Retrospective Evaluation of the Diagnostic Accuracy of the Modified Alvarado Scoring System (MASS) in a Croatian Hospital

Branko Bakula¹, Vanja Radišić Biljak², Mirko Bakula³, Fran Rašić⁴, Ana-Maria Šimundić²

¹Surgery Clinic, University Hospital “Sveti Duh”, Zagreb, Croatia, ²Department of Medical Laboratory Diagnostics, University Hospital “Sveti Duh”, Zagreb, Croatia, ³Department of Urology, University Hospital Centre Zagreb, Zagreb, Croatia, ⁴School of Medicine, University of Zagreb, Zagreb, Croatia

Correspondence: vanja.radisic@gmail.com
Tel.: + 385 1 3712 118

Received: 10 December 2018
Accepted: 26 April 2019

Key Words: Acute Appendicitis • Diagnostic Accuracy • Modified Alvarado Scoring System (MASS).

Introduction

Acute appendicitis (AA) represents the most common indication for emergency abdominal surgery (1). The lifetime risk of developing appendicitis is approximately 7%. When diagnosed early and treated on time, morbidity from acute appendicitis is very low. On the other hand, unrecognized appendicitis with delayed surgical intervention can lead to major complications or even death (2, 3). This is the reason why surgeons usually have a defensive approach when encountering patients with suspected appendicitis, which results in relatively high rate of negative appendectomies (8-28%) (4-6).

Diagnosing appendicitis is not always simple. The typical clinical presentation with migration of pain to the lower right quadrant of the abdomen or localized peritonitis is only found in less than half of the patients with acute appendicitis (7, 8). Besides el-
evated inflammatory parameters, such as leukocytes and CRP, no specific laboratory marker for appendicitis exists yet. An abdominal CT-scan recognizes acute appendicitis with high accuracy, but the high dosage of harmful radiation makes it absolutely unacceptable for routine use in patients suspected of having appendicitis, especially considering the high rate in the younger female population (9, 10).

All these issues have encouraged surgeons to develop a scoring system that would potentially decrease the rate of misdiagnosis in patients with suspected appendicitis. Although a variety of different scoring systems for acute appendicitis have been proposed, the most widely used today is the Alvarado scoring system (11). It was first devised in 1986 for pregnant women, and later it was also used in the general population. This scoring system, with a maximum total score of ten, is based on six clinical and two laboratory parameters, where leukocytosis and right iliac fossa tenderness are considered to be the most important factors, and are therefore assigned two points (Table 1). In many medical centers differential blood count is not part of routine laboratory investigations for patients with abdominal pain. For this reason Kalan et al. proposed a modified Alvarado scoring system (MASS) that uses the same value categories but without the left shift of leukocytes, so the scores range from 0 to 9 (Table 1) (12).

Considering the fact that the sensitivity and specificity of MASS for diagnosing appendicitis at the recommended cut off point of ≥7 varies a great deal between studies (13-18), the present situation at our University Hospital does not include an initial evaluation of patients with suspected acute appendicitis using any of the proposed scoring systems. Our hypothesis includes reversing the original paradigm of the Alvarado scoring system – we propose applying the Alvarado scoring system to evaluate patients who have already been designated to undergo surgery, and so assist surgeons in supporting their initial decision.

Therefore we aimed retrospectively to assess the diagnostic accuracy of the modified Alvarado score in operated patients. Additionally, we wanted to propose a different MASS cut-off point which would be a useful tool to decrease the number of false negatives to the lowest possible number.

Table 1. Alvarado Score and Modified Alvarado Scoring System (MASS) for Diagnosis of Acute Appendicitis (11, 12)

<table>
<thead>
<tr>
<th>Alvarado Score</th>
<th>MASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
<td>Point</td>
</tr>
<tr>
<td>Migratory right iliac fossa pain</td>
<td>1</td>
</tr>
<tr>
<td>Anorexia</td>
<td>1</td>
</tr>
<tr>
<td>Nausea/vomiting</td>
<td>1</td>
</tr>
<tr>
<td>Signs</td>
<td></td>
</tr>
<tr>
<td>Tenderness in right iliac fossa</td>
<td>2</td>
</tr>
<tr>
<td>Rebound tenderness</td>
<td>1</td>
</tr>
<tr>
<td>Elevated temperature (&gt;37.3°C)</td>
<td>1</td>
</tr>
<tr>
<td>Laboratory findings</td>
<td></td>
</tr>
<tr>
<td>Leukocytosis (WBC &gt;10x10^9/L)</td>
<td>2</td>
</tr>
<tr>
<td>Shift to the left of neutrophiles (&gt;70%)</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
</tr>
</tbody>
</table>

MASS=Modified Alvarado Scoring system.
Materials and Methods

Subjects

We retrospectively analyzed all adult patients who underwent emergency appendectomy due to clinically suspected acute appendicitis at the University Hospital “Sveti Duh” (Zagreb, Croatia) in the period between July 2015 and March 2017. The decision to operate was made by the senior attending surgeon, based on clinical judgment (physical examination, imaging and blood test). No diagnostic scoring systems were used pre-operatively. The exclusion criterion was the absence of the data needed to calculate the MASS. The Hospital Ethical Committee granted approval for the retrospective analysis of the study data.

Methods

The MASS was retrospectively calculated for all patients based on the clinical signs and laboratory investigations obtained during initial admission in the emergency room.

Statistical Analysis

All clinical variables, except age, are given in absolute number and percentages. Age is given as the median and range, and was tested between groups with the Mann-Whitney test. The clinical and demographic variables were compared with a comparison of the proportions test. The difference between categorical data was tested with Fisher’s exact test. Receiver operating characteristic (ROC) curve analysis was used to determine the optimal MASS score cut-off value for the amended purpose, and to calculate the area under the curve (AUC), sensitivity, and the specificity of the MASS, with a respective 95% confidence interval (CI). Statistical analyses were performed using MedCalc Statistical Software version 16.2.0 (MedCalc Software bvba, Ostend, Belgium). P<0.05 was defined as the threshold of significance.

Results

During the study period, from July 2015 to March 2017, 153 patients underwent emergency appendectomy. Seven patients did not have all the relevant medical data to calculate the MASS score, thus leaving 146 patients eligible for inclusion in the study (76 males and 70 females).

In 118 out of 146 operated patients, acute appendicitis was confirmed intra-operatively, yielding a negative appendectomy rate of 19.2% at our Hospital. Sixty-eight patients underwent laparoscopic appendectomy, 75 patients underwent open appendectomy, and in 3 patients conversion from laparoscopic to open procedure was performed.

The study population was divided into two groups, depending on the rejected (negative appendectomies) or confirmed diagnosis of AA. Comparison of the basic demographic characteristics of the study subjects is shown in Table 2.

Among 28 patients with negative appendectomy, 4 cases had pelvic inflammatory disease, 2 cases had mesenterial lymphadenitis, 2 cases had enterocolitis and 1 case each had coecal diverticulitis, sigmoidal diverticulitis, Meckel’s diverticulitis, terminal ileitis, inflamed sigmoidal cancer and epiploic appendicitis. In 14 cases the cause of symptoms was unknown (Table 3).

The retrospectively calculated MASS score for each patient, showed a statistically significant difference between the average MASS score in the positive and negative appendectomy groups (6 v. 4, respectively), P<0.001. Although there were significantly more women in the negative appendectomy group (Table 2), no difference was observed in the average MASS score between genders (0.292).
According to the current recommendations regarding the MASS score cut-off values (≥7), the sensitivity and specificity measured were 44.1% (95% CI: 34.9 – 53.5) and 89.3% (95% CI: 71.8 – 97.7), respectively. However, the ROC curve analysis identified a MASS cut-off value of >5 to achieve the optimum combination of both sensitivity (68.6%, 95% CI: 59.5 – 76.9) and specificity (71.4%, 95% CI: 51.3 – 86.8) in identifying the patients with AA. The area under the ROC curve revealed good diagnostic accuracy ((AUC) 0.745 (95% CI: 0.666 – 0.814), P<0.001). If the cut-off is lowered to values as low as ≥3, the sensitivity of the MASS score increases to 97.45% (95% CI: 92.7 – 99.5), thus obtaining a very low false negative rate of merely 2.55%.

### Discussion

This retrospective diagnostic accuracy study confirmed the higher average MASS score in the group of patients with confirmed AA diagnosis compared to the negative appendectomy group. The ROC analysis identified a lower optimal cut-off value of ≥3 for the amended purpose of the Modified Alvarado Scoring System, thus increasing the sensitivity to a respectable 97.45%.

Our study confirmed the literature data about the difficulties in diagnosing AA in women, especially in childbearing age. There were significantly more women in the negative appendectomy group, however no difference was observed in the average MASS between men and women in that particular group. Perhaps it would be interesting to explore this point through prospective evaluation in female patients only.

When deciding about the optimal cut-off point, the real question is whether this scoring system could assist surgeons reduce the risk of missing diagnosis by identifying patients with suspected acute appendicitis. Considering the variable literature data about the diagnostic accuracy of MASS, it

### Table 2. Comparison of Basic Demographic Characteristics Between Groups of Patients with Confirmed / Negative Appendectomies Diagnosis of Acute Appendicitis

<table>
<thead>
<tr>
<th>Patient (N=146)</th>
<th>Appendectomies</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Negative (N=28)</td>
<td>Confirmed (N=118)</td>
</tr>
<tr>
<td>Male, N (ratio)</td>
<td>3 (0.04)</td>
<td>73 (0.96)</td>
</tr>
<tr>
<td>Female, N (ratio)</td>
<td>25 (0.36)</td>
<td>45 (0.64)</td>
</tr>
<tr>
<td>Age, years</td>
<td>38.5 (20 – 79)</td>
<td>37 (20 – 85)</td>
</tr>
</tbody>
</table>

The results are presented as absolute numbers and ratio, except age which is given in median and range (min-max); P<0.05 is considered statistically significant; The P value corresponds to the difference between the groups in columns, except for age. There was no statistically significant difference in mean age between the patients subdivided according to the postoperative outcome (P=0.720). However, there was a significantly higher number of females among the negative appendectomies (P<0.001), and a significantly higher number of males in the AA group (P<0.001).

### Table 3. Causes of Acute Abdominal Pain in the Group of Negative Appendectomy Patients

<table>
<thead>
<tr>
<th>Negative appendectomy group</th>
<th>Diagnosis</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pelvic inflammatory disease</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Enterocolitis</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Lymphadenitis mesenterialis</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Coecal diverticulitis</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sigmoidal diverticulitis</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Meckel’s diverticulitis</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Terminal ileitis</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Inflamed sigmoidal cancer</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Appendicitis epiploica</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28</td>
</tr>
</tbody>
</table>
is not likely that the MASS will become the
dominant criteria in management of patients
with acute abdominal pain at our Hospital.
More likely the final decision will always be
in the domain of the attending surgeon. In
that light, and considering the results ob-
tained, perhaps it would be wiser to aim at
higher sensitivity. To increase the sensitivity
of the MASS score further and thus help
the attending surgeons in ruling out the true
negative patients, we wanted to calculate the
measure of diagnostic accuracy at a very
low cut-off value of ≥3 MASS points. When
applying this proposed lower cut-off value
of ≥3 on our retrospective study group, we
obtained a very low false negative rate of
merely 2.55%.

Conclusion

Although our study has some shortcomings,
it represents a valuable insight into the initial
official data about the diagnostic accuracy of
MASS at one Croatian University Hospital.
In conclusion, we propose the use of MASS
at a cut-off of ≥3 as a useful tool to help sur-
geons ruling in patients with AA in order to
reduce the risk of missing diagnosis.

What Is Already Known on this Topic
Unrecognized appendicitis with delayed surgical intervention can lead to major complications or even death. Diagnosing appendicitis is not always simple. A variety of different scoring systems for acute appendicitis have been proposed, and the most widely used today is the Modified Alvarado scoring system. The sensitivity and specificity of the Modified Alvarado scoring system for diagnosing appendicitis at a recommended cut-off point of ≥7 varies a great deal between studies.

What this Study Adds
We aimed to assess retrospectively the diagnostic accuracy of the modified Alvarado score in operated patients and to pro-
duce a different MASS cut-off point for the amended purpose. Following the current recommendations regarding the MASS score cut-off values (≥7), the measured sensitivity and specificity were 44.1% (95% CI: 34.9 – 53.5) and 89.3% (95% CI: 71.8 – 97.7), respectively. We identified a MASS cut-off value of >5 to achieve the optimum combination of both sensitivity (68.6%, 95% CI: 59.5 – 76.9) and specificity (71.4%, 95% CI: 51.3 – 86.8) in identifying the patients with AA. By lowering the cut-off to values as low as ≥3, the sensitivity of the MASS score increased to a respectable 97.45% (95% CI: 92.7 – 99.5), thus obtaining a very low false negative rate of merely 2.55%.

Authors’ Contributions: Conception and design: BB and VRB; Acquisition, analysis and interpretation of data: BB, VRB, MB, FR and AMŠ; Drafting the article: BB and VRB; Revising it critically for important intellectual content: AMŠ; Approved final version of the manuscript: BB, VRB, MB, FR and AMŠ.

Conflict of Interest: The authors declare that they have no conflict of interest.

References