Received Date : 13-Nov-2016
Revised Date : 17-Jan-2017
Accepted Date : 06-Feb-2017
Article type : Evidence-based Diagnostics

Title of Manuscript: Diagnostic Accuracy of History, Physical Exam, Laboratory Tests and Point-of-Care-Ultrasound for Pediatric Acute Appendicitis in the Emergency Department: A Systematic Review and Meta-Analysis

Authors:

-Roshanak Benabbas, MD

1. SUNY Downstate Medical Center
   Department of Emergency Medicine
   450 Clarkson Avenue, Brooklyn, NY 11203

2. Kings County Hospital Center
   Department of Emergency Medicine
   451 Clarkson Avenue, Brooklyn, NY 11203

Roshanak.Benabbas@downstate.edu

-Mark Hanna, MD

1. SUNY Downstate Medical Center
   Department of Pediatrics
   450 Clarkson Avenue, Brooklyn, NY 11203

2. Kings County Hospital Center
   Department of Pediatrics
   451 Clarkson Avenue, Brooklyn, NY 11203

abdou1one@gmail.com

-Jay Shah

1. Kings County Hospital Center
   Department of Emergency Medicine
   451 Clarkson Avenue, Brooklyn, NY 11203

jay4shah@gmail.com

-Richard Sinert, DO

1. SUNY Downstate Medical Center
   Department of Emergency Medicine
   450 Clarkson Avenue, Brooklyn, NY 11203

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/acem.13181-16-858

2. Kings County Hospital Center
Department of Emergency Medicine
451 Clarkson Avenue, Brooklyn, NY 11203
nephron1@gmail.com

Corresponding Author: Richard Sinert, Do

Running Title: Diagnosing Acute Appendicitis in Children in the ED

Keywords: Pediatrics, Appendicitis, Ultrasonography, Computed Tomography, Magnetic Resonance Imaging

Prior Presentations: None
Funding Sources/Disclosures: None

Acknowledgments: The authors would like to thank Christopher Stewart, senior assistant librarian at SUNY Downstate Medical Center, for his help in conducting the literature searches.

ABSTRACT

Background: Acute appendicitis (AA) is the most common surgical emergency in children. Accurate and timely diagnosis is crucial but challenging due to atypical presentations and the inherent difficulty of obtaining a reliable history and physical examination in younger children.

Objectives: To determine the utility of history, physical exam, laboratory tests, Pediatric Appendicitis Score (PAS) and Emergency Department-Point-of-Care Ultrasound (ED-POCUS) in the diagnosis of AA in ED pediatric patients. We performed a systematic review and meta-analysis and used a test-treatment threshold model to identify diagnostic findings that could rule in/out AA and obviate the need for further imaging studies specifically, CT scan, MRI and Radiology Department Ultrasound (RUS).

Methods: We searched PUBMED, EMBASE, and SCOPUS up to October 2016 for studies on ED pediatric patients with abdominal pain. Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS-2) was used to evaluate the quality and applicability of included studies. Positive and negative Likelihood Ratios (LR+ and LR-) for diagnostic modalities were calculated and when appropriate data

was pooled using Meta-DiSc. Based on the available literature on the test characteristics of different imaging modalities and applying Pauker-Kassirer method we developed a test-treatment threshold model.

Results: Twenty-one studies were included encompassing 8,605 patients with AA prevalence of 39.2%. Studies had variable quality using the QUADAS-2 tool with most studies at high risk of partial verification bias. We divided studies based on their inclusion criteria into two groups of "undifferentiated abdominal pain" and abdominal pain "suspected of AA". In patients with "undifferentiated abdominal pain" history of "pain migration to RLQ" (LR+ 4.81, 95% CI 4.81-6.44) and presence of "cough/hop pain" in the physical exam (LR+ 7.64, 95% CI 5.94-9.83) were most strongly associated with AA. In patients "suspected of AA" none of the history or laboratory findings were strongly associated with AA. Rovsing's sign was the physical exam finding most strongly associated with AA (LR+ 3.52, 95% CI 2.65-4.68). Among different PAS cutoff points PAS≥ 9 (LR+ 5.26, 95% CI 3.34-8.29) was most associated with AA. None of the history, physical exam, lab tests findings or PAS alone could rule in or rule out AA in patients with "undifferentiated abdominal pain" or those "suspected of AA". Emergency Department Point-of-Care Ultrasound (ED-POCUS) had LR+ 9.24 (95% CI 6.24-13.28) and LR- 0.17 (95% CI 0.09-0.30). Using our test-treatment threshold model, positive ED-POCUS could rule in AA without the use of CT and MRI, but negative ED-POCUS could not rule out AA.

Conclusion: Presence of AA is more likely in patients with undifferentiated abdominal pain migrating to the RLQ or when cough/hop pain is present in the physical exam. Once AA is suspected, no single history, physical exam, lab finding or score attained on PAS can eliminate the need for imaging studies. Test characteristics of ED-POCUS are similar to those reported for RUS in literature for diagnosis of AA. In ED patients suspected of AA, a positive ED-POCUS is diagnostic and obviates the need for CT or MRI while negative ED-POCUS is not enough to rule out AA.

INTRODUCTION

Abdominal pain is one of the most common chief complaints among Emergency Department (ED) pediatric patients with over 1,000,000 annual visits in patients younger than 15 years old.[1]Although many cases are benign, it is crucial to correctly and timely identify those requiring further workup, imaging studies, or surgical intervention. Acute Appendicitis (AA) is the most common surgical emergency in children with 72,000 hospital discharges per year.[2] Diagnosing AA in children remains challenging due to atypical presentations[3-5] and difficulty of obtaining a reliable history and physical exam, especially in younger children.[6,7] Prompt diagnosis of acute appendicitis can prevent complications such as perforation and abscess formation.[8-11]

In everyday clinical practice, physicians combine their clinical suspicion of AA with laboratory tests findings and imaging studies to make a final diagnosis. Acute Appendicitis scoring systems such as Pediatric Appendicitis Score (PAS)[12] and Alvarado Score[13] use elements of history, physical exam, and lab test findings to identify patients with a high risk of having AA. However, the reported sensitivity and specificity of these scoring systems vary widely between studies[14-17]and neither scoring system (PAS and Alvarado[13]) integrate imaging studies despite the increasing use of CT scan, in the ED.[18-20]

Given the concern of exposing children to ionizing radiation by using CT scan, American College of Emergency Physician[21] and American College of Radiology[22] recommend considering Ultrasonography (US) as the initial radiologic modality for pediatric AA.

In an attempt to integrate imaging and clinical and lab findings, Bachur et al[23] calculated PAS in a cohort of ED pediatric patients suspected of AA who had undergone US study. Bachur et al[23] suggested in their conclusion that patients with high-risk PAS (PAS 7-10) but negative US, or low-risk PAS (PAS 0-3) but positive US benefit from serial examination or further workup. In patients with medium-risk PAS (PAS=

4-6), Bachur et al[23] suggested appendectomy in those with positive US and observation in those with negative US. However, this single-center study used Radiology Department US (RUS) as opposed to ED Point-Of-Care Ultrasound (ED-POCUS)

Not all EDs have a radiology department sonographer available 24/7 which can delay the diagnosis of AA and increase the risk of complications. Emergency Department Point-Of-Care Ultrasound of appendix can provide valuable, real-time information to the treating physician while decreasing ED length of stay.[24,25]

We decided to use a systematic review and meta-analysis methodology to evaluate which element(s) of history, physical exam, lab tests, PAS, or ED-POCUS are most useful in the diagnosis of AA in ED pediatric patients. Specifically, we were interested in investigating if any of these findings could obviate the need for radiology department resources (CT, MRI, or RUS) and therefore, expedite patient disposition.

Prior systematic reviews of pediatric appendicitis by Bundy et al[6] and Dahabreh et al[26], both included studies that were heterogeneous in design and study population. However, in order to evaluate the test-characteristics of index tests in ED patients, it is more useful to examine studies limited to the ED population. Therefore, we limited our population to pediatric patients presenting to the ED with abdominal pain. We also limited our assessment of US to ED-POCUS, performed and interpreted by ED physicians.

METHODS

Study Design

We conducted a systematic review and meta-analysis of studies on the diagnosis of AA in ED pediatric patients with abdominal pain. The design of this systematic review and Meta-Analysis follows the

recommendations of the Preferred Reporting Items for Systematic Review and Meta-analyses (PRISMA)[27] guideline and Meta-analysis of Observational Studies in Epidemiology (MOOSE) statement.[28]

Search Strategy

In conjunction with a medical librarian, we searched the medical literature in PUBMED, EMBASE and SCOPUS from their inception up to October 2016 for search terms diagnosis and appendicitis. (See Appendix-1 for search strategies and MeSH terms used.) The PUBMED, EMBASE, and SCOPUS searches were combined and limited to human subjects and English language for three separate search topics: History and physical exam, Laboratory tests and ED-POCUS. Narrative reviews, case control studies, and case reports were excluded.

Study Selection and Data Abstraction

Two authors (RB, MH) independently selected articles for each index test category (History and physical exam, Laboratory tests, and ED-POCUS) from the combined PubMed, EMBASE, and SCOPUS search for the full-text review. Each reviewer independently selected potentially eligible studies. Studies eligible for inclusion were those that described patients with the maximum age of 21 years presenting to the ED with either "undifferentiated abdominal pain" or abdominal pain "suspected of AA". In studies with both adult and pediatric participants, we included only those that either presented the data from their pediatric participants separately or could provide us with that data upon contacting the author. Studies on ED-POCUS were included only if performed and interpreted by an ED physician. Studies were included only if provided sufficient data to construct 2 by 2 tables either in the text or after contacting the author. Included studies were those which described positive and negative index test along with the final diagnosis using a gold standard, histopathologic diagnosis of AA. Among trials on clinical scores, we decided to include only those dedicated to PAS, as the most broadly studied clinical score for the

diagnosis of AA in children. We decided not to review Alvarado score as it was originally developed for identifying adult patient at high risk of AA.

A meeting was held and any disagreement in study selection were resolved by consensus before a final list of included studies was made. Reference list of included studies was reviewed to look for additional studies that could be included.

Data Analysis

Sensitivities, specificities, and likelihood ratios (LR) were calculated based on construction of two by two tables for findings of each included study. When more than one study reported a variable, we pooled the data using Meta-DiSc software with random-effects model.[29] Inter-study heterogeneity was assessed using the DerSimonian-Laird random effect model. We pooled data only when I-square was less than 50% and reported point estimates for variables demonstrating high heterogeneity.

Quality Assessment

Two authors (RB, MH) independently assessed the quality and applicability of each included study using the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS-2)[30]. Agreement between the two reviewers was assessed by calculating kappa using SPSS (IBM® SPSS®, Version 21, IBM Corp., Armonk, NY)

Four domains were assessed for biases: 1) Patient selection 2) Index test 3) Reference test 4) Flow and timing. Several considerations were established prior to assessing the quality of individual studies, and a set of signaling questions were developed for each section of the QUADAS-2.

The ideal study population would be patients who presented to the ED with abdominal pain and subsequently received both index test and reference test with the interpreter of each test blinded to the

results of the other. An exclusion based on previous history of abdominal surgeries other than appendectomy, abdominal trauma, or previous work up such as surgical consultation, imaging studies, or lab tests were judged to be inappropriate.

If the execution of the index test was not clearly defined or the index test was performed after knowing the result of the reference test, that portion of QUADAS-2 would be at high risk for bias. For lab tests and ED-POCUS the criteria for a positive test needed to be clearly specified for that study to be at low risk of bias. For physical exam, the decision to qualify a finding as positive or negative was left to physician discretion. We assessed the index-reference test interval as appropriate if patients received the index test upon presentation to the ED and were sent to the operating room in a timely manner if indicated. If performance of the reference test was not clearly defined or the interpreter of the reference test was not blinded of the result of the index test, then that portion of QUADAS-2 would be at risk of bias. Concerns regarding the applicability of the results of index or reference test were raised if these tests were conducted in a manner that differed from routine clinical practice.

Test-Treatment Threshold Estimates

We used the Pauker-Kassirer method [31] to assess the testing and treatment thresholds for CT scan and MRI and investigated which element(s) of History and Physical exam, lab tests, PAS, or ED-POCUS could have sufficient discriminatory power to eliminate the need for CT Scan, MRI or Radiology Department US (RUS) in the diagnosis of AA and facilitate patient disposition.

We used the accuracy and risk associated with each diagnostic modality as well as the risk and benefit of treatment to estimate thresholds for testing and treatment for each imaging modality (CT, MRI, and RUS). Then, using the operating characteristics of each index test (history, physical exam, lab finding, PAS, and ED-POCUS) we estimated the post-test probability of AA applying the Bayes theorem. Post-test probabilities of AA were then compared to the testing and treatment thresholds of CT, MRI, and RUS

to investigate if presence or absence of each index test can rule in or rule out AA without the use of radiology department resources.

RESULTS

The PUBMED, EMBASE and SCOPUS searches identified 846 citations for History and Physical (H&P), 485 citations for Laboratory studies, and 1,586 citations for ED-POCUS. Upon review of the bibliography of the reviewed articles, 4 more citations were found.

We decided to remove all the retrospective studies, [32-68] from our review. All of these studies had issues related to reliability of their retrospectively abstracted data. All of the retrospective studies failed to document methods suggested by Gilbert et al[69] to improve their accuracy and minimize inconsistencies in data acquisition. If an article did not provide adequate data to reproduce 2 by 2 tables it was excluded from the review [70-72] We also excluded any article that did not mention "Emergency Department", "ED" or "Emergency Room" as their setting [12,73-78], or trials that included pre or post appendectomy [79-86], admitted [87-89] or referral patients. [90-93] We found four articles with possibility of having overlapping patient population: Schneider et al[94], Becker et al[95] Colvin et al[49], and Kharbanda et al[96]. We decided to include only Schneider et al[94] and exclude the other three. Our decision was based on the fact that Schneider et al [94], unlike other three, provided data on PAS. Trials on the accuracy of clinical scores, other than PAS, were only included if they reported the test-characteristics of all of their variables. Articles that only provided the final score with limited or no data on the variables were excluded.[97-101]

Twenty-one studies[14-16, 23-25, 94, 102-115] were included in our review. Table-1.1 and Table-1.2 describe the included studies.

Prevalence

The combined population from the 21 studies included in this review was 8,605 patients of which 3,344 were diagnosed with AA (38.86%, 95% CI 37.80-39.89). Prevalence of AA ranged from 9.8% [110] to 63%[109] with weighted prevalence of 39.23 % (95% CI 38.20-40.27). The inclusion criteria were not uniform across studies. Therefore, we decided to group studies based on their inclusion criteria.

The weighted prevalence of AA was significantly ($p<0.001$) higher in studies that used either "abdominal pain suspected of AA" or "RLQ pain" as their inclusion criteria (19 studies [15,16,23-25, 94,102-109, 111-115], N=7,510 , AA prevalence= 42.8% ,95%CI 41.68-43.92) compared to those that used "abdominal pain" or "undifferentiated abdominal pain" (2 studies[14, 110] , N=1,095, AA prevalence=13.4%, 95%CI 11.53-15.58).

History and Physical Exam

Fifteen studies provided data on history and physical exam findings (Table 1.1). All except for two[105,115] were single-center studies. Inclusion criteria were not uniform across reviewed studies. Goldman et al[14] and O'shea et al[110] used "undifferentiated abdominal pain" as inclusion criteria while 11 trials included patients with "suspected of AA" and two (Wu et al[109] & Khan et al[113]) used "RLQ pain". Sample size varied considerably between studies ranging from 40 (Doniger[111]) to 2,133 (Bachur[115]). The mean age of study participants ranged from means of 9.4 to 12 years. While some studies excluded patients with history of abdominal surgery at any time [23,94,104,114,115] or in the previous year[102], others [14,15,105,106] only excluded patients with history of appendectomy. All included studies defined AA as positive histopathologic findings in patients who underwent appendectomy and follow-up by phone call in non-surgical patients. The follow-up length ranges from 5 days (Cayrol[106]) to 6 months (Sivitz[25]). Zuniga et

al[15] used a database search for non-surgical patients and in Huckins et al [105] discharge diagnosis was considered as the final diagnosis.

We decided to report the pooled data only when I-square was less than 50% [116] and only report point estimates for variables that showed moderate to high heterogeneity. In Table 2.1, Table 2.2 and Table 2.3 I-square is mentioned when pooling of the data was not possible due to high heterogeneity (I^2>50%).

Across studies on patients "suspected of AA" (Table 2.1), none of the history findings had strong test-characteristics with "Migration of pain" demonstrating the highest LR+ (1.75, 95% CI 1.58-1.94) and "Anorexia" having the lowest LR- (0.58 , 95% CI 0.52-0.65). In patients with "undifferentiated abdominal pain" (Table 2.2) and based on the results from a single study (O'shea et al[110]) fever was most suggestive of AA, LR+ 3.4 (95% CI 2.42-4.76) while absence of fever decreased the probability of AA the most, LR- 0.32 (95% CI 0.16-0.64)

Among studies on patients "suspected of AA" (Table 2.3), Rovsing's sign was the physical exam finding most suggestive of AA (LR+ 3.52, 95% CI 2.65-4.68) while it's absence had minimal effect on probability of AA (LR- 0.72, 95% CI 0.66-0.78). Presence or absence of fever in physical exam (as measured in the ED), did not significantly alter the probability of AA (LR+ 1.13 and LR-0.94). Grouping studies based on cutoff point used to define fever did not improve the test-characteristics of this variable (Table 2.4). Only one study (Goldman et al [14]) investigated physical exam findings in patients with "undifferentiated abdominal pain" (Table 2.2). In Goldman et al[14] study "Cough/Hop pain" (LR+ 7.64, 95% CI 5.94-9.83) and "Right Iliac Fossa tenderness" (LR+ 4.74, 95% CI 3.94-5.7) were most suggestive of AA.

Laboratory Tests

Eleven studies provided data on laboratory tests: Ten studies[15,23,94,102,104-106,112,113,115] were done on patients "suspected of AA" and one study[14]on patients with "undifferentiated abdominal pain" (Table 1.1). All except for two[105,115], were single-center studies. White blood count (WBC) was the most commonly reported lab test (10 studies) followed by Absolute Neutrophil Count (ANC) or percentage of Neutrophils (Neut) (7 studies). The cutoff point for WBC was 10,000 cells/mm^3 in nine studies and 12,000 cells/ mm^3 in one study (Kwan[102]). The cut off point for ANC varied among studies: ≥ 7,500 cells/ mm^3 (Zuniga[15], Goldman[14]) and ≥ 6,750 cells/ mm^3 (Sivitz[25]). One study (Bachur et al[115]) investigated various cutoff points of WBC and ANC in three different age groups (less than 5, 5-12 and older than 12 years old). Several studies reported percentage of neutrophils. Bachur[23], Mandeville[104] and Schneider[94]reported Neutrophil≥ 75% and Santillanes[112] reported Neutrophil≥67%. Other variables studied were D-dimer, CRP, Urine ketones, and Procalcitonin.

Across studies on patients "suspected of AA", most laboratory tests had poor test-characteristics (Table 2.5). Of laboratory tests reported in more than one article, CRP>3 was most suggestive of AA (LR+ 2.10, 95% CI 1.61-2.76) while WBC<10,000 was most associated with absence of AA (LR- 0.21, 95% CI 0.19-0.25). A combination of WBC≥12,000 and CRP>3 had the highest LR+ (4.36, 95% CI 2.26-8.42) however, this combination was only described in Kwan et al[102] study (Table 2.5).

One study, Bachur et al [115], on patients with "suspected AA" investigating the effect of age on WBC and ANC in AA patients and found these variables to have a better diagnostic performance in older children. Across reported cutoff points of WBC count, (≥5,000 to ≥ 15,000) LR+ was 1.05-1.91 in children younger than 5 vs. LR+ 1.05-5.25 in children older than 12. Similarly, across different cutoff points of ANC (≥5,000 to ≥ 15,000), LR+ was 1.25-1.87 in children younger than 5 vs 1.7-5.85 in children older than 12.

Only one study (Goldman et al[14]) reported lab tests in patients with 'undifferentiated abdominal pain". (Table 2.2) In Goldman et al[14] study, ANC ≥7,500 was most associated with diagnosis of AA (LR+ 2.33, 95% CI 1.89-2.88) while absence of leukocytosis was most associated with absence of AA (LR- 0.22, 95% CI 0.13-0.36).

Pediatric Appendicitis Score (PAS)

Seven studies evaluated PAS at different cutoff points. Five studies included patients suspected of AA, Wu et al,[109] only included patients with RLQ pain and Goldman et al[14]included patients with "undifferentiated abdominal pain". Sample size ranged from 99[16] to 1,395[109] and the mean age ranged from 9.8 to 11.9 years.

Two studies, Mandeville et al[104] and Goldman et al[14], reported data on every possible cutoff point for PAS (0-10) while three studies (Schneider [94], Khanafer[114] and Escriba[16]) described data on PAS cutoff points 1 to 10. Bachur et al[23] categorized the results in the following groups: PAS<4, PAS=4-6 and PAS≥7 and. Wu[109] described PAS≥ 7 cutoff point on the day 1 to 3 of presentation. For the purpose of this review we only included data collected on Day 1 (at presentation).

We decided to exclude one study, Escriba et al[16], from final analysis. Escriba et al[16]used "more than" (>) as the definition of cutoff point. It's unclear what the authors mean by PAS>10 as PAS=10 is the maximum score a patient can get on PAS. We unsuccessfully tried to contact the author to clarify this point. In the article by Schneider et al[94], "1-Specificity" was incorrectly reported in place of "Specificity" . After contacting one of the authors and confirming that the data presented was in fact an error in print we recalculated the reported specificities and used the data.

Data from six studies were included in the meta-analysis, five on patients "suspected of AA" and Goldman et al[14]on patients with "undifferentiated abdominal pain". Across studies on patients "suspected of AA (Table 2.6), the highest LR+ was for PAS=10 (LR+ 5.80), PAS≥9 (LR+ 5.26), and PAS≥8 (LR+ 4.40) making only highest scores (PAS= 8, 9 and 10) good predictors of AA. The results were very heterogeneous (I^2=77.2%-85.6%) for the lowest cutoff points of PAS (PAS 1,2 and 3) and therefore pooled data could not be calculated. In one study on patients with "undifferentiated abdominal pain" (Table 2.7), moderate PAS cutoff points had the highest LR+ (PAS≥5 with LR+ 4.56, and PAS≥6 with LR+ 4.07). In Goldman et al[14], PAS≥0 (LR- 0.02) and PAS≥1 (LR- 0.24) had the lowest LR- and therefore PAS=0 and PAS<1 were most suggestive of absence of AA.

Emergency Department Point-of-Care Ultrasonography (ED-POCUS)

Five studies[24,25,107,108,110] met our inclusion criteria. All five studies included patients "suspected of AA". (Table 1.2) While four studies were done exclusively on pediatric patients, Fox et al[107] included both adult and pediatric patients; however, they presented data for their pediatric population (n=42) separately.

Sivitz et al[25]reported number of positive and negative ED-POCUS studies, which was slightly higher than the sample size suggesting that some patients received more than one ED-POCUS study. Since the number of scans was close to the sample size, (264 ED-POCUS studies in 231 patients) we decided to include this study in our review. In one study (Kim et al[108]) both EM residents and attending physicians (either on-site or via tele-Ultrasonography) performed ED-POCUS on the same group of participants. Given that resident-performed POCUS, tele-Ultrasonography, and attending-performed POCUS were performed on the same study population and to avoid any overlapping data, we decided to include only the ED-POCUS scans done by residents.

Sample size varied widely between studies ranging from 40 (Doniger et al[111]) to 264 (Sivitz et al [25]). Four studies[24,25, 107, 111]excluded patients who had unstable vital signs. Sivitiz et al [25]excluded patients with history of abdominal surgery while Fox[107]and Doniger[111]excluded pregnant patients. Doniger[111]also excluded those with recent abdominal imaging and Kim et al [108]did not specify any exclusion criteria. The duration of follow-up in patients who were managed non-surgically varied from 2 weeks[111]to 6 months.[25] One study[108] did not specify the duration of follow up. In four studies [24,25,107,111], the treating and the enrolling physician could be the same. Kim et al[108]did not provide details about their treating physicians. In all studies ED residents and attending physicians or Pediatric Emergency Medicine (PEM) fellows and attending physicians obtained and interpreted the ED-POCUS. Fox et al[107] limited the duration of ED-POCUS to 5 minutes, whereas the rest of the studies did not use such limitation.

Elikashvili et al [24] reported significantly higher percentage of equivocal results compared to all other studies in this group. High number of equivocal results in Elikashvili et al[24] can be attributed to the fact that full visualization of a normal appendix was mandatory to consider an ED-POCUS scan negative whereas other studies in this group did not mandate a full visualization. Due to high prevalence of equivocal results in Elikashvili et al [24] , we decided to exclude this study from our final analysis. Across the remaining four studies [25,107,108,110], ED-POCUS had a sensitivity of 86% (95% CI 79%-90%), specificity of 91% (95% CI 87%-94%), LR+ 9.24 (95% CI 6.42-13.28), and LR- 0.17 (95% CI 0.09-0.30) making positive ED-POCUS a good predictor of AA while negative ED-POCUS considerably decreases the probability of AA (Table 2.8, Figure 3.1). A sensitivity analysis adding Elikashvili et al [24]results did not change LR+ (9.56 vs 9.24) drastically while significantly increased the heterogeneity.

QUADAS-2 Analysis of Included Studies

Initial inter-rater reliability among the two QUADAS reviewers was substantial (kappa 0.75, 95% CI 0.61-0.82). A meeting was held between two reviewers (RB, MH) and the third author (RS) to resolve any disagreements by consensus and all authors agreed 100% on the final QUADAS-2 scoring. (Figure 2)

Patient Selection: We found several studies at risk of bias due to inappropriate exclusion: Several studies[23, 25,94,102,104] excluded patients with history of any abdominal surgery (and not exclusively appendectomy), abdominal trauma. [105,110]or abdominal imaging [94,104].Inappropriate exclusion reduces the generalizability of the results. Certain exclusions can result in missing milder cases of AA. For instance, exclusion of all patients without surgical consultation, imaging studies, or lab tests[15, 23, 103,114,115]. Several studies [15,16,23,94] excluded all patients in whom missing data prevented investigators from calculating PAS which can introduce significant bias. It is unclear how this bias skews the results; more severe cases of AA might have received less workup and were sent straight to surgery and therefore miss data for calculating PAS. On the other hand, patients least suspected of having AA might have been discharged without any lab tests and therefore miss data for calculating PAS.

Index Test: In all studies, the interpreter of the index test (H&P, Labs, PAS and ED-POCUS) was blinded to the reference test (Operative report and histopathology). In all of the studies on lab tests or ED-POCUS, criteria for a positive test were pre-specified. For physical exam, whether to qualify a finding as positive or negative was left to physician discretion.

Reference Test: Surgery and histopathologic examination of removed tissue was used as the reference test in all included studies. With the exception of three studies, (Kentsis [103], Khanafer[114], Fox [107]) all trials failed to specify if the interpreter of the reference test (pathologist) was blinded to the results of the index

test; this can introduce incorporation bias. Incorporation bias is likely when the result of the index test can determine whether the reference test classifies patients as disease-positive or disease-negative. In several studies the pathologist was solely blinded to the main index test under investigation.For instance, in Schneider [94], Goldman[14] and Bachur[23] the interpreter of the reference test (pathologist) was blinded to the final calculated PAS but not necessarily to the components of PAS (History, Physical exam, and Labs) Similarly, in studies by Khan[113] Cayrol[106] and Huckins [105]pathologist was blinded to the main lab test studied but not to the history, physical exam findings, or other lab tests and in Sivitz et al [25]pathologist was only blinded to ED-POCUS.

Flow and Timing: Studies that examine the accuracy of history, physical exam, laboratory tests, and PAS in patients suspected of AA, are at a high risk of partial verification bias. Partial verification bias, as described by Kohn et al [117] occurs when the result of the index test determines who receives the reference test. In studies that included only patients suspected of AA, the index tests (history, physical exam, and labs) are already used to decide who enters the study and later receives the reference standard. Although PAS was not used independently as an inclusion criterion, the risk of partial verification bias is still high given that PAS is calculated using a combination of history, physical exam, and lab findings. Partial verification bias could inflate estimates of sensitivity while underestimating the specificity of index tests.

We found all studies at high risk of differential verification bias. Differential verification bias, also called double gold standard bias, can occur when patients with a positive index test are more likely to receive an immediate reference test whereas those with negative index test receive only clinical follow-up[117] . All trials used follow-up as an alternate for histopathology in non-surgical patients. Although no case of AA was reported in the follow-up group, cases of self-resolving AA are reported in the literature. [118,119] Furthermore, most included studies did not specify if they discharged any patient on antibiotics, which can result in resolution of milder cases of AA[120]. Differential verification bias can falsely increase both sensitivity and specificity of the index test. The risk of bias is higher in Huckins[105] et al study due to lack

of follow up or in those studies[16,25,94,102,104,107-109, 112, 114] that lost patients to the standard follow-up. Although some of these studies tried alternative follow up such as searching electronic records [102,114] or contacting patient's pediatrician,[94] the risk of missing AA still exists. Loss to follow-up even in small numbers can introduce significant bias. [121,122]

In four studies[24,25,107,111]on ED-POCUS, the treating and the enrolling physician were the same in all or some patients which introduces differential verification bias since the results of ED-POCUS could influence further testing and determine who receives the gold standard. Sivitz et al[25] reduced this risk by blinding the treating physician to the results of ED-POCUS when possible. Doniger et al[111]and Fox et al[107] decreased the risk of differential verification bias by ensuring that the treating physician made any decision regarding diagnosis and treatment approach before performing the ED-POCUS. Kim et al[108]did not specify whether their enrolling and treating physician were the same.

Several studies were at risk of bias due to exclusion of subgroups of their enrolled patients from final analysis: Khanafer[114], Bachur[23], Schneider[94], Wu[109], Zuniga[15], Khan[113]and Escriba[16] excluded patients due to incomplete data or missing data. Fox[107], Mandeville[104], Santillanes[112], Wu[109], Escriba[16]and Kim[108] excluded patients who were lost to follow up from the final analysis, and Kentsis et al[103]excluded patients with perforated AA who underwent interval appendectomy.

Test-Treatment Threshold Estimates

We used the Pauker and Kassirer method [31] to estimate thresholds for testing or treatment when caring for a pediatric patient with abdominal pain in the ED. Operative characteristics for ED-POCUS (Table 2.8 sensitivity 86%, specificity 91%) were very similar to those of RUS (sensitivity 88%, specificity 94%) reported in the literature [123] and therefore we decided to remove RUS from our test-treatment threshold model. This model utilizes the unique operating characteristics of each diagnostic modality (CT scan and

MRI) while controlling for the risk of treatment of patients without AA (Rrx), the risk of each diagnostic modality (Rt), and the benefit of treatment of AA (Brx). Variables are presented as probabilities.

In Figure 4, we created two test treatment threshold models (1. CT scan 2. MRI) to diagnose AA. The top half of Figure 4 describes the variables and calculations used to produce the test and treatment thresholds of CT and MRI illustrated in the graphic below. We used the operating characteristics of CT scan and MRI documented in recent systematic reviews. In a review of 26 articles with a total population of 9,356 pediatric patients, Doria et al [123]reported 94% sensitivity and 95% specificity for CT scan. Moore et al[124] reviewed 11 studies, encompassing 1,698 pediatric patients and found MRI to be 96.5% sensitive and 96.1% specific in diagnosis of AA.

We defined risk of treatment (Rrx) as the risk of mortality and morbidity following appendectomy. Aziz et al [125] reviewed 23 studies with a total population of 6,477 and found the risk of complications, including wound infection, to be 1.5 %-4.9% depending on the technique used (Open and laparoscopic appendectomy respectively). In a review of 9 studies with a total population of 65,995 Healy at al[126] reported the overall risk of intraabdominal collection formation, wound infection, and readmission to be 2.4%, 1.8% and 1.5 % respectively. Based on the literature we estimated Rrx to be 0.05

We judged the risk of the diagnostic test (Rt) as 0.00026 for CT scan and zero for MRI. The risk of CT scan is based on the lifetime risk of radiation-induced cancer following a single abdominal CT scan in a 5-year-old. (20/100,000 in males and 26/100,000 in females) [127]

Finally, the benefit of treatment (Brx) of patients with AA has never been, nor ever will be tested by a randomized double-blinded placebo-controlled methodology; it would be unethical to study the spontaneous recovery rate of AA without antibiotics or surgery. Without available evidence we used a conservative estimate for the benefit of treating AA (Brx = 0.90).

In the lower half of the Figure-4, the test thresholds are depicted as the left-most open arrow for each diagnostic modality (CT scan 0.3 % and MRI 0.2%). The treatment thresholds are represented by the right-most open arrow for each diagnostic modality (CT scan 46.5% and MRI 60.4%).

The vertical dashed lines represent the post-test probabilities of AA in the presence of a negative or a positive ED-POCUS. Applying Bayes theorem and using the AA prevalence in patients suspected of AA (42.8%) as the pre-test probability, a negative ED-POCUS (LR- 0.17) would result in a decrease in the post-test probability from 42.8% to 11% represented by the left most vertical dashed-line. In case of a positive ED-POCUS (LR+ 9.24), the post-test probability of AA would increase from 42.8% to 87% represented by the right most vertical dashed-line.

As seen in Figure-4, the far right vertical dashed-line (the probability of AA in presence of a positive ED-POCOS, 87%), falls to the right of the treatment threshold for both diagnostic modalities. (46.5% for CT scan and 60.4% for MRI) Therefore, in a patient "suspected of AA", a positive ED-POCUS could obviate the need to perform CT scan or MRI before treatment can be initiated. The far left vertical dashed-line (the probability of AA after a negative ED-POCOS, 11%), falls to the right of the test-threshold for both CT scan (0.3%) and MRI (0.2%) and therefore, negative ED-POCUS is not sufficient to rule out AA without the need for CT scan or MRI.

Based on this model, in patients with "undifferentiated abdominal pain" (pretest probability of 13.4%) a test needs to have LR+ > 5.8 and LR+>11 to establish the diagnosis of AA without the need for CT and MRI, respectively. In the same population, a test with LR- < 0.03 can rule out AA obviating the need for CT and MRI (Post-test probability = 0%).

Across history and physical exam findings, only "Cough/Hop pain showed high enough LR+ 7.64 to obviate the need for CT scan (but not MRI) in patients with "undifferentiated AA". However, this finding is from a single study and therefore the results may not be generalizable. No history, physical exam, lab, or PAS cutoff point had a low enough LR- to exclude AA without use of CT or MRI. In patients "suspected of AA", it's inaccurate to estimate the post-test probability of AA based on test-characteristics of history, physical exam, lab tests and PAS due to high risk of bias mentioned earlier.

The test-treatment model presented here is an interactive tool and some variables can be modified in the Microsoft Excel calculator published online (Appendix-2). For instance, we judged the risk of treatment (Rrx) to be 5%. Assigning a greater risk to appendectomy, based on the clinical judgment, will result in an increased treatment threshold and therefore additional testing may be needed in such patients.

DISCUSSION

Our systematic review examined the utility of History, Physical exam, Lab results, PAS and ED-POCUS for the diagnosis of AA in ED pediatric patients. Twenty-one studies met the inclusion criteria with most studies dedicated to evaluating history and physical exam, lab tests, and PAS at high risk of bias. We found no single history, physical exam, lab test finding or PAS cutoff point to be sufficiently robust enough to rule out AA and eliminate the need for using CT scan or MRI. This is not to imply that history, physical exam or lab tests are not valuable in the diagnosis of AA as the presence of these findings is necessary to suspect AA in a patient presenting to the ED with abdominal pain.

Compared to the two previously published systematic reviews on pediatric AA by Bundy et al[7] and Dehabreh et al [26] ,we used more rigorous inclusion/exclusion criteria and included only prospective studies dedicated to ED patients. Although reviews such as Bundy[7]and Dehabreh [26]are more comprehensive and provide the reader with a summary of *all* available literature on pediatric AA, their

results cannot be used to answer our clinical question: Which element(s) of history, physical exam, lab tests, or imaging studies in ED pediatric patients could rule in/out AA obviating the need for CT and MRI?

Our results were in concordance with Dehabreh et al[26] who also found history, physical exam, and lab test findings to have low sensitivity and specificity when used in isolation for the diagnosis of AA. Our results are also similar to Bundy et al[7] systematic review "Does this child have appendicitis?". Bundy et al [7] found history of fever to be "the most useful" but not diagnostic finding associated with AA. Although Bundy et al reviewed five studies for the variable fever, they derived their conclusion from only a single study on "undifferentiated abdominal pain" patients (O'Shea et al[110]) also included in our review. Our review does not support Bundy et al's statement about fever. In patients with "undifferentiated abdominal pain", using the pre-test probability of 9.8% (AA prevalence in O'shea et al[110]) and applying Bayes theorem, history of fever increases the probability of AA to only 27% which is below the treatment threshold for both CT (46.5%) and MRI (60.4%). Absence of fever decreases the probability of AA to 3% which is above the testing threshold of both CT (0.3%) and MRI (0.2%).

One criticism of previous systematic reviews [7,26] , is pooling of data from studies with different inclusion criteria. In an attempt to compare studies with similar inclusion methodology, we separated our reviewed studies into two groups: Goldman[14] and O'shea [110] on patients with "undifferentiated abdominal pain" and 19 studies [15,16,23-25, 94,102-109, 111-115]on patients "suspected of AA".

Studies that used either "suspected appendicitis" or "right lower quadrant abdominal pain" as their inclusion criteria are at high risk of partial verification bias given that the index tests (history and physical exam findings) were already used as part of their inclusion criteria. The same logic applies to PAS studies that used patients "suspected of appendicitis" as their inception cohort. Partial verification bias falsely increases sensitivity and decreases specificity and therefore alters the calculated likelihood ratios.

Studies on patients with "undifferentiated abdominal pain", may provide less biased estimates of the operating characteristics of history and physical exam findings. Comparing the test characteristics of "Cough /Hop Pain" from Goldman[14] study (N=849) which included patients with "undifferentiated abdominal pain" to studies on patients "suspected of AA" (5 studies, N=1,935), demonstrates the effect of partial verification bias. For this variable, LR+ is lower (1.61 vs 7.46) and LR- is higher (0.52 vs 0.31) in "suspected of AA" studies compared to Goldman et al[14] on "undifferentiated abdominal pain". Unfortunately, few studies are available on patients with "undifferentiated abdominal pain".

In a cohort of patients with "undifferentiated abdominal pain", Goldman et al[14] found the presence of "Cough/ hop pain" and "Pain migration to RLQ" to be most suggestive of AA (LR+ 7.64 and LR+ 4.81, respectively). Using our test-treatment model, in patients with "undifferentiated abdominal pain" (pretest probability of AA=13.4%), presence of Cough/Hop pain (LR+ 7.64) could obviate the need for CT scan but not MRI (Post-test probability of AA=54% compared to the treatment threshold of 46.5% for CT scan and 60.4% for MRI). However, this finding should be interpreted with caution as it is derived from one, single-center study and therefore may not be generalizable.

One approach used to increase the power of history, physical exam and labs in diagnosing AA is to combine them into a scoring system. One of the most studied of these scoring systems is Pediatric Appendicitis Score (PAS) which we used as an example to evaluate the bias and heterogeneity in validation studies. Pediatric Appendicitis Score (PAS) was developed by Samuel et al[12] as a clinical decision rule to identify high-risk patients for AA (Table 3). In a prospective single-center study, Samuel et al[12] evaluated 1,170 patients with a very high AA prevalence of 63% and suggested that PAS≥6 is a good predictor of AA with sensitivity of 100% and specificity of 92%. However, numerous studies aiming to validate PAS have not found such favorable operating characteristics (Sensitivity 82-88%, Specificity 50-65%) [94,104, 114]. Moreover, studies that aim to validate PAS in a cohort of patients

"suspected of AA" are at considerable risk of partial verification bias. Patients "suspected of AA" are more likely to be positive for many variables of PAS (Table-3) and consequently assigned a higher final score. The use of the test-characteristics of PAS, derived from a series of studies with such bias and their reapplication on patients "suspected of AA" substantially increases the risk of bias. A similar logic applies to any other scoring system [13,93,96,99] derived and validated on patients "suspected of AA". Studies on patients with "undifferentiated abdominal pain" are at lower risk of such bias.

We could only find one study that met our inclusion criteria and tested PAS in patients with "undifferentiated abdominal pain" and thus at lower risk of partial verification bias. Goldman et al [14] suggested PAS≥7 to be diagnostic of AA and PAS≤2 to have high validity for ruling out AA. However, using the data from Goldman et al[14] (pre-test probability of 14.5%) and applying Bayes theorem, in patients with "undifferentiated abdominal pain" and PAS≥7, the post-test probability of AA would increase to 32% which is lower than the treatment threshold for both CT scan (46.5%) and MRI (60.4%). In the same population but with PAS≤2, the post-test probability of AA would decrease to 7% which is above the test threshold for both CT scan (0.3%) and MRI (0.2%). Therefore, neither PAS≥7 nor PAS≤2 can eliminate the need for CT scan or MRI. According to our test-treatment model only PAS=0 decreases the post-test probability of AA low enough to obviate the need for CT or MRI. In other words, if a patient with "undifferentiated abdominal pain" is negative for *all* variables of PAS, the probability of AA is nearly 0% and patient can be discharged without further investigation.

Overall we found very heterogeneous results for most cutoff points of PAS. Our results are similar to Ebell et al [128]. In a systematic review of six studies, Ebell et al [128] did not find any cutoff point that can rule in/out AA. Of articles reviewed by Ebell et al [128], five were included in our review along with three additional studies that were not reviewed by Ebell et al [128] (Wu[109], Bachur[23]and Khanafer[114]) for a total 8 studies (N=4,128). One criticism to the Ebell et al [128] review is inappropriately pooling of the data from studies with different inclusion criteria and heterogeneous results [129] (I^2=91-96%). Although, we grouped

studies with similar inclusion criteria together, we still observed heterogeneous results for most cutoff points of PAS (I^2=75%-94%). Even studies with similar settings and AA prevalences had heterogeneous results. We unsuccessfully attempted to contact the authors of the included studies to utilize their raw data to calculate interval likelihood ratios instead of using arbitrary cutoff point, and therefore decrease heterogeneity.

One possible explanation for high heterogeneity observed across PAS studies is that 6 out of 8 variables composing the PAS are history and physical exam findings and therefore examiner-dependent. In one recent study focusing solely on the inter-examiner reliability of history and physical exam findings in pediatric abdominal pain patients with and without AA[130] , only vomiting showed high inter-examiner reliability (k=0.82) with other findings failing to show acceptable inter-examiner reliabilities (0.14-0.54). Yen et al[131] also found poor inter-examiner reliability of physical exam findings in pediatric patients with abdominal pain (k=0.13-0.54). When most components of a clinical score are inherently at risk of low reproducibility, it is only natural for the end results to be heterogeneous as well.

Since history, physical exam, lab tests, and PAS all had high heterogeneity and poor test characteristics, we decided to investigate the operating characteristics of ED-POCUS. Test-characteristics of ED-POCUS were similar to those of Radiology Ultrasound (RUS) as reported by Doria et al[123] (Sensitivity 86% vs 88% and Specificity91% vs 94%) Therefore, we decided not to include RUS in our test-treatment threshold model.

Using our test-treatment threshold model, (Figure 4) in patients suspected of AA (pre-test probability of 42.8%) and a positive ED-POCUS, the physician can assume the diagnosis of AA (post-test probability of 87%) without the need for CT scan or MRI. A negative ED-POCUS decreases the post-test probability of AA from 42.8% to 11% exceeding the testing threshold for CT (0.3%) and MRI (0.2%) and further investigation is recommended. The low testing threshold of CT scan and MRI can be attributed to a

combination of high benefit of treatment (Brx), low risk associated with these diagnostic modalities (Rt), and the low risk of appendectomy (Rrx). In fact, in order to obviate the need for CT and MRI using ED-POCUS, the pretest probability must be lower than 2.8 %, which is extremely unlikely in a patient evaluated for abdominal pain "suspected of AA". Factors such as availability of radiology modalities, cost of treatment, and emergency department length of stay undeniably play important roles in real life setting. However, such factors are beyond the scope of this review and are not accounted for within the aforementioned calculations.

In a patient with negative ED-POCUS, the decision to use further imaging studies, observe in the ED, obtain surgical consultation, or discharge from the ED is based on treating physician's clinical judgment. An alternative management approach to performing CT or MRI in patients with negative ED-POCUS could be antibiotic therapy. While non-operative management of AA is not the main focus of this review, we briefly address it here given the increasing evidence supporting this approach. Antibiotic therapy is a common practice in pediatric patients with complicated AA but less studied in uncomplicated AA. Few studies on this subject [132-134] , demonstrated no difference in the rate of post-operative complications between children who underwent appendectomy after failure of antibiotic therapy and those who were treated surgically upon first presentation of AA. This is similar to the findings from the adult population. In a systematic review of 5 RCTs in adult patients with uncomplicated AA, Rollins et al[135] showed that risk of complications was lower in those who had appendectomy following "failure" of antibiotic therapy compared to those who underwent appendectomy upon their first presentation of AA (10.9% vs 17.9%). In one RCT in children, Svensson et al[132] treated 24 pediatric patients suspected of AA with antibiotics alone while sent 26 to surgery. During the follow-up period of 1 year, two patients (8%) had appendectomy with pathologically proven AA. This is similar to the recurrence rate in antibiotic therapy in pediatric patients with complicated AA with or without interval appendectomy.[136] Based on the literature [132-134,137-139], we hypothesized risk of treatment failure to be 10-20 %. Therefore, the risk of recurrence in *ED-POCUS-negative patients discharged on antibiotics* will be 1-2%. This is a rough

estimate as most trials on treating AA patients with antibiotics are not RCTs and therefore, at substantial risk of selection bias. Whether or not to accept this risk depends on physician's clinical judgment and the setting in which patient is being treated. If patient can be followed and transferred to a hospital in case of a recurrence, 1-2% may be a reasonable risk to accept. However, if the treating physician has high clinical suspicion for AA despite negative ED-POCUS or anticipates poor medication compliance or loss to follow-up, using CT scan or MRI to make the final diagnosis may be a better approach.

Limitations

We excluded all studies in languages other than English which can decrease the generalizability of the findings. We did not have access to patient-level data therefore, it was impossible to evaluate the effect of factors such as symptom duration, severity of the disease, ethnicity, or socioeconomics. All studies on ED-POCUS included in this review were performed in academic settings decreasing the generalizability of the results to other settings.

CONCLUSIONS

While the aim of this review is not to provide a practice guideline, the results can be used by physicians to make decisions in the ED when caring for a child with abdominal pain. The following is the summary of our findings:

- In a patient presenting to the ED with "undifferentiated abdominal pain", migration of pain to the RLQ or presence of "cough/hop pain" in physical exam increases the probability of AA and a diagnosis of AA should be suspected in such patient. Once physician suspects AA, no single history, physical exam, lab finding, or PAS result can establish the diagnosis of AA without the need for imaging studies.

- The pooled operating characteristics for ED-POCUS in this review are similar to those reported for RUS in literature. If operator of ED-POCUS has similar expertise and training as operators in our included studies, ED-POCUS can replace RUS for the diagnosis of AA.
- In a pediatric patient "suspected of AA", a positive ED-POCUS is diagnostic of AA. However, a negative ED-POCUS is not sufficient to rule out AA without the use of CT Scan or MRI.

Areas of Future Research

Clinical scores are perhaps more valuable if constructed using variables that are well studied in multiple settings and with a design that is less prone to partial verification bias. A multi-center study on patients with undifferentiated abdominal pain that evaluates not only the test characteristics of history and physical exam findings but also their inter-examiner reliability is suggested and would be very beneficial in developing future clinical scores or guidelines. To decrease the risk of differential verification bias, it's suggested that in future studies patients observed without surgery undergo a complete follow up.
Also, the test-characteristics and inter-rater reliability of ED-POCUS should be tested in various settings, including non-academic EDs.

References:

1. National Hospital Ambulatory Medical Care Survey: 2011 Emergency Department Summary Tables https://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2011_ed_web_tables.pdf
2. Hall MJ, DeFrances CJ, Williams SN, Golosinskiy A, Schwartzman A. National Hospital Discharge Survey: 2007 summary. National health statistics reports 2010:1-20, 4.
3. Becker T, Kharbanda A, Bachur R. Atypical clinical features of pediatric appendicitis. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2007;14:124-9.
4. Nance ML, Adamson WT, Hedrick HL. Appendicitis in the young child: a continuing diagnostic challenge. Pediatric emergency care 2000;16:160-2.
5. Davenport M. Acute abdominal pain in children. BMJ (Clinical research ed) 1996;312:498-501.
6. Marzuillo P, Germani C, Krauss BS, Barbi E. Appendicitis in children less than five years old: A challenge for the general practitioner. World journal of clinical pediatrics 2015;4:19-24.
7. Bundy DG, Byerley JS, Liles EA, Perrin EM, Katznelson J, Rice HE. Does this child have appendicitis? Jama 2007;298:438-51.

8. Bratton SL, Haberkern CM, Waldhausen JH. Acute appendicitis risks of complications: age and Medicaid insurance. Pediatrics 2000;106:75-8.

9. Peng YS, Lee HC, Yeung CY, Sheu JC, Wang NL, Tsai YH. Clinical criteria for diagnosing perforated appendix in pediatric patients. Pediatric emergency care 2006;22:475-9.

10. Narsule CK, Kahle EJ, Kim DS, Anderson AC, Luks FI. Effect of delay in presentation on rate of perforation in children with appendicitis. The American journal of emergency medicine 2011;29:890-3.

11. Rothrock SG, Pagane J. Acute appendicitis in children: emergency department diagnosis and management. Annals of emergency medicine 2000;36:39-51.

12. Samuel M. Pediatric appendicitis score. Journal of pediatric surgery 2002;37:877-81.

13. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Annals of emergency medicine 1986;15:557-64.

14. Goldman RD, Carter S, Stephens D, Antoon R, Mounstephen W, Langer JC. Prospective validation of the pediatric appendicitis score. The Journal of pediatrics 2008;153:278-82.

15. Zuniga RV, Arribas JL, Montes SP, et al. Application of Pediatric Appendicitis Score on the emergency department of a secondary level hospital. Pediatric emergency care 2012;28:489-92.

16. Escriba A, Gamell AM, Fernandez Y, Quintilla JM, Cubells CL. Prospective validation of two systems of classification for the diagnosis of acute appendicitis. Pediatric emergency care 2011;27:165-9.

17. Bhatt M, Joseph L, Ducharme FM, Dougherty G, McGillivray D. Prospective validation of the pediatric appendicitis score in a Canadian pediatric emergency department. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2009;16:591-6.

18. 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 2012;263:778-85.

19. Broder J, Fordham LA, Warshauer DM. Increasing utilization of computed tomography in the pediatric emergency department, 2000-2006. Emergency radiology 2007;14:227-32.

20. 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 2011;259:793-801.

21. Howell JM, Eddy OL, Lukens TW, Thiessen ME, Weingart SD, Decker WW. Clinical policy: Critical issues in the evaluation and management of emergency department patients with suspected appendicitis. Annals of emergency medicine 2010;55:71-116.

22. Rosen MP, Ding A, Blake MA, et al. ACR Appropriateness Criteria(R) right lower quadrant pain--suspected appendicitis. Journal of the American College of Radiology : JACR 2011;8:749-55.

23. Bachur RG, Callahan MJ, Monuteaux MC, Rangel SJ. Integration of ultrasound findings and a clinical score in the diagnostic evaluation of pediatric appendicitis. The Journal of pediatrics 2015;166:1134-9.

24. Elikashvili I, Tay ET, Tsung JW. The effect of point-of-care ultrasonography on emergency department length of stay and computed tomography utilization in children with suspected appendicitis. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2014;21:163-70.

25. Sivitz AB, Cohen SG, Tejani C. Evaluation of acute appendicitis by pediatric emergency physician sonography. Annals of emergency medicine 2014;64:358-64 e4.

26. Dahabreh IJ, Adam GP, Halladay CW, et al. AHRQ Comparative Effectiveness Reviews. Diagnosis of Right Lower Quadrant Pain and Suspected Acute Appendicitis. Rockville (MD): Agency for Healthcare Research and Quality (US); 2015.

27. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ (Clinical research ed) 2015;349:g7647.

28. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. Jama 2000;283:2008-12.

29. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC medical research methodology 2006;6:31.

30. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine 2011;155:529-36.

31. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. The New England journal of medicine 1980;302:1109-17.

32. McCloskey K, Grover S, Vuillermin P, Babl FE. Ovarian torsion among girls presenting with abdominal pain: a retrospective cohort study. Emergency medicine journal : EMJ 2013;30:e11.

33. Mallick MS. Appendicitis in pre-school children: a continuing clinical challenge. A retrospective study. International journal of surgery (London, England) 2008;6:371-3.

34. Kutasy B, Laxamanadass G, Puri P. Is C-reactive protein a reliable test for suspected appendicitis in extremely obese children? Pediatric surgery international 2010;26:123-5.

35. Laurell H, Hansson LE, Gunnarsson U. Manifestations of acute appendicitis: a prospective study on acute abdominal pain. Digestive surgery 2013;30:198-206.

36. Henneman PL, Marcus CS, Inkelis SH, Butler JA, Baumgartner FJ. Evaluation of children with possible appendicitis using technetium 99m leukocyte scan. Pediatrics 1990;85:838-43.

37. Dunning PG, Goldman MD. The incidence and value of rectal examination in children with suspected appendicitis. Annals of the Royal College of Surgeons of England 1991;73:233-4.

38. Toprak H, Kilincaslan H, Ahmad IC, et al. Integration of ultrasound findings with Alvarado score in children with suspected appendicitis. Pediatrics international : official journal of the Japan Pediatric Society 2014;56:95-9.

39. Lin CH, Chen JH, Li TC, Ho YJ, Lin WC. Children presenting at the emergency department with right lower quadrant pain. The Kaohsiung journal of medical sciences 2009;25:1-9.

40. Ang A, Chong NK, Daneman A. Pediatric appendicitis in "real-time": The value of sonography in diagnosis and treatment. Pediatric Emergency Care 2001;17:334-40.

41. Bates MF, Khander A, Steigman SA, Tracy Jr TF, Luks FI. Use of white blood cell count and negative appendectomy rate. Pediatrics 2014;133:e39-e44.

42. Anandalwar SP, Callahan MJ, Bachur RG, et al. Use of white blood cell count and polymorphonuclear leukocyte differential to improve the predictive value of ultrasound for suspected appendicitis in children. Journal of the American College of Surgeons 2015;220:1010-7.

43. Cerón M, Guerrero D, Bustos E, Martínez O, Bañuelos C, Olivar V. Most frequent clinical manifestations of acute appendicitis in children under the age of three: Fifteen years of experience in the hospital infantil de mexico federico gomez emergency department. Acta Paediatrica, International Journal of Paediatrics 2011;100:46.

44. Gendel I, Gutermacher M, Buklan G, et al. Relative value of clinical, laboratory and imaging tools in diagnosing pediatric acute appendicitis. European Journal of Pediatric Surgery 2011;21:229-33.

45. Okamoto T, Sano K, Ogasahara K. Receiver-operating characteristic analysis of leukocyte counts and serum c-reactive protein levels in children with advanced appendicitis. Surgery Today 2006;36:515-8.

46. Salö M, Ohlsson B, Arnbjörnsson E, Stenström P. Appendicitis in children from a gender perspective. Pediatric Surgery International 2015;31:845-53.

47. Cesur O, Benli AR, Koyuncu M. Analyses of laboratory tests in cases with appendicitis in childhood. Konuralp Tip Dergisi 2016;8:5-8.

48. Kim E, Subhas G, Mittal VK, Golladay ES. C-reactive protein estimation does not improve accuracy in the diagnosis of acute appendicitis in pediatric patients. International journal of surgery (London, England) 2009;7:74-7.

49. Colvin JM, Bachur R, Kharbanda A. The presentation of appendicitis in preadolescent children. Pediatric emergency care 2007;23:849-55.

50. Mekhail P, Naguib N, Yanni F, Izzidien A. Appendicitis in paediatric age group: correlation between preoperative inflammatory markers and postoperative histological diagnosis. African journal of paediatric surgery : AJPS 2011;8:309-12.

51. Yazici M, Ozkisacik S, Oztan MO, Gursoy H. Neutrophil/lymphocyte ratio in the diagnosis of childhood appendicitis. The Turkish journal of pediatrics 2010;52:400-3.

52. Hsiao KH, Lin LH, Chen DF. Application of the MANTRELS scoring system in the diagnosis of acute appendicitis in children. Acta paediatrica Taiwanica = Taiwan er ke yi xue hui za zhi 2005;46:128-31.

53. Dado G, Anania G, Baccarani U, et al. Application of a clinical score for the diagnosis of acute appendicitis in childhood: a retrospective analysis of 197 patients. Journal of pediatric surgery 2000;35:1320-2.

54. Graham JM, Pokorny WJ, Harberg FJ. Acute appendicitis in preschool age children. American journal of surgery 1980;139:247-50.

55. Pearl RH, Hale DA, Molloy M, Schutt DC, Jaques DP. Pediatric appendectomy. Journal of pediatric surgery 1995;30:173-8; discussion 8-81.

56. Paajanen H, Somppi E. Early childhood appendicitis is still a difficult diagnosis. Acta paediatrica (Oslo, Norway : 1992) 1996;85:459-62.

57. Wu HP, Chang CF, Lin CY. Predictive inflammatory parameters in the diagnosis of acute appendicitis in children. Acta paediatrica Taiwanica = Taiwan er ke yi xue hui za zhi 2003;44:227-31.

58. Harland RN. Diagnosis of appendicitis in childhood. Journal of the Royal College of Surgeons of Edinburgh 1991;36:89-90.

59. Rodriguez-Sanjuan JC, Martin-Parra JI, Seco I, Garcia-Castrillo L, Naranjo A. C-reactive protein and leukocyte count in the diagnosis of acute appendicitis in children. Diseases of the colon and rectum 1999;42:1325-9.

60. Bonello JC, Abrams JS. The significance of a "positive" rectal examination in acute appendicitis. Diseases of the colon and rectum 1979;22:97-101.

61. Mollitt DL, Mitchum D, Tepas JJ, 3rd. Pediatric appendicitis: efficacy of laboratory and radiologic evaluation. Southern medical journal 1988;81:1477-9.

62. Davidson PM, Douglas CD, Hosking CS. Graded compression ultrasonography in the assessment of the "tough decision" acute abdomen in childhood. Pediatric surgery international 1999;15:32-5.

63. Klein MD, Rabbani AB, Rood KD, et al. Three quantitative approaches to the diagnosis of abdominal pain in children: practical applications of decision theory. Journal of pediatric surgery 2001;36:1375-80.

64. Tseng YC, Lee MS, Chang YJ, Wu HP. Acute abdomen in pediatric patients admitted to the pediatric emergency department. Pediatrics and neonatology 2008;49:126-34.

65. Wilson D, Sinclair S, McCallion WA, Potts SR. Acute appendicitis in young children in the Belfast urban area: 1985-1992. The Ulster medical journal 1994;63:3-7.

66. Fox JC, Hunt MJ, Zlidenny AM, Oshita MH, Barajas G, Langdorf MI. Retrospective analysis of emergency department ultrasound for acute appendicitis. The California journal of emergency medicine / California Chapter of the American Academy of Emergency Medicine 2007;8:41-5.

67. Macco S, Vrouenraets BC, de Castro SM. Evaluation of scoring systems in predicting acute appendicitis in children. Surgery 2016.

68. Wang LT, Prentiss KA, Simon JZ, Doody DP, Ryan DP. The use of white blood cell count and left shift in the diagnosis of appendicitis in children. Pediatric emergency care 2007;23:69-76.

69. Gilbert EH, Lowenstein SR, Koziol-McLain J, Barta DC, Steiner J. Chart reviews in emergency medicine research: Where are the methods? Annals of emergency medicine 1996;27:305-8.

70. Abbas PI, Zamora IJ, Elder SC, et al. How Long Does it Take to Diagnose Appendicitis? Time Point Process Mapping in the Emergency Department. Pediatric emergency care 2016.

71. Prada-Arias M, Vazquez JL, Salgado-Barreira A, Gomez-Veiras J, Montero-Sanchez M, Fernandez-Lorenzo JR. Diagnostic accuracy of fibrinogen to differentiate appendicitis from nonspecific abdominal pain in children. The American journal of emergency medicine 2016.

72. Benito J, Acedo Y, Medrano L, Barcena E, Garay RP, Arri EA. Usefulness of new and traditional serum biomarkers in children with suspected appendicitis. The American journal of emergency medicine 2016;34:871-6.

73. Turkyilmaz Z, Sonmez K, Karabulut R, et al. Sequential cytokine levels in the diagnosis of appendicitis. Scandinavian journal of clinical and laboratory investigation 2006;66:723-31.

74. Scholer SJ, Pituch K, Orr DP, Dittus RS. Use of the rectal examination on children with acute abdominal pain. Clinical pediatrics 1998;37:311-6.

75. Dixon JM, Elton RA, Rainey JB, Macleod DA. Rectal examination in patients with pain in the right lower quadrant of the abdomen. BMJ (Clinical research ed) 1991;302:386-8.

76. Groselj-Grenc M, Repse S, Vidmar D, Derganc M. Clinical and laboratory methods in diagnosis of acute appendicitis in children. Croatian medical journal 2007;48:353-61.

77. Ozguner I, Kizilgun M, Karaman A, et al. Are neutrophil CD64 expression and interleukin-6 early useful markers for diagnosis of acute appendicitis? European journal of pediatric surgery : official journal of Austrian Association of Pediatric Surgery 2014;24:179-83.

78. Sack U, Biereder B, Elouahidi T, Bauer K, Keller T, Trobs RB. Diagnostic value of blood inflammatory markers for detection of acute appendicitis in children. BMC surgery 2006;6:15.

79. Lau WY, Ho YC, Chu KW, Yeung C. Leucocyte count and neutrophil percentage in appendicectomy for suspected appendicitis. The Australian and New Zealand journal of surgery 1989;59:395-8.

80. Chakhunashvili L, Inasaridze A, Svanidze S, Samkharadze J, Chkhaidze I. Procalcitonin as the biomarker of inflammation in diagnostics of pediatric appendicular peritonitis and for the prognosis of early postoperative complications. Georgian medical news 2005:78-81.

81. Mikaelsson C, Arnbjornsson E. The value of C-reactive protein (CRP) determinations in patients with suspected acute appendicitis. Annales chirurgiae et gynaecologiae 1984;73:281-4.

82. Chen CY, Zhao LL, Lin YR, Wu KH, Wu HP. Different urinalysis appearances in children with simple and perforated appendicitis. The American journal of emergency medicine 2013;31:1560-3.

83. Kafetzis DA, Velissariou IM, Nikolaides P, et al. Procalcitonin as a predictor of severe appendicitis in children. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology 2005;24:484-7.

84. Beltran MA, Almonacid J, Vicencio A, Gutierrez J, Cruces KS, Cumsille MA. Predictive value of white blood cell count and C-reactive protein in children with appendicitis. Journal of pediatric surgery 2007;42:1208-14.

85. Lycopoulou L, Mamoulakis C, Hantzi E, et al. Serum amyloid A protein levels as a possible aid in the diagnosis of acute appendicitis in children. Clinical chemistry and laboratory medicine 2005;43:49-53.

86. Stefanutti G, Ghirardo V, Gamba P. Inflammatory markers for acute appendicitis in children: are they helpful? Journal of pediatric surgery 2007;42:773-6.

87. Miskowiak J, Burcharth F. The white cell count in acute appendicitis. A prospective blind study. Danish medical bulletin 1982;29:210-1.

88. Owen TD, Williams H, Stiff G, Jenkinson LR, Rees BI. Evaluation of the Alvarado score in acute appendicitis. Journal of the Royal Society of Medicine 1992;85:87-8.

89. Yap TL, Chen Y, Low WW, et al. A new 2-step risk-stratification clinical score for suspected appendicitis in children. Journal of pediatric surgery 2015;50:2051-5.

90. van den Broek WT, van der Ende ED, Bijnen AB, Breslau PJ, Gouma DJ. Which children could benefit from additional diagnostic tools in case of suspected appendicitis? Journal of pediatric surgery 2004;39:570-4.

91. Rubin SZ, Martin DJ. Ultrasonography in the management of possible appendicitis in childhood. Journal of pediatric surgery 1990;25:737-40.

92. Kosloske AM, Love CL, Rohrer JE, Goldthorn JF, Lacey SR. The diagnosis of appendicitis in children: outcomes of a strategy based on pediatric surgical evaluation. Pediatrics 2004;113:29-34.

93. Fleischman RJ, Devine MK, Yagapen MA, et al. Evaluation of a novel pediatric appendicitis pathway using high- and low-risk scoring systems. Pediatric emergency care 2013;29:1060-5.

94. Schneider C, Kharbanda A, Bachur R. Evaluating appendicitis scoring systems using a prospective pediatric cohort. Annals of emergency medicine 2007;49:778-84, 84.e1.

95. Becker T, Kharbanda A, Bachur R. Atypical clinical features of pediatric appendicitis. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2007;14:124-9.

96. Kharbanda AB, Taylor GA, Fishman SJ, Bachur RG. A clinical decision rule to identify children at low risk for appendicitis. Pediatrics 2005;116:709-16.

97. Macklin CP, Radcliffe GS, Merei JM, Stringer MD. A prospective evaluation of the modified Alvarado score for acute appendicitis in children. Annals of the Royal College of Surgeons of England 1997;79:203-5.

98. Bond GR, Tully SB, Chan LS, Bradley RL. Use of the MANTRELS score in childhood appendicitis: a prospective study of 187 children with abdominal pain. Annals of emergency medicine 1990;19:1014-8.

99. Lintula H, Pesonen E, Kokki H, Vanamo K, Eskelinen M. A diagnostic score for children with suspected appendicitis. Langenbeck's archives of surgery / Deutsche Gesellschaft fur Chirurgie 2005;390:164-70.

100. Saucier A, Huang EY, Emeremni CA, Pershad J. Prospective evaluation of a clinical pathway for suspected appendicitis. Pediatrics 2014;133:e88-95.

101. Kharbanda AB, Dudley NC, Bajaj L, et al. Validation and refinement of a prediction rule to identify children at low risk for acute appendicitis. Archives of pediatrics & adolescent medicine 2012;166:738-44.

102. Kwan KY, Nager AL. Diagnosing pediatric appendicitis: usefulness of laboratory markers. The American journal of emergency medicine 2010;28:1009-15.

103. Kentsis A, Ahmed S, Kurek K, et al. Detection and diagnostic value of urine leucine-rich alpha-2-glycoprotein in children with suspected acute appendicitis. Annals of emergency medicine 2012;60:78-83.e1.

104. Mandeville K, Pottker T, Bulloch B, Liu J. Using appendicitis scores in the pediatric ED. The American journal of emergency medicine 2011;29:972-7.

105. Huckins DS, Simon HK, Copeland K, Spiro DM, Gogain J, Wandell M. A novel biomarker panel to rule out acute appendicitis in pediatric patients with abdominal pain. The American journal of emergency medicine 2013;31:1368-75.

106. Cayrol J, Miguez MC, Guerrero G, Tomatis C, Simal I, Maranon R. Diagnostic accuracy and prognostic utility of D Dimer in acute appendicitis in children. European journal of pediatrics 2016;175:313-20.

107. Fox JC, Solley M, Anderson CL, Zlidenny A, Lahham S, Maasumi K. Prospective evaluation of emergency physician performed bedside ultrasound to detect acute appendicitis. European Journal of Emergency Medicine 2008;15:80-5.

108. Kim C, Kang BS, Choi HJ, Lim TH, Oh J, Chee Y. Clinical application of real-time tele-ultrasonography in diagnosing pediatric acute appendicitis in the ED. American Journal of Emergency Medicine 2015;33:1354-9.

109. Wu HP, Yang WC, Wu KH, Chen CY, Fu YC. Diagnosing appendicitis at different time points in children with right lower quadrant pain: comparison between Pediatric Appendicitis Score and the Alvarado score. World journal of surgery 2012;36:216-21.

110. O'Shea JS, Bishop ME, Alario AJ, Cooper JM. Diagnosing appendicitis in children with acute abdominal pain. Pediatric emergency care 1988;4:172-6.

111. Doniger SJ, Kornblith A. Point-of-Care Ultrasound Integrated Into a Staged Diagnostic Algorithm for Pediatric Appendicitis. Pediatric emergency care 2016.

112. Santillanes G, Simms S, Gausche-Hill M, et al. Prospective evaluation of a clinical practice guideline for diagnosis of appendicitis in children. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2012;19:886-93.

113. Khan AN, Sawan A, Likourezos A, Schnellinger M, Garcia E. The usefulness of procalcitonin in the diagnosis of appendicitis in children: a pilot study. Emergency medicine international 2012;2012:317504.

114. Khanafer I, Martin DA, Mitra TP, et al. Test characteristics of common appendicitis scores with and without laboratory investigations: a prospective observational study. BMC pediatrics 2016;16:147.

115. Bachur RG, Dayan PS, Dudley NC, et al. The Influence of Age on the Diagnostic Performance of White Blood Cell Count and Absolute Neutrophil Count in Suspected Pediatric Appendicitis. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2016.

116. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ (Clinical research ed) 2003;327:557-60.

117. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2013;20:1194-206.

118. Migraine S, Atri M, Bret PM, Lough JO, Hinchey JE. Spontaneously resolving acute appendicitis: clinical and sonographic documentation. Radiology 1997;205:55-8.

119. Heller MB, Skolnick ML. Ultrasound documentation of spontaneously resolving appendicitis. The American journal of emergency medicine 1993;11:51-3.

120. Di Saverio S, Sibilio A, Giorgini E, et al. The NOTA Study (Non Operative Treatment for Acute Appendicitis): prospective study on the efficacy and safety of antibiotics (amoxicillin and clavulanic acid) for treating patients with right lower quadrant abdominal pain and long-term follow-up of conservatively treated suspected appendicitis. Annals of surgery 2014;260:109-17.

121. Dettori JR. Loss to follow-up. Evidence-Based Spine-Care Journal 2011;2:7-10.

122. Bhandari M, Guyatt GH, Swiontkowski MF. User's guide to the orthopaedic literature: how to use an article about a surgical therapy. The Journal of bone and joint surgery American volume 2001;83-a:916-26.

123. Doria AS, Moineddin R, Kellenberger CJ, et al. US or CT for Diagnosis of Appendicitis in Children and Adults? A Meta-Analysis. Radiology 2006;241:83-94.

124. Moore MM, Kulaylat AN, Hollenbeak CS, Engbrecht BW, Dillman JR, Methratta ST. Magnetic resonance imaging in pediatric appendicitis: a systematic review. Pediatric radiology 2016;46:928-39.

125. Aziz O, Athanasiou T, Tekkis PP, et al. Laparoscopic versus open appendectomy in children: a meta-analysis. Annals of surgery 2006;243:17-27.

126. Healy DA, Doyle D, Moynagh E, et al. Systematic Review and Meta-Analysis on the Influence of Surgeon Specialization on Outcomes Following Appendicectomy in Children. Medicine 2015;94:e1352.

127. Wan MJ, Krahn M, Ungar WJ, et al. Acute appendicitis in young children: cost-effectiveness of US versus CT in diagnosis--a Markov decision analytic model. Radiology 2009;250:378-86.

128. Ebell MH, Shinholser J. What are the most clinically useful cutoffs for the Alvarado and Pediatric Appendicitis Scores? A systematic review. Annals of emergency medicine 2014;64:365-72.e2.

129. Kharbanda AB. Appendicitis: do clinical scores matter? Annals of emergency medicine 2014;64:373-5.

130. Kharbanda AB, Fishman SJ, Bachur RG. Comparison of pediatric emergency physicians' and surgeons' evaluation and diagnosis of appendicitis. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2008;15:119-25.

131. Yen K, Karpas A, Pinkerton HJ, Gorelick MH. Interexaminer reliability in physical examination of pediatric patients with abdominal pain. Archives of pediatrics & adolescent medicine 2005;159:373-6.

132. Svensson JF, Patkova B, Almstrom M, et al. Nonoperative treatment with antibiotics versus surgery for acute nonperforated appendicitis in children: a pilot randomized controlled trial. Annals of surgery 2015;261:67-71.

133. Minneci PC, Mahida JB, Lodwick DL, et al. Effectiveness of Patient Choice in Nonoperative vs Surgical Management of Pediatric Uncomplicated Acute Appendicitis. JAMA surgery 2016;151:408-15.

134. Tanaka Y, Uchida H, Kawashima H, et al. Long-term outcomes of operative versus nonoperative treatment for uncomplicated appendicitis. Journal of pediatric surgery 2015;50:1893-7.

135. Rollins KE, Varadhan KK, Neal KR, Lobo DN. Antibiotics Versus Appendicectomy for the Treatment of Uncomplicated Acute Appendicitis: An Updated Meta-Analysis of Randomised Controlled Trials. World journal of surgery 2016.

136. Andersson RE, Petzold MG. Nonsurgical treatment of appendiceal abscess or phlegmon: a systematic review and meta-analysis. Annals of surgery 2007;246:741-8.

137. Abes M, Petik B, Kazil S. Nonoperative treatment of acute appendicitis in children. Journal of pediatric surgery 2007;42:1439-42.

138. Armstrong J, Merritt N, Jones S, Scott L, Butter A. Non-operative management of early, acute appendicitis in children: is it safe and effective? Journal of pediatric surgery 2014;49:782-5.

139. Minneci PC, Sulkowski JP, Nacion KM, et al. Feasibility of a nonoperative management strategy for uncomplicated acute appendicitis in children. Journal of the American College of Surgeons 2014;219:272-9.

PUBMED

History and Physical Exam

("appendicitis"[MeSH Terms] OR "appendicitis"[All Fields]) AND (("sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "Predictive Value of Tests"[Mesh])) AND (("child"[MeSH Terms] OR "child"[All Fields]) OR ("adolescent"[MeSH Terms] OR "adolescent"[All Fields])) AND (("physical examination"[MeSH Terms] OR (("physical"[All Fields] OR “clinical”[All Fields]) AND ("examination"[All Fields] OR “feature*”[All Fields] OR “finding*”[All Fields])) OR "physical examination"[All Fields]) OR (("medical"[All Fields] AND "history"[All Fields]) OR "medical history"[All Fields] OR "Medical History Taking"[Mesh]))

Ultrasonography

("appendicitis"[MeSH Terms] OR "appendicitis"[All Fields]) AND (("sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "Predictive Value of Tests"[Mesh])) AND (("child"[MeSH Terms] OR "child"[All Fields]) OR ("adolescent"[MeSH Terms] OR "adolescent"[All Fields])) AND (("ultrasonography"[Subheading] OR "ultrasonography"[All Fields] OR "ultrasound"[All Fields] OR "ultrasonography"[MeSH Terms] OR "ultrasound"[All Fields] OR "ultrasonics"[MeSH Terms] OR "ultrasonics"[All Fields]) OR ("tomography, x-ray computed"[MeSH Terms] OR ("tomography"[All Fields] AND "x-ray"[All Fields] AND "computed"[All Fields]) OR "x-ray computed tomography"[All Fields] OR ("ct"[All Fields] AND "scan"[All Fields]) OR "ct scan"[All Fields]))

Lab Tests

("appendicitis"[MeSH Terms] OR "appendicitis"[All Fields]) AND (("sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "Predictive Value of Tests"[Mesh])) AND (("child"[MeSH Terms] OR "child"[All Fields]) OR ("adolescent"[MeSH Terms] OR "adolescent"[All Fields])) AND ((("laboratories"[MeSH Terms] OR "laboratories"[All Fields] OR "laboratory"[All Fields]) AND (characteristics[All Fields]] OR tests[All Fields])) OR ("clinical laboratory techniques"[MeSH Terms] OR ("clinical"[All Fields] AND "laboratory"[All Fields] AND "techniques"[All Fields]) OR "clinical laboratory techniques"[All Fields] OR ("laboratory"[All Fields] AND "diagnosis"[All Fields]) OR "laboratory diagnosis"[All Fields]))

EMBASE

History and Physical Exam

(('appendicitis'/exp OR appendicitis) AND ((sensitivity AND specificity) OR 'sensitivity and specificity'/exp OR 'diagnostic value'/exp OR 'prediction'/exp OR prediction OR 'predictive value'/exp) AND ('child'/exp OR child OR 'adolescent'/exp OR adolescent)) AND ('physical examination'/exp OR 'physical examination' OR 'clinical examination'/exp OR 'clinical examination' OR 'clinical feature'/exp OR 'clinical feature' OR physical OR clinical) AND ('examination'/exp OR 'examination' OR examination/exp OR examination OR findings OR features)

Ultrasonography

(('appendicitis'/exp OR appendicitis) AND ((sensitivity AND specificity) OR 'sensitivity and specificity'/exp OR 'diagnostic value'/exp OR 'prediction'/exp OR prediction OR 'predictive value'/exp) AND ('child'/exp OR child OR 'adolescent'/exp OR adolescent)) AND (('ultrasound'/exp OR ultrasound) OR ('echography'/exp OR echography) OR ('ultrasonics'/exp OR ultrasonics) OR 'computer assisted tomography'/exp OR (ct AND scan))

Lab Tests

(('appendicitis'/exp OR appendicitis) AND ((sensitivity AND specificity) OR 'sensitivity and specificity'/exp OR 'diagnostic value'/exp OR 'prediction'/exp OR prediction OR 'predictive value'/exp) AND ('child'/exp OR child OR 'adolescent'/exp OR adolescent)) AND (('laboratory'/exp OR laboratory) AND (characteristics OR tests OR 'diagnosis'/exp OR diagnosis) OR 'laboratory diagnosis'/exp)

SCOPUS

History and Physical Exam

Appendicitis AND Physical Examination

Appendicitis AND Physical findings

Appendicitis AND history

Appendicitis AND medical history

Ultrasonography

Appendicitis AND ultrasonography

Appendicitis AND ultrasound

Lab Tests

Appendicitis AND laboratory tests

Table 1.1- Description of Reviewed Studies: History, Physical Exam and Lab findings

Study	Design and Participants	Potential Predictors of Appendicitis	Gold Standard	Prevalence (95% CI)
O'Shea, 1988	**Design:** Prospective **Inclusion:** Abdominal Pain < 1 W Age: 3-18 YR **Exclusion:** Recent Trauma Recurrent abdominal pain Care taker not knowing English **Sample Size:** 246 **Median Age:** 11 (3-18) YR **Gender:** 48% M	Anorexia Nausea/Vomiting Fever Diarrhea Dysuria Lethargy	-Histopathology -No surgery, F/u* 3-6 D	9.8% (6.6-14.11)
Schneider, 2007	**Design:** Prospective **Inclusion:** Suspected appendicitis Age: 3-21 YR **Exclusion:** Pregnancy Previous abdominal surgery Chronic medical conditions Abdominal imaging in < 2 W **Sample Size:** 588 **Median Age:** 11.9 (IQR 8.5-14.9) YR **Gender:** 54% M	Anorexia Nausea/Vomiting Pain migration to RLQ RLQ tenderness Cough/hop pain Rebound tenderness T≥37.3 WBC≥10,000 Neutrophil≥75%	-Histopathology -No surgery, F/u 2 W or contacting patient's pediatrician	34% (30-37)
Goldman ,2008	**Design:** Prospective **Inclusion:** Abdominal pain < 1 W Age:1-17 YR **Exclusion:** -Previous appendectomy **Sample Size:** 849 **Mean Age:** NS **Gender:** NS	Anorexia Nausea/Vomiting Pain migration to RLQ Cough/hop pain Fever (T≥38° C) RLQ tenderness WBC≥10,000 Neutrophil ≥ 7500	-Histopathology -No surgery, F/u 5-7 D	14.5% (12-17%)

Study	Design and Participants	Potential Predictors of Appendicitis	Gold Standard	Prevalence (95% CI)
Kwan, 2010	**Design:** Prospective **Inclusion:** Suspected appendicitis Age:1-18 YR **Exclusion:** Pregnancy Chronic medical conditions Abdominal surgery in < 1 YR **Sample Size:** 209 **Mean Age:** 10.5 ± 3.7 YR **Gender:** 59% M	RLQ Tenderness RLQ Rebound tenderness LLQ Tenderness Periumbilical Tenderness RUQ Tenderness LUQ Tenderness Epigastric Tenderness Suprapubic Tenderness WBC≥12,000 CRP≥3 mg/dL	-Histopathology -No surgery, F/u 2-6 W	55% (48-61%)
Mandeville ,2011	**Design:** Prospective **Inclusion:** Suspected appendicitis Age: 4-17 YR **Exclusion:** Pregnancy Previous abdominal surgery Chronic medical conditions Abdominal imaging in < 2 W **Sample Size:** 287 **Mean Age:** 9.8 ±3.1 YR **Gender:** 52.6% M	Anorexia Nausea/Vomiting Pain Migration to RLQ RLQ tenderness T≥37.3 Rebound tenderness Cough/hop pain WBC≥10,000 Neutrophil≥75%	-Histopathology -No surgery, F/u 2 W	54% (48-59%)
Escriba, 2011	**Design:** Prospective **Inclusion:** Suspected appendicitis Age: 4-18 YR **Exclusion:** Not having lab tests **Sample Size:** 99 **Mean Age:** 11.2 ± 3.7 YR **Gender:** 62.6% M	PAS	-Histopathology -No surgery, F/u 10 D	42% (33-52%)

Study	Design and Participants	Potential Predictors of Appendicitis	Gold Standard	Prevalence (95% CI)
Wu, 2012	**Design:** Prospective **Inclusion:** RLQ pain Age: 3-18 YR **Exclusion:** Pain duration ≥3 D Loss to F/u **Sample Size:** 1,395 **Mean Age:** 11.1± 4.2 YR **Gender:** 46.2 % M	PAS	-Histopathology -No surgery, F/u 2 W	63% (60-65%)
Khan, 2012	**Design:** Prospective **Inclusion:** RLQ pain Age: 5-17 YR **Exclusion:** Obvious signs of Gastroenteritis Chronic medical conditions Pregnancy **Sample Size:** 50 **Mean Age:** 11±3.2 YR **Gender:**44% M	Anorexia Nausea/Vomiting Fever RLQ Pain LLQ Pain Epigastric Pain RLQ Tenderness RLQ Rebound tenderness Guarding Bowel Sounds Procalcitonin	-Histopathology -No surgery, F/u 24 HRs and 2 W	44% (30-58)
Santillanes, 2012	**Design:** Prospective **Inclusion:** Suspected appendicitis **Exclusion:** NS **Sample Size:** 475 **Mean Age:** 11 (IQR 7-15)YR **Gender:** 50% M	Fever Nausea/Vomiting Anorexia RLQ pain Periumbilical pain Obstipation Diarrhea RLQ tenderness RLQ Rebound tenderness Guarding Psoas Sign Obturator sign Rosving's sign WBC≥10,000	-Histopathology -No surgery, F/u 1 W	41% (36-45%)

Study	Design and Participants	Potential Predictors of Appendicitis	Gold Standard	Prevalence (95% CI)
Kentsis ,2012	**Design:** Prospective **Inclusion:** Suspected Appendicitis Age <18 Surgical consult or Imaging requested **Exclusion:** Chronic Medical Conditions Pregnancy **Sample Size:**49 **Mean Age:** 10.9 ± 4.3 YR **Gender:** 53% M	Nausea/Vomiting Fever Pain Migration to RLQ Pain Duration RLQ Pain or Tenderness	-Histopathology -No surgery, F/u 6-8 W	49% (35-62%)
Zuniga, 2012	**Inclusion:** Suspected Appendicitis Age< 14 **Exclusion:** Pain ≥7 D Previous appendectomy No lab test available **Sample size:** 101 **Mean age:** 9.51 (± 2.76) years **Gender:** 54.5% M	Anorexia Nausea/Vomiting Pain Migration to RLQ T≥37.3 RLQ Tenderness Rebound Tenderness Cough/Hop Pain WBC≥ 10,000 Neutrophil ≥ 7500	-Histopathology -No surgery, F/u 7 D through database search	28% (20-37%)

Study	Design and Participants	Potential Predictors of Appendicitis	Gold Standard	Prevalence (95% CI)
Huckins, 2013	**Design:** Prospective **Inclusion:** Suspected appendicitis Age:2-20 YR Pain< 72 HR **Exclusion:** Previous appendectomy Chronic Medical conditions Abdominal trauma Invasive abdominal procedures Participation in any other research protocol in the past 2 W **Sample Size**: 503 **Median Age:** 12 (8-16) YR **Gender:** 43% M	Anorexia Nausea/Vomiting Pain migration to RLQ RLQ tenderness Fever ≥ 37.5°C Rebound tenderness Rigidity and guarding Rovsing sign	-Histopathology -No surgery, Discharge diagnosis, No F/u	29% (25-32%)
Sivitz, 2014	**Design:** Prospective **Inclusion:** Suspected appendicitis **Exclusion:** Previous abdominal surgery Unstable Vital Signs **Sample Size**: 231 **Median Age:** 10.3 (IQR 7.8-16.1) YR **Gender:** 60% M	Anorexia Nausea/Vomiting Pain migration to RLQ Fever Rebound Tenderness Cough/Hop pain WBC≥10,000 Neutrophil≥ 6750 Urine Ketones	-Histopathology -No surgery, F/u ,6 Mo	33% (27-39)
Bachur,2015	**Design:** Prospective **Inclusion:** Suspected appendicitis Age: 3-18 YR Having RLQ Ultrasound **Exclusion:** Previous abdominal surgery Current Antibiotic use Chronic medical conditions **Sample Size:** 728	Anorexia Nausea Pain migration to RLQ T≥38 Maximal pain in RLQ Guarding Rebound Cough/Hop Pain RLQ Pain duration	-Histopathology -No surgery, F/u 1-2 W	29% (26-32%)

Study	Design and Participants	Potential Predictors of Appendicitis	Gold Standard	Prevalence (95% CI)
	Median Age: 11.7 (IQR 7.8-14.9) YR **Gender:** 44% M	WBC≥10,000 Neutrophil ≥75%		
Cayrol ,2016	**Design:** Prospective **Inclusion:** Suspected appendicitis Age: 1-16 YR Having lab tests **Exclusion:** Previous appendectomy Chronic medical conditions Pregnancy Anticoagulant treatment **Sample Size:** 135 **Mean Age:** 9.44 ± 3.3 YR **Gender:** 51.1%	Nausea and Vomiting T≥37.5 Pain Migration to RLQ RLQ pain Tenderness Diffuse Abdominal Tenderness Rebound Tenderness WBC≥10,000 D-dimer CRP	-Histopathology -No surgery, F/u 5 D	38% (30-47%)
Bachur, 2016	**Design:** Prospective **Inclusion:** Suspected appendicitis Age: 3-18 YR Pain< 72 HR Having WBC and ANC in their workup **Exclusion:** Previous abdominal surgery Pregnancy Chronic GI conditions **Sample Size:** 2,133 **Median Age:** 10.9 (IQR 8-13.9) YR **Gender:** 42% M	Symptom Duration WBC ANC	-Histopathology -No surgery, F/u 1-2 W, 3 M Medical record review	41% (39-43)
Khanfer, 2016	**Design:** Prospective **Inclusion:** Suspected appendicitis Age: 5-17 YR **Exclusion:**	PAS	-Histopathology -No surgery, F/u 1 M	30.6% (24-37%)

Study	Design and Participants	Potential Predictors of Appendicitis	Gold Standard	Prevalence (95% CI)
	Previous appendectomy Previous Abdominal Surgery Established dx of AA Pregnancy Non-verbal **Sample Size:** 180 **Mean Age:** 11.2 ± 3.1 years YR **Gender:** 43.3% M			
Doniger, 2016	**Design:** Prospective **Inclusion:** Suspected appendicitis Age 2-18 **Exclusion:** Pregnancy Previous abdominal imaging **Sample Size:** 40 **Mean Age:** 9.26 YR **Gender:** 50% M	Anorexia Nausea/Vomiting Fever Fever RLQ Rebound Tenderness	-Histopathology -No surgery, F/u >2 W	40% (26-55)

*Follow-up: Defined as contacting patient's care giver unless otherwise specified
YR: Year, Mo: Month, D: Day, HR: Hour, M: Male, T: Temperature, M:Male
RLQ: Right Lower Quadrant, LLQ: Left Lower Quadrant, PAS: Pediatric Appendicitis Score
IQR: Interquartile Range, NS: Non Specified
WBC: White Blood Cells, CRP:C-Reactive Protein

Table 1.2- Description of Reviewed Studies: Emergency Department Point-of-Care Ultrasound (ED-POCUS)

Study	Design and Participants	Gold Standard	ED-POCUS	Prevalence (95% CI)
Fox et al. , 2007	**Design:** Prospective **Inclusion:** Suspected appendicitis **Exclusion:** -Pregnancy -Unable to consent **Sample Size:**42 **Mean Age:** NS **Median Age:** NS **Gender:** NS	-Histopathology -No surgery, F/u 2 W and 3 M	Operator: EM faculty physicians and residents Interpreter: Same Training: Lecture	54 % (40-68%)
Eliskashvili et al. ,2014	**Design:** Prospective **Inclusion:** Suspected appendicitis Age<21 **Exclusion:** -Unstable Vital Signs -Dx of AA or IBD -Prior abdominal CT or US **Sample size:** 150 **Mean age:** 12 ± 5.2 YR **Gender:** 44% M	-Histopathology -No surgery, F/u 3 W	Operator: Pediatric EM faculty and fellows Interpreter: Same Training: 30 minutes lecture plus 30 minutes hands-on session	33% (26-41%)
Sivitz et al. , 2014	**Design:** Prospective **Inclusion:** Suspected appendicitis **Exclusion:** -Previous abdominal surgery -Unstable Vital Signs **Sample Size:** 231 **Median Age:** 10.2 (2-20.9) YR **Gender:** 53% M	-Histopathology -No surgery, F/u, 6 M	Operator: Pediatric EM faculty and fellows Interpreter: Same Training: 45 minutes lecture plus 5 supervised scans	33% (27-39)

Study	Design and Participants	Gold Standard	ED-POCUS	Prevalence (95% CI)
Kim et al., 2015	**Design:** Prospective **Inclusion:** Suspected appendicitis Age< 19 **Exclusion:** -Patients lost to F/u -Previous CT scan **Sample Size:** 115 **Mean Age:** 10.6 ± 3.3 YR **Gender:** 56.5% M	-Histopathology -No surgery, F/u for undetermined duration	Operator: EM residents Interpreter: Same Training: 1-2 YR experience with ultrasound plus 20 minutes simulation training session	31% (23-40%)
Doniger et al , 2016	**Design:** Prospective **Inclusion:** Suspected appendicitis Age 2-18 **Exclusion:** Pregnancy Previous abdominal imaging **Sample Size:** 40 **Mean Age:** 9.26 YR **Gender:** 50% M	-Histopathology -No surgery, F/u >2 W	Operator: EM resident and attending, PEM attending Interpreter: Same Training: 30 minutes appendicitis ultrasound tutorial plus 40 supervised scans	40% (26-55)

YR: Year, M: Male ,F/u: Follow-up, IBD: Inflammatory Bowel Disease, CT: Computed Tomography US: Ultrasound , EM: Emergency Pediatric Emergency Medicine

Table 2.1- History Findings in Children Suspected of Appendicitis

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI
Fever	Kentsis et al, 2012	49	46% (26-67)	48% (28-69)	0.88 (0.5
	Khan et al, 2012	50	41%(21-63)	64% (44-81)	1.15 (0.5
	Sivitz et al, 2014	231	26% (17-38)	72% (65-79)	0.95 (0.6
	Doniger et al, 2016	40	38% (15-64)	42% (22-63)	0.64 (0.3
Pooled Data		**370**	**33% (26-42)**	**I^2=75.3%**	**0.90 (0.67-1**
Nausea /Vomiting	Schneider et al, 2007	588	86% (80-90)	35% (31-40)	1.33 (1.2
	Mandeville et al, 2011	287	75% (68-82)	36% (28-44)	1.18 (1.0
	Santillanes et al, 2012	475	73% (66-79)	38% (33-45)	1.19 (1.0
	Kentsis et al, 2012	49	59% (36-79)	64% (43-82)	1.64 (0.8
	Zuniga et al, 2012	101	79% (59-92)	44% (32-56)	1.40 (1.0
	Khan et al , 2012	50	73% (50-89)	39% (21-59)	1.19 (0.8
	Huckins et al, 2013	503	64% (56-72)	63% (58-68)	1.74 (1.4
	Sivitz et al, 2014	231	74% (62-83)	37% (29-45)	1.17 (0.9
	Bachur et al, 2015	728	70% (63-76)	40% (35-44)	1.16 (1.0
	Cayrol et al, 2016	134	42% (29-57)	77% (66-85)	1.82 (1.1
	Doniger et al, 2016	40	75% (71-76)	33% (16-55)	1.13 (0.7
Pooled Data		**3,186**	**I^2=79.9%**	**I^2=91.6%**	**1.30 (1.19-1**
Anorexia	Schneider et al, 2007	588	73% (66-79)	44% (39-49)	1.29 (1.1
	Mandeville et al, 2011	287	74% (67-81)	38% (30-47)	1.19 (1.0
	Santillanes et al, 2012	475	80% (74-86)	43% (37-50)	1.42 (1.2
	Zuniga et al, 2012	101	89% (72-98)	33% (22-45)	1.33 (1.0
	Khan et al, 2012	50	86% (65-97)	32% (16-52)	1.27 (0.9
	Huckins et al, 2013	503	73% (65-80)	45% (40-50)	1.32 (1.1
	Sivitz et al, 2014	231	74% (62-83)	51% (43-59)	1.50 (1.2
	Bachur et al, 2015	728	71% (64-77)	47% (42-40)	1.31 (1.1
	Doniger et al, 2016	40	75% (47-93)	54% (33-74)	1.64 (0.9
Pooled Data		**3,003**	**75% (72-78)**	**44% (42-46)**	**1.33 (1.26-1**
Pain Migration to RLQ	Schneider et al, 2007	588	49% (42-56)	73% (69-78)	1.85 (1.4
	Mandeville et al, 2011	287	45% (37-53)	64% (56-73)	1.27 (0.9
	Zuniga et al, 2012	101	46% (28-66)	77% (65-86)	1.99 (1.1

Predictors of Pediatric Appendicitis		Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI
		Kentsis et al, 2012	49	63% (41-81)	72% (51-88)	2.23 (1.1
		Huckins, 2013	503	69% (61-77)	61% (56-66)	1.77 (1.4
		Sivitz et al, 2014	231	50% (38-62)	68% (60-75)	1.55 (1.1
		Bachur et al, 2015	728	50% (43-57)	75% (71-79)	1.99 (1.6
		Cayrol et al, 2016	134	33% (20-47)	80% (70-88)	1.68 (0.9
Pooled Data			**2,621**	**I^2=77.6%**	**I^2=77%**	**1.75 (1.58-1**
Pain	**RLQ**	Santillanes et al, 2012	475	91% (86-95)	36% (30-42)	1.42 (1.2
		Khan et al, 2012	50	82% (60-95)	39% (22-59)	1.35 (0.9
		Cayrol et al, 2016	134	75% (61-86)	38% (27-49)	1.21 (0.9
		Pooled Data	**703**	**I^2=77.4%**	**37% (32-42)**	**1.38 (1.26-1**
	Periumbilical	Santillanes et al, 2012	475	50% (43-57)	50% (44-56)	1.0 (0.8
	LLQ	Khan et al, 2012	50	4% (0.1-22)	99% (82-100)	1.27 (0.08
	Epigastric	Khan et al, 2012	50	9% (1-29)	82% (62-93)	0.51 (0.1
	Diffuse	Santillanes et al, 2012	475	30% (24-37)	72% (67-78)	1.08 (0.8
		Khan et al, 2012	50	4.5% (1-23)	96% (82-100)	1.27 (0
		Pooled Data	**525**	**I^2=88.3%**	**I^2=90.6%**	**1.09 (0.81-1**
Symptom Duration <12hrs		Kentsis et al, 2012	49	29% (13-51)	55% (35-76)	0.66 (0.3
		Huckins et al, 2013	503	24% (17-31)	72% (67-76)	0.83 (0.5
		Bachur et al, 2015	728	19% (14-25)	77% (73-81)	0.86 (0.6
		Bachur et al, 2016	2,133	24% (21-27)	67% (64-70)	0.73 (0.6
Pooled Data			**3,413**	**23% (21-26)**	**I^2=86%**	**0.76 (0.68-0**
Symptom Duration 12-24hrs		Kentsis et al, 2012	49	29% (13-51)	84% (64-95)	1.82 (0.6
		Huckins et al, 2013	503	38% (30-47)	66% (61-71)	1.12 (0.8
		Bachur et al, 2015	728	33% (27-40)	72% (67-75)	1.17 (0.9
		Bachur et al, 2016	2,133	34% (31-37)	68% (66-71)	1.07 (0.9
Pooled Data			**3,413**	**35% (30-40)**	**69% (67-71)**	**1.10 (1.00-1**
Symptom Duration 24-48hrs		Kentsis et al, 2012	49	38% (19-59)	68% (46-85)	1.17 (0.5
		Huckins et al, 2013	503	22% (15-29)	80% (75-84)	1.07 (0.7
		Bachur et al, 2015	728	28% (22-35)	78% (75-82)	1.30 (1
		Bachur et al, 2016	2,133	30% (27-34)	75% (73-78)	1.23 (1.0

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI
Pooled Data		**3,413**	**29% (27-32)**	**77% (75-79)**	**1.23 (1.09-1**
Symptom Duration 48-72hrs	Kentsis et al, 2012	49	4% (0.0-21)	96% (80-100)	1.04 (0.0
	Huckins et al, 2013	503	17% (11-24)	82% (78-86)	0.95 (0.6
	Bachur et al, 2015	728	10% (7-15)	92% (89-94)	1.30 (0.7
	Bachur et al, 2016	2,133	11% (9-14)	89% (87-91)	1.06 (0.8
Pooled Data		**3,413**	**12% (10-14)**	**I^2=85.5%**	**1.07 (0.88-1**
Symptom Duration >72hrs	Bachur et al, 2015	728	5% (2-8)	89% (85-91)	0.41 (0.2
Obstipation	Santillanes et al, 2012	475	17% (11-23)	91% (87-95)	1.96 (1.1
Diarrhea	Santillanes et al, 2012	475	22% (16-28)	82% (77-86)	1.21 (0.8

See Table 4 for test characteristics of "Fever" based on temperature cutoff point used

Pooled data is reported only when I-square (I^2) ≤ 50%

Table 2.2- History, Physical Exam and Lab Tests Findings in Children with Undifferentiated Abdominal Pain

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Fever	O'Shea et al, 1988	246	75% (53-90)	78% (72-83)	3.4 (2.42-4.76)	0.32 (0.16-0.64)
Nausea or Vomiting	O'Shea et al, 1988	246	79% (58-93)	64% (57-70)	2.2 (1.68-2.88)	0.33 (0.15-0.71)
	Goldman et al, 2008	849	75% (66-82)	54% (50-57)	1.62 (1.42-1.84)	0.47 (0.34-0.64)
Pooled Data		**1,095**	**75% (68-82)**	**I^2=86%**	**I^2=75%**	**0.45 (0.33-0.60)**
Anorexia	O'Shea et al, 1988	246	21% (7-42%)	73% (66-78%)	0.77(0.34-1.7)	1.08 (.087-1.35)
	Goldman et al, 2008	849	68% (59-76)	64% (61-68)	1.92 (1.65-2.24)	0.49 (0.38-0.64)
Pooled Data		**1,095**	**I^2=94.7%**	**I^2=82.4%**	**I^2=80.7%**	**I^2=96.6%**
Pain Migration to RLQ	Goldman et al, 2008	849	46% (37-56)	90% (88-92)	4.81 (3.59-6.44)	0.59 (0.50-0.70)
Cough/Hop Pain	Goldman et al, 2008	849	72% (63-79)	91% (88-93%)	7.64 (5.94-9.83)	0.31 (0.24-0.42)
Right Iliac Fossa tenderness	Goldman et al, 2008	849	79% (71-86)	83% (80-86)	4.74 (3.94-5.7)	0.24 (0.17-0.35)
T>38 °C	Goldman et al, 2008	849	59% (50-68)	79% (76-82)	2.80 (2.28-3.43)	0.52 (0.42-0.64)
Lethargy	O'Shea et al, 1988	246	4%(0.11-21)	95% (91-98)	0.84 (0.11-6.23)	1.01 (0.92-1.1)
Diarrhea	O'Shea et al, 1988	246	33% (15-55)	87% (82-941)	2.55(1.32-4.94)	0.77 (0.58-1.02)
Dysuria	O'Shea et al, 1988	246	0% (0.0-13)	96% (93-98)	0.59 (0.35-10)	1.0 (0.95-1.07)

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Nausea without Vomiting	O'Shea et al, 1988	246	29% (13-51)	79% (73-84)	1.38 (0.7-2.7)	0.9 (0.69-1.17)
WBC≥10,000	Goldman et al, 2008	308*	88% (80-93)	57% (50-64)	2.04 (1.71-2.43)	0.22 (0.13-0.36)
ANC≥7500	Goldman et al, 2008	289**	83% (74-90)	64% (57-71)	2.33 (1.89-2.88)	0.26 (0.17-0.41)

Pooled data is reported only when I-square (I^2) ≤50%

*Data available only for a subgroup of total population (308/849)

**Data available only for a subgroup of total population (289/849)

RLQ: Right Lower Quadrant, T: Temperature

WBC: White Blood Cells, ANC: Absolute Neutrophil Count reported as cells per mm^3

Table 2.3- Physical Exam Findings in Children Suspected of Appendicitis

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)
Fever*	Schneider et al, 2007	588	46% (39-53)	55% (50-60)	1.03 (0.85
	Mandeville et al, 2011	287	50% (42-58)	64% (55-72)	1.37 (1.04
	Santillanes et al, 2012	475	31%(25-38)	67%(61-73)	0.95 (0.73
	Zuniga et al, 2012	101	46% (28-66)	68% (57-79)	1.47 (0.87
	Huckins et al, 2013	503	25% (18-33)	84% (79-87)	1.52 (1.05
	Bachur et al, 2015	728	31% (25-37)	71% (67-75)	1.05 (0.82
	Cayrol et al, 2016	134	42% (29-57)	60% (48-70)	1.05 (0.70
Pooled Data		**2,816**	**I^2=83%**	**I^2=92.4%**	**1.13 (0.99-1.29**
RLQ Tenderness	Schneider et al, 2007	588	83% (77-88)	37% (32-42)	1.32 (1.19
	Kwan et al, 2010	209	78% (70-85)	23% (15-33)	1.02 (0.88
	Mandeville et al, 2011	287	92% (87-96)	11% (6-17)	1.03 (0.96
	Santillanes et al, 2012	475	95% (90-97)	32% (27-38)	1.40 (1.28
	Zuniga et al, 2012	101	89% (72-98)	33% (22-45)	1.33 (1.08
	Khan et al, 2012	50	100% (84-100)	7% (0.9-23)	1.07 (0.96
	Huckins, 2013	503	99% (96-100)	14% (11-18)	1.15 (1.10
	Sivitz et al, 2014	231	96% (89-99)	8% (4-13)	1.04 (0.98
	Bachur et al, 2015	728	87% (82-91)	40% (35-44)	1.45 (1.32
	Doniger et al, 2016	40	100% (79-100)	25% (10-47)	1.31 (1.02
Pooled Data		**3,212**	**I^2=86.3%**	**I^2=94.5%**	**I^2=90.5%**
Cough/Hop Pain	Schneider et al, 2007	588	68% (60-74)	64% (59-69)	1.89 (1.60
	Mandeville et al, 2011	287	83% (76-89)	46% (38-55)	1.55 (1.30
	Zuniga et al, 2012	101	71% (51-87)	48% (36-60)	1.37 (1.00
	Sivitz et al, 2014	231	64% (53-75)	66% (58-73)	1.89 (1.43
	Bachur et al, 2015	728	69% (62-75)	52% (48-57)	1.44 (1.27
Pooled Data		**1,935**	**I^2=74.6%**	**I^2=84.8%**	**1.61 (1.42**
RLQ Rebound Tenderness	Schneider et al, 2007	588	48% (41-55)	77% (72-81)	2.05 (1.63
	Kwan et al, 2010	209	15% (9-23)	90% (83-96)	1.54 (0.72

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)
	Mandeville et al, 2011	287	47% (39-55)	71% (62-78)	1.59 (1.17
	Santillanes et al, 2012	475	36% (29-43)	87% (83-91)	2.83 (1.97
	Zuniga et al, 2012	101	50% (31-69)	71% (59-81)	1.74 (1.04
	Khan et al, 2012	50	68% (45-86)	68% (47-84)	2.12 (1.15
	Huckins, 2013	503	51% (42-59)	80% (75-84)	2.49 (1.92
	Sivitz et al, 2014	231	62% (50-73)	79% (71-85)	2.91 (1.05
	Bachur et al, 2015	728	38% (32-45)	84% (80-87)	2.33 (1.80
	Cayrol et al, 2016	134	58% (43-71)	67% (56-71)	1.75 (1.19
	Doniger et al, 2016	40	50% (25-75)	92% (73-99)	6.0 (1.46
Pooled Data		**3,346**	**I^2=86.9%**	**I^2=78.6%**	**2.19 (1.91-2.51**
Guarding	Khan et al, 2012	50	59% (36-79)	64% (44-81)	1.65 (0.90
	Santillanes et al , 2012	475	70% (63-76)	69% (63-74)	2.25 (1.84
	Huckins, 2013	503	74% (66-81)	68% (63-73)	2.32 (1.93
	Bachur et al, 2015	728	65% (59-72)	65% (61-69)	1.87 (1.60
Pooled Data		**1,756**	**69% (65-73)**	**67% (64-69)**	**2.09 (1.83-2.37**
RUQ Tenderness	Kwan et al, 2010	209	6% (2-12)	89% (81-95)	0.57 (0.23
LLQ Tenderness	Kwan et al, 2010	209	13% (7-20)	85% (76-91)	0.87 (0.44
LUQ Tenderness	Kwan et al, 2010	209	2% (0-6)	97% (91-99)	0.54 (0.09
Epigastric Tenderness	Kwan et al, 2010	209	14% (8-21)	86% (77-92)	1.00 (0.51
Suprapubic Tenderness	Kwan et al, 2010	209	5% (2-11)	96% (89-98)	1.23 (0.36
Periumbilical Tenderness	Kwan et al, 2010	209	22% (15-30)	72% (62-81)	0.79 (0.49
	Santillanes et al, 2012	475	42% (35-49)	63% (57-69)	1.13 (0.90
Pooled Data		**684**	**I^2=92.4%**	**I^2=64%**	**1.00 (0.72**
Diffuse Tenderness	Cayrol et al, 2016	134	71% (57-83)	45% (34-57)	1.30 (1.00
	Santillanes et al, 2012	475	27% (21-34)	78% (72-83)	1.22 (0.88
Absent/decreased Bowel Sounds	Sanillanes et al, 2012	475	40% (32-47)	87% (82-91)	3.06 (2.14
	Khan et al, 2012	50	14% (3-35)	98% (84-100)	7.64 (0.4
Psoas Sign	Santillanes et al, 2012	475	38% (31-46)	88% (83-91)	3.15 (2.17
Obturator Sign	Santillanes et al, 2012	475	34% (27-41)	90% (86-94)	3.52 (2.30

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)
Rovsing's Sign	Santillanes et al, 2012	475	34% (27-42)	91% (87-94)	3.94 (2.54
	Huckins, 2013	503	36% (28-44)	89% (85-92)	3.24 (2.23
Pooled Data		**978**	**35% (30-40)**	**90% (87-92)**	**3.52 (2.65-4.68**

*See Table 2.4 for test characteristics of "Fever" based on temperature cutoff point used

Pooled data is reported only when I-square (I^2) ≤ 50%

Table 2.4 Fever in Patients Suspected of Appendicitis

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
Objective Fever in ED	Santillanes et al, 2012	475	31%(25-38)	67%(61-73)	0.95(0.73-1-24)	1.02(0.93-1.13)
T>38 °C	Bachur et al, 2015	728	31% (25-37)	71% (67-75)	1.05 (0.82-1.33)	0.98 (0.89-1.09)
T>37.3-37.5 °C	Schneider et al, 2007	588	46% (39-53)	55% (50-60)	1.03 (0.85-1.24)	0.98 (0.84-1.15)
	Mandeville et al, 2011	287	50% (42-58)	64% (55-72)	1.37 (1.04-1.80)	0.79 (0.65-0.97)
	Zuniga et al, 2012	101	46% (28-66)	69% (57-79)	1.47 (0.87-2.48)	0.78 (0.54-1.14)
	Huckins et al, 2013	503	25% (18-33)	84% (79-87)	1.52 (1.05-2.20)	0.90 (0.81-1.00)
	Cayrol et al, 2016	134	42% (29-57)	60% (48-70)	1.05 (0.70-1.59)	0.97 (0.72-1.29)
Pooled Data		**1,613**	**I^2=82.9**	**I^2=94.8%**	**1.22 (1.02-1.46)**	**0.90 (0.83-0.97)**

Pooled data is reported only when I-square (I^2) ≤ 50%

Table 2.5- Laboratory Tests in Patients Suspected of Appendicitis

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95%
WBC≥10,000	Schneider et al, 2007	588	90% (85-94)	52% (47-57)	1.88 (1.6
	Mandeville et al, 2011	287	83% (76-89)	63% (54-71)	2.24 (1.7
	Zuniga et al, 2012	101	93% (77-91)	49% (37-61)	1.83 (1.4
	Santillanes et al, 2012	475	89% (84-93)	49% (42-55)	1.73 (1.5
	Sivitz et al, 2014	231	86% (76-93)	64% (56-71)	2.37 (1.8
	Bachur et al, 2015	728	83% (77-88)	64% (60-68)	2.29 (2.0
	Cayrol et al, 2016	134	88% (77-96)	60% (48-70)	2.20 (1.6
	Bachur et al, 2016	2,133	90% (88-92)	54% (51-57)	1.96 (1.8
Pooled Data		**4,677**	**88% (87-90)**	**56% (54-58)**	**2.01 (1.86**
WBC>12,000	Kwan et al, 2010	209	71% (62-79)	66% (55-75)	2.09 (1.5
ANC≥6,750	Sivitz et al, 2014	231	91% (82-96)	57% (49-65)	2.10 (1.7
ANC≥7,500	Zuniga et al, 2012	101	96% (82-100)	56% (44-68)	2.20 (1.6
Neut≥67%	Santillanes et al,2012	475	96% (92-98)	39% (33-45)	1.57 (1.4
Neut≥75%	Schneider et al, 2007	588	84% (78-89)	57% (52-62)	1.95 (1.7
	Mandeville et al, 2011	287	77% (69-83)	64% (55-72)	2.11 (1.6
	Bachur et al, 2015	728	75% (68-80)	64% (60-68)	2.09 (1.9
Pooled Data		**1,603**	**I^2=67.9%**	**I^2=64.1%**	**2.02 (1.85**
CRP>3 mg/dL	Kwan et al, 2010	209	70% (60-79)	65% (53-75)	1.98 (1.4
	Cayrol et al, 2016	134	38% (25-53)	85% (76-92)	2.63 (1.4
Pooled Data		**343**	**I^2=93.3%**	**I^2=90%**	**2.10 (1.61**
WBC>12,000+CRP>3	Kwan et al, 2010	209	42% (33-51)	91% (86-97)	4.36 (2.2
Positive Urine Ketone	Sivitz et al, 2014	231	37% (26-49)	75% (67-82)	1.46 (0.9
D-Dimer>230 ng/dL	Cayrol et al, 2016	134	40% (27-55)	80% (70-88)	2.07 (1.1
Procalcitonin > 0.39 ng/dL	Khan et al ,2012	50	25% (8-45)	92% (76-99)	3.25 (0
3 Marker Panel*	Huckins et al, 2013	503	96% (92-99)	43% (38-48)	1.70 (1.5

Pooled data is reported only when I-square (I^2) ≤ 50%

WBC: White Blood Cells, ANC: Absolute Neutrophil Count reported as cells per mm^3 Neut: Neutrophil

***** Mathematical combination of WBC, CRP and Myeloid Related Protein 8/14 (MRP 8/14)

Table 2.6-Pediatric Appendicitis Score (PAS) in Patients Suspected of AA

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI
PAS≥0	Mandeville et al, 2011	287	100% (98-100)	0% (0-0)	1.00 (0.99-
PAS≥1	Schneider et al, 2007	588	100% (98-100)	1% (0.6-3)	1.01(1-
	Mandeville et al, 2011	287	94% (89-97)	1% (1-4)	0.95 (0.91-
	Khanafer et al, 2016	180	100% (93-100)	1% (0.2-5)	1.01 (0.99-
Pooled Data		**1,055**	**I^2=88.2%**	**1% (1-3)**	**I^2=85.6%**
PAS≥2	Schneider et al, 2007	588	100% (98-100)	4% (2-6)	1.03 (0.97-
	Mandeville et al, 2011	287	94% (89-97)	2% (0-7)	0.96 (0.92-
	Khanafer et al, 2016	180	100% (93-100)	4.8% (2-10)	1.04 (0.99-
Pooled Data		**1,055**	**I^2=88.8**	**3.7% (2-5)**	**I^2=79.6%**
PAS≥3	Schneider et al, 2007	588	100% (98-100)	12% (0.9-16)	1.14 (1.1-
	Mandeville et al, 2011	287	93% (89-97)	8% (4-14)	1.02 (0.96-
	Khanafer et al, 2016	180	98% (90-100)	13% (7-20)	1.13 (1.04-
Pooled Data		**1,055**	**I^2=88.3%**	**11.7% (9-14)**	**I^2=77.2%**
PAS≥4	Schneider et al, 2007	588	96% (92-98)	26% (21-30)	1.29 (1.21-
	Mandeville et al, 2011	287	93% (89-97)	15% (10-22)	1.10 (1.02-
	Khanafer et al, 2016	180	91% (80-97)	23% (16-31)	1.18 (1.04-
Pooled Data		**1,055**	**94% (91-98)**	**23% (20-26)**	**I^2=76.1%**
PAS≥5	Schneider et al, 2007	588	92% (88-96)	46% (41-51)	1.73 (1.57-
	Mandeville et al, 2011	287	90% (84-94)	30% (23-39)	1.29 (1.14-
	Khanafer et al, 2016	180	91% (80-97)	39% (31-48)	1.50 (1.24-
Pooled Data		**1,055**	**91% (88-94)**	**I^2=82%**	**I^2=85.5%**
PAS≥6	Schneider et al, 2007	588	82% (77-88)	65% (60-70)	2.38 (2.05-
	Mandeville et al, 2011	287	88% (82-93)	50% (41-59)	1.77 (1.48-
	Khanafer et al, 2016	180	82% (69-91)	56% (47-65)	1.88 (1.48-
Pooled Data		**1,055**	**85% (81-88)**	**I^2=80.9%**	**2.01 (1.64-2**
PAS≥7	Schneider et al, 2007	588	70% (63-76)	78% (74-82)	3.16 (2.57-
	Mandeville et al, 2011	287	80% (73-86)	67% (58-75)	2.40 (1.86-
	Wu et al, 2012	1,395	82% (79-85)	82% (78-86)	4.68 (3.87-

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI
	Bachur et al, 2015	728	59% (52-66)	81% (77-84)	3.11 (2.52-
	Khanafer et al, 2016	180	69% (55-81)	72% (63-80)	2.47 (1.77-
Pooled Data		**3,178**	**I^2=92.7%**	**I^2=79.5%**	**I^2=83.8%**
PAS≥8	Schneider et al, 2007	588	50% (43-58)	90% (86-93)	5.04 (3.63-
	Mandeville et al, 2011	287	66% (58-74)	80% (73-87)	3.37 (2.35-
	Zuniga et al, 2012	101	61% (41-78)	93% (85-98)	8.86 (3.62-2
	Khanafer et al, 2016	180	56% (42-70)	86% (78-91)	3.91 (2.41-
Pooled Data		**1,156**	**I^2=62.6%**	**I^2=72%**	**4.40 (3.26-5**
PAS≥9	Schneider et al, 2007	588	28% (22-35)	96% (94-98)	7.37 (4.28-1
	Mandeville et al, 2011	287	44% (36-52)	92% (86-96)	5.27 (2.91-
	Khanafer et al, 2016	180	26% (15-39)	92% (86-96)	3.18 (1.50-
Pooled Data		**1,055**	**I^2=82%**	**94% (92-96)**	**5.26 (3.34-8**
PAS≥10	Schneider et al, 2007	588	9% (5-13)	99% (98-100)	16.49 (3.84-
	Mandeville et al, 2011	287	13% (8-19)	98% (94-100)	5.68 (1.73-1
	Khanafer et al, 2016	180	7% (2-17)	97% (92-99)	2.27 (0.59-
Pooled Data		**1,055**	**10% (7-13)**	**98.6% (97-99)**	**5.80 (1.97-1**
PAS 0-3	Bachur et al, 2015	728	5% (3-9)	71% (67-75)	0.18 (0.10-
PAS 4-6	Bachur et al, 2015	728	36% (29-43)	48% (44-53)	0.69 (0.57-
PAS≥1	Escriba et al, 2011	99	100% (92-100)	0% (0-0.06)	0.99 (0.98-
PAS>1	Escriba et al, 2011	99	100% (91-100)	10.5% (4-21)	1.10 (1.0-
PAS>2	Escriba et al, 2011	99	100% (92-100)	19% (10-32)	2.36 (1.07-
PAS>3	Escriba et al, 2011	99	100% (92-100)	39% (26-52)	1.63 (1.33-
PAS>4	Escriba et al, 2011	99	98% (87-100)	67% (53-79)	2.93 (2.02-
PAS>5	Escriba et al, 2011	99	93% (81-99)	86% (74-94)	6.62 (3.46-1
PAS>6	Escriba et al, 2011	99	88% (74-96)	98% (91-100)	50.21 (7.17
PAS>7	Escriba et al, 2011	99	69% (53-82)	100% (94-100)	79.58 (5.00-1
PAS>8	Escriba et al, 2011	99	43% (28-59)	100% (94-100)	49.91 (3.09
PAS>9	Escriba et al, 2011	99	7% (2-19)	100% (94-100)	9.44 (0.50
PAS>10	Escriba et al, 2011	99	0% (0-8)	100% (94-100)	0.46 (0.02-1

Pooled data is reported only when I-square (I^2) ≤ 50%

Table 2.7- Pediatric Appendicitis Score (PAS) in Patients with Abdominal Pain

Predictors of Pediatric Appendicitis	Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)
PAS≥0	Goldman et al, 2008	849	100% (97-100)	27% (24-30)	1.37 (1.30-1
PAS≥1	Goldman et al, 2008	849	87% (80-92)	55% (51-59)	1.93 (1.74-2
PAS≥2	Goldman et al, 2008	849	68% (59-76)	73% (70-76)	2.53 (2.14-3
PAS≥3	Goldman et al, 2008	849	50% (41-60)	83% (80-86)	2.98 (2.35-3
PAS≥4	Goldman et al, 2008	849	40% (31-49)	90% (88-92)	3.96 (2.91-5
PAS≥5	Goldman et al, 2008	849	28% (20-36)	94% (92-96)	4.56 (3.04-6
PAS≥6	Goldman et al, 2008	849	16% (10-24)	96% (94-97)	4.07 (2.38-6
PAS≥7	Goldman et al, 2008	849	6% (2-11)	98% (97-99)	2.75 (1.15-6
PAS≥8	Goldman et al, 2008	849	2% (0-6)	99% (98-100)	1.69 (0.35-8
PAS≥9	Goldman et al, 2008	849	0% (0-5)	100% (100-100)	5.86 (0.12-294
PAS≥10	Goldman et al, 2008	849	0% (0-5)	100% (100-100)	5.86 (0.12-294

Pooled data is reported only when I-square (I^2) ≤ 50%

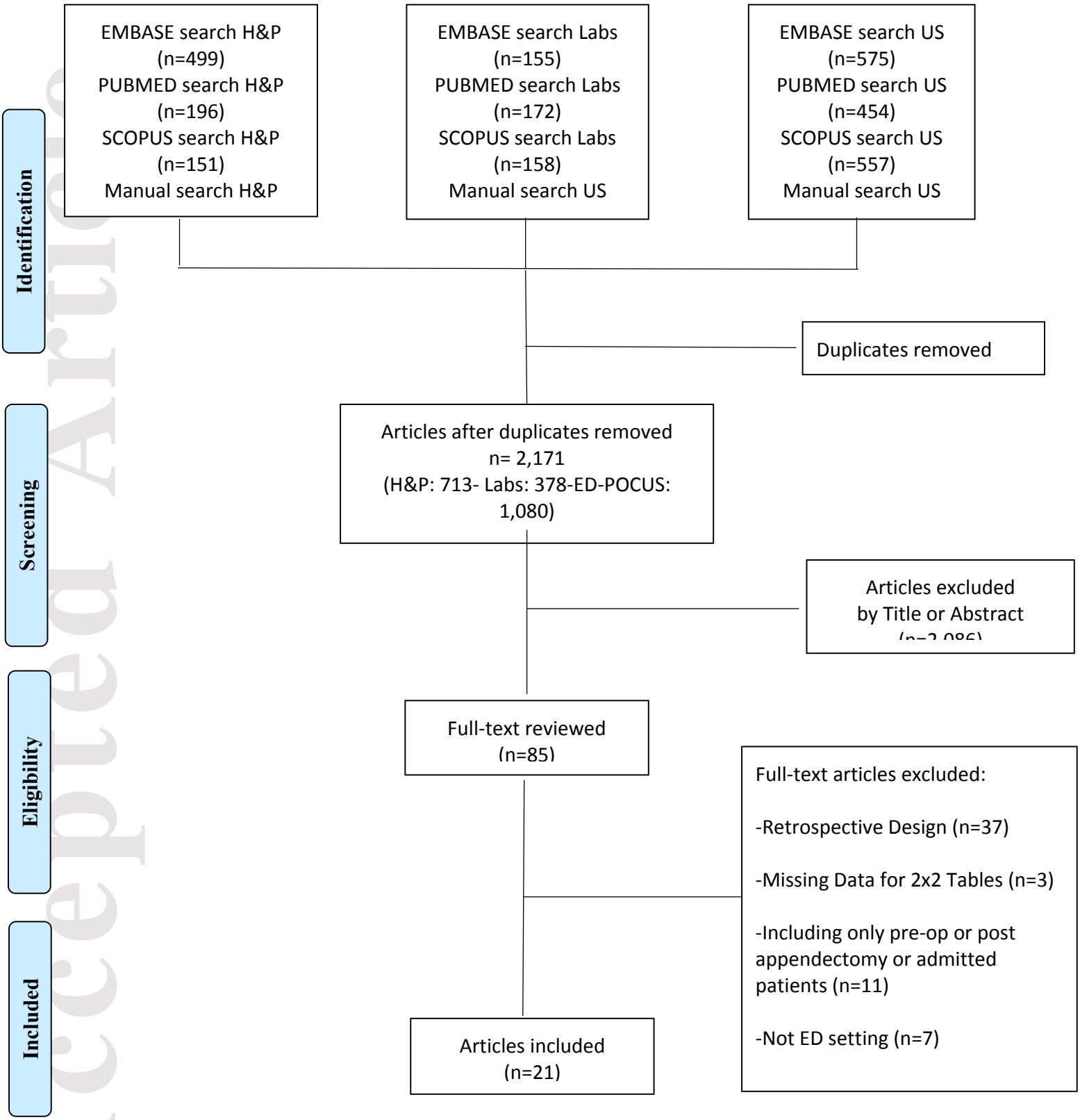
Table 2.8- Emergency Department-Point of Care Ultrasound (ED-POCUS) in patients suspected of Appendicitis

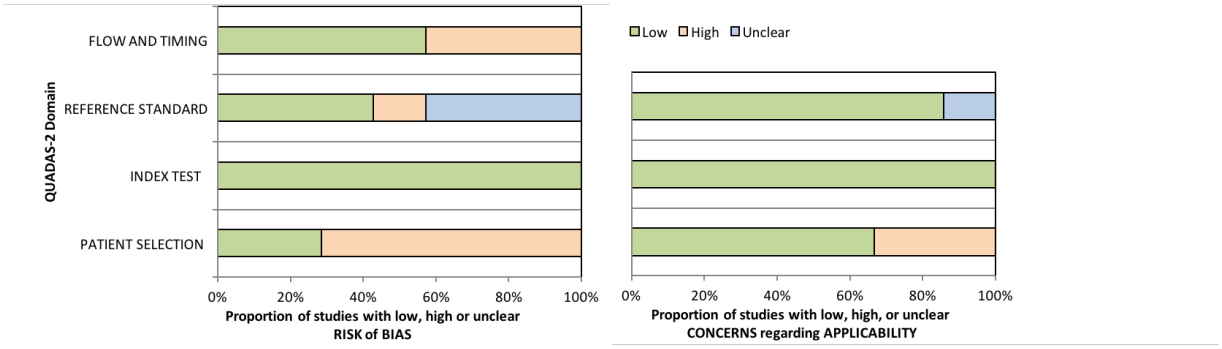
Studies	Sample Size	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	
Fox et al, 2007	42	74% (52-90)	85% (60-97)	4.68 (1.61-13.60)	
Sivitz et al, 2014	264	85% (75-92)	93% (88-96)	11.66 (6.86-19.84)	
Kim et al, 2015	115	92% (78-98)	90% (81-96)	9.05 (4.66-17.59)	
Doniger et al, 2016	40	94% (70-100)	88%(68-97)	7.50 (2.58-21.78)	
Pooled Data	**461**	**86% (79-91)**	**91%(87-94)**	**9.24 (6.42-13.28)**	**0.1**
Eliskashvilli et al, 2014	150	60% (45-74)	95% (89-98)	12.00 (4.96-29.04)	

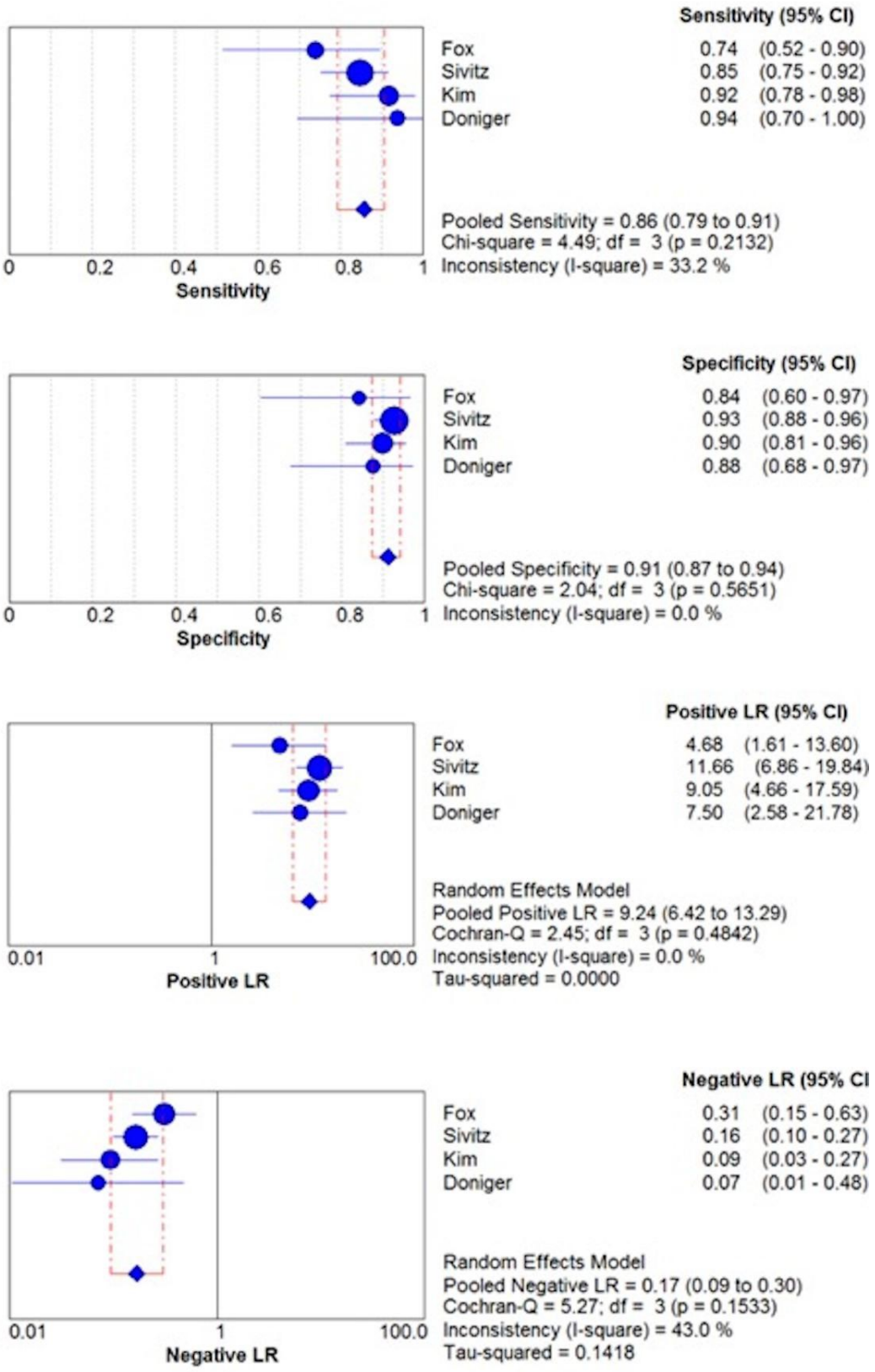
Pooled data is reported only when I-square (I^2) ≤ 50%

Table 3. Variable of Pediatric Appendicitis Score (PAS)

Variable	Point Value
Migration of Pain	1
Anorexia	1
Nausea/Vomiting	1
RLQ tenderness	2
Pain with cough/hopping/percussion	2
Fever	1
Leukocytosis	1
Differential WBC count with a left shift	1







Testing threshold= [(Ppos/nd) x (Rrx) + Rt] ÷ [(Ppos/nd x Rrx) + (Ppos/d x Brx)] = 0.3% (CT scan), 0.2% (MRI)

Treatment threshold= [(Pneg/nd) x (Rrx) - Rt] ÷ [(Pneg/nd x Rrx) + (Pneg/d x Brx)]= 46.5% (CT scan), 60.4% (MRI)

Ppos/nd = Probability of a positive result in patients without disease = 1-specificity = 0.05 (CT scan), 0.039 (MRI)

Pneg/nd = Probability of a negative result in patients without disease = specificity= 0.95 (CT scan), 0.961 (MRI)

Rrx = Risk of treatment in patients without disease= 0.05

Rt = Risk of diagnostic test= 0.00026 (CT scan), 0 (MRI)

Ppos/d = Probability of a positive result in patients with disease = sensitivity= 0.94 (CT scan), 0.965 (MRI)

Pneg/d = Probability of a negative result in patients with disease = 1 – sensitivity = 0.06 (CT scan), 0.035 (MRI)

Brx = Benefit of treatment in patients with disease= 0.9