BJS Open, 2024, zrae120
https://doi.org/10.1093/bjsopen/zrae120
Original Article

Nationwide prospective audit for the evaluation of appendicitis risk prediction models in adults: right iliac fossa treatment (RIFT)—Turkey

Ali Yalcinkaya[1,2,*], Ahmet Yalcinkaya[3,4], Bengi Balci[5], Can Keskin[6], Ibrahim Erkan[6], Alp Yildiz[7], Erdinc Kamer[8], Sezai Leventoglu[1] and RIFT TURKEY Study Collaboration

[1]Department of General Surgery, Faculty of Medicine, Gazi University, Ankara, Turkey
[2]Department of Clinical Medicine, Faculty of Medicine, Aalborg University, Aalborg, Denmark
[3]Department of Medical Biochemistry, Faculty of Medicine, Hacettepe University, Ankara, Turkey
[4]Department of Medical Biochemistry and Microbiology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden
[5]Department of General Surgery, ASV Yasam Hospital, Antalya, Turkey
[6]Tibbi Akademik, Ankara, Turkey
[7]Department of General Surgery, Ankara Yenimahalle Training and Research Hospital, Ankara, Turkey
[8]Department of General Surgery, Izmir Katip Celebi University Ataturk Training and Research Hospital, Izmir, Turkey

**Correspondence to:* Ali Yalcinkaya, Department of Clinical Medicine, Faculty of Medicine, Aalborg University, Aalborg 9000, Denmark (e-mail: aliyalcink@gmail.com; @AliYalcinkayaMD, @RiftTurkey)

Members of the RIFT TURKEY Study Collaboration are co-authors of this study and are listed under the heading Collaborators.

Abstract

Background: Appendicitis is the most prevalent surgical emergency. The negative appendicectomy rate and diagnostic uncertainty are important concerns. This study aimed to assess the effectiveness of current appendicitis risk prediction models in patients with acute right iliac fossa pain.

Methods: A nationwide prospective observational study was conducted, including all consecutive adult patients who presented with right iliac fossa pain. Diagnostic, clinical and negative appendicectomy rate data were recorded. The Alvarado score, Appendicitis Inflammatory Response (AIR), Raja Isteri Pengiran Anak Saleha Appendicitis (RIPASA) and Adult Appendicitis Score systems were calculated with collected data to classify patients into risk categories. Diagnostic value and categorization performance were evaluated, with use of risk category-based metrics including 'true positive rate' (percentage of appendicitis patients in the highest risk category), 'failure rate' (percentage of patients with appendicitis in the lowest risk category) and 'categorization resolution' (true positive rate/failure rate).

Results: A total of 3358 patients from 84 centres were included. Female patients were less likely to undergo surgery than men (71.5% *versus* 82.5% respectively; relative risk 0.866, 95% c.i. 0.834 to 0.901, $P < 0.001$); with a three-fold higher negative appendicectomy rate (11.3% *versus* 4.1% respectively; relative risk 2.744, 95% c.i. 2.047 to 3.677, $P < 0.001$). Ultrasonography was utilized in 56.8% and computed tomography in 75.2% of all patients. The Adult Appendicitis Score had the best diagnostic performance for the whole population; however, only RIPASA was significant in men. All scoring systems were successful in females patients, but Adult Appendicitis Score had the highest area under the receiver operating characteristic curve value. The RIPASA and the Adult Appendicitis Score had the best categorization resolution values, complemented by their exceedingly low failure rates in both male and female patients. Alvarado and AIR had extremely high failure rates in men.

Conclusion: The negative appendicectomy rate was low overall, but women had an almost three-fold higher negative appendicectomy rate despite lower likelihood to undergo surgery. The overuse of imaging tests, best exemplified by the 75.2% frequency of patients undergoing computed tomography, may lead to increased costs. Risk-scoring systems such as RIPASA and Adult Appendicitis Score appear to be superior to Alvarado and AIR.

Introduction

Acute appendicitis is one of the most commonly encountered surgical emergencies[1,2]. Despite decades of experience and studies, diagnosis remains challenging, particularly in young women for whom acute abdominal pain necessitates the assessment of a broader range of differential diagnoses[3,4]. A key concern is overtreatment, which increases unnecessary surgeries (negative appendicectomy) and can be associated with postoperative complications, prolonged hospital stays and unnecessary healthcare costs[4–6]. Although many international guidelines recommend the routine use of risk prediction models

Received: November 09, 2023. **Revised:** August 19, 2024. **Accepted:** August 21, 2024

in patients with acute abdominal pain, the negative appendicectomy rate (NAR) has been reported to be as high as 28% in female patients and 12% in male patients[5,7–9].

To decrease NAR, numerous risk scoring systems that aid in diagnosis have been introduced, including the Alvarado score, Appendicitis Inflammatory Response (AIR), Raja Isteri Pengiran Anak Saleha Appendicitis (RIPASA) and Adult Appendicitis Score (AAS). However, the positive predictive values (PPV) and negative predictive values (NPV) of these scoring systems are often poor, and inconsistent performance metrics have been demonstrated by many studies[10–12]. These limitations have resulted in limited clinical implementation of these scores in patients with right iliac fossa (RIF) pain, and there is an apparent need to establish their utility by comparing outcomes in different populations. Nonetheless, it is crucial to recognize that these scoring systems are particularly limited when assessing female patients, in whom differential diagnoses are various, or while evaluating immigrants, who often experience problems in accessing healthcare and are often affected by a language barrier. Today, around 5–7% of Turkey's population is comprised of immigrants, with the majority being of Syrian origin[13]. Therefore, auditing patients with acute appendicitis and appendicitis scoring systems with respect to sex and immigrant status have become crucial to establish problems in patient management, ensure high-quality healthcare, and reduce the financial burden caused by unnecessary imaging studies and surgeries.

This study presents the first analysis of the Right Iliac Fossa Treatment-Turkey (RIFT-TR) study, which aimed to: identify optimal risk prediction models for acute right iliac fossa (RIF) pain in Turkey's population according to age and sex, to assess whether these scores have similar efficacy in immigrants and to reveal nationwide clinical trends, thereby establishing issues and facilitating debates that may lead to possible improvements.

Methods

Study design

The present study is a continuum of the study conducted by the RIFT study group in the UK and several other European countries[9]. It was designed as a nationwide multicentre prospective observational cohort study to examine risk prediction models used in the assessment of acute appendicitis. Several problems identified in the data collection and analysis processes of the original RIFT study were corrected, and new physical and electronic forms were prepared. The study protocol, including outcome measures and data collection methods, were disseminated through a readily available network of Department Chairs at secondary and tertiary healthcare centres throughout the country (a total of 84 centres).

All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Ethical approval was obtained from the Clinical Research Ethics Committee of Gazi University Faculty of Medicine (7 September, 2020) and the study was registered on clinicaltrials.gov, registration number NCT04614649. All patients received information regarding the study and the use of their data, and written informed consent forms were obtained. The study was reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies[14].

Eligibility criteria and study groups

All consecutive patients older than 18 years of age who were referred by a general practitioner or emergency physician to the on-call surgical team with acute RIF pain or suspected appendicitis were included in this study. The timeline for the inclusion of patients was from September 2020 to December 2020. Patients who reported they had previously undergone appendicectomy were not included in the analyses. All pregnant women were also excluded from the study because pregnant women usually cannot be assessed with scoring systems and have considerably different clinical needs compared with other patients. Pregnancy was identified by patients' self-reports or urine analysis upon suspicion.

To evaluate scoring systems, patients were divided into two groups: surgical and non-surgical treatment—the latter of which included patients treated with conservative management. Surgical treatment groups were further classified into four subgroups based on their surgical pathology results: those diagnosed with acute appendicitis; those without acute appendicitis but with non-obstetric/gynaecological surgical pathologies; those with obstetric or gynaecological pathologies such as ovarian cyst rupture and pelvic inflammatory disease; and those with negative laparotomy or laparoscopic findings.

Study variables and data collection

All data were collected by use of a standardized case report form and this information was then swiftly transferred to the online database. These data were collected prospectively by the attending physicians at the time of admission, after obtaining the imaging results, completing the care process and at the end of the follow-up interval. Demographic and clinical features including age, sex, symptoms, examination findings, blood tests, and imaging reports obtained from ultrasonography (USG), computed tomography (CT) and magnetic resonance imaging (MRI) were recorded. In addition to these, the operative procedure, operative findings and histopathological results were noted for patients in the surgical treatment group. A supervising consultant surgeon at each hospital oversaw study conduct and was responsible for the overall quality assurance of data submission. Local lead investigators were contacted with specific details of missing data when necessary. Also, ten participating centres were randomly identified to validate their data. Where incorrect data were identified, validators were asked to amend those data points on the study database. If any data points were missing despite aforementioned measures, these data points were excluded from the analysis without excluding the patients themselves. Owing to the exceptionally low level of missing data, this approach was deemed sufficient.

Diagnostic analysis and risk score assessment

Negative appendicectomy rate

The NAR value was calculated as the percentage of patients with normal appendix histology who had undergone appendicectomy. Patients with appendix pathology other than appendicitis (such as appendix tumour) were included in the denominator but not the numerator.

Calculation of risk scores

Collected data were also used to calculate the four most commonly used adult risk prediction models: Alvarado[15], AIR[16], RIPASA[17] and AAS[18]. If these scores had been calculated by the attending physicians, the values were recorded after being checked for accuracy. However, in the majority of patients, attending physicians had not calculated scores despite data availability. As such, Alvarado, AIR, RIPASA and AAS scores were calculated for each patient by the primary researchers of the RIFT study based on readily recorded information in order to create a hypothetical clinical scenario for risk prediction models. The information used to calculate these scoring systems was collected prospectively before the results of the imaging studies were available. Of note, the original RIFT study was known to have problems with the calculation of scores and the cut-off values, as well as issues in data collection[9]. These problems were corrected before initiating the present study, both in the printed forms used for bedside data collection and the digital forms/files used for data entry and storage. Scoring systems are summarized in Table 1.

Table 1 AIR, Alvarado, RIPASA and AAS score characteristics

	Alvarado	AIR	RIPASA	AAS
Vomiting		1		
Nausea or vomiting	1		1	
Anorexia	1		1	
Pain in RIF		1	0.5	2
Migration of pain to the RIF	1		0.5	2
Rovsing's sign			2	
RIF tenderness	2		1	
Women ≥50 years or men (any age)				3
Women <50 years				1
Rebound tenderness or muscular defence/guarding	1		1 + 2	
Light		1		2
Medium		2		4
Strong		3		4
Body temperature				
>37.5°C	1			
≥38.5°C		1		
>37 to <39°C			1	
White blood cell count				
>10.0 × 10^9/l	2		1	
10.0–14.9 × 10^9/l		1		
≥15.0 × 10^9/l		2		
≥7.2 and <10.9 × 10^9/l				1
≥10.9 and <14.0 × 10^9/l				2
≥14.0 × 10^9/l				3
Leucocytosis shift	1			
Polymorphonuclear leucocytes				
70–84%		1		
≥85%		2		
≥62% and <75%				2
≥75% and <83%				3
≥83%				4
CRP concentration				
10–49 mg/l		1		
≥50 mg/l		2		
Symptoms <24 h and CRP (C-reactive protein) concentration				
≥4 and <11 mg/l				2
≥11 and <25 mg/l				3
≥25 and <83 mg/l				5
≥83 mg/l				1
Symptoms >24 h and CRP (C-reactive protein) concentration				
≥12 and <53 mg/l				2
≥53 and <152 mg/l				2
≥152 mg/l				1
Sex				
Male			1	
Female			0.5	
Age				
<40 years			1	
≥40 years			0.5	
Duration of symptoms				
<48 h			1	
>48 h			0.5	
Negative urinalysis			1	
Foreign NRIC (national registration identity card)			1	
Highest possible total score	10	12	16	23

Alvarado score/acute appendicitis response score (AIR), low risk (<5), intermediate risk (5–8), high risk (>8). Raja Isteri Pengiran Anak Saleha Appendicitis (RIPASA) score, unlikely (<5), low risk (5–7), high risk (7.5–11.5), definite (≥12). Adult Appendicitis Score (AAS), low risk (0–10), intermediate risk (11–15), high risk (≥16). RIF, right iliac fossa; NRIC, national registration identity card.

Assessment of diagnostic performance

The diagnostic performance of scores was assessed with two approaches: by examining receiver operating characteristic (ROC) analyses that create a single cut-off value, and by evaluating predictive classification results. For the primary analyses, individuals who had undergone non-surgical treatment were excluded from the ROC analysis. While this approach artificially reduced negative predictive values, it was deemed necessary to simulate real-world scenarios. In the clinical setting, risk stratification systems are useful if they perform well among patients that present with typical RIF pain and are seriously being considered for surgery, not patients who can be excluded by examination or preliminary methods. As such, the primary ROC analyses examined in this study were performed on only the operated on patients ($n = 2610$). It should be mentioned that non-operated on appendicitis patients were also excluded with this approach, and this may be considered as a limitation. However, this decision aimed to assess the diagnostic accuracy of the scores in situations where they would be most crucial—when surgeons were faced with a challenging decision regarding the necessity of appendicectomy. Nonetheless, ROC analyses with the entire cohort ($n = 3358$) were also performed in the interest of providing complete outcome evaluations (Supplementary data).

In the second approach, risk prediction models were assessed by evaluating their classification performance by including all subjects enrolled into the study ($n = 3358$). 'Appendicitis ratio', 'categorization resolution' (CR) and 'failure rate' (the latter widely known as false omission rate) were calculated. The 'appendicitis ratio' was the percentage of patients with acute appendicitis in each category of each risk scoring system. For the highest risk category of each prediction model, this value was accepted to describe the 'true positive rate'. The CR was calculated as a ratio, by dividing the appendicitis ratio in the highest risk category (true positive rate) by the appendicitis ratio in the lowest risk category (the percentage of patients with appendicitis despite being predicted to have the lowest risk = failure rate). This latter value (failure rate) will be used in this study, similar to prior research[4]. The failure rate is defined as the proportion of patients with appendicitis who were stratified into the low-risk group of a given risk score (false negatives/(true negatives + false negatives)). Finally, despite the fact that CR values would be impacted by the uneven distribution of patients into each risk category, this approach was chosen as it would present data for the real-world utility of scoring systems.

Follow-up

Patients were followed up to assess any subsequent hospital admissions within 30 days of initial presentation, and then a further 30 days of follow-up was performed using a combination of phone calls and electronic hospital records. Thus, all participants were evaluated for a 60-day interval after the initial admission that caused study inclusion.

Statistical analysis

All analyses were subject to a significance threshold of $P < 0.05$ and were performed on SPSS v.25 (IBM, NY, USA). Histogram and Q-Q plots were used to determine whether variables conformed to a normal distribution. Data are given as mean(s.d.) or as median (1st quartile–3rd quartile) for continuous variables according to the normality of distribution, and as an absolute and relative frequency for categorical variables. Normally distributed variables were analysed with the one-way analysis of variance (ANOVA). Non-normally distributed variables were analysed with the Kruskal–Wallis test. Categorical variables were analysed with chi-square tests (Pearson, Yate's correction, Fisher's exact, Fisher–Freeman–Halton). Pairwise comparisons were adjusted by the Bonferroni correction. Risk prediction models were assessed if patients could be scored with the readily available data points, and the prediction performances were evaluated based on sex and immigrant status. The overall ability of the risk prediction models to discriminate between patients with and without acute appendicitis was determined by calculation of the area under the curve (AUC) for ROC analyses, and optimal cut-off points were determined using the Youden *J* index.

Superscripted letters in tables refer to pairwise comparison results. The letter notation approach was used to describe statistical similarities between columns (groups) that were found to be statistically similar. For each row, the presence of the same letters in different columns indicates that the denoted variables were statistically similar between those groups, as assessed by Bonferroni correction. For example, 'a', 'ab', and 'b' in the same row indicate that there was a significant difference between the first and third groups and that the second group was similar to both the first and third groups.

Results

Clinical characteristics and outcomes

A total of 3358 patients (1911 men and 1447 women) from 84 centres nationwide were included in this study. The median age was 32 (range 18–91) years. A total of 196 (5.8%) immigrant patients were identified. A prior history of RIF pain was reported by 594 (17.8%) patients, while 533 (15.9%) patients had previously undergone abdominal surgery (*Table 2*).

Table 2 Demographic features and abdominal history of the cohort

Age (years), median (1st–3rd quartile)	32 (24–44)
<45	2540 (75.6)
≥45	818 (24.4)
Sex	
Male	1911 (56.9)
Female	1447 (43.1)
Nationality	
Turkish	3162 (94.2)
Syrian	107 (3.2)
Other immigrants	89 (2.7)
RIF pain history	
Once	401 (12.0)
Twice or more	193 (5.8)
Abdominal surgery history	533 (15.9)

Values are *n* (%) unless otherwise indicated. RIF, right iliac fossa.

At admission, symptom duration was less than or equal to 24 h in 2021 (60.3%) patients, greater than 24–48 h in 596 (17.7%) patients and longer than 48 h in 741 (22%) patients. RIF pain was reported as the primary complaint by 3245 (96.6%) patients, while the other common symptoms of acute appendicitis, nausea and loss of appetite were present in 57.3% and 50.4% of patients respectively. The most common findings on physical examination were 'tenderness without rebound' and 'localized guarding' in 44.8% and 45.2% of patients respectively. USG was utilized in 1906 (56.8%) patients and confirmed appendicitis in 809 of these patients (42.4%), CT was utilized in 2524 (75.2%) patients and confirmed appendicitis in 1766 of these patients

Table 3 Clinical features of the study cohort

Variables	Overall cohort (*n* = 3358)
Duration of symptoms	
≤24 h	2021 (60.3)
24–48 h	596 (17.7)
>48 h	741 (22.0)
Patient-reported complaints/symptoms	
RIF pain	3245 (96.6)
Nausea	1924 (57.3)
Vomiting	921 (27.4)
Loss of appetite	1686 (50.2)
Pain migration to RIF	1640 (48.8)
RIF examination	
No tenderness	171 (5.1)
Tenderness without rebound	1506 (44.8)
Localized guarding	1518 (45.2)
Generalized guarding	163 (4.9)
RIF rebound tenderness	1948 (58.0)
Rovsing's sign	789 (23.5)
Urine analysis	
No investigation	1345 (40.1)
Negative	1565 (46.6)
Positive	448 (13.3)
Laboratory results, median (1st–3rd quartiles)	
White blood cell count ($\times 10^3$)	13.05 (9.93–16.11)
Neutrophil ($\times 10^3$)	9.92 (6.85–13.00)
CRP	18.4 (4.2–66.8)
USG findings	
Appendicitis confirmed	809 (24.1)
Appendicitis excluded	149 (4.4)
Appendicitis suspicious	814 (24.2)
Another pathology	134 (4.0)
Not applied	1452 (43.2)
CT findings	
Appendicitis confirmed	1766 (52.6)
Appendicitis excluded	249 (7.4)
Appendicitis suspicious	318 (9.5)
Another pathology	191 (5.7)
Not applied	834 (24.8)
MRI findings	
Appendicitis confirmed	7 (0.2)
Appendicitis excluded	8 (0.2)
Appendicitis suspicious	1 (0.0)
Another pathology	12 (0.4)
Not applied	3330 (99.2)
Risk prediction scores, mean(s.d.)	
Alvarado score	6.23(1.91)
AIR score	6.38(2.02)
RIPASA score	8.85(2.55)
AAS score	14.63(3.30)

Values are *n* (%) unless otherwise indicated. RIF, right iliac fossa; CRP, C-reactive protein; USG, ultrasound; CT, computed tomography; MRI, magnetic resonance imaging; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

(70%), MRI was utilized in only 28 (0.8%) patients and confirmed appendicitis in seven of these patients (25%) (Table 3).

When the whole cohort was assessed with respect to discharge diagnoses ($n = 3358$), the most frequent were appendicitis in 2614 (77.8%), non-specific pain in 157 (4.7%) and other gastrointestinal pathologies in 111 (3.3%) patients. Among the 2614 patients with appendicitis, 186 (7.1%) had received only conservative treatment (antibiotic therapy). Surgery was performed in 2610 (77.7%) patients and 2428 (93%) of the subjects underwent surgery within 24 h. Discharge diagnoses showed that 2428 (93%) of patients had acute appendicitis, while 182 (7.0%) had normal appendices (NAR: 7%). Despite the fact that female patients with RIF pain were found to be less likely to undergo surgery than men (71.5% *versus* 82.5% respectively; RR 0.866, 95% c.i. 0.834 to 0.901, $P < 0.001$), NAR in women was almost three-fold higher than in men (11.3% *versus* 4.1%; RR 2.744, 95% c.i. 2.047 to 3.677). In the immigrant group, NAR was 7.8%.

A total of 229 surgeries were performed for surgical conditions other than appendicitis. This number included patients in which surgery was begun with a preliminary diagnosis of acute appendicitis, but the surgery was then changed due to intraoperative findings. Open surgery was utilized in 1128 (54.3%) patients, while surgeries were carried out with the laparoscopic approach in 915 (44.1%) patients. In 34 patients (1.6%), the operation was started with the laparoscopic approach but conversion to open surgery was deemed necessary (Table 4).

Table 4 Operative features of the surgical cohort

Variables	Surgical cohort (n = 2610)
Time between admission and surgery	
<24 h	2428 (93.0)
24–48 h	148 (5.7)
48–72 h	16 (0.6)
72–96 h	8 (0.3)
96–168 h	3 (0.1)
>168 h	7 (0.3)
Discharge diagnosis with respect to subgroups	
Acute appendicitis	2428 (93)
Patients with normal appendicectomy (NAR)	182 (7)
Surgical pathology except OB/GYN	61 (2.3)
OB/GYN pathologies	51 (2.0)
No pathology detected	70 (2.7)
Operative approach (reported in 2077 patients)	
Open—RIF incision	1042 (50.2)
Open—midline incision	86 (4.1)
Laparoscopic	915 (44.1)
Conversion to open surgery	34 (1.6)
Procedure(s) other than appendicectomy (performed in 229 patients)	
Diagnostic	32 (14.0)
Right hemicolectomy	17 (7.5)
Meckel's diverticulum resection	3 (1.3)
Small intestine resection	4 (1.7)
Other intestinal surgery	22 (9.7)
Gynaecological	48 (20.9)
Urological	2 (0.8)
Other	101 (44.1)
Duration of hospital stay in days, median (1st–3rd quartile)	1 (1–2)

Values are *n* (%) unless otherwise indicated. RIF, right iliac fossa, NAR, normal appendicectomy rate, OB/GYN, obstetric/gynaecological.

Patient characteristics compared across groups

Among the 2610 patients who underwent surgery, 2428 (93.0%) patients had received surgical treatment for acute appendicitis, with the majority being men (62.2%). In this group, surgical intervention was performed in 2370 (97.6%) patients at first admission, while 58 (2.4%) underwent surgery at readmission. Simple appendicitis was diagnosed in 1768 and complex appendicitis in 596, while definitive results were not reported in 51 patients. Surgical patients with acute appendicitis had the highest white blood cell (WBC) values (13.89 (11.00–16.66)), whereas the mean C-reactive protein (CRP) values (75.8 (14–172)) were found to be highest in patients with other surgical abnormalities (except obstetric and gynaecological). Interestingly, mean CRP values were similar between surgically treated acute appendicitis patients (21.82 (5.2–71.6)) and those with negative laparotomy/laparoscopy findings (21.74 (4.9–63.9)). The mean Alvarado, RIPASA and AAS scores were significantly higher in patients with surgically treated acute appendicitis, whereas the mean AIR score was found to be higher in patients with other surgical abnormalities (Table 5).

Diagnostic performance of risk scoring systems

A total of 1125 (33.5%) patients were recorded as having been formally risk scored at admission by their clinical/surgical team. The Alvarado score was by far the most frequently used scoring system by attending physicians (1106 of 1125, 98.3%), while AIR scoring was performed in the remaining 19 patients. For the remaining 66.5% of patients, physicians did not formally utilize any of the scores in the initial patient management process, despite availability of data. These scores were calculated by RIFT researchers with use of the data collected as part of the study, given that all data allowing calculation were available. In the comparison of the risk scoring systems, it was found that the AAS score had the highest AUC and demonstrated 76.4% sensitivity and 49.4% specificity for its optimal cut-off point (more than 13 points). All scoring systems had significance for prediction, but AUC values were somewhat poor (<0.700) (Table 6). The primary analyses reported in Table 6 only included the surgical cohort ($n = 2610$). This approach was preferred as these patients represent the subgroup in which risk scoring would be most useful in a real-world scenario. However, in the interests of providing comprehensive data, ROC analysis results including the entire cohort ($n = 3358$; Table S1) have also been reported. ROC curves for all four risk scoring systems were created by the analysis of the surgical subgroup and the entire cohort (Figs. S1, S2).

Among the 1576 men who underwent surgery, 1511 (95.9%) had acute appendicitis and 65 (4.1%) had normal appendices. Only RIPASA score performance was significant to predict acute appendicitis in men (cut-off: more than 6.5, AUC: 0.580, $P = 0.036$). On the other hand, despite being marginally non-significant, AAS achieved the highest specificity (77.8%) in men (Table 7). Data including non-surgical men are detailed in Table S2.

A total of 1034 women had undergone surgery and 917 (88.7%) of these were diagnosed with acute appendicitis, while 117 (11.3%) had normal appendices. All scoring systems were successful in predicting acute appendicitis in women. AAS had the highest AUC (0.707) and a cut-off value of more than 13 points, whereas Alvarado achieved the highest specificity (72.4%) (Table 8). Data including non-surgical women are detailed in Table S3.

Finally, a total of 154 immigrants underwent surgery and 142 (92.2%) of these were ultimately discharged with a diagnosis of

Table 5 Clinical and operative features stratified according to pathology results

	Surgical cohort ($n = 2610$)				Non-surgical treatment ($n = 748$)	P
	Acute appendicitis ($n = 2428$)	Other surgical abnormality except for OB/GYN ($n = 61$)	OB/GYN surgical pathology ($n = 51$)	Negative laparotomy/ laparoscopy ($n = 70$)		
Age	32 (24–44)[a]	44 (30–59)[b]	32 (26–38)[a]	30.5 (24–41)[a]	33 (25–45)[a]	**<0.001**
<45	1850 (76.2)	32 (52.5)	44 (86.3)	55 (78.6)	559 (74.7)	**<0.001**
≥45	578 (23.8)[a]	29 (47.5)[b]	7 (13.7)[a]	15 (21.4)[a]	189 (25.3)[a]	
Sex						
Male	1511 (62.2)[a]	33 (54.1)[a,b]	0 (0.0)[c]	32 (45.7)[a,b]	335 (44.8)[b]	**<0.001**
Female	917 (37.8)	28 (45.9)	51 (100.0)	38 (54.3)	413 (55.2)	
Nationality						
Turkish	2286 (94.2)	57 (93.4)	45 (88.2)	68 (97.1)	706 (94.4)	**0.041**
Syrian	77 (3.2)	4 (6.6)	1 (2.0)	2 (2.9)	23 (3.1)	
Other immigrants	65 (2.7)[a]	0 (0.0)[a]	5 (9.8)[b]	0 (0.0)[a]	19 (2.5)[a]	
RIF pain history						
One	275 (11.4)	7 (11.5)	6 (12.2)	12 (17.1)	101 (13.5)	**<0.001**
Two or more	105 (4.3)[a]	6 (9.8)[b]	2 (4.1)[a]	7 (10.0)[b]	73 (9.8)[b]	
Abdominal surgery history	333 (13.7)[a]	14 (23.0)[b]	9 (17.6)[ab]	15 (21.4)[b]	162 (21.7)[b]	**<0.001**
White blood cell count ($\times10^3$), median (1st-3rd quartile)	13.89 (11.00–16.66)[a]	11.70 (8.70–15.21)[a,b]	9.34 (7.45–12.29)[b]	11.45 (8.58–15.09)[b]	10.52 (8.05–13.87)[b]	**<0.001**
Neutrophil ($\times10^3$), median (1st-3rd quartile)	10.80 (8.00–13.55)[a]	8.80 (6.84–12.20)[a,b]	6.15 (4.82–9.73)[c]	8.38 (5.00–11.48)[b,c]	7.20 (4.80–10.50)[c]	**<0.001**
CRP, median (1st-3rd quartile)	21.82 (5.2–71.6)[a]	75.8 (14–172)[b]	21.5 (5.45–64.7)[a]	21.74 (4.9–63.9)[a]	9.5 (2.48–40)[c]	**<0.001**
Surgery performed at						
First admission	2370 (97.6)[a]	54 (88.5)[b]	45 (88.2)[b]	66 (94.3)[a,b]	–	**<0.001**
Re-admission	58 (2.4)	7 (11.5)	6 (11.8)	4 (5.7)	–	
Operative approach						
Open—RIF incision	996 (52.0)[a]	19 (35.2)[a]	4 (8.3)[b]	23 (39.7)[a]	–	**<0.001**
Open—midline incision	69 (3.6)	11 (20.4)	2 (4.2)	4 (6.9)	–	
Laparoscopic	825 (43.0)[a]	19 (35.2)[a]	42 (87.5)[b]	29 (50.0)[a]	–	
Conversion to open	27 (1.4)	5 (9.3)	0 (0.0)	2 (3.4)	–	
Duration of hospital stay, median (1st-3rd quartile)	2 (1–2)[a]	3 (2–6)[b]	3 (2–3)[b]	1 (1–3)[a]	1 (0–1)[c]	**<0.001**
Alvarado score, mean(s.d.)	6.63(1.75)[a]	6.30(1.68)[a,b]	4.82(1.58)[c]	5.85(1.85)[b]	5.07(1.92)[c]	**<0.001**
AIR score, mean(s.d.)	6.75(1.88)[a]	7.09(2.09)[a]	4.75(2.11)[b]	6.07(1.88)[c]	5.24(2.01)[b]	**<0.001**
RIPASA score, mean(s.d.)	9.47(2.34)[a]	8.52(2.68)[b]	7.15(2.07)[c,d]	8.24(2.30)[b,c]	7.02(2.27)[d]	**<0.001**
AAS, mean(s.d.)	15.46(2.90)[a]	14.75(2.84)[a,b]	10.78(2.46)[c]	13.84(3.00)[b]	12.31(3.34)[d]	**<0.001**

Values are *n* (%) unless otherwise stated. Superscripted letters refer to pairwise comparison results, same letters indicate statistically similar values in the denoted groups. Bold values indicate statistically significant results. RIF, right iliac fossa; WBC, white blood cell count; CRP, C-reactive protein; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score; OB/GYN, obstetrics and gynecology.

Table 6 Performance of the scores to predict acute appendicitis in all operated on patients

	Alvarado	AIR	RIPASA	AAS
Cut-off	>5	>5	>7	>13
Sensitivity	74.5%	74.6%	81.0%	76.4%
Specificity	46.1%	44.3%	43.7%	49.4%
Accuracy	72.5%	72.4%	78.4%	74.4%
PPV	94.8%	94.4%	95.0%	95.0%
NPV	12.0%	12.2%	14.7%	14.3%
AUC (95% c.i.)	0.645 (0.603,0.686)	0.602 (0.555,0.650)	0.669 (0.626,0.712)	0.686 (0.645,0.727)
P	**<0.001**	**<0.001**	**<0.001**	**<0.001**

Bold values indicate statistically significant results. PPV, positive predictive value; NPV, negative predictive value; AUC, area under receiver operating characteristic (ROC) curve; c.i., confidence intervals; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

acute appendicitis. Twelve (7.8%) immigrants had normal appendices, yielding a NAR value of 7.8%. AAS was the only scoring system that demonstrated predictive significance among immigrants, with an AUC of 0.826 and a cut-off value of more than 13 points (*Table 9*). Analyses including non-surgical immigrants are reported in *Table S4*.

Analysis of risk categories

In addition to diagnostic performance analyses that were based on ROC results in operated on patients, each scoring system was evaluated according to the distribution of the entire population into risk categories based on literature-defined cut-off values. In overall analysis, it was found that being categorized as 'high

Table 7 Performance of the scores to predict acute appendicitis in operated on male patients

	Alvarado	AIR	RIPASA	AAS
Cut-off	>5	>6	>6.5	>17
Sensitivity	73.4%	54.2%	87.3%	33.3%
Specificity	35.9%	54.1%	30.0%	77.8%
Accuracy	71.9%	54.2%	85.0%	35.2%
PPV	96.4%	96.3%	96.8%	97.1%
NPV	5.5%	5.0%	8.9%	5.0%
AUC (95% c.i.)	0.562 (0.491,0.634)	0.536 (0.460,0.612)	0.580 (0.501,0.659)	0.567 (0.499,0.635)
P	0.092	0.345	**0.036**	0.072

Bold value indicate statistically significant results. PPV, positive predictive value; NPV, negative predictive value; AUC, area under receiver operating characteristic (ROC) curve; c.i., confidence intervals; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

Table 8 Performance of the scores to predict acute appendicitis in operated on female patients

	Alvarado	AIR	RIPASA	AAS
Cut-off	>6	>5	>7	>13
Sensitivity	55.8%	75.1%	77.1%	64.0%
Specificity	72.4%	50.4%	50.0%	66.1%
Accuracy	57.7%	72.1%	74.0%	64.2%
PPV	94.0%	91.7%	92.2%	93.3%
NPV	17.4%	21.7%	22.3%	19.8%
AUC (95% c.i.)	0.695 (0.645,0.744)	0.640 (0.580,0.700)	0.692 (0.642,0.742)	0.707 (0.657,0.756)
P	**<0.001**	**<0.001**	**<0.001**	**<0.001**

Bold values indicate statistically significant results. PPV, positive predictive value; NPV, negative predictive value; AUC, area under receiver operating characteristic (ROC) curve; c.i., confidence intervals; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

Table 9 Performance of the scores to predict acute appendicitis in operated on immigrants

	Alvarado	AIR	RIPASA	AAS
Cut-off	>5	>5	>7	>13
Sensitivity	82.4%	82.3%	87.4%	85.0%
Specificity	50.0%	50.0%	41.7%	63.6%
Accuracy	79.9%	79.9%	83.7%	83.2%
PPV	95.1%	95.3%	94.4%	96.2%
NPV	19.4%	18.5%	22.7%	28.0%
AUC (95% c.i.)	0.655 (0.504,0.806)	0.585 (0.375,0.795)	0.613 (0.434,0.792)	0.826 (0.714,0.937)
P	0.076	0.372	0.195	**<0.001**

Bold value indicate statistically significant results. PPV, positive predictive value; NPV, negative predictive value; AUC, area under receiver operating characteristic (ROC) curve; c.i., confidence intervals; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

Table 10 The frequencies of acute appendicitis diagnoses according to risk category in the whole study group

	Acute appendicitis		True appendicitis %	CR (fold)
	Yes	No		
Alvarado (*n*)	2395	915		
High (>8)	336 (14.0)	32 (3.5)	91.3	1.93
Intermediate (5–8)	1787 (74.6)	580 (63.3)	75.5	
Low (<5)	272 (11.3)	303 (33.1)	47.3*	
AIR (*n*)	2182	831		
High (>8)	383 (17.5)	57 (6.8)	87.0	1.87
Intermediate (5–8)	1542 (70.6)	480 (57.7)	76.3	
Low (<5)	257 (11.7)	294 (35.3)	46.6*	
RIPASA (*n*)	2314	887		
Definite (≥12)	419 (18.1)	41 (4.6)	91.1	3.88
High (7.5–11.5)	1456 (62.9)	338 (38.1)	81.2	
Low (5–7)	408 (17.6)	407 (45.8)	50.1	
Unlikely (<5)	31 (1.3)	101 (11.3)	23.5*	
AAS (*n*)	2236	874		
High (≥16)	1161 (51.9)	178 (20.3)	86.7	2.61
Intermediate (11–15)	950 (42.4)	444 (50.8)	68.1	
Low (0–10)	125 (5.6)	252 (28.8)	33.2*	

Values are *n* (%) unless otherwise indicated. CR was calculated by dividing the appendicitis percentage in the highest risk category by the percentage in the lowest risk category (true positive rate/failure rate). Percentages in parentheses are row percentages, showing the distribution of patients in each risk category. *Corresponds to failure rate. CR, categorization resolution; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

Table 11 The frequencies of acute appendicitis diagnoses according to risk category, divided by sex

Sex	Scoring system	Acute appendicitis		True appendicitis %	CR (fold)
		Yes	No		
Male	**Alvarado (*n*)**	1490	392		
	High (>8)	203 (13.6)	16 (4.0)	92.7	1.58
	Intermediate (5–8)	1112 (74.6)	253 (64.5)	81.5	
	Low (<5)	175 (11.7)	123 (31.3)	58.7*	
	AIR (*n*)	1355	355		
	High (>8)	239 (17.6)	29 (8.1)	89.2	1.54
	Intermediate (5–8)	956 (70.5)	210 (59.1)	82.0	
	Low (<5)	160 (11.8)	116 (32.6)	58.0*	
	RIPASA (*n*)	1445	377		
	Definite (≥12)	279 (19.3)	28 (7.4)	90.9	3.89
	High (7.5–11.5)	926 (64.0)	158 (41.9)	85.4	
	Low (5–7)	229 (15.8)	155 (41.1)	59.6	
	Unlikely (<5)	11 (0.7)	36 (9.5)	23.4*	
	AAS (*n*)	1384	377		
	High (≥16)	848 (61.2)	120 (31.8)	87.6	2.76
	Intermediate (11–15)	498 (35.9)	175 (46.4)	74.0	
	Low (0–10)	38 (2.7)	82 (21.7)	31.7*	
Female	**Alvarado (*n*)**	905	523		
	High (>8)	133 (14.6)	16 (3.1)	89.3	2.55
	Intermediate (5–8)	675 (74.5)	327 (62.5)	67.4	
	Low (<5)	97 (10.7)	180 (34.4)	35.0*	
	AIR (*n*)	827	476		
	High (>8)	144 (17.4)	28 (5.8)	83.7	2.37
	Intermediate (5–8)	586 (70.8)	270 (56.7)	68.5	
	Low (<5)	97 (11.7)	178 (37.3)	35.3*	
	RIPASA (*n*)	869	510		
	Definite (≥12)	140 (16.1)	13 (2.5)	91.5	3.89
	High (7.5–11.5)	530 (60.9)	180 (35.3)	74.6	
	Low (5–7)	179 (20.5)	252 (49.4)	41.5	
	Unlikely (<5)	20 (2.3)	65 (12.7)	23.5*	
	AAS (*n*)	852	497		
	High (≥16)	313 (36.7)	58 (11.6)	84.4	2.49
	Intermediate (11–15)	452 (53.0)	269 (54.1)	62.7	
	Low (0–10)	87 (10.2)	170 (34.2)	33.9*	

Values are *n* (%) unless otherwise indicated. CR was calculated by dividing the appendicitis percentage in the highest risk category by the percentage in the lowest risk category (true positive rate/failure rate). Percentages in parentheses are row percentages, showing the distribution of patients in each risk category. *Corresponds to failure rate. CR, categorization resolution; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

risk' by the Alvarado score resulted in the highest true positive rate (91.3%), followed by the 'definite risk' category of RIPASA (91.1%). When the lowest risk categories were examined, it was found that the RIPASA and the AAS had the lowest (best) failure rates (23.5% and 33.2% respectively). RIPASA and AAS also had the highest CR (3.88 and 2.61) values, indicating better performance in differentiating the likelihood of appendicitis in the highest and lowest risk groups (*Table 10*).

When assessed in only male patients, the 'high risk' category of Alvarado again proved to have the highest true positive rate (92.7%), whereas in female patients, the 'definite' category of RIPASA had the highest true positive rate (91.5%). Interestingly, for men, failure rates of the Alvarado and AIR scoring systems were 58.7% and 58.0% respectively (both were non-significant in ROC analyses), whereas the corresponding values in women were 35% and 35.3% respectively. Notwithstanding the fact that AAS was non-significant in the ROC analysis for men, the RIPASA and AAS systems again had the highest CR values in men (3.89 and 2.76). In women, RIPASA was also leading in this respect (3.89); however, Alvarado (2.55) had a greater CR compared with AAS (2.49) (*Table 11*).

Finally, in immigrants, all patients (100%) in the 'high risk' category of Alvarado were ultimately diagnosed with acute appendicitis, showing perfect true positive performance. RIPASA had 100% rule-out accuracy (0% failure rate) among patients in the 'unlikely' category, and therefore, the best CR. However, due to the limited number of patients in categories and the 0% failure rate for RIPASA, the CR value was calculated by combining the 'low' and 'unlikely' categories to determine failure rate. As described previously, AAS was the only scoring system that demonstrated diagnostic significance among immigrants. AAS also proved to be successful in terms of CR, as demonstrated by the fact that CR values were higher in immigrants compared with overall analysis (3.38 *versus* 2.61) (*Table 12*).

Discussion

This study found that women with acute RIF pain in Turkey had a lower likelihood of undergoing surgery than men (71.5% *versus* 82.5%); however, NAR in women was almost three-fold higher than in men (11.3% *versus* 4.1%). In the immigrant group, NAR was found to be slightly higher compared with NAR in the entire population (7.8% *versus* 7%). All scoring systems had successful prediction performance in overall analyses with variations based on sex and immigrant status; however, performance metrics, particularly sensitivity values, were notably low. In ROC analyses, RIPASA score was the only scoring system that was significant for men. In the immigrant group, AAS was the only significant predictor of acute appendicitis (cut-off point of more than 13) and had relatively high CR compared with the overall

Table 12 The frequencies of acute appendicitis diagnoses according to risk category among immigrants

	Acute appendicitis		True appendicitis %	CR (fold)
	Yes	No		
Alvarado (*n*)	142	53		
High (>8)	17 (11.9)	0 (0.0)	100.0	1.89
Intermediate (5–8)	107 (75.3)	37 (69.8)	74.3	
Low (<5)	18 (12.6)	16 (30.2)	52.9*	
AIR (*n*)	124	43		
High (>8)	31 (25.0)	4 (9.3)	88.6	1.94
Intermediate (5–8)	82 (66.1)	26 (60.4)	75.9	
Low (<5)	11 (8.87)	13 (30.2)	45.8*	
RIPASA (*n*)	135	53		
Definite (≥12)	31 (22.9)	2 (3.7)	93.9	2.71†
High (7.5–11.5)	87 (64.4)	19 (35.8)	82.1	
Low (5–7)	17 (12.5)	28 (52.8)	37.8	
Unlikely (<5)	0 (0.0)	4 (7.5)	0.0*	
AAS (*n*)	120	48		
High (≥16)	73 (60.8)	8 (16.6)	90.1	3.38
Intermediate (11–15)	43 (35.8)	29 (60.4)	59.7	
Low (0–10)	4 (3.33)	11 (22.9)	26.7*	

Values are *n* (%) unless otherwise indicated. CR was calculated by dividing the appendicitis percentage in the highest risk category by the percentage in the lowest risk category (true positive rate/failure rate). Percentages in parentheses are row percentages, showing the distribution of patients into each risk category. *Corresponds to failure rate. †Calculated by combining the 'low' and 'unlikely' categories due to 0% failure rate in the 'unlikely' category. CR, categorization resolution; AIR, Appendicitis Inflammatory Response; RIPASA, Raja Isteri Pengiran Anak Saleha Appendicitis; AAS, Adult Appendicitis Score.

results, indicating an important advantage in this population. The poor performance of Alvarado and AIR in this respect was associated with the high failure rates of these scoring systems among men (over 50% for both), which showed high failure and very low CR.

The first RIFT study was a dramatic example of different clinical trends regarding NAR, which was found to be 20% (392 of 1957) in the UK and 6.2% (54 of 868) in other participating countries[4]. One of the major factors with an impact on NAR is the use of diagnostic imaging[19,20], as exemplified by a study showing that NAR was reduced from 19% to 3.5% when preoperative imaging was employed[21]. Another nationwide audit reported a NAR value as low as 3.2% when cross-sectional imaging was employed in 99.7% of subjects[22]. The literature also exhibits the effect of female sex on NAR, which ranges from 4 to 35% throughout the world[21,23–30]. Studies from Turkey report NARs of 15.8% to 43.9%, and also demonstrate the higher frequencies among women and the positive impact of imaging[31–35]. In this study, USG (with or without CT) was used in 98.6% of patients who underwent surgery and overall NAR was found to be 7%.

Routine risk scoring has been found to be associated with reduced need for imaging and hospital admission, as well as reduced NAR[4,18,36]. This study revealed that the attending surgeons/clinical teams had used these models in only 33.8% of patients, despite the fact that necessary clinical data for each scoring system had been gathered. The most frequently used scoring system in clinical practice was Alvarado, followed by AIR in a very small number of subjects; however, based on the results, RIPASA and AAS outperformed these models in almost every metric. Since ROC analyses were employed to determine optimal cut-off values in the assessment of diagnostic performance, it must be kept in mind that these singular cut-off values should not be used to classify risk[7,17,18]. Further analysis of data with a purposefully selected study group is necessary to create risk categories, which will also necessitate the input of experts in the field. The current results, however, show that specific improvements are required to address the limitations of these models in the Turkish population. These efforts can enhance the value of risk prediction models and increase their utilization in clinical practice, thereby improving patient management and reducing costs.

In addition to determining cut-off values via ROC analyses, risk categories and calculated CR were assessed to further examine classification performances. Overall, the 'high risk' category of Alvarado resulted in the highest true positive rate (91.3%), while RIPASA, likely owing to its four-category classification approach, had the best CR values in all analyses. AAS also exhibited better CR compared with Alvarado and AIR, despite all three systems having three risk categories—revealing an important superiority. Therefore, utilizing RIPASA or AAS (and preferring AAS for immigrants) could be beneficial in the clinical setting, especially in high-volume centres where risk stratification can simplify and expedite patient management. It is also critical to note that the Alvarado and AIR systems, which were utilized by attending physicians, had the highest (worst) failure rates in overall analysis (47.3% and 46.6% respectively). These poor performances were largely associated with the exceptionally high failure rates in men (58.7 and 58% respectively).

An important factor to consider in this study was the restrictions on laparoscopic surgery due to the coronavirus disease (COVID)-19 pandemic[37]. The impact is evident through the substantial high frequency of open surgeries (54.3%) when compared with the recommended laparoscopic approach[7]. During the initial waves of the pandemic, open surgery was prioritized in an attempt to ensure the safety of healthcare workers[37]. Compared with the Prospective Observational Study on acute Appendicitis Worldwide (POSAW) study published in the pre-COVID era (116 surgical departments from 44 countries), the present study demonstrated a 10–12% shift towards open surgery preference. The RIFT Turkey data shows that 44.1% of the appendicectomies were performed laparoscopically (51.7% in POSAW), and 55.4% of the appendicectomies were open (42.2% in POSAW)[38].

Diagnostic imaging data shows that 56.8% of patients admitted for RIF pain had undergone USG and 75.2% had undergone CT. Overuse of imaging could lead to long-term issues with regard to healthcare costs. A high proportion of USG examinations is relatively acceptable in RIF pain, due to patient-related needs and also the interobserver variability of USG in different centres and in the hands of different operators (radiologists, emergency physicians or surgeons). Nevertheless, it is rather evident that CT should not be the first choice to rule out acute appendicitis; it should be reserved only for select patients, particularly when USG results are uncertain. The present study demonstrates an extremely high frequency of CT use in Turkey, but this could again be associated with the reluctance to perform USG examinations during the pandemic. Indeed, another UK-based study has drawn attention to the increased use of CT during the COVID-19 pandemic (from 36.3% to 85.9%), which resulted in a significant decrease in NAR (from 21.7% to 7.1%)[39]. Therefore, the overuse of CT may be another factor that reduced NAR in the present study.

Ultimately, the utilization of risk prediction models has been well documented to have positive effects in reducing unnecessary imaging tests, hospital admissions and surgeries[4,36]. Therefore, the authors recommend implementation of appropriate risk prediction models in routine clinical practice for patients presenting with acute RIF pain or suspected acute appendicitis. A careful/sparing incorporation of imaging tests is necessary, and it appears that the RIPASA and AAS systems could yield considerable benefits in this regard if they were to be utilized more widely. The fact that the Alvarado and AIR systems had greater than 50% failure rates in men is a critical concern, given that these models are quite clearly the only models that have been formally applied to patients. It is worth noting that AAS was the only system that showed significance in ROC analysis for immigrants, which is a critical finding with potential implications for risk assessment and management strategies in diverse healthcare settings.

A total of 84 hospitals contributed to data collection, almost all of them being the highest volume centres from the largest cities of Turkey. The findings are therefore broadly generalizable across Turkey. Nonetheless, there are several limitations that deserve mention. The present study aimed to include all patients admitted with RIF pain, but inclusion was not done at triage. This creates potential for selection bias towards relatively typical RIF pain and could explain the high rate of surgical intervention (77.7%) and the failure of scoring systems in terms of ruling out appendicitis. Triage may also lead to the underrepresentation of women, as they may have been diverted for obstetric/gynaecological assessment. However, it is critical to reiterate that all centres were secondary or tertiary healthcare institutions, ensuring uninterrupted availability of surgical teams. Second, epidemiological evidence suggests a weak but notable seasonal variation in appendicitis incidence[40]. Collecting year-long data could improve generalizability, but maintaining high-quality data collection for extended intervals may prove challenging. Third, the primary ROC analyses were restricted to the surgical cohort ($n = 2610$) to create a single cut-off for testing the utility of scores in a real-world scenario, and thus, patients excluded based on clinical examination and those who received non-surgical appendicitis treatment were not included in this analysis. ROC data concerning the entire population has been provided in the *Supplementary data*. Finally, it is essential to acknowledge that the classification analyses may have obscured the utility of intermediary risk categories, because CR was calculated based on the highest and lowest risk groups. However, this approach creates a reliable ratio to assess the discriminatory capabilities of each system, particularly since it is feasible to suggest that the intermediary groups may not alleviate diagnostic uncertainty.

The present nationwide audit offers valuable insights regarding the characteristics of patients with RIF pain in Turkey, as well as the importance of using appendicitis risk scoring systems. This information can assist clinicians in making informed decisions, improving diagnostic accuracy and optimizing patient care. NAR was low overall (7%) with a similar result in immigrants (7.8%); however, NAR was around three-fold greater in women compared with men (11.3% *versus* 4.1%), indicating the need for improvement in the evaluation of women. Clinicians in Turkey may be overutilizing imaging tests, particularly CT, leading to increased healthcare costs. Follow-up studies are needed to determine whether the increased use of CT due to the COVID-19 pandemic is becoming an unnecessary trend. Based on the results, risk scoring systems could alleviate the financial burden associated with imaging tests, particularly if the superior RIPASA and AAS scoring systems are adopted. Moreover, AAS may serve as the best risk stratification tool for immigrants; however, further evidence is necessary before drawing generalizable conclusions for diverse populations.

Collaborators

Ali Yalcinkaya (Gazi University Hospital, Ankara, Turkey); Ahmet Yalcinkaya (Uppsala University, Uppsala, Sweden); Can Keskin (Tibbi Akademik, Ankara, Turkey); Ibrahim Erkan (Tibbi Akademik, Ankara, Turkey); Sezai Leventoglu (Gazi University Hospital, Ankara, Turkey); Mehmet Caglikulekci (Istanbul Yeni Yuzyıl University Gaziosmanpasa Hospital, Istanbul, Turkey); Elbrus Zarbaliyev (Istanbul Yeni Yuzyıl University Gaziosmanpasa Hospital, Istanbul, Turkey); Murat Sevmis (Istanbul Yeni Yuzyıl University Gaziosmanpasa Hospital, Istanbul, Turkey); Yigit Ulgen (Bagcilar Training and Research Hospital, Istanbul, Turkey); Yuksel Altinel (Bagcilar Training and Research Hospital, Istanbul, Turkey); Serhat Meric (Bagcilar Training and Research Hospital, Istanbul, Turkey); Ahmet Akbas (Bagcilar Training and Research Hospital, Istanbul, Turkey); Nadir Adnan Hacim (Bagcilar Training and Research Hospital, Istanbul, Turkey); Talar Vartanoglu Aktokmanyan (Bagcilar Training and Research Hospital, Istanbul, Turkey); Yunus Emre Aktimur (Bagcilar Training and Research Hospital, Istanbul, Turkey); Fikret Calikoglu (Bagcilar Training and Research Hospital, Istanbul, Turkey); Hasim Furkan Gullu (Bagcilar Training and Research Hospital, Istanbul, Turkey); Ahmet Guray Durma (Bagcilar Training and Research Hospital, Istanbul, Turkey); Sami Acar (Zeynep Kamil Women's and Children's Diseases Training and Research Hospital, Istanbul, Turkey); Erman Ciftci (Zeynep Kamil Women's and Children's Diseases Training and Research Hospital, Istanbul, Turkey); Emre Balik (Koc University Hospital, Istanbul, Turkey); Cemil Burak Kulle (Koc University Hospital, Istanbul, Turkey); Ibrahim Halil Ozata (Koc University Hospital, Istanbul, Turkey); Tutku Tufekci (Koc University Hospital, Istanbul, Turkey); Cihad Tatar (Istanbul Training and Research Hospital, Istanbul, Turkey); Mert Mahsuni Sevinc (Istanbul Training and Research Hospital, Istanbul, Turkey); Husnu Sevik (Istanbul Training and Research Hospital, Istanbul, Turkey); Candeniz Ertürk (Istanbul Training and Research Hospital, Istanbul, Turkey); Irem Nur Kiraz (Istanbul Training and Research Hospital, Istanbul, Turkey); Volkan Ozben (Acibadem Mehmet Ali Aydinlar University Atakent Hospital, Istanbul, Turkey); Erman Aytac (Acibadem Mehmet Ali Aydinlar University Atakent Hospital, Istanbul, Turkey); Zumrud Aliyeva (Acibadem Mehmet Ali Aydinlar University Atakent Hospital, Istanbul, Turkey); Arda Ulas Mutlu

(Acibadem Mehmet Ali Aydinlar University Atakent Hospital, Istanbul, Turkey); Mert Tanal (University of Health Sciences, Sisli Hamidiye Etfal Training and Research Hospital, Istanbul, Turkey); Mustafa Fevzi Celayir (University of Health Sciences, Sisli Hamidiye Etfal Training and Research Hospital, Istanbul, Turkey); Emre Bozkurt (University of Health Sciences, Sisli Hamidiye Etfal Training and Research Hospital, Istanbul, Turkey); Sitki Gurkan Yetkin (University of Health Sciences, Sisli Hamidiye Etfal Training and Research Hospital, Istanbul, Turkey); Emin Ergin (University of Health Sciences, Sisli Hamidiye Etfal Training and Research Hospital, Istanbul, Turkey); Wafi Attaallah (Marmara University Hospital, Istanbul, Turkey); Tevfik Kivilcim Uprak (Marmara University Hospital, Istanbul, Turkey); Ahmet Omak (Marmara University Hospital, Istanbul, Turkey); Oguzhan Simsek (Marmara University Hospital, Istanbul, Turkey); Mehmet Abdussamet Bozkurt (Kanuni Sultan Suleyman Training and Research Hospital, Istanbul, Turkey); Yasin Kara (Kanuni Sultan Suleyman Training and Research Hospital, Istanbul, Turkey); Emre Bozdag (Kanuni Sultan Suleyman Training and Research Hospital, Istanbul, Turkey); Hakan Yirgin (Kanuni Sultan Suleyman Training and Research Hospital, Istanbul, Turkey); Adem Ozcan (Kanuni Sultan Suleyman Training and Research Hospital, Istanbul, Turkey); Nuri Okkabaz (Medipol University Medipol Mega Hospital, Istanbul, Turkey); Yasar Ozdenkaya (Medipol University Medipol Mega Hospital, Istanbul, Turkey); Mustafa Celalettin Haksal (Medipol University Medipol Mega Hospital, Istanbul, Turkey); Caglar Kazim Pekuz (Medipol University Medipol Mega Hospital, Istanbul, Turkey); Sila Duru (Medipol University Medipol Mega Hospital, Istanbul, Turkey); Emre Sivrikoz (Acıbadem Bakirkoy Hospital, Istanbul, Turkey); Yavuz Ozdemir (Pendik Yuzyil Private Hospital, Istanbul, Turkey); Necati Tan (Pendik Yuzyil Private Hospital, Istanbul, Turkey); Feza Yarbug Karayali (Baskent University Istanbul Hospital, Istanbul, Turkey); Abdulla Taghiyeva (Baskent University Istanbul Hospital, Istanbul, Turkey); Ismail Tirnova (Baskent University Istanbul Hospital, Istanbul, Turkey); Ilknur Erenler Bayraktar (Istanbul Florence Nightingale Hospital, Istanbul, Turkey); Onur Bayraktar (Istanbul Florence Nightingale Hospital, Istanbul, Turkey); Emine Zulal Emsal (Istanbul Florence Nightingale Hospital, Istanbul, Turkey); Munevver Irem Dalkilic (Istanbul Florence Nightingale Hospital, Istanbul, Turkey); Metin Yesiltas (Professor Cemil Tascioglu City Hospital, Istanbul, Turkey); Hasan Tok (Professor Cemil Tascioglu City Hospital, Istanbul, Turkey); Dursun Ozgur Karakas (Professor Cemil Tascioglu City Hospital, Istanbul, Turkey); Ali Pusane (Professor Cemil Tascioglu City Hospital, Istanbul, Turkey); Ali Ilbey Demirer (Professor Cemil Tascioglu City Hospital, Istanbul, Turkey); Hasan Berk Sahin (Professor Cemil Tascioglu City Hospital, Istanbul, Turkey); Ali Fuat Kaan Gok (Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey); Halil Alper Bozkurt (Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey); Mehmet Iskender Yildirim (Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey); Gorkem Uzunyolcu (Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey); Hakan Teoman Yanar (Istanbul University, Istanbul Faculty of Medicine, Istanbul, Turkey); Sefa Ergun (Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Istanbul, Turkey); Fadime Kutluk (Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Istanbul, Turkey); Server Sezgin Uludag (Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Istanbul, Turkey); Abdullah Kagan Zengin (Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Istanbul, Turkey); Mehmet Faik Ozcelik (Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Istanbul, Turkey); Ahmet Necati Sanli (Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Istanbul, Turkey); Yunus Emre Altuntas (Kartal Dr. Lutfi Kirdar City Hospital, Istanbul, Turkey); Ecem Memisoglu (Kartal Dr. Lutfi Kirdar City Hospital, Istanbul, Turkey); Ramazan Sari (Kartal Dr. Lutfi Kirdar City Hospital, Istanbul, Turkey); Osman Akdogan (Kartal Dr. Lutfi Kirdar City Hospital, Istanbul, Turkey); Hasan Fehmi Kucuk (Kartal Dr. Lutfi Kirdar City Hospital, Istanbul, Turkey); Omer Faruk Ozkan (Istanbul Umraniye Training and Research Hospital, Istanbul, Turkey); Hanife Seyda Ulgur (Istanbul Umraniye Training and Research Hospital, Istanbul, Turkey); Emre Furkan Kirkan (Istanbul Umraniye Training and Research Hospital, Istanbul, Turkey); Sema Yuksekdag (Istanbul Umraniye Training and Research Hospital, Istanbul, Turkey); Ahmet Rencuzogullari (Cukurova University Hospital, Adana, Turkey); Melik Kagan Aktas (Cukurova University Hospital, Adana, Turkey); Murat Aba (Cukurova University Hospital, Adana, Turkey); Ahmet Onur Demirel (Cukurova University Hospital, Adana, Turkey); Ismail Cem Eray (Cukurova University Hospital, Adana, Turkey); Burak Aydogan (Cukurova University Hospital, Adana, Turkey); Suleyman Cetinkunar (University of Health Sciences, Adana City Training and Research Hospital, Adana, Turkey); Kemal Yener (University of Health Sciences, Adana City Training and Research Hospital, Adana, Turkey); Alper Sozutek (University of Health Sciences, Adana City Training and Research Hospital, Adana, Turkey); Oktay Irkorucu (University of Health Sciences, Adana City Training and Research Hospital, Adana, Turkey); Mehmet Bayrak (Adana Ortadogu Private Hospital, Adana, Turkey); Yasemin Altintas (Adana Ortadogu Private Hospital, Adana, Turkey); Omer Alabaz (Adana Ortadogu Private Hospital, Adana, Turkey); Ahmet Atasever (Kahramanmaras Afsin District State Hospital, Kahramanmaras, Turkey); Guven Erdogrul (Kahramanmaras Afsin District State Hospital, Kahramanmaras, Turkey); Aydın Hakan Kupeli (Necip Fazil City Hospital, Kahramanmaras, Turkey); Bahtiyar Muhammedoglu (Necip Fazil City Hospital, Kahramanmaras, Turkey); Suleyman Kokdas (Necip Fazil City Hospital, Kahramanmaras, Turkey); Murat Kaya (Necip Fazil City Hospital, Kahramanmaras, Turkey); Erkan Uysal (Necip Fazil City Hospital, Kahramanmaras, Turkey); Ali Cihat Yildirim (University of Health Sciences, Kutahya Evliya Celebi Training and Research Hospital, Kutahya, Turkey); Sezgin Zeren (University of Health Sciences, Kutahya Evliya Celebi Training and Research Hospital, Kutahya, Turkey); Mehmet Fatih Ekici (University of Health Sciences, Kutahya Evliya Celebi Training and Research Hospital, Kutahya, Turkey); Mustafa Cem Algin (University of Health Sciences, Kutahya Evliya Celebi Training and Research Hospital, Kutahya, Turkey); Gultekin Ozan Kucuk (University of Health Sciences, Samsun Training and Research Hospital, Samsun, Turkey); Huseyin Eraslan (University of Health Sciences, Samsun Training and Research Hospital, Samsun, Turkey); Engin Aybar (University of Health Sciences, Samsun Training and Research Hospital, Samsun, Turkey); Suleyman Polat (University of Health Sciences, Samsun Training and Research Hospital, Samsun, Turkey); Alper Ceylan (University of Health Sciences, Samsun Training and Research Hospital, Samsun, Turkey); Ozgen Isik (Uludag University Hospital, Bursa, Turkey); Said Kural (Uludag University Hospital, Bursa, Turkey); Ahmet Aktas (Uludag University Hospital, Bursa, Turkey); Burak Bakar (Uludag University Hospital, Bursa, Turkey); Mustafa Yener Uzunoglu (Kestel State Hospital, Bursa,

Turkey); Baris Gulcu (Medicana Bursa Hospital, Bursa, Turkey); Ersin Ozturk (Medicana Bursa Hospital, Bursa, Turkey); Ali Onder Devay (Medicana Bursa Hospital, Bursa, Turkey); Ersoy Taspinar (Medicana Bursa Hospital, Bursa, Turkey); Ozkan Balcin (Bursa City Hospital, Bursa, Turkey); Fuat Aksoy (Bursa City Hospital, Bursa, Turkey); Gokhan Garip (Bursa City Hospital, Bursa, Turkey); Omer Yalkin (Bursa City Hospital, Bursa, Turkey); Nidal Iflazoglu (Bursa City Hospital, Bursa, Turkey); Direnc Yigit (Bursa City Hospital, Bursa, Turkey); Rumeysa Betul Kaya (Bursa City Hospital, Bursa, Turkey); Mustafa Ugur (Hatay Mustafa Kemal University Hospital, Hatay, Turkey); Erol Kilic (Hatay Mustafa Kemal University Hospital, Hatay, Turkey); Akin Dedemoglu (Hatay Mustafa Kemal University Hospital, Hatay, Turkey); Rasim Ersin Arslan (Hatay Mustafa Kemal University Hospital, Hatay, Turkey); Muhyittin Temiz (Hatay Mustafa Kemal University Hospital, Hatay, Turkey); Cengiz Aydin (University of Health Sciences, Tepecik Training and Research Hospital, Izmir, Turkey); Semra Demirli Atici (University of Health Sciences, Tepecik Training and Research Hospital, Izmir, Turkey); Tayfun Kaya (University of Health Sciences, Tepecik Training and Research Hospital, Izmir, Turkey); Selen Ozturk (University of Health Sciences, Tepecik Training and Research Hospital, Izmir, Turkey); Bulent Calik (University of Health Sciences, Tepecik Training and Research Hospital, Izmir, Turkey); Gizem Kilinc (University of Health Sciences, Tepecik Training and Research Hospital, Izmir, Turkey); Erdinc Kamer (Izmir Katip Celebi University Ataturk Training and Research Hospital, Izmir, Turkey); Turan Acar (Izmir Katip Celebi University Ataturk Training and Research Hospital, Izmir, Turkey); Nihan Acar (Izmir Katip Celebi University Ataturk Training and Research Hospital, Izmir, Turkey); Fevzi Cengiz (Izmir Katip Celebi University Ataturk Training and Research Hospital, Izmir, Turkey); Orhan Ureyen (University of Health Sciences, Izmir Bozyaka Training and Research Hospital, Izmir, Turkey); Sedat Tan (University of Health Sciences, Izmir Bozyaka Training and Research Hospital, Izmir, Turkey); Mehmet Yildirim (University of Health Sciences, Izmir Bozyaka Training and Research Hospital, Izmir, Turkey); Enver Ilhan (University of Health Sciences, Izmir Bozyaka Training and Research Hospital, Izmir, Turkey); Yigit Turk (Bakırçay University Cigli Training and Research Hospital, Izmir, Turkey); Ahmet Turan Durak (Urla State Hospital, Izmir, Turkey); Mehmet Yilmaz (Buca Seyfi Demirsoy Training and Research Hospital, Izmir, Turkey); Metin Mercan (Buca Seyfi Demirsoy Training and Research Hospital, Izmir, Turkey); Recep Atci (Buca Seyfi Demirsoy Training and Research Hospital, Izmir, Turkey); Selman Sokmen (Dokuz Eylul University Hospital, Izmir, Turkey); Tayfun Bisgin (Dokuz Eylul University Hospital, Izmir, Turkey); Tufan Egeli (Dokuz Eylul University Hospital, Izmir, Turkey); Yasemin Yildirim (Dokuz Eylul University Hospital, Izmir, Turkey); Turugsan Safak (Dokuz Eylul University Hospital, Izmir, Turkey); Kazim Celik (Tire State Hospital, Izmir, Turkey); Eyup Murat Yilmaz (Aydın Adnan Menderes University Hospital, Aydin, Turkey); Mahir Kirnap (Aydın Adnan Menderes University Hospital, Aydin, Turkey); Ahmet Ender Demirkiran (Aydın Adnan Menderes University Hospital, Aydin, Turkey); Ulas Utku Sekerci (Aydın Adnan Menderes University Hospital, Aydin, Turkey); Erkan Karacan (Aydin State Hospital, Aydin, Turkey); Ethem Bilgic (Didim State Hospital, Aydin, Turkey); Mehmet Mahir Ozmen (Liv Hospital Ankara, Ankara, Turkey); Cem Emir Guldogan (Liv Hospital Ankara, Ankara, Turkey); Emre Gundogdu (Liv Hospital Ankara, Ankara, Turkey); Munevver Moran (Liv Hospital Ankara, Ankara, Turkey); Timucin Erol (Hacettepe University Hospital, Ankara, Turkey); Hilmi Anil Dincer (Hacettepe University Hospital, Ankara, Turkey); Busenur Kirimtay (Hacettepe University Hospital, Ankara, Turkey); Sumeyye Yilmaz (Hacettepe University Hospital, Ankara, Turkey); Omer Cennet (Hacettepe University Hospital, Ankara, Turkey); Alp Yildiz (Yildirim Beyazit University Yenimahalle Training and Research Hospital, Ankara, Turkey); Aybala Yildiz (Yildirim Beyazit University Yenimahalle Training and Research Hospital, Ankara, Turkey); Can Sahin (Yildirim Beyazit University Yenimahalle Training and Research Hospital, Ankara, Turkey); Cihangir Akyol (Ankara University Ibni Sina Hospital, Ankara, Turkey); Mehmet Ali Koc (Ankara University Ibni Sina Hospital, Ankara, Turkey); Siyar Ersoz (Ankara University Ibni Sina Hospital, Ankara, Turkey); Anil Turhan (Ankara University Ibni Sina Hospital, Ankara, Turkey); Can Konca (Ankara University Ibni Sina Hospital, Ankara, Turkey); Tugan Tezcaner (Baskent University Ankara Hospital, Ankara, Turkey); Murathan Erkent (Baskent University Ankara Hospital, Ankara, Turkey); Onur Aydin (Baskent University Ankara Hospital, Ankara, Turkey); Tevfik Avci (Baskent University Ankara Hospital, Ankara, Turkey); Saygin Altiner (Ankara Training and Research Hospital, Ankara, Turkey); Igbal Osmanov (Memorial Ankara Hospital, Ankara, Turkey); Ahmet Cihangir Emral (Sincan State Hospital, Ankara, Turkey); Gokay Cetinkaya (Sincan State Hospital, Ankara, Turkey); Emin Lapsekili (Gulhane Training and Research Hospital, Ankara, Turkey); Merve Sakca (Gulhane Training and Research Hospital, Ankara, Turkey); Sebnem Cimen (Gulhane Training and Research Hospital, Ankara, Turkey); Dogan Ozen (Gulhane Training and Research Hospital, Ankara, Turkey); Erdem Baran Kozan (Gulhane Training and Research Hospital, Ankara, Turkey); Lutfi Dogan (Ankara Oncology Training and Research Hospital, Ankara, Turkey); Elifcan Haberal (Ankara Oncology Training and Research Hospital, Ankara, Turkey); Bengi Balci (Ankara Oncology Training and Research Hospital, Ankara, Turkey); Okan Kayhan (Ankara Oncology Training and Research Hospital, Ankara, Turkey); Bulent Aksel (Ankara Oncology Training and Research Hospital, Ankara, Turkey); Harun Karabacak (University of Health Sciences, Ankara Diskapi Yildirim Beyazid Training and Research Hospital, Ankara, Turkey); Cem Azili (University of Health Sciences, Ankara Diskapi Yildirim Beyazid Training and Research Hospital, Ankara, Turkey); Faruk Yazici (University of Health Sciences, Ankara Diskapi Yildirim Beyazid Training and Research Hospital, Ankara, Turkey); Muhammed Apaydin (University of Health Sciences, Ankara Diskapi Yildirim Beyazid Training and Research Hospital, Ankara, Turkey); Ismail Oskay Kaya (University of Health Sciences, Ankara Diskapi Yildirim Beyazid Training and Research Hospital, Ankara, Turkey); Erdinc Cetinkaya (Ankara City Hospital, Ankara, Turkey); Tezcan Akin (Ankara City Hospital, Ankara, Turkey); Gizem Gunes (Ankara City Hospital, Ankara, Turkey); Huseyin Turap (Ankara City Hospital, Ankara, Turkey); Deniz Aslan (Ankara City Hospital, Ankara, Turkey); Ali Eba Demirbag (Ankara City Hospital, Ankara, Turkey); Basak Bolukbasi (Gazi University Hospital, Ankara, Turkey); Berkay Enes Karaca (Gazi University Hospital, Ankara, Turkey); Ece Ozturk (Gazi University Hospital, Ankara, Turkey); Elif Ozeller (Gazi University Hospital, Ankara, Turkey); Gulsum Sueda Kayacan (Gazi University Hospital, Ankara, Turkey); Alp Ozgun Borcek (Gazi University Hospital, Ankara, Turkey); Ilhan Ece (Selcuk University Hospital, Konya, Turkey); Serdar Yormaz (Selcuk University Hospital, Konya, Turkey); Bayram Colak (Selcuk University Hospital, Konya, Turkey); Akin Calisir (Selcuk University Hospital, Konya, Turkey); Mustafa

Sahin (Selcuk University Hospital, Konya, Turkey); Kemal Arslan (Konya City Hospital, Konya, Turkey); Ismail Hasirci (Konya City Hospital, Konya, Turkey); Mehmet Esref Ulutas (Konya City Hospital, Konya, Turkey); Sukru Hakan Metin (Konya City Hospital, Konya, Turkey); Fatma Ayca Gultekin (Zonguldak Bulent Ecevit University Hospital, Zonguldak, Turkey); Zeynep Ozkan (Elazig Fethi Sekin City Hospital, Elazig, Turkey); Onur Ilhan (Elazig Fethi Sekin City Hospital, Elazig, Turkey); Tamer Gundogdu (Elazig Fethi Sekin City Hospital, Elazig, Turkey); Rumeysa Kevser Liman (Elazig Fethi Sekin City Hospital, Elazig, Turkey); Burhan Hakan Kanat (Elazig Fethi Sekin City Hospital, Elazig, Turkey); Altan Aydin (University of Health Sciences, Trabzon Kanuni Training and Research Hospital, Trabzon, Turkey); Ugur Sungurtekin (Pamukkale University Hospital, Denizli, Turkey); Utku Ozgen (Pamukkale University Hospital, Denizli, Turkey); Muhammed Rasid Aykota (Pamukkale University Hospital, Denizli, Turkey); Fatih Altintoprak (Sakarya University Training and Research Hospital, Sakarya, Turkey); Emre Gonullu (Sakarya University Training and Research Hospital, Sakarya, Turkey); Guner Cakmak (Sakarya University Training and Research Hospital, Sakarya, Turkey); Ugur Can Dulger (Sakarya University Training and Research Hospital, Sakarya, Turkey); Baris Mantoglu (Sakarya University Training and Research Hospital, Sakarya, Turkey); Hakan Demir (Sakarya University Training and Research Hospital, Sakarya, Turkey); Emrah Akin (Sakarya University Training and Research Hospital, Sakarya, Turkey); Erhan Eroz (Toyotasa Emergency Hospital, Sakarya, Turkey); Okay Nazli (Mugla Sitki Kocman University Training and Research Hospital, Mugla, Turkey); Ozcan Dere (Mugla Sitki Kocman University Training and Research Hospital, Mugla, Turkey); Mustafa Aykut Dadasoglu (Mugla Sitki Kocman University Training and Research Hospital, Mugla, Turkey); Eray Kara (Manisa Celal Bayar University Hospital, Manisa, Turkey); Semra Tutcu (Manisa Celal Bayar University Hospital, Manisa, Turkey); Ilhami Solak (Manisa Celal Bayar University Hospital, Manisa, Turkey); Ilayda Gencer (Manisa Celal Bayar University Hospital, Manisa, Turkey); Alperen Dalkiran (Manisa Celal Bayar University Hospital, Manisa, Turkey); Baris Sevinc (Usak Training and Research Hospital, Usak, Turkey); Omer Karahan (Usak Training and Research Hospital, Usak, Turkey); Nurullah Damburaci (Usak Training and Research Hospital, Usak, Turkey); Erdem Sari (Bandirma State Hospital, Balikesir, Turkey); Tamer Akay (Bandirma State Hospital, Balikesir, Turkey); Alpaslan Fedayi Calta (Bandirma State Hospital, Balikesir, Turkey); Abdullah Ozdemir (Bandirma State Hospital, Balikesir, Turkey); Nurian Ohri (Balikesir State Hospital, Balikesir, Turkey); Ilker Ermis (Kirikkale Yuksek Ihtisas Hospital, Kirikkale, Turkey); Osman Bozbiyik (Ege University Hospital, Izmir, Turkey); Murat Ozdemir (Ege University Hospital, Izmir, Turkey); Berk Goktepe (Ege University Hospital, Izmir, Turkey); Batuhan Demir (Ege University Hospital, Izmir, Turkey); Ozgur Kilincarslan (Ege University Hospital, Izmir, Turkey); Umut Riza Gunduz (Antalya Training and Research Hospital, Antalya, Turkey); Mehmet Olcum (Antalya Training and Research Hospital, Antalya, Turkey); Onur Ilkay Dincer (Antalya Training and Research Hospital, Antalya, Turkey); Remzi Can Cakir (Antalya Training and Research Hospital, Antalya, Turkey); Bulent Dinc (Antalya Training and Research Hospital, Antalya, Turkey); Enes Sahin (Kocaeli State Hospital, Kocaeli, Turkey); Emrah Uludag (Kocaeli State Hospital, Kocaeli, Turkey); Yusuf Arslan (Kocaeli State Hospital, Kocaeli, Turkey); Gokhan Posteki (Kocaeli State Hospital, Kocaeli, Turkey); Ahmet Oktay (Kocaeli State Hospital, Kocaeli, Turkey); Ozan Can Tatar (Kocaeli University Hospital, Kocaeli, Turkey); Sertac Ata Guler (Kocaeli University Hospital, Kocaeli, Turkey); Nihat Zafer Utkan (Kocaeli University Hospital, Kocaeli, Turkey); Serkan Tayar (Erzurum Regional Training and Research Hospital, Erzurum, Turkey); Yasar Copelci (Erzurum Regional Training and Research Hospital, Erzurum, Turkey); Murat Kartal (Erzurum Regional Training and Research Hospital, Erzurum, Turkey); Tolga Kalayci (Erzurum Regional Training and Research Hospital, Erzurum, Turkey); Mustafa Yeni (Erzurum Regional Training and Research Hospital, Erzurum, Turkey); Ahmet Cagri Buyukkasap (Siirt State Hospital, Siirt, Turkey); Selahattin Vural (Giresun University Faculty of Medicine Training and Research Hospital, Giresun, Turkey); Tugrul Kesicioglu (Giresun University Faculty of Medicine Training and Research Hospital, Giresun, Turkey); Ismail Aydin (Giresun University Faculty of Medicine Training and Research Hospital, Giresun, Turkey); Mehmet Gulmez (Giresun University Faculty of Medicine Training and Research Hospital, Giresun, Turkey); Can Saracoglu (Giresun University Faculty of Medicine Training and Research Hospital, Giresun, Turkey); Omer Topcu (Sivas Cumhuriyet University Hospital, Sivas, Turkey); Atilla Kurt (Sivas Cumhuriyet University Hospital, Sivas, Turkey); Sinan Soylu (Sivas Cumhuriyet University Hospital, Sivas, Turkey); Begum Kurt (Sivas Cumhuriyet University Hospital, Sivas, Turkey); Musa Serin (Sivas Cumhuriyet University Hospital, Sivas, Turkey); Salim Ilksen Basceken (Diyarbakir Gazi Yasargil Training and Research Hospital, Diyarbakir, Turkey); Ebubekir Gundes (Diyarbakir Gazi Yasargil Training and Research Hospital, Diyarbakir, Turkey); Mervan Savda (Diyarbakir Gazi Yasargil Training and Research Hospital, Diyarbakir, Turkey); Ali Zeynel Abidin Balkan (Diyarbakir Gazi Yasargil Training and Research Hospital, Diyarbakir, Turkey); Mehmet Nuri Yildiz (Diyarbakir Gazi Yasargil Training and Research Hospital, Diyarbakir, Turkey); Ali Uzunkoy (Harran University Training and Research Hospital, Sanliurfa, Turkey); Emre Karaca (Harran University Training and Research Hospital, Sanliurfa, Turkey); Ahmet Berkan (Harran University Training and Research Hospital, Sanliurfa, Turkey); Arda Isik (Erzincan University Hospital, Erzincan, Turkey); Yasin Alper Yildiz (Kastamonu Training and Research Hospital, Kastamonu, Turkey); Zafer Ergul (Kastamonu Training and Research Hospital, Kastamonu, Turkey); Necdet Fatih Yasar (Eskisehir Osmangazi University Hospital, Eskisehir, Turkey); Bartu Badak (Eskisehir Osmangazi University Hospital, Eskisehir, Turkey); Ata Ozen (Eskisehir Osmangazi University Hospital, Eskisehir, Turkey); Melih Velipasaoglu (Eskisehir Osmangazi University Hospital, Eskisehir, Turkey); Iyimser Ure (Eskisehir Osmangazi University Hospital, Eskisehir, Turkey).

Funding

The authors have no funding to declare.

Acknowledgements

Ali Yalcinkaya and Ahmet Yalcinkaya contributed equally to this work and share first authorship. We thank Tibbi Akademik (www.tibbiakademik.com) for excellent statistical and scientific guidance.

Disclosure

The authors declare no conflict of interest.

Supplementary material

Supplementary material is available at *BJS Open* online.

Data availability

All data are available from the corresponding author upon reasonable request.

Author contributions

Ali Yalcinkaya (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Writing—original draft, Writing—review & editing), Ahmet Yalcinkaya (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review & editing), Bengi Balci (Conceptualization, Data curation, Methodology, Writing—original draft, Writing—review & editing), Can Keskin (Data curation, Formal analysis, Investigation, Validation, Visualization, Writing—original draft, Writing—review & editing), Ibrahim Erkan (Data curation, Formal analysis, Investigation, Software, Validation, Writing—original draft, Writing—review & editing), Alp Yildiz (Data curation, Validation, Writing—review & editing), Erdinc Kamer (Conceptualization, Project administration, Resources, Supervision, Writing—review & editing) and Sezai Leventoglu (Conceptualization, Project administration, Resources, Supervision, Writing—review & editing).

References

1. Teng TZJ, Thong XR, Lau KY, Balasubramaniam S, Shelat VG. Acute appendicitis–advances and controversies. *World J Gastrointest Surg* 2021;**13**:1293–1314
2. Moris D, Paulson EK, Pappas TN. Diagnosis and management of acute appendicitis in adults, a review. *JAMA* 2021;**326**: 2299–2311
3. Abdelhadi MS. Acute abdominal pain in women of child-bearing age remains a diagnostic dilemma. *J Family Community Med* 2001; **8**:45–50
4. Bhangu A, RIFT Study Group on behalf of the West Midlands Research Collaborative. Evaluation of appendicitis risk prediction models in adults with suspected appendicitis. *Br J Surg* 2020;**107**:73–86
5. Alhamdani YF, Rizk HA, Algethami MR, Algarawi AM, Albadawi RH, Faqih SN *et al.* Negative appendectomy rate and risk factors that influence improper diagnosis at King Abdulaziz University Hospital. *Mater Sociomed* 2018;**30**:215–220
6. Ahmed HO, Muhedin R, Boujan A, Aziz AHS, Abdulla AM, Hardi RA *et al.* A five-year longitudinal observational study in morbidity and mortality of negative appendectomy in Sulaimani teaching Hospital/Kurdistan Region/Iraq. *Sci Rep* 2020;**10**:2028
7. Di Saverio S, Birindelli A, Kelly MD, Catena F, Weber DG, Sartelli M *et al.* WSES Jerusalem guidelines for diagnosis and treatment of acute appendicitis. *World J Emerg Surg* 2016;**11**:34
8. Gorter RR, Eker HH, Gorter-Stam MA, Abis GS, Acharya A, Ankersmit M *et al.* Diagnosis and management of acute appendicitis. EAES consensus development conference 2015. *Surg Endosc* 2016;**30**:4668–4690
9. RIFT Study Group On behalf of the West Midlands Research Collaborative. Right iliac fossa pain treatment (RIFT) study, protocol for an international, multicentre, prospective observational study. *BMJ Open* 2018;**8**:e017574
10. Schneider C, Kharbanda A, Bachur R. Evaluating appendicitis scoring systems using a prospective pediatric cohort. *Ann Emerg Med* 2007;**49**:778–84.e1
11. Mantoglu B, Gonullu E, Akdeniz Y, Yigit M, Firat N, Akin E *et al.* Which appendicitis scoring system is most suitable for pregnant patients? A comparison of nine different systems. *World J Emerg Surg* 2020;**15**:34
12. Shuaib A, Shuaib A, Fakhra Z, Marafi B, Alsharaf K, Behbehani A. Evaluation of modified Alvarado scoring system and RIPASA scoring system as diagnostic tools of acute appendicitis. *World J Emerg Med* 2017;**8**:276–280
13. Akin-Sabuncu S, Chen Lee C, Knight-Manuel M. Opening Pandora's box: a culturally relevant perspective on elementary teachers' approaches to educating immigrant and refugee students in Turkey. *Eur J Teacher Educ* 2023:1–19. doi: 10.1080/02619768.2023.2181072
14. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;**61**:344–349
15. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med* 1986;**15**:557–564
16. 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**: 1843–1849
17. Chong CF, Adi MI, Thien A, Suyoi A, Mackie AJ, Tin AS *et al.* Development of the RIPASA score: a new appendicitis scoring system for the diagnosis of acute appendicitis. *Singapore Med J* 2010;**51**:220–225
18. Sammalkorpi HE, Mentula P, Savolainen H, Leppäniemi A. The introduction of adult appendicitis score reduced negative appendectomy rate. *Scand J Surg* 2017;**106**:196–201
19. Raja AS, Wright C, Sodickson AD, Zane RD, Schiff GD, Hanson R *et al.* Negative appendectomy rate in the era of CT: an 18-year perspective. *Radiology* 2010;**256**:460–465
20. Bakker OJ, Go PM, Puylaert JB, Kazemier G, Heij HA; Werkgroup en klankbordgroup "Richtlijn acute appendicitis". Guideline on diagnosis and treatment of acute appendicitis: imaging prior to appendectomy is recommended. *Ned Tijdschr Geneeskd* 2010; **154**:A303
21. Tseng J, Cohen T, Melo N, Alban RF. Imaging utilization affects negative appendectomy rates in appendicitis: an ACS-NSQIP study. *Am J Surg* 2019;**217**:1094–1098
22. van Rossem CC, Bolmers MD, Schreinemacher MH, van Geloven AA, Bemelman WA, Snapshot Appendicitis Collaborative Study Group. Prospective nationwide outcome audit of surgery for suspected acute appendicitis. *Br J Surg* 2016;**103**:144–151
23. Seetahal SA, Bolorunduro OB, Sookdeo TC, Oyetunji TA, Greene WR, Frederick W *et al.* Negative appendectomy: a 10-year review of a nationally representative sample. *Am J Surg* 2011;**201**:433–437
24. Hatem F, Baig H, Khaldas F, Lucocq J. Negative appendicectomy rates in females of childbearing age, a retrospective analysis and literature review. *Cureus* 2022;**14**:e27412
25. Wagner PL, Eachempati SR, Soe K, Pieracci FM, Shou J, Barie PS. Defining the current negative appendectomy rate: for whom is preoperative computed tomography making an impact? *Surgery* 2008;**144**:276–282

26. Malkomes P, Edmaier F, Liese J, Reinisch-Liese A, El Youzouri H, Schreckenbach T *et al.* DIALAPP: a prospective validation of a new diagnostic algorithm for acute appendicitis. *Langenbecks Arch Surg* 2021;**406**:141–152
27. Flum DR, McClure TD, Morris A, Koepsell T. Misdiagnosis of appendicitis and the use of diagnostic imaging. *J Am Coll Surg* 2005;**201**:933–939
28. Liese J, Halbinger TM, Ulrich F, Bechstein WO, Strey CW. Appendicitis—the balance between cost effectiveness and safety remains challenging. *Langenbecks Arch Surg* 2014;**399**: 493–501
29. Drake FT, Florence MG, Johnson MG, Jurkovich GJ, Kwon S, Schmidt Z *et al.* Progress in the diagnosis of appendicitis, a report from Washington state's surgical care and outcomes assessment program. *Ann Surg* 2012;**256**:586–594
30. Coursey CA, Nelson RC, Patel MB, Cochran C, Dodd LG, Delong DM *et al.* Making the diagnosis of acute appendicitis, do more preoperative CT scans mean fewer negative appendectomies? A 10-year study. *Radiology* 2010;**254**: 460–468
31. Engin O, Calik B, Yildirim M, Coskun A, Coskun GA. Gynecologic pathologies in our appendectomy series and literature review. *J Korean Surg Soc* 2011;**80**:267–271
32. Koç C, Akbulut S, Coşkun EI, Sarıcı B, Yılmaz S. Comparison of the demographic and clinical features of pregnant and non-pregnant patients undergoing appendectomy. *Ulus Travma Acil Cerrahi Derg* 2020;**26**:555–562
33. Cakmak G, Mantoglu B, Gonullu E, Ozdemir K, Kamburoglu B. The comparison of clinical features between patients with positive and negative appendectomy. *Van Med J* 2021;**28**:479–485
34. Memisoglu K, Karip B, Mestan M, Onur E. The value of preoperative diagnostic tests in acute appendicitis, retrospective analysis of 196 patients. *World J Emerg Surg* 2010;**5**:5
35. Özsoy Z, Yenidoğan E. Evaluation of the Alvarado scoring system in the management of acute appendicitis. *Turk J Surg* 2017;**33**:200–204
36. Andersson M, Kolodziej B, Andersson RE; STRAPPSCORE Study Group. Randomized clinical trial of Appendicitis Inflammatory Response score-based management of patients with suspected appendicitis. *Br J Surg* 2017;**104**:1451–1461
37. Seehra J, Isherwood J, Verma A. Does the risk of SARS-COVID-19 at laparoscopy justify the precautions? *Br J Surg* 2020;**107**:e525
38. Sartelli M, Baiocchi GL, Di Saverio S, Ferrara F, Labricciosa FM, Ansaloni L *et al.* Prospective Observational Study on acute Appendicitis Worldwide (POSAW). *World J Emerg Surg* 2018;**13**:19
39. Mowbray NG, Hurt L, Powell-Chandler A, Reeves N, Chandler S, Walters E *et al.* Where have all the appendicectomies gone? *Ann R Coll Surg Engl* 2021;**103**:250–254
40. Fares A. Summer appendicitis. *Ann Med Health Sci Res* 2014;**4**: 18–21