The Research of Medical Science Review
ISSN: 3007-1208 & 3007-1216 Volume 4, Issue 2, 2026

DIAGNOSTIC ACCURACY OF ULTRASOUND IN THE DIAGNOSIS OF ACUTE APPENDICITIS IN PATIENTS WITH RIGHT ILIAC FOSSA PAIN TAKING PER-OPERATIVE FINDINGS AS GOLD STANDARD

Rabia Anwar[*1], Ghazala Naz[2], Tahira Nishtar[3], Hamza Javed[4]

[*1,2]*Residnet Radiologist, Lady Reading Hospital- Peshawar*
[3]*Professor of Radiology, Lady Reading Hospital–Peshawar*
[4]*Resident Radiologist, Ayub Teaching Hospital- Abbottabad*

DOI: https://doi.org/10.5281/zenodo.19396097

Keywords
Acute appendicitis; Alvarado score; Diagnostic accuracy; Sensitivity; Specificity; Abdominal Ultrasound

Article History
Received: 27 December 2025
Accepted: 11 February 2026
Published: 26 February 2026

Corresponding Author: *
Tahira Nishtar

Abstract

Introduction: *Acute appendicitis (AA) is a leading cause of acute abdomen requiring emergency surgery. Prompt diagnosis is vital to prevent perforation, related morbidity and mortality. Ultrasonography (US) offers a non-invasive, cost-effective, and radiation-free diagnostic option, especially in low-resource settings. This study assessed the diagnostic accuracy of ultrasound in detecting AA, taking per-operative findings as the gold standard.*

Methodology: *This cross-sectional validation study was conducted at Ayub Teaching Hospital, Abbottabad, and Lady Reading Hospital, Peshawar (January 2024 to February 2025). A total of 139 patients aged ≥ 18 years with right iliac fossa pain and clinical suspicion of AA were included through consecutive sampling. Abdominal ultrasound was performed by consultant radiologists using defined criteria (appendiceal diameter > 6 mm, non-compressibility, peri-appendiceal fluid, and ± appendicolith). Data were analyzed in SPSS v.27 to calculate sensitivity, specificity, predictive values, and overall accuracy.*

Results: *Ultrasound correctly identified appendicitis in 87 patients and excluded it in 31 patients ($p < 0.001$). Sensitivity was 85.3%, specificity 83.8%, PPV 93.6%, NPV 67.4%, and overall accuracy 84.9%. Diagnostic performance was comparable across genders, with slightly higher PPV among females. For Alvarado < 7, ultrasound achieved 100% sensitivity and NPV, while PPV approached 99% in Alvarado ≥ 7 cases.*

Conclusion: *Ultrasound has a high accuracy for diagnosing AA, reliably excluding the disease in low-risk and confirming it in high-risk patients. It remains a safe, accessible, and cost-effective first-line imaging modality, supporting its integration into diagnostic algorithms in resource-limited healthcare settings.*

INTRODUCTION

Acute appendicitis (AA), inflamed vermiform appendix is a quite prevalent source of abdominal pain and it remains a real dilemma in health care practice on an international scale, taking into consideration the variety of patients and contexts 1. It typically presents in young adults and adolescents and predominantly appears as that typical right lower quadrant pain, and therefore a timely accurate diagnosis must be made to prevent such risks as perforation that is an escalator of morbidity and mortality 2. It has a lifetime risk of 7-8, it is a significant health concern worldwide with its prevalence highest among individuals aged 10 -30 although it can occur at any age 3. Although the incidence declined in the Western world mid-century, in the 21 st century the newly industrialized countries have

experienced an increase in incidence 4, and appendicitis is currently the eighth most common cause of death in the world, with the death rate expected to increase in the following years 5. Physical tests, laboratory tests, and imaging have been a combination of tests that we traditionally used to diagnose it 6. Nowadays medical age provides a great number of tools to nail down AA and the choice of the tool is determined by such aspects as the availability of the tools, their cost, radiation level, and the precision of each of the methods 7, 8. Ultrasound has emerged as a leader of the rest since it is non-invasive, inexpensive, and is readily available. It has giant benefits in diagnosing AA 9 particularly in low-resource environments that you do not receive radiation and you receive real-time images 4, 7. Given the possibility to identify various abdominal problems, it is a useful element of the diagnostic arsenal 4, 5. Ultrasound gives you the opportunity to dynamically evaluate the appendix, identifying such features as appendicoliths, thickened walls, and peri-appendiceal fluid 4, 7. This dynamic perception becomes of particular use when the symptoms are suspicious or when the distinction between AA and other issues of the abdominal region is difficult. The quality of images and interpretation can however rely on the skill of the operator and can be very inaccurate 10. Although the ultrasound is promising, there is an actual vacuum in the literature regarding the use and performance of ultrasound particularly in certain areas. The utility of current findings may depend on the local demographics, resources, and the prevalence of atypical presentations. The proposed research will address this gap by conducting an extensive research to determine the diagnostic accuracy of ultrasound in Hazara division of Khyber Pakhtunkhwa (KPK). We expect that by comparing the ultrasound outcomes with the naked-eye inspection by the surgeon during surgery, we will be able to provide region-specific, evidence-based information, that would benefit the care approach of patients suffering right lower quadrant pain and suspected AA. The outcomes of the findings to be expected may assist in streamlining the diagnostic strategies, reduction of costs, better outcomes, and less exposure to radiations.

MATERIALS AND METHODS

This study was conducted as a cross-sectional validation study in the Radiology Department of Ayub Teaching Hospital (ATH), Abbottabad and Lady Reading Hospital, Peshawar. Data collection took place following approval of the study protocols from ethical committee of Ayub Medical and Teaching Institution- Abbottabad (RC-EA-2023/204). Data collection spanned from January 2024 - September 2024 and in LRH from December 2024 - February 2025, for a total period of one year. At this time data collection as stopped as sample size was achieved.

A total of 139 patients presenting with right iliac fossa pain and a clinical suspicion of AA were enrolled through non-probability consecutive sampling. Both male and female patients aged 18 years and above were included. Patients who were managed conservatively without surgery were excluded from the study (25 cases of mesenteric adenitis mimicking AA and six patients of appendiceal mass). The sample size was calculated by formula for validation study using sensitivity and specificity of ultrasonography to diagnose AA to be 84.8% and 83.3 % and prevalence of AA in patients presenting with right iliac fossa pain to be 73.3% [11]. The confidence level was set at 95% and precision at 12% for estimation of sample size.

After obtaining informed consent, each participant's demographic and clinical details including age, gender, duration of pain, Alvarado score, and relevant clinical findings for ALVARADO score were recorded on a predesigned proforma. All enrolled patients underwent abdominal ultrasonography performed by a consultant radiologist with a minimum of five years of post-fellowship experience. A Toshiba Xario 100 ultrasound machine equipped with a linear high frequency probe (8-14 MHz) was used. The sonographic criteria for AA included a non-compressible, aperistaltic, blind ending tubular structure in the right iliac fossa with a diameter greater than 6 mm, wall thickness ≥ 3 mm, peri-appendiceal fluid, or the presence of an appendicolith.

Following ultrasound, decision to undergo surgery were made as per the departmental protocols of the gynecology department and surgery department, where the patient was referred appropriately to the clinical situation. The intra-operative findings served as the gold standard for diagnosis, as mentioned in the patient operation theater notes by a consultant surgeon/ gynecologist. Patients were followed in their respective wards for operative findings mentioned on their files to obtain a per-operative diagnosis of appendicitis or alternative diagnosis of appendicitis mimicker. Features such as a macroscopically inflamed appendix, presence of pus or excess peritoneal fluid, or omental adhesions were taken as confirmatory evidence of appendicitis.

Data were entered into SPSS version 27.0 for analysis. Descriptive statistics were generated for demographic and clinical variables. Diagnostic accuracy of ultrasound was calculated by comparing its findings with intraoperative results. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall diagnostic accuracy were determined using standard formulae from a 2×2 contingency table. Effect modifiers such as age, gender, and duration of pain were explored through stratification, and chi-square or Fisher's exact test was applied where appropriate, with a significance level set at 0.05.

RESULTS

This study enrolled 139 participants after evaluation of 145 patients. Six participants, who were evaluated but not included in study because of conservative management for appendiceal mass. In these cases, per-operative findings could not be obtained to confirm or rule out the diagnosis of AA.

The table 1 shows the socio-demographic and clinical characteristics of the study participants. The average age of the subjects was 29 years, and the number of males and females was nearly the same (58.3 vs. 41.7). The mean length of pain during the time of presentation in ultrasound was approximately 29 hours and the median Alvarado score was 7 indicating that the majority of the patients were at moderate-to-high risk of appendicitis. Almost three quarters of patients were diagnosed with appendicitis with the majority being uncomplicated (46%). It is worth noting that approximately one-fourth (26.6%) were provided with other diagnoses and gynecological problems were the primary mimickers. This brings out the diagnostic problem of right iliac fossa pain in surgery.

Table 1: Baseline demographic and clinical features of patients with suspected appendicitis.

Characteristics	**Mean ± Standard Deviation**	**Median (IQR)**
Age (Years)	28.14 ± 9.89	27 (21 - 35)
Duration of pain (Hours)	29.22 ± 11.79	29 (21 – 37)
Alvarado Score	6.81 ± 1.98	7 (6 – 8)
Characteristics	Categories	Frequency (Percentage)
Gender	• Male • Female	• 81/139 (58.3 %) • 58/139 (41.7%)
Per-operative diagnoses	• Uncomplicated Appendicitis • Gangrenous Appendicitis • Perforated Appendicitis • Appendicitis Mimicker	• 64/139 (46%) • 13/139 (9.4%) • 25/139 (18%) • 37/139 (26.6%)
Appendicitis Mimicker	• Gynecological issues • Meckel's diverticulum • Mesenteric adenitis with intussusception • Gut perforation (Other than appendiceal)	• 16/139 (11.5) • 4/139 (2.9%) • 8/139 (5.8%) • 9/139 (6.5%)

Ultrasound properly detected appendicitis in 87 patients (true positives) and ruled it out in 31 patients (true negatives) as indicated in Table 2 and shows a high conformity with operative findings ($p < 0.001$). It was accurate in both genders with a little bit higher performance in the female gender. There were striking differences when stratified by Alvarado score. There were no false negatives in the patients with Alvarado <7 and ultrasound were able to detect all the appendicitis cases (14 true positives, 28 true negatives). Conversely, in individuals with Alvarado, ≥7, ultrasound showed 73 true positives and 15 false negatives, which highlights the more difficult diagnosis in this clinically high-risk group despite an exceedingly low false positive (1 case). The above pattern illustrates that ultrasound is most reliable in the cases when there is a low level of suspicion that the patient has appendicitis and in high-suspicion cases where ultrasound confirms the disease in most cases, but a few cases may be missed.

Table 2: Diagnostic performance of ultrasound stratified by gender and Alvarado score.

		Per-operative Diagnosis of Appendicitis		Total	P value
		Yes	No		
Diagnosis of Appendicitis on Ultrasound (DAU)	Yes	87	6	93	<0.001
	No	15	31	46	
Gender wise DAU- Male	Yes	48	4	52	<0.001
	No	9	20	29	
Gender wise DAU- Female	Yes	39	2	41	<0.001
	No	6	11	17	
Alvarado score wise DAU- Score < 7	Yes	14	5	19	<0.001
	No	0	28	28	
Alvarado score wise DAU- Score ≥ 7	Yes	73	1	74	0.023
	No	15	3	18	

Table 3: Sensitivity, specificity, and predictive values of ultrasound across subgroups

Group	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
Overall	85.29	83.78	93.55	67.39	84.89
Male	84.21	83.33	92.31	68.97	83.95
Female	86.67	84.62	95.12	64.71	86.21
Alvarado <7	100.00	84.85	73.68	100.00	89.36
Alvarado ≥7	82.95	75.00	98.65	16.67	82.61

As shown in the table 3, ultrasound showed a high sensitivity 85%, specificity 84%, and accuracy 84.89% do diagnose appendicitis in patients with clinically suspicion ($p < 0.001$). Similarly, ultrasound showed a high diagnostic validity in patients of both genders. For patients with Alvarado score ≥ 7, PPV approached 99%, indicating that a positive ultrasound in high-suspicion cases almost always confirmed appendicitis. Conversely, in patients with Alvarado score < 7, sensitivity and NPV approached 100%, reinforcing ultrasound's utility in safely excluding appendicitis when suspicion is low.

Table 4: Association of appendicitis diagnosis with sex and Alvarado score.

		Diagnosis of Appendicitis		Total	P value
		Yes	No		
Gender	Female	45	13	58	0.343
	Male	57	24	81	
Alvarado Score	Score < 7	14	33	47	<0.001
	Score ≥ 7	88	4	92	

As shown in the table 4, gender was not significantly associated with appendicitis (p = 0.343) as it occurs almost equally in both genders. However, Alvarado score ≥ 7 showed a strong correlation with a diagnosis of AA (p < 0.001). Patients with scores ≥ 7 were overwhelmingly more likely to have appendicitis 88/ 139 (63.31%), confirming the clinical value of this scoring system as a predictor of disease.

DISCUSSION:

In this cross-sectional validation study, ultrasound demonstrated a sensitivity of 85.3%, specificity of 83.8%, PPV of 93.6%, NPV of 67.4%, and overall accuracy of 84.9% for diagnosing AA. These findings are strongly aligned with per-operative outcomes ($p < 0.001$), with 87 true positives and 31 true negatives against only 6 false positives and 15 false negatives. The results confirm ultrasound's reliability as a first-line imaging modality for suspected appendicitis in resource-limited emergency settings.

Our findings are consistent with published evidence. A study conducted in Saudi Arabia reported the sensitivity and specificity of ultrasound in the diagnosis of AA to be 36% and 83% respectively in children younger than ten years and 52% sensitivity and 100% specificity in adults [8]. A meta-analysis of 2841 participants, the sensitivity and specificity of ultrasound for the diagnosis of AA was reported to be 69% and 81% respectively [10]. Another recent meta-analysis reported the sensitivity and specificity of ultrasound for the diagnosis of AA to be 77.2% and 60% respectively [9]. Another study of 180 patients undergoing appendectomy, the sensitivity and specificity of AA reported were 84.8% and 83.3% respectively. The prevalence of AA in patients presenting with right inferior fossa pain in that study was 73.3% [11]. The close similarity between these published values and our observed sensitivity (85.3%) and specificity (83.8%) strengthens the external validity of our findings.

Our sensitivity (85.3%) and specificity (83.8%) align well with the ranges reported across systematic reviews, where sensitivity varies between 70–97.9% and specificity from 33–100% [9, 10]. Roberts et al. (2023) reported much higher values in pediatric cohorts, often exceeding 96%, highlighting that US performs optimally in children, whereas our adult cohort produced moderately lower values, consistent with global adult data [12]. Farooq et al. (2020) reported lower performance (sensitivity 80%, specificity 60%), and Vatsa and Pathak (2022) noted false negatives of 11%, comparable to our 10.8% FN rate [13, 14]. Shah et al. (2011) demonstrated improved performance when ultrasound was combined with the Alvarado score (sensitivity 97.9%, specificity 75%), a pattern echoed in our subgroup analysis, with 100% sensitivity and NPV in patients with Alvarado <7, and 98.7% PPV in patients with Alvarado ≥7 [15].

Our study also rule out appendicitis in mimickers such as mesenteric lymphadenitis with intussusception, gynecological causes, gut perforation and Meckel's diverticulitis, demonstrating ultrasound's role in reducing negative appendectomies. Gender-stratified analysis showed comparable diagnostic accuracy in males and females, with slightly higher PPV in females (95.1% vs. 92.3%). Although gynecological pathology can mimic appendicitis and increase false positives, our results (only 6 total FP cases) indicate reliable sonographic discrimination in both sexes.

The Alvarado score indeed indicates the presence of differences when it comes to clinical risk. Ultrasound nailed had a 100% sensitivity and NPV for patients with an Alvarado <7, which is that it correctly excluded

all who did not have appendicitis. However, in the case of those whose Alvarado 7 or higher positive, the ultrasound possessed a PPV of approximately 99, thus a positive scan in a high-risk man virtually confirms appendicitis. Nevertheless, we also had few false negatives (15 cases) and we must remain especially vigilant on high-suspicion patients. In fact, these data support us in applying ultrasound to the exclusion of disease in low-risk individuals and to confirm in high-risk individuals, so that they are consistent with the integrated imaging-clinical pathways in recent papers 7, 9, 16. Our intraoperative profile of 63% uncomplicated, 25% perforated, 13% gangrenous is similar to international reports that indicate 20-30% of patients present with complicated appendicitis at surgery 6, 9. In addition to this, the non-appendicitis mimickers, mesenteric lymphadenitis with intussusception, gynecological issues, Meckel and gut perforation have been observed in 26.6% of the cases, therefore demonstrating the value of ultrasound to differentiate other causes of right iliac fossa pain 4. The guidelines of the World Society of Emergency Surgery (WSES) 2020 discuss ultrasound as a firm first-line method, largely since it does not use radiation as opposed to CT 7. This suggestion is supported in our research in the Pakistani environment where ultrasound is unexpectedly available and affordable. Ultrasound dependency on operators and possible verification bias was the primary limitations as we did not include the cases that were conservatively managed. These are rather common problems that have been observed in previous systematic reviews 9, 10.

Clinical implications

Our findings put the US-first diagnostic pathway in consideration:

Patients with low risk (Alvarado <7): A negative US excludes appendicitis, and we can afford to treat the patient conservatively and obtain no unnecessary CT scans.

High-risk (Alvarado 7 or more): A positive US is basically a confirmation of appendicitis and early surgery should be involved. In case of negative or inconclusive scan the next US or CT should be pursued to rule out perforation.

Mimickers surgical: The identifications of the other diagnoses reduce negative appendectomies, patient outcome, and utilization of operating resources.

Conclusion

The ultrasound is a valid and safe diagnostic tool that can be used to examine suspected appendicitis with high accuracy irrespective of gender and risk group. It is particularly effective in the exclusion of appendicitis among low-risk patients as well as the confirmation of its presence among high-risk patients. These results correspond with the world literature, and the use of ultrasound as the center of diagnostic algorithms in the locations where we desire to reduce radiations and save resources is supported.

REFERENCES:

1. Raffa A, Abduljabbar A, Alharthy A. Comparison of ultrasound and computed tomography scanning accuracy in diagnosing acute appendicitis at King Abdulaziz university hospital. Cureus. 2022 Nov 25;14(11). https://doi.org/10.7759/cureus.31880
2. del Olmo JC, López JR, Esteban ML, Cutoli PC, Martín MA, Puerta CV, Mestanza CL. Clinical Prediction Rules in acute appendicitis: Which combination of variables is more effective at predicting? (2020). https://doi.org/10.21203/rs.3.rs-22581/v1
3. Guan L, Liu Z, Pan G, Zhang B, Wu Y, Gan T, Ouyang G. The global, regional, and national burden of appendicitis in 204 countries and territories, 1990–2019: a systematic analysis from the Global Burden of Disease Study 2019. BMC gastroenterology. 2023 Feb 22;23(1):44. https://doi.org/10.1186/s12876-023-02678-7
4. Echevarria S, Rauf F, Hussain N, Zaka H, Ahsan N, Broomfield A, Akbar A, Khawaja UA, Farwa UE. Typical and atypical presentations of appendicitis and their implications for diagnosis and treatment: a literature review. Cureus. 2023 Apr 2;15(4). https://doi.org/10.7759/cureus.37024

5. Amnedya TT, Wijaya S. Comparison of Sensitivity and Specificity of Alvarado Score and Tzanakis Score in Enforcement of Acute Appendicitis Diagnosis. Journal of Social Science. 2023;4(3):739-44. https://doi.org/10.46799/jss.v4i3.583
6. Torun M, Subaşı İE, Özbay DK, Özbay MA, Özdemir H. Utilizing non-invasive biomarkers for early and accurate differentiation of uncomplicated and complicated acute appendicitis: a retrospective cohort analysis. Scientific Reports. 2025 Feb 20;15(1):6177. https://doi.org/10.1038/s41598-025-90591-2
7. 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 Apr 15;15(1):27. https://doi.org/10.1186/s13017-020-00306-3
8. Alelyani M, Hadadi I, Shubayr N, Alashban Y, Alqahtani M, Adam M, Almater H, Alamri S. Evaluation of ultrasound accuracy in acute appendicitis diagnosis. Applied Sciences. 2021 Mar 17;11(6):2682. https://doi.org/10.3390/app11062682
9. Fu J, Zhou X, Chen L, Lu S. Abdominal ultrasound and its diagnostic accuracy in diagnosing acute appendicitis: a meta-analysis. Frontiers in Surgery. 2021 Jun 28;8:707160. https://doi.org/10.3389/fsurg.2021.707160
10. Giljaca V, Nadarevic T, Poropat G, Nadarevic VS, Stimac D. Diagnostic accuracy of abdominal ultrasound for diagnosis of acute appendicitis: systematic review and meta-analysis. World journal of surgery. 2017 Mar;41(3):693-700. https://doi.org/10.1007/s00268-016-3792-7
11. Al-Ajerami Y. Sensitivity and specificity of ultrasound in the diagnosis of acute appendicitis. Eastern Mediterranean Health Journal. 2012;18(1):66.
12. Roberts K, Moore H, Raju M, Gent R, Piotto L, Taranath A, Ee M, Linke R, Goh DW. Diagnostic ultrasound for acute appendicitis: the gold standard. Journal of Pediatric Surgery. 2024 Feb 1;59(2):235-9. https://doi.org/10.1016/j.jpedsurg.2023.10.028
13. Farooq A, Zameer S, Khadim R. Diagnostic accuracy of ultrasound in acute appendicitis in comparison with alvarado score keeping histopathological correlation as gold standard. Pakistan Armed Forces Medical Journal. 2020 Jun 30(3):807.
14. Vatsa A, Pathak S. ROLE OF ULTRASOUND IN DIAGNOSIS OF ACUTE APPENDICITIS: SHOULD YOU 'SOUND'BEFORE YOU 'SCOPE'?. Int J Acad Med Pharm. 2022;4(5):817-23. DOI: 10.47009/jamp.2022.4.5.170
15 Shah NA, Islam M, Sabir IA, Mehreen T, Khan M. Diagnostic accuracy of clinical examination versus combination of abdominal ultrasound and Alvarado score, in patients with acute appendicitis. Journal of Postgraduate Medical Institute. 2008;22(1).
16. Tayebi A, Olamaeian F, Mostafavi K, Khosravi K, Tizmaghz A, Bahardoust M, Zakaryaei A, Mehr DE. Assessment of Alvarado criteria, ultrasound, CRP, and their combination in patients with suspected acute appendicitis: a single centre study. BMC gastroenterology. 2024 Jul 31;24(1):243. https://doi.org/10.1186/s12876-024-03333-5