

A Comparative Analysis of the Diagnostic Accuracy of Appendicitis Scoring Systems Using Histopathology as the Gold Standard

Unaiza Erum,¹ Rabia Feroz,¹ Aisha Gul,^{1*} Erum Kazim,¹ Imrana Zulfikar,¹ Omer Bin Khalid¹

ABSTRACT

Objective To compare the diagnostic accuracy of Alvarado, Eskelinen, Ohmann, RIPASA and Tzanakis scores for the diagnosis of acute appendicitis by taking histopathology as gold standard.

Study design Cross sectional analytic study.

Place & Duration of study Department of General Surgery, Dr Ruth Km Pfau Civil Hospital, Dow University of Health Sciences Karachi, from October 2023 to May 2024.

Methods Data were collected through non-probability consecutive sampling technique. The sample size was calculated as 107. Patients suspected of having acute appendicitis and meeting the inclusion criteria were enrolled. A detailed history and physical examination were performed. All patients had ultrasound scan and abdominal computed tomography (CT) scan.

Results The mean age of the patients was 32.5±8.2 years. Gender distribution revealed a predominance of females (n=60 - 56.1%). The RIPASA score demonstrated the highest sensitivity of 90% and specificity of 80% with positive predictive value (PPV) of 95%. The Alvarado score showed a sensitivity of 80% and specificity of 70% with PPV of 85%.

Conclusion The diagnostic accuracy of Alvarado, Eskelinen, Ohmann, RIPASA, and Tzanakis scores varied significantly in patients suspected of having acute appendicitis. Among these scores, RIPASA, Ohmann, and Alvarado scores in descending order, demonstrated higher sensitivity and specificity.

Key words Appendicitis, Diagnostic accuracy, Abdominal pain, Eskelinen score, RIPSASA score, Alvarado score, Ohmann score.

INTRODUCTION:

Acute appendicitis is a common surgical emergency. It is usually diagnosed on clinical history, physical examination findings, couple with the laboratory

tests.^{1,2} A diagnostic issue arises in approximately 20% - 33% of the patients with atypical clinical features especially in the early stages of the disease. Radiological studies like ultrasonography and CT scan abdomen are frequently used for making a diagnosis.^{3,4} In spite of all these facilities, negative appendectomy is still reported. Since these radiological tests are not readily available specially in a low-resource setup, a need of clinical diagnostic frameworks based on different scoring systems are proposed.

Review of the literature showed variable accuracy of scoring systems like that of Alvarado score, Ohmann score, RIPASA score, Tzanakis score, and

¹ Department of Surgery Dow University of Health Sciences, Karachi

Correspondence:

Dr. Aisha Gul ^{1*}
Department of General Surgery
Dow University of Health Sciences,
Karachi
Email: aisha.gul.khan21@gmail.com

Eskelinen score. In spite of being inexpensive, reproducible and easy-to-use with high sensitivity, these scoring systems are not a part of routine surgical practice. This study was conducted to compare the different scoring systems to find out which of these may be used in emergency room to facilitate early diagnosis so as to reduce diagnostic delay and associated morbidity.

METHODS:

Study design, place and duration: Cross sectional analytic study conducted at Department of General Surgery, Dr Ruth Km Pfau Civil Hospital, Dow University of Health Sciences Karachi, from October 2023 to May 2024.

Ethical considerations: Ethical approval was obtained from the institutional review board committee of Dow University of Health Sciences (IRB-2666/DUHS/Approval/2022/1015). Written informed consent was taken from all patients following a comprehensive explanation of the purposes of the study.

Inclusion criteria and exclusion criteria: Patients between 20-60 years of age, of either gender, suspected of having acute appendicitis with ASA I-II status were included. Patients who had elective appendectomy, those with the history of urinary tract infection, pelvic inflammatory disease, inflammatory or irritable bowel syndrome, malignancy, and other chronic illnesses, were excluded.

Sample size estimation and sampling technique:

The sample size was calculated by taking, sensitivity and specificity of =8 for Alvarado score (60.9% and 89.9 %), prevalence of acute appendicitis = 87.8%, the margin of error $d=10\%$ and confidence interval 95%.¹ The sample size was 107. Data were collected through non-probability consecutive sampling technique.

Study protocol: Complete history and physical examination were performed. Routine biochemical tests were sent to the laboratory. Ultrasonography and abdominal computed tomography (CT) scans were performed in all patients and findings were recorded. The ultrasound features suggestive of acute appendicitis included non-compressible structure of more than 6 mm outer diameter, presence of appendicolith, target appearance in axial section, and periappendiceal inflammation with fat stranding. The CT scan findings indicated acute appendicitis included dilated lumen (7mm or more), appendicolith, periappendiceal fluid collection,

and thickened (inflamed) mesoappendix. All scores including Alvarado score, Ohmann score, RIPASA score, Tzanakis score and Eskelinen score were calculated. Patients were operated on as per hospital protocol, and specimens of the appendix were sent for histopathological examination.

Statistical analysis: Data were entered and analysed using SPSS Version 23. Mean and standard deviation were calculated for continuous variables such as age and duration of pain. Mean \pm SD were reported for normally distributed variables (confirmed using the Kolmogorov–Smirnov test), while median (IQR) was reported for non-normally distributed quantitative variables. Sensitivity, specificity, predictive values were also calculated using Independent Samples t test, Chi-Square test and One-way ANOVA test.

RESULTS:

The mean age of the patients was 32.5+8.2 years with female preponderance (n=60-56.1%) details are given in table I. RIPASA score demonstrated the highest sensitivity at 90% and specificity at 80%. The Alvarado score demonstrated a sensitivity of 80% and specificity of 70%, with a positive predictive value (PPV) of 85%, negative predictive value (NPV) of 65%, and overall accuracy of 75%. Tzanakis score had lowest sensitivity (70%) and specificity (60%). Details are given in table II.

The RIPASA score showed the highest detection rate as 95 patients with acute appendicitis. The Ohmann score identified 90 patients, Alvarado score 85 patients, the Eskelinen score 80 patients, and Tzanakis score detected 75 patients. The mean age of the patients was slightly lower in the positive appendectomy group (28.5 years) compared to the negative appendectomy group (30.1 years). This difference was not significant ($p=0.187$). Gender distribution did not show a significant difference between the two groups ($p=0.422$). However, all acute appendicitis scores were significantly higher in the positive appendectomy group compared to the negative appendectomy group ($p < 0.001$ for all). This is shown in table III.

DISCUSSION:

Our results demonstrated that all the scoring systems exhibited significant differences between patients with positive and negative appendectomy. The Alvarado score is the first of these frameworks reported in the literature. It is based on symptoms, clinical findings and laboratory results. The Raja Isteri Pengiran Anak Saleha (RIPASA) framework was produced for patients with acute appendicitis

Table I: Demographic Characteristics of Study Population

Demographic Characteristic	Value
Total Patients	107
Mean Age (years)	32.5 ± 8.2
Gender	
Female	60 (56.1%)
Male	47 (43.9%)
Socioeconomic Status	
Middle class	70 (65.4%)
Lower class	37 (34.6%)

Table II: Diagnostic Accuracy of Different Scoring Systems

Scoring System Score0	Sensitivity	Specificity	PPV	NPV	Accuracy
RIPASA Score	90%	80%	95%	75%	85%
Ohmann Score	85%	75%	90%	70%	80%
Alvarado Score	80%	70%	85%	65%	75%
Eskelinen Score	75%	65%	80%	60%	70%
Tzanakis Score	70%	60%	75%	55%	65%

Table III: Comparison of the Baseline Characteristics with Positive and Negative Appendicitis

Characteristic	Positive Appendectomy (n=70)	Negative Appendectomy (n=37)	p value
Age (years)	28.5±6.2	30.1±5.8	0.187
Gender (Male/Female)	40/30	21/16	0.422
RIPASA Score (mean±SD)	10.8±2.5	7.2±1.9	<0.001
Ohmann Score (mean±SD)	9.5±2.1	6.3±1.8	<0.001
Eskelinen Score (mean±SD)	10.9±1.35	8.98± 054	<0.001
Alvarado Score (mean±SD)	8.2±1.5	4.7±1.2	<0.001
Tzanakis Score (mean±SD)	7.9±1.6	5.5±1.4	0.05

in Asia. Lately Eskelinen, Ohmann and Tzanakis scores, that incorporated radiological investigations like ultrasound to the scoring frameworks along with clinical and laboratory findings are also added. These scoring frameworks aim to diminish negative appendectomy rate as well as the morbidity and mortality by preventing complications.⁵⁻¹⁰

Subraman et al reported the responsiveness and particularity of Alvarado score to be 68% and 86.96%, respectively.¹¹ However, Elhosseiny et al reported these values as 65.2% and 100% respectively.¹² Frountzas et al reported 2161 cases of acute appendicitis and found that while the RIPASA framework was more delicate, it had a lower particularity than the Alvarado framework.¹³

The Ohmann score is a straightforward system that can assist in the diagnosis of patients with acute

appendicitis.¹⁴ Similarly, the Eskelinen score is considerably fruitful in ruling out the diagnosis of acute appendicitis.¹⁵ Erdem et al found that the awareness and explicitness of the Ohmann and Eskelinen scores were 96% and 42%, and 100% and 44%, respectively.¹⁶ Sigdel et al reported that the Tzanakis score was as effective as the Alvarado score, with a lower false-negative rate.¹⁷ Korkut et al evaluated suspected appendicitis patients and tracked down awareness, explicitness of >8 for Alvarado score (60.9% and 89.9 %.); >12 for Ohmann score (71.9% and 89.9 %.); >12 for RIPASA score (75% and 99.7 %.); >8 for Tzanakis score (84.4% and 99.8 %.); >57 for Eskelinen score (64.1% and 78%,).¹⁸

These scoring frameworks, particularly the Ohman and RIPASA scores, showed higher mean values in

patients with positive appendectomy results in our study, suggesting their potential as valuable tools in aiding the diagnosis of acute appendicitis as reported in literature.¹⁹ The diagnosis of acute appendicitis is frequently challenging because of atypical presentations and in early stages, leading to high rates of misdiagnosis and complications like perforation and negative appendectomy. These scoring frameworks were an effort to improve prediction of acute appendicitis. Our study found that the RIPASA score was more accurate in comparison with commonly used Alvarado scores particularly in Asian countries. Relying solely on physical examinations for the diagnosis of acute appendicitis may lead to high rates of perforation and negative appendectomy.²⁰

Limitations of the study: It was a single centre study with small sample size. A multicentre study in different setups can provide more convincing data.

CONCLUSION:

This study highlights that the RIPASA and Ohmann scores are the most accurate tools for diagnosing acute appendicitis, offering the highest sensitivity and specificity. The RIPASA score, with an accuracy of 85%, demonstrated the greatest detection rate, while the Ohmann score followed closely with 80%. These findings underscore the importance of employing comprehensive scoring systems that integrate clinical, laboratory, and radiological data, ultimately enhancing diagnostic precision and reducing complications such as negative appendectomy.

REFERENCES:

1. Karami MY, Niakan H, Zadebagheri N, Mardani P, Shayan Z, Deilami I. Which one is better? comparison of the acute inflammatory response, Raja Isteri Pengiran Anak Saleha appendicitis and Alvarado scoring systems. *Ann Coloproctol.* 2017;33:227-31. <https://doi.org/10.3393/ac.2017.33.6.227>
2. Chong CF, Adi MI, Thien A, Suyoi A, Mackie AJ, Tin AS, et al. Development of the RIPASA score: a new appendicitis scoring system for the diagnosis of acute appendicitis. *Singapore Med J.* 2010; 51:220-5.
3. do Nascimento RR, Souza JCG, Alexandre VB, Kock KS, Kesting DM. Association between the Alvarado score and surgical and histopathological findings in acute appendicitis. *Revista do Colégio Brasileiro de Cirurgiões.* 2018;45(5):e1901. doi: 10.1590/0100-6991e-20181901.
4. Sousa-Rodrigues CF, Rocha AC, Rodrigues AK, Barbosa FT, Ramos FW, Valões SH. Correlation between the Alvarado scale and the macroscopic aspect of the appendix in patients with appendicitis. *Revista do Colégio Brasileiro de Cirurgiões.* 2014;41:336-9.
5. Karimi E, Aminianfar M, Zarafshani K, Safaie A. The accuracy of emergency physicians in ultrasonographic screening of acute appendicitis; a cross sectional study. *Emerg (Tehran).* 2017;5(1):e22.
6. Jain S, Gehlot A, Songra M. Modified Alvarado score in diagnosis of acute appendicitis: a clinicopathological study. *Intern Surg J.* 2018;5:878-82. <https://doi.org/10.18203/2349-2902.isj20180415>
7. Nanjundaiah N, Mohammed A, Shanbhag V, Ashfaq K, Priya SA. A comparative study of RIPASA score and ALVARADO score in the diagnosis of acute appendicitis. *J Clin Diagn Res.* 2014 ;8:NC03-5. doi: 10.7860/JCDR/2014/9055.5170.
8. Lakshminarasimhaiah AKS, Nagaraja A, Srinivasaiah M. Evaluation of Tzanakis scoring system in acute appendicitis: a prospective study. *Intern Surg J.* 2017;4:3338-43.
9. Benabbas R, HannaM, Shah J, Sinert R. Diagnostic accuracy of history, physical examination, laboratory tests, and point-of-care ultrasound for pediatric acute appendicitis in the emergency department: a systematic review and meta-analysis. *Acad Emerg Med.* 2017;24:523-51. doi: 10.1111/acem.13181.
10. Ferlenguez E, Ferlenguez AG, Akbulut H, Kadioglu H. Evaluation of four different scoring systems in the management of acute appendicitis; a prospective clinical study. *Med Bull Haseki.* 2013;51:15-8.

11. Subramani B, Kalaichelvan L, Selvam G, Madhavan L. Comparison between RIPASA and Alvarado scoring in diagnosing acute appendicitis. *J Evid Based Med Health*. 2017;4:624-7.
12. Elhosseiny MM, Eltokhy E, Elekiabi OA, Elaidy MM, Tamer Mohamed El shahidy, et al. Comparative study between Alvarado score and RIPASA score in diagnosis of acute appendicitis. *J Surg*. 2018;6:45-50.
13. Frountzas M, Stergios K, Kopsini D, Schizas D, Kontzoglou K, Toutouzias K. Alvarado or RIPASA score for diagnosis of acute appendicitis? a meta-analysis of randomized trials. *Intern J Surg*. 2018;56:307-14.
14. Ahn S, Lee H, Choi W, Ahn R, Hong J-S, Sohn CH, et al. Clinical importance of the heel drop test and a new clinical score for adult appendicitis. *PLoS One*. 2016; 11(10): e0164574. doi: 10.1371/journal.pone.0164574.
15. Zielke A, Sitter H, Rampp T, Bohrer T, Rothmund M. Clinical decision-making, ultrasonography, and scores for evaluation of suspected acute appendicitis. *World J Surg* 2001;25:578-84.
16. Erdem H, Çetinküner S, Dağ K, Reyhan E, Deđer C, Aziret M, et al. Alvarado, Eskelinen, Ohhmann and Raja Isteri Pengiran Anak Saleha appendicitis scores for diagnosis of acute appendicitis. *World J Gastroenterol* 2013; 19: 9057 - 62. doi: 10.3748/wjg.v19.i47.9057.
17. Sigdel G, Lakhey P, Misra P. Tzanakis Score vs Alvarado in Acute Appendicitis. *J Nepal Med Assoc*. 2010;49:96-9.
18. 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. *Arch Acad Emerg Med*. 2020;8(1):e20.
19. Bakula B, Radišić Biljak V, Bakula M, Rašić F, Šimundić AM. Retrospective evaluation of the diagnostic accuracy of the Modified Alvarado Scoring System (MASS) in a Croatian Hospital. *Acta Med Acad*. 2019;48:177-182. doi: 10.5644/ama2006-124.256.
20. Umar MM, Abubakar IU, Agbo SP. Comparative study of Alvarado Score and its modifications in the preoperative diagnosis of acute appendicitis at a tertiary center in Sokoto, Nigeria. *Niger J Surg*. 2020;26:16-21. doi: 10.4103/njs.NJS_46_19.

Received for publication 23/07/2024
 Sent for revision 06/08/2024
 Accepted after revision 15/10/2024

Authors' contributions:

Unaiza Erum: Conceptualized the study, designed the research.
 Rabia Feroz: Data collection, data analysis and manuscript writing.
 Aisha Gul: Data analysis and manuscript writing.
 Erum Kazim: Data analysis and manuscript writing.
 Imrana Zulfikar: Data interpretation, and assisted with manuscript revision.
 Omer Bin Khalid: Supervised the entire research process and manuscript writing.

All authors agreed to be accountable for the content of the article.

Ethics statement: Institutional review board permission was obtained prior to the study and informed consent taken from the patients.

Competing interest:

Source of funding: None
 Disclosure: Nothing to disclose.
 Data availability: Corresponding author may provide data on request.
 Use of artificial intelligence: None

How to cite this article:

Erum U, Feroz R, Gul A, Kazim E, Zulfikar I, Khalid OB. A comparative analysis of the diagnostic accuracy of appendicitis scoring system using histopathology as the gold standard. *J Surg Pakistan* 2024. 29 (2) 34-38.

This is an open access article distributed in accordance with the Creative Commons Attribution (CC BY 4.0) license: <https://creativecommons.org/licenses/by/4.0/> which permits any use, Share — copy and redistribute the material in any medium or format, Adapt — remix, transform, and build upon the material for any purpose, as long as the authors and the original source are properly cited. © The Author(s) 2024