Comparing the Diagnostic Accuracy of AIR and Alvarado Scores in Acute Appendicitis

Amanpreet Singh[1], Rajkamal Jenaw[2], Pankaj Porwal[3]

Abstract: _Background: Acute appendicitis (AA) is one of the most common causes of an acute abdomen. Clinical scoring systems, such as the Alvarado and Acute inflammatory response (AIR) score, were developed with the goal of reducing the negative appendectomy rate. Though the Alvarado scoring system is widely used, differences in diagnostic accuracy have been observed when applied to varied populations._ _Methods: This study aims to compare the diagnostic accuracy of the Alvarado and Acute Inflammatory Response AIR scoring systems in diagnosing acute appendicitis in an Indian population. A prospective analysis was conducted on 77 patients with right iliac fossa pain, applying both scoring systems and correlating postoperative pathology reports._ _Results: The AIR score showed a sensitivity of 74.29 and correctly classified 76.62 of confirmed appendicitis cases, while the Alvarado score had a sensitivity of 65.71 and classified 68.83 accurately. Both scores had 100 specificity. The findings suggest that the AIR score has better diagnostic accuracy compared to the Alvarado score._ _Conclusion: The AIR scoring system had more sensitivity, better NPV compared with Alvarado score while both the scores had same specificity and PPV._

Keywords: Acute appendicitis, Alvarado score, AIR score, diagnostic accuracy, Indian population

1. Introduction

Acute appendicitis is one of the most common emergency conditions in general surgical wards. Diagnosis of acute appendicitis is still a difficult entity clinically and involves complex decision making in the management as it involves surgical exploration. An earlier, effective and quick diagnosis of acute appendicitis can avoid complications arising from perforation with subsequent early appendicectomy[1]. Several clinical parameters and laboratory tests can lead to an appropriate diagnosis, which is really important in choosing the appropriate treatment. Recently, the Alvarado, the Raja Isteri Pengiran Anak Saleha Appendicitis (RIPASA), and the acute inflammatory response (AIR) scores are very commonly used clinical diagnostic scoring systems for the diagnosis of acute appendicitis[2, 3].

Acute appendicitis incidence is highest in young adults, but the incidence of complicated appendicitis shows subtle variation between different age groups[4]. It is difficult to diagnosis acute appendicitis in very young, females of reproductive age and elderly patients due to atypical presentation, and many other conditions mimicing acute appendicitis in these patients[5]. Various radiological studies like pelvi-abdominal ultrasound and Computerized Tomograph (CT) scan can help in the diagnosis of acute appendicitis[6]. Some studies reported that the use of CT has lead to changes in the decisions of treatment of many patients evaluated for appendicitis[7, 8.] It has resulted in decrease in negative appendectomy rates and has sensitivity and specificity of 95% and 94%, respectively, as per data from a recent meta-analysis[9]. Laboratory parameters like neutrophilia and increased leucocyte counts and increased C reactive protein (CRP) levels also provide diagnostic value: sensitivities for the latter range between 38 and 70% (specificities 85 and 65%, respectively) [10, 11, 12.] Presently, the gold standard for AA diagnosis is the histopathology study. Thus, we have designed this prospective cohort study to compare the diagnoses based on the Alvarado and AIR scoring systems with those obtained from histopathology and to evaluate several predictive diagnostic values.

2. Methods

This prospective cohort study was conducted at SMS Medical College and attached Hospitals. Ethical Permission in accordance with International Helsinki protocols was obtained prior to the beginning of this study. Informed consent was obtained from all patients. The study population included all patients with acute (lasting less than 4 days) non traumatic right iliac fossa pain aged 12 .

Table 1: Alvarado appendicitis scoring system

Variable	Score
Symptoms	
Migratory right iliac fossa pain	1
Anorexia	1
Nausea/vomiting	1
Signs	
Right iliac fossa tenderness	2
Rebound tenderness	1
Elevated temperature $> 37.3\ °C$	1
Laboratory tests	
Leukocytosis $> 10.0 \times 10^9/L$	2
Neutrophils $> 75\%$ or left shift	1

Guidelines for management according to the total score: <4, probability of acute appendicitis (AA) unlikely; 4–7, AA suspected; >7, definite AA.

To 70 years consistent with a diagnosis of appendicitis during the period from June 2023 to November 2023 and who were seen at SMS Medical College and attached Hospital. Patients who had a right iliac fossa mass, did not provide informed consent, had a history of urolithiasis or pelvic inflammatory disease, and children below 12 years of age and elderly above 70 years were excluded from the study. The sample size was calculated at 95 confidence and 10 precision to verify the diagnostic sensitivity of the Alvarado and AIR scores separately. The highest value of 77 cases of suspected appendicitis was calculated with the Alvarado score, which is considered the final sample size for our study. A total of 77

patients were qualified for the study during the study period. The ages of the patients ranged from 12 to 50 years. All 77 patients were scored using the AIR and Alvarado scoring systems.

Table 2: Acute inflammatory response scoring system

Diagnosis	Score
Vomiting	1
Pain in RIF	1
Abdominal defense	
Low	1
Mild	2
Severe	3
Temperature >38.5ºC	1
Segmented neutrophils	
70%–84%	1
≥85%	2
Leukocytes (×109/L)	
10.0–14.9	1
≥15.0	2
CRP (g/L)	
10–49	1
≥50	2
Total	12

AIR; sum 0–4, low probability; sum 5–8, mild probability; Sum9-12, high probability.
RIF, right iliac fossa; CRP, C-reactive protein

The Alvarado score contains 8 parameters whereas the AIR score consists of 5 parameters. The scores for each of the parameters ranged from 1 to 2 for the Alvarado system, and 1 to 3 for the AIR system, as shown in Tables 1 and 2 respectively. Scoring charts were completed by a resident on presentation. A score of more than 7 was taken as a high probability of AA for the Alvarado scoring system whereas the scores for the AIR scoring systems was more than 8. The positive (PPVs) and negative predictive values (NPVs) and diagnostic accuracy of the two scoring systems were assessed.

The patients were monitored from the time of admission until discharge from the hospital. Daily follow-up included the monitoring of vital signs twice a day and systemic examination once a day. Histopathology findings on the surgical cases were collected and correlated with the scores. Scores were tabulated, and the sensitivities, specificities, positive and NPVs, and positive and NLRs were calculated. Also, the area under the receiver operating characteristic (ROC) curves were obtained using IBM SPSS Statistics ver.21.0 (IBM Co., Armonk, NY, USA) and MedCalc ver.15 (MedCalc, Ostend, Belgium).

3. Result

The mean ± standard deviation (SD) of the patient ages in our study group was 32 ± 15 years old. Of the participants, 74% were male patients compared with 26% of female patients. Clinically, Clinically, all patients presented with acute right iliac fossa tenderness, rebound tenderness 22%, nausea and vomiting 65%, and 78% had elevated white blood count, as shown in Table 3. Histopathological analysis of appendices of the studied patients showed that 72% of the patients had acute appendicitis and only 18% had complicated perforated appendicitis. Meanwhile, about 9% had normal (negative) appendix as shown in (Fig.2).

In this study, most of the patients with acute appendicitis were found to have ALVARDO score 7-8 (52.86%) followed by score 9-10 (30%) Only 11 (15.71%) had score 5-6, while only 1 patient (1.43%) had score 3-4. Among patients without acute appendicitis, most were found to have ALVARDO score of 5-6 (42.86%) followed by 7-8 (28.57%) while none of the patient had score 9-10, while most of the patients with acute appendicitis were found to have AIR score 9-12 (74.29%) followed by score 5-8 (25.71%), while none of the patient had score <5. Among patients without acute appendicitis, most were found to have AIR score of 5-8 (85.71%) followed by score <5 (14.29%). Of the 77 patients who underwent surgery, 91% were positive for appendicitis on the histopathological report. At the optimal cutoff point of >8 for the AIR scoring system, the sensitivity and the specificity were 74.29% and 100%, respectively. Also, at the optimal cutoff point of >7 for the Alvarado scoring system, the sensitivity and the specificity were 65.71% and 100%, respectively. The PPVs and NPVs and diagnostic accuracy of the 2 scoring systems are presented in Table 4. According to the ROC curve, the area under the curve (AUC) was 0.924 for the

**Table 3:** Clinical and laboratory measures of the study sample (N = 77)

Variables	N (%)
Manifestations	
Anorexia	56 (73)
Nausea and vomiting	50 (65
Fever	49 (64)
Rt. iliac fossa tenderness	75 (97)
Guarding	53 (69)
Rebound tenderness	17 (22)
Right iliac fossa pain	77 (100)
Investigation	
Elevated WBCs	60 (78)
Neutrophilia	65 (84)

Table 4: Diagnostic predictive values with individual RIPASA, Alvarado, and AIR scoring systems

Statistic	AIR score (>8)	Alvarado score (>7)
Sensitivity (%)	74.29%	65.71%
Specificity (%)	100%	100%
PPV	100%	100%
NPV	28%	22.58%
Diagnostic accuracy	76.62%	68.83%
PLR	---	--- -
NLR	0.257	0.343

AIR, acute inflammatory response; PPV, positive predictive value; NPV, negative predictive value; PLR, positive likelihood ratio; NLR, negative likelihood ratio

AIR scoring system, which is greater than the AUC of 0.896 for the Alvarado scoring system (Fig.1).
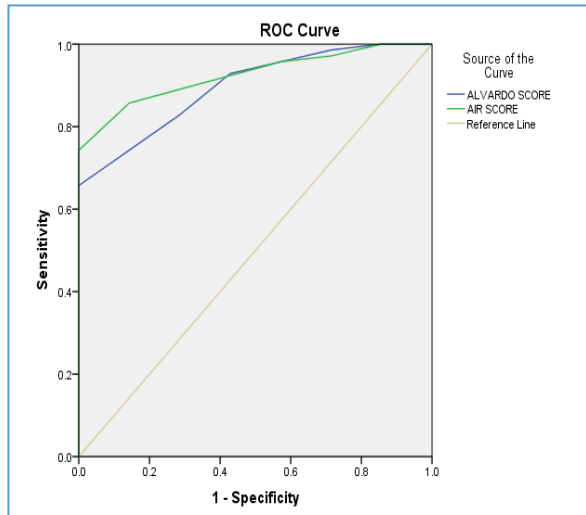
Figure 1: The receiver operating characteristic (ROC) curves demonstrate the sensitivity vs. specificity of the Alvarado and acute inflammatory response (AIR) scoring systems in the diagnoses of appendicitis

The difference in the AUCs of 11.3% between the RIPASA and the AIR scoring systems was significant (P = 0.0020), as was the difference in the AUCs of 7.48% between the RIPASA and the Alvarado scoring systems (P = 0.0026).

4. Discussion

Acute appendicitis is most common surgical emergency with 8% incidence and seen in early adult life. Over past 100 years, the morbidity and mortality rates related to this condition have markedly decreased[13]. This is because of the recognition of deleterious effects of appendiceal perforation. The diagnostic accuracy of clinical assessment of acute appendicitis varies from 50%-80%. The clinical diagnosis is especially difficult in the very young, the elderly and in the women of reproductive age group. Radiological methods such as ultrasonography and computed tomography, as well as invasive procedure like laparoscopy are all methods that have been investigated previously[14]. Many diagnostic scores have seen advocated but most are complex and difficult to implement in a clinical situation. The Alvarado score, first described in1988, is a simple scoring system. It is a scoring system that can be instituted easily in the outpatient setting and a cheap and quick tool to apply in the emergency room but it has low sensitivity in oriental population in comparison to western population[15-17]. To overcome this limitation, in 2008 a new scoring system named Acute inflammatory response (AIR) score was introduced by Andersson et al consisting of clinical and laboratory parameters. In achieving diagnostic accuracy, if surgery is delayed, there are high chances of complications like appendicular perforation and sepsis with high mortality and in contrast with reduced diagnostic accuracy rate of negative appendectomy increases which is generally reported to be approximately 20-40%[18, 19]. Our study compared the sensitivities, the specificities, the PPVs, the NPVs, the PLRs, the NLRs, and diagnostic accuracy between the Alvarado, and AIR scoring systems. The true positive rate (sensitivity) is the proportion of actual positives that are appropriately recognized; this is the percentage of sick people who are properly diagnosed as having appendicitis. The true negative rate (specificity) is the

proportion of negatives that are properly recognized; this is the percentage of healthy people who are correctly recognized as not having appendicitis[20-22]. The PPV and the NPV are the proportion of patients with positive criteria who really have the disease and the proportion of patients with negative criteria who are actually free of the disease, respectively. In the present study, the sensitivity of the AIR score with a score >8 (74.29%) was significantly better than that of Alvarado score with a score >7 (65.71%). The specificity of both the Alvarado score and AIR score were both same i. e.100%. The positive predictive value obtained was 100% for both the Alvarado score and AIR score while the negative predictive value was found to be 28% for AIR score and 22.5% for Alvarado score. Out of 77 patients who had undergone surgery 70 patients were confirmed as having appendicitis or appendicular perforation by histopathological examination, so the observed negative appendicectomy rate was 9%. This study also shows that application of AIR scoring system in the diagnosis of acute appendicitis can provide a high degree of sensitivity and thus diagnostic accuracy. In the study of Memon et al[23] in an Indian population, the sensitivity and the specificity of the Alvarado scoring system were found to be 93.5% and 80.6%, respectively. The PPV and the NPV were 92.3% and 83.3%, respectively. In the present study, the Alvarado scoring system's sensitivity and NPV were less than those mentioned above while the specificity and the PPV were more. Based on these results, the Alvarado score can be used effectively to reduce the incidence of negative appendectomies. The area under the ROC curve with the AIR scoring systems was significantly larger than it was with the Alvarado scoring systems. The AIR score is fast and perfect in categorizing patients with suspected appendicitis and reduces the need for diagnostic imaging. Overall, a higher sensitivity, NPV, and equal PPV indicate that the AIR score is a much better diagnostic tool than the Alvarado score for the diagnosis of AA. Thus, upon seeing patients with right iliac fossa pain, the operating surgeon can make a quick decision by using the AIR scoring system, with a score >8 indicating a need for surgery.

5. Conclusion

Both the Alvarado and AIR scoring systems are effective, simple, and reliable tools for diagnosing acute appendicitis. However, the AIR score demonstrates higher sensitivity and diagnostic accuracy compared to the Alvarado score, making it a superior modality for clinical diagnosis in the studied population.

References

[1] Meer M. Chisthi, Anilkumar Surendran, Jiju Therumpurathu Narayanan. RIPASA and air scoring systems are superior to alvarado scoring in acute appendicitis: Diagnostic accuracy study. Annals of Medicine and Surgery 59 (2020); 138–142.

[2] 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.

[3] Mohammad Yasin Karami, Hadi Niakan1, Navid Zadebagheri, 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.

[4] Henna E Sammalkorpi*, Panu Mentula and Ari Leppäniemi. A new adult appendicitis score improves diagnostic accuracy of acute appendicitis – a prospective study. BMC Gastroenterology 2014; 14: 114.

[5] Paola Fugazzola, Marco Ceresoli, Vanni Agnoletti, et al. The SIFIPAC/WSES/SICG/SIMEU guidelines for diagnosis and treatment of acute appendicitis in the elderly (2019 edition). World Journal of Emergency Surgery (2020); 15: 19.

[6] Emmanuel S Kanumba, Joseph B Mabula, Peter Rambau, Phillipo L Chalya. Modified Alvarado Scoring System as a diagnostic tool for Acute Appendicitis at Bugando Medical Centre, Mwanza, Tanzania, Kanumba et al. BMC Surgery 2011; 11: 4.

[7] Bassant Sayed Moussa, Mohamed Amin Ali, Dina Abdu lRahman Ramadan Mohamed. Comparing the diagnostic accuracy of modified RIPASA and MASS in patients diagnosed with acute appendicitis in Suez Canal University Hospital Emergency Department: a cross-sectional study. BMC Emergency Medicine (2022); 22: 142.

[8] Michael D. Repplinger, MD, MS, Joseph F. Levy, BS, Erica Peethumnongsin, et al. Systematic Review and Meta-Analysis of the Accuracy of MRI to Diagnose Appendicitis in the General Population. J Magn Reson Imaging.2016 June; 43 (6): 1346–1354.

[9] van Randen A, Laméris W, van Es HW, et al. A comparison of the accuracy of ultrasound and computed tomography in common diagnoses causing acute abdominal pain. Eur Radiol 2011; 21 (7): 1535–1545.

[10] Winson Jianhong Tan, Sanchalika Acharyya. Randomized control trial comparing an Alvarado Score-based management algorithm and current best practice in the evaluation of suspected appendicitis. World Journal of Emergency Surgery (2020); 15: 30.

[11] Josephine Reismann, Alessandro Romualdi, Natalie Kiss et al. Diagnosis and classification of pediatric acute appendicitis by artificial intelligence methods: An investigator-independent approach. PLOS ONE (2019) https //doi. org/10.1371/ journal. pone.0222030.

[12] Kwan KY, Nager AL. Diagnosing pediatric appendicitis: usefulness of laboratory markers. Am J Emerg Med.2010; 28: 1009–1015.

[13] N N, Mohammed A, Shanbhag V, Ashfaque K, S A P. A Comparative Study of RIPASA Score and ALVARADO Score in the diagnosis of acute appendicitis. J Clin Diagn Res 2014; 8: NC03-5.

[14] Birnbaum B, Jeffrey RB. CT and Sonographic evaluation of acute right lower quadrant abdominal pain. *Am J Radiol* 1998; 170: 361-371.

[15] Owen TD, Williams H, Stiff G, Jenkinson LR, Rees BI. Evaluation of the Alvarado score in acute appendicitis. J R Soc Med 1992; 85: 87-8.

[16] Jang SO, Kim BS, Moon DJ. Application of alvarado score in patients with suspected appendicitis. Korean J Gastroenterol 2008; 52: 27-31.

[17] Khan I, ur Rehman A. Application of alvarado scoring system in diagnosis of acute appendicitis. J Ayub Med Coll Abbottabad 2005; 17: 41-4.

[18] 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.

[19] de Castro SM, 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: 1540-5.

[20] Altman DG, Bland JM. Diagnostic tests.1: sensitivity and specificity. BMJ 1994; 308: 1552.

[21] Sedighi I. Interpretation of diagnostic tests: likelihood ratio vs. predictive value. Iran J Pediatr 2013; 23: 717.

[22] Kharbanda AB, Monuteaux MC, Bachur RG, Dudley NC, Bajaj L, Stevenson MD, et al. A clinical score to predict appendicitis in older male children. Acad Pediatr 2017; 17: 261-6.

[23] Memon ZA, Irfan S, Fatima K, Iqbal MS, Sami W. Acute appendicitis: diagnostic accuracy of Alvarado scoring system. Asian J Surg2013; 36: 144-9.