Review

ANNALS OF MEDICINE & SURGERY

OPEN

A systematic review and meta-regression for validation of the Alvarado score as a tool for predicting acute appendicitis

Sapna Gupta, MRCS[a,*], Venkata S. Kolli, MRCS[a], Kimberly Da Costa, MRCS[a], Sariya Javed, MRCS[b], Ahmed Ammar, MRCS[c], Ashraf Rasheed, FRCS[a]

Background: The Alvarado score (AS) has not been widely used for diagnosing acute appendicitis although it has shown to be a good predictor for diagnosing appendicitis. The aim was to perform a systematic review of the available literature and synthesize the evidence.

Methods: A systematic review was performed as per the PRISMA guidelines using search engines like Ovid, PubMed, and Google Scholar with predefined, strict inclusion and exclusion criteria. The quality assessment of included studies was performed using the QUADAS 2 tool. Summary statistics were performed for all variables. A linear regression model was performed between dependent and independent variables using STATA software. Heterogeneity testing showed significant heterogeneity within the included studies; hence, a forest plot with pooled estimates could not be constructed, and therefore a meta-regression was performed.

Results: Seventeen full-text articles met inclusion and exclusion criteria. Ten of which were identified as low-risk studies. Five studies were included in final data pooling with total patients being 2239 and mean age of 31.9 years. (1) Linear regression demonstrated an association between 'histological appendicitis' and 'AS 7–0' with patients receiving intervention, with a significant *P* value of less than 0.005. (2) Meta-regression demonstrated a positive coefficient (0.298), a positive *Z* score of 2.20 with a significant *P* value of 0.028 for patients with 'high AS' who received interventions that were significantly proven to be 'histologically appendicitis', indicating a cause-and-effect relationship.

Conclusion: High AS (7 and above) is a significant predictor of acute appendicitis. The authors recommend further prospective randomized clinical trials to establish a cause-and-effect relationship.

Keywords: acute appendicitis, Alvarado score, appendicitis scoring

HIGHLIGHTS

- Alvarado score for predicting or diagnosing appendicitis.
- Histological confirmation of appendicitis is more likely if the patient has a high.
- Alvarado score.

Introduction

Acute appendicitis (AA) is one of the most common causes of acute abdominal pain requiring surgical intervention. Historically, a description of the appendix was first provided by an anatomist, Berengario de Carpi, in 1521[1]. In 1880, Robert Lawson Tait made the first diagnosis of appendicitis and surgically removed the appendix[2]. In 1886, pathologist-physician Reginald Fitz gave a public lecture and published a study on the traditional signs and symptoms of appendicitis and named the procedure appendectomy[1–4]. Initially, the diagnosis of appendicitis was made by surgeons clinically in patients presenting with right lower quadrant pain. Charles McBurney defined McBurney's point in 1891 as the exact point on the abdomen where tenderness is maximal in cases of AA[5,6], to aid in the diagnosis of appendicitis. AA occurs year-round, but certain months have higher rates. Some authors have linked AA, demographics, and seasons [7].

Since 1891, many advancements have been made in diagnosing AA. Many studies have suggested that blood tests like white blood cell count, neutrophil count, and C-reactive protein, as well as imaging modalities like ultrasonography scan and computed tomography (CT) scan, have helped improve the diagnosis of AA over the years[8–11]. However, there were also concerns about the accuracy of these modalities[10,11]. Raised white cell count had high diagnostic sensitivity but poor specificity, consistent with its

[a]Royal Gwent Hospital, Newport, Wales, [b]Queen Elizabeth University Hospital, Glasgow and [c]Royal Derby Hospital, Derby, UK

Alvarado score for predicting or diagnosing appendicitis.

Using a scoring system may help reducing negative appendicectomy rate.

Histological confirmation of appendicitis is more likely if the patient has a high Alvarado score.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding author. Address: Royal Gwent Hospital Newport, Wales NP20 2UB, UK E-mail address: sapnagupta@doctors.net.uk (S. Gupta).

Annals of Medicine & Surgery (2023) 85:111–121

Received 2 October 2022; Accepted 25 December 2022

Published online 7 February 2023

http://dx.doi.org/10.1097/MS9.0000000000000238

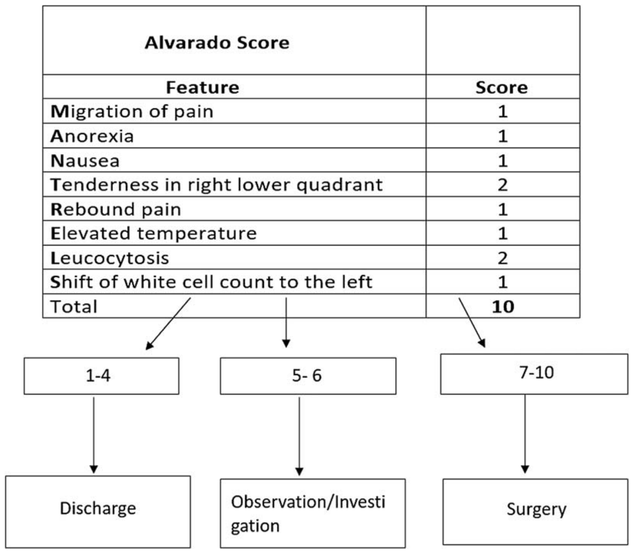

Figure 1. Alvarado score 10-point scoring system with suggested outcome.

nonspecific inflammatory marker[12–15]. C-reactive protein is considered a poor marker of early uncomplicated appendicitis; however, it is a better indicator for complicated or late appendicitis[13–16]. Whereas ultrasonography scan has problems such as operator dependent variability, difficult visualization of the appendix due to high BMI, anatomical variation, and overlying bowel gas, CT has its own limitations such as radiologist reporting being a limiting factor, high radiation exposure, contrast-related complications, and high cost[17,18].

With a reported lifetime risk of AA of 8.6% in males and 6.7% in females[19–21], an accurate diagnosis of AA is important. Despite its prevalence, diagnosis is difficult due to nonspecific symptoms and atypical manifestations[22]. In 1986, Alfredo Alvarado from Florida, USA, designed a 10-point scoring system based on signs, symptoms, and blood tests in patients presenting with suspected appendicitis[23] (Fig. 1).

Over the years, many other scoring systems have been discovered in order to aid in predicting the risk of AA. The Alvarado score (AS) is the most studied scoring system and has reported sensitivity of between 94.7 and 99%, specificity of 94.4 and 100%[24,25] and a negative predictive value of 97.4% for patients with suspected appendicitis[26,27]. AS helps in risk stratification for patients by dividing them into groups like 'discharge', 'observation', or 'surgical intervention' based on their score [23].

Recent international guidelines suggest routine use of scoring systems to improve the diagnosis of AA[28,29]. AS and appendicitis inflammatory response score have been recommended most frequently, but none of them have been widely accepted[28–30]. Another document from the American College of Physicians has recommended only AS to be more useful in predicting the presence and absence of appendicitis[31]. Even though this evidence is not very strong, it is still helpful.

The above-mentioned guidelines suggest discharging low-risk patients with worsening advice. However, as these guidelines are not in routine use, patients end up overstaying, being over investigated, and undergoing diagnostic laparoscopy with or without appendicectomy, which can subsequently turn out to be negative findings for appendicitis after histopathology confirmation. All this will have a negative impact on the patients with all the obvious risks of undergoing surgery, increased cost to the healthcare system, or being unnecessary and expensive treatment for the patients in the countries where patients fund their own treatment.

Most recent studies on negative appendicectomy rate (NAR) have reported NAR between 15 and 34%[32–35]. The NAR of the UK has been reported to be higher than other developed countries like the USA or the Netherlands[36,37], which is not acceptable in the modern era with all the technologies available. Multiple studies have reported an improvement in the NAR when a combination of AS and imaging has been used for preoperative assessment[38,39].

The AS was designed more than 3 decades ago and has been reported in several studies as a useful tool[38,39]. Our main aim of this study is to perform a systematic review and meta-analysis to validate the 'AS' as a tool for predicting appendicitis.

Methods

Search strategy

An electronic search was performed using search engines like Ovid, PubMed, and Google Scholar. MeSH terms such as 'suspected appendicitis' OR 'Alvarado score' OR 'prediction models' OR 'algorithm' were used in combination. The search strategy was in line with PRISMA guidelines and is presented in the PRISMA flow chart (Fig. 2). This systematic review has been registered with the Research Registry, with a unique identification number (UIN)-'reviewregistry1496' [40].

The work has been reported in line with the PRISMA, Supplemental Digital Content 1, http://links.lww.com/MS9/A18 criteria[41] (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and AMSTAR-2 (Assessing the methodological quality of systematic reviews) guidelines[42], Supplemental Digital Content 2, http://links.lww.com/MS9/A19.

Based on the AMSTAR-2 checklist the compliance of this systematic review is high.

Study selection

Inclusion and exclusion criteria were predefined and a consensual agreement was reached by all the co-authors in selecting studies.

Studies with patients data who were 16 years of age or older, had suspected appendicitis, had no prior investigations, and which were published in English exclusively met the inclusion criteria for this review.

Studies were excluded if they have included patients younger than 16, women who were pregnant, or used a scoring system other than the AS. Previously published systematic reviews or survey articles, case reports, opinions or letter to editors were also excluded.

Two reviewers completed the review process. After assessing the study titles and abstracts independently, a decision was made regarding the studies that should be reviewed along with full-text. Full-text articles were reviewed by the same reviewers, and any disagreements were resolved by discussion and by strictly following inclusion and exclusion criteria.

Quality assessment, data pooling, and statistical analysis

The quality assessment of included studies was performed using the 'Quality assessment of studies of diagnostic accuracy included in systematic reviews' (QUADAS 2) tool, which included the risk

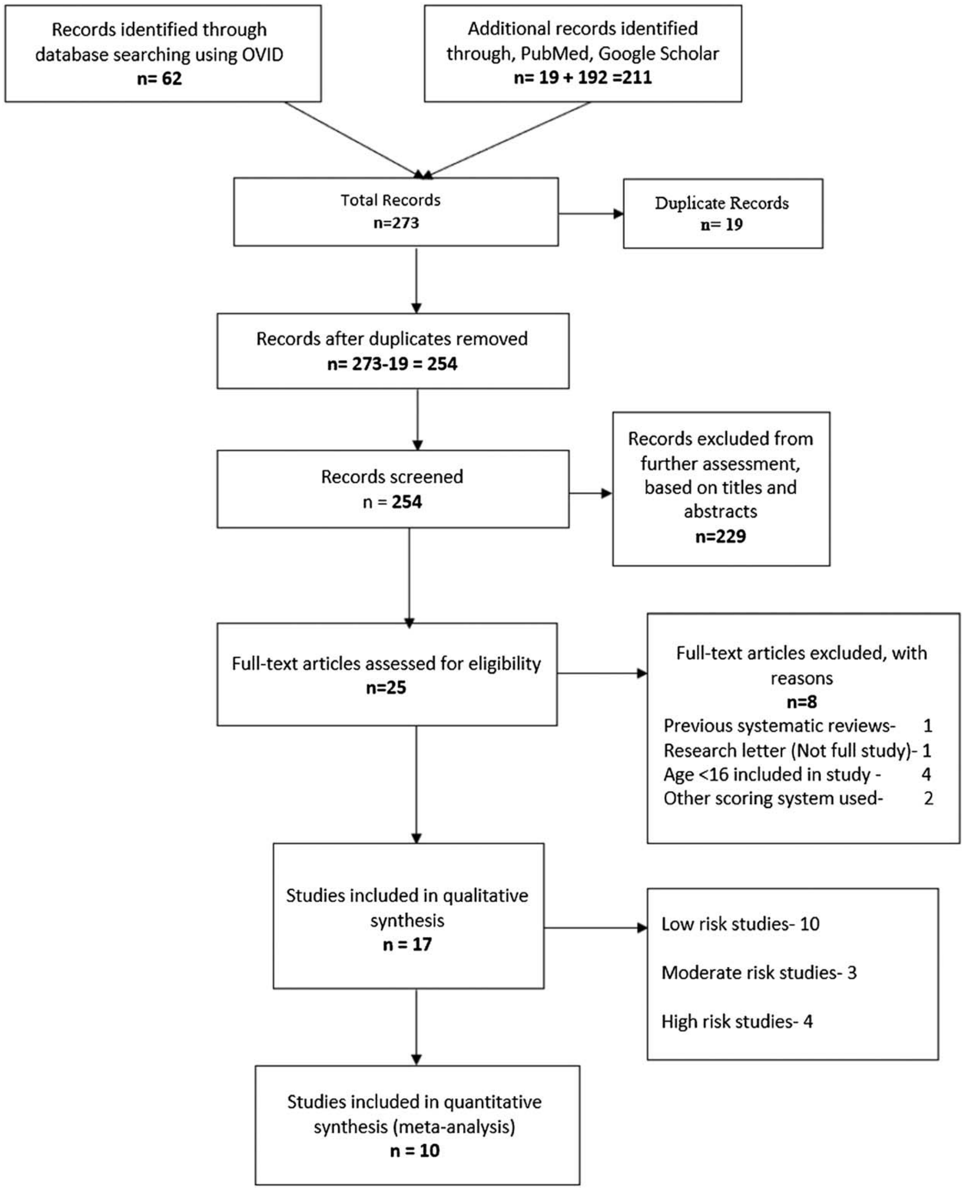

Figure 2. PRISMA flow chart.

of bias. A summary of QUADAS 2 of the included papers is presented as Table 1.

A separate assessment of the risk of bias in included studies was conducted using the ROBINS-I tool, which is also presented as Table 2.

Quality assessment and risk of bias assessment were performed by the same two reviewers, and any disagreements were resolved by discussion with the third reviewer.

Only low-risk studies were included in the data pooling. After pooling the data of low-risk studies with different variables like total number of patients included in studies, males, females, mean age, patients with AS 0–6, patients with AS 7–10, patients with no surgical intervention, patients with surgical intervention, histologically normal appendix, histologically proven appendicitis. Studies with a lot of missing data in the above-described variables were removed from final analysis.

Summary statistics were performed using all the included variables. A linear regression model was applied between the dependent variables like patients with no surgical intervention, patients with surgical intervention, histological appendicitis, and histological normal appendix, and independent variables like AS 0–6, AS 7–10, and patients who received surgical intervention in different combinations of statistical equations.

Table 1

QUADADS-2 quality assessment of studies of diagnostic accuracy included in systematic reviews

References	Risk of bias				Applicability concerns			Level of risk	Level of evidence of study
	Patient selection	Index test	Reference standards	Flow and timing	Patient selection	Index test	Reference standards		
Mariadason *et al.*[43]	😐	🙂	🙂	🙁	😐	🙂	🙂	Moderate risk study	2
Ramez *et al.*[44]	😐	🙂	🙂	🙂	😐	🙂	🙂	Low-risk study	1
Coleman *et al.*[45]	🙂	🙂	🙂	🙁	🙂	🙂	🙂	Low-risk study	2
Apisarnthanarak *et al.*[46]	🙂	🙂	🙂	🙁	🙂	🙂	🙂	Low-risk study	2
Tan *et al.*[47]	😐	🙂	🙂	🙁	😐	🙂	🙂	Moderate risk study	2
Meltzer *et al.*[48]	🙂	🙂	🙂	🙂	🙂	🙂	🙂	Low-risk study	2
Wang *et al.*[49]	😐	😐	🙂	🙂	😐	😐	🙂	High-risk study	2
Mishra *et al.*[50]	😐	🙂	🙂	🙂	😐	🙂	🙂	Low-risk study	2
Waheed *et al.*[51]	😐	🙂	🙂	🙁	😐	🙂	🙂	Moderate risk study	3
de *et al.*[52]	😐	🙁	🙁	🙂	😐	🙁	🙁	High-risk study	2
Neupane *et al.*[53]	😐	🙂	🙂	🙂	😐	🙂	🙂	Low-risk study	2
Tan *et al.*[54]	😐	🙂	🙂	🙂	😐	🙂	🙂	Low-risk study	1
Mukhopadhyay *et al.*[55]	🙂	🙂	🙂	🙂	🙂	🙂	🙂	Low-risk study	2
Lamparelli *et al.*[56]	🙁	🙂	🙂	🙂	🙁	🙂	🙂	High-risk study	1
Lopez *et al.*[57]	🙂	🙂	🙂	🙂	🙂	🙂	🙂	Low-risk study	1
Tan *et al.*[58]	🙂	🙂	🙂	🙂	🙂	🙂	🙂	Low-risk study	1
Seshadri *et al.*[59]	🙂	🙂	🙁	🙁	🙂	🙂	🙁	High-risk study	3

🙂, low-risk of bias; 😐, moderate risk of bias; 🙁, high-risk of bias.

Heterogeneity was assessed with 'I^2', 'τ^2' and 'Q' statistics during the meta-regression using STATA software. Significant heterogeneity was identified within the included studies; hence, a forest plot would not be meaningful to demonstrate effect size.

A meta set was built with sample size (n) as 'patients with histologically proven appendicitis' and the mean ('μ') in the model was AS 7–10. The SD in this meta set was 'patients with normal appendix on histology'. With this meta set, a meta-regression analysis was performed to see if there were any differences between patients who received intervention and those who did not.

Results were presented as summary statistics, linear regression models, and meta-regression.

Results

Literature search

Seventeen full-text articles met the inclusion criteria and were appraised following the literature search as shown in the PRISMA flow chart (Fig. 2).

Risk of bias assessment of studies

The risk of bias in nonrandomized studies of interventions (ROBINS-I) tool was used for risk of bias assessment in the included studies in our systematic review (Table 2). Both reviewers who did the initial review agreed to exclude the ROBINS-I 'serious' risk of bias studies from our systematic review and only included the 'moderate' risk of bias studies.

Quality assessment of studies

The QUADAS 2 tool (Table 1) was applied to 17 studies that were included in the qualitative synthesis. Seven out of 17 (42%) studies were moderate or high-risk studies, due to specific inclusion and exclusion criteria and were excluded from the final analysis.

Results of qualitative synthesis

Data pooling was performed from 10 low-risk studies. Due to lot of missing data in five studies, they were excluded from the qualitative synthesis (Table 3). This led to five studies being used for final pooled data analysis and estimates.

Studies included for meta-regression analysis

After doing a qualitative synthesis, only five studies were considered adequate for pooling data and further testing for heterogeneity. Data from the included studies is compiled and presented as Table 4.

Summary statistics

Summary statistics were performed for all the included variables. Linear regression statistical modeling was performed between dependent and independent variables using STATA statistical software.

Total patients in five studies were 2239. The mean age of the patients was 31.9 years, with a skewness of –0.26.

The mean number of patients included in the analysis was 427.8 (minimum 118, maximum 1021). The analysis included an average of 183 males and 245 females. The male to female ratio was around 1 : 1.35.

Table 2

Summary of the risk of bias assessment of the 20 included studies using ROBINS-I tool

References	Confounding bias	Selection bias	Bias in classification of intervention	Bias due to derivations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of reported results	Overall risk of bias
Mariadason *et al.*[43]	Moderate	Moderate	Serious	Low	Low	NI	Low	Serious
Ramez *et al.*[44]	Moderate	Low	Moderate	Moderate	Moderate	NI	Low	Moderate
Coleman *et al.*[45]	Moderate	Low	Moderate	Moderate	Low	Low	Low	Moderate
Apisarnthanarak *et al.*[46]	Moderate	Low	Moderate	Moderate	Low	NI	Low	Moderate
Tan *et al.*[58]	Moderate	Low	Moderate	Moderate	Moderate	NI	Low	Moderate
Meltzer *et al.*[48]	NI	Moderate	Low	Low	Moderate	NI	Low	Moderate
Wang *et al.*[49]	Moderate	Low	Moderate	Serious	Low	NI	Low	Serious
Mishra *et al.*[50]	NI	Low	Moderate	Moderate	Moderate	Low	Low	Moderate
Waheed *et al.*[51]	Moderate	Low	Serious	Moderate	Serious	NI	Low	Serious
de *et al.*[52]	Moderate	Low	Low	Moderate	Low	NI	Low	Serious
Neupane *et al.*[53]	NI	Moderate	Low	Low	Low	NI	Low	Moderate
Tan *et al.*[54]	Moderate	Low	Moderate	Moderate	Low	NI	Low	Moderate
Mukhopadhyay *et al.*[55]	Moderate	Moderate	Low	Low	Low	NI	Low	Moderate
Lamparelli *et al.*[56]	Moderate	Moderate	Serious	Moderate	Low	NI	Low	Serious
Lopez *et al.*[57]	Moderate	Moderate	Low	Low	Moderate	Low	Low	Moderate
Tan *et al.*[47]	NI	Moderate	Serious	Low	Low	NI	Low	Serious
Seshadri *et al.*[59]	Moderate	Moderate	Moderate	Serious	Moderate	Low	Low	Serious

The seven studies with 'serious' risk of bias were excluded from our review.

The mean number of patients included from studies with AS 0–6 was 223 (minimum 70 and maximum 382) with SD 150.39, resulting in a wider confidence interval, skewness of 0.81 and kurtosis of 1.26.

The mean number of patients with AS 7–10 was 205 (minimum 48 and maximum 668) with SD 259.39, skewness of 1.46, and kurtosis of 3.20.

Other variables tested for summary statistics are shown in Table 5 along with their characteristics (Supplementary Fig. 1, Supplemental Digital Content 3, http://links.lww.com/MS9/A20).

Table 3

Included and excluded studies with reason for exclusion

References	Excluded studies	Reason for exclusion
Neupane *et al.*[53]	Seshadri *et al.*[59]	High-risk of bias
Tan *et al.*[54]	Lamparelli *et al.*[56]	High-risk of bias
Mukhopadhyay *et al.*[55]	de *et al.*[52]	High-risk of bias
Coleman *et al.*[45]	Wang *et al.*[49]	High-risk of bias
Apisarnthanarak *et al.*[46]	Mariadason *et al.*[43]	High-risk of bias
	Waheed *et al.*[51]	High-risk of bias
	Tan *et al.*[47]	High-risk of bias
	Mishra *et al.*[50]	Incomplete data on outcome measures of interest
	Lopez *et al.*[57]	Incomplete data on outcome measures of interest
	Tan *et al.*[58]	Incomplete data on outcome measures of interest
	Meltzer *et al.*[48]	Incomplete data on outcome measures of interest
	Ramez *et al.*[44]	Incomplete data on outcome measures of interest

Linear regression analysis

A linear regression equation was constructed between the dependent variable 'histological appendicitis' and the independent variable 'patient receiving intervention'. The regression coefficient of the equation was 0.923 with a 95% CI (0.899–0.947) and a '*P*' value of less than 0.001 with an adjusted 'R^2' of 0.999.

Linear regression between dependent variable 'histological appendicitis' and independent variable 'AS 7–10 and patients receiving intervention' shows a regression coefficient of 0.945 with a 95% CI: 0.647–1.244 and a '*P*' value less than 0.005 with an adjusted 'R^2' of 0.999.

A similar linear regression equation between 'histological appendicitis' as dependent variable and 'AS 0–6 and patients receiving intervention' as independent variable shows a regression coefficient of 0.915 with a 95% CI: 0.898–0.932 and a '*P*' value of less than 0.001 with an adjusted 'R^2' of 0.999.

The linear regression equation between dependent variable 'patients with no intervention' and independent variable 'AS 0–6' demonstrated a regression coefficient of 0.818 with 95% CI: 0.071–1.564 and a '*P*' value of less than 0.040 with an adjusted 'R^2' of 0.736; and between dependent variable 'patients with no intervention' and independent variable 'AS 7–10' has a regression coefficient of 0.113 with 95% CI: – 0.837–1.06 and '*P*' value of less than 0.729 with an adjusted 'R^2' of – 0.272). Therefore, the model has failed.

The consolidated results of linear regression analysis are demonstrated in Table 6. (Supplementary Fig. 2, Supplemental Digital Content 4, http://links.lww.com/MS9/A21; Supplementary Fig. 3, Supplemental Digital Content 5, http://links.lww.com/MS9/A22; Supplementary Fig. 4, Supplemental Digital Content 6, http://links.lww.com/MS9/A23).

Table 4

Final five studies included in analysis with their respective data.

	References	Total patients in study	Patients included in analysis	Males [*n* (%)]	Females [*n* (%)]	Patients with AS 0–6 [*n* (%)]	Patients with AS 7–10 [*n* (%)]	Patients with no intervention [*n* (%)]	Patients received intervention [*n* (%)]
Included studies	Neupane *et al.*[53]	1021	1021	498 (48.7)	523 (51.2)	354 (34.6)	667 (65.3)	203 (19.8)	818 (80.1)
	Tan *et al.*[54]	450	350	134 (38.2)	216 (61.7)	240 (68.5)	110 (31.4)	182 (52)	168 (48)
	Mukhopadhyay *et al.*[55]	118	118	58 (49.1)	60 (50.8)	70 (59.3)	48 (40.6)	0	118 (100)
	Coleman *et al.*[45]	492	492	157 (31.9)	355 (72.1)	382 (77.6)	110 (22.3)	369 (75)	123 (25)
	Apisarnthanarak *et al.*[46]	158	158	66 (41.7)	92 (58.2)	67 (42.4)	91 (57.6)	94 (59.4)	64 (40.5)

Table 5

Summary statistics of different variables

Variable	Mean	SD	Variance	Skewness	Kurtosis
Patients included in analysis	427.8	364.29	132 707.2	0.920	2.455
Male	182.6	181.394	32 903.8	1.298	2.980
Female	245.2	189.643	35 964.7	0.499	1.872
Mean age	31.906	7.400	54.772	– 0.265	1.5
Patients with AS 0–6	222.6	150.39	222 618.8	– 0.0818	1.265
Patients with AS 7–10	205.2	259.39	67 284.7	1.463	3.204
Patients had USS	345	498.867	248 869	0.662	1.5
Patients had CT scan	203.2	139.825	19 551.2	0.0342	1.458
Threshold AS for CT scan	4.75	3.304	10.916	– 0.9003	2.089
Patients with no intervention	169.6	137.394	18 877.3	0.286	2.141
Patients with intervention	258.2	315.103	99 290.2	1.448	3.182
Histologically normal appendix	22.6	24.419	596.3	1.450	3.189
Histological appendicitis	235.6	291.060	84 716.3	1.438	–

AS, Alvarado score; CT, computed tomography; USS, ultrasonography scan.

Subgroup analysis

We performed subgroup analysis with the dependent variable as histologically proven appendicitis and independent variables as males and females separately. The other independent variable was the number of patients who received intervention. Two separate equations are shown in Table 7. They demonstrated that although there are no true gender differences, the female gender has shown a significant number of interventions with high 'z' scores, almost three times as compared to males.

In a multivariate analysis, the Poisson regression equations for both genders have shown that gender is a significant variable when we add high AS 7–10 into the equation, as demonstrated in Table 8.

Meta-regression analysis

A meta set was built with sample size (n) as patients with histologically proven appendicitis, and the mean (μ) in the model was AS 7–10. The SD in this meta set was patients with normal appendix on histology. With this meta set, a meta-regression analysis was performed to see if there were any differences between patients who received intervention and those who did not.

This meta-regression equation showed the coefficient for patients who received intervention was positive (0.029) whereas the coefficient for those who did not receive intervention was negative (– 0.111) and P values for both were significant independently as shown in Table 4. As we noticed, 'I^2' was 99.81, which means there is high heterogeneity within the included studies. Also, there is high variance ('H^2'-519.86), therefore a forest plot is not valid for this data to conclude pooled estimates. Overall, the meta-regression model had 'R^2' value of 76.99 with 'P' less than 0.0005, which means patients with high AS have received interventions significantly higher with histologically proven appendicitis as compared to those without histological evidence of appendicitis. Therefore, high AS will help in reducing the NAR. But results should be cautiously interpreted as the regression coefficient is weakly positive (0.029). The meta-regression analysis is shown in Table 9.

Table 6

Linear regression analysis

Equation	Dependent variable	Independent variables	Regression coefficient	95% CI	*P* value	Adjusted R^2
1	Histological appendicitis	Patient received intervention	0.923	0.899–0.947	< 0.001	0.999
2	Histological normal appendix	Patient received intervention	0.763	0.052–0.100	< 0.002	0.962
3	Histological appendicitis	AS 7–10	0.495	0.647–1.244	< 0.005	0.999
4	Histological appendicitis	AS 0–6	0.915	0.898–0.932	< 0.001	0.999
5	Patients with no intervention	AS 0–6	0.818	0.071–1.564	< 0.040	0.736
6	Patients with no intervention	AS 7–10	0.113	– 0.837–1.06	< 0.729	– 0.272
7	Histological appendicitis	CT scan	0.425	– 4.169–3.318	< 0.742	– 0.277
8	Histological appendicitis	US scan	0.709	– 1.018–2.436	< 0.121	0.929

AS, Alvarado score; CT, computed tomography; US, ultrasound.

Discussion

The AS is an easy and simple predictive tool to carry out the risk prediction of AA. It can be carried out at any time of the day as it is easy, cheap, and reproducible[27,60]. A recent article on the Jerusalem Guidelines update in 2020 strongly recommended the use of clinical scores such as AS, appendicitis inflammatory response score, and the new adult appendicitis score to exclude AA and identify intermediate-risk patients who require imaging diagnostics[60,61].This approach of avoiding investigation in all patients with suspected appendicitis is not only beneficial in low-resource countries where diagnostic tests are scarce[62], but it is

Table 7

Subgroup analysis

1. **Poisson Regression:**

Males with histological appendicitis and male patients received intervention

No. of obs = 5

LR chi^2 (2) = 797.35

Prob > chi2 = 0.0000

Psedo R2= 0.9427

Males with intervention	Coef.	SE	*Z*	*P*>[*z*]	95% CI
Male	0.000923	0.0020324	0.45	0.650	−0.0030603–0.0049063
Patients receiving intervention	0.0026793	0.0011475	2.33	0.020	0.0004302–0.0049283
_cons	3.517691	0.1139974	30.86	0.000	3.29426–3.741122

2. **Poisson regression:**

Females with histological appendicitis and female patients received intervention

No. of obs = 5

LR chi^2 (2) = 813.89

Prob > chi2 = 0.0000

Psedo R2= 0.9535

Females with intervention	Coef.	SE	*Z*	*P*>[*z*]	95% CI
Female	0.0025624	0.0003321	7.72	0.000	0.0019114–0.0032133
Patients receiving intervention	0.0012297	0.0006536	1.88	0.060	−0.0000513 to 0.0025108
_cons	3.459559	0.1189916	29.07	0.000	3.22634–3.692778

Table 8

Multivariate analysis

1. Poisson Regression:

Males with histological appendicitis and male patients received intervention and AS 7-10

No. of obs = 5

LR chi^2 (3) = 815.06

Prob > chi2 = 0.0000

Psedo R2= 0.9636

Males with histological intervention	Coef.	Std Err.	Z	P>[z]	95% conf. interval
Male patients received intervention	.0094158	.0020588	4.57	0.000	.0053805 .013451
male	.0075879	.0026689	2.84	0.004	.002357 .0128187
AS 7-10	-.0125958	.003306	-3.81	0.000	-.0190755 -.0061161
_Cons	3.083934	.1726194	17.87	0.000	2.745606 3.422262

2. Poisson Regression:

Females with histological appendicitis and female patients received intervention and AS 7-10

No. of obs = 5

LR chi^2 (3) = 820.33

Prob > chi2 = 0.0000

Psedo R2= 0.9611

Females with histological intervention	Coef.	Std Err.	Z	P>[z]	95% conf. interval
Female patients received intervention	.0074297	.0020186	3.68	0.000	.0034733 .0113861
male	.002043	.0007479	2.73	0.006	.0005771 .0035088
AS 7-10	-.0063439	.0026079	-2.43	0.015	-.0114553 -.0012324
_Cons	3.28095	.1469081	22.33	0.000	2.993015 3.568884

AS, Alvarado score.

also beneficial in developed countries by allowing for more efficient resource utilization.

The studies included in our systematic review have highlighted the effectiveness of the use of AS in various ways. The study published by Tan *et al.*[54] in 2014 is a prospective study with 450 patients, and they evaluated the subset of patients that could benefit from further diagnostic evaluation. This study concluded that categories for AS are different in males and females, and male patients with AS 7 or above and female patients with AS 9 or above are least likely to benefit from a CT scan and can be counseled for surgery without further imaging.

Coleman *et al.*[45] calculated the AS retrospectively in 492 patients who underwent CT scan for suspected appendicitis and concluded that patients with AS 9–10 should go straight to surgery without imaging and those with AS 2 or less can be safely discharged home. This study, however, did not classify male and female patients separately. A large cross-sectional observational study by Neupane *et al.* in 2019[53] with 1021 patients suggested that AS is significantly higher in patients with AA as compared to patients with other diseases. They identified the large variation in presentation, severity of disease, radiological work-up, and surgical management of AA.

Another retrospective cohort study by Apisarnthanarak *et al.*[46] with 158 patients concludes that patients with AS 7 or more are likely to have complicated appendicitis. However, this study also says that AS cannot be used as an independent tool and should be used in conjunction with a CT scan, as demonstrated by previous studies too.

Table 9

Meta-regression

Meta-regression model:

Effect-size label: Hedge's *g*
Sample size (*n*): patients with histologically proven appendicitis
Mean (μ) : AS 7–10
SD: patients with histologically normal appendix

Model statistics:

Random-effects meta-regression
Method: REML

Number of studies	5
τ^2	70.75
I^2 (%)	99.81
H^2	519.86
R^2 (%)	76.99
Wald χ^2 (2)	15.05
Prob $>\chi^2$	0.0005

Results of the meta-regression model:

_meta_es	Co-effiecient	SD	*Z*	*P* > IzI	95% CI
Patient receiving intervention	0.298764	0.135763	2.20	0.028	0.3332674–0.564855
Patients with no intervention	−0.1112007	0.314728	−3.53	0.000	−0.1728864 to −0.0495151
Constant	2.437063	6.92916	0.35	0.725	−11.14384 to 16.01797

Mukhopadhyay *et al.*[55] in their prospective observational study with 118 patients found that combined use of imaging with AS can prevent surgical complications and minimize NARs.

Although all these studies[45,46,54–56] support the use of AS in combination with imaging, there are limitations to these studies in view of the heterogeneity. The classification of AS was different in every study, which made statistical modeling more difficult. Some of the studies[51–46,53] have calculated AS retrospectively, which underestimates the true performance of AS (raises the possibility of performance bias). Variables collected in each study were different from one another, and there was a lot of missing data when we combined the data, leading to statistical heterogeneity. Hence, the meta-analysis could not be performed and therefore, meta-regression was performed instead.

The results of our study show a significant co-relation between patients with AS 7–10 receiving surgical intervention and histological appendicitis, which was also verified in the meta-regression analysis performed during this study. Our meta-regression revealed that patients with high AS (7–10) have a significant cause-and-effect relationship with histologically proven AA ('*Z*' score: 2.20 and '*P*' = 0.028), whereas patients with histologically normal appendix showed a negative '*Z*' score and a significant negative cause-and-effect relationship between AS and normal appendix ('*Z*' score: 3.53 and '*P*' = 0.001), which means that patients with high AS will have AA. Our results also confirm that patients with high AS who underwent surgery had histologically proven appendicitis, which suggests that high AS can be used as a tool to avoid negative appendicectomy.

Limitations

There are some limitations to our study. The main limitations were the limited number of studies available for analysis and, moreover, significant heterogeneity was noted across the included studies. This has made analysis of the existing data spurious, leading to inconclusive results and false discoveries (type 1 error). The possible sources of heterogeneity include the case mix of men and women and observer variation in the measurement of signs and symptoms.

When we conducted subgroup analysis to look for any gender differences, we found a threefold increase in the number of interventions in the female group, although the overall equation has shown that there are no gender differences with the addition of the higher AS group into the equation. This could be due to the small number of patients with high AS.

The other limitation was that most of the studies included in our systematic review were single-center, retrospective studies. Two of the five studies calculated AS retrospectively from the electronic medical records, leading to missing information, misclassification, and the true performance of AS was likely to be underestimated. Finally, although we used an up-to-date search strategy for our systematic review, we understand that there may be some articles that have been missed, as it happens in most systematic reviews.

Future recommendations

For AS, it is recommended that a more rigorous method of documenting signs and symptoms be used prospectively as opposed to retrospectively in order to reduce the errors of inter-rater agreements.

A head-to-head comparative study between the AS and the intraoperative findings is also recommended to assess the feasibility of the development of an algorithm (protocol).

The above two recommendations are mandatory before rolling a flow chart of the algorithm into the public domain using AS.

Conclusion

We conclude that having a high AS is a significant predictor of AA. High ASs have a positive association with histologically proven appendicitis. We found statistical heterogeneity in the included studies, which was our major limitation.

Ethical approval

Not applicable.

Consent

Not applicable.

Source of funding

No funding (external or internal) was used for this manuscript.

Author contribution

S.G.: study design and conception, data collection, writing. M.U.: data analysis and writing. V.S.K.: data collection and analysis. K.D.C.: data collection. A.R.: study design, writing.

Conflict of interest disclosure

No conflict of interest declared by any author.

Research registration unique identifying number (UIN)

1. Name of the registry: NA.
2. Unique Identifying number or registration ID: NA.
3. Hyperlink to your specific registration (must be publicly accessible and will be checked): NA.

Guarantor

Professor Ashraf Rasheed, Consultant Upper GI and General Surgeon, Royal Gwent Hospital, Newport, Wales NP20 2 UB, UK

Provenance and peer review

Not commissioned, externally peer-reviewed.

References

[1] Lukáš K. Příběh apendixu [the story of Appendix] (Czech). Cas Lek Cesk 2015;154:189–93.
[2] Meljnikov I, Radojcić B, Grebeldinger S, *et al.* History of surgical treatment of appendicitis (Serbian). Med Pregl 2009;62:489–92.
[3] Williams GR. Presidential Address: a history of appendicitis. With anecdotes illustrating its importance. Ann Surg 1983;197:495.
[4] Al Awayshih MM, Nofal MN, Yousef AJ. Evaluation of Alvarado score in diagnosing acute appendicitis. Pan Afr Med J 2019;34.
[5] McBURNEY CH II . The indications for early laparotomy in appendicitis. Ann Surg 1891;13:233.
[6] Grover CA, Sternbach G. Charles McBurney: McBurney's point. J Emerg Med 2012;42:578–81.
[7] Akbulut S, Bahçe ZS, Öztaş T, *et al.* Assessment of demographic, clinical and histopathological features of patients who underwent appendectomy due to a presumed diagnosis of acute appendicitis. Turk J Trauma Emerg Surg 2021;27:315–24.
[8] Rypins EB, Evans DG, Hinrichs W, *et al.* Tc-99m-HMPAO white blood cell scan for diagnosis of acute appendicitis in patients with equivocal clinical presentation. Ann Surg 1997;226:58.
[9] Horton MD, Counter SF, Florence MG, *et al.* A prospective trial of computed tomography and ultrasonography for diagnosing appendicitis in the atypical patient. Am J Surg 2000;179:379–81.
[10] Franke C, Böhner H, Yang Q, *et al.* Ultrasonography for diagnosis of acute appendicitis: results of a prospective multicenter trial. World J Surg 1999;23:141.
[11] Weyant MJ, Eachempati SR, Maluccio MA, *et al.* Interpretation of computed tomography does not correlate with laboratory or pathologic findings in surgically confirmed acute appendicitis. Surgery 2000;128: 145–52.
[12] Cardall T, Glasser J, Guss DA. Clinical value of the total white blood cell count and temperature in the evaluation of patients with suspected appendicitis. Acad Emerg Med 2004;11:1021–7.
[13] Andersson RE, Hugander AP, Ghazi SH, *et al.* Diagnostic value of disease history, clinical presentation, and inflammatory parameters of appendicitis. World J Surg 1999;23:133.
[14] Yu CW, Juan LI, Wu MH, *et al.* Systematic review and meta-analysis of the diagnostic accuracy of procalcitonin, C-reactive protein and white blood cell count for suspected acute appendicitis. J Br Surg 2013;100: 322–9.
[15] Al-Gaithy ZK. Clinical value of total white blood cells and neutrophil counts in patients with suspected appendicitis: retrospective study. World J Emerg Surg 2012;7:1–7.
[16] Wu HP, Lin CY, Chang CF, *et al.* Predictive value of C-reactive protein at different cutoff levels in acute appendicitis. Am J Emerg Med 2005;23: 449–53.

[17] Hernanz-Schulman M. CT and US in the diagnosis of appendicitis: an argument for CT. Radiology 2010;255:3–7.
[18] Kim K, Kim YH, Kim SY, *et al*. Low-dose abdominal CT for evaluating suspected appendicitis. N Engl J Med 2012;366:1596–605.
[19] Birnbaum BA, Wilson SR. Appendicitis at the millennium. Radiology 2000;215:337–48.
[20] Jaschinski T, Mosch C, Eikermann M, *et al*. Laparoscopic versus open appendectomy in patients with suspected appendicitis: a systematic review of meta-analyses of randomised controlled trials. BMC Gastroenterol 2015;15:48.
[21] Addiss DG, Shaffer N, Fowler BS, *et al*. The epidemiology of appendicitis and appendectomy in the United States. Am J Epidemiol 1990;132: 910–25.
[22] Akbulut S, Koç C, Şahin TT, *et al*. An investigation into the factors predicting acute appendicitis and perforated appendicitis. Turk J Trauma Emerg Surg 2021;27:434–2.
[23] Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med 1986;15:557–64.
[24] Karami MY, Niakan H, Zadebagheri N, *et al*. Which one is better? Comparison of the acute inflammatory response, Raja Isteri Pengiran Anak Saleha Appendicitis and Alvarado Scoring Systems. Ann Coloproctol 2017;33:227–31.
[25] Gujar NN, Mudhol Sajid, Choudhari Ravi, *et al*. Determination of sensitivity and specificity of modified Alvarado score and ultrasonography in patients with acute appendicitis. J Krishna Inst Med Sci Univ 2015;4: 89–99.
[26] Brigand C, Steinmetz JP, Rohr S. De l'intérêt des scores en matière de diagnostic d'appendicite [The usefulness of scores in the diagnosis of appendicitis] (French). J Chir (Paris) 2009;146(Spec No 1):2–7.
[27] Ohle R, O'Reilly F, O'Brien KK, *et al*. The Alvarado score for predicting acute appendicitis: a systematic review. BMC Med 2011;9:139.
[28] Di Saverio S, Birindelli A, Kelly MD, *et al*. WSES Jerusalem guidelines for diagnosis and treatment of acute appendicitis. World J Emerg Surg 2016;11:1–25.
[29] Gorter RR, Eker HH, Gorter-Stam MA, *et al*. Diagnosis and management of acute appendicitis. EAES consensus development conference 2015. Surg Endosc 2016 ;30:4668–90.
[30] Gregory S, Kuntz K, Sainfort F, *et al*. Cost-effectiveness of integrating a clinical decision rule and staged imaging protocol for diagnosis of appendicitis. Value Health 2016;19:28–35.
[31] Howell JM, Eddy OL, Lukens TW, *et al*. American College of Emergency Physicians. Clinical policy: critical issues in the evaluation and management of emergency department patients with suspected appendicitis. Ann Emerg Med 2010;55:71–116.
[32] Flum DR, Morris A, Koepsell T, *et al*. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. JAMA 2001;286: 1748–53.
[33] Perez J, Barone JE, Wilbanks TO, *et al*. Liberal use of computed tomography scanning does not improve diagnostic accuracy in appendicitis. Am J Surg 2003;185:194–7.
[34] Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: nationwide analysis. Arch Surg 2002;137: 799–804; discussion 804.
[35] Blisard D, Rosenfeld JC, Estrada F, *et al*. Institutioning a clinical practice guideline to decrease the rate of normal appendectomies. Am Surg 2003;69:796–8.
[36] D'Souza N, Marsden M, Bottomley S, *et al*. Cost-effectiveness of routine imaging of suspected appendicitis. Ann R Coll Surg Engl 2018;100: 47–51.
[37] Lahaye MJ, Lambregts DM, Mutsaers E, *et al*. Mandatory imaging cuts costs and reduces the rate of unnecessary surgeries in the diagnostic work-up of patients suspected of having appendicitis. Eur Radiol 2015;25: 1464–70.
[38] Kundiona I, Chihaka OB, Muguti GI. Negative appendicectomy: evaluation of ultrasonography and Alvarado score. Cent Afr J Med 2015;61: 66–73.
[39] Nautiyal H, Ahmad S, Keshwani NK, *et al*. Combined use of modified Alvarado score and USG in decreasing negative appendicectomy rate. Indian J Surg 2010;72:42–8.
[40] https://www.researchregistry.com/browse-the-registry#registryofsystematicreviewsmeta-analyses/
[41] Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021;10:1.
[42] Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008.
[43] Mariadason JG, Wang WN, Wallack MK, *et al*. Negative appendicectomy rate as a quality metric in the management of appendicitis: impact of computed tomography, Alvarado score and the definition of negative appendicectomy. Ann Royal Coll Surg Engl 2012;94:395–401.
[44] Ramez MW, Akhnokh SG, Kamel S. Evaluation of Alvarado score in the diagnosis of acute appendicitis. Med J Cairo Univ 2020;88(September): 1663–72.
[45] Coleman Jamie J, Carr Bryan W, Rogers Tyrone, *et al*. The Alvarado score should be used to reduce emergency department length of stay and radiation exposure in select patients with abdominal pain. J Trauma Acute Care Surg 2018;84:946–50.
[46] Apisarnthanarak P, Suvannarerg V, Pattaranutaporn P, *et al*. Alvarado score: can it reduce unnecessary CT scans for evaluation of acute appendicitis? Am J Emerg Med 2015;33:266–70.
[47] Tan WJ, Pek W, Kabir T, *et al*. A lvarado score: a guide to computed tomography utilization in appendicitis. ANZ J Surg 2013;83:748–52.
[48] Meltzer AC, Baumann BM, Chen EH, *et al*. Poor sensitivity of a modified Alvarado score in adults with suspected appendicitis. Ann Emerg Med 2013;62:126–31.
[49] Wang SY, Fang JF, Liao CH, *et al*. Prospective study of computed tomography in patients with suspected acute appendicitis and low Alvarado score. Am J Emerg Med 2012;30:1597–601.
[50] Mishra BM, Panigrahi PB, Mishra S, *et al*. Alvarado score and computerized tomography scan as impact indicator in intervening negative appendectomy rate. Int Surg J 2020;7:1223–8.
[51] Waheed KB, Shah WJ, Alshehri AS, *et al*. Usefulness of CT scan in diagnosing acute appendicitis in patients with low Alvarado scores. Imaging 2019;6:8.
[52] de J, Wagner P, Haroon M, *et al*. Does CT reduce the rate of negative laparoscopies for acute appendicitis? A single-center retrospective study. J Med Life 2020;13:26.
[53] Neupane BR, Paudel SR, Shrestha A, *et al*. Management of acute abdomen: acute appendicitis. J Gandaki Med Coll Nepal 2019;12:3–9.
[54] Tan WJ, Acharyya S, Goh YC, *et al*. Prospective comparison of the Alvarado score and CT scan in the evaluation of suspected appendicitis: a proposed algorithm to guide CT use. J Am Coll Surg 2015;220:218–4.
[55] Mukhopadhyay A, Majhi RC, Bhattachariya U, *et al*. A study to evaluate the combined use of Alvarado score and imaging for diagnosis of acute appendicitis. J Evol Med Dent Sci 2020;9:851–6.
[56] Lamparelli MJ, Hoque HM, Pogson CJ, *et al*. A prospective evaluation of the combined use of the modified Alvarado score with selective laparoscopy in adult females in the management of suspected appendicitis. Ann R Coll Surg Engl 2000;82:192.
[57] Lopez PP, Cohn SM, Popkin CA, *et al*. Appendicitis Diagnostic Group. The use of a computed tomography scan to rule out appendicitis in women of childbearing age is as accurate as clinical examination: a prospective randomized trial. Am Surg 2007;73:1232–6.
[58] Tan WJ, Acharyya S, Chew MH, *et al*. Randomized control trial comparing an Alvarado Score-based management algorithm and current best practice in the evaluation of suspected appendicitis. World J Emerg Surg 2020;15:1–8.
[59] Seshadri LN, Mohan LN. Alvarado score in acute appendicitis: revisited. Int Surg J 2019;6:1637–40.
[60] Pouget-Baudry Y, Mucci S, Eyssartier E, *et al*. The use of the Alvarado score in the management of right lower quadrant abdominal pain in the adult. J Visc Surg 2010;147:e40–4.
[61] Di Saverio S, Podda M, De Simone B, *et al*. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. World J Emerg Surg 2020;15:1–42.
[62] Tade AO. Evaluation of Alvarado score as an admission criterion in patients with suspected diagnosis of acute appendicitis. West Afr J Med 2007;26:210–2.