Effect of surgeon’s judgement on the diagnosis of acute appendicitis

Mustafa Hasbahçeci¹, Cengiz Erol², Mustafa Törü³, Mehmet Şeker⁴

ABSTRACT

Objective: The accuracy of a surgeon’s judgement still remains to be controversial in the diagnosis of acute appendicitis, which is a diagnosis usually based on laboratory data and imaging tests.

Material and Methods: Patients with a possible diagnosis of acute appendicitis were reviewed retrospectively with regard to demographic variables, laboratory and imaging results, and treatment modalities.

Results: There were 128 patients with a mean age of 31.2±14 years. The mean white blood cell count and the proportion of polymorphonuclear leukocytes were 11403±4669/mm³ and 75±11%, respectively. Appendectomy was performed on 66 (51.6%) patients. Conservative management was applied to 62 (48.4%) patients. Statistical analysis showed that patients with appendicitis have a higher white blood cell count (p=0.015) and a higher proportion of polymorphonuclear leukocytes (p=0.023). Sensitivity, specificity and accuracy rates were 84.6%, 63.7% and 74.3% for ultrasound and 100%, 86.7% and 92.2% for computed tomography, respectively.

Conclusion: Diagnosis based on patients’ laboratory and imaging data, in combination with, the surgeon’s judgement appears to yield the best outcomes in patients with suspicion of acute appendicitis.

Key Words: Acute appendicitis, diagnosis, ultrasound, tomography, appendectomy

INTRODUCTION

Acute appendicitis (AA) is the most frequently suspected disorder in patients presenting with acute abdominal pain and is the most common indication for urgent abdominal surgery. Diagnosis of AA based only on clinical and laboratory data results in high negative appendectomy rates and missed diagnoses with increased morbidity (1-4). Imaging tests such as graded compression ultrasound (US) with or without color Doppler evaluation and computed tomography (CT) have been used to improve diagnostic performance for the last several decades (1, 5-8). The lower sensitivity of US compared to CT and the great variability caused by operator dependency may result in a higher number of false negative diagnoses if US is used as the only imaging technique (1). However, CT is associated with considerable ionizing radiation exposure, which discourages its use. Quality assurance for patients with suspected AA should aim to minimize the negative appendectomy rate, without delaying the treatment of perforated AA, by optimal diagnostic use of US and CT (3, 9).

The purpose of this retrospective study was to evaluate the optimal use of laboratory investigation, imaging techniques and surgeon’s judgement to diagnose patients presenting with acute abdominal pain and possible AA.

MATERIAL AND METHODS

This study was a descriptive study based on retrospectively collected data. Patients who presented with right lower quadrant pain and possible AA between June 2006 and June 2010 in a private hospital were included. The institutional review board approved the study protocol (Istanbul 29 Mayis Hospital-04.02.2011/4). A clinical worksheet list was used to collect patient’s demographic information.

On admission (Figure 1), all patients underwent physical examination and blood testing to determine white blood cell (WBC) count and proportion of polymorphonuclear (PMN) leukocytes. For the standard values of our laboratory, the WBC count and proportion of PMN leukocytes were considered normal when lower than 10400/mm³ and 75%, respectively. All patients underwent diagnostic imaging techniques (Figure 1), including US, CT or both depending on the discretion of the on-call surgeon at the general surgical unit. The imaging techniques were performed and reviewed during office hours (8 am-6 pm). If patients were admitted out of office hours, their imaging evaluations were performed and reviewed on the next day. Within the first 24 hours after admission, the attending surgeon decided between surgical treatment and conservative management based on clinical and laboratory findings.
and imaging results. Conservative management included pain control, restricting oral intake except fluids and active observation with serial clinical examinations performed by the same attending surgeon on either an in-patient or outpatient basis. During the conservative management, intravenous or oral antibiotics were not given. Pregnant patients and patients without complete data were excluded from the study.

Diagnosis of AA or perforated AA at the time of operation was based on macroscopic findings. Neither diagnostic laparoscopy nor laparoscopic appendectomy was used. Surgeons with at least 5 years of experience performed all operations. Normal-looking appendixes discovered during laparotomy were removed via split-muscle McBurney incision. All excised appendixes were microscopically analyzed by pathology using paraffin sections. Histological diagnosis of appendicitis was based on infiltration of the muscularis propria by PMN leukocytes. The proportion of patients with perforated appendixes and negative appendectomy were identified by pathologic determination of perforated AA and the proportion of normal-looking appendixes noted during surgery, respectively. Patients in whom conservative management was initially intended and then, within 48 hours, AA was diagnosed, were considered as missed AA cases.

Imaging Techniques
Senior radiologists performed all US examinations. For patients examined by US (Logiq 9, GE Healthcare, Milwaukee, Wisconsin, USA), both convex and linear probes were used. A routine US examination of the upper abdomen and pelvis using a 3-5-MHz convex transducer was initially performed. The scan was started after a 70-80 s delay. All images were reconstructed as 2.5 mm axial sections. A diagnosis of acute appendicitis was determined via CT if thickening of the appendix (≥7 mm) and associated inflammation of the periappendiceal fat (fat stranding) were observed. No inconclusive CT assessments were reported. Both US and CT were assessed for their ability to determine the best mode of treatment-surgery or conservative management-compared to reference standards and based on the findings of the operation, follow-up period, and pathological analysis. When evaluating the correlation between treatment modality decided by the attending surgeon and imaging results, if both US and CT were performed on the same patient, the CT results were regarded as the final diagnosis. Otherwise, whichever technique was performed was accepted as the final diagnosis.

Follow-up information was acquired from all patients over-the-phone or in-person follow-up appointments six months after discharge.

Statistical Analysis
Sensitivity, specificity, positive and negative predictive values, positive and negative likelihood values, and accuracy were calculated for each imaging technique to compare their diagnostic accuracy rates.

The collected data were entered in an electronic database (Microsoft Excel for Windows, Microsoft Corporation, Redmond, WA). Statistical calculations were performed using NCSS (Number Cruncher Statistical System, 2007) and PASS Statistical software (Utah, USA, 2008). Normally distributed continuous variables were expressed as mean-standard deviation (SD). Categorical variables were expressed as frequencies and percentages of an appropriate denominator. The Student’s t-test was used for analysis of normally distributed, descriptive continuous
variables, which were expressed as mean ± SD. The chi-square test and McNemar’s test were used to compare qualitative variables. The sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and accuracy of diagnostic screening tests were also used to compare the diagnostic efficiency of imaging techniques. Differences were considered statistically significant if the p value was equal to or less than 0.05 with a 95% confidence interval.

RESULTS

One hundred forty three patients were admitted with right lower quadrant pain complaint. Fifteen patients had incomplete data and were excluded from the study; a total of 128 patients were included. The mean age was 31.2±14 years, with a range of 10 to 83 years. There were 58 (45.3%) female and 70 (57.7%) male patients. Demographic findings and laboratory parameters are shown in Table 1. Comparative analysis of laboratory parameters revealed that sixty-five patients (51%) had WBC counts above 10400/mm³, and 73 (57%) had PMN leukocyte percentages above 75%.

For diagnostic purposes, 101 US and 51 CT examinations were performed. 77 only underwent a US examination, and 27 only underwent CT; 24 patients required both US and CT examinations. US results were positive for appendicitis in 62 patients. Five of them underwent additional CT at the discretion of the surgeon. Although three of the patients were regarded as positive for AA, appendectomy was performed in only two of them. Fifteen patients with a diagnosis of AA after US, including the other two patients who underwent additional CT, were managed conservatively. Surgical treatment was performed on 44 patients diagnosed with AA by US. Appendicitis was confirmed by histological examination in 42 patients, and two patients had normal-looking appendices.

Computed tomography was applied to 27 patients without an initial US examination. All 14 of the patients whose CT results were negative for AA were managed conservatively. Of the remaining 13 patients with a diagnosis of AA, one patient was followed in a conservative manner. All of the other patients were operated on, but AA was only confirmed in 11; one patient had a normal-looking appendix. Three perforated appendicitis cases were diagnosed preoperatively using US (one patient) and CT (two patients).

Sixty-six (51.6%) patients diagnosed with AA were operated. Three cases of perforated AA and three normal-looking appendices were confirmed by histological examination after the operation. Therefore, both perforated AA and negative appendectomy occurred at a rate of 4.5%. The actual prevalence of AA was 49% after exclusion of the three negative appendectomy cases. Conservative management was applied to 62 (48.4%) patients who recovered without any complaints. Four alternative diagnoses (three cases of ovarian cyst rupture and one case of pyelonephritis) were determined using imaging techniques. During the follow-up period, no patients presented with missed AA requiring surgical treatment or additional hospitalization.

The treatment modality correlated with the final radiological diagnosis in 87.5% of the cases (112 out of 128 patients); there were three cases of negative appendectomy. The remaining 16 (12.5%) patients (13 US patients and 3 CT patients) did not undergo surgical treatment, although all the patients had a radiological diagnosis of AA. Conservative management with active serial examinations by the attending surgeon was performed with no complications.

Patients with and without proven AA diagnoses (n=63 and n=65, respectively) were compared with regards to age, gender, WBC count, proportion of PMN leukocytes, and the number of patients with WBC counts and PMN leukocyte percentages higher than the cut-off value (Table 2, 3). Patients with proven AA diagnoses had significantly higher WBC counts and PMN leukocyte percentages (p<0.05). AA was more frequently seen in patients with WBC counts higher than the cut-off value (p<0.05).

Eighteen of the 62 patients diagnosed with AA using US were not confirmed as having AA (18 false positive results), and eight of the 39 patients not diagnosed with AA were confirmed as having AA (8 false negative results). Four of the 25 patients diagnosed with AA using CT were not confirmed as having AA (4 false positive results). However, all patients not diagnosed with AA by CT were confirmed as not having AA (no false negative results; Table 4). There was no difference in the diagnostic performance of the imaging techniques for either female or male patients (p>0.05).

DISCUSSION

Although appendectomy is the most common surgical procedure performed throughout the world, diagnoses based on clinical history, physical examination, and routine laboratory tests are not always accurate. The role of imaging in confirming the diagnosis of AA and detecting alternative AA-like disorders has been well demonstrated (3). Therefore, imaging techniques have been used to increase diagnostic accuracy and to lower negative appendectomy and perforated AA discharge of patients (48.4%) patients who recovered without any complaints. Four alternative diagnoses (three cases of ovarian cyst rupture and one case of pyelonephritis) were determined using imaging techniques. During the follow-up period, no patients presented with missed AA requiring surgical treatment or additional hospitalization.

The treatment modality correlated with the final radiological diagnosis in 87.5% of the cases (112 out of 128 patients); there were three cases of negative appendectomy. The remaining 16 (12.5%) patients (13 US patients and 3 CT patients) did not undergo surgical treatment, although all the patients had a radiological diagnosis of AA. Conservative management with active serial examinations by the attending surgeon was performed with no complications.

Patients with and without proven AA diagnoses (n=63 and n=65, respectively) were compared with regards to age, gender, WBC count, proportion of PMN leukocytes, and the number of patients with WBC counts and PMN leukocyte percentages higher than the cut-off value (Table 2, 3). Patients with proven AA diagnoses had significantly higher WBC counts and PMN leukocyte percentages (p<0.05). AA was more frequently seen in patients with WBC counts higher than the cut-off value (p<0.05).

Eighteen of the 62 patients diagnosed with AA using US were not confirmed as having AA (18 false positive results), and eight of the 39 patients not diagnosed with AA were confirmed as having AA (8 false negative results). Four of the 25 patients diagnosed with AA using CT were not confirmed as having AA (4 false positive results). However, all patients not diagnosed with AA by CT were confirmed as not having AA (no false negative results; Table 4). There was no difference in the diagnostic performance of the imaging techniques for either female or male patients (p>0.05).

DISCUSSION

Although appendectomy is the most common surgical procedure performed throughout the world, diagnoses based on clinical history, physical examination, and routine laboratory tests are not always accurate. The role of imaging in confirming the diagnosis of AA and detecting alternative AA-like disorders has been well demonstrated (3). Therefore, imaging techniques have been used to increase diagnostic accuracy and to lower negative appendectomy and perforated AA
rates without causing any negative effects (5). As a policy, we performed at least one kind of imaging technique on each patient presenting with right lower quadrant pain. Many different studies have examined the efficacy of diagnosing AA using clinical history, physical examination, laboratory tests, and/or imaging techniques (2, 3, 5-8, 10). After review of these studies, it was clear that guidelines regarding the optimal diagnostic approach to AA were lacking and that each center should find its own algorithm after consideration of current AA protocols and local factors, such as the availability of imaging techniques (11).

Although previous studies had conflicting results regarding the inflammatory parameters of AA, we aimed to use both the WBC count and the proportion of PMN leukocytes in every patient (7, 12-15). In one study, 15.6% of the patients presenting with right lower quadrant pain had normal WBC counts and C-reactive protein levels (12). However, our results showed that a higher WBC count was significantly correlated with the presence of AA (p<0.05). Although the proportion of PMN leukocytes was higher in AA patients, the correlation lacked statistical significance. AA was also more frequently seen in patients with WBC counts higher than the cut-off value (p<0.05). Overall, laboratory parameters related to inflammation should be regarded as adjuncts to clinical and imaging findings, and any increase in these parameters should raise suspicion of AA.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Patients with AA (n=63)</th>
<th>Patients without AA (n=65)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)‡</td>
<td>32.4±15</td>
<td>30±12.6</td>
<td>0.345</td>
</tr>
<tr>
<td>WBC count (/mm³)§</td>
<td>12417±5369</td>
<td>10420±3654</td>
<td>0.015*</td>
</tr>
<tr>
<td>Proportion of PMN leukocytes (%)*</td>
<td>77±10</td>
<td>73±11</td>
<td>0.023*</td>
</tr>
</tbody>
</table>

*: Number of patients in each group
§: Mean±SD (median)
†: Statistical significance

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Patients with AA (n=63)</th>
<th>Patients without AA (n=65)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>28 (44)</td>
<td>30 (46)</td>
<td>0.846</td>
</tr>
<tr>
<td>Male</td>
<td>35 (56)</td>
<td>35 (54)</td>
<td></td>
</tr>
<tr>
<td>Number of patients (n (%))§</td>
<td>24 (38)</td>
<td>39 (60)</td>
<td>0.013*</td>
</tr>
<tr>
<td>WBC count &lt;10 400</td>
<td>24 (38)</td>
<td>39 (60)</td>
<td></td>
</tr>
<tr>
<td>WBC count ≥10 400</td>
<td>39 (62)</td>
<td>26 (40)</td>
<td></td>
</tr>
<tr>
<td>Proportion of PMN leukocytes &lt;75</td>
<td>23 (37)</td>
<td>32 (49)</td>
<td>0.146</td>
</tr>
<tr>
<td>Proportion of PMN leukocytes ≥75</td>
<td>40 (63)</td>
<td>33 (51)</td>
<td></td>
</tr>
</tbody>
</table>

*: Number of patients in each group
§: Mean±SD (median)
*: Statistical significance

Table 2. Comparison of continuous parameters with regard to the presence or absence of AA

Table 3. Comparison of categorical parameters with regard to the presence or absence of AA

Table 4. Statistical analysis of the ability of US and CT to diagnose AA

US and CT were the most commonly used imaging techniques for AA. Previous studies have shown that the two techniques diagnose AA with different accuracies (5). CT is recognized as the most accurate imaging method for the detection of AA in patients with right lower quadrant pain (16). Several factors of imaging techniques must be taken into consideration, including the operator dependency inherent to US, the length of time required to perform the imaging techniques, the radiation dosage during CT examination, and any correlation with clinical findings (1-3, 17).

Ultrasound performance has been shown to be more accurate during the day, when senior radiologists perform US, than during the night, when on-duty residents are responsible for US examinations (3). In accordance with the literature, senior radiologists performed all US examinations in the present study during the day with an acceptable diagnostic accuracy rate. In the present study, sensitivity, specificity, and positive and negative predictive values for US were found to be 0.85, 0.64, 0.71, and 0.80, respectively. Although all cases in which the appendix could not be visualized were regarded as negative, the results were within the reported ranges (3, 5, 7-10, 18, 19).

Previous studies reported several pitfalls and limitations which lowered the diagnostic accuracy of US examinations, such as obesity, excessive bowel gas, unusual appendix location, appendicitis in the tip of the appendix, right lower quadrant abscess without visualization of the appendix, and edematous incompressible terminal ileum or cecum (3, 8). In general, very obese patients suspected of having AA are sent directly to CT because of the difficulty to penetrate tissues during US (3, 18). We usually referred our obese patients to CT without US. However, we were not able to give the body mass index values because of the retrospective nature of our study.

Mean sensitivities and specificities are substantially in favor of CT over US. CT has a sensitivity approaching 100%, is not operator dependent, and can be performed on patients on which US is difficult to perform. However, contrast administration, ionizing radiation, and cost are limiting factors to CT, and we used CT as the primary imaging technique for diagnosing AA.
CONCLUSION

Laboratory data demonstrating the severity of inflammation should be used for the clinical diagnosis of AA. US should be the primary imaging technique performed on every patient presenting with right lower quadrant pain. However, CT should be chosen as the primary imaging technique in selected cases at the discretion of the attending surgeon. CT should also be performed in addition to US on patients who are highly suspected of having AA but showing negative US results. A good clinical approach including the skillful use of laboratory data, as well as US and CT on selected patients per the surgeon’s judgement, should be considered as part of an effective diagnostic algorithm for AA.

Ethics Committee Approval: Ethics committee approval was received for this study from the Institutional Review Board of Ethical Committee, Istanbul 29 Mayis Hospital.

Informed Consent: Due to retrospective design of the study, written informed consent could not be taken from the patients. Institutional Review Board from the Ethical Committee of the Hospital has been considered.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept - M.H.; Design - M.H., C.E.; Supervision - M.Ş.; Funding - M.H., C.E., M.T., M.Ş.; Materials - M.H., C.E., M.T., M.Ş.; Data Collection and/or Processing - M.H.; Analysis and/or Interpretation - M.H., C.E., M.T., M.Ş.; Literature Review - M.H.; Writer - M.H., C.E.; Critical Review - M.T., M.Ş.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES


only 27 patients (21%) despite its higher diagnostic accuracy (5, 19, 20).

It is known that positive and negative likelihood ratios above 10 and below 0.1, respectively, imply strong effects, whereas a likelihood ratio of 1 implies no effect. Mean sensitivity and specificity for CT were reported as high as 0.91 and 0.90 (5). Given these results, sensitivity, negative predictive value and negative likelihood ratio were accepted as powerful indicators of CT effectiveness in this study, especially for patients without AA.

Previous studies demonstrated that routine referral of patients suspected of having acute appendicitis to US, and limited referral to CT based on the US results and clinical judgement, improves diagnostic accuracy and therapeutic management (1-3, 10, 16, 18, 19, 21). This approach reduces the necessity of CT examinations and has been shown to be cost effective (1). However, the use of imaging was largely successful because of multidisciplinary cooperative imaging algorithms involving the emergency department, physicians, surgeons, and radiologists. The staged protocol was also successful, especially in children, because of the effective use of US (2, 21). Therefore, in the present study, US were regarded as the primary imaging technique in most of the cases. However, CT was performed as the primary or adjunctive imaging technique in select cases as decided by the attending surgeon.

Current practice has a high negative appendectomy rate of 10%-20% but is regarded as acceptable for preventing a missed appendicitis rate of up to 12%, which is considered clinically more important (4, 5, 14). In missed appendicitis cases, perforation and abscess leading to increased morbidity and mortality, may occur more frequently (4). However, according to comparative studies published in the last decade, a mean uniform decrease in the negative appendectomy rate has occurred due to widespread implementation of CT (18, 22). A meta-analysis has similarly demonstrated an overall significant decrease in the negative appendectomy rate due to CT: from 16.7% to 8.6% for all patients, and from 27.3% to 9.6% for female patients, respectively (22). However, CT may cause more appendiceal perforations due to the delay before the scan. While the difference in the appendiceal perforation rate between the CT group and the clinical evaluation group was not statistically significant (CT: 23.4%; clinical evaluation: 16.7%), further studies should be done for clarification (22). Others and we believe that the lower rates of negative appendectomy and perforated AA were due to clinical judgement, in addition to the laboratory and imaging results (23). It was also observed that in 16 cases (12.5%), treatment modality was decided based on the clinical judgement of the attending surgeon, although all 16 had a radiological diagnosis of AA. Our findings also showed that surgical treatment should be avoided for patients with negative CT results for AA. Therefore, it should be kept in mind that clinical judgement is especially important in atypical cases with positive imaging findings for AA.

Study Limitations

The present study was limited, as the absence of AA could not be absolutely confirmed in patients who did not undergo surgery. The retrospective design and small number of patients were other limiting factors.


13. Kim HC, Yang DM, Lee CM, Jin W, Nam DH, Song JY. Acute appendicitis: relationships between CT-determined severities and serum white blood cell counts and C-reactive protein levels. Br J Radiol 2011; 84: 1115-1120. [CrossRef]


