International Surgery Journal
Yuen A. Int Surg J. 2025 Apr;12(4):534-538
http://www.ijsurgery.com
pISSN 2349-3305 | eISSN 2349-2902

DOI: https://dx.doi.org/10.18203/2349-2902.isj20250811

Original Research Article

Validation of the APPEND clinical decision tool for the diagnosis of acute appendicitis

Alexander Yuen*

Department of Surgery, Mater Hospital Brisbane, Brisbane, Queensland, Australia

Received: 25 February 2025
Revised: 18 March 2025
Accepted: 19 March 2025

***Correspondence:**
Dr. Alexander Yuen,
E-mail: alexanderyuen@outlook.com

ABSTRACT

Background: The APPEND score is a clinical decision tool (CDT) for stratifying the risk of patients presenting with right lower quadrant (RLQ) abdominal pain having acute appendicitis. The aim of this study was to validate the APPEND score in an Australian metropolitan hospital population.
Methods: A retrospective review of consecutive patients from January to December 2024 at least 16 years of age that were referred to the general surgery service with acute RLQ pain in an Australian metropolitan hospital. The diagnostic properties including sensitivity, specificities, positive predictive value (PPV), negative predictive values (NPV) and area under the curve (AUC) of the APPEND score were assessed and then compared with the Alvarado score. Agreement between the scores were also assessed.
Results: Of the 229 patients referred for acute RLQ pain, the APPEND score demonstrated high sensitivity of 98% in the low risk group identified with a score of 1 or less and a PPV of 100% in the high risk group (APPEND score ≥5. It has good agreement with the Alvarado score when identifying the low risk group (p=1) but low agreement when identifying the high risk group (p<0.001). The AUC was reasonable at 0.746 compared to the 0.823, p=0.02.
Conclusions: The APPEND score shows promise as a CDT for acute appendicitis in Australian metropolitan populations, potentially reducing negative appendectomy rates and radiological resources. While the APPEND score offers simplicity and strong performance, further multicenter validation is crucial before widespread clinical adoption in Australia.

Keywords: Acute appendicitis, APPEND score, Clinical decision tool, Negative appendicectomy rate

INTRODUCTION

Acute appendicitis is one of the most frequent surgical emergency presentations, demanding significant healthcare resources. In Australia alone, the 2022-2023 period saw 35,917 emergency admissions for appendicitis, resulting in 94,115 hospital bed days.[1] Despite its prevalence, accurate and timely diagnosis remains a clinical challenge due to the overlapping symptomatology with a wide range of other pathologies.[2] Historically, a certain rate of negative appendicectomies (NAR) has been accepted to minimize the risk of missed or perforated appendicitis. While historical NARs ranged from 10-20%, the increased availability of imaging like computed tomography (CT) has prompted calls for a reduction in this rate.[3-5] Negative appendicectomies are associated with potential postoperative complications, including wound infection and collections.[6]

Clinical decision tools (CDT) offer a promising avenue for improving diagnostic accuracy and reducing NAR, thereby optimizing patient care and resource utilization.[7] The Alvarado score is the most well-known and this has demonstrated a reduction in NAR to 6-8%.[8] Recently, the APPEND score, developed at Middlemore Hospital in South Auckland, New Zealand, has emerged as an alternative.[9] This score utilizes six equally weighted clinical and laboratory parameters - anorexia, migratory

pain, localized peritonitis, elevated C-reactive protein (CRP), neutrophilia, and male gender (dude). Patients are then categorized into high-risk (5-6), intermediate-risk (2-4), or low-risk (0-1) groups, guiding clinical decision-making towards immediate operative intervention, further observation/investigation, or discharge, respectively. The APPEND score's simplicity and reported high sensitivity and specificity make it an attractive alternative to the Alvarado score.

However, the diagnostic performance of CDTs can vary across different populations. For example, the Alvarado score has been reported to exhibit suboptimal performance in Asian and Middle Eastern populations, where the RIPASA score has demonstrated superior accuracy.[10] Therefore, this study aims to evaluate the performance of the APPEND score in an Australian population. We will compare the APPEND score's diagnostic accuracy with that of the Alvarado score and assess its efficacy within the Australian healthcare context.

METHODS

Study design and setting

This retrospective cohort study examined patients who were referred to the general surgical service with acute right lower quadrant (RLQ) pain at a metropolitan hospital, the Mater Hospital Brisbane in Brisbane, Australia. Data were collected for a 12-month period, from 01 January 2024 to 31 December 2024. Inclusion criteria were patients aged 16 years or older presenting with acute RLQ pain of less than 7 days' duration. Exclusion criteria were patients with a history of prior appendicectomy. All patients meeting these criteria within the general surgical service's referral database during the study period were included and no sampling was performed.

Data collection

Patient data were extracted from electronic medical records in February 2025 by the researcher. Collected variables included age, gender, clinical symptoms, laboratory test results and surgical procedures performed.

APPEND score

This score was calculated by assigning 1 point to each of the following: anorexia, migratory pain, peritoneal signs, elevated CRP, neutrophilia, and male gender. The total APPEND score was the sum of these points.

Alvarado score

This score was calculated by assigning points to the following: right lower quadrant tenderness (2 points), elevated temperature (>37.3°C; 1 point), rebound tenderness (1 point), migration of pain to the RLQ (1 point), anorexia (1 point), nausea or vomiting (1 point), leukocytosis (WBC >10,000/µl; 2 points), and a left shift (>75% of neutrophils; 1 point). The total Alvarado score was the sum of these points.

Statistical analysis

Statistical analyses were performed using R (version 4.4.2; The R Foundation for Statistical Computing, Vienna, Austria), with the Tidyverse (version 2.0.0) and pROC (version 1.18.5) packages. Diagnostic performance of the APPEND and Alvarado scores was evaluated by calculating sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) at each score threshold. McNemar's test was used to determine statistical differences between the scores' low and high risk groups. Receiver operating characteristic (ROC) curve analysis was used to assess the overall diagnostic accuracy of each scoring system. Differences between the ROC curves were assessed using DeLong's test. A p value of less than 0.05 was considered statistically significant.

RESULTS

A total of 229 patients were consecutively referred to our general surgical service with acute right lower quadrant abdominal pain with no history of appendicectomy during the period of 1 January 2024 and 31 December 2024. Their demographics and characteristics are summarised in Table 1. Thirteen patients (5.68%) had parameters missing and were excluded. The most common missing value was the CRP (n=9 or 3.9%) and then neutrophils (n=1 or 0.004%). The rest of these did not have adequate or accessible notes (n=3 or 1.3%).

Of the remaining 216 patients left, 158 patients received an operation and 134 (84.8%) of patients were confirmed to have appendicitis which leaves 24 patients who had a negative appendicectomy, a rate of 15.2%. Of note, this included one with a diagnosis of high-grade appendiceal mucinous neoplasm, one with low-grade appendiceal mucinous neoplasm and one patient with diverticulitis in the appendix. Hence, the nonproductive exploration operation rate was 13.3% (n=21).

Three patients (2.19%) did not receive an operation. Two of the patients were treated conservatively and one patient discharged against medical advice.

Tables 2 and 3 displays the rate of appendicitis and the diagnostic properties of the APPEND score. Using a cut off value of ≥5, the positive predictive value was 1 (i.e. all of these patients will have appendicitis) which suggests that those scoring 5 and 6 will almost definitely have appendicitis. In fact, the positive predictive value (PPV) in scores ≥3 was quite high at 0.89. The negative predictive value (NPV) with a cutoff of ≥1 is 0.81 though which suggests that 19% would have a diagnosis of appendicitis.

Table 1: Demographics and characteristics of patients referred for acute right lower quadrant abdominal pain including age, gender, patients excluded from the study and patients receiving an operation. These are separated into a final diagnosis of appendicitis versus non-appendicitis.

Variables	Total	Appendicitis	Non-appendicitis
Total, N	229	147	82
Age, mean	35.3	37.3	31.8
Gender, N (%)			
Male	134 (58.5)	76 (51.7)	20 (24.4)
Female	95 (41.5)	71 (48.3)	62 (75.6)
Excluded, N (%)	13 (5.68)	10 (6.80)	3 (3.66)
Missing CRP, N (%)	9 (3.93)	7 (4.76)	2 (2.44)
Missing neut, N (%)	1(0.00437)	1 (0.68)	0 (0)
No notes, N (%)	3 (1.31)	2 (1.36)	1 (1.22)
Operation, N (%)			
Yes	158 (73.1)	134 (97.8)	24 (30.4)
No	58 (26.9)	3 (2.19)	55 (69.6)

Table 2: Rate of appendicitis based on APPEND score.

APPEND score	Appendicitis (n=137) (%)	Non-appendicitis (n=79) (%)
0	3 (19)	13 (81)
1	25 (49)	26 (51)
2	37 (54)	31 (46)
3	33 (89)	4 (11)
4	29 (85)	5 (15)
5	10 (100)	0 (0)
6	0 (0)	0 (0)

Table 3: Diagnostic properties including sensitivity, specificity, positive and negative predictive values for each cut off value of APPEND score. Note that no patient in the study received a score of 6.

Parameters	APPEND score				
	≥1	**≥2**	**≥3**	**≥4**	**≥5**
Sensitivity	0.98	0.8	0.53	0.28	0.073
Specificity	0.16	0.49	0.89	0.94	1.00
Positive predictive value	0.67	0.73	0.89	0.89	1.00
Negative predictive value	0.81	0.58	0.52	0.43	0.38

Comparison with Alvarado score

Table 4 displays the comparison of diagnostic properties at different cut off points to rule out (exclude) and rule in appendicitis using the APPEND and Alvarado scores. The APPEND score demonstrated a higher sensitivity than the Alvarado score at its rule out score of 0.98 (APPEND ≤1) versus 0.88 (Alvarado ≤4). The negative predictive value was 0.81 compared to 0.75. It also demonstrated a very high positive predictive value of 1.0 (APPEND ≥5) compared to 0.91 (Alvarado ≥7).

Table 4: Diagnostic properties including sensitivity, specificity, positive and negative predictive values for each cut off value of APPEND score. Note that no patient in the study received a score of 6.

Variables	APPEND score	Alvarado score
Rule out	≤1	≤4
Sensitivity	0.98	0.88
Specificity	0.16	0.63
Positive predictive value	0.67	0.81
Negative predictive value	0.81	0.75
Rule in	≥5	≥7
Sensitivity	0.073	0.55
Specificity	1.0	0.95
Positive predictive value	1.0	0.91
Negative predictive value	0.38	0.44

Table 5 summarises the agreement between the APPEND and Alvarado scores when identifying low risk group (≤1 and ≤4, respectively). This demonstrated good agreement (p=1) with n=56 (27.3%) of patients in disagreement.

Table 5: Agreement between the APPEND and Alvarado scores when identifying the low risk groups.

Low risk groups (APPEND score)	Alvarado score	
	≤4	**≥5**
≤1	39	28
≥2	28	121

Table 6 summarises the agreement between the APPEND and Alvarado scores when identifying high risk groups (≥5 and ≥7, respectively). This demonstrated a significant difference (p<0.0001) with n=77 (35.6%) of patients in disagreement.

Figure 1 demonstrates the paired receiver operating characteristic (ROC) curves of the APPEND (blue) and the

Alvarado (red) scores. Importantly, the APPEND score shows very similar sensitivities and specificities and the extremes of the score, indicating good performance at ruling in and out appendicitis at the extremes. The area under the curve (AUC) is 0.746 (95% CI; 0.682-0.809) for the APPEND score compared to 0.823 (95% CI; 0.764-0.881) for the Alvarado score. This was significantly different with a p value of 0.02.

Table 6: Agreement between the APPEND and Alvarado scores when identifying the low risk groups.

High risk groups (APPEND score)	Alvarado score	
	≤4	≥5
≥5	9	1
≤4	76	130

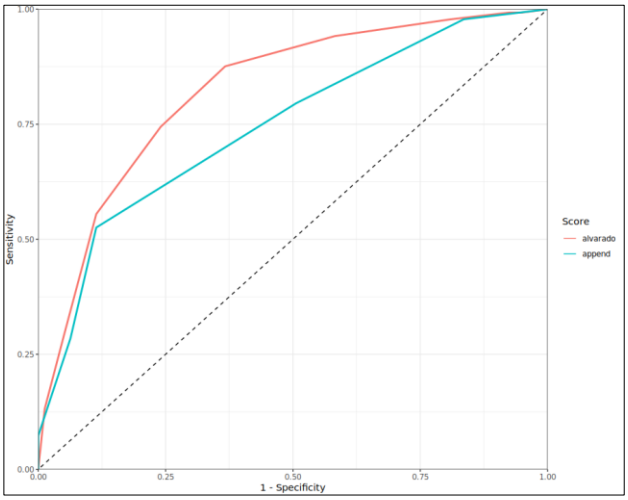

Figure 1: Receiver operating characteristic (ROC) curves of the APPEND score (AUC=0.746) and the Alvarado score (AUC=0.823) with a reference line (dash). This was significantly different with a p value=0.02.

DISCUSSION

This retrospective cohort study is the first study to externally evaluate the diagnostic properties of the APPEND score for stratifying the risk of patients presenting with right lower quadrant pain having acute appendicitis in an Australian metropolitan hospital setting.

The APPEND score was demonstrated to be an excellent CDT with a PPV of 100% and specificity of 100% when the APPEND score was ≥5. Compared to studies in the New Zealand population, this performed better at this cut off with reported PPVs of 85-88% and specificities of 94-96%.[9,11] Scores of ≥3 appeared to have high PPVs of 89% and specificities of 89%, higher than the 65-71% and 61-64% previously recorded.[9,11] This could be considered a reasonable cutoff to indicate appendicitis in our cohort but further studies may be required to confirm this.

A cutoff of the APPEND score of ≤1, which has been suggested as the cutoff value to rule out appendicitis, had a high sensitivity of 98%, slightly lower but comparable to previous studies of 99.8-100%.[9,11] However, the NPV at this cutoff was 81%, lower than the 98.6% reported in the derivation study.[9] Hence, a cutoff score of ≤1 can reasonably be used to rule out appendicitis albeit with slightly less confidence.

In the comparison of the APPEND score with the Alvarado score, the results of this study were consistent with previous studies showing higher sensitivities (98% versus 88%) and higher NPVs (81% versus 75%) at the low cut off APPEND scores of ≤1.[9] The Alvarado score performed very well in this cohort at ruling in with a specificity of 95% and PPV of 91% which is mostly higher than the previously reported specificity of 37.5-87.5% but slightly lower than the PPV of 97.3-98.86% reported in the same studies.[13-15]

The area under the receiver operating characteristic curve (AUC) for the APPEND score was 0.746. This was lower than the 0.84 reported in the derivation study and 0.83 in Pasifika population but similar to one particular study of 0.747.[9,11,12] This confirms that there may be differences in the diagnostic properties of the APPEND score in the Australian population with unique genetic, dietary, environmental and financial factors. The AUC for the Alvarado score in this study was 0.823 which is comparable but mostly higher than other reviews of 0.725-0.829.[9,16-18] This explains why we were unable to verify that the APPEND score performed better in this cohort as others have in theirs.[9]

The clinical implication of these findings is the potential to reduce negative appendectomy rates (NAR) without relying on costly radiological investigations. Our NAR was 15.2% and a non-productive operation rate of 13.3% which is well within the 10-20% which is widely accepted.[3] Patients in the high-risk group (APPEND score ≥5) could potentially proceed directly to surgery based on clinical assessment, as our study observed a 0% NAR in this group. This is much lower than the study on the New Zealand population of 14.8%.[9] The intermediate-risk group with an APPEND score of 2-4 would benefit from a combination of clinical and serial assessment with further investigations. APPEND scores of ≤1 can be used to rule out appendicitis and alternative diagnosis should be considered.

Limitations

However, several limitations must be acknowledged. The retrospective design may possibly introduce potential biases and confounding factors. The limited sample size, especially the small number of patients in the high-risk group (n=10), restricts statistical power and subgroup analyses. Additionally, the single-center nature of the study limits the generalizability of the findings to other

populations and regions, particularly outside of Australian metropolitan settings.

CONCLUSION

The APPEND score demonstrates promise as a tool to aid the clinical diagnosis of acute appendicitis in an Australian metropolitan population. Its simplicity and strong performance in ruling in appendicitis compared to the Alvarado score, suggest potential for reducing NAR and radiological resources. However, larger, multicenter, prospective validation studies are essential to confirm these findings and establish the score's broader applicability within the Australian population before widespread clinical implementation may be considered.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. Australian Institute of Health and Welfare. 2024. Admitted patients. Available at: https://www.aihw.gov.au/reports-data/myhospitals/sectors/admitted-patients. Accessed on 25 February 2025.
2. Jearwattanakanok K, Yamada S, Suntornlimsiri W, Smuthtai W, Patumanond J. Clinical scoring for diagnosis of acute lower abdominal pain in female of reproductive age. Emerg Med Int. 2013;2013:730167.
3. Colson M, Skinner KA, Dunnington G. High negative appendectomy rates are no longer acceptable. Am J Surg. 1997;174(6):723-6.
4. Mariadason JG, Wang WN, Wallack MK, Belmonte A, Matari H. 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 R Coll Surg Engl. 2012;94(6):395-401.
5. Drake FT, Florence MG, Johnson MG, Jurkovich GJ, Kwon S, Schmidt Z, et al; SCOAP Collaborative. Progress in the diagnosis of appendicitis: a report from Washington State's Surgical Care and Outcomes Assessment Program. Ann Surg. 2012;256(4):586-94.
6. Lee M, Paavana T, Mazari F, Wilson TR. The morbidity of negative appendicectomy. Ann R Coll Surg Engl. 2014;96(7):517-20.
7. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med. 2006;144(3):201-9.
8. Awayshih MMA, Nofal MN, Yousef AJ. Evaluation of Alvarado score in diagnosing acute appendicitis. Pan Afr Med J. 2019;34:15.
9. Mikaere H, Zeng I, Lauti M, Kularatna M, MacCormick AD. Derivation and validation of the APPEND score: an acute appendicitis clinical prediction rule. ANZ J Surg. 2018;88(4):E303-7.
10. Malik MU, Connelly TM, Awan F, Pretorius F, Fiuza-Castineira C, El Faedy O, et al. The RIPASA score is sensitive and specific for the diagnosis of acute appendicitis in a western population. Int J Colorectal Dis. 2017;32(4):491-7.
11. Pitesa R, Hill AG, MacCormick AD. Modified APPEND Score for the Diagnosis of Acute Appendicitis in a New Zealand Pasifika Population. World J Surg. 2025;19.
12. Fagan PVB, Stanfield B, Korduke O, Henderson N, Kodeda K. Validation of the appendicitis inflammatory response score in a provincial population. Int Surg J. 2020;7:3540-5.
13. Pifeleti S, Hansell D, Kaspar A. Sensitivity and specificity of the Alvarado Score for the timely differential diagnosis of acute appendicitis for a case series in Samoa. Ann Med Surg (Lond). 2022;73:103219.
14. Awayshih MMA, Nofal MN, Yousef AJ. Evaluation of Alvarado score in diagnosing acute appendicitis. Pan Afr Med J. 2019;34:15.
15. Al-Tarakji M, Zarour A, Singh R, Ghali MS. The Role of Alvarado Score in Predicting Acute Appendicitis and Its Severity in Correlation to Histopathology: A Retrospective Study in a Qatar Population. Cureus. 2022;14(7):e26902.
16. Polat Düzgün A, Pehlevan Özel H, Şahingöz E, Dinç T. Identification of complicated and non-complicated appendicitis: a new alvarado-based scoring system. Ulus Travma Acil Cerrahi Derg. 2024;30(2):101-6.
17. Polat Düzgün A, Pehlevan Özel H, Şahingöz E, Dinç T. Identification of complicated and non-complicated appendicitis: a new alvarado-based scoring system. Ulus Travma Acil Cerrahi Derg. 2024;30(2):101-6.
18. Amnedya TT, Wijaya S. Comparison of sensitivity and specificity of Alvarado score and TZANAKIS score in enforcement of acute appendicitis diagnosis. J Soc Sci. 2023;4(3):739-44.

Cite this article as: Yuen A. Validation of the APPEND clinical decision tool for the diagnosis of acute appendicitis. Int Surg J 2025;12:534-8.