6)	<b>0.0015<sup>b</sup></b>
Pain migration to RLQ, n (%)	73 (33.5)	49 (67.1)	24 (32.9)	<b>0.0002<sup>b</sup></b>
Duration of symptoms $\geq$ 48 h, n (%)	163 (74.8)	89 (54.6)	74 (45.4)	<b>0.0051<sup>b</sup></b>
Physical examination findings				
Guarding, n (%)	128 (58.7)	101 (78.9)	27 (21.1)	<b>&lt;0.0001<sup>b</sup></b>
Rebound tenderness, n (%)	106 (48.6)	73 (68.9)	33 (31.1)	<b>&lt;0.0001<sup>b</sup></b>
Rovsing's sign, n (%)	90 (41.3)	73 (81.1)	17 (18.9)	<b>&lt;0.0001<sup>b</sup></b>
Hell-drop test, n (%)	107 (49.1)	74 (69.2)	33 (30.8)	<b>&lt;0.0001<sup>b</sup></b>
Laboratory findings				
CRP (mg/L), median (IQR)	0.84 (0.12 to 2.86)	1.31 (0.35 to 3.71)	0.59 (0.06 to 2.15)	<b>0.0069<sup>a</sup></b>
MPV (fL), median (IQR)	7.20 (6.60 to 7.80)	7.00 (6.50 to 7.70)	7.30 (6.70 to 8.10)	<b>0.0403<sup>a</sup></b>
Thrombocyte (mm <sup>3</sup> /mL), median (IQR)	340.00 (278.00 to 377.00)	345.00 (278.00 to 383.25)	324.00 (278.00 to 369.25)	0.2838 <sup>a</sup>
PMN leucocyte (%), median (IQR)	70.05 (60.90 to 79.10)	76.60 (67.85 to 81.90)	63.50 (57.50 to 72.90)	<b>&lt;0.0001<sup>a</sup></b>
WBC (mm <sup>3</sup> /mL), median (IQR)	11.35 (8.50 to 15.10)	14.10 (10.85 to 16.58)	9.40 (7.63 to 11.60)	<b>&lt;0.0001<sup>a</sup></b>

Group AA: acute appendicitis group; Group N-A: non-appendicitis group; IQR: interquartile range; CRP: C-reactive protein; AA: acute appendicitis; MPV: mean platelet volume; PMN: polymorphonuclear neutrophils; WBC: white blood cell, N&V: nausea and vomiting, RLQ: right lower quadrant.

<sup>a</sup>Mann-Whitney U test.

<sup>b</sup>Chi-square test.

**Table 2.** Final ED diagnosis of the patients participated to the study.

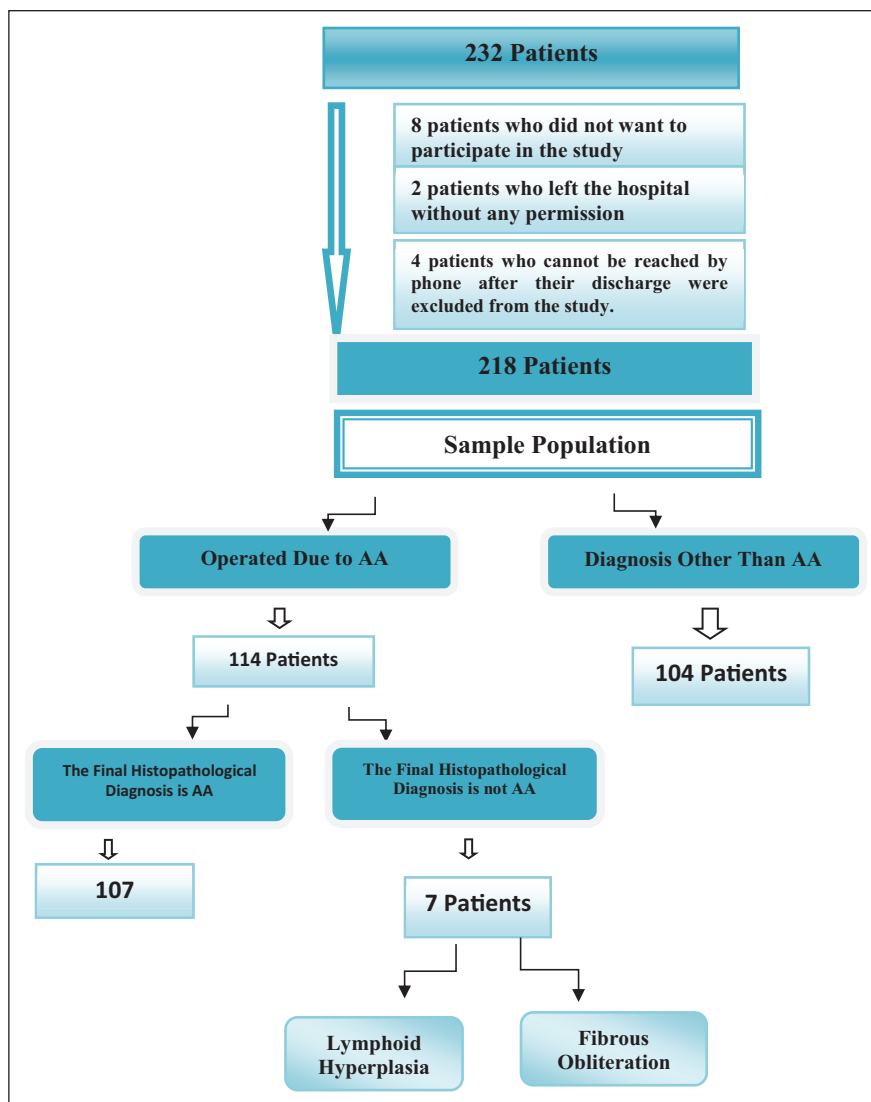
Final diagnosis	n (%)
Histopathological diagnosis	
Acute appendicitis	107 (49.1)
No appendicitis	7 (3.2)
Radiological diagnosis	
NSAP	33 (15.1)
Acute diverticulitis	5 (2.3)
Over cyst rupture	4 (1.8)
Colon tumor	1 (0.5)
Acute gastroenteritis	9 (4.1)
PID	5 (2.3)
Renal colic	13 (6.0)
Orchitis/epididymitis	1 (0.5)
Urinary tract infection	28 (12.8)
Acute vaginitis	2 (0.9)
Rectus sheath hematoma	1 (0.5)
Median arcuate ligament syndrome	1 (0.5)
Ectopic pregnancy	1 (0.5)
Total	218

ED: emergency department; NSAP: non-specific abdominal pain; PID: pelvic inflammatory disease.

sensitivity and specificity of Alvarado score for  $\geq 5$  points was 72.90% and 54.05%; and the sensitivity and specificity of RIPASA score for  $\geq 7.5$  points was 91.59% and 65.77%, respectively (Table 5). If a cut-off threshold score of  $\geq 7$  was taken as the highest discriminative power for the RIPASA, the sensitivity of the scoring system to rule-out AA would be 99%. However, the specificity would decrease to 50%.

According to histopathological results of the appendectomy patients, negative appendectomy rates for RIPASA, AIR, and Alvarado were 14.28%, 14.29%, and 71.43%, respectively. In 85.7% (n=6) of the negative appendectomy patients, a score of  $\geq 5$  for AIR, a score of  $\geq 7.5$  for RIPASA, and a score of  $\geq 5$  for Alvarado score were reported.

The National Registration Identity Card (NRIC) is an additional parameter of the RIPASA scoring system. NRIC is used only in Singapore. In previous studies, Shuaib et al.<sup>1</sup> and Malik et al.<sup>15</sup> evaluated the utility of RIPASA score in different populations without using this parameter and also reported a good sensitivity, specificity, and diagnostic accuracy. We also performed a sensitivity analysis for the two foreign patients and found that, even if we excluded this parameter, there would be no significant changes in the results.



**Figure 1.** The flowchart of patients.

### Sensitivity analysis

To evaluate the effect of the four patients lost to follow-up, a sensitivity analysis was performed. Even it was assumed that all four patients lost-to-follow-up had very low pretest probability despite being AA, the AUCs, sensitivity, and specificity values of the scoring systems changed slightly without any statistically significant difference ( $p > 0.05$ ).

### Discussion

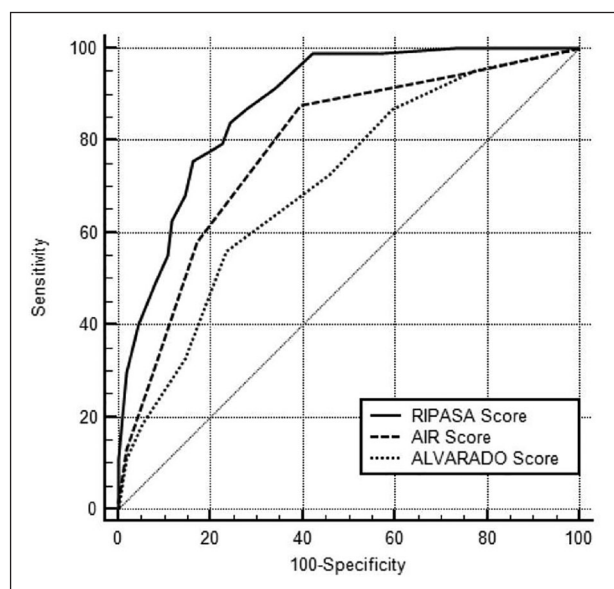
The diagnosis of AA can be challenging for physicians. A delay in appendectomy may increase the risk of perforation.<sup>16</sup> The improvement of the diagnostic accuracy and reducing the negative appendectomy rates can be achieved with the help of diagnostic tools (such as clinical signs, symptoms, or laboratory findings). Although radiological imaging procedures (ultrasound or CT) increase the

accuracy of the diagnosis of AA, the interpreter difference, high costs, and unavailability are unfortunately still the most important disadvantages.<sup>16</sup> Several scoring systems have been developed to increase the sensitivity and specificity of the diagnosis of AA.<sup>1</sup> The important thing in risk scoring systems is to minimize the unnecessary tests and operations, and protect the patients from unnecessary interventional procedures. The most famous scoring systems the Alvarado and the modified Alvarado have good sensitivity and specificity when performed in Western populations, but they have less sensitivity and specificity in Asian population. The AIR score, which has a better diagnostic accuracy than the Alvarado score, is another well-known scoring system used for AA during last decade. Because of the low sensitivity and specificity rates of these scoring systems for Asian populations, the RIPASA score, with a better sensitivity and specificity, was developed.<sup>1,11</sup>

**Table 3.** Cut-off values for the maximum sensitivity, specificity, +LR, and -LR of the scoring systems.

	AUC (95% CI)	p	Threshold	Sensitivity (95% CI)	Specificity (95% CI)	+LR	-LR
AIR score	0.79 (0.73–0.84)	<0.0001	≥5	94.39 (88.19 to 97.92)	26.13 (18.25 to 35.32)	1.28 (1.13 to 1.44)	0.22 (0.09 to 0.50)
Alvarado score	0.71 (0.64–0.77)	<0.0001	≥5	72.90 (63.45 to 81.04)	54.05 (44.33 to 63.55)	1.59 (1.26 to 2.00)	0.50 (0.35 to 0.72)
RIPASA score	0.88 (0.83–0.92)	<0.0001	≥7.5	91.59 (84.63 to 96.08)	65.77 (56.16 to 74.51)	2.68 (.05 to 3.48)	0.13 (0.07 to 0.24)

AIR: Acute Inflammatory Response; RIPASA: Raja Isteri Pengiran Anak Saleha Appendicitis; 95% CI: 95% confidence interval.

**Figure 2.** The ROCs of the AIR, Alvarado, and RIPASA scores.**Table 4.** The diagnostic accuracy of the scores and the comparison of the diagnostic accuracy with each other.

Comparison of scores	Delta AUCs (95% CI)	p
AIR score versus Alvarado score	0.08 (0.02 to 0.14)	<b>0.0117</b>
AIR score versus RIPASA score	0.10 (0.05 to 0.15)	<b>0.0002</b>
Alvarado score versus RIPASA score	0.18 (0.12 to 0.23)	<b>&lt;0.0001</b>

AUC: area under the curve; 95% CI: 95% confidence interval; AIR: Acute Inflammatory Response; RIPASA: Raja Isteri Pengiran Anak Saleha Appendicitis.

The diagnostic accuracy of the RIPASA, AIR, and Alvarado scoring systems are compared. The RIPASA score with a diagnostic accuracy of 0.88 was remarkably better than the AIR (AUC=0.79) or the Alvarado (AUC=0.71) score in diagnosing AA. The sensitivity of the AIR score (94.39%) was higher than those of the RIPASA (91.59%) and the Alvarado (72.90%). Besides, the

specificity of the RIPASA (65.77%) score was better than the Alvarado (54.05%) and the AIR (26.13%) score. In this study, a high probability cutoff score of ≥7 for the RIPASA showed a better sensitivity (99%).

In our study, the Alvarado cut-off point was set at ≥5 and yielded a sensitivity of 72.90%, a specificity of 54.05%, and the diagnostic accuracy of 0.71. In the study of Shuaib et al.,<sup>1</sup> the cut-off threshold was set at 7 and a sensitivity of 82.8% and a specificity of 56% were found for the modified Alvarado score. Memon et al.<sup>17</sup> reported a sensitivity of 93.5% and a specificity of 80.6% for the Alvarado score at a cut-off threshold of 6. Also, Gwynn<sup>18</sup> investigated the diagnostic utility of the Alvarado score and found a sensitivity of 91.6% and a specificity of 84.7%. It was assumed that the reason of the large difference in the sensitivity rates between the study of Gwynn and this study was the method used for the AA diagnosis. In the previous study, radiological imaging findings were used for the AA diagnosis, not histopathological results. CT is a valuable tool in the evaluation of AA. Based on published reports the sensitivity and the specificity of CT for the diagnosis of AA were 94.1% and 96.4%, respectively. In our study, 4 of 107 AA patients had negative CT findings, and 6 of 7 negative appendectomy patients had positive CT findings. In the systematic review by Ohle et al.,<sup>19</sup> a sensitivity of 99% and a specificity of 43% were reported for Alvarado score at a cut-off point of 5. Also, there are still some other factors that may have influenced the presence of lower utility measures within our study. In our study, all patients were actively followed-up. However, as mentioned in the systematic review, a bunch of studies had no follow-up at all. The lack of an active follow-up may have led the patients to be misclassified and may have inflated the estimated sensitivity and specificity. Also, the review has not searched the gray literature, this also may lead to inflated sensitivity and specificity values.<sup>19</sup> Finally, ethnicity may also play an important role in this difference. Since scoring systems were created in Western populations, the sensitivity and specificity levels may be found to be lower when applied to Asian, Oriental, or Middle-Eastern populations.

In our study, the AIR score showed a higher sensitivity but a lower specificity than the other scoring systems. A score of ≥5 showed 94% sensitivity for the patients with

**Table 5.** Distribution of acute appendicitis scores according to the last diagnosis.

Scores		All patients, n (%) (n = 218)	Acute appendicitis (n = 107), n (%)	Other diagnoses (n = 111), n (%)	p
AIR score	<5	35 (16.1)	6 (5.6)	29 (26.1)	<0.0001
	≥5	183 (83.9)	101 (94.4)	82 (73.9)	
Alvarado score	<5	89 (40.8)	29 (27.1)	60 (54.1)	<0.0001
	≥5	129 (59.2)	78 (72.9)	51 (45.9)	
RIPASA score	<7.5	82 (37.6)	9 (8.4)	73 (65.8)	<0.0001
	≥7.5	136 (62.4)	98 (91.6)	38 (34.2)	

AIR: Acute Inflammatory Response; RIPASA: Raja Isteri Pengiran Anak Saleha Appendicitis. Percentages are calculated from column total.

mild to high probability of AA. Andersson and Andersson<sup>10</sup> demonstrated a sensitivity of 96% and a specificity of 73% for an AIR score of ≥5. In our study, the diagnostic power of the AIR score (AUC=0.79) was significantly better than the Alvarado score (AUC=0.71) ( $p<0.05$ ). De Castro et al.<sup>14</sup> compared the diagnostic utility of the AIR score and the Alvarado score in their study and showed similar results with our study.

In this study, in a Middle-East population, the RIPASA score ≥7.5 showed 91.59% sensitivity and 65.77% specificity. The RIPASA scoring system had a significantly higher AUC value of 0.88 than the other scoring systems. In the study by Chong et al.,<sup>13</sup> in an Asian population, patients with a RIPASA score higher than 7 were classified as high risk, and the sensitivity and the specificity values were found to be 88% and 67%, respectively. Nanjundiah et al.<sup>20</sup> compared the diagnostic power of the RIPASA and the Alvarado scores, the sensitivity and the specificity of the scoring systems were 96.2% and 58.9% and 90.5% and 85.7%, respectively. In the study of Shuaib et al.,<sup>1</sup> a sensitivity of 94.5% and a specificity of 88% were reported.

### Limitations

The major limitation of our study was the one-centered design. In our study, the establishment of the diagnosis of AA was based on the clinical judgment of the surgical specialist, combined with the laboratory and/or radiological imaging findings. Physical examination findings, which constituted the parameters of the scoring system, were subjective variables and could be interpreted differently according to the experience of the applying physician. The laboratory findings that constituted the scoring systems (e.g. WBC, CRP, and BUN) could influence the outcome according to the time of presentation. Laboratory findings also, could be easily affected by the accompanying diseases and the time of onset of pain, and thus, could have the potential to affect the accuracy of the calculated results. Moreover, the criteria of foreign patient, which differentiated the RIPASA scoring system from other scoring systems, could vary for every region and ethnicity. Due to the low number of foreign patients, presenting to our hospital,

the data obtained were limited. Different results could have been obtained, if the same study was performed at another region.

### Conclusion

The diagnostic accuracy of the RIPASA scoring system is better than the Alvarado and the AIR scores. It was concluded that the RIPASA score is a useful, basic, fast, and non-invasive diagnostic tool. We found a sensitivity of 91.59% and a specificity of 65.77% for the RIPASA score at a cut-off threshold score of 7.5. We also demonstrated a greater sensitivity than that was reported in Eastern populations (92% vs 88%), but a similar specificity (66% vs 66%).

### Author contributions

R.A., F.D., C.A.T., E.K., and A.B.U. contributed to the literature search; E.U.A., R.A., H.A., and O.O. contributed to the study design; R.A., E.U.A., H.A., A.B.U., and O.O. contributed to the legislative applications; F.D., R.A., A.B.U., E.K., and C.A.T. contributed to the data collection; E.U.A., H.A., and O.O. contributed to the supervision and quality control; E.U.A. and HA contributed to the statistical advice and statistical data analysis; R.A., E.U.A., and H.A. contributed to the data interpretation; R.A., F.D., E.K., E.U.A., and H.A. contributed to the drafting the manuscript. All authors were involved in the writing and critical revision of the manuscript and approved the final version. R.A. and E.U.A. take responsibility for the paper as a whole.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### Availability of data and materials

The authors agree to the conditions of publication including the availability of data and materials in our manuscript.

### Ethical approval

This study was approved by the local ethics committee.

### Human rights

The principles outlined in the Declaration of Helsinki have been followed.

### Informed consent


Informed consent was obtained from the participants or their legally authorized representatives.

### Supplemental material

Supplemental material for this article is available online.

### ORCID iDs

Rohat Ak  <https://orcid.org/0000-0002-8324-3264>

Ebru Unal Akoğlu  <https://orcid.org/0000-0003-3674-133X>

### References

- Shuaib A, Shuaib A, Fakhra Z, et al. Evaluation of modified Alvarado scoring system and RIPASA scoring system as diagnostic tools of acute appendicitis. *World J Emerg Med* 2017; 8(4): 276–280.
- Jones PF. Active observation in management of acute abdominal pain in childhood. *Br Med J* 1976; 2(6035): 551–553.
- Zouari M, Jallouli M, Louati H, et al. Predictive value of C-reactive protein, ultrasound and Alvarado score in acute appendicitis: a prospective pediatric cohort. *Am J Emerg Med* 2016; 34(2): 189–192.
- Golden SK, Haringa JB, Pickhardt PJ, et al. Prospective evaluation of the ability of clinical scoring systems and physician-determined likelihood of appendicitis to obviate the need for CT. *Emerg Med J* 2016; 33(7): 458–464.
- Di Saverio S, Birindelli A, Kelly MD, et al. WSES Jerusalem guidelines for diagnosis and treatment of acute appendicitis. *World J Emerg Surg* 2016; 11: 34.
- Owen TD, Williams H, Stiff G, et al. Evaluation of the Alvarado score in acute appendicitis. *J R Soc Med* 1992; 85(2): 87–88.
- Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med* 1986; 15(5): 557–564.
- Khan I and ur Rehman A. Application of Alvarado scoring system in diagnosis of acute appendicitis. *J Ayub Med Coll Abbottabad* 2005; 17(3): 41–44.
- Jang SO, Kim BS and Moon DJ. Application of Alvarado score in patients with suspected appendicitis. *Korean J Gastroenterol* 2008; 52(1): 27–31.
- Andersson M and 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(8): 1843–1849.
- 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(6): 227–231.
- Butt MQ, Chatha SS, Ghumman AQ, et al. RIPASA score: a new diagnostic score for diagnosis of acute appendicitis. *J Coll Physicians Surg Pak* 2014; 24(12): 894–897.
- Chong CF, Adi MI, Thien A, et al. Development of the RIPASA score: a new appendicitis scoring system for the diagnosis of acute appendicitis. *Singapore Med J* 2010; 51(3): 220–225.
- De Castro SM, Unlu C, Steller EP, et al. Evaluation of the appendicitis inflammatory response score for patients with acute appendicitis. *World J Surg* 2012; 36(7): 1540–1545.
- Malik MU, Connelly TM, Awan F, et al. The RIPASA score is sensitive and specific for the diagnosis of acute appendicitis in a western population. *Int J Colorectal Dis* 2017; 32(4): 491–497.
- Erdem H, Cetinkunar S, Das K, et al. Alvarado, Eskelinen, Ohlmann and Raja Isteri Pengiran Anak Saleha Appendicitis scores for diagnosis of acute appendicitis. *World J Gastroenterol* 2013; 19(47): 9057–9062.
- Memon ZA, Irfan S, Fatima K, et al. Acute appendicitis: diagnostic accuracy of Alvarado scoring system. *Asian J Surg* 2013; 36(4): 144–149.
- Gwynn LK. The diagnosis of acute appendicitis: clinical assessment versus computed tomography evaluation. *J Emerg Med* 2001; 21(2): 119–123.
- 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.
- Nanjundaiah N, Mohammed A, Shanbhag V, et al. A comparative study of RIPASA score and Alvarado score in the diagnosis of acute appendicitis. *J Clin Diagn Res* 2014; 8(11): NC03–NC15.