

C-reactive protein and white blood cell count do not improve clinical decision-making in acute appendicitis

Sofie Tind¹, Annmarie Touborg Lassen², Erik Zimmerman-Nielsen³ & Niels Qvist⁴

ABSTRACT

INTRODUCTION: Acute appendicitis (AA) remains a diagnostic challenge as indicated by the high rate of unnecessary surgery. Blood samples, primarily C-reactive protein (CRP) and leucocyte counts, are used as a diagnostic supplement despite their relatively low sensitivities and specificities. However, their influence on diagnostic decision-making has not previously been investigated. The aim of the present study was to investigate if the results of CRP and leucocytes had any positive or negative influence on the decision-making of surgeons handling patients with suspected AA.

METHODS: This was a prospective, observational cohort study including patients (≥ 15 years of age) admitted on suspicion of AA. The surgeons were instructed to perform their physical examination and to register whether they found the patient more or less than 50% likely to have AA. Thereafter, the surgeons had to assess the blood results and re-evaluate their diagnosis. The surgeon's diagnosis before and after was compared with the final diagnosis defined by surgical findings or follow-up. The gold standard was any degree of appendicitis on histology.

RESULTS: A total of 226 patients were included of whom 91 (40.3%) had appendicitis on histology. The surgeons changed their diagnosis in nine cases after assessing blood samples. The changes in the proportion of correct diagnoses, sensitivity, specificity and predictive values after assessing blood samples were not significant.

CONCLUSIONS: The results of CRP and leucocyte counts did not influence clinical decision-making.

FUNDING: Financial support was obtained from the Danish Council for Independent Research (Ref. no. 12-132020).

TRIAL REGISTRATION: Clinicaltrials.gov: NCT02304653.

Acute appendicitis (AA) is the most common abdominal surgical emergency [1]. It is predominantly a clinical diagnosis, but the diagnosis is not definitive until confirmed by surgery or histology. Albeit common, AA is a difficult diagnosis to make because the symptoms of AA vary greatly as reflected in a high negative laparoscopy rate [2-4]. Thus there is a need for diagnostic tests with a high sensitivity and specificity. Computed tomography (CT) has proven to be the most accurate test, but it is associated with radiation risks [5] which make its use less appropriate in a target population consisting mainly of children and adolescents [1]. Blood samples such as C-

reactive protein (CRP) and number of leucocytes have been investigated extensively; however, such samples have a disappointingly low diagnostic sensitivity and specificity in the 65-85% and 43-73% range, respectively [6-10]. Nevertheless, they are widely used in the diagnostic process. Many emergency departments (ED) have introduced standardised blood sample programmes for specific patient groups. CRP and leucocytes are among the blood samples taken in patients with suspected AA, but it has never been investigated whether this diagnostic approach is beneficial or if it might, in fact, be misleading. The primary aim of this study was to investigate whether the results of CRP and a leucocyte count in peripheral blood had any influence on the surgeon's clinical decision-making.

METHODS

Study design and setting

We carried out a prospective observational cohort study including patients (≥ 15 years of age) admitted to the ED at Odense University Hospital, Denmark, with suspected AA in the period from 1 February 2013 to 31 October 2013. Suspected AA was defined as patients with a tentative diagnosis of AA according to the referring general practitioner or attending surgeon.

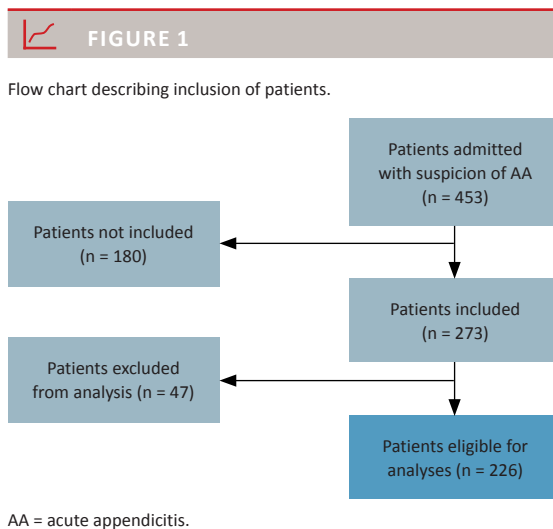
Study protocol

Upon arrival, the patients had a standardised blood sample drawn and were examined by a surgeon. According to hospital protocol, the examination should be performed within 30 minutes after admission. Immediately after the physical examination, the surgeon had to indicate on a registration form whether the patient was more or less than 50% likely to have AA. Furthermore, the surgeon's primary action was registered. The results of the blood samples were available 1-2 hours later, and at that time the surgeon was asked to re-evaluate the likelihood of AA. All registrations were made during the patient's stay in the ED and before operation. Registration forms were available in folders located in all offices used by the surgeons in the ED, and the surgeons were responsible for filling out and returning the finished registration forms. Precautions were taken to ensure that protocol was followed. Surgeons were thoroughly informed about the study and inclusion protocol, and a

ORIGINAL ARTICLE

- 1) Emergency Department, Kolding Hospital
- 2) Department of Emergency Medicine, Odense University Hospital
- 3) Department of Surgery, Odense University Hospital
- 4) Department of Surgery, Odense University Hospital, Denmark

Dan Med J
2015;62(12):A5167



one-week pilot study was performed. At the beginning of the inclusion period, information and instructions were emailed to all surgeons working in the ED. During the entire inclusion period, the registration process and the inclusion rate were controlled weekly, and any necessary reminders were either sent by email or handed out at the department's morning meetings.

The results of the blood samples were presented to the surgeons according to the hospital standard in an electronic database with standardised normal values from the local laboratory ($3.50\text{--}8.80 \times 10^9/\text{l}$ for leucocytes, $< 6 \text{ mg/l}$ for CRP).

The patients were followed until discharge, at which time a final diagnosis was assigned to each patient. This final diagnosis was based on surgical and histological findings in those who underwent surgery, and on the clinical course and result of additional investigation in those who were not operated. The non-operated patients had a 30-day follow-up where any hospital re-admission was registered and evaluated. Finally, each patient included was classified into one of two groups: Patients with histology positive for any degree of AA were classified as "appendicitis" and all others as "not appendicitis". This final diagnosis was compared with the surgeon's diagnosis before and after assessing the blood samples.

Occasionally, the surgeon did not evaluate the blood samples, but acted based on clinical findings alone. These patients were excluded from the analyses unless the primary action was immediate operation on the indication of AA. If patients underwent surgery based on clinical findings alone, the clinical indications were considered strong enough to overrule any given blood result. These patients ($n = 48$) were therefore included as "no change in diagnosis after assessing blood samples". A diagnostic laparoscopy was performed in all

patients referred to surgery. Additional exclusion criteria were: known cancer illness, pregnancy, inflammatory bowel disease, chronic hepatitis, duration of symptoms of more than seven days, if the results from blood samples were known before the physical examination and if a full anamnesis could not be obtained. Finally, patients were excluded if the surgeon chose radiologic examination before the blood samples were available because the result of the radiologic examination was expected to be a significant confounder. Patients undergoing radiologic examination after evaluation of the blood samples were included in the study.

Patients could be included more than once if the time between admissions exceeded one week. If re-admission happened within the same week, only the first admission was considered eligible.

Data analysis

The influence of the inflammatory markers in blood samples on clinical decision-making was evaluated by comparing the proportion of correct diagnoses before and after assessing blood samples using the exact McNemar's test to compare paired proportions.

A p-value below 0.05 was considered significant. Proportions were presented with 95% confidence intervals (CI) based on a binomial distribution. Furthermore, sensitivity, specificity and predictive values of the surgeons' diagnoses before and after assessing the results of the blood samples were calculated. As a secondary endpoint, the diagnostic accuracy of CRP and leucocyte counts were evaluated using receiver-operating characteristics (ROC) curves.

Ethics

The study was approved by the Danish data protection agency (Ref. no. 2008-58-0035). Treatment of the patients was according to standard; thus, patient consent was not required, which was confirmed by the National Committee on Health Research Ethics.

Trial registration: Clinicaltrials.gov: NCT02304653.

RESULTS

During the recruitment period, two senior surgeons and 20 surgical residents were involved in the evaluation of the patients. In total, 453 patients met the inclusion criteria. Of these, 60.3% ($n = 273$) were included in the study and 226 were eligible for analysis (Figure 1). Four patients were included twice. Table 1 summarises baseline information regarding eligible patients, patients included and those not included. Of the 226 eligible patients, 91 (40.3%) fulfilled the gold standard for appendicitis, and the most frequent diagnoses besides AA were "nonspecific abdominal pain", ($n = 102, 45.1\%$) and

ovarian pathology (n = 11, 4.9%). In total, 48 patients were included without having blood samples evaluated as the clinical findings alone were sufficient indication for diagnostic laparoscopy, and 47 of these patients underwent surgery.

Comparing the proportion of correct diagnoses made by the surgeons before and after assessing the results from the blood samples showed no significant changes. Based on clinical findings alone, 76.1% (95% CI: 70.0-81.5) of the diagnoses were correct, and this was only marginally improved to 78.3% (95% CI: 72.4-83.5) after taking into consideration the results of the blood samples (p = 0.1797). **Table 2** summarises the diagnostic status before and after assessing the blood samples. The table shows that 4.0% (95% CI: 1.8-7.4, n = 9) of the diagnoses were changed after assessing the blood results and 77.7% (95% CI: 40.0-97.2, n = 7) of these were changed from incorrect to correct diagnoses.

Table 3 shows the sensitivity, specificity and predictive values as well as the proportion of correctly diagnosed AA and not-AA before and after assessing the blood samples. Slight increases in sensitivity, specificity and predictive values were seen after assessing the blood samples, but the 95% CI intervals largely overlapped. ROC curves showed an area under the curve (AUC) of 0.7189 and 0.7514 for the different values of CRP and leucocyte counts, respectively.

DISCUSSION

The results of CRP and leucocyte counts had little influence on the surgeons' clinical decision-making when diagnosing AA. Only a small proportion (4.0%) of the diagnoses was changed after assessment of the blood results, which shows that the surgeons rely strongly on the physical examination and the patient's history. We observed no significant change in the proportion of correct diagnoses from before to after, as well as no change in the sensitivity, specificity and predictive values of the surgeons' diagnoses. This indicates that the blood samples did not contribute with additional diagnostic information. However, on the other hand, they also had no negative influence on the decision concerning surgery for suspected acute appendicitis.

A study comparable to ours [11] showed similar results, finding that the results of leucocyte counts only led to alterations in management in two out of 100 cases. Furthermore, they found inconsistency in the clinical diagnosis and the number of leucocytes in 39 cases. In 74% of those cases, the leucocyte count would have led to incorrect management if the results had informed the decision-making. More recent studies have focused on the sensitivity and specificity of CRP and leucocyte counts, and the results have been varied [6-10]. A recent meta-analysis found that the AUC for CRP and leuco-

TABLE 1

Baseline demographics of patients suspected of acute appendicitis.

	Eligible (N = 226)	Included (N = 273)	Not included (N = 180)
Age, yrs, median (25-75 percentile)	27 (21-42)	29 (21-44)	28.5 (20-44.5)
Symptom period ^a , h, median (25-75 percentile)	24 (14-48)	28 (14-50)	24 (12-72)
Temperature, °C, median (25-75 percentile)	37.2 (36.7-37.8)	37.2 (36.7-37.8)	37.2 (36.8-37.7)
Male, n (%)	85 (37.6)	97 (35.6)	58 (32.2)
Operated, n (%)	128 (56.6)	146 (53.7)	72 (40.0)
Diagnosed appendicitis, n (%)	91 (40.3)	102 (37.5)	49 (27.2)
Negative laparoscopy ^b , n (%)	37 (28.9)	44 (30.1)	23 (31.9)

a) Duration of symptoms at the time of the physical examination.

b) Including normal appendices left in situ and normal histology results.

TABLE 2

Diagnosis before blood samples	Diagnosis after blood samples		Total
	correct	incorrect	
Correct	170	2	172
Incorrect	7	47	54
Total	177	49	226

Diagnostic status before and after assessing blood samples. The values are n.

TABLE 3

	Before	After	Proportion of correct diagnoses made by the surgeons and sensitivity, specificity and predictive values for acute appendicitis before and after blood sample evaluation. The values are % (95% CI).
Correct AA diagnosis	66.6 ^b (57.1-75.3)	68.4 ^d (59.1-76.8)	
Correct non-AA diagnosis	85.2 ^c (77.4-91.1)	88.4 ^e (81.0-93.7)	
Sensitivity	81.3 (71.8-88.7)	85.7 (76.8-92.2)	
Specificity	72.6 (64.3-79.9)	73.3 (65.0-80.6)	
Positive predictive value ^a	66.7 (57.1-75.3)	68.4 (59.1-76.8)	
Negative predictive value ^a	85.2 (77.4-91.1)	88.4 (81.0-93.7)	

AA = acute appendicitis; CI = confidence interval.

a) Based on a population of patients suspected of AA with a prevalence of AA of 0.403.

b) n = 74.

c) n = 98.

d) n = 78.

e) n = 99.

cytes, respectively, were 0.75 and 0.72. This is similar to our findings, although we found a higher value for leucocytes than for CRP [12].

Decision-making is an important part of surgery. Yet, the process is not well understood. Surgeons continuously make subconscious as well as conscious decisions with the primary goal of choosing the treatment that offers the greatest chance of success [13]. The decision-making process leading to appendectomy has been described by Larsson et al [14] as an interplay between two categories of influences – *the medical assessment of*

Acute appendicitis; surgeons rely more on clinical examination than on results of blood samples.



the patient's condition and contextual characteristics, such as working conditions in the department and the surgeon's own experience. Laboratory tests were only one of many factors in the decision algorithm, and they were regarded as having only a low informational value. Jacklin et al evaluated surgical decision-making by identifying the different steps in the handling of symptomatic gallstones [15]. This model involved aspects similar to those in the above-mentioned model for AA, and the model also showed that personal experience and intuition play a great role in surgical decision-making. In general, the available literature offers no unique answer as to how decision-making in surgery is carried out. However, one important factor seems to be the surgeon's own experience and interpretation of the symptoms. This is consistent with our findings where the surgeon's decision based on the physical examination and patient history was not influenced by laboratory findings. Furthermore, the complex decision pathway combined with varying symptoms in AA and individual interpretation explains why it is so challenging to make the correct diagnosis.

In total, 60.3% of the eligible patients were included. A larger proportion of the included patients were diagnosed with AA compared to the non-included patients. This may be explained by less severe symptoms among non-included patients. This was not registered and is a possible confounder. Therefore, the present results have to be interpreted with this in mind. Apart from this, the two groups of patients were similar. In the group of patient included as "no change" without blood sample evaluation, 47 out of 48 patients underwent surgery as planned. We therefore assume that blood samples would not have altered the clinical decision. A limitation of the present study design is that it was not possible to fully control whether the surgeons performed the physical examination before assessing the blood samples. A randomised design or controlling the viewing of the blood samples would have strengthened the results, but this was not possible due to ethical and logistic considerations. As described in the Method section, several precautions were taken to ensure that pro-

col was followed, and we found no indications of protocol violation. Another unknown factor was that the time between the physical examination and assessment of the blood results varied between cases; this could have been a confounder as the surgeon might gain information from a prolonged observation time. We have no information on the influence of the specific results of the blood samples on differential diagnoses.

AA is a difficult diagnosis as 28.9% of the patients operated on suspicion of AA had normal findings, and alternative diagnostic tools are still warranted for this most common condition. Routine measurement of conventional inflammatory markers is not the answer when diagnosing patients suspected of AA. Most importantly, the results of inflammatory markers do not seem to lead to harmful treatment of the patients.

CORRESPONDENCE: Sofie Tind. E-mail: fietind@hotmail.com

ACCEPTED: 9 October 2015

CONFLICTS OF INTEREST: Disclosure forms provided by the authors are available with the full text of this article at www.danmedj.dk

LITERATURE

1. Buckius MT, McGrath B, Monk J et al. Changing epidemiology of acute appendicitis in the United States: study period 1993-2008. *J Surg Res* 2012;175:185-90.
2. Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: nationwide analysis. *Arch Surg* 2002;137:799-804; discussion 804.
3. Flum DR, Morris A, Koepsell T. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. *JAMA* 2001;286:1748-53.
4. Drake FT, Florence MG, Johnson MG et al. Progress in the diagnosis of appendicitis: a report from Washington State's Surgical Care and Outcomes Assessment Program. *Ann Surg* 2012;256:586-94.
5. Parks NA, Schroeppel TJ. Update on imaging for acute appendicitis. *Surg Clin North Am* 2011;91:141-54.
6. Birchley D. Patients with clinical acute appendicitis should have pre-operative full blood count and C-reactive protein assays. *Ann R Coll Surg Engl* 2006;88:27-32.
7. Panagiotopoulou IG, Parashar D, Lin R et al. The diagnostic value of white cell count, C-reactive protein and bilirubin in acute appendicitis and its complications. *Ann R Coll Surg Engl* 2013;95:215-21.
8. Sengupta A, Bax G, Paterson-Brown S. White cell count and C-reactive protein measurement in patients with possible appendicitis. *Ann R Coll Surg Engl* 2009;91:113-5.
9. Tepel J, Sommerfeld A, Klomp HJ et al. Prospective evaluation of diagnostic modalities in suspected acute appendicitis. *Langenbecks Arch Surg* 2004; 389:219-24.
10. Vaughan-Shaw PG, Rees JR, Bell E et al. Normal inflammatory markers in appendicitis: evidence from two independent cohort studies. *JRSM Short Rep* 2011;2:43.
11. Bolton JP, Craven ER, Croft RJ et al. An assessment of the value of the white cell count in the management of suspected acute appendicitis. *Br J Surg* 1975;62:906-8.
12. Yu CW, Juan LI, Wu MH et al. Systematic review and meta-analysis of the diagnostic accuracy of procalcitonin, C-reactive protein and white blood cell count for suspected acute appendicitis. *Br J Surg* 2013;100:322-9.
13. Szatmary P, Arora S, Sevdalis N. To operate or not to operate? A multi-method analysis of decision-making in emergency surgery. *Am J Surg* 2010;200:298-304.
14. Larsson G, Weibull H, Larsson BW. Analysis of the decision-making process leading to appendectomy: a grounded theory study. *Scand J Psychol* 2004; 45:449-54.
15. Jacklin R, Sevdalis N, Darzi A, Vincent C. Mapping surgical practice decision making: an interview study to evaluate decisions in surgical care. *Am J Surg* 2008;195:689-96.