How much Reliable Is Alvarado Scoring System in Reducing Negative Appendectomy?

Mehmet Tahsin Tekeli¹ · Enver Ilhan¹ · Orhan Ureyen¹ · Abdullah Senlikci¹ · Eyup Yeldan¹ · Mucteba Ozturk¹ · Mehmet Akif Ustuner¹ · Ugur Gokcelli¹ · Emrah Dadali¹ · Fevzi Cengiz¹ · Savas Yakan¹

Received: 22 July 2015 / Accepted: 23 December 2015 / Published online: 11 January 2016
© Association of Surgeons of India 2016

Abstract There is still an ongoing debate, especially regarding early diagnosis of acute appendicitis. Early surgery leads to inadequate evaluation of acute abdominal pain and negative appendectomy, whereas delayed surgery leads to appendicitis perforation complications. The diagnosis of this condition is considerably difficult, especially due to subtle early symptoms and clinical condition. The aim of the present study was to identify whether the Alvarado scoring system could reduce the incidence of negative appendectomy in patients who will undergo surgery for acute appendicitis. Patients who underwent surgery with acute appendicitis prediagnosis were retrospectively classified as negative appendectomies (group A) and positive appendectomies (group B) according to histological diagnosis. All groups were evaluated for age, gender, Alvarado scores, and parameters. Two hundred eighty-one patients were included in the study. Group A contained 71 (25.3 %) patients, and group B contained 210 (74.7 %) patients. There was a significant difference in WBC, left shift, rebound, and change of pain localization between the groups ($p=0.002$, $p<0.001$, $p<0.001$, and $p=0.023$, respectively). Alvarado scores were significantly different between the groups ($p<0.001$). In logistic model examination, the major factor was the Alvarado score (7 or above) and the minor factor was spreading pain. The Alvarado scoring system can be used to reduce negative appendectomy in patients who will undergo surgery with acute appendicitis.

Keywords Negative appendectomy · Acute appendicitis · Alvarado score · Appendectomy

Background

Acute appendicitis is among the most common causes of acute abdominal pain which require surgery, and the probability of having this condition in lifetime is approximately 7 % [1]. There is still an ongoing debate, especially regarding early diagnosis of this disease. Early surgery leads to inadequate evaluation of acute abdominal pain and negative appendectomy, whereas delayed surgery leads to appendicitis perforation complications [2]. The diagnosis of this condition is considerably difficult, especially due to subtle early symptoms and clinical condition [3]. Negative appendectomy is observed in 15–30 % of cases, where a decision for surgery is made based on the clinical symptoms and findings [2, 4]. Delayed diagnosis leads to various complications, including perforation, periappendicular abscess, wound infection, and intra-abdominal adhesion [5]. Different scoring systems are used for early diagnosis of acute appendicitis and its appropriate management. These systems are valuable and valid tools to distinguish between acute appendicitis and nonspecific abdominal pain [6]. The probability of appendicitis can be determined by using a scoring system, namely the Alvarado scoring system [7]. This system consists of the following parameters: right lower quadrant pain, rebound, change of pain localization, nausea and/or vomiting, loss of appetite, fever, leukocytosis, and left shift [2, 7]. More than two decades ago, this scoring system has been designed as a diagnostic tool; however, there is still an ongoing debate since its performance and routine use remain unclear [8]. In addition, the Alvarado scoring system was modified over time (over 9 points, without calculating the neutrophil count); however, studies were

✉ Orhan Ureyen
dureyen@yahoo.com

¹ Department of Surgery, Izmir Bozyaka Training and Research Hospital, Izmir, Turkey

unable to prove the practicality of its clinical use compared to the normal Alvarado scoring system [9]. Furthermore, appendicitis inflammatory response (AIR) is an actual scoring system formed by the addition of new parameters and change of some current parameters of the Alvarado scoring system [10]. Although AIR is demonstrated as an effective method avoiding negative appendectomy [11, 12], clinical studies about the effectiveness of AIR are in progress.

For the Alvarado scoring system, patient history and clinical examination serve as the basis in the classical approach; patients with a score between 1 and 4 may have no appendicitis and can be discharged, patients with a score between 5 and 6 may have appendicitis and can be monitored/hospitalized, and patients with a score between 7 and 10 may have appendicitis and surgical intervention is recommended [7]. The parameters for the Alvarado scoring system are presented in Table 1.

Objectives

The present study aimed to investigate whether the Alvarado scoring system could reduce the incidence of negative appendectomy in patients who would undergo surgery with acute appendicitis diagnosis.

Patients and Methods

Preoperative data for patients who were operated with acute appendicitis prediagnosis were evaluated between September 2011 and January 2014 retrospectively. Patients who were 16 years old or older were included in the study, as our institution was for adults only. All fertile female patients were

examined by a gynecologist in the preoperative period. Ultrasonography (US) was performed preoperatively in all patients. Alvarado scores were calculated for all patients. The decision for surgery was based on clinical symptoms and physical examination. US, laboratory findings, and Alvarado scoring system were used as supplementary means during the decision-making process for surgery. All specimens were evaluated histopathologically by an experienced pathologist. Patients were evaluated with respect to age, gender, Alvarado score, Alvarado parameters, and postoperative histopathology results.

Patients who underwent surgery with acute appendicitis prediagnosis were retrospectively classified as negative appendectomies (group A) or positive appendectomies (group B) according to histological diagnosis. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and diagnostic accuracy were calculated for Alvarado score.

Statistical Methods

Statistical Package for the Social Sciences (SPSS) 21.0 software was used for data analysis. Shapiro-Wilk test was used to determine whether the data showed normal distribution. Parametric tests were used for variables that were normally distributed, whereas nonparametric test were used for variables that were not normally distributed. Independent-samples *t* test and Mann-Whitney *U* (exact) tests were used to compare two independent groups. Pearson chi-square (exact) and linear-by-linear association (exact) tests were used to compare categorical variables, and Cohen's kappa coefficient was used to test inter-rater agreement. Logistic regression test was used to determine the cause-effect relationship with descriptive variables in diatom and multinomial categories. The sensitivity and the specificity of the association between the classification according to the cutoff value (calculated for the pathology result) and the actual classification were examined and expressed using receiver operating curve (ROC) sensitivity analysis. Quantitative variables were expressed as mean $\pm$ standard deviation (SD) and median $\pm$ interquartile range (IQR) within the tables. Categorical variables were expressed as the number (*n*) and percentages (%). Data were analyzed with 95 % CI, and *p* values <0.05 were considered as statistically significant.

Results

Two hundred eighty-one patients were included in the study. Ninety patients (32 %) were female, and 191 patients (68 %) were male. The average age of group A and group B was 29.2 and 31.7, respectively. The female/male ratio in group A and group B was 28:43 and 62:148, respectively. There is no

Table 1 Detailed analysis of the Alvarado scoring system

Component	Score
Symptoms	
Change of pain localization	1
Loss of appetite	1
Nausea/vomiting	1
Clinical findings	
Pain in the right iliac fossa	2
Rebound	1
Fever (>37.3 °C)	1
Laboratory findings	
Leukocyte count $> 10 \times 10^9/l$	2
Elevated neutrophil count > 75 %	1
Total score	10

Table 2 Distribution of cases with respect to Alvarado score subgroups (total score)

		Pathology		<i>p</i> value
		No <i>n</i> (%)	Yes <i>n</i> (%)	
Right lower quadrant tenderness	No/yes	0 (0)/71 (100)	5 (2.4)/205 (97.6)	0.335
Rebound	No/yes	27 (38)/ 44 (62)	40 (19)/170 (81)	0.002
Leukocyte	Normal/high	35 (49.3)/36 (50.7)	35 (16.7)/175 (83.3)	<0.001
Left shift	No/yes	35 (49.3)/36 (50.7)	32 (15.2)/178 (84.8)	<0.001
Nausea/vomiting	No/yes	18 (25.4)/53 (74.6)	35 (16.7)/175 (83.3)	0.116
Loss of appetite	No/yes	26 (36.6)/45 (63.4)	80 (38.1)/130 (61.9)	0.888
Fever	No/yes	62 (87.3)/9 (12.7)	161 (76.7)/49 (23.3)	0.062
Spreading pain	No/yes	58 (81.7)/13 (18.3)	140 (66.7)/ 70 (33.3)	0.023
Total Alvarado score	1	0 (0)	1 (0.5)	<0.001*
	3	1 (1.4)	2 (1)	
	4	9 (12.7)	8 (3.8)	
	5	19 (26.8)	14 (6.7)	
	6	25 (35.2)	27 (12.9)	
	7	9 (12.7)	48 (22.9)	
	8	5 (7)	67 (31.9)	
	9	3 (4.2)	40 (19)	
	10	0 (0)	3 (1.4)	
	Median ± IQR	6 ± 1	8 ± 1	

Independent *t* test-Mann-Whitney *U* test (exact)-Pearson chi-square test (exact)*IQR* interquartile range^a Linear-by-linear association (exact)

difference between the groups in terms of age and gender (*p* value was 0.177 and 0.082, respectively). Group A (negative laparotomy) contained 71 (25.3 %) patients, and group B contained 210 (74.7 %) patients. The negative appendectomy rate in females and males was 31.1 and 22.5 %, respectively. There was a significant difference in WBC, left shift, rebound, and change of pain localization between the groups (*p* < 0.001, *p* < 0.001, *p* = 0.002, and *p* = 0.023, respectively) (Table 2). When the Alvarado score cutoff value was set as 7, the sensitivity was 75.2 %, the specificity was 76.1 %, and the AUC was 0.780 ± 0.031 (*p* < 0.001) (Table 3, Fig. 1). The diagnostic accuracy was 75.4 %.

In logistic model examination, the major factor was the Alvarado score (7 or above) and the minor factor was

spreading pain. The odds ratio of patients with an Alvarado score of 7 or above in group B was 4.716-fold higher compared to patients with an Alvarado score of 7 or above in group A (*p* < 0.001). The odds ratio of having spreading pain was 2.691-fold higher in patients in group B compared to group A (*p* = 0.044) (Table 4). There was no significant difference in the median Alvarado score between females and males (*p* = 0.224).

Discussion

For the diagnosis of acute appendicitis, anamnesis, clinical examination, and laboratory findings can be

Table 3 Sensitivity and specificity distribution of cases according to Alvarado

Alvarado score	Pathology		Sensitivity/specificity (%)	AUC ± SE	<i>p</i> value
	No <i>n</i> (%)	Yes <i>n</i> (%)			
<7	54 (76.1 %)	52 (24.8 %)	75.2/76.1	0.780 ± 0.031	<0.001
≥7	17 (23.9 %)	158 (75.2 %)			

AUC area under the ROC curve, *SE* standard error

Fig. 1 Sensitivity and specificity according to the ROC curve

sufficient for cases with classical presentation. In atypical presentations, however, US, computer-assisted tomography, magnetic resonance imaging, scoring systems, and laparoscopy are supplementary methods for diagnosis [6]. The Alvarado scoring system was developed in 1986 by Alvarado [7]. Clinical symptoms, physical examination, and laboratory findings create the basis of this system. It is a highly sensitive scoring system in cases of diagnostic difficulty regarding acute appendicitis [2]. Different prospective studies have suggested that the Alvarado score on its own is insufficient as a diagnostic test [13, 14].

Memon et al. carried out a study of 110 patients and found that the Alvarado scoring system had a sensitivity of 93.5 %, specificity of 80.6 %, PPV of 92.3 %, NPV of 83.3 %, and diagnostic accuracy of 89.8 % [6]. In a similar study by Kirkil et al. [5], the sensitivity, specificity, PPV, NPV, and diagnostic accuracy values were 95.5 %, 67.5 %, 77.8, 92.6, and 82.7 %, respectively. According to the review by Ohle et al. for a cutoff score of 7, the sensitivity and specificity of the Alvarado scoring system for all studies were 82 and 81 %, respectively [8]. In the present study, sensitivity, specificity, PPV, and NPV values were 75.2 %, 76.1 %, 90.2, and 50.9, respectively. The diagnostic accuracy was 75.4 %.

Abou et al. [15] carried out a study on pediatric patients and found a significant difference in certain Alvarado scoring parameters, including loss of appetite, rebound, and left shift. While sensitivity and specificity are comparably low in the present study, multivariate analysis showed that rebound, leukocytosis, change of pain localization, and left shift were significant for accurate diagnosis. In the present study, logistic model examination of these parameters showed that only an Alvarado score of 7 or above and spreading pain were significant. Some of the recent studies have shown that the

Table 4 Logistic model examination

	<i>p</i> value	Odds ratio	95 % CI for odds ratio Lower limit
Rebound	0.322	1.489	0.677
Leukocytosis	0.453	1.629	0.456
Left shift	0.522	1.490	0.440
Spreading pain (yes)	0.044	2.265	1.024
Alvarado score (≥ 7)	0.001	4.716	1.865

Logistic regression (method = enter)

CI confidence interval

diagnostic value of the Alvarado scoring system is higher for a cutoff score of 8 or higher [16].

According to the literature, the incidence of negative appendectomy ranges between 11 and 40 % [16–18]. Some studies have indicated that the Alvarado scoring system can significantly reduce the incidence of negative appendectomy [19]. According to the study by Kanumba et al. [20], the incidence of negative appendectomy was 33.1 %, which is also high like that of another study (28.7 %) [6]; on the other hands, some authors have reported low negative appendectomy rates about 10 % [2]. The incidence of negative appendectomy in the present case (25.3 %) was consistent with that in the literature. The Alvarado score was 6 or lower in 76.1 % of these cases. While several studies have pointed out that the use of the Alvarado scoring system in females is limited [6, 8]. There was no significant difference in the present study.

Although the negative appendectomy rate of our study is compatible with that in the literature, it is high. Our study is a retrospective clinical study. The greater amount of our negative appendectomy cases is in <7 Alvarado scoring system group (76.1 %) (Table 3). If the Alvarado scoring system was taken into consideration with the other clinical findings, negative appendectomy rate of our cases would decrease, in the period of diagnosis.

Our study has several limitations. Firstly, our study is a retrospective clinical study. Further, operation decision of cases were made by several surgeons in one institution.

Conclusion

The Alvarado scoring system is an effective scoring system in reducing negative appendectomy rate and, especially for scores of 7 or above, indicates patients with appendicitis. A score below 7 means that the negative appendectomy could be reduced. In the diagnosis period, physicians would regard the Alvarado scoring system of the patient and the options of follow-up or additional imaging procedures would be evaluated in the Alvarado scoring system scores below 7, before the decision of surgery. However, the fact that very low negative predictive value can cause more serious problems in appendicitis cases is the cause of misdiagnosis and delay in treatment.

Compliance with Ethical Standards This article was prepared in compliance with ethics rules.

Conflict of Interest The authors declare that they have no competing interests.

References

1. Ilhan E, Senlikci A, Kizanoglu H et al (2013) Do intestinal parasitic infestations in patients with clinically acute appendicitis increase the rate of negative laparotomy? Analysis of 3863 cases from Turkey. *Prz Gastroenterol* 8:366–369
2. Nasiri S, Mohebbi F, Sodagari N et al (2012) Diagnostic values of ultrasound and the modified Alvarado scoring system in acute appendicitis. *Int J Emerg Med* 5:26
3. Dey S, Mohanta PK, Baruah AK et al (2010) Alvarado scoring in acute appendicitis—a clinicopathological correlation. *Indian J Surg* 72:290–293
4. Flum DR, McClure TD, Morris A et al (2005) Misdiagnosis of appendicitis and the use of diagnostic imaging. *J Am Coll Surg* 201:933–939
5. Kırkal C, Karabulut K, Aygen E et al (2013) Appendicitis scores may be useful in reducing the costs of treatment for right lower quadrant pain. *Turkish J Trauma Emerg Surg* 19:13–19
6. Memon AZ, Irfan S, Fatima K et al (2013) Acute appendicitis: diagnostic accuracy of Alvarado scoring system. *Asian J Surg* 36: 144–149
7. Alvarado A (1986) A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med* 15:557–564
8. Ohle R, O'Reilly F, O'Brien K et al (2011) The Alvarado score for predicting acute appendicitis: a systematic review. *BMC Med* 9:139
9. Al-Hashemy AM, Seleem MI (2004) Appraisal of the modified Alvarado score for acute appendicitis in adults. *Saudi Med J* 25: 1229–1231
10. Andersson M, Andersson RE (2008) The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. *World J Surg* 32:1843–1849
11. Scott AJ, Mason SE, Arunakirinathan M et al (2015) Risk stratification by the appendicitis inflammatory response score to guide decision-making in patients with suspected appendicitis. *Br J Surg* 102:563–572
12. Kollár D, McCartan DP, Bourke M et al (2015) Predicting acute appendicitis? A comparison of the Alvarado score, the appendicitis inflammatory response score and clinical assessment. *World J Surg* 39:104–109
13. Yıldırım E, Karagülle E, Kırbaş İ et al (2008) Alvarado scores and pain onset in relation to multislice CT findings in acute appendicitis. *Diagn Interv Radiol* 14:14–18
14. Ohmann C, Yang Q, Franke C (1995) Diagnostic scores for acute appendicitis. Abdominal Pain Study Group. *Eur J Surg* 161:273–281
15. Abou MB, Khalil M, Daoud N (2014) Comparison of Alvarado score evaluation and clinical judgment in acute appendicitis. *Med Arh* 68:10–13
16. Yüksel Y, Dinç B, Yüksel D et al (2014) How reliable is the Alvarado score in acute appendicitis? *Ulus Travma Acil Cerrahi Derg* 20:12–18
17. Chong CF, Thien A, Mackie AJ et al (2011) Comparison of RIPASA and Alvarado scores for the diagnosis of acute appendicitis. *Singapore Med J* 52:340–345
18. Lintula H, Pesonen E, Kokki H et al (2005) A diagnostic score for children with suspected appendicitis. *Langenbecks Arch Surg* 390: 164–170
19. Limpawattanasiri C (2011) Alvarado score for the acute appendicitis in a provincial hospital. *J Med Assoc Thai* 94:441–448
20. Kanumba ES, Mabula JB, Rambau P et al (2011) Modified Alvarado scoring system as a diagnostic tool for acute appendicitis at Bugando Medical Centre, Mwanza, Tanzania. *BMC Surg* 11:1–5