

## A simplified appendicitis score in the diagnosis of acute appendicitis

### 診斷急性闌尾炎的一個簡化闌尾炎指數

PL Goh 吳柏良

**Introduction:** Scoring systems such as the Modified Alvarado Score (MAS) can help in the early diagnosis of acute appendicitis (AA) but is cumbersome to use and has not found widespread popularity. A more robust, user-friendly Simplified Appendicitis Score (SAS) was investigated. **Methods:** Patients presenting with suspected AA were prospectively enrolled. The performance of the SAS (using only 5 variables – migratory pain, right lower quadrant tenderness, rebound pain, fever >37.3 degrees Celsius and leucocytosis >12,000/uL) was analysed and compared to the MAS. **Results:** Out of 238 patients enrolled over four months, 95 (39.9%) patients underwent appendectomy, of which 81 patients had histologically proven AA. A MAS of  $\geq 8$  was 92.4% specific for ruling in AA, while a score of  $< 5$  was 91.4% sensitive in ruling out AA. A SAS of  $\geq 6$  was 91.7% specific for ruling in AA, while a score of  $< 4$  was 90.1% sensitive in ruling out AA. **Conclusions:** The performance of the MAS can be maintained by omitting the two subjective variables – "anorexia" and "nausea/vomiting", and increasing the leucocytosis cut-off level. A SAS using only 5 variables performed as well as the original MAS. (*Hong Kong j.emerg.med.* 2010;17:230-235)

**引言：**分數系統如「阿爾瓦拉多修訂指數」可幫助及早診斷急性闌尾炎，但難以使用及不流行。故研究一個較穩健及容易使用的「簡化闌尾炎指數」。 **方法：**前瞻性地招募懷疑闌尾炎的求診病人。分析簡化闌尾炎指數（只使用5個變數：痛楚轉移、右下腹有觸痛、回彈痛、發燒高於攝氏37.3度及白細胞增多高於每微升12000）的表現及與阿爾瓦拉多修訂指數比較。 **結果：**於4個月的時間招募238名病人，95名（39.9%）病人經受闌尾切除術，其中81名病人在組織學上證明為急性闌尾炎。「阿爾瓦拉多修訂指數」 $\geq 8$ 確定急性闌尾炎的特異性為92.4%，而指數 $< 5$ 排除急性闌尾炎的敏感度為91.4%。「簡化闌尾炎指數」 $\geq 6$ 確定急性闌尾炎的特異性為91.7%，而指數 $< 4$ 排除急性闌尾炎的敏感度為90.1%。 **結論：**刪除兩個主觀的變數：「厭食」及「作嘔/嘔吐」及提高白細胞增多的截止水平，仍可保持「阿爾瓦拉多修訂指數」的表現。只使用5個變數的「簡化闌尾炎指數」的表現與原本的「阿爾瓦拉多修訂指數」相同。

**Keywords:** Acute abdomen, early diagnosis, ROC curve

**關鍵詞：**急腹症、及早診斷、接收運行特性曲線

## Introduction

Acute appendicitis (AA) is a common cause of abdominal pain, yet even in this day and age, it can be difficult to

diagnose, especially in the early stages of the disease process. Failure to make an early, correct diagnosis contributes to appreciable morbidity and mortality.

Various adjuncts to increase diagnostic accuracy are available but they are by no means ideal. Computer-aided diagnoses often require a proprietary and costly program and doctors are often reluctant to use such aids because they 'impinge' upon the 'art' of clinical judgement. The use of laparoscopy enables the

Correspondence to:

PL Goh, MBBS(Singapore), FRCSEd(A&E), FAMS(Singapore)

Changi General Hospital, Accident and Emergency Department,  
2 Simei Street 3, Singapore 529889, Singapore

Email: pak\_liang\_goh@cgh.com.sg

abdominal cavity to be visualised directly, but is disadvantaged by its invasiveness. Advanced imaging using ultrasound and computed tomography has shown much promise, but the accurate interpretation of the results is not easy without specialised radiological training. The use of scoring systems is attractive because it requires no special equipment, is user-friendly and easily comprehensible to the clinician. Numerous scoring systems have been developed over the years and promising results have been reported, yet many surgeons remain skeptical of their results and they have yet to find widespread clinical use.<sup>1-4</sup> Their use in primary care clinics and emergency departments is also limited, perhaps by the difficulty in remembering the various components of each score.

In this study, the authors sought to validate the widely used Modified Alvarado Score (MAS) and to investigate whether simple modification of its component variables could result in a more robust and user-friendly score.

## Methods

This prospective study was conducted at Changi General Hospital, an 800-bed tertiary hospital in Singapore. Over a 4-month period (July to October 2006), all patients presenting to the emergency department with features suggestive of AA were admitted to the hospital. A checklist of clinical features likely to be present in AA was filled by the attending physician independently. Upon discharge, their inpatient charts were reviewed for the following information: inpatient progress, relevant investigations, definitive surgical procedures done (if any), and eventual discharge diagnosis.

With reference to the MAS, the authors felt that the variables "anorexia" and "nausea/vomiting", being subjective symptoms, were prone to cultural and ethnic variation in reporting by the patient, and therefore less reliable.<sup>5</sup> We postulated that the omission of these two variables would result in a more objective score. In addition, the use of 10,000/uL as the cut-off for leucocytosis was felt to be too low as most reference sources already consider 10,800/uL to 11,100/uL to be the upper end of the normal range of white blood

cell count (WBC).<sup>6,7</sup> In addition, smoking can cause WBC to be elevated to 11,200/uL in women and 11,500/uL in men.<sup>8</sup> For this reason, the authors postulated that a raised WBC cut-off of 12,000/uL might be more appropriate.

The performance of a Simplified Appendicitis Score (SAS) using only 5 variables (migratory pain, right lower quadrant [RLQ] tenderness, rebound pain, fever >37.3 degrees Celsius and WBC >12,000/uL) was analyzed and compared to the original MAS. As with the original MAS, the presence of RLQ tenderness and WBC >12,000/uL scored double compared to the rest of the variables.

Using the histologically proven cases of AA as the standard, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and area under the curve (AUC) with the associated 95% confidence intervals of the MAS and the SAS were calculated. The AUC of the two scores were compared. The statistical significance was set at  $p < 0.05$ . All the analyses were performed using STATA version 9.1.

## Results

Within the four-month study period, a total of 238 patients were admitted with suspected AA. After inpatient surgical review, serial assessments and necessary investigations, 143 (60.1%) patients were eventually discharged without undergoing appendectomy; 95 (39.9%) patients underwent appendectomy, of which 81 had histologically proven AA, while 14 patients had normal appendectomy.

The ability of the MAS in predicting AA at various cut-offs was then charted in the form of a receiver operating characteristic (ROC) curve. The performance of the MAS  $\geq 7$  in predicting AA was as follows: sensitivity 64.2%, specificity 80.9%, PPV 63.4%, and NPV 81.4% (Table 1). A score of  $\geq 8$  was more than 90% specific for ruling in AA (specificity 92.4%, sensitivity 29.6%), whereas a score of  $< 5$  was more than 90% sensitive in ruling out AA (sensitivity 91.4%, specificity 36.9%).

A combination omitting both "anorexia" and "nausea/vomiting" with leucocytosis cut-off of 12,000/uL yielded the following results: a score of  $\geq 5$  predicted AA with sensitivity 77.8%, specificity 73.3%, PPV 60%, NPV 86.5% (Table 2). A score of  $\geq 6$  was more than 90% specific in ruling in AA (specificity 91.7%, sensitivity 44.4%), while a score of  $< 4$  was more than 90% sensitive in ruling out AA (sensitivity 90.1%, specificity 52.2%).

A comparison of the AUC of MAS and SAS did not reveal any significant differences (Table 3).

### Discussion

The original score derived by Alvarado utilised 8 predictive factors, with 10 as the maximum score.<sup>9</sup> Kalan later described the MAS where one of the factors (neutrophil left shift) was dropped, thus it utilised only 7 predictive variables, with 9 as the maximum score.<sup>10</sup> Since then, there have been numerous validation studies of both the original Alvarado score as well as the MAS, not surprisingly with vastly differing ranges of sensitivity and specificity.<sup>11-14</sup> In a review of the various diagnostic scores for AA, Ohmann et al found

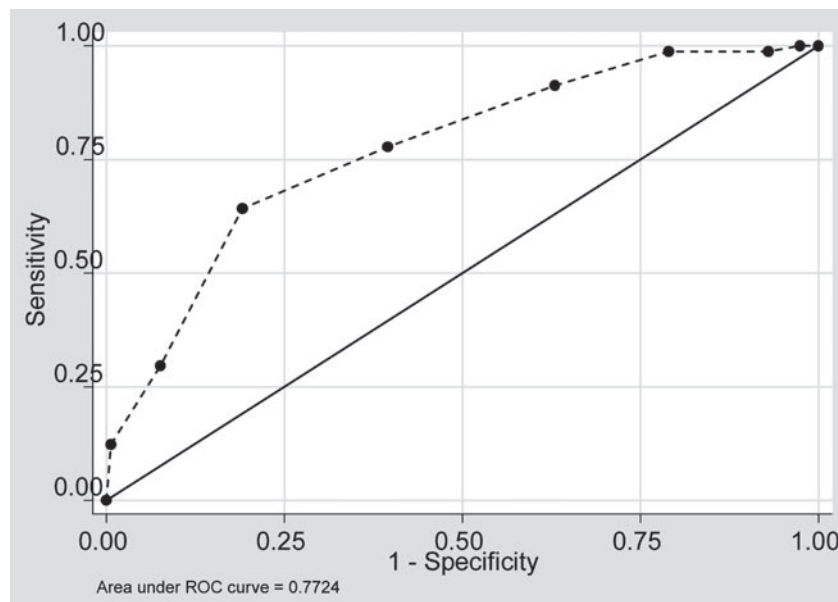
Table 1. Modified Alvarado Score (MAS) and ROC curve

Histology proven	MAS									Total
	1	2	3	4	5	6	7	8	9	
No	4	7	22	25	37	32	18	11	1	157
Yes	0	1	0	6	11	11	28	14	10	81
Total	4	8	22	31	48	43	46	25	11	238

	MAS							
	$\geq 2$	$\geq 3$	$\geq 4$	$\geq 5$	$\geq 6$	$\geq 7$	$\geq 8$	$\geq 9$
Sensitivity (%)	100	98.77	98.77	91.36	77.78	64.20	29.63	12.35
Specificity (%)	2.55	7.01	21.02	36.94	60.51	80.89	92.36	99.36
PPV (%)	34.62	35.4	39.22	42.77	50.4	63.41	66.67	90.91
NPV (%)	100	91.67	97.06	89.23	84.07	81.41	71.78	68.72
Correctly classified (%)	35.71	38.24	47.48	55.46	66.39	75.21	71.01	69.75

NPV=negative predictive value; PPV=positive predictive value



the Alvarado score to have sensitivity of 64% and specificity of 84%, figures similar to the performance of the MAS in our present study.<sup>15</sup> A more recent systemic review of clinical decision tools for acute abdominal pain by Liu et al found sensitivities ranging from 53-99% and specificities ranging from 30-99%.<sup>16</sup>

Perhaps not surprisingly, the use of scoring systems in the diagnosis of acute appendicitis has never found widespread popularity. However, most studies quote their performance at the point of optimal balance between sensitivity and specificity, but in practical

terms, this is not the ideal way to utilise such scores. A much better way to use such scoring systems is to choose a point that optimises either sensitivity or specificity, depending on the clinical goal. For example, a primary care physician or an admitting emergency physician is mainly concerned about not missing AA, and therefore would be most interested in the lower end of the score (which optimises sensitivity). On the other hand, a surgeon would likely be most interested in considering those patients with a higher score (which optimises specificity) for surgery. In this manner, different practitioners may use the same score to achieve different objectives.

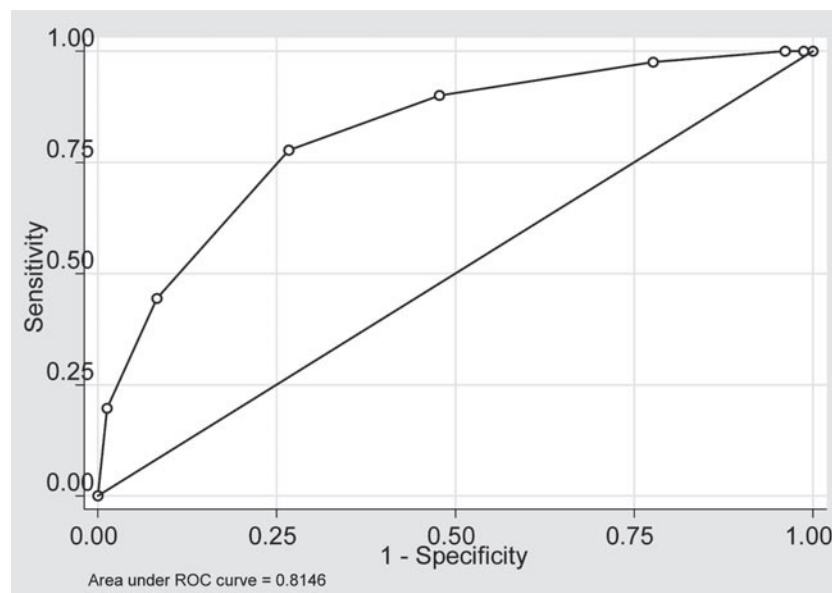
Table 2. Simplified Appendicitis Score (SAS): migration pain + tenderness + rebound + fever + WBC >12,000/uL; and ROC curve

Histology proven	SAS								Total
	0	1	2	3	4	5	6	7	
No	2	4	29	47	33	29	11	2	157
Yes	0	0	2	6	10	27	20	16	81
<b>Total</b>	<b>2</b>	<b>4</b>	<b>31</b>	<b>53</b>	<b>43</b>	<b>56</b>	<b>31</b>	<b>18</b>	<b>238</b>

	SAS						
	≥1	≥2	≥3	≥4	≥5	≥6	≥7
Sensitivity (%)	100	100	97.53	90.12	77.78	44.44	19.75
Specificity (%)	21.27	3.82	22.29	52.23	73.25	91.72	98.73
PPV (%)	34.32	34.91	39.30	49.32	60	73.47	88.89
NPV (%)	100	100	94.59	91.11	86.47	76.19	70.45
Correctly classified (%)	34.87	36.55	47.90	65.13	74.79	75.63	71.85

NPV=negative predictive value; PPV=positive predictive value; WBC=white blood cell count



In our study, we have found that a simplified appendicitis score (SAS) using only 5 variables had similar performance to the MAS. As discussed above, different SAS cut-offs may be used for different purposes. A SAS of  $\geq 5$  yielded an optimal balance of 77.8% sensitivity and 73.3% specificity. A SAS of  $< 4$  may be used as a 'screening' tool to 'exclude' AA with a sensitivity of 90.1%, while a SAS of  $\geq 6$  'rules in' AA with a specificity of 91.7%. Should more 'stringency' be required, a SAS of  $< 3$  would exclude AA with a sensitivity of 97.5%, while a SAS of 7 rules in AA with a specificity of 98.7%.

It is reasonable to expect that a simpler score with fewer variables will translate in practical terms to be more frequently used, but this remains unproven. This new scoring system must also be prospectively validated before it can be considered as a replacement for the MAS. A follow-up study will attempt to address these two issues.

The study has several limitations. As discussed, the performance of such diagnostic scores is very much dependent on the differing thresholds for admission and surgery for such patients, and will therefore vary widely amongst various institutions and practices. The use of any diagnostic score, regardless of performance, can at best only be a diagnostic aid, and should never be used to replace astute and careful clinical judgement.

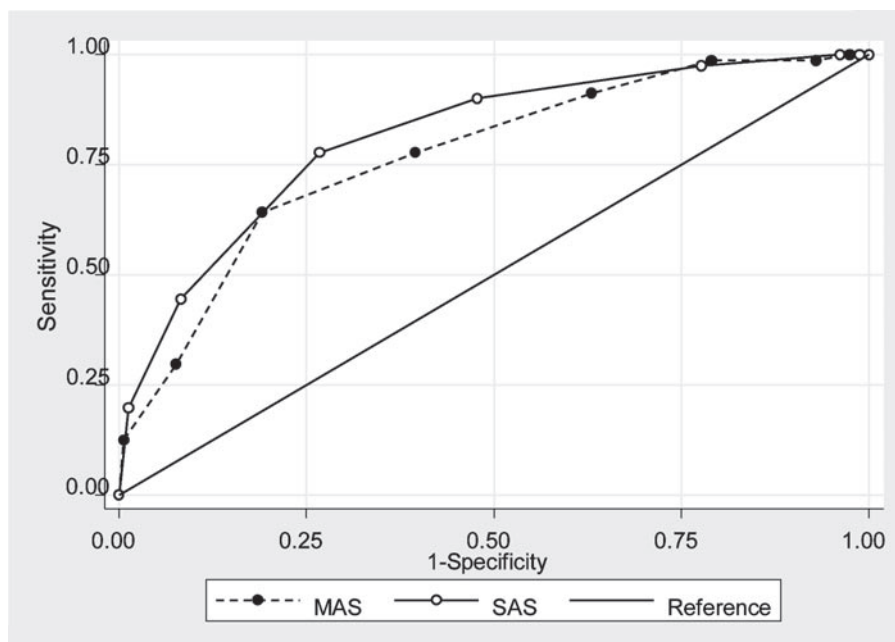
## Conclusion

The performance of the MAS can be maintained by omitting the subjective variables (anorexia and nausea/vomiting), and increasing the leucocytosis cut-off level. A simplified appendicitis score (SAS) using only 5 variables (migratory pain, RLQ tenderness, rebound pain, fever  $> 37.3$  degrees Celsius and WBC  $> 12,000/\mu\text{L}$ ) performed as well as the original MAS.

**Table 3.** Comparison of area under the ROC curve (AUC) of Modified Alvarado Score (MAS) and Simplified Appendicitis Score (SAS)

Scores	<i>n</i>	AUC	Standard Error of AUC	95% Confidence Interval of AUC
MAS	238	0.77	0.03	0.71 to 0.83
SAS	238	0.81	0.03	0.76 to 0.87

There was no significant difference in the area under ROC between MAS and SAS ( $p=0.05$ ).



## References

- Ramirez JM, Deus J. Practical score to aid decision making in doubtful cases of appendicitis. *Br J Surg* 1994;81(5):680-3.
- Eskelinen M, Ikonen J, Lipponen P. Sex-specific diagnostic scores for acute appendicitis. *Scand J Gastroenterol* 1994;29(1):59-66.
- Fenyo G, Lindberg G, Blind P, Enochsson L, Oberg A. Diagnostic decision support in suspected acute appendicitis: validation of a simplified scoring system. *Eur J Surg* 1997;163(11):831-8.
- Ohmann C, Franke C, Yang Q. Clinical benefit of a diagnostic score for appendicitis: results of a prospective interventional study. *German Study Group of Acute Abdominal Pain. Arch Surg* 1999;134(9):993-6.
- Flint A. Symptomatology. In: Flint A, editor. *A treatise on the principles and practice of medicine*. 3rd ed. Philadelphia: Henry C. Lea;1868:p.108.
- Maedel LB, Doig K. Examination of the peripheral blood smear and correlation with the complete blood count. In: Rodak BF, Fritsma GA, Doig K, editors. *Hematology: clinical principles and applications*. 3rd ed. St. Louis, Missouri: Saunders Elsevier;2007:p.187.
- Bain BJ. Normal ranges. In: Bain BJ, editor. *Blood cells: a practical guide*. 3rd ed. Oxford: Blackwell Publishers; 2002:p.180.
- Steinvil A, Berliner S, Herishanu Y, Paran Y, Justo D, Cohen M, et al. Range and determinants of white blood cell count in a large survey of Israelis without inflammation. *Isr Med Assoc J* 2009;11(6):363-6.
- Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med* 1986;15(5):557-64.
- Kalan M, Talbot D, Cunliffe WJ, Rich AJ. Evaluation of the modified Alvarado score in the diagnosis of acute appendicitis: a prospective study. *Ann R Coll Surg Engl* 1994;76(6):418-9.
- Chan MY, Teo BS, Ng BL. The Alvarado score and acute appendicitis. *Ann Acad Med Singapore* 2001;30(5): 510-2.
- Bhattacharjee PK, Chowdhury T, Roy D. Prospective evaluation of modified Alvarado score for diagnosis of acute appendicitis. *J Indian Med Assoc* 2002;100(5): 310-1, 314.
- Chan MY, Tan C, Chiu MT, Ng YY. Alvarado score: an admission criterion in patients with right iliac fossa pain. *Surgeon* 2003;1(1):39-41.
- Al-Hashemy AM, Seleem MI. Appraisal of the modified Alvarado score for acute appendicitis in adults. *Saudi Med J* 2004;25(9):1229-31.
- Ohmann C, Yang Q, Franke C. Diagnostic scores for acute appendicitis. *Abdominal Pain Study Group. Eur J Surg* 1995;161(4):273-81.
- Liu JL, Wyatt JC, Deeks JJ, Clamp S, Keen J, Verde P, et al. Systematic reviews of clinical decision tools for acute abdominal pain. *Health Technol Assess* 2006;10 (47):1-167.