International Journal of Life Sciences, Biotechnology and Pharma Research Vol. 13, No. 12, December 2024 Online ISSN: 2250-3137
Print ISSN: 2977-0122

DOI: 10.69605/ijlbpr_13.12.2024.32

ORIGINAL RESEARCH

To assess the sensitivity and specificity of Appendicitis Inflammatory Response (AIR) scores in suspected appendicitis cases by comparing them with histopathological reports

Dr. Deepak[1], Dr. Subodh Gupta[2], Dr. R N Mittal[3]

[1]Junior Resident-III, [2]Professor & HOD, [3]Professor, Department of General & Minimal Access Surgery, Jaipur Golden Hospital, New Delhi, India

Corresponding author
Dr. Deepak
Junior Resident-III, Department of General & Minimal Access Surgery, Jaipur Golden Hospital, New Delhi, India
Email: yadavdeepak529@gmail.com

Received Date: 22 October, 2024 Accepted Date: 25 November, 2024

ABSTRACT

Appendicitis is a disorder that causes inflammation of the appendix, a little finger-shaped organ found in the lower right abdomen. While appendicitis appears to be a somewhat common medical issue, the complexities of the condition, as well as the difficulties in diagnosing it, have sparked significant medical attention. **Aim:** The aim of the present is to evaluate the Appendicitis Inflammatory Response (AIR) score in the case of suspected appendicitis and compare the scores with histopathological report to find out the sensitivity and specificity of the scoring system. **Materials & methods:** Researchers from Jaipur Golden Hospital's General and Minimal Invasive Surgery Department set out to determine the efficacy of the AIR Score for the diagnosis of acute appendicitis. Ultrasonography, laboratory investigations, and clinical examination were used to examine 110 male and female patients in the study. The pathologists analyzed samples retrieved from the appendix after patients with verified acute appendicitis underwent either an open or laparoscopic appendectomy. Approval from the IEC was granted to the investigation. **Results:** Among the 110 patients, 96 patients confirmed to have appendicitis through HPE, the AIR score classified 91 (94.79 %) as high-risk subjects were confirmed to have appendicitis by HPE. We observed a statistically significant difference between the AIR score, and the HPE diagnosis of appendicitis. The sensitivity of the AIR score was 95 % and the specificity of The AIR score was 82 %. Correspondingly, both false-positive was 21 % and false-negative was 5 % rate. The positive and negative predictive values of the AIR score were 97 % and 69 %. **Conclusion:** The AIR score enhances diagnostic capabilities for urgent surgical intervention, with superior validity and reliability, but further research is needed to understand variations within specific population groups.

Keywords: Appendicitis Inflammatory Response score, Nausea, Vomiting, White blood cells, Reliability, Accuracy.

INTRODUCTION

The human body is a marvel of intricacy, with various organs and systems functioning in tandem to sustain life. However, despite its great resistance, it is susceptible to a variety of diseases and disorders. Appendicitis is one such condition that has perplexed physicians for decades. Appendicitis is a disorder that causes inflammation of the appendix, a little finger-shaped organ found in the lower right abdomen. While appendicitis appears to be a somewhat common medical issue, the complexities of the condition, as well as the difficulties in diagnosing it, have sparked significant medical attention.

Appendicitis is notorious for its harmful qualities, such as its rapid progression and onset [1]. Delays in diagnosing the condition, resulting in a delay in early surgical surgery, can be devastating and contribute to the sick person's chronically high death rate [2]. These infections can be highly serious and fatal if not treated within 48 hours, and they have a high mortality rate; thus, early detection is critical and priority [3-5]. To avoid errors in managing these patients, clinicians must always be available. The appendicitis

DOI: 10.69605/ijlbpr_13.12.2024.32

inflammatory response (AIR) was initially reported by researchers in 2008 [6]. We identified and extensively analyzed multiple clinical studies [8-10] published after 2008 to establish the validity, reliability, and application of each piece of data. We developed these practice suggestions based on a summative study of the data. Diagnosing appendicitis can be difficult due to its varying presentation and the potential for serious complications if misdiagnosed or left untreated. The current study is to evaluate the most effective ways and procedures used by healthcare providers to accurately diagnose appendicitis and provide prompt treatment. Clinicians have developed a variety of scoring systems to aid in the diagnosis process, with the AIR score acting as a prominent example in practice [1-4].

In this era of technological growth, our goal is to use machine learning techniques to improve the prediction capabilities of both scoring systems [11]. To improve the accuracy of acute appendicitis diagnosis, machine learning approaches could be combined with other methodologies [1-7]. This would lower the number of false positive and negative outcomes. As a result, the current study aims to assess the Appendicitis Inflammatory Response (AIR) score in cases of suspected appendicitis and compare it to the histopathological report to ascertain the scoring system's sensitivity and specificity.

MATERIALS & METHODS

Following approval from the IEC, we set out to conduct the current study. This research used an observational prospective design. The research was carried out by the General and Minimal Invasive Surgery Department at Jaipur Golden Hospital. From 2022 to 2024, data was gathered for the project. The participants in the study were all those who, after a thorough evaluation, received an appendectomy after visiting the emergency room for symptoms consistent with acute appendicitis. All patients admitted for general surgery with a suspicion of acute appendicitis during the study period were included in the study, provided they gave their consent. Patients did not qualify for inclusion in the study if they met any of the following criteria: they were pregnant, had a history of abdominal cancer, refused surgery or received conservative treatment, had an appendectomy (incidental or interval), had a lump in the appendix, or did not give their informed consent. By following the rules laid out for inclusion and exclusion, the study "Comparison of the AIR Score and Alvarado Score in Diagnosis of Acute Appendicitis" comprised 110 patients of both sexes. The surgical specialists at Jaipur Golden Hospital arrived at the conclusion. All eligible people were evaluated using ultrasonography, necessary laboratory investigations, and a thorough clinical examination as part of the study approach. For every individual, we calculated their AIR score [13]. Each subject had abdominal ultrasonography. Patients with proven acute appendicitis were sent to the operating room for an open or laparoscopic appendectomy according to the institute's approved protocol. Pathologists examined the tissues extracted from the appendix.

Each case's AIR score is determined by weighing the clinical and laboratory features noted at the time of patient admission. Regurgitation, pain in the right iliac fossa, stomachache, fever, white blood cell count, white blood cell count, and C-reactive protein levels are some of the characteristics that can be considered.

Table 1: Scoring scheme for AIR score

Variable	Appendicitis Inflammatory Response (AIR) score
Anorexia	
Nausea or vomiting	
Vomiting	*1*
Migration of pain to right lower quadrant (RLQ)	
RLQpain	
Pain in the right lower quadrant	*1*
Rebound tenderness	*Light-1; Medium – 2; Strong – 3*
Elevated temperature (> 37.3)	
White blood cells (WBC) (> 10000 cells/cu. mm)	
WBC (10000 – 14999 cells/cu. mm)	*1*
WBC(> 15000 cells/cu. mm)	*2*
Leukocytosis shift	
C-reactive protein (10 – 49 mg/l)	*1*
C-reactive protein (≥ 50 mg/l)	*2*
Polymorphonuclear leucocytes (70% -84 %)	*1*
Polymorphonuclear leucocytes (≥ 85%)	*2*
Total score	***12***

The diagnosis of acute appendicitis is considered to be consistent with an AIR score sums ranging from 0 to 4 are indicative of a low probability, whereas sums ranging from 5 to 8 suggest a moderate probability, and sums ranging from 9 to 12 indicate a high possibility. The study will involve the examination of several factors related to the patients, including their

DOI: 10.69605/ijlbpr_13.12.2024.32

features, clinical presentations, underlying diseases, and hospital regimens.

Statistical analysis

This research evaluated the correlation between HPE findings and quantitative and categorical variables, divided patients into high-risk and low-risk groups, and provided a summary of these factors. By comparing AIR scores to the results of histological examinations, we were able to determine their predictive validity, sensitivity, specificity, predictive values, and diagnostic accuracy through the use of ROC analysis.

RESULTS

The table 2 presents the association risk scores and histopathological examination (HPE) results in a study sample. For the AIR Score, a high score was found in 91 positive and 3 negative cases, and a low score in 5 positive and 11 negative cases. The chi-square value for the AIR Score is 47.17 with a p-value of 0.0001. These findings suggest that AIR scores is significantly associated with the HPE results, indicating their potential effectiveness in predicting positive histopathological outcomes.

Table 2: ROC analysis variables of risk scores in study sample

Risk Score	Area under the curve (AUC)	Asymptomatic 95 % Confidence Interval		P Value
		Lower Bound	Upper Bound	
AIR Score	0.89	0.74	0.98	< 0.05

The Area Under the Curve (AUC) for the AIR Score is 0.89, with a 95% confidence interval ranging from 0.74 to 0.98, and a p-value of less than 0.05. This reflects a statistically significant and high discriminative ability of the AIR Score. Overall, the AIR Score shows superior predictive performance.

Table 3: Sensitivity and False Positive Rate of AIR score

Variable	AIR Score (95 % CI)
Sensitivity	95 %
Specificity	82 %
False positive rate	21%
False negative rate	5%
Positive Predictive Value	97%
Negative Predictive Value	69%

Table 3 presents the Appendicitis Inflammatory Response (AIR) Score in diagnosing appendicitis. The AIR Score demonstrates higher sensitivity (95% vs. 85%) and specificity (82% vs. 57%) compared to the AS, indicating that the AIR Score is more accurate in correctly identifying both positive and negative cases. Additionally, the AIR Score has a lower false positive rate (21% vs. 43%) and false negative rate (5% vs. 15%), making it more reliable in minimizing diagnostic errors. The Positive Predictive Value (PPV) is slightly higher for the AIR Score (97% vs. 93%), suggesting better accuracy in predicting true positives. The Negative Predictive Value (NPV) is significantly higher for the AIR Score (69% vs. 40%), indicating a greater likelihood of correctly identifying true negatives.

Table 4: Reliability of the two risk scores in study sample

Risk scores	Kappa Statistic	Standard error	P value
AIR Score	0.69	0.24	< 0.05

Table 4 shows the reliability of the AIR Score using the Kappa statistic, standard error, and P value. The Kappa statistic measures agreement between the risk scores and the HPE observations beyond chance, with values closer to 1 indicating higher agreement. The AIR Score shows a higher Kappa statistic (0.69) indicating that the AIR Score has substantial agreement with HPE observations. The scores have P values less than 0.05, indicating that the observed agreements are statistically significant and not due to chance. The standard errors are relatively similar (0.24 for AIR), suggesting consistent reliability in the measurement. Overall, the AIR Score is more reliable and consistent in predicting outcomes.

DISCUSSION

Due to limited availability of advanced diagnostic techniques such as CT scans in resource-constrained regions, risk stratification scores serve as valuable tools for facilitating the diagnosis of acute appendicitis [12]. Nevertheless, the utilization of these methods in clinical practice has been limited due to issues about their inadequate validity and reliability, which subsequently lead to unfavorable rates of appendectomy [13]. The ease

DOI: 10.69605/ijlbpr_13.12.2024.32

of use, ability to predict clinical outcomes, and capacity to minimize unnecessary inquiry and treatment are important factors to consider while designing a clinical score. Consequently, a number of scoring systems have been developed to predict the occurrence of acute appendicitis. The primary objective of these scores is to decrease dependence on imaging and minimize the occurrence of negative appendectomies, while also maintaining perforation rates within an acceptable range.

In light of the emergence of numerous novel scoring systems purporting to surpass existing scores in terms of effectiveness, it is crucial to subject these claims to rigorous examination across various subgroups within the population prior to endorsing their implementation as standard practice [14-16]. The present study conducted a comparative analysis of the validity and reliability of the recently developed AIR score. The present study evaluated the predictive validity of the AIR score using the AUC metric. The AIR score demonstrated a higher AUC value of 0.89 (95% CI 0.74–0.98). Several studies have conducted comparisons between the two scores, and the findings consistently indicate that the AIR score has more accuracy across various patient demographics [6,7,9,17]. In the initial prospective study conducted by Andersson and Andersson, a total of 545 patients were included. The AUC values were reported as 0.93 for AIR [7]. The cross-sectional observational study conducted by Jose et al. involved a sample size of 130 patients. The study found that the AUC for AIR was 0.90 [17]. De Castro et al. [9] observed that the AUC for AIR was 0.96. There has been a prevailing belief that a decrease in the percentage of negative appendectomies is accompanied with an increase in perforation rates, leading to elevated morbidity because of delayed identification or treatment. The present study sensitivity of the AIR score was found to be 95 % and the specificity of the AIR score was found to be 82%, In a similar vein, it is noteworthy that the AIR score exhibited positive and negative predictive values of 97% and 69%, respectively. A study carried out by researchers [18]. According to previous research conducted by authors [6] and [7], a significant proportion of patients, namely sixty-three percent, were accurately categorized into either the low- or high-probability groups, achieving a high level of precision at 97.2%. Nevertheless, several studies have raised concerns over the validity of these scores, particularly when used to young patients. In a multicentric prospective study [10,11], it was shown that among a group of 661 children, the AIR exhibited higher sensitivity and specificity, but only within the low-risk subgroup. More specifically, the AIR exhibited a specificity of 90% and a sensitivity of 95%.

Among the 110 patients, 96 patients confirmed to have appendicitis through HPE, the AIR score classified 91 (94.79 %) as high-risk subjects were confirmed to have appendicitis by HPE. We observed a statistically significant difference between the AIR score, and the HPE diagnosis of appendicitis. The AIR score exhibited superior diagnostic performance in challenging patient populations, including the elderly, infants, and women, when assessing appendicitis, as per a study conducted by [18]. In comparison to the AIR score, a distinct investigation conducted by [11] demonstrated substantially higher specificity (97%) and positive predictive value (88%).

The kappa statistic was used to assess the reliability of the risk scores. The AIR score demonstrated considerably greater reliability (0.842). According to a study [9], the AIR score demonstrated an AUC of 0.96. In a separate investigation, the authors of a previous study [18] documented an AUC of 0.90 for the AIR score. Previous research studies [6,7] that have introduced the AIR score have documented an AUC of 0.97 for advanced appendicitis and 0.93 for all cases of appendicitis.

The limitations of the study encompass several factors, including the relatively limited sample size of patients, the cross-sectional and observational design of the study, and the restriction to just those patients who underwent appendicectomy. Additional research that is more comprehensive and forward-looking, focusing on specific demographics with greater detail, is necessary. This would facilitate the determination of ideal threshold values for specific scores and demographics, hence enabling the utilization of targeted imaging approaches through risk stratification.

CONCLUSION

The study shows that the AIR score has improved diagnostic capabilities indicating potential for urgent surgical intervention. Its superior validity and reliability make it a more effective tool for treatment decisions. However, further research is needed to understand variations in validity and reliability within specific population groups, such as females, pediatrics, obese individuals, and the elderly.

Conflict of interest

There is no conflict of interest among the present study authors.

REFERENCES

1. Di Saverio S, Podda M, De Simone B, Ceresoli M, Augustin G, Gori A, Boermeester M, Sartelli M, Coccolini F, Tarasconi A, De'Angelis N. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. World journal of emergency surgery. 2020 Dec;15(1):1-42.
2. Pouli S, Kozana A, Papakitsou I, Daskalogiannaki M, Raissaki M. Gastrointestinal perforation: clinical and MDCT clues for identification of aetiology. Insights into Imaging. 2020 Dec;11(1):1-9.

DOI: 10.69605/ijlbpr_13.12.2024.32

3. Smith S, Schreiber MA. Evaluating the Acute Abdomen. In Emergency General Surgery 2019 (pp. 13-26). Springer, Cham.
4. Bayne JE, Edmondson MA. Diseases of the gastrointestinal system. Sheep, goat, and cervid medicine. 2021:63.
5. Gossage JA, Burnand KM. The abdomen. In Browse's Introduction to the Symptoms & Signs of Surgical Disease 2021 May 3 (pp. 491-547). CRC Press.
6. Andersson M, Andersson RE. The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. World journal of surgery. 2008 Aug;32(8):1843-9.
7. Andersson M, Kolodziej B, Andersson RE, Andersson RE, Andersson M, Eriksson T, Ramsing A, Westman L, Björkman J, Håkansson HO, Lundström T. Randomized clinical trial of Appendicitis Inflammatory Response score-based management of patients with suspected appendicitis. Journal of British Surgery. 2017 Oct;104(11):1451-61.
8. Vaziri M, Nafissi N, Jahangiri F, Nasiri M. Comparison of the appendicitis inflammatory response and Alvarado scoring systems in the diagnosis of acute appendicitis in children. Journal of Medicine and Life. 2021 Jan;14(1):75.
9. 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 journal of surgery. 2012 Jul;36:1540-5.
10. Vaziri M, Nafissi N, Jahangiri F, Nasiri M. Comparison of the appendicitis inflammatory response and Alvarado scoring systems in the diagnosis of acute appendicitis in children. Journal of Medicine and Life. 2021 Jan;14(1):75.
11. 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 journal of surgery. 2015 Jan;39:104-9.
12. Bom WJ, Scheijmans JC, Salminen P, Boermeester MA. Diagnosis of uncomplicated and complicated appendicitis in adults. Scandinavian Journal of Surgery. 2021 Jun;110(2):170-9.
13. Korkut M, Bedel C, Karancı Y, Avcı A, Duyan M. Accuracy of Alvarado, Eskelinen, Ohmann, RIPASA and Tzanakis scores in diagnosis of acute appendicitis; a cross-sectional study. Archives of academic emergency medicine. 2020;8(1).
14. Chisthi MM, Surendran A, Narayanan JT. RIPASA and air scoring systems are superior to alvarado scoring in acute appendicitis: Diagnostic accuracy study. Annals of Medicine and Surgery. 2020 Nov 1;59:138-42.
15. Gupta S, Kolli VS, Da Costa K, Javed S, Ammar A, Rasheed A. A systematic review and meta-regression for validation of the Alvarado score as a tool for predicting acute appendicitis. Annals of Medicine and Surgery. 2023 Feb;85(2):111.
16. Favara G, Maugeri A, Barchitta M, Ventura A, Basile G, Agodi A. Comparison of RIPASA and ALVARADO scores for risk assessment of acute appendicitis: A systematic review and meta-analysis. Plos one. 2022 Sep 30;17(9):e0275427.
17. Jose T, Rajesh PS. Appendicitis inflammatory response score in comparison to Alvarado score in acute appendicitis. The Surgery Journal. 2021 Jul 19;7(03):e127-31.
18. Macco S, Vrouenraets BC, de Castro SM. Evaluation of scoring systems in predicting acute appendicitis in children. Surgery. 2016 Dec 1;160(6):1599-604.