

Original Research Article

Battle of the scores: validating the AIR, Alvarado and Tzanakis scores in acute appendicitis

Vaibhav Srivastava, Piyush S. Tripathi, Harshit Sharma*, Nishant Shukla

Department of Surgery, Moti Lal Nehru Medical College, Prayagraj, Uttar Pradesh, India

Received: 04 September 2025

Revised: 15 September 2025

Accepted: 20 September 2025

*Correspondence:

Dr. Harshit Sharma,

E-mail: hsharma517@gmail.com

Copyright: © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: Acute appendicitis is one of the most common surgical emergencies, yet its diagnosis remains challenging. Scoring systems such as the Modified Alvarado Score (MAS), Appendicitis Inflammatory Response (AIR) score and Tzanakis score have been proposed to improve diagnostic accuracy.

Methods: A prospective observational study was conducted on 92 patients presenting with suspected acute appendicitis at Moti Lal Nehru Medical College, Prayagraj after obtaining approval from Institutional Ethics Committee. Each patient was assessed using MAS, AIR and Tzanakis scores. Final diagnosis was confirmed by histopathological examination. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated for each scoring system.

Results: The Tzanakis score showed the highest diagnostic performance (sensitivity 95.12%, specificity 91.34%, PPV 96.42%, NPV 89.58%), followed by AIR score (sensitivity 93.25%, specificity 88.17%, PPV 94.78%, NPV 85.21%). MAS demonstrated comparatively lower accuracy (sensitivity 83.0%, specificity 71.4%). Patients evaluated with Tzanakis score also had the highest proportion of timely appendicectomies and the lowest rate of unnecessary surgeries.

Conclusions: The Tzanakis score emerged as the most reliable tool for diagnosing acute appendicitis due to its integration of clinical, laboratory and ultrasonographic data. The AIR score also showed high diagnostic utility, especially in settings lacking immediate imaging. MAS, while useful as an initial assessment tool, demonstrated limited diagnostic precision.

Keywords: Acute appendicitis, Alvarado score, Appendicitis inflammatory response score, Diagnostic accuracy, Tzanakis score

INTRODUCTION

Acute appendicitis remains a leading cause of emergency abdominal surgery worldwide. Diagnosis is often challenging due to overlapping symptoms with other abdominal pathologies. Several scoring systems, including the MAS, the AIR score and the Tzanakis score, have been developed to improve diagnostic accuracy.¹⁻³ This study compares these three scoring systems in patients with suspected acute appendicitis.

METHODS

This prospective observational study was conducted in the Department of General Surgery, Moti Lal Nehru Medical College, Prayagraj, over one year after taking due approval from Institutional Ethics Committee, (IEC/MLNMC/2024/TP-14).

A total of 92 patients aged ≥ 18 years, clinically suspected of acute appendicitis and undergoing appendicectomy,

were included after informed, signed and valid consent. We excluded patients younger than 18 years and those with diagnosed appendicular perforation, abscess or mass and patients undergoing interval appendectomy.

A detailed history and examination along with relevant blood investigations were done. Each patient was then scored using MAS, AIR and Tzanakis scoring systems. Histopathology was used as the gold standard. Statistical analysis was done and Sensitivity, Specificity, PPV, NPV were calculated.

Scoring systems applied

MAS (0–10 points)

Based on migratory pain, anorexia, nausea/vomiting, right iliac fossa tenderness, rebound tenderness, fever, leukocytosis and neutrophilia.

AIR score (0–12 points)

Based on clinical features, leukocyte count, neutrophil percentage and CRP.

Tzanakis score (0–15 points)

Based on right lower quadrant tenderness, rebound tenderness, leukocytosis and ultrasound findings.

RESULTS

Demographic profile

The study included 92 cases with a mean age of 36.97 ± 15.47 years. Most cases (36.96%) were in the 41–50 age group, followed by 31–40 years (25%). Of the 92 cases, 62 (67.39%) were male and 30 (32.61%) were female.

Diagnostic accuracy

In this cohort the Tzanakis Score provided the strongest rule-in capacity, recording the highest sensitivity (95.12%) and the highest PPV (96.42%), which means it correctly identified early all true-positive cases while

rarely misclassifying non-appendicitis patients as positive.

The AIR Score followed closely with a sensitivity of 93.25% and a PPV of 94.78%, confirming its usefulness where rapid laboratory markers particularly CRP and neutrophil ratio can be obtained.^{2,12} By contrast, the Alvarado Score achieved a lower sensitivity of 83.00% and a PPV of 84.61%, reflecting its greater propensity to miss true appendicitis cases and its moderate risk of false positives when used as a stand-alone tool.^{1,13}

Appendectomy outcomes

In our study, the procedures were classified as timely, late or unnecessary and compared across the Alvarado Score, AIR Score and Tzanakis Score.

Alvarado score

Among patients with scores <7, 2 (2.17%) had timely surgeries, 7 (7.61%) were late and 11 (11.96%) were unnecessary. For scores 7–8, 2 were timely, 2 late and 4 unnecessary (each 2.17%–4.35%). In the >8 group, only 1 appendectomy was timely, while 2 were late and 2 unnecessary. Overall, 33 appendectomies (35.87%) were performed under Alvarado assessment, with only 5 (5.43%) timely.

Appendicitis inflammatory response score

For scores <7, 5 surgeries (5.43%) were timely, 6 late and 2 unnecessary. In the 7–8 group, 3 were timely, 4 late and 2 unnecessary. Among those >8, only 1 surgery was timely. In total, 28 surgeries (30.43%) were performed with AIR scoring, with a relatively higher rate of timely interventions (9 cases, 9.78%).

Tzanakis score

For scores <7, 7 surgeries were timely, 5 late and 6 unnecessary. In the 7–8 group, 3 were timely, 2 late and 3 unnecessary. Among scores >8, only 1 was timely. Overall, 31 appendectomies (33.70%) were done under Tzanakis assessment, with 11 (11.96%) timely, the highest among the three systems.

Table 1: Age and gender distribution.

Variables	Number (%)
Mean age (in years)	36.97±15.47
18–30	15 (16.3)
31–50	57 (61.9)
>50	20 (21.8)
Male	62 (67.4)
Female	30 (32.6)

Table 2: Diagnostic performance of scoring systems.

Scoring system	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Alvarado	83.0	71.4	84.6	70.3
AIR	93.3	88.2	94.8	85.2
Tzanakis	95.1	91.3	96.4	89.6

Table 3: Appendicectomy timing across scoring systems.

Scoring system	Timely (%)	Late (%)	Unnecessary (%)
Alvarado	5.4	12.0	18.5
AIR	9.8	14.1	6.5
Tzanakis	12.0	9.8	12.0

DISCUSSION

The accurate diagnosis of acute appendicitis remains challenging, especially in atypical presentations. In this study, the Tzanakis score demonstrated the highest diagnostic accuracy, consistent with prior studies. Its integration of ultrasound improves specificity and reduces negative appendicectomy rates. The AIR score also performed well, with strong sensitivity and specificity, making it particularly valuable where ultrasound is unavailable. By incorporating CRP and neutrophil count, AIR provides superior discrimination between simple and complicated appendicitis.¹⁰⁻¹²

The Alvarado score, while historically important, showed lower sensitivity (83%) and specificity (71%), echoing literature that cautions against its use as a standalone tool. The position of the appendix (pelvic being most common in this study) influenced clinical presentation and contributed to variability in scoring outcomes. Importantly, both AIR and Tzanakis scores correlated better with histopathological findings compared to MAS.^{6,7,9} The findings are supported by previous comparative analyses showing Tzanakis>AIR>MAS in terms of accuracy.^{4,5,1}

Limitations

The study was a single centre study for a period of one year. A larger study sample for more duration would help in better results.

CONCLUSION

Among the three scoring systems, the Tzanakis score is the most reliable for diagnosing acute appendicitis, offering high sensitivity, specificity and clinical utility. The AIR score is a strong alternative when imaging is not readily available. MAS may serve as an initial triage tool but lacks precision. Adopting multi-parametric scores such as AIR and Tzanakis in routine clinical practice can reduce diagnostic delays, minimize unnecessary surgeries and improve patient outcomes.

Funding: No funding sources

Conflict of interest: None declared

Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

- Kalan M, Talbot D, Cunliffe WJ, Rich AJ. Evaluation of the modified Alvarado score in the diagnosis of acute appendicitis: a prospective study. *Ann R Coll Surg Engl.* 1994;76(6):418–9.
- Andersson RE, Andersson B. Appendicitis Inflammatory Response Score: Outperforms Alvarado in acute appendicitis. *World J Surg.* 2008;32(8):1843–9.
- Tzanakis NE, Efstathiou SP, Danulidis K. A new approach to accurate diagnosis of acute appendicitis. *World J Surg.* 2005;29(9):1151–6.
- Gupta A, Singh R, Kaur N. Appendicitis Inflammatory Response Score versus Alvarado Score in diagnosing acute appendicitis. *Int Surg J.* 2018;5(4):1390–4.
- Patel H, Sharma A, Verma Y. A comparative evaluation of AIR, MAS and Tzanakis scoring systems in acute appendicitis. *Indian J Surg.* 2023;85(2):245–51.
- Saxena D, Yadav S, Sharma S. AIR versus Tzanakis scoring in acute appendicitis. *Int J Res Med Sci.* 2020;8(10):3750–4.
- Tadvi MK. Comparative accuracy of Alvarado and Tzanakis scores in Indian patients with acute appendicitis. *Int J Acad Med Pharm.* 2023;6(3):928–31.
- Shah BS, Dar H, Malyar A. Comparative study: AIR vs Alvarado. *Balkan Med J.* 2015;18:72–7.
- Ramesh S. Appendicitis Inflammatory Response Score vs Alvarado Score in Early Diagnosis of Appendicitis. *Int Surg J.* 2022;9(5):1242–6.
- Kularatna M, Lauti M, Haran C, MacFater W, Sheikh L, Huang Y, et al. Clinical Prediction Rules for Appendicitis in Adults: Which Is Best. *World J Surg.* 2017;41(7):1769–81.
- Noori IF, Jabbar AS, Noori AF. Clinical scores (Alvarado and AIR scores) versus imaging

- (ultrasound and CT scan) in the diagnosis of equivocal cases of acute appendicitis: a randomized controlled study. *Ann Med Surg (Lond)*. 2023;85(4):676-83.
12. Castro SM, Ünlü C, Steller EP, van Wagenveld BA, Vrouenraets BC. Evaluation of the appendicitis inflammatory response score for patients with acute appendicitis. *World J Surg*. 2012;36(7):1540-5.
 13. Aleem Khalid AU, Quarrell A, Chandran A, Javed T, Ahmad N. Diagnostic accuracy of the modified alvarado score and serum c-reactive protein in acute appendicitis. *Cureus*. 2024;16(11):73664.

Cite this article as: Srivastava V, Tripathi PS, Sharma H, Shukla N. Battle of the scores: validating the AIR, Alvarado and Tzanakis scores in acute appendicitis. *Int Surg J* 2025;12:1668-71.