



Ruling out Appendicitis in Children: Can We Use Clinical Prediction Rules?

Paul van Amstel¹ · Ramon R. Gorter¹ · Johanna H. van der Lee² · Huib A. Cense³ · Roel Bakx¹ · Hugo A. Heij¹

Received: 2 June 2018 / Accepted: 23 September 2018 / Published online: 29 October 2018
© 2018 The Author(s)

Abstract

Purpose To identify available clinical prediction rules (CPRs) and investigate their ability to rule out appendicitis in children presenting with abdominal pain at the emergency department, and accordingly select CPRs that could be useful in a future prospective cohort study.

Methods A literature search was conducted to identify available CPRs. These were subsequently tested in a historical cohort from a general teaching hospital, comprising all children (< 18 years) that visited the emergency department between 2012 and 2015 with abdominal pain. Data were extracted from the electronic patient files and scores of the identified CPRs were calculated for each patient. The negative likelihood ratios were only calculated for those CPRs that could be calculated for at least 50% of patients.

Results Twelve CPRs were tested in a cohort of 291 patients, of whom 87 (29.9%) suffered from acute appendicitis. The Ohmann score, Alvarado score, modified Alvarado score, Pediatric Appendicitis score, Low-Risk Appendicitis Rule Refinement, Christian score, and Low Risk Appendicitis Rule had a negative likelihood ratio < 0.1. The Modified Alvarado Scoring System and Lintula score had a negative likelihood ratio > 0.1. Three CPRs were excluded because the score could not be calculated for at least 50% of patients.

Conclusion This study identified seven CPRs that could be used in a prospective cohort study to compare their ability to rule out appendicitis in children and investigate if clinical monitoring and re-evaluation instead of performing additional investigations (i.e., ultrasound) is a safe treatment strategy in case there is low suspicion of appendicitis.

Keywords Clinical prediction rules · Appendicitis · Children

Introduction

The diagnosis of acute appendicitis in children remains challenging as symptoms can vary from mild abdominal pain to generalized peritonitis and septicemia. Historically, the diagnosis of appendicitis is mainly based upon clinical examination in combination with biochemical variables indicative for

inflammation. A disadvantage of this diagnostic strategy was the relatively high negative appendectomy rate of 12.3–19%.¹ To reduce this, an evidence-based guideline was proposed in 2010 by the Association of Surgeons of the Netherlands, which makes preoperative imaging mandatory in patients with suspected appendicitis.³ Ultrasound is the preferred initial diagnostic imaging modality in both the adult and pediatric

✉ Ramon R. Gorter
r.r.gorter@amc.uva.nl

Paul van Amstel
p.vanamstel@vumc.nl

Johanna H. van der Lee
j.h.vanderlee@amc.uva.nl

Huib A. Cense
hcense@rkz.nl

Roel Bakx
r.bakx@amc.uva.nl

Hugo A. Heij
hugo.heij@icloud.com

¹ Paediatric Surgical Centre of Amsterdam, Emma Children's Hospital Amsterdam University Medical Centre, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands

² Division Woman and Child, Amsterdam University Medical Centre, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands

³ Division of Surgery, Red Cross Hospital, Vondellaan 13, 1942 LE Beverwijk, The Netherlands

population.³ Implementation of this guideline resulted in a significant decrease of negative appendectomies to 2.2%–5%.^{2,4} Currently in the Netherlands, in 99.7% of the adult patients' preoperative imaging studies are performed.⁴

A consequence of the abovementioned policy is that the threshold to perform additional imaging studies is low in children presenting at the ER, especially since ultrasonography (US) can be performed quickly with minimal burden and harm for the patient. The downside of this lower threshold is the risk of potential inconclusive results from ultrasound, which may lead to exposure of children to harmful and expensive diagnostic procedures, such as CT scans, MRIs, or even diagnostic laparoscopies.^{5–7} Instead of these invasive diagnostic procedures, literature suggests that watchful waiting could be considered after non-visualization of the appendix on ultrasound.⁸ Selection of patients with high probability of acute appendicitis would help to reduce exposure to abovementioned invasive diagnostic procedures. Clinical prediction rules (CPR), such as the Alvarado score,⁹ were initially designed to diagnose appendicitis, but may also be used to rule out appendicitis. CPRs mostly consist of variables from medical history, physical examination, and biochemical testing. Large heterogeneity exists between CPRs in terms of included variables and cutoff values. Several studies showed that the value of these CPRs to diagnose appendicitis is low, reflected by positive likelihood ratios ranging from 1.7–8.5.^{10–15} Data regarding their value in ruling out appendicitis in the pediatric population are scarce.^{10,16}

The first objective of this study was to identify commonly applied CPRs through a literature search. The aim of the second part of the study was to investigate the value of the identified CPRs in ruling out appendicitis in the pediatric population in the Netherlands based on the negative likelihood ratios and thereby select CPRs that could potentially be used in a future prospective cohort study. Additionally, in order to determine if the use of imaging modalities could be reduced by adopting CPRs to rule out appendicitis, we determined the number of imaging procedures performed in patients that were qualified as low risk for the disease according to these CPRs.

Methods

Identification of the CPRs: Literature Review

Initially, a literature search (according to the PRISMA guidelines) was performed in the PubMed database to identify potential usable CPRs.¹⁷ (Appendix 1) Studies were screened for title and abstract and subsequently assessed for full text by two independent reviewers. Disagreements were solved by consensus. In addition, references from the included articles were screened to identify other CPRs. No other databases than PubMed were screened for potential CPRs. Studies about CPRs that were developed to diagnose or exclude appendicitis

were included in this review. A CPR was excluded if it contained variables only applicable to the adult population (e.g., points attributed to age > 50 years). CPRs consisting of more than 15 variables or variables that needed multiplication were considered as impractical in an emergency department and therefore were excluded. CPRs described in other languages than English and CPRs containing variables that were not routinely determined in our hospital, such as rectal-axillary temperature difference, were also excluded.

Study Design and Selection of Participants

A single-center historical cohort study was conducted in a general teaching hospital. All children younger than 18 years presenting at the emergency department between January 1st 2012 and December 31st 2014 with abdominal pain were eligible for inclusion. A consecutive sample of patients with a differential diagnosis of appendicitis, identified using the international classification of diseases (ICD) codes for acute abdomen, acute appendicitis, and general abdominal complaints was used. The treating physician assigned these codes at the time of presentation at the emergency department. Children with abdominal pain due to trauma, presentation of another main complaint than abdominal pain, those not cooperating with physical examination, and those referred to another hospital were excluded.

Data Extraction

Data were extracted from electronic patient files using a standardized form (Appendix 2), based upon the variables used in the identified CPRs. One author (PA) performed the data extraction and 10% of the database was randomly reviewed for completeness by another author (RG). Information on the following variables was extracted:

General Gender, age (years), and date of presentation.

Clinical Variables Duration of abdominal pain (hours), location of pain in right iliac fossa (RIF), migration of pain, anorexia, nausea, vomiting, intensity of pain (Numeric Rating Scale (NRS)), progression of pain, steady/constant pain, hopping tenderness, coughing tenderness, percussion tenderness, dysuria, rebound tenderness, guarding, Rovsing's sign (contralateral palpation tenderness), presence of bowel sounds, rigidity, tenderness inside/outside the right iliac fossa (RIF), and temperature.

Biochemical Variables Leucocytes ($\times 10^9/L$), C-reactive protein (CRP) (mg/L), leucocytes differential count, and urinalysis (for detection of urinary tract infection (UTI)).

Imaging Variables Free fluid on ultrasonography (US), appendicolith on US, appendicular wall thickening (wall thickness > 0.7 cm) on US, appendicular abscess or suppuration on US, performance of computed tomography (CT) abdomen, and performance of magnetic resonance imaging (MRI) abdomen.

The following definitions were used in this study.

Appendicitis Intraoperative diagnosis made by the treating surgeon in combination with pathologically proven inflammation of the appendix was used as the reference standard. Patients with radiographically documented appendicitis who were managed by antibiotics alone did not have pathology reports and were therefore excluded from this study.¹⁸

Non-appendicitis No recurrence of abdominal pain or diagnosis/treatment for appendicitis by 30 days after initial presentation without any specific treatment for appendicitis. Readmission was checked for the follow-up of all cases of non-appendicitis. Telephone follow-up was not performed and we included any patient that did not subsequently return to our hospital for re-evaluation as non-appendicitis. Children with negative appendectomy were classified as non-appendicitis as well.

CPR scores were only calculated if all of the required variables were included in the patients' records. CPRs were excluded from the analysis if the score could be calculated in less than 50% of the patients in the cohort. Cutoff values to rule out appendicitis, as presented in the original manuscript, were used to calculate the performance of the CPRs. When several cutoff values were reported, patients with a negative test result according to the lower original cutoff value of the CPRs were classified as low suspicion of appendicitis.

Data Analysis

IBM SPSS statistics version 22.0 was used for descriptive analysis of our data. The likelihood ratio of a negative test with its 95% CI was calculated for each CPR and displayed as value with 95% CI. A CPR with a value < 0.1 is considered as adequate to rule out appendicitis.¹⁹

Secondary outcomes in terms of sensitivity and negative predictive value are displayed as % with 95% CI. Performed imaging studies are displayed as numbers and percentages.

Results

Identification of CPRs: Literature Review

In total, 19 CPRs were identified, of which seven were excluded (Fig. 1). Reasons for exclusion were multiplication of

variables ($n = 2$),^{20,21} consisting of > 15 variables ($n = 1$),²² not applicable in children ($n = 1$),²³ and included variables not obtained routinely ($n = 3$).^{24–26} Variables that were not routinely obtained in our hospital were a priori suspicion of appendicitis (low, intermediate, high), rectal-axillary temperature difference, and classification of rebound tenderness into light, medium, and strong.

The following 12 CPRs were included: the Alvarado score, Christian score, Fenyö score, Lintula score, Low Risk Appendicitis Rule (LRAR), Low Risk Appendicitis Rule Refinement (LRARR), Modified Alvarado Scoring System (MASS), Modified Alvarado score by Shera (MAS-Shera), modified Lindberg score, Ohmann score, Pediatric Appendicitis score (PAS), and Raja Isteri Penigran Anak Saleha Appendicitis (RIPASA) score.^{9,27–37} (Appendix 3).

Results of the CPRs Retrospectively Tested in our Cohort

311 patients were identified in the defined time period of which 20 were excluded for the following reasons: abdominal pain caused by trauma (13 patients), presentation of another main complaint other than abdominal pain (four patients), transfer to an academic hospital (two patients), and no cooperation with physical examination (one patient).

The general characteristics of the 291 included patients are listed in Table 1. In total, 87 (29.9%) patients were diagnosed with acute appendicitis.

Table 2 shows the number of patients (%) for which each CPR could be calculated. The RIPASA, modified Lindberg, and Fenyö score were excluded from further analysis as less than 50% of patients' scores could be calculated mainly due to missing data. Aggravation with cough, progression of pain, and Rovsing's sign were the variables with most frequently missing values for the Fenyö score, modified Lindberg, and RIPASA score, respectively. The negative likelihood ratio, sensitivity, and negative predictive value for the CPRs are presented in Table 3, which divides the CPRs into those that are developed for the pediatric population and those for the adult population. The point estimate of the negative likelihood ratio of seven CPRs was < 0.1. These were the Ohmann score (0), Alvarado score (0.03, 95% CI, 0.00–0.20), MAS-Shera (0.03, 95% C, 0.00–0.23), PAS (0.07, 95% CI, 0.00–0.22), LRARR (0.07, 95% CI, 0.02–0.23), Christian score (0.08, 95% CI, 0.00–0.22), and LRAR (0.09, 95% CI, 0.04–0.25).

Table 4 presents numbers of patients with low suspicion of appendicitis according to each of the seven CPRs with a negative likelihood ratio < 0.1 in whom additional imaging was performed. In 30–46% of these patients, additional imaging studies had been performed during diagnostic work-up to exclude appendicitis. Nine patients had a false negative test result according to at least one of these CPRs and were

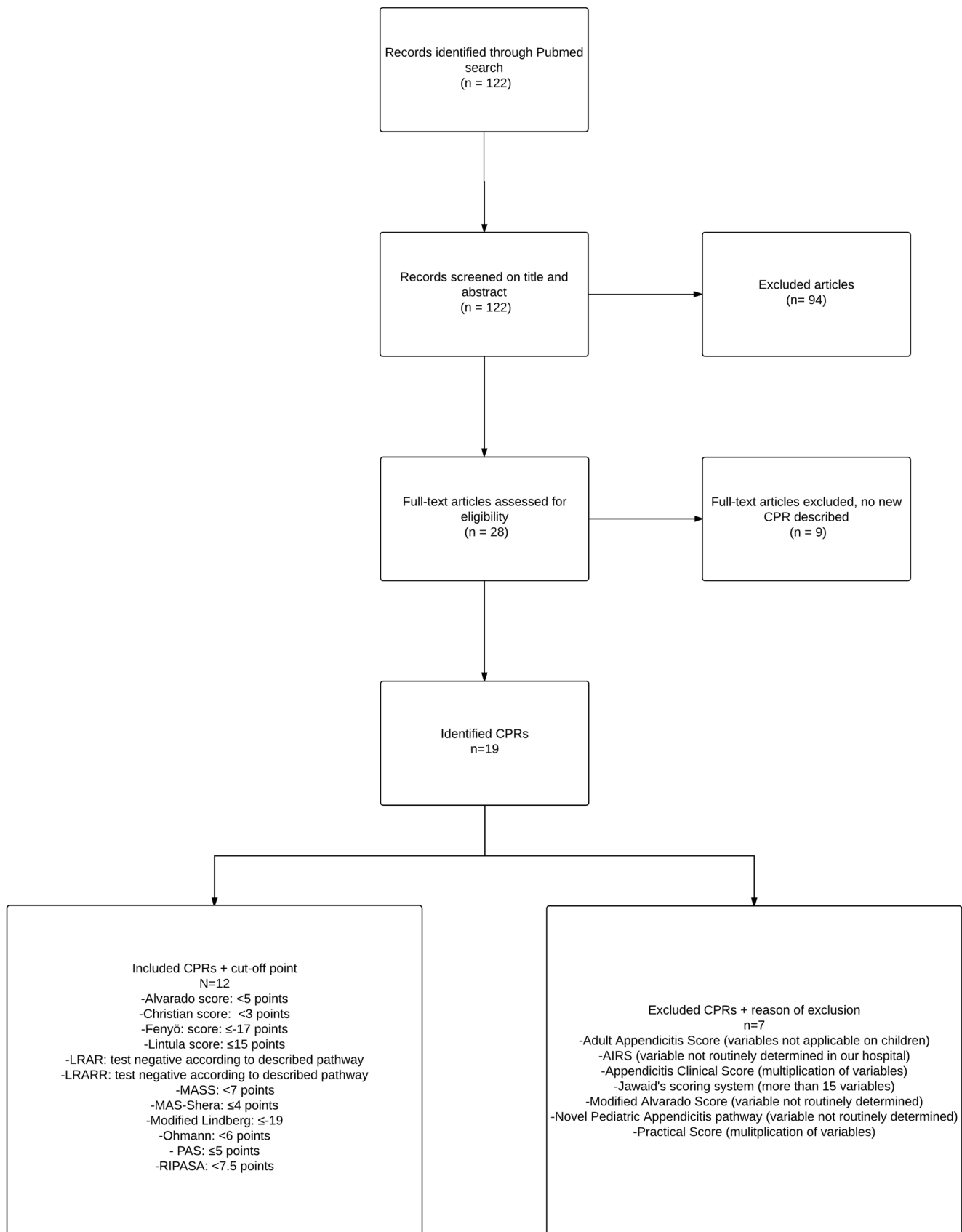


Fig. 1 Flowchart PubMed search

Table 1 General characteristics

Characteristic	Appendicitis	Non-appendicitis
Number of patients	87	204
Gender (male)	49 (56.3)	99 (48.5)
Age	10.8 ± 3.1*	9.9 ± 3.6*
Diagnosis		
Appendicitis	87 (100)	
Other surgical diagnosis		
Intussusception		1 (0.5)
Cholecystitis		1 (0.5)
Testicular torsion		1 (0.5)
Small bowel obstruction		1 (0.5)
Crohn’s disease		1 (0.5)
Follicular cyst		3 (1.5)
Other non-surgical diagnosis		
General abdominal pain		115 (56.4)
Gastro-enteritis		27 (13.2)
Mesenteric lymphadenitis		20 (9.8)
Fecal impaction		19 (9.3)
Urinary tract infection		6 (2.9)
Upper respiratory tract infection		4 (2.0)
Pneumonia		3 (1.5)
Enterobius vermicularis		1 (0.5)
Otitis media		1 (0.5)
Imaging		
Ultrasound	76 (87.4)	86 (42.2)
US + CT	5 (5.7)	4 (2.0)
US + MRI	4 (4.6)	10 (4.9)
No imaging	2 (2.3)	104 (51.0)

Characteristics are shown as *N* (%)

*Mean age (standard deviation)

diagnosed with appendicitis despite a low suspicion. In all nine patients, imaging showed inflammation of the appendix.

Discussion

The aim of this study was to investigate the value of CPRs in ruling out appendicitis in our retrospective cohort in terms of negative likelihood ratio in order to select CPRs that could potentially be included in a future prospective cohort study.

In this study, seven CPRs had a negative likelihood ratio point estimate < 0.1, which therefore could impact clinical decision-making.¹⁹ Therefore, these CPRs might be used in a future prospective cohort study comparing their ability to rule out appendicitis in children presenting with abdominal pain at the emergency department. Depending on the used CPR, in no more than 4% of the patients with a low suspicion of appendicitis, appendicitis was diagnosed within 30 days. In

Table 2 Number of patients from which the CPR score could be calculated

CPR	<i>n/N</i> (%)
LRAR	262/291 (90.0)
LRARR	260/291 (89.3)
Christian	242/291 (83.2)
Lintula	219/291 (75.3)
MASS	206/291 (70.8)
Alvarado	203/291 (69.8)
PAS	199/291 (68.4)
MAS-Shera	199/291 (68.4)
Ohmann	146/291 (50.2)
RIPASA	105/291 (36.1)
Modified Lindberg	102/291 (35.1)
Fenyö	101/291 (34.7)

30–46% of patients with a low suspicion of appendicitis, additional imaging studies had been undertaken.

Only a few studies have investigated the value of CPRs in ruling out appendicitis in children and they mostly expressed this value by sensitivity. The discriminatory power of a diagnostic test can best be displayed by likelihood ratios in our opinion, as it is not influenced by disease prevalence.³⁸

Recent systematic reviews, comprising 10–12 prospective derivation and validation studies with a total of around 4000 children, investigated the Alvarado score and PAS in the pediatric population and found negative likelihood ratios for these CPRs that were similar to our results; for the Alvarado score, negative likelihood ratios between 0.03 (95% CI, 0–0.36) and 0.38 (95% CI, 0.21–0.70) were found. Regarding the PAS, negative likelihood ratios ranging between 0 and 0.27 (95% CI, 0.20–0.43) have been reported.^{10,16}

Differences in negative likelihood ratios regarding the Alvarado score in the published negative likelihood ratios might be caused by different cutoff values that were used in the systematic reviews.^{11,39} Furthermore, daily practice concerning the use of additional imaging might differ between countries.³⁹ Regarding the PAS, modest differences in negative likelihood ratio compared to the results in our study could be explained by the prospective nature of the included studies (versus our retrospective study) and by different inclusion criteria of the included population.

To our knowledge, we are the first to present negative likelihood ratios of other CPRs in addition to the Alvarado score and PAS in the same cohort. Furthermore, this study included multiple CPRs that do not incorporate extensive laboratory parameters. Multiple biochemical variables that are included in most CPRs, such as neutrophil count and leukocyte differentiation, are not routinely tested in the Netherlands when a child presents at the emergency department. Because of the identification and inclusion of both CPRs with and without extensive laboratory parameters, we were able to present a

Table 3 Accuracy statistics

CPRs	Negative likelihood ratio	Sensitivity*	Negative predictive value*
Pediatric CPRs			
LRAR (<i>n</i> = 262)	0.09 (0.04–0.25)	95.3 (87.9–98.5)	95.6 (88.5–98.6)
LRARR (<i>n</i> = 260)	0.07 (0.02–0.23)	96.5 (89.4–99.1)	96.5 (89.3–99.1)
PAS (<i>n</i> = 199)	0.07 (0.00–0.22)	95.3 (86.0–98.8)	96.7 (90.0–99.1)
Adult CPRs			
Christian (<i>n</i> = 242)	0.08 (0.00–0.22)	94.8 (86.5–98.3)	96.3 (90.2–98.8)
Lintula (<i>n</i> = 219)	0.50 (0.38–0.65)	54.0 (41.0–66.4)	83.2 (76.6–88.3)
MASS (<i>n</i> = 206)	0.30 (0.20–0.46)	73.0 (60.1–83.1)	88.2 (81.5–92.8)
Alvarado (<i>n</i> = 203)	0.03 (0.00–0.20)	98.4 (90.0–99.9)	98.8 (92.5–99.9)
MAS-Shera (<i>n</i> = 199)	0.03 (0.00–0.23)	98.4 (90.5–99.9)	98.5 (90.9–99.9)
Ohmann (<i>n</i> = 146)	0	100 (90.2–100.0)	100 (88.6–100.0)

Data is displayed as value (95% CI)

*Data is displayed as percentage (95% CI)

complete overview of all CPRs that can potentially be used in future prospective studies comparing their ability in ruling out appendicitis. In order to present a complete overview of all potential CPRs, we determined a low cutoff value of at least 50% of available data per CPR for inclusion in our analysis. We do realize that this cutoff value is low, but the aim of this study was to identify CPRs, investigate their potential in ruling out appendicitis, and investigate their appropriateness in the current diagnostic work-up as performed in the Netherlands in order to select them for a prospective cohort study. This cutoff value was determined prior to the identification of CPRs, and we realize that the use of a higher cutoff value might have led to a more stringent selection of only those CPRs that are most appropriate in our population.

The evidence-based guideline regarding the diagnosis and treatment of appendicitis, introduced by the Association of Surgeons of the Netherlands in 2010, emphasized reduction of the negative appendectomy rate. Imaging procedures are advocated to improve diagnostic accuracy and the consequence of this change has been the increased utilization of ultrasound as the initial imaging modality to evaluate abdominal pain in children in the Netherlands.⁴⁰ In 2010, preoperative imaging procedures

were performed in 44% of patients presenting at the emergency room with abdominal pain in the Netherlands, compared to only 22% of the patients a decade earlier.^{3,41} Currently in the Netherlands, in 99.7% of patients preoperative imaging is performed.⁴ A recent study conducted in the USA found that 99.7% of pediatric patients underwent preoperative-imaging studies as well.⁴² This differs significantly from the performance of preoperative imaging in the UK, where preoperative ultrasound and computed tomography (CT) were performed in 19.9 and 12.9% of patients respectively.⁴³ Ultrasound has a high frequency of inconclusive results, reported to range between 37 to 51% in the pediatric population.^{7,44} Increased performance of ultrasound therefore results in increased use of costly and potentially harmful imaging studies, such as CT and MRI in pediatric patients.

In this study, in 30–46% of patients with a low suspicion of appendicitis according to these CPRs, additional imaging studies had been undertaken, whereas in no more than 4% of these patients (depending on the used CPR) acute appendicitis was diagnosed within 30 days. Nonetheless, because of the retrospective nature of this study, it might be possible that these additional imaging studies have not been solely

Table 4 Radiological examinations in patients with low suspicion of appendicitis according to the CPR

Low suspicion of appendicitis	Children who only underwent ultrasound	Children who underwent CT after ultrasound	Children who underwent MRI after ultrasound	Total number of children with low suspicion undergoing any radiological test	Children with appendicitis (false negative test result)
Ohmann (<i>N</i> = 38)	12 (32%)	0 (0%)	0 (0%)	12 (32%)	0 (0%)
Alvarado (<i>N</i> = 83)	27 (33%)	3 (4%)	0 (0%)	30 (36%)	1 (1%)
MAS-Shera (<i>N</i> = 67)	18 (27%)	1 (1%)	1 (1%)	20 (30%)	1 (1%)
PAS (<i>N</i> = 88)	30 (34%)	3 (3%)	2 (2%)	35 (40%)	3 (3%)
LRARR (<i>N</i> = 85)	32 (38%)	0 (0%)	0 (0%)	32 (38%)	3 (3%)
Christian (<i>N</i> = 108)	45 (42%)	3 (3%)	2 (2%)	50 (46%)	4 (4%)
LRAR (<i>N</i> = 91)	38 (42%)	2 (2%)	0 (0%)	40 (44%)	4 (4%)

Data is displayed as value (% of total patients with low suspicion of appendicitis per CPR)

performed to diagnose appendicitis, but also to exclude other potential diagnoses. Still, it raises the question whether or not watchful waiting should be considered for children with a low suspicion of appendicitis instead of additional imaging studies to rule out appendicitis. Opponents of this less aggressive diagnostic work-up mostly fear perforation of the appendix in case of complicated appendicitis.⁴⁵ However, several studies have not found clinical observation or re-evaluation to be associated with a significantly higher incidence of complicated appendicitis and perforation.^{46,47} Time to presentation at the emergency department appears to be the main factor associated with perforation in children with appendicitis.^{46,48} Furthermore, literature suggests that perforation can rarely be prevented, implicating that a correct diagnosis is more important than a rapid treatment strategy.⁴⁹

This study has several limitations. First, due to the single-center nature, generalizability might be reduced, although it was performed in a general teaching hospital. Second, the retrospective nature of this study might have led to selection bias and information bias. In case of wrong ICD code classification, patients might have been missed. We do realize that inclusion of 291 patients in 3 years' time seems to be low for a large teaching hospital. This low number of patients could be explained by the fact that we only used ICD codes of acute appendicitis, acute abdomen, and general abdominal pain, because inclusion of children presenting with, for example, mainly symptoms of urinary tract infection would artificially decrease negative likelihood ratios and overestimate the value of the CPRs. Patient files with missing data were left out of the analysis. As a result, for only 34–90% of the patients the CPRs could be calculated. However, the aim of this study was not only to determine the ability of CPRs to rule out appendicitis in our cohort but also to investigate their appropriateness within the current diagnostic work-up as performed in the Netherlands. As mentioned previously, this study was performed in order to select appropriate and useful CPRs for a future prospective cohort study that can compare their value in ruling out appendicitis in our cohort. Nonetheless, missing data could have led to a selection bias, whereby the results of a CPR may have been inflated, leading to a low negative likelihood ratio. For example, in our population the CPR with the optimal negative likelihood ratio (Ohmann score) could only be calculated for 50.2% of the total population. Third, the small sample size causes wide confidence intervals for the calculated accuracy statistics.

In addition, CPRs are prone to subjective interpretation by the treating physician (e.g., variables of physical examination). In a prospective study by Mandeville, an interobserver agreement was found of 88 and 83.5% for the Alvarado score and the PAS, respectively.⁴⁸ Another problem is the potential partial verification bias. Patients were classified in the non-appendicitis group if there was no recurrence of abdominal pain during 30 days after initial presentation. Patients who went to another hospital during these 30 days could have been

missed. On the other hand, there is no nearby facility that is comparable to the general teaching hospital where this study was conducted. Therefore, it can be expected that patients attend the same emergency department as during their initial presentation. Another issue is that a number of patients were included in the appendicitis group despite the fact that intraoperative or histopathological findings were not obtained, potentially also leading to misclassification. However, these patients were included in the APAC study, in which a radiologically confirmed simple appendicitis was an inclusion criterion.

In conclusion, we identified seven CPRs that could potentially be used in a future prospective cohort study to compare their ability to rule out appendicitis in the pediatric population in the Netherlands and other countries with comparable diagnostic work-up. Further prospective studies are needed to investigate if imaging studies could safely be omitted and be replaced by clinical monitoring or re-evaluation in children with a low CPR score.

Author Contributions PA: Study concept and design; acquisition of the data, analysis, and interpretation of the data; and drafting of the manuscript.

RG: Study concept and design, analysis and interpretation of the data, drafting of the manuscript, and critical revision of the manuscript.

JL: Critical revision of the manuscript for important intellectual content and statistical expertise.

HC: Drafting of the manuscript and critical revision of the manuscript for important intellectual content.

RB: Drafting of the manuscript and critical revision of the manuscript for important intellectual content.

HH: Drafting of the manuscript and critical revision of the manuscript for important intellectual content.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflicts of interest.

Appendix 1

Pubmed database search

((“Appendix”[Mesh] OR “Appendicitis”[Mesh] OR “Appendectomy”[Mesh] OR appendix[tiab] OR appendic*[tiab] OR appendec*[tiab]) AND (Clinical prediction rule*[tiab] OR prediction rule*[tiab] OR scoring system*[tiab])) AND (child*[tw] OR schoolchild*[tw] OR infan*[tw] OR adolescen*[tw] OR pediatri*[tw] OR paediatr*[tw] OR neonat*[tw] OR boy[tw] OR boys[tw] OR boyhood[tw] OR girl[tw] OR girls[tw] OR girlhood[tw] OR youth[tw] OR youths[tw] OR baby[tw] OR babies[tw] OR toddler*[tw] OR teen[tw] OR teens[tw] OR teenager*[tw] OR newborn*[tw] OR postneonat*[tw] OR postnat*[tw] OR perinat*[tw] OR puberty[tw] OR preschool*[tw] OR suckling*[tw] OR picu[tw] OR nicu[tw])

Appendix 2 Data extraction form

CPRA

Studynumber:
Date of presentation:

Gender: male / female
Dutch residence: yes / no

HISTORY				PHYSICAL EXAMINATION			
Code	Variables	Yes	No	Code	Variables	Yes	No
1	Hours of abdominal pain:			1	Coughing tenderness (RLQ):		
2	Right lower quadrant pain:			2	Percussion tenderness (RLQ):		
3	Migration of pain:			3	Rebound tenderness (RLQ):		
4	Anorexia:			4	Guarding:		
5	Nausea:			5	Rovsing's sign:		
6	Vomiting:			6	Bowel sounds absent, tinkling or high pitched:		
7	Intensity of pain severe (VAS) ¹ :			7	Rigidity:		
8	Progression of pain:			8	Temperature: °C		
9	Steady pain:			9	Tenderness in RLQ:		
10	Hopping tenderness:			10	Tenderness outside RLQ:		
11	Dysuria/frequent miction:			LABORATORY RESULTS			
				Code	Variables	Yes	No
				1	Number of Leukocytes:		
				2	Leukocyten differentiation: Neutrophils: Eosinophils: Basophils: Lymfocytes: Monocytes: Shift to the left:		
				4	CRP:		
				5	Acetone in urine:		
				6	Negative urinalysis:		

1. Visual Analogue scale

ADDITIONAL IMAGING	
Ultrasound abdomen:	
CTabdomen:	
MRlabdomen:	

Appendix 3 Clinical prediction rules

Alvarado score (cutoff value < 5: no appendicitis)

Variable	Points
Migration	1
Anorexia/acetone	1
Nausea/vomiting	1
Tenderness in right lower quadrant	2
Rebound pain	1
Elevation of temperature	1
Leukocytosis	2
Shift to the left	1
Total	10

Christian score (cutoff value < 3: no appendicitis)

Variable	Points
Abdominal pain	1
Vomiting	1
Right lower quadrant tenderness	1
Low grade fever (≤ 38.8 °C)	1
Polymorphonuclear leucocytosis (TC $\geq 10,000$ with polymorphs $\geq 75\%$)	1
Total	5

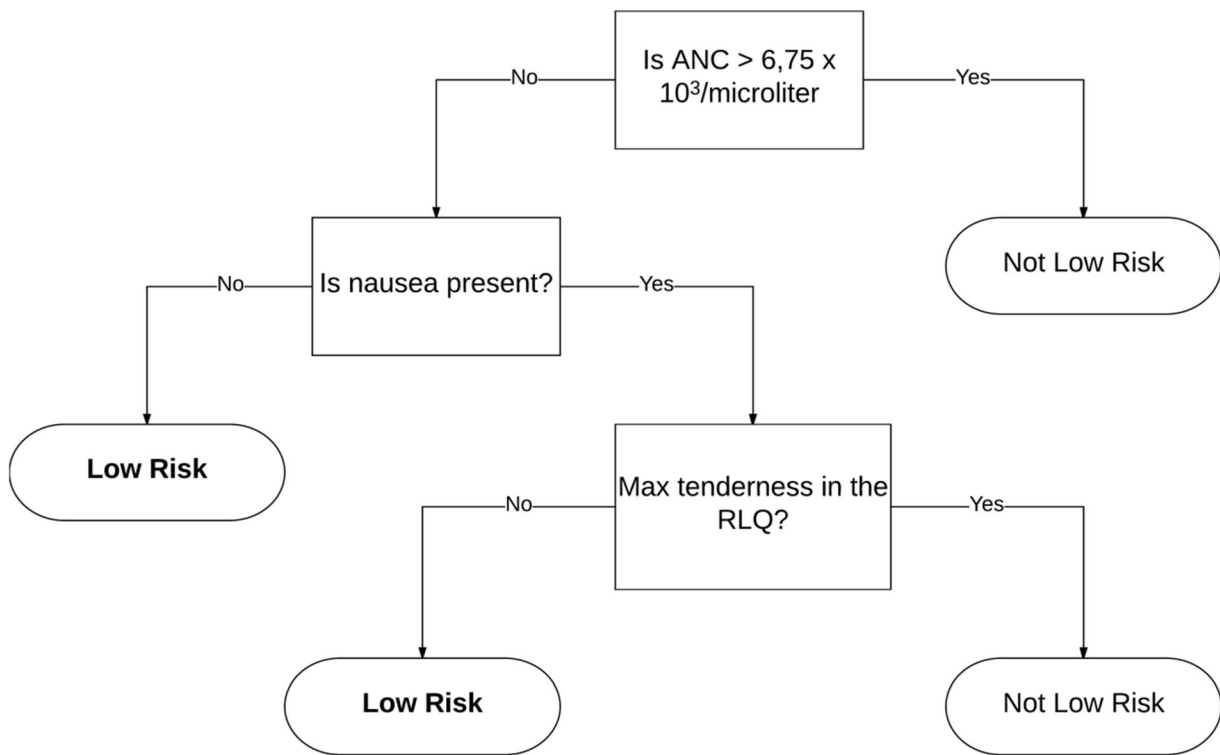
Fenyő score (cutoff value ≤ -17 : no appendicitis)

Variable	Points
Constant	-10
Male	8
Female	-8
White blood cell count < 8.9	-15
White blood cell count 8.9–13.9	2
White blood cell count > 14.0	10
Pain duration < 24 h	3
Pain duration 24–48 h	0
Pain duration > 48 h	-12
Progression of pain	3
No progression of pain	-4
Relocation of pain	7
No relocation of pain	-9
Vomiting	7
No vomiting	-5
Aggravation with cough	4
No aggravation with cough	-11
Rebound tenderness	5
No rebound tenderness	-10
Rigidity	15
No rigidity	-4
Tenderness outside right lower quadrant	-6
No tenderness outside right lower quadrant	4
Total	56

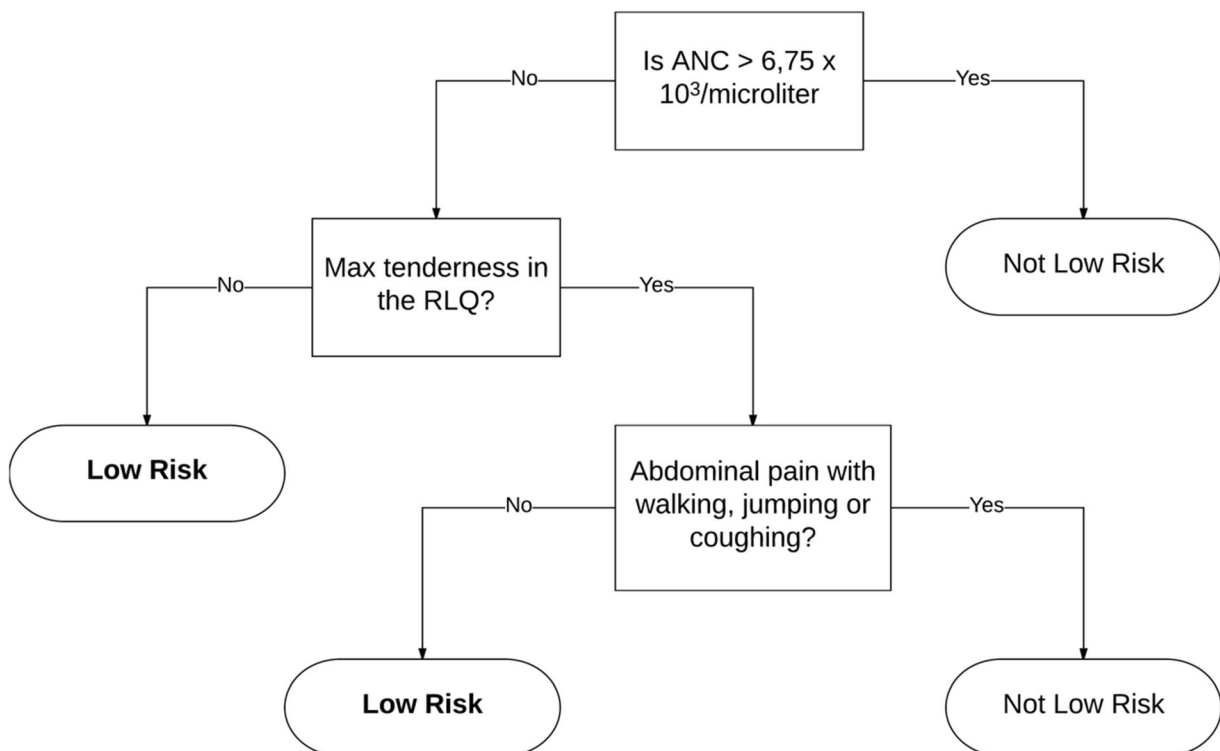
Lintula score (cutoff value ≤ 15 : no appendicitis)

Variable	Points
Male	2
Female	0
Intensity of pain severe	2
Intensity of pain mild or moderate	0
Relocation of pain	4
Pain in the right-lower abdominal quadrant	4
Vomiting	2
Body temperature ≥ 37.5 °C	3
Guarding	4
Bowel sounds absent, tinkling, or high pitched	4
Rebound tenderness	7
Total	32

Low-Risk Appendicitis Rule



Low-Risk Appendicitis Rule Refinement.



Modified Alvarado Scoring System (MASS) (Cutoff value < 7: no appendicitis)

Variable	Points
Migratory right iliac fossa pain	1
Nausea/vomiting	1
Anorexia	1
Tenderness in right iliac fossa	2
Rebound tenderness in right iliac fossa	1
Elevated temperature	1
Leukocytosis	2
Total	9

Modified Lindberg score (cutoff value ≤ -19: no appendicitis)

Variable	Points
Constant	-10
Male	8
Female	-8
White blood cell count < 12.0	-15
White blood cell count 12.0–20.0	2
White blood cell count 20.0	10
Pain duration < 24 h	3
Pain duration 24–48 h	0
Pain duration > 48 h	-12
Progression of pain	3
No progression of pain	-4
Temperature ≥ 37.5 °C	7
Temperature < 37.5 °C	-9
Vomiting	7
No vomiting	-5
Migration of pain	4
No migration of pain	-11
Rebound tenderness	5
No rebound tenderness	-10
Rigidity	15
No rigidity	-4
Tenderness outside right lower quadrant	-6
No tenderness outside right lower quadrant	4
Total	56

Ohmann score (cutoff value < 6: no appendicitis)

Variable	Points
Tenderness, right lower quadrant	4.5
Rebound tenderness	2.5
No micturition difficulties	2.0
Steady pain	2.0
Leukocyte count ≥ 10.0 × 10 ⁹ /L	1.5
Age < 50 years	1.5
Relocation of pain to right lower quadrant	1.0
Rigidity	1.0
Total	16.0

Pediatric Appendicitis Score (PAS) (cutoff value ≤ 5: no appendicitis)

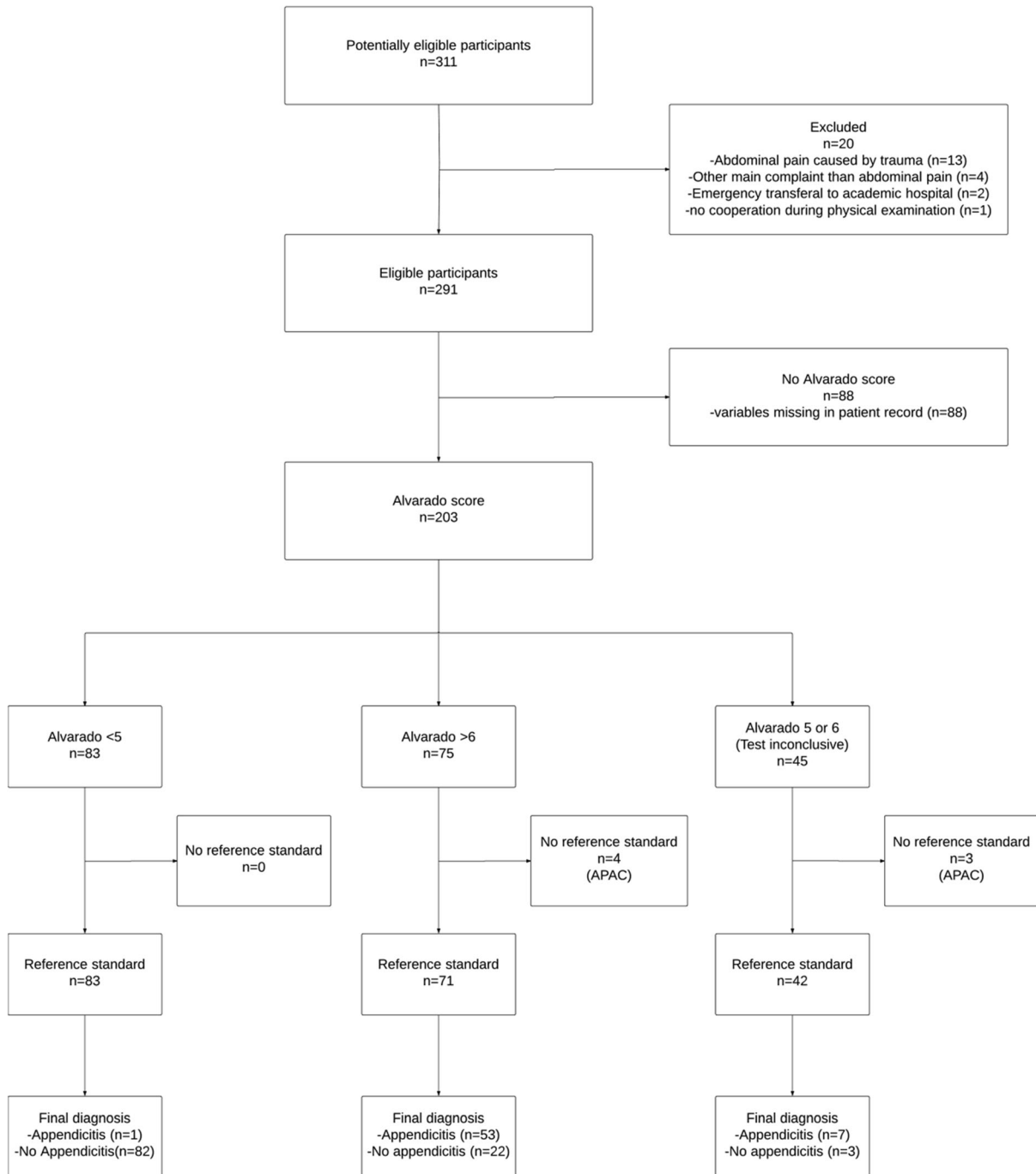
Variable	Points
Cough/percussion/hopping tenderness	2
Anorexia	1
Pyrexia	1
Nausea/emesis	1
Tenderness in right lower quadrant	2
Leukocytosis	1
Polymorphonuclear neutrophilia	1
Migration of pain	1
Total	10

RIPASA score (cutoff value < 7.5: no appendicitis)

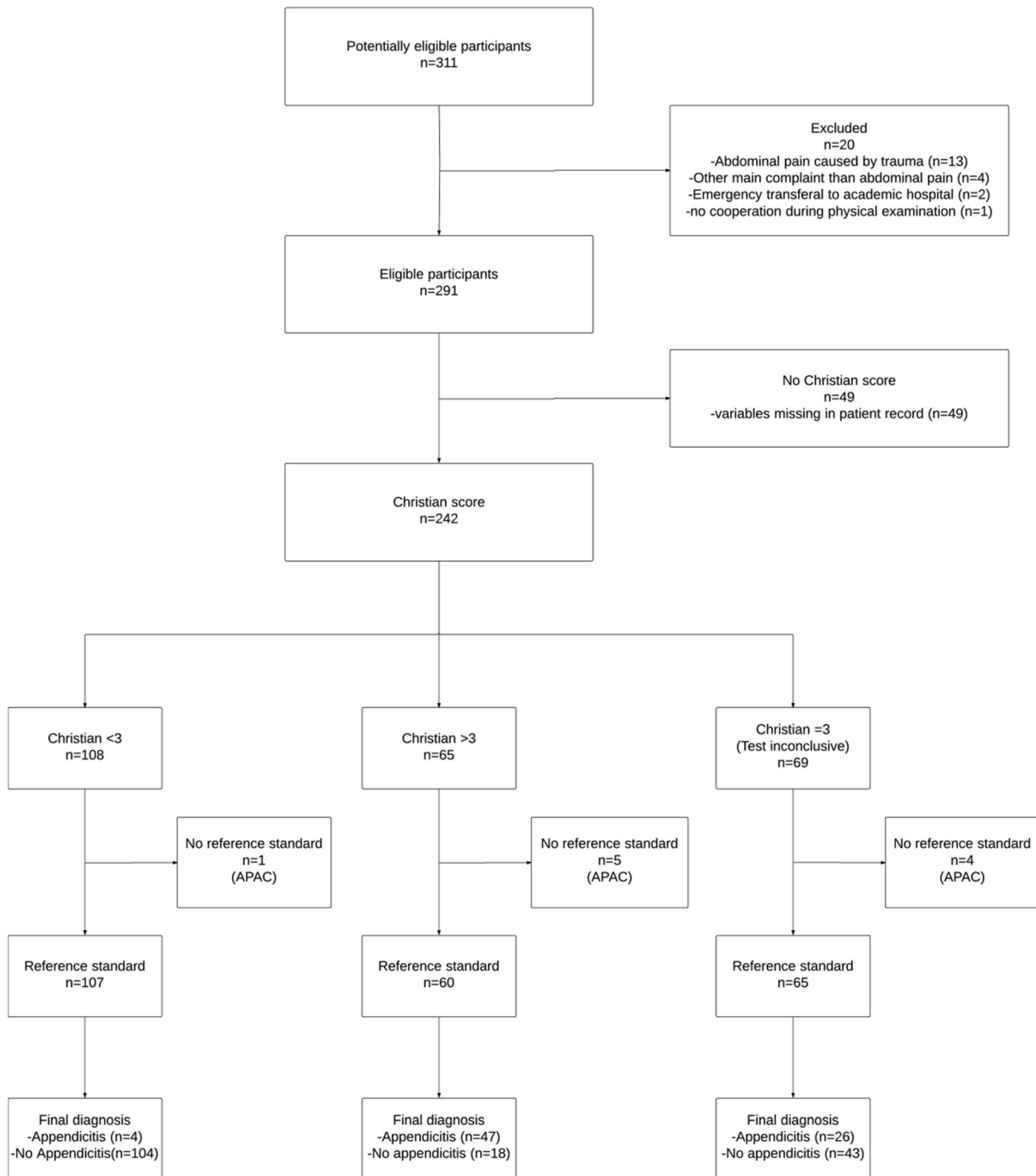
Variable	Points
Age < 40 years	1
Age > 40 years	0.5
Male	1
Female	0.5
Right iliac fossa pain	0.5
Migration of pain	0.5
Nausea and vomiting	1
Anorexia	1
Duration of symptoms < 48 h	1
Duration of symptoms > 48 h	0.5
Right iliac fossa tenderness	1
Guarding	2
Rebound tenderness	1
Rovsing's sign	2
Fever	1
Raised white cell count	1
Negative urinalysis	1
Foreign national registration identity card	1
Total	16

Appendix 4 Flow of participants for the CPRs

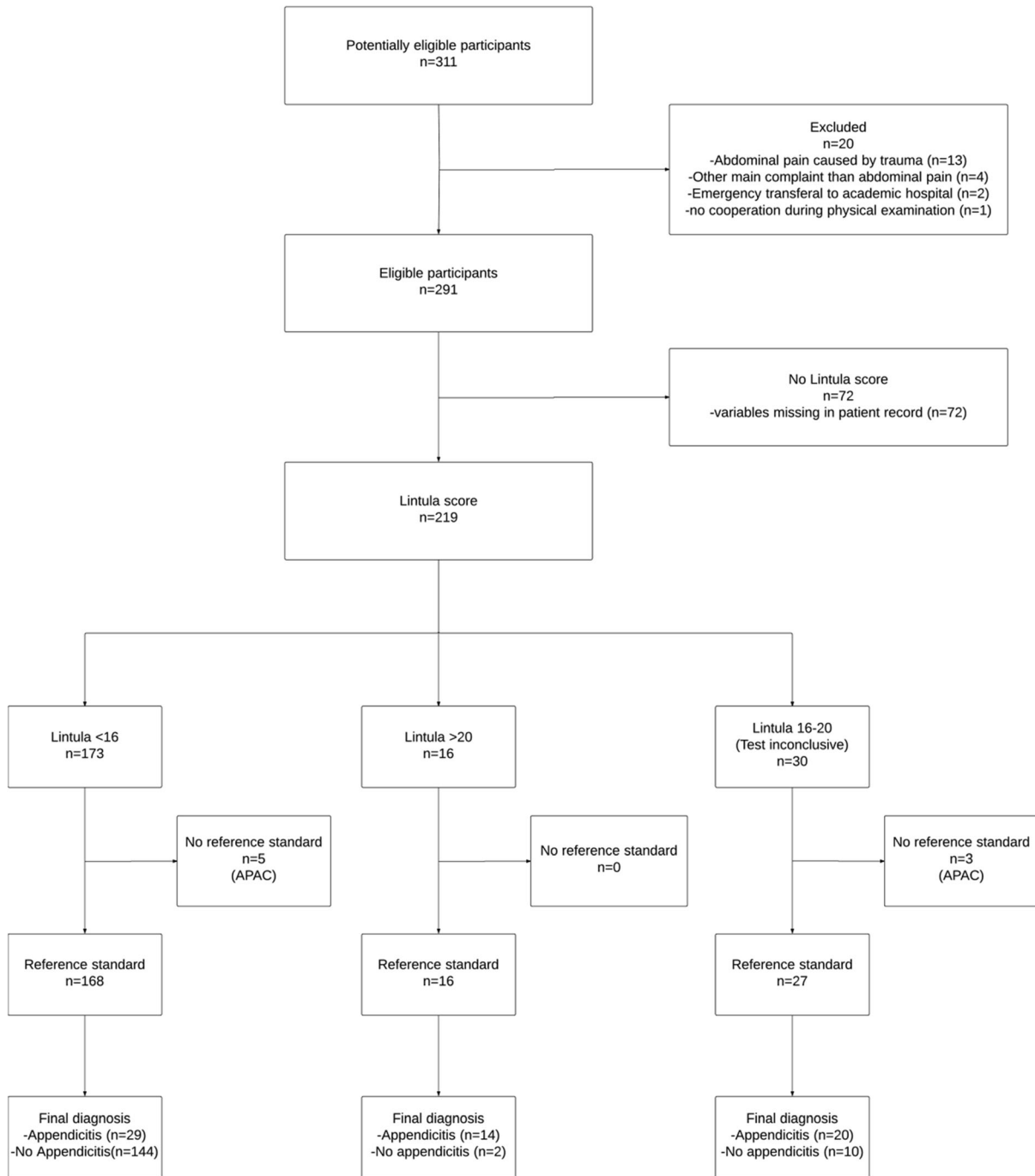
Alvarado score



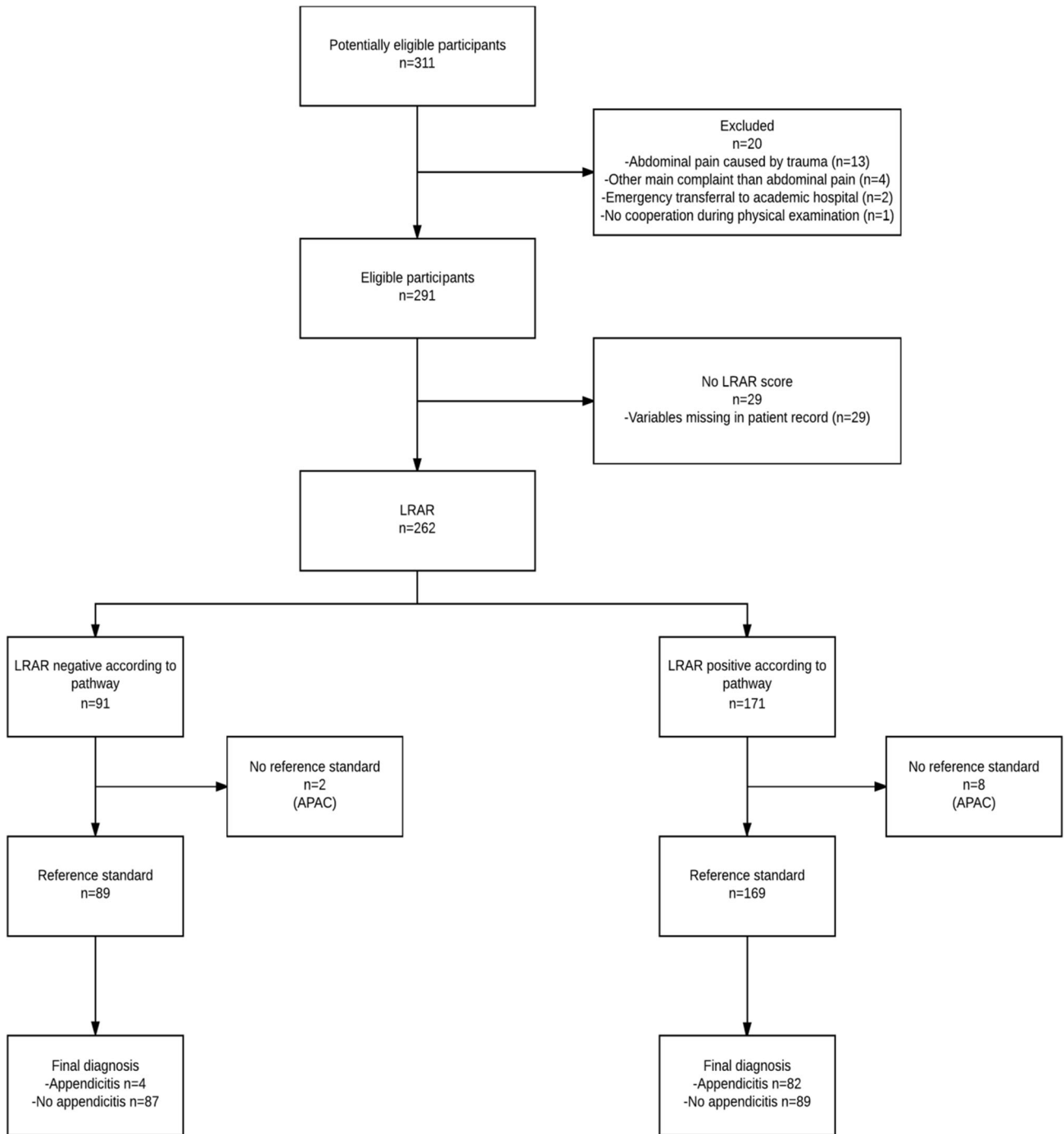
Christian score



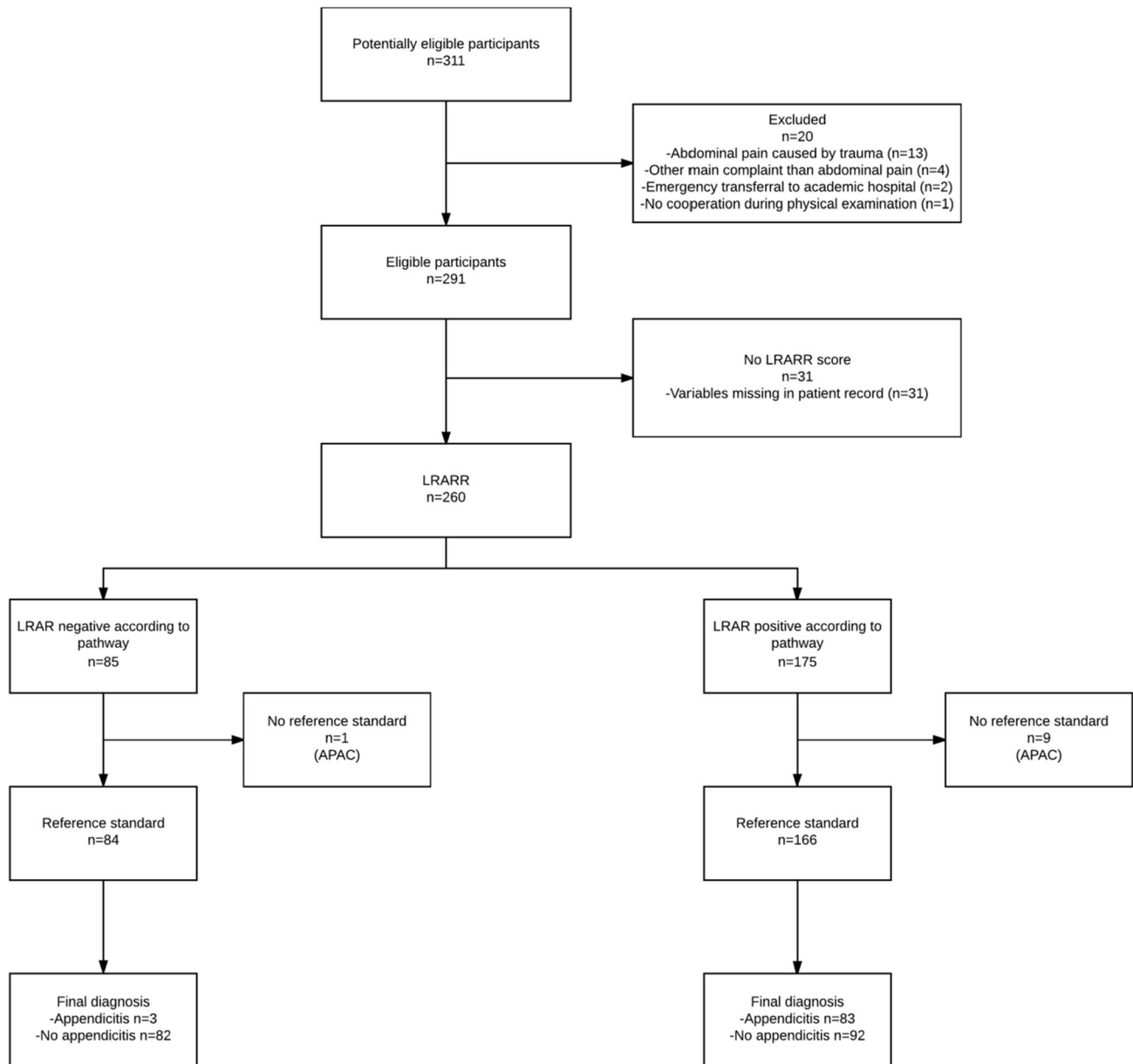
Lintula score



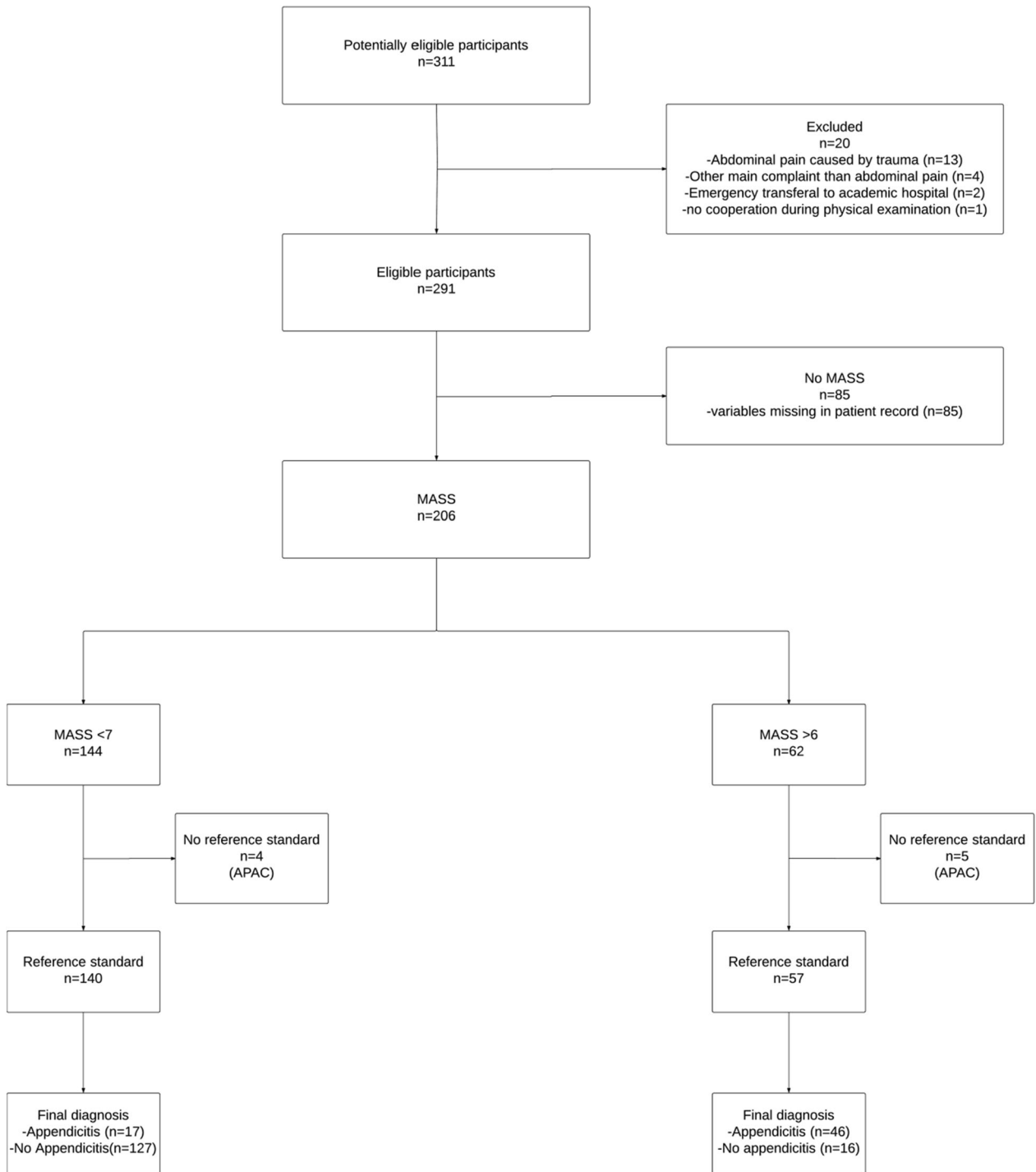
Low-Risk Appendicitis Rule



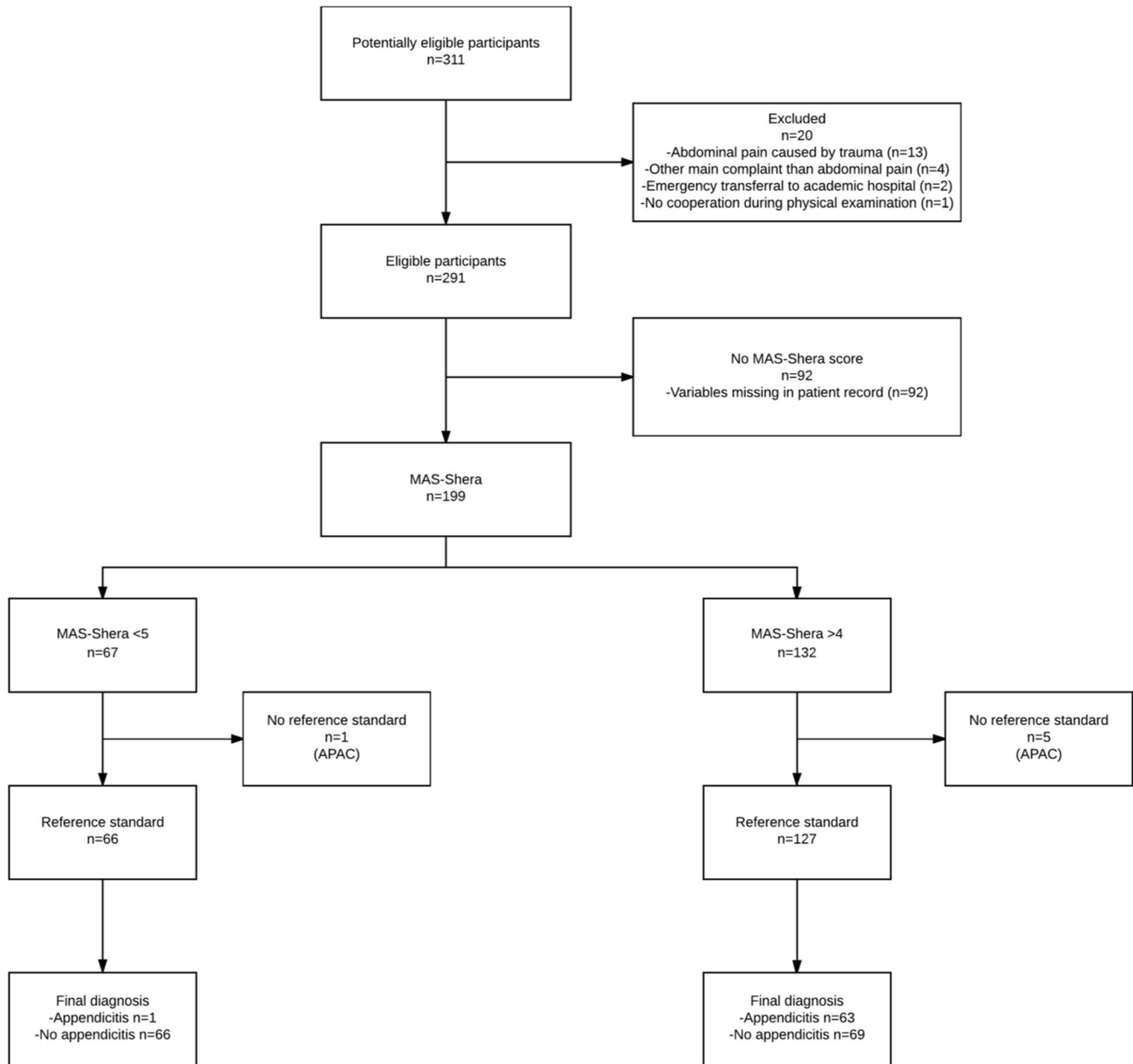
Low-Risk Appendicitis Rule Refinement



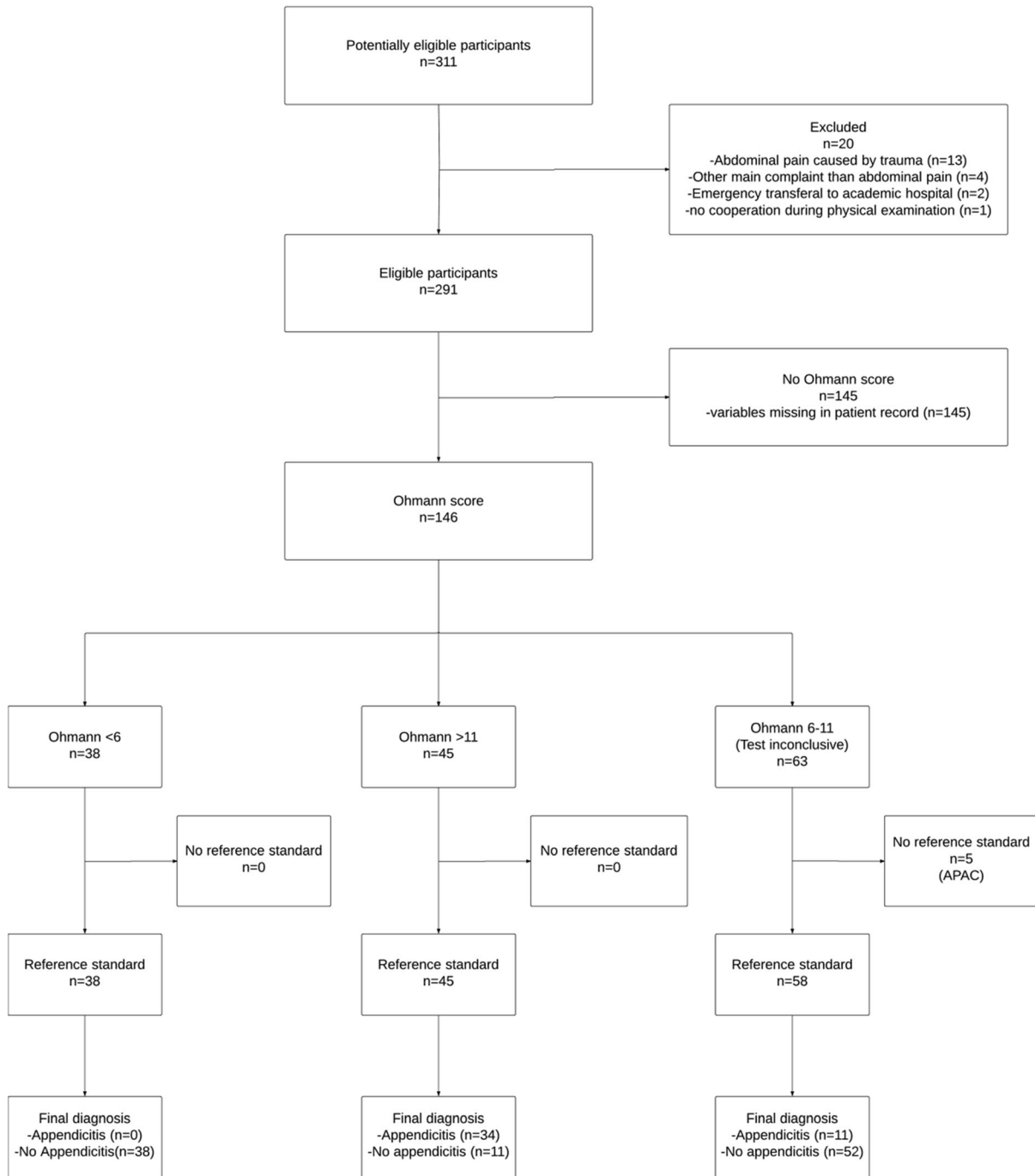
Modified Alvarado Scoring System



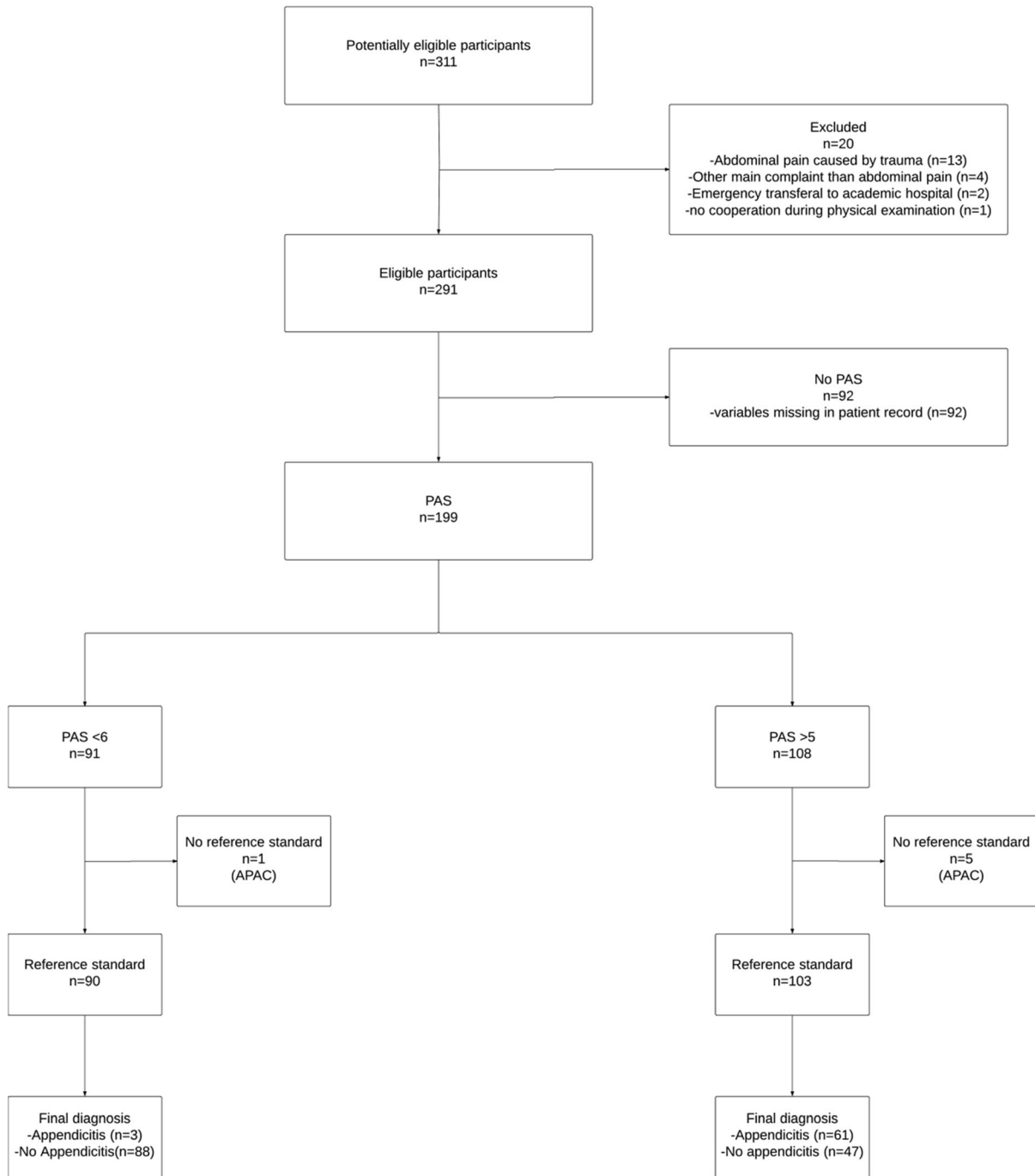
Modified Alvarado Score by Shera



Ohmann score



Pediatric Appendicitis Score



Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Warner BW, Kulick RM, Stoops MM, Mehta S, Stephan M, Kotagal UR. An evidenced-based clinical pathway for acute appendicitis decreases hospital duration and cost. *J Pediatr Surg.* 1998;33(9):1371–1375. [https://doi.org/10.1016/S0022-3468\(98\)90010-0](https://doi.org/10.1016/S0022-3468(98)90010-0).
- Boonstra PA, van Veen RN, Stockmann HB. Less negative appendectomies due to imaging in patients with suspected appendicitis. *Surg Endosc.* 2015;29(8):2365–2370. <https://doi.org/10.1007/s00464-014-3963-2>.
- Bakker OJ, Go PMNYH, Puylaert JBCM, Kazemier G, Heij HA. Guideline on diagnosis and treatment of acute appendicitis: imaging prior to appendectomy is recommended. *Ned Tijdschr Geneeskd.* 2010;154:A303.
- Van Rossem CC, Bolmers MDM, Schreinemacher MHF, Van Geloven AAW. Prospective nationwide outcome audit of surgery for suspected acute appendicitis. *Br J Surg.* 2016;103:144–151. <https://doi.org/10.1002/bjs.9964>.
- Blitman NM, Anwar M, Brady KB, Taragin BH, Freeman K. Value of focused appendicitis ultrasound and alvarado score in predicting appendicitis in children: Can we reduce the use of CT? *Am J Roentgenol.* 2015;204(6):W707–W712. <https://doi.org/10.2214/AJR.14.13212>.
- Schok T, Simons PCG, Janssen-Heijnen MLG, Peters NALR, Konsten JLM. Prospective evaluation of the added value of imaging within the dutch national diagnostic appendicitis guideline - Do we forget our clinical eye? *Dig Surg.* 2014;31(6):436–443. <https://doi.org/10.1159/000369587>.
- Mittal MK, Dayan PS, Macias CG, Bachur RG, Bennett J, Dudley NC, Bajaj L, Sinclair K, Stevenson MD, Kharbanda AB. Performance of ultrasound in the diagnosis of appendicitis in children in a multicenter cohort. *Acad Emerg Med.* 2013;20(7):697–702. <https://doi.org/10.1111/acem.12161>.
- Stewart JK, Olcott EW, Jeffrey BR. Sonography for appendicitis: Nonvisualization of the appendix is an indication for active clinical observation rather than direct referral for computed tomography. *J Clin Ultrasound.* 2012;40(8):455–461. <https://doi.org/10.1002/jcu.21928>.
- Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* 1986;15(5):557–564.
- Ebell MH, Shinholser J. What Are the Most Clinically Useful Cutoffs for the Alvarado and Pediatric Appendicitis Scores? A Systematic Review. *Ann Emerg Med.* 2014;64(4):365–372.e2. <https://doi.org/10.1016/j.annemergmed.2014.02.025>.
- Schneider C, Kharbanda A, Bachur R. Evaluating Appendicitis Scoring Systems Using a Prospective Pediatric Cohort. *Ann Emerg Med.* 2007;49(6):778–84. <https://doi.org/10.1016/j.annemergmed.2006.12.016>.
- Mandeville K, Pottker T, Bulloch B, Liu J. Using appendicitis scores in the pediatric ED. *Am J Emerg Med.* 2011;29(9):972–977. <https://doi.org/10.1016/j.ajem.2010.04.018>.
- Goldman RD, Carter S, Stephens D, Antoon R, Mounstephen W, Langer JC. Prospective Validation of the Pediatric Appendicitis Score. *J Pediatr.* 2008;153(2):278–282. <https://doi.org/10.1016/j.jpeds.2008.01.033>.
- Bhatt M, Joseph L, Ducharme FM, Dougherty G, McGillivray D. Prospective validation of the pediatric appendicitis score in a Canadian Pediatric Emergency Department. *Acad Emerg Med.* 2009;16(7):591–596. <https://doi.org/10.1111/j.1553-2712.2009.00445.x>.
- Zúñiga RV, Arribas JLF, Montes SP, Fernandez MNC, Abad CG, Martin LG, Gonzalez-Sagrado M. Application of Pediatric Appendicitis Score on the Emergency Department of a Secondary Level Hospital. *Pediatr Emerg Care.* 2012;28(6):489–492. <https://doi.org/10.1097/PEC.0b013e3182586d34>.
- Kulik DM, Uleryk EM, Maguire JL. Does this child have appendicitis? A systematic review of clinical prediction rules for children with acute abdominal pain. *J Clin Epidemiol.* 2013;66(1):95–104. <https://doi.org/10.1016/j.jclinepi.2012.09.004>.
- Moher D, Liberati A, Tetzlaff J, Altman DG. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Ann Intern Med.* 2009;151(4):264–269. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>.
- Gorter RR, van der Lee JH, Cense HA, Kneepkens CMF, Wijnen MHWA, In 't Hof KH, Offringa M, Heij HA. Initial antibiotic treatment for acute simple appendicitis in children is safe: Short-term results from a multicenter, prospective cohort study. *Surgery.* 2015;157(5):916–923. <https://doi.org/10.1016/j.surg.2015.01.008>.
- Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA.* 1994;271(5):389–391. <https://doi.org/10.1186/1752-1505-2-9>.
- Kharbanda AB, Monuteaux MC, Bachur RG, Dudley NC, Bajaj L, Stevenson MD, Macias CG, Mittal MK, Bennett JE, Sinclair K, Dayan PS. A Clinical Score to Predict Appendicitis in Older Male Children. *Acad Pediatr.* 2017;17(3):261–266. <https://doi.org/10.1016/j.acap.2016.11.014>.
- Ramirez JM, Deus J. Practical score to aid decision making in doubtful cases of appendicitis. *Br J Surg.* 1994;81(5):680–683. <https://doi.org/10.1002/bjs.1800810516>.
- Jawaid A, Asad A, Motie A, Munir A, Bhutto E, Choudry H, Idrees K, Durrani K, Rahman M, Ahuja M, Nawab Q, Ahmed R, Ali S, Aslam S, Abbasi S, Feerasta S, Alam S, Ahmed U, Jehan I. Clinical scoring system: a valuable tool for decision making in cases of acute appendicitis. *J Pak Med Assoc.* 1999;49(10):254–259.
- Sammalkorpi HE, Mentula P, Leppäniemi A. A new adult appendicitis score improves diagnostic accuracy of acute appendicitis—a prospective study. *BMC Gastroenterol.* 2014;14(1):114. <https://doi.org/10.1186/1471-230X-14-114>.
- Fleischman RJ, Devine MK, Yagapen M-AN, Zigman A, Spiro DM. Evaluation of a novel pediatric appendicitis pathway using high- and low-risk scoring systems. *Pediatr Emerg Care.* 2013;29(10):1060–1065. <https://doi.org/10.1097/PEC.0b013e3182a5c9b6>.
- Mán E, Simonka Z, Varga Á, Ráosi F, Lázár G. Impact of the alvarado score on the diagnosis of acute appendicitis: Comparing clinical judgment, alvarado score, and a new modified score in suspected appendicitis: A prospective, randomized clinical trial. *Surg Endosc Other Interv Tech.* 2014;28(8):2398–2405. <https://doi.org/10.1007/s00464-014-3488-8>.
- Andersson M, Andersson RE. The Appendicitis Inflammatory Response Score: A Tool for the Diagnosis of Acute Appendicitis that Outperforms the Alvarado Score. *World J Surg.* 2008;32(8):1843–1849. <https://doi.org/10.1007/s00268-008-9649-y>.

27. Christian F, Christian GP, Grogono JL. A simple scoring system to reduce the negative appendectomy rate. *Ann R Coll Surg Engl*. 1992;74(4):281–285.
28. Fenyo G, Lindberg G, Blind P, Enochsson L, Oberg A. Diagnostic Decision Support in Suspected Acute Appendicitis: Validation of a Simplified Scoring System. *Eur J Surg*. 1997;831–838.
29. Lintula H, Kokki H, Kettunen R, Eskelinen M. Appendicitis score for children with suspected appendicitis. A randomized clinical trial. *Langenbeck's Arch Surg*. 2009;394(6):999–1004. <https://doi.org/10.1007/s00423-008-0425-0>.
30. Kharbanda AB, Taylor GA, Fishman SJ, Bachur RG. A Clinical Decision Rule to Identify Children at Low Risk for for Appendicitis. *Pediatrics*. 2005;116(3):709–716. <https://doi.org/10.1542/peds.2005-0094>.
31. Kalan M, Talbot D, Cunliffe W, Rich A. Evaluation of the modified Alvarado score in the diagnosis of acute appendicitis: a prospective study. *Ann R Coll Surg Engl*. 1994;76(6):418–419.
32. Shera AH, Nizami FA, Malik AA, Naikoo ZA, Wani MA. Clinical scoring system for diagnosis of acute appendicitis in children. *Indian J Pediatr*. 2011;78(3):287–290. <https://doi.org/10.1007/s12098-010-0285-9>.
33. Dado G, Anania G, Baccarani U, Marcotti E, Donini A, Risaliti A, Pasqualucci A, Bresadola F. Application of a clinical score for the diagnosis of acute appendicitis in childhood: A retrospective analysis of 197 patients. *J Pediatr Surg*. 2000;35(9):1320–1322. <https://doi.org/10.1053/jpsu.2000.9316>.
34. Ohmann C, Franke C, Yang Q. Clinical Benefit of a Diagnostic Score for Appendicitis. *Arch Surg*. 1999;134:993–996.
35. Samuel M. Pediatric appendicitis score. *J Pediatr Surg*. 2002;37(6):877–881. <https://doi.org/10.1053/jpsu.2002.32893>.
36. Chong CF, Adi MIW, Thien A, Suyoi A, Mackie AJ, Tin AS, Tripathi S, Jaman NH, Tan KK, Kok KY, Mathew VV, Paw O, Chua HB, Yapp SK. Development of the RIPASA score : a new appendicitis scoring system for the diagnosis of acute appendicitis. *Singapore Med J*. 2010;51(3):220–225.
37. Kharbanda AB, Dudley NC, Bajaj L, Dudley NC, Bajaj L, Stevenson MD, Macias CG, Mittal MK, Bachur RG, Bennett JE, Sinclair K, Huang C, Dayan PS. Validation and refinement of a prediction rule to identify children at low risk for acute appendicitis. *Arch Pediatr Adolesc Med*. 2012;166(8):738–744. <https://doi.org/10.1001/archpediatrics.2012.490>.
38. Attia J. Moving beyond sensitivity and specificity: Using likelihood ratios to help interpret diagnostic tests. *Aust Prescr*. 2003;26(5):111–113.
39. Escribá A, Gamell AM, Fernández Y, Quintillá JM, Cubells CL. Prospective validation of two systems of classification for the diagnosis of acute appendicitis. *Pediatr Emerg Care*. 2011;27(3):165–169. <https://doi.org/10.1097/PEC.0b013e31820d6460>.
40. Tompane T, Bush R, Dansky T, Huang JS. Diagnostic Imaging Studies Performed in Children Over a Nine-Year Period. *Pediatrics*. 2013;131(1):e45–e52. <https://doi.org/10.1542/peds.2012-1228>.
41. van Laarhoven S, Eeftinck Schattenkerk M. Nieuwe richtlijn acute apendicitis: routinematige preoperatieve beeldvorming is te simpel. *Ned Tijdschr Geneesk*. 2010;154:A1571.
42. Kotagal M, Richards MK, Flum DR, Acierno SP, Weinsheimer RL, Goldin AB. Use and accuracy of diagnostic imaging in the evaluation of pediatric appendicitis. *J Pediatr Surg*. 2015;50(4):642–646. <https://doi.org/10.1016/j.jpedsurg.2014.09.080>.
43. Bhangu A, Richardson C, Torrance A, Pinkney T, Battersby C, Beral D, Cornish J, Dent H, Hall NJ, Palser T, Panagiotopoulou IG, Strong S, Velineni R. Multicentre observational study of performance variation in provision and outcome of emergency appendectomy. *Br J Surg*. 2013;100(9):1240–1252. <https://doi.org/10.1002/bjs.9201>.
44. Estey A, Poonai N, Lim R. Appendix Not Seen: the Predictive Value of Secondary Inflammatory Sonographic Signs. *Pediatr Emerg Care*. 2013;29(4):435–439. <https://doi.org/10.1097/PEC.0b013e318289e8d5>.
45. Teo AT, Lefter LP, Zarrouk AJ, Merrett ND. Institutional review of patients presenting with suspected appendicitis. *ANZ J Surg*. 2015;85(6):420–424. <https://doi.org/10.1111/ans.12531>.
46. Bergeron E. Clinical judgment remains of great value in the diagnosis of acute appendicitis. *Can J Surg*. 2006;49(2):96–100.
47. Jones PF. Suspected acute appendicitis: trends in management over 30 years. *Br J Surg*. 2001;88(12):1570–1577. <https://doi.org/10.1046/j.0007-1323.2001.01910.x>.
48. Mandeville K, Monuteaux M, Pottker T, Bulloch B. Effects of Timing to Diagnosis and Appendectomy in Pediatric Appendicitis. *Pediatr Emerg Care*. 2015;31(11):753–758. <https://doi.org/10.1097/PEC.0000000000000596>.
49. Andersson RE. The natural history and traditional management of appendicitis revisited: Spontaneous resolution and predominance of prehospital perforations imply that a correct diagnosis is more important than an early diagnosis. *World J Surg*. 2007;31(1):86–92. <https://doi.org/10.1007/s00268-006-0056-y>.