Comparative analysis of diagnostic scales of acute appendicitis: Alvarado, RIPASA and AIR

Análisis comparativo de escalas diagnósticas de apendicitis aguda: Alvarado, RIPASA y AIR

Martín A. Bolívar-Rodríguez†, Benny A. Osuna-Wong†, Ana B. Calderón-Alvarado†, Jaime Matus-Rojas†, Edgar Dehesa-López§ and Felipe de Jesús Peraza-Garay§

†Department of General Surgery; ‡Department of Research; §Department of Statistics. Center of Health Sciences Research and Teaching, Universidad Autónoma de Sinaloa, Culiacán, Sin., Mexico

Abstract

Introduction: Acute appendicitis is the most common surgical disease in emergency surgery, however, it remains a diagnostic problem and represents a challenge despite the experience and the different clinical and paraclinical diagnostic methods.

Objective: To evaluate in a comparative way the scale of Alvarado, AIR and RIPASA to determine which one is best as a diagnostic test of acute appendicitis in our population in order to arrive to an accurate diagnosis in the shortest possible time and cost.

Method: Observational, prospective, transversal and comparative study of 137 patients to whom the scale of Alvarado, AIR and RIPASA was applied, who entered the emergency service of the Civil Hospital of Culiacán (México) with abdominal pain syndrome suggestive of acute appendicitis.

Results: The Alvarado scale presented sensitivity 97.2% and specificity of 27.6%. AIR presented sensitivity of 81.9% and specificity of 89.5%. RIPASA showed the same results as Alvarado. All tests showed diagnostic accuracy above 80%.

Conclusions: Alvarado and RIPASA presented good sensitivity, however, AIR is more specific, and has better accuracy for the diagnosis of acute appendicitis, making a better screening and thus reducing unnecessary surgeries. Therefore, it is recommended to use more AIR than Alvarado and RIPASA.

KEY WORDS: Appendicitis. Diagnosis. Alvarado. RIPASA. AIR.

Correspondence: *Martín A. Bolivar-Rodríguez
Eustaquio Buelna 91,
Col. Gabriel Leyva
C.P. 80039 Culiacán, Sin., México
E-mail: bolivarmartin64@hotmail.com

Date of reception: 02-02-2018
Date of acceptance: 16-04-2018
DOI: 10.24875/CIRUE.M18000024
Cir Cir. 2018;86:151-155
Contents available at PubMed
www.cirugiaycirujanos.com
Introduction

Acute appendicitis is the surgical condition that most often causes hospitalization and emergency surgery; however, even after such a long time elapsed since its first description, sometimes it remains a diagnostic problem and represents a challenge for all doctors who treat patients with symptoms suggestive of this pathological process, despite the experience and the different clinical and paraclinical diagnostic methods.

Appendectomy is currently the most common emergency surgical procedure in the world. The risk of developing acute appendicitis throughout life is estimated at approximately 7%. Its incidence is 11 cases per 10,000 people per year in developed countries. In the USA, 250,000 cases of appendicitis are reported annually, representing one million patient-hospital days per year. In our population, estimates indicate that one in every 15-20 people will develop acute appendicitis at some point in life, with a reported incidence of 1.5-1.9 cases per 1000 population.

Incorrect or late diagnosis increases the risk of complications, and the rate of appendicitis misdiagnosis was thus reported to range between 15 and 25%, which is considered acceptable to reduce the incidence of complications. However, in recent years, the frequency of negative appendectomies has been decreasing in a sustained manner in association with the use of diagnostic imaging studies. "Now there is clear evidence that (previously used ranges) should not be considered acceptable anymore". However, in our setting, routine medical history and physical examination continue to be the most practical diagnostic modalities. Today, a number of authors claim that acute appendicitis diagnosis continues to be clinical, as Rebollar et al. refer: "a well-structured medical history, as well as a good physical examination, provide the diagnosis in most cases".

Despite being a common condition, it remains a difficult-to-establish diagnosis, especially in the population groups of young people, the elderly and women of childbearing age, in whom various inflammatory conditions of the gynecological and genitourinary tract can exhibit signs and symptoms that are similar to those of acute appendicitis. A delay in the performance of appendectomy in order to improve diagnostic accuracy increases the risk of appendicular perforation and sepsis, as well as morbidity and mortality.

Currently, there are numerous tools that can be used in the assessment of patients with suspected appendicitis. How these tools are combined in clinical practice largely depends on the framework, availability of resources and clinical objectives. For example, in rural areas of Kenya, clinical evaluation remains the cornerstone of diagnosis for everybody. In Washington, the diagnostic goals vary depending on the patient: for children, minimizing ionizing radiation exposure can mitigate the desire for more diagnostic information; for elderly patients, in whom radiation is less risky, the clarity and the quantity of information provided by computed tomography can justify its systematic use.

In consequence, several scales (Alvarado, RIPASA and AIR) include acute appendicitis classic signs and symptoms plus laboratory tests.

The Alvarado scale (Table 1) is the most widely known and the one that until a few years ago showed the best performance in validation studies. It was developed in 1986 by Dr. Alfredo Alvarado. The Appendicitis Inflammatory Response (AIR) scale (Table 2) is a scoring system, created in Sweden in 2008, developed by Andersson and Andersson. More recently, the RIPASA scale (Table 3) has been published in the Raja Isteri Pengiran Anak Saleha Hospital (RIPAS), by Chong, et al. Since this scale was developed in Asia in 2010, its applicability and effectiveness in other populations is still under study.

<table>
<thead>
<tr>
<th>Table 1. The Alvarado Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alvarado Scale</strong></td>
</tr>
<tr>
<td>Symptoms</td>
</tr>
<tr>
<td>Migratory pain</td>
</tr>
<tr>
<td>Anorexia/ketonuria</td>
</tr>
<tr>
<td>Nausea/vomiting</td>
</tr>
<tr>
<td>Signs</td>
</tr>
<tr>
<td>Right iliac fossa pain</td>
</tr>
<tr>
<td>Rebound</td>
</tr>
<tr>
<td>Temperature &gt; 37.3°C</td>
</tr>
<tr>
<td>Laboratory</td>
</tr>
<tr>
<td>Leukocytosis &gt; 100,000 cells/mm³</td>
</tr>
<tr>
<td>Neutrophilia &gt; 75%</td>
</tr>
</tbody>
</table>

Low risk: 0-4 points. There is a very low probability of appendicitis, since very rare cases have occurred with less than 4 points. Intermediate risk: 5-6 points. The patient presents with probable appendicitis and serial evaluations, both clinical and laboratory, as well as some imaging studies (ultrasonography, computed tomography) will be required. High risk: 7 points or more. The patient requires surgery, since he/she is considered to be suffering from acute appendicitis.

Method

A prospective, cross-sectional, comparative, observational study was carried out in 137 patients who were admitted to the Hospital Civil de Culiacán emergency...
M.A. Bolívar-Rodríguez, et al.: Comparative analysis of diagnostic department with abdominal pain syndrome suggestive of acute appendicitis and to whom the Alvarado, AIR and RIPASA scales were applied.

### Inclusion criteria
- Patients admitted to the emergency department with abdominal pain suggestive of acute appendicitis.
- Both genders.
- Age between 15 and 70 years.
- Informed consent signature.

### Exclusion criteria
- Known cause of abdominal pain other than acute appendicitis.
- Patients who develop right iliac fossa pain after admission.

### Censoring criteria
- Incomplete medical record.
- Voluntary discharge prior to completing the diagnostic protocol.

### Appendicitis diagnostic confirmation
Ratification of the acute appendicitis diagnosis by histopathological study (reference method).

### Statistical analysis
Statistical analysis was carried out with a sample size calculation for a sensitivity of 90% and specificity of 80%, with an accuracy of 10% and a prevalence of 50%, with a calculated N of 125 patients.

Data were analyzed using the statistical package for the social sciences (SPSS), version 20.

Descriptive statistics were calculated with measures of central tendency for general data, and diagnostic test analysis (sensitivity, specificity, predictive values, diagnostic accuracy) to assess each scale.

### Results
The research was conducted in a total sample of 137 patients, with three study groups being formed, which were assigned the Alvarado group, RIPASA group and AIR group names, where all the patients were included, and the results compared with the reference method, namely, histopathological analysis.
The Alvarado group was composed of 137 patients to whom the diagnostic scale was applied, with a total of 100 (72.99%) patients being found to be at high risk, among whom histopathological findings were positive in 84 (61.3%) and negative in 16 (11.67%). There were 30 subjects (21.89%) who were classified at intermediate risk, among whom histopathological findings were positive in 20 (14.5%) and negative in 10 (7.29%); and there were seven patients (5.10%) classified at low risk, in whom histopathological findings were positive in 4 (2.91%) and negative in 3 (2.18%). When all data were pooled, 105 (76.64%) true positives, 21 patients (15.32%) with positive scale and negative biopsy, 8 (5.83%) true negatives and 3 patients (2.18%) with negative scale and positive biopsy were obtained. Thus, this diagnostic scale shows a sensitivity of 97.2%, specificity of 27.6%, a positive predictive value of 83.3%, a negative predictive value of 72.7% and a diagnostic accuracy of 82.5%.

The RIPASA group consisted of 137 patients to whom the diagnostic scale was applied, with a total of 68 patients (49.63%) being found with high probability, among whom histopathological findings were positive in 53 (38.68%) and negative in 15 (10.94%). There were nine patients (6.56%) with low probability, among whom histopathological findings were positive in 2 (1.45%) and negative in 7 (5.10%). Among the 58 patients (42.33%) in the acute appendicitis diagnosis category, the histopathological study tested positive in 52 (37.95%) and negative in 6 (4.37%); and among the 2 patients in the unlikely category, 1 (0.72%) had a positive histopathological result and 1 (0.72%) negative. Pooling all data, there were 105 (76.64%) true positives, 21 (15.32%) patients with positive scale and negative histopathological study, 8 (5.83%) true negatives and 3 (2.18%) with negative scale and positive histopathological study. Thus, this diagnostic scale shows a sensitivity of 97.2%, specificity of 27.6%, a positive predictive value of 83.3%, a negative predictive value of 72.7% and a diagnostic accuracy of 82.5%.

The AIR group consisted of 137 patients, out of which 48 were eliminated because they did not comply with the parameters required for the diagnostic scale. Finally, 91 patients were included, to whom the diagnostic scale was applied, with a total of 13 (9.48%) being found with high probability, among whom histopathological findings were positive in 11 (8.02%) and negative in 2 (1.45%). Among 24 (17.51%) who were classified with low probability, the histopathological study tested positive in 13 (9.48%) and negative in 11 (8.02%); and out of 52 (37.95%) with intermediate risk, the histopathological study was positive in 46 (33.57%) and negative in 6 (4.37%). When all data were pooled, there were 59 (46.83%) true positives, 2 (1.95%) patients with positive scale and negative histopathological study, 17 (18.66%) true negative and 13 (14.28%) patients with negative scale and positive histopathology. Thus, this diagnostic scale shows a sensitivity of 81.9%, specificity of 89.5%, a positive predictive value of 96.7%, a false positive rate of 10.5%, a false negative rate of 1.81% and a diagnostic accuracy of 83.5% (Table 4).

### Table 4. Diagnostic scales results

<table>
<thead>
<tr>
<th>Results</th>
<th>Alvarado, %</th>
<th>Ct, %</th>
<th>RIPASA, %</th>
<th>Ct, %</th>
<th>AIR, %</th>
<th>Ct, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>97.2</td>
<td>92.1</td>
<td>97.2</td>
<td>92.1</td>
<td>91.9</td>
<td>71.5</td>
</tr>
<tr>
<td>Specificity</td>
<td>27.6</td>
<td>14.7</td>
<td>27.6</td>
<td>14.7</td>
<td>89.5</td>
<td>68.6</td>
</tr>
<tr>
<td>PPV</td>
<td>83.3</td>
<td>75.9</td>
<td>83.3</td>
<td>75.9</td>
<td>96.7</td>
<td>88.8</td>
</tr>
<tr>
<td>NPV</td>
<td>72.7</td>
<td>43.4</td>
<td>72.7</td>
<td>43.4</td>
<td>10.5</td>
<td>2.9</td>
</tr>
<tr>
<td>FPR</td>
<td>72.4</td>
<td>54.3</td>
<td>72.4</td>
<td>54.3</td>
<td>85.3</td>
<td>85.3</td>
</tr>
<tr>
<td>FNR</td>
<td>2.8</td>
<td>0.9</td>
<td>2.8</td>
<td>0.9</td>
<td>7.9</td>
<td>7.9</td>
</tr>
<tr>
<td>Accuracy</td>
<td>82.5</td>
<td>75.3</td>
<td>82.5</td>
<td>75.3</td>
<td>83.5</td>
<td>74.6</td>
</tr>
</tbody>
</table>

Ct: confidence interval; FNR: false negative rate; FPR: false positive rate; NPV: negative predictive value; PPV: positive predictive value.

### Discussion

As previously described, in our review, no studies comparing these three diagnostic scales were found in the world literature, but there are numerous studies comparing Alvarado and AIR, or Alvarado and RIPASA scales. Comparing our results with those available in the world literature, we find that in the studies by Andersson and Andersson [12], regarding AIR and Alvarado, there is a complete contrast, but the study by Sammalkorpi et al. [14] shows similarity with ours in terms of specificity, with better sensitivity for ours in the comparison between AIR and Alvarado. As for the studies carried out by Reyes-García et al. [15] in 2012, Nanjundaiah et al. [16] in 2014, and Walczak DA et al. [17] in 2015, they report results similar to those found for the RIPASA scale in the present study, with these results being shared with the world literature.
The first study published on the subject was carried out in Sweden in 2008, with the authors, Andersson and Andersson, obtaining a diagnostic accuracy of 93% for advanced appendicitis and 97% for all appendicitis cases, versus 92% \((p = 0.0027)\) and 88% \((p = 0.0007)\), respectively, for the Alvarado scale. As for diagnostic accuracy, 82.5% of diagnostic precision was obtained for Alvarado and RIPASA, versus 83.5% for AIR, with similar values to those available in the literature.

Burgos-Oliveros conducted a study with the purpose to evaluate the differences between these scales, which included 352 patients aged 14 to 60 years diagnosed with acute appendicitis who underwent appendectomy. The RIPASA scale shows superiority versus the Alvarado scale; in our results, the Alvarado and RIPASA scales share the same diagnostic values. On the same subject, Walczak et al. carried out a study that included 92 patients who underwent laparotomy under suspicion of acute appendicitis. In this work, five different scales were compared, including RIPASA and Alvarado. The authors concluded that, based on their findings, the scale systems have a limited value for the diagnosis of acute appendicitis. Most of the used systems showed high sensitivity and a positive predictive value, which allows selecting truly ill patients and decreasing unnecessary laparotomies. On the other hand, low specificity and negative predictive value were evident, which could be associated with late diagnosis and subsequent complications. These results differ from those of the rest of the literature, particularly in regards to specificity. Our positive predictive value was high for all 3 scales, with values of 83.3% for Alvarado and RIPASA, and 96.7% for AIR; negative predictive values were low.

Conclusions

In the present study, we concluded that the Alvarado scale and the RIPASA scale have a very good margin to detect truly ill patients when they exhibit high risk, with a sensitivity of 97.2%, and a high positive predictive value of 83.3%, with a diagnostic accuracy that reaches 82.5%.

As for the AIR scale, it shows an adequate specificity of 89.5% and a sensitivity of 81.9% when it classifies patients with high probability and intermediate probability, with a positive predictive value as high as 96.7%, and a diagnostic accuracy of 83.5%.

However, the research for pre-surgical diagnosis of the patient with suspected acute appendicitis continues to be challenging, even after the introduction of imaging studies, which is why these tools should be selectively used. Although acute appendicitis diagnosis is considered to be relatively easy, appendicitis classic signs can sometimes be difficult to obtain, and are unequivocally present only in 50% of patients.

This study sought to determine which scale has higher accuracy as a diagnostic test, with the AIR scale obtaining more accuracy. It can be used in our population as an objective method to support decision-making in the management of patients with suspected acute appendicitis.

As for the rate of negative appendectomies, we firmly believe that we are within an acceptable range (13.6%) and that we should focus more on the short-term follow-up of patients under suspicion.

Conflicts of interest

The authors declare not having any conflicts of interest.

References