Evaluation of the Alvarado scoring system in the management of acute appendicitis

Zeki Özsoy, Erdinç Yenidoğan

Objective: In this study, we aimed to show the effectiveness of Alvarado score and its components to predict the correct diagnosis of acute appendicitis and to find an optimum cut-off value for Alvarado score.

Material and Methods: The patients who underwent surgical operation between January 2011 and January 2012 with the suspicion of acute appendicitis were included in the study. Their demographic and clinical features and histopathological results were retrieved from the medical records. They were divided into three groups according to their Alvarado scores. With the use of "receiver operating characteristic" curve analysis, the optimum cut-off value needed to make a correct diagnosis of acute appendicitis was determined.

Results: In all, 156 patients were included into the study. The mean age was 31.41±13.27 years. Histopathologically, acute appendicitis was detected in 125 (80.1%) patients, and negative appendectomy was found in 31 patients (19.8%). Mean Alvarado score was 6.44±1.49. There was a significant correlation between negative appendectomy and low Alvarado score (p<0.001). The main component of Alvarado score that makes the difference was rebound. Fever higher than 37.3°C, rebound, loss of appetite, and existence of shifting pain were statistically differential components (p=0.042, p<0.001, p=0.045, p<0.001, respectively). The rate of correct diagnosis of acute appendicitis was maximum in group 3 (100%) and minimum in group 1 (21.7%). Optimum cut-off value for Alvarado score was 7.

Conclusion: Patients with an Alvarado score of over 7 can be taken into surgical operation without the need of imaging methods.

Keywords: Acute appendicitis, Alvarado Score, cut-off value

INTRODUCTION
Acute appendicitis (AA) is the most common emergency condition requiring surgery in emergency admissions (1). Lifetime prevalence of the disease for individuals is 8% (2). Even with the frequent occurrence and the technical improvements in diagnostic modalities, the diagnosis of AA is challenging. It is mainly based on careful clinical history and physical examination (3). Many authors use imaging modalities when the diagnosis is doubtful. However, in rural areas of the country, it is difficult to evaluate the patient with ultrasonography or abdominal computed tomography under emergency conditions. Besides, unnecessary use of imaging modalities might cause a delay in diagnosis and also increase the economic burden (1, 4). Around 15%-30% of the patients who had previously undergone surgery with a diagnosis of AA were reported to have had negative appendectomy and 20% had perforation (5-8). Correct diagnosis is important to prevent perforation and negative appendectomy.

Therefore, clinical studies have aimed to find out the most cost-efficient, noninvasive, and practical diagnostic procedures. To this end several scoring systems have been proposed for AA (9). Alvarado score (AS) is the first and most widely known scoring method, the accuracy of which has been clinically approved (10). It uses mainly clinical findings and laboratory values (i.e., migration of pain, anorexia, nausea and/or vomiting, right lower quadrant tenderness, rebound, elevation of temperature ≥37.3°C, leukocytosis ≥10×10³/mm³, and polymorphonuclear neutrophilia ≥75%) to assess the presence of AA (3) (Table 1). It has been shown that AS prevents the delay in therapy and reduced the negative appendectomy rate (11). Although AS has a high accuracy rate, there is no sufficient study about the contribution of individual components to correct diagnosis. In this study, we aimed to show the effectiveness of AS and its components individually on accurately diagnosing patients with AA. Besides, we would like to suggest a “management algorithm” based on AS.

MATERIAL AND METHODS
This study was conducted in accordance with The Helsinki Declaration. Patients who underwent appendectomy due to AA between January 2011 and January 2012 at Gaziosmanpaşa University School of Medicine, Department of General Surgery were included in the study. Patients who have gynecological
pathology requiring surgical intervention with a diagnosis of AA, those whose histories could not be taken due to communication problems, pregnant patients, those taking painkillers, those who have comorbidities (such as diabetes mellitus, hypertension, and cardiac disease), and those who have incidental appendix tumors were excluded. Patients were clinically evaluated by the emergency medical team on call. Ultrasonography was requested based on physical examination and laboratory findings. In cases where the accuracy of diagnosis was uncertain, abdominal computed tomography was performed. Subsequently, the team consulted surgery for the diagnosis. Written informed consent was obtained from patients participated in this study. The operation was performed by the same surgical team laparoscopically or as open surgery. The demographic and clinical findings and histopathological characteristics were all retrieved from patient files. As were calculated retrospectively. The patients were divided into three groups depending on AS values (12).

**Group 1**: Patients with AS of ≤4, appendicitis unlikely
**Group 2**: Patients with AS of 5-7, appendicitis with moderate possibility
**Group 3**: Patients with AS of ≥8, appendicitis with high possibility

For individual components of AS, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were separately determined.

**Statistics Analysis**

For statistical evaluation, Statistical Package for the Social Sciences version 15.0 (SPSS Inc.; Chicago, IL, USA) was used. For descriptive data, numeric values were given as mean±standard deviation (SD) and categorical values as rate. Data distribution was analyzed with Kolmogorov-Smirnov. If the distribution was normal, Student T test was used, and if not, Mann-Whitney U test was used. Categorical values were compared using Chi-Square Tests. Performance of AS for determining AA was evaluated using “receiver operating characteristic” (ROC) curve. The field under ROC curve and its standard error were calculated. Optimum cut-off value for AA diagnosis was found. Sensitivity and specificity rates were determined for all ASs. p<0.05 was accepted as statistically significant.

**RESULTS**

**Demographic and Clinical Findings**

A total of 156 patients were included. Of these, 98 patients (62%) were male and 58 patients (38%) were female. Demographic and clinical features of the patients were shown in Table 2. There were 23 patients in Group 1, 86 patients in Group 2, and 47 patients in Group 3. The mean age of patients was 31.41±13.27 years. The mean AS was 6.44±1.49 in patients. Although histopathologically appendicitis was detected in 125 patients (80.1%), 31 patients (19.8%) had negative appendectomy. There was difference among AS groups in terms of average age (p<0.001). The post-hoc analysis showed a difference between Group 1 and Group 2. There was no difference in terms of sex between AS groups (p=0.513). The negative appendectomy rates in the three groups were 78%, 15%, and 0%, respectively. The difference among the groups was statistically significant (p<0.001).

**The Efficacy of AS Components in the Diagnosis**

Sensitivity, specificity, PPV, and NPV of each Alvarado components are shown in Table 3. Pain in the right lower quadrant was positive in all patients. Fever higher than 37.3°C, rebound, loss of appetite, and existence of migrating pain were the differential components that were statistically significant (p=0.042, p<0.001, p=0.045, and p<0.001, respectively).

**ROC Curve Analysis and Optimum Cut-Off Value**

The ROC curve for AS is shown in Figure 1. The area under the curve for AS was 0.853 (standard error (SE), 0.038; 95% confidence interval (CI), 0.779-0.927) (p<0.001). AS performance at each cut-off value is shown in Table 4. Cut-off value was found significant for 6 and 7 (p<0.05). For these cut-off values, sensitivity, specificity, PPV, and NPV with histopathological findings are shown in Table 5. ROC curve indicated that the optimum cut-off value of AS for correct diagnosis of appendicitis is 7.

---

**Table 1. Alvarado Scoring system**

<table>
<thead>
<tr>
<th>Component</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
<td></td>
</tr>
<tr>
<td>Migration of pain</td>
<td>1</td>
</tr>
<tr>
<td>Anorexia</td>
<td>1</td>
</tr>
<tr>
<td>Nausea and/or vomiting</td>
<td>1</td>
</tr>
<tr>
<td>Signs</td>
<td></td>
</tr>
<tr>
<td>Right lower quadrant tenderness</td>
<td>2</td>
</tr>
<tr>
<td>Rebound</td>
<td>1</td>
</tr>
<tr>
<td>Elevation of temperature ≥37.3°C</td>
<td>1</td>
</tr>
<tr>
<td>Laboratory</td>
<td></td>
</tr>
<tr>
<td>Leukocytosis ≥10×10³/mm³</td>
<td>2</td>
</tr>
<tr>
<td>Polymorphonuclear neutrophilia ≥75%</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 2. Demographic and clinical characteristics of the patients**

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Total</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients (n/%)</td>
<td>23/14.7</td>
<td>86/55.1</td>
<td>47/30.1</td>
<td>156/100</td>
<td></td>
</tr>
<tr>
<td>Gender (n/%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.513</td>
</tr>
<tr>
<td>Male</td>
<td>12/7.6</td>
<td>55/35.2</td>
<td>31/19.8</td>
<td>98/62.8</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>11/7</td>
<td>31/19.8</td>
<td>16/10.2</td>
<td>58/37.2</td>
<td></td>
</tr>
<tr>
<td>Age (mean±SD)</td>
<td>25.7±11.5</td>
<td>34.46±14.2</td>
<td>30.6±13.1</td>
<td>31.41±13.27</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AS (mean±SD)</td>
<td>3.60±0.58</td>
<td>6.09±0.80</td>
<td>8.48±0.68</td>
<td>6.44±1.49</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Appendicitis (-/+)(n)</td>
<td>18/5</td>
<td>13/73</td>
<td>0/47</td>
<td>31/125</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AS: Alvarado score; SD: standard deviation
Acute appendicitis is the most frequent reason for acute abdomen in all age groups (12). Diagnosis can often be made solely with physical examination and laboratory findings. In suspicious cases, radiological methods are used. However, they have some disadvantages such as they are not available under rural conditions, some use radiation, and they are not cost-effective. To reduce the cost and prevent delay in diagnosis, some scoring systems such as Lintula, Madan, Ohmann, Eskelinen, De Dombal, and Alvarado were suggested. Common features of these scoring systems are that they aimed to be practical, noninvasive, and do not require much expertise to use (9). AS is a scoring system that includes clinical and laboratory findings of patient. AS has been shown to predict appendicitis accurately in many clinical studies. In their prospective study, Kariman et al. (13) evaluated AS in 300 patients who were admitted with abdominal pain and randomized them into two groups depending on their score with the cut-off value of 7. They found that AS can be used in patients with abdominal pain reliably and a score of more than 7 indicates AA possibility at the rate of 93% and a negative test (i.e., AS of ≤7) indicates unlikeliness of AA at the rate of 26%. As a result, they stated that AS is a reliable test for the diagnosis of appendicitis but insufficient to exclude the diagnosis. Chan et al. (14) found that the incidences of AA in patients with an AS of 9-10, 7-8, and ≤6 were 100%, 68%, and 5%, respectively. When AS and histopathological findings in the groups in our study were compared, AA possibility for groups shows a resemblance to the literature.

Eskelinen, De Dombal, and Alvarado were suggested. Common features of these scoring systems are that they aimed to be practical, noninvasive, and do not require much expertise to use (9). AS is a scoring system that includes clinical and laboratory findings of patient. AS has been shown to predict appendicitis accurately in many clinical studies. In their prospective study, Kariman et al. (13) evaluated AS in 300 patients who were admitted with abdominal pain and randomized them into two groups depending on their score with the cut-off value of 7. They found that AS can be used in patients with abdominal pain reliably and a score of more than 7 indicates AA possibility at the rate of 93% and a negative test (i.e., AS of ≤7) indicates unlikeliness of AA at the rate of 26%. As a result, they stated that AS is a reliable test for the diagnosis of appendicitis but insufficient to exclude the diagnosis. Chan et al. (14) found that the incidences of AA in patients with an AS of 9-10, 7-8, and ≤6 were 100%, 68%, and 5%, respectively. When AS and histopathological findings in the groups in our study were compared, AA possibility for groups shows a resemblance to the literature.
Table 5. Alvarado score performance at cutoff values of 6 and 7 with regard to histopathology

<table>
<thead>
<tr>
<th>Positive if greater than or equal to</th>
<th>AA+</th>
<th>AA-</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>PPV %</th>
<th>NPV %</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>100</td>
<td>9</td>
<td>80</td>
<td>70.9</td>
<td>91.7</td>
<td>46.8</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>7*</td>
<td>74</td>
<td>5</td>
<td>59.2</td>
<td>83.9</td>
<td>93.7</td>
<td>33.8</td>
<td></td>
</tr>
</tbody>
</table>

AA: acute appendicitis; PPV: positive predictive value; NPV: negative predictive value

*Optimum cut-off value

The authors concluded that patients with AS of $\leq 4$ can be discharged from emergency department with the suggestion of distant follow-up. Yildirim et al. (3) stated that patients with AS between 5 and 7 should be further examined using imaging modalities; however, only AS is not enough for the decision of surgery. Similarly, McKay and Shepherd (18) concluded that imaging methods should be used in patients with an AS of 5-7 for diagnosis and patients with an AS of $\geq 7$ should directly undergo operation without any more workup. Alvarado himself defends that patients with a score of 7 should undergo surgery. In our study, the optimum cut-off value for AS in AA diagnosis after ROC curve analysis was found to be 7. All patients with an AS of $> 7$ were found to have histopathological appendicitis. The positive predictive value of AA for optimum cut-off value was 93.7%. Similar to ours, in the study of Pogorelic et al. (19) the optimum cut-off value for AS was 7 and the PPV for this value was 93.1%.

In our study, the negative appendectomy rate was 19.8%. This rate was found to be 32% by Menteş et al. (20), and 7% by Jo et al. (21). In the literature, there are few studies giving negative appendectomy rates for AS groups separately. In a previous study, Yüksel et al. (12) found negative appendectomy rates in the groups (with increasing order with AS score) to be 43.5%, 24.1%, and 11%, respectively. Although negative appendectomy rates were similar in our study and Yüksel et al. (12) study, it was higher than the average in the groups with AS of $\leq 4$.

Based on our findings, we agree with the recommendations put forward in the literature. In short, patients with an AS score of $\leq 4$ can be discharged and followed at home with the necessary information from the patient. Patients with an AS of 5-7 should be followed closely using imaging methods as they probably have AA. We believe that the patients with an AS of $\geq 8$ can undergo operation without imaging methods (Figure 2). Through the literature data and our findings, this approach has been shown to increase the accuracy of diagnosis (especially in case of unavailability of radiological imaging), decrease the negative appendectomy rates, and provide an advantage in terms of cost and duration.

According to the results of this study, the main component of AS exhibiting the difference was rebound. Among the scoring components, fever higher than 37.3°C, rebound, loss of appetite, high fever, and high neutrophil count into account during decision-making may reduce the negative appendectomy rate in the patients with AA diagnosis.

Our study has some limitations. First, the study was retrospective in nature. Second, AS was evaluated only in patients who underwent operation. However, there were also some patients who presented at the emergency unit with nonspecific abdominal pain without proceeding to the operation. We have no data about their follow-up and treatment. In this study, physical examination, laboratory findings, and imaging methods were used for surgical decision-making. We have not compared AS solely with any of the radiological imaging methods. AS components were examined, but they were not effective for surgical decision-making. The results of our study showed that AS system can be used in the diagnosis of AA just like all the other diagnostic methods. The effectiveness of individual AS components on making a diagnosis and defining an optimum cut-off value was defined.

**CONCLUSION**

Thus, there may be difficulties in diagnosis of AA sometimes and delayed diagnosis may affect treatment negatively. AS system is one of the most frequently used scoring systems. Among the scoring components, fever higher than 37.3°C, rebound, loss of appetite, and existence of shifting pain make a positive contribution to the diagnosis. We believe that patients with an AS over 7 should be operated without the use of imaging methods.

**Ethics Committee Approval:** Authors declared that the research was conducted according to the principles of the World Medical Association Declaration of Helsinki “Ethical Principles for Medical Research Involving Human Subjects” (amended in October 2013).

**Informed Consent:** Written informed consent was obtained from patients who participated in this study.

**Peer-review:** Externally peer-reviewed.

**Author Contributions:** Concept - Z.O., E.Y.; Design - Z.O., E.Y.; Supervision - Z.O., E.Y.; Resource - Z.O.; Materials - Z.O.; Data Collection and/or Processing - Z.O.; Analysis and/or Interpretation - Z.O.; Literature Search - Z.O.; Writing Manuscript - Z.O.; Critical Reviews - E.Y.

**Acknowledgements:** Due to his contributions to the study we thank Abdullah Özugür Yeniova from, Gaziosmanpaşa University School of Medicine, Gastroenterology Department.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** The authors declared that this study has received no financial support.

**REFERENCES**


12. Yüksel Y, Dinç B, Yüksel D, Dinç SE, Mesci A. How reliable is the alvarado score in acute appendicitis?. Ulus Travma Acil Cerrahi Derg 2014; 20: 12-18. [CrossRef]


16. Denizbasi A, Unluer EE. The role of the emergency medicine resident using the alvarado score in the diagnosis of acute appendicitis compared with the general surgery resident. Eur J Emerg Med 2003; 10: 296-301. [CrossRef]


