We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,400 Open access books available
117,000 International authors and editors
130M Downloads

154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE
Selection of our books indexed in the Book Citation Index in Web of Science Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected. For more information visit www.intechopen.com
Chapter 3

Diagnostic Scores in Acute Appendicitis

Alfredo Alvarado

Additional information is available at the end of the chapter
http://dx.doi.org/10.5772/intechopen.77230

Abstract
Diagnostic scores should be part of the initial evaluation of patients suspected of acute appendicitis. This approach could be very helpful in order to make an early diagnosis and to stratify the cases for observation, further investigation, or surgical intervention.

Keywords: diagnostic scores in acute appendicitis, Alvarado score, clinical approach

1. Introduction

Several scoring systems have been developed to help clinicians in the diagnosis of acute appendicitis. The best-known scores are the Alvarado score, the modified Alvarado score, the Pediatric Appendicitis Score, the Appendicitis Inflammatory Response score, and the RIPASA score. These tools not only can be used for diagnostic purposes but also for stratification, separating those patients who require observation and workup from those who can be assigned for certain specific treatment. The aim of these scores is to reduce the number of negative appendectomies without increasing the number of perforations.

The Alvarado score was described in 1986 [1] and since then has been evaluated and validated in many studies. It consists of three symptoms, three clinical signs, and two laboratory tests. This system uses a simple mnemonics (MANTRELS) that is easy to remember and can be applied in many settings without the need of a computer. The symptoms are migration (one point), anorexia-acetonuria (one point), and nausea/vomiting (one point). The clinical signs are tenderness in the right lower quadrant (two points), rebound pain (one point), and elevation of oral temperature (37.3°C or more) (one Point).
The basic laboratory tests are a complete blood count (CBC) to look for leukocytosis (>10,000 cells/mm$^3$) and a differential white blood count (WBC) looking for left shift (increased stabs >5% or segmented neutrophils >75%). A urinalysis is useful to determine if there is acetone, which indicates the presence of a fasting state related to anorexia, and also, it may show many red cells due to an inflammatory process around the appendix. If the urine shows too many red cells, it may point to a ureteral calculus, and further investigation should be done. The C-reactive protein (CRP) test is not included in the score because it is a nonspecific test that detects an inflammatory process only and is not diagnostic for any particular condition. Besides this, it would be a redundancy since the shift to the left and leukocytosis are doing the same thing. Furthermore, it will not help in the initial stages of acute appendicitis because it will defeat the purpose of the score, that is to say, to make an early diagnosis of acute appendicitis.

Direct tenderness on the right lower quadrant can be replaced by direct percussion with the fist, as a mallet, on the right lumbar area in cases of retrocecal appendicitis which occurs in 75–85% of cases.

Rebound pain can be replaced by other indirect signs such as the Rovsing sign, Dunphy sign (cough test) or the Markle’s test (heel-drop jarring test), pain on walking, pain with jolts or bumps in the road, and the inspiration test. Uncommon tests of peritoneal irritation such as the psoas and the obturator tests can replace the rebound pain test also. In children who are unable to communicate well, cutaneous hyperesthesia can be added to replace the migration symptom.

In order of decreasing importance, the best predictive factors proved to be localized tenderness on the right lower quadrant, leukocytosis, migration of pain, shift to the left, temperature elevation, nausea or vomiting, anorexia or acetone in the urine, and direct rebound pain. Two points are assigned to the more important factors (tenderness and leukocytosis) and a value of 1 for each one of the others, for a possible total score of 10. A score of 4–5 is compatible with the diagnosis of acute appendicitis, a score of 7 or 8 indicates a probable appendicitis, and a score of 9 or 10 indicates a very probable appendicitis. To this score the clinician could subtract two points if the patient complains of headache because this symptom is very rare in cases of acute appendicitis. In this particular situation, the patient may need further investigation to rule out a different disorder.

Scores of 5 or 6 are in a gray area, and in this case, the clinician may want to observe the patient for a short time (reevaluate every 4–6 hours) for 12–24 hours, and if the score remains, the same consider other tests such as ultrasound or diagnostic laparoscopy. When the score is 3 or 4, the clinician has two options: the patient could be kept under observation and repeat the tests or, even more, order additional tests such as an US or a CT scan if they are available in that particular setting. Another option is to rely on the clinical impression of the examiner because, as I already mentioned in my original article, “there is always an intangible ingredient in the diagnosis of acute appendicitis.”

The modified Alvarado score (MAS) [2] is a simplification of the Alvarado score by eliminating the neutrophil count because a differential WBC count is not available in certain facilities. The results are similar to the original score but with less capacity to detect the early stages of acute appendicitis.
The Pediatric Appendicitis Score (PAS), developed by Samuel in 2002 [3], is a modification of the Alvarado score in which the rebound sign has been replaced by cough/percussion/hopping tenderness in the right lower quadrant, and the elevation of temperature has been increased to 38°C. In this score the sign of tenderness in the right lower quadrant, the most relevant feature of the score, was given one point only.

The Appendicitis Inflammatory Response (AIR) score [4] is based along the same principles of the Alvarado score assigning patients to low, medium, or high probability of acute appendicitis. It was developed by Andersson and Andersson in 2008 and was constructed from eight independent variables (right lower quadrant pain, rebound tenderness, muscular defense, WBC count, proportion of neutrophils, CRP, body temperature, and vomiting). The AIR score contains rebound tenderness or muscular defense that is divided in three groups—light, medium, and strong—which makes these signs subjective and very difficult to evaluate, and this may deviate the final score one way or another. Besides this, the AIR score omits the symptom of migration of pain which is a very important and specific symptom in the diagnosis of acute appendicitis.

The Raja Isteri Pengiran Anak Saleha Appendicitis (RIPASA) score [5] was developed for the diagnosis of acute appendicitis in Brunei, Darussalam, in 2008. It contains 14 patient characteristics: gender, age, and symptoms, right iliac fossa (RIF) pain, migration to the RIF, anorexia, nausea and vomiting, duration of symptoms, and clinical signs RIF tenderness, guarding, rebound tenderness, Rovsing sign, and fever. It also contains two laboratory tests (WBC and urinalysis) and an additional parameter related to a foreign national card record. Some authors found that the Alvarado score was disappointing in the diagnosis of acute appendicitis in Asian and Mid-Eastern populations, so they decided to have a different score more suitable to them. Chong [6] found that the RIPASA score of >7.5 correctly classified 98% patients confirmed with histological findings of acute appendicitis in comparison with 68.3% patients with an Alvarado score of >7. However, RIPASA and Alvarado scores correctly classified 81.3% and 87.9% patients without acute appendicitis into the true negative groups with scores of >7.5 and <7, respectively. The negative appendectomy rate was 14.66% for the RIPASA score and 13.75% for the Alvarado score.

Khadda et al. [7] found that the RIPASA score has a sensitivity of 97.7% and a specificity of 77.4% and a negative appendectomy rate of 13.7% which is higher than many reports that had used the Alvarado score such as Menon et al. [8], in Pakistan, who reported a negative appendectomy rate of 1.9%. In other study, Pouget-Baudry et al. [9], in France, reported 3 out of 174 patients with a normal appendix on histological examination which equals to 1.72%. The good thing is that Khadda recognized that the Alvarado score is the simplest of all the scores used in current practice. Furthermore, Gaikwad et al. [10], in India, found that the false-positive rate is reduced to zero when ultrasonography is added to the Alvarado score.

Goel et al. [11], in India, evaluated the efficacy of the Alvarado score and the RIPASA score finding that the Alvarado score has a better specificity than the RIPASA score (100 vs. 50%) and also a better negative appendectomy rate (0 vs. 5%). Similar results were reported by Karami et al. [12], in Iran, who found that the Alvarado score was 100% specific as compared with the RIPASA and the AIR scores (91.6% for both).
Malik et al. [13], in Ireland, found that the RIPASA score has a PPV of 84.06% and a NPV of 72.86% with a negative appendectomy rate of 15.94% and an accuracy of 80%. This is the first study evaluating the utility of the RIPASA score predicting acute appendicitis in a Western population. However, Rodrigues and Sindhu [14], in India, found that the Alvarado score had a greater specificity, PPV, and positive likelihood than the RIPASA score. The negative appendectomy in this study was quite high (18.09%) as compared to different negative appendectomy rates reported with the Alvarado score that range between 0 and 10%. Similar results were reported by Rathod et al. [15] with a negative appendectomy rate of 20.69% and a perforated appendicitis of 8.05%. This indicates that the RIPASA score can reduce the number of complicated appendectomies at the expense of a high negative appendectomy rate.

In a recent study in India, Regar et al. [16] found that the Alvarado score is more specific (80%) than the RIPASA score (60%). The PPV of the Alvarado score was 98.46% as compared to 97.83% of the RIPASA score. The negative appendectomy rate for the Alvarado score was lower that the RIPASA score (1.54 vs. 2.17%).

In another recent study, Sinnet et al. [17], in India, found that the RIPASA score has more sensitivity than the Alvarado score (95.5 vs. 65%) but has less specificity (65 vs. 90%). The PPV was 92.89% for the RIPASA score and 96.6% for the Alvarado score which indicates that the negative appendectomy rate is higher for the RIPASA score than the Alvarado score (7.61 vs. 3.33%).

In a study to assess the reliability and practical application of the Alvarado, Eskelinen, Ohmann, and RIPASA scoring systems, Erdem et al. [18], in Turkey, found that the Alvarado score had the best negative appendectomy rate (12%) than the RIPASA score (25%). The negative appendectomy rate for the Ohmann and the Eskelinen scores was 22 and 21%, respectively.

Diaz-Barrientos et al. [19], in Mexico, found that the RIPASA score showed no advantage over the Modified Alvarado score taking into consideration that the ROC curve area was 0.59 for the RIPASA score vs. 0.71 for the modified Alvarado score.

In another study, in Mexico, Reyes-Garcia et al. [20] found 15.7% cases of necrotic appendicitis and 14.3% cases of perforated appendix when using the RIPASA score. The negative appendectomy rate was also high (18.6%).

Golden et al. [21] compared the physician-determined decision with the RIPASA, the Alvarado, and the modified Alvarado score systems in order to measure the physician gestalt in the diagnosis of acute appendicitis. They found that at the higher “rule-in” cutoff threshold, the RIPASA score had a high sensitivity (78%) but a low specificity (36%). Conversely, the modified Alvarado score had a low sensitivity (47%) and a high specificity (81%). The original Alvarado score had test characteristics between these two values. They also calculated the test characteristics for the clinical scoring systems at lower “rule-out” threshold. The NPV for each score varied from 75% for the modified Alvarado score to 89% for the RIPASA score. The NPV for the physician-determined decision was 83%. The area under the curve (AUC) was greatest for the Alvarado score and the physician-determined decision (72% for both), 70% for
the MAS score, and 67% for the RIPASA score. These authors concluded that the physician-determined probability estimates were accurate as these scoring systems, which proves that the physician gestalt works well in the diagnosis of acute appendicitis.

All of these findings on the RIPASA score indicate that we need more studies to find out why the differences among the Western and South Asian and Middle Eastern populations. It is possible that these differences have to do with the anatomical position of the appendix and not precisely with the physiopathological process of acute appendicitis or the cultural differences of these populations.

2. Other scores

There are other less-known scores similar to the Alvarado score such as the Adult Appendicitis score of Sammalkorpi et al. [22] that was constructed by logistic regression analysis using multiple imputations for missing values. This score contains four symptoms and clinical signs including the sign of guarding divided into three graduations (mild, moderate, and severe) which is in reality a very subjective sign. It also contains two laboratory tests (WBC and CRP) divided at different levels that are very difficult to memorize. They reported sensitivities and specificities similar to the Alvarado score and areas under the ROC curve of 0.882 for the new score and 0.790 for the Alvarado score. The negative appendectomy rate for this new score is 18.2% which is much higher than the usual reported rates with the Alvarado score.

The Tzanakis scoring system [23] is a very simplified score that contains two clinical signs only: right abdominal tenderness (four points) and rebound tenderness (three points). The only laboratory test is a white blood cell count (WBC) greater than 12,000 cells/mm$^3$ (two points). The score relies on positive ultrasound scan findings (six points).

Sigdel et al. [24] carried out a prospective study of the Tzanakis score to compare this score with the Alvarado score and reported a sensitivity of 91.4% for the Tzanakis score and 81% for the Alvarado score. The specificity for both scores was the same (66.6%). The ROC curve gave an AUC of 0.867 for the Tzanakis score and 0.81 for the Alvarado score. The negative appendectomy rate was reported as 6% which is certainly low and is due to the addition of the ultrasound studies that are not available in many health facilities. The overall diagnostic accuracy for the Tzanakis score was 91.48% vs. 81.91% for the Alvarado score.

In a study to compare the sensitivity, specificity, and diagnostic accuracy of the Tzanakis score (TS) and the modified Alvarado score (MAS), Sharma et al. [25], in India, found that the sensitivity for the MAS was higher than the TS score (97.7 vs. 82.0%), but the specificity for the TS was higher (36.38 vs. 18%). The PPV for both scores was the same (19%), and the accuracy for the MAS was better than the TS (89 vs. 79%). They concluded that the MAS was better than the TS since in the TS there are chances of observer bias. Besides this, they could not wait till a leukocyte count goes up to 12,000 cells/cm$^3$ if clinical suspicion is present.
Kumar et al. [26], in India, found that the Tzanakis score is an effective modality in the establishment of accuracy in the diagnosis of acute appendicitis, but the limitation is observer bias which may vary the scoring results.

The Lintula score [27] was developed from 35 symptoms and clinical signs recorded for 131 Finnish children with abdominal pain and was modeled using logistic regression. This complicated score uses gender, intensity of pain, relocation of pain, vomiting, pain in the right lower quadrant, fever, guarding, bowel sounds, and rebound tenderness with different grades. Some of these signs are very difficult to evaluate which may alter the final scoring.

Konan et al. [28], in a study to compare the Alvarado and the Lintula scores in patients older than 65 years of age, found that the Alvarado score was better predictor than the Lintula score. Both scores have a high sensitivity and specificity in the diagnosis of acute appendicitis.

Ojuka and Sangoro [29], in a prospective study, carried out at Kenyatta National Hospital, found that the ROC curves for Lintula and Alvarado scores are almost identical (0.6824 and 0.6966), respectively. However, the sensitivity for the Lintula score is lower than the Alvarado score (60.8 vs. 83.3%), and the overall accuracy for the Lintula score was also lower (69.6 vs. 70.4%).

The Ohmann score [30] was developed in Germany using a computer-aided diagnosis. The variables of the score are tenderness, no micturition difficulties, steady pain, leukocytosis count >10,000 cells/mm³, age >50 years, relocation of pain to the right lower quadrant, and rigidity. In spite of this computerized system, there was no improvement in the number of perforations or complications.

In an analysis of scores in the diagnosis of acute appendicitis in women, Horzic et al. [31] compared the modified Alvarado score, Ohmann score, and Eskelinen score finding that all patients with the modified Alvarado score of 7 or more had acute appendicitis (100% specificity) which can be used to determine the need for immediate appendectomy.

Recently, Wilasrusmee et al. [32] developed a new appendicitis score for patients with suspected appendicitis and compared it with the Alvarado score. This score, also known as RAMA-AS, includes seven variables (migration of pain, progression of pain, pain aggravation by cough or movement, temperature of 37.8°C or more, and rebound tenderness). Also, it includes two laboratory tests (WBC >10,000 cells/mm³ and neutrophils <75%). In the evaluation of the variables of the score, there are serious questions. For example, they gave great importance to rebound tenderness (the only sign of the score) which contradicts the literature that always mentions direct tenderness in the right lower quadrant as the main variable. Besides this, their own statistic shows that rebound tenderness is present in 23.9%, whereas tenderness in the right lower quadrant is present in 88.4% of their cases. Another significant discrepancy is that they gave more importance to pain aggravation than anorexia (56.3 vs. 76.1%). Another objectionable symptom is progression of pain since this is a very subjective symptom that is difficult to evaluate. The C-statistics reported by Wilasrusmee et al. are better than the Alvarado score, but the RAMA-AS score did not perform well in the external data when compared to the derived data. Using the score in practice is not as easy as claimed by this group since it requires the use of the Fagan nomogram. In addition, the calculation of the score is difficult to obtain because the evaluation of the parameters is given in fractional numbers. For all of these reasons, the new score will need external evaluations to establish its usefulness in the real practice.
Khanafer et al. [33] made some modifications to the Alvarado score (AS) and the Pediatric Appendicitis score (PAS) to screen children at low risk for appendicitis who could be carefully observed at home without the need for laboratory investigation. In this study, a total of 180 children were enrolled with an average age of 11.2 years of which 56.7% were female. According to their findings, children with a score of >7 for the modified PAS and AS may be safely sent home with close follow-up, while those above this cut-off would benefit from a referral for further evaluation in the ED. They found similar sensitivities for all the scores but reduced specificities and predictive values for the modified PAS and AS scores. As expected, the ROC curves showed a reduced AUC using the modified scores. The negative appendectomy rate was 5.2% only.

3. Conclusion

A good diagnostic score for acute appendicitis should be simple, easy to memorize, repeatable, economical, and easy to apply in an emergency setting. It should contain elements with a good statistical significance. Also, a good diagnostic score for acute appendicitis could be useful for statistical purposes by providing a more precise indexing of the disease. For example, it could be used, as a clinical indicator, in the International Classification of Diseases at a fifth digit level.

Author details

Alfredo Alvarado1,2,3,4,5*

*Address all correspondence to: alfredoalvara@hotmail.com
1 Alma Mater Universidad Nacional de Colombia, Bogotá, Colombia
2 Certified by the American Board of Surgeons and the American Board of Emergency Medicine, Miami, FL, USA
3 International College of Surgeons, Chicago, IL, USA
4 Society of Laparoscopic Surgeons, Chicago, USA
5 American College of Emergency Physicians, Irving TX, USA

References


