

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

Branko Bakula<sup>1</sup>, Vanja Radišić Biljak<sup>2</sup>, Mirko Bakula<sup>3</sup>, Fran Rašić<sup>4</sup>, Ana-Maria Šimundić<sup>2</sup>

<sup>1</sup>Surgery Clinic, University Hospital "Sveti Duh", Zagreb, Croatia, <sup>2</sup>Department of Medical Laboratory Diagnostics, University Hospital "Sveti Duh", Zagreb, Croatia, <sup>3</sup>Department of Urology, University Hospital Centre Zagreb, Zagreb, Croatia, <sup>4</sup>School 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

**Objective.** Diagnosing acute appendicitis (AA) is challenging and this has stimulated surgeons to develop scoring systems that could potentially decrease the rate of misdiagnosis in patients with suspected appendicitis. One of the most widely used today is the Modified Alvarado scoring system (MASS), however its sensitivity and specificity varies a great deal between studies. As a result, we wanted to assess the diagnostic accuracy of MASS retrospectively at our institution to achieve the highest possible value of sensitivity and decrease the number of false negative patients. **Material and Methods.** We retrospectively calculated MASS for all subsequent patients who had undergone an appendectomy at our institution between July 2015 and March 2017. **Results.** In 118 out of 146 operated patients, AA was confirmed intraoperatively. There was a statistically significant difference between the average MASS score in the positive and negative appendectomy groups (6 v. 4, respectively,  $P < 0.001$ ), with a significantly higher number of females among the negative appendectomies ( $P < 0.001$ ). When lowering the cut-off to a value as low as  $\geq 3$ , the sensitivity of the MASS score increased to 97.45% (95% CI: 92.7 – 99.5), thus obtaining a very low false negative rate of merely 2.55%. **Conclusion.** This retrospective diagnostic accuracy study confirmed the higher average MASS score in the group of patients with confirmed AA diagnosis. A MASS score above the proposed low cut-off value ( $\geq 3$ ) can be a useful tool to help surgeons ruling in patients with AA in order to reduce the risk of missing diagnosis.

(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 in-

vestigations 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  $\geq 7$  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)

Alvarado Score		MASS	
Symptoms	Point	Symptoms	Point
Migratory right iliac fossa pain	1	Migratory right iliac fossa pain	1
Anorexia	1	Anorexia	1
Nausea/vomiting	1	Nausea/vomiting	1
Signs		Signs	
Tenderness in right illiac fossa	2	Tenderness in right illiac fossa	2
Rebound tenderness	1	Rebound tenderness	1
Elevated temperature ( $>37,3^{\circ}\text{C}$ )	1	Elevates temperature ( $>37,3^{\circ}\text{C}$ )	1
Laboratory findings		Laboratory findings	
Leukocytosis ( $\text{WBC} >10 \times 10^9/\text{L}$ )	2	Leukocytosis ( $\text{WBC} >10 \times 10^9/\text{L}$ )	2
Shift to the left of neutrophils ( $>70\%$ )	1	-	-
Total	10	Total	9

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).

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

Patient (N=146)	Appendectomies		P*
	Negative (N=28)	Confirmed (N=118)	
Male, N (ratio)	3 (0.04)	73 (0.96)	-
Female, N (ratio)	25 (0.36)	45 (0.64)	-
P*	<0.001	<0.001	-
Age, years	38.5 (20 – 79)	37 (20 – 85)	0.720

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

Negative appendectomy group	
Diagnosis	N
Pelvic inflammatory disease	4
Enterocolitis	2
Lymphadenitis mesenterialis	2
Coecal diverticulitis	1
Sigmoidal diverticulitis	1
Meckel's diverticulitis	1
Terminal ileitis	1
Inflamed sigmoidal cancer	1
Appendicitis epiploica	1
Unknown	14
Total	28

According to the current recommendations regarding the MASS score cut-off values ( $\geq 7$ ), 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  $\geq 3$ , 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  $\geq 3$  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

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 obtained, 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  $\geq 3$  MASS points. When applying this proposed lower cut-off value of  $\geq 3$  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  $\geq 3$  as a useful tool to help surgeons 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  $\geq 7$  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 propose a different MASS cut-off point for the amended purpose. Following the current recommendations regarding the MASS score cut-off values ( $\geq 7$ ), 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  $\geq 3$ , 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

1. Ferris M, Quan S, Kaplan BS, Molodecky N, Ball CG, Chernoff GW, et al. The Global Incidence of Appendicitis: A Systematic Review of Population-based Studies. *Ann Surg.* 2017;266(2):237-41.
2. Asad S, Ahmed A, Ahmad S, Ali S, Ahmed S, Ghaffar S, et al. Causes of delayed presentation of acute appendicitis and its impact on morbidity and mortality. *J Ayub Med Coll Abbottabad.* 2015;27(3):620-3.
3. Adamu A, Maigatari M, Lawal K, Iliyasu M. Waiting time for emergency abdominal surgery in Zaria, Nigeria. *Afr Health Sci.* 2010;10(1):46-53.
4. SCOAP Collaborative, Cuschieri J, Florence M, Flum DR, Jurkovich GJ, Lin P, et al. Negative appendectomy and imaging accuracy in the Washington State Surgical Care and Outcomes Assessment Program. *Ann Surg.* 2008;248(4):557-63.
5. Antevil J, Rivera L, Langenberg B, Brown CV. The influence of age and gender on the utility of computed tomography to diagnose acute appendicitis. *Am Surg.* 2004;70(10):850-3.
6. Körner H, Söndena K, Söreide JA, Andersen E, Nysted A, Lende TH, et al. Incidence of acute nonperforated and perforated appendicitis: age-specific and sex-specific analysis. *World J Surg.* 1997;21(3):313-7.
7. Laméris W, van Randen A, Go PM, Bouma WH, Donkervoort SC, Bossuyt PM, et al. Single and combined diagnostic value of clinical features and laboratory tests in acute appendicitis. *Acad Emerg Med.* 2009;16(9):835-42.
8. Castro ADAE, Skare TL, Yamauchi FI, Tachibana A, Ribeiro SPP, Fonseca EKUN, et al. Diagnostic value of C-reactive protein and the influence of visceral fat in patients with obesity and acute appendicitis. *Arq Bras Cir Dig.* 2018;31(1):e1339.
9. Anderson SW, Soto JA, Lucey BC, Ozonoff A, Jordan JD, Ratevosian J, et al. Abdominal 64-MDCT

- for suspected appendicitis: the use of oral and IV contrast material versus IV contrast material only. *AJR Am J Roentgenol.* 2009;193(5):1282-8.
10. Costello JE, Cecava ND, Tucker JE, Bau JL. CT radiation dose: current controversies and dose reduction strategies. *AJR Am J Roentgenol.* 2013;201(6):1283-90.
  11. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* 1986;15(5):557-64.
  12. Kalan M, Talbot D, Cunliffe WJ, Rich AJ. Evaluation of the modified Alvarado score in the diagnosis of acute appendicitis: a prospective study. *Ann R Coll Surg Engl.* 1994;76(6):418-9.
  13. Kanumba ES, Mabula JB, Rambau P, Chalya PL. Modified Alvarado Scoring System as a diagnostic tool for acute appendicitis at Bugando Medical Centre, Mwanza, Tanzania. *BMC Surg.* 2011;11:4.
  14. Shuaib A, Shuaib A, Fakhra Z, Marafi B, Alsharaf K, Behbehani A. Evaluation of modified Alvarado scoring system and RIPASA scoring system as diagnostic tools of acute appendicitis. *World J Emerg Med.* 2017;8(4):276-80.
  15. Díaz-Barrientos CZ, Aquino-González A, Heredia-Montaña M, Navarro-Tovar F, Pineda-Espinosa MA, Espinosa de Santillana IA. The RIPASA score for the diagnosis of acute appendicitis: A comparison with the modified Alvarado score. *Rev Gastroenterol Mex.* 2018;83(2):112-6.
  16. Kamran H, Naveed D, Asad S, Hameed M, Khan U. Evaluation of modified Alvarado score for frequency of negative appendicectomies. *J Ayub Med Coll Abbottabad.* 2010;22(4):46-9.
  17. Meltzer AC, Baumann BM, Chen EH, Shofer FS, Mills AM. Poor sensitivity of a modified Alvarado score in adults with suspected appendicitis. *Ann Emerg Med.* 2013;62(2):126-31.
  18. Sobnach S, Ede C, Van Der Linde G, Klopper J, Thomson S, Bhyat A, et al. A retrospective evaluation of the Modified Alvarado Score for the diagnosis of acute appendicitis in HIV-infected patients. *Eur J Trauma Emerg Surg.* 2018;44(2):259-63.