Indian J Pediatr
DOI 10.1007/s12098-016-2106-2

ORIGINAL ARTICLE

The Heidelberg Appendicitis Score Simplifies Identification of Pediatric Appendicitis

Michael Boettcher[1,2] · **Thomas Breil**[2] · **Patrick Günther**[2]

Received: 22 November 2015 / Accepted: 4 April 2016

Abstract
Objective To identify the factors that facilitate the diagnosis of pediatric appendicitis.
Methods Institutionally approved retrospective, single center analysis of all patients with acute abdominal pain was done. Medical history, symptoms, laboratory and radiologic findings of all children presenting with abdominal pain were evaluated. To identify the best predictors, uni- and multi-variate analysis were used.
Results In 2 years, 431 patients fulfilled the inclusion criteria. Data was complete in all subjects. Of these, 156 (36.2 %) suffered from appendicitis. The best discriminators for appendicitis were clinical and ultrasound features. The four best factors were identified by CART analysis (continuous abdominal pain, tenderness on the right lower quadrant, rebound tenderness and conspicuous ultrasound) and combined to the Heidelberg Appendicitis score. A positive score (>3 features) is highly predictive for acute appendicitis (PPV 89.3 %, NPV 94.9 %) and includes all cases of perforated appendicitis.
Conclusions It is possible to predict acute appendicitis in children. The decision making process can be simplified by the proposed Heidelberg Appendicitis score, which is comprised of four factors. It has great potential to facilitate and accelerate the diagnosis of pediatric appendicitis.

Michael Boettcher and Thomas Breil contributed equally to this work.

✉ Michael Boettcher
michboettcher@gmail.com

1 Department of Pediatric Surgery, UKE Medical School, Martinistr. 52, 20246 Hamburg, Germany

2 Department of Surgery, Section of Pediatric Surgery, University Hospital Heidelberg, Im Neuenheimer Feld 110, 69120 Heidelberg, Germany

Published online: 26 April 2016

Keywords Prediction · Appendicitis · Children · Appendicitis scores · Heidelberg appendicitis score

Abbreviations

WBC	White blood count
CRP	C-reactive protein
US	Ultrasound
CT	Computerised tomography
SD	Standard deviation
CART	Classification and regression tree analysis
OR	Odds ratio
CI	Confidence intervals
PPV	Positive predictive value
NPV	Negative predictive value
PAS	Pediatric appendicitis score

Introduction

Appendectomy is one of most commonly performed emergency operations worldwide with a lifetime risk for appendicitis of 8.6 % in males and 6.7 % in females [1]. Despite the relatively high incidence of this emergency, it remains a complex diagnosis. Especially in young children, appendicitis may present atypically and many non-surgical conditions such as gastroenteritis, mesenteric lymphadenitis, respiratory tract or urinary tract infections may mimic acute appendicitis [2].

Advanced imaging like ultrasound (US), magnetic resonance imaging (MRI) and computerised tomography (CT) offer high sensitivity (>70 %) and specificity rates (>95 %) [3, 4]. However, despite dramatic increases in imaging the outcome has not been improved substantially; the rate of negative appendectomies has decreased, but perforation rate has remained high (25 %–50 %) [5]. Over-utilization of these modalities has resulted in prolonged emergency department visits, increased costs, need for anesthesia for MRI in young children and in case of CT, exposure to ionizing radiation [3]. The latter is especially fretful, as it has been reported that the ionizing radiation of abdominal CTs induces solid cancer in every 300 to 390 scans in children [6]. This indicates a need to re-evaluate the diagnostic assessment for this disease [7].

The aim of this current study was to identify factors or a set of factors that help to predict appendicitis and to facilitate the management of this disease.

Material and Methods

An institutionally approved retrospective analysis of all children younger than 16 y with suspected appendicitis that presented at the Pediatric Surgery Department of the University Hospital, Heidelberg, Germany from January 2009 until December 2010 was performed.

Patients were selected from the hospital database. All data were collected using patient charts, operation theater records, office notes as well as ICD-9 and CPT codes. Children with chronic medical conditions (like cystic fibrosis, inflammatory bowel disease, sickle cell anemia) were excluded from the study.

Data gathered for medical history included duration of symptoms, progression of symptoms, associated symptoms, menstruation, previous episodes of symptoms, previous operations, medication, and psychiatric diseases. All participants had been examined physically by a resident and/or an attendent of pediatric surgery. Various aspects of the physical examination were recorded including abdominal distension, abdominal pain and tenderness, guarding, rebound tenderness and psoas sign. Recorded laboratory values included white blood count (WBC) and C-reactive protein (CRP) at admission. Ultrasound studies were performed by a resident or attending radiologist. The US report was checked in retrospect for the following factors: appendix outer diameter, prominent inflamed perifocal fat and appendix wall hyper-perfusion. Moreover, in all children Tzanakis scoring consisting of four variables (6 points for ultrasound demonstrating appendicitis, 4 points for tenderness in the right lower quadrant, 3 points for rebound tenderness and 2 points for leukocyte count >12,000/ul) was calculated [8].

Finally, operation method, conversion rate, complications, time until discharge, re-admission rate and findings of pathology were documented. Distinction between appendicitis and non-appendicitis was based on the pathology report. Furthermore, distinction between perforated and non-perforated appendicitis relied upon intraoperative assessment (visible perforation) and pathology report. In case of discrepancy between surgical and pathological findings, the more severe case was chosen.

Statistical analysis was performed using SPSS 21.0. Data are presented as mean (standard deviation, SD). Differences between groups were calculated using Mann-Whitney test and chi-square or ANOVA and Kruskal Wallis test when comparing more than two groups and are expressed by value as well as 95 % confidence intervals (CI). Receiver operation characteristic (ROC) analysis was performed, before converting continuous variables to categorical variables for univariate analysis. Finally, to identify the best predictors for appendicitis a classification and regression tree analysis (CART) was performed. CART has some advantages over multivariable regression analysis like the ability to utilize large numbers of predictor variables and non-reliance on the underlying distributions for statistical inference [9]. The level of significance was set at 0.05.

Results

In 2 y 431 children with suspected appendicitis were included in the study. In all the children, all the features of history, clinical examination, laboratory testing and ultrasound examination were available. In total, 157 children (36.4 %) suffered from appendicitis, of whom 47 (29.9 %) had perforated appendicitis. Children that ultimately had no appendicitis suffered mostly from gastroenteritis (44.0 %) or mesenteric lymphadenopathy (33.8 %). There was a slight male dominance (52.0 % males *vs*. 48.0 % females). The mean age was 10.12 (3.26) y.

In 171 (36.5 %) of all children with abdominal pain, a surgical intervention was performed: mostly appendectomy (95.1 %) followed by ovarian cystotomy (2.9 %) and Meckel's diverticulum resection (1.2 %). Analysis of pathology reports revealed a negative appendectomy in 6 children (negative appendectomy rate 4.8 %).

Children with appendicitis suffered significantly more often from continuous abdominal pain (appendicitis 89.1 % *vs*. non-appendicitis 10.5 %; $p < 0.001$) and showed significantly less bowel dysfunctions or pyrexia than non-appendicitis. Physical examination showed besides abdominal pain on the right lower quadrant, significantly more often rebound tenderness (appendicitis 80.0 % *vs*. non-appendicitis 20.7 %; $p < 0.001$), guarding or rigidity (appendicitis 23.9 % *vs*. non-appendicitis 1.5 %; $p < 0.001$) and a positive psoas sign (appendicitis 38.7 % *vs*. non-appendicitis 4.0 %; $p < 0.001$) in children with appendicitis.

In order to establish transferability into other clinical settings, a threshold value that is close to a typical limit was chosen. ROC analysis showed good results for typical threshold limits of WBC ($12 \times 10^3/\mu L$), CRP (20 mg/dl) and appendix size (6 mm diameter). Laboratory analysis showed significantly higher WBC and CRP levels for appendicitis compared to non-appendicitis, but they were not elevated in all children. Urine analysis was rarely pathological, but showed significantly more often a (nitrite-negative) leucocytouria in cases of appendicitis (appendicitis 7.75 % *vs*. non-appendicitis 1.5 %; $p = 0.002$).

In children with appendicitis, appendix was visualized significantly more often (appendicitis 86.5 % *vs*. non-appendicitis 16.0 %; $p < 0.001$). Moreover, the appendix size was significantly larger (appendicitis 8.99 *vs*. non-appendicitis 5.76 mm) and the threshold of 6 mm appendix size was exceeded significantly more frequent (both, $p < 0.001$). The other appendix specific ultrasound signs like appendix wall hyperemia (appendicitis 66.7 % *vs*. non-appendicitis 2.5 %; $p < 0.001$) or surrounding hyperechoic fat (appendicitis 56.4 % *vs*. non-appendicitis 1.8 %; $p < 0.001$) were less common, yet significantly more frequent in appendicitis compared to non-appendicitis. US diagnosis of appendicitis was made significantly more often in children with appendicitis (appendicitis 83.3 % *vs*. non-appendicitis 12.0 %; $p < 0.001$). Moreover, free fluid was found more often in the appendicitis group (appendicitis 45.5 % *vs*. non-appendicitis 20.0 %; $p < 0.001$).

Using univariate analysis several factors were identified that are predictive of appendicitis. In terms of patient history, continuous pain is highly predictive and intermittent pain makes appendicitis unlikely (Table 1). Moreover, guarding and a positive psoas sign are highly predictive of appendicitis. The highest predictors were ultrasound signs like appendix wall hyperemia and hyperechoic perifocal fat (PPV >93.7 %), whereas appendix visualization and US demonstrating appendicitis had very high NPV's (Table 1).

As pain intensity as well as presence or absence of pain migration and anorexia were rarely documented (<10 %), evaluation of the Alvarado score, the Pediatric Appendicitis score (Samuel score) and the Lintula score was not reasonable in the current series [10–12]. Data for the Tzanakis score was available in all cases. The score was significantly higher in children with appendicitis [appendicitis 13.12 (2.94) *vs*. non-appendicitis 6.16 (2.92) points] and the cut-off of 8 was exceeded significantly more frequent (appendicitis 91.0 % *vs*. 20.4 % non-appendicitis) than in non-appendicitis. Thus a score ≥8 was highly predictive of pediatric appendicitis: PPV 71.6 % (67.7–74.4 %), NPV 94.0 % (90.6–96.4 %).

Using CART analysis the best factors to diagnose appendicitis within a group of children with acute abdominal pain were identified: 1. continuous abdominal pain, 2. tenderness on the right lower quadrant, 3. rebound tenderness and 4. conspicuous US (appendix diameter >6 mm, prominent inflamed perifocal fat and appendix wall hyper-perfusion). These four factors are the basis of the Heidelberg Appendicitis score (HAS). The best discrimination threshold for the score was 3 out of 4 points with an area under the curve of 0.94 (0.91–0.97; $p < 0.001$). The score was highly predictive for appendicitis: OR 153.92 (69.78–347.70), PPV 89.3 % (85.1–92.4 %), NPV 94.9 % (92.4–96.6 %), Sensitivity 91.0 % (86.8–94.1 %), Specificity 93.8 % (91.4–95.6 %). Additionally, it performed very well for advanced appendicitis (phlegmonosa, gangranosa and perforata): OR 133.74 (43.97–455.13), PPV 76.1 % (72.5–82.9 %), NPV 97.7 % (94.3–99.2 %), Sensitivity 96.8 % (92.2–98.9 %), Specificity 81.6 % (78.8–82.9 %). All cases of perforated appendicitis had a positive score (>3 points).

If only children with a positive HAS (>3 points) had been operated, 5.1 % of appendicitis cases (but none with perforation) would have been missed. A HAS of 4 includes 63.6 % of all appendicitis cases and only three cases without appendicitis (all of whom were older than 10-y-old). Moreover, the score performs well for various age subgroups: 0–4 y PPV 83.3 % (44.9–83.3 %), NPV 100 % (86.4–100 %); 4–8 y PPV 85.3 % (73.5–91.5 %), NPV 94.9 (88.1–98.5 %); 8–12 y PPV 92.6 % (85.6–96.9 %), NPV 93.6 % (89.7–95.9 %); 12–16 y PPV 88.0 % (79.5–92.3 %), NPV 96.3 % (91.0–98.9 %).

Table 1 Univariate analysis for factors predicting appendicitis using chi-square

	PPV (%)	NPV (%)
Female	30.9 (26.1–35.9)	58.9 (54.5–63.5)
Continuous pain	82.7 (78.3–86.1)	93.5 (90.7–95.7)
Nausea	38.5 (35.4–41.5)	69.2 (61.9–76.0)
Vomiting	39.1 (33.3–44.9)	65.8 (61.9–69.7)
Frequent vomiting (≥3×)	29.7 (16.7–46.4)	63.2 (62.0–64.8)
Rare bowels (<1/3 d)	6.3 (0.3–31.8)	62.7 (62.4–63.6)
Frequent bowels (≥3×/d)	1.3 (0.2–4.5)	88.7 (88.1–90.5)
Dysuria	66.7 (24.2–94.0)	64.4 (63.8–64.8)
Pyrexia (>38.5°)	27.2 (19.7–35.8)	61.0 (58.6–63.7)
Fever before abdominal pain	6.3 (1.1–21.7)	61.4 (61.0–62.6)
Rebound tenderness	68.5 (63.5–72.8)	87.2 (83.6–90.3)
Guarding/rigidity	90.2 (76.6–96.8)	69.5 (68.1–70.2)
Positive psoas sign	84.5 (75.5–91.5)	73.3 (71.4–74.7)
Blood: WBC ≥12 × 10^3 /μL	58.5 (53.9–62.6)	84.0 (79.8–87.7)
Blood: CRP ≥20 mg/dl	56.6 (49.8–63.0)	74.1 (70.7–77.4)
Urine: leucocytes	75.0 (47.9–91.6)	65.5 (64.4–66.1)
US: appendix visualization	75.4 (70.9–79.0)	91.7 (88.4–94.2)
US: appendix ≥6 mm	81.4 (76.2–85.7)	89.5 (86.5–91.9)
US: wall hyperemia	93.7 (87.8–97.1)	83.8 (81.7–84.9)
US: hyperechoic fat	94.6 (88.0–98.0)	79.9 (78.1–80.8)
US: free fluid	56.3 (48.8–63.6)	72.1 (69.0–75.1)
US: demonstrating appendix	79.3 (74.3–83.4)	90.3 (87.2–92.8)
Tzanakis score ≥8	71.6 (67.7–74.4)	94.0 (90.6–96.4)
Heidelberg score ≥3	89.3 (85.1–92.4)	94.9 (92.4–96.6)

Data is presented as positive predictive value (PPV) and negative predictive value (NPV) as value (95 % Confidence Interval)

Discussion

The aim of the study was to identify factors that facilitate the diagnosis of pediatric appendicitis. Advanced imaging techniques like MRI or CT provide excellent diagnostic value but have several disadvantage: they are time consuming, costly and require anesthesia in case of MRI in young children or exposure to ionizing radiation in case of CT [3]. Consequently, several scoring systems that bundle clinical features like Pediatric Appendicitis score, Alvarado Appendicitis score and Lintula Appendicitis score have been described [10–12]. These scores have been validated in a wide variety of pediatric populations: in validation studies the median sensitivity has been reported to be 88–92 % with specificity between 49 and 91 % [13].

However, due to their complexity, deployment in clinical practice is very limited [13]. The scores rely on different weighting factors for variable features, thereby, complicating instead of facilitating the process (Table 2). Moreover, scores like Tzanakis score (that was validated by the results of this current study) have the disadvantage that they comprise of laboratory tests that may delay diagnosis up to an hour. A simpler score like the HAS consisting of only four features (each with the same importance) that can be accessed by the pediatrician or surgeon himself and in real time, enables immediate decision-making in the ER or on the ward. In this study the HAS showed excellent predictive capabilities for pediatric appendicitis. It was particularly sensitive for advanced appendicitis (97 %) and picked up all cases of perforated appendicitis. A negative score almost excludes appendicitis (NPV 95 %). This simple score has thus the potential to facilitate diagnosis of pediatric appendicitis and to speed up the decision-making process in children with acute abdominal pain.

Routine diagnostic workup for appendicitis includes white blood count, CRP and urine analysis. As in previous studies, WBC and CRP were elevated in children with appendicitis and the two features are thus predictive of appendicitis [14–16]. However, normal CRP and WBC levels may not exclude appendicitis [14, 15]. In the present study, almost 25 % of the children had normal WBC and CRP levels. This is in accordance with a previous report that concluded that due to the non-specific nature of most inflammatory variables, it is not possible to determine the need for further radiological investigation or surgical intervention, just by relying on laboratory tests [17]. Consequently, as blood sampling can be demanding and may be traumatizing for younger children, one might consider skipping it in patients that are very likely to have appendicitis, like children with a HAS of four.

Most limitations of the current study are inherent in a retrospective review. It is possible that patients sought help at another institution after discharge. Moreover, US is user dependent and trained Radiologists, Pediatrician or Pediatric Surgeons are needed to diagnose appendicitis in children to use the HAS.

In conclusion, the Heidelberg Appendicitis score and the proposed algorithm simplifies the decision-making process. It has the potential to facilitate diagnosis of pediatric appendicitis and to speed up decision-making substantially. It might be safe to operate only on patients with a positive HAS and re-evaluate the remaining children continuously. However, prospective validation of this concept is needed.

Table 2 The differences between the common scoring systems for pediatric appendicitis [8, 10–13]. It illustrates the major disadvantage of the previous scoring systems: due to the use of different factors, that are weighted differently, clinical utilization of these scores is complicated. WBC >10/ml for Alvarado and Pediatric Appendicitis score and >12/ml for Tzanakis, Neutrophilia >75 % of WBC for Alvarado and >7.500/mm^2 for Pediatric Appendicitis score. US demonstrating appendicitis which includes appendix diameter >6 mm, surrounding echogenic inflamed fat and hyperemia in the wall on color Doppler

	Alvarado	Pediatric	Lintula	Tzanakis	Heidelberg
Continuous pain					1
Tenderness right lower quadrant	2	2	4	4	1
Rebound tenderness	1		7	3	1
US demonstrating appendicitis				6	1
WBC (>10/>12)	2	1		2	
Guarding			4		
Pain migration	1	1	4		
Nausea/Vomiting	1	1	2		
Temperature (>38°)	1	1	3		
Anorexia	1	1			
Neutrophilia (>7 5 %/>7.500/mm^2)	1	1			
Male			2		
Percussion tenderness		2			
Intensity of pain			2		
Decreased bowel sounds			4		
Positive, if equal or higher than	5/10	6/10	21/32	8 /15	3/4

Contributions MB: Conceptualized and designed the study, acquired the data, drafted the initial manuscript, and approved the final manuscript as submitted; TB: Conceptualized and designed the study, acquired the data, and approved the final manuscript as submitted; PG: Conceptualized and designed the study, approved the final manuscript as submitted and will act as guarantor for the paper. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Compliance with Ethical Standards

Conflict of Interest None.

Source of Funding None.

References

1. Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: nationwide analysis. Arch Surg. 2002;137:799–804; discussion 804.
2. Becker T, Kharbanda A, Bachur R. Atypical clinical features of pediatric appendicitis. Acad Emerg Med. 2007;14:124–9.
3. Karakas SP, Guelfguat M, Leonidas JC, Springer S, Singh SP. Acute appendicitis in children: comparison of clinical diagnosis with ultrasound and ct imaging. Pediatr Radiol. 2000;30:94–8.
4. Aspelund G, Fingeret A, Gross E, et al. Ultrasonograph/MRI versus CT for diagnosing appendicitis. Pediatrics. 2014;133:586–93.
5. Cheong LH, Emil S. Outcomes of pediatric appendicitis: an international comparison of the United States and Canada. JAMA Surg. 2014;149:50–5.
6. Miglioretti DL, Johnson E, Williams A, et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. JAMA Pediatr. 2013;167:700–7.
7. Flum DR, Morris A, Koepsell T, Dellinger EP. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. J Am Med Assoc. 2001;286:1748–53.
8. Tzanakis NE, Efstathiou SP, Danulidis K, et al. A new approach to accurate diagnosis of acute appendicitis. World J Surg. 2005;29: 1151–6; discussion 7.
9. Terrin N, Schmid CH, Griffith JL, D'Agostino RB, Selker HP. External validity of predictive models: a comparison of logistic regression, classification trees, and neural networks. J Clin Epidemiol. 2003;56:721–9.
10. Samuel M. Pediatric appendicitis score. J Pediatr Surg. 2002;37: 877–81.
11. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986;15:557–64.
12. Lintula H, Pesonen E, Kokki H, Vanamo K, Eskelinen M. A diagnostic score for children with suspected appendicitis. Langenbeck's Arch Surg/Deutsche Gesellschaft fur Chirurgie. 2005;390:164–70.
13. Kulik DM, Uleryk EM, Maguire JL. Does this child have appendicitis? A systematic review of clinical prediction rules for children with acute abdominal pain. J Clin Epidemiol. 2013;66:95–104.
14. Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. Br J Surg. 2004;91:28–37.
15. Yang HR, Wang YC, Chung PK, Chen WK, Jeng LB, Chen RJ. Laboratory tests in patients with acute appendicitis. ANZ J Surg. 2006;76:71–4.
16. Kwan KY, Nager AL. Diagnosing pediatric appendicitis: Usefulness of laboratory markers. Am J Emerg Med. 2010;28: 1009–15.
17. Scheinfeld MH, Mahadevia S, Stein EG, Freeman K, Rozenblit AM. Can lab data be used to reduce abdominal computed tomography (CT) usage in young adults presenting to the emergency department with nontraumatic abdominal pain? Emerg Radiol. 2010;17:353–60.