MHA MEDIHEALTH ACADEMY

Surgery on Children

Original Article

doi:10.51271/SOC-0076

Diagnostic performance of the Alvarado score for identifying low-risk pediatric patients with suspected acute appendicitis

Şenay Kurtuluş

Department of Pediatric Surgery, Faculty of Medicine, Çanakkale Onsekiz Mart University, Çanakkale, Turkiye

Received: 08/06/2026 ◆ **Accepted**: 15/07/2026 ◆ **Published**: 29/07/2026

Cite this article: Kurtuluş Ş. Diagnostic performance of the Alvarado score for identifying low-risk pediatric patients with suspected acute appendicitis. *Surg Child.* 2026;3(3):82-87. doi:10.51271/SOC-0076

Corresponding Author: Şenay Kurtuluş, senaykurtulus@yahoo.com

ABSTRACT

Aims: Acute appendicitis (AA) remains a major diagnostic challenge in children because of its variable clinical presentation and overlap with other causes of abdominal pain. This study aimed to evaluate the diagnostic performance of the Alvarado score (AS) in children with suspected AA.

Methods: This retrospective diagnostic accuracy study included 737 pediatric patients presenting with abdominal pain between January 2019 and December 2023. Patients were classified into the AA and nonspecific abdominal pain (NAP) groups. The diagnostic performance of the AS was assessed using receiver operating characteristic (ROC) analysis. Clinically relevant rule-out and rule-in thresholds based on previously published literature were evaluated.

Results: A total of 104 patients were diagnosed with AA and 633 with NAP. The median AS was significantly higher in the AA group than in the NAP group (8 [IQR 7-8] vs 3 [IQR 2-4]; p<0.001). ROC analysis demonstrated excellent diagnostic performance of the AS, with an area under the curve (AUC) of 0.979 (95% CI 0.961-0.996; p<0.001). An AS threshold of ≤4 demonstrated a sensitivity of 98.1% and a negative predictive value (NPV) of 99.6%, whereas a threshold of ≥6 demonstrated a specificity of 96.5% and a positive likelihood ratio of 27.2.

Conclusion: The AS demonstrated excellent diagnostic performance and may serve as a useful adjunctive tool for the initial risk stratification of children with suspected AA. In particular, an AS ≤4 was associated with a very low probability of AA and may help identify selected patients who could potentially avoid immediate imaging. Prospective multicenter studies are needed to validate these findings.

Keywords: Appendicitis, child, emergency service, diagnostic imaging, decision support techniques

INTRODUCTION

Acute abdominal pain is one of the most common reasons for pediatric emergency department visits and may result from self-limiting conditions or from acute appendicitis (AA). The diagnosis of AA in children remains challenging because of overlapping clinical features and variable presentation. Inadequate evaluation may lead to negative appendectomy, whereas delayed diagnosis increases the risk of complications.[1,2]

Several clinical scoring systems have been developed to assist in identifying children at risk for AA. These tools are considered sufficiently sensitive to exclude AA in low-risk patients, thereby reducing unnecessary imaging and negative appendectomy rates.[3,4] However, the routine use of abdominal and pelvic computed tomography (CT) in children remains controversial because increased CT utilization has not consistently reduced negative appendectomy rates, while concerns regarding ionizing radiation exposure persist.[5] Ultrasonography (US) is generally recommended as the preferred first-line imaging modality in children because it avoids radiation exposure. Nevertheless, its diagnostic performance may vary depending on operator experience, patient characteristics, and appendix visualisation rates.[6,7]

The Alvarado score (AS) is a simple, widely used clinical tool that incorporates symptoms, physical examination findings, and laboratory parameters.[8,9] Despite its practicality, the diagnostic accuracy of the AS in pediatric populations remains variable across studies. A meta-analysis by Bai et al.[10] reported pooled sensitivity and specificity values of 76% and 71%, respectively, suggesting moderate overall diagnostic accuracy and supporting its role primarily as an adjunctive diagnostic tool.

Therefore, this study aimed to evaluate the diagnostic performance of the AS in a pediatric population and to investigate clinically applicable cut-off values for risk stratification in children with suspected AA.

METHODS

Ethics

The study protocol was approved by the Ethics Committee for Non-interventional Clinical Researches at Çanakkale Onsekiz Mart University (Date: 10.12.2025, Decision No: 2025-382) and conducted in accordance with the Declaration of Helsinki.

Study Design and Setting

This retrospective diagnostic accuracy study was conducted at the pediatric emergency and pediatric surgery department between January 2019 and December 2023.

Patients

A total of 737 children were included in the final analysis (**Figure 1**). Children aged 0-17 years presenting with acute abdominal pain and evaluated for suspected AA were eligible. Patients with incomplete clinical or laboratory data, incomplete medical records, or an alternative acute abdominal diagnosis were excluded from the study. Because this was a retrospective diagnostic accuracy study, all consecutive eligible patients presenting during the predefined study period were included in the analysis. Therefore, no formal a priori sample size calculation was performed.

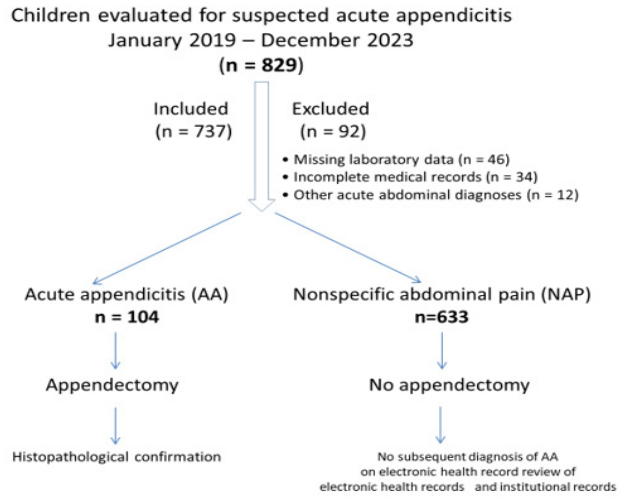

Figure 1. STARD flow diagram of patient selection

STARD flow diagram illustrating patient selection, reasons for exclusion, final study groups, and the reference standards used for diagnosis. AA: Acute appendicitis, NAP: Nonspecific abdominal pain

The patients were then divided into two groups: nonspecific abdominal pain (NAP) (n=633) and AA (n=104). AA was histopathologically confirmed in patients with appendectomy. Negative appendectomy was defined by postoperative histopathologic examination as the absence of inflammation in the appendix wall or polymorphonuclear leukocytes, including normal appendix and lymphoid hyperplasia. Non-operated patients were classified as NAP if they improved without surgery and had no subsequent diagnosis of AA based on review of the available electronic health records and institutional medical records.

The patient data analysed included age, gender, presence of migratory pain, anorexia, nausea/vomiting, right lower quadrant (RLQ) tenderness, and rebound pain. Body temperature, white blood cell count (WBC), neutrophil and lymphocyte percentages, neutrophil-to-lymphocyte ratio (NLR), C-reactive protein (CRP) levels, and US findings were also analysed. AS was calculated according to the Alvarado scoring system and is summarised in **Table 1**.

Table 1. The Alvarado scoring system (ASS)

Component		Score
Clinical symptoms	Migration of pain to the RLQ	1
	Anorexia	1
	Nausea/vomiting	1
Physical examination	RLQ tenderness	2
	Rebound tenderness	1
	Fever (>37.5 °C)	1
Laboratory findings	Leukocytosis (WBC >10×10^9/L)	2
	Elevated neutrophil count >75%	1
Total		10

RLQ: Right lower quadrant, WBC: White blood count

US Evaluation

US reports were retrospectively reviewed and categorised as positive or negative based on the documented radiological findings. A normal appendix was defined as a compressible, blind-ended tubular structure with a maximal outer diameter of <6 mm, displaying preserved wall stratification and no evidence of hyperemia or surrounding inflammatory changes. An inflamed appendix was defined as a non-compressible tubular structure with a maximal outer diameter ≥6 mm, with or without the presence of an appendicolith. Secondary sonographic signs of appendicitis included increased echogenicity of the periappendiceal fat, localised free fluid, phlegmon, abscess formation, or an inflammatory mass.[11-13]

US was not performed routinely in all patients. The decision to perform US was made at the discretion of the treating pediatric surgeon or emergency physician according to the patient's clinical presentation and routine clinical practice. All US examinations were performed by board-certified general radiologists as part of routine clinical care.

A US was classified as negative when a normal appendix was clearly visualised without sonographic features of inflammation, or when the appendix was not visualised. No secondary signs suggestive of appendicitis were documented.[14-16] US was considered positive when either an inflamed appendix was directly visualised or when secondary inflammatory findings were present, regardless of direct visualisation of the appendix.

Statistical Analysis

The data were analysed using the Statistical Package for the Social Sciences (SPSS) 18.0 software (SPSS Inc., Chicago, IL). The normality of continuous variables was assessed using the Shapiro–Wilk test. As the variables were not normally distributed ($p<0.05$), continuous variables were expressed as median and interquartile range (IQR), along with minimum

and maximum values.The Mann–Whitney U-test was used for non-normally distributed variables. Categorical variables were analysed using the chi-square test. A two-tailed p-value of <0.05 was considered statistically significant.

The diagnostic performance of the AS was evaluated using cutoff values of ≤4 and ≥6, selected based on previously published literature evaluating rule-out and rule-in thresholds in pediatric appendicitis. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), overall accuracy, and likelihood ratios (LR+ and LR–) were calculated using cross-tabulation analysis. Receiver operating characteristic (ROC) curve analysis was conducted to assess the discriminatory performance of the AS, and the area under the curve (AUC) with 95% confidence intervals (CI) was reported.

RESULTS

A total of 737 patients were included, comprising 633 with NAP and 104 with AA. As shown in **Table 2**, patients with AA were significantly older (median 12 (IQR 9-14.75) vs 10 (IQR 7-13) years; p=0.001), and male gender was more common in this group (63.5% vs 49.1%; p=0.007).

Table 2. Demographic characteristics of the study population

Variable	NAP (n=633)	AA (n=104)	p value
Age, years (median, IQR)	10 (7-13)	12 (9-14.75)	0.001†
Male gender, n (%)	311 (49.1%)	66 (63.5%)	0.007‡

NAP: Nonspecific abdominal pain, AA: Acute appendicitis, IQR: Interquartile range, † Mann–Whitney U test, ‡ Pearson Chi-square test

Laboratory parameters and clinical scores demonstrated marked intergroup differences (**Table 3**). The median AS was significantly higher in the AA group (8 (IQR 7-8)) than in the NAP group (3 (IQR 2-4); p<0.001). Leukocyte count, neutrophil percentage, NLR, and CRP levels were significantly elevated in AA (all p<0.001), whereas lymphocyte percentage was significantly lower (p<0.001).

Table 3. Comparison of laboratory parameters and Alvarado scores between the NAP and AA groups

Variable	NAP (n=633) median (IQR)	AA (n=104) median (IQR)	p value
Alvarado score (AS)	3 (2-4)	8 (7-8)	<0.001
WBC (×10^9/L)	10.03 (7.8-12.5)	14.59 (12.1-17.4)	<0.001
Neutrophil (%)	68.5 (55.2-80.1)	81.3 (70.2-87.4)	<0.001
Lymphocyte (%)	21.8 (12.4-33.7)	11.0 (6.9-18.6)	<0.001
NLR	3.13 (1.64-6.57)	7.16 (4.08-12.49)	<0.001
CRP (mg/dl)	0.36 (0.10-1.50)	3.97 (1.31-10.14)	<0.001

NAP: Nonspecific abdominal pain, AA: Acute appendicitis, IQR: Interquartile range, WBC: White blood cell count, NLR: Neutrophil-to-lymphocyte ratio, CRP: C-reactive protein. Values are presented as median (25th–75th percentile [IQR]).

ROC curve analysis demonstrated high diagnostic performance of the AS, with an AUC of 0.979 (95% CI 0.961-0.996; p<0.001) (**Figure 2**).

Using previously reported clinically relevant thresholds, an AS value of ≤4 demonstrated a sensitivity of 98.1% and NPV of 99.6%. Given the relatively low prevalence of AA in the study cohort, the observed NPV should be interpreted in the context of disease prevalence. In contrast, a threshold of ≥6 demonstrated a specificity of 96.5% and a positive likelihood ratio (LR+) of 27.2. Detailed diagnostic performance metrics, including sensitivity, specificity, positive and NPVs, likelihood ratios (LR+ and LR–), and overall accuracy, are presented in **Supplementary Table S1**.

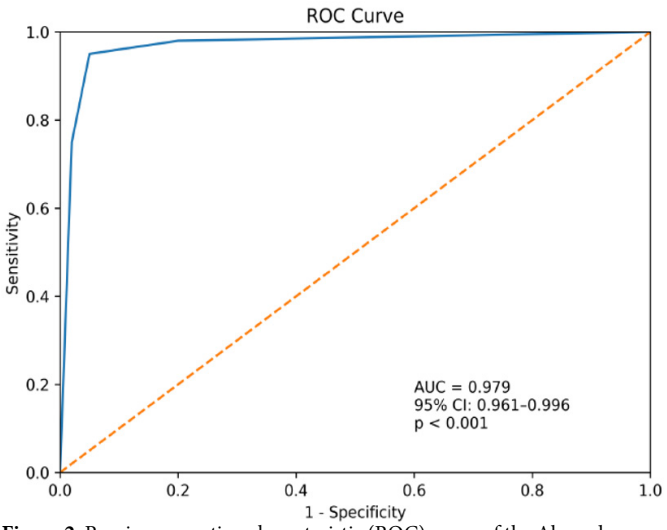

Figure 2. Receiver operating characteristic (ROC) curve of the Alvarado score (AS)

The ROC analysis demonstrated high diagnostic performance of the AS, with an area under the curve (AUC) of 0.979 (95% CI 0.961–0.996; p<0.001). The dashed diagonal line represents the line of no discrimination.

Among patients who underwent US (n=240), imaging demonstrated a sensitivity of 91.1% (95% CI 83.2-96.1) and a specificity of 38.5% (95% CI 13.9-68.4).

DISCUSSION

In the present study, the AS demonstrated high diagnostic performance for identifying AA in children, with an AUC of 0.979. An AS threshold of ≤4 was associated with high sensitivity and a very high NPV, whereas a threshold of ≥6 demonstrated high specificity and a strong positive likelihood ratio. These findings support the potential utility of the AS as a risk stratification tool in the initial evaluation of children with suspected AA. The excellent diagnostic performance observed in our study was higher than that reported in most previous pediatric validation studies. This discrepancy may be explained by differences in patient selection, disease spectrum, and study design. In addition, the retrospective nature of the study, selective patient inclusion, and the use of different reference standards for operated and non-operated patients may have introduced selection bias, spectrum bias, and partial verification bias, potentially contributing to the observed diagnostic performance. Therefore, the reported diagnostic accuracy should be interpreted with appropriate caution until validated in prospective multicenter cohorts.

Previous studies have reported variable diagnostic performance of the AS in children. Reported sensitivities of approximately 72% and specificities of 79% suggest that an AS ≥6 alone may be insufficient for definitive diagnosis of AA, whereas lower scores may help identify patients at lower

Supplementary Table S1. Diagnostic performance of the Alvarado score at clinically relevant rule-in and rule-out thresholds

Test	Cut-off/ definition	Sensitivity % (95% CI)	Specificity % (95% CI)	LR+	LR–	PPV % (95% CI)	NPV % (95% CI)	Accuracy % (95% CI)	AUC (95% CI)
AS	≥6 (rule-in)	95.2 (89.1–98.4)	96.5 (94.8–97.8)	27.2	0.05	81.8 (73.8–88.2)	99.2 (98.1–99.7)	96.3 (94.7–97.6)	0.979 (0.967–0.991)
AS	≤4 (rule-out)	98.1 (93.2–99.8)	78.7 (75.3–81.8)	4.61	0.02	43.0 (36.6–49.6)	99.6 (98.6–100.0)	81.4 (78.4–84.2)	0.979 (0.967–0.991)

AS: Alvarado score, LR+: Positive likelihood ratio, LR–: Negative likelihood ratio, PPV: Positive predictive value, NPV: Negative predictive value, CI: Confidence interval, AUC: Area under the curve. The 95% confidence intervals were calculated using the exact (Clopper–Pearson) binomial method.

risk.[4,17] In addition, a comprehensive meta-analysis evaluating the diagnostic value of the AS in children concluded that the score demonstrates moderate overall accuracy and should primarily be considered an adjunctive diagnostic tool rather than a standalone diagnostic method.[10]

Although other clinical scoring systems, such as the Pediatric Appendicitis Score (PAS) and the Appendicitis Inflammatory Response (AIR) score, have also demonstrated good diagnostic performance in children, we focused on the AS because it remains one of the most widely used and externally validated clinical prediction tools in both pediatric and adult populations. The PAS was specifically developed for pediatric patients, whereas the AIR score incorporates inflammatory biomarkers and may provide improved specificity in selected populations.[17,18,26] However, comparative studies have generally shown similar overall diagnostic performance among these scoring systems, and no single score has consistently demonstrated clear superiority across different clinical settings. Therefore, the AS remains a practical and clinically applicable tool for the initial evaluation of children with suspected AA.

Compared with these reports, the diagnostic performance observed in the present study was notably higher. An AS threshold of ≤4 was associated with high sensitivity (98.1%) and a high NPV (99.6%), suggesting that low AS values may help identify children with a low probability of AA. However, the high NPV should be interpreted in the context of the relatively low prevalence of AA in the study population, since predictive values are prevalence-dependent. Therefore, the diagnostic performance observed in this study may not be directly generalizable to populations with different disease prevalence. Similarly, a threshold of ≥6 demonstrated high specificity and a strong positive likelihood ratio, supporting its potential utility in identifying higher-risk patients.[18]

Careful physical examination remains a cornerstone of the evaluation of children with suspected AA. Previous studies have emphasized that inadequate physical examination may contribute to unnecessary imaging and radiation exposure.[19] The high diagnostic performance of the AS observed in the present study likely reflects the integration of key clinical findings, physical examination, and laboratory parameters into a structured risk stratification tool. Taken together, these findings suggest that structured clinical assessment remains an important component of pediatric appendicitis evaluation alongside laboratory and imaging findings.

Laboratory parameters also served as important adjunctive findings in our cohort. Leukocytosis, neutrophil predominance, and elevated NLR were significantly associated with AA. Previous studies have similarly reported moderate-to-high diagnostic performance for NLR in pediatric appendicitis.[20]

US remains the preferred first-line imaging modality in children with suspected AA because it avoids ionizing radiation exposure. However, its diagnostic performance may vary depending on operator experience, patient characteristics, and the rate of appendix visualisation.[16,21] In our study, only a subset of patients underwent US evaluation, most likely reflecting clinician-directed selection of patients with higher clinical suspicion. Consequently, the observed diagnostic performance of US should be interpreted cautiously because substantial selection and work-up bias may have been present. Furthermore, the relatively low specificity observed in our cohort may have been influenced by our retrospective classification strategy, whereby examinations with non-visualization of the appendix and no secondary inflammatory findings were categorized as negative. Because non-visualization of the appendix is generally considered a non-diagnostic rather than a negative US finding in pediatric practice, this approach may have underestimated the specificity of US. Together with the operator-dependent nature of US, these methodological considerations likely contributed to the observed diagnostic performance and should be taken into account when comparing our findings with previous pediatric studies.[17,22] Therefore, the diagnostic performance of US reported in this study should not be interpreted as the true diagnostic accuracy of US in children with suspected AA. Rather, it reflects the performance of US within the context of our retrospective study design and selective imaging strategy.

Previous studies have reported that CT does not improve diagnostic accuracy beyond that achieved through history, physical examination, and laboratory findings alone. They emphasised that a negative CT scan does not reliably exclude AA when clinical suspicion is strong.[23,24] At the same time, increasing CT utilization in pediatric populations has not consistently reduced negative appendectomy rates.[5,25] Given the potential long-term risks associated with ionizing radiation exposure in children, selective and clinically guided imaging strategies remain important.

Overall, our findings support the use of the AS as an adjunctive tool for the initial risk stratification of children with suspected AA. However, given the retrospective design and the potential for partial verification bias, these findings should be interpreted with caution.

Limitations

This study has several strengths. First, it included a relatively large pediatric cohort evaluated using routinely collected clinical and laboratory parameters. Second, the diagnostic

performance of the AS was comprehensively assessed using ROC analysis and clinically relevant rule-out and rule-in thresholds. Third, the study evaluated practical cut-off values that may assist in risk stratification during the initial assessment of children with suspected appendicitis.

However, several limitations should be acknowledged. First, the retrospective single-center design may have introduced selection bias and limits the generalizability of the findings. Second, the use of different reference standards for operated and non-operated patients may have introduced partial verification (work-up) bias. Whereas AA was confirmed histopathologically in patients undergoing appendectomy, non-operated patients were classified based on clinical follow-up and review of medical records. Consequently, the true diagnostic accuracy of the AS cannot be determined with complete certainty, and the reported diagnostic performance should be interpreted in light of this limitation. Third, US was performed selectively rather than systematically in all patients, introducing potential work-up and selection bias in the imaging subgroup analysis. Fourth, the duration of abdominal pain before presentation was not consistently documented because of the retrospective study design and therefore could not be evaluated. Fifth, the relatively low prevalence of AA in the study population may have inflated the NPV, limiting the generalizability of this finding to populations with different disease prevalence.

Finally, the findings have not been externally validated. Prospective multicenter studies with external validation are warranted to confirm the generalizability and clinical applicability of these results.

CONCLUSION

In this retrospective diagnostic accuracy study, the AS demonstrated high diagnostic performance and may serve as a useful adjunctive tool for the initial risk stratification of children with suspected AA. In particular, an AS of ≤4 was associated with a very low probability of AA and may help identify selected low-risk patients in whom immediate imaging could potentially be deferred. However, given the retrospective design and the use of different reference standards for operated and non-operated patients, these findings should be interpreted with caution. Imaging decisions should remain individualized and integrated with clinical assessment, laboratory findings, and physician judgment. Further prospective multicenter studies are required to confirm these findings before widespread clinical implementation.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study protocol was approved by the Ethics Committee for Non-interventional Clinical Researches at Çanakkale Onsekiz Mart University (Date: 10.12.2025, Decision No: 2025-382).

Informed Consent

This retrospective study used pre-existing anonymized patient data. No additional intervention was performed, and there was no direct patient contact. The study was approved by the Ethics Committee, and the requirement for written informed consent was waived by the ethics committee.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The author declare no conflicts of interest related to this study.

Financial Disclosure

The author received no financial support for the conduct or publication of this research.

Author Contributions

The author is solely responsible for the entirety of conception, execution, analysis, and writing of the manuscript.

REFERENCES

1. Nasiri S, Mohebbi F, Sodagari N, et al. Diagnostic values of ultrasound and the modified Alvarado scoring system in acute appendicitis. *Int J Emerg Med*. 2012;5(1):26. doi:10.1186/1865-1380-5-26
2. Kırkıl C, Karabulut K, Aygen E, et al. Appendicitis scores may be useful in reducing the costs of treatment for right lower quadrant pain. *Ulus Travma Acil Cerrahi Derg*. 2013;19:13-19. doi:10.5505/tjtes.2013.88714.
3. Neilson IR, Laberge JM, Nguyen LT, et al. Appendicitis in children: current therapeutic recommendations. *J Pediatr Surg*. 1990;25(11):1113-1116. doi:10.1016/0022-3468(90)90742-r
4. Podda M, Pisanu A, Sartelli M, et al. Diagnosis of acute appendicitis based on clinical scores: is it a myth or reality? *Acta Biomed*. 2021;92: e2021231. doi:10.23750/abm.v92i4.11666
5. Martin AE, Vollman D, Adler B, et al. CT scans may not reduce the negative appendectomy rate in children. *J Pediatr Surg*. 2004;39(6):886-890. doi:10.1016/j.jpedsurg.2004.02.034
6. Harel S, Mallon M, Langston J, Blutstein R, Kassutto Z, Gussow J. Factors contributing to nonvisualisation of the appendix on ultrasound in children with suspected appendicitis. *Pediatr Emerg Care*. 2022;38(2): e678-e682. doi:10.1097/PEC.0000000000002394
7. Bal A, Anil M, Nartürk M, et al. Importance of clinical decision making by experienced pediatric surgeons when children are suspected of having acute appendicitis: the reality in a high-volume pediatric emergency department. *Pediatr Emerg Care*. 2017;33(9):e38-e42. doi:10.1097/PEC.0000000000000763
8. Memon AZ, Irfan S, Fatima K, et al. Acute appendicitis: diagnostic accuracy of Alvarado scoring system. *Asian J Surg*. 2013;36(4):144-149. doi:10.1016/j.asjsur.2013.04.004
9. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med*. 1986;15(5):557-564. doi:10.1016/s0196-0644(86)80993-3
10. Bai S, Hu S, Zhang Y, et al. The value of the Alvarado score for the diagnosis of acute appendicitis in children: a systematic review and meta-analysis. *J Pediatr Surg*. 2023;58(10):1886-1892. doi:10.1016/j.jpedsurg.2023.02.060
11. Expert Panel on Pediatric Imaging:, Koberlein GC, Trout AT, et al. ACR Appropriateness Criteria® suspected appendicitis-child. *J Am Coll Radiol* 2019;16(5S)S252-S263 doi:10.1016/j.jacr.2019.02.022
12. Fallon SC, Kim ME, Hallmark CA, et al. Correlating surgical and pathological diagnoses in pediatric appendicitis. *J Pediatr Surg*. 2015; 50(4):638-641. doi:10.1016/j.jpedsurg.2014.11.001
13. Estey A, Poonai N, Lim R. Appendix not seen: the predictive value of secondary inflammatory sonographic signs. *Pediatr Emerg Care*. 2013; 29(4):435-439. doi:10.1097/PEC.0b013e318289e8d5
14. Puylaert JBCM. Acute appendicitis: US evaluation using graded compression. *Radiology*. 1986;158(2):355-360. doi:10.1148/radiology.158.2.2934762
15. Wiersma F, Srámek A, Holscher HC. US features of the normal appendix and surrounding area in children. *Radiology*. 2005;235(3):1018-1022. doi:10.1148/radiol.2353040086
16. Pernía J, Cancho T, Segovia I, et al. Predictive values of indirect ultrasound signs for low risk of acute appendicitis in paediatric patients without visualisation of the appendix on ultrasound. *Emerg Med J*. 2024;41(8):475-480. doi:10.1136/emermed-2023-213466

17. Gorter RR, Eker HH, Gorter-Stam MA, et al. Diagnosis and management of acute appendicitis. EAES consensus development conference 2015. *Surg Endosc.* 2016;30(11):4668-4690. doi:10.1007/s00464-016-5245-7
18. Noori IF, Jabbar AS, Nasser AA. 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):676-683. doi:10.1097/MS9.0000000000000270
19. Verghese A, Charlton B, Kassirer JP, Ramsey M, Ioannidis JPA. Inadequacies of physical examination as a cause of medical errors and adverse events: a collection of vignettes. *Am J Med.* 2015;128(12):1322-1324.e3. doi:10.1016/j.amjmed.2015.06.004
20. Khan SA, Ashraf R, Hassaan N, Naseer M, Azad MH, Javed H. The role of neutrophil-to-lymphocyte ratio in the diagnosis of acute appendicitis. *Cureus.* 2023;15(12):e51164. doi:10.7759/cureus.51164
21. Zouari M, Issaoui A, Hbaieb M, et al. Predictive factors of acute appendicitis in children with non-visualised appendix on ultrasound: a prospective cohort study. *Surg Infect (Larchmt).* 2024;25(1):26-31. doi:10.1089/sur.2023.295
22. Téoule P, Laffolie J, Rolle U, et al. Acute appendicitis in childhood and adulthood. *Dtsch Arztebl Int.* 2020;117(45):764-774. doi:10.3238/arztebl.2020.0764
23. El-Gohary Y, Molina M, Chang J, et al. The use of computed tomography versus clinical acumen in diagnosing appendicitis in children: a two-institution international study. *J Pediatr Surg.* 2021;56(8):1356-1361. doi:10.1016/j.jpedsurg.2020.09.061
24. Stephen AE, Segev DL, Ryan DP, et al. The diagnosis of acute appendicitis in a pediatric population: to CT or not to CT. *J Pediatr Surg.* 2003;38(3):367-371. doi:10.1053/jpsu.2003.50110
25. Russell WS, Schuh AM, Hill JG, et al. Clinical practice guidelines for pediatric appendicitis evaluation can decrease computed tomography utilization while maintaining diagnostic accuracy. *Pediatr Emerg Care.* 2013;29(5):568-573. doi:10.1097/PEC.0b013e31828e5718
26. Samuel M. Pediatric appendicitis score. *J Pediatr Surg.* 2002;37(6):877-881. doi:10.1053/jpsu.2002.32893