The use of diagnostic laparoscopy in patients with suspected appendicitis
van den Broek, W.T.

Citation for published version (APA):
van den Broek, W. T. (2002). The use of diagnostic laparoscopy in patients with suspected appendicitis
Which children could benefit from additional diagnostic tools in case of suspected appendicitis?

W.T. van den Broek
E.D. van der Ende
A.B. Bijnen
P.J. Breslau
D.J. Gouma

1 Department of Surgery, Medical Center Alkmaar, Alkmaar, The Netherlands
2 Department of Surgery, Rode Kruis Ziekenhuis, The Hague, The Netherlands
3 Department of Surgery, Academic Medical Center, Amsterdam, The Netherlands

Submitted for publication
ABSTRACT

Introduction: new diagnostic tools as ultrasound, CT-scan and diagnostic laparoscopy, have become available for children with suspected appendicitis. Because routine application of these tools might cause unnecessary delay, costs, patient discomfort and morbidity they should be reserved for equivocal cases in which it can contribute for diagnosis. The aim of this study is to develop a scoring system to identify this subgroup of children.

Patients and Methods: in two different periods: period 1, 99 consecutive children (group 1) and in period 2, 62 consecutive children (group 2) with suspected appendicitis were prospectively evaluated. Variables predicting appendicitis were obtained from group 1. By means of a regression analysis, a scoring system was created and applied to the patients of group 2. The clinical outcome, in terms of delayed and negative appendectomy rates obtained by clinical practice were compared with the results that would have been accomplished based on the scoring system. Hereafter the scoring system was applied and so externally validated in a group of children presented in another hospital (group 3, n=114).

Results: The variables: leukocytes count \( \geq 10.10^9/L \) (2 points); rebound tenderness (2 points) and temperature \( \geq 38^\circ C \) (1 point) correlated significantly with the diagnosis appendicitis. The scoring system was used to categorize patients into three groups: appendicitis unlikely, doubtful appendicitis and suspected appendicitis. The specificity and sensitivity of the scoring system were respectively 85% and 89%. Applying the scoring system would lead to comparable negative appendectomy rates of 8%, versus 6% using clinical judgement and a comparable number of performed laparoscopies: 26 vs. 31%. However, it could lead to a lower delayed appendectomy rate: (1 vs. 6%) and a lower perforation rate (0 vs. 11%). External validation showed comparable performed laparoscopies (32%) and delayed appendectomies (2%) rates but a higher negative appendectomy rate (19%), possibly due to a lower percentage of appendicitis in hospital 2 (47%) compared to hospital 1 (71%).

Leaving out temperature (1 point) would not change further strategy. Thus children can be observed if both leukocyte count is normal and rebound tenderness is absent. A laparoscopy should be performed if one of these two variables is present, and if both are present one could perform an appendectomy without other diagnostic tools.

Conclusion: The indication for a more selective use of laparoscopy can be established in children by using a simplified scoring system.
Which children could benefit from additional diagnostic tools in case of suspected appendicitis?

INTRODUCTION

The clinical diagnosis acute appendicitis remains difficult in children. Therefore initial misdiagnosis has been reported between 28 and 57% for children ≤ 12 year to nearly 100% for those 2 years or younger. From one hand, delay in diagnosis, especially in children, can lead to perforation and peritonitis and subsequently will increase morbidity and mortality. On the other hand, a negative appendectomy is associated with morbidity and costs due to hospital admission and should therefore be avoided. To increase diagnostic accuracy new modalities as ultrasound, CT-scan, and diagnostic laparoscopy have been introduced.

Although laparoscopic appendectomy is feasible in children and has shown subtle advantages (less postoperative pain and shorter hospital stay) over open appendectomy, appendectomy is still mostly performed by open surgery in the Netherlands as well as in many other European countries. Diagnostic laparoscopy for acute abdominal pain however, has been used for many years.

Performing diagnostic laparoscopy routinely, followed by an open procedure if the appendix appears to be inflamed, would lead to prolonged operation time and extra costs for that group of patients. Because of the relatively high costs and possible complications from this invasive diagnostic procedure, we reported previously that laparoscopy should be applied selectively on patients where it can most effectively contribute to obtain the diagnosis. Therefore, a method should be developed to define the subgroup of children with a doubtful diagnosis of appendicitis that can benefit from diagnostic laparoscopy. For this purpose a new scoring system was created, using a simple and feasible set of variables, which allows determining whether appendicitis is unlikely, equivocal or suspected.

Several scoring systems have already been developed in adults, achieving negative appendectomy rates between 5 to 17.5% and false negative rates between 0 to 7%. Applying these systems on patients in other hospitals however, lead to considerable higher negative appendectomy rates ranging between 13 to 37% and false negative rates ranging between 0.9 to 13% and are therefore not recommended as a standard tool for diagnostic decision-making in suspected appendicitis.

Therefore this study aims to define a new scoring system in an attempt to identify a group of children with a doubtful diagnosis of appendicitis that could optimally benefit from extra diagnostic tools such as diagnostic laparoscopy, ultrasonography or CT-scan. The scoring system is tested in another group of patients and finally externally validated by the application on a group of children presented in another hospital and referred by other general practitioners.
PATIENTS AND METHODS

The study was performed in two hospitals with each a different referral area from general practitioners: the Medisch Centrum Alkmaar (hospital 1), Alkmaar, The Netherlands and the Rode Kruis Ziekenhuis (hospital 2), The Hague, the Netherlands.

In hospital 1 were consecutive children (age ≤ 11 years), referred to the hospital by general practitioners for suspected appendicitis, prospectively evaluated in a consecutive period of 4 years. These children were divided into 2 groups. Group 1, patients included in the first 2 years. Group 2, the patients included in the second period of 2 years.

Amongst others, the variables: age, sex, duration of symptoms, rectal temperature, rebound tenderness and white cell count (WCC) were collected. Based on clinical judgment, patients were categorized as unlikely, doubtful or highly suspected having appendicitis, and accordingly they were observed, or a diagnostic laparoscopy or open procedure was performed. Imaging techniques such as CT-scan and ultrasound were not used.

Children were observed clinically or on an outpatient basis if this was socially and practically acceptable. In case of increasing symptoms they were re-evaluated immediately or otherwise routinely after one day at the emergency ward. Depending on the actual symptoms, these primarily observed patients were then discharged, underwent diagnostic laparoscopy, or an open appendectomy.

At laparoscopy, a normal appendix was left in place as a routine; an inflamed appendix was removed laparoscopically or by muscle-splitting incision, depending on the experience of the surgeon. During an open procedure the appendix was always removed irrespective of the macroscopic appearance. Histological examination was performed on all removed appendices.

Group 3 represents children (age ≤ 15 years) presented in hospital 2 who were prospectively evaluated for acute abdominal complaints as part of another study.

The same variables as used in hospital 1 were collected. There were no diagnostic laparoscopies performed in hospital 2 but they all underwent an ultrasound, which determined further treatment. Except for a lower percentage of patients with appendicitis in group 3, there were not differences between groups, as shown in table 1.

Table 1:

<table>
<thead>
<tr>
<th>Patient's characteristics</th>
<th>Group 1: patients hospital 1 period 1994-5</th>
<th>Group 2: patients hospital 1 period 1996-7</th>
<th>Group 3: patients hospital 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>99</td>
<td>62</td>
<td>114</td>
</tr>
<tr>
<td>Age (SE)</td>
<td>9.3 (2.0)</td>
<td>9.1 (2.0)</td>
<td>9.0 (3.3)</td>
</tr>
<tr>
<td>Male:</td>
<td>51 (51%)</td>
<td>35 (57%)</td>
<td>62 (54%)</td>
</tr>
<tr>
<td>Female:</td>
<td>48 (49%)</td>
<td>27 (43%)</td>
<td>52 (46%)</td>
</tr>
<tr>
<td>Appendicitis:</td>
<td>65 (66%)</td>
<td>50 (81%)</td>
<td>54 (47%) $p &lt; 0.001$</td>
</tr>
</tbody>
</table>
Which children could benefit from additional diagnostic tools in case of suspected appendicitis?

Data were analysed with the use of the SPSS computer program. Negative appendectomy rates, delayed (i.e. missed appendicitis) and perforation rates were calculated. In order to compare these results with the scoring system, the patients of whom we could retrieve all variables for the scoring system: 143/161 (89%) in hospital 1 and 114/114 (100%) in hospital 2, were considered.

Creating the scoring system

From the children of group 1, variables were obtained that correlated significantly with the prevalence of appendicitis by using $\chi^2$-square tests. Sensitivity and specificity rates of these variables were calculated.

Hereafter we performed a regression analysis to establish how strongly the individual variables contribute to the diagnosis appendicitis. Based on the regression coefficients, each significant variable was given a score. The final score was created by the sum of these scores, varying from 0 (appendicitis not likely) to 5 (appendicitis very likely).

Validation and clinical testing of the scoring system

The odds ratio (ratio of odds of positive test result in case of appendicitis with the odds of a positive test result in case of absence of appendicitis) is a measure that indicates the increased chance of having appendicitis per score. The scoring system was validated by comparing odds ratios for score of group 1 with the independent patient group 2. To evaluate the clinical value of the scoring system, data of both group 1 and group 2 were combined.

By choosing two cut-off points in our scoring system, patients were divided into three groups: low probability, amenable to observation; intermediate probability, necessitating further diagnostic tools such as diagnostic laparoscopy; and high probability of appendicitis, justifying immediate appendectomy by muscle-splitting incision. As the final diagnosis in all cases is known, the sensitivity, specificity rates can be calculated. Finally, the results that would have been obtained from applying the system were compared to the actual results of clinical practice, performed in the same series.

External validation of the scoring system

Hereafter the scoring system was applied to children from hospital 2. Number of performed laparoscopies, delayed appendectomies and negative appendectomy rates that would have been obtained if the scoring system was used in hospital 2 were calculated and compared with the results of hospital 1.
RESULTS

Creating the scoring system:

The following two objective variables: temperature (≥ 38°C) and White Cell Count (WCC) (≥ 10.10⁹/L) and the subjective variable: rebound tenderness, correlated significantly (p ≤ 0.01) with appendicitis. The other variables were not significant. By means of a logistic regression model these variables were allotted a certain weight by deriving them from the regression coefficients (RC) instead of using the exact RC, in order to make the scoring system workable in a practical setting (table 2).

Table 2:

Variables predicting appendicitis in group 1 (n = 99) obtained by χ² - tests.
Score based on regression coefficient (RC) of variable in regression analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>χ²</th>
<th>P-value: univariate</th>
<th>Regression: P-value: multivariate</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rebound tenderness:</td>
<td>0.87</td>
<td>0.78</td>
<td>40.195</td>
<td>&lt; 0.001</td>
<td>2.8102</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Leukocyte count: (≥ 10.10⁹/L)</td>
<td>0.80</td>
<td>0.76</td>
<td>29.816</td>
<td>&lt; 0.001</td>
<td>2.1302</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Temperature (≥ 38°C):</td>
<td>0.72</td>
<td>0.50</td>
<td>4.861</td>
<td>0.024</td>
<td>0.5536</td>
<td>0.3696</td>
</tr>
<tr>
<td>Duration symptoms: (&lt; 48 hrs)</td>
<td>0.77</td>
<td>0.41</td>
<td>3.531</td>
<td>0.051</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sex (male):</td>
<td>0.50</td>
<td>0.44</td>
<td>0.395</td>
<td>0.339</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Maximal score: 5

94
Which children could benefit from additional diagnostic tools in case of suspected appendicitis?

Validation of the scoring system:

The relative incidence of appendicitis for each score is shown in figure 1. As expected, the percentage of appendicitis increases with higher scores (score 0: 0% to score 5: 93%). The odds ratio of score was 2.4 in this group, which means that a difference in score of 1 point contributes to a 2.4 higher chance of having appendicitis. The odds ratio of score in group 2 was comparable, namely: 3.2 which validates it as a reproducible scoring system in this setting. As the findings in group 1 and group 2 do not differ, they have been combined for further analysis to evaluate its efficiency in the clinical setting.

Figure 1:

Distribution of appendicitis among the different scores in group 1, 2 and 3.
Chapter 7

Observed results:

From 143/161 (89%) children, all evaluated variables could be retrieved and were amenable for comparing with the scoring system. The other 18 (11%) children were not amenable because the variable temperature was lacking. Clinical practice led to an overall 6% negative appendectomy rate and 7% delayed operation rate. 31% of the patients underwent primary diagnostic laparoscopy. In the 3 different groups created by clinical judgment the percentage appendicitis was respectively 17%, 82% and 97%, shown in table 3.

Table 3: children hospital 1 (group 1 + 2)

Outcome of treatment of patients by clinical practice.

<table>
<thead>
<tr>
<th></th>
<th>Patients</th>
<th>Appendectomy</th>
<th>Appendicitis</th>
<th>Normal appendix removed</th>
<th>Perforation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation:</td>
<td>35</td>
<td>7 (20%)</td>
<td>6 (17%)</td>
<td>1 (3%)</td>
<td>4 (11%)</td>
</tr>
<tr>
<td>Laparoscopy:</td>
<td>45</td>
<td>41 (91%)</td>
<td>37 (82%)</td>
<td>4 (9%)</td>
<td>4 (9%)</td>
</tr>
<tr>
<td>Open operation:</td>
<td>63</td>
<td>63 (100%)</td>
<td>61 (97%)</td>
<td>2 (3%)</td>
<td>9 (14%)</td>
</tr>
<tr>
<td>Total:</td>
<td>43</td>
<td>111 (78%)</td>
<td>104 (73%)</td>
<td>7 (6%)</td>
<td>17 (12%)</td>
</tr>
</tbody>
</table>

Delayed operation rate: 6/104 (6%)
Negative appendectomy rate: 7/111 (6%)

Clinical testing of the scoring system:

The scoring system was used with two cut-off points to determine a treatment strategy: observation, laparoscopy or appendectomy. The low cut-off point will be determined by the sensitivity rate (indication for the capability of the test to rule out appendicitis), as the specificity rate (indication for the capability of the test to demonstrate appendicitis) will be useful in determining the level of the high cut-off point. Sensitivity and specificity rates of the different cut-off points for the scoring system applied to all patients are presented in figure 2.
Which children could benefit from additional diagnostic tools in case of suspected appendicitis?

Figure 2: Sensitivity and specificity rates for different cut-off points.

Varying the low and high cut-off points would affect the different numbers of delayed diagnosed appendicitis, performed laparoscopies and negative appendectomy rates. Considering figure 2, the optimal low cut-off point could be 2, i.e. patients with the scores: 0 and 1 would be observed (sensitivity rate: 0.89), in combination with a high cut-off point of 3 (specificity rate: 0.85), i.e. patients with scores: 4 and 5 will undergo directly appendectomy by muscle-splitting incision. This would (theoretically) lead to a delayed appendectomy rate of 1/104 (1%) and a negative appendectomy rate of 6/110 (5%). In this scenario 37/143 (26%) diagnostic laparoscopies would be performed, shown in table 4. As in our series, after diagnostic laparoscopy, in 9% a normal appendix was incorrectly removed, the overall negative appendectomy rate might in practice be higher, namely (3+6)/(30+83) = 8%, shown in table 4.
In the 3 different groups created by our cut-off points 3 and 6, the percentage appendicitis was respectively 4%, 73% and 93%, shown in table 4.

Table 4: children hospital 1 (group 1 + 2)
Theoretical results of outcome of treatment according to the scoring system, using cut-off points 2 and 4.

Low cut-off point = 2: patients with scores: 0 and 1 will be observed.
High cut-off point = 4: patients with scores 4 and 5 will undergo directly appendectomy by muscle-splitting incision.

Patients with score 2 and 3 will undergo diagnostic laparoscopy.

<table>
<thead>
<tr>
<th></th>
<th>Patients</th>
<th>Appendectomy</th>
<th>Appendicitis</th>
<th>Normal appendix removed</th>
<th>Perforation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation:</td>
<td>24</td>
<td>1 (4%)</td>
<td>1 (4%)</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Laparoscopy:</td>
<td>37</td>
<td>30 (81%)</td>
<td>27 (73%)</td>
<td>3 (9%)</td>
<td>5 (14%)</td>
</tr>
<tr>
<td>Open operation:</td>
<td>82</td>
<td>82 (100%)</td>
<td>76 (93%)</td>
<td>6 (7%)</td>
<td>12 (15%)</td>
</tr>
<tr>
<td>Total:</td>
<td>143</td>
<td>113 (79%)</td>
<td>104 (73%)</td>
<td>9 (8%)</td>
<td>17 (12%)</td>
</tr>
<tr>
<td>Delayed operation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative appendectomy:</td>
<td></td>
<td>1/104 (1%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

External validation of the scoring system:

Applying the scoring system on children from group 3 would lead to 2% delayed operation rate with 36 (32%) performed diagnostic laparoscopies. When an expected 9% negative appendectomy rate of the children who undergo diagnostic laparoscopy is taken into account, the overall negative appendectomy rate would be 19%. In the 3 different groups created by the scoring system the percentage appendicitis was respectively 4%, 36% and 80%, shown in table 5.
Which children could benefit from additional diagnostic tools in case of suspected appendicitis?

Table 5: Group 3 (children hospital 2)

Theoretical results of outcome of treatment according to the scoring system, using cut-off points 2 and 4.

<table>
<thead>
<tr>
<th></th>
<th>Patients</th>
<th>Appendectomy</th>
<th>Appendicitis</th>
<th>Normal appendix removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation:</td>
<td>28</td>
<td>1 (4%)</td>
<td>1 (4%)</td>
<td>-</td>
</tr>
<tr>
<td>Laparoscopy:</td>
<td>36</td>
<td>15 (41%)</td>
<td>13 (36%)</td>
<td>3 (9%)</td>
</tr>
<tr>
<td>Open operation:</td>
<td>50</td>
<td>50 (100%)</td>
<td>40 (80%)</td>
<td>10 (20%)</td>
</tr>
<tr>
<td>Total:</td>
<td>114</td>
<td>66 (58%)</td>
<td>54 (47%)</td>
<td>13 (20%)</td>
</tr>
<tr>
<td>Delayed operation:</td>
<td></td>
<td>1/53 (2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative appendectomy:</td>
<td></td>
<td>13/67 (19%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DISCUSSION

The present study showed that improved results could be obtained by applying laparoscopy selectively, in case of clinical doubt in the diagnosis, as there are subsets of children in which the diagnosis appendicitis is almost certain, respectively very unlikely. This study demonstrates that such subsets can be identified by good clinical judgment but even more effective by a simple, reproducible scoring system since the scoring system would lead to lower delayed operation rates (1 vs. 6%) and lower perforation rates in the primary observed patient group (0 vs. 11%) as shown in tables 3 and 4. Because using the scoring system would also reduce the number of performed laparoscopies (26 vs. 31%) it would lead to lower costs for children with suspected appendicitis. Negative appendectomy rates (8% vs. 6%) will not improve by using the scoring system.

Although the variable temperature was significant in predicting appendicitis (p = 0.024) when using the (univariate) \( \chi^2 \)-test, in the multivariate analysis temperature was not significant (0.3696). Therefore temperature contributed for only 1 point in our scoring system. When using cut-off points 1 and 4, temperature does not contribute in determining the further strategy as both scores 0 and 1 will be observed, both scores 3 and 4 will undergo diagnostic laparoscopy and both scores 5 and 6 direct appendectomy. So the scoring system could be simplified by concluding that if both elevated leukocyte count and rebound tenderness are not present, the child can be observed, if one of the two variables is present a laparoscopy should be performed and if both variables are present one could directly proceed with an appendectomy.

Due to lack of the variable temperature, only 89% of our children could be evaluated. Our simplified scoring system could now be applied on all children (n = 161), which would lead to a
delayed appendectomy rate of 1/115 (1%), 45 (28%) performed laparoscopies and a negative appendectomy rate of 7/122 = 6%. Our clinical practise in all children lead to 6/115 (5%) delayed appendectomies, 52 (32%) performed laparoscopies and a negative appendectomy rate of 11/126 = 9%. Because in 5/52 (10%) of children who underwent a laparoscopy a normal appendix was removed, the corrected negative appendectomy by our simplified scoring system would be 11/126 = 9%.

So also the simplified scoring system could lead to improved patient care by lowering delayed appendectomy and perforation rates, compared to our clinical practise. Scoring could therefore be of help in the pre-clinical setting and for less experienced physicians by preventing unnecessary delay in applying further diagnostic tools on these children. Unfortunately, a relatively large number of normal appendices (9%) were removed in the laparoscopy group for unclear reasons and this was taken into account in evaluating the scoring system. The overall negative appendectomy rate of 8% could therefore considerably be improved, namely to 5%, if the indication for appendectomy after laparoscopy would be more thoughtful imposed.

Outcome of scoring systems could be dependent on the prevalence of appendicitis in the referred patient population, which depend on different referral policies by general practitioners. Therefore, it was also applied on children referred to another hospital by other general practitioners. The percentage of appendicitis was higher in hospital 1 (71%) compared to hospital 2 (47%). As a result, applying the scoring system led to a comparable delayed operation rate (1%) and comparable performed laparoscopies (35%) but also to a relative high negative appendectomy rate of 19%. So the scoring system performed well for excluding appendicitis but performed less for demonstrating appendicitis.

Scoring systems for adults have previously been devised by others to decide whether or not to operate. In the present study the scoring system is used to define a group of patients in whom additional laparoscopy or other diagnostic tools can be helpful. The correlation between score and likelihood of appendicitis allows us to choose cut-off points to increase the diagnostic yield of additional diagnostic tests. The number of needed diagnostic laparoscopies will be determined by the percentage of negative appendectomies that is regarded acceptable. Cost-efficiency studies (laparoscopies versus perforation risk and negative appendectomies) should optimally determine the choice of those points, but are not yet available.

In conclusion, the subset of patients in which laparoscopy can contribute in obtaining the diagnosis appendicitis can be identified by good clinical judgment but also by a simple, reproducible scoring system.

Acknowledgements

We would like to thank Ms. E.S.M. de Lange from the bio-statistics department of the VU Medical Center, Amsterdam, the Netherlands for her contribution in the statistical analysis.
Which children could benefit from additional diagnostic tools in case of suspected appendicitis?

REFERENCES


Which children could benefit from additional diagnostic tools in case of suspected appendicitis?


