LINK MEDICAL INTERFACE

JHWCR

Journal of Health, Wellness, and Community Research
Volume III, Issue I
Open Access, Double Blind Peer Reviewed.
Web: https://jhwcr.com, **ISSN**: 3007-0570
DOI: https://doi.org/10.61919/pqmgcv78

Article

Comparative Analysis of Appendicitis Inflammatory Response Score Versus ALVARADO Score as a Diagnostic Tool for Acute Appendicitis

Zahoor Muhi-ud-Deen[1], Muhammad Shoaib[1], Hafiz Sohail Akram[1], Hira Khalid[2], Muhammad Imran[1], Najam ul Hassan[1]

1 Combined Military Hospital (CMH), Rawalpindi, Pakistan
2 Radiology Department, Combined Military Hospital (CMH), Rawalpindi, Pakistan

Correspondence:
geniuszmd@gmail.com

Cite this Article

Received	2025-02-13
Revised	2025-03-11
Accepted	2025-03-18
Published	2025-03-27
Authors' Contributions	ZMD, MS, HSA: study concept and design; MI, NUH: data collection and interpretation; HK: radiological assessment and analysis; ZMD, MS, HK: manuscript drafting and critical revision.
Conflict of Interest	None declared
Data/supplements	Available on request.
Funding	None
Ethical Approval	Respective Ethical Review Board
Informed Consent	Obtained from all participants
Study Registration	-
Acknowledgments	N/A

© 2025 by the Authors. This is an Open Access double blind peer reviewed publication licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

ABSTRACT

Background: Acute appendicitis remains a prevalent surgical emergency, yet its diagnosis is often complicated by overlapping symptoms and limited access to imaging in resource-constrained settings. While the Alvarado score is widely utilized, it is limited by subjective criteria; the Appendicitis Inflammatory Response (AIR) score incorporates objective inflammatory markers such as CRP, but comparative data from local clinical settings are scarce. **Objective**: This study aimed to compare the diagnostic accuracy of the AIR score versus the Alvarado score in predicting acute appendicitis, using histopathological findings as the gold standard. **Methods**: This prospective observational cohort study was conducted at Combined Military Hospital (CMH) Rawalpindi over six months, including 185 patients with clinical suspicion of acute appendicitis who underwent surgery. Patients under 6 years, pregnant, or with abdominal trauma or malignancy were excluded. Alvarado and AIR scores were computed preoperatively, blinded to surgical decision-making. Histopathology confirmed diagnoses. Ethical approval was obtained from the CMH Ethical Review Committee per the Declaration of Helsinki. Data were analyzed using SPSS v27, with diagnostic accuracy, sensitivity, specificity, PPV, and NPV calculated. **Results**: Of 185 patients (mean age: 28.21 ± 12.76 years), histopathology confirmed appendicitis in 162 (87.6%). The AIR score (cut-off ≥6) demonstrated higher accuracy (89.19%) compared to the Alvarado score (cut-off ≥7, accuracy: 69.73%), with superior sensitivity (90.12% vs. 70.37%), specificity (82.61% vs. 65.22%), PPV (97.33% vs. 93.44%), and NPV (54.29% vs. 23.81%). **Conclusion**: The AIR score outperformed the Alvarado score in diagnostic precision, offering a more objective and reliable tool for acute appendicitis diagnosis. Its integration into clinical protocols may reduce unnecessary imaging and surgeries, enhancing decision-making in emergency and low-resource settings.

Keywords: Acute Appendicitis, Alvarado Score, Appendicitis Inflammatory Response Score, Diagnostic Accuracy, Histopathology, Clinical Scoring System, Appendectomy

INTRODUCTION

Acute appendicitis is one of the most frequently encountered surgical emergencies worldwide. The appendix, a blind-ended tubular structure at the ileocecal junction, resembles a worm in shape, hence termed the vermiform appendix. Acute appendicitis refers to the acute inflammation of the inner lining of the appendix (1). The condition has a lifetime prevalence of 8.6% in males and 6.7% in females, with the highest incidence observed in the second and third decades of life (2). Despite being a common clinical presentation, the diagnosis of acute appendicitis remains challenging due to its variable clinical manifestations. Since the first successful appendectomy performed by Tait in 1880 in England (3), advances in diagnostic tools have evolved, yet achieving diagnostic accuracy continues to be problematic even after more than a century. The high negative appendectomy rate—up to 17.5% in patients suspected of having acute appendicitis—underscores the difficulty in establishing a definitive diagnosis before surgical intervention (4).

This diagnostic uncertainty has significant clinical and economic implications. On one hand, unnecessary surgeries increase hospital stay, patient discomfort, and burden on limited healthcare resources; on the other hand, delayed or missed diagnosis raises the risk of complications such as perforation, peritonitis, and increased morbidity and mortality (5). Imaging modalities such as ultrasound and computed tomography (CT)

have been recommended to aid diagnosis; however, in resource-constrained settings like Pakistan, these investigations are not always feasible due to financial limitations, time constraints, and dependence on operator expertise (6). Given these challenges, clinical scoring systems have been developed to enhance diagnostic accuracy and reduce reliance on imaging.

Among these, the Alvarado score, introduced in 1986, has gained widespread use. It incorporates common clinical features and laboratory findings—including differential leukocyte count—to assess the likelihood of acute appendicitis (7)(8). However, the Alvarado score does not consider biomarkers of inflammation such as C-reactive protein (CRP), which may limit its diagnostic precision (9). In contrast, the Appendicitis Inflammatory Response (AIR) score includes CRP, and studies have shown it to have a higher predictive value for diagnosing acute appendicitis. Several comparative studies globally have evaluated the diagnostic utility of these two scores across different populations and clinical settings, with histopathological examination currently regarded as the gold standard for confirmation (10).

Despite international data supporting the value of both scoring systems, there is a paucity of local evidence comparing their diagnostic performance in the Pakistani population. This gap highlights the need for region-specific research to validate the applicability of these tools in our healthcare context. Therefore, the present study aims to compare the predictive value of the Alvarado score and AIR score against histopathological findings in patients presenting with suspected acute appendicitis. The central hypothesis is that the AIR score, owing to the inclusion of CRP, will demonstrate higher diagnostic accuracy compared to the Alvarado score in our study population.

MATERIAL AND METHODS:

This prospective observational cohort study was conducted at Combined Military Hospital (CMH) Rawalpindi over a six-month period from 1st July 2022 to 31st December 2022, following approval from the hospital's Ethical Review Committee. The primary objective was to compare the predictive value of the Alvarado and Appendicitis Inflammatory Response (AIR) scores against histopathological findings, the gold standard for diagnosis. A sample size of 185 was calculated using a 95% confidence interval, based on an estimated 14% incidence of acute appendicitis in the population (8). Participants were recruited from patients presenting to the emergency department with clinical features suggestive of acute appendicitis, including migratory right lower quadrant abdominal pain, nausea, vomiting, fever, and localized tenderness. Inclusion criteria encompassed all patients with a high clinical suspicion of acute appendicitis, while exclusion criteria included children under six years of age, individuals with abdominal trauma, known malignancies, pregnancy, history of urolithiasis, and those unwilling to provide informed consent.

All patients meeting the inclusion criteria underwent thorough clinical and laboratory evaluations necessary to compute both the Alvarado and AIR scores. The Alvarado score incorporated parameters such as migratory pain, anorexia, nausea/vomiting, right lower quadrant tenderness, rebound pain, elevated temperature, leukocytosis, and neutrophilia, while the AIR score included additional components such as C-reactive protein (CRP), offering an inflammatory biomarker-based approach. A threshold of ≥7 for Alvarado and ≥6 for AIR was used to classify patients as having a high probability of acute appendicitis. These scores were calculated solely for research purposes and were not used to guide clinical decisions. The on-duty surgical consultant, who determined the patient's management, remained blinded to the scores to prevent bias in clinical decision-making. Final diagnosis was confirmed by histopathological examination of the resected appendix, with acute appendicitis defined by neutrophilic infiltration into the muscularis propria.

The primary outcome of the study was the diagnostic accuracy of the Alvarado and AIR scores as compared to histopathological confirmation. Secondary outcomes included the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of both scoring systems. All data were collected using structured proformas, and no imaging modalities were mandated as part of the study protocol to ensure real-world applicability, especially in resource-constrained settings. Missing data were minimized through direct clinical observation and immediate documentation; however, any incomplete entries were excluded from final analysis. Potential confounding variables, such as age and sex, were recorded and considered during statistical evaluation.

This study adhered to the ethical principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the CMH Rawalpindi Ethical Review Committee (approval number not applicable/pending). Written informed consent was obtained from all participants or their legal guardians prior to inclusion in the study, and confidentiality of participant data was maintained by assigning unique identifiers and restricting access to the dataset. All statistical analyses were performed using IBM SPSS version 27. Descriptive statistics were used to summarize demographic and clinical characteristics, while inferential analysis included chi-square tests for categorical variables and independent sample t-tests for continuous variables, where appropriate. Receiver operating characteristic (ROC) curves were plotted to assess the predictive value of both scores, and area under the curve (AUC) was compared. A p-value of less than 0.05 was considered statistically significant.

RESULTS:

A total of 185 patients who underwent surgery for clinical suspicion of acute appendicitis were included in this study. The mean age of participants was 28.21 ± 12.76 years. Of these, 108 (58.4%) were males and 77 (41.6%) were females, indicating a male predominance. Surgical approaches included open appendectomy in 58 patients (31.4%), laparoscopic appendectomy in 120 patients (64.9%), while 7 cases (3.8%) required conversion from laparoscopic to open procedure due to intraoperative complications or unclear anatomy.

The mean Alvarado score across the cohort was 5.97 ± 1.31, while the mean AIR score was 7.04 ± 2.06. Histopathological

 © Authors. CC BY 4.0. **DOI**: https://doi.org/10.61919/pqmgcv78

confirmation of acute appendicitis was found in 162 cases (87.6%), whereas 21 patients (11.4%) had a normal appendix, and 2 patients (1.1%) demonstrated other pathological findings. These results confirmed a high overall diagnostic accuracy of 87.6% based on histopathological analysis.

Table 1. Baseline and Diagnostic Characteristics of the Study Population (N=185)

Variable	Value
Total Patients	185
Gender – Male, n (%)	108 (58.4%)
Gender – Female, n (%)	77 (41.6%)
Age (years), Mean ± SD	28.21 ± 12.76
Alvarado Score, Mean ± SD	5.97 ± 1.31
AIR Score, Mean ± SD	7.04 ± 2.06
Surgical Procedure, n (%)	
– Open	58 (31.4%)
– Laparoscopic	120 (64.9%)
– Laparoscopic converted to open	7 (3.8%)
Histopathological Diagnosis, n (%)	
– Acute Appendicitis	162 (87.6%)
– Normal Appendix	21 (11.4%)
– Miscellaneous Pathologies	2 (1.1%)

Diagnostic performance of both Alvarado and AIR scores was evaluated using histopathology as the gold standard. For the Alvarado score (cut-off ≥7), sensitivity was 70.37%, specificity 65.22%, positive predictive value (PPV) 93.44%, and negative predictive value (NPV) 23.81%, resulting in an overall diagnostic accuracy of 69.73%. This indicates that although the score has high PPV, it lacks in NPV, thereby risking underdiagnosis in borderline or atypical presentations.

Table 2. Diagnostic Performance of Alvarado Score (Cut-off ≥7)

	Histopathology Positive	Histopathology Negative
Alvarado Score Positive	114	8
Alvarado Score Negative	48	15
Sensitivity	70.37%	
Specificity		65.22%
Positive Predictive Value (PPV)	93.44%	
Negative Predictive Value (NPV)	23.81%	
Overall Accuracy		69.73%

Conversely, the AIR score (cut-off ≥6) demonstrated significantly superior diagnostic characteristics. Sensitivity was 90.12%, specificity 82.61%, PPV 97.33%, and NPV 54.29%, yielding a diagnostic accuracy of 89.19%. These findings suggest that the AIR score is more robust in ruling in and ruling out acute appendicitis in suspected cases, largely due to the inclusion of C-reactive protein (CRP) as an inflammatory marker.

Table 3. Diagnostic Performance of AIR Score (Cut-off ≥6)

	Histopathology Positive	Histopathology Negative
AIR Score Positive	146	4
AIR Score Negative	16	19
Sensitivity	90.12%	
Specificity		82.61%
Positive Predictive Value (PPV)	97.33%	
Negative Predictive Value (NPV)	54.29%	
Overall Accuracy		89.19%

Clinically, the AIR score appears to outperform the Alvarado score in diagnostic precision. Its higher sensitivity and specificity suggest it can more accurately identify true positive and true negative cases, which is critical in emergency settings where diagnostic uncertainty may lead to either unnecessary surgery or delayed intervention. Notably, while the Alvarado score showed excellent PPV (93.44%), it had poor NPV (23.81%), indicating that a low score should not confidently exclude the diagnosis. In contrast, the AIR score, with a PPV of 97.33% and NPV of 54.29%, not only supports positive cases more reliably but also shows greater reliability in excluding disease.

These results underscore the clinical utility of the AIR score in resource-limited and high-volume emergency departments, especially where imaging modalities are unavailable or delayed. The inclusion of CRP, a marker of systemic inflammation, may explain the improved discriminatory power of the AIR score. Although both scoring systems have their roles in clinical

 © Authors. CC BY 4.0. DOI: https://doi.org/10.61919/pqmgcv78

decision-making, the AIR score may serve as a more reliable tool in guiding early management of suspected acute appendicitis.

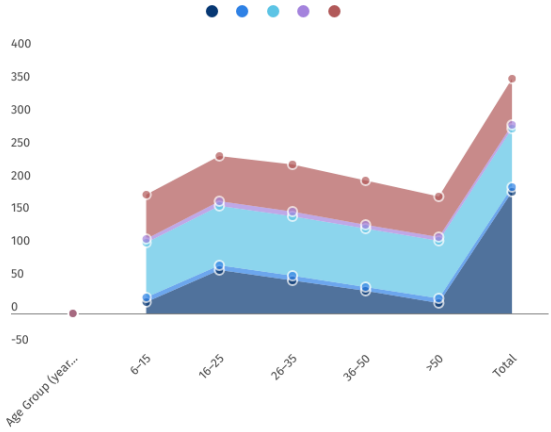

Figure 1 Diagnostic Accuracy of AIR and Alvarado Scores Stratified by Age Groups

DISCUSSION:

Despite being one of the most commonly encountered surgical emergencies, the diagnosis of acute appendicitis continues to pose a clinical challenge due to its overlap with a variety of other abdominal conditions that mimic its presentation, such as gastroenteritis, urinary tract infections, and gynecological disorders in women (11,12). Early and accurate diagnosis is critical, as delayed intervention is associated with an increased risk of complications including perforation, abscess formation, prolonged hospital stays, exposure to general anesthesia, and a higher burden on healthcare systems (13). Although radiological advancements such as ultrasonography and computed tomography (CT) have enhanced diagnostic accuracy, their routine use is limited by factors such as cost, radiation exposure, operator dependency, and delayed availability in resource-constrained settings (14). These limitations underscore the importance of clinical scoring systems that are fast, cost-effective, and based on readily accessible clinical and laboratory parameters.

Among these, the Alvarado score has been extensively used worldwide due to its simplicity and applicability in diverse clinical environments. However, its reliance on subjective symptoms such as nausea, vomiting, and anorexia, which are nonspecific and variably reported, especially in children, compromises its diagnostic precision (15). In contrast, the Appendicitis Inflammatory Response (AIR) score, a relatively newer tool, incorporates objective parameters including C-reactive protein (CRP), temperature, and leukocyte counts, and intentionally excludes vague symptoms. This makes the AIR score more reliable and reproducible, particularly in emergency settings where rapid triage is essential. Based on these theoretical advantages, our study was conducted to compare the diagnostic performance of the Alvarado and AIR scores using histopathological confirmation as the reference standard. In our findings, the AIR score demonstrated superior diagnostic accuracy (89.19%) compared to the Alvarado score (69.73%). This is consistent with previous literature; a study by Mumtaz et al. reported diagnostic accuracies of 86.2% for AIR and 71.6% for Alvarado score, supporting the enhanced discriminatory ability of AIR (16). Additionally, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the AIR score in our cohort were all substantially higher, indicating better performance both in identifying true positives and ruling out non-appendicitis cases. In contrast, although the Alvarado score showed a high PPV (93.44%), it had a poor NPV (23.81%), highlighting its limitation in confidently excluding the diagnosis, a drawback particularly important in borderline or atypical presentations.

Our results also echoed findings from Hesham et al., who noted that normal appendices on histopathology were more commonly reported in female patients. This is often attributed to gynecological conditions such as pelvic inflammatory disease, ruptured ovarian cysts, and ectopic pregnancies, which may mimic appendicitis and lower the diagnostic threshold in women (1). Similarly, our study observed a higher prevalence of acute appendicitis in males, a well-established epidemiological trend documented across multiple populations (17,18). These gender-based patterns emphasize the need for more refined diagnostic algorithms, especially for women of reproductive age, where overlapping conditions often complicate clinical judgment.

The major strength of our study lies in its prospective design, adequate sample size, and blinding of the surgical team to score-based categorization, thereby minimizing selection and observer bias. The use of histopathology as the definitive diagnostic reference adds robustness to the validity of our findings. However, the study is not without limitations. The single-center design may limit the generalizability of results to broader or more diverse populations. Although efforts were made to standardize clinical assessments and laboratory measurements, variability in symptom reporting and laboratory processing could still introduce minor inconsistencies. Additionally, while the AIR score performed well, its reliance on CRP may still limit utility in settings where such biochemical testing is unavailable or delayed.

Future research should focus on multicentric studies with larger and more heterogeneous populations to validate these findings across different healthcare settings. Moreover, integrating these scores into combined diagnostic algorithms alongside selective imaging could help optimize the diagnostic pathway for acute appendicitis, balancing accuracy with resource efficiency. Development and validation of age-specific or gender-specific modifications to existing scores may also enhance diagnostic sensitivity in subpopulations, particularly pediatric and female patients. In conclusion, the AIR score, by virtue of its higher sensitivity, specificity, and overall accuracy, emerges as a more reliable diagnostic tool than the Alvarado score for suspected acute appendicitis, and its clinical adoption may contribute to reduced negative appendectomy rates and improved patient outcomes.

CONCLUSION:

This study, aimed at comparing the predictive value of the Alvarado and Appendicitis Inflammatory Response (AIR) scores

 © Authors. CC BY 4.0. **DOI**: https://doi.org/10.61919/pqmgcv78

against histopathological outcomes in suspected cases of acute appendicitis, found that the AIR score demonstrated superior diagnostic performance in terms of accuracy (89.19%), sensitivity (90.12%), specificity (82.61%), positive predictive value (97.33%), and negative predictive value (54.29%) compared to the Alvarado score. The inclusion of objective markers such as C-reactive protein and the exclusion of subjective symptoms enhances the AIR score's reliability, making it a valuable tool in guiding early clinical decision-making. Its adoption can potentially reduce unnecessary imaging and negative appendectomies, particularly in resource-constrained settings. Clinically, the AIR score offers a safer, faster, and cost-effective alternative for stratifying patients based on risk, while future research should focus on its integration into standardized diagnostic protocols and validation across different populations to strengthen evidence-based practice in surgical emergencies.

REFERENCES

1. Ohle R, O'Reilly F, O'Brien KK, Fahey T, Dimitrov BD. The Alvarado Score for Predicting Acute Appendicitis: A Systematic Review. BMC Med. 2011;9:139.
2. Ferris M, Quan S, Kaplan BS, Molodecky N, Ball CG, Chernoff GW, et al. The Global Incidence of Appendicitis: A Systematic Review of Population-Based Studies. Ann Surg. 2017;266(2):237–41.
3. Tait IS. The First Appendicectomy. J R Coll Surg Edinb. 1995;40(6):392–6.
4. Flum DR, Morris A, Koepsell T, Dellinger EP. Has Misdiagnosis of Appendicitis Decreased Over Time? A Population-Based Analysis. JAMA. 2001;286(14):1748–53.
5. Bickell NA, Aufses AH Jr, Rojas M, Bodian C. How Time Affects the Risk of Rupture in Appendicitis. J Am Coll Surg. 2006;202(3):401–6.
6. Terasawa T, Blackmore CC, Bent S, Kohlwes RJ. Systematic Review: Computed Tomography and Ultrasonography to Detect Acute Appendicitis in Adults and Adolescents. Ann Intern Med. 2004;141(7):537–46.
7. Alvarado A. A Practical Score for the Early Diagnosis of Acute Appendicitis. Ann Emerg Med. 1986;15(5):557–64.
8. 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(8):1843–9.
9. Sammalkorpi HE, Mentula P, Leppäniemi A. A New Adult Appendicitis Score Improves Diagnostic Accuracy of Acute Appendicitis—A Prospective Study. BMC Gastroenterol. 2014;14:114.
10. Mohamed AA, Mohamed MA, Omer MM. Diagnostic Accuracy of Alvarado and AIR Scores in Acute Appendicitis: A Comparative Study. Int J Surg Open. 2020;25:38–42.
11. Humes DJ, Simpson J. Acute Appendicitis. BMJ. 2006;333(7567):530–4.
12. Snyder MJ, Guthrie M, Cagle S. Acute Appendicitis: Efficient Diagnosis and Management. Am Fam Physician. 2018;98(1):25–33.
13. Bickell NA, Szabat C, Aufses AH Jr. Racial Disparities in Outcomes after Appendectomy. J Am Coll Surg. 2008;207(3):410–6.
14. Doria AS, Moineddin R, Kellenberger CJ, Epelman M, Beyene J, Schuh S, et al. US or CT for Diagnosis of Appendicitis in Children and Adults? A Meta-Analysis. Radiology. 2006;241(1):83–94.
15. Goldman RD, Carter S, Stephens D, Antoon R, Mounstephen W, Langer JC. Prospective Validation of the Pediatric Appendicitis Score. J Pediatr. 2008;153(2):278–82.
16. Mumtaz S, Ahmed M, Chohan U, Anwar M. Diagnostic Accuracy of Appendicitis Inflammatory Response (AIR) Score Compared to Alvarado Score in Acute Appendicitis. Pak J Med Sci. 2020;36(4):773–7.
17. Ilves I, Fagerström A, Herzig KH, Juvonen P, Miettinen P. Seasonal Variations of Acute Appendicitis and Nonspecific Abdominal Pain in Finland. World J Gastroenterol. 2014;20(14):4037–42.
18. Moris D, Paulson EK, Pappas TN. Diagnosis and Management of Acute Appendicitis in Adults: A Review. JAMA. 2021;326(22):2299–311.

Disclaimer: The views and data in articles are solely those of the authors. The journal disclaims liability for any use of the published content