



Many diagnostic tools for appendicitis: a scoping review

Nadir Noureldin Abdella Bahta¹ · Jacob Rosenberg¹ · Siv Fonnes¹

Received: 8 August 2022 / Accepted: 15 January 2023 / Published online: 3 February 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Background We aimed to provide an overview of all diagnostic tools developed to diagnose appendicitis with their reported accuracy and to further characterize these including their need for diagnostic equipment.

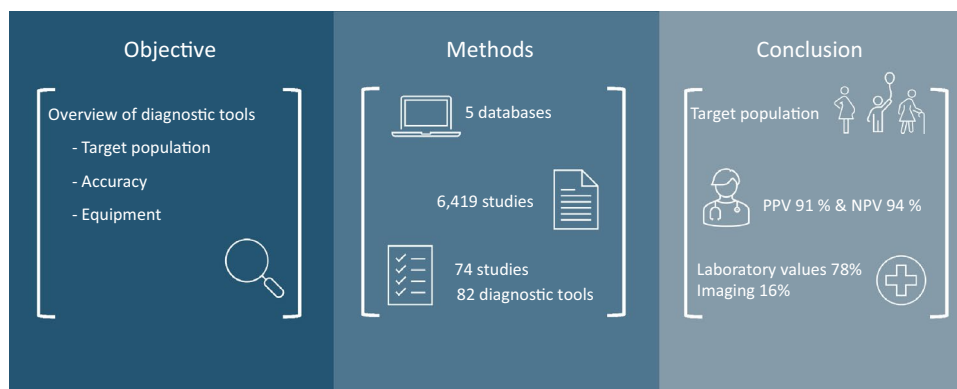
Methods This scoping review followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews and a protocol was registered at Open Science Framework. We performed a systematic literature search in PubMed, Embase, China National Knowledge Infrastructure, Índice Bibliográfico Espanhol de Ciências da Saúde, and Latin American and Caribbean Health Sciences Literature. We included original articles of all languages with the purpose to derive an accessible diagnostic tool. We extracted data regarding study- and diagnostic tool characteristics, and the accuracy of each diagnostic tool.

Results The search led to 6419 records, where 74 studies were included, yielding 82 diagnostic tools reported in seven different languages. Among these tools, 35% included patient characteristics, 85% symptoms, 93% physical examinations, 37% vital signs, 78% laboratory values, and 16% imaging. Among the diagnostic tools, 35% relied on a medical doctor/surgeon with access to a laboratory, and six diagnostic tools did not require a bedside medical doctor/surgeon. The median positive predictive value, negative predictive value, sensitivity, and specificity across diagnostic tools were 91%, 94%, 89%, and 86%, respectively.

Conclusions We identified 82 diagnostic tools that most frequently were based on symptoms and physical examinations. Most diagnostic tools relied on a medical doctor/surgeon with access to laboratory values. The accuracy was high across the diagnostic tools.

Graphical abstract

Many diagnostic tools for appendicitis: a scoping review



Bahta et al. Surgical Endoscopy 2023

@SurgEndosc

Keywords Appendicitis · Diagnostic tools · Scoping review · Score · Index

Extended author information available on the last page of the article

Appendicitis is the most common cause for abdominal surgery in children and it is the commonest surgical emergency worldwide [1] with an estimated 17.7 million cases in 2019 [2], but the incidence differs across geographical regions [3]. Appendicitis can be a challenging diagnosis to make due to the various clinical presentations including diarrhea or non-specific symptoms, hence, often misdiagnosed as gastroenteritis [4]. The risk of misdiagnosing appendicitis is higher in specific populations such as women, certain ethnicities [5], and older patients due to their comorbidities and various differential diagnoses [6]. Conversely, a misdiagnosed appendicitis can lead to a negative appendectomy, where women [7] and patients aged over 40 years [8] are at greater risk. In addition, patients undergoing a negative appendectomy have a higher risk of postoperative complications such as wound infections and incisional hernia compared with patients operated for uncomplicated appendicitis [9]. In addition, patients undergoing negative appendectomy have higher short- and long-term mortality in comparison with patients with uncomplicated appendicitis [10]. This indicates that there is a need for better preoperative diagnostic methods to avoid both the negative appendectomies and the missed diagnoses of appendicitis. One method is the use of a diagnostic tool, as incorporation of a diagnostic tool was shown to reduce the number of admissions and the surgical rate [11]. Typically, the appendicitis diagnosis through these tools is made from a predetermined combination of criteria in one or multiple categories such as patient characteristics, symptoms, physical examination, laboratory values, and imaging. One category alone, e.g. patient characteristics, physical examination, or laboratory values may not be sufficient in the diagnosis of appendicitis [12]. In combination, however, they have a higher discriminatory power. Conversely, the incorporation of diagnostic tools could require accessibility to special equipment possibly unavailable at the hospital [13, 14] making the diagnostic tool unusable.

No overview of all the available diagnostic tools for diagnosing appendicitis including non-English publications currently exist. This scoping review aimed to provide an overview of all existing diagnostic tools to diagnose appendicitis. Furthermore, we wanted to characterize these with respect to the target population, accuracy, and their need for diagnostic equipment.

Materials and methods

Protocol and registration

This review was reported according to Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) [15]. The protocol was

uploaded prior to data extraction at Open Science Framework (OSF) [16].

Eligibility criteria

This study investigated the diagnostic tools specific for diagnosing appendicitis. The primary outcomes were to characterize the incorporated variables in the diagnostic tools into categories including patient characteristics, symptoms, physical examination, vital signs, laboratory values, and/or imaging. Additionally, we characterized the needed hospital access to staff and special equipment to utilize these diagnostic tools. This could be a medical doctor/surgeon, other health professionals, thermometer, laboratory-, and/or imaging equipment. The secondary outcomes were to investigate, whether the diagnostic tools were targeted at specific populations, e.g. with regards to age (children/adults/other), sex, ethnicity, etc. and to note the geographical regions of these populations according to the United Nations [17]. In addition, we wanted to illustrate the accuracy of these diagnostic tools such as sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), where variables were either extracted or calculated [18, 19].

Inclusion criteria

The studies included in this review were based on the following eligibility criteria. Original articles that stated their purpose to derive an accessible diagnostic tool (a score, an app, a website, etc.) that approximates the risk of appendicitis or propose a cut-off score for performing an appendectomy. The variables incorporated in the diagnostic tool had to be presented, however, the individual weight of the incorporated variables was not required to be presented. Modified diagnostic tools and subsequent variations were only included if they had recalculated the risk of appendicitis with the new variables incorporated. A minimum of three predictive variables were required in either of the following diagnostic categories: patient characteristics, symptoms, physical examination, vital signs, laboratory values, and/or imaging. Articles of all languages and publication years were included.

Exclusion criteria

Articles were excluded if the diagnostic tool only depended on imaging modality (e.g. computed tomography, ultrasound, or magnetic resonance imaging) or only added an imaging modality to an existing diagnostic tool. Articles that aimed to differentiate between simple and complex appendicitis were also excluded, since they did not align with the aim of this study.

Information sources

The author group developed a search string in collaboration with a research librarian and subsequently adapted it specifically to the individual databases. The following five databases were searched: PubMed (1966 to present), China National Knowledge Infrastructure (1951 to present), Latin American and Caribbean Health Sciences Literature (1982 to present), Índice Bibliográfico Espanhol de Ciências da Saúde (2000 to present), and Embase (1974 to present). The search was performed on 3rd of March, 2022. Scientific journals and libraries were contacted if the full text articles were not available online. Furthermore, the reference lists of the included studies were screened using a snowball search [20]. For PubMed the following search string was used: (“appendicitis”[MeSH Terms] OR “appendicitis” OR “appendectomy”[MeSH Terms] OR “appendectomy” OR “appendectomies” OR “appendicectomy” OR “appendicectomies”) AND (score OR classify OR “index”). The other adapted search strings can be viewed in the protocol [16].

Data processing

Records were uploaded to Mendeley (version 1.19.8, Elsevier, UK) where duplicates were removed. Afterwards, the records were uploaded to and screened in Covidence [21], which is an online screening tool. Firstly, two authors independently screened the titles and abstracts. Secondly, two authors independently screened the relevant reports by full text. Conflicts were resolved by discussion within the author group. The data to be extracted were explored during the data charting process, in which the first author chose ten studies reporting different diagnostic tools. A template for the data charting process was made in Microsoft Excel (Microsoft, Redmond, Washington, USA) with variables detected in the full-text articles. Data charting was then performed independently by two authors and conflicts were resolved by discussion within the author group. Double data entry for the data extraction was performed by the first author. The variables extracted were study characteristics such as study period and country of population, and if the article reported the ages of the patients suffering from appendicitis and not suffering from appendicitis separately. We extracted the characteristics of the diagnostic tools such as the target population (age group, sex, and ethnicity), type of diagnostic tool (decision tree, score, etc.), and the accuracy of diagnostic tools (positive predictive value, negative predictive value, sensitivity, and specificity) for the validation cohort. However, if these were not reported for a validation cohort, the accuracy variables were extracted for the derivation cohort. If necessary, simplifications were made on the variables so they could be categorized or grouped together.

Synthesis of results

Data were plotted in histograms, and the median and range were calculated for the population size and accuracy across studies. Data were categorized into six diagnostic groups: patient characteristics, symptoms (where initial pain located elsewhere than the right lower quadrant was grouped together with migration, and appetite was grouped together with anorexia), physical examination, vital signs, laboratory values (where absolute neutrophil count, neutrophil–lymphocyte ratio, and polymorph neutrophils were grouped as neutrophils), and imaging modalities (computed tomography, ultrasound, and/or magnetic resonance imaging).

Results

After the removal of duplicates, the literature search identified 6419 unique records that underwent title and abstract screening. We then screened 159 records in full texts, and we included 77 eligible studies [13, 14, 22–96], see Fig. 1. Six of these studies [22, 23, 25, 26, 56, 57] reported on one or more similar tools (Supplementary Table 1), thus, 74 unique studies remained, and 82 diagnostic tools were found (Fig. 1 and Supplementary Table 1).

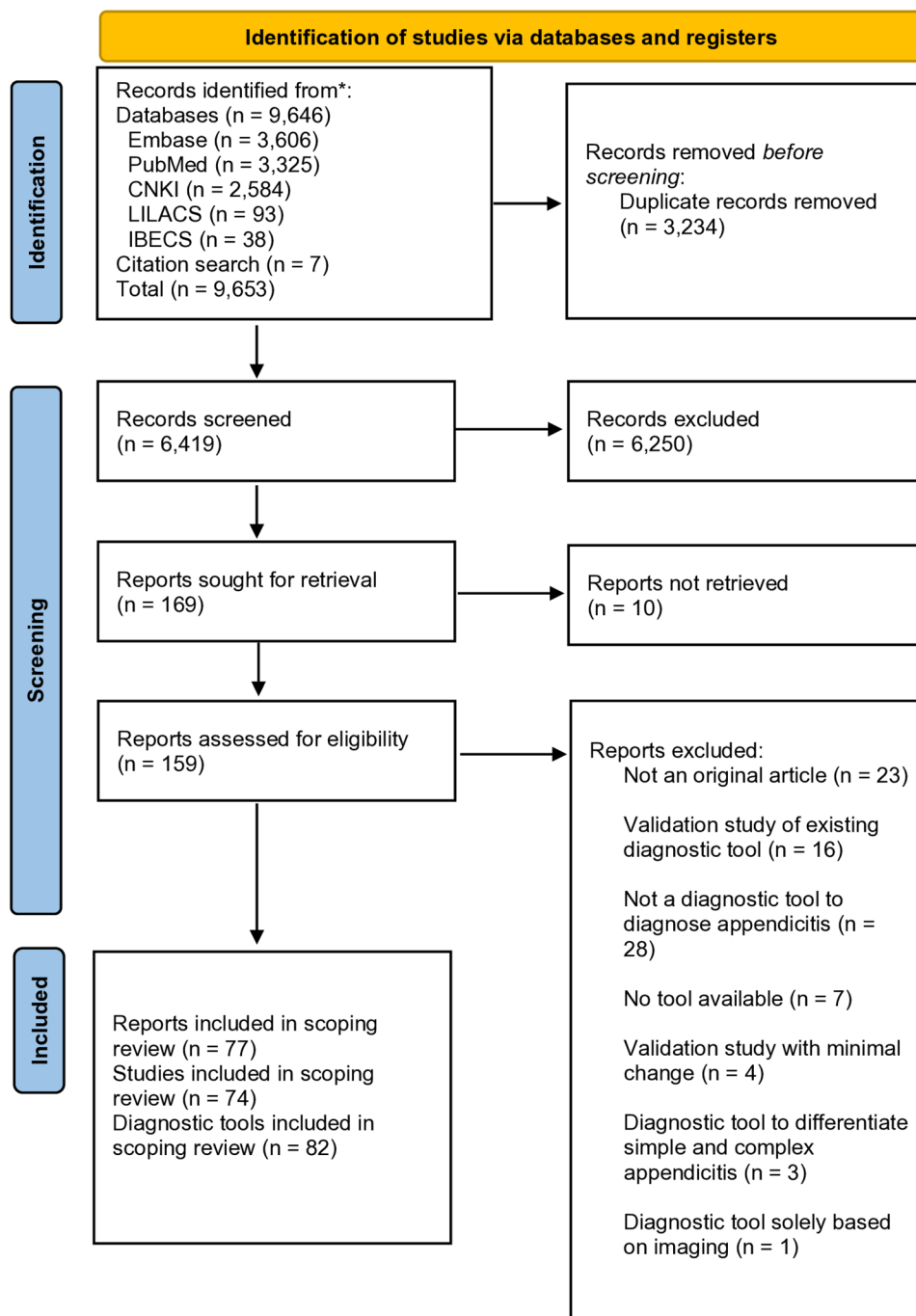
Study characteristics

The studies reported on various types of diagnostic tools, comprising 54 scoring systems (66%), 12 Eqs. (15%), eight diagnostic trees (10%) and eight others (nomogram, website, app, or desktop software) (10%), see Table 1 and Supplementary Table 1. Most of these studies were reported in English (84%) and the populations of the derivation cohorts were primarily from Europe (43%) and Asia (32%). For the respective country, region, and population cohort of the individual studies, see Supplementary Table 1. Across studies, the derivation cohort had a median size (range) of 315 (49–2423) patients, while the validation cohort had a median size of 171 (40–1426) patients.

Target population

The studies’ target populations were mainly categorized by age, sex, and ethnicity, see Table 1 and Supplementary Table 1. The remaining studies did not specify or report on the target population. Age was the most frequently used target population with a total of 28 studies and 33 diagnostic tools. Twenty-nine of these diagnostic tools were developed for children [29, 30, 34, 35, 37, 40, 42, 45, 53, 56, 70, 71, 73, 79–82, 87, 91–94] where age ranged from 0 to 20 years and the median of the mean and median ages was 10 years. The diagnostic tools consisted of 20 scoring systems, two

Fig. 1 PRISMA 2020 flow diagram of study selection process. *CNKI* China National Knowledge Infrastructure, *LILACS* Latin American and Caribbean Health Sciences Literature, *IBECs* Índice Bibliográfico Espanhol de Ciências da Saúde, *n* number



websites, one equation, and one nomogram (Supplementary Table 1). Furthermore, two studies and two diagnostic tools targeted the adolescent- and adult patients [43, 55], with age ranging from 12 to 58 years and aimed at females in their reproductive years. One was a scoring system and one a diagnostic tree (Supplementary Table 1). For older patients with the lowest delimiter being over 50 years, there were two studies and four diagnostic tools, which all consisted of equations [24–26].

For the studies differentiating the presentation of appendicitis between sexes, there were a total of six studies and nine diagnostic tools [22, 28, 38, 41, 43, 55]. Three of the studies included only females [28, 43, 55] in the derivation cohort, one study only included males [38], and two studies included both females and males with 48% males [22] and 42% males [41]. The diagnostic tools consisted of four equations, three scoring systems, and two diagnostic trees (Table 1 and Supplementary Table 1).

Table 1 Summary of characteristics of included studies and diagnostic tools

	<i>n</i>	%
Characteristics of studies (<i>n</i> = 74)		
Region		
Europe	32	43
Asia	24	32
North America	9	12
Oceania	3	4
South America	3	4
Africa	2	3
Not applicable ^a	1	1
Language		
English	62	84
Spanish	4	5
Chinese	2	3
French	2	3
German	2	3
Other ^b	2	3
Publication year		
2015–2021	26	35
2010–2014	14	19
2005–2009	7	9
2000–2004	7	9
1995–1999	4	5
1990–1994	10	14
< 1990	6	8
Characteristics for tools (<i>n</i> = 82)		
Diagnostic tools		
Scoring system	54	66
Equation	12	15
Diagnostic tree	8	10
Nomogram	3	4
Website	3	4
Other ^c	2	2
Age category		
Children	29	35
Adolescents ^d and adults	18	22
Adults	7	9
All	15	18
Not reported	13	16
Target population		
Age	33	40
Sex	9	11
Ethnicity	3	4
Not reported or applicable	37	45

^aliterature search^bincluding Korean (*n* = 1) and Russian (*n* = 1)^cincluding app (*n* = 1) and figure (*n* = 1)^dcohorts with patients aged > 10 years, *n* number

Lastly, studies targeting ethnicity as their desired population were the least frequent with a total of three studies and three diagnostic tools. Their methodology was to use patients from their region, by either making use of a local database or having a derivation cohort from their region and calculating the best predictors to diagnose appendicitis. The specific populations were from Brunei [65], Pakistan [74], and New Zealand [33], respectively. The diagnostic tools consisted of three scoring systems (Table 1 and Supplementary Table 1).

Diagnostic tools characteristics

The diagnostic tools' characteristics were categorized into six groups, see Table 2. Patient characteristics were included in 35% of the diagnostic tools, the most frequent of these were sex (32%) and age (16%). Symptoms were included in 85% of the diagnostic tools where the most frequent were nausea/vomiting (41%) and migration (40%). Physical examinations were included in 93% of the diagnostic tools, the most frequent ones being right lower quadrant tenderness (51%) and rebound pain (50%). Laboratory values were included in 78% of the diagnostic tools, the most frequent being leucocytes (66%) and neutrophils (34%). The imaging category was included in 16% of the diagnostic tools and involved ultrasound (15%) and CT (1%). None of the diagnostic tools included magnetic resonance imaging.

To characterize the needed hospital access to staff and equipment to utilize these 82 diagnostic tools under different circumstances, we developed a flow diagram (Fig. 2). Many diagnostic tools (35%) relied solely on a medical doctor/surgeon who could perform a physical examination and access to laboratory tests (*n* = 29). Often (26%) a thermometer was needed in addition (*n* = 21) and only 7% of diagnostic tools did not require a medical doctor/surgeon (*n* = 6) to perform a physical examination. However, for all of these (*n* = 6) another health professional was deemed necessary to collect patient characteristics and/or symptoms.

Accuracy

The diagnostic accuracy of the tools is depicted in Fig. 3 and the exact values can be seen in Supplementary Table 2. The total number of patients in the derivation cohort across all diagnostic tools was 34,603 patients with a median (range) of 320 (49–2423), while the validation cohort across all diagnostic tools comprised 6034 patients with a median of 176 (40–1426). A large dispersion in the accuracy was also observed, see Fig. 3. PPV was reported or possible to calculate in 56% of diagnostic tools with a median value (range) of 91% (34–100%), and > 90% in 48% of the reported diagnostic tools. NPV was reported or possible to calculate in 48% of diagnostic tools with a median of 94% (0–100%),

Table 2 Characteristics of categories used in the diagnostic tools. Characteristics used in only one or two diagnostic tools was group under other, and details can be found in footnotes, where ($n=1$) unless otherwise stated

Diagnostic categories	<i>n</i> tools	%
Patient characteristics	29	35
Sex	26	32
Age	13	16
Other ^a	5	6
Symptoms	69	85
Nausea/vomiting	34	41
Migration to right lower quadrant	33	40
Anorexia/appetite	23	28
Right lower quadrant pain	22	27
Duration of symptoms	18	22
Aggravation of pain by cough or movement	10	12
Pain character	9	11
Continuous pain	7	9
No previously similar pain	6	7
Diarrhea	4	5
Genitourinary signs	3	4
Other ^b	6	7
Physical examination	74	93
Right lower quadrant tenderness	42	51
Rebound pain	41	50
Guarding	27	33
Rigidity	12	15
Rectal tenderness	8	10
Rovsing's sign	6	7
Percussion pain	6	7
Previous abdominal surgery	5	6
Unable to walk or walk with a limb	5	6
Peritonitis	4	5
Rectal mass, right side	3	4
Tenderness outside right lower quadrant	3	4
Other ^c	13	16
Vital signs	30	37
Temperature	29	35
Pulse	1	1
Laboratory values	64	78
Leucocytes	54	66
Neutrophils	28	34
C-reactive protein	18	22
Other ^d	10	12
Imaging	13	16
Ultrasound	12	15
Computed tomography	1	1

^achronic illness, foreign national registration identity card, obesity, and risk of pregnancy ($n=2$)

^babdominal pain (not right iliac fossa alone), muscle spasm-right lower quadrant, pain in movement ($n=2$), and pain qualification ($n=2$)

^cabdominal mass; right lower quadrant, bowel sounds ($n=2$), gurgling, heel drop test ($n=2$), left lower quadrant tenderness, mesenteric lymphadenopathy, Murphy's sign, renal tenderness, obturator sign, and psoas sign

Table 2 (continued)

^dcalprotectin, hemoglobin ($n=2$), ketone bodies, leucocytic index of intoxication, leucine rich α -2 glycoprotein /creatinine, negative urinalysis, stool analyze, and shift to left ($n=2$); *n* number

and > 90% in 62% of the reported diagnostic tools. The sensitivity was reported or possible to calculate in 76% of diagnostic tools with a median of 89% (15–100%), and > 90% in 52% of the reported diagnostic tools. The specificity was reported or possible to calculate in 76% of diagnostic tools with a median of 86% (34–100%), and > 90% in 39% of the reported diagnostic tools.

Discussion

A total of 82 diagnostic tools were included in this scoping review with 12 tools reported in another language than English. Both symptoms (85%) and physical examination (93%) were included in the majority of the diagnostic tools, and one third of tools relied on a medical doctor/surgeon with access to laboratory equipment. Six diagnostic tools did not require a medical doctor/surgeon, thus no physical examination of the patient was needed. The accuracy was high for most of the diagnostic tools, meaning a diagnostic tool categorizing the patient with high or low risk for appendicitis, had a high correlation to the patient having/not-having appendicitis.

We used a comprehensive search string with the assistance of a research librarian, and the literature search was performed in five databases from three different regions. Furthermore, there were no restrictions on publication year or language. This led to an increased number of studies included and identified diagnostic tools compared with previous reviews on this subject [7, 97]. Prior to the data extraction a protocol was uploaded at Open Science Framework and the study was reported according to PRISMA-ScR [15]. A scoping review approach was chosen as the method, as our aim was to provide an overview of diagnostic tools used in diagnosing appendicitis and if diagnostic equipment was required. The scoping method allowed us to group data in variables meaningfully [98]. However, our review had some limitations. Even though we searched in multiple databases, we did not cover all regions, e.g. the Middle East, which could lead to bias if studies were published and not indexed in the included databases. However, the included databases cover most of the studies published worldwide. The accuracy of the diagnostic tools was generally high. Nevertheless, there is a need for some of the diagnostic tools to be reexamined further and revalidated in other cohort than the cohort of the original study.

In conclusion, this scoping review provided an overview of 82 diagnostic tools including 12 tools reported in other

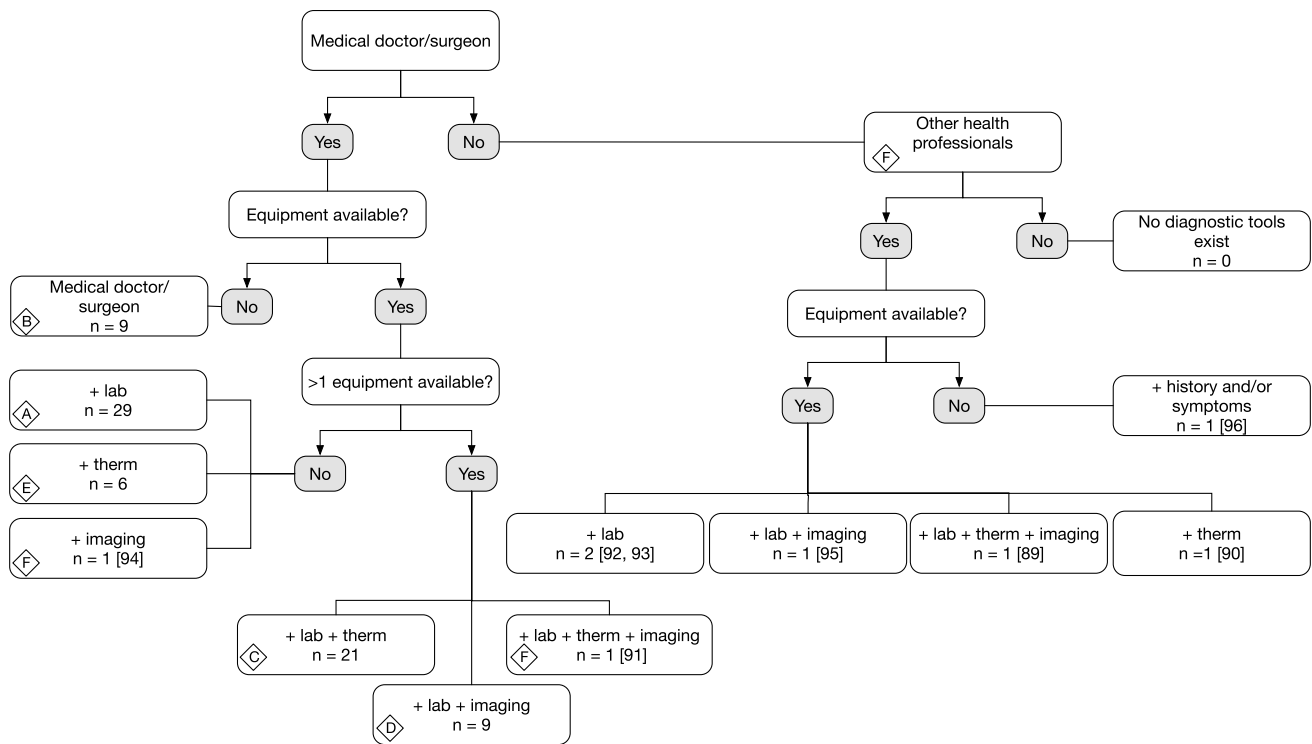
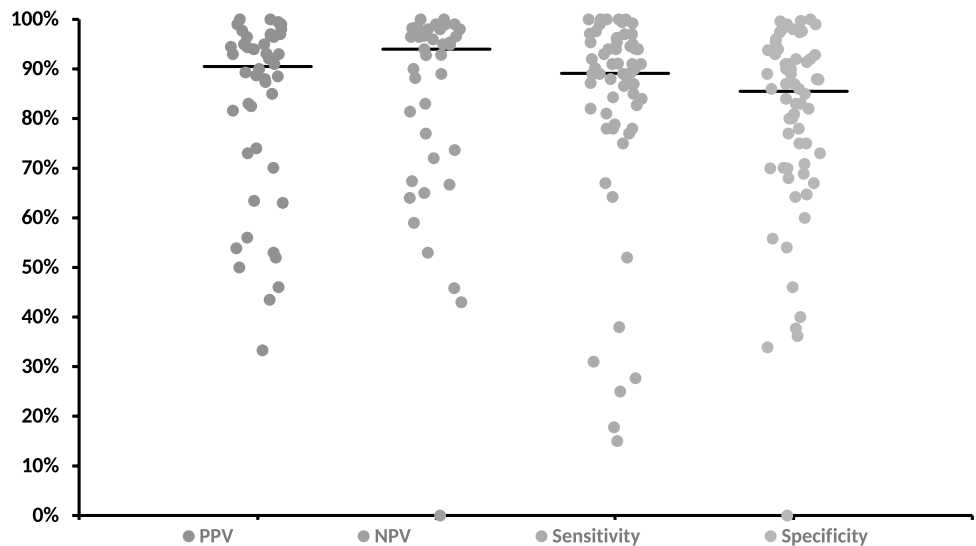


Fig. 2 Flow diagram of needed equipment and hospital staff for the 82 diagnostic tools. For further details on each diagnostic tool see Supplementary Table 1 (language, target population, and type of tool) and Supplementary Table 2 (accuracy). References to diagnostic tools in each equipment category are: A: [22–24, 28–54], B: [13, 22–27, 55–58], C: [22, 23, 59–78], D: [79–86], E: [14, 22, 23, 25–27, 87,

88], and F [89–96]. *n* number; medical doctor/surgeon: e.g. physical examination needed; other health professionals: e.g. patient characteristics and/or symptoms needed; lab: laboratory values e.g. urine, stool, or blood test; imaging: computed tomography or ultrasound; therm: thermometer

Fig. 3 Scatterplot of the accuracy in percent (y-axis) of the diagnostic tools where each grey dot represents a tool. The black horizontal lines represent the medians. The exact values can be seen in Supplemental Table 2. Sensitivity and specificity (*n*=62 tools). *PPV* positive predictive value (*n*=46 tools), *NPV* negative predictive value (*n*=39 tools), *n* number



languages than English. Some diagnostic tools were developed for specific target populations such as age, sex, and ethnicity. Furthermore, most diagnostic tools relied on a medical doctor/surgeon with access to laboratory values. The accuracy of diagnostic tools showed a large variation

but an overall good accuracy with a total median of *PPV* and *NPV* above 90% and a total median sensitivity and specificity above 85%. This study can be used as a guide for clinicians worldwide to choose a fitting diagnostic tool

according to the patient population, hospital staff, and hospital equipment.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00464-023-09890-2>.

Acknowledgements None.

Funding This research did not receive any specific Grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Disclosures Nadir Bahta, Jacob Rosenberg, and Siv Fonnes have no conflicts of interest or financial ties to disclose.

References

- Kabir SA, Kabir SI, Sun R, Jafferbhoy S, Karim A (2017) How to diagnose an acutely inflamed appendix; a systematic review of the latest evidence. *Int J Surg* 40:155–162
- Wickramasinghe DP, Xavier C, Samarasekera DN (2021) The worldwide epidemiology of acute appendicitis: an analysis of the global health data exchange dataset. *World J Surg* 45:1999–2008
- Ferris M, Quan S, Kaplan BS, Molodecky N, Ball CG, Chernoff GW, Bhala N, Ghosh S, Dixon E, Ng S, Kaplan GG (2017) The global incidence of appendicitis: a systematic review of population-based studies. *Ann Surg* 266:237–241
- Song CW, Kang JW, Kim JY (2018) Different clinical features and lower scores in clinical scoring systems for appendicitis in preschool children: comparison with school age onset. *Pediatr Gastroenterol Hepatol Nutr* 21:51–58
- Mahajan P, Basu T, Pai CW, Singh H, Petersen N, Bellolio MF, Gadepalli SK, Kamdar NS (2020) Factors associated with potentially missed diagnosis of appendicitis in the emergency department. *JAMA Netw Open* 3:1–13
- Lapsa S, Ozolins A, Strumfa I, Gardovskis J (2021) Acute appendicitis in the elderly: a literature review on an increasingly frequent surgical problem. *Geriatrics* 6:1–8
- Bhangu A (2020) Rift Study Group on behalf of the West Midlands research collaborative study evaluation of appendicitis risk prediction models in adults with suspected appendicitis. *Br J Surg* 107:73–86
- Kotaluoto S, Ukkonen M, Pauniah SL, Helminen M, Sand J, Rantanen T (2017) Mortality related to appendectomy; a population based analysis over two decades in Finland. *World J Surg* 41:64–69
- Jeon BG (2017) Predictive factors and outcomes of negative appendectomy. *Am J Surg* 213:731–738
- Andersson RE (2013) Short and long-term mortality after appendectomy in Sweden 1987 to 2006. Influence of appendectomy diagnosis, sex, age, co-morbidity, surgical method, hospital volume, and time period. A national population-based cohort study. *World J Surg* 37:974–981
- Andersson M, Kolodziej B, Andersson RE (2017) Randomized clinical trial of appendicitis inflammatory response score-based management of patients with suspected appendicitis. *Br J Surg* 104:1451–1461
- Andersson REB (2013) Meta-analysis of the clinical and laboratory diagnosis of appendicitis. *Br J Surg* 91:28–37
- Chattopadhyay S, Rabhi F, Acharya UR, Joshi R, Gajendran R (2012) An approach to model right iliac fossa pain using pain-only-parameters for screening acute appendicitis. *J Med Syst* 36:1491–1502
- Gajbhiye AS, Meshram MM, Kathod AP (2013) Utility of practical score in early diagnosis of acute appendicitis. *J Evol Med Dent Sci* 2:2548–2557
- Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MDJ, Horsley T, Weeks L, Hempel S, Akl EA, Chang C, McGowan J, Stewart L, Hartling L, Aldcroft A, Wilson MG, Garrity C, Lewin S, Godfrey CM, MacDonal MT, Langlois EV, Soares-Weiser K, Moriarty J, Clifford T, Tunçalp Ö, Straus SE (2018) PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 169:467–473
- Bahta N, Fonnes S, Rosenberg J (2022) Diagnostic tools to diagnose appendicitis - a scoping review. <https://doi.org/10.17605/OSF.IO/WT94B>
- United Nations (1999) Standard country or area codes for statistical use. Available at <https://unstats.un.org/unsd/methodology/m49/>. Accessed 02 August 2022.
- Altman DG, Bland JM (1994) Statistics notes: diagnostic tests 1: sensitivity and specificity. *BMJ* 308:1552
- Altman DG, Bland JM (1994) Statistics notes: diagnostic tests 2: predictive values. *BMJ* 309:102
- Greenhalgh T, Peacock R (2005) Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* 331:1064–1065
- Covidence (2014) Better systematic review management. Available at <https://www.covidence.org/>. Accessed 02 August 2022.
- Eskelinen M, Ikonen J, Lipponen P (1994) Sex-specific diagnostic scores for acute appendicitis. *Scand J Gastroenterol* 29:59–66
- Meklin J, Eskelinen M, Syrjänen K, Eskelinen M (2020) Gender-specific performance of a diagnostic score in acute appendicitis. *In Vivo* 34:3687–3703
- Eskelinen M, Ikonen J, Lipponen P (1994) Acute appendicitis in patients over the age of 65 years; comparison of clinical and computer based decision making. *Int J Biomed Comput* 36:239–249
- Eskelinen M, Ikonen J, Lipponen P (1995) The value of history-taking, physical examination, and computer assistance in the diagnosis of acute appendicitis in patients more than 50 years old. *Scand J Gastroenterol* 30:349–355
- Eskelinen M, Meklin J, Syrjänen K, Eskelinen M (2021) A Diagnostic Score (DS) is a powerful tool in diagnosis of acute appendicitis in elderly patients with acute abdominal pain. *Anticancer Res* 41:1459–1469
- Klein MD, Rabbani AB, Rood KD, Durham T, Rosenberg NM, Bahr MJ, Thomas RL, Langenburg SE, Kuhns LR (2001) Three quantitative approaches to the diagnosis of abdominal pain in children: practical applications of decision theory. *J Pediatr Surg* 36:1375–1380
- Arab M, Honar NB, Ghavami B, Ghodssi-Ghassemabadi R, Aghaei M, Yousefi N, Sheibani K (2021) A new decision making model for diagnosing acute appendicitis among non-pregnant women and its comparison with alvarado scoring system. *Med J Islam Repub Iran* 35:1–5
- Dadeh AA, Puitong K (2021) Predictive factors to diagnose appendicitis in children in the emergency department. *Open Access Emerg Med* 13:363–372
- Lima M, Persichetti-Proietti D, Di Salvo N, Antonellini C, Libri M, Randi B, Maffi M, Gargano T, Ruggeri G, Catania VD (2019) The APpendicitis PEdiatric (APPE) score: a new diagnostic tool in suspected pediatric acute appendicitis. *Pediatr Medica e Chir* 41:10–16

31. Gudelis M, Garcia JD, Trujillano Cabello JJ (2019) Diagnosis of pain in the right iliac fossa. A new diagnostic score based on decision-tree and artificial neural network methods. *Cir Esp* 97:329–335
32. Dzib-Calan EA, Hernández-Rojas MA, Murata C, Morales-Pérez JL, Larracilla-Salazar I, Mérida GL (2019) Modification of Alvarado score with pain semiology for the diagnosis of acute appendicitis. *Cir Cir* 87:674–681
33. Mikaere H, Zeng I, Lauti M, Kularatna M, MacCormick AD (2018) Derivation and validation of the APPEND score: an acute appendicitis clinical prediction rule. *ANZ J Surg* 88:303–307
34. Yap TL, Fan JD, Chen Y, Ho MF, Choo CS, Allen J, Low Y, Jacobsen AS, Nah SA (2019) A novel noninvasive appendicitis score with a urine biomarker. *J Pediatr Surg* 54:91–96
35. Kharbanda AB, Vazquez-Benitez G, Ballard DW, Vinson DR, Chettipally UK, Kene MV, Dehmer SP, Bachur RG, Dayan PS, Kuppermann N, O'Connor PJ, Kharbanda EO (2018) Development and validation of a novel pediatric appendicitis risk calculator (pARC). *Pediatrics* 141:1–9
36. Karaman K, Ercan M, Demir H, Yalkın Ö, Uzunoğlu Y, Gündoğdu K, Zengin İ, Aksoy YE, Bostancı EB (2018) The Karaman score: a new diagnostic score for acute appendicitis. *Ulus Travma ve Acil Cerrahi Derg* 24:545–551
37. Altali K, Ruiz-Artacho P, Trenchs V, Zárata MMO, Navarro C, Fernández C, Bodas-Pinedo A, Castillo JG, Martín-Sánchez FJ (2017) Hospital emergency room diagnosis of acute appendicitis in patients aged 2 to 20 years: the INFURG-SEMES score from the emergency infections study of the Spanish society of emergency medicine. *SEMES* 29:231–236
38. Kharbanda AB, Monuteaux MC, Bachur RG, Dudley NC, Bajaj L, Stevenson MD, Macias CG, Mittal MK, Bennett JE, Sinclair K, Dayan PS (2017) A clinical score to predict appendicitis in older male children. *Acad Pediatr* 17:261–266
39. Ahn S, Lee H, Choi W, Ahn R, Hong JS, Sohn CH, Seo DW, Lee YS, Lim KS, Kim WY (2016) Clinical importance of the heel drop test and a new clinical score for adult appendicitis. *PLoS ONE* 11:1–10
40. Yap TL, Chen Y, Low WWX, Ong CCP, Nah SA, Jacobsen AS, Shen L, Low Y (2015) A new 2-step risk-stratification clinical score for suspected appendicitis in children. *J Pediatr Surg* 50:2051–2055
41. Sammalkorpi HE, Mentula P, Leppäniemi A (2014) A new adult appendicitis score improves diagnostic accuracy of acute appendicitis - a prospective study. *BMC Gastroenterol* 14:1–7
42. Ertürk A, Tuncer IS, Balci Ö, Karaman İ, Karaman A, Çe A, Yılmaz E, Özgüner İF, Cavusoglu YH, Erdogan D (2014) The value of pediatric appendicitis score and laboratory findings on the diagnosis of pediatric appendicitis. *Turkish J Pediatr Dis* 2015:184–188
43. Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J (2013) Clinical scoring for diagnosis of acute lower abdominal pain in female of reproductive age. *Emerg Med Int* 2013:1–6
44. Kim JY, Choi SH, Cha JW, Byun CK, Koh YT, Seo DY (2010) Application of a scoring system to the diagnosis of acute appendicitis. *J Korean Soc Coloproctol* 26:34–38
45. Kharbanda AB, Taylor GA, Fishman SJ, Bachur RG (2005) A clinical decision rule to identify children at low risk for appendicitis. *Pediatrics* 116:709–716
46. Birkhahn RH, Briggs M, Datillo PA, van Deusen SK, Gaeta TJ (2006) Classifying patients suspected of appendicitis with regard to likelihood. *Am J Surg* 191:497–502
47. Jahn H, Mathiesen FK, Neckelmann K, Hovendal CP, Bellstrøm T, Gottrup F (1997) Comparison of clinical judgment and diagnostic ultrasonography in the diagnosis of acute appendicitis: experience with a score-aided diagnosis. *Eur J Surg* 163:433–443
48. Ohmann C, Franke C, Yang Q, Margulies M, Chan M, van Elk PJ, Dombal FT, Röher HD (1995) Diagnostic score for acute appendicitis. *Chirurg* 66:135–141
49. Ramirez JM, Deus J (1994) Practical score to aid decision making in doubtful cases of appendicitis. *Br J Surg* 81:1829–1830
50. Eskelinen M, Ikonen J (1992) A computer-based diagnostic score to aid in diagnosis of acute appendicitis: a prospective study of 1333 patients with acute abdominal pain. *Theor Surg* 7:86–90
51. Izbicki JR, Knoefel WT, Wilker DK, Mandelkow HK, Müller K, Siebeck M, Schweiberer L (1992) Accurate diagnosis of acute appendicitis: a retrospective and prospective analysis of 686 patients. *Eur J Surg* 158:227–231
52. François Y, Bonvoisin S, Descos L, Vignal J (1991) Prospective study of a predictive scoring system for the diagnosis of appendicitis in patients with right lower quadrant pain. Long-term outcome. *Gastroenterol Clin Biol* 15:794–799
53. Scheye T, Vanneville G (1988) Evaluation of a diagnostic score for painful abdominal syndromes suggestive of appendicitis in children over 3 years. *J Chir* 125:166–169
54. Teicher I, Landa B, Cohen M, Kabnick LS, Wise L (1983) Scoring system to aid in diagnoses of appendicitis. *Ann Surg* 198:753–759
55. Morishita K, Gushimiyagi M, Hashiguchi M, Stein GH, Tokuda Y (2007) Clinical prediction rule to distinguish pelvic inflammatory disease from acute appendicitis in women of childbearing age. *Am J Emerg Med* 25:152–157
56. Eskelinen M, Ikonen J, Lipponen P (1993) The value of history and clinical examination in the diagnosis of acute appendicitis in childhood, with special reference to computer-based decision making. *Theor Surg* 8:203–209
57. Eskelinen M, Meklin J, Syrjänen K, Eskelinen M (2021) Pediatric acute appendicitis score in children with acute abdominal pain (AAP). *Anticancer Res* 41:297–306
58. Lehmann K, Villiger P, Jenny M (1992) Negative appendectomies can be decreased by improved clinical assessment alone. *Helv Chir Acta* 58:837–840
59. Chuanjin W, Shuguang Z, Weiguo F (2020) Establishment and application of rapid score for acute appendicitis in patients with acute abdomen. *J Clin Surg* 28:1033–1035
60. Wilasrusmee C, Siribumrungwong B, Phuwapraisirisan S, Poprom N, Woratanarat P, Lertsithichai P, Attia J, Thakkinstantian A (2017) Developing and validating of Ramathibodi Appendicitis Score (RAMA-AS) for diagnosis of appendicitis in suspected appendicitis patients. *World J Emerg Surg* 12:1–10
61. Faruquzzaman MSK, Hoque J (2014) Implementation of newly proposed Appendicular Scoring System (ASS) for diagnosis of acute appendicitis in surgical practice-a clinical study. *Dinajpur Med Col J* 7:47–51
62. Kasimov R (2012) Integral diagnosis of acute appendicitis. *Sovrem Tehnol v Med* 2012:112–114
63. Shera AH, Nizami FA, Malik AA, Naikoo ZA, Wani MA (2011) Clinical scoring system for diagnosis of acute appendicitis in children. *Indian J Pediatr* 78:287–290
64. Kim BS, Ryu DH, Kim TH, Jeong IU, Song JH, Cho S, Kim JK, Jeong YS, Lee SJ (2010) Diagnosis of acute appendicitis using scoring system: compared with the Alvarado Score. *J Korean Surg Soc* 79:207–214
65. 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 (2010) Development of the RIPASA score: a new appendicitis scoring system for the diagnosis of acute appendicitis. *Singapore Med J* 51:220–225
66. Goh P (2010) A simplified appendicitis score in the diagnosis of acute appendicitis. *Hong Kong J Emerg Med* 17:230–235

67. Beltran M, Almonacid J, Gutierrez J (2009) Diagnostic score for appendicitis in children by emergency unit paediatricians. *Rev Chil Pediatr* 78:584–591
68. Andersson M, Andersson RE (2008) The appendicitis inflammatory response score: a tool for the diagnosis of appendicitis that outperforms the Alvarado score. *World J Surg* 32:1843–1849
69. Beltran M, Ra'ul V, Tapia T, (2004) Score diagnóstico de apendicitis: estudio prospectivo, doble ciego, no aleatorio. *Chil Cir* 56:550–557
70. van den Broek WT, van der Ende ED, Bijnen AB, Breslau PJ, Gouma DJ (2004) Which children could benefit from additional diagnostic tools in case of suspected appendicitis? *J Pediatr Surg* 39:570–574
71. Zhi-bin N, Qun Z, Hui-min J (2003) Establishment and preliminary application of scoring system to aid in diagnosis of acute appendicitis in children. *J Clin Pediatr Surg* 2:406–409
72. van den Broek WT, Bijnen BB, Rijbroek B, Gouma DJ (2003) Scoring and diagnostic laparoscopy for suspected appendicitis. *Eur J Surg* 168:349–354
73. Dado G, Anania G, Baccarani U, Marcotti E, Donini A, Risaliti A, Pasqualucci A, Bresadola F (2000) Application of a clinical score for the diagnosis of acute appendicitis in childhood: a retrospective analysis of 197 patients. *J Pediatr Surg* 35:1320–1322
74. Jawaid A, Asad A, Motiei A, Munir A, Bhutto E, Choudry H, Idrees K, Durrani K, Rahman M, Ahuja M, Ahmed NQ, R, Ali S, Aslam S, Abbasi S, Feerasta S, Alam S, Ahmed U, Jehan I, (1999) Clinical scoring system: a valuable tool for decision making in cases of acute appendicitis. *J Pak Med Assoc* 49:254–259
75. Kalan M, Talbot D, Cunliffe WJ, Rich AJ (1994) Evaluation of the modified Alvarado score in the diagnosis of acute appendicitis: a prospective study. *Ann R Coll Surg Engl* 76:418–419
76. Grogono JL, Christian F, Christian GP (1992) A simple scoring system to reduce the negative appendectomy rate. *Ann R Coll Surg Engl* 74:281–285
77. Fenyő G (1987) Routine use of a scoring system for decision-making in suspected acute appendicitis in adults. *Acta Chir Scand* 153:545–551
78. Alvarado A (1986) A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med* 15:557–564
79. Marcinkevics R, Wolfertstetter PR, Wellmann S, Knorr C, Vogt JE (2021) Using machine learning to predict the diagnosis, management and severity of pediatric appendicitis. *Front Pediatr* 9:1–11
80. Stiel C, Elrod J, Klinke M, Herrmann J, Junge CM, Ghadban T, Reinshagen K, Boettcher M (2020) The modified Heidelberg and the AI appendicitis score are superior to current scores in predicting appendicitis in children: a two-center cohort study. *Front Pediatr* 8:1–6
81. Pop G, Costea F, Lungeanu D, Iacob E, Popoiu C (2022) Ultrasonographic findings of child acute appendicitis incorporated into a scoring system. *Singapore Med* 63:35–41
82. Dokumcu Z, Kurtmen TB, Divarci E, Tamay PB, Kose T, Sezak M, Ozok G, Ergun O, Celik A (2018) Retrospective multivariate analysis of data from children with suspected appendicitis: a new tool for diagnosis. *Emerg Med Int* 2018:1–9
83. Lamture YR, Ramteke H, Shinde RK, Shahapurkar VV, Gajbhiye VP (2017) Clinico sonological and laboratory co-relation with histopathology of acute appendicitis to develop new diagnostic scoring system (Yash scoring system). *Int Surg J* 4:2556–2564
84. Mán E, Simonka Z, Varga A, Ráosi F, Lázár G (2014) 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* 28:2398–2405
85. Siddiqui ZR, Khaliq T, Shah SA (2011) A new simple scoring system for the diagnosis of acute appendicitis. *J Pioneer Med Sci* 1:32–37
86. Tzanakis NE, Efstathiou SP, Danulidis K, Rallis GE, Tsioulos DI, Chatzivasiliou A, Peros G, Nikiteas NI (2005) A new approach to accurate diagnosis of acute appendicitis. *World J Surg* 29:1151–1156
87. Lintula H, Pesonen E, Kokki H, Vanamo K, Eskelinen M (2005) A diagnostic score for children with suspected appendicitis. *Langenbeck's Arch Surg* 390:164–170
88. Arnbjörnsson E (1985) Scoring system for computer-aided diagnosis of acute appendicitis. The value of prospective versus retrospective studies. *Ann Chir Gynaecol* 74:159–166
89. Ghareeb WM, Emile SH, Elshobaky A (2022) Artificial Intelligence compared to Alvarado scoring system alone or combined with ultrasound criteria in the diagnosis of acute appendicitis. *J Gastrointest Surg* 26:655–658
90. Al Masud F, Islam RR, Sajal MHK (2020) Smart risk prediction tools of appendicitis patients: a machine learning approach. *Biointerface Res Appl Chem* 11:7804–7813
91. Abo-Sherif A, Arafa M, Hashish M, Metwally M (2020) Evaluation of a new scoring system for early diagnosis and management of acute appendicitis in pediatric patients. *Egypt J Surg* 39:1060–1068
92. Benito J, Fernandez S, Gendive M, Santiago P, Perez-Garay R, Arana-Arri E, Mintegi S (2020) A new clinical score to identify children at low risk for appendicitis. *Am J Emerg Med* 38:554–561
93. Lu YT, Chen PC, Huang YH, Huang FC (2020) Making a decision between acute appendicitis and acute gastroenteritis. *Children* 7:1–11
94. Boettcher M, Breil T, Günther P (2016) The Heidelberg appendicitis score simplifies identification of pediatric appendicitis. *Indian J Pediatr* 83:1093–1097
95. Antevil JL, Rivera L, Langenberg BJ, Hahm G, Favata MA, Brown CVR (2006) Computed tomography-based clinical diagnostic pathway for acute appendicitis: prospective validation. *J Am Coll Surg* 203:849–856
96. Van Way CW 3rd, Murphy JR, Dunn EL, Elerding SC (1982) A feasibility study of computer aided diagnosis in appendicitis. *Surg Gynecol Obstet* 155: 685–688
97. Kularatna M, Lauti M, Haran C, MacFater W, Sheikh L, Huang Y, McCall J, MacCormick AD (2017) Clinical prediction rules for appendicitis in adults: which is best? *World J Surg* 41:1769–1781
98. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E (2018) Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol* 18:1–7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Nadir Nouredin Abdella Bahta¹  · Jacob Rosenberg¹ · Siv Fonnes¹

✉ Nadir Nouredin Abdella Bahta
nadirbuhta@hotmail.com

of Copenhagen, Borgmester Ib Juuls Vej 1, 2730 Herlev,
Denmark

¹ Department of Surgery, Center for Perioperative
Optimization, Herlev and Gentofte Hospital, University