Prospective Validation of Alvarado Score and Pediatric Appendicitis Score for the Diagnosis of Acute Appendicitis in Children

Zenon Pogorelić, MD, PhD, Stella Rak, MD,† Ivana Mrklić, MD, PhD,‡ and Ivo Jurić, MD, PhD**

Objectives: The purpose of this study was to compare the results of the Alvarado and Pediatric Appendicitis Score (PAS) scoring systems and to establish which one is more reliable in setting the diagnosis of acute appendicitis in children.

Methods: All children operated on because of acute appendicitis from October 2011 to May 2013 were enrolled in this prospective study. Both clinical scoring systems have been compared over the same patients, and cutoff values were determined by the receiver operating characteristic curve analysis.

Results: A total of 311 patients were included in the study, and 265 (85.2%) of them had acute appendicitis. Mean Alvarado score for patients with appendicitis was 8.2 and 6.7 for those without ($P < 0.001$). Mean PAS for patients with appendicitis was of 7.8 and 6.6 for those without ($P < 0.001$). Based on the ROC curve analysis, a cutoff value for both scoring systems was 7. In patients with acute appendicitis and Alvarado score of 7 or higher, the correct diagnosis would have been set in 236 patients (sensitivity, 89%; specificity, 59%; positive predictive value, 93.1%), whereas in patients with acute appendicitis and a PAS of 7 or higher, the correct diagnosis would have been set in 228 patients (sensitivity, 86%; specificity, 50%; positive predictive value, 90.1%). No significant difference was found in sensitivity and specificity between the observed scoring systems.

Conclusions: Both scoring systems can be of assistance in setting the diagnosis of acute appendicitis, but none has adequate predictive values in assessing acute appendicitis and none can be used as an exclusive standard in setting the diagnosis of acute appendicitis in children. The final decision still remaines on the opinion of an expert pediatric surgeon.

Key Words: appendicitis, Alvarado score, Pediatric Appendicitis Score

(*Pediatr Emer Care* 2015;31: 164–168)

Acute appendicitis is the most common intra-abdominal surgical emergency, and appendectomy is one of the most commonly performed surgical procedures in abdominal surgery.[1] Approximately 7% to 10% of the population is expected to have acute appendicitis during their lifetime.[1–3] In most cases, the clinical data and physical and laboratory examinations are sufficient to establish the presumed diagnosis.[1,2,4] On the other hand, in some cases, distinguishing appendicitis from other disorders is difficult, particularly in young preverbal children and can lead to delays in the diagnosis and an increase in the percentage of appendicular perforations.[3,5,6] Differential diagnosis of an acute appendicitis is extensive, and many clinical conditions can mimic acute appendicitis.[1,5–8] Computed tomography, with sensitivity of 95% to 97% and specificity of 93% to 99%, has improved diagnostic accuracy of acute appendicitis but, on the other hand, exposured children to ionizing radiation.[9] Abdominal ultrasound, with sensitivity of 75% and specificitiy of 90%, has become more popular and negates the risk of radiation, but ultrasound use is operator dependent.[10] In an effort to reduce the time needed for diagnosis and the number of inappropriate appendectomies, different clinical scoring systems aiding in the diagnosis of appendicitis have been developed.[2,11–13] These scoring systems are based on the presence or absence of symptoms. The Alvarado score is the most frequently used system for classification of acute appendicitis. It was oriented toward an adult population, although it has been validated in multiple studies that also included patients of pediatric age (Table 1).[12] Recently, Samuel[13] created another scoring system oriented to the pediatric population, the Pediatric Appendicitis Score (PAS) (Table 2). For each score, the author identified a cut point at which surgery was recommended versus observation.

The aim of this study was to prospectively evaluate the Alvarado and PAS scoring systems in a pediatric population.

METHODS

Patients

A total number of 311 patients who underwent appendectomy because of suspected appendicitis between October 2011 and May 2013 at the Department of Pediatric Surgery, Split University Hospital Centre, were enrolled in the study. There were 172 males (55.3%) and 139% (44.7%) females. Median age was 11.7 years (range, 3–17 years).

The study was carried out as a prospective trial. Informed consent was obtained from all parents of the patients, and the ethical committees of Split University Hospital Centre approved the study protocol.

All patients who underwent appendectomy because of suspected appendicitis in the age group of 0 to 18 years were enrolled in the study. From the total number of patients who entered the study, 186 patients (59.8%) were seen in an emergency department and referred to a surgeon and 125 patients (40.2%) were seen by a primary care physician and sent to a surgeon. Exclusion criteria were patients older than 18 years, pregnant patients, patients with a history of previous appendectomy, patients with chronic diseases, and patients who refused surgery.

Methods

The pediatric surgeon completed a data sheet that included the patient's name, age, sex, laboratory analysis (leukocyte, C-reactive protein, and neutrophils), duration of symptoms, and all the items that make up the Alvarado score and the PAS. To determine interobserver reliability of clinical findings whenever possible, we had 2 independent assessments. In all patients, open or laparoscopic appendectomy was performed. Indication for appendectomy was made by the pediatric surgeon based on clinical features and laboratory measurements. Ultrasound was performad

From the *Department of Pediatric Surgery, Split University Hospital Centre; †University of Split, School of Medicine; and ‡Department of Pathology, Split University Hospital Centre, Split, Croatia.

Disclosure: The authors declare no conflict of interest.

Reprints: Zenon Pogorelić, MD, PhD, Department of Pediatric Surgery, Split University Hospital Centre, Spinčićeva 1, 21 000 Split, Croatia (e-mail: zenon@vip.hr).

ISSN: 0749-5161

TABLE 1. Alvarado Scoring System

	Variable	Value
Symptoms	Migration of pain	1
	Anorexia	1
	Nausea/vomiting	1
Signs	Right lower quadrant tenderness	2
	Rebound pain	1
	Elevation of temperature ≥37.3°C	1
Laboratory	Leukocytosis ≥10 × 10^9/L	2
	Polymorphonuclear neutrophilia ≥75%	1

in 69 (22.2%) of the patients usually when clinical presentation was uncommon. Computed tomography was performd in only 2 patients (0.6%).

Intraoperative findings were recorded. All specimens were examined by a pathologist. Depending on pathohistological findings, patients were divided in 2 groups: acute appendicitis (AA) and nonappendicitis (non-AA).

Outcome Measures

The primary outcome measures were Alvarado score and PAS. Secondary outcome variables were specificity and sensitivity on Alvarado score and PAS.

Statistical Analysis

The data were analyzed using Microsoft Excel for Windows Version 11.0 (Microsoft Corporation, USA) and Statistica for Windows Release 12.0 (Statsoft Inc, Tulsa, Okla). A Student *t* test was used to analyze continuous sizes, and χ^2 test was used for a statistical analysis of the categorical sizes. A receiver operating characteristic (ROC) curve was constructed for both the Alvarado score and the PAS, and for each value of the score, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. The optimal cutoff point to discriminate between AA and non-AA patients was determinated. All values of $P < 0.05$ were considered to indicate statistical significance.

RESULTS

During the study period, 311 patients who underwent appendectomy and gave consent and had complete data for both the PAS and Alvarado scores were included in the study. Mean age was 11.7 years (range, 3–17 years). Mean duration of symptoms at the time of surgery was 29.8 hours (range, 2–240 hours). The characteristics of the 2 groups, along with the most relevant comparative results, are shown in Table 3.

The histopathologic analysis revealed a positive appendicitis in 265 patients (85.2%). There were 28 patients (9.0%) with innocent appendix and 18 patients (5.8%) with other pathologies (Table 4).

The Alvarado score and PAS were calculated for each of the 311 patients from the data collected. Patients with appendicitis had a mean PAS of 7.8 (SD ± 1.4), whereas those without appendicitis had a mean PAS of 6.4 (SD ± 1.6) ($P < 0.001$). Patients with appendicitis had a mean Alvarado score of 8.2 (SD ± 1.5), and patients without appendicitis had a mean Alvarado score of 6.2 (SD ± 1.8) ($P < 0.001$).

The ROC curves for both scores are shown in Figure 1. The area under the curve for Alvarado score was 0.74 (95% confidence interval [95% CI], 0.662-0.818) and that for PAS was 0.73 (95% CI, 0.649–0.811). No significant differences were found between the 2 scores. The optimal cutoff point was 7 for both Alvarado score and PAS (Table 5).

Of the patients with appendicitis, an Alvarado cutoff score of 7 or greater would have led to 236 correct appendicitis diagnoses among 265 appendicitis patients (Table 6). With this cutoff point, the Alvarado score showed a sensitivity of 89%, a specificity of 59%, a PPV of 93.1%, and an NPV of 46% (Table 5).

Of the patients with appendicitis, a PAS cutoff score of 7 or greater would have led to 228 correct appendicitis diagnoses among 265 appendicitis patients (Table 6). With the cutoff point of 7, the PAS showed a sensitivity of 86%, a specificity of 50%, a PPV of 90.1%, and an NPV of 38% (Table 5).

DISCUSSION

Abdominal pain is one of the most common symptoms of patients seeking medical attention. Acute appendicitis is the most common cause of acute abdominal pain, and distinguishing appendicitis from other disorders is sometimes difficult, particularly in young preverbal children.[1,5–8] Early and accurate diagnosis of acute appendicitis is required to reduce the morbidity and mortality associated with delayed diagnosis and its complications. On the other hand, it is very important to reduce the number of unnecessary appendectomies. The rate of unnecessary appendectomies in literature is 10% to 30%; in our study, it was 14.8%.[14] Nowadays, numerous scoring systems for appendicitis have been designed and published in the literature, and 2 most used are Alvarado score and PAS.[12,13]

Alvarado score was originally developed by Alfredo Alvarado in 1986 as an aid to the diagnosis of patients with appendicitis. The score was based on a cohort of 305 patients based at the Nazareth Hospital in Philadelphia in the United States who presented with suspected appendicitis. The charts of these patients were reviewed retrospectively, and the sensitivity and specificity of 8 predictive factors were assessed. Alvarado recommended that the patients with a score less than 5 can be discharged as nonappendicitis, those with a score of 5 or 6 required observation, whereas those with a score of 7 or higher needed to proceed to surgery because it was likely that they had appendicitis.[12] Schneider et al[15] observed 588 patients aged 3 to 21 years and, using the Alvarado that recommended a score of 7 as a cutoff value for having appendicitis, they found a PPV of 65%, an NPV of 46%, a sensitivity of 72%, and a specificity of 81%. Mandeville et al[16] in their study on 287 patients found that an Alvarado cutoff score of 7 or higher would give 118 correct diagnoses; sensitivity, 76%; specificity, 72%; and PPV, 76%. On the basis of these findings, the Alvarado score saw universal use, and there have been many studies showing it as a useful tool in the diagnosis of acute appendicitis.[17–22] In our study, the results obtained with the Alvarado score are quite similar to other reported studies (specificity of 89%,

TABLE 2. Pediatric Appendicitis Scoring System

	Variable	Value
Symptoms	Migration of pain	1
	Anorexia	1
	Nausea/vomiting	1
Signs	Right lower quadrant tenderness	2
	Rebound pain	2
	Elevation of temperature ≥38°C	1
Laboratory	Leukocytosis ≥10 × 10^9/L	1
	Polymorphonuclear neutrophilia ≥75%	1

TABLE 3. The Characteristics of the Groups Along With the Most Relevant Comparative Results

	Group AA (n = 265)	Group Non-AA (n = 46)	*P*
Demographic data			
Age, y	11.6 (3–17)	12 (5–16)	ns
Sex (male/female), %	57/43	55/45	ns
Duration of symptoms, h	29.4 (6–96)	31.9 (2–240)	ns
Preoperative laboratory values			
Leukocytes ($\times 10^9$/L)	15.4 (5.5–34.1)	12 (3–23.9)	<0.001
C-reactive protein, mg/dL	40 (0.5–280)	26 (0.2–243)	<0.001
Neutrophils, %	81 (58–94)	70 (48.5–91)	<0.001
Scoring systems			
Alvarado score, mean ± SD	8.2 ± 1.5	6.7 ± 1.8	<0.001
PAS, mean ± SD	7.8 ± 1.4	6.6 ± 1.6	<0.001

sensitivity of 59%, PPV of 92%, NPV of 46%). In our study, we did not find adequate predictive values for the Alvarado score because 19 (5.7 %) patients with an Alvarado score of 7 or higher had a negative appendix on pathohistological examination and 29 (8.8 %) patients with an Alvarado score less than 7 had a positive appendix on pathohistological examination.

Samuel[13] designed the PAS in 2002 on the basis of a cohort of 1170 children 4 to 15 years old. When tested on the same population, the sensitivity was 100%, specificity was 92%, PPV was 96%, and NPV was 99%. Conclusion from Samuel's study was that patients with a score of 5 or less do not have appendicitis, and a score of 6 or higher was highly associated with appendicitis.[13] Schneider et al[15] found that the same cutoff score of 6 or greater had a PPV of 54%, a sensitivity of 82%, and a specificity of 65%. Bhatt et al[18] found a sensitivity of 92.8% and a specificity of 69.3%. These results are similar to ours. We found that a PAS of 7 or higher (rather than the 6 originally proposed by Samuel) gave a sensitivity of 86%, a specificity of 50%, a PPV of 90.1%, and an NPV of 38%. In contrast, Goldman et al[23] found that a PAS of 7 or greater gave a sensitivity of 94% and a specificity of 98% and suggested that, by using this cutoff for selection of the patients for the operating room, only 4% of appendices with misdiagnosed appendicitis would have been taken out. In our study, we did not find an adequate predictive value for PAS because 23 (7.15%) patients with an Alvarado score of 7 or higher had a negative appendix on pathohistological examination and 37 (11.8%) patients with an Alvarado score less than 7 had a positive appendix on pathohistological examination.

TABLE 4. Pathohistological Findings in Patients Who Underwent Appendectomy

Pathohistological Finding	Total
Phlegmonous (n)	133
Gangrenous (n)	126
No pathology (n)	28
Chronic (n)	6
Other pathologies (n)	18
Mesenteric lymphadenitis	6
Enterobiasis	5
Meckel diverticulitis	2
Salphingitis	1
Torsion of ovarian cyst	4

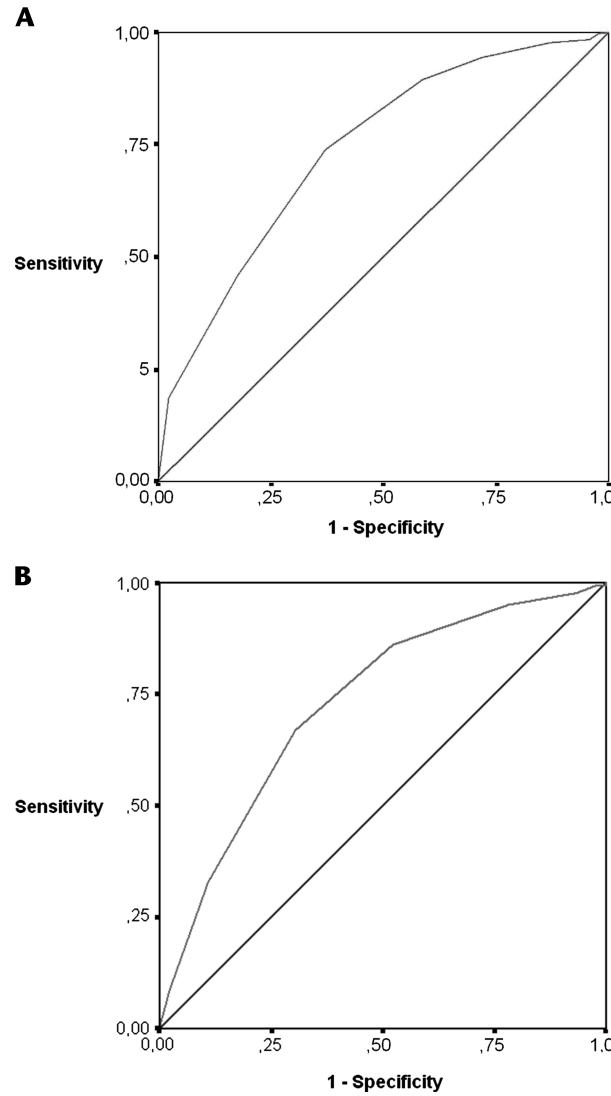

FIGURE 1. ROC analysis. A, Area under the ROC curve for Alvarado score 0.74; 95% CI, 0.662-0.818; $P < 0.001$. B, Area under the ROC curve for PAS 0.73; SE, 0.041; 95% CI, 0.649-0.811; $P < 0.001$.

TABLE 5. PAS and Alvarado Score Performance at Each Cutoff Value

	PAS				Alvarado			
Score	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Sensitivity, %	Specificity, %	PPV, %	NPV, %
0	100	0	85.2	0	100	0	85.2	0
1	100	0	85.2	0	100	0	85.2	0
2	99.4	1.4	85.2	10	100	0	85.2	0
3	99.2	2.2	85.4	33.3	99.6	2.2	85.3	55
4	99.2	2.2	85.4	33.3	98.5	4.3	85.9	50
5	97.7	6.5	85.8	34	97.7	13	86	50
6	95	21.1	90	43	94.3	28.2	87	46.4
7*	86	50	90.1	38	89	59	93.1	46
8	66.8	69.5	92.7	26	73.9	63	93.5	29
9	32.8	89.1	94.5	18.7	45.6	82.6	93.8	20.9
10	6.9	97.8	95.8	15.7	19.5	95.6	96	16.9

*Cut-off value.

The differences in these studies may be explained by the different populations of patients enrolled. Schneider et al[15] and Bhatt et al[18] enrolled only patients with suspected appendicitis, and both of their populations had a rate of appendicitis of approximately 34%. Mandeville et al[16] in their study had a rate of appendicitis of 54%. In contrast, Goldman et al[23] applied the scoring system to all children who presented with abdominal pain. In their sample, only 14.5% had appendicitis. We enrolled only patients who underwent appendectomy because only for these patients was pathohistological analysis performed, and 85.2% of patients in our study had appendicitis proven by pathologic examination. Neither the PAS nor the Alvarado score had an adequate predictive value in the diagnosis of appendicitis. The predictive values obtained in our study were not sufficient to be used solely for making the diagnosis of appendicitis in children. Similar conclusions were made in studies of Schneider et al[15] and Escribá et al.[3]

It seems evident that the opinion of an expert surgeon can never be replaced by a scoring system, and that the final decision whether to operate or not must rest on his criteria. It is also true that the first contact with most cases of acute appendicitis is not made by the surgeon but rather by the clinician who may or may not be familiar with the appendicular pathology. And this is precisely where the protocolization of the physical and clinical findings may be useful in ruling out or detecting acute appendicitis as early as possible.

TABLE 6. Cutoff Values for Alvarado Score and PAS Regarding Pathohistology

	Pathohistology		
	AA	Non-AA	Total
PAS			
≥7	228	23	251
<7	37	23	60
Total	265	46	311
Alvarado			
≥7	236	19	255
<7	29	27	56
Total	265	46	311

CONCLUSIONS

Both scoring systems can be of assistance in setting the diagnosis of acute appendicitis, but none has adequate predictive values in assessing acute appendicitis, and none can be used as an exclusive standard in setting the diagnosis of acute appendicitis in children. The final decision still remains on the opinion of an expert pediatric surgeon.

REFERENCES

1. Kulik DM, Uleryk EM, Maguire JL. Does this child have appendicitis? A systematic review of clinical prediction rules for children with acute abdominal pain. *J Clin Epidemiol*. 2013;66:95–104.
2. Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: Nationwide analysis. *Arch Surg*. 2002;137:799–804.
3. Escribá A, Gamell AM, Fernández Y, et al. Prospective validation of two systems of classification for the diagnosis of acute appendicitis. *Pediatr Emerg Care*. 2011;27:165–169.
4. Groselj-Grenc M, Repse S, Vidmar D, et al. Clinical and laboratory methods in diagnosis of acute appendicitis in children. *Croat Med J*. 2007; 48:353–361.
5. Bansal S, Banever GT, Karrer FM, et al. Appendicitis in children less than 5 years old: influence of age on presentation and outcome. *Am J Surg*. 2012;204:1031–1035.
6. Pogorelić Z, Biočić M, Jurić I, et al. Acute appendicitis as a complication of varicella. *Acta Medica (Hradec Králové)*. 2012;55:150–152.
7. Pogorelić Z, Stipić R, Druzijanić N, et al. Torsion of epiploic appendage mimic acute appendicitis. *Coll Antropol*. 2011;35:1299–1302.
8. Perko Z, Bilan K, Pogorelić Z, et al. Acute appendicitis and ileal perforation with a toothpick treated by laparoscopy. *Coll Antropol*. 2008;32:307–309.
9. 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;29:568–573.
10. Estey A, Poonai N, Lim R. Appendix not seen: the predictive value of secondary inflammatory sonographic signs. *Pediatr Emerg Care*. 2013;29: 435–439.
11. Chong CF, Thien A, Mackie AJA, et al. Comparison of RIPASA and Alvarado scores for the diagnosis of acute appendicitis. *Singapore Med J*. 2011;52:340–345.
12. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med*. 1986;15:557–564.

13. Samuel M. Pediatric appendicitis score. *J Pediatr Surg*. 2002;37:872–881.
14. Kalan M, Talbot D, Cunliffe WJ, et al. Evaluation of the modified Alvarado score in the diagnosis of acute appendicitis: a prospective study. *Ann R Coll Surg*. 1994;76:418–419.
15. Schneider C, Kharbanda A, Bachur R. Evaluating appendicitis scoring systems using a prospective pediatric cohort. *Ann Emerg Med*. 2007; 49:778–784.
16. Mandeville K, Pottker T, Bulloch B, et al. Using appendicitis scores in the pediatric ED. *Am J Emerg Med*. 2011;29:972–977.
17. McKay R, Shepherd J. The use of the clinical scoring system by Alvarado in the decision to perform computed tomography for acute appendicitis in the ED. *Am J Emerg Med*. 2007;25:489–493.
18. Bhatt M, Joseph L, Ducharme F. Prospective validationof the pediatric appendicitis score in a Canadian pediatric emergencydepartment. *Acad Emerg Med*. 2009;16:591–596.
19. Subotic AM, Sijacki AD, Dugalic VD, et al. Evaluation of the Alvarado score in the diagnosis of acute appendicitis. *Acta Chir Iugos*. 2008;55: 55–61.
20. Arian GM, Sohu KM, Ahmad E, et al. Role of Alvarado score in diagnosis of acute appendicitis. *Pak J Surg*. 2001;17:41–46.
21. Hsiao KH, Lin LH, Chen DF. Application of the MANTRELS scoring system in the diagnosis of acute appendicitis in children. *Acta Paediatr Taiwan*. 2005;46:128–131.
22. Sun JS, Noh HW, Min YG, et al. Receiver operating characteristic analysis of the diagnostic performance of a computed tomographic examination and the Alvarado score for diagnosing acute appendicitis: emphasis on age and sex of the patients. *J Comput Assist Tomogr*. 2008;32: 386–391.
23. Goldman RD, Carter S, Stephens D, et al. Prospective validation of the pediatric appendicitis score. *J Pediatr*. 2008;153:278–282.