Clinical scores for prediction of acute appendicitis in children in a hospital of Lima, Perú

Edson Guzmán[a,c,d,e] and Nadia García[b]

Objective To determine the usefulness of the Alvarado score and the Pediatric Appendicitis score (PAS) in the Pediatric Emergency of the National Hospital Daniel A. Carrion.

Materials and methods A prospective observational study was carried out of patients younger than 15 years of age with abdominal pain and suspected acute appendicitis (AA) attending the Pediatric Emergency in a Hospital of Lima, Peru. These patients underwent a survey to assess the parameters of the Alvarado score and PAS.

Results Three hundred and seventeen patients with abdominal pain and suspected of AA were recruited over a study period of 12 months. Of the patients, 232 were considered to have AA clinically and underwent surgery. 85.3% were confirmed by pathology and 14.7% were normal. The mean Alvarado score was 8.27 ± 1.31; the mean Surgical Procedure Assessment (SPA) score was 8.08 ± 1.47. Sensitivity and specificity for both scores are equivalent. The area under the curve for the Alvarado score and SPA were 0.887 and 0.901, respectively. Alvarado score higher than 6 had a sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of 88.9, 75.6, 97.4, 68.1, and 86.4%, respectively. SPA higher than 6 points had sensitivity, specificity, PPV, NPV, and accuracy of 84.3, 80.7, 94.7, 73.1, and 86.7%, respectively.

Conclusion Alvarado score and the PAS are scores with high sensitivity, specificity, PPV, and accuracy for the diagnosis of AA when the score is higher than 6 points. The results found in our study justify their use in emergency services, but they should not be used as the only means of clinically determining the need for surgery. *Ann Pediatr Surg* 10:35–38 © 2014 Annals of Pediatric Surgery.

Introduction

Acute appendicitis (AA) is one of the most common surgical pathologies and the most important cause of acute abdomen in childhood [1–4]. Most of the time, the diagnosis of appendicitis is easy, but, in some cases, clinical features are atypical, leading to a wrong diagnosis and a high risk of complications. The correct diagnosis of appendicitis varies between 50 and 70% in adults in the first visit to the Emergency Services [5,6]. Otherwise, some studies have shown that the risk of performing inappropriate appendectomies can be as high as 10–30% [1,6–8].

For these reasons, many clinical scores have been created for the diagnosis of AA. One of these is the Alvarado score [9], created in 1986, mostly used in adult populations, and one of the scores for the pediatric population is the Pediatric Appendicitis score (PAS), created by Samuel in 2012 and validated recently [10–13].

The main objective of the present study was to compare the prediction of AA with the Alvarado score and PAS in pediatric patients presenting in emergency with abdominal pain.

Materials and methods

This study was carried out in the emergency service of Daniel Alcides Carrión Hospital, Callao, Peru. The study was prospective, carried out from the June 2011 to May 2012, and including all patients younger than 15 years of age who presented with abdominal pain and suspected of having AA.

AA was suspected in all patients with acute abdominal pain that, according to the interview with the parents and/or the patient himself/herself, required examination for a probable AA. Medical history was assessed and physical examination was performed, and patients were subjected to blood tests and/or ultrasonography. Patients with abdominal pain but not subjected to blood analysis or incomplete serology, assuming that they did not have AA, and those referred from other centers with a previous diagnosis of AA were excluded. Informed written consent was obtained from all parents or legal guardians.

For all patients included in the study, with or without a clinical diagnosis of appendicitis defined, we used the prediction scores of appendicitis, Alvarado score (Table 1) and PAS (Table 2), only to determine the diagnostic accuracy of these scores in predicting AA. Demographic features were recollected (which included the patient's age, sex, symptoms, and all the items that make up the Alvarado score and the PAS). As mentioned by Escribá *et al.* [1], the PAS did not define exactly the percentages for polymorphonuclear neutrophilia and fever, so that in our study, we defined the percentage of polymorphonuclear neutrophilia as 75% or higher and fever as 38°C or higher under the arm. For those operated on, the confirmed diagnosis of AA was made by pathological anatomy. Both physical examination and data collection were performed by a third year resident of the Pediatric Department in her final year, supervised by a Pediatric specialist. Patients without a diagnosis of AA were discharged by the

pediatrician from the observation room of the Pediatric Emergency Service about 8–12 h after admission.

The data obtained after the selection and analyses of the sample were analyzed using the program PASW Statistics 18 (SPSS Inc., Chicago, Illinois, USA). Clinical and analytical variables of the AA and non-AA groups were compared. In the comparison of categorical variables, unless otherwise noted, any test of a hypothesis was two sided and the level of significance was set at 5%. Clinical parameters were tested by univariate analysis using the Student t-test or the χ^2-test. The results were expressed as means ± SDs. Patients' demographic and other characteristics were analyzed using the Student t-test or the Mann–Whitney U-test for continuous variables depending on the normality or the non-normality of the distribution of each variable and the χ^2-test (with the Yates correction) or Fisher's exact test for categorical variables. A receiver operating characteristic curve was constructed to assess sensitivity and specificity and optimal cut points for the Alvarado score and PAS to diagnose appendicitis. Area under the curve and the corresponding 95% confidence interval were calculated, and for each value of the score, we calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The study was approved by the hospital ethics committee.

Results
Initially, 317 patients were included in the study, admitted to the Pediatric Emergency Service with a diagnosis of abdominal pain syndrome with suspected AA. These patients were recruited over a period of 12 months, from June 2011 to May 31 2012. Of the 317 patients, 53.6% were men (170 patients) and 46.4% were women (147 patients). The mean age of the patients was 9.6 years (SD ± 3.05 years; range, 2–14 years) and the mean evolution of symptoms at the time of presentation

in Pediatric Emergency Service was 38.43 h (SD ± 34.7; range, 2–140 h).

Of the 317 patients with abdominal pain, 73.2% were diagnosed with AA (232 patients) and underwent surgery. Eighty-five patients (26.8%) were discharged with other final diagnoses and excluded from the study. Of the 232 patients who were operated, 198 (85.3%) had appendicitis confirmed by pathology and 34 patients (14.7%) had negative appendectomies (normal appendix). Of the 198 patients with appendicitis, only five (2.5%) presented with a perforated appendix.

The most frequent location of abdominal pain in the patients with appendicitis confirmed by pathology was the right lower quadrant in 93, 1% (216 of 232 patients); in the four, in 3% (10 patients), the location was diffuse. Excluding abdominal pain, the most common symptoms presented by patients on admission to the emergency were nausea and/or vomiting (78.2%) and anorexia (75.7%), whereas the most frequent signs were right lower quadrant tenderness (82.6%) and migration of pain (72.5%).

The patients were divided into two groups. The first group (group A) included patients with histological confirmation of AA (198 patients) and the second group (group B) included patients without AA (119 patients). The characteristics of the two groups along with the most relevant comparative results are shown in Table 3. On analysis by sex, it was observed that in the group with AA, 86 were women (43.4%) and 112 were men (56.6%), whereas in the group without AA, 61 were women (51.2%) and 58 were men (48,8%) ($P = 0.2$).

The Alvarado score and PAS were calculated for all 317 patients. The means of the Alvarado score and PAS were 8.27 and 8.08 in group A and 4.43 and 3.99 in group B. These data were statistically significant ($P < 0.001$). The area under the curve for the Alvarado score was 0.887, whereas for the PAS, it was 0.901. There were no significant differences between the two scores. The receiver operating characteristic curves for both scores are shown in Fig. 1. The optimal cutoff point was 6 for both the Alvarado score and PAS (Table 4). With this cutoff point, the Alvarado score showed a sensitivity of 88.9%, a specificity of 75.6%, a PPV of 97.4%, and an NPV of 68.1%. With the cutoff point of 6, the PAS showed a sensitivity of 84.3%, a specificity of 80.7%, a PPV of 94.7%, and an NPV of 73.1%. The accuracy for Alvarado score at least 6 is 86.4%, whereas that for PAS is 86.7%.

Thirty-four patients were operated for suspected appendicitis, but the histopathology was normal. In these patients, the mean Alvarado score and PAS were 6.76 and 6.5, respectively. The means Alvarado score and PAS of

Table 1 Alvarado score

Alvarado score	
Migration of pain	1
Anorexia	1
Nausea/vomiting	1
Right lower quadrant tenderness	2
Rebound pain	1
Elevation in temperature ($>37.3°C$)	1
Leukocytes $>10\,000/\mu l$	2
Polymorphonuclear neutrophilia $>75\%$	1
Total	10

Table 2 Pediatric Appendicitis score (PAS)

Pediatric Appendicitis score	
Migration of pain	1
Anorexia	1
Nausea/vomiting	1
Right lower quadrant tenderness	2
Cough/hopping/percussion tenderness in the right lower quadrant	2
Elevation in temperature ($38°C$)	1
Leukocytes $>10\,000/\mu l$	1
Polymorphonuclear neutrophilia 75%	1
Total	10

Table 3 Characteristics of the two groups along with the most relevant comparative results

	Group A	Group B	P
Age (mean ± SD) (years)	9.54 ± 3.1	9.69 ± 2.9	0.6
Sex (males/females)	112/86	58/61	0.17
Evolution of symptoms (mean ± SD) (h)	34.54 ± 25.8	44.91 ± 45.2	0.01
Alvarado score (mean)	8.27 ± 1.31	4.43 ± 2.58	<0.001
Pediatric Appendicitis score (mean)	8.08 ± 1.47	3.99 ± 2.48	<0.001
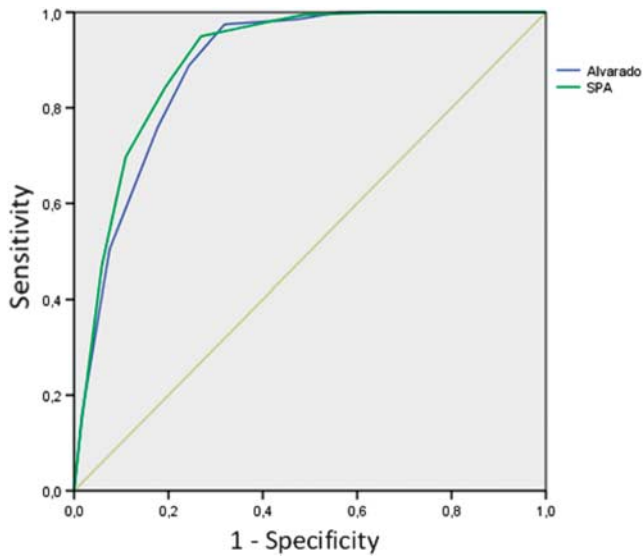
Area under the receiver operating characteristic curve for the Alvarado score and Pediatric Appendicitis score. SPA, Surgical Procedure Assessment.

Table 4 Alvarado score and PAS

Criterion	Alvarado score		PAS	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
≥ 1	100	10.9	100	15.1
≥ 2	100	30.3	100	35.3
≥ 3	100	43.7	99.5	51.3
≥ 4	98.5	52.9	97.5	61.3
≥ 5	97.5	68.1	94.9	73.1
≥ 6	88.9	75.6	84.3	80.7
≥ 7	75.8	82.4	69.7	88.1
≥ 8	50.5	92.4	47.0	94.1
≥ 9	16.2	98.3	15.2	98.3
10	0	100	0	100

Each cutoff point for all possible scores. Criterion values and coordinates of the receiver operating characteristic curve.
PAS, Pediatric Appendicitis score.

the three groups were determined. Statistically significant differences were found in both cases (P < 0.001; Table 5).

Abdominal ultrasonography was not performed in all patients. Of the 317 patients with abdominal pain, only 57% (181 patients) underwent abdominal ultrasonography, 93 patients (29.3%) had normal abdominal ultrasound, 45 had appendicitis (14.2%), four had appendicular plastron (1.3%), and 39 patients (12.3%) had ileum. Of the 80 patients without AA who underwent abdominal ultrasound, five were reported to have appendicitis on abdominal ultrasonography (6.2%), whereas of the 101 patients with appendicitis, only 40 (39.6%) were reported to have AA on abdominal ultrasonography.

Discussion

Most of the clinical scores for the prediction of AA have been used in adult populations. A retrospective study concluded that the use of these clinical scores reduced the rate of negative appendectomies by one-third [14].

Table 5 Comparison between the means of the scores in patients with appendicitis, without appendicitis, and SAS

	Appendicitis (n=198)	No appendicitis (n=85)	SAS (n=34)	P
Alvarado score (mean)	8.27	3.49	6.76	<0.001
Pediatric Appendicitis score (mean)	8.08	2.98	6.5	<0.001

Alvarado [9], in 1986, designed a score with eight predictive factors applied to a retrospective sample of 305 hospitalized patients. In his article, he recommended that patients with less than 5 points be discharged as non-AA, those with 5–6 be placed under observation as possible AA, and those with 7 or higher be operated on as likely AA. In our study, the optimal cutoff point was 6 for both the Alvarado score and the PAS (Table 4). With this cutoff point, the Alvarado score showed a sensitivity of 88.9%, a specificity of 75.6%, a PPV of 97.4%, an NPV of 68.1%, and an accuracy of 86.4%. In our study, the accuracy with a cutoff of at least 6 is higher than that obtained with a cutoff point of at least 7.

On the basis of these studies, it has been proven that the Alvarado score is a useful tool for the diagnosis of AA [1]. In our study, the results obtained with the Alvarado score are quite similar to those described most recently in the literature.

In the study designed by Samuel [10] on the PAS, it is recommended that a score of 5 or lower does not help establish a diagnosis of AA, whereas 6 or more points may point to a diagnosis of AA and these patients should be operated [1,10]. In our study, with the cutoff point of 6, the PAS showed a sensitivity of 84.3%, a specificity of 80.7%, a PPV of 94.7%, and an NPV of 73.1%. This scoring system has been validated recently in other studies, with similar results [1,11–13].

There are differences between the PAS and the Alvarado scores. The PAS assigns 2 points to cough/percussion/hopping tenderness and to right lower quadrant tenderness, whereas the Alvarado scoring system assigns 2 points each to tenderness in the right iliac fossa and white blood cells count greater than $10\,000/mm^3$ [10,15]. Both scores are useful tools to predict the diagnosis of AA.

In our study, it is important to note that only five patients with an Alvarado score of 5 points or less were finally confirmed to have AA after undergoing surgical treatment; three of them had a score of 4 points and one had 1 point. In the case of the PAS, 10 patients had a score of 5 points or less who underwent surgery for persistent symptoms; one of them had a PAS less than 3, four had 4 points, and five had 5 points. This means that although the scores have good sensitivity and specificity, clinical judgment should govern whether a patient needs surgery.

Misdiagnosis can lead to unnecessary appendectomy; in some series, this has been reported to be between 10 and 30% [1,6–8]. In our study of 232 patients, 34 were classified histologically as normal, which represented

14.7% of the total operated patients. These data are within the ranges observed in the literature previously.

In our study, the mean age of patients with a diagnosis of AA was 9.6 years (range 2–14 years). In the study of Escribá et al. [1], the mean age of the patients was 11.2 years; this can probably be explained by the fact that in the latter study, patients up to 18 years of age were included, whereas in our study, all patients were younger than 15 years of age.

In the present study, only 2.5% of patients had a perforated appendicitis (only five of the 198 patients were diagnosed with AA). These five patients had persistence of abdominal pain of at least 24 h; two of them even had 3 days (> 72 h) of abdominal pain. No dependent factors analyzed this delay of these patients to go to the emergency. It is likely that the low rate of perforations observed in patients diagnosed with AA is because of the fact that surgery was not delayed after a suspected diagnosis of appendicitis was established in these patients.

Gomez et al. [16], in a study carried out in Brazil, validated the Alvarado score in children and adolescents; it was found that a score of at least 6 points had a sensitivity of 81.5% and a specificity of 72.7%. They agreed that a score Alvarado of at least 5 points can be used as a tool of high prognostic value for the diagnosis of AA. Meanwhile, the study of Bhatt et al. [13], published in 2009 in Canada, concluded that the PAS is useful for the evaluation of possible appendicitis in children; they concluded that patients with a score of 4 points or less can be safely discharged and that appendicitis can be ruled out. Similarly, a score of at least 8 points for the PAS can predict AA. In this study, a score of at least 7 points had a sensitivity of 73.5% and a specificity of 85.3%. In the study by Beltran et al. [5], published in Chile in 2007, all nonoperated patients had an Alvarado score of 0–6 points and most operated patients had a score between 9 and 10 points. For these patients, the score had a high sensitivity, specificity, and accuracy. Our study showed that the means of the two scores were greater than 6 points in the patients with SAS; therefore, our recommendation for patients with this score (6 points) would be careful monitoring so that they are not subjected to unnecessary surgery.

Moreover, Goulder et al. [17], in a study published in 2008, found that the PAS cannot be recommended as it would later lead to an unacceptable risk of high medical or surgical delays in up to 13% of patients with AA.

Our study has certain limitations, including the size of the selected sample; however, this sample is larger than those of other studies [1]. Nevertheless, given the highly significant results that we obtained, we believe that our study allows for relevant conclusions to be drawn. Another limitation was that ultrasound studies were not carried out in our study; nevertheless, we believe that clinical judgment is sufficient to make a diagnosis of AA.

Conclusion

The Alvarado score and the PAS were useful for the evaluation of patients suspected to have AA because of their heightened sensitivity and specificity when the score is equal to or higher than 6 points, and without statistical differences between them.

Alvarado score and PAS lower than 6 points should not be the criteria for discharging patients with suspected appendicitis because there is a percentage of patients in whom the diagnosis of the AA may be missed.

The results found in our study justify their use in emergency services, although they should not be used as the only means of clinically determining the need for surgery.

Acknowledgements

Conflicts of interest
There are no conflicts of interest.

References

1 Escribá A, Gamell AM, Fernández Y, Quintillá JM, Cubells CL. Prospective validation of two systems of classification for the diagnosis of acute appendicitis. Pediatr Emerg Care 2011; 27:165–169.
2 Carty MG, Azizkhan RG. Acute surgical conditions of the abdomen. Pediatr Ann 1994; 23:192–194.
3 Wagner JM, McKinney WP, Carpenter JL. Does this patient have appendicitis? JAMA 1996; 276:1589–1594.
4 Graffeo CS, Counselman FL. Appendicitis. Emerg Med Clin North Am 1996; 14:653–671.
5 Beltran M, Almonacid J, Gutierrez J, Cruces K. Puntuación diagnóstica de apendicitis aguda en niños realizada por pediatras de las Unidades de Emergencia [Diagnostic score for appendicitis in children by Emergency Unit paediatricians]. Rev Chil Pediatr 2007; 78:584–591.
6 Surana R, O'Donnell N, Puri P. Appendicitis diagnosed following active observation does not increase morbidity in children. Pediatr Surg Int 1995; 10:76–78.
7 Douglas CD, Macpherson NE, Davidson PM, Gani JS. Randomised controlled trial of ultrasonography in diagnosis of acute appendicitis incorporating the Alvarado score. BMJ 2000; 321:919–922.
8 Benjamin IS, Patel AG. Managing acute appendicitis: laparoscopic surgery is particularly useful in women. BMJ 2002; 325:505–506.
9 Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med 1986; 15:557–564.
10 Samuel M. Pediatric appendicitis score. J Pediatr Surg 2002; 37:877–881.
11 Schneider C, Kharbanda A, Bachur R. Evaluating appendicitis scoring system using prospective pediatric cohort. Ann Emerg Med 2007; 49:778–784.
12 Goldman RD, Cartes S, Stephens D, Antoon R, Mounstephen W, Langer JC. Prospective validation of the pediatric appendicitis score. J Pediatr 2008; 153:278–282.
13 Bhatt M, Joseph L, Ducharme F, Ducharme FM, Dougherty G, McGillivray D. Prospective validation of the pediatric appendicitis score in a Canadian pediatric emergency department. Acad Emerg Med 2009; 16:591–596.
14 Teicher I, Landa B, Cohen M, Kabnick LS, Wise L. Scoring system to aid in diagnoses of appendicitis. Ann Surg 1983; 198:753–759.
15 Wu HP, Yang WC, Wu KH, Chen CY, Fu YC. Diagnosing appendicitis at different time points in children with right lower quadrant pain: comparison between Pediatric Appendicitis Score and the Alvarado Score. World J Surg 2012; 36:216–221.
16 Gomez PS, De Carvalho M, Falbo G. The Alvarado score validation in diagnosing acute appendicitis in children and teenagers at the Instituto Materno Infantil at Pernambuco. Rev Bras Saude Matern Infant 2003; 4:439–445.
17 Goulder F, Simpsom T. Pediatric appendicitis score: a retrospective analysis. J Indian Assoc Pediatr Surg 2008; 13:125–127.