

Original Research Article

Diagnostic accuracy of acute appendicitis by comparing clinical judgment of surgeons and Alvarado scoring in suspected acute appendicitis

Ishory Bhusal^{1*}, Chandra Shekhar Agarwal², Rakesh Kumar Gupta², Suresh Prasad Sah²

¹Department of Surgery, Rapti Academy of Health Science, Ghorahi, Dang, Nepal

²Department of Surgery, BPKIHS, Dharan, Nepal

Received: 26 June 2021

Revised: 04 August 2021

Accepted: 09 August 2021

*Correspondence:

Dr. Ishory Bhusal,

E-mail: bhusalishory2008@gmail.com

Copyright: © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: The Alvarado score is a clinical scoring system used in the diagnosis of acute appendicitis. The aim of the study was to compare the clinical judgment of surgeons and Alvarado score in diagnosing acute appendicitis within Nepalese people and to refine the score and suggest a new score to make a more accurate diagnosis of acute appendicitis.

Methods: In this prospective, parallel-group, quasi-randomized study of patients presenting at a tertiary hospital in eastern Nepal with suspected appendicitis during 1 year were assigned in weekly alternation to either group A or group B. The group A patients were treated on the basis of their Alvarado score, and the group B patients based on the clinical judgment. The correctness of the methods was assessed by the final histology.

Results: In this study, the mean age of patients in Alvarado group was 26.45 years and in clinical judgment was 28.68 years. The sensitivity, the specificity, the diagnostic accuracy, the positive predictive value and negative appendectomies in Alvarado group were 95.5%, 68.9%, 90.91%, 93.4% and 6.56% respectively whereas in clinical judgment group were 98.51%, 85.71%, 496.4%, 97.04% and 3.59% respectively.

Conclusions: This study showed clinical judgment to be more reliable in the diagnosis of acute appendicitis than the Alvarado score, but the score is a useful diagnostic aid, especially for young colleagues. The use of the new scoring system has become easier. It includes fewer criteria as well as an important and sensitive predictor: the ultrasound investigation.

Keywords: Appendix, Abdominal, Ultrasonography, Alvarado score, Clinical judgment

INTRODUCTION

Appendicitis is caused by obstruction of the appendiceal lumen leading to a progressive cycle of pathologic changes.¹ Fecoliths, parasites, tumours, foreign bodies and viral and bacterial agents have all been associated with the development of appendiceal obstruction. Acute appendicitis is a sudden and severe inflammation of the appendix manifested by stomach pain, fever, loss of appetite, and if left untreated can have fatal complications. It is one of the most common abdominal emergencies that call for surgical treatment.²

On a global scale, acute appendicitis has been a common surgical emergency comprising with a lifetime risk of 1 in 7, which means that around 6% of the individuals suffer an attack during their lifetime.³ In India, the occurrence of acute appendicitis was the highest in the 11-20 years age group which constituted 44.6%.⁴ Diagnosis of acute appendicitis is based on medical history, physical examination, laboratory analysis and imaging techniques (ultrasound and CT scan).⁵ Its differential diagnosis is extremely difficult, especially for the elderly, children, and fertile-age women, in whom it can mimic numerous gynecologic and urogenital conditions. The gold standard

for diagnosis of appendicitis is the histopathology of the surgically removed appendix.⁶ However, surgical removal of a normal appendix is a burden both on patients and health resources.

Therefore, early and accurate recognition of the condition and prompt operation have been the most important factors in reducing morbidity and possible mortality, length of stay, and overall cost of treatment.⁷ Improved diagnostic accuracy not only helps in taking early management decisions but also curtails negative appendectomy rates. Numerous clinical scoring systems have been created to make the diagnosis of acute appendicitis easier.^{8,9} The Alvarado score is a clinical scoring system used in the diagnosis of acute appendicitis. This score was created in 1986 by Alvarado, who processed the data of appendectomy patients retrospectively. This scoring system includes eight diagnostic criteria (historical data, physical examination, and laboratory values).¹⁰

The reported sensitivity and specificity for the Alvarado and the modified Alvarado scores range from 53%-88% and 75%-80%, respectively.¹¹ However, these scoring systems were developed in western countries, and several studies have reported very low sensitivity and specificity when these scores were applied to a population with a completely different ethnic origin and diet.¹² The main drawback of the Alvarado score is its application in the patients with atypical appendix position. That may conclude that even in the presence of lower values of the Alvarado score 5 and 6, a careful examination should be carried out. Clinical judgment of the surgeon towards screening acute appendicitis from suspected cases is the mainstay in diagnosing acute appendicitis. Studies have been undertaken in several countries to compare the clinical judgment of surgeons and Alvarado scoring in suspected acute appendicitis.^{13,14}

There has been no comparative study done to assess the diagnostic accuracy of acute appendicitis between clinical judgment of surgeons and Alvarado scoring in suspected acute appendicitis in Nepal as far as best of our literature search. Therefore, objective of this study was to compare the clinical judgment of surgeons and Alvarado score in diagnosing acute appendicitis within Nepalese hospital setting. Further, this study also aimed to refine the score and suggest a new score to make a more accurate diagnosis of acute appendicitis.

METHODS

Between 16th February 2016 to 15th February 2017, this hospital-based prospective, open, parallel-group, quasi-randomized study enrolled 278 patients presenting with right lower abdominal pain with suspicion of acute appendicitis at the emergency and surgery outpatient department of B. P. Koirala Institute of Health Sciences, Nepal. The study protocol was performed in accordance with the principle of the declaration of Helsinki and after approval by the Institutional ethical review board. After

signing a consent form, the patients were divided into two groups. Quasi randomization was done by assigning patient either of group A or B based on weekly alternation. In group A, the treatment decision was based on the Alvarado score as follows: 1-4 points (discharge), 5-6 points (observation, with scoring repeated in 12 h), 7-10 points (urgent surgery). Further treatment of the patients in group B was based on the decision made by the head surgeon on duty, who did not know the Alvarado score of the patient.

Only those patients within the age group of 10 to 60 years and of both sexes with right lower abdominal pain with suspicion of acute appendicitis were included in this study. Whereas, those patients who were managed exclusively by conservative management and did not undergo appendectomy after 24 hours, aged <10 years and >60 years, suffering from recurrent appendicitis, appendicular perforation with peritonitis, appendicular lump, diagnosed with HIV/immuno-compromised/transplant recipient/chronic oral steroid use, having acute appendicitis with pregnancy and who have not provided consent to participate in this study were excluded.

Sample size calculation

A sample size was estimated on assumption of sensitivity of Alvarado score as 81.91% and clinical judgement of surgeons having 93.0% sensitivity.^{28,29} A sample size of 139 in each group A and B (total 278) of suspected appendicitis case was needed to reject null hypothesis at 80% power (2 sided) to detect at least 11.09% difference between groups using an alpha error of 5%. It was calculated by using following formula:

$$N = \frac{[Z_{1-\frac{\alpha}{2}}\sqrt{2\pi_0(1-\pi_0)} + Z_{1-\beta}\sqrt{\pi_1(1-\pi_1) + \pi_2(1-\pi_2)}]^2}{(\pi_2 - \pi_1)^2}$$

$$\pi_0 = \frac{(\pi_1 + \pi_2)}{2}$$

Here, sensitivity of the new test=81.91%, sensitivity of reference test=93%, alpha=5%, 1-beta=80%. Putting this value in above formula we got N=139 in each group, which was required sample size.

Statistical analysis

The data was entered; tabulated and statistical analysis was performed by using Statistical Package for the Social Sciences (SPSS 20). Descriptive statistics included mean, standard deviation for quantitative variable, number and percentage for categorical data. The inferential statistics were calculated using Pearson's Chi square test. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated. The logistic regression analysis for each parameter was done. The receiver operating curve (ROC) at the optimal cut-off

threshold score for the new scoring system was derived. A value of $p < 0.05$ was considered significant.

RESULTS

During a period of 1 year, 278 patients were enrolled in our study (139 in group A and 139 in group B). Their mean age was 27.56 years (range, 11-60 years; group A, 26.45 years; group B, 28.68 years).

Only 21 patients fell within 51-60 years. Most of the patients were 11-20 years of age. In terms of gender distribution, group A had 46 women and 93 men, whereas group B had 48 women and 91 men ($p=0.78$). Based on these data, the demographics of the two groups was considered to be homogeneous.

Age distribution

In our study, patient age ranged from 10 years to 60 years in both the groups (Table 1).

Sex distribution

In our study, among 278 patients, 184 (66.42%) were male and 94 (33.5%) were female, (M: F: 1.95: 1). In Alvarado group among 139 patients, 93 (66.90%) were male and 46 (33.1%) were female (Table 2).

Comparison between both the groups

Comparing both groups there was no statistically significant difference in terms of age and gender (Table 3).

Geographic distribution

Most of the participants in our study were from Sunsari district (Table 4) with total patients of 162 (58%), followed by 32 (11.5%), 27 (10%), 25 (9%) and 32 (11.5%) from Dhankutta, Saptari, Morang and other districts, respectively.

Symptoms

Migratory pain

In Alvarado group: sensitivity: 93.5% (87.2%-96.95%), specificity: 6.2% (0.3%-32.2%), Positive predictive value (PPV)- 88.46%, Negative predictive value (NPV): 11.11%. In clinical judgment group: sensitivity: 97.5% (90.5%-99.67%), specificity: 5.17% (1.3%-15.2%), PPV-58.9%, NPV: 60%. In both groups: sensitivity: 95.1% (90.9%-97.4%), specificity: 8.4% (1.7%-13.9%), PPV-73.48%, NPV: 28.57%. In our study, migratory pain was present in 204 patients (73.38%) and absent in 74 patients (36.7%). The sensitivity and specificity was 95.1% and 8.4% respectively. In Alvarado group (Table 5), migratory pain was present in 123 patients (88.4%) and absent in 16 patient (11.6%). In clinical judgment group (Table 5), migratory pain was present in 81 patients (58.27%) and

absent in 58 patients (41.73%). Sensitivity was slightly better in Alvarado group (97.5% vs 93.5%) but specificity was better in clinical judgment group (6.2% vs 5.17 %).

Nausea or vomiting

In Alvarado group: sensitivity: 92.6% (86.06%-96.35%), specificity: 0% (0%-27%), PPV: 86.94%, NPV: 0%. In clinical judgment group: sensitivity: 97.91% (91.25%-99.63%), specificity: 6.97% (1.8%-20.12%), PPV: 70.14%, NPV: 60%. In both group: sensitivity: 94.95% (90.9%-97.32%), specificity: 5% (1.3%-14.8%) 50, PPV:78.4%, NPV: 21.42%. In our study, nausea or vomiting was present in 218 patients (78.4%) and absent in 50 patients (21.6%) (Table 6). The sensitivity and specificity were 94.05% and 5% respectively. In Alvarado group (Table 6), nausea or vomiting was present in 122 patients (88.4%) and absent in 17 patient (12.3%). In clinical judgment group (Table 6), nausea or vomiting was present in 96 patients (69%) and absent in 43 patients (31%). Sensitivity (97.1% vs 92.6%) and specificity (6.97% vs 0%) of clinical judgment group was better than Alvarado group.

Anorexia

In Alvarado group: sensitivity: 91.30% (83.10%-95.89%), specificity: 2.1% (0.1%-12%), PPV: 64.61%, NPV: 11.11%. In clinical judgment group: sensitivity: 97.95% (87.7%-99.89%), specificity: 4.44% (1.4%-11.62%), PPV: 35.82%, NPV: 80%. In both group: sensitivity: 93.61% (87.87%-96.8%), specificity: 3.6% (1.3%-0.7%), PPV: 50%, NPV: 35.71%. Anorexia was present in 141 (50.7%) patient and absent in 137 (49.3%) patient (Table 7). The sensitivity and specificity were 93.6% and 3.6% respectively. In Alvarado group (Table 7), anorexia was present in 92 patients (66.2%) and absent in 47 patients (33.8%). In clinical judgment group (Table 7), anorexia was present in 49 patients (35.2%) and absent in 90 patients (64.8%). Sensitivity (97.15% vs 91.30%) and specificity (4.44% vs 2.1%) of clinical judgment group was better than Alvarado group.

Fever

In Alvarado group: sensitivity: 94.91% (84.94%-98.6%), specificity: 7.5% (3%-16.9%), PPV: 43.07%, NPV: 66.67%. In clinical judgment group: sensitivity: 97.73% (86.5%-99.89%), specificity: 4.2% (1.3%-11%), PPV: 32.08%, NPV: 80%. In both groups: sensitivity: 96.11% (89.88%-98.7%) 53, specificity: 5.7% (2.9%-10.55%), PPV: 37.5%, NPV: 71.4%.

In our study (Table 8), fever was present in 103 patients (32.2%) and absent in 174 patients (62.8%). The sensitivity and specificity were 96.11% and 5.7% respectively. In Alvarado group (Table 8), fever was present in 59 patients (42.5%) and absent in 80 patients (57.5%). In clinical judgment group (Table 8), fever was present in 44 patient (31.6%) and absent in 95 patients

(68.4%). Sensitivity (97.73% vs 94.91%) of clinical judgment group was better than Alvarado group and the specificity (7.5% vs 4.2%) of Alvarado group was better than clinical judgment group.

Table 1: Age distribution.

Age range (years)	Alvarado group (group A)	Clinical judgment group (group B)	Total
11-20	60	55	115
21-30	32	31	63
31-40	27	25	52
41-50	14	13	27
51-60	6	15	21
Total	139	139	278
Mean age	26.45	28.68	27.56

Table 2: Sex distribution.

Groups	Male (%)	Female (%)
Alvarado	93 (66.90)	46 (33.1)
Clinical judgment	91 (65.46)	48 (34.53)

Table 3: Patient characteristics.

Variables	Alvarado group	Clinical judgment group
Age (years)	26.45	28.68
Gender		
Female	46	48; p=0.78
Male	93	91

Table 4: Geographic distribution.

Districts	Alvarado group	Clinical judgment group	Total	Percentage (%)
Sunsari	70	92	162	58
Dhankutta	18	14	32	11.5
Saptari	15	12	27	10
Morang	14	11	25	9
Others	22	10	32	11.5

Table 5: Migratory pain in Alvarado group, clinical judgment group and both groups.

Migratory pain in Alvarado group			
HPE	Present	Absent	Total
Positive	115	15	130
Negative	8	1	9
Total	123	16	139
Migratory pain in clinical judgment group			
Positive	79	55	134
Negative	2	3	5
Total	81	58	139
Migratory pain in both groups			
Positive	194	70	265
Negative	10	4	14
Total	204	74	278

Table 6: Nausea or vomiting in Alvarado group, clinical judgment group and both groups.

Nausea or vomiting in Alvarado group			
HPE	Present	Absent	Total
Positive	113	17	130

Continued.

Nausea or vomiting in Alvarado group			
Negative	9	0	9
Total	122	17	139
Nausea or vomiting in clinical judgment group			
Positive	94	40	134
Negative	2	3	5
Total	96	43	139
Nausea or vomiting in both groups			
Positive	207	57	264
Negative	11	3	14
Total	218	60	278

Table 7: Anorexia in Alvarado group, clinical judgment group and both groups.

Anorexia in Alvarado group			
HPE	Present	Absent	Total
Positive	84	46	130
Negative	8	1	9
Total	92	47	139
Anorexia in clinical judgment group			
Positive	48	86	134
Negative	1	4	5
Total	49	90	139
Anorexia in both groups			
Positive	132	132	264
Negative	9	5	14
Total	141	137	278

Table 8: Fever in Alvarado group, clinical judgment group and both groups.

Fever in Alvarado group			
HPE	Present	Absent	Total
Positive	56	74	130
Negative	3	6	9
Total	59	80	139
Fever in clinical judgment group			
Positive	43	91	134
Negative	1	4	5
Total	44	95	139
Fever in both groups			
Positive	99	165	264
Negative	4	10	14
Total	103	175	278

Table 9: Coordinates of the curve.

Total score	Sensitivity	Specificity
6	1.0	1.0
7.5	0.75	0.22
8.5	0.41	0.0
9.5	0.08	0.0

When cut off value of Alvarado score was decreased to 6, sensitivity and specificity increased to maximum level.

Table 10: New score made by logistic regression, Receiver operating characteristic curve (ROC) analysis and clinical evaluation of all parameters.

Parameters	Score
Migratory pain	1

Continued.

Parameters	Score
Nausea/vomiting	1
Fever (>38.3°C)	2
Right iliac fossa tenderness	2
Guarding in right iliac fossa	1
Leukocytosis	2
Negative urinalysis	1
USG abdomen and pelvis positive for acute appendicitis	2
Total score	12

DISCUSSION

In the present study 278 appendicectomic patients, group A (139) and group B (139) were studied in two sections in an alternate week.

Age distribution

The age of our sample patients varied between 10 and 60 years (Table 1). In Alvarado group, the mean age of patients was 26.45±12.2 years and 28.68±13.86 years in clinical assessment.

The average age group was significantly smaller than the research carried out by Man et al in 2014.¹³ In his research, the average was 33, 33 years for the Alvaradic group and 35, 52 years for the therapeutic assessment group, because we were in the 60. Our age group analysis was equivalent to Korkut et al study in 2020. The mean age of his sample was 27, 5 years for a similar age group (17 to 68 years).¹⁵

Sex distribution

In our sample, the figures were 184 (66.42%) men and 94 (33.5%) female. The ratio of men to women was 2:1. In Community Alvarado 93 were men (66.90%) and women (33.1%). On a clinical basis, 91 (65.94%) were males and 48 (34.06%) were women. Our thesis in Korkut et al in 2020 compares our results in his analysis, 72 men and 28 women had a ratio of 2.6: 1 between men and women.¹⁵

Primary outcome

Alvarado group

Sensitivity, precision of the diagnosis, positive predictive value and negative predictive value in Alvarado community is 95.5%, 68.9%, 90.91% 93.4% and 76.95%. The negative appendectomy average in the Alvarado party amounted to 6.56 percent (N=9) as compared to 3.59 percent (N=5). With a sensitivity, precision and diagnostic precision of 81.91, 66.66 and 81 percent respectively, our Alvarado findings were marginally superior to Korkut et al in 2020.¹⁵ The results in our research were better than his analysis, since we have repeatedly taken Alvarado's score based on appendectomy based on Alvarado scoring but only on the clinical judgment during entry. Our analysis showed that the negative appendectomy rate was 6 percent.¹⁵ In comparison with the Puttaraju analysis in

2020, the sensitivity and specificity of Alvarado group is very high.

This study demonstrated a 59 percent and 23 percent sensitivity and precision with a negative 15.6 percent appendectomy.¹⁶ This may be attributed to a wider sample of our research and a tertiary treatment center where most patients are referred to after suspicion of acute appendicitis by a trained MBBS doctor. Our analysis was similar to Noor's 2020 study, which showed 93.5% and 80.6% sensitivity and specificity respectively. The predictive positive and negative values were 92.3% and 83.3%, and the diagnostic exactness was 89.8%, but the average negative appendectomy figure was 28.7% (men: 28.2% women: 30%), higher than our sample.¹⁷ Compared to the 2020 analysis by Koujalagi et al, we performed well, showing sensitivity, accuracy, optimistic predictive value and negative Alvarado performance 76.0%, 75.0%, 97.2% and 21.4%.¹⁸ A negative appendectomy was 8.0%, similar to our research in Sharma and Koujalagi et al since their study was observational, non-randomized, observative study and there was no direct association with the patients.¹⁸

The sensitivity, specificities, PPV, NPV and Alvarado community diagnostics scores were respectively 68.3 percent, 87.9 percent, 86.3%, 71.4 percent and 86.5 percent, respectively, in 2018, for another prospective analysis to evaluate the RIPASA and the Alvarado scores at a cut-off threshold of 7.0.¹⁹ This importance was lower than our sample due to multiple differences in geography and patient ethnicity.

Clinical judgment

The therapeutic rating community had 98.51, 85.71%, 96.4% and 97.04% of sensitivities, specificity, diagnostic precision and positive predictive performance. The negative incidence of appendectomy was 3.59%. In the category of clinical judgments, the negative appendectomy rates were similar to the Guaitoli et al studies for assessment of clinical judgments with a negative appendectomy rate of 7.6 percent that was fair but negative, compared with a prospective study in 2020 which was based on a clinical evaluation with a negative appendectomy rate of 28.0%.^{20,21} Our best value was when ultrasound was used, as it was not included in the later analysis. In a study conducted by Khan et al in 2020 to determine Lintula's scoring in adults, the clinical judgment community had 100%, 84%, 100% and 16%

respectively, for sensitivity, accuracy, predictive positive, and negative, predictive values. The diagnostic precision was 91%.²² This was similar to our clinical assessment community with the exception of a higher negative Appendicectomy Score. The stronger outcome was the use of ultrasound in our analysis.

Symptoms

In our sample, 204 patients (73.38 percent) had migratory pain and 74 patients were missing (36.7 percent). 95.1% and 8.4% respectively, were sensitive and special. As with Mantoglu in 2020, with a sensitivity and accuracy of around 84 per cent, a positive gender balance of 3.18 was achieved and a negative advantage of chance of 0.5.²³ is achieved. In our sample 218 patients (78.4 per cent) reported nausea or vomiting and 50 were absent (21.6 percent). There were 94.05 percent with sensitivity and accuracy, 5 percent. This was better than the 59.3 percent vomiting report by Podda et al.²⁴ The lower importance in Podda et al is the product of his research on conservative acute appendicitis and operational control is known for most patients with frequent vomiting.²⁴ Likewise, 103 (32.2%) and 174 patients absent fever in our sample were present in fever (62.8 percent). Furthermore, sensitivity and precision were respectively 96.11% and 5.7%. In 46.7% and in 21.8% cases, this finding was higher than the Podda et al report of nausea and fever.

Laboratory investigations

In our sample, 159 (57.2 percent) and natural leucocytosis were observed in eleventy-nine patients (42.8 percent). Increase in neutrophil to 78,2% and regular to 21,2%. Leucocytosis sensitivity and specificity were respectively 99.37 and 10.9%. Neutrophilia was, however, sensitive and specific, respectively 94.9% and 4.9%. Studies also reliably shown that the bulk of appendicitis patients have high white blood cell levels with neutrophils (polymorphs >75%).

Prospective BPKIHS analysis of assessment function in diagnosis of acute appendicitis of C-reactive protein and leucocyte counts reveals that WBC's sensitivity and specificity are 78.6 and 54.8% respectively. The predictive values were positive and negative, respectively 81.0% and 51.1%. The WBC estimate was considered to be diagnostic correct at 71.7%.²¹ This was similar to our research. Likewise, in all cases C-reactive protein (N=278) was sent; in 71,9% it was lifted from 14 cases of negative appendicectomy; 4 cases were regular and 10 were raised, suggesting acute appendicitis. 94.94% and 5% respectively were sensitive and specific. In Koyuncu et al a study showed a substantial increase in the amount of C-reactive protein among strongly inflamed appendices. C-reactive protein was 84,31% sensitive, 66,66% specific in nature, 97,72% positive for predictive benefit and 20% negative for the acute appendicitis diagnosis.²⁵ C-reactive elevated susceptibility was higher in our study, but in our study, the specialty was minimal.

Imaging

The abdominal ultrasound was performed in all cases (N=278), 7 cases (50%) were found to have regular scans and 7 cases (50%) to have appendicitis, out of 14 cases of negative appendicectomy. The sensitivity and the accuracy of the PPV value was 95.65%, 5.9%, 58.3% and 50% respectively. NPV was 95.65%.

In analysis by Elsherbiny in 2020, ultrasound sensitivity 86 was 71.2%, with a species content of 46.7%, a PPV of 82.2%, a NPV of 31.8% and a precision rate of 65.7%.²⁶ The better outcome of our research was that our study did not blind radiologists to clinical and laboratory parameters. In the Elsherbiny report, they did not know the parameters of the clinic and laboratory. In a future study, ultrasound sensitivity was 89.5%, a favorable predictive value of 77, close to that of our study, when diagnosing acute appendicitis.²⁷ This was close to our research. In all cases of possible appendicitis, an abdominal ultrasound scan was done. The Alvarado score has been refined with the inclusion of the abdominal ultrasound findings, improving its reliability (AUC increased from 0.810 to 0.869). Currently, without imaging procedures we cannot foresee abdominal pain diagnostics.

Analysis for new score

The curve area (AUC) of ≥ 7 for acute appendicitis cut off from Alvarado was 0.811 and reduced to 6 for cut off from Alvarado (Table 9), Increased to maximum standard sensitivity and accuracy. In our analysis, the accuracy of the traditional clinical assessment was higher than the classification scheme Alvarado, but the gap ($p=0.297$) was not important. We found a discrepancy in the negative appendicectomy rate of group A (6.59 percent) and group B, however (3.52 percent). Six of the 20 patients went to their hospital with stomach pain in the subsequent time. Group A completed nine negative appendectomies (6.59 percent) and group B ($p=0.297$) included 5 (3.52 percent).

By weighing those parameters we have attempted to build a new score using multiple logistic regressions. The standards structure included parameters not previously included but which are relevant on the basis of our clinical experience (susception, negative urinalysis, abdominal ultrasound). The AUC increased to 0.869 by ROC analysis (95 percent CI 0.729-1.0). The initial score structure has now been refined successfully (Table 10).

Any less important predictors have been omitted from the scoring system to further refine the scoring. This also boosted the AUC. The AUC was improved due to improvements in the Alvarado ranking. In our diagnosis of acute appendicitis, we considered the Alvarado score accurate, helping the doctor in emergencies mostly in deciding on admittance and discharge and consulting a specialist. It has become easier to use the current scoring system. The outcome of the ultrasound scan requires fewer

parameters and the inclusion of a large and responsive forecast.

This score is made better by the more sensitive predictors of the Alvarado scoring, as well as by a system of imagery available in most emergency departments at BPKIHS and most emergency departments in Nepal. This new score makes decision-making efficient. Consequently, increased contrast in abdominal and pelvic tomography is only needed when diagnosis is incorrect in problem situations. It limits the patient's exposure to radiation and saves hospital expenses. The research was limited to the fact that multiple individuals carried out clinical and ultrasound assessment, which allows for differences in the results between viewers.

CONCLUSION

This study proved that the diagnostic accuracy of clinical judgment was better than Alvarado score. The new scoring system included fewer criteria as well as the addition of an important and sensitive predictor: the result of the ultrasound scan. With the help of this new score, decision making is more reliable because it contains the most sensitive predictors from the original Alvarado score and new added parameters. Contrast enhanced computed tomography of abdomen and pelvis investigation is needed only in problematic cases when the diagnosis is equivocal. It means less radiation exposure for the patient and cost saving for the hospital. The new score advised for appendectomy when score is more than equal to 6 and observation for score below 6 with repetition every 12 hourly.

Recommendations

The recommendations of new scoring system developed in our study is when the new score is more or equivalent to 6 appendectomy should be considered and observed when score is 3 to 5 with re-scoring every 12 hourly. The patient with score less than 2 may be considered for alternative diagnosis/discharge. We recommend the use of this scoring system in decision making in all the patients with suspected acute appendicitis.

Funding: No funding sources

Conflict of interest: None declared

Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. Souza N, Nugent K. Appendicitis. Am Fam Physician. 2016;93(2):142-3.
2. Stringer MD. Acute appendicitis. J Paediatr Child Health. 2017;53(11):1071-6.
3. Sippola S, Haijanen J, Viinikainen L, Grönroos J, Paajanen H, Rautio T, et al. Quality of Life and Patient Satisfaction at 7-Year Follow-up of Antibiotic Therapy vs Appendectomy for

- Uncomplicated Acute Appendicitis: A Secondary Analysis of a Randomized Clinical Trial. JAMA Surg. 2020;155(4):283-9.
4. Lohar HP, Calcuttawala MA, Nirhale DS, Athavale VS, Malhotra M, Priyadarshi N. Epidemiological aspects of appendicitis in a rural setup. Med J DY Patil Univ. 2014;7(6):753.
5. Saverio S, Podda M, Simone B, Ceresoli M, Augustin G, Gori A, et al. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. World J Emerg Surg. 2020;15(1):27.
6. Khan U, Kitar M, Krichen I, Maazoun K, Ali AR, Khalif M, et al. To determine validity of ultrasound in predicting acute appendicitis among children keeping histopathology as gold standard. Ann Med Surg. 2018;38:22-7.
7. Pinto F, Pinto A, Russo A, Coppolino F, Bracale R, Fonio P, et al. Accuracy of ultrasonography in the diagnosis of acute appendicitis in adult patients: review of the literature. Crit Ultrasound J. 2013;5(1):2.
8. Macco S, Vrouwenraets BC, Castro SM. Evaluation of scoring systems in predicting acute appendicitis in children. Surgery. 2016;160(6):1599-604.
9. Erdem H, Cetinkunar S, Daş K, Reyhan E, Deger C, Aziret M, et al. Alvarado, Eskelinen, Ohmann and Raja Isteri Pengiran Anak Saleha Appendicitis scores for diagnosis of acute appendicitis. World J Gastroenterol. 2013;19(47):9057-62.
10. Pogorelić Z, Rak S, Mrkljić I, Jurić I. Prospective validation of Alvarado score and Pediatric Appendicitis Score for the diagnosis of acute appendicitis in children. Pediatr Emerg Care. 2015;31(3):164-8.
11. Khan I, Rehman A. Application of alvarado scoring system in diagnosis of acute appendicitis. J Ayub Med Coll Abbottabad. 2005;17(3):41-4.
12. Karami MY, Niakan H, Zadebagheri N, Mardani P, Shayan Z, Deilami I. Which One is Better? Comparison of the Acute Inflammatory Response, Raja Isteri Pengiran Anak Saleha Appendicitis and Alvarado Scoring Systems. Ann Coloproctol. 2017;33(6):227-31.
13. Mán E, Simonka Z, Varga A, Rárosi F, Lázár G. Impact of the Alvarado score on the diagnosis of acute appendicitis: comparing clinical judgment, Alvarado score, and a new modified score in suspected appendicitis: a prospective, randomized clinical trial. Surg Endosc. 2014;28(8):2398-405.
14. Kollár D, McCartan DP, Bourke M, Cross KS, Dowdall J. Predicting acute appendicitis? A comparison of the Alvarado score, the Appendicitis Inflammatory Response Score and clinical assessment. World J Surg. 2015;39(1):104-9.
15. Korkut M, Bedel C, Karancı Y, Avcı A, Duyan M. Accuracy of Alvarado, Eskelinen, Ohmann, RIPASA and Tzanakis Scores in Diagnosis of Acute Appendicitis; a Cross-sectional Study. Arch Acad Emerg Med. 2020;8(1):20.

16. Puttaraju S. Study of evaluation of Alvarado scoring system in acute appendicitis. *Int Surg J.* 2019;7(1):64-8.
17. Noor S, Wahab A, Afridi G, Ullah K. Comparing Ripasa Score And Alvarado Score In An Accurate Diagnosis Of Acute Appendicitis. *J Ayub Med Coll Abbottabad.* 2020;32(1):38-41.
18. Sharma D, Koujalagi RS. A comparative study to assess efficacy of Tzanakis score and Alvarado score for effective diagnosis of patients with acute appendicitis at a tertiary care centre in North Karnataka: a one-year prospective analytical study. *Int Surg J.* 2020;7(6):1742-5.
19. Frountzas M, Stergios K, Kopsini D, Schizas D, Kontzoglou K, Toutouzas K. Alvarado or RIPASA score for diagnosis of acute appendicitis? A meta-analysis of randomized trials. *Int Surg J.* 2018;56:307-14.
20. Gwaitoli E, Gallo G, Cardone E, Conti L, Famularo S, Formisano G, et al. Consensus Statement of the Italian Polispecialistic Society of Young Surgeons (SPIGC): Diagnosis and Treatment of Acute Appendicitis. *J Invest Surg.* 2020;1-15.
21. Sumanth G. Significance of C-Reactive Protein in Diagnosis of Acute Appendicitis. *J Med Sci Clin Res.* 2020;8.
22. Khan S, Usama M, Basir Y, Muhammad S, Jawad M, Khan T, et al. Evaluation Of Modified Alvarado, Ripasa And Lintula Scoring System As Diagnostic Tools For Acute Appendicitis. *J Ayub Med Coll Abbottabad.* 2020;32(1):46-50.
23. 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(1):34.
24. Podda M, Poillucci G, Pacella D, Mortola L, Canfora A, Aresu S, et al. Appendectomy versus conservative treatment with antibiotics for patients with uncomplicated acute appendicitis: a propensity score-matched analysis of patient-centered outcomes (the ACTUAA prospective multicenter trial). *Int J Colorectal Dis.* 2021;36(3):589-98.
25. Koyuncu S, Ismail O. The role of C-reactive protein to lymphocyte ratio in the differentiation of acute and perforated appendicitis. *Ulus Travma Acil Cerrahi Derg.* 2020;26(5):760-4.
26. Elsherbiny MW, Emile SH, Abdelnaby M, Khafagy W, Elshobaky A. Assessment of the Diagnostic Accuracy of Alvarado Scoring System Combined with Focused Ultrasound in the Diagnosis of Acute Appendicitis. *Br J Surg.* 2020;107(12):594-5.
27. Bappayya S, Chen F, Alderuccio M, Xu E, Vootukuru N, Lee JC. Non-diagnostic sonography may reduce negative appendectomy rate in women when combined with abbreviated Alvarado score. *ANZ J Surg.* 2021;91(4):609-15.
28. Sigdel GS, Lakhey PJ, Mishra PR. Tzanakis score vs. Alvarado score in acute appendicitis. *JNMA J Nepal Med Assoc.* 2010;49(178):96-9.
29. Meltzer AC, Baumann BM, Chen EH, Shofer FS, Mills AM. Poor sensitivity of a modified Alvarado score in adults with suspected appendicitis. *Ann Emerg Med.* 2013;62(2):126-31.

Cite this article as: Bhusal I, Agarwal CS, Gupta RK, Sah SP. Diagnostic accuracy of acute appendicitis by comparing clinical judgment of surgeons and Alvarado scoring in suspected acute appendicitis. *Int Surg J* 2021;8:2725-33.