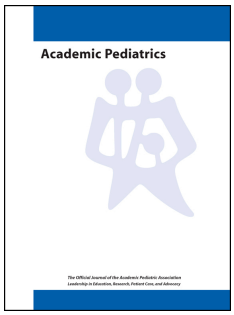
Accepted Manuscript



A Clinical Score to Predict Appendicitis in Older Male Children

Anupam B. Kharbanda, MD, MSc, Michael C. Monuteaux, ScD, Richard G. Bachur, MD, Nanette C. Dudley, MD, Lalit Bajaj, MD, MPH, Michelle D. Stevenson, MD, MS, Charles G. Macias, MD, MPH, Manoj K. Mittal, MD, Jonathan E. Bennett, MD, Kelly Sinclair, MD, Peter S. Dayan, MD, MSc

PII: S1876-2859(16)30504-6

DOI: 10.1016/j.acap.2016.11.014

Reference: ACAP 945

To appear in: *Academic Pediatrics*

Received Date: 11 July 2016

Revised Date: 14 November 2016

Accepted Date: 17 November 2016

Please cite this article as: Kharbanda AB, Monuteaux MC, Bachur RG, Dudley NC, Bajaj L, Stevenson MD, Macias CG, Mittal MK, Bennett JE, Sinclair K, Dayan PS, for the Pediatric Emergency Medicine Collaborative Research Committee of the American Academy of Pediatrics, A Clinical Score to Predict Appendicitis in Older Male Children, *Academic Pediatrics* (2016), doi: 10.1016/j.acap.2016.11.014.

A Clinical Score to Predict Appendicitis in Older Male Children

Anupam B. Kharbanda, MD, MSc[1], Michael C. Monuteaux, ScD[2], Richard G. Bachur MD[2], Nanette C. Dudley, MD[3], Lalit Bajaj, MD, MPH[4], Michelle D. Stevenson, MD, MS[5], Charles G. Macias, MD, MPH[6], Manoj K. Mittal, MD[7], Jonathan E. Bennett MD[8], Kelly Sinclair, MD[9], Peter S. Dayan, MD, MSc[10] for the Pediatric Emergency Medicine Collaborative Research Committee of the American Academy of Pediatrics

[1] Department of Pediatrics Emergency Medicine, Children's Hospitals and Clinics of Minnesota
[2] Division of Emergency Medicine, Boston Children's Hospital, Harvard Medical School, Boston, MA
[3] Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT
[4] Department of Pediatrics, University of Colorado School of Medicine, Denver, CO
[5] Department of Pediatrics, University of Louisville, Louisville, KY
[6] Department of Pediatrics, Baylor College of Medicine, Houston, TX
[7] Department of Pediatrics, The Children's Hospital of Philadelphia and Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
[8] Department of Pediatrics, Alfred I. duPont Hospital for Children, Jefferson Medical College, Wilmington, DE
[9] Department of Pediatrics, University of Missouri, Kansas City, MO
[10] Department of Pediatrics, Columbia University, NY, NY

Address Correspondence to: Anupam Kharbanda, Children's Hospitals and Clinics of Minnesota, 2525 Chicago Avenue South, Minneapolis, MN 55404 Anupam.kharbanda@childrensmn.org, 612-813-6903

Short Title: Appendicitis Score for Male Patients

Abbreviations:
ANC- Absolute neutrophil count;
CRF- Case report forms;
CT-Computerized tomography;
US – Ultrasound
IRB-Institutional review board;
PEM-CRC- Pediatric Emergency Medicine Collaborative Research Committee;
PPV- Positive predictive value;
RLQ- Right lower quadrant;
WBC – White Blood cell count;
AUC – Area Under the Curve

Key Words: Appendicitis, pediatrics, clinical scoring systems

Word count: Abstract: 225, main text: 2604

Funding Source: This study was supported by grant UL1RR024156 from the National Center for Research Resources, a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research.

Financial Disclosure: The authors have indicated they have no financial relationship relevant to this article to disclose

Conflict of Interest: The authors have indicated they have no potential conflicts of interest.

Abstract

Objective: To develop a clinical score to predict appendicitis among older, male children who present to the emergency department (ED) with suspected appendicitis.

Methods: Patients with suspected appendicitis were prospectively recruited at 9 pediatric EDs. A total 2,625 patients enrolled; a subset of 961 males, age 8-18 were analyzed in this secondary analysis. Outcomes determined by pathology, operative reports and follow-up calls. Clinical and laboratory predictors with < 10% missing data and Kappa > 0.4 were entered into a multivariable model. Resultant beta-coefficients were used to develop a clinical score. Test performance was assessed by calculating the sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV], and likelihood ratios [LRs].

Results: The mean age was 12.2 years; 49.9% (480) had appendicitis, 22.3% (107) had perforation, and the negative appendectomy rate was 3%. In patients with and without appendicitis, overall imaging rates were 68.6% (329) and 84.4% (406), respectively. Variables retained in the model included maximum tenderness in RLQ, pain with walking/coughing or hopping, and the absolute neutrophil count. A score ≥ 8.1 had a sensitivity of 25% [95% CI 20-29%], specificity of 98% [96-99%] and PPV of 93% [86-97%] for ruling in appendicitis.

Conclusion: We developed an accurate scoring system for predicting appendicitis in older boys. If validated, the score may allow clinicians to manage a proportion of these male patients without diagnostic imaging.

What's New

Given recent emphasis on reducing diagnostic imaging as well as healthcare costs, we believe that our high risk clinical score could provide clinicians a more judicious and standardized approach to the care of male children with possible appendicitis.

Background

Appendicitis represents a common and challenging diagnosis within pediatric emergency medicine. A clinician's ability to diagnosis appendicitis based on historical and physical examination findings alone is variable, with a sensitivity of 75% and specificity of 78%.[1] This diagnostic uncertainty, coupled with a desire to reduce negative appendectomy rates, has led to a heavy reliance on diagnostic imaging such as CT, ultrasound, and MRI.[2-4] Recent data has indicated a reduction in the utilization of CT at Children's Hospitals, with an increase in total diagnostic imaging rates (use of ultrasound and MRI above and beyond the declines in CT use.[5] These trends may be problematic and although not associated with direct exposure to ionizing radiation, the mixed test performance of ultrasound, could potentially lead to unnecessary testing and increased health care expenditures.[6] For this reason, a more nuanced approach in which risk for appendicitis is more accurately determined may offer clinical benefit.

Clinical scoring systems can help in identifying patients at high or low risk for appendicitis.[7-9] Unfortunately, prospective validation of these scores has shown mixed test performance, and thus limited their acceptance as alternatives to diagnostic imaging.[10,11] The heterogeneous presentation of children with possible appendicitis, especially among females and young children, may be an important reason for the lack of success of these rules.[12] In comparison, male patients are known to present with more typical findings for appendicitis and have fewer alternative etiologies for right lower quadrant pain, and thus may serve as better target populations for an appendicitis clinical scoring system.[13] Therefore, in this study we sought to identify which male patients were at highest risk for appendicitis The ultimate benefit

for such a rule may be to identify a sub-population of patients that require urgent referral for surgical evaluation or for whom diagnostic imaging is not required to confirm the diagnosis.

Methods

Study Design and Setting

We conducted a planned secondary analysis of a prospective, observational study of patients with suspected appendicitis at 9 pediatric emergency departments (PED) located in Children's Hospitals. Study subjects were enrolled from March 2009 through April 2010. All enrolling sites were members of the Pediatric Emergency Medicine Collaborative Research Committee (PEMCRC) of the American Academy of Pediatrics. The PEM-CRC reviewed and approved the final study protocol. Each participating site's Institutional Review Board (IRB) also approved the study. Six IRBs granted a waiver of written informed consent/assent and instead obtained verbal consent. At the three remaining sites, written consent from the guardians and assent from children seven years of age and older was obtained.

Study Patients

In the parent study, we enrolled children and adolescents between 3 and 18 years of age who presented to the ED with acute abdominal pain of < 96 hours duration and were being evaluated for suspected appendicitis. "Suspected appendicitis" was defined as those patients who were being evaluated by blood tests (e.g. complete blood count), radiologic studies (CT and/or US) and/or a surgical consultation for the purpose of diagnosing appendicitis. In the current analysis, we limited our analytic sample to males between the ages of 8-18 years. We excluded patients with any of the following conditions: prior abdominal surgery (e.g. gastrostomy tube, abdominal hernia repair), chronic gastrointestinal illness or abdominal pain

(e.g. inflammatory bowel disease, chronic pancreatitis, chronic/recurrent appendicitis), sickle cell anemia, cystic fibrosis, a medical condition affecting the provider's ability to obtain an accurate history (e.g. significant language/developmental delay), or history of abdominal trauma within 7 days of evaluation. We also excluded patients who had radiologic studies (CT or US) of the abdomen performed prior to ED arrival. Study procedures related to training of site staff, patient enrollment, standardized data collection, and transmission to the central data management warehouse have been described previously.[14]

Study Procedure and Data Collection

Site principal investigators received standardized training, a detailed manual of operations and instructions on the proper completion of case report forms (CRF's). CRF's were completed by a pediatric emergency medicine attending/ fellow or resident physician with attending oversight. CRF's were completed prior to knowledge of CT or US results. The decision to obtain laboratory studies, radiologic studies or surgical consultation was not dictated by study protocol. We conducted telephone follow-up (in English or Spanish, as appropriate) within 2 weeks of the index ED visit to determine resolution of signs and symptoms, visits to other sites of care and need for surgery. If we were unable to contact the guardian, research coordinators reviewed the medical record for 90 days after the index PED visit to determine if the patient underwent a CT, US, or operation at that specific facility.

Outcome Measures

The primary outcome was presence or absence of appendicitis. Final diagnosis of appendicitis was determined by pathology, operative reports or by telephone follow-up. For those who underwent an appendectomy, we determined the presence or absence of

appendicitis by pathology reports. The presence or absence of perforation was determined from the attending surgeon's written operative report. In cases where a non-surgical diagnosis was assigned, we contacted the family between 14-21 days after the ED visit to assess for resolution of signs and symptoms, visits to other sites of care and need for surgery.

Data Analysis

We used standard descriptive statistics to describe our two groups (patients with and without appendicitis). Potential predictors were selected from review of the prior literature and were collected prospectively during patient enrollment. For the present analysis, we only included predictors with less than 10% missing data and at least moderate inter-rater reliability (Kappa>0.4) in the male subgroup.[15] The predictors analyzed were (coded as binary variables unless otherwise indicated): age (in years), duration of pain (categorized as <12, 12-23, 24-35, 36-47, 48-71, and ≥72 hours), history of anorexia, history of nausea, history of emesis migration of pain to right lower quadrant (RLQ), focal pain in the RLQ, tenderness (coded as mild, moderate, or severe), right-sided abdominal tenderness, maximum tenderness in the RLQ, presence of rebound tenderness, guarding, and pain with walking, coughing or hopping. For this analysis, "unsure" or "don't know" responses were coded as missing data. Additionally, we included in the regression analysis the white blood cell count (WBC) and absolute neutrophil count (ANC), both as continuous measures.

Logistic Regression

We estimated a multivariable logistic regression model with appendicitis diagnosis as the dependent variable and the predictor variables as described above as the independent variables. To determine the most parsimonious model without sacrificing discriminative ability,

we eliminated the least predictive covariate (as determined by the covariates' p value) and statically compared the area under the curve (AUC) of the full model to that of the reduced model (i.e., the null hypothesis held that the AUCs were equal). A significant reduction in AUC would indicate that the eliminated variable contributed significantly to the predictive ability of the model and should be retained. A non-significant AUC would indicate that the variable could be eliminated without any corresponding loss of predictive ability. We conducted an iterative model reduction procedure, repeating the steps described above until the AUC comparison test was not significant. The resulting model was considered our final model for subsequent analyses.

We developed a clinical score by using the adjusted β coefficients from the final model and multiplying them by a patient's own values for the corresponding variables, and then summing over the resulting set of products to arrive at a patient-specific final score. Thus, this score represented a sum of a patient's risk factors, with each factor weighted by its predictive ability in the multivariable model.

Finally, we assessed the test performance of the clinical score at several high risk cut-points. Because we had no *a priori* reason to use a specific value for our clinical score as a cut-point to define appendicitis case status, we report the test performance across a range of values for the positive predictive value (PPV) that we believe may influence clinical management. Thus, we created three cut-points corresponding to PPVs of 85, 90, and 95%. For each cut-point, a patient was classified as a case (i.e., with appendicitis) if the calculated score for the given patient was greater than or equal to the cut-point, and a non-case otherwise. Test performance of each cut-point was assessed by calculating the sensitivity,

specificity, and likelihood ratios [LRs]. All analyses were conducted with STATA version 13.1 (College Station, TX).

Results

Participant Characteristics

From 2625 patients in the parent study, our analytic dataset consisted of 961 males with suspected appendicitis (Figure 1). The mean (SD) patient age in the study sample was 12.2 years (standard deviation: 2.6 years). Overall, 49.9% (n = 480) of patients had appendicitis, 22.3% (n = 107) had perforated appendicitis and the negative appendectomy rate was 3% (n = 17) (Table 1). In patients with appendicitis, the overall imaging rate was 68.6% (n = 329), the rate of CT utilization was 65.0% (n = 214), and the rate of US was 42.9% (n = 141). In patients without appendicitis, the overall imaging rate was 84.4% (n = 406), the CT utilization rate was 72.7% (n = 295), and the US rate was 42.6% (n = 173).

Development of the Appendicitis Clinical Score

The bivariate association between each potential predictor and the diagnosis of appendicitis is displayed in Table 2. All predictors except for age and duration of pain were statistically associated with appendicitis. The number of patients with missing data varied by predictor (e.g. n=2 for history of emesis; n=72 for ANC). The number of patients with complete data on all potential predictors was 719, which constituted the sample for the remainder of the analyses.

We first estimated the full model predicting appendicitis with all available predictors (AUC = 0.89, 95% CI = 0.86, 0.91). Four predictors (maximum tenderness in RLQ, rebound

tenderness, pain with walking/coughing/hopping, and ANC) emerged as statistically significant. After conducting the model reduction procedure, twelve predictors were eliminated without any significant reduction in the AUC. Three predictors (maximum tenderness in RLQ, pain with walking/coughing/hopping, and ANC) could not be dropped without a significant reduction in the AUC and so were retained. Table 3 displays the results from the final multivariable model. The area under the curve for this model was 0.88 (95% CI = 0.85, 0.90). The resulting appendicitis clinical score, calculated from the adjusted beta coefficients provided by this multivariable model ranged from 0.29 to 12.6 (mean ± standard deviation = 5.7 ± 2.1). We created three cut-points of this clinical score, corresponding to our a priori determined values for the PPV. These cut-points and corresponding PPVs were 6.2 (85%), 7.2 (90%), and 8.1 (93%), respectively. Although we intended to use a PPV of 95%, the maximum PPV provided by the risk score was risk score was 93%.

Application of Risk Score

Table 4 displays the performance of the clinical score to classify patients according to the appendicitis diagnosis. Among patients who screened positive using the cut point corresponding to a PPV of 93%, (n=96), 89 of these patients had a diagnosis of appendicitis. The use of advanced imaging among this cohort was 63% (n=60).

Discussion

Using data from a large, prospective, multi-site study of patients with suspected appendicitis, we developed an accurate clinical score to predict appendicitis in males 8-18 years. The clinical score allows for identification of population of male patients at high risk,

which may help facilitate more appropriate utilization of advanced diagnostic imaging and resource allocation for these patients.

Several previous studies have developed appendicitis prediction models.[7-9,16] The most widely cited models to risk stratify children for appendicitis were derived by Samuel [9] and Alvarado.[7] Both studies reported strong test performance, with the Alvarado's study reporting a sensitivity of 81% and specificity of 74% while the PAS study reported a sensitivity of 100% and a specificity of 92%. Independent validation studies, however, have noted varied performance, thus limiting their widespread acceptance.[1,10,16,17] In one recent publication, the authors attempted to improve the performance of the Samuels score through alternative cut-points, unfortunately the PPV was no better than 85% with this approach.[11] A more recent meta-analysis examined a range of cut-points for the Samuels and Alvarado score, which unfortunately also concluded that no high risk score adequately predicted appendicitis among pediatric patients (maximum PPV of 85% with a pre-test probability of 40%).[18]

Rather than attempting to develop a rule that would be applicable all patients, we elected to design a rule that was age and gender specific. We hypothesized this narrow population would have less variability in presentation and a more limited differential diagnosis. Accordingly, our developed clinical scoring system provided a higher specificity and PPV than is reported in the previous studies.[10,11,16] It is important to note that similar historical and physical examination variables have been found to be significant across all of the previous prediction models. In our study, maximum tenderness in RLQ, rebound tenderness, pain with walking/coughing/hopping, as well as the ANC were the most important factors. The major

differentiation we note is that we did not select a discrete cut-point for the ANC in our study. By calculating a risk score across utilizing a wide range of values for the ANC, we were able to develop a better performing rule.

Ultimately, in order for a high risk appendicitis score to be accepted by surgeons and ED physicians, it will need to perform as well as current management options when integrated into a clinical setting. In this regard, several studies have tested clinical pathways that include existing scoring systems (i.e. PAS score) to identify low, high and moderate risk patients and recommend management.[19-21] Unfortunately, when compared to judgment of experienced clinicians, these clinical pathways may not perform significantly better than clinicians at predicting appendicitis.[1,19] However, these studies have indicated that utilization of clinical pathways can standardize care and lower rates of CT.[19-22] Our clinical score may improve on the currently established methods for risk stratifying patients, and in the case of male patients, provide a more nuanced approach to patient care. Given the unclear benefit of advanced imaging in males over age 5 in reducing rates of negative appendectomies,[23] elevated values of our clinical score would have sufficient test performance to provide surgeons confidence in deferring the need for imaging and taking high risk children directly to the operating room.

Our study is subject to several limitations. Most notably, this study was conducted within a network of tertiary care, academic children's hospitals; therefore our findings may not be applicable in other settings. Although we collected data prospectively from multiple different hospital systems, external validation of our results, especially within non-academic settings, will be important prior to implementation. The overall rate of appendicitis in our

cohort was high (50%). As such, the performance of our clinical score will need to be assessed among populations with lower background rates of appendicitis. Finally, the developed score is difficult to calculate without the assistance of a computer or other aid to remind the clinician of the score elements. Given this limitation, we did consider other methods for developing and deploying clinical decision rules, which can be used without the aid of electronic calculation.[24] However, we believe that developments in bioinformatics and other technologies could facilitate the integration of algorithms into the electronic medical record, automatically accessing necessary data elements and generating risk scores. This approach may be most useful as part of larger risk stratification algorithms for care or for community clinicians to determine the urgency of referral.

Conclusions

We have developed a highly accurate scoring system for predicting appendicitis in boys 8-18 years. This rule has the potential to facilitate a more standardized and judicious utilization of diagnostic imaging while maintaining high quality care. Further work to externally validate this rule is necessary before application should be considered.

References

1. Bundy DG, Byerley JS, Liles EA, Perrin EM, Katznelson J, Rice HE. Does this child have appendicitis? *JAMA.* Jul 25 2007;298(4):438-451.
2. Fahimi J, Herring A, Harries A, Gonzales R, Alter H. Computed tomography use among children presenting to emergency departments with abdominal pain. *Pediatrics.* Nov 2012;130(5):e1069-1075.
3. Hryhorczuk AL, Mannix RC, Taylor GA. Pediatric abdominal pain: use of imaging in the emergency department in the United States from 1999 to 2007. *Radiology.* Jun 2012;263(3):778-785.
4. Larson DB, Johnson LW, Schnell BM, Goske MJ, Salisbury SR, Forman HP. Rising use of CT in child visits to the emergency department in the United States, 1995-2008. *Radiology.* Jun 2011;259(3):793-801.
5. Parker MW, Shah SS, Hall M, Fieldston ES, Coley BD, Morse RB. Computed Tomography and Shifts to Alternate Imaging Modalities in Hospitalized Children. *Pediatrics.* Sep 2015;136(3):e573-581.
6. Richards MK, Kotagal M, Goldin AB. Campaigns Against Ionizing Radiation and Changed Practice Patterns for Imaging Use in Pediatric Appendicitis. *JAMA pediatrics.* Aug 2015;169(8):720-721.
7. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* May 1986;15(5):557-564.
8. Kharbanda AB, Taylor GA, Fishman SJ, Bachur RG. A clinical decision rule to identify children at low risk for appendicitis. *Pediatrics.* Sep 2005;116(3):709-716.
9. Samuel M. Pediatric appendicitis score. *J Pediatr Surg.* Jun 2002;37(6):877-881.
10. Schneider C, Kharbanda A, Bachur R. Evaluating appendicitis scoring systems using a prospective pediatric cohort. *Ann Emerg Med.* Jun 2007;49(6):778-784, 784 e771.
11. Bhatt M, Joseph L, Ducharme FM, Dougherty G, McGillivray D. Prospective validation of the pediatric appendicitis score in a Canadian pediatric emergency department. *Acad Emerg Med.* Jul 2009;16(7):591-596.
12. Forcier M. Emergency department evaluation of acute pelvic pain in the adolescent female. *Clinical Pediatric Emergency Medicine.* 2009;10(1):20-30.
13. Bachur RG, Hennelly K, Callahan MJ, Monuteaux MC. Advanced radiologic imaging for pediatric appendicitis, 2005-2009: trends and outcomes. *J Pediatr.* Jun 2012;160(6):1034-1038.
14. Kharbanda AB, Dudley NC, Bajaj L, et al. Validation and refinement of a prediction rule to identify children at low risk for acute appendicitis. *Arch Pediatr Adolesc Med.* Aug 2012;166(8):738-744.
15. Kharbanda AB, Stevenson MD, Macias CG, et al. Interrater reliability of clinical findings in children with possible appendicitis. *Pediatrics.* Apr 2012;129(4):695-700.
16. 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.* Jan 2013;66(1):95-104.

17. Meltzer AC, Baumann BM, Chen EH, Shofer FS, Mills AM. Poor sensitivity of a modified Alvarado score in adults with suspected appendicitis. *Ann Emerg Med.* Aug 2013;62(2):126-131.
18. Ebell MH, Shinholser J. What are the most clinically useful cutoffs for the Alvarado and Pediatric Appendicitis Scores? A systematic review. *Ann Emerg Med.* Oct 2014;64(4):365-372 e362.
19. Fleischman RJ, Devine MK, Yagapen MA, et al. Evaluation of a novel pediatric appendicitis pathway using high- and low-risk scoring systems. *Pediatr Emerg Care.* Oct 2013;29(10):1060-1065.
20. Saucier A, Huang EY, Emeremni CA, Pershad J. Prospective evaluation of a clinical pathway for suspected appendicitis. *Pediatrics.* Jan 2014;133(1):e88-95.
21. Kharbanda AB, Madhok M, Krause E, et al. Implementation of Electronic Clinical Decision Support for Pediatric Appendicitis. *Pediatrics.* May 2016 2016;137(5):e20151745.
22. Santillanes G, Simms S, Gausche-Hill M, et al. Prospective evaluation of a clinical practice guideline for diagnosis of appendicitis in children. *Acad Emerg Med.* Aug 2012;19(8):886-893.
23. Bachur RG, Hennelly K, Callahan MJ, Chen C, Monuteaux MC. Diagnostic imaging and negative appendectomy rates in children: effects of age and gender. *Pediatrics.* May 2012;129(5):877-884.
24. Sullivan LM, Massaro JM, D'Agostino RB, Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Statistics in medicine.* May 30 2004;23(10):1631-1660.

Figure Legend

Figure 1: Flow Diagram of Study Population and Final Diagnosis

Figure 2: Full Distribution of Scores

Table 1: Patient Demographics

Sex	Male 100% (961)
Age (years)	Mean 12.2, SD 2.6
Race	
American Indian/Alaska Native	0.1% (1)
Asian	1.9% (18)
Native Hawaiian or Pacific Islander	0.3% (3)
Black or African American	12.4% (119)
White	65.9% (633)
Other	17.1% (164)
Unknown/not reported	2.4% (23)
Ethnicity	
Not Hispanic or Latino	69.5% (668)
Hispanic or Latino	30.3% (291)
Unknown/not reported	0.2% (2)
Language	
English	82.1% (789)
Spanish	17.2% (165)
Other	0.7% (7)
Appendicitis Rate	49.9% (480)
Perforation Rate	22.3% (107)
Negative Appendectomy Rate	3% (17)

Table 2: Predictors Included in Model and Bivariate Association with Appendicitis

Predictor	N	beta	OR (95% CI)	Sensitivity	Specificity	PPV	NPV
Age (>12 years)	961	<0.00	1.00 (0.96, 1.05)	49	53	51	51
Duration of pain (>24 hours)	961	-0.04	0.96 (0.89, 1.04)	45	56	51	51
History of anorexia	934	0.59	1.80 (1.38, 2.35)	68	46	56	59
History of nausea	943	0.31	1.37 (1.06, 1.76)	56	52	53	54
History of emesis	959	0.71	2.03 (1.57, 2.63)	65	52	58	60
Migration of pain to RLQ	921	0.89	2.43 (1.87, 3.17)	60	62	62	60
Focal pain in the right lower quadrant	950	0.98	2.66 (1.85, 3.82)	90	24	54	69
Tenderness on examination	955						
mild (referent)		---	1.00	---	---	---	---
moderate		1.46	4.30 (3.00, 6.18)	68	43	55	57
severe		2.29	9.91 (6.06, 16.21)	22	92	74	54
Right sided tenderness	956	1.75	5.76 (2.55, 13.03)	99	8	51	84
Maximal tenderness in RLQ	932	1.58	4.87 (3.42, 6.94)	90	36	58	78
Presence of rebound tenderness	936	1.13	3.09 (2.34, 4.09)	49	76	67	60
Presence of guarding	947	0.90	2.46 (1.89, 3.22)	70	51	59	63
Pain with walking, coughing or hopping	906	1.61	5.01 (3.50, 7.19)	90	36	59	78
WBC (> 10 x 10^3/µL)	900	0.27	1.31 (1.26, 1.37)	89	65	72	86
ANC (> 6.75 x 10^3/µL)	889	0.26	1.30 (1.25, 1.35)	91	61	70	87
Patients with complete data	719						

Table 3: Final multivariable model predicting appendicitis (n=719)

Predictor	Adjusted beta	Adjusted OR (95% CI)
Maximal tenderness in RLQ	2.00	7.40 (4.38, 12.49)
Pain with walking, coughing or hopping	1.70	5.50 (3.29, 9.20)
ANC ($10^3/\mu L$)[1]	0.30	1.35 (1.28, 1.41)

[1] A one unit change corresponds to an increase of 1,000 cells per mcL

Final Risk Score Formula

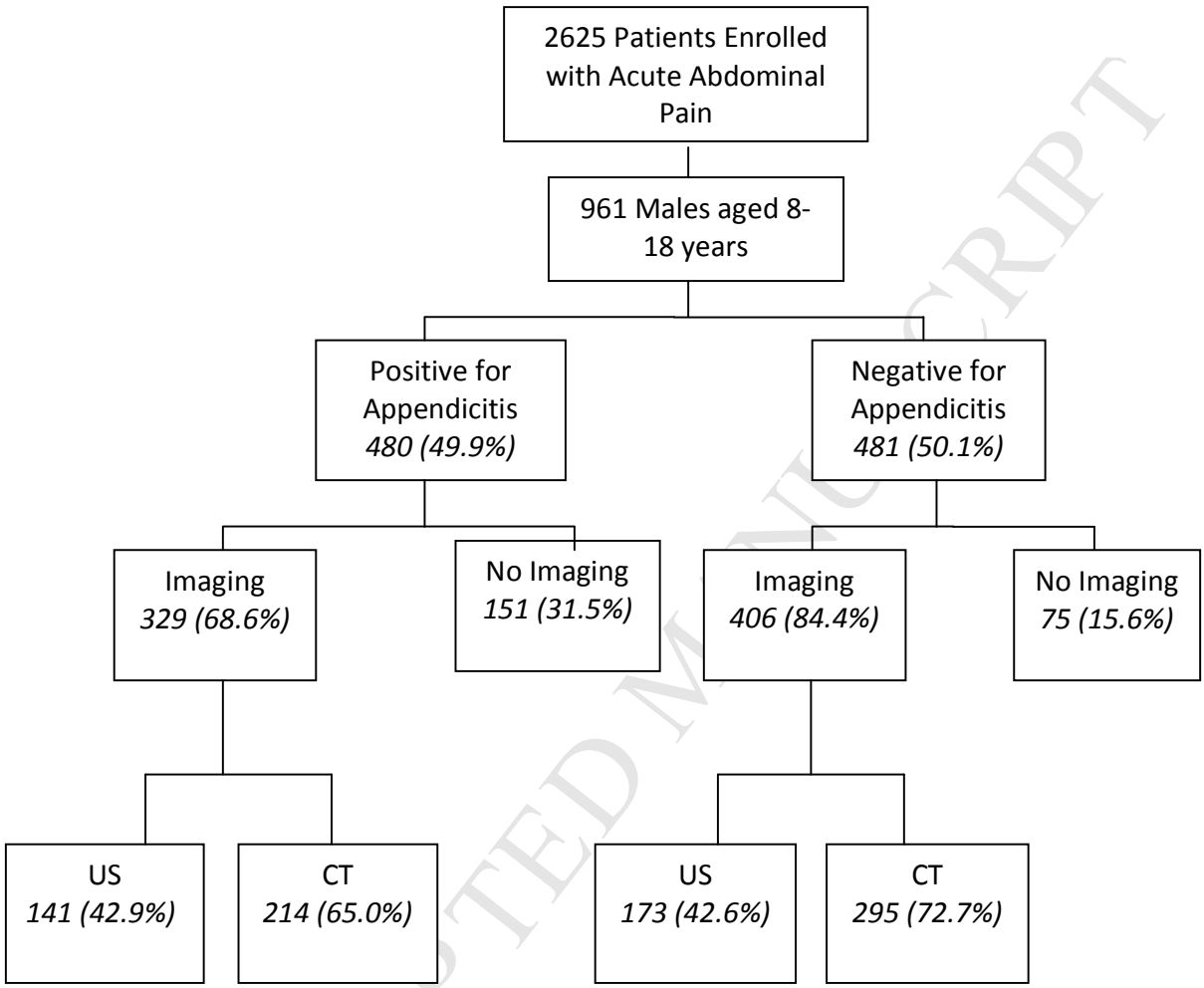
(maximum tenderness[y=1/n=0]* 2.0) + (pain with walking, coughing or hopping[y=1/n=0]*1.70) + ANC [WBC*neutrophil%]*0.30)

Table 4: Test performance of Empirically Derived Appendicitis Clinical Score (n=719)

Risk Score	Sens	Specificity	PPV	+ LR	Patients with a positive screen /patients diagnosed with Appendicitis
≥6.2	69 (64, 74)	87 (83, 91)	85 (80, 89)	5.3	299/254
≥7.2	44 (39, 50)	95 (92, 97)	90 (85, 94)	8.8	180/162
≥8.1	25 (20, 29)	98 (96, 99)	93 (86, 97)	12.0	96/89

Values for sensitivity, specificity and PPV represent *percentages (95% confidence interval)*

Figure 1: Flow Diagram of Study Population and Final Diagnosis



Proportions exceed 100% as patients may have had more than one study

Figure 2: Distribution of Scores