Validation of Clinical Scoring System in the Primary Care of Children with Suspected Acute Appendicitis

Niramol Tantemsapya, M.D., FRCST*, Mongkol Laohapensang, M.D., FRCST*, Suthipol Udompunturak, M.S.**

**Department of Surgery, **Clinical Epidemiology Unit, Office of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*

ABSTRACT

Objective: To establish a scoring system for the diagnosis of acute appendicitis in children in order to decrease the rate of negative and delayed appendectomies and delayed.

Methods: Record charts of pediatric patients admitted with suspected appendicitis at the Division of Pediatric Surgery, Department of Surgery, Faculty of Medicine Siriraj Hospital during July 2000 - September 2001 were retrospectively reviewed. Clinical predictors were listed out and weighed as coefficients, which were later formed into a score equation. The score was then prospectively studied in the following year from October 2001-September 2002. The score was used to aid decision making in the management of cases suspected of appendicitis. The out come was compared to the conventional management based on histologically confirmed diagnosis.

Results: The retrospectively collected data listed out 14 significant variables. Using logistic regression, the parameters were weighed as coefficients, which became scores to each parameter. The score was applied to a consecutive series in the following year. By using the score, the patients were categorized into three groups: discharged, observed and operated. The accuracy of the score is 85.83% with a positive predictive value of 96.55% and negative predictive value of 75.81%. From this study group, the negative appendectomy rate is 4.17% and the perforation rate is 6.67%. Compared to the previous year during which 133 children were treated under the conventional management, the negative appendectomy rate was 4.51% and the perforation rate was 21.8%

Conclusion: The scoring system is an accurate diagnostic tool, which proves the importance of history taking, physical examination in coordination with laboratory investigations and close observation. Altogether these parameters can reduce the rate of misdiagnosis and delayed diagnosis of pediatric appendicitis.

Keywords: Pediatric appendicitis score; Clinical score; Child

Siriraj Med J 2005;57:481-485

Acute appendicitis accounts for 1/3 of all pediatric hospital admissions with abdominal pain;[1,2] and it is the most common cause of surgical emergency in childhood.[3,4] The diagnosis of appendicitis is based on clinical findings in correlation with laboratory investigations. Currently, there is no appropriate diagnostic tool that would make a certain preoperative diagnosis. The rate of negative appendectomy ranges between 10-46%[1-4] whereas the perforation rate is 18-30% increasing the morbidity rate to 5-11%[5] with prolonged hospital stay and postoperative complications in up to 15% of patients.[2] Attempts to increase diagnostic accuracy have included computer-aided, ultrasonic imaging and laparoscopy. However, all these methods need special equipments and a radiologist who is 24-hour available, which do not appear cost effective.[5-9] Intensive inpatient observation can significantly reduce the negative appendectomy rates,[1-3,7,10] but there are no clear criteria as to foretell exactly which patient needs immediate surgery, or inpatient observation, or simply to be advised and discharged for outpatient follow up.

Clinical scoring systems to aid diagnosis of acute appendicitis have been described since 1986 as Alvarado Score and have subsequently been validated in surgical practices for adults.[1,7] Moreover, there were modified Alvarado,[11-13] Ohmann,[14] Eskelinen[9] and Lindberg Scores[3] being published later but not as well known as the former.[9,14,15] Some studies used computer-based or imaging-aided to add to the scoring in order to increase its sensitivity and specificity.[8,9,15-17] These reports have included small numbers of children, which have brought about the interest to establish this study. Current reported studies were mostly based on clinical diagnosis, which is difficult to attain in young children and varies amongst clinicians, depending on individual experiences and skills.[18] Migratory right iliac fossa pain maybe overlooked by a distressed child and rebound tenderness or rectal tender-

Correspondence to: Niramol Tantemsapya
E-mail: tulob@yahoo.com

TABLE 1. The clinical predictors and their score.

Clinical predictors	Score
Onset : Insidious	5
Acute	0
Duration > 12 hours	2
Gastrointestinal symptoms (nausea, vomiting or anorexia)	5
Constipation	-2
History of previous visits	4
Migration of pain	2
Progression of pain	6
Temperature > 37.5°C	3
Leukocytosis: WBC < 12,000	0
12,000-15,000	1
> 15,000	2
Polymorphonuclearcytosis > 80%	3
Urine white cells > 10/hp	7
Per-rectal exam: No impact feces	7
Guarding	5
Rebound tenderness	6

Cut points of the score combined : 1-18 Advise and follow-up 24-48 hours
19-34 Admit with active inpatient observation
35-57 Prepare for surgery

ness can be difficult to interpret accurately.[1,5,18] Some physical examinations that are too painful to be carried out in children may no longer be recommended including the rebound tenderness. Recently, a study was coincidently carried on in London with a 5-year study of 1,170 children prospectively. The study was able to reduce the rate of negative appendectomy to lower than 5%.[19]

The purpose of this research is to create a scoring system to aid the diagnosis of acute appendicitis in children and to emphasize in history taking more than the symptom of pain and tenderness. Despite a high statistic significance of the physical signs, painful examinations result in loss of trust and cooperation of the child. The score is expected to provide as a useful, practical tool to help reduce the rate of misdiagnosis of acute appendicitis in children including negative explorations and delayed treatment causing complicated appendicitis.

MATERIALS AND METHODS

Record charts of pediatric patients ages 2-15 years old admitted with acute abdominal pain at the Division of Pediatric Surgery, Department of Surgery, Faculty of Medicine Siriraj Hospital from July 2000 to September 2001 were retrospectively reviewed. There were 133 cases included and their data were collected, i.e. age, sex, onset, and duration of illness, medical history of visits, previous treatment, signs and symptoms such as location, migration, progression and character of pain, nausea, vomiting, anorexia and fever. Associated constipation or diarrhea, physical examination and laboratory investigations including leucocytosis, neutrophilia and urinalysis were all assembled. The final diagnosis was based on histology. The clinical predictors were calculated, comparing the diagnosis, using Mann-Whitney and Chi-Square tests, the significant predictors with $p < 0.05$ were listed out. By logistic regression the parameters were weighed as coefficients, which represented the relation of the clinical predictors to the diagnosis, i.e. whether or not it was acute appendicitis. The coefficients were formed into a group of scores and equations. The score used, was from the equation with the highest sensitivity and specificity proved by the most area under ROC (Receiver Operating Characteristic) curve. The score was first evaluated by applying to the retrospective group and the accuracy was then calculated. A questionnaire was then created, according to the significant parameters and scores, and was again used to collect the information to calculate each individual score. After being applied to the study group in the retrospective chart review, the score was then compared to the final diagnosis again and the cut points of the results were calculated by using cross-tabs, the patients were grouped into three categories, namely:

Group 1: Patients who were to be advised and appointed within 24-48 hours

Group 2: Patients who needed active inpatient observation

Group 3: Patients who needed surgery

Thereafter, the score was prospectively applied to the study group of the same age and department in the following year from October 2001-September 2002. The score was assessed separately with respect to the decision of the attending physician. The outcome was compared to the conventional management based on histologically confirmed diagnosis.

RESULTS

The retrospective data collected listed out 19 clinical predictors from a study group 133 cases (87 boys and 46 girls). By the Mann-Whitney test, 14 variables were significant with $p < 0.05$. The parameters were onset, duration, gastrointestinal symptoms, constipation, history of previous visits, pyrexia, migration, progression of pain, leukocytosis, neutrophilia, urine white cells, per-rectal examination, abdominal guarding and rebound tenderness. The score of each parameter and the cut points of the sum were summarized in table 1.

After applying the score to the retrospective study group, of the 133 cases listed out, there were 33 cases with scores categorized into group 1 and 7 cases into group 2, all were doing well after follow-up and none had acute appendicitis making no false negative. Of the 93 cases that were categorized into group 3, all were operated, having six negative appendectomies and 29 ruptured appendicitis.

During the prospective study period, 120 patients (63 boys and 57 girls) with a mean age of 8.4 years old (range 2-14) were included. From the score calculated, 19 patients were categorized into group 1, who should be advised and followed-up within 24-48 hours; 43 cases into group 2, who should be admitted for close observation; and, 58 in group 3, who were to be prepared for surgery. With respect to the attending physician's decision, 12 patients were discharged home and were doing well after two days with no abdominal pain, 52 were observed in which 15 were later operated with no delay in the diagnosis; and 56 underwent surgery right away. Of the seven patients assumed to be sent home but were admit-

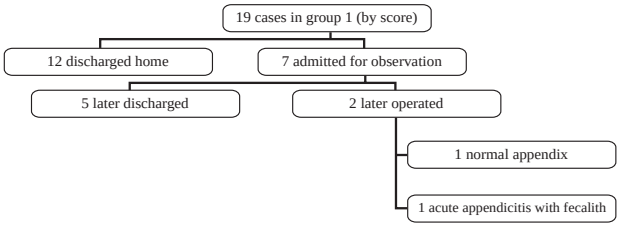

Fig 1. Patients in group 1 (discharge home).

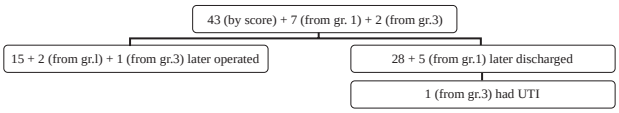

Fig 2. Patients in group 2 (observation).

ted according to the physicians' decisions, five were later discharged and doing well; whilst two were subsequently operated due to unremitting abdominal pain (Fig 1-3). Of the two cases referred to, one had normal appendix while the other acute appendicitis with fecalith.

There were two cases in the operable group from the score who were closely observed and one was consequently operated while the other had severe urinary tract infection. However, the patient was treated after appendicitis, and was later excluded and the symptoms subsided.

Overall, there were 71 operated cases with five cases having lymphoid hyperplasia assumed to be normal appendices and 8-ruptured appendices. In the observed group, none had ruptured appendicitis, which means that there was no delayed diagnosis in this group. Of the five negative appendectomies, one had a score of 35 with insidious constant abdominal pain, one had a score of 15 with unremitting pain migrated to the right iliac fossa; the other three cases were in the observed group with progressive pain. Two of the mentioned negative appendectomies specimens were impacted with fecal material and fecalith, which may be the cause of progressive unremitting pain making a future operation necessary.

According to the previous year of 133 children (87 boys and 46 girls), 93 of them underwent surgery and 40 were observed and discharged. From the scores calculated, the outcome was analyzed to have had an accuracy of 94.73%, sensitivity of 100% and specificity of 82.5%; the patients were in the study group. With conventional management, the negative appendectomy rate was 4.51% and ruptured appendicitis 21.8%.

After applying the score to the prospective study group, the accuracy of the score is 85.83% with 95% confident interval between 78.44% - 90.96%, sensitivity of 97.18%, specificity of 95.92%, positive predictive value of 96.55% and negative predictive value of 75.81%. The negative appendectomy rate is 4.17% with 95% confident interval between 3.05%-15.45% (Wilson Method) and the perforation rate is 6.67%.

DISCUSSION

Clinical scoring system is proved to be useful to aid decision-making in approaching children suspected of appendicitis.[20] During the retrospective period, most collectable data from reviewed charts were mainly focused on physical examination. But the problem in diagnosing pediatric cases is that sometimes it is difficult to perform a thorough physical examination under many

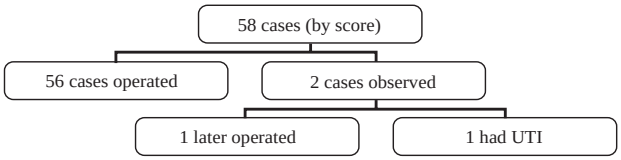

Fig 3. Patients in group 3 (surgery).

certain circumstances. Moreover, some physical signs are not as obvious as in adults or maybe misled by many other signs. That is why we conducted this study to find some other ways to support the data to reach the final diagnosis. The physical examinations that were proved to be a significant parameter in this study including per-rectal examination, abdominal guarding and rebound tenderness, may sometimes be difficult to attain and diversify relying on the child's cooperation and the physician's experience. In young children pain results in loss of confidence, trust and cooperation of the child. However, they were included owing to a significant relation to the diagnosis. Therefore, these parameters had to be coordinated with history taking and laboratory investigation.[21-24] Both the misdiagnosis and the delayed ones are related to errors in clinical assessment which the physicians had put too strong emphasis on pain and tenderness, and too little attention to the duration of the symptoms and objective signs of the inflammation. The score pointed out that history taking should be emphasized in pediatric patients more than pain in physical examination.

Constipation is often a cause of non-specific abdominal pain in children. Owing to the fact that it is unlikely to coincide with appendicitis, constipation was weighed minus 2. Children with constipation often fell into the observation group, and their symptoms usually do not progress but subsided after rectal enema or saline irrigation.

Per- rectal examination having no fecal impaction was a highly specific indicator more than rectal tenderness, weighing as much as 7. The sign was seldom paid attention to by the physcian. We were taught to perform the per-rectal examination always in every case suspected of appendicitis. But we mainly focused on rectal tenderness on the right side and the bulging cul-de-sac. Fortunately, we were able to collect the data from the reviewed charts, noting the empty rectum or otherwise fecal impaction to correlate with the diagnosis to find the outstanding correlation.

Other than history of constipation, we also collected information concerning history of previous visits which were later proved to have a highly correspondence to the diagnosis with a score of 4. Because the difficulty in diagnosis of acute appendicitis in children making the parents "shop around" seeking the right therapy for their child. Whereas some medications such as antibiotics or antispasmodics may relieve the signs and symptoms but worsen the disease process causing a more complicated appendicitis. With a history of previous visits the score would be raised to a closer number for observation or maybe surgery.

According to this study, there was one child whose score fell into the third group. He was not operated right away because his abdominal signs were weak and his symptoms did not progress. Afterwards, he was proved to have urinary tract infection. Since the presence of urine white cells were positively high in severe cases of appendicitis making the parameter highly sensitive, patients with

urinary tract infection tend to have high scores. The laboratory results should be considered along with other parameters. Urinary analysis was not statistically significant in some studies, however, the test should be performed routinely. Thus urine exam with abnormally high white cells may have a false positive result from the score resulting in less specificity therefore a further study with a larger group of population should be carried on.

There were two cases that fell into the first group to be discharged home but were subsequently operated due to progressive unremitting pain. And one case turned out to have a normal appendix. This is one false negative case, which the child had an abnormally long insidious onset more like a chronic abdominal pain. His pain was localized with no guarding and no associated fever along with an almost normal lab results so his score fell into group 1 but we decided to observe his symptoms making a future appendectomy necessary due to unremitting pain. His histology report revealed a fecalith which is suspected to be the cause of his chronic unremitting pain. In suspected cases, we still prefer close observation to sending the child home and having a complication later. Moreover, of the five negative appendectomies two had impact fecal material and fecalith which in the literature there is no proof of whatsoever whether such cases need a future operation or not but the unremitting pain consequently disappeared.

No test is ideal with 100% certainty to exact the diagnosis of appendicitis. As long as constipation, fecal impaction and other non-specific abdominal pain can present indistinguishable symptoms and signs to acute appendicitis, recurrent clinical evaluation is of utmost importance.

CONCLUSION

A thorough physical examination is difficult to access in pediatric patients. Both misdiagnosis and delayed diagnosis are related to errors in the clinical assessment with which the physician paid too strong emphasis on tenderness and too little attention to history taking. The scoring system is a useful diagnostic tool proved to be accurate and signifies the importance of history taking, physical examination in coordination with laboratory investigations and close observation. Altogether these clinical predictors can reduce the rate of misdiagnosis and delayed diagnosis of appendicitis in children.

REFERENCES

1. Macklin CP, Radcliffe GS, Merei JM, Stringer MD. A prospective evaluation of the modified Alvarado score for acute appendicitis in children. Ann R Coll Surg Engl 1997;79:203-5.
2. Christian F, Christian GP. A simple scoring system to reduce the negative appendicectomy rate. Ann R Coll Surg Engl 1992;74:281-5.
3. Dado G, Anania G, Baccarani U, Marcotti E, Donini A, Risaliti A, et al. Application of a clinical score for the diagnosis of acute appendicitis in childhood: a retrospective analysis of 197 patients. J Pediatr Surg 2000;35:1320-2.
4. Sally L, Reynolds SL. Missed appendicitis in a pediatric emergency department. Pediatr Emerg Care 1993;9:1-3.
5. Ravitch MM, Welch KJ, Benson CD. In: Ravitch MM. Pediatric Surgery, 3rd ed. Chicago: Year book Medical Publisher: Chapter Abdomen,1986:470-7.
6. Jawaid A, Asad A, Motei A, Munir A, Bhulto E, Choudry H, et al. Clinical scoring system: a valuable tool for decision making in cases of acute appendicitis. J Pak Med Assoc 1999;49:254-9.
7. Jahn H, Mathiesen FK, Neckelmann K, Hovendal CP, Bellstrom T, Gottrup F. Comparison of clinical judgement and diagnostic ultrasonography in the diagnosis of acute appendicitis: experience with a score-aided diagnosis. Eur J Surg 1997;163:433-43.
8. Arnbjornsson E. Scoring system for computer-aided diagnosis of acute appendicitis. The value of prospective versus retrospective studies. Ann Chir Gynaecol 1985;74:159-66.
9. Eskelinen M, Ikonen J, Lipponen P. The value of history-taking, physical examination, and computer assistance in the diagnosis of acute appendicitis in patients more than 50 years old. Scand J Gastroenterol 1995;30:349-55.
10. Graff L, Radford MJ, Werne C. Probability of appendicitis before and after observation. Ann Emerg Med 1991;20:503-7.
11. Impellizzeri P, Centonze A, Antonuccio P, Turiaco N, Cifala S, Basile M, et al. Utility of a scoring system in the diagnosis of acute appendicitis in pediatric age. A retrospective study. Minerva Chir 2002;57:341-6.
12. Owen TD, Williams H, Stiff G, Jenkinson LR, Rees BI. Evaluation of the Alvarado score in acute appendicitis. J R Soc Med 1992;85:87-8.
13. Malik AA, Wani NA. Continuing diagnostic challenge of acute appendicitis: evaluation through modified Alvarado score. Aust N Z J Surg 1998;68:504-5.
14. Ohmann C, Yang Q, Franke C. Diagnostic scores for acute appendicitis. Abdominal Pain Study Group. Eur J Surg 1995;161:273-81.
15. Fenyo G. Routine use of a scoring system for the decision-making in suspected acute appendicitis in adults. Acta Chir Scand 1987;153:545-51.
16. Fenyo G, Linberg G, Blind P, Enochsson L, Oberg A. Diagnostic decision support in suspected acute appendicitis: validation of a simplified scoring system. Eur J Surg 1997;163:831-8.
17. Teicher I, Landa B, Cohen M, Kabnick LS, Wise L. Scoring system to aid in diagnosis of appendicitis. Ann Surg 1983;198:753-9.
18. Andersson RE, Hugander AP, Ghazi SH, Ravn H, Offenbartl SK, Nystrom PO, et al. Why does the clinical diagnosis fail in suspected appendicitis? Eur J Surg 2000;166:796-802.
19. Samuel M. Pediatric appendicitis score. J Pediatr Surg 2002;37:877-81.
20. Ramirez JM, Deus J. Practical score to aid decision making in doubtful cases of appendicitis. Br J Surg 1994;81:1829-30.
21. Andersson RE, Hugander AP, Ghazi SH, Ravn H, Offenbartl SK, Nystrom PO, et al. Diagnostic value of disease hitory, clinical presentation and inflammatory parameters of appindicitis. World J Surg 1999;23:133-40.
22. Delany HM. Appendicitis: trends and risks. J Assoc Acad Minor Phys 1996;7:70-7.
23. Korner H, Sondenaa K, Soreide JA, Nysted A, Vatten L. The history is important in patients with suspected acute appendicitis. Dig Surg 2000;17:364-8.
24. Stringer MD, Drake DP. Diagnosis and management of acute abdominal pain in childhood. Curr Pediatr 1991;1:2-7.

บทคัดย่อ

การใช้ระบบการให้คะแนนเพื่อช่วยในการวินิจฉัยผู้ป่วยเด็กที่สงสัยภาวะไส้ติ่งอักเสบ

นิรมล ตันเต็มทรัพย์ พ.บ., มงคล เลาหเพ็ญแสง พ.บ., สุทธิพล อุดมพันธุรัก วท.ม.

**ภาควิชาศัลยศาสตร์, **สถานส่งเสริมการวิจัย, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กทม 10700, ประเทศไทย.*

วัตถุประสงค์ : เพื่อสร้างระบบการให้คะแนนมาช่วยในการวินิจฉัยภาวะไส้ติ่งอักเสบในเด็กที่จะสามารถลดอัตราการผ่าตัดโดยไม่จำเป็น ตลอดจนการวินิจฉัยที่ล่าช้าจนไส้ติ่งแตก

วิธีการ: ศึกษาย้อนหลังประวัติผู้ป่วยเด็กมาด้วยอาการปวดท้อง สงสัยไส้ติ่งอักเสบ โดยรวบรวมข้อมูลที่มีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับการวินิจฉัยมาให้คะแนนตามน้ำหนักความสำคัญต่อการวินิจฉัยและสร้างเป็นสูตรสมการเพื่อคิดค่าคะแนน นำคะแนนที่ได้ไปใช้คำนวณในกลุ่มศึกษาปีต่อไป ได้ค่าคะแนนที่สามารถแบ่งผู้ป่วยเป็น 3 กลุ่ม ได้แก่ กลุ่มที่ให้กลับบ้าน, สังเกตอาการ และผ่าตัด

ผลการศึกษา: การใช้ระบบการให้คะแนนมีความแม่นยำ 85.83% และสามารถลดอัตราการผ่าตัดโดยไม่จำเป็นจาก 4.51% เป็น 4.17% และสามารถลดอุบัติการณ์การเกิดไส้ติ่งแตกลงจาก 21.8% เป็น 6.67%

สรุป: การศึกษานี้แสดงให้เห็นว่า ระบบการให้คะแนนเป็นเครื่องมือที่มีประโยชน์ช่วยในการวินิจฉัยภาวะไส้ติ่งอักเสบในเด็กโดยเน้นความสำคัญของการซักประวัติร่วมกับผลการตรวจร่างกายและการตรวจทางห้องปฏิบัติการ ตลอดจนการสังเกตอาการอย่างใกล้ชิดจะสามารถลดภาวะแทรกซ้อนที่เกิดจากการผ่าตัดโดยไม่จำเป็นและการวินิจฉัยล่าช้าลงได้