SURGEON’S CLINICAL VALUATION AND ACCURACY OF ULTRASOUND IN THE DIAGNOSIS OF ACUTE APPENDICITIS: A COMPARISON WITH INTRAOPERATIVE EVALUATION. FIVE YEARS EXPERIENCE.

Running title: accuracy of ultrasound in the diagnosis of acute appendicitis

Alessia Ferrarese¹, Alessandro Falcone¹, Mario Solej¹, Dario Bono¹, Paolo Moretto², Najada Dervishi², Veltri Andrea², Stefano Enrico¹, Mario Nano¹, Valter Martino¹

¹University of Turin - Department of Oncology – School of Medicine – Teaching Hospital “San Luigi Gonzaga” – Section of General Surgery - Orbassano – Turin

²University of Turin - Department of Oncology – School of Medicine – Teaching Hospital “San Luigi Gonzaga” – Section of Radiology - Orbassano – Turin

e-mails: AF: alessia.ferrarese@gmail.com, AF: alessandrofalcone.md@gmail.com, MS: mariosolej@gmail.com, DB: bonodario9@gmail.com, PM: morettopaulo1978@gmail.com, , ND:najadadervishi@yahoo.it, VE: andrea.veltri@unito.it, SE: stefano_e@libero.it, MN:mario.nano@unito.it, VM: valtermartino.md@gmail.com

AF: Assistant of General Surgery at General Surgery Section – Orbassano- University of Torino
AF: Assistant of General Surgery at General Surgery Section – Orbassano- University of Torino
MS: General Surgeon at General Surgery Section – Orbassano- University of Torino
DB: Assistant of General Surgery at General Surgery Section – Orbassano- University of Torino
PM: Radiologist of Radiology Section – Orbassano – University of Torino
ND: Radiologist of Radiology Section – Orbassano – University of Torino
AV: Associated Professor of Radiology – Orbassano – University of Torino
SE: General Surgeon at General Surgery Section – Orbassano- University of Torino
MN: Ordinary Professor of General Surgery - University of Torino
VM: General Surgeon at General Surgery Section – Orbassano- University of Torino

Correspondence to:
Alessia Ferrarese MD, Department of Oncology, University of Turin, Section of General Surgery, San Luigi Gonzaga Teaching Hospital, Regione Gonzole 10, 10043 Orbassano – Turin (Italy), e-mail alessia.ferrarese@gmail.com
The Authors have no conflict of interest or any financial support
Abstract

Introduction

Acute appendicitis is the most common cause of acute abdomen in adolescents, with an overall incidence of 7%. Two such tools are used to diagnose acute appendicitis: ultrasound and Computered Tomography imaging. End point of this study was to verify the accuracy of ultrasound imaging in the diagnosis of acute appendicitis with respect to intraoperative observations and the respective clinical and laboratory findings in young and in the elderly.

Methods

We considered all the appendectomies for acute appendicitis performed between 1 January 2010 and 1 January 2015. We evaluated clinical symptoms, laboratory findings, ultrasound findings, intraoperative signs, and anatomical and pathological findings. In the study we compared the ultrasound and intraoperative findings and then compared these with the respective clinical and laboratory data.

Results

In a comparison of diagnostic accuracy, the difference between clinical and ultrasound examinations was not significant. The differences between the diagnostic accuracy of clinical and laboratory findings and between ultrasound and laboratory investigations were statistically significant.

Conclusion

We defined white blood cells and C protein levels as non-diagnostic of the type of acute inflammation but rather as indicators of the severity of the inflammatory process.

We also agree with the authors who proposed the incorporation of ultrasonography into routine practice in the diagnosis of acute appendicitis, but only and exclusively to support other diagnostic procedures and preferably within emergency departments. A thorough clinical examination of patients with suspected acute appendicitis is still the best diagnostic procedure available to us.

Keywords: ultrasound, appendicitis, acute appendicitis

Abbreviations: CT=Computed tomography, US= ultrasound
1. Introduction

Acute appendicitis is the most common cause of acute abdomen in adolescents [1-4], with an overall incidence of 7% in young and elderly, as reported in the literature [5].

An important predictor in the clinical diagnosis of acute appendicitis is the classic migration of pain described by Murphy in 1905 [6]; according to the literature, this alone has a diagnostic accuracy of up to 95% [7,8]. The positivity of McBurney’s sign increases suspicion of acute appendicitis [9]. If presentation is typical, the diagnosis of acute appendicitis is based on clinical and laboratory findings with no need for any further investigations; however, in 35 to 40% of cases the clinical features are aspecific and unclear [10]. According to some studies, the discriminatory power of clinical and laboratory findings alone is not strong enough to diagnose acute inflammation of the appendix [11-16], and the use of a first-level diagnostic tool is essential for early diagnosis [17].

Two such tools are used to diagnose acute appendicitis: ultrasound and CT imaging [18-21].

The use of ultrasonography to visualize the appendix was first described by Deutsch and Leopold in 1981 [22], and in 1986 Puylaert described the use of graded compression during ultrasound examination in the diagnosis of patients with suspected acute appendicitis [23].

Ultrasound imaging is currently the diagnostic examination of choice for patients admitted to the emergency department with acute inflammation [24,25]. CT imaging has been found to have better diagnostic accuracy than ultrasonography, but is also more expensive [18-21,26].

End point of this study was to verify the accuracy of ultrasound imaging in the diagnosis of acute appendicitis with respect to intraoperative observations and the respective clinical and laboratory findings.

2. Methods

The retrospective study was performed at the San Luigi Gonzaga University Hospital General Surgery Unit, in collaboration with the University Radiology Unit, in Orbassano, Turin, Italy, and took into consideration all the appendectomies for acute appendicitis performed between 1 January 2010 and 1 January 2015. The cohort comprised a total of 157 patients.

Of these, the following were excluded from the study: 44 patients in whom a certain diagnosis was made on the basis of clinical and laboratory findings and surgery was performed without preoperative imaging, and nine patients with particularly serious clinical and biohumoral symptoms, all of whom underwent a preoperative CT scan in the first instance. In the latter group of patients, ultrasound scans were not performed prior to surgery. Our study sample thus comprised 104 patients.

The following parameters were evaluated: clinical symptoms (pain, nausea, vomiting, body temperature, McBurney’s sign, guarding of right iliac fossa), laboratory findings (WBC, CRP), ultrasound findings (visualization of the appendix,
appendiceal peristalsis, appendiceal wall thickening, compression of the viscus by application of the probe, periappendiceal effusion and lymphadenopathy), intraoperative signs (appendiceal erythema-edema, appendiceal phlegmon, gangrene of the appendix, perforation, gangrene and effusion), and anatomical and pathological findings (perivisceritis, edema, serositis, necrosis). For each group, a final overall rating of the “typicality of findings” for acute appendicitis was assigned. Typical clinical symptoms included fever and localized right iliac fossa pain, with or without nausea and vomiting. As regards laboratory variables, typical symptoms included a WBC of > 13,000 and CRP of > 5. Ultrasound variables included visibility of the appendix with thickening of the walls, or the simultaneous presence of two or more of the following secondary characteristics: adipose inflammation, periappendiceal lymphadenopathy, peripappendiceal effusion. Typical anatomical and pathological findings confirmed the presence of lymphocytic infiltration associated with one or more of the following characteristics: perivisceritis, exudative peritonitis, edema, serositis, necrosis or polymorphonuclear inclusions. For the ultrasound diagnosis only, the “doubtful finding” parameter was included when just one of the secondary symptoms was present.

Intraoperatively, a positive diagnosis of acute appendicitis was made if the surgeon identified one of the following symptoms: appendiceal erythema, erythema-edema, phlegmon, necrosis. The simultaneous presence of free fluid or visceral perforation with diffuse peritonitis was recognized as characteristic of acute appendicitis but not as an actual diagnostic variable.

All ultrasound scans were performed by a team of radiologists from the same school. All the appendectomies were performed by laparoscopy, with access Veress assisted, through umbilical incision and disposition of two operative trocars: one in the left iliac fossa and one in the suprapubic area. All the operations were performed by three surgeons with similar experience in laparoscopy (more than 100 emergency laparoscopic procedures and more than 200 laparoscopic cholecystectomies).

Results of the anatomical and pathological evaluations were found to be fully in agreement with intraoperative observations. The latter were therefore taken as the valid finding.

In the study we compared the ultrasound and intraoperative findings and then compared these with the respective clinical and laboratory data.

3. Results

Demographic characteristics of our study sample are described in Table 1: the patients were statistically comparable. Table 2 shows the results in terms of the “typical findings” of the evaluations performed. 24 uncertain diagnoses were made with ultrasonography; of these, 20 were found to be acute appendicitis during surgery and four were normal.
Uncertain diagnoses were based on the identification of a single positive finding and were therefore classified as positive, albeit only faintly.

Table 3 shows the definitions and the stratification of the true positive, true negative, false positive and false negative results for each parameter evaluated. As regards true positives, clinical examinations identified 70 cases, laboratory investigations 35 and ultrasound imaging 75. Clinical examinations produced false negative results in 31 cases, laboratory investigations in 62 and ultrasonography in 22. Clinical examinations produced no false positives and four true negatives. Laboratory investigations also produced four true negatives and four false positives. Ultrasound imaging produced five false positives and three true negatives.

Specificity, sensitivity, positive predictive value, negative predictive value and diagnostic accuracy are shown in Table 4.

Significance was: 100% for clinical examinations, 50% for laboratory investigations, 37.5% for ultrasound imaging.

Sensitivity was: 67.9% for clinical examinations, 77.3% for ultrasound imaging, 36.1% for laboratory investigations.

Negative predictive values were low for all the methods used; the least predictive were laboratory findings (6.1%) and clinical examinations (5.7%) compared to ultrasound, which had a predictive rate of 12%.

Clinical examinations produced a positive predictive value of 100% compared to 98.7% for ultrasound and 98.7% for laboratory investigations.

Overall diagnostic accuracy (DA-Table 5) was 74.3% for ultrasound, 68.6% for clinical examinations and 37.1% for laboratory investigations.

The concordance between clinical and laboratory findings and between clinical and ultrasound findings are shown in table 6.

The diagnostic accuracy of clinical examinations associated with laboratory findings was 54%. The diagnostic accuracy of clinical examinations associated with ultrasound imaging was 72%.

In a comparison of DA, the difference between clinical and ultrasound examinations was not significant. The differences between the DA of clinical and laboratory findings and between ultrasound and laboratory investigations were statistically significant.

In our study the rate of appendectomies in patients with a normal appendix was 3.8% (four out of a total of 105 cases). In each case of a normal appendix, the intraoperative diagnosis was: pelvic inflammatory disease in female patients. There were no cases of normal appendix in male patients. In three of these cases, ultrasound produced a negative preoperative diagnosis. In all four cases the findings of the laboratory tests and clinical examinations were rated at the lower end of our scale.

4. Conclusion
The origins of surgical treatment of appendicitis date back a long way. The first open appendectomy was performed by McBurney in 1894 [27] and Kurt Semm performed the first laparoscopic appendectomy in 1983 [28].

Appendectomy is currently the surgical procedure most commonly performed by trainee surgeons [2-4, 29].

Etiologic mechanism of acute appendicitis appears to be multifactorial and seems to be caused by the combination of an ischemic event and a bacterial superinfection after luminal obstruction [30,31].

There are two main clinical scoring systems used in the diagnosis of acute appendicitis, the RIPASA score and the ALVARADO score, which consider clinical, physical and laboratory data. The former has been found to show better diagnostic accuracy than the latter [32,33].

We did not use a formal diagnostic scoring system in our study, but performed the conventional complete physical examination and laboratory tests as proposed in the literature [34-39].

Following a review of the literature, we chose to use WBC and CRP levels: previous multivariate analyses have suggested that a preoperative white blood count of less than or equal to 13.5 x 10^9/L is a negative predictive factor for acute appendicitis [40] and according to another multivariate analysis, a CRP level of more than 7.05 is a positive independent positive predictor for acute appendicitis (especially in the elderly and children when clinical examination is less accurate) [41].

The results of our retrospective statistical analysis of the data produced are discussed below.

The 100% specificity of the clinical examination refers to the ability of this procedure performed by an expert to correctly diagnose the condition. On the other hand, this type of examination has a sensitivity of 67.9% in that, especially in female patients, it is less accurate in distinguishing between right acute abdomen and gynecological disorders.

Laboratory findings do not achieve the specificity of clinical examinations, as they only give a general measure of the inflammation but are never specific. Ultrasound has a specificity of 50%, due to the fact that the appendix is not always visible, even in patients with acute appendicitis; however, sensitivity is high as this type of investigation is able to evaluate the consequences of inflammatory events (e.g. effusion or lymphadenopathy) with extreme accuracy.

Negative predictive values were low for all methods. All produced a small number of true negatives and a large number of false negatives. The false negative results produced by laboratory investigations can be accounted for by the aspecific nature of inflammatory values and the fixed lower limit of our scale. False negative clinical diagnoses referred to patients in whom the findings of the clinical examination were not fully in agreement with the classification and whose abdominal symptoms were less clear and aspecific.

Ultrasound identified 22 false negative cases in which the appendix was not visible and the overall diagnosis was atypical. In all of these 22 cases the clinical and laboratory tests showed a positive diagnosis for acute appendicitis, which was confirmed during surgery.
Positive predictive values were high for all methods: the 100% for physical examinations was due to the negative nature of the false positive diagnoses these produced. The high positive predictive values of laboratory findings and ultrasound were also due to the small number of false positives produced. The five false positive diagnoses produced by ultrasound imaging were attributable to the identification of two minor diagnostic factors, namely pelvic effusion and periappendiceal lymphadenopathy, which are not specific to acute appendicitis. The four false positive diagnoses produced by laboratory findings were due to the aspecific nature of high inflammatory marker levels.

Laboratory tests had a diagnostic accuracy of 37.1%. This is in line with the findings reported in the literature and is due to the aspecific nature of evaluation based on inflammatory markers only.

Ultrasound imaging was found to have the highest level of diagnostic accuracy (74.3%); the comparison with the accuracy of clinical examinations alone (68.6%) was not statistically significant.

As stated by some authors, the experience of the person performing the ultrasound scan is a factor that can affect the accuracy of the examination [42].

The statistical non-significance of the comparison between clinical examinations and ultrasound in terms of DA can be explained by the high level of accuracy of both methods in diagnosing acute appendicitis. The statistical significance of the comparison between clinical and ultrasound examinations versus laboratory findings reflects the much greater accuracy of clinical and ultrasound examinations in diagnosing the specific condition, compared to the accuracy of laboratory findings alone.

A cross-study comparison between clinical and laboratory findings showed a reduction in diagnostic accuracy in relation to clinical examinations alone (Table 6); this is due to the low specificity and sensitivity of laboratory data.

A cross-study comparison between clinical and ultrasound findings revealed an increase in diagnostic accuracy in relation to clinical examinations alone (Table 6); this is due to the large number of true diagnoses (positive and negative) based on ultrasound imaging, associated with the high specificity and positive predictive value of clinical examinations.

Some studies have reported that an estimated 15 to 25% of removed appendices are normal; this can occur in up to 40% of cases in women, given the difficulty of distinguishing the symptoms of gynecological diseases from those of acute appendicitis [7,43-45].

Other authors have reported a 32.5% rate of negative appendectomies in patients with a high white blood count and 13.4% in patients whose blood tests were normal at the time of diagnosis [46].

In our study, a normal appendix was removed in four out of a total of 105 patients (3.8% of cases), an appreciable result in terms of overall diagnostic accuracy. The fact that all four patients were females supports the hypothesis that a certain diagnosis of acute appendicitis is more difficult to make in females [47].
In one of these four patients, the ultrasound examination was positive for right pelvic effusion and thickening of the periappendiceal fatty tissue, with an intraoperative diagnosis of right PID.

In this case the physical examination and laboratory findings were not specific for acute appendicitis. In the other three patients the clinical and laboratory findings were at the lower end of our scale, but they presented with aspecific pelvic pain, a WBC of between 10,000 and 12,000 and CRP of between 3 and 5. Moreover, all three patients had previously been admitted to the emergency department with the same symptoms.

In our study, all the appendectomies were performed laparoscopically since we believe that the diagnostic power of this minimally invasive technique is fundamental for a disease that can pose some difficulty in terms of differential diagnosis. In line with the studies reported in the literature, we also agree that the laparoscopic approach gives the best overall results in terms of postoperative stay in hospital, level of postoperative pain and incidence of complications [48].

In accordance with a number of authors, taking into account cost effectiveness and availability of equipment, we believe that the most appropriate imaging technique for use in the first instance when acute appendicitis is suspected is ultrasound imaging [24,25].

We defined WBC and CRP levels as non-diagnostic of the type of acute inflammation but rather as indicators of the severity of the inflammatory process [34,38].

We also agree with the authors who proposed the incorporation of ultrasonography into routine practice in the diagnosis of acute appendicitis, but only and exclusively to support other diagnostic procedures and preferably within emergency departments [24,25].

Furthermore, apart from specific cases in which diagnosis is particularly problematic and uncertain, we believe it is important for ultrasound scanning to be performed by the surgeon so that he or she can gain as complete a picture as possible in order to make a definitive diagnosis [49].

A thorough clinical examination of patients with suspected acute appendicitis is still the best diagnostic procedure available to us.
References


### TABLE 1. Patient Baseline Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male [n° (%)]</td>
<td>59 (56.19%)</td>
</tr>
<tr>
<td>Female [n° (%)]</td>
<td>46 (43.81%)</td>
</tr>
<tr>
<td>Mean age (yr), mean (± SD)</td>
<td>35 (± 12.3)</td>
</tr>
</tbody>
</table>

SD: Standard deviation

### TABLE 2. Results of evaluation

<table>
<thead>
<tr>
<th>Opinion</th>
<th>Clinic (n°)</th>
<th>Laboratory (n°)</th>
<th>Ultrasound (n°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical</td>
<td>70</td>
<td>39</td>
<td>56</td>
</tr>
<tr>
<td>Not Typical</td>
<td>35</td>
<td>66</td>
<td>25</td>
</tr>
<tr>
<td>Dubt</td>
<td>Ø</td>
<td>0</td>
<td>24</td>
</tr>
</tbody>
</table>

### TABLE 3. Statistical analysis

<table>
<thead>
<tr>
<th>Method</th>
<th>Specificity</th>
<th>Sensibility</th>
<th>Negative predictive value</th>
<th>Positive predictive value</th>
<th>Diagnostic Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinic</td>
<td>100</td>
<td>67.9</td>
<td>5.7</td>
<td>100</td>
<td>68.6</td>
</tr>
<tr>
<td>Laboratory</td>
<td>50</td>
<td>36.1</td>
<td>6.1</td>
<td>89.7</td>
<td>37.1</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>37.5</td>
<td>77.3</td>
<td>12</td>
<td>98.7</td>
<td>74.3</td>
</tr>
</tbody>
</table>
### TABLE 4. Diagnostic Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Diagnostic Accuracy</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinic</td>
<td>68.6</td>
<td>0.372</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>74.3</td>
<td></td>
</tr>
<tr>
<td>Clinic</td>
<td>68.6</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Laboratory Data</td>
<td>37.1</td>
<td></td>
</tr>
<tr>
<td>Ultrasound</td>
<td>74.3</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Laboratory Data</td>
<td>37.1</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5. Definition of true and false value

<table>
<thead>
<tr>
<th></th>
<th>True positive value</th>
<th>True negative value</th>
<th>False positive value</th>
<th>False negative value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinic</td>
<td>- Clinic positive value</td>
<td>- Clinic negative value</td>
<td>- Clinic positive value</td>
<td>- Clinic negative value</td>
</tr>
<tr>
<td></td>
<td>- Acute appendicitides at surgical evaluation</td>
<td>- Normal appendix at surgical evaluation</td>
<td>- Acute appendicitides at surgical evaluation</td>
<td>- Normal appendix at surgical evaluation</td>
</tr>
<tr>
<td>Laboratory Data</td>
<td>- Laboratoristic positive value</td>
<td>- Laboratoristic negative value</td>
<td>- Laboratoristic positive value</td>
<td>- Laboratoristic negative value</td>
</tr>
<tr>
<td></td>
<td>- Acute appendicitides at surgical evaluation</td>
<td>- Normal appendix at surgical evaluation</td>
<td>- Acute appendicitides at surgical evaluation</td>
<td>- Normal appendix at surgical evaluation</td>
</tr>
<tr>
<td>US</td>
<td>- US positive value</td>
<td>- US negative value</td>
<td>- US positive value</td>
<td>- US negative value</td>
</tr>
<tr>
<td></td>
<td>- Acute appendicitides at surgical evaluation</td>
<td>- Normal appendix at surgical evaluation</td>
<td>- Acute appendicitides at surgical evaluation</td>
<td>- Normal appendix at surgical evaluation</td>
</tr>
</tbody>
</table>

|                      | Clinic (n°) 70                                                                 | 4                                                                                   | 0                                                                                   | 31                                                                                  |
| Laboratory Data (n°) | 35                                                                                   | 4                                                                                   | 4                                                                                   | 62                                                                                  |
| US (n°)              | 75                                                                                   | 3                                                                                   | 5                                                                                   | 22                                                                                  |

US: Ultrasound
### TABLE 6. Clinical-US and clinical-laboratoristic evaluation

<table>
<thead>
<tr>
<th></th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical-US evaluation</td>
<td>145</td>
<td>87</td>
<td>5</td>
<td>53</td>
</tr>
<tr>
<td>Clinical-laboratoristic evaluation</td>
<td>105</td>
<td>8</td>
<td>4</td>
<td>93</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Specificity</th>
<th>Sensibility</th>
<th>Negative predictive value</th>
<th>Positive predictive value</th>
<th>Diagnostic Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical-US evaluation</td>
<td>58</td>
<td>73</td>
<td>11</td>
<td>96</td>
<td>72</td>
</tr>
<tr>
<td>Clinical-laboratoristic evaluation</td>
<td>66</td>
<td>53</td>
<td>7.9</td>
<td>96</td>
<td>54</td>
</tr>
</tbody>
</table>

US: ultrasound
TP: true positive value
TN: true negative value
FP: false positive value
FN: false positive value